

WebSphere Application Server



Load Balancer Administratorhandbuch

Version 6.1

WebSphere Application Server



Load Balancer Administratorhandbuch

Version 6.1

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die allgemeinen Informationen in Anhang E, „Bemerkungen“, auf Seite 495 gelesen werden.

Erste Auflage (Mai 2006)

Diese Auflage gilt für:

WebSphere Application Server Version 6.1

und, sofern nicht anders angegeben, für alle nachfolgenden Releases und Modifikationen.

Veröffentlichungen können über den zuständigen IBM Ansprechpartner oder die zuständige IBM Geschäftsstelle bezogen werden.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs

IBM WebSphere Application Server Load Balancer Administrator's Guide, Version 6.1,

IBM Form GC31-6921-00,

herausgegeben von International Business Machines Corporation, USA

© Copyright International Business Machines Corporation 2006

© Copyright IBM Deutschland GmbH 2006

Informationen, die nur für bestimmte Länder Gültigkeit haben und für Deutschland, Österreich und die Schweiz nicht zutreffen, wurden in dieser Veröffentlichung im Originaltext übernommen.

Möglicherweise sind nicht alle in dieser Übersetzung aufgeführten Produkte in Deutschland angekündigt und verfügbar; vor Entscheidungen empfiehlt sich der Kontakt mit der zuständigen IBM Geschäftsstelle.

Änderung des Textes bleibt vorbehalten.

Herausgegeben von:

SW TSC Germany

Kst. 2877

Mai 2006

Inhaltsverzeichnis

Tabellen	xiii
---------------------------	-------------

Abbildungsverzeichnis	xv
--	-----------

Zu diesem Handbuch	xvii
-------------------------------------	-------------

Zielgruppe	xvii
Referenzliteratur	xvii
Zugriffsmöglichkeiten	xvii
Senden von Kommentaren	xviii

Referenzliteratur und zugehörige Websites	xix
--	------------

Teil 1. Einführung in Load Balancer	1
--	----------

Kapitel 1. Load Balancer im Überblick.	3
---	----------

Was ist Load Balancer?	3
Welche Komponenten von Load Balancer sollen verwendet werden?	3
Welche Vorteile bringt die Verwendung von Load Balancer?	4
Bereitstellung der hohen Verfügbarkeit durch Load Balancer	6
Dispatcher	6
CBR	6
Cisco CSS Controller oder Nortel Alteon Controller	6
Neue Features	7

Kapitel 2. Komponenten von Load Balancer im Überblick	9
--	----------

Komponenten von Load Balancer	9
Dispatcher im Überblick	9
Lokale Server mit dem Dispatcher verwalten	10
Serververwaltung mit Dispatcher und Metric Server	11
Lokale und ferne Server mit Dispatcher verwalten	11
Content Based Routing (CBR) im Überblick.	12
Lokale Server mit CBR verwalten	13
Site Selector im Überblick	13
Lokale und ferne Server mit Site Selector und Metric Server verwalten	14
Cisco CSS Controller im Überblick.	15
Nortel Alteon Controller im Überblick	16

Kapitel 3. Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer	19
---	-----------

Manager, Advisor-Funktionen und Metric Server (für Dispatcher, CBR und Site Selector)	19
Features der Komponente Dispatcher.	19
Fernverwaltung	19

Verknüpfung	19
Hohe Verfügbarkeit.	20
Client-Server-Affinität	20
Regelbasierter Lastausgleich	20
Inhaltsabhängiges Routing mit der Dispatcher-Weiterleitungsmethode cbr	21
Lastausgleich im WAN	22
Port-Zuordnung	22
Dispatcher in einem privaten Netz konfigurieren	22
Platzhaltercluster und Platzhalter-Port	23
Erkennung von DoS-Attacken	23
Binäres Protokollieren	23
Alerts	23

Features der Komponente CBR (Content Based Routing).	23
Komponente CBR und Dispatcher-Weiterleitungsmethode cbr im Vergleich.	24
Fernverwaltung	24
Verknüpfung	24
CBR mit mehreren Instanzen von Caching Proxy	24
Inhaltsabhängiges Routing für SSL-Verbindungen	24
Serverpartitionierung	24
Regelbasierter Lastausgleich	25
Client-Server-Affinität	25
Hohe Verfügbarkeit bei Verwendung von Dispatcher und CBR	26
Binäres Protokollieren	26
Alerts	26

Features der Komponente Site Selector	26
Fernverwaltung	26
Verknüpfung	26
Hohe Verfügbarkeit.	26
Client-Server-Affinität	26
Regelbasierter Lastausgleich	27
Lastausgleich im WAN	27
Alerts	27
Features der Komponente Cisco CSS Controller	27
Fernverwaltung	28
Verknüpfung	28
Hohe Verfügbarkeit.	28
Binäres Protokollieren	28
Alerts	28

Features der Komponente Nortel Alteon Controller	28
Fernverwaltung	29
Verknüpfung	29
Hohe Verfügbarkeit.	29
Binäres Protokollieren	29
Alerts	29

Kapitel 4. Load Balancer installieren	31
--	-----------

Systemvoraussetzungen und Installation unter AIX	31
Voraussetzungen für AIX-Systeme	31
Installation auf AIX-Systemen	32
Installation vorbereiten	32
Installationsschritte	33

Systemvoraussetzungen und Installation unter HP-UX	35
Voraussetzungen für HP-UX-Systeme.	35
Installation auf HP-UX-Systemen	35
Installation vorbereiten	35
Installationsschritte	35
Systemvoraussetzungen und Installation unter Linux	37
Voraussetzungen für Linux-Systeme	37
Installation auf Linux-Systemen	37
Installation vorbereiten	37
Installationsschritte	37
Systemvoraussetzungen und Installation unter Solaris	39
Voraussetzungen für Solaris	39
Installation unter Solaris	39
Installation vorbereiten	39
Installationsschritte	39
Systemvoraussetzungen und Installation unter Windows	40
Voraussetzungen für Windows-Systeme	40
Installation auf Windows-Systemen	40
Installation vorbereiten	41
Installationsschritte	41

Teil 2. Dispatcher 43

Kapitel 5. Schnellkonfiguration 45

Voraussetzungen	45
Vorbereitungen	46
Dispatcher konfigurieren	47
Konfiguration von der Befehlszeile aus	47
Konfiguration testen	48
Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus	48
Konfigurationsassistent	48
Arten von Cluster-, Port- und Serverkonfigurationen	48

Kapitel 6. Planung für Dispatcher 51

Überlegungen bei der Planung	51
Weiterleitungsmethoden	53
Dispatcher-Weiterleitungsmethode mac	53
Dispatcher-Weiterleitungsmethode nat	53
Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)	55
Beispielschritte für das Konfigurieren der Dispatcher-Weiterleitungsmethoden nat und cbr	57
Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)	58
Serverpartitionierung mit der HTTP- oder HTTP-PS-Advisor-Funktion	58
Beispiel für das Konfigurieren von logischen Servern auf einem physischen Server	59
Hohe Verfügbarkeit.	60
Einfache hohe Verfügbarkeit.	60
Gegenseitige hohe Verfügbarkeit	61

Kapitel 7. Dispatcher konfigurieren 63

Konfigurations-Tasks im Überblick	63
Konfigurationsmethoden	63

Befehlszeile	64
Scripts	64
GUI	65
Konfiguration mit dem Konfigurationsassistenten	66
Dispatcher-Maschine konfigurieren	66
Schritt 1. Serverfunktion starten	68
Schritt 2. Executor-Funktion starten	68
Schritt 3. NFA definieren (falls vom Hostnamen abweichend)	69
Schritt 4. Cluster definieren und Clusteroptionen festlegen	69
Schritt 5. Aliasnamen auf der Netzschnittstellenkarte erstellen	69
Schritt 6. Ports definieren und Port-Optionen festlegen	70
Schritt 7. Am Lastausgleich beteiligte Servermaschinen definieren	71
Schritt 8. Manager-Funktion starten (optional)	71
Schritt 9. Advisor-Funktion starten (optional)	71
Schritt 10. Clusterproportionen festlegen.	72
Servermaschinen für Lastausgleich konfigurieren	72
Schritt 1. Aliasnamen auf der Loopback-Einheit festlegen	72
Schritt 2. Überprüfung auf zusätzliche Route	76
Schritt 3. Zusätzliche Routen löschen	77
Schritt 4. Serverkonfiguration prüfen	77

Alternativen für die Festlegung eines Loopback-Aliasnamens unter Linux bei Verwendung der Load-Balancer-Weiterleitungsmethode mac	78
---	----

Kapitel 8. Deployment des Dispatchers von Load Balancer für IPV4 und IPV6 81

Unterstützte Plattformen für Load Balancer für IPV4 und IPV6	82
Unterstützte Plattformen für den Lastausgleich im Benutzerbereich	82
Spezielle Hinweise zur Linux-Plattform	82
Einschränkungen für Back-End-Server	82
Load Balancer für IPV4 und IPV6 installieren	83
Hinweise und Einschränkungen für Load Balancer für IPV4 und IPV6	83
Verbindungslokale IPV6-Adresse konfigurieren	84
Homogene Cluster-Server-Paare	84
Nicht unterstützte Dispatcher-Features	84
Advisor-Funktionen konfigurieren.	84
Hohe Verfügbarkeit konfigurieren	85
Server verknüpfen	86
Affinitätsfeature für im Benutzerbereich arbeitende Systeme (Linux).	86
Metric Server konfigurieren	87
Verarbeitung von IPV6-Paketen in Load Balancer für IPV4 und IPV6 aktivieren.	88
Aliasnamen auf der Schnittstelleneinheit für Load Balancer für IPV4 und IPV6 festlegen.	88
Erforderliche Schritte für die Clusterkonfiguration unter Linux auf zSeries	92
Dispatcher-Befehle (dscontrol) für Load Balancer für IPV4 und IPV6	92
Unterschiede in der Befehlssyntax	92
Unterstützte dscontrol-Befehle	93
Nicht unterstützte dscontrol-Befehle	96

Teil 3. CBR (Content Based Routing) 97

Kapitel 9. Schnellkonfiguration 99

Voraussetzungen	99
Vorbereitungen	100
CBR konfigurieren.	100
Konfiguration von der Befehlszeile aus	100
Konfiguration testen	102
Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus	102
Konfiguration mit dem Konfigurationsassistenten	102
Arten von Cluster-, Port- und Serverkonfigurationen	102

Kapitel 10. Planung für Content Based Routing 105

Überlegungen bei der Planung	105
Client-Anfragen nach verschiedenen Inhalten verteilen	106
Siteinhalt für kürzere Antwortzeiten aufteilen	106
Webserverinhalt sichern	107
CPU-Nutzung mit mehreren Caching-Proxy-Prozessen verbessern	107
Regelbasierter Lastausgleich mit CBR	107
Lastausgleich für sichere Verbindungen (SSL)	107
Lastausgleich für SSL-Datenverkehr vom Client zum Proxy und HTTP-Datenverkehr vom Proxy zum Server	108

Kapitel 11. Content Based Routing konfigurieren 109

Konfigurations-Tasks im Überblick	109
Konfigurationsmethoden	109
Befehlszeile	110
Scripts.	111
GUI	112
Konfigurationsassistent	113
CBR-Maschine konfigurieren	113
Schritt 1. Caching Proxy für die Verwendung von CBR konfigurieren	114
Schritt 2. Serverfunktion starten	115
Schritt 3. Executor-Funktion starten	115
Schritt 4. Cluster definieren und Clusteroptionen festlegen	116
Schritt 5. Aliasnamen auf der Netzschnittstellenkarte erstellen (optional).	116
Schritt 6. Ports definieren und Port-Optionen festlegen	117
Schritt 7. Am Lastausgleich beteiligte Servermaschinen definieren	117
Schritt 8. Regeln zur Konfiguration hinzufügen	118
Schritt 9. Server zu den Regeln hinzufügen	118
Schritt 10. Manager-Funktion starten (optional)	118
Schritt 11. Advisor-Funktion starten (optional)	118
Schritt 12. Clusterproportionen festlegen	118
Schritt 13. Caching Proxy starten	119
CBR-Konfigurationsbeispiel.	119

Teil 4. Site Selector 121

Kapitel 12. Schnellkonfiguration 123

Voraussetzungen	123
Vorbereitungen	123
Site Selector konfigurieren	124
Konfiguration von der Befehlszeile aus	124
Konfiguration testen	125
Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus	125
Konfiguration mit dem Konfigurationsassistenten	125

Kapitel 13. Planung für Site Selector 127

Überlegungen bei der Planung	127
Hinweise zu TTL	129
Netzproximität verwenden	130

Kapitel 14. Site Selector konfigurieren 131

Konfigurations-Tasks im Überblick	131
Konfigurationsmethoden	131
Befehlszeile	131
Scripts.	132
GUI	133
Konfigurationsassistent	134
Maschine mit Site Selector konfigurieren	134
Schritt 1. Serverfunktion starten	134
Schritt 2. Namensserver starten	134
Schritt 3. Sitenamen definieren und Optionen für Sitenamen festlegen	134
Schritt 4. Am Lastausgleich beteiligte Servermaschinen definieren.	135
Schritt 5. Manager-Funktion starten (optional)	135
Schritt 6. Advisor-Funktion starten (optional)	135
Schritt 7. Systemmesswert definieren (optional)	135
Schritt 8. Proportionen für den Sitenamen festlegen.	135
Servermaschinen für Lastausgleich konfigurieren	136

Teil 5. Cisco CSS Controller 137

Kapitel 15. Schnellkonfiguration 139

Voraussetzungen	139
Vorbereitungen	139
Cisco CSS Controller konfigurieren	140
Konfiguration von der Befehlszeile aus	140
Konfiguration testen	141
Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus	141

Kapitel 16. Planung für Cisco CSS Controller 143

Systemvoraussetzungen	143
Überlegungen bei der Planung	143
Position des Consultant im Netz	144
Hohe Verfügbarkeit	146
Wertigkeiten berechnen	147
Fehlerbestimmung.	147

Kapitel 17. Cisco CSS Controller konfigurieren 149

Konfigurations-Tasks im Überblick	149
Konfigurationsmethoden	149
Befehlszeile	149
XML	150
GUI	151
Maschine mit Controller für Cisco CSS Switches konfigurieren	152
Schritt 1. Serverfunktion starten	152
Schritt 2. Befehlszeilenschnittstelle aufrufen	152
Schritt 3. Consultant konfigurieren	153
Schritt 3. Eignerangaben konfigurieren	153
Schritt 4. Konfiguration der Services prüfen	153
Schritt 5. Messwerte konfigurieren	153
Schritt 6. Consultant starten	153
Schritt 7. Metric Server starten (optional)	153
Schritt 8. Hohe Verfügbarkeit konfigurieren (optional).	154
Konfiguration testen	154

Teil 6. Nortel Alteon Controller. . . 155

Kapitel 18. Schnellkonfiguration . . . 157

Voraussetzungen	157
Vorbereitungen	158
Nortel Alteon Controller konfigurieren	158
Konfiguration von der Befehlszeile aus	158
Konfiguration testen	159
Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus	159

Kapitel 19. Planung für Nortel Alteon Controller 161

Systemvoraussetzungen	161
Überlegungen bei der Planung	161
Position des Consultant im Netz	162
Vom Controller festgelegte Serverattribute auf dem Switch	165
Ausweichserver konfigurieren	165
Gruppen konfigurieren	166
Hohe Verfügbarkeit	167
Optimierung	169
Fehlerbestimmung.	169

Kapitel 20. Nortel Alteon Controller konfigurieren 171

Konfigurations-Tasks im Überblick	171
Konfigurationsmethoden	171
Befehlszeile	171
XML	172
GUI	173
Nortel Alteon Controller konfigurieren	174
Schritt 1. Serverfunktion starten	174
Schritt 2. Befehlszeilenschnittstelle aufrufen	174
Schritt 3. Consultant für den Nortel Alteon Web Switch definieren	175
Schritt 4. Service zum Switch-Consultant hinzufügen	175

Schritt 5. Messwerte konfigurieren	175
Schritt 6. Consultant starten	175
Schritt 7. Hohe Verfügbarkeit konfigurieren (optional).	175
Schritt 8. Metric Server starten (optional)	175
Schritt 9. Konfiguration für Nortel Alteon Controller aktualisieren.	176
Konfiguration testen	176

Teil 7. Funktionen und erweiterte Features von Load Balancer . . . 177

Kapitel 21. Manager, Advisor-Funktionen und Metric Server für Dispatcher, CBR und Site Selector 179

Lastausgleich mit Load Balancer optimieren	180
Proportionale Bedeutung von Statusinformationen	180
Wertigkeiten.	181
Manager-Intervalle	183
Sensitivitätsschwelle	183
Glättungsfaktor.	184
Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden	184
Advisor-Funktionen	185
Arbeitsweise der Advisor-Funktionen	186
Advisor-Funktion starten und stoppen	186
Advisor-Intervalle	187
Berichtszeitlimit für Advisor-Funktion	188
Serververbindungs- und -empfangszeitlimit der Advisor-Funktion	188
Wiederholungsversuche der Advisor-Funktion	188
Liste der Advisor-Funktionen	189
Option 'Anforderung/Antwort (URL)' der HTTP- oder HTTPS-Advisor-Funktion konfigurieren	191
Advisor-Funktion 'self' in einer Client/Server-WAN-Konfiguration	192
Kundenspezifische (anpassbare) Advisor-Funktion erstellen	193
WAS-Advisor-Funktion	194
Namenskonvention	194
Kompilierung	194
Ausführung	195
Erforderliche Routinen	195
Suchreihenfolge	196
Benennung und Pfad.	196
Beispiel-Advisor-Funktion	196
Metric Server	196
WLM-Einschränkung.	197
Vorbedingungen	197
Metric Server verwenden	197
Advisor-Funktion Workload Manager	199
Einschränkung für Metric Server	200

Kapitel 22. Erweiterte Funktionen für Dispatcher, CBR und Site Selector . . 201

Verknüpfte Server verwenden	202
Für Dispatcher	203

Für CBR	204
Für Site Selector	204
Hohe Verfügbarkeit	204
Hohe Verfügbarkeit konfigurieren	205
Fehlererkennung mit Hilfe von Überwachungs- signal und Erreichbarkeitsziel	208
Wiederherstellungsstrategie.	208
Scripts verwenden.	209
Verknüpfung und hohe Verfügbarkeit konfigu- rieren (Windows-Systeme)	212
Regelbasierten Lastausgleich konfigurieren	212
Wie werden Regeln ausgewertet?.	213
Auf der Client-IP-Adresse basierende Regeln verwenden	214
Auf dem Client-Port basierende Regeln verwen- den.	214
Auf der Uhrzeit basierende Regeln verwenden	214
Auf der Serviceart basierende Regeln verwen- den.	215
Regeln auf der Basis der Verbindungen pro Sekunde verwenden	215
Regeln auf der Basis der Summe aktiver Verbin- dungen verwenden	215
Regeln auf der Basis der reservierten und gemeinsam benutzten Bandbreite verwenden.	216
Regel 'Metrik gesamt'	218
Regel 'Metrik Durchschnitt'	218
Immer gültige Regeln verwenden	219
Auf dem Inhalt der Anforderung basierende Regeln verwenden.	219
Port-Affinität außer Kraft setzen	220
Regeln zur Konfiguration hinzufügen	220
Regeloption für Serverauswertung	220
Funktionsweise der Affinität für Load Balancer	222
Verhalten bei inaktiver Affinität	222
Verhalten bei aktivierter Affinität	222
Port-übergreifende Affinität	223
Affinitätsadressmaske (stickymask)	223
Bearbeitung von Serververbindungen stilllegen	224
Affinitätsoption der Regel ausgehend vom Inhalt der Client-Anfrage	225
Aktive Cookie-Affinität	225
Passive Cookie-Affinität	227
URI-Affinität	228
Dispatcher-WAN-Unterstützung konfigurieren	229
Befehlssyntax	230
Ferne Advisor-Funktionen mit der Dispatcher- WAN-Unterstützung verwenden	231
Konfigurationsbeispiel	233
Unterstützung für GRE (Generic Routing Encap- sulation)	235
Explizite Verbindungen benutzen.	236
Konfiguration für ein privates Netz verwenden	236
Platzhaltercluster zum Zusammenfassen von Serverkonfigurationen verwenden	237
Platzhaltercluster für den Lastausgleich von Fire- walls verwenden	238
Platzhaltercluster mit Caching Proxy für transpa- rente Weiterleitung verwenden	239
Platzhalter-Port für die Übertragung von Daten- verkehr mit nicht konfiguriertem Port verwenden	239

Platzhalter-Port für FTP-Datenverkehr verwen- den.	239
Erkennung von DoS-Attacken	240
Binäre Protokolle für die Analyse von Server- statistiken verwenden	241
Verknüpften Client verwenden	243

Kapitel 23. Erweiterte Funktionen für Cisco CSS Controller und Nortel

Alteon Controller	245
Verknüpfung	245
Hohe Verfügbarkeit	245
Konfiguration	246
Ausfallerkennung	247
Wiederherstellungsstrategie.	247
Beispiele	248
Lastausgleich mit Load Balancer optimieren	248
Bedeutung von Messwerten	248
Wertigkeiten.	249
Ruhezeiten für Wertigkeitsberechnung	249
Sensitivitätsschwelle	250
Advisor-Funktionen	250
Arbeitsweise der Advisor-Funktionen	250
Advisor-Ruhezeiten	251
Serververbindungs- und -empfangszeitlimit der Advisor-Funktion	251
Wiederholungsversuche der Advisor-Funktion	251
Kundenspezifische (anpassbare) Advisor-Funktion erstellen	252
Namenskonvention	253
Kompilierung	253
Ausführung	254
Erforderliche Routinen	254
Suchreihenfolge	255
Benennung und Pfad.	255
Beispiel-Advisor-Funktion	255
Metric Server	255
Vorbedingungen	255
Metric Server verwenden	255
Advisor-Funktion Workload Manager	257
Binäre Protokolle für die Analyse von Server- statistiken verwenden	258
Scripts zum Generieren eines Alerts oder Protokol- lieren eines Serverausfalls verwenden	259

Teil 8. Verwaltung von Load Balancer und Fehlerbehebung 261

Kapitel 24. Betrieb und Verwaltung von Load Balancer 263

Fernverwaltung von Load Balancer	263
Remote Method Invocation (RMI)	264
Webgestützte Verwaltung	265
Protokolle von Load Balancer verwenden	267
Für Dispatcher, CBR und Site Selector	267
Für Cisco CSS Controller und Nortel Alteon Controller	269
Komponente Dispatcher verwenden.	270
Dispatcher starten und stoppen	270

Inaktivitätszeitlimit verwenden	270
Bereinigung von Verbindungssätzen mit fintimeout und staletimeout steuern	271
Berichte der GUI — Menüoption 'Überwachen'	271
Simple Network Management Protocol mit Dispatcher verwenden	271
Gesamten Datenverkehr zur Sicherheit der Load-Balancer-Maschine mit ipchains oder iptables zurückweisen (Linux-Systeme).	278
Komponente Content Based Routing verwenden	279
CBR starten und stoppen	279
CBR steuern.	280
CBR-Protokolle verwenden.	280
Site Selector verwenden	280
Site Selector starten und stoppen.	280
Site Selector steuern	280
Protokolle von Site Selector verwenden.	280
Cisco CSS Controller verwenden	280
Cisco CSS Controller starten und stoppen	280
Cisco CSS Controller steuern	281
Protokolle von Cisco CSS Controller verwenden	281
Nortel Alteon Controller verwenden.	281
Nortel Alteon Controller starten und stoppen	281
Nortel Alteon Controller steuern	281
Protokolle von Nortel Alteon Controller verwenden.	281
Metric Server verwenden	282
Metric Server starten und stoppen	282
Protokolle von Metric Server verwenden	282

Kapitel 25. Fehlerbehebung 283

Informationen zur Fehlerbehebung abrufen	283
Basisinformationen (immer erforderlich)	283
Probleme mit der hohen Verfügbarkeit	284
Advisor-Fehler	285
Fehler beim Content Based Routing	286
Cluster nicht erreichbar	286
Alle Versuche sind gescheitert.	287
Upgrades.	287
Java-Code	287
Nützliche Links	287
Fehlerbehebungstabellen	287
Port-Nummern für Dispatcher überprüfen.	301
Port-Nummern für CBR überprüfen.	302
Port-Nummern für Site Selector überprüfen	302
Port-Nummern für Cisco CSS Controller überprüfen	303
Port-Nummern für Nortel Alteon Controller überprüfen.	304
Allgemeine Probleme lösen — Dispatcher	305
Problem: Dispatcher wird nicht ausgeführt	305
Problem: Dispatcher und Server antworten nicht	305
Problem: Dispatcher-Anforderungen werden nicht verteilt.	305
Problem: Die Dispatcher-Funktion für hohe Verfügbarkeit kann nicht ausgeführt werden	305
Problem: Es kann kein Überwachungssignal hinzugefügt werden (Windows-Plattform)	306
Problem: Zusätzliche Routen (Windows 2000)	306
Problem: Advisor-Funktionen arbeiten nicht korrekt.	306

Problem: Dispatcher, Microsoft IIS und SSL funktionieren nicht (Windows-Plattform)	306
Problem: Dispatcher-Verbindung zu einer fernen Maschine.	307
Problem: Der Befehl dscontrol oder lbadm scheitert	307
Problem: Fehlernachricht 'Datei nicht gefunden...' beim Anzeigen der Onlinehilfe (Windows-Plattform)	308
Problem: Die grafische Benutzerschnittstelle (GUI) wird nicht richtig gestartet.	308
Problem: Fehler bei der Ausführung von Dispatcher mit installiertem Caching Proxy	308
Problem: Die grafische Benutzerschnittstelle (GUI) wird nicht richtig angezeigt	308
Problem: Auf der Windows-Plattform sind die Hilfefenster manchmal von anderen offenen Fenstern verdeckt	308
Problem: Load Balancer kann Rahmen nicht verarbeiten und weiterleiten	309
Problem: Beim Starten des Executors von Load Balancer erscheint eine blaue Anzeige	309
Problem: Automatische Pfaderkennung verhindert Datenrückfluss mit Load Balancer	309
Problem: Keine hohe Verfügbarkeit im Weitverkehrsmodus von Load Balancer	310
Problem: Beim Laden einer großen Konfigurationsdatei blockiert die GUI oder verhält sich nicht erwartungsgemäß	311
Problem: lbadm trennt nach dem Aktualisieren der Konfiguration die Verbindung zum Server	311
Problem: IP-Adressen werden über die Fernverbindung nicht richtig aufgelöst	312
Problem: Auf der koreanischen Schnittstelle von Load Balancer werden bei AIX- und Linux-Systemen überlappende oder unpassende Schriftarten angezeigt	312
Problem: Auf Windows-Systemen wird beim Absetzen von Befehlen wie hostname an Stelle der lokalen Adresse die Aliasadresse zurückgegeben	312
Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten	313
Problem: Unerwartetes Verhalten bei Ausführung von 'rmmod ibmlb' (Linux-Systeme).	313
Problem: Lange Antwortzeiten beim Ausführen von Befehlen auf der Dispatcher-Maschine	313
Problem: Bei Verwendung der Weiterleitungsmethode mac registriert die Advisor-Funktion SSL oder HTTPS keine Serverlast.	314
Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung.	314
Problem: Bei aktiviertem Socket-Pooling wird der Webserver an 0.0.0.0 gebunden	314
Problem: Auf Windows-Systemen erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1).	315
Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads	315

Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv.	316	Allgemeine Probleme lösen — CBR	328
Problem: Auflösung von IP-Adressen in Hostna- men auf Windows-Systemen, wenn für einen Adapter mehrere Adressen konfiguriert sind . . .	316	Problem: CBR wird nicht ausgeführt	328
Problem: Advisor-Funktionen können auf Win- dows-Systemen in einer Konfiguration für hohe Verfügbarkeit nach einem Netzwerkausfall nicht ausgeführt werden	317	Problem: Der Befehl cbrcontrol oder lbadm- in scheitert	328
Problem: Beim Festlegen von Aliasnamen für mehrere Cluster auf der Loopback-Einheit von Linux-Systemen nicht den Befehl 'IP address add' verwenden	318	Problem: Anforderungen werden nicht verteilt	329
Problem: Fehlermeldung 'Router-Adresse nicht angegeben oder nicht gültig für Port-Methode' erscheint	318	Problem: Auf Solaris-Systemen scheitert der Befehl cbrcontrol executor start	329
Problem: Auf Solaris-Systemen werden Load- Balancer-Prozesse beim Verlassen des Terminal- fensters, in dem die Prozesse gestartet wurden, beendet	319	Problem: Syntax- oder Konfigurationsfehler . . .	329
Problem: Verzögerung beim Laden einer Load- Balancer-Konfiguration	319	Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten	330
Problem: Fehlermeldung zu einem IP-Adressen- konflikt auf Windows-Systemen	319	Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung.	330
Problem: In einer Konfiguration mit hoher Ver- fügbarkeit sind sowohl primäre Maschinen als auch Ausweichmaschinen aktiv	320	Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nati- onalen Sonderzeichen (Latin-1)	330
Problem: Client-Anforderungen scheitern, wenn die Rückgabe von Antworten in Form großer Seiten versucht wird	320	Problem: Java-Fehler unter HP-UX wegen unzu- reichender Speicherkapazität/Threads	330
Auf Windows-Systemen wird beim Absetzen eines Befehls dscontrol oder lbadm in der Fehler 'Der Server antwortet nicht' angezeigt	321	Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv.	330
Problem: Dispatcher-Maschinen mit hoher Ver- fügbarkeit können unter Linux auf S/390 mit qeth-Treibern möglicherweise nicht synchroni- siert werden.	321	Problem: Auflösung von IP-Adressen in Hostna- men auf Windows-Systemen, wenn für einen Adapter mehrere Adressen konfiguriert sind . . .	331
Problem: Konfigurationstipps für hohe Verfüg- barkeit	321	Allgemeine Fehler beheben — Site Selector . . .	331
Problem: Einschränkungen für die Dispatcher- Konfiguration unter Linux auf zSeries- oder S/390-Servern mit OSA-Karten	323	Problem: Site Selector wird nicht ausgeführt . . .	331
Problem: Speicherverlust unter einigen Linux- Versionen bei Ausführung von Dispatcher mit Manager und Advisor-Funktionen	325	Problem: Site Selector verteilt den Datenverkehr von Solaris-Clients nicht nach der RoundRobin- Methode	331
Problem: Unter SUSE Linux Enterprise Server 9 leitet Dispatcher Pakete weiter, die den Back- End-Server jedoch nicht erreichen	325	Problem: Der Befehl sscontrol oder lbadm scheitert	331
Problem: Fehlermeldung zu einem IP-Adressen- konflikt auf Windows-Systemen während der Funktionsübernahme bei hoher Verfügbarkeit	326	Problem: ssserver wird auf der Windows-Platt- form nicht gestartet	332
Problem: Mögliche Störung der Paket- weiterleitung durch iptables unter Linux . . .	327	Problem: Site Selector führt bei duplizierten Routen den Lastausgleich nicht korrekt durch . .	332
Problem: Auf Solaris-Systemen kann kein IPV6- Server zur Load-Balancer-Konfiguration hinzu- gefügt werden	327	Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten	333
Java-Warnung bei Installation von Service- korrekturen	328	Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung.	333
Upgrade der Java-Dateigruppe einer Load-Ba- lancer-Installation	328	Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nati- onalen Sonderzeichen (Latin-1)	333
		Problem: Java-Fehler unter HP-UX wegen unzu- reichender Speicherkapazität/Threads	333
		Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv.	333
		Allgemeine Probleme lösen — Cisco CSS Control- ler	334
		Problem: ccoserver wird nicht gestartet.	334
		Problem: Der Befehl ccocontrol oder lbadm scheitert	334
		Problem: Für Port 13099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden . .	334
		Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten	335

Problem: Beim Hinzufügen eines Consultant wird ein Verbindungsfehler empfangen.	335
Problem: Auf dem Switch werden die Wertigkeiten nicht aktualisiert	335
Problem: Befehl refresh aktualisiert die Consultant-Konfiguration nicht.	336
Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung.	336
Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)	336
Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads	336
Allgemeine Probleme lösen — Nortel Alteon Controller	336
Problem: nalserver wird nicht gestartet.	336
Problem: Der Befehl nalcontrol oder lbadm scheitert	337
Problem: Für Port 14099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden	337
Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten	337
Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung.	338
Problem: Beim Hinzufügen eines Consultant wird ein Verbindungsfehler empfangen.	338
Problem: Auf dem Switch werden die Wertigkeiten nicht aktualisiert	338
Problem: Befehl refresh aktualisiert die Consultant-Konfiguration nicht.	338
Problem: Auf Windows-Systemen erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1).	339
Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads	339
Allgemeine Fehler beheben — Metric Server	339
Problem: IOException für Metric Server auf der Windows-Plattform bei Ausführung von benutzerdefinierten Messwertdateien mit der Erweiterung .bat oder .cmd.	339
Problem: Metric Server meldet die Last nicht an die Load-Balancer-Maschine	339
Problem: Metric-Server-Protokoll meldet, dass für den Zugriff auf den Agenten eine Kennung erforderlich ist	340
Problem: Bei Ausführung von Metric Server auf AIX-Systemen kann die Ausgabe des Befehls ps -vg beschädigt werden	340
Problem: Konfigurieren von Metric Server in einer zweistufigen Konfiguration mit Site Selector für die Lastverteilung auf Dispatcher mit hoher Verfügbarkeit	340
Problem: Unerwünschte Konsolnachrichten bei Ausführung von Scripts auf Solaris-Maschinen mit mehreren CPUs	342
Bei Ausführung von Load Balancer für IPV6 ist auf Linux-Systemen kein Abruf der Werte von Metric Server möglich.	342

Problem: Nach dem Start von Metric Server wird der Wert -1 zurückgegeben	343
--	-----

Teil 9. Befehlsreferenz 345

Kapitel 26. Syntaxdiagramm lesen . . 347

Symbole und Interpunktion	347
Parameter	347
Beispiele für die Syntax	347

Kapitel 27. Befehlsreferenz für Dispatcher und CBR 349

Konfigurationsunterschiede bei CBR und Dispatcher	350
dscontrol advisor — Advisor-Funktion steuern	351
dscontrol binlog — Binäre Protokolldatei steuern	356
dscontrol cluster — Cluster konfigurieren	357
dscontrol executor — Executor steuern	361
dscontrol file — Konfigurationsdateien verwalten	366
dscontrol help — Hilfetext für diesen Befehl anzeigen oder drucken	368
dscontrol highavailability — Hohe Verfügbarkeit steuern	369
dscontrol host — Ferne Maschine konfigurieren	373
dscontrol logstatus — Protokolleinstellungen des Servers anzeigen	374
dscontrol manager — Manager steuern.	375
dscontrol metric — Systemmesswerte konfigurieren	381
dscontrol port — Ports konfigurieren	382
dscontrol rule — Regeln konfigurieren	388
dscontrol server — Server konfigurieren	394
dscontrol set — Serverprotokoll konfigurieren	400
dscontrol status — Aktivitätsanzeige für Manager und Advisor-Funktionen	401
dscontrol subagent — SNMP-Subagenten konfigurieren	402

Kapitel 28. Befehlsreferenz für Site Selector. 405

sscontrol advisor — Advisor-Funktion steuern	406
sscontrol file — Konfigurationsdateien verwalten	411
sscontrol help — Hilfetext für diesen Befehl anzeigen oder drucken	413
sscontrol logstatus — Protokolleinstellungen des Servers anzeigen	414
sscontrol manager — Manager steuern	415
sscontrol metric — Systemmesswerte konfigurieren	420
sscontrol nameserver — Namensserver steuern	421
sscontrol rule — Regeln konfigurieren	422
sscontrol server — Server konfigurieren	425
sscontrol set — Serverprotokoll konfigurieren	427
sscontrol sitename — Sitenamen konfigurieren	428
sscontrol status — Aktivitätsanzeige für Manager und Advisor-Funktionen	431

Kapitel 29. Befehlsreferenz für Cisco CSS Controller. 433

cococontrol consultant — Consultant konfigurieren und steuern	434
---	-----

cococontrol controller — Controller steuern	437
cococontrol file — Konfigurationsdateien verwalten	439
cococontrol help — Hilfetext für diesen Befehl anzeigen oder drucken	440
cococontrol highavailability — Hohe Verfügbarkeit steuern	441
cococontrol metriccollector — Messwerterfassung konfigurieren	444
cococontrol ownercontent — Eigernamen und content-Regel steuern	446
cococontrol service — Service konfigurieren	449

Kapitel 30. Befehlsreferenz für Nortel

Alteon Controller 451

nalcontrol consultant — Consultant konfigurieren und steuern	452
nalcontrol controller — Controller steuern	455
nalcontrol file — Konfigurationsdateien verwalten	457
nalcontrol help — Hilfetext für diesen Befehl anzeigen oder drucken	458
nalcontrol highavailability — Hohe Verfügbarkeit steuern	459
nalcontrol metriccollector — Messwerterfassung konfigurieren	462
nalcontrol server — Server konfigurieren	464
nalcontrol service — Service konfigurieren	466

Anhang A. Allgemeine Anweisungen

zur GUI 469

Anhang B. Syntax der content-Regel 477

Syntax der content-Regel	477
Reservierte Schlüsselwörter	477

Anhang C. Beispielkonfigurations-

dateien 481

Beispielkonfigurationsdateien für Load Balancer	481
Dispatcher-Konfigurationsdatei — AIX-, Linux- und Solaris-Systeme	481
Dispatcher-Konfigurationsdatei — Windows-Systeme	484
Beispiel-Advisor-Funktion	487

Anhang D. Beispiel für eine Client/Server-Konfiguration mit hoher Verfügbarkeit unter Verwendung von Dispatcher, CBR und Caching Proxy . . . 491

Servermaschine einrichten	491
-------------------------------------	-----

Anhang E. Bemerkungen 495

Marken	497
------------------	-----

Glossar 499

Index 509

Tabellen

1.	installp-Images für AIX	32	11.	Konfigurations-Tasks für Nortel Alteon Controller	171
2.	AIX-Installationsbefehle	34	12.	Erweiterte Konfigurations-Tasks für Load Balancer	179
3.	Details zur Paketinstallation von Load Balancer unter HP-UX	35	13.	Erweiterte Konfigurations-Tasks für Load Balancer	201
4.	Konfigurations-Tasks für Dispatcher	63	14.	Tabelle zur Fehlerbehebung für Dispatcher	288
5.	Befehle zum Festlegen eines Aliasnamens auf der Loopback-Einheit (lo0) für Dispatcher	73	15.	Tabelle zur Fehlerbehebung für CBR	294
6.	Befehle zum Löschen zusätzlicher Routen für Dispatcher	77	16.	Tabelle zur Fehlerbehebung für Site Selector	295
7.	Konfigurations-Tasks für die Komponente CBR	109	17.	Tabelle zur Fehlerbehebung für Controller für Cisco CSS Switches	297
8.	Befehle zum Erstellen eines Aliasnamens auf der NIC	116	18.	Tabelle zur Fehlerbehebung für Nortel Alteon Controller	298
9.	Konfigurations-Tasks für Site Selector	131	19.	Tabelle zur Fehlerbehebung für Metric Server	299
10.	Konfigurations-Tasks für Cisco CSS Controller	149			

Abbildungsverzeichnis

1. Beispiel für die physische Darstellung einer Site mit Dispatcher für die Verwaltung lokaler Server	10
2. Beispielsite mit Dispatcher und Metric Server für die Serververwaltung	11
3. Beispiel für eine Site mit Dispatcher für die Verwaltung lokaler und ferner Server	11
4. Beispielsite mit CBR für die Verwaltung lokaler Server	13
5. Beispielsite mit Site Selector und Metric Server für die Verwaltung lokaler und ferner Server	14
6. Beispielsite mit Cisco CSS Controller und Metric Server für die Verwaltung lokaler Services	16
7. Beispielsite mit Nortel Alteon Controller für die Verwaltung lokaler Server	17
8. Einfache lokale Dispatcher-Konfiguration	45
9. Dispatcher-Beispielkonfiguration mit einem Cluster und zwei Ports.	48
10. Dispatcher-Beispielkonfiguration mit zwei Clustern mit jeweils einem Port.	49
11. Dispatcher-Beispielkonfiguration mit zwei Clustern mit jeweils zwei Ports	50
12. Beispiel für die Dispatcher-Weiterleitungsmethoden nat und cbr	57
13. Beispiel für einen Dispatcher mit einfacher hoher Verfügbarkeit.	60
14. Beispiel für einen Dispatcher mit gegenseitiger hoher Verfügbarkeit.	61
15. Beispiel der für die Dispatcher-Maschine erforderlichen IP-Adressen	68
16. Einfache lokale CBR-Konfiguration.	99
17. CBR-Beispielkonfiguration mit einem Cluster und zwei Ports	102
18. CBR-Beispielkonfiguration mit zwei Clustern mit jeweils einem Port	103
19. CBR-Beispielkonfiguration mit zwei Clustern mit jeweils zwei Ports.	104
20. CBR-Konfigurationsdatei für AIX-, Linux- und Solaris-Systeme	115
21. CBR-Konfigurationsdatei für HP-UX-Systeme	115
22. CBR-Konfigurationsdatei für Windows-Systeme	115
23. Einfache Site-Selector-Konfiguration	123
24. Beispiel für eine DNS-Umgebung.	127
25. Einfache Konfiguration mit Cisco CSS Controller	139
26. Beispiel für einen Consultant, der hinter den Switches mit dem Netz verbunden ist	145
27. Beispiel für einen Consultant (optional mit einem Partner für hohe Verfügbarkeit) hinter dem Switch mit einer Benutzerschnittstelle vor dem Switch.	146
28. Einfache Konfiguration mit Nortel Alteon Controller.	157
29. Beispiel für einen Consultant, der hinter dem Switch mit dem Netz verbunden ist	163
30. Beispiel für einen Consultant, der über ein Intranet vor dem Switch mit dem Netz verbunden ist.	164
31. Beispiel für einen Consultant hinter dem Switch mit einer Benutzerschnittstelle vor dem Switch	164
32. Beispiel für einen Consultant mit Ausweichservern	166
33. Beispiel für hohe Verfügbarkeit mit einem Nortel Alteon Controller und einem Nortel Alteon Web Switch.	168
34. Beispiel für eine Client/Server-WAN-Konfiguration mit Advisor-Funktion self	192
35. Beispiel einer Konfiguration mit einem LAN-Segment	229
36. Beispiel einer Konfiguration mit lokalen und fernen Servern	230
37. WAN-Beispielkonfiguration mit fernen Load-Balancer-Maschinen	233
38. WAN-Beispielkonfiguration mit einer Serverplattform, die GRE unterstützt	235
39. Beispiel für ein privates Netz mit dem Dispatcher	237
40. SNMP-Befehle für Linux- und UNIX-Systeme	273
41. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Dispatcher	470
42. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente CBR	471
43. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Site Selector	472
44. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Cisco CSS Controller	473
45. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Nortel Alteon Controller.	474
46. Beispiel für eine Client/Server-Konfiguration mit hoher Verfügbarkeit unter Verwendung von Dispatcher, CBR und Caching Proxy	491

Zu diesem Handbuch

In diesem Handbuch sind die Planung, Installation, Konfiguration, Verwendung und Fehlerbehebung für IBM WebSphere Application Server Load Balancer für die Betriebssysteme AIX, HP-UX, Linux, Solaris und Windows beschrieben. Zuvor hatte dieses Produkt den Namen Edge Server Network Dispatcher, SecureWay Network Dispatcher, eNetwork Dispatcher und Interactive Network Dispatcher.

Zielgruppe

Das *Load Balancer Administratorhandbuch* ist für erfahrene Netz- und Systemadministratoren geschrieben, die sich mit den von ihnen verwendeten Betriebssystemen und dem Bereitstellen von Internet-Services auskennen. Vorkenntnisse zu Load Balancer sind nicht erforderlich.

Dieses Handbuch bietet keine Unterstützung für frühere Releases von Load Balancer.

Referenzliteratur

Die Website mit dem Edge Components Informationszentrum enthält Links zur aktuellen Version dieses Handbuchs in den Formaten HTML und PDF.

Die letzten Aktualisierungen zu Load Balancer finden Sie auf der Support-Seite der Website. Klicken Sie auf dieser Seite auf den Link zur Technote-Site.

Die URLs dieser und weiterer Webseiten sind im Abschnitt „Referenzliteratur und zugehörige Websites“ auf Seite xix aufgelistet.

Zugriffsmöglichkeiten

Features zur Erleichterung des Zugriffs helfen körperbehinderten Benutzern (mit eingeschränkter Beweglichkeit oder Sehschwäche), Softwareprodukte erfolgreich anzuwenden. Load Balancer bietet im Wesentlichen die folgenden Features für verbesserte Zugriffsmöglichkeiten an:

- Sie können ein Bildschirmleseprogramm und einen digitalen Sprachsynthesizer verwenden, um zu hören, was auf dem Bildschirm angezeigt wird. Für die Dateneingabe und die Navigation auf der Benutzerschnittstelle können Sie Spracherkennungssoftware wie IBM ViaVoice einsetzen.
- Für die Ausführung von Funktionen können Sie an Stelle der Maus die Tastatur benutzen.
- Für die Konfiguration und Verwaltung der Features von Load Balancer können Sie an Stelle der bereitgestellten grafischen Oberflächen auch Standardtexteditoren oder Befehlszeilenschnittstellen verwenden. Weitere Informationen zu den Zugriffsmöglichkeiten für bestimmte Features finden Sie in der Dokumentation zu diesen Features.

Senden von Kommentaren

Ihre Rückmeldung ist uns wichtig, damit wir möglichst genaue und hochwertige Informationen bieten können. Falls Sie Kommentare zum vorliegenden Handbuch oder einem anderen Dokument zu Edge Components abgeben möchten:

- Senden Sie Ihren Kommentar per E-Mail an fsdoc@us.ibm.com. Geben Sie dabei Folgendes an: Handbuchtitel, Teilenummer des Handbuchs, Version und wenn möglich die Position der Textstelle, auf die sich Ihr Kommentar bezieht (z. B. Seitenzahl oder Tabellenummer).

Referenzliteratur und zugehörige Websites

- *Edge Components Konzepte, Planung und Installation*, IBM Form GC12-3678-00
- *Programming Guide for Edge Components*, IBM Form GC31-6919-00
- *Caching Proxy Administratorhandbuch*, IBM Form GC12-31679-00
- IBM Homepage: www.ibm.com/
- Produktsite zu IBM WebSphere Application Server: www.ibm.com/software/webservers/appserv/
- Library-Website zu IBM WebSphere Application Server: www.ibm.com/software/webservers/appserv/was/library/
- Support-Website zu IBM WebSphere Application Server: www.ibm.com/software/webservers/appserv/was/support/
- IBM WebSphere Application Server Information Center: www.ibm.com/software/webservers/appserv/infocenter.html
- IBM WebSphere Application Server Edge Components Information Center: www.ibm.com/software/webservers/appserv/ecinfocenter.html

Teil 1. Einführung in Load Balancer

Dieser Teil gibt einen Überblick über Load Balancer und die zugehörigen Produktkomponenten. Er enthält außerdem eine kurze Beschreibung der verfügbaren Konfigurationsfunktionen, eine Auflistung der Hardware- und Softwarevoraussetzungen sowie Installationsanweisungen. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 1, „Load Balancer im Überblick“, auf Seite 3
- Kapitel 2, „Komponenten von Load Balancer im Überblick“, auf Seite 9
- Kapitel 3, „Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer“, auf Seite 19
- Kapitel 4, „Load Balancer installieren“, auf Seite 31

Kapitel 1. Load Balancer im Überblick

Dieses Kapitel gibt einen Überblick über Load Balancer und ist in die folgenden Abschnitte gegliedert:

- „Was ist Load Balancer?“
- „Welche Komponenten von Load Balancer sollen verwendet werden?“
- „Welche Vorteile bringt die Verwendung von Load Balancer?“ auf Seite 4
- „Bereitstellung der hohen Verfügbarkeit durch Load Balancer“ auf Seite 6
- „Neue Features“ auf Seite 7

Eine detaillierte Liste der Konfigurationsfunktionen der einzelnen Komponenten von Load Balancer, die Sie zur Planung Ihrer Netzverwaltung heranziehen können, finden Sie in Kapitel 3, „Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer“, auf Seite 19.

Was ist Load Balancer?

Load Balancer ist eine Softwarelösung für die Verteilung eingehender Client-Anforderungen auf mehrere Server. Dieses Produkt verbessert die Leistung von Servern erheblich, indem es TCP/IP-Sitzungsanforderungen auf verschiedene Server einer Gruppe verteilt. Dieser Lastausgleich ist für Benutzer und andere Anwendungen transparent. Load Balancer ist vor allem bei Anwendungen wie E-Mail-Servern, WWW-Servern, verteilten Abfragen paralleler Datenbanken und anderen TCP/IP-Anwendungen nützlich.

Beim Einsatz mit Webservern kann Load Balancer zur optimalen Nutzung des Potenzials Ihrer Site beitragen, da er eine leistungsfähige, flexible und skalierbare Lösung für Probleme bietet, die durch eine sehr hohe Belastung auftreten können. Haben Besucher zu Zeiten höchster Belastung Schwierigkeiten, auf Ihre Site zuzugreifen, können Sie mit Load Balancer automatisch den optimalen Server zur Bearbeitung eingehender Anforderungen suchen.

Welche Komponenten von Load Balancer sollen verwendet werden?

Wichtiger Hinweis: Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, steht nur die Dispatcher-Komponente zur Verfügung. Weitere Informationen finden Sie in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

Load Balancer besteht aus den folgenden fünf Komponenten, die separat oder zusammen verwendet werden können, um bessere Ergebnisse beim Lastausgleich zu erzielen:

- Sie können allein mit der Komponente **Dispatcher** einen Lastausgleich auf Servern in einem lokalen Netz oder Weitverkehrsnetz durchführen, indem Sie die von Dispatcher dynamisch festgelegten Wertigkeiten und Messungen verwenden. Diese Komponente führt den Lastausgleich auf der Ebene bestimmter Dienste aus, beispielsweise für HTTP, FTP, SSL, NNTP, IMAP, POP3, SMTP, SIP und Telnet. Die Komponente benutzt für die Zuordnung von Domännennamen zu IP-Adressen keinen Domännennamensserver.

Wenn Sie mit dem Protokoll HTTP arbeiten, können Sie auch die Dispatcher-Funktion für inhaltsabhängige Weiterleitung verwenden, um die Last ausgehend

vom Inhalt der Client-Anfrage zu verteilen. Der ausgewählte Server ist das Ergebnis des Abgleichs des URL mit einer angegebenen Regel. Für die inhaltsabhängige Weiterleitung von Dispatcher (Weiterleitungsmethode cbr) ist *kein* Caching Proxy erforderlich.

- Wenn Sie mit den Protokollen HTTP und HTTPS (SSL) arbeiten, können Sie die Komponente **Content Based Routing (CBR)** für einen Lastausgleich ausgehend vom Inhalt der Client-Anfrage verwenden. Ein Client sendet eine Anfrage an Caching Proxy, und Caching Proxy sendet diese Anfrage an einen geeigneten Server. Der ausgewählte Server ist das Ergebnis des Abgleichs des URL mit einer angegebenen Regel.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode cbr der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

- Mit der Komponente **Site Selector** können Sie einen Lastausgleich für Server in einem lokalen oder Weitverkehrsnetz durchführen. Sie können dazu nach einer DNS-RoundRobin-Methode oder nach einer komplexeren benutzerdefinierten Methode vorgehen. Site Selector ordnet zusammen mit einem Namensserver IP-Adressen DNS-Namen zu.
- Mit der Komponente **Cisco CSS Controller** oder **Nortel Alteon Controller** können Sie Serverwertigkeiten generieren, die dann zur Erreichung einer optimalen Serverauswahl sowie von Lastoptimierung und Fehlertoleranz an den Cisco CSS Switch bzw. den Nortel Alteon Web Switch gesendet werden.

Weitere Informationen zu den Komponenten Dispatcher, CBR, Site Selector, Cisco CSS Controller und Nortel Alteon Controller finden Sie im Abschnitt „Komponenten von Load Balancer“ auf Seite 9.

Welche Vorteile bringt die Verwendung von Load Balancer?

Die Anzahl von Benutzern und Netzen, die mit dem globalen Internet verbunden sind, wächst mit rasanter Geschwindigkeit. Dieses Wachstum verursacht Probleme hinsichtlich der Skalierbarkeit, da der Benutzerzugriff auf attraktive Sites bei einem hohen Anforderungsaufkommen möglicherweise eingeschränkt wird.

Derzeit benutzen Netzadministratoren verschiedene Methoden zur Optimierung des Zugriffs. Bei einigen dieser Methoden können Sie nach dem Zufallsprinzip einen anderen Server auswählen, wenn der vorher ausgewählte Server zu langsam oder überhaupt nicht antwortet. Diese Vorgehensweise ist jedoch mühsam und ineffektiv. Eine weitere Methode ist die Standard-RoundRobin-Methode, bei der der Domänennamensserver der Reihe nach Server zur Bearbeitung von Anforderungen auswählt. Dieser Ansatz ist zwar besser, aber immer noch ineffizient, da der Datenverkehr ohne Berücksichtigung der Serverauslastung weitergeleitet wird. Zudem werden bei dieser Methode auch dann noch Anforderungen an einen Server gesendet, wenn er ausgefallen ist.

Der Bedarf an einer leistungsfähigeren Lösung hat zur Entwicklung von Load Balancer geführt. Dieses Produkt bietet gegenüber früheren Lösungen und Lösungen anderer Anbieter eine Vielzahl von Vorteilen:

Skalierbarkeit

Wenn die Anzahl der Client-Anforderungen steigt, können Sie Server dynamisch hinzufügen und Millionen von Anforderungen pro Tag auf Hunderten von Servern unterstützen.

Effektive Nutzung der Ausrüstung

Der Lastausgleich gewährleistet, dass jede Servergruppe die zugehörige Hardware optimal nutzen kann, da die bei einer Standard-RoundRobin-Methode häufig auftretenden Spitzenbelastungen auf ein Minimum reduziert werden.

Problemlose Integration

Load Balancer benutzt TCP/IP-Standardprotokolle oder UDP/IP-Protokolle. Das Produkt kann zu einem vorhandenen Netz hinzugefügt werden, ohne dass physische Änderungen am Netz erforderlich sind. Es ist leicht zu installieren und zu konfigurieren.

Geringer Systemaufwand

Bei Anwendung der einfachen MAC-Weiterleitung achtet der Dispatcher nur auf den beim Server eingehenden Datenverkehr vom Client, nicht aber auf den vom Server zum Client abgehenden Datenverkehr. Dies führt im Vergleich zu anderen Methoden zu einer erheblichen Reduzierung der Auswirkungen auf die Anwendung und zu einer verbesserten Leistung im Netz.

Hohe Verfügbarkeit

Der Dispatcher, der Cisco CSS Controller und der Nortel Alteon Controller sind Komponenten mit integrierter hoher Verfügbarkeit, für die eine Partnermaschine benutzt wird, die jederzeit die Weiterleitung von Paketen übernehmen kann, wenn die primäre Servermaschine ausfällt. Sollte einer der Server ausfallen, werden die Anfragen vom anderen Server bedient. Dies bedeutet, dass keiner der Server als Single Point of Failure eingesetzt werden muss. Die Site zeichnet sich somit durch hohe Verfügbarkeit aus.

Weitere Informationen hierzu finden Sie im Abschnitt „Bereitstellung der hohen Verfügbarkeit durch Load Balancer“ auf Seite 6.

Content Based Routing (mit der Komponente CBR oder Dispatcher)

Zusammen mit Caching Proxy kann die Komponente CBR HTTP- und HTTPS-Anforderungen (SSL) ausgehend vom angefragten Inhalt an bestimmte Server weiterleiten. Wenn eine Anforderung im Verzeichnisabschnitt des URL beispielsweise die Zeichenfolge `"/cgi-bin/"` enthält und der Servername ein lokaler Server ist, kann CBR die Anforderung an den besten Server einer speziell für die Bearbeitung von cgi-Anforderungen zugeordneten Servergruppe übertragen.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode `cbr` der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Die Komponente Dispatcher erlaubt auch eine inhaltsabhängige Weiterleitung, erfordert jedoch nicht die Installation von Caching Proxy. Da die inhaltsabhängige Weiterleitung der Komponente Dispatcher beim Empfang von Paketen im Kernel ausgeführt wird, ist sie *schneller* als die der Komponente CBR. Die Komponente Dispatcher führt das Content Based Routing für HTTP (unter Verwendung des Regeltyps content) und für HTTPS (unter Verwendung der Affinität von SSL-Sitzungs-IDs) durch.

Anmerkung: Für HTTPS (SSL) kann die Komponente CBR den Regeltyp content nur für den Lastausgleich von Datenverkehr verwenden, der auf dem Inhalt der HTTP-Anforderung basiert, was das Entschlüsseln und erneute Verschlüsseln von Nachrichten erfordert.

Bereitstellung der hohen Verfügbarkeit durch Load Balancer

Dispatcher

Die Komponente Dispatcher stellt eine integrierte Funktion für hohe Verfügbarkeit bereit, so dass der Dispatcher nicht mehr als Single Point of Failure in Ihrem Netzwerk eingesetzt werden muss. Für diese Funktion ist eine zweite Dispatcher-Maschine erforderlich, die die primäre Maschine überwacht und den Lastausgleich übernehmen kann, wenn die primäre Maschine ausfällt. Die Komponente Dispatcher gewährleistet außerdem eine gegenseitige hohe Verfügbarkeit, so dass sowohl die primäre als auch die sekundäre Maschine die jeweils andere Maschine als Ausweichmaschine nutzen kann. Lesen Sie hierzu die Informationen im Abschnitt „Hohe Verfügbarkeit konfigurieren“ auf Seite 205.

CBR

Wenn Sie eine Client/Server-Konfiguration mit einer Dispatcher-Maschine verwenden, bei der der Datenverkehr auf mehrere Server mit CBR verteilt wird, können Sie auch mit der Komponente CBR ein hohes Maß an Verfügbarkeit erreichen.

Cisco CSS Controller oder Nortel Alteon Controller

Die Controller stellen eine Funktion für hohe Verfügbarkeit bereit, so dass der Controller nicht mehr als Single Point of Failure eingesetzt werden muss. Ein Controller auf einer Maschine kann als primärer Controller und ein Controller auf einer anderen Maschine als Ausweichcontroller konfiguriert werden. Der Ausweichcontroller überwacht den primären Controller und ist bereit, bei einem Ausfall des primären Controllers dessen Aufgaben zu übernehmen und den Switches Serverwertigkeiten zur Verfügung zu stellen. Weitere Informationen finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 245.

Neue Features

Load Balancer für IBM WebSphere Application Server Version 6.1 ist durch eine Reihe neuer Features gekennzeichnet. Die wichtigsten neuen Features sind nachfolgend aufgelistet.

- **Unterstützung für die Ausführung von Lastausgleichsprozessen im Benutzerbereich auf Linux-Systemen**

Installationen von Load Balancer für IPV4 und IPV6 unterstützen jetzt die Ausführung von Lastausgleichsprozessen im Benutzerbereich an Stelle der Ausführung im Kernel-Bereich. Auf Linux-Systemen besteht keine Abhängigkeit mehr vom Kernel-Modul.

Die aktuellsten Informationen zur Unterstützung der (Kernel-unabhängigen) Verarbeitung im Benutzerbereich finden Sie auf der folgenden Website:

<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

Weitere Informationen hierzu finden Sie im Abschnitt „Unterstützte Plattformen für Load Balancer für IPV4 und IPV6“ auf Seite 82.

- **Unterstützung für HP 11iv2 auf PA-RISC (Unterstützung für HP 11iv1 fällt weg)**

Informationen zu den Hardware- und Softwarevoraussetzungen finden Sie auf der folgenden Website: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

- **Unterstützung für Linux auf zSeries (64-Bit-Systeme)**

Nur Installationen von Load Balancer für IPV4 und IPV6 bieten Unterstützung für Linux auf zSeries (64-Bit-Systeme).

Informationen zu Load Balancer für IPV4 und IPV6 und spezielle Hinweise zur Ausführung von Linux auf zSeries (64-Bit-Systeme) enthält Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

Informationen zu den Hardware- und Softwarevoraussetzungen finden Sie auf der folgenden Website: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

- **Unterstützung für den SIP-Advisor**

Die Unterstützung für einen SIP-Advisor (Session Initiation Protocol) ist neu hinzugekommen. Der unterstützte SIP-Advisor funktioniert nur mit dem Protokoll TCP.

Weitere Informationen hierzu finden Sie auf Seite 189.

- **Unterstützung verknüpfter Client-Konfigurationen auf Linux-Systemen**

Dieses Feature ist für alle Komponenten von Load Balancer verfügbar.

Clients, die sich auf derselben Maschine wie Load Balancer befinden, werden nur auf Linux-Systemen unterstützt.

Weitere Informationen hierzu finden Sie im Abschnitt „Verknüpften Client verwenden“ auf Seite 243.

- **Unterstützung für den Browser Firefox**

Informationen zu den unterstützten Versionen von Firefox und zu allen anderen unterstützten Browsern finden Sie auf der folgenden Website:

<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

Kapitel 2. Komponenten von Load Balancer im Überblick

Dieses Kapitel gibt einen Überblick über die Komponenten von Load Balancer und ist in die folgenden Abschnitte gegliedert:

- „Komponenten von Load Balancer“
- „Dispatcher im Überblick“
- „Content Based Routing (CBR) im Überblick“ auf Seite 12
- „Site Selector im Überblick“ auf Seite 13
- „Cisco CSS Controller im Überblick“ auf Seite 15
- „Nortel Alteon Controller im Überblick“ auf Seite 16

Eine detaillierte Liste der Konfigurationsfunktionen der einzelnen Komponenten von Load Balancer, die Sie zur Planung Ihrer Netzverwaltung heranziehen können, finden Sie in Kapitel 3, „Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer“, auf Seite 19.

Komponenten von Load Balancer

Die fünf Komponenten von Load Balancer sind Dispatcher, Content Based Routing (CBR), Site Selector, Cisco CSS Controller und Nortel Alteon Controller. Load Balancer bietet Ihnen die Möglichkeit, die Komponenten flexibel entsprechend der Konfiguration Ihrer Site einzeln oder zusammen zu verwenden. Dieses Kapitel gibt einen Überblick über die Komponenten.

Wichtiger Hinweis: Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, steht nur die Dispatcher-Komponente zur Verfügung. Weitere Informationen finden Sie in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

Dispatcher im Überblick

Die Komponente Dispatcher verteilt den Datenverkehr mit einer Kombination von Lastausgleichs- und Verwaltungssoftware auf Ihre Server. Der Dispatcher kann auch einen ausgefallenen Server erkennen und den Datenverkehr entsprechend umleiten. Dispatcher unterstützt HTTP, FTP, SSL, SMTP, NNTP, IMAP, POP3, Telnet, SIP sowie alle anderen TCP-basierten bzw. kontextlosen UDP-basierten Anwendungen.

Alle an die Dispatcher-Maschine gesendeten Client-Anforderungen werden auf der Basis dynamisch festgelegter Wertigkeiten an den "besten" Server übertragen. Für diese Wertigkeiten können Sie Standardwerte benutzen oder die Werte während des Konfigurationsprozesses ändern.

Dispatcher kann drei Weiterleitungsmethoden anwenden (die für den Port angegeben werden):

- Weiterleitungsmethode **mac**. Bei dieser Methode der Weiterleitung führt der Dispatcher einen Lastausgleich für die beim Server eingehenden Anforderungen durch. Der Server gibt die Antwort direkt, d. h. ohne Eingreifen des Dispatchers, an den Client zurück.
- NAT/NAPT-Weiterleitungsmethode (**nat**). Bei Verwendung der Dispatcher-Methode NAT (Konvertierung von Netzadressen) bzw. NAPT (Port-Umsetzung für

Netzadressen) entfällt die Einschränkung, dass sich die Back-End-Server in einem lokal angeschlossenen Netz befinden müssen. Falls Sie Server an fernen Standorten haben, sollten Sie an Stelle einer GRE/WAN-Kapselungstechnik die NAT-Technik anwenden. Bei der Weiterleitungsmethode nat verteilt der Dispatcher die eingehenden Anforderungen auf den Server. Der Server gibt die Antwort an den Dispatcher zurück. Die Dispatcher-Maschine gibt die Antwort dann an den Client zurück.

- Inhaltsabhängige Weiterleitungsmethode (**cbr**). Ohne Caching Proxy können Sie mit der Komponente Dispatcher ein Content Based Routing für HTTP (unter Verwendung des Regeltyps content) und für HTTPS (unter Verwendung der Affinität von SSL-Sitzungs-IDs) durchführen. Für HTTP- und HTTPS-Datenverkehr ist das Content Based Routing der Komponente Dispatcher *schneller* als das der Komponente CBR. Bei der Weiterleitungsmethode cbr verteilt der Dispatcher die eingehenden Anforderungen auf den Server. Der Server gibt die Antwort an den Dispatcher zurück. Die Dispatcher-Maschine gibt die Antwort dann an den Client zurück.

Die Komponente Dispatcher ist der Schlüssel für eine stabile, effiziente Verwaltung eines großen skalierbaren Servernetzes. Mit dem Dispatcher können Sie viele einzelne Server so verbinden, dass sie ein virtueller Server zu sein scheinen. Besucher sehen Ihre Site daher unter einer IP-Adresse. Der Dispatcher arbeitet unabhängig von einem Domännennamensserver. Alle Anforderungen werden an die IP-Adresse der Dispatcher-Maschine gesendet.

Der Dispatcher bringt klare Vorteile bei der Lastverteilung auf eine Gruppe von Servern und ermöglicht daher eine stabile und effiziente Verwaltung Ihrer Site.

Lokale Server mit dem Dispatcher verwalten

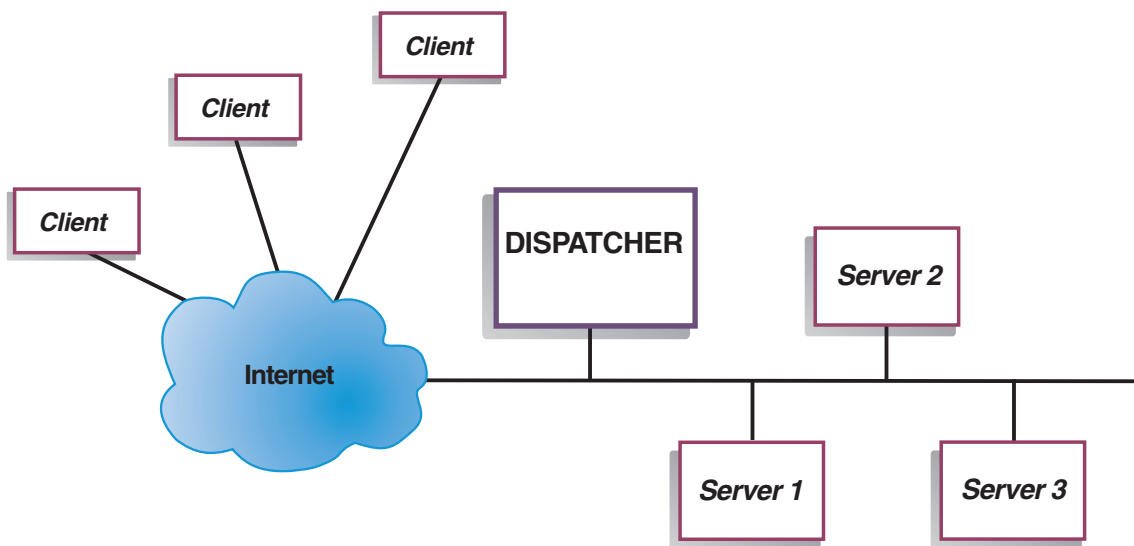


Abbildung 1. Beispiel für die physische Darstellung einer Site mit Dispatcher für die Verwaltung lokaler Server

In Abb. 1 wird die physische Darstellung der Site mit einer Ethernet-Netzkonfiguration gezeigt. Die Dispatcher-Maschine kann installiert werden, ohne dass physische Änderungen am Netz erforderlich sind. Nachdem der Dispatcher eine Client-Anforderung an den optimalen Server übertragen hat, wird die Antwort mit der Weiterleitungsmethode mac ohne Eingriff des Dispatchers direkt vom Server an den Client gesendet.

Serververwaltung mit Dispatcher und Metric Server

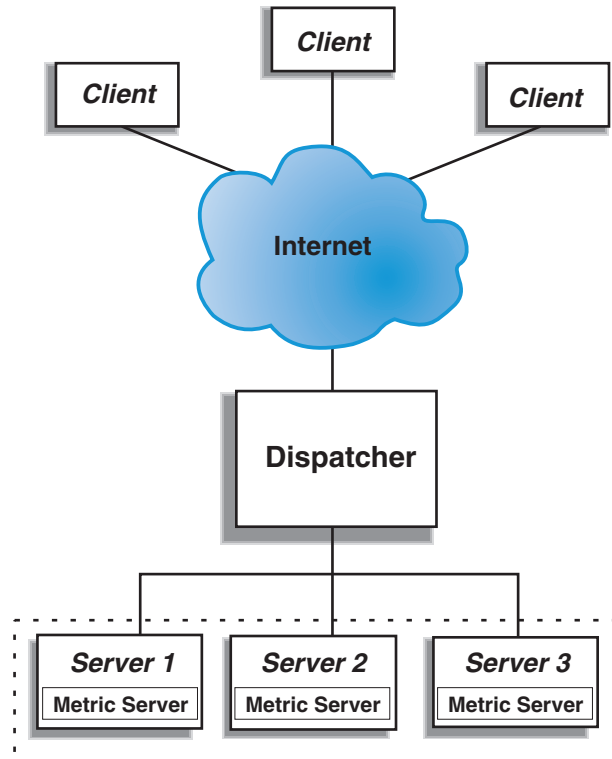


Abbildung 2. Beispielsite mit Dispatcher und Metric Server für die Serververwaltung

In Abb. 2 wird eine Site gezeigt, bei der sich alle Server in einem lokalen Netz befinden. Die Komponente Dispatcher leitet Anforderungen weiter und Metric Server stellt der Dispatcher-Maschine Informationen zur Systembelastung zur Verfügung.

In diesem Beispiel ist der Metric-Server-Dämon auf allen Back-End-Servern installiert. Sie können Metric Server zusammen mit der Komponente Dispatcher oder einer beliebigen anderen Komponente von Load Balancer verwenden.

Lokale und ferne Server mit Dispatcher verwalten

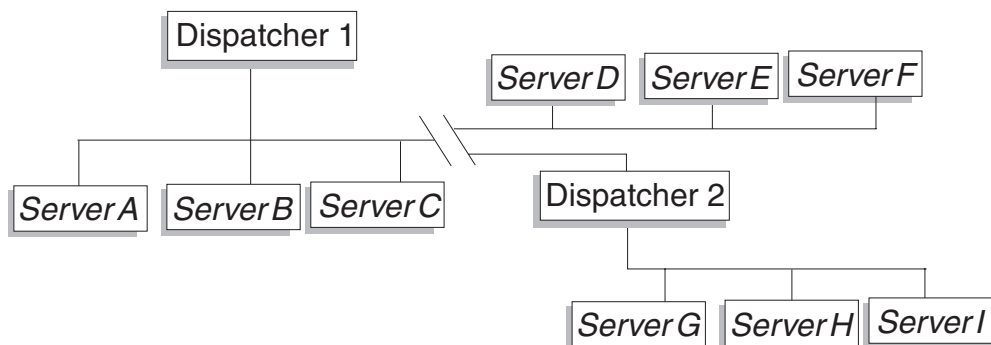


Abbildung 3. Beispiel für eine Site mit Dispatcher für die Verwaltung lokaler und ferner Server

Die Weitverkehrsunterstützung von Dispatcher ermöglicht Ihnen die Verwendung lokaler und ferner Server (d. h. Server in anderen Teilnetzen). Abb. 3 auf Seite 11 zeigt eine Konfiguration, bei der ein lokaler Dispatcher (Dispatcher 1) als Eingangspunkt für alle Anforderungen dient. Er verteilt diese Anforderungen auf seine lokalen Server (ServerA, ServerB, ServerC) und den fernen Dispatcher (Dispatcher 2), der die Last auf seine lokalen Server (ServerG, ServerH, ServerI) verteilt.

Wenn Sie die Weiterleitungsmethode nat oder die GRE-Unterstützung von Dispatcher nutzen, können Sie eine Weitverkehrsunterstützung auch ohne Verwendung eines Dispatchers am fernen Standort (wo sich ServerD, ServerE und ServerF befinden) erreichen. Weitere Informationen hierzu finden Sie in den Abschnitten „Dispatcher-Weiterleitungsmethode nat“ auf Seite 53 und „Unterstützung für GRE (Generic Routing Encapsulation)“ auf Seite 235.

Content Based Routing (CBR) im Überblick

CBR arbeitet mit Caching Proxy zusammen, um Client-Anforderungen an angegebene HTTP- oder HTTPS-Server (SSL) weiterzuleiten. Diese Komponente ermöglicht die Bearbeitung von Caching-Angaben für ein schnelleres Abrufen von Webdokumenten mit geringen Anforderungen an die Netzbandbreite. CBR überprüft zusammen mit Caching Proxy HTTP-Anforderungen anhand angegebener Regeltypen.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode cbr der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Bei Verwendung von CBR können Sie eine Gruppe von Servern angeben, die eine Anforderung ausgehend von der Übereinstimmung eines regulären Ausdrucks mit dem Inhalt der Anforderung bearbeiten. Da CBR die Angabe mehrerer Server für jede Art von Anforderung zulässt, können die Anforderungen so verteilt werden, dass eine optimale Client-Antwortzeit erreicht wird. CBR erkennt auch, wenn ein Server in einer Gruppe ausgefallen ist. In diesem Fall werden keine weiteren Anforderungen an diesen Server weitergeleitet. Der von der Komponente CBR verwendete Lastausgleichsalgorithmus ist mit dem bewährten Algorithmus identisch, der von der Komponente Dispatcher verwendet wird.

Wenn Caching Proxy eine Anfrage empfängt, wird diese mit den Regeln, die für die Komponente CBR definiert wurden, abgeglichen. Wird eine Übereinstimmung gefunden, wird einer der Server, die dieser Regel zugeordnet sind, für die Bearbeitung der Anforderung ausgewählt. Caching Proxy führt dann die normale Verarbeitung aus, um die Anfrage an den ausgewählten Server weiterzuleiten.

CBR stellt mit Ausnahme der hohen Verfügbarkeit, des SNMP-Subagenten, der Weitverkehrsunterstützung und einiger anderer Konfigurationsbefehle dieselben Funktionen wie der Dispatcher bereit.

CBR kann erst mit dem Lastausgleich für Client-Anfragen beginnen, wenn Caching Proxy aktiv ist.

Lokale Server mit CBR verwalten

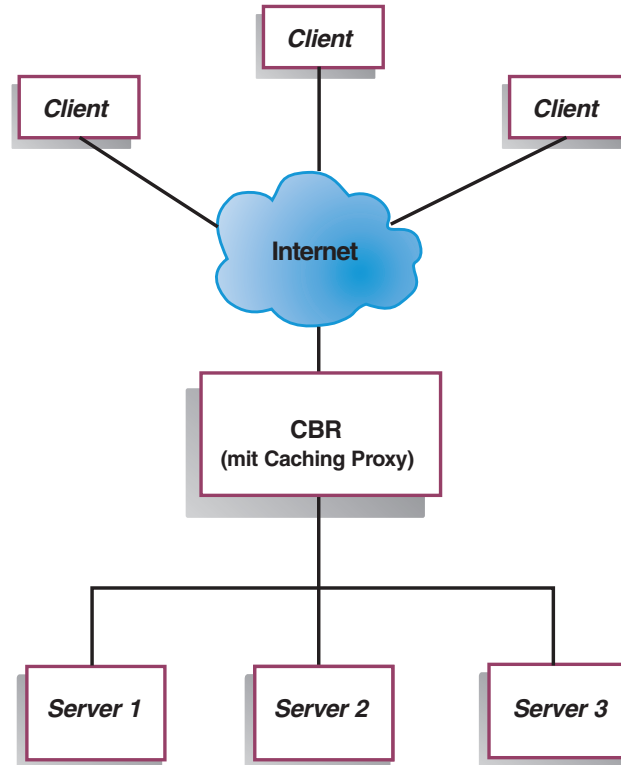


Abbildung 4. Beispielsite mit CBR für die Verwaltung lokaler Server

Abb. 4 zeigt die logische Darstellung einer Site, bei der ein Teil der Inhalte von lokalen Servern mit CBR weitergeleitet wird. Die Komponente CBR leitet mit Caching Proxy Client-Anfragen (HTTP oder HTTPS) ausgehend vom Inhalt des URL an die Server weiter.

Site Selector im Überblick

Site Selector fungiert als Namensserver und führt zusammen mit anderen Namensservern in einem Domännennamenssystem auf der Grundlage abgerufener Messungen und Wertigkeiten einen Lastausgleich für Servergruppen durch. Sie können eine Sitekonfiguration erstellen, bei der die Last innerhalb einer Servergruppe auf der Grundlage des für eine Client-Anfrage verwendeten Domännennamens verteilt wird.

Ein Client fordert die Auflösung eines Domännennamens bei einem Namensserver innerhalb seines Netzes an. Der Namensserver leitet die Anforderung an die Site-Selector-Maschine weiter. Site Selector löst den Domännennamen dann in die IP-Adresse eines der Server auf, die für den Sitenamen konfiguriert wurden. Anschließend gibt Site Selector die IP-Adresse des ausgewählten Servers an den Namensserver zurück. Der Namensserver liefert die IP-Adresse an den Client.

Metric Server ist eine Systemüberwachungskomponente von Load Balancer, die auf jedem am Lastausgleich beteiligten Server innerhalb der Konfiguration installiert sein muss. Mit Metric Server kann Site Selector das Aktivitätsniveau eines Servers überwachen, den Server mit der geringsten Auslastung ermitteln und einen ausgefallenen Server erkennen. Die Last ist ein Maß für das Arbeitsaufkommen eines

Servers. Durch Anpassung der Script-Dateien für Systemmesswerte können Sie steuern, auf welche Art die Last gemessen wird. Sie können Site Selector an die Anforderungen der eigenen Umgebung anpassen und dabei Faktoren wie die Zugriffshäufigkeit, die Gesamtzahl der Benutzer und die Zugriffsarten (beispielsweise kurze Abfragen, lange Abfragen, Transaktionen mit hoher CPU-Belastung) berücksichtigen.

Lokale und ferne Server mit Site Selector und Metric Server verwalten

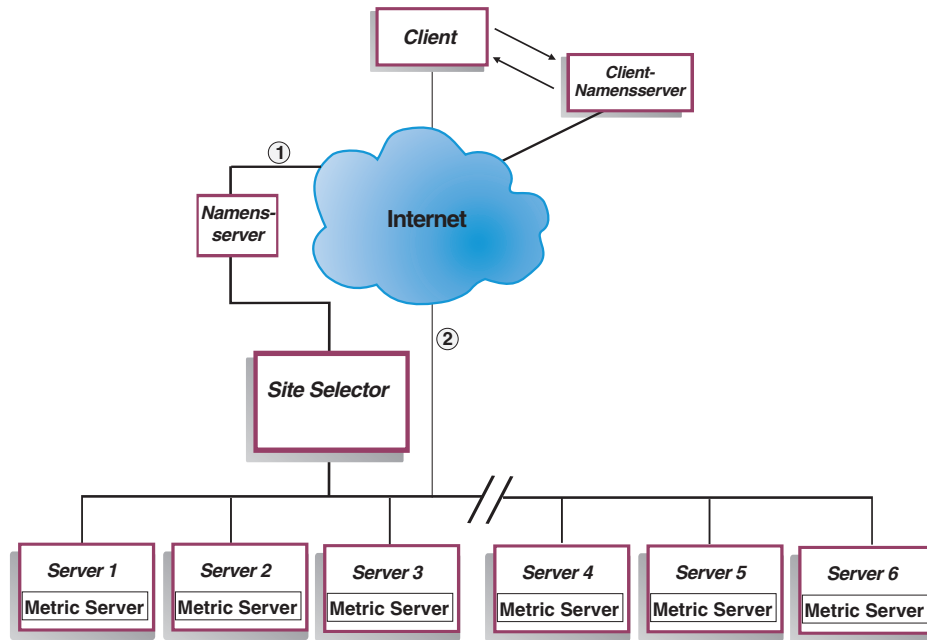


Abbildung 5. Beispielsite mit Site Selector und Metric Server für die Verwaltung lokaler und ferner Server

Abb. 5 stellt eine Site dar, bei der die Komponente Site Selector Anfragen beantwortet. Server 1, Server 2 und Server 3 sind lokale Server. Server 4, Server 5 und Server 6 sind ferne Server.

Ein Client fordert die Auflösung eines Domännennamens bei einem Client-Namensserver an. Der Client-Namensserver leitet die Anfrage über den DNS an die Site-Selector-Maschine weiter (Pfad 1). Site Selector löst den Domännennamen dann in die IP-Adresse eines der Server auf. Anschließend gibt Site Selector die IP-Adresse des ausgewählten Servers an den Client-Namensserver zurück. Der Namensserver liefert die IP-Adresse an den Client.

Sobald der Client die IP-Adresse des Servers empfangen hat, leitet er Anwendungsanforderungen direkt an den ausgewählten Server weiter (Pfad 2).

Anmerkung: In diesem Beispiel liefert Metric Server Informationen zur Systembelastung an die Site-Selector-Maschine. Der Agent Metric Server ist auf jedem Back-End-Server installiert. Verwenden Sie Site Selector zusammen mit Metric Server, da Site Selector sonst nur eine Round-Robin-Auswahlmethode für den Lastausgleich anwenden kann.

Cisco CSS Controller im Überblick

Cisco CSS Controller ist eine ergänzende Lösung für die CSS 11000 Series Switches von Cisco. Die kombinierte Lösung verbindet die zuverlässige Paket- und Inhaltsweiterleitung der CSS 11000 Series mit den ausgeklügelten Erkennungsalgorithmen von Load Balancer, um die Ladedaten und die Verfügbarkeit des *Services* (Anwendung oder Datenbank auf dem Back-End-Server) festzustellen. Cisco CSS Controller verwendet den Algorithmus für Wertigkeitsberechnung, die standardmäßigen und angepassten Advisor-Funktionen sowie den Metric Server von Load Balancer, um die Messwerte, den Zustand und die Auslastung des *Services* zu ermitteln. Aus diesen Informationen generiert der Cisco CSS Controller Servicewertigkeiten, die dann zur Erreichung einer optimalen Serviceauswahl sowie von Lastoptimierung und Fehlertoleranz an den Cisco CSS Switch gesendet werden.

Der Cisco CSS Controller protokolliert zahlreiche Kriterien. Dazu gehören unter anderem:

- aktive Verbindungen und Verbindungsrate (Anzahl neuer Verbindungen innerhalb eines Wertigkeitsberechnungszyklus)
- Verfügbarkeit von Anwendungen und Datenbanken (was durch standardmäßige und angepasste Advisor-Funktionen sowie durch serviceresidente und für bestimmte Anwendungen maßgeschneiderte Agenten erleichtert wird)
- CPU-Auslastung
- Speicherauslastung
- vom Benutzer anpassbare Systemmesswerte.

Wenn ein Cisco CSS Switch ohne Cisco CSS Controller den Zustand eines Inhalte bereitstellenden *Services* ermittelt, greift er dabei auf die Antwortzeiten für Inhaltsanfragen und andere Netzmesswerte zurück. Wird der Cisco CSS Controller verwendet, gehen diese Aktivitäten vom Cisco CSS Switch auf den Cisco CSS Controller über. Der Cisco CSS Controller beeinflusst die Fähigkeit des *Services*, Inhalte bereitzustellen, und aktiviert einen Service als geeigneten Service, wenn dieser verfügbar ist, bzw. stellt ihn als geeigneten Service zurück, wenn er nicht mehr verfügbar ist.

Cisco CSS Controller:

- empfängt über eine veröffentlichte SNMP-Schnittstelle Verbindungsdaten vom Cisco CSS Switch
- analysiert anhand der Vorgaben von Advisor-Funktionen die Verfügbarkeit und Reaktion von *Services*
- analysiert die Systembelastung anhand der Informationen von Metric Server
- generiert Wertigkeiten für jeden Service der Konfiguration.

Wertigkeiten gelten für alle *Services* an einem Port. An einem bestimmten Port werden die Anfragen ausgehend von einem Vergleich der Wertigkeiten der einzelnen *Services* verteilt. Wenn ein Service beispielsweise die Wertigkeit 10 und ein anderer die Wertigkeit 5 hat, erhält der Service mit der Wertigkeit 10 doppelt so viele Anfragen wie der Service mit der Wertigkeit 5. Diese Wertigkeiten werden dem Cisco CSS Switch mit SNMP zur Verfügung gestellt. Wird ein Service mit einer höheren Wertigkeit eingestuft, überträgt der Cisco CSS Switch mehr Anfragen an diesen Service.

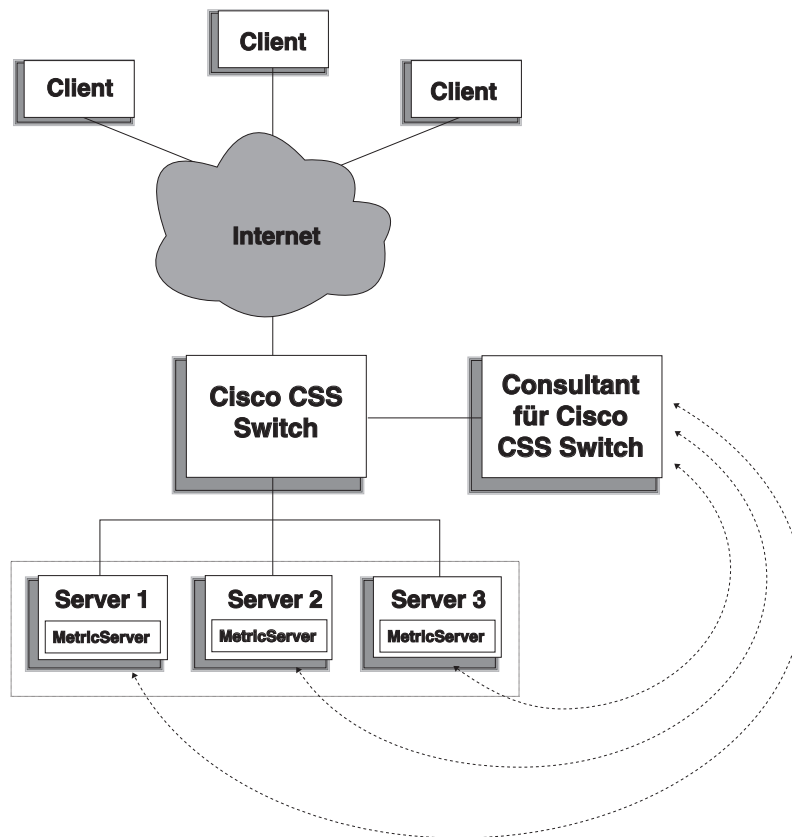


Abbildung 6. Beispielsite mit Cisco CSS Controller und Metric Server für die Verwaltung lokaler Services

Cisco CSS Controller ist in Verbindung mit dem Cisco CSS Switch eine Lösung, die das "Beste aus zwei Welten" auf sich vereint, super schnelles Content-Switching und ausgeklügelte Anwendungserkennung, Fehlertoleranz sowie optimale Serviceauslastung. Cisco CSS Controller ist Bestandteil einer ergänzenden Lösung für den Cisco CSS Switch und IBM WebSphere Application Server Load Balancer.

Nortel Alteon Controller im Überblick

Nortel Alteon Controller ist in Verbindung mit der Alteon-Web-Switch-Familie von Nortel eine ergänzende Lösung, die die Geschwindigkeit und Kapazität der Switches im Bereich der Paketweiterleitung mit den ausgeklügelten Erkennungsalgorithmen von Load Balancer zur Bestimmung von Serverwertigkeiten verbindet.

Mit Nortel Alteon Controller können Sie angepasste Advisor-Funktionen entwickeln, die eine intelligentere und mehr auf die Anwendung zugeschnittene Bewertung der Verfügbarkeit und Auslastung von Anwendungen für die Implementierung von Services vornehmen können.

Metric Server stellt Informationen zur Systembelastung, z. B. zur Auslastung von CPU und Speicher, sowie ein Gerüst für die Entwicklung benutzerdefinierter Messungen zur Systembelastung bereit.

Nortel Alteon Controller stellt die verschiedensten Arten von Messdaten zusammen, um die Wertigkeit von Servern zu bestimmen, deren Arbeitslast durch Nortel Alteon Web Switches verteilt wird. Dazu gehören unter anderem:

- aktive und neue Verbindungen
- Verfügbarkeit von Anwendungen und Datenbanken (was durch standardmäßige und angepasste Advisor-Funktionen sowie durch serverresidente und für bestimmte Anwendungen maßgeschneiderte Agenten erleichtert wird)
- CPU-Auslastung
- Speicherauslastung
- vom Benutzer anpassbare Servermesswerte
- Erreichbarkeit.

Der Nortel Alteon Controller kommuniziert mit dem Switch über SNMP. Die Komponente ruft Konfigurations-, Status- und Verbindungsdaten vom Switch ab. Wenn der Controller Serverwertigkeiten berechnet hat, werden diese auf dem Switch festgelegt. Der Switch verwendet die vom Controller definierten Wertigkeiten, um den Server auszuwählen, der Client-Anfragen nach einem Service am besten bearbeiten kann.

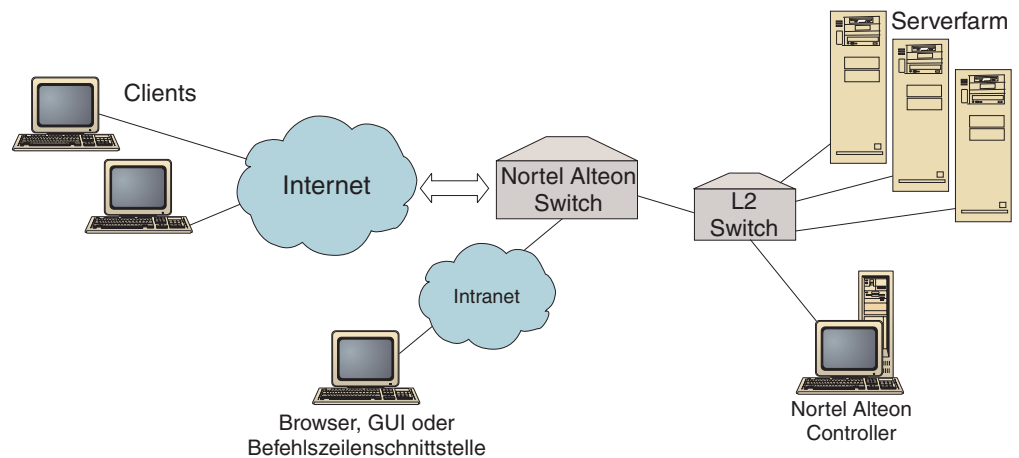


Abbildung 7. Beispielsite mit Nortel Alteon Controller für die Verwaltung lokaler Server

Sie können den Controller von einem Browser, einer fernen GUI oder einer fernen Befehlszeilenschnittstelle aus verwalten.

Nortel Alteon Controller ist in Verbindung mit der Alteon-Web-Switch-Familie von Nortel eine Lösung, die das "Beste aus zwei Welten" auf sich vereint: super schnelle Paketvermittlung und ausgeklügelte Anwendungserkennung, Fehler-toleranz sowie optimale Serverauslastung. Nortel Alteon Controller ist Teil einer ergänzenden Lösung mit den Alteon-Web-Switches von Nortel und IBM WebSphere.

Kapitel 3. Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer

Die folgenden Abschnitte listen die Konfigurationsfunktionen der Komponenten von Load Balancer auf, so dass Sie bestimmen können, welche Features Sie für die Verwaltung Ihres Netzes benötigen:

- „Manager, Advisor-Funktionen und Metric Server (für Dispatcher, CBR und Site Selector)“
- „Features der Komponente Dispatcher“
- „Features der Komponente CBR (Content Based Routing)“ auf Seite 23
- „Features der Komponente Site Selector“ auf Seite 26
- „Features der Komponente Cisco CSS Controller“ auf Seite 27
- „Features der Komponente Nortel Alteon Controller“ auf Seite 28

Wichtiger Hinweis: Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, steht nur die Dispatcher-Komponente zur Verfügung. Weitere Informationen finden Sie in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

Manager, Advisor-Funktionen und Metric Server (für Dispatcher, CBR und Site Selector)

Wenn Sie die Last optimal auf mehrere Server verteilen und sicherstellen möchten, dass stets der "richtige" Server ausgewählt wird, lesen Sie die folgenden Abschnitte:

- „Lastausgleich mit Load Balancer optimieren“ auf Seite 180
- „Advisor-Funktionen“ auf Seite 185
- „Metric Server“ auf Seite 196

Features der Komponente Dispatcher

Dispatcher unterstützt den Lastausgleich auf Servern für HTTP, FTP, SSL, SMTP, NNTP, IMAP, POP3, Telnet, SIP sowie alle anderen TCP-basierten bzw. kontextlosen UDP-basierten Anwendungen.

Fernverwaltung

- Wenn Sie Load Balancer nicht von der Maschine aus konfigurieren möchten, auf der Load Balancer installiert ist, lesen Sie den Abschnitt „Fernverwaltung von Load Balancer“ auf Seite 263.
(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, ist dieses Feature nicht verfügbar.)

Verknüpfung

- Falls Sie Dispatcher auf derselben Maschine ausführen möchten wie einen Webserver, dessen Last verteilt werden soll, lesen Sie die Informationen im Abschnitt „Verknüpfte Server verwenden“ auf Seite 202.

Hohe Verfügbarkeit

- Falls Sie den Dispatcher einsetzen, um die Einschränkungen eines Single Point of Failure für Ihr Netz zu beseitigen, lesen Sie die Informationen im Abschnitt „Einfache hohe Verfügbarkeit“ auf Seite 60 und im Abschnitt „Gegenseitige hohe Verfügbarkeit“ auf Seite 61.
(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, können Sie die einfache hohe Verfügbarkeit, jedoch nicht die gegenseitige hohe Verfügbarkeit nutzen.)

Client-Server-Affinität

Lastausgleich für SSL-Datenverkehr (HTTPS):

- Wenn Sie sicherstellen möchten, dass der Client für mehrere Verbindungen denselben SSL-Server verwendet, lesen Sie den Abschnitt „Funktionsweise der Affinität für Load Balancer“ auf Seite 222.
- Wenn Sie sicherstellen möchten, dass der Client für HTTP- und SSL-Datenverkehr denselben SSL-Server verwendet, lesen Sie den Abschnitt „Port-übergreifende Affinität“ auf Seite 223.
(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, ist die Port-übergreifende Affinität nicht verfügbar.)
- Wenn Sie sicherstellen möchten, dass der Client für mehrere Verbindungen denselben Server verwendet, lesen Sie den Abschnitt „Funktionsweise der Affinität für Load Balancer“ auf Seite 222.
- Wenn Sie sicherstellen möchten, dass eine Gruppe von Clients für mehrere Verbindungen denselben Server verwendet, lesen Sie den Abschnitt „Affinitätsadressmaske (stickymask)“ auf Seite 223.
(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, ist das Feature stickymask nicht verfügbar.)
- Wenn Sie einen Server aus der Konfiguration entfernen möchten (z. B. für eine Wartung), ohne den Client-Datenverkehr zu unterbrechen, lesen Sie den Abschnitt „Bearbeitung von Serververbindungen stilllegen“ auf Seite 224.

Regelbasierter Lastausgleich

Sie können Clients für dieselbe Webadresse zu verschiedenen Servergruppen dirigieren, indem Sie Regeln zu Ihrer Dispatcher-Konfiguration hinzufügen. Weitere Informationen hierzu finden Sie im Abschnitt „Regelbasierten Lastausgleich konfigurieren“ auf Seite 212.

- Falls Sie Clients ausgehend von der Quellen-IP-Adresse des Clients zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf der Client-IP-Adresse basierende Regeln verwenden“ auf Seite 214.
- Falls Sie Clients ausgehend vom Client-Port zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf dem Client-Port basierende Regeln verwenden“ auf Seite 214.
- Falls Sie Clients ausgehend von der Uhrzeit zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf der Uhrzeit basierende Regeln verwenden“ auf Seite 214.
- Wenn Sie Clients ausgehend von den TOS-Bits (Type of Service) in Netzpaketen zu bestimmten Servern dirigieren möchten, lesen Sie die Informationen im Abschnitt „Auf der Serviceart basierende Regeln verwenden“ auf Seite 215.
- Falls Sie Clients ausgehend vom Datenverkehr der Site zu verschiedenen Servergruppen dirigieren möchten, haben Sie die folgenden Möglichkeiten:

- Sie können die Anzahl der Verbindungen pro Sekunde verwenden. Lesen Sie dazu den Abschnitt „Regeln auf der Basis der Verbindungen pro Sekunde verwenden“ auf Seite 215.
- Sie können die Summe der aktiven Verbindungen verwenden. Lesen Sie dazu den Abschnitt „Regeln auf der Basis der Summe aktiver Verbindungen verwenden“ auf Seite 215.
- Sie können für verschiedene Webadressen Bandbreite gemeinsam benutzen und reservieren. Lesen Sie dazu den Abschnitt „Regeln auf der Basis der reservierten und gemeinsam benutzten Bandbreite verwenden“ auf Seite 216.
- Sie können sicherstellen, dass der Datenverkehr für jede Servergruppe exakt gemessen wird. Lesen Sie dazu den Abschnitt „Regeloption für Serverauswertung“ auf Seite 220.
- Falls Sie überlaufenden Datenverkehr zu einer Standardgruppe von Servern dirigieren möchten (z. B. zu Servern, die die Antwort "Site ausgelastet" ausgeben), lesen Sie den Abschnitt „Immer gültige Regeln verwenden“ auf Seite 219.
- Wenn Sie die Client-Affinität außer Kraft setzen möchten, um sicherzustellen, dass ein Client nicht an einen Überlaufserver gebunden bleibt, lesen Sie die Informationen im Abschnitt „Port-Affinität außer Kraft setzen“ auf Seite 220.

(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, ist der regelbasierte Lastausgleich nicht verfügbar.)

Inhaltsabhängiges Routing mit der Dispatcher-Weiterleitungsmethode cbr

Wenn Sie sicherstellen möchten, dass SSL-Clients ausgehend von der SSL-ID in der Client-Anfrage zu demselben SSL-Server zurückkehren, lesen Sie die Informationen auf Seite 56.

Falls Sie HTTP-Clients mit Regeln, die auf dem URL-Inhalt der Client-Anforderung basieren, zu verschiedenen Servergruppen dirigieren möchten, lesen Sie die Informationen im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55 und im Abschnitt „Auf dem Inhalt der Anforderung basierende Regeln verwenden“ auf Seite 219.

- Wenn Sie zwischen bestimmten URLs und ihren Serviceanwendungen unterscheiden möchten, lesen Sie den Abschnitt „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58.
- Wenn Sie durch Cookies von Ihren Webservern sicherstellen möchten, dass Clients zu demselben Server zurückkehren, wenn sie in mehreren Verbindungen ähnliche Inhalte anfragen, lesen Sie die Informationen im Abschnitt „Passive Cookie-Affinität“ auf Seite 227.
- Falls Sie den Webdatenverkehr auf Caching-Proxy-Server verteilen möchten, um das Zwischenspeichern eindeutiger Inhalte auf jedem einzelnen Server zu ermöglichen (und dadurch den Cache Ihrer Site vergrößern, da eine redundante Zwischenspeicherung von Inhalten auf mehreren Maschinen vermieden wird), lesen Sie den Abschnitt „URI-Affinität“ auf Seite 228.

(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, ist die Dispatcher-Weiterleitungsmethode cbr nicht verfügbar.)

Dispatcher-Weiterleitungsmethode cbr und Komponente CBR im Vergleich

Der Vorteil der Dispatcher-Weiterleitungsmethode cbr besteht darin, dass sie schneller auf Client-Anfragen reagiert als die Komponente CBR. Die Dispatcher-Weiterleitungsmethode cbr erfordert auch *nicht* die Installation und Verwendung von Caching Proxy.

Bei einem Netz mit sicherem SSL-Datenverkehr vom Client zum Server liegt der Vorteil der Komponente CBR (in Verbindung mit Caching Proxy) darin, dass CBR die für ein inhaltsabhängiges Routing erforderliche Verschlüsselung/Entschlüsselung durchführen kann. Bei vollständig gesicherten Verbindungen kann die Dispatcher-Weiterleitungsmethode cbr nur mit SSL-ID-Affinität konfiguriert werden, da sie nicht in der Lage ist, die für ein vom Inhalt des URL abhängiges Routing von Client-Anforderungen erforderliche Verschlüsselung/Entschlüsselung vorzunehmen.

Lastausgleich im WAN

Ein Lastausgleich im Weitverkehrsnetz kann mit verschiedenen Methoden erreicht werden.

- Falls Sie Last mit dem WAN-Feature von Dispatcher auf ferne Server verteilen möchten, lesen Sie die Informationen im Abschnitt „Dispatcher-WAN-Unterstützung konfigurieren“ auf Seite 229 und im Abschnitt „Unterstützung für GRE (Generic Routing Encapsulation)“ auf Seite 235.

Anmerkung: Falls der ferne Standort keine GRE-Unterstützung bietet, ist ein zusätzlicher Dispatcher am fernen Standort erforderlich.

- Wenn Sie Last mit der Dispatcher-Weiterleitungsmethode nat auf ferne Server verteilen möchten, lesen Sie den Abschnitt „Dispatcher-Weiterleitungsmethode nat“ auf Seite 53.

Anmerkung: Bei Verwendung der Weiterleitungsmethode nat ist am fernen Standort *kein* zusätzlicher Dispatcher erforderlich.

(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, ist das Feature für WAN-Lastausgleich nicht verfügbar.)

Port-Zuordnung

- Wenn Sie die Last einer Webadresse auf mehrere Serverdämonen auf einer Maschine verteilen möchten, die jeweils an einem eindeutigen Port empfangsbereit sind, lesen Sie die Informationen im Abschnitt „Dispatcher-Weiterleitungsmethode nat“ auf Seite 53.

(Falls Sie Load Balancer für IPV4 und IPV6 verwenden, ist dieses Feature nicht verfügbar.)

Dispatcher in einem privaten Netz konfigurieren

- Falls der Dispatcher-Datenverkehr in einem anderen Netz als der Client-Datenverkehr bearbeitet werden soll (um Konkurrenzsituationen im externen Netz zu reduzieren und so den Durchsatz zu erhöhen), lesen Sie den Abschnitt „Konfiguration für ein privates Netz verwenden“ auf Seite 236.

Platzhaltercluster und Platzhalter-Port

- Falls Sie mehrere Webadressen in einer Konfiguration kombinieren möchten, lesen Sie den Abschnitt „Platzhaltercluster zum Zusammenfassen von Serverkonfigurationen verwenden“ auf Seite 237.
- Informationen zum Lastausgleich für Firewalls finden Sie im Abschnitt „Platzhaltercluster für den Lastausgleich von Firewalls verwenden“ auf Seite 238.
- Wenn Sie den Datenverkehr für alle Ziel-Ports lenken möchten, lesen Sie den Abschnitt „Platzhalter-Port für die Übertragung von Datenverkehr mit nicht konfiguriertem Port verwenden“ auf Seite 239.

Erkennung von DoS-Attacken

- Informationen zur Erkennung möglicher DoS-Attacken (Denial of Service) finden Sie im Abschnitt „Erkennung von DoS-Attacken“ auf Seite 240.

Binäres Protokollieren

- Informationen zur Analyse des Serverdatenverkehrs können Sie dem Abschnitt „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 241 entnehmen.

Alerts

- Falls Alerts generiert werden sollen, wenn Server als aktiv oder inaktiv markiert werden, lesen Sie die Informationen im Abschnitt „Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden“ auf Seite 184.

Features der Komponente CBR (Content Based Routing)

CBR integriert den Lastausgleich in Caching Proxy von WebSphere Application Server, um Client-Anforderungen an angegebene HTTP- oder HTTPS-Server (SSL) weiterzuleiten. Für die Verwendung von CBR muss auf derselben Maschine Caching Proxy installiert und konfiguriert sein. Informationen zum Konfigurieren von Caching Proxy für die Verwendung von CBR finden Sie unter „Schritt 1. Caching Proxy für die Verwendung von CBR konfigurieren“ auf Seite 114.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode `cbr` der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (`cbr`)“ auf Seite 55.

Bei Verwendung der Komponente CBR (oder der Dispatcher-Weiterleitungsmethode `cbr`) ergeben sich für Ihre Clients die folgenden Vorteile:

- Sie können Client-Anfragen nach verschiedenen Inhalten auf Servergruppen verteilen. (Lesen Sie hierzu den Abschnitt „Client-Anfragen nach verschiedenen Inhalten verteilen“ auf Seite 106.)
- Sie können die Antwortzeit verkürzen, indem Sie den Inhalt der Site optimal auf mehrere Webserver verteilen. (Lesen Sie hierzu den Abschnitt „Siteinhalt für kürzere Antwortzeiten aufteilen“ auf Seite 106.)

- Sie können bei einem Serverausfall einen unterbrechungsfreien Client-Datenverkehr gewährleisten, indem Sie jedem Inhaltstyp mehrere Server zuordnen. (Lesen Sie hierzu den Abschnitt „Webserverinhalt sichern“ auf Seite 107.)

Komponente CBR und Dispatcher-Weiterleitungsmethode cbr im Vergleich

Bei einem Netz mit sicherem SSL-Datenverkehr vom Client zum Server liegt der Vorteil der Komponente CBR (in Verbindung mit Caching Proxy) darin, dass CBR die für ein inhaltsabhängiges Routing erforderliche Verschlüsselung/Entschlüsselung durchführen kann.

Bei vollständig gesicherten SSL-Verbindungen kann die Dispatcher-Weiterleitungsmethode cbr nur mit SSL-ID-Affinität konfiguriert werden, da sie nicht in der Lage ist, die für ein vom Inhalt des URL abhängiges Routing von Client-Anforderungen erforderliche Verschlüsselung/Entschlüsselung vorzunehmen.

Für HTTP-Datenverkehr besteht der Vorteil der Dispatcher-Weiterleitungsmethode cbr darin, dass sie schneller auf Client-Anfragen reagiert als die Komponente CBR. Die Dispatcher-Weiterleitungsmethode cbr erfordert auch *nicht* die Installation und Verwendung von Caching Proxy.

Fernverwaltung

- Wenn Sie Load Balancer nicht von der Maschine aus konfigurieren möchten, auf der Load Balancer installiert ist, lesen Sie den Abschnitt „Fernverwaltung von Load Balancer“ auf Seite 263.

Verknüpfung

- CBR kann auf derselben Maschine wie ein am Lastausgleich beteiligter Server ausgeführt werden. Weitere Informationen hierzu finden Sie im Abschnitt „Verknüpfte Server verwenden“ auf Seite 202.

CBR mit mehreren Instanzen von Caching Proxy

- Wie Sie die CPU-Ausnutzung durch mehrere Caching-Proxy-Prozesse verbessern können, erfahren Sie im Abschnitt „CPU-Nutzung mit mehreren Caching-Proxy-Prozessen verbessern“ auf Seite 107.

Inhaltsabhängiges Routing für SSL-Verbindungen

Wenn Sie für SSL-Datenverkehr ein inhaltsabhängiges Routing ermöglichen wollen, können Sie wie folgt vorgehen:

- Sie können auf beiden Seiten (Client-zu-Proxy und Proxy-zu-Server) sichere Verbindungen verwenden. Lesen Sie dazu den Abschnitt „Lastausgleich für sichere Verbindungen (SSL)“ auf Seite 107.
- Sie können sichere Verbindungen ausschließlich vom Client zum Proxy verwenden. Lesen Sie dazu den Abschnitt „Lastausgleich für SSL-Datenverkehr vom Client zum Proxy und HTTP-Datenverkehr vom Proxy zum Server“ auf Seite 108.

Serverpartitionierung

- Wenn Sie zwischen bestimmten URLs und ihren Serviceanwendungen unterscheiden möchten, lesen Sie den Abschnitt „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58.

Regelbasierter Lastausgleich

Sie können Clients für dieselbe Webadresse zu verschiedenen Servergruppen dirigieren, indem Sie Regeln zu Ihrer CBR-Konfiguration hinzufügen. Weitere Informationen hierzu finden Sie im Abschnitt „Regelbasierten Lastausgleich konfigurieren“ auf Seite 212.

- Falls Sie Clients ausgehend vom Inhalt des angefragten URL zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf dem Inhalt der Anforderung basierende Regeln verwenden“ auf Seite 219.
- Falls Sie Clients ausgehend von der Quellen-IP-Adresse des Clients zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf der Client-IP-Adresse basierende Regeln verwenden“ auf Seite 214.
- Falls Sie Clients ausgehend von der Uhrzeit zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf der Uhrzeit basierende Regeln verwenden“ auf Seite 214.
- Falls Sie Clients ausgehend vom Datenverkehr der Site zu verschiedenen Servergruppen dirigieren möchten, haben Sie die folgenden Möglichkeiten:
 - Sie können die Anzahl der Verbindungen pro Sekunde verwenden. Lesen Sie dazu den Abschnitt „Regeln auf der Basis der Verbindungen pro Sekunde verwenden“ auf Seite 215.
 - Sie können die Summe der aktiven Verbindungen verwenden. Lesen Sie dazu den Abschnitt „Regeln auf der Basis der Summe aktiver Verbindungen verwenden“ auf Seite 215.
- Falls Sie überlaufenden Datenverkehr zu einer Standardgruppe von Servern dirigieren möchten (z. B. zu Servern, die die Antwort "Site ausgelastet" ausgeben), lesen Sie den Abschnitt „Immer gültige Regeln verwenden“ auf Seite 219.
- Wenn Sie die Client-Affinität außer Kraft setzen möchten, um sicherzustellen, dass ein Client nicht an einen Überlaufserver gebunden bleibt, lesen Sie die Informationen im Abschnitt „Port-Affinität außer Kraft setzen“ auf Seite 220.

Client-Server-Affinität

- Wenn Sie sicherstellen möchten, dass ein Client für mehrere Verbindungen zu demselben Server zurückkehrt, lesen Sie den Abschnitt „Funktionsweise der Affinität für Load Balancer“ auf Seite 222.
- Wenn Sie einen Server aus der Konfiguration entfernen möchten (z. B. für eine Wartung), ohne den Client-Datenverkehr zu unterbrechen, lesen Sie den Abschnitt „Bearbeitung von Serververbindungen stilllegen“ auf Seite 224.
- Falls Sie unabhängig von der Erstellung von Cookies durch Ihre Webserver sicherstellen möchten, dass Clients zu demselben Server zurückkehren, wenn sie in mehreren Verbindungen ähnliche Inhalte anfragen, lesen Sie die Informationen im Abschnitt „Aktive Cookie-Affinität“ auf Seite 225.
- Wenn Sie durch Cookies von Ihren Webservern sicherstellen möchten, dass Clients zu demselben Server zurückkehren, wenn sie in mehreren Verbindungen ähnliche Inhalte anfragen, lesen Sie die Informationen im Abschnitt „Passive Cookie-Affinität“ auf Seite 227.
- Falls Sie den Webdatenverkehr auf Caching-Proxy-Server verteilen möchten, um das Zwischenspeichern eindeutiger Inhalte auf jedem einzelnen Server zu ermöglichen (und dadurch den Cache Ihrer Site vergrößern, da eine redundante Zwischenspeicherung von Inhalten auf mehreren Maschinen vermieden wird), lesen Sie den Abschnitt „URI-Affinität“ auf Seite 228.

Hohe Verfügbarkeit bei Verwendung von Dispatcher und CBR

- Falls Sie den Dispatcher in einer Client/Server-Konfiguration mit CBR einsetzen, um die Einschränkungen eines Single Point of Failure für Ihr Netz zu beseitigen, lesen Sie die Informationen im Abschnitt „Bereitstellung der hohen Verfügbarkeit durch Load Balancer“ auf Seite 6.

Binäres Protokollieren

- Informationen zur Analyse des Serverdatenverkehrs können Sie dem Abschnitt „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 241 entnehmen.

Alerts

- Falls Alerts generiert werden sollen, wenn Server als aktiv oder inaktiv markiert werden, lesen Sie die Informationen im Abschnitt „Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden“ auf Seite 184.

Features der Komponente Site Selector

Site Selector verteilt die Last einer Namensserveranforderung auf eine Gruppe von Servern.

Fernverwaltung

- Wenn Sie Load Balancer nicht von der Maschine aus konfigurieren möchten, auf der Load Balancer installiert ist, lesen Sie den Abschnitt „Fernverwaltung von Load Balancer“ auf Seite 263.

Verknüpfung

- Site Selector kann auf derselben Maschine wie ein am Lastausgleich beteiligter Server ausgeführt werden. Dazu sind keine zusätzlichen Konfigurationsschritte erforderlich.

Hohe Verfügbarkeit

- Die hohe Verfügbarkeit ist bei Verwendung mehrerer redundanter Site-Selector-Komponenten durch DNS-Methoden (Domain Name System) inhärent, sofern der Elternnamensserver korrekt konfiguriert ist und normale DNS-Wiederherstellungsmethoden zur Anwendung kommen. Beispiele für DNS-Wiederherstellungsmethoden sind die erneute Übertragung von Abfragen und die Wiederholung von Zonenübertragungen.
- Falls Sie den Dispatcher in einer Client/Server-Konfiguration mit Site Selector einsetzen, um die Einschränkungen eines Single Point of Failure für Ihr Netz zu beseitigen, lesen Sie die Informationen im Abschnitt „Bereitstellung der hohen Verfügbarkeit durch Load Balancer“ auf Seite 6.

Client-Server-Affinität

- Wenn Sie sicherstellen möchten, dass der Client für mehrere Namensserveranfragen denselben Server verwendet, lesen Sie den Abschnitt „Funktionsweise der Affinität für Load Balancer“ auf Seite 222.

- Falls Sie die Standard-DNS-Methode anwenden und die Lebensdauer (TTL, Time To Live) festlegen möchten, um die Client-Server-Affinität zu gewährleisten, lesen Sie die Informationen im Abschnitt „Hinweise zu TTL“ auf Seite 129.

Regelbasierter Lastausgleich

Sie können Client-Anforderungen nach einer Domänennamensauflösung zu verschiedenen Servergruppen dirigieren, indem Sie Regeln zu Ihrer Site-Selector-Konfiguration hinzufügen. Weitere Informationen hierzu finden Sie im Abschnitt „Regelbasierten Lastausgleich konfigurieren“ auf Seite 212.

- Falls Sie Clients ausgehend von der Quellen-IP-Adresse des Clients zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf der Client-IP-Adresse basierende Regeln verwenden“ auf Seite 214.
- Falls Sie Clients ausgehend von der Uhrzeit zu verschiedenen Servergruppen dirigieren möchten, lesen Sie den Abschnitt „Auf der Uhrzeit basierende Regeln verwenden“ auf Seite 214.
- Falls Sie Clients ausgehend von den gemessenen Lastwerten zu verschiedenen Servergruppen dirigieren möchten, lesen Sie die folgenden Abschnitte:
 „Regel 'Metrik gesamt'“ auf Seite 218
 „Regel 'Metrik Durchschnitt'“ auf Seite 218
- Falls Sie überlaufenden Datenverkehr zu einer Standardgruppe von Servern dirigieren möchten (z. B. zu Servern, die die Antwort "Site ausgelastet" ausgeben), lesen Sie den Abschnitt „Immer gültige Regeln verwenden“ auf Seite 219.

Lastausgleich im WAN

Site Selector kann in einem lokalen Netz (LAN) oder in einem Weitverkehrsnetz (WAN) ausgeführt werden.

WAN-Umgebung:

- Falls Sie die Namensserveranforderungen von Clients mit einer gewichteten RoundRobin-Auswahlmethode verteilen möchten, sind keine zusätzlichen Konfigurationsschritte erforderlich.
- Wenn die Nähe des Client-Namensservers im Netz zu den Servern, die die angeforderte Anwendung bereitstellen (den Zielservers), berücksichtigt werden soll, lesen Sie die Informationen im Abschnitt „Netzproximität verwenden“ auf Seite 130.

Alerts

- Falls Alerts generiert werden sollen, wenn Server als aktiv oder inaktiv markiert werden, lesen Sie die Informationen im Abschnitt „Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden“ auf Seite 184.

Features der Komponente Cisco CSS Controller

Cisco CSS Controller gestaltet den Serverlastausgleich von Cisco-Switches anwendungs- und systemspezifisch. Der Controller verwendet für die dynamische Berechnung von Serverwertigkeiten mehr anwendungs- und systemspezifische Messwerte. Die Wertigkeiten werden dem Switch über SNMP bereitgestellt. Der Switch nutzt diese Wertigkeiten für die Verarbeitung von Client-Anforderungen. Auf diese Weise wird die Auslastung optimiert und die Fehlertoleranz verbessert.

Wenn Sie die Last optimal auf mehrere Server verteilen und sicherstellen möchten, dass stets der "richtige" Server ausgewählt wird, lesen Sie die folgenden Abschnitte:

- „Lastausgleich mit Load Balancer optimieren“ auf Seite 248
- „Advisor-Funktionen“ auf Seite 250 und „Kundenspezifische (anpassbare) Advisor-Funktion erstellen“ auf Seite 252
- „Metric Server“ auf Seite 255

Fernverwaltung

- Wenn Sie Load Balancer nicht von der Maschine aus konfigurieren möchten, auf der Load Balancer installiert ist, lesen Sie den Abschnitt „Fernverwaltung von Load Balancer“ auf Seite 263.

Verknüpfung

- Cisco CSS Controller kann auf derselben Maschine wie ein am Lastausgleich beteiligter Server ausgeführt werden. Dazu sind keine zusätzlichen Konfigurationsschritte erforderlich.

Hohe Verfügbarkeit

- Der Cisco CSS Switch und der Cisco CSS Controller bieten eine Funktion für hohe Verfügbarkeit an, um die Einschränkungen eines Single Point of Failure für Ihr Netz zu beseitigen. Für den Switch können Sie die hohe Verfügbarkeit mit dem CSS-Redundanzprotokoll nutzen. Für den Cisco CSS Controller müssen Sie ein internes Protokoll verwenden, das die Konfiguration von zwei Controllern im fehlertoleranten Modus ermöglicht.
Weitere Informationen zum Konfigurieren der hohen Verfügbarkeit finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 146.

Binäres Protokollieren

- Informationen zur Analyse des Serverdatenverkehrs können Sie dem Abschnitt „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 258 entnehmen.

Alerts

- Falls Alerts generiert werden sollen, wenn Server als aktiv oder inaktiv markiert werden, lesen Sie die Informationen im Abschnitt „Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden“ auf Seite 259.

Features der Komponente Nortel Alteon Controller

Nortel Alteon Controller gestaltet den Serverlastausgleich für Alteon-Switches von Nortel anwendungs- und systemspezifisch. Der Controller verwendet für die dynamische Berechnung von Serverwertigkeiten mehr anwendungs- und systemspezifische Messwerte. Die Wertigkeiten werden dem Switch über SNMP bereitgestellt. Der Switch nutzt diese Wertigkeiten für die Verarbeitung von Client-Anforderungen. Auf diese Weise wird die Auslastung optimiert und die Fehlertoleranz verbessert.

Wenn Sie die Last optimal auf mehrere Server verteilen und sicherstellen möchten, dass stets der "richtige" Server ausgewählt wird, lesen Sie die folgenden Abschnitte:

- „Lastausgleich mit Load Balancer optimieren“ auf Seite 248
- „Advisor-Funktionen“ auf Seite 250 und „Kundenspezifische (anpassbare) Advisor-Funktion erstellen“ auf Seite 252
- „Metric Server“ auf Seite 255

Fernverwaltung

- Wenn Sie Load Balancer nicht von der Maschine aus konfigurieren möchten, auf der Load Balancer installiert ist, lesen Sie den Abschnitt „Fernverwaltung von Load Balancer“ auf Seite 263.

Verknüpfung

- Nortel Alteon Controller kann auf derselben Maschine wie ein am Lastausgleich beteiligter Server ausgeführt werden. Dazu sind keine zusätzlichen Konfigurationsschritte erforderlich.

Hohe Verfügbarkeit

- Der Alteon Web Switch und der Alteon Controller von Nortel bieten eine Funktion für hohe Verfügbarkeit an, um die Einschränkungen eines Single Point of Failure für Ihr Netz zu beseitigen. Für den Switch können Sie die hohe Verfügbarkeit nutzen, indem Sie für Verbindungen zu Servern und für Services ein Redundanzprotokoll verwenden. Für die hohe Verfügbarkeit des Nortel Alteon Controller müssen Sie ein internes Protokoll verwenden, das die Konfiguration von zwei Controllern im fehlertoleranten Modus ermöglicht.
Weitere Informationen zum Konfigurieren der hohen Verfügbarkeit finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 167.

Binäres Protokollieren

- Informationen zur Analyse des Serverdatenverkehrs können Sie dem Abschnitt „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 258 entnehmen.

Alerts

- Falls Alerts generiert werden sollen, wenn Server als aktiv oder inaktiv markiert werden, lesen Sie die Informationen im Abschnitt „Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden“ auf Seite 259.

Kapitel 4. Load Balancer installieren

Dieses Kapitel enthält Informationen zur Installation von Load Balancer mit den System-Tools für Paketinstallation und zu den Voraussetzungen für alle unterstützten Betriebssysteme.

- „Systemvoraussetzungen und Installation unter AIX“
- „Systemvoraussetzungen und Installation unter HP-UX“ auf Seite 35
- „Systemvoraussetzungen und Installation unter Linux“ auf Seite 37
- „Systemvoraussetzungen und Installation unter Solaris“ auf Seite 39
- „Systemvoraussetzungen und Installation unter Windows“ auf Seite 40

Anweisungen für die Installation mit dem Installationsprogramm des Produkts können Sie dem Dokument *Edge Components - Konzepte, Planung und Installation* entnehmen.

Das Java 2 SDK wird auf allen Plattformen automatisch mit Load Balancer installiert.

Wenn Sie vor der Installation eine frühere Version von Load Balancer migrieren oder ein Betriebssystem erneut installieren möchten, können Sie Ihre bisherigen Konfigurationsdateien oder Script-Dateien für Load Balancer sichern.

- Stellen Sie Ihre Konfigurationsdateien nach der Installation in das Verzeichnis `.../ibm/edge/lb/servers/configurations/Komponente`. (Hier steht *Komponente* für `dispatcher`, `cbr`, `ss`, `cco`, oder `nal`.)
- Stellen Sie Ihre Script-Dateien (wie `goIdle` und `goStandby`) nach der Installation in das Verzeichnis `.../ibm/edge/lb/servers/bin`, damit sie ausgeführt werden.

Die in diesem Abschnitt aufgelisteten Pakete mit Load-Balancer-Komponenten stehen nicht für alle Installationsarten zur Verfügung.

- Für Edge-Components-Installationen mit Load Balancer und Caching Proxy sind alle Load-Balancer-Installationskomponenten verfügbar.
- Bei Edge-Components-Installationen, die Load Balancer bereitstellen, jedoch nicht Caching Proxy, ist das Paket mit der Komponente CBR nicht in Load Balancer enthalten.
- Bei Installationen von Edge Components für IPV6 (Load Balancer für IPV4 und IPV6) gehört zu Load Balancer das Paket mit der Komponente Dispatcher. Die Komponentepakete CBR, Site Selector und Controller sind nicht enthalten. Die für die Installation der Pakete von Load Balancer für IPV4 und IPV6 empfohlene Reihenfolge ist im Abschnitt „Load Balancer für IPV4 und IPV6 installieren“ auf Seite 83 angegeben.

Systemvoraussetzungen und Installation unter AIX

Voraussetzungen für AIX-Systeme

Informationen zu den Hardware- und Softwarevoraussetzungen, einschließlich der unterstützten Browser, finden Sie auf der folgenden Webseite:

<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation auf AIX-Systemen

In Tabelle 1 sind die installp-Images für Load Balancer in der Reihenfolge aufgelistet, in der sie mit dem System-Tool für Paketinstallation installiert werden sollten.

Tabelle 1. *installp-Images für AIX*

Basisprodukt	ibmlb.base.rte
Administration (mit Nachrichten)	<ul style="list-style-type: none">• ibmlb.admin.rte• ibmlb.msg.<i>Sprache</i>.admin
Einheitentreiber	ibmlb.lb.driver
Lizenz	ibmlb.lb.license
Load-Balancer-Komponenten (mit Nachrichten)	<ul style="list-style-type: none">• ibmlb.<i>Komponente</i>.rte• ibmlb.msg.<i>Sprache</i>.lb
Dokumentation (mit Nachrichten)	<ul style="list-style-type: none">• ibmlb.doc.rte• ibmlb.msg.en_US.doc
Metric Server	ibmlb.ms.rte

Komponente kann folgende Werte annehmen: disp (Dispatcher), cbr (CBR), ss (Site Selector), cco (Cisco CSS Controller) oder nal (Nortel Alteon Controller). Wählen Sie (optional) die Komponenten aus, die Sie installieren möchten.

Sprache kann folgende Werte annehmen:

- en_US
- de_CH
- de_DE
- es_ES
- fr_CA
- fr_CH
- fr_FR
- it_CH
- it_IT
- ja_JP
- Ja_JP
- ko_KR
- pt_BR
- zh_CN
- ZH_CN
- zh_TW
- Zh_TW

Das Dokumentationspaket enthält nur englische Dokumente. Übersetzungen des Load-Balancer-Dokumentationssets finden Sie auf der Website www.ibm.com/software/webservers/appserv/ecinfocenter.html.

Installation vorbereiten

Falls Sie eine frühere Version installiert haben, deinstallieren Sie diese Kopie vor Installation der aktuellen Version. Vergewissern Sie sich zunächst, dass alle Steuerprogramme und Server gestoppt wurden. Geben Sie dann **installp -u ibmlb** (oder

den früheren Namen, z. B. **intnd**) ein, um die Installation des gesamten Produkts zu entfernen. Wenn Sie bestimmte Dateigruppen deinstallieren möchten, listen Sie diese an Stelle des Paketnamens einzeln auf.

Bei der Installation des Produkts können Sie auswählen, ob Sie nur bestimmte oder alle der in der folgenden Liste aufgeführten Optionen installieren wollen:

- Basisprodukt
- Administration (mit Nachrichten)
- Einheitentreiber (erforderlich)
- Lizenz (erforderlich)
- Komponente Dispatcher (mit Nachrichten)
- Komponente CBR (mit Nachrichten)
- Komponente Site Selector (mit Nachrichten)
- Komponente Cisco CSS Controller (mit Nachrichten)
- Komponente Nortel Alteon Controller (mit Nachrichten)
- Dokumentation (mit Nachrichten)
- Metric Server

Installationsschritte

Führen Sie die folgenden Schritte aus, um Load Balancer auf AIX-Systemen zu installieren:

1. Melden Sie sich als root an.
2. Legen Sie den Datenträger mit dem Produkt ein oder, falls Sie das Produkt aus dem Web installieren, kopieren Sie die Installationsimages in ein Verzeichnis.
3. Installieren Sie das Installationsimage. Installieren Sie Load Balancer für AIX mit SMIT, da in diesem Fall alle Nachrichten automatisch installiert werden.

Verwendung von SMIT:

Auswahl

Softwareinstallation und Wartung

Auswahl

Software installieren und aktualisieren

Auswahl

Aus gesamter verfügbarer Software installieren und aktualisieren

Eingabe

Die Einheit oder das Verzeichnis mit den installp-Images

Eingabe

In der Zeile '*Zu installierende SOFTWARE' die entsprechende Optionsangabe (oder wählen Sie 'Liste' aus)

Taste OK

Ist der Befehl vollständig ausgeführt, drücken Sie die Taste für **Ende**. Wählen Sie dann im Menü "Beenden" den Eintrag **SMIT beenden** aus oder drücken Sie die Taste **F12**. Drücken Sie bei Verwendung von SMITTY die Taste **F10**, um das Programm zu verlassen.

Verwendung der Befehlszeile:

Wenn Sie die Installation von einer CD ausführen, müssen Sie die folgenden Befehle eingeben, um die CD anzuhängen:

```
mkdir /cdrom
mount -v cdrfs -p -r /dev/cd0 /cdrom
```

Stellen Sie anhand der folgenden Tabelle fest, welche Befehle Sie eingeben müssen, um die gewünschten Pakete von Load Balancer auf AIX-Systemen zu installieren:

Tabelle 2. AIX-Installationsbefehle

Basisprodukt	<code>installp -acXgd <i>Einheit</i> ibmlb.base.rte</code>
Administration (mit Nachrichten)	<code>installp -acXgd <i>Einheit</i> ibmlb.admin.rte ibmlb.msg.<i>Sprache</i>.admin</code>
Einheitentreiber	<code>installp -acXgd <i>Einheit</i> ibmlb.lb.driver</code>
Lizenz	<code>installp -acXgd <i>Einheit</i> ibmlb.lb.license</code>
Load-Balancer-Komponenten (mit Nachrichten). Dazu gehören Dispatcher, CBR, Site Selector, Cisco CSS Controller und Nortel Alteon Controller.	<code>installp -acXgd <i>Einheit</i> ibmlb.<i>Komponente</i>.rte ibmlb.msg.<i>Sprache</i>.lb</code>
Dokumentation (mit Nachrichten)	<code>installp -acXgd <i>Einheit</i> ibmlb.doc.rte ibmlb.msg.en_US.lb</code>
Metric Server	<code>installp -acXgd <i>Einheit</i> ibmlb.ms.rte</code>

Einheit steht hier für Folgendes:

- `/cdrom`, wenn die Installation von einer CD erfolgt.
- `/dir` (das Verzeichnis mit den `installp`-Images), wenn die Installation von einem Dateisystem aus erfolgt.

Achten Sie darauf, dass die Ergebnisspalte in der Zusammenfassung für alle installierten Komponenten von Load Balancer jeweils die Angabe **ERFOLGREICH** enthält. Fahren Sie erst fort, wenn alle ausgewählten Komponenten erfolgreich installiert wurden.

Anmerkung: Wenn Sie für ein `installp`-Image eine Liste der Dateigruppen einschließlich aller verfügbaren Nachrichtenkataloge generieren möchten, geben Sie Folgendes ein:

```
installp -ld Einheit
```

Einheit steht hier für Folgendes:

- `/cdrom`, wenn die Installation von einer CD erfolgt.
- `/dir` (das Verzeichnis mit den `installp`-Images), wenn die Installation von einem Dateisystem aus erfolgt.

Geben Sie Folgendes ein, um die CD abzuhängen:

```
umount /cdrom
```

4. Überprüfen Sie, ob das Produkt installiert ist. Geben Sie den folgenden Befehl ein:

```
lsllpp -h | grep ibmlb
```

Wurde das gesamte Produkt installiert, gibt dieser Befehl Folgendes zurück:

```
ibmlb.base.rte
ibmlb.admin.rte
ibmlb.lb.driver
ibmlb.lb.license
ibmlb.<Komponente>.rte
ibmlb.doc.rte
ibmlb.ms.rte
ibmlb.msg.Sprache.admin
ibmlb.msg.en_US.doc
ibmlb.msg.Sprache.lb
```


Für Load Balancer gelten die folgenden Installationspfade:

- Administration - **/opt/ibm/edge/lb/admin**
- Komponenten von Load Balancer - **/opt/ibm/edge/lb/servers**
- Metric Server - **/opt/ibm/edge/lb/ms**
- Dokumentation (*Administratorhandbuch*) - **/opt/ibm/edge/lb/documentation**

Für die Fernverwaltung von Load Balancer mit Remote Method Invocation (RMI) müssen Sie auf dem Client die Pakete "Administration", "Basisprodukt", "Komponente" und "Lizenz" installieren. Weitere Informationen zu RMI finden Sie im Abschnitt „Remote Method Invocation (RMI)“ auf Seite 264.

Systemvoraussetzungen und Installation unter HP-UX

Voraussetzungen für HP-UX-Systeme

Informationen zu den Hardware- und Softwarevoraussetzungen, einschließlich der unterstützten Browser, finden Sie auf der folgenden Webseite:

<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation auf HP-UX-Systemen

In den folgenden Abschnitten wird erklärt, wie Load Balancer auf HP-UX-Systemen von der Produkt-CD installiert wird.

Installation vorbereiten

Vergewissern Sie sich vor Beginn des Installationsverfahrens, dass Sie die root-Berechtigung für die Installation der Software haben.

Haben Sie eine frühere Version installiert, sollten Sie die Installation dieser Kopie entfernen, bevor Sie die aktuelle Version installieren. Vergewissern Sie sich zunächst, dass das Steuerprogramm und der Server gestoppt wurden. Fahren Sie dann mit dem Abschnitt „Anweisungen für die Deinstallation der Pakete“ auf Seite 36 fort, um Load Balancer zu deinstallieren.

Installationsschritte

In Tabelle 3 sind die Namen der Installationspakete für Load Balancer in der Reihenfolge aufgelistet, in der sie mit dem System-Tool für Paketinstallation installiert werden sollten.

Tabelle 3. Details zur Paketinstallation von Load Balancer unter HP-UX

Paketbeschreibung	Name des HP-UX-Pakets
Basisprodukt	ibmlb.base
Administration und Nachrichten	ibmlb.admin ibmlb.nlv-Sprache
Lizenz für Load Balancer	ibmlb.lic
Load-Balancer-Komponenten	ibmlb.Komponente
Dokumentation	ibmlb.doc
Metric Server	ibmlb.ms

Tabelle 3. Details zur Paketinstallation von Load Balancer unter HP-UX (Forts.)

Anmerkungen:

1. Die Variable *Sprache* muss durch einen der folgenden sprachspezifischen Codes ersetzt werden: de_DE, en_US, es_ES, fr_FR, it_IT, ja_JP, ko_KR, zh_CN, zh_TW.
2. Die Variable *Komponente* muss durch einen der folgenden Werte ersetzt werden: disp (Dispatcher), cbr (CBR), ss (Site Selector), cco (Cisco CSS Controller) oder nal (Nortel Alteon Controller).
3. Das Dokumentationspaket (ibmlb.doc) enthält nur englische Dokumente. Übersetzungen des Load-Balancer-Dokumentationssets finden Sie auf der Website www.ibm.com/software/webervers/appserv/ecinfocenter.html.

Anmerkung: HP-UX-Systeme bieten keine Unterstützung für brasilianisches Portugiesisch (pt_BR). Folgende Landessprachen werden von HP-UX-Systemen unterstützt:

- de_DE.iso88591
- en_US.iso88591
- es_ES.iso88591
- fr_FR.iso88591
- it_IT.iso88591
- ja_JP.SJIS
- ko_KR.eucKR
- zh_CN.hp15CN
- zh_TW.big5

Anweisungen für die Installation der Pakete

Nachfolgend sind die Schritte aufgeführt, die zur vollständigen Ausführung dieser Task erforderlich sind.

1. Melden Sie sich als lokaler Superuser root an.

```
su - root
Password: Kennwort
```

2. Setzen Sie den folgenden Installationsbefehl für die Installation der Pakete ab:

```
swinstall -s /Quelle Paketname
```

Hier steht *Quelle* für den absoluten Verzeichnispfad zur Position des Pakets und *Paketname* für den Namen des Pakets.

Der folgende Befehl installiert nur das Basispaket für Load Balancer (ibmlb.base) vom Ausgangsverzeichnis der CD aus:

```
swinstall -s /Quelle ibmlb.base
```

Setzen Sie den folgenden Befehl ab, wenn Sie alle Pakete für Load Balancer vom Ausgangsverzeichnis der CD aus installieren möchten:

```
swinstall -s /Quelle ibmlb
```

3. Überprüfen Sie die Installation der Load-Balancer-Pakete.

Setzen Sie den Befehl **swlist** ab, um alle installierten Pakete aufzulisten.

Beispiel:

```
swlist -l fileset ibmlb
```

Anweisungen für die Deinstallation der Pakete

Verwenden Sie zum Deinstallieren der Pakete den Befehl **swremove**. Entfernen Sie die Pakete in der umgekehrten Reihenfolge, in der sie installiert wurden. Setzen Sie beispielsweise die folgenden Befehle ab:

- Für die Deinstallation aller Load-Balancer-Pakete:

```
swremove ibmlb
```

Für die Deinstallation eines einzelnen Pakets, z. B. der Komponente Dispatcher:

```
swremove ibmlb.disp
```

Systemvoraussetzungen und Installation unter Linux

Voraussetzungen für Linux-Systeme

Die Hardware- und Softwarevoraussetzungen, einschließlich der unterstützten Browser, finden Sie auf der folgenden Webseite: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation auf Linux-Systemen

In den folgenden Abschnitten wird erklärt, wie Load Balancer unter Linux von der Produkt-CD installiert wird.

Installation vorbereiten

Vergewissern Sie sich vor Beginn des Installationsverfahrens, dass Sie die root-Berechtigung für die Installation der Software haben.

Falls Sie eine frühere Version installiert haben, sollten Sie diese Kopie vor Installation der aktuellen Version deinstallieren. Vergewissern Sie sich zunächst, dass alle Steuerprogramme und Server gestoppt wurden. Geben Sie anschließend **rpm -e Paketname** ein, um das gesamte Produkt zu deinstallieren. Bei der Deinstallation ist die Reihenfolge umzukehren, die für die Installation der Pakete verwendet wurde. Damit wird sichergestellt, dass die Administrationspakete zuletzt deinstalliert werden.

Installationsschritte

Gehen Sie wie folgt vor, um Load Balancer zu installieren:

1. Bereiten Sie die Installation vor.

- Melden Sie sich als root an.
- Legen Sie den Produktdatenträger ein oder laden Sie das Produkt von der Website herunter und installieren Sie das Installationsimage mit Hilfe von RPM (Red Hat Packaging Manager).

Das Installationsimage ist eine Datei im Format **eLBLX-Version:tar.z**.

- Entpacken Sie die tar-Datei im temporären Verzeichnis, indem Sie Folgendes eingeben: **tar -xf eLBLX-Version:tar.z** Das Ergebnis ist eine Gruppe von Dateien mit der Erweiterung .rpm.

Die folgende Liste enthält die mit RPM installierbaren Pakete.

- *ibmlb-base-Releaseversion.Hardware.rpm* (Basisprodukt)
- *ibmlb-admin-Releaseversion.Hardware.rpm* (Administration)
- *ibmlb-lic-Releaseversion.Hardware.rpm* (Lizenz)
- *ibmlb-Komponente-Releaseversion.Hardware.rpm* (LB-Komponente)
- *ibmlb-doc-Releaseversion.Hardware.rpm* (Dokumentation)
- *ibmlb-ms-Releaseversion.Hardware.rpm* (Metric Server)

Für diese Angaben gilt Folgendes:

- *Releaseversion* ist das aktuelle Release, z. B. 6.1-0.

- *Hardware* steht für einen der folgenden Werte: i386, ppc64, ppc, s390, s390x, x86_64
- *Komponente* steht für einen der folgenden Werte: disp (Komponente Dispatcher), cbr (Komponente CBR), ss (Komponente Site Selector), cco (Cisco CSS Controller), nal (Nortel Alteon Controller)

Das Dokumentationspaket enthält nur englische Dokumente. Übersetzungen des Load-Balancer-Dokumentationssets finden Sie auf der Website www.ibm.com/software/webervers/appserv/ecinfocenter.html.

- Die Reihenfolge, in der die Pakete installiert werden, ist wichtig. Die folgende Liste zeigt die erforderlichen Pakete und die Reihenfolge, in der sie installiert werden müssen:

- Basisprodukt (base)
- Administration (admin)
- Lizenz (lic)
- Load-Balancer-Komponenten (disp, cbr, ss, cco, nal)
- Metric Server (ms)
- Dokumentation (doc).

Der Befehl zum Installieren der Pakete sollte von dem Verzeichnis mit den RPM-Dateien aus abgesetzt werden. Setzen Sie zum Installieren der einzelnen Pakete den Befehl **rpm -i Paket.rpm** ab.

Red-Hat-Linux-Systeme: Wegen eines bekannten Problems von Red Hat Linux müssen Sie auch die *_db**-RPM-Dateien löschen, da andernfalls ein Fehler auftritt.

- Für Load Balancer gelten die folgenden Installationspfade:
 - Administration - **/opt/ibm/edge/lb/admin**
 - Komponenten von Load Balancer - **/opt/ibm/edge/lb/servers**
 - Metric Server - **/opt/ibm/edge/lb/ms**
 - Dokumentation - **/opt/ibm/edge/lb/documentation**
- Bei der Deinstallation der Pakete ist die Reihenfolge umzukehren, die für die Installation der Pakete verwendet wurde. Damit wird sichergestellt, dass das Administrationspaket zuletzt deinstalliert wird.

2. Überprüfen Sie, ob das Produkt installiert ist. Geben Sie den folgenden Befehl ein:

rpm -qa | grep ibmlb

Wurde das gesamte Produkt installiert, wird eine Liste wie die folgende generiert:

- *ibmlb-base-Releaseversion*
- *ibmlb-admin-Releaseversion*
- *ibmlb-lic-Releaseversion*
- *ibmlb-dsp-Releaseversion*
- *ibmlb-cbr-Releaseversion*
- *ibmlb-ss-Releaseversion*
- *ibmlb-cco-Releaseversion*
- *ibmlb-nal-Releaseversion*
- *ibmlb-doc-Releaseversion*
- *ibmlb-ms-Releaseversion*

Für die Fernverwaltung von Load Balancer mit Remote Method Invocation (RMI) müssen Sie auf dem Client die Pakete "Administration", "Basisprodukt", "Komponente" und "Lizenz" installieren. Weitere Informationen zu RMI finden Sie im Abschnitt „Remote Method Invocation (RMI)“ auf Seite 264.

Systemvoraussetzungen und Installation unter Solaris

Voraussetzungen für Solaris

Informationen zu den Hardware- und Softwarevoraussetzungen, einschließlich der unterstützten Browser, finden Sie auf der folgenden Webseite:

<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation unter Solaris

In den folgenden Abschnitten wird erklärt, wie Load Balancer auf Solaris-Systemen von der Produkt-CD installiert wird.

Installation vorbereiten

Vergewissern Sie sich vor Beginn des Installationsverfahrens, dass Sie die root-Berechtigung für die Installation der Software haben.

Haben Sie eine frühere Version installiert, entfernen Sie die Installation dieser Kopie, bevor Sie die aktuelle Version installieren. Vergewissern Sie sich zunächst, dass alle Steuerprogramme (Executors) und Server gestoppt wurden. Geben Sie dann zur Deinstallation von Load Balancer **pkgrm** *Paketname* ein.

Installationsschritte

Gehen Sie wie folgt vor, um Load Balancer zu installieren:

1. Bereiten Sie die Installation vor.

- Melden Sie sich als Benutzer root an.
- Legen Sie die CD-ROM mit der Load-Balancer-Software in das entsprechende Laufwerk ein.

Geben Sie an der Eingabeaufforderung **pkgadd -d** *Pfadname* ein. Dabei ist *Pfadname* der Einheitenname des CD-ROM-Laufwerks oder das Verzeichnis auf dem Festplattenlaufwerk, in dem sich das Paket befindet. Beispiel: **pkgadd -d /cdrom/cdrom0/**.

Die folgende Liste zeigt die erforderlichen Pakete und die Reihenfolge, in der sie installiert werden sollten.

- **ibmlbbase** (Basisprodukt)
- **ibmlbadm** (Administration)
- **ibmlblic** (Lizenz)
- **ibmlbdisp** (Komponente Dispatcher)
- **ibmlbcbr** (Komponente CBR)
- **ibmlbss** (Komponente Site Selector)
- **ibmlbcc** (Komponente Cisco CSS Controller)
- **ibmlbna** (Komponente Nortel Alteon Controller)
- **ibmlbdoc** (Dokumentation)
- **ibmlbms** (Metric Server)

Das Dokumentationspaket (ibmlbdoc) enthält nur englische Dokumente. Übersetzungen des Load-Balancer-Dokumentationssets finden Sie auf der Website www.ibm.com/software/webervers/appserv/ecinfocenter.html.

Sollen alle Pakete installiert werden, geben Sie einfach "all" ein und drücken Sie die Rückföhrtaste. Sollen einzelne Komponenten installiert werden, geben Sie die Namen der zu installierenden Pakete durch ein Leerzeichen oder Komma getrennt ein und drücken Sie die Rückföhrtaste. Möglicherweise werden Sie aufgefordert, Berechtigungen für vorhandene Verzeichnisse oder Dateien zu ändern. Drücken Sie einfach die Rückföhrtaste oder antworten Sie mit "yes". Sie müssen vorausgesetzte Pakete installieren (da die Installation in alphabetischer Reihenfolge und nicht in der Reihenfolge der vorausgesetzten Pakete erfolgt). Haben Sie "all" eingegeben, antworten Sie auf alle Eingabeaufforderungen mit "yes". Die Installation wird dann erfolgreich ausgeführt.

Wenn Sie beispielsweise nur die Komponente Dispatcher mit der Dokumentation und Metric Server installieren möchten, müssen Sie die Pakete ibmlbbase, ibmlbadm, imblblic, ibmdisp, ibmlbdoc und ibmlbms installieren.

Für die Fernverwaltung von Load Balancer mit Remote Method Invocation (RMI) müssen Sie auf dem Client die Pakete "Administration", "Basisprodukt", "Komponente" und "Lizenz" installieren. Weitere Informationen zu RMI finden Sie im Abschnitt „Remote Method Invocation (RMI)“ auf Seite 264.

Für Load Balancer gelten die folgenden Installationspfade:

- Die Load-Balancer-Komponenten befinden sich im Installationsverzeichnis **/opt/ibm/edge/lb/servers**.
- Die installierte Administrationskomponente befindet sich im Verzeichnis **/opt/ibm/edge/lb/admin**.
- Der installierte Metric Server befindet sich im Verzeichnis **/opt/ibm/edge/lb/ms**.
- Die installierte Dokumentation befindet sich im Verzeichnis **/opt/ibm/edge/lb/documentation**.

2. Überprüfen Sie, ob das Produkt installiert ist. Setzen Sie den folgenden Befehl ab: **pkginfo | grep ibm**

Systemvoraussetzungen und Installation unter Windows

Voraussetzungen für Windows-Systeme

Informationen zu den Hardware- und Softwarevoraussetzungen, einschließlich der unterstützten Browser, finden Sie auf der folgenden Webseite:

<http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Installation auf Windows-Systemen

In den folgenden Abschnitten wird erklärt, wie Load Balancer auf Windows-Systemen von der Produkt-CD installiert wird.

Installationspakete

Es wird eine Liste mit Paketen angezeigt, die installiert werden können:

- Administration
- Lizenz
- Dispatcher
- Content Based Routing
- Site Selector
- Cisco CSS Controller

- Nortel Alteon Controller
- Dokumentation
- Metric Server

Für die Fernverwaltung von Load Balancer mit Remote Method Invocation (RMI) müssen Sie auf dem Client die Pakete "Administration" und "Lizenz" sowie die Komponentepakete installieren. Weitere Informationen zu RMI finden Sie im Abschnitt „Remote Method Invocation (RMI)“ auf Seite 264.

Installation vorbereiten

Einschränkungen: Die Windows-Version von Load Balancer kann nicht auf derselben Maschine wie IBM Firewall installiert werden.

Vergewissern Sie sich vor Beginn der Installation, dass Sie als Administrator oder Benutzer mit Administratorberechtigung angemeldet sind.

Falls Sie eine frühere Version installiert haben, sollten Sie diese Kopie vor Installation der aktuellen Version deinstallieren. Gehen Sie zum Deinstallieren mit der Option **Software** wie folgt vor:

1. Klicken Sie auf **Start** > **Einstellungen** (für Windows 2000) > **Systemsteuerung**.
2. Klicken Sie doppelt auf **Software**.
3. Wählen Sie *IBM WebSphere Edge Components* (oder den früheren Namen, z. B. *IBM Edge Server*) aus.
4. Klicken Sie auf die Schaltfläche **Ändern/Entfernen**.

Installationsschritte

Gehen Sie wie folgt vor, um Load Balancer zu installieren:

1. Legen Sie die CD-ROM mit Load Balancer in das CD-ROM-Laufwerk ein. Das Installationsfenster sollte automatisch angezeigt werden.
2. Der folgende Schritt ist nur erforderlich, wenn die automatische Ausführung der CD auf Ihrem Computer nicht funktioniert. Verwenden Sie für die folgenden Tasks die erste (linke) Maustaste:
 - Klicken Sie auf **Start**.
 - Wählen Sie **Ausführen** aus.
 - Geben Sie das CD-ROM-Laufwerk gefolgt von setup.exe an. Beispiel:
`E:\setup`
3. Wählen Sie die **Sprache** aus, die für den Installationsprozess verwendet werden soll.
4. Klicken Sie auf **OK**.
5. Befolgen Sie die Anweisungen des Installationsprogramms.
6. Wollen Sie das Ziellaufwerk oder -verzeichnis ändern, klicken Sie auf **Durchsuchen**.
7. Sie können "Das gesamte Produkt Load Balancer" oder "Ihre Auswahl der Komponenten" auswählen.
8. Nach Abschluss der Installation erscheint eine Nachricht, in der Sie aufgefordert werden, vor der Benutzung von Load Balancer einen Warmstart auszuführen. Dies ist erforderlich, um sicherzustellen, dass alle Dateien installiert sind und die Umgebungsvariable IBMLBPATH zur Registrierungsdatenbank hinzugefügt wurde.

Für Load Balancer gelten die folgenden Installationspfade:

- Administration – **C:\Programme\IBM\edge\lb\admin**
- Load-Balancer-Komponenten – **C:\Programme\IBM\edge\lb\servers**
- Metric Server – **C:\Programme\IBM\edge\lb\ms**
- Dokumentation (Administratorhandbuch) – **C:\Programme\IBM\edge\lb\documentation**

Anmerkung: Die Dokumentation im Installationsverzeichnis enthält nur englische Dokumente. Übersetzungen des Load-Balancer-Dokumentationssets finden Sie auf der Website www.ibm.com/software/webservers/appserv/ecinfocenter.html.

Teil 2. Dispatcher

Dieser Teil enthält Informationen zu einer schnellen Erstkonfiguration sowie zur Planung und beschreibt die Konfigurationsmethoden für die Komponente Dispatcher von Load Balancer. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 5, „Schnellkonfiguration“, auf Seite 45
- Kapitel 6, „Planung für Dispatcher“, auf Seite 51
- Kapitel 7, „Dispatcher konfigurieren“, auf Seite 63
- Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81

Kapitel 5. Schnellkonfiguration

Dieses Beispiel zeigt die Konfiguration von drei lokal angeschlossenen Workstations, die die Weiterleitungsmethode `mac` der Komponente Dispatcher verwenden, um den Webdatenverkehr auf zwei Webserver zu verteilen. Für die Verteilung des Datenverkehrs einer anderen TCP-Anwendung oder einer kontextlosen UDP-Anwendung würde die Konfiguration im Wesentlichen genauso aussehen.

Wichtiger Hinweis: Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, lesen Sie auch die Informationen in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

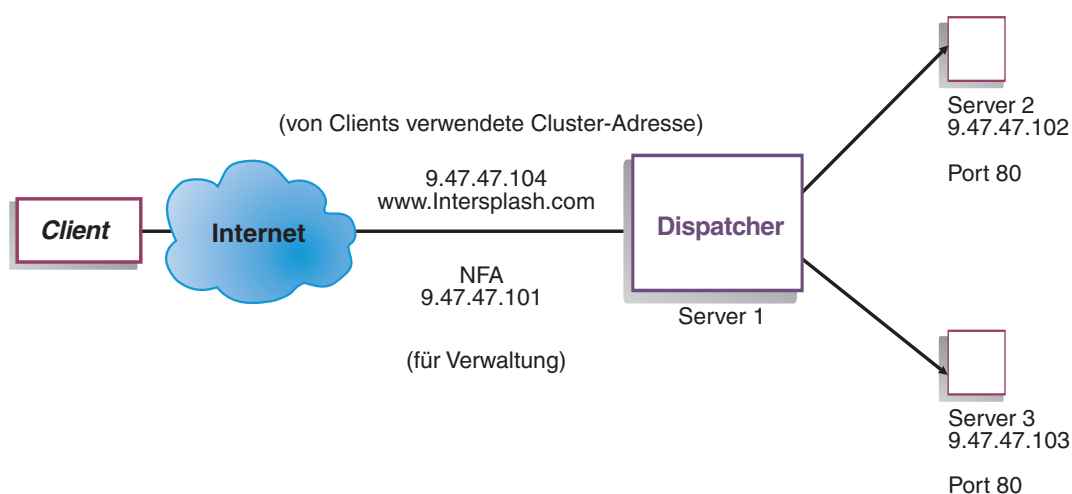


Abbildung 8. Einfache lokale Dispatcher-Konfiguration

Die `mac`-Weiterleitung ist die Standardweiterleitungsmethode, bei der der Dispatcher eingehende Anforderungen auf die Server verteilt und der jeweilige Server die Antwort direkt an den Client zurückgibt. Weitere Informationen zur Dispatcher-Weiterleitungsmethode `mac` finden Sie im Abschnitt „Dispatcher-Weiterleitungsmethode `mac`“ auf Seite 53.

Anmerkung: Für die Konfiguration reichen zwei Workstations. Der Dispatcher befindet sich dabei auf einer der Webserverworkstations. Dies wäre dann eine verknüpfte Konfiguration. Prozeduren für das Erstellen komplexer Konfigurationen finden Sie im Abschnitt „Dispatcher-Maschine konfigurieren“ auf Seite 66.

Voraussetzungen

In dem Beispiel für einen schnellen Start werden drei Workstations und vier IP-Adressen benötigt. Eine Workstation ist die CBR-Maschine, und die beiden anderen Workstations sind die Webserver. Jeder Webserver benötigt eine IP-Adresse. Die Dispatcher-Workstation muss zwei Adressen haben: die NFA (Non-Forwarding Address) und die Clusteradresse (für den Lastausgleich). Beide Adressen werden den Clients für den Zugriff auf Ihre Website zur Verfügung gestellt.

Anmerkung: Die NFA ist die Adresse, die vom Befehl **hostname** zurückgegeben wird. Sie ist für Verwaltungszwecke (z. B. für die Fernkonfiguration) bestimmt.

Vorbereitungen

1. Konfigurieren Sie Ihre Workstations für dieses Beispiel mit lokalem Anschluss so, dass sie sich innerhalb eines LAN-Segments befinden. Stellen Sie sicher, dass der Datenaustausch im Netz zwischen den drei Maschinen nicht über Router oder Brücken erfolgen muss. (Informationen zu Konfigurationen mit fern-Servern finden Sie im Abschnitt „Dispatcher-WAN-Unterstützung konfigurieren“ auf Seite 229.)
2. Konfigurieren Sie die Netzwerkadapter der drei Workstations. In diesem Beispiel wird die folgende Netzkonfiguration angenommen:

Workstation	Name	IP-Adresse
1	server1.Intersplashx.com	9.47.47.101
2	server2.Intersplashx.com	9.47.47.102
3	server3.Intersplashx.com	9.47.47.103
Netzmaske = 255.255.255.0		

Jede Workstation enthält nur eine Standard-Ethernet-Netzschnittstellenkarte.

3. Stellen Sie sicher, dass server1.Intersplashx.com ping-Aufrufe an server2.Intersplashx.com und server3.Intersplashx.com senden kann.
4. Stellen Sie sicher, dass server2.Intersplashx.com und server3.Intersplashx.com ping-Aufrufe an server1.Intersplashx.com senden können.
5. Stellen Sie sicher, dass der Inhalt auf den beiden Webservern (Server 2 und 3) identisch ist. Dies kann durch die Vervielfältigung der Daten auf beiden Workstations, durch die Verwendung eines gemeinsamen Dateisystems wie beispielsweise NFS, AFS oder DFS oder durch eine andere für Ihre Site geeignete Methode erreicht werden.
6. Stellen Sie sicher, dass die Webserver auf server2.Intersplashx.com und server3.Intersplashx.com betriebsbereit sind. Fordern Sie mit einem Webbrowser Seiten direkt von **http://server2.Intersplashx.com** und **http://server3.Intersplashx.com** an.
7. Definieren Sie eine andere gültige IP-Adresse für dieses LAN-Segment. Dies ist die Adresse, die Sie den Clients zur Verfügung stellen, die auf Ihre Site zugreifen möchten. In diesem Beispiel wird folgende Adresse verwendet:

Name=www.Intersplashx.com
IP=9.47.47.104

8. Konfigurieren Sie die beiden Webserverworkstations so, dass sie Datenverkehr für www.Intersplashx.com akzeptieren.

Fügen Sie zur **Loopback**-Schnittstelle von server2.Intersplashx.com und server3.Intersplashx.com einen Aliasnamen für www.Intersplashx.com hinzu.

- Für AIX-Systeme:
ifconfig lo0 alias www.Intersplashx.com netmask 255.255.255.0
- Für Solaris-9-Systeme:
ifconfig lo0:1 plumb www.Intersplashx.com netmask 255.255.255.0 up
- Für andere Betriebssysteme: siehe Tabelle 5 auf Seite 73.

9. Löschen Sie alle zusätzlichen Routen, die unter Umständen durch die Festlegung von Aliasnamen auf der Loopback-Schnittstelle erstellt wurden. Weitere Informationen hierzu finden Sie in „Schritt 2. Überprüfung auf zusätzliche Route“ auf Seite 76.

Sie haben jetzt alle für die beiden Webserverworkstations erforderlichen Konfigurationsschritte ausgeführt.

Dispatcher konfigurieren

Für den Dispatcher können Sie eine Konfiguration unter Verwendung der Befehlszeile, des Konfigurationsassistenten oder der grafischen Benutzerschnittstelle (GUI) erstellen.

Anmerkung: Die Parameterwerte müssen mit Ausnahme der Parameterwerte für Hostnamen und Dateinamen in englischen Zeichen eingegeben werden.

Konfiguration von der Befehlszeile aus

Führen Sie folgende Schritte aus, wenn Sie die Befehlszeile verwenden:

1. Starten Sie wie folgt den dserver für Dispatcher:
 - Führen Sie auf AIX-, HP-UX-, Linux- oder Solaris-Systemen den folgenden Befehl als Benutzer root aus: **dserver**
 - Auf Windows-Systemen ist dserver ein Dienst, der automatisch gestartet wird.
2. Starten Sie wie folgt die Executor-Funktion des Dispatchers:
dscontrol executor start
3. Fügen Sie wie folgt die Clusteradresse zur Dispatcher-Konfiguration hinzu:
dscontrol cluster add www.Intersplashx.com
4. Fügen Sie wie folgt den Port für das Protokoll HTTP zur Dispatcher-Konfiguration hinzu:
dscontrol port add www.Intersplashx.com:80
5. Fügen Sie wie folgt alle Webserver zur Dispatcher-Konfiguration hinzu:
dscontrol server add www.Intersplashx.com:80:server2.Intersplashx.com
dscontrol server add www.Intersplashx.com:80:server3.Intersplashx.com
6. Konfigurieren Sie wie folgt die Workstation, so dass sie den Datenverkehr für die Clusteradresse akzeptiert:
dscontrol executor configure www.Intersplashx.com
7. Starten Sie wie folgt die Manager-Funktion des Dispatchers:
dscontrol manager start
Der Dispatcher führt den Lastausgleich jetzt ausgehend von der Serverleistung durch.
8. Starten Sie wie folgt die Advisor-Funktion des Dispatchers:
dscontrol advisor start http 80
Der Dispatcher stellt jetzt sicher, dass keine Client-Anforderungen an einen ausgefallenen Webserver gesendet werden.

Die Basiskonfiguration mit lokal angeschlossenen Servern ist damit vollständig.

Konfiguration testen

Testen Sie wie folgt, ob die Konfiguration korrekt ist:

1. Rufen Sie mit einem Webbrowser die Adresse **http://www.Intersplashx.com** auf. Wenn eine Seite angezeigt wird, funktioniert die Konfiguration.
2. Laden Sie die Seite erneut im Webbrowser.
3. Überprüfen Sie die Ergebnisse des folgenden Befehls: **dscontrol server report www.Intersplashx.com:80:**. Die Einträge der Spalte "Summe Verbindungen" für beide Server sollten addiert "2" ergeben.

Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus

Informationen zur Verwendung der Dispatcher-GUI finden Sie im Abschnitt „GUI“ auf Seite 65 und in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Konfigurationsassistent

Informationen zur Verwendung des Konfigurationsassistenten finden Sie im Abschnitt „Konfiguration mit dem Konfigurationsassistenten“ auf Seite 66.

Arten von Cluster-, Port- und Serverkonfigurationen

Es gibt viele Möglichkeiten, Load Balancer für die Unterstützung Ihrer Site zu konfigurieren. Wenn Sie für Ihre Site nur einen Hostnamen haben, zu dem alle Kunden eine Verbindung herstellen, können Sie einen Cluster mit Servern definieren. Für jeden dieser Server konfigurieren Sie einen Port, über den Load Balancer kommuniziert. Siehe Abb. 9.

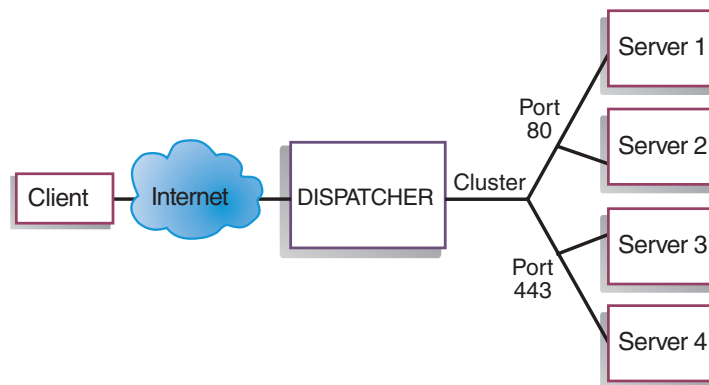


Abbildung 9. Dispatcher-Beispielkonfiguration mit einem Cluster und zwei Ports

In diesem Beispiel ist für die Komponente Dispatcher ein Cluster mit der Adresse www.productworks.com definiert. Dieser Cluster hat zwei Ports: Port 80 für HTTP und Port 443 für SSL. Ein Client, der eine Anforderung an <http://www.productworks.com> (Port 80) richtet, wird einem anderen Server zugeordnet als ein Client, der eine Anforderung an <http://www.productworks.com> (Port 443) richtet.

Wenn Ihre Site sehr groß ist und Sie für jedes unterstützte Protokoll mehrere dedizierte Server haben, sollten Sie Load Balancer auf andere Weise konfigurieren. In diesem Fall könnten Sie für jedes Protokoll einen Cluster mit nur einem Port, aber mehreren Servern definieren (siehe Abb. 10).

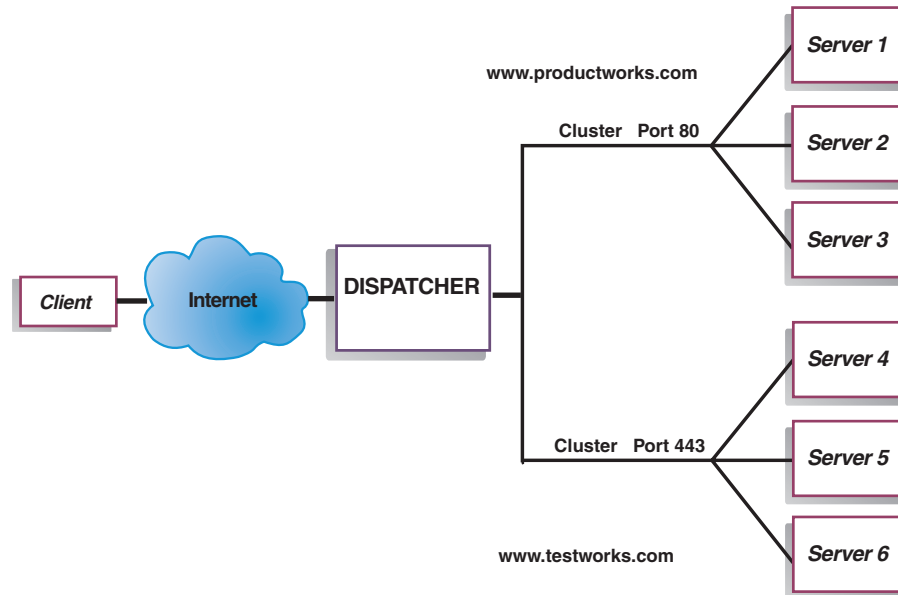


Abbildung 10. Dispatcher-Beispielkonfiguration mit zwei Clustern mit jeweils einem Port

In diesem Beispiel für die Komponente Dispatcher sind zwei Cluster definiert: `www.productworks.com` für Port 80 (HTTP) und `www.testworks.com` für Port 443 (SSL).

Wenn Ihre Site Inhalte für mehrere Unternehmen oder Abteilungen bereitstellt, die jeweils mit einem eigenen URL auf Ihre Site zugreifen, muss Load Balancer auf eine dritte Art konfiguriert werden. In diesem Fall könnten Sie für jede Firma oder Abteilung einen Cluster definieren und anschließend die Ports, an denen Verbindungen mit dem jeweiligen URL empfangen werden sollen (siehe Abb. 11 auf Seite 50).

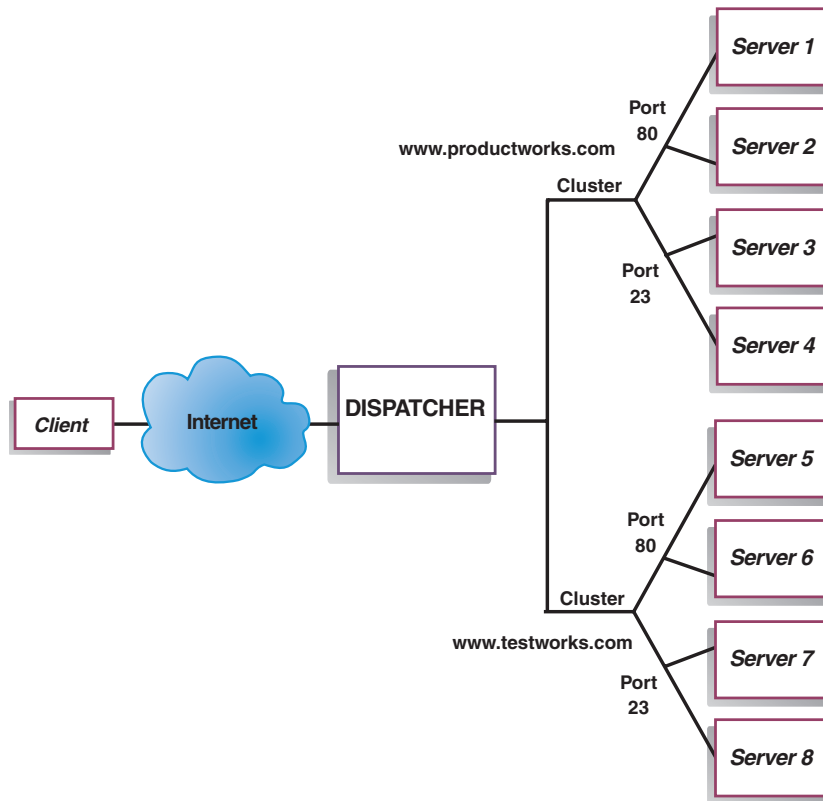


Abbildung 11. Dispatcher-Beispielkonfiguration mit zwei Clustern mit jeweils zwei Ports

In diesem Beispiel für die Komponente Dispatcher wurden für die Sites `www.productworks.com` und `www.testworks.com` jeweils zwei Cluster mit Port 80 (HTTP) und Port 23 (Telnet) definiert.

Kapitel 6. Planung für Dispatcher

In diesem Kapitel wird beschrieben, was die für die Planung des Netzes zuständige Person vor der Installation und Konfiguration der Komponente Dispatcher berücksichtigen muss.

- In Kapitel 3, „Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer“, auf Seite 19 finden Sie einen Überblick über die Features, die für die Verwaltung Ihres Netzes verfügbar sind.
- Informationen zum Konfigurieren der Lastausgleichsparameter von Dispatcher finden Sie in Kapitel 7, „Dispatcher konfigurieren“, auf Seite 63.
- Falls Sie Load Balancer für IPV4 und IPV6 verwenden, lesen Sie Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.
- Informationen zum Konfigurieren von Load Balancer für erweiterte Funktionen finden Sie in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Dieses Kapitel enthält die folgenden Abschnitte:

- „Überlegungen bei der Planung“
- „Dispatcher-Weiterleitungsmethode mac“ auf Seite 53
- „Dispatcher-Weiterleitungsmethode nat“ auf Seite 53
- „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55
- „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58
- „Hohe Verfügbarkeit“ auf Seite 60

Anmerkung: Frühere Versionen dieses Produkts liefen unter dem Namen Network Dispatcher. In diesen Versionen war der Dispatcher-Steuerbefehl `ndcontrol`. Jetzt ist der Dispatcher-Steuerbefehl **`dscontrol`**.

Überlegungen bei der Planung

Wichtiger Hinweis: Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, lesen Sie auch die Informationen in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

Der Dispatcher stellt die folgenden Funktionen bereit:

- Der **`dsserver`** bearbeitet Anforderungen von der Befehlszeile an den Executor, den Manager und die Advisor-Funktionen.
- Der **`Executor`** unterstützt die Verteilung von TCP- und UDP-Verbindungen auf Port-Basis. Der Executor kann Verbindungen ausgehend vom Typ der empfangenen Anforderung (HTTP, FTP, SSL usw.) an Server weiterleiten. Er wird immer ausgeführt, wenn die Komponente Dispatcher für den Lastausgleich verwendet wird.

- Der **Manager** definiert Wertigkeiten, die vom Executor verwendet werden und auf folgenden Kriterien basieren:
 - interne Zähler des Executors
 - von den Advisor-Funktionen bereitgestellte Rückmeldungen von den Servern
 - Rückmeldungen von einem Systemüberwachungsprogramm wie Metric Server oder WLM.

Die Benutzung des Managers ist optional. Ohne den Manager wird der Lastausgleich nach einer gewichteten RoundRobin-Zeitplanung und ausgehend von den aktuellen Serverwertigkeiten durchgeführt. Es stehen keine Advisor-Funktionen zur Verfügung.

- Die **Advisor-Funktionen** richten Abfragen an die Server und analysieren die Ergebnisse nach Protokoll, bevor sie den Manager zum Festlegen der entsprechenden Wertigkeiten aufrufen. Derzeit sind Advisor-Funktionen für die folgenden Protokolle verfügbar: HTTP, FTP, SSL, SMTP, NNTP, IMAP, POP3, SIP und Telnet.

Dispatcher bietet außerdem Advisor-Funktionen an, die keine protokollspezifischen Informationen austauschen. Dazu gehören unter anderem die DB2-Advisor-Funktion, die Angaben zum Status von DB2-Servern macht, und die Advisor-Funktion ping, die meldet, ob der Server auf ein gesendetes ping antwortet. Eine vollständige Liste der Advisor-Funktionen finden Sie im Abschnitt „Liste der Advisor-Funktionen“ auf Seite 189.

Sie können auch eigene Advisor-Funktionen schreiben. (Lesen Sie hierzu die Informationen im Abschnitt „Kundenspezifische (anpassbare) Advisor-Funktion erstellen“ auf Seite 193.)

Die Benutzung der Advisor-Funktionen ist optional, wird jedoch empfohlen.

- Zum Konfigurieren und Verwalten des Executors, der Advisor-Funktionen und des Managers können Sie die Befehlszeile (**dscontrol**) oder die grafische Benutzerschnittstelle (**lbadmin**) verwenden.
- Für die Konfiguration und Verwaltung der Dispatcher-Maschine steht eine **Beispielkonfigurationsdatei** bereit (siehe Anhang C, „Beispielkonfigurationsdateien“, auf Seite 481). Nach der Installation des Produkts finden Sie diese Datei im Unterverzeichnis **...ibm/edge/lb/servers/samples** des Verzeichnisses mit Load Balancer.
- Mit dem **SNMP-Subagenten** kann eine SNMP-gestützten Verwaltungsanwendung den Status des Dispatchers überwachen.

Die drei Schlüsselfunktionen des Dispatchers (Executor, Manager und Advisor) kommunizieren miteinander, um die eingehenden Anforderungen auf die Server zu verteilen. Neben Lastausgleichsanforderungen überwacht der Executor die Anzahl neuer, aktiver und beendeter Verbindungen. Der Executor übernimmt auch die Garbage Collection für beendete oder zurückgesetzte Verbindungen und stellt diese Informationen dem Manager zur Verfügung.

Der Manager stellt Informationen vom Executor, von den Advisor-Funktionen und von einem Systemüberwachungsprogramm wie Metric Server zusammen. Der Manager passt anhand der erhaltenen Informationen die Wertigkeit der Servermaschinen an den einzelnen Ports an und teilt dem Executor die neue Wertigkeit mit, die dieser dann beim Lastausgleich für neue Verbindungen verwendet.

Die Advisor-Funktionen überwachen die einzelnen Server am zugeordneten Port, um Antwortzeit und Verfügbarkeit der Server zu ermitteln, und übergeben diese Informationen an den Manager.

Die Advisor-Funktionen überwachen zudem, ob ein Server aktiv oder inaktiv ist. Ohne Manager und Advisor-Funktionen wendet der Executor eine RoundRobin-Zeitplanung auf der Basis der aktuellen Serverwertigkeiten an.

Weiterleitungsmethoden

Der Dispatcher stellt drei Weiterleitungsmethoden auf Port-Ebene bereit: MAC-Weiterleitung, NAT/NAPT-Weiterleitung und CBR (inhaltsabhängige Weiterleitung).

Dispatcher-Weiterleitungsmethode mac

Wenn der Dispatcher seine Standardweiterleitungsmethode, die MAC-Weiterleitung, anwendet, werden die eingehenden Anforderungen an den ausgewählten Server weitergeleitet. Der Server gibt die Antwort *direkt*, d. h. ohne Eingreifen des Dispatchers, an den Client zurück. Bei dieser Methode der Weiterleitung achtet der Dispatcher nur auf den beim Server eingehenden Datenfluss vom Client, nicht aber auf den abgehenden Datenfluss vom Server zum Client. Dies führt zu einer erheblichen Reduzierung der Auswirkungen auf die Anwendung und zu einem verbesserten Durchsatz im Netz.

Sie können die Weiterleitungsmethode auswählen, wenn Sie mit dem Befehl **dscontrol port add Cluster:Port method Wert** einen Port hinzufügen. Der Wert für die Standardweiterleitungsmethode ist **mac**. Die Parameter für die Methode können Sie nur beim Hinzufügen des Ports angeben. Ist der Port hinzugefügt, können Sie die Einstellung für die Weiterleitungsmethode nicht mehr ändern. Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol port — Ports konfigurieren“ auf Seite 382.

Einschränkung für Linux: Linux-Systeme nutzen ein hostgestütztes Modell, um Hardwareadressen über ARP für IP-Adressen zugänglich zu machen. Dieses Modell ist nicht mit den Anforderungen der Load-Balancer-Weiterleitungsmethode mac an einen Back-End-Server oder einen verknüpften Server mit hoher Verfügbarkeit kompatibel. Der Abschnitt „Alternativen für die Festlegung eines Loopback-Aliasnamens unter Linux bei Verwendung der Load-Balancer-Weiterleitungsmethode mac“ auf Seite 78 beschreibt eine Reihe von Lösungen, wie Sie das Verhalten des Linux-Systems so ändern können, dass es mit der Load-Balancer-Weiterleitungsmethode mac kompatibel ist.

Einschränkung für Linux auf zSeries- oder S/390-Servern: Wenn Sie zSeries- oder S/390-Server mit OSA-Karten (Open System Adapter) verwenden, gibt es Einschränkungen. Mögliche Strategien zur Lösung von Problemen finden Sie im Abschnitt „Problem: Einschränkungen für die Dispatcher-Konfiguration unter Linux auf zSeries- oder S/390-Servern mit OSA-Karten“ auf Seite 323.

Dispatcher-Weiterleitungsmethode nat

Bei Verwendung der Dispatcher-Methode NAT (Konvertierung von Netzadressen) bzw. NAPT (Port-Umsetzung für Netzadressen) entfällt die Einschränkung, dass sich die am Lastausgleich beteiligten Server in einem lokal angeschlossenen Netz befinden müssen. Falls Sie Server an fernen Standorten haben, sollten Sie an Stelle einer GRE/WAN-Kapselungstechnik die Weiterleitungsmethode nat anwenden. Mit NAPT können Sie außerdem auf mehrere Serverdämonen zugreifen, die sich auf den einzelnen am Lastausgleich beteiligten Servermaschinen befinden und jeweils an einem eindeutigen Port empfangsbereit sind.

Einen Server mit mehreren Dämonen können Sie auf die beiden folgenden Arten konfigurieren:

- Mit NAT können Sie mehrere Serverdämonen für die Beantwortung von Anfragen, die an verschiedene IP-Adressen gerichtet sind, konfigurieren. Dieses Konfiguration wird auch als Bindung eines Serverdämons an eine IP-Adresse bezeichnet.
- Mit NATP können Sie mehrere Serverdämonen (die auf einem physischen Server aktiv sind) so konfigurieren, dass sie an unterschiedlichen Port-Nummern empfangsbereit sind.

Diese Anwendung funktioniert gut mit höheren Anwendungsprotokollen wie HTTP, SSL, IMAP, POP3, NNTP, SMTP, Telnet usw.

Einschränkungen:

- Die NAT/NAPT-Implementierung durch den Dispatcher ist eine *einfache* Implementierung dieser Funktion. Sie analysiert lediglich den Inhalt der Header von TCP/IP-Paketen und kann nur auf diese angewendet werden. Der Inhalt des Datenabschnitts der Pakete kann nicht analysiert werden. Der Dispatcher kann NAT/NAPT nicht für Anwendungsprotokolle wie FTP verwenden, die die Adressen oder Port-Nummern in den Datenabschnitt von Nachrichten einbetten. Dies ist eine allgemein bekannte Einschränkung für die Header-bezogene NAT/NAPT.
- Die NAT/NAPT-Funktion von Dispatcher kann nicht zusammen mit Platzhalter-clustern oder -Ports verwendet werden.

Für die Dispatcher-Maschine benötigen Sie drei IP-Adressen: NFA, Clusteradresse und Rückkehradresse. Sie können NAT/NAPT wie folgt implementieren (vergleichen Sie hierzu auch den Abschnitt „Beispielschritte für das Konfigurieren der Dispatcher-Weiterleitungsmethoden nat und cbr“ auf Seite 57):

- Geben Sie den Befehl **dscontrol executor set** mit dem Parameter **clientgateway** an. Der Parameter **clientgateway** ist eine IP-Adresse, die als Router-Adresse verwendet wird, über die Load Balancer den Antwortdatenverkehr an die Clients weiterleitet. Sie können NAT/NAPT erst verwenden, wenn dieser Wert auf eine IP-Adresse ungleich null gesetzt ist. Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol executor — Executor steuern“ auf Seite 361.
- Fügen Sie mit dem Befehl **dscontrol port add Cluster:Port method Wert** einen Port hinzu. Der Wert für die Weiterleitungsmethode sollte auf **nat** gesetzt werden. Die Parameter für die Methode können Sie nur beim Hinzufügen des Ports angeben. Ist der Port hinzugefügt, können Sie die Einstellung für die Weiterleitungsmethode nicht mehr ändern. Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol port — Ports konfigurieren“ auf Seite 382.

Anmerkung: Wenn Sie die Client-Gateway-Adresse nicht auf einen Wert ungleich null gesetzt haben, kann als Weiterleitungsmethode nur **mac** (MAC-basierte Weiterleitung) angegeben werden.

- Fügen Sie mit den Parametern **mapport**, **returnaddress** und **router** des Befehls **dscontrol** einen Server hinzu. Beispiel:

```
dscontrol server add Cluster:Port:Server mapport Wert returnaddress  
Rückkehradresse router Router-Adresse
```

– **mapport** (optional)

Dieser Parameter ordnet die (für den Dispatcher bestimmte) Nummer des Ziel-Ports für die Client-Anforderung der Nummer des Server-Ports zu, an dem der Dispatcher die Client-Anforderungen verteilt. Mit **mapport** kann

Load Balancer die Anforderung eines Clients an einem Port empfangen und an einen anderen Port der Servermaschine übertragen. Mit `mapport` können Sie den Lastausgleich für die Anforderungen eines Clients auf einer Servermaschine mit mehreren Serverdämonen durchführen. Der Standardwert für `mapport` ist die Nummer des Ziel-Ports für die Client-Anforderung.

- **returnaddress**

Die Rückkehradresse ist eine eindeutige Adresse oder ein Hostname, die bzw. den Sie auf der Dispatcher-Maschine konfigurieren. Der Dispatcher verwendet die Rückkehradresse beim Lastausgleich für die Client-Anforderung auf dem Server als Quellenadresse. Auf diese Weise wird sichergestellt, dass der Server das Paket an die Dispatcher-Maschine zurückgibt und es nicht direkt an den Client sendet. (Der Dispatcher leitet das IP-Paket dann an den Client weiter.) Sie müssen den Wert für die Rückkehradresse beim Hinzufügen des Servers angeben. Die Rückkehradresse kann nur geändert werden, wenn Sie den Server entfernen und dann erneut hinzufügen. Die Rückkehradresse darf nicht mit dem Wert für Cluster, Server oder NFA übereinstimmen.

- **router**

Die Adresse des Routers zum fernen Server. Falls dies ein lokal angeschlossener Server ist, der sich nicht auf derselben Maschine wie Load Balancer befindet, geben Sie die Serveradresse ein. Befindet sich der Server auf derselben Maschine wie Load Balancer, verwenden Sie weiter die reale Router-Adresse.

Weitere Informationen zum Befehl **dscontrol server** mit den Parametern `mapport`, `returnaddress` und `router` finden Sie im Abschnitt „dscontrol server — Server konfigurieren“ auf Seite 394.

Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)

Mit der Komponente Dispatcher können Sie das Content Based Routing für HTTP (unter Verwendung des Regeltyps `content`) und für HTTPS (unter Verwendung der Affinität von SSL-Sitzungs-IDs) ohne Caching Proxy ausführen. Für HTTP- und HTTPS-Datenverkehr ist das Content Based Routing der Dispatcher-Weiterleitungsmethode `cbr` schneller als das der Komponente CBR, die Caching Proxy erfordert.

Für HTTP: Die Serverauswahl für die inhaltsabhängige Weiterleitung basiert auf dem Inhalt eines URL oder eines HTTP-Headers. Sie wird mit dem Regeltyp `content` konfiguriert. Wenn Sie die `content`-Regel konfigurieren, geben Sie für die Regel den Suchbegriff (das Muster) und eine Gruppe von Servern an. Beim Verarbeiten einer neu eingehenden Anforderung vergleicht diese Regel die angegebene Zeichenfolge mit dem URL des Clients oder mit dem in der Client-Anforderung angegebenen HTTP-Header.

Findet der Dispatcher die Zeichenfolge in der Client-Anforderung, leitet er diese an einen der für die Regel definierten Server weiter. Anschließend gibt der Dispatcher die Antwortdaten vom Server an den Client zurück (Weiterleitungsmethode `cbr`).

Findet der Dispatcher die Zeichenfolge nicht in der Client-Anforderung, wählt er *keinen* der für die Regel definierten Server aus.

Anmerkung: Die `content`-Regel wird für die Komponente Dispatcher auf die gleiche Weise wie für die Komponente CBR konfiguriert. Der Dispatcher kann die `content`-Regel für HTTP-Datenverkehr verwenden. Die Komponente CBR kann die `content`-Regel für HTTP- und HTTPS-Datenverkehr (SSL) verwenden.

Für HTTPS (SSL): Bei der inhaltsabhängigen Weiterleitung von Dispatcher erfolgt der Lastausgleich ausgehend vom Feld für die SSL-Sitzungs-ID in der Client-Anforderung. Bei Verwendung von SSL enthält eine Client-Anforderung die SSL-Sitzungs-ID einer früheren Sitzung und Server speichern ihre früheren SSL-Verbindungen im Cache. Durch die Dispatcher-Funktion für Affinität der SSL-Sitzungs-ID können Client und Server eine neue Verbindung aufbauen und dafür die Sicherheitsparameter der vorherigen Verbindung zum Server verwenden. Da SSL-Sicherheitsparameter wie gemeinsam verwendete Schlüssel und Verschlüsselungsalgorithmen nicht neu ausgehandelt werden müssen, benötigen die Server weniger CPU-Zyklen und der Client erhält schneller eine Antwort. Zum Aktivieren der Affinität von SSL-Sitzungs-IDs muss für den Port als **protocol** (Protokolltyp) **SSL** angegeben und **stickytime** auf einen Wert ungleich null gesetzt werden. Wenn die Haltezeit (stickytime) abgelaufen ist, wird der Client unter Umständen an einen anderen als den vorherigen Server verwiesen.

Für die Dispatcher-Maschine benötigen Sie drei IP-Adressen: NFA, Clusteradresse und Rückkehradresse. Sie können die inhaltsabhängige Weiterleitung folgt implementieren (vergleichen Sie hierzu auch den Abschnitt „Beispielschritte für das Konfigurieren der Dispatcher-Weiterleitungsmethoden nat und cbr“ auf Seite 57):

- Geben Sie den Befehl **dscontrol executor set** mit dem Parameter **clientgateway** an. Der Parameter clientgateway ist eine IP-Adresse, die als Router-Adresse verwendet wird, über die der Dispatcher den Antwortdatenverkehr an die Clients weiterleitet. Der Standardwert für clientgateway ist null. Sie können eine CBR-Weiterleitungsmethode erst verwenden, wenn dieser Wert auf eine IP-Adresse ungleich null gesetzt ist. Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol executor — Executor steuern“ auf Seite 361.
- Verwenden Sie den Befehl **dscontrol port add** mit den Parametern **method** und **protocol**, um einen Port hinzuzufügen. Der Wert für die Weiterleitungsmethode sollte auf **cbr** gesetzt werden. Der Protokolltyp für den Port kann HTTP oder SSL sein. Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol port — Ports konfigurieren“ auf Seite 382.

Anmerkung: Wenn Sie die Client-Gateway-Adresse nicht auf einen Wert ungleich null gesetzt haben, kann als Weiterleitungsmethode nur **mac** angegeben werden.

- Fügen Sie mit den Parametern **mapport**, **returnaddress** und **router** einen Server hinzu.

dscontrol server add *Cluster:Port:Server* **mapport** *Wert* **returnaddress** *Rückkehradresse* **router** *Router-Adresse*

Anmerkung: Informationen zum Konfigurieren des Servers mit den Parametern **mapport** (optional), **returnaddress** und **router** finden Sie auf Seite 54.

- **Für HTTP:** Verwenden Sie für die Konfiguration Regeln, die auf dem Inhalt der Client-Anforderung basieren (Regeltyp **content**). Beispiel:

dscontrol rule 125.22.22.03:80:content-Regel1 **type** content **pattern** *Muster*

Muster gibt hier das für den Regeltyp content zu verwendende Muster an. Weitere Informationen zum Regeltyp content finden Sie im Abschnitt „Auf dem Inhalt der Anforderung basierende Regeln verwenden“ auf Seite 219. Weitere Informationen zu gültigen Ausdrücken für *Muster* können Sie Anhang B, „Syntax der content-Regel“, auf Seite 477 entnehmen.

Anmerkung: Die für eine hohe Verfügbarkeit ausgeführte Vervielfältigung von Verbindungseinträgen stellt sicher, dass die Verbindung eines Clients nicht unterbrochen wird, wenn eine Ausweich-Dispatcher-Maschine die Aufgaben der primären Maschine übernimmt. Diese Vervielfältigung wird bei der inhaltsabhängigen Weiterleitung durch den Dispatcher *nicht* unterstützt.

Beispielschritte für das Konfigurieren der Dispatcher-Weiterleitungsmethoden nat und cbr

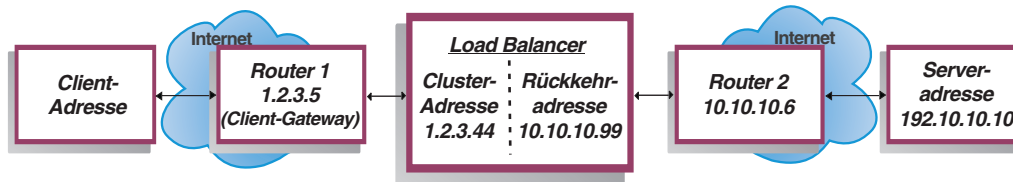


Abbildung 12. Beispiel für die Dispatcher-Weiterleitungsmethoden nat und cbr

Sie benötigen mindestens drei IP-Adressen für die Dispatcher-Maschine. Für das Beispiel in Abb. 12 sind die folgenden Schritte notwendig, um eine Minimal-konfiguration der Dispatcher-Weiterleitungsmethoden nat und cbr zu erzielen:

1. Starten Sie den Executor.
`dscontrol executor start`
2. Definieren Sie das Client-Gateway.
`dscontrol executor set clientgateway 1.2.3.5`
ANMERKUNG: Wenn es in Ihrem Teilnetz keinen lokalen Router gibt, müssen Sie eine Maschine für die IP-Weiterleitung konfigurieren und diese als Client-Gateway verwenden. Anweisungen für das Aktivieren der IP-Weiterleitung finden Sie in der Dokumentation zum Betriebssystem.
3. Definieren Sie die Clusteradresse.
`dscontrol cluster add 1.2.3.44`
4. Konfigurieren Sie die Clusteradresse.
`dscontrol executor configure 1.2.3.44`
5. Definieren Sie den Port mit der Methode nat oder cbr.
`dscontrol port add 1.2.3.44:80 method nat`
 oder
`dscontrol port add 1.2.3.44:80 method cbr protocol http`
6. Konfigurieren Sie eine Aliasrückkehradresse für Load Balancer (mit Ethernet-Karte 0).
Anmerkung: Auf Linux-Systemen benötigen Sie keine Aliasrückkehradresse, wenn Sie auf einer verknüpften Maschine die Weiterleitungsmethode nat verwenden.
`dscontrol executor configure 10.10.10.99`

 Sie können auch den Befehl `ifconfig` verwenden (nur für Linux oder UNIX):
 AIX: `ifconfig en0 alias 10.10.10.99 netmask 255.255.255.0`
 HP-UX: `ifconfig lan0:1 10.10.10.99 netmask 255.255.255.0 up`
 Linux: `ifconfig eth0:1 10.10.10.99 netmask 255.255.255.0 up`
 Solaris: `ifconfig eri0 addif 10.10.10.99 netmask 255.255.255.0 up`
7. Definieren Sie die Back-End-Server.
`dscontrol server add 1.2.3.4:80:192.10.10.10`
`router 10.10.10.6 returnaddress 10.10.10.99`

Das Client-Gateway (1.2.3.5) ist die Adresse für Router 1 zwischen Load Balancer und dem Client. Der Router (10.10.10.6) ist die Adresse für Router 2 zwischen Load Balancer und dem Back-End-Server. Falls Sie die Adresse für das Client-Gateway und für Router 2 nicht kennen, können Sie ein Routenverfolgungsprogramm (traceroute) mit der Client-Adresse (oder Serveradresse) verwenden, um die Router-Adresse zu bestimmen. Die genaue Syntax des Programms richtet sich nach dem von Ihnen verwendeten Betriebssystem. Weitere Informationen zu diesem Programm können Sie der Dokumentation zum Betriebssystem entnehmen.

Wenn sich der Server in demselben Teilnetz wie Load Balancer befindet (so dass traceroute keine Router zurückgibt), geben Sie die Serveradresse als Router-Adresse ein. Befindet sich der Server jedoch auf derselben Maschine wie Load Balancer, sollten Sie im Feld für die Router-ID an Stelle der Serveradresse die Router-Adresse eingeben. Die Router-Adresse ist die in Schritt 7 auf der Load-Balancer-Maschine für den Befehl "server add" verwendete Adresse.

Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)

Bei Anwendung der Serverpartitionierung können Sie zwischen URLs und ihren spezifischen Anwendungen unterscheiden. Ein Webserver kann beispielsweise JSPs, HTML-Seiten und GIF-Dateien bereitstellen, Datenbankabfragen bedienen usw. Load Balancer bietet jetzt die Möglichkeit, einen cluster- und Port-spezifischen Server in mehrere logische Server zu partitionieren. Dadurch können Sie einen bestimmten Dienst auf der Maschine anweisen festzustellen, ob eine Servlet-Steuerkomponente oder eine Datenbankabfrage schneller oder gar nicht ausgeführt wird.

Mit der Serverpartitionierung kann Load Balancer z. B. erkennen, dass der HTML-Dienst Seiten schnell bereitstellt, die Datenbankverbindung jedoch nicht mehr aktiv ist. Dadurch können Sie die Last mit größerer detaillierter und dienstspezifisch verteilen und müssen sich nicht auf die Wertigkeit des Gesamtserver verlassen.

Serverpartitionierung mit der HTTP- oder HTTPS-Advisor-Funktion

Die Serverpartitionierung kann in Verbindung mit der HTTP- und der HTTPS-Advisor-Funktion hilfreich sein. Wenn Sie beispielsweise einen HTML-Server für HTML- und GIF-Seiten sowie JSPs haben und den Server einmal unter Port 80 definieren (hinzufügen), erhalten Sie für den gesamten HTTP-Server nur einen Lastwert. Dies kann irreführend sein, weil es möglich ist, dass der GIF-Service auf dem Server nicht funktioniert. Der Dispatcher leitet GIF-Seiten unverändert an den Server weiter. Der Client sieht jedoch eine Zeitlimitüberschreitung oder einen Ausfall.

Wenn Sie den Server dreimal unter dem Port definieren (ServerHTML, ServerGIF, ServerJSP) und den Serverparameter **advisorrequest** für jeden logischen Server mit einer anderen Zeichenfolge definieren, können Sie den Zustand des jeweiligen Service auf dem Server abfragen. ServerHTML, ServerGIF und ServerJSP repräsentieren die drei logischen Server, die durch die Partitionierung eines physischen Servers entstanden sind. Für ServerJSP können Sie eine Zeichenfolge für advisorrequest definieren, um auf der Maschine zur Bearbeitung von JSPs den Service abzufragen. Für ServerGIF können Sie die Zeichenfolge advisorrequest so definieren, dass der GIF-Service abgefragt wird. Und für ServerHTML können Sie advisorrequest so definieren, dass der HTML-Service abgefragt wird. Empfängt der Client keine Antwort von advisorrequest zur Abfrage des GIF-Services, registriert der Dispatcher, dass der logische Server (ServerGIF) inaktiv ist, die beiden anderen

logischen Server jedoch weiterhin in gutem Zustand sein können. Der Dispatcher leitet keine weiteren GIFs an den physischen Server weiter, sendet jedoch unverändert JSP- und HTML-Anfragen an den Server.

Weitere Informationen zum Parameter **advisorrequest** finden Sie im Abschnitt „Option 'Anforderung/Antwort (URL)' der HTTP- oder HTTPS-Advisor-Funktion konfigurieren“ auf Seite 191.

Beispiel für das Konfigurieren von logischen Servern auf einem physischen Server

Innerhalb der Dispatcher-Konfiguration können Sie einen physischen oder logischen Server mit der Hierarchie *Cluster:Port:Server* darstellen. Der Server kann eine eindeutige IP-Adresse der Maschine (physischer Server) sein, die als symbolischer Name oder im IP-Adressenformat angegeben wird. Wenn Sie den Server als partitionierten Server definieren, müssen Sie den Befehl **dscontrol server add** mit einer auflösbaren Serveradresse des physischen Servers für den Parameter **address** angeben. Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol server — Server konfigurieren“ auf Seite 394.

Nachfolgend sehen Sie ein Beispiel für die Partitionierung physischer Server in logische Server zur Bearbeitung von Anforderungen verschiedenen Typs.

```
Cluster: 1.1.1.1
  Port: 80
    Server: A (IP-Adresse 1.1.1.2)
             HTML-Server
    Server: B (IP-Adresse 1.1.1.2)
             GIF-Server
    Server: C (IP-Adresse 1.1.1.3)
             HTML-Server
    Server: D (IP-Adresse 1.1.1.3)
             JSP-Server
    Server: E (IP-Adresse 1.1.1.4)
             GIF-Server
    Server: F (IP-Adresse 1.1.1.4)
             JSP-Server
  Rule1: /*.htm
    Server: A
    Server: C
  Rule2: /*.jsp
    Server: D
    Server: F
  Rule3: /*.gif
    Server: B
    Server: E
```

In diesem Beispiel wird der Server 1.1.1.2 in zwei logische Server partitioniert: "A" (zur Bearbeitung von HTML-Anforderungen) und "B" (zur Bearbeitung von GIF-Anforderungen). Server 1.1.1.3 wird in zwei logische Server partitioniert: "C" (zur Bearbeitung von HTML-Anforderungen) und "D" (zur Bearbeitung von JSP-Anforderungen). Server 1.1.1.4 wird in zwei logische Server partitioniert: "E" (zur Bearbeitung von GIF-Anforderungen) und "F" (zur Bearbeitung von JSP-Anforderungen).

Hohe Verfügbarkeit

Einfache hohe Verfügbarkeit

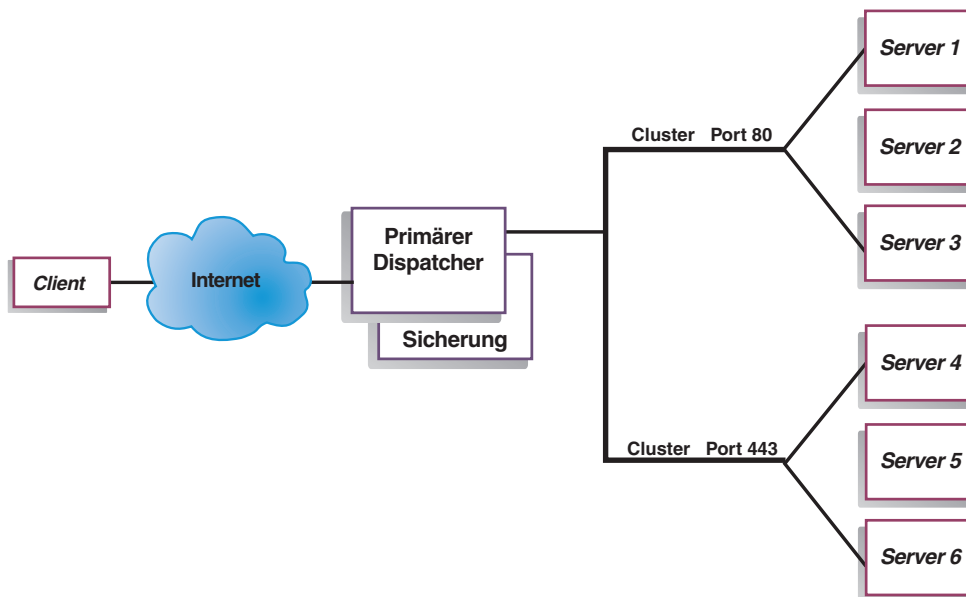


Abbildung 13. Beispiel für einen Dispatcher mit einfacher hoher Verfügbarkeit

Die Funktion für hohe Verfügbarkeit erfordert eine zweite Dispatcher-Maschine. Die erste Dispatcher-Maschine führt den Lastausgleich für den gesamten Client-Datenverkehr aus, wie dies in einer Konfiguration mit einem einzelnen Dispatcher geschehen würde. Die zweite Dispatcher-Maschine überwacht den "Zustand" der ersten Maschine und übernimmt die Task des Lastausgleichs, wenn sie erkennt, dass die erste Dispatcher-Maschine ausgefallen ist.

Jeder der beiden Maschinen wird eine bestimmte Rolle zugewiesen, entweder die der primären Maschine (*primary*) oder die der Ausweichmaschine (*backup*). Die primäre Maschine sendet ständig Verbindungsdaten an die Partnermaschine. Während die primäre Maschine *aktiv* ist (den Lastausgleich durchführt), befindet sich die Partnermaschine in *Bereitschaft*. Sie wird ständig aktualisiert und ist bereit, den Lastausgleich zu übernehmen, falls dies erforderlich ist.

In den Übertragungssitzungen zwischen den beiden Maschinen werden *Überwachungssignale* ausgetauscht. Mit Hilfe der Überwachungssignale kann jede Maschine den Zustand der anderen Maschine überwachen.

Stellt die Ausweichmaschine fest, dass die aktive Maschine ausgefallen ist, übernimmt sie deren Aufgaben und beginnt mit dem Lastausgleich. An diesem Punkt kehrt sich der *Status* der beiden Maschinen um: die Partnermaschine wird zur *aktiven Maschine* und die primäre Maschine wird zur *Maschine in Bereitschaft*.

In der Konfiguration mit hoher Verfügbarkeit müssen sich die primäre Maschine und die Partnermaschine innerhalb eines Teilnetzes befinden und identisch konfiguriert sein.

Informationen zum Konfigurieren der hohen Verfügbarkeit finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 204.

Gegenseitige hohe Verfügbarkeit

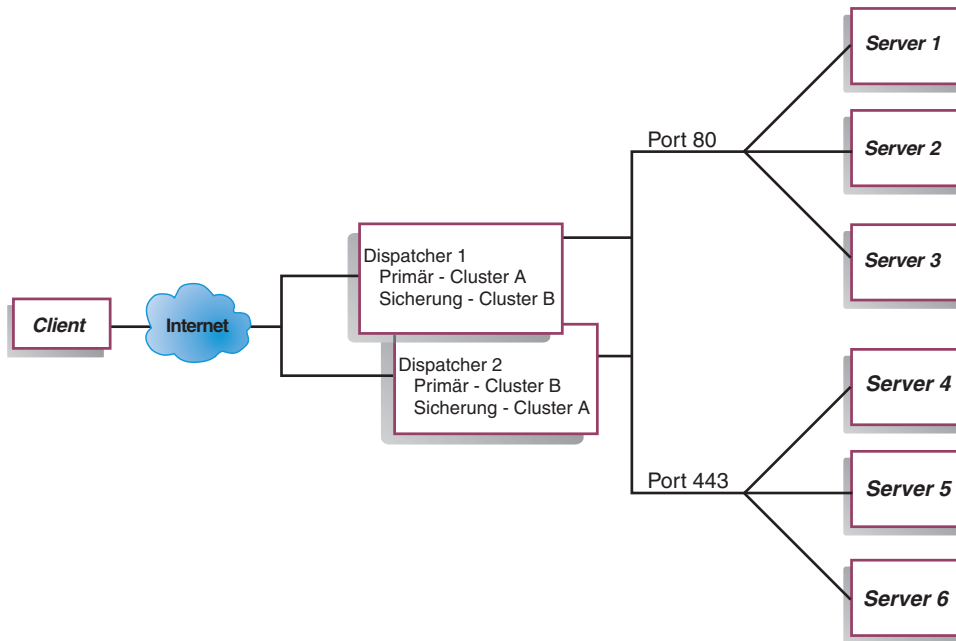


Abbildung 14. Beispiel für einen Dispatcher mit gegenseitiger hoher Verfügbarkeit

Für die gegenseitige hohe Verfügbarkeit sind zwei Dispatcher-Maschinen erforderlich. Beide Maschinen führen aktiv den Lastausgleich des Client-Datenverkehrs aus und sind gleichzeitig Partnermaschinen. In einer Konfiguration mit einfacher hoher Verfügbarkeit führt nur eine Maschine den Lastausgleich durch. In einer Konfiguration mit gegenseitiger hoher Verfügbarkeit verteilen beide Maschinen einen Teil des Client-Datenverkehrs.

Bei der gegenseitigen hohen Verfügbarkeit wird der Client-Datenverkehr den Dispatcher-Maschinen auf der Basis einer Clusteradresse zugeordnet. Jeder Cluster kann mit der nicht für Weiterleitungszwecke bestimmten Adresse (NFA, nonforwarding Address) seines primären Dispatchers konfiguriert werden. Die primäre Dispatcher-Maschine führt normalerweise den Lastausgleich für diesen Cluster durch. Fällt die Maschine aus, führt die andere Maschine den Lastausgleich für ihren eigenen Cluster und für den Cluster des ausgefallenen Dispatchers durch.

Abb. 14 zeigt eine Beispielskonfiguration mit gegenseitiger hoher Verfügbarkeit, bei der die "Clustergruppe A" und die "Clustergruppe B" gemeinsam benutzt werden. Jeder Dispatcher kann aktiv für seinen *primären* Cluster bestimmte Pakete weiterleiten. Fällt einer der Dispatcher aus, so dass er nicht länger aktiv für seinen primären Cluster bestimmte Pakete weiterleiten kann, übernimmt der andere Dispatcher die Weiterleitung der Pakete zu seinem *Ausweichcluster*.

Anmerkung: Auf beiden Maschinen müssen die gemeinsam benutzten Clustergruppen identisch konfiguriert werden. Das heißt, die in beiden Konfigurationen verwendeten Ports und die unter den Ports konfigurierten Server müssen identisch sein.

Informationen zum Konfigurieren der hohen Verfügbarkeit und der gegenseitigen hohen Verfügbarkeit finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 204.

Kapitel 7. Dispatcher konfigurieren

Lesen Sie vor Ausführung der in diesem Kapitel beschriebenen Schritte Kapitel 6, „Planung für Dispatcher“, auf Seite 51. Dieses Kapitel erläutert das Erstellen einer Basiskonfiguration für die Komponente Dispatcher von Load Balancer.

- Falls Sie Load Balancer für IPV4 und IPV6 verwenden, lesen Sie Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.
- Komplexere Konfigurationen für Load Balancer finden Sie in Kapitel 21, „Manager, Advisor-Funktionen und Metric Server für Dispatcher, CBR und Site Selector“, auf Seite 179 und in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Anmerkung: Frühere Versionen dieses Produkts liefen unter dem Namen Network Dispatcher. In diesen Versionen war der Dispatcher-Steuerbefehl `ndcontrol`. Jetzt ist der Dispatcher-Steuerbefehl **`dscontrol`**.

Konfigurations-Tasks im Überblick

Wichtiger Hinweis: Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, lesen Sie die Informationen in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

Vergewissern Sie sich vor Ausführung der Konfigurationsschritte in dieser Tabelle, dass die Dispatcher-Maschine und alle Servermaschinen mit dem Netz verbunden sind, gültige IP-Adressen haben und sich gegenseitig mit ping-Aufrufen erreichen können.

Tabelle 4. Konfigurations-Tasks für Dispatcher

Task	Beschreibung	Referenzinformationen
Dispatcher-Maschine konfigurieren.	Definieren Sie Ihre Lastausgleichskonfiguration.	„Dispatcher-Maschine konfigurieren“ auf Seite 66
Am Lastausgleich beteiligte Maschinen konfigurieren.	Definieren Sie einen Aliasnamen auf der Loopback-Einheit. Überprüfen Sie, ob eine zusätzliche Route vorhanden ist und löschen Sie zusätzliche Routen.	„Servermaschinen für Lastausgleich konfigurieren“ auf Seite 72

Konfigurationsmethoden

Es gibt vier grundlegende Methoden für die Konfiguration des Dispatchers:

- Befehlszeile
- Scripts
- grafische Benutzerschnittstelle (GUI)
- Konfigurationsassistent.

Befehlszeile

Dies ist die direkte Methode für die Konfiguration des Dispatchers. Die Werte der Befehlsparameter müssen in englischen Zeichen eingegeben werden. Die einzige Ausnahme hiervon bilden Hostnamen (die in den Befehlen `cluster`, `server` und `highavailability` verwendet werden) und (die in Dateibefehlen verwendeten) Dateinamen.

Starten Sie den Dispatcher wie folgt von der Befehlszeile aus:

1. Setzen Sie an der Eingabeaufforderung den Befehl **dsserver** ab. Geben Sie zum Stoppen des Services **dsserver stop** ein.

Anmerkung: Klicken Sie auf Windows-Systemen nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf **IBM Dispatcher** und wählen Sie **Starten** aus. Zum Stoppen des Services müssen Sie dieselben Schritte ausführen und **Beenden** auswählen.

2. Setzen Sie anschließend Dispatcher-Steuerbefehle ab, um die Konfiguration einzurichten. Bei den in diesem Handbuch beschriebenen Prozeduren wird von der Verwendung der Befehlszeile ausgegangen. Der Befehl lautet **dscontrol**. Weitere Informationen zu Befehlen finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Für die Parameter des Befehls `dscontrol` können Sie die Kurzform verwenden. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **dscontrol he f** an Stelle von **dscontrol help file** angeben.

Wenn Sie die Befehlszeilenschnittstelle starten möchten, setzen Sie den Befehl **dscontrol** ab, um die Eingabeaufforderung `dscontrol` aufzurufen.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie den Befehl **exit** oder **quit** absetzen.

Scripts

Die Befehle zum Konfigurieren von Dispatcher können in eine Konfigurations-Script-Datei eingegeben und dann zusammen ausgeführt werden. Lesen Sie hierzu die Informationen im Abschnitt „Beispielkonfigurationsdateien für Load Balancer“ auf Seite 481.

Anmerkung: Falls Sie den Inhalt einer Script-Datei (z. B. `meinScript`) schnell ausführen möchten, verwenden Sie einen der folgenden Befehle:

- Führen Sie zum Aktualisieren der derzeitigen Konfiguration die folgenden ausführbaren Befehle Ihrer Script-Datei aus:
dscontrol file appendload *meinScript*
- Führen Sie zum vollständigen Ersetzen der derzeitigen Konfiguration die folgenden ausführbaren Befehle Ihrer Script-Datei aus:
dscontrol file newload *meinScript*

Führen Sie den folgenden Befehl aus, um die aktuelle Konfiguration in einer Script-Datei (z. B. `sicherungsscript`) zu speichern:

dscontrol file save *sicherungsscript*

Dieser Befehl speichert die Script-Datei mit der Konfiguration im Verzeichnis `...ibm/edge/lb/servers/configurations/dispatcher`.

GUI

Abb. 41 auf Seite 470 zeigt ein Beispiel für die grafische Benutzerschnittstelle (GUI) mit allgemeinen Anweisungen.

Gehen Sie zum Starten der GUI wie folgt vor:

1. Vergewissern Sie sich, dass `dsserver` aktiv ist.
 - Führen Sie auf AIX-, HP-UX-, Linux- oder Solaris-Systemen den folgenden Befehl als Benutzer `root` aus:
dsserver
 - Auf Windows-Systemen ist `dsserver` ein Dienst, der automatisch gestartet wird.
2. Führen Sie je nach Betriebssystem eine der folgenden Aktionen aus:
 - Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Geben Sie **lbadm** ein.
 - Auf Windows-Systemen: Klicken Sie nacheinander auf **Start > Programme > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**.

Zum Konfigurieren von Dispatcher auf der GUI müssen Sie zunächst in der Baumstruktur **Dispatcher** auswählen. Sie können den Executor und den Manager starten, nachdem Sie eine Verbindung zu einem Host hergestellt haben. Sie können auch Cluster mit Ports und Servern erstellen und Advisor-Funktionen für den Manager starten.

Mit der GUI können Sie dieselben Tasks wie mit dem Befehl **dscontrol** ausführen. Wenn Sie beispielsweise einen Cluster von der Befehlszeile aus konfigurieren möchten, müssten Sie den Befehl **dscontrol cluster add Cluster** eingeben. Zum Definieren eines Clusters von der GUI aus müssen Sie mit der rechten Maustaste auf "Executor" klicken und im daraufhin angezeigten Popup-Menü mit der linken Taste auf **Cluster hinzufügen**. Geben Sie die Clusteradresse in das Dialogfenster ein und klicken Sie dann auf **OK**.

Bereits vorhandene Dispatcher-Konfigurationsdateien können Sie mit der im Popup-Menü **Host** angezeigten Option **Neue Konfiguration laden** (zum vollständigen Ersetzen der derzeitigen Konfiguration) oder **An aktuelle Konfiguration anfügen** (zum Aktualisieren der derzeitigen Konfiguration) laden. Sie sollten Ihre Dispatcher-Konfiguration von Zeit zu Zeit mit der Option **Konfigurationsdatei sichern unter** in einer Datei sichern. Diese Option ist ebenfalls im Popup-Menü **Host** enthalten. Das oben auf der GUI befindliche Menü **Datei** bietet Ihnen die Möglichkeit, die aktuellen Hostverbindungen in einer Datei zu speichern oder Verbindungen aus vorhandenen Dateien für alle Komponenten von Load Balancer wiederherzustellen.

Die Konfigurationsbefehle können auch auf einem fernen System ausgeführt werden. Weitere Informationen hierzu finden Sie im Abschnitt „Remote Method Invocation (RMI)“ auf Seite 264.

Wenn Sie von der GUI aus einen Befehl ausführen möchten, gehen Sie wie folgt vor: Heben Sie in der GUI-Baumstruktur den Hostknoten hervor und wählen Sie im Popup-Menü "Host" **Befehl senden...** aus. Geben Sie im Befehlseingabefeld den gewünschten Befehl ein, z. B. **executor report**. In einem Fenster sehen Sie die Ergebnisse und die Historie der in der aktuellen Sitzung ausgeführten Befehle.

Sie können auf **Hilfe** zugreifen, indem Sie auf das Fragezeichen in der oberen rechten Ecke des Fensters von Load Balancer klicken.

- **Hilfe: Feldebene** — beschreibt jedes Feld und gibt die Standardwerte an.
- **Hilfe: Wie funktioniert das** — listet Tasks auf, die von dieser Anzeige aus ausgeführt werden können.
- **InfoCenter** — ermöglicht den zentralen Zugriff auf Produktinformationen.

Weitere Informationen zur Verwendung der GUI finden Sie in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Konfiguration mit dem Konfigurationsassistenten

Führen Sie folgende Schritte aus, wenn Sie den Konfigurationsassistenten verwenden:

1. Starten Sie wie folgt dserver für den Dispatcher:
 - Führen Sie auf AIX-, HP-UX-, Linux- oder Solaris-Systemen den folgenden Befehl als Benutzer root aus:
dserver
 - Auf Windows-Systemen ist dserver ein Dienst, der automatisch gestartet wird.
2. Starten Sie den Assistenten des Dispatchers, **dswizard**.

Der Assistent führt Sie schrittweise durch den Prozess zum Erstellen einer Basis-konfiguration für die Komponente Dispatcher. Der Assistent stellt Ihnen Fragen zu Ihrem Netz. Sie erhalten eine Anleitung für die Konfiguration eines Clusters, bei der der Dispatcher den Datenverkehr auf eine Gruppe von Servern verteilt.

Dispatcher-Maschine konfigurieren

Vor dem Konfigurieren der Dispatcher-Maschine müssen Sie (auf AIX-, HP-UX-, Linux- oder Solaris-Systemen) als Benutzer root oder (auf Windows-Systemen) als Administrator registriert sein.

Der Load Balancer kann auf allen unterstützten Plattformen einen **verknüpften** Server haben. Dies bedeutet, dass sich Load Balancer physisch auf einer Servermaschine befinden kann, für die er einen Lastausgleich durchführt.

Wenn Sie die Weiterleitungsmethode mac anwenden möchten, benötigen Sie für die Dispatcher-Maschine mindestens zwei gültige IP-Adressen. Bei Verwendung der Weiterleitungsmethode cbr oder nat sind mindestens drei gültige IP-Adressen erforderlich:

- Eine IP-Adresse speziell für die Dispatcher-Maschine
Diese IP-Adresse ist die primäre IP-Adresse der Dispatcher-Maschine und wird als NFA (Non-Forwarding Address) bezeichnet. Dies ist standardmäßig dieselbe Adresse wie die vom Befehl **hostname** zurückgegebene. Benutzen Sie diese Adresse, wenn Sie zu Verwaltungszwecken (z. B. für eine Fernkonfiguration über Telnet oder für den Zugriff auf den SNMP-Subagenten) eine Verbindung zur Maschine herstellen möchten. Kann die Dispatcher-Maschine bereits andere Maschinen im Netz über ping-Aufrufe erreichen, sind keine weiteren Aktionen zum Konfigurieren der NFA erforderlich.
- Eine IP-Adresse pro Cluster
Eine Clusteradresse ist eine Adresse, die einem Hostnamen zugeordnet ist (beispielsweise www.IhreFirma.com). Diese IP-Adresse wird von einem Client

benutzt, um die Verbindung zu den Servern in einem Cluster herzustellen. An dieser Adresse führt der Dispatcher den Lastausgleich durch.

- Für die Weiterleitung mit cbr oder nat eine IP-Adresse für die Rückkehradresse
Der Dispatcher verwendet die Rückkehradresse beim Lastausgleich für die Client-Anforderung auf dem Server als Quellenadresse. Auf diese Weise wird sichergestellt, dass der Server das Paket an die Dispatcher-Maschine zurückgibt und es nicht direkt an den Client sendet. (Der Dispatcher leitet das IP-Paket dann an den Client weiter.) Sie müssen den Wert für die Rückkehradresse beim Hinzufügen des Servers angeben. Die Rückkehradresse kann nur geändert werden, wenn Sie den Server entfernen und dann erneut hinzufügen.

Nur Solaris-Systeme:

- Der Dispatcher ist standardmäßig für den Lastausgleich für Datenverkehr auf 100-Mbit/s-Ethernet-Netzschnittstellenkarten konfiguriert. Der standardmäßige 100-Mbit/s-Ethernet-Adapter ist in der Datei `ibmlb.conf` als `eri` angegeben. Es werden aber auch andere Arten von Schnittstellenkarten unterstützt, unter anderem `le`, `ce`, `ge`, `hme`, `eri`, `bge`, `vge`, `qfe`, `dfme`, `fji` und `fje`.

Editieren Sie beispielsweise die Datei `/opt/ibm/edge/lb/servers/ibmlb.conf` wie folgt, um die Standardeinstellung zu ändern:

- Wenn Sie einen 10-Mbit/s-Ethernet-Adapter verwenden, ersetzen Sie `eri` durch `le`.
- Für einen 1-Gbit/s-Ethernet-Adapter müssen Sie `eri` durch `ge` ersetzen.
- Wenn Sie einen Adapter mit mehreren Anschlüssen verwenden, ersetzen Sie `eri` durch `qfe`.

Sollen mehrere Adaptertypen unterstützt werden, kopieren Sie die Zeile in der Datei `ibmlb.conf` und passen Sie die einzelnen Zeilen an den Einheitentyp an.

Wenn Sie vorhaben, zwei 100-Mbit/s-Ethernet-Adapter zu verwenden, muss die Datei `ibmlb.conf` eine Zeile mit der Einheitenangabe `eri` enthalten.

Falls Sie einen 10-Mbit/s-Ethernet-Adapter und einen 100-Mbit/s-Ethernet-Adapter verwenden möchten, müssen Sie in der Datei `ibmlb.conf` zwei Zeilen angeben, eine Zeile für die Einheit `le` und eine für die Einheit `eri`.

Anmerkung: Die Datei `ibmlb.conf` stellt Vorgaben für den Solaris-Befehl **auto-push** bereit und muss mit dem Befehl `autopush` kompatibel sein.

- Setzen Sie an der Solaris-Eingabeaufforderung den folgenden Befehl ab, um den Typ der Ethernet-Netzschnittstelle Ihrer Maschine zu bestimmen:

```
ifconfig -a
```

Überprüfen Sie, ob die Ausgabe wie folgt aussieht:

```
lo0: flags=2001000849<UP,LOOPBACK,RUNNING,MULTICAST,IPv4,VIRTUAL>
    mtu 8232 index 1 inet 127.0.0.1 netmask ff000000
eri0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4>
    mtu 1500 index 2 inet 9.42.93.208
    netmask fffffc00 broadcast 9.42.95.255 ether 0:3:ba:2d:24:45
```

Bei einer solchen Ausgabe müssen Sie die Datei `ibmlb.conf` wie folgt editieren:

```
eri -1 0 ibmlb
```

- Beim Starten oder Stoppen des Dispatcher-Executors werden alle Aliasnamen für die in der Datei `ibmlb.conf` aufgelisteten Adapter aus der Konfiguration entfernt. Wenn Sie die Aliasnamen auf diesen Adaptern (mit Ausnahme der von der Komponente Dispatcher von Load Balancer verwendeten Adapter) automatisch neu konfigurieren möchten, verwenden Sie die Script-Datei **goAliases**. Im Verzeichnis `.../ibm/edge/lb/servers/samples` finden Sie ein Beispiel-Script, das Sie

vor der Ausführung in das Verzeichnis `...ibm/edge/lb/servers/bin` verschieben *müssen*. Das Script `goAliases` wird beim Starten oder Stoppen des Dispatcher-Executors automatisch ausgeführt.

Sind die beiden Cluster X und Y für einen der in `ibmlb.conf` aufgelisteten Adapter beispielsweise für die Komponente CBR konfiguriert, werden die Cluster X und Y aus der Konfiguration entfernt, sobald der Befehl **`dscontrol executor start`** oder **`dscontrol executor stop`** abgesetzt wird. Dieses Ergebnis ist unter Umständen nicht erwünscht. Wenn die Cluster X und Y im Script `goAliases` konfiguriert sind, werden Sie nach dem Starten oder Stoppen des Dispatcher-Executors automatisch rekonfiguriert.

Vergewissern Sie sich, dass die IP-Weiterleitung für das TCP/IP-Protokoll nicht aktiviert ist.

Abb. 15 zeigt ein Beispiel für einen mit einem Cluster, zwei Ports und drei Servern konfigurierten Dispatcher.

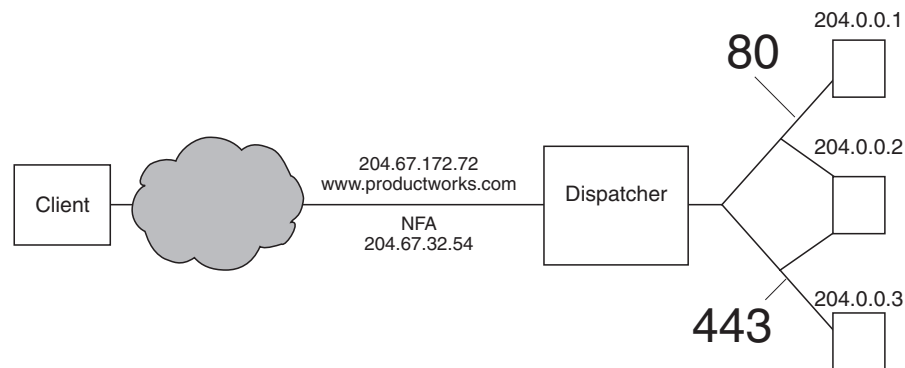


Abbildung 15. Beispiel der für die Dispatcher-Maschine erforderlichen IP-Adressen

Hilfe zu den in dieser Prozedur verwendeten Befehlen finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Eine Beispielkonfigurationsdatei finden Sie im Abschnitt „Beispielkonfigurationsdateien für Load Balancer“ auf Seite 481.

Schritt 1. Serverfunktion starten

AIX-, HP-UX-, Linux- oder Solaris-Systeme: Geben Sie zum Starten der Serverfunktion **`dsserver`** ein.

Windows-Systeme: Die Serverfunktion wird automatisch als Dienst gestartet.

Anmerkung: Eine Standardkonfigurationsdatei (`default.cfg`) wird beim Starten von `dsserver` automatisch geladen. Entscheidet der Benutzer, dass die Dispatcher-Konfiguration in `default.cfg` gesichert werden soll, werden alle in dieser Datei gesicherten Daten automatisch geladen, wenn `dsserver` das nächste Mal gestartet wird.

Schritt 2. Executor-Funktion starten

Geben Sie zum Starten der Executor-Funktion den Befehl **`dscontrol executor start`** ein. Sie können jetzt auch verschiedene Executor-Einstellungen ändern. Weitere Informationen hierzu finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Schritt 3. NFA definieren (falls vom Hostnamen abweichend)

Die NFA wird benutzt, um zu Verwaltungszwecken (z. B. für die Verwendung von Telnet oder SMTP) eine Verbindung zur Maschine herzustellen. Standardmäßig ist diese Adresse der Hostname.

Geben Sie zum Definieren der NFA den Befehl **dscontrol executor set nfa** *IP-Adresse* ein oder editieren Sie die Beispielkonfigurationsdatei. Die *IP-Adresse* ist entweder ein symbolischer Name oder die IP-Adresse.

Schritt 4. Cluster definieren und Clusteroptionen festlegen

Der Dispatcher verteilt die Last der an die Clusteradresse gesendeten Anforderungen auf die für die Ports dieses Clusters konfigurierten Server.

Der Cluster ist entweder der symbolische Name, die Adresse in Schreibweise mit Trennzeichen oder die spezielle Adresse 0.0.0.0, die einen Platzhaltercluster definiert. Setzen Sie zum Definieren eines Clusters den Befehl **dscontrol cluster add** ab. Setzen Sie zum Definieren von Clusteroptionen den Befehl **dscontrol cluster set** ab oder verwenden Sie die GUI zum Absetzen von Befehlen. Platzhaltercluster können verwendet werden, wenn mehrere IP-Adressen für den Lastausgleich eingehender Pakete in Frage kommen. Weitere Informationen hierzu finden Sie in den Abschnitten „Platzhaltercluster zum Zusammenfassen von Serverkonfigurationen verwenden“ auf Seite 237, „Platzhaltercluster für den Lastausgleich von Firewalls verwenden“ auf Seite 238 und „Platzhaltercluster mit Caching Proxy für transparente Weiterleitung verwenden“ auf Seite 239.

Schritt 5. Aliasnamen auf der Netzschnittstellenkarte erstellen

Nachdem der Cluster definiert wurde, müssen Sie normalerweise die Clusteradresse auf einer der Netzschnittstellenkarten der Dispatcher-Maschine konfigurieren. Setzen Sie dazu den Befehl **dscontrol executor configure** *Clusteradresse* ab. Damit wird nach einem Adapter mit einer vorhandenen Adresse gesucht, die zu demselben Teilnetz wie die Clusteradresse gehört. Anschließend wird der Adapterkonfigurationsbefehl des Betriebssystems für die Clusteradresse unter Verwendung des gefundenen Adapters und der Netzmaske für die auf diesem Adapter vorhandene Adresse abgesetzt. Beispiel:

```
dscontrol executor configure 204.67.172.72
```

In manchen Fällen soll die Clusteradresse möglicherweise nicht konfiguriert werden. Dies gilt für Cluster, die zu einem Bereitschaftsserver im Modus für hohe Verfügbarkeit hinzugefügt wurden, oder für Cluster, die zu einem Weitverkehrs-Dispatcher hinzugefügt wurden, der als ferner Server dient. Sie müssen den Befehl **executor configure** auch nicht ausführen, wenn Sie im Standalone-Modus das Beispiel-Skript **goldle** verwenden. Informationen zum Skript **goldle** finden Sie im Abschnitt „Scripts verwenden“ auf Seite 209.

In seltenen Fällen haben Sie möglicherweise eine Clusteradresse, die mit keinem Teilnetz für vorhandene Adressen übereinstimmt. Verwenden Sie in diesem Fall die zweite Form des Befehls **executor configure** und geben Sie explizit den Schnittstellennamen und die Netzmaske an. Verwenden Sie **dscontrol executor configure** *Clusteradresse Schnittstellename Netzmaske*.

Beispiele:

```
dscontrol executor configure 204.67.172.72 en0 255.255.0.0
(AIX-Systeme)
dscontrol executor configure 204.67.172.72 eth0:1 255.255.0.0
(Linux-Systeme)
```

```
dscontrol executor configure 204.67.172.72 eri0 255.255.0.0
(Solaris-Systeme)
dscontrol executor configure 204.67.172.72 en1 255.255.0.0
(Windows-Systeme)
```

Windows-Systeme

Für die zweite Form des Befehls `executor configure` müssen Sie auf Windows-Systemen den zu verwendenden Schnittstellennamen ermitteln. Befindet sich in Ihrer Maschine nur eine einzige Ethernet-Karte, lautet der Schnittstellename `en0`. Befindet sich in Ihrer Maschine nur eine einzige Token-Ring-Karte, lautet der Schnittstellename `tr0`. Befinden sich in Ihrer Maschine mehrere Karten beider Typen, müssen Sie die Zuordnung der Karten festlegen. Gehen Sie wie folgt vor:

1. Starten Sie den Executor wie folgt in der Befehlszeile: `dscontrol executor start`
2. Führen Sie den folgenden Befehl aus: `dscontrol executor xm 1`

In der Anzeige erscheint eine Ausgabe. Suchen Sie in den Zeilen, die auf `Number of NIC records` folgen, nach der IP-Adresse Ihrer Load-Balancer-Maschine, um den für die Load-Balancer-Konfiguration zu verwendenden Schnittstellennamen zu bestimmen.

Die IP-Adresse Ihrer Load-Balancer-Maschine wird wie folgt aufgelistet:
`ia->ia_addr`. Der zugehörige Schnittstellename wird mit `ifp->if_name` angegeben.

Die mit dem Befehl `executor configure` zugewiesenen Schnittstellennamen werden den in diesem Befehl aufgelisteten Schnittstellennamen zugeordnet.

Nachdem Sie diese Zuordnungsinformationen erhalten haben, können auf der Netzschnittstelle einen Aliasnamen für die Clusteradresse festlegen.

Clusteraliasnamen mit `ifconfig`-Befehlen konfigurieren

Auf Linux- oder UNIX-Systemen führt der Befehl `"executor configure"` `ifconfig`-Befehle aus.

Solaris- und HP-UX-Systeme: Bei Verwendung von bindungsspezifischen Serveranwendungen, die an eine Liste von IP-Adressen ohne die IP-Adresse des Servers gebunden werden, verwenden Sie an Stelle von `ifconfig` den Befehl **`arp publish`**, um auf der Load-Balancer-Maschine dynamisch eine IP-Adresse festzulegen. Beispiel:

```
arp -s <Cluster> <MAC-Adresse von Load Balancer> pub
```

Schritt 6. Ports definieren und Port-Optionen festlegen

Zum Definieren eines Ports können Sie den Befehl **`dscontrol port add Cluster:Port`** eingeben, die Beispielfunktionsdatei editieren oder die GUI verwenden. *Cluster* ist entweder der symbolische Name oder die IP-Adresse. *Port* ist die Nummer des Ports, den Sie für dieses Protokoll verwenden. Sie können jetzt auch verschiedene Port-Einstellungen ändern. Sie müssen alle Server für einen Port definieren und konfigurieren. Lesen Sie hierzu die Informationen in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Mit der Port-Nummer 0 (null) wird ein Platzhalter-Port angegeben. Dieser Port akzeptiert Datenverkehr, der nicht für einen der definierten Ports eines Clusters bestimmt ist. Der Platzhalter-Port wird zum Konfigurieren von Regeln und Servern für alle Ports verwendet. Diese Funktion kann auch verwendet werden, wenn Sie eine identische Server-/Regelkonfiguration für mehrere Ports haben. Der Datenverkehr an einem Port könnte dann die Lastausgleichsentscheidungen für Daten-

verkehr an anderen Ports beeinflussen. Weitere Informationen zur Verwendung eines Platzhalter-Ports finden Sie im Abschnitt „Platzhalter-Port für die Übertragung von Datenverkehr mit nicht konfiguriertem Port verwenden“ auf Seite 239.

Schritt 7. Am Lastausgleich beteiligte Servermaschinen definieren

Geben Sie zum Definieren einer am Lastausgleich beteiligten Servermaschine den Befehl **dscontrol server add** *Cluster:Port:Server* ein. Sie können auch die Beispielkonfigurationsdatei editieren oder die GUI verwenden. *Cluster* und *Server* sind entweder symbolische Namen oder IP-Adressen. *Port* ist die Nummer des Ports, den Sie für dieses Protokoll verwenden. Für einen Port eines Clusters müssen Sie mehrere Server definieren, um einen Lastausgleich durchführen zu können.

Bindungsspezifische Server: Wenn die Komponente Dispatcher die Last auf bindungsspezifische Server verteilt, *müssen* die Server so konfiguriert werden, dass sie an die Clusteradresse gebunden werden. Da der Dispatcher Pakete ohne Änderung der Ziel-IP-Adresse weiterleitet, enthalten die beim Server eingehenden Pakete noch immer die Clusteradresse als Ziel. Wenn ein Server für die Bindung an eine andere IP-Adresse als die Clusteradresse konfiguriert ist, kann der Server für den Cluster bestimmte Anforderungen nicht akzeptieren.

Mit dem Befehl `netstat -an` können Sie feststellen, ob der Server bindungsspezifisch ist. Achten Sie auf den `Server:Port`. Wenn der Server nicht bindungsspezifisch ist, gibt der Befehl `0.0.0.0:80` zurück. Andernfalls sehen Sie eine Adresse wie `192.168.15.103:80`.

Anmerkung: Für Solaris- und Linux-Systeme: Bei Verwendung von Advisor-Funktionen dürfen bindungsspezifische Server nicht verknüpft werden.

Verknüpfung mehrerer Adressen: In einer verknüpften Konfiguration muss die Adresse der verknüpften Servermaschine *nicht* mit der NFA übereinstimmen. Wenn Ihre Maschine mit mehreren IP-Adressen definiert wurde, können Sie eine andere Adresse verwenden. Für die Komponente Dispatcher muss die verknüpfte Servermaschine mit dem Befehl **dscontrol server** als **verknüpft** definiert werden. Weitere Informationen zu verknüpften Servern finden Sie im Abschnitt „Verknüpfte Server verwenden“ auf Seite 202.

Weitere Informationen zur Syntax des Befehls `dscontrol server` können Sie dem Abschnitt „`dscontrol server` — Server konfigurieren“ auf Seite 394 entnehmen.

Schritt 8. Manager-Funktion starten (optional)

Die Manager-Funktion verbessert den Lastausgleich. Soll der Manager gestartet werden, geben Sie den Befehl **dscontrol manager start** ein, editieren Sie die Beispielkonfigurationsdatei oder verwenden Sie die GUI.

Schritt 9. Advisor-Funktion starten (optional)

Die Advisor-Funktionen liefern dem Manager weitere Informationen über die Fähigkeit der am Lastausgleich beteiligten Servermaschinen, auf Anforderungen zu antworten. Advisor-Funktionen sind protokollspezifisch. Soll beispielsweise die HTTP-Advisor-Funktion gestartet werden, setzen Sie den folgenden Befehl ab:

```
dscontrol advisor start http Port
```

Eine Liste der Advisor-Funktionen mit den zugehörigen Standard-Ports finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Eine Beschreibung der einzelnen Advisor-Funktionen können Sie dem Abschnitt „Liste der Advisor-Funktionen“ auf Seite 189 entnehmen.

Schritt 10. Clusterproportionen festlegen

Wenn Sie Advisor-Funktionen starten, können Sie die Wichtigkeit ändern, die in Entscheidungen für den Lastausgleich einfließenden Informationen von Advisor-Funktionen beigemessen wird. Setzen Sie zum Festlegen von Clusterproportionen den Befehl **dscontrol cluster set Cluster proportions** ab. Weitere Informationen hierzu finden Sie im Abschnitt „Proportionale Bedeutung von Statusinformationen“ auf Seite 180.

Servermaschinen für Lastausgleich konfigurieren

Führen Sie die nachstehenden Schritte aus, wenn eine der folgenden Bedingungen erfüllt ist:

- Sie verwenden die Weiterleitungsmethode `mac` und der Server ist eine Back-End-Servermaschine.
- Sie verwenden die Weiterleitungsmethode `mac` und der Server ist ein verknüpfter Server, der als Bereitschaftsmaschine für hohe Verfügbarkeit konfiguriert ist.

Anmerkungen:

1. Prozeduren wie das Löschen von Aliasnamen auf der Loopback-Einheit müssen für den Fall, dass die Maschine in den aktiven Modus wechselt, in die `go*`-Scripts gestellt werden.
2. Ist diese Maschine als aktive Maschine mit hoher Verfügbarkeit konfiguriert, müssen Prozeduren wie das Festlegen eines Aliasnamens auf der Loopback-Einheit für den Fall, dass die Maschine in den Bereitschaftsmodus wechselt, in die `go*`-Scripts gestellt werden.

Wird die Weiterleitungsmethode `mac` verwendet, verteilt der Dispatcher die Last nur auf Server, bei denen der Loopback-Adapter mit einer zusätzlichen IP-Adresse konfiguriert werden kann, da der Back-End-Server nicht auf ARP-Anforderungen (Adressauflösungsprotokoll) reagiert. Führen Sie die Schritte in diesem Abschnitt aus, um die am Lastausgleich beteiligten Servermaschinen zu konfigurieren.

Schritt 1. Aliasnamen auf der Loopback-Einheit festlegen

Damit die am Lastausgleich beteiligten Servermaschinen arbeiten können, müssen Sie die Loopback-Einheit (die häufig als `lo0` bezeichnet wird) auf die Clusteradresse setzen (oder bevorzugen einen Aliasnamen für die Clusteradresse angeben). Bei Verwendung der Weiterleitungsmethode `mac` ändert die Komponente Dispatcher nicht die Ziel-IP-Adresse des TCP/IP-Pakets, bevor sie dieses an eine TCP-Servermaschine weiterleitet. Wird die Loopback-Einheit auf die Clusteradresse gesetzt oder diese Adresse als Aliasname der Loopback-Einheit festgelegt, akzeptieren die am Lastausgleich beteiligten Servermaschinen ein an die Clusteradresse gerichtetes Paket.

Falls Ihr Betriebssystem Aliasnamen auf Netzschnittstellen unterstützt (wie es bei AIX-, HP-UX-, Linux-, Solaris- oder Windows-Systemen der Fall ist), sollten Sie auf der der Loopback-Einheit einen Aliasnamen für die Clusteradresse festlegen. Ein Betriebssystem mit Unterstützung für Aliasnamen bringt den Vorteil, dass die am Lastausgleich beteiligten Servermaschinen für mehrere Clusteradressen konfiguriert werden können.

WICHTIGER HINWEIS: Lesen Sie die Informationen zu Linux-Systemen im Abschnitt „Alternativen für die Festlegung eines Loopback-Aliasnamens unter Linux bei Verwendung der Load-Balancer-Weiterleitungsmethode mac“ auf Seite 78.

Wenn das Betriebssystem Ihres Servers keine Aliasnamen unterstützt, müssen Sie die Loopback-Einheit auf die Clusteradresse setzen.

Verwenden Sie den in Tabelle 5 angegebenen betriebssystemspezifischen Befehl, um die Adresse der Loopback-Einheit oder einen Aliasnamen für die Einheit zu definieren.

Tabelle 5. Befehle zum Festlegen eines Aliasnamens auf der Loopback-Einheit (lo0) für Dispatcher

AIX bis Version 4.3	ifconfig lo0 alias Clusteradresse netmask Netzmaske Anmerkung: Verwenden Sie die Netzmaske des Primäradapters.
AIX 5.x	ifconfig lo0 alias Clusteradresse netmask 255.255.255.255
HP-UX	ifconfig lo0:1 Clusteradresse up
Linux	Wählen Sie einen der folgenden Befehle aus: <ul style="list-style-type: none"> • ip -4 addr add Clusteradresse /32 dev lo • ifconfig lo:1 Clusteradresse netmask 255.255.255.255 up <p>WICHTIGER HINWEIS: Wenn Sie einen der Konfigurationsbefehle auf Ihrer Maschine abgesetzt haben, müssen Sie denselben Befehl konsistent verwenden (ip oder ifconfig), da es andernfalls zu unerwarteten Ergebnissen kommen kann.</p>
OS/2	ifconfig lo Clusteradresse
OS/390	Loopback-Aliasnamen auf einem OS/390-System konfigurieren <ul style="list-style-type: none"> • In der Teildatei mit den IP-Parametern muss ein Administrator einen Eintrag in der Liste der Ausgangsadressen erstellen. Beispiel: <pre>HOME ;Address Link 192.168.252.11 tr0 192.168.100.100 1tr1 192.168.252.12 loopback</pre> • Für die Loopback-Einheit können mehrere Adressen definiert werden. • Standardmäßig ist die Loopback-Adresse 127.0.0.1 konfiguriert.
Solaris 7	ifconfig lo0:1 Clusteradresse 127.0.0.1 up
Solaris 8, Solaris 9 und Solaris 10	ifconfig lo0:1 plumb Clusteradresse netmask Netzmaske up

Tabelle 5. Befehle zum Festlegen eines Aliasnamens auf der Loopback-Einheit (lo0) für Dispatcher (Forts.)

Windows Server 2003	<ol style="list-style-type: none"> 1. Klicken Sie auf Start und dann auf Systemsteuerung. 2. Fügen Sie den MS Loopback Adapter Driver hinzu (falls noch nicht erfolgt). <ol style="list-style-type: none"> a. Klicken Sie auf Hardware. Damit wird der "Hardware-Assistent" gestartet. b. Klicken Sie auf "Weiter". c. Wählen Sie die Option Ja, die Hardware wurde bereits angeschlossen aus. Klicken Sie dann auf "Weiter". d. Wenn der MS Loopback Adapter in der Liste aufgeführt ist, ist er bereits installiert. Klicken Sie auf "Abbrechen", um die Anzeige zu verlassen. e. Ist der MS Loopback Adapter <i>nicht</i> aufgelistet, wählen Sie Neues Gerät hinzufügen aus und klicken Sie auf "Weiter". f. Falls Sie Hardware aus einer Liste auswählen möchten, wählen Sie für die Frage Soll nach neuen Hardwarekomponenten gesucht werden? als Antwort "Nein, die Hardwarekomponenten selbst in der Liste auswählen" aus und klicken Sie auf "Weiter". g. Wählen Sie Netzwerkadapter aus und klicken Sie auf "Weiter". h. Wählen Sie in der Anzeige Netzwerkadapter wählen unter "Hersteller" Microsoft und dann Microsoft Loopback Adapter aus. i. Klicken Sie auf "Weiter". Klicken Sie dann erneut auf "Weiter", um die Standardeinstellungen zu installieren (oder wählen Sie Datenträger aus, legen Sie die CD ein und installieren Sie von der CD). j. Klicken Sie auf "Fertig stellen", um die Installation zu beenden. 3. Klicken Sie in der Systemsteuerung doppelt auf Netzwerk- und DFÜ-Verbindungen. 4. Wählen Sie die Verbindung mit dem Einheitennamen "Microsoft Loopback Adapter" aus. 5. Wählen Sie im angezeigten Menü Eigenschaften aus. 6. Wählen Sie Internetprotokoll (TCP/IP) aus und klicken Sie auf Eigenschaften. 7. Klicken Sie auf Folgende IP-Adresse verwenden. Geben Sie für <i>IP-Adresse</i> die Clusteradresse und für <i>Subnetzmaske</i> die Teilnetzmaske des Back-End-Servers ein. Anmerkung: Geben Sie keine Router-Adresse ein. Verwenden Sie den lokalen Host als Standard-DNS-Server.
---------------------	---

Tabelle 5. Befehle zum Festlegen eines Aliasnamens auf der Loopback-Einheit (lo0) für Dispatcher (Forts.)

Windows 2000	<ol style="list-style-type: none"> 1. Klicken Sie auf Start, Einstellungen und dann auf Systemsteuerung. 2. Fügen Sie den MS Loopback Adapter Driver hinzu (falls noch nicht erfolgt). <ol style="list-style-type: none"> a. Klicken Sie doppelt auf Hardware. Damit wird der Assistent zum Hinzufügen/Entfernen von Hardware gestartet. b. Klicken Sie auf Weiter, wählen Sie Gerät hinzufügen bzw. Problem beheben aus und klicken Sie dann auf Weiter. c. Die Anzeige blinkt. Anschließend erscheint die Anzeige Gerät wählen. d. Wenn der MS Loopback Adapter in der Liste aufgeführt ist, ist er bereits installiert. Klicken Sie auf "Abbrechen", um die Anzeige zu verlassen. e. Ist der MS Loopback Adapter <i>nicht</i> aufgelistet, wählen Sie Neues Gerät hinzufügen aus und klicken Sie auf "Weiter". f. Falls Sie Hardware aus einer Liste auswählen möchten, wählen Sie für die Frage Soll nach neuen Hardwarekomponenten gesucht werden? als Antwort "Nein, die Hardwarekomponenten selbst in der Liste auswählen" aus und klicken Sie auf "Weiter". g. Wählen Sie Netzwerkadapter aus und klicken Sie auf "Weiter". h. Wählen Sie in der Anzeige Netzwerkadapter wählen unter "Hersteller" Microsoft und dann Microsoft Loopback Adapter aus. i. Klicken Sie auf "Weiter". Klicken Sie dann erneut auf "Weiter", um die Standardeinstellungen zu installieren (oder wählen Sie Datenträger aus, legen Sie die CD ein und installieren Sie von der CD). j. Klicken Sie auf "Fertig stellen", um die Installation zu beenden. 3. Klicken Sie in der Systemsteuerung doppelt auf Netzwerk- und DFÜ-Verbindungen. 4. Wählen Sie die Verbindung mit dem Einheitennamen "Microsoft Loopback Adapter" aus und klicken Sie mit der rechten Maustaste auf den Namen. 5. Wählen Sie im angezeigten Menü Eigenschaften aus. 6. Wählen Sie Internetprotokoll (TCP/IP) aus und klicken Sie auf Eigenschaften. 7. Klicken Sie auf Folgende IP-Adresse verwenden. Geben Sie für <i>IP-Adresse</i> die Clusteradresse und für <i>Subnetzmaske</i> die Standardteilnetzmaske (255.0.0.0) ein. Anmerkung: Geben Sie keine Router-Adresse ein. Verwenden Sie den lokalen Host als Standard-DNS-Server.
--------------	---

Tabelle 5. Befehle zum Festlegen eines Aliasnamens auf der Loopback-Einheit (lo0) für Dispatcher (Forts.)

Windows NT	<ol style="list-style-type: none"> 1. Klicken Sie auf Start und dann auf Einstellungen. 2. Klicken Sie auf Systemsteuerung und dann doppelt auf Netzwerk. 3. Fügen Sie den MS Loopback Adapter Driver hinzu (falls noch nicht erfolgt). <ol style="list-style-type: none"> a. Klicken Sie im Fenster "Netzwerk" auf Netzwerkkarte. b. Wählen Sie MS Loopback Adapter aus und klicken Sie auf OK. c. Legen Sie bei der entsprechenden Aufforderung die Installations-CD bzw. den Installationsdatenträger ein. d. Klicken Sie im Fenster "Netzwerk" auf Protokolle. e. Wählen Sie TCP/IP-Protokoll aus und klicken Sie auf Eigenschaften. f. Wählen Sie MS Loopback Adapter aus und klicken Sie auf OK. 4. Legen Sie als Loopback-Adresse die Clusteradresse fest. Übernehmen Sie die Standardteilnetzmaske (255.0.0.0) und geben Sie keine Gateway-Adresse ein. <p>Anmerkung: Möglicherweise müssen Sie die Netzwerkeinstellungen verlassen und erneut aufrufen, bevor der MS Loopback Driver unter der TCP/IP-Konfiguration angezeigt wird.</p>
------------	--

Schritt 2. Überprüfung auf zusätzliche Route

Unter einigen Betriebssystemen wurde möglicherweise eine Standardroute erstellt, die entfernt werden muss.

- Überprüfen Sie mit dem folgenden Befehl, ob unter Windows-Betriebssystemen eine zusätzliche Route vorhanden ist:

```
route print
```

WICHTIGER HINWEIS: Unter Windows 2003 müssten alle zusätzlichen Routen ignoriert werden. Falls nach dem Konfigurieren des Aliasnamens Probleme beim Routing auftreten, entfernen Sie den Aliasnamen und fügen Sie ihn dann mit einer anderen Netzmaske wieder hinzu.

- Überprüfen Sie auf allen Linux- und UNIX-Systemen mit dem folgenden Befehl, ob eine zusätzliche Route vorhanden ist:

```
netstat -nr
```

Beispiel für Windows:

1. Nachdem **route print** eingegeben wurde, wird eine ähnliche Tabelle wie die folgende angezeigt. (Dieses Beispiel veranschaulicht das Auffinden und Entfernen einer zusätzlichen Route zu Cluster 9.67.133.158 mit der Standardnetzmaske 255.0.0.0.)

Aktive Routen:

Netzwerkziel	Netzwerkmaske	Gateway	Schnittstelle	Anzahl
0.0.0.0	0.0.0.0	9.67.128.1	9.67.133.67	1
9.0.0.0	255.0.0.0	9.67.133.158	9.67.133.158	1
9.67.128.0	255.255.248.0	9.67.133.67	9.67.133.67	1
9.67.133.67	255.255.255.255	127.0.0.1	127.0.0.1	1
9.67.133.158	255.255.255.255	127.0.0.1	127.0.0.1	1
9.255.255.255	255.255.255.255	9.67.133.67	9.67.133.67	1
127.0.0.0	255.0.0.0	127.0.0.1	127.0.0.1	1
224.0.0.0	224.0.0.0	9.67.133.158	9.67.133.158	1
224.0.0.0	224.0.0.0	9.67.133.67	9.67.133.67	1
255.255.255.255	255.255.255.255	9.67.133.67	9.67.133.67	1

- Suchen Sie die Clusteradresse in der Spalte "Gateway". Ist eine zusätzliche Route vorhanden, wird die Clusteradresse zweimal aufgeführt. In diesem Beispiel erscheint die Clusteradresse (9.67.133.158) in Zeile 2 und Zeile 8.
- Ermitteln Sie für jede Zeile, in der die Clusteradresse erscheint, die Netzadresse. Sie benötigen eine dieser Routen und müssen die überschüssige Route löschen. Die zu löschende zusätzliche Route ist die Route, deren Netzwerkadresse mit der ersten Ziffer der Clusteradresse beginnt, gefolgt von drei Nullen. In diesem Beispiel erscheint die zusätzliche Route in Zeile 2. Diese Route hat die Netzadresse **9.0.0.0**:

```
9.0.0.0    255.0.0.0    9.67.133.158    9.67.133.158    1
```

Schritt 3. Zusätzliche Routen löschen

Die zusätzliche Route muss gelöscht werden. Löschen Sie die zusätzliche Route mit dem in Tabelle 6 angegebenen betriebssystemspezifischen Befehl.

Beispiel: Geben Sie zum Löschen einer zusätzlichen Route wie in der Beispielaufstellung "Aktive Routen" für Schritt 2 Folgendes ein:

```
route delete 9.0.0.0 9.67.133.158
```

Tabelle 6. Befehle zum Löschen zusätzlicher Routen für Dispatcher

HP-UX	<code>route delete</code> Clusteradresse Clusteradresse
Windows	<p>route delete Netzadresse Clusteradresse (an einer MS-DOS-Eingabeaufforderung)</p> <p>Anmerkung: Die zusätzliche Route müssen Sie bei jedem Neustart des Servers löschen.</p> <p>Unter Windows 2003 können keine Routen gelöscht werden. Alle zusätzlichen Routen müssten ignoriert werden. Falls nach dem Konfigurieren des Aliasnamens Probleme beim Routing auftreten, entfernen Sie den Aliasnamen und fügen Sie ihn dann mit einer anderen Netzmaske wieder hinzu.</p>

Wenn Sie für das in Abb. 15 auf Seite 68 gezeigte Beispiel eine Servermaschine mit AIX konfigurieren, würde der Befehl wie folgt lauten:

```
route delete -net 204.0.0.0 204.67.172.72
```

Schritt 4. Serverkonfiguration prüfen

Führen Sie zum Überprüfen der Konfiguration eines Back-End-Servers auf einer anderen Maschine im selben Teilnetz bei nicht aktivem Load Balancer und nicht konfiguriertem *Cluster* die folgenden Schritte aus:

- Setzen Sie den folgenden Befehl ab:
`arp -d Cluster`
- Setzen Sie den folgenden Befehl ab:
`ping Cluster`

Sie sollten keine Antwort empfangen. Falls Sie eine Antwort auf das ping erhalten, vergewissern Sie sich, dass Sie nicht mit `ifconfig` die Schnittstelle auf die Clusteradresse gesetzt haben. Vergewissern Sie sich, dass keine Maschine einen veröffentlichten Eintrag `arp` für die Clusteradresse hat.

- Senden Sie ein ping an den Back-End-Server und setzen Sie unmittelbar darauf den folgenden Befehl ab:
`arp -a`

Die Ausgabe des Befehls sollte die MAC-Adresse Ihres Servers enthalten. Setzen Sie den folgenden Befehl ab:

```
arp -s Cluster MAC-Adresse_des_Servers
```

4. Senden Sie ein ping an den Cluster. Sie sollten eine Antwort empfangen. Setzen Sie http, telnet oder eine andere an den Cluster adressierte Anfrage ab, die Ihr Back-End-Server verarbeiten können müsste. Vergewissern Sie sich, dass der Server ordnungsgemäß arbeitet.
5. Setzen Sie den folgenden Befehl ab:

```
arp -d Cluster
```
6. Senden Sie ein ping an den Cluster. Sie sollten keine Antwort empfangen.

Anmerkung: Falls Sie eine Antwort empfangen, setzen Sie die Anweisung `arp Cluster` ab, um die MAC-Adresse der falsch konfigurierten Maschine zu ermitteln. Wiederholen Sie dann die Schritte 1 bis 6.

Alternativen für die Festlegung eines Loopback-Aliasnamens unter Linux bei Verwendung der Load-Balancer-Weiterleitungsmethode mac

Einige Linux-Versionen setzen für jede auf der Maschine konfigurierte IP-Adresse auf allen Schnittstellen der Maschine eine ARP-Antwort ab. Linux kann außerdem für ARP-Abfragen `who-has` ausgehend von allen auf der Maschine vorhandenen IP-Adressen eine ARP-Quellen-IP-Adresse auswählen, ohne dabei die Schnittstellen zu berücksichtigen, für die diese Adressen konfiguriert sind. Dies führt dazu, dass der gesamte Clusterdatenverkehr auf unkontrollierte Weise an nur einen Server weitergeleitet wird.

Wenn Sie die Dispatcher-Weiterleitungsmethode `mac` verwenden, müssen Sie durch einen Mechanismus sicherstellen, dass die Stacks der Back-End-Server an den Cluster adressierten Datenverkehr akzeptieren können. Wenn sowohl die hohe Verfügbarkeit als auch die Verknüpfung aktiviert ist, gehört zu diesen Back-End-Servern auch die verknüpfte Bereitschaftsmaschine für hohe Verfügbarkeit.

In den meisten Fällen müssen Sie auf der Loopback-Einheit einen Aliasnamen für die Clusteradresse festlegen. Bei Back-End-Servern muss daher auf der Loopback-Einheit ein Aliasname für den Cluster definiert sein. Falls Sie mit hoher Verfügbarkeit und Verknüpfung arbeiten, müssen Sie auf der Loopback-Einheit der Bereitschaftsserver für Lastausgleich Clusteraliasnamen festlegen.

Um sicherzustellen, dass Linux-Systeme Adressen nicht auf der Loopback-Einheit zugänglich machen, können Sie eine der folgenden vier Lösungen nutzen, die Linux-Systeme mit der Dispatcher-Weiterleitungsmethode `mac` kompatibel machen.

1. Verwenden Sie einen Kernel, der die Adresse nicht verfügbar macht. Dies ist die bevorzugte Option, da sie keinen Systemaufwand pro Paket verursacht und keine Konfiguration für jeden Kernel erfordert.
 - United Linux 1 / SLES8 mit SP2(x86) oder SP3 (alle anderen Architekturen) und alle Folgeversionen enthalten den Julian Hidden Patch für das ARP-Problem. Vergewissern Sie sich, dass dieser Patch immer wirksam ist, bevor Sie einen Aliasnamen für die Clusteradresse festlegen. Verwenden Sie dazu folgenden Befehl:

```
# sysctl -w net.ipv4.conf.all.hidden=1 net.ipv4.conf.lo.hidden=1
```

Anschließend können Sie ganz normal Aliasnamen für Cluster festlegen.
Beispiel:

```
# ifconfig lo:1 $CLUSTER_ADDRESS netmask 255.255.255.255 up
```

- Verwenden Sie das in den Versionen 2.4.25 und 2.6.5 sowie allen Folgeversionen verfügbare `arp_ignore` sysctl. Beachten Sie jedoch, dass einige Distributionen manchmal Backports für Features durchführen. Stellen Sie sicher, dass diese Option aktiviert ist, bevor Sie Aliasnamen für die Clusteradressen festlegen. Verwenden Sie dazu die folgenden Befehle:

```
# sysctl -w net.ipv4.conf.all.arp_ignore=3
net.ipv4.conf.all.arp_announce=2
```

Anschließend müssen Sie die Aliasnamen für die Cluster mit folgendem Befehl definieren:

```
# ip addr add $CLUSTER_ADDRESS/32 scope host dev lo
```

In Konfigurationen mit hoher Verfügbarkeit muss ein ähnlicher Befehl in den `go*`-Scripts enthalten sein.

- Anmerkung: Vergewissern Sie sich bei Verwendung von `sysctl`, dass diese Einstellungen auch noch nach einem Warmstart gültig sind. Fügen Sie daher die Einstellungen zu `/etc/sysctl.conf` hinzu.
2. Leiten Sie mit IP-Tabellen den gesamten ankommenden Clusterdatenverkehr an den lokalen Host um. Falls Sie sich für diese Methode entscheiden, konfigurieren Sie auf dem Loopback-Adapter keinen Aliasnamen. Verwenden Sie stattdessen den folgenden Befehl:

```
# iptables -t nat -A PREROUTING -d $CLUSTER_ADDRESS -j REDIRECT
```

Dieser Befehl veranlasst Linux-Systeme, für jedes Paket ein DNAT (Destination NAT) durchzuführen, bei dem die Clusteradressen in die Schnittstellenadressen konvertiert werden. Die Kehrseite dieser Methode ist jedoch, dass bei den Verbindungen pro Sekunde ein Durchsatzrückgang von ca. 6,4 % zu verzeichnen ist. Diese Methode funktioniert nur bei unterstützten Stock Distributionen. Es ist weder ein Kernel-Modul noch ein Patch-Build oder eine Patch-Installation erforderlich.

3. Wenden Sie das Modul `noarp` ab Version 1.2.0 an. Die Kernel-Quelle muss verfügbar und ordnungsgemäß konfiguriert sein. Außerdem müssen die Entwicklungstools (`gcc`, `gnu make` usw.) verfügbar sein. Bei jedem Kernel-Upgrade müssen Sie einen Build für das Modul erstellen und das Modul installieren. Es ist unter <http://www.masarlabs.com/noarp/> verfügbar. Da der eigentliche Kernel-Code nicht modifiziert wird, ist diese Methode kein so starker Eingriff wie die (nachfolgende) Lösung Nr. 4 und auch nicht so fehleranfällig. Sie müssen das Modul konfigurieren, bevor Sie auf der Loopback-Einheit Aliasnamen für Clusteradressen festlegen. Beispiel:

```
# modprobe noarp
# noarpctl add $CLUSTER_ADDRESS Adresse_der_primären_NIC
```

Hier steht *Adresse_der_primären_NIC* für eine Adresse in demselben Teilnetz wie die Clusteradresse. Anschließend können Sie ganz normal Aliasnamen für Cluster festlegen. Beispiel:

```
# ifconfig lo:1 Clusteradresse netmask 255.255.255.255 up
```

Anmerkung: Bei Verknüpfungskonfigurationen mit hoher Verfügbarkeit müssen `noarpctl adds` und `dels` in die `go*`-Scripts aufgenommen werden, um sicherzustellen, dass der aktive Load Balancer ARP für die Clusteradressen verwenden kann und der als Server fungierende Bereitschafts-Load-Balancer nicht versehentlich den gesamten Clusterdatenverkehr empfängt.

4. Laden Sie den Julian Patch von der folgenden Website herunter: <http://www.ssi.bg/~ja/#hidden>. Lesen Sie die Anweisungen zu Ihrer Distribution, damit Sie den Patch auf einen Kernel anwenden und einen Kernel kompilieren, der für diese Distribution geeignet ist. Falls dies ein verknüpfter Load Balancer mit hoher Verfügbarkeit ist, vergewissern Sie sich, dass `uname -r` mit dem von der Distribution bereitgestellten Kernel übereinstimmt. Stellen Sie außerdem sicher, dass beim Start die Datei `.config` des Distributions-Kernels verwendet wird. Wenn Sie für Ihren Kernel einen Build mit dem Julian Hidden Patch erstellt, den Kernel installiert haben und nun ausführen, führen Sie die Anweisungen zur ersten Lösung aus, um den Patch zu aktivieren.

Anmerkung: Die Ausführung eines angepassten Kernels kann Probleme bei der Unterstützung für die Distribution nach sich ziehen.

Kapitel 8. Deployment des Dispatchers von Load Balancer für IPV4 und IPV6

Load Balancer für IPV4 und IPV6 bietet Unterstützung für das erweiterte IP-Adressierungsschema von IPV6. Load Balancer für IPV4 und IPV6 ist ein gesondertes Installationsimage, das nur die Komponente Dispatcher umfasst. Bei Auswahl dieses Installationstyps kann IPV4- und IPV6-Datenverkehr mit der MAC-basierten Paketweiterleitung des Dispatchers auf die in Ihrem Netz konfigurierten Server verteilt werden.

Das vorliegende Kapitel beschreibt die Konfigurationsunterschiede und Einschränkungen für Dispatcher bei Installation von Load Balancer für IPV4 und IPV6 und ist in folgende Abschnitte unterteilt:

- „Unterstützte Plattformen für Load Balancer für IPV4 und IPV6“ auf Seite 82
- „Load Balancer für IPV4 und IPV6 installieren“ auf Seite 83
- „Hinweise und Einschränkungen für Load Balancer für IPV4 und IPV6“ auf Seite 83
- „Verarbeitung von IPV6-Paketen in Load Balancer für IPV4 und IPV6 aktivieren“ auf Seite 88
- „Aliasnamen auf der Schnittstelleneinheit für Load Balancer für IPV4 und IPV6 festlegen“ auf Seite 88
- „Erforderliche Schritte für die Clusterkonfiguration unter Linux auf zSeries“ auf Seite 92
- „Dispatcher-Befehle (dscontrol) für Load Balancer für IPV4 und IPV6“ auf Seite 92

Allgemeine Informationen zur Komponente Dispatcher finden Sie in folgenden Abschnitten und Kapiteln:

- Im Abschnitt „Features der Komponente Dispatcher“ auf Seite 19 finden Sie einen Überblick über die Dispatcher-Features, die für die Verwaltung Ihres Netzes verfügbar sind.
- Informationen zur Planung der Lastausgleichsparameter von Dispatcher finden Sie in Kapitel 6, „Planung für Dispatcher“, auf Seite 51.
- Informationen zum Konfigurieren der Lastausgleichsparameter von Dispatcher finden Sie in Kapitel 7, „Dispatcher konfigurieren“, auf Seite 63.
- Informationen zum Konfigurieren von Load Balancer für erweiterte Funktionen finden Sie in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201.
- Informationen zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Beachten Sie, dass es bei Verwendung von Load Balancer für IPV4 und IPV6 eine Syntaxänderung für den Dispatcher-Befehl (dscontrol) gibt. Als Begrenzer für den Befehl dscontrol wird an Stelle des Doppelpunkts (:) ein at-Symbol (@) verwendet. Wenn Sie die Befehlsreferenzen in anderen Kapiteln dieses Handbuchs lesen, denken Sie daran, dass Sie den Begrenzer von dscontrol-Befehlen (:) durch @ ersetzen müssen.

Unterstützte Plattformen für Load Balancer für IPV4 und IPV6

Load Balancer für IPV4 und IPV6 kann auf allen unterstützten Plattformen mit Ausnahme von Windows 2000 installiert werden.

Informationen zu den Hardware- und Softwaresystemvoraussetzungen finden Sie auf der folgenden Webseite: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Unterstützte Plattformen für den Lastausgleich im Benutzerbereich

Auf einigen unterstützten Plattformen, z. B. auf allen Linux-Architekturen, führen Installationen von Load Balancer für IPV4 und IPV6 Lastausgleichsprozesse im Benutzerbereich und nicht im Kernel-Bereich aus. Diese Systeme hängen nicht mehr vom Kernel-Modul ab.

Die aktuellsten Informationen zu den Plattformen, die den (Kernel-unabhängigen) Lastausgleich im Benutzerbereich unterstützen, finden Sie auf der folgenden Webseite: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

Unterstützte Systeme mit Lastausgleichsprozessen im Benutzerbereich haben z. T. andere Konfigurationsprozeduren als Systeme, bei denen die Lastausgleichsprozesse im Kernel-Bereich ausgeführt werden. Der verbleibende Teil dieses Abschnitts zu Load Balancer für IPV4 und IPV6 beschäftigt sich mit diesen Unterschieden.

Spezielle Hinweise zur Linux-Plattform

Linux auf zSeries-Systemen

- **Für Linux auf zSeries-Systemen ist libstdc++.so.5 erforderlich.** Für eine ordnungsgemäße Installation benötigen zSeries-Systeme mit Linux das RPM-Paket libstdc++.so.5. Ohne dieses Paket schlägt die Installation fehl.
- **Einschränkungen bei der Benutzung einer qeth/OSA-Schnittstelle.** Für zSeries-Systeme mit Linux gelten Einschränkungen hinsichtlich der Verwendung einer qeth/OSA-Schnittstelle. Die native Weiterleitung von einer qeth/OSA-Schnittstelle aus wird nicht unterstützt. Es gibt jedoch eine Ausweichmaßnahme, denn Linux-Systeme arbeiten im Benutzerbereich und können die Linux-Tunnelung unterstützen.

Unterstützung für Linux-Tunnelung

Installationen von Load Balancer für IPV4 und IPV6 haben auf Linux-Systemen die Möglichkeit zur Weiterleitung durch Tunnel, z. B. durch IPIP und IPGRE. Wenn Sie Linux-zSeries-Maschinen mit einer qeth/OSA-Schnittstelle verwenden, können Sie einen Linux-Tunnel definieren, der diese Schnittstelle durchquert. Bei Linux-Systemen ist eine Weiterleitung zwischen Maschinen möglich, die sich im Bereich derselben qeth/OSA-Einheit oder anderer qeth/OSA-Einheiten bzw. an beliebiger anderer Stelle im Netz befinden.

Einschränkungen für Back-End-Server

Solaris-Systeme: Der Lastausgleich für IPV6-Datenverkehr auf Back-End-Servern mit Solaris 5.8 wird nicht unterstützt. Der IPV6-Stack von Solaris 5.8 ist nicht mit IPV6-Paketen kompatibel, die mit mac weitergeleitet wurden. Wenn der Cluster mit dem Befehl ifconfig lo0 (Loopback) auf einem Back-End-Server mit Solaris 5.8 konfiguriert ist, kommt das Paket zwar beim Solaris-5.8-Knoten an, wird jedoch

nicht akzeptiert. Mit Installationen von Load Balancer für IPV4 und IPV6 können Sie jedoch einen Lastausgleich für IPV4-Datenverkehr auf Back-End-Servern mit Solaris 5.8 durchführen.

z/OS-Systeme: Der Lastausgleich für IPV6-Datenverkehr auf z/OS-Back-End-Servern wird nicht unterstützt. Mit Installationen von Load Balancer für IPV4 und IPV6 können Sie jedoch einen Lastausgleich für IPV4-Datenverkehr auf z/OS-Back-End-Servern durchführen.

Load Balancer für IPV4 und IPV6 installieren

Die Installationsschritte und Paketnamen für Load Balancer für IPV4 und IPV6 stimmen mit denen überein, die für Load Balancer mit ausschließlicher Unterstützung für IPV4-Serveradressen gelten. Im Lieferumfang sind jedoch weniger Load-Balancer-Komponentenpakete enthalten, weil nur die Komponente Dispatcher verfügbar ist.

Wenn Sie System-Tools für Paketinstallation verwenden, wird für die Installation von Load Balancer für IPV4 und IPV6 eine etwas andere Reihenfolge empfohlen. Das Paket mit der Administrationskomponente muss nach dem Paket mit der Komponente Dispatcher installiert werden. Installieren Sie die Pakete für Load Balancer für IPV4 und IPV6 mit den System-Tools für Paketinstallation in dieser Reihenfolge: Basisprodukt, Lizenz, Komponente Dispatcher, Administration, Dokumentation, Metric Server.

Die folgende Beispielliste zeigt die empfohlene Installationsreihenfolge für die Pakete von Load Balancer für IPV4 und IPV6 auf AIX-Systemen:

- `ibmlb.base.rte` (Basispaket)
- `ibmlb.lb.license` (Lizenzpaket bei Installation von einer CD)
- `ibmlb.lb.driver` (spezifisches Einheits-treiberpaket für AIX)
- `ibmlb.disp.rte` und `ibmlb.msg.Sprache.lb` (Komponentenpaket für Dispatcher mit Nachrichtenpaket)
- `ibmlb.admin.rte` und `ibmlb.msg.Sprache.admin` (Komponentenpaket für Administration mit Nachrichtenpaket)
- `ibmlb.doc.rte` und `ibmlb.msg.en_US.doc` (Dokumentationspaket mit Nachrichtenpaket)
- `ibmlb.ms.rte` (Metric-Server-Paket)

Es ist wichtig, dass Sie vor der Installation von Load Balancer für IPV4 und IPV6 alle früheren Versionen von Load Balancer deinstallieren. Zwei Load Balancer dürfen nicht auf einer Maschine installiert sein.

Anweisungen für die Installation des Produkts finden Sie in Kapitel 4, „Load Balancer installieren“, auf Seite 31.

Hinweise und Einschränkungen für Load Balancer für IPV4 und IPV6

Die Komponente Dispatcher stellt viele, aber nicht alle Funktionen bereit, die die Dispatcher-Komponente von Load-Balancer-Installationen mit ausschließlicher IPV4-Unterstützung bietet. In den folgenden Abschnitten sind die einzelnen Konfigurationsunterschiede und Funktionseinschränkungen für den Dispatcher von Load Balancer für IPV4 und IPV6 erläutert.

Verbindungslokale IPV6-Adresse konfigurieren

Bei der IPV6-Adressierung muss jede Maschine der Load-Balancer-Konfiguration eine verbindungslokale IPV6-Adresse haben.

Die verbindungslokale Adresse wird beim IPV6-Datenverkehr für die Nachbarerkennung verwendet. Ohne diese Adresse auf der Load-Balancer-Maschine und auf den Back-End-Servern findet keine Nachbarerkennung statt, so dass die Maschinen nichts voneinander wissen. Load Balancer für IPV6 kann keinen Datenverkehr weiterleiten, wenn nicht für eine Schnittstelle jeder Maschine in der Load-Balancer-Konfiguration eine verbindungslokale IPV6-Adresse konfiguriert ist.

Homogene Cluster-Server-Paare

Wenn Sie Load Balancer für IPV4 und IPV6 konfigurieren, müssen alle Server innerhalb des Clusters homogen sein. Ist beispielsweise Cluster1 mit einer IPV4-Adresse definiert, müssen alle Server in Cluster1 IPV4-Server sein. Wenn Cluster2 mit einer IPV6-Adresse definiert ist, müssen alle für Cluster2 definierten Server IPV6-Server sein. Darüber hinaus muss das vom Client für das Senden von IP-Paketen verwendete Protokoll mit dem Cluster-IP-Format übereinstimmen.

Die Unterstützung für eine gemischte Client-Umgebung mit IPV4 und IPV6 erfordert, dass es für jeden Cluster zwei logische Clusterdefinitionen geben muss, eine Definition als IPV4-Cluster und eine als IPV6-Cluster. Clients, die IPV4-Pakete senden, werden von Load Balancer zu dem logischen Cluster weitergeleitet, der für die Verwendung von IPV4-Adressen konfiguriert ist. Clients, die IPV6-Pakete senden, werden von Load Balancer zu dem logischen Cluster weitergeleitet, der für die Verwendung von IPV6-Adressen konfiguriert ist.

Nicht unterstützte Dispatcher-Features

Viele der in Kapitel 6, „Planung für Dispatcher“, auf Seite 51 und in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201 beschriebenen Dispatcher-Features sind in Load Balancer für IPV4 und IPV6 verfügbar.

Nachfolgend sind die Dispatcher-Features aufgelistet, die von Load Balancer für IPV4 und IPV6 *nicht* unterstützt werden:

- Weiterleitungsmethode cbr
- Weiterleitungsmethode nat
- Fernverwaltung
- Regelbasierter Lastausgleich
- SNMP-Subagent
- WAN-Lastausgleich
- Protokollunterstützung für UDP

Im Abschnitt „Features der Komponente Dispatcher“ auf Seite 19 finden Sie eine kurze Beschreibung der Dispatcher-Features, die für die Verwaltung Ihres Netzes verfügbar sind.

Advisor-Funktionen konfigurieren

Falls Sie auf Ihrer Maschine mit dem Protokoll IPV6 arbeiten und Advisor-Funktionen verwenden möchten, müssen Sie sicherstellen, dass die **Protokolldatei** die folgende Zeile enthält.

```
ipv6-icmp 58 IPV6-ICMP
```

Auf Linux- und UNIX-Systemen befindet sich die Protokolldatei im Verzeichnis `/etc/protocols`. Auf Windows-Systemen finden Sie die Protokolldatei im Verzeichnis `C:\windows\system32\drivers\etc\`.

Einschränkung für die Verwendung von Advisor-Funktionen: Wenn Load Balancer auf einem Computer mit mehreren Netzwerkadaptern ausgeführt wird und der Advisor-Datenverkehr über einen bestimmten Adapter fließen soll, können Sie für die Pakete keine bestimmte Quellen-IP-Adresse erzwingen. (Für Installationen von Load Balancer für IPV4 und IPV6 steht das Merkmal `-DLB_ADV_SRC_ADDR` nicht zur Verfügung.)

Weitere Informationen zu Advisor-Funktionen finden Sie im Abschnitt „Advisor-Funktionen“ auf Seite 250.

Hohe Verfügbarkeit konfigurieren

Falls Sie auf Ihrer Maschine mit dem Protokoll IPV6 arbeiten und die hohe Verfügbarkeit nutzen möchten, müssen Sie überprüfen, ob `protocol 58` in der **Protokolldatei** als ICMPv6 definiert ist. Informationen zum Editieren der Protokolldatei finden Sie im Abschnitt „Advisor-Funktionen konfigurieren“ auf Seite 84.

Wenn Sie Load Balancer für IPV4 und IPV6 installiert haben, gelten für die Konfiguration einer Dispatcher-Maschine mit hoher Verfügbarkeit die folgenden Einschränkungen und speziellen Hinweise:

- Die gegenseitige hohe Verfügbarkeit wird nicht unterstützt.
- Die Überwachungssignalaare müssen beide das IPV4- oder das IPV6-Format haben. (Diese Signale werden zwischen dem primären Dispatcher und den Bereitschafts-Dispatchern ausgetauscht, um den Ausfall eines Dispatchers festzustellen.)
- Systeme, die im Benutzerbereich arbeiten (z. B. Linux-Systeme): In einer Umgebung mit hoher Verfügbarkeit oder einer Standalone-Umgebung dürfen Sie auf dem Netzadapter keinen Aliasnamen für die Clusteradresse verwenden.
- Systeme, die im Benutzerbereich arbeiten (z. B. Linux-Systeme): Die `go*`- und `highavailChange`-Scripts können vom Verzeichnis `.../ibm/edge/lb/servers/samples` in das Verzeichnis `.../ibm/edge/lb/servers/bin` verschoben werden, um die Änderungen des Status der hohen Verfügbarkeit für die Dispatcher-Maschine zu protokollieren. Sie müssen diese Scripts nicht ändern.
- Linux-zSeries-Systeme mit einer qeth/OSA-Schnittstelle: Dies ist der einzige Netzschnittstellentyp, für den das allgemeine Verbot, Schnittstellenaliasnamen für Clusteradressen zu verwenden, nicht gilt. Sie können vielmehr wie folgt sicherstellen, dass der Clusterdatenverkehr über eine OSA-Schnittstelle zur Linux-Gastmaschine gelangt:
 - Sie benötigen die `go*`-Scripts, die mit den im Abschnitt „Erforderliche Schritte für die Clusterkonfiguration unter Linux auf zSeries“ auf Seite 92 beschriebenen Befehlen wie folgt modifiziert werden müssen:
 - `goActive`: Fügen Sie die Befehle `ip` und `iptables/ip6tables` hinzu, um die Clusteradresse zu konfigurieren und die `iptables`-Regel hinzuzufügen.
 - `goStandby`: Fügen Sie die Befehle `ip` und `iptables/ip6tables` hinzu, um die Clusteradresse aus der Konfiguration zu entfernen und die `iptables`-Regel zu entfernen.
 - `goInOp`: Fügen Sie die Befehle `ip` und `iptables/ip6tables` hinzu, um die Clusteradresse aus der Konfiguration zu entfernen und die `iptables`-Regel zu entfernen.
 - `goIdle`: Dieses Script sollte nicht erstellt werden.

Weitere Informationen zum Feature für hohe Verfügbarkeit finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 204.

Server verknüpfen

Die Verknüpfung ist eine Konfiguration, bei der sich Load Balancer auf derselben Maschine befinden kann wie ein Server, für dessen Anforderungen er einen Lastausgleich durchführt.

Wenn Sie Load Balancer für IPV4 und IPV6 installiert haben, ist das Verknüpfungsfeature für alle unterstützten Betriebssysteme verfügbar, außer auf Windows-Systemen und auf Systemen, die im Benutzerbereich arbeiten, z. B. auf Linux-Systemen.

Weitere Informationen zur Verknüpfung von Servern finden Sie im Abschnitt „Verknüpfte Server verwenden“ auf Seite 202.

Affinitätsfeature für im Benutzerbereich arbeitende Systeme (Linux)

Das Load-Balancer-Affinitätsfeature für im Benutzerbereich arbeitende Systeme (z. B. für Linux-Systeme) unterscheidet sich in seiner Funktionsweise vom Affinitätsfeature für andere Betriebssysteme, die im Kernel-Bereich ausgeführt werden.

Bei Systemen, die im Benutzerbereich arbeiten, ordnet Load Balancer einem Back-End-Server eine Client-IP-Adresse zu. Die Affinität wird hergestellt, wenn die Ziel-IP-Adresse eines Pakets mit dem Cluster übereinstimmt, der Ziel-Port mit dem Load-Balancer-Port übereinstimmt und die Quellen-IP-Adresse passt.

Bei bestehender Affinität werden alle Pakete an denselben Back-End-Server gesendet. Wenn die Affinität durch das Herunterfahren oder Entfernen eines Servers unterbrochen wird, kommt es zu einer Unterbrechung der gesamten Affinität und damit aller Verbindungen zu diesem Server.

In der Befehlszeile oder auf GUI-Clients werden auch keine Verbindungsinformationen ausgegeben. Es wird nur die Anzahl der aktiven Affinitätssätze verwendet.

Mit dieser Herangehensweise wird eine feste Affinität etabliert, die für Load Balancer effizienter ist.

Bei Systemen, die den Lastausgleich im Kernel verarbeiten, hat die IP-Affinität den Nachteil einer erhöhten CPU-Belastung und Speicherbelegung durch den Verbindungsweiterleitungsmechanismus. Bei Systemen, die den Lastausgleich im Benutzerbereich verarbeiten, bewirkt die Affinitätsmethode verglichen mit der Verbindungsweiterleitung eine geringere Speicherbelegung und CPU-Belastung.

Bei diesem Eindatensatzmodell von im Benutzerbereich arbeitenden Systemen wurden die der Affinität zugeordneten Werte `stickytime` und `staletimeout` außerdem zu einem Wert, `staletimeout`, zusammengefasst. Wenn Sie ein System mit Verarbeitung im Kernel-Bereich auf ein System mit Verarbeitung im Benutzerbereich migrieren, sollten Sie die Werte für das maximale Inaktivitätszeitlimit (`maximum staletimeout`) und für die Haltezeit (`stickytime`) als neues Inaktivitätszeitlimit (`staletimeout`) für Load Balancer verwenden, da die Entfernung eines Affinitätssatzes auch Verbindungen unterbricht.

Allgemeine Informationen zum Affinitätsfeature für Systeme mit Verarbeitung im Kernel-Bereich und für Systeme mit Verarbeitung im Benutzerbereich finden Sie im Abschnitt „Funktionsweise der Affinität für Load Balancer“ auf Seite 222.

Metric Server konfigurieren

Falls Sie auf Ihrer Maschine mit dem Protokoll IPV6 arbeiten und Metric Server verwenden möchten, müssen Sie überprüfen, ob protocol 58 in der **Protokolldatei** als ICMPv6 definiert ist. Informationen zum Editieren der Protokolldatei finden Sie im Abschnitt „Advisor-Funktionen konfigurieren“ auf Seite 84.

In einer Load-Balancer-Konfiguration mit Unterstützung für IPV4- und IPV6-Cluster können Server, die die Funktion Metric Server ausführen, nur als IPV4-Server oder nur als IPV6-Server konfiguriert werden. Wenn Sie durchsetzen möchten, dass Metric Server ein bestimmtes Protokoll verwendet (IPV4 oder IPV6), geben Sie im Script metricserver das Java-Merkmal `java.rmi.server.hostname` an.

WICHTIGER HINWEIS: Der im Java-Merkmal angegebene `hostname` muss die physische IP-Adresse von Metric Server sein.

UNIX- oder Linux-Systeme: Wenn Metric Server über die IPV6-Adresse 2002:92a:8f7a:162:9:42:92:67 kommunizieren soll, geben Sie im metricserver-Start-Script (im Verzeichnis `/usr/bin`) das Java-Merkmal nach `$LB_CLASSPATH` an.

```
/opt/ibm/edge/lb/java/jre/bin/java ..... $LB_CLASSPATH
-Djava.rmi.server.hostname=2002:92a:8f7a:162:9:42:92:67
com.ibm.internet.nd.sma.SMA_Agent
$LB_RMIPORT $LOG_LEVEL $LOG_SIZE $LOG_DIRECTORY $KEYS_DIRECTORY
$SCRIPT_DIRECTORY &
```

Windows-Systeme: Wenn Metric Server über die IPV6-Adresse 2002:92a:8f7a:162:9:42:92:67 kommunizieren soll, müssen Sie die Datei `metricserver.cmd` (im Verzeichnis `C:\winnt\system32`) wie folgt bearbeiten:

```
start/min /wait %IBMLBPATH%\java\jre\bin\java
-Djava.rmi.server.hostname=2002:92a:8f7a:162:9:42:92:67
-Djava.net.preferIPv4Stack=false
-Djava.net.preferIPv6Stack=true
-Xrs -cp
%LB_CLASSPATH% com.ibm.internet.nd.sma.SMA_Agent
%RMI_PORT% %LOG_LEVEL% %LOG_SIZE% %LOG_DIRECTORY% %KEYS_DIRECTORY%
%SCRIPT_DIRECTORY%
goto done

:stop
%IBMLBPATH%\java\jre\bin\java
-Djava.rmi.server.hostname=2002:92a:8f7a:162:9:42:92:67
-Djava.net.preferIPv4Stack=false
-Djava.net.preferIPv6Stack=true
-cp %LB_CLASSPATH% com.ibm.internet.nd.sma.SMA_AgentStop %RMI_PORT%
:done
```

Weitere Informationen hierzu finden Sie im Abschnitt „Metric Server“ auf Seite 196.

Verarbeitung von IPV6-Paketen in Load Balancer für IPV4 und IPV6 aktivieren

AIX-, Linux- und Windows-Systeme: Bevor Sie den Executor (mit `dscontrol executor start`) starten, müssen Sie als root in der Befehlszeile Folgendes absetzen:

- Auf AIX-Systemen: `autoconf6`
Wenn IPV6-Pakete (auch nach einem Systemwarmstart) unterbrechungsfrei verarbeitet werden sollen, bearbeiten Sie die Datei `/etc/rc.tcpip`. Entfernen Sie das Kommentarzeichen in der folgenden Zeile und fügen Sie die Option `-A` hinzu:
`start /usr/sbin/autoconf6 " " -A`
- Auf Linux-Systemen: `modprobe ipv6`
- Auf Windows-Systemen: `netsh interface ipv6 install`

Diese Befehle aktivieren die Verarbeitung von IPV6-Paketen unter dem jeweiligen Betriebssystem. Dieser Befehl muss nur einmal abgesetzt werden. Danach können Sie den Executor so oft starten und stoppen, wie es erforderlich ist.

Wenn Sie den Befehl zur Aktivierung von IPV6-Paketen auf diesen Systemen nicht absetzen, wird der Executor nicht gestartet.

HP-UX- und Solaris-Systeme: Mit dem Befehl `ifconfig` müssen IPV6-Adressen gebunden und eine Schnittstelle konfiguriert werden, damit der Dispatcher IPV6-Pakete überprüft. Setzen Sie vor dem Starten des Executors (mit `dscontrol executor start`) als root Folgendes in der Befehlszeile ab:

- Auf HP-UX-Systemen:
`ifconfig device inet6 up`
- Auf Solaris-Systemen:
`ifconfig device inet6 plumb`
`ifconfig device inet6 Adresse/Präfix up`

Wenn Sie diese Befehle nicht absetzen, wird der Executor zwar gestartet, es können jedoch keine IPV6-Pakete angezeigt werden.

Aliasnamen auf der Schnittstelleneinheit für Load Balancer für IPV4 und IPV6 festlegen

Zum Konfigurieren der Clusteradresse auf einer Netzschnittstellenkarte (NIC, Network Interface Card) der Dispatcher-Maschine können Sie den Befehl `dscontrol executor configure Clusteradresse` absetzen. Der Befehl `dscontrol executor configure` führt die Betriebssystembefehle für Adapterkonfiguration aus (z. B. den Befehl `ifconfig`, `dsconfig` (nur IPV6) oder `ip`). Falls Sie auf der NIC der Dispatcher-Maschine einen Aliasnamen angeben möchten, können Sie die Betriebssystembefehle für Adapterkonfiguration direkt angeben, anstatt den Befehl `executor configure` zu verwenden.

Anmerkung: Bei Systemen, die im Benutzerbereich arbeiten (z. B. Linux-Systemen), dürfen Sie die Clusteradresse nicht mit dem Befehl `dscontrol executor configure`, `ip` oder `ifconfig` konfigurieren. Load Balancer macht die Clusteradressen im Netz nativ zugänglich. Die Clusteradresse erscheint auch nicht als Aliasname irgendeiner Schnittstelle. Dies ist normal.

Dies gilt jedoch **nicht** für Linux auf zSeries mit einer qeth/OSA-Schnittstelle. Auf dieser Plattform muss die Clusteradresse konfiguriert werden. Weitere Details hierzu finden Sie im Abschnitt „Erforderliche Schritte für die Clusterkonfiguration unter Linux auf zSeries“ auf Seite 92.

Zum Erstellen eines Aliasnamens auf der Loopback-Einheit (lo0) der am Lastausgleich beteiligten Server müssen Sie die Betriebssystembefehle für Adapterkonfiguration verwenden.

Wenn Sie Load Balancer für IPV4 und IPV6 installiert haben, können Sie mit den folgenden Befehlen auf der Netzchnittstelle und der Loopback-Einheit (*Schnittstellename*) einen Aliasnamen angeben.

AIX-Systeme (5.x):

- Für IPV6-Adressen:

```
ifconfig Schnittstellename inet6  
Clusteradresse/Präfixlänge alias
```

Beispiel für das Festlegen eines Aliasnamens auf der Loopback-Einheit von Servern, die am Lastausgleich teilnehmen:

```
ifconfig lo0 inet6 2002:4a::541:56/128 alias
```

- Für IPV4-Adressen: Der Prozess hat sich nicht geändert. Tabelle 5 auf Seite 73 enthält die Befehle zum Erstellen eines Aliasnamens auf der Loopback-Einheit der am Lastausgleich beteiligten Server.

HP-UX-Systeme:

- Für IPV6-Adressen:

```
ifconfig Schnittstellename:Aliasname  
inet6 Clusteradresse up prefix Präfixlänge
```

Beispiel für das Festlegen eines Aliasnamens auf der Loopback-Einheit von Servern, die am Lastausgleich teilnehmen:

```
ifconfig lo0:1 inet6 3ffe:34::24:45 up prefix 128
```

- Für IPV4-Adressen: Der Prozess hat sich nicht geändert. Tabelle 5 auf Seite 73 enthält die Befehle zum Erstellen eines Aliasnamens auf der Loopback-Einheit der am Lastausgleich beteiligten Server.

Linux-Systeme:

- Für IPV6- oder IPV4-Adressen:

```
ip -Version addr add Clusteradresse/Präfixlänge dev lo
```

Beispiel für das Festlegen eines Aliasnamens auf der Loopback-Einheit von Servern, die am Lastausgleich teilnehmen:

```
ip -6 addr add 3ffe:34::24:45/128 dev lo  
ip -4 addr add 12.42.38.125/32 dev lo
```

Anmerkung: Sie können auch den Befehl `ifconfig` verwenden. Tabelle 5 auf Seite 73 enthält die Informationen zum Erstellen eines Aliasnamens auf der Loopback-Einheit mit dem Befehl `ifconfig`.

Wenn Sie einen der Konfigurationsbefehle auf Ihrer Maschine abgesetzt haben, müssen Sie denselben Befehl unbedingt konsistent verwenden (**ip** oder **ifconfig**), da es andernfalls zu unerwarteten Ergebnissen kommen kann.

Systeme mit Solaris 8, 9 und 10:

- Für IPV6-Adressen:

```
ifconfig Schnittstellename inet6 addif  
Clusteradresse/Präfixlänge up
```

Beispiel für das Festlegen eines Aliasnamens auf der Loopback-Einheit von Servern, die am Lastausgleich teilnehmen:

```
ifconfig lo0 inet6 addif 3ffe:34::24:45/128 up
```

- Für IPV4-Adressen: Der Prozess hat sich nicht geändert. Tabelle 5 auf Seite 73 enthält die Befehle zum Erstellen eines Aliasnamens auf der Loopback-Einheit der am Lastausgleich beteiligten Server.

Windows-2003-Systeme (Windows 2000 und Windows NT bieten keine Unterstützung für IPV6):

- Für IPV4-Adressen: Der Prozess hat sich nicht geändert. Tabelle 5 auf Seite 73 enthält die Befehle zum Erstellen eines Aliasnamens auf der Loopback-Einheit der am Lastausgleich beteiligten Server.
- Für IPV6-Adressen:

1. Bestimmen Sie mit dem Befehl `ipconfig /all` den Schnittstellennamen für die Loopback-Einheit. Dieser Befehl lokalisiert die Verbindung mit einer Beschreibung des Microsoft Loopback Adapters. Bei der folgenden Beispielausgabe des Befehls `ipconfig /all` ist der Microsoft Loopback Adapter 'Ethernet-Adapter "LAN-Verbindung 2"'. Die Verbindung ist demzufolge "LAN-Verbindung 2":

Windows-IP-Konfiguration

```
Hostname . . . . . : ndserv10  
Primäres DNS-Suffix . . . . . : rtp.raleigh.ibm.com  
Knotentyp . . . . . : unbekannt  
IP-Routing aktiviert . . . . . : Nein  
WINS-Proxy aktiviert . . . . . : Nein  
DNS-Suffixsuchliste . . . . . : rtp.raleigh.ibm.com
```

Ethernet-Adapter "LAN-Verbindung 2":

Verbindungsspezifisches DNS-Suffix:

```
Beschreibung . . . . . : Microsoft Loopback Adapter  
Physikalische Adresse . . . . . : 02-00-4C-4F-4F-50  
DHCP aktiviert . . . . . : Nein  
IP-Adresse . . . . . : 9.42.92.158  
Subnetzmaske . . . . . : 255.255.252.0  
IP-Adresse . . . . . : 9.42.92.159  
Subnetzmaske . . . . . : 255.255.252.0  
IP-Adresse . . . . . : 2002:92a:8f7a:162:9:42:92:160  
IP-Adresse . . . . . : 2002:92a:8f7a:162:9:42:92:159  
IP-Adresse . . . . . : fe80::4cff:fe4f:4f50%4  
Standardgateway . . . . . :  
DNS-Server . . . . . : 127.0.0.1  
fec0:0:0:ffff::1%1  
fec0:0:0:ffff::2%1  
fec0:0:0:ffff::3%1
```


2. Fügen Sie die Clusteradresse mit dem Befehl netsh zur Loopback-Einheit hinzu. Beispiel:

```
netsh interface ipv6 add address "LAN-Verbindung 2"
2002:92a:8f7a:162:9:42:92:161
```

3. Setzen Sie den Befehl ipconfig /all erneut ab. Jetzt müssten Sie die auf dem Loopback-Adapter hinzugefügte Adresse sehen. Beispiel:

Ethernet-Adapter "LAN-Verbindung 2":

```
Verbindungsspezifisches DNS-Suffix:
Beschreibung . . . . . : Microsoft Loopback Adapter
Physikalische Adresse . . . . . : 02-00-4C-4F-4F-50
DHCP aktiviert . . . . . : Nein
IP-Adresse . . . . . : 9.42.92.158
Subnetzmaske . . . . . : 255.255.252.0
IP-Adresse . . . . . : 9.42.92.159
Subnetzmaske . . . . . : 255.255.252.0
IP-Adresse . . . . . : 2002:92a:8f7a:162:9:42:92:161
IP-Adresse . . . . . : 2002:92a:8f7a:162:9:42:92:160
IP-Adresse . . . . . : 2002:92a:8f7a:162:9:42:92:159
IP-Adresse . . . . . : fe80::4cff:fe4f:4f50%4
Standardgateway . . . . . :
DNS-Server . . . . . : 127.0.0.1
                        fec0:0:0:ffff::1%1
                        fec0:0:0:ffff::2%1
                        fec0:0:0:ffff::3%1
```

4. Aktivieren Sie mit dem Befehl netsh interface ipv6 show interface die Weiterleitung für alle Schnittstellen der Maschine. Für alle Schnittstellen, die mit dem Namen "LAN-Verbindung" aufgelistet sind, muss die IP-Weiterleitung aktiviert sein. Beispiel:

```
netsh interface ipv6>show interface
Aktivitätsstatus wird abgefragt...
```

Idx	Met	MTU	Status	Name
---	----	-----	-----	-----
6	2	1280	Getrennt	Teredo-Tunnelung Pseudoschnittstelle
5	0	1500	Verbunden	LAN-Verbindung
4	0	1500	Verbunden	LAN-Verbindung 2
3	1	1280	Verbunden	6-zu-4 Pseudoschnittstelle
2	1	1280	Verbunden	Automatische Tunnelung Pseudoschnittstelle
1	0	1500	Verbunden	Loopback Pseudoschnittstelle

```
netsh interface ipv6>set interface "LAN-Verbindung"
forwarding=enabled
Ok.
```

```
netsh interface ipv6>set interface "LAN-Verbindung 2"
forwarding=enabled
Ok.
```

OS/2-Systeme:

- Für IPV6- und IPV4-Adressen: Der Prozess hat sich nicht geändert. Tabelle 5 auf Seite 73 enthält die Befehle zum Erstellen eines Aliasnamens auf der Loopback-Einheit der am Lastausgleich beteiligten Server.

Erforderliche Schritte für die Clusterkonfiguration unter Linux auf zSeries

Unter Linux auf zSeries sind die folgenden zusätzlichen Konfigurationsschritte für Load Balancer erforderlich:

1. Konfigurieren Sie die Clusteradresse mit dem Befehl `ip` oder `ifconfig`.

Für IPV6- oder IPV4-Adressen:

```
ip -Version addr  
add Clusteradresse/Präfixlänge dev Einheit
```

Beispiel:

```
ip -4 addr add 12.42.38.125/24 dev eth0  
ip -6 addr add 3ffe:34::24:45/64 dev eth0
```

2. Fügen Sie wie folgt eine iptables-Regel für das Übergehen von Paketen hinzu, die an die Clusteradresse gerichtet sind:

Für IPV4-Adressen:

```
iptables -t filter -A INPUT -d Clusteradresse -j DROP
```

Für IPV6-Adressen:

```
ip6tables -t filter -A INPUT -d Clusteradresse -j DROP
```

Beispiel:

```
iptables -t filter -A INPUT -d 12.42.38.125 -j DROP  
ip6tables -t filter -A INPUT -d 3ffe:34::24:45 -j DROP
```

Mit den folgenden Befehlen können Sie die obige Konfiguration rückgängig machen:

```
ip -Version addr  
del Clusteradresse/Präfixlänge dev Einheit  
iptables -t filter -D INPUT -d Clusteradresse -j DROP  
ip6tables -t filter -D INPUT -d Clusteradresse -j DROP
```

Dispatcher-Befehle (dscontrol) für Load Balancer für IPV4 und IPV6

Da Load Balancer für IPV4 und IPV6 nicht alle Komponentenfeatures unterstützt, sind für diese Installation nur einige der `dscontrol`-Befehle gültig, die für Load Balancer mit ausschließlicher Unterstützung für IPV4 gelten. Dieser Abschnitt beschäftigt sich mit den Syntaxunterschieden und listet alle für die Komponente Dispatcher von Load Balancer für IPV4 und IPV6 unterstützten `dscontrol`-Befehle auf.

Unterschiede in der Befehlssyntax

Bei Verwendung von Load Balancer für IPV4 und IPV6 gibt es eine wichtige Syntaxänderung für den Dispatcher-Befehl (`dscontrol`). Als Begrenzer für den Befehl `dscontrol` wird an Stelle des Doppelpunkts (`:`) ein `at`-Symbol (`@`) verwendet.

Es war notwendig, einen anderen Begrenzer zu definieren, weil das IPV6-Format den Doppelpunkt (`:`) in seinem Adressierungsschema verwendet.

Beispiele für den Befehl `dscontrol` mit einem `at`-Begrenzer (`@`):

- Hinzufügen eines IPV6-Servers (30::200) am Port 80 zu einem IPV6-Cluster (30::100):

```
dscontrol server add 30::100@80@30::200
```

- Hinzufügen eines IPV4-Servers (192.4.40.35) am Port 80 zu einem IPV4-Cluster (192.4.40.30):
`dscontrol server add 192.4.40.30@80@192.4.20.35`

Wichtiger Hinweis: Wenn Sie die Befehlsreferenzen in diesem Handbuch lesen, denken Sie daran, dass Sie den Begrenzer von dscontrol-Befehlen (:) durch @ ersetzen müssen.

Unterstützte dscontrol-Befehle

Ausführliche Informationen zu allen dscontrol-Befehlen und Beispiele für die Syntax finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Zusammenfassung der unterstützten Befehle für den Dispatcher von Load Balancer für IPV4 und IPV6:

- `dscontrol advisor`
 - Alle Argumente und ihre Schlüsselwerte sind gültig.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol advisor — Advisor-Funktion steuern“ auf Seite 351.
- `dscontrol binlog`
 - Alle Argumente und ihre Schlüsselwerte sind gültig.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol binlog — Binäre Protokolldatei steuern“ auf Seite 356.
- `dscontrol cluster`
 - Alle Argumente sind gültig. Die einzigen gültigen Schlüsselwerte sind `address` und `proportions`.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol cluster — Cluster konfigurieren“ auf Seite 357.
- `dscontrol executor`
 - Alle Argumente sind gültig. Für das Argument `set` sind nur die Schlüsselwerte `nfa`, `hatimeout` und `hasynctimeout` gültig.

Im Benutzerbereich arbeitende Systeme (z. B. Linux-Systeme):

- Alle Argumente sind gültig bis auf `configure` und `unconfigure`. Beachten Sie unbedingt, dass im System-Stack kein Aliasname für die Clusteradresse konfiguriert werden darf.
- Für das Argument `set` sind nur die Schlüsselwerte `nfa` und `hatimeout` gültig.
- Für das Argument `configure` müssen Sie an Stelle der *Netzmaske* die *Präfixlänge* angeben.

In IPV6 gibt die Präfixlänge die Anzahl der Bits im Netzabschnitt der IPV6-Adresse an. Durch die Präfixlänge ist erkennbar, wo die Netzadresse endet und die Hostadresse beginnt.

Bestimmen Sie für IPV4 wie folgt die Präfixlänge: Falls Ihre Teilnetzmaske 255.255.252.0 lautet, ist das Hexadezimaläquivalent FF.FF.FC.0. Der entsprechende Binärwert ist 11111111 11111111 11111100 00000000. Zählen Sie die Einsen in der Teilnetzmaske, um die Präfixlänge zu ermitteln. Wenn die Teilnetzmaske 22 Einsen hat, lautet das Präfix 22.

Syntax für `executor configure`:

`dscontrol executor configure Schnittstellenadresse Schnittstellename Präfixlänge`

Beispiel mit IPV6-Adressierung:

```
dscontrol executor configure 2002:092a:8f7a:4226:9:37:240:99 en0 112
```

Beispiel mit IPV4-Adressierung und der Teilnetzmaske 255.255.252.0:

```
dscontrol e config 191.60.20.20 en1 22
```

Beachten Sie unbedingt, dass der Befehl `executor configure` nicht für Installationen von Load Balancer für IPV4 und IPV6 auf Systemen verwendet wird, die im Benutzerbereich arbeiten, wie es bei Linux-Systemen der Fall ist.

- Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol executor — Executor steuern“ auf Seite 361.
- `dscontrol file`
 - Alle Argumente und ihre Schlüsselwerte sind gültig.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol file — Konfigurationsdateien verwalten“ auf Seite 366.
- `dscontrol help`
 - Alle Argumente sind gültig bis auf `host` (zum Konfigurieren einer fernen Maschine), `rule` (zum Konfigurieren von Regeln) und `subagent` (zum Konfigurieren des SNMP-Subagenten). Die Befehle `host`, `rule` und `subagent` werden nicht unterstützt.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol help — Hilfetext für diesen Befehl anzeigen oder drucken“ auf Seite 368.
- `dscontrol highavailability`
 - Alle Argumente sind gültig. Alle Schlüsselwerte sind gültig bis auf `both`, weil die gegenseitige hohe Verfügbarkeit nicht unterstützt wird.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol highavailability — Hohe Verfügbarkeit steuern“ auf Seite 369.
- `dscontrol logstatus`
 - Alle Argumente und ihre Schlüsselwerte sind gültig.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol logstatus — Protokolleinstellungen des Servers anzeigen“ auf Seite 374.
- `dscontrol manager`
 - Alle Argumente sind gültig bis auf `version`. Alle Schlüsselwerte sind gültig.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol manager — Manager steuern“ auf Seite 375.
- `dscontrol metric`
 - Alle Argumente und ihre Schlüsselwerte sind gültig.
 - Eine ausführliche Beschreibung der Befehlssyntax enthält der Abschnitt „dscontrol metric — Systemmesswerte konfigurieren“ auf Seite 381.
- `dscontrol port`
 - Es sind alle Argumente mit Ausnahme des nicht unterstützten Arguments `halfopenaddressreport` gültig.
Die folgenden Schlüsselwerte sind für die Argumente `add` und `set` des Befehls `dscontrol port` gültig:
 - `staletimeout`
 - `weightbound`
 - `stickymask`

Im Benutzerbereich arbeitende Systeme (z. B. Linux-Systeme): Für die Argumente `add` und `set` des Befehls `dscontrol port` sind die folgenden Schlüsselwerte gültig:

- `staletimeout`
- `weightbound`
- `selectionalgorithm`

Optionen für den Serverauswahlalgorithmus `selectionalgorithm`:

- `connection` – Der Server wird mit einer einfachen RoundRobin-Methode ausgewählt (Standardoption).
- `affinity` – Der Server wird ausgehend von der Client-Affinität ausgewählt.

Beispiel:

```
dscontrol port add Cluster@Port selectionalgorithm affinity
```

- Eine ausführliche Beschreibung der Befehlssyntax finden Sie im Abschnitt „`dscontrol port` — Ports konfigurieren“ auf Seite 382.

- `dscontrol server`

- Alle Argumente sind gültig.

Die folgenden Schlüsselwerte sind für das Argument `add` des Befehls `dscontrol server` gültig:

- `address`
- `advisorrequest`
- `advisorresponse`
- `collocated`

Das Schlüsselwort `collocated` ist unter allen unterstützten Betriebssystemen verfügbar, außer unter Windows und auf Systemen, die im Benutzerbereich arbeiten, z. B. auf Linux-Systemen.

- `fixedweight`
- `weight`

Die folgenden Schlüsselwerte sind für das Argument `set` des Befehls `dscontrol server` gültig:

- `advisorrequest`
- `advisorresponse`
- `collocated`

Das Schlüsselwort `collocated` ist unter allen unterstützten Betriebssystemen verfügbar, außer unter Windows und auf Systemen, die im Benutzerbereich arbeiten, z. B. auf Linux-Systemen.

- `fixedweight`
- `weight`

- Eine ausführliche Beschreibung der Befehlssyntax finden Sie im Abschnitt „`dscontrol server` — Server konfigurieren“ auf Seite 394.

- `dscontrol set`

- Alle Argumente und ihre Schlüsselwerte sind gültig.

- Eine ausführliche Beschreibung der Befehlssyntax finden Sie im Abschnitt „`dscontrol set` — Serverprotokoll konfigurieren“ auf Seite 400.

- `dscontrol status`
 - Alle Argumente und ihre Schlüsselwerte sind gültig.
 - Eine ausführliche Beschreibung der Befehlssyntax finden Sie im Abschnitt „`dscontrol status` — Aktivitätsanzeige für Manager und Advisor-Funktionen“ auf Seite 401.

Nicht unterstützte `dscontrol`-Befehle

Die folgenden Befehle sind für den Dispatcher von Load Balancer für IPV4 und IPV6 *nicht* verfügbar:

- `dscontrol host` (zum Konfigurieren einer fernen Maschine)
- `dscontrol rule` (zum Konfigurieren von Regeln)
- `dscontrol subagent` (zum Konfigurieren des SNMP-Subagenten)

Teil 3. CBR (Content Based Routing)

Dieser Teil enthält Informationen zu einer schnellen Erstkonfiguration sowie zur Planung und beschreibt die Konfigurationsmethoden für die Komponente CBR von Load Balancer. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 9, „Schnellkonfiguration“, auf Seite 99
- Kapitel 10, „Planung für Content Based Routing“, auf Seite 105
- Kapitel 11, „Content Based Routing konfigurieren“, auf Seite 109

Kapitel 9. Schnellkonfiguration

Dieses Beispiel zeigt die Konfiguration von drei lokal angeschlossenen Workstations, die CBR mit Caching Proxy verwenden, um den Webdatenverkehr auf zwei Webserver zu verteilen. (Der Einfachheit halber zeigt dieses Beispiel die Server innerhalb desselben LAN-Segments. Bei der Verwendung von CBR müssen sich die Server jedoch nicht in demselben LAN befinden.)

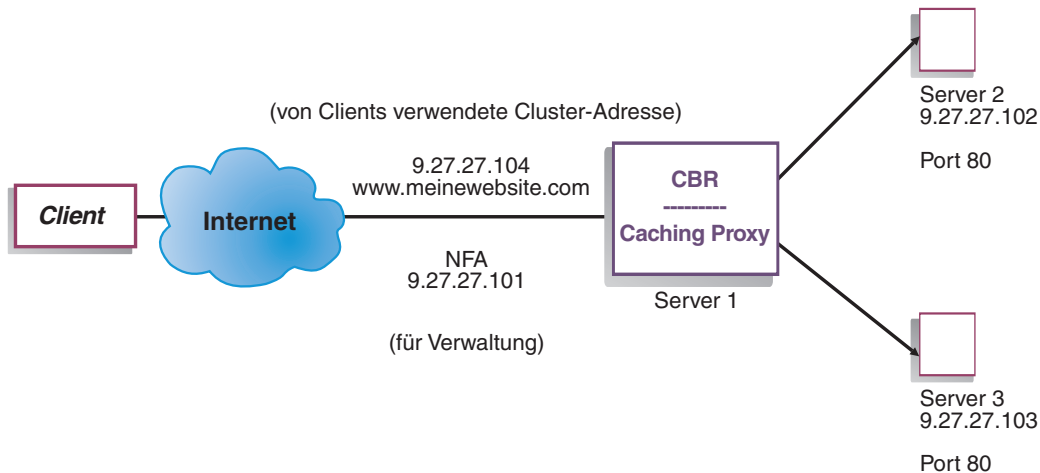


Abbildung 16. Einfache lokale CBR-Konfiguration

Voraussetzungen

In dem Beispiel für einen schnellen Start werden drei Workstations und vier IP-Adressen benötigt. Eine Workstation wird als CBR-Maschine verwendet; die beiden anderen Workstations werden als Webserver verwendet. Jeder Webserver benötigt eine IP-Adresse. Die CBR-Workstation benötigt eine eigene Adresse und eine Adresse für den Lastausgleich.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode `cbr` der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (`cbr`)“ auf Seite 55.

Für die Verwendung von CBR muss auf demselben Server Caching Proxy installiert sein. Informationen zum Konfigurieren von Caching Proxy für CBR finden Sie unter „Schritt 1. Caching Proxy für die Verwendung von CBR konfigurieren“ auf Seite 114.

Vorbereitungen

1. Konfigurieren Sie Ihre Workstations für dieses Beispiel so, dass sie sich innerhalb eines LAN-Segments befinden. Stellen Sie sicher, dass der Datenaustausch im Netz zwischen den drei Maschinen nicht über Router oder Brücken erfolgen muss.
2. Konfigurieren Sie die Netzwerkadapter der drei Workstations. In diesem Beispiel wird die folgende Netzkonfiguration angenommen:

Workstation	Name	IP-Adresse
1	server1.meinewebsite.com	9.27.27.101
2	server2.meinewebsite.com	9.27.27.102
3	server3.meinewebsite.com	9.27.27.103
Netzmaske = 255.255.255.0		

- Jede Workstation enthält nur eine Standard-Ethernet-Netzschnittstellenkarte.
3. Stellen Sie sicher, dass server1.meinewebsite.com ping-Aufrufe an server2.meinewebsite.com und server3.meinewebsite.com senden kann.
 4. Stellen Sie sicher, dass server2.meinewebsite.com und server3.meinewebsite.com ping-Aufrufe an server1.meinewebsite.com senden können.
 5. Stellen Sie sicher, dass die Webserver von server2.meinewebsite.com und server3.meinewebsite.com betriebsbereit sind. Fordern Sie mit einem Webbrowser Seiten direkt von **http://server2.meinewebsite.com** (z. B. .../member/index.html) und von **http://server3.meinewebsite.com** (z. B. .../guest/index.html) an.
 6. Definieren Sie eine andere gültige IP-Adresse für dieses LAN-Segment. Dies ist die Clusteradresse, die Sie den Clients zur Verfügung stellen, die auf Ihre Site zugreifen möchten. In diesem Beispiel wird folgende Adresse verwendet:
Name=www.meinewebsite.com
IP=9.27.27.104

CBR konfigurieren

Für CBR können Sie eine Konfiguration unter Verwendung der Befehlszeile, des Konfigurationsassistenten oder der grafischen Benutzerschnittstelle (GUI) erstellen. Dieses Beispiel für schnellen Start zeigt die Ausführung der Konfigurationsschritte in der Befehlszeile.

Anmerkung: Die Parameterwerte müssen mit Ausnahme der Parameterwerte für Hostnamen und Dateinamen in englischen Zeichen eingegeben werden.

Konfiguration von der Befehlszeile aus

Führen Sie an einer Eingabeaufforderung die folgenden Schritte aus:

1. Starten Sie den cbrserver. Führen Sie den folgenden Befehl als Benutzer root oder als Administrator aus: **cbrserver**

Anmerkung: Windows-Plattform: Starten Sie cbrserver (Content Based Routing) von der Anzeige "Dienste" aus, indem Sie nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste** klicken.

2. Starten Sie wie folgt die Executor-Funktion von CBR:
cbrcontrol executor start

3. Starten Sie Caching Proxy. (Caching Proxy kann nach dem Starten der Executor-Funktion jederzeit gestartet werden.)

ibmproxy

Anmerkung: Windows-Plattform: Sie können Caching Proxy auch von der Anzeige "Dienste" aus starten. Klicken Sie dazu nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**.

4. Fügen Sie wie folgt den Cluster (Hostname und Website, zu denen Clients eine Verbindung herstellen) zur CBR-Konfiguration hinzu:

cbrcontrol cluster add www.meinewebsite.com

5. Fügen Sie die Clusteradresse (9.27.27.104) für die Website zur Netz-schnittstellenkarte der CBR-Maschine hinzu. Weitere Informationen hierzu finden Sie unter „Schritt 5. Aliasnamen auf der Netzschnittstellenkarte erstellen (optional)“ auf Seite 116.

6. Fügen Sie wie folgt den Port für das Protokoll HTTP zur CBR-Konfiguration hinzu:

cbrcontrol port add www.meinewebsite.com:80

7. Fügen Sie wie folgt alle Webserver zur CBR-Konfiguration hinzu:

cbrcontrol server add www.meinewebsite.com:80:server2.meinewebsite.com

cbrcontrol server add www.meinewebsite.com:80:server3.meinewebsite.com

8. Fügen Sie content-Regeln zu Ihrer CBR-Konfiguration hinzu. (Mit einer content-Regel wird definiert, wie zwischen URL-Anforderungen unterschieden wird und wie eine Anforderung an einen der Server oder eine der Servergruppen gesendet wird.)

cbrcontrol rule add www.meinewebsite.com:80:memberRule type content pattern uri=*/member/*

cbrcontrol rule add www.meinewebsite.com:80:guestRule type content pattern uri=*/guest/*

In diesem Beispiel werden Client-Anforderungen an die Website www.meinewebsite.com bei Anwendung der content-Regel ausgehend von einem Verzeichnis in ihrem URI-Anforderungspfad an verschiedene Server gesendet. Weitere Informationen finden Sie in Anhang B, „Syntax der content-Regel“, auf Seite 477.

9. Fügen Sie wie folgt Server zu Ihren Regeln hinzu:

cbrcontrol rule useserver www.meinewebsite.com:80:memberRule server2.meinewebsite.com

cbrcontrol rule useserver www.meinewebsite.com:80:guestRule server3.meinewebsite.com

CBR führt den Lastausgleich jetzt ausgehend von der content-Regel durch. Ein Client mit einer URL-Anforderung, die **/member/** enthält, wird zu server2.meinewebsite.com dirigiert. Ein Client mit einer URL-Anforderung, die **/guest/** enthält, wird zu server3.meinewebsite.com dirigiert.

10. Starten Sie wie folgt die Manager-Funktion von CBR:

cbrcontrol manager start

11. Starten Sie wie folgt die Advisor-Funktion von CBR:

cbrcontrol advisor start http 80

CBR stellt jetzt sicher, dass keine Client-Anforderungen an einen ausgefallenen Webserver gesendet werden.

Die Basiskonfiguration mit lokal angeschlossenen Servern ist damit vollständig.

Konfiguration testen

Testen Sie wie folgt, ob die Konfiguration korrekt ist:

1. Rufen Sie mit einem Webbrowser die Seite **http://www.meinewebsite.com/member/index.htm** auf. Wenn eine Seite angezeigt wird, funktioniert die Konfiguration.
2. Laden Sie die Seite erneut im Webbrowser.
3. Überprüfen Sie die Ergebnisse des folgenden Befehls:
cbrcontrol server report www.meinewebsite.com:80:

Die Einträge der Spalte "Summe Verbindungen" für beide Server sollten addiert "2" ergeben.

Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus

Informationen zur Verwendung der CBR-GUI finden Sie im Abschnitt „GUI“ auf Seite 112 und in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Konfiguration mit dem Konfigurationsassistenten

Informationen zur Verwendung des CBR-Assistenten können Sie dem Abschnitt „Konfigurationsassistent“ auf Seite 113 entnehmen.

Arten von Cluster-, Port- und Serverkonfigurationen

Es gibt viele Möglichkeiten, CBR für die Unterstützung Ihrer Site zu konfigurieren. Wenn Sie für Ihre Site nur einen Hostnamen haben, zu dem alle Kunden eine Verbindung herstellen, können Sie einen Cluster mit Servern definieren. Für jeden dieser Server konfigurieren Sie einen Port, über den CBR kommuniziert. Vergleichen Sie hierzu Abb. 9 auf Seite 48.

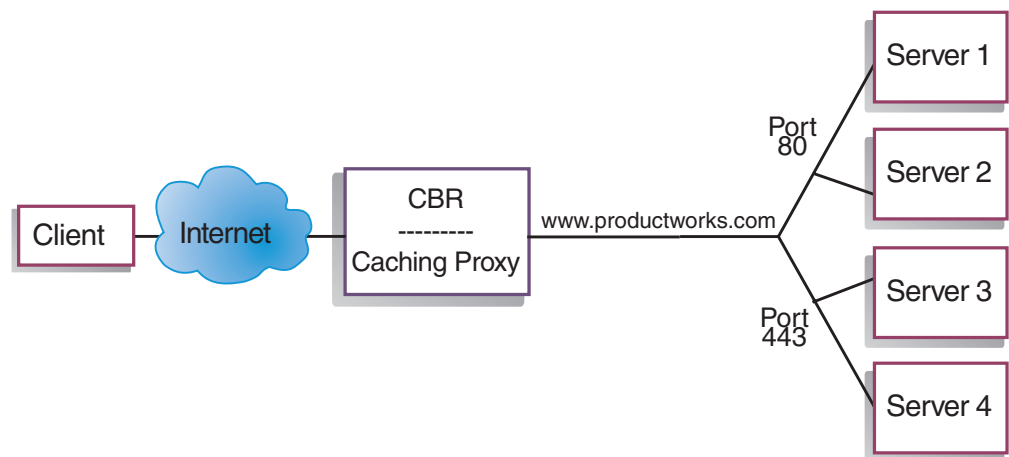


Abbildung 17. CBR-Beispielkonfiguration mit einem Cluster und zwei Ports

In diesem Beispiel ist für die Komponente CBR ein Cluster mit der Adresse **www-productworks.com** definiert. Dieser Cluster hat zwei Ports: Port 80 für HTTP und Port 443 für SSL.

Ein Client, der eine Anforderung an `http://www.productworks.com` (Port 80) richtet, wird einem anderen Server zugeordnet als ein Client, der eine Anforderung an `http://www.productworks.com` (Port 443) richtet.

Wenn Ihre Site sehr groß ist und Sie für jedes unterstützte Protokoll mehrere dedizierte Server haben, sollten Sie CBR auf andere Weise konfigurieren. In diesem Fall könnten Sie für jedes Protokoll einen Cluster mit nur einem Port, aber mehreren Servern definieren (siehe Abb. 10 auf Seite 49).

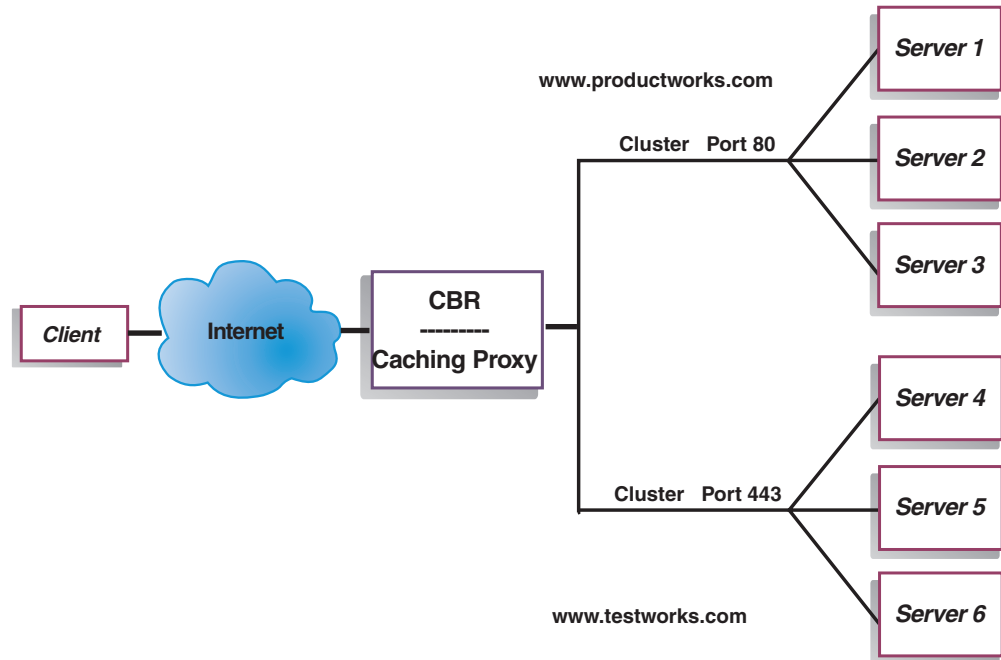


Abbildung 18. CBR-Beispielkonfiguration mit zwei Clustern mit jeweils einem Port

In diesem Beispiel für die Komponente CBR sind zwei Cluster definiert: `www.productworks.com` für Port 80 (HTTP) und `www.testworks.com` für Port 443 (SSL).

Wenn Ihre Site Inhalte für mehrere Unternehmen oder Abteilungen bereitstellt, die jeweils mit einem eigenen URL auf Ihre Site zugreifen, muss CBR auf eine dritte Art konfiguriert werden. In diesem Fall könnten Sie für jede Firma oder Abteilung einen Cluster definieren und anschließend die Ports, an denen Verbindungen mit dem jeweiligen URL empfangen werden sollen (siehe Abb. 11 auf Seite 50).

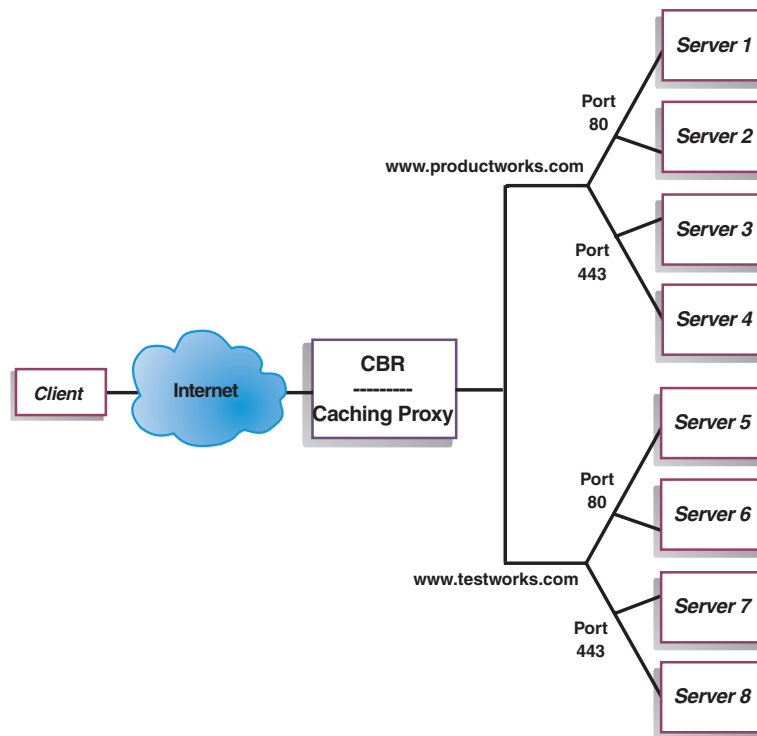


Abbildung 19. CBR-Beispielkonfiguration mit zwei Clustern mit jeweils zwei Ports

In diesem Beispiel für die Komponente CBR wurden für die Sites `www.productworks.com` und `www.testworks.com` jeweils zwei Cluster mit Port 80 (HTTP) und Port 443 (SSL) definiert.

Kapitel 10. Planung für Content Based Routing

In diesem Kapitel wird beschrieben, was die für die Planung des Netzes zuständige Person vor der Installation und Konfiguration der Komponente CBR mit Caching Proxy berücksichtigen muss.

- In Kapitel 3, „Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer“, auf Seite 19 finden Sie einen Überblick über die Features, die für die Verwaltung Ihres Netzes verfügbar sind.
- Informationen zum Konfigurieren der Lastausgleichsparameter von CBR finden Sie in Kapitel 11, „Content Based Routing konfigurieren“, auf Seite 109.
- Informationen zum Konfigurieren von Load Balancer für erweiterte Funktionen finden Sie in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Dieses Kapitel enthält die folgenden Abschnitte:

- „Überlegungen bei der Planung“
- „Regelbasierter Lastausgleich mit CBR“ auf Seite 107
- „Lastausgleich für sichere Verbindungen (SSL)“ auf Seite 107
- „Lastausgleich für SSL-Datenverkehr vom Client zum Proxy und HTTP-Datenverkehr vom Proxy zum Server“ auf Seite 108

Überlegungen bei der Planung

Mit der Komponente CBR können Sie unter Verwendung von Caching Proxy zum Weiterleiten der Anforderung HTTP- und SSL-Datenverkehr verteilen. Mit CBR können Sie einen Lastausgleich für Server durchführen, die Sie in der CBR-Konfigurationsdatei mit `cbrcontrol`-Befehlen konfiguriert haben.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode `cbr` der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

CBR ist dem Dispatcher hinsichtlich der Komponentenstruktur sehr ähnlich. CBR umfasst die folgenden Funktionen:

- **cbrserver** bearbeitet Anforderungen von der Befehlszeile an den Executor, den Manager und die Advisor-Funktionen.
- Der **Executor** unterstützt die Verteilung von Client-Anforderungen. Vor Verwendung der Komponente CBR muss der Executor gestartet sein.
- Der **Manager** definiert Wertigkeiten, die vom Executor verwendet werden und auf folgenden Kriterien basieren:
 - interne Zähler des Executors

- von den Advisor-Funktionen bereitgestellte Rückmeldungen von den Servern
- Rückmeldungen von einem Systemüberwachungsprogramm wie Metric Server.

Die Benutzung des Managers ist optional. Ohne den Manager wird der Lastausgleich nach einer gewichteten RoundRobin-Zeitplanung und ausgehend von den aktuellen Serverwertigkeiten durchgeführt. Es stehen keine Advisor-Funktionen zur Verfügung.

- Die **Advisor-Funktionen** richten Abfragen an die Server und analysieren die Ergebnisse nach Protokoll, bevor sie den Manager zum Festlegen der entsprechenden Wertigkeiten aufrufen. Es ist nicht immer sinnvoll, einige dieser Advisor-Funktionen in einer typischen Konfiguration zu verwenden. Sie können auch eigene Advisor-Funktionen schreiben. Die Benutzung der Advisor-Funktionen ist optional, wird jedoch empfohlen. Load Balancer stellt eine Caching-Proxy-Advisor-Funktion (cachingproxy) bereit. Weitere Informationen hierzu finden Sie im Abschnitt „Advisor-Funktionen“ auf Seite 185.
- Zum Konfigurieren und Verwalten des Executors, der Advisor-Funktionen und des Managers können Sie die Befehlszeile (**cbrcontrol**) oder die grafische Benutzerschnittstelle (**lbadmin**) verwenden.

Die drei wichtigsten Funktionen der Komponente CBR (Executor, Manager und Advisor-Funktionen) arbeiten gemeinsam an der Verteilung der eingehenden Anforderungen auf die Server. Neben dem Verteilen von Anforderungen überwacht der Executor die Anzahl neuer und aktiver Verbindungen. Diese Informationen stellt er anschließend dem Manager zur Verfügung.

Client-Anfragen nach verschiedenen Inhalten verteilen

Die Komponente CBR gibt Ihnen die Möglichkeit, eine Gruppe von Servern anzugeben, die eine Anforderung auf der Basis des Abgleichs eines regulären Ausdrucks mit dem Inhalt der Client-Anforderung bearbeiten. Mit CBR können Sie Ihre Site partitionieren, so dass verschiedene Inhalte oder Anwendungsdienste von unterschiedlichen Servergruppen bearbeitet werden. Diese Partitionierung ist für Clients, die auf Ihre Site zugreifen, transparent.

Siteinhalt für kürzere Antwortzeiten aufteilen

Eine Möglichkeit, Ihre Site zu partitionieren, besteht darin, einige Server ausschließlich für die Bearbeitung von cgi-Anforderungen und eine andere Gruppe von Servern für die Bearbeitung aller anderen Anforderungen zuzuordnen. Dies verhindert, dass die Server aufgrund der Verarbeitung umfangreicher cgi-Scripts für den normalen HTML-Datenverkehr zu langsam werden, und resultiert in einer insgesamt besseren Antwortzeit für die Clients. Mit Hilfe dieses Schemas könnten Sie auch leistungsstärkere Workstations für normale Anforderungen zuordnen. Dadurch würde die Antwortzeit für Clients verbessert, ohne dass alle Ihre Server aufgerüstet werden müssen. Sie könnten auch für cgi-Anforderungen leistungsstärkere Workstations zur Verfügung stellen.

Eine andere Möglichkeit zur Partitionierung Ihrer Site besteht darin, Clients, die auf Seiten mit erforderlicher Registrierung zugreifen, an eine Servergruppe zu verweisen, und alle anderen Anforderungen an eine zweite Servergruppe zu senden. Damit würde verhindert, dass Browser Ihrer Site Ressourcen binden, die von bereits registrierten Clients verwendet werden könnten. Außerdem könnten Sie leistungsstärkere Workstation verwenden, um Services für die registrierten Clients zur Verfügung zu stellen.

Sie könnten natürlich die oben genannten Methoden kombinieren, um eine noch größere Flexibilität und einen noch besseren Service zu erreichen.

Webserverinhalt sichern

Da CBR die Angabe mehrerer Server für jede Art von Anforderung zulässt, können die Anforderungen so verteilt werden, dass eine optimale Client-Antwortzeit erreicht wird. Aufgrund der Möglichkeit, jedem Inhaltstyp mehrere Server zuzuordnen, sind Sie abgesichert, wenn eine Workstation oder ein Server ausfällt. CBR erkennt den Ausfall und verteilt die Client-Anforderungen auf die übrigen Server der Gruppe.

CPU-Nutzung mit mehreren Caching-Proxy-Prozessen verbessern

Caching Proxy kommuniziert über die zugehörige Plug-in-Schnittstelle mit einem CBR-Prozess. Voraussetzung dafür ist, dass CBR auf der lokalen Maschine aktiv ist. Da dies zwei separate Prozesse sind, können mehrere Instanzen von Caching Proxy aktiv sein und mit einer Instanz von CBR zusammenarbeiten. Mit dieser Konfiguration können Sie die Adressen oder Funktionen unter den Caching Proxies aufteilen oder die Ressourcennutzung der Maschine verbessern, weil der Client-Datenverkehr von mehreren Caching Proxies bearbeitet wird. Die Proxy-Instanzen können, je nach den Erfordernissen des Datenverkehrs, an verschiedenen Ports empfangsbereit sein oder an eindeutige IP-Adressen eines Ports gebunden werden.

Regelbasierter Lastausgleich mit CBR

CBR überprüft zusammen mit Caching Proxy HTTP-Anforderungen anhand angegebener Regeltypen. Wenn Caching Proxy aktiv ist, akzeptiert es Client-Anforderungen und fragt bei CBR den besten Server an. Bei dieser Abfrage gleicht CBR die Anforderung mit einer Gruppe von Regeln mit bestimmten Prioritäten ab. Wenn eine Regel erfüllt ist, wird aus einer vorkonfigurierten Servergruppe ein geeigneter Server ausgewählt. Abschließend teilt CBR Caching Proxy mit, welcher Server ausgewählt wurde. Die Anforderung wird dann an diesen Server weitergeleitet.

Nachdem Sie einen Cluster für den Lastausgleich definiert haben, müssen Sie sicherstellen, dass es für alle Anforderungen an diesen Cluster eine Regel für die Auswahl eines Servers gibt. Wird keine Regel gefunden, die zu einer bestimmten Anforderung passt, empfängt der Client von Caching Proxy eine Fehlerseite. Das Erstellen einer in allen Fällen gültigen Regel ("always true") mit einer sehr hohen Prioritätsnummer ist der einfachste Weg zu gewährleisten, dass alle Anforderungen mit einer Regel übereinstimmen. Vergewissern Sie sich, dass die von dieser Regel verwendeten Server alle Anforderungen bearbeiten können, die nicht explizit von den Regeln mit einer kleineren Prioritätsnummer bearbeitet werden. (Anmerkung: Die Regeln mit kleinerer Prioritätsnummer werden zuerst ausgewertet.)

Weitere Informationen hierzu finden Sie im Abschnitt „Regelbasierten Lastausgleich konfigurieren“ auf Seite 212.

Lastausgleich für sichere Verbindungen (SSL)

CBR mit Caching Proxy kann SSL-Übertragungen vom Client zum Proxy empfangen und Übertragungen vom Proxy zu einem SSL-Server unterstützen. Wenn Sie für einen Server der CBR-Konfiguration einen SSL-Port für den Empfang der SSL-Anforderung vom Client definieren, können Sie den Datenverkehr mit CBR auf sichere Server (SSL-Server) verteilen und die Sicherheit Ihrer Site gewährleisten.

Zusätzlich zu den Änderungen an der Datei `ibmproxy.conf` für CBR müssen Sie zur Datei `ibmproxy.conf` für Caching Proxy eine Konfigurationsanweisung hinzufügen, um die SSL-Verschlüsselung für Datenverkehr vom Proxy zum Server zu aktivieren. Diese Anweisung muss das folgende Format haben:

```
proxy
uri-Muster url-Muster Adresse
```

Hier ist *uri-Muster* ein zu suchendes Muster (z. B. `/secure/*`), *url-Muster* ein Austausch-URL (z. B. `https://ClusterA/secure/*`) und *Adresse* die Clusteradresse (z. B. `ClusterA`).

Lastausgleich für SSL-Datenverkehr vom Client zum Proxy und HTTP-Datenverkehr vom Proxy zum Server

Die Komponente CBR mit Caching Proxy kann auch SSL-Übertragungen vom Client empfangen und die SSL-Anfrage vor der Weiterleitung an einen HTTP-Server entschlüsseln. Für den Befehl `cbrcontrol server` gibt es das optionale Schlüsselwort **mapport**, damit CBR SSL-Datenverkehr vom Client zum Proxy und HTTP-Datenverkehr vom Proxy zum Server unterstützen kann. Verwenden Sie dieses Schlüsselwort, wenn der Port auf dem Server ein anderer als der vom Client eingehende Port ist. Nachfolgend sehen Sie ein Beispiel für das Hinzufügen eines Ports mit dem Schlüsselwort `mapport`. Der Client-Port ist 443 (SSL) und der Server-Port 80 (HTTP):

```
cbrcontrol server add Cluster:443 mapport 80
```

Die Port-Nummer für `mapport` kann eine beliebige positive ganze Zahl sein. Die Standard-Port-Nummer ist der Wert des vom Client eingehenden Ports.

Da CBR in der Lage sein muss, Empfehlungen zu einer HTTP-Anforderung für einen am Port 443 (SSL) konfigurierten Server zu geben, gibt es die spezielle Advisor-Funktion `ssl2http`. Diese Advisor-Funktion wird an (dem vom Client eingehenden) Port 443 gestartet und gibt Empfehlungen zu den für diesen Port konfigurierten Servern. Wenn zwei Cluster konfiguriert sind und jeder der Cluster den Port 443 und die Server mit einem anderen `mapport` konfiguriert hat, kann eine Instanz der Advisor-Funktion den entsprechenden Port öffnen. Nachfolgend ist ein Beispiel dieser Konfiguration aufgeführt:

```
Executor
  Cluster1
    Port:443
      Server1 mapport 80
      Server2 mapport 8080
  Cluster2
    Port:443
      Server3 mapport 80
      Server4 mapport 8080
  Manager
    Advisor ssl2http 443
```

Kapitel 11. Content Based Routing konfigurieren

Lesen Sie vor Ausführung der in diesem Kapitel beschriebenen Schritte Kapitel 10, „Planung für Content Based Routing“, auf Seite 105. In diesem Kapitel wird erklärt, wie eine Basiskonfiguration für die Komponente CBR von Load Balancer erstellt wird.

- Komplexere Konfigurationen für Load Balancer finden Sie in Kapitel 21, „Manager, Advisor-Funktionen und Metric Server für Dispatcher, CBR und Site Selector“, auf Seite 179 und in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Konfigurations-Tasks im Überblick

Vergewissern Sie sich vor Ausführung der Konfigurationsschritte in dieser Tabelle, dass die CBR-Maschine und alle Servermaschinen mit dem Netz verbunden sind, gültige IP-Adressen haben und sich mit ping-Aufrufen erreichen können.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode `cbr` der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (`cbr`)“ auf Seite 55.

Tabelle 7. Konfigurations-Tasks für die Komponente CBR

Task	Beschreibung	Referenzinformationen
CBR-Maschine konfigurieren	Stellen Sie fest, welche Voraussetzungen zu erfüllen sind.	„CBR-Maschine konfigurieren“ auf Seite 113
Am Lastausgleich beteiligte Maschinen konfigurieren	Definieren Sie Ihre Lastausgleichskonfiguration.	„Schritt 7. Am Lastausgleich beteiligte Servermaschinen definieren“ auf Seite 117

Konfigurationsmethoden

Es gibt im Wesentlichen vier Methoden für das Erstellen einer Basiskonfiguration für die Komponente CBR von Load Balancer:

- Befehlszeile
- Scripts
- grafische Benutzerschnittstelle (GUI)
- Konfigurationsassistent.

Voraussetzung für die Verwendung von CBR ist die Installation von Caching Proxy.

Anmerkung: Caching Proxy ist ein Dienst, der nach der Installation standardmäßig automatisch gestartet wird. Vor dem Starten der CBR-Serverfunktion (cbrserver) müssen Sie Caching Proxy stoppen und den Dienst Caching Proxy so modifizieren, dass er manuell gestartet wird.

- Für Linux- und UNIX-Systeme: Stoppen Sie Caching Proxy. Stellen Sie dazu mit dem Befehl `ps -ef|grep ibmproxy` die Prozesskennung des Dienstes fest. Beenden Sie dann den Prozess mit dem Befehl `kill Prozess-ID`.
- Für Windows-Systeme: Stoppen Sie Caching Proxy im Fenster "Dienste".

Befehlszeile

Dies ist die direkte Methode für die Konfiguration von CBR. Die Werte der Befehlsparameter müssen in englischen Zeichen eingegeben werden. Die einzige Ausnahme hiervon bilden Hostnamen (die z. B. in den Befehlen `cluster` und `server` verwendet werden) und Dateinamen.

Starten Sie CBR wie folgt von der Befehlszeile aus:

- Auf Linux- oder UNIX-Systemen: Setzen Sie als Benutzer `root` an der Eingabeaufforderung den Befehl **cbrserver** ab. (Mit dem Befehl **cbrserver stop** können Sie den Dienst stoppen.)

Auf Windows-Systemen: Klicken Sie nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf **IBM Content Based Routing** und wählen Sie **Starten** aus. Zum Stoppen des Services müssen Sie dieselben Schritte ausführen und **Beenden** auswählen.

- Setzen Sie anschließend die gewünschten CBR-Steuerbefehle ab, um die Konfiguration einzurichten. Bei den in diesem Handbuch beschriebenen Prozeduren wird von der Verwendung der Befehlszeile ausgegangen. Der Befehl lautet **cbrcontrol**. Weitere Informationen zu Befehlen finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.
- Starten Sie Caching Proxy. Setzen Sie an der Eingabeaufforderung den Befehl **ibmproxy** ab. (Vor dem Starten von Caching Proxy müssen Sie den Executor starten.)

Anmerkung: Auf Windows-Plattformen können Sie Caching Proxy wie folgt in der Anzeige "Dienste" starten. Klicken Sie dazu auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**.

Sie können eine gekürzte Version der Parameter für den Befehl `cbrcontrol` eingeben. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **cbrcontrol he f** an Stelle von **cbrcontrol help file** angeben.

Wenn Sie die Befehlszeilenschnittstelle starten möchten, setzen Sie den Befehl **cbrcontrol** ab, um die Eingabeaufforderung `cbrcontrol` aufzurufen.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie den Befehl **exit** oder **quit** absetzen.

Anmerkungen:

1. Auf der Windows-Plattform wird dsserver für die Komponente Dispatcher automatisch gestartet. Falls Sie nur mit CBR und nicht mit der Komponente Dispatcher arbeiten, können Sie den automatischen Start von dsserver wie folgt unterbinden:
 - a. Klicken Sie im Fenster "Dienste" mit der rechten Maustaste auf "IBM Dispatcher".
 - b. Wählen Sie den Menüeintrag "Eigenschaften" aus.
 - c. Wählen Sie im Feld **Starttyp** die Option "Manuell" aus.
 - d. Klicken Sie auf OK und schließen Sie das Fenster "Dienste".
2. Wenn Sie Content Based Routing (CBR) nicht von der Eingabeaufforderung `cbrcontrol>>` aus, sondern lieber von der Eingabeaufforderung des Betriebssystems aus konfigurieren möchten, seien Sie bei Verwendung der folgenden Zeichen vorsichtig:
 - () linke und rechte runde Klammer
 - & Et-Zeichen
 - | vertikaler Balken
 - ! Ausrufezeichen
 - * Stern.

Die Shell des Betriebssystems könnte diese Zeichen als Sonderzeichen interpretieren und in alternativen Text konvertieren, bevor sie von `cbrcontrol` ausgewertet werden.

Die oben aufgelisteten Sonderzeichen sind optionale Zeichen für den Befehl **cbrcontrol rule add** und werden zum Angeben eines Musters für eine content-Regel verwendet. Der folgende Befehl ist deshalb unter Umständen nur bei Verwendung der Eingabeaufforderung `cbrcontrol>>` gültig.

```
rule add 10.1.203.4:80:cbr_prod_rule_ek type content
pattern uri=/nipoek/*
```

Wenn dieser Befehl an der Eingabeaufforderung des Betriebssystems funktionieren soll, müssen Sie das Muster wie folgt in Anführungszeichen setzen:

```
cbrcontrol rule add 10.1.203.4:80:cbr_prod_rule_ek type content
pattern "uri=/nipoek/*"
```

Fehlen die Anführungszeichen, könnte beim Speichern der Regel in CBR ein Teil des Musters abgeschnitten werden. An der Eingabeaufforderung `cbrcontrol>>` wird die Verwendung von Anführungszeichen nicht unterstützt.

Scripts

Die Befehle zum Konfigurieren von CBR können in eine Konfigurations-Script-Datei eingegeben und dann zusammen ausgeführt werden.

Anmerkung: Falls Sie den Inhalt einer Script-Datei (z. B. `meinScript`) schnell ausführen möchten, verwenden Sie einen der folgenden Befehle:

- Führen Sie zum Aktualisieren der derzeitigen Konfiguration die folgenden ausführbaren Befehle Ihrer Script-Datei aus:
cbrcontrol file appendload *meinScript*
- Führen Sie zum vollständigen Ersetzen der derzeitigen Konfiguration die folgenden ausführbaren Befehle Ihrer Script-Datei aus:
cbrcontrol file newload *meinScript*

Führen Sie den folgenden Befehl aus, um die aktuelle Konfiguration in einer Script-Datei (z. B. *sicherungsscript*) zu speichern:

```
cbrcontrol file save sicherungsscript
```

Dieser Befehl speichert die Script-Datei mit der Konfiguration im Verzeichnis **...ibm/edge/lb/servers/configurations/cbr**.

GUI

Abb. 41 auf Seite 470 zeigt ein Beispiel für die grafische Benutzerschnittstelle (GUI) mit allgemeinen Anweisungen.

Gehen Sie zum Starten der GUI wie folgt vor:

1. Vergewissern Sie sich, dass **cbrserver** aktiv ist. Setzen Sie an einer Eingabeaufforderung als Benutzer **root** oder **Administrator** den Befehl **cbrserver** ab.
2. Führen Sie je nach Betriebssystem eine der folgenden Aktionen aus:
 - Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Geben Sie **lbadmin** ein.
 - Auf Windows-Systemen: Klicken Sie nacheinander auf **Start > Programme > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**.
3. Starten Sie Caching Proxy. (Wenn Sie die GUI verwenden, müssen Sie zunächst eine Hostverbindung herstellen und vor dem Start von Caching Proxy den Executor für die Komponente CBR starten.) Führen Sie einen der folgenden Schritte aus:
 - Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Geben Sie zum Starten von Caching Proxy **ibmproxy** ein.
 - Auf Windows-Systemen: Rufen Sie die Anzeige "Dienste" auf, indem Sie nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste** klicken.

Zum Konfigurieren der Komponente CBR von der GUI aus müssen Sie zunächst in der Baumstruktur **Content Based Routing** auswählen. Sie können den Manager starten, nachdem Sie eine Verbindung zu einem Host hergestellt haben. Sie können auch Cluster mit Ports und Servern erstellen und Advisor-Funktionen für den Manager starten.

Mit der GUI können Sie dieselben Tasks wie mit dem Befehl **cbrcontrol** ausführen. Wenn Sie beispielsweise einen Cluster von der Befehlszeile aus konfigurieren möchten, müssten Sie den Befehl **cbrcontrol cluster add Cluster** eingeben. Zum Definieren eines Clusters von der GUI aus müssen Sie mit der rechten Maustaste auf "Executor" klicken und in dem daraufhin angezeigten Popup-Menü mit der linken Maustaste auf **Cluster hinzufügen**. Geben Sie die Clusteradresse in das Dialogfenster ein und klicken Sie dann auf **OK**.

Bereits vorhandene CBR-Konfigurationsdateien können Sie mit der im Popup-Menü **Host** angezeigten Option **Neue Konfiguration laden** (zum vollständigen Ersetzen der derzeitigen Konfiguration) oder **An aktuelle Konfiguration anfügen** (zum Aktualisieren der derzeitigen Konfiguration) laden. Sie sollten Ihre CBR-Konfiguration von Zeit zu Zeit mit der Option **Konfigurationsdatei sichern unter** in einer Datei sichern. Diese Option ist ebenfalls im Popup-Menü **Host** enthalten. Das oben auf der GUI befindliche Menü **Datei** bietet Ihnen die Möglichkeit, die aktuellen Hostverbindungen in einer Datei zu speichern oder Verbindungen aus vorhandenen Dateien für alle Komponenten von Load Balancer wiederherzustellen.

Sie können auf **Hilfe** zugreifen, indem Sie auf das Fragezeichen in der oberen rechten Ecke des Fensters von Load Balancer klicken.

- **Hilfe: Feldebene** — beschreibt jedes Feld und gibt die Standardwerte an.
- **Hilfe: Wie funktioniert das** — listet Tasks auf, die von dieser Anzeige aus ausgeführt werden können.
- **InfoCenter** — ermöglicht den zentralen Zugriff auf Produktinformationen.

Wenn Sie von der GUI aus einen Befehl ausführen möchten, gehen Sie wie folgt vor: Heben Sie in der GUI-Baumstruktur den Hostknoten hervor und wählen Sie im Popup-Menü "Host" **Befehl senden...** aus. Geben Sie im Befehlseingabefeld den gewünschten Befehl ein, z. B. **executor report**. In einem Fenster sehen Sie die Ergebnisse und die Historie der in der aktuellen Sitzung ausgeführten Befehle.

Weitere Informationen zur Verwendung der GUI finden Sie in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Konfigurationsassistent

Führen Sie folgende Schritte aus, wenn Sie den Konfigurationsassistenten verwenden:

1. Starten Sie cbrserver, indem Sie an der Eingabeaufforderung als root oder Administrator den Befehl **cbrserver** absetzen.
2. Starten Sie wie folgt die Assistentenfunktion von CBR:
Starten Sie den Assistenten von der Eingabeaufforderung aus, indem Sie den Befehl **cbrwizard** absetzen. Sie können den Konfigurationsassistenten auch im CBR-Komponentenmenü auswählen, das auf der GUI angezeigt wird.
3. Starten Sie Caching Proxy zum Verteilen des HTTP- oder HTTPS-Datenverkehrs (SSL).
Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Geben Sie zum Starten von Caching Proxy **ibmproxy** ein.
Auf Windows-Systemen: Rufen Sie die Anzeige "Dienste" auf, indem Sie nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste** klicken.

Der CBR-Assistent führt Sie schrittweise durch den Prozess zum Erstellen einer Basiskonfiguration für die Komponente CBR. Der Assistent stellt Ihnen Fragen bezüglich Ihres Netzes und führt Sie durch die Konfiguration eines Clusters, mit dem CBR den Datenverkehr auf eine Gruppe von Servern verteilen kann.

CBR-Maschine konfigurieren

Vor dem Konfigurieren der Maschine mit CBR müssen Sie (auf AIX-, HP-UX-, Linux- oder Solaris-Systemen) als Benutzer root oder (auf Windows-Systemen) als Administrator registriert sein.

Sie benötigen für jeden zu konfigurierenden Servercluster eine IP-Adresse. Eine Clusteradresse ist eine Adresse, die einem Hostnamen zugeordnet ist (beispielsweise www.firma.com). Diese IP-Adresse wird von einem Client benutzt, um die Verbindung zu den Servern in einem Cluster herzustellen. Diese Adresse ist in der URL-Anforderung von dem Client enthalten. CBR verteilt alle Anforderungen, die an dieselbe Clusteradresse gerichtet sind.

Nur für Solaris-Systeme: Vor Verwendung der Komponente CBR müssen die Systemstandardwerte für die prozessübergreifende Kommunikation (Inter-process Communication) geändert werden. Die maximale Größe des gemeinsam benutzten Speichersegments und die Anzahl von Semaphore-Kennungen müssen erhöht werden. Sie können Ihr System auf die Unterstützung für CBR einstellen, indem Sie die Datei `/etc/system` auf Ihrem System editieren und die folgenden Anweisungen hinzufügen, bevor Sie dann einen Warmstart durchführen:

```
set shmsys:shminfo_shmmax=0x02000000
set semsys:seminfo_semmap=750
set semsys:seminfo_semmni=30
set semsys:seminfo_semmns=750
set semsys:seminfo_semmnu=30
set semsys:seminfo_semtime=30
```

Wenn Sie das gemeinsam benutzte Speichersegment nicht auf die oben gezeigten Werte vergrößern, kann der Befehl **cbrcontrol executor start** nicht ausgeführt werden.

Schritt 1. Caching Proxy für die Verwendung von CBR konfigurieren

Voraussetzung für die Verwendung von CBR ist die Installation von Caching Proxy.

Anmerkung: Caching Proxy ist ein Dienst, der nach der Installation standardmäßig automatisch gestartet wird. Vor dem Starten der CBR-Serverfunktion müssen Sie Caching Proxy stoppen und den Dienst Caching Proxy so modifizieren, dass er manuell gestartet wird.

- Für AIX-, HP-UX-, Linux- und Solaris-Systeme: Stoppen Sie Caching Proxy. Stellen Sie dazu mit dem Befehl `ps -ef|grep ibmproxy` die Prozesskennung des Dienstes fest. Beenden Sie dann den Prozess mit dem Befehl `kill Prozess-ID`.
- Für Windows-Systeme: Stoppen Sie Caching Proxy im Fenster "Dienste".

Die Konfigurationsdatei für Caching Proxy (`ibmproxy.conf`) müssen Sie wie folgt ändern:

Vergewissern Sie sich, dass die eingehende URL-Anweisung **CacheByIncomingUrl** auf `off` (Standardeinstellung) gesetzt ist.

Fügen Sie in der Konfigurationsdatei für jeden Cluster zum Abschnitt mit den Zuordnungsregel eine ähnliche Zuordnungsregel wie die folgende hinzu:

```
Proxy /* http://cluster.domain.com/* cluster.domain.com
```

Anmerkung: CBR legt das Protokoll, den Server und den Ziel-Port später fest.

Für das CBR-Plug-in müssen Sie vier Einträge editieren:

- `ServerInit`
- `PostAuth`
- `PostExit`
- `ServerTerm`

Jeder Eintrag muss sich jeweils in einer neuen Zeile befinden. Die Datei `ibmproxy.conf` enthält mehrere Einträge "ServerInit", einen für jedes Plug-in. Die Einträge für das CBR-Plug-in müssen editiert werden. Außerdem müssen Sie das Kommentarzeichen löschen.

Nachfolgend sehen Sie die spezifischen Zusätze zur Konfigurationsdatei für die einzelnen Betriebssysteme.

Abbildung 20. CBR-Konfigurationsdatei für AIX-, Linux- und Solaris-Systeme

```
ServerInit /opt/ibm/edge/lb/servers/lib/liblbcbr.so:ndServerInit
PostAuth /opt/ibm/edge/lb/servers/lib/liblbcbr.so:ndPostAuth
PostExit /opt/ibm/edge/lb/servers/lib/liblbcbr.so:ndPostExit
ServerTerm /opt/ibm/edge/lb/servers/lib/liblbcbr.so:ndServerTerm
```

Abbildung 21. CBR-Konfigurationsdatei für HP-UX-Systeme

```
ServerInit /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndServerInit
PostAuth /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndPostAuth
PostExit /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndPostExit
ServerTerm /opt/ibm/edge/lb/servers/lib/liblbcbr.sl:ndServerTerm
```

Abbildung 22. CBR-Konfigurationsdatei für Windows-Systeme

```
ServerInit C:\Programme\IBM\edge\lb\servers\lib\liblbcbr.dll:ndServerInit
PostAuth C:\Programme\IBM\edge\lb\servers\lib\liblbcbr.dll:ndPostAuth
PostExit C:\Programme\IBM\edge\lb\servers\lib\liblbcbr.dll:ndPostExit
ServerTerm C:\Programme\IBM\edge\lb\servers\lib\liblbcbr.dll:ndServerTerm
```

Schritt 2. Serverfunktion starten

Geben Sie zum Starten der CBR-Serverfunktion in der Befehlszeile **cbrserver** ein.

Eine Standardkonfigurationsdatei (`default.cfg`) wird beim Starten von `cbrserver` automatisch geladen. Wenn Sie die CBR-Konfiguration in `default.cfg` sichern, werden alle in dieser Datei gesicherten Angaben beim nächsten Starten von `cbrserver` automatisch geladen.

Schritt 3. Executor-Funktion starten

Geben Sie zum Starten der Executor-Funktion den Befehl **cbrcontrol executor start** ein. Sie können jetzt auch verschiedene Executor-Einstellungen ändern. Lesen Sie hierzu die Informationen im Abschnitt „`dscontrol executor` — Executor steuern“ auf Seite 361.

Schritt 4. Cluster definieren und Clusteroptionen festlegen

CBR verteilt die an den Cluster gesendeten Anforderungen auf die entsprechenden Server, die für die Ports dieses Clusters konfiguriert wurden.

Der Cluster ist der symbolische Name im Hostabschnitt des URL und muss mit dem in der Proxy-Anweisung der Datei `ibmproxy.conf` verwendeten Namen übereinstimmen.

In CBR definierte Cluster sollten so definiert werden, dass sie mit der eingehenden Anforderung übereinstimmen. Ein Cluster muss mit dem Hostnamen oder der IP-Adresse definiert werden, der bzw. die in der eingehenden Anforderung enthalten ist. Wird die Anforderung beispielsweise mit der IP-Adresse eingegeben, muss der Cluster mit der IP-Adresse definiert sein. Falls mehrere Hostnamen in eine IP-Adresse aufgelöst werden (und Anforderungen mit jedem dieser Hostnamen eingehen können), sollten alle Hostnamen als Cluster definiert sein.

Setzen Sie zum Definieren eines Clusters den folgenden Befehl ab:

```
cbrcontrol cluster add Cluster
```

Setzen Sie zum Festlegen von Clusteroptionen den folgenden Befehl ab:

```
cbrcontrol cluster set Wert_der_Clusteroption
```

Weitere Informationen hierzu finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Schritt 5. Aliasnamen auf der Netzschnittstellenkarte erstellen (optional)

Wenn Sie Caching Proxy als Reverse Proxy konfiguriert haben und einen Lastausgleich für mehrere Websites durchführen, müssen Sie die Clusteradresse jeder Website zu mindestens einer Netzschnittstellenkarte der Load-Balancer-Maschine hinzufügen. Andernfalls kann dieser Schritt übergangen werden.

Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Fügen Sie die Clusteradresse mit dem Befehl `ifconfig` zur Netzschnittstellenkarte hinzu. Den Befehl für das von Ihnen verwendete Betriebssystem können Sie Tabelle 8 entnehmen.

Tabelle 8. Befehle zum Erstellen eines Aliasnamens auf der NIC

AIX	<code>ifconfig Schnittstellenname alias Clusteradresse netmask Netzmaske</code>
HP-UX	<code>ifconfig Schnittstellenname Clusteradresse netmask Netzmaske up</code>
Linux	<code>ifconfig Schnittstellenname Clusteradresse netmask Netzmaske up</code>
Solaris 8, Solaris 9 und Solaris 10	<code>ifconfig Schnittstellenname addif Clusteradresse netmask Netzmaske up</code>

Anmerkung: Auf Linux- und HP-UX-Systemen muss der *Schnittstellenname* für jede hinzugefügte Clusteradresse eine eindeutige Nummer haben, z. B. `eth0:1`, `eth0:2` usw.

Für Windows 2000: Gehen Sie zum Hinzufügen der Clusteradresse zur Netzschnittstelle wie folgt vor:

1. Klicken Sie nacheinander auf **Start > Einstellungen > Systemsteuerung**.
2. Klicken Sie doppelt auf **Netzwerk- und DFÜ-Verbindungen**.

3. Klicken Sie mit der rechten Maustaste auf **LAN-Verbindung**.
4. Wählen Sie den Eintrag **Eigenschaften** aus.
5. Wählen Sie **Internetprotokoll (TCP/IP)** aus und klicken Sie auf **Eigenschaften**.
6. Wählen Sie **Folgende IP-Adresse verwenden** aus und klicken Sie auf **Erweitert**.
7. Klicken Sie auf **Hinzufügen**. Geben Sie dann die **IP-Adresse** und die **Subnetzmaske** für den Cluster ein.

Für Windows 2003: Gehen Sie zum Hinzufügen der Clusteradresse zur Netz-schnittstelle wie folgt vor:

1. Klicken Sie auf **Start > Systemsteuerung > Netzwerkverbindungen > LAN-Verbindung**.
2. Klicken Sie auf **Eigenschaften**.
3. Wählen Sie **Internetprotokoll (TCP/IP)** aus und klicken Sie auf **Eigenschaften**.
4. Wählen Sie **Folgende IP-Adresse verwenden** aus und klicken Sie auf **Erweitert**.
5. Klicken Sie auf **Hinzufügen**. Geben Sie dann die IP-Adresse und die Subnetzmaske für den Cluster ein.

Schritt 6. Ports definieren und Port-Optionen festlegen

Die Port-Nummer bezeichnet den Port, an dem die Serveranwendungen empfangsbereit sind. Für HTTP-Datenverkehr ist dies bei Verwendung von CBR mit Caching Proxy in der Regel Port 80.

Setzen Sie den folgenden Befehl ab, um für den im vorherigen Schritt definierten Cluster einen Port zu definieren:

```
cbrcontrol port add Cluster:Port
```

Setzen Sie zum Festlegen von Port-Optionen den folgenden Befehl ab:

```
cbrcontrol port set Cluster:Wert_der_Port-Option
```

Weitere Informationen hierzu finden Sie in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349.

Schritt 7. Am Lastausgleich beteiligte Servermaschinen definieren

Die Servermaschinen sind die Maschinen, auf denen die Anwendungen ausgeführt werden, deren Last verteilt werden soll. Für den *Server* wird der symbolische Name der Servermaschine oder deren Adresse in Schreibweise mit Trennzeichen angegeben. Setzen Sie den folgenden Befehl ab, um für Cluster und Port einen Server zu definieren:

```
cbrcontrol server add Cluster:Port:Server
```

Für einen Cluster müssen Sie pro Port mehrere Server definieren, um einen Lastausgleich durchführen zu können.

Schritt 8. Regeln zur Konfiguration hinzufügen

Dies ist der wichtigste Schritt beim Konfigurieren von CBR mit Caching Proxy. Mit einer Regel wird definiert, wie zwischen URL-Anforderungen unterschieden wird und wie eine Anforderung an die entsprechende Gruppe von Servern gesendet wird. Der spezielle von CBR verwendete Regeltyp ist 'content'. Setzen Sie zum Definieren einer content-Regel den folgenden Befehl ab:

```
cbrcontrol rule  
add Cluster:Port:Regel type content pattern Muster
```

Der Wert *Muster* ist der reguläre Ausdruck, der mit dem URL in den einzelnen Client-Anforderungen verglichen wird. Weitere Informationen über zum Konfigurieren des Musters finden Sie in Anhang B, „Syntax der content-Regel“, auf Seite 477.

Einige der anderen in Dispatcher definierten Regeltypen können ebenfalls für CBR verwendet werden. Weitere Informationen hierzu finden Sie im Abschnitt „Regelbasierten Lastausgleich konfigurieren“ auf Seite 212.

Schritt 9. Server zu den Regeln hinzufügen

Wenn für eine Client-Anforderung eine Übereinstimmung mit einer Regel gefunden wird, wird bei der der Regel zugeordneten Servergruppe der beste Server abgefragt. Die der Regel zugeordnete Servergruppe ist eine Untergruppe der Server, die für den Port definiert sind. Setzen Sie den folgenden Befehl ab, um Server zur Servergruppe einer Regel hinzuzufügen:

```
cbrcontrol rule useserver Cluster:Port:Regel server
```

Schritt 10. Manager-Funktion starten (optional)

Die Manager-Funktion verbessert den Lastausgleich. Setzen Sie zum Starten des Managers den folgenden Befehl ab:

```
cbrcontrol manager start
```

Schritt 11. Advisor-Funktion starten (optional)

Die Advisor-Funktionen liefern dem Manager weitere Informationen über die Fähigkeit der am Lastausgleich beteiligten Servermaschinen, auf Anforderungen zu antworten. Advisor-Funktionen sind protokollspezifisch. Soll beispielsweise die HTTP-Advisor-Funktion gestartet werden, setzen Sie den folgenden Befehl ab:

```
cbrcontrol advisor start http Port
```

Schritt 12. Clusterproportionen festlegen

Wenn Sie Advisor-Funktionen starten, können Sie die Wichtigkeit ändern, die in Entscheidungen für den Lastausgleich einfließenden Informationen von Advisor-Funktionen beigemessen wird. Setzen Sie zum Festlegen von Clusterproportionen den Befehl **cbrcontrol cluster set *Cluster* proportions** ab. Weitere Informationen hierzu finden Sie im Abschnitt „Proportionale Bedeutung von Statusinformationen“ auf Seite 180.

Schritt 13. Caching Proxy starten

- AIX-Systeme: Fügen Sie Folgendes zur Umgebungsvariablen LIBPATH hinzu:
/opt/ibm/edge/lb/servers/lib
- Linux-, HP-UX- oder Solaris-Systeme: Fügen Sie Folgendes zur Umgebungsvariablen LD_LIBRARY_PATH hinzu:
/opt/ibm/edge/lb/servers/lib
- Windows-Systeme: Fügen Sie Folgendes zur Umgebungsvariablen PATH hinzu:
C:\Programme\IBM\edge\lb\servers\lib

Starten Sie Caching Proxy in der neuen Umgebung, indem Sie an der Eingabeaufforderung den Befehl **ibmproxy** absetzen.

Anmerkung: Auf Windows-Systemen können Sie Caching Proxy wie folgt in der Anzeige "Dienste" starten. Klicken Sie dazu auf **Start -> Einstellungen** (für Windows 2000) -> **Systemsteuerung -> Verwaltung -> Dienste**.

CBR-Konfigurationsbeispiel

Führen Sie die folgenden Schritte aus, um CBR zu konfigurieren:

1. Starten Sie CBR durch Absetzen des Befehls **cbrserver**.
2. Starten Sie die Befehlszeilenschnittstelle. Setzen Sie dazu den Befehl **cbrcontrol** ab.
3. Die Eingabeaufforderung **cbrcontrol** wird angezeigt. Setzen Sie die folgenden Befehle ab (*Cluster(c),Port(p),Regel(r),Server(s)*)
 - :executor start
 - cluster add c
 - port add c:p
 - server add c:p:s
 - rule add c:p:r type content pattern uri=*
 - rule useserver c:p:r s
4. Starten Sie Caching Proxy durch Absetzen des Befehls **ibmproxy**. (Auf der Windows-Plattform müssen Sie Caching Proxy von der Anzeige "Dienste" aus starten.)
5. Entfernen Sie alle Proxy-Konfigurationen aus dem Browser.
6. Laden Sie <http://c/> in Ihren Browser. Hier steht "c" für den Cluster, den Sie zuvor konfiguriert haben.
 - Server "s" wird aufgerufen.
 - Es wird die Webseite <http://s/> angezeigt.

Teil 4. Site Selector

Dieser Teil enthält Informationen zu einer schnellen Erstkonfiguration sowie zur Planung und beschreibt die Konfigurationsmethoden für die Komponente Site Selector von Load Balancer. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 12, „Schnellkonfiguration“, auf Seite 123
- Kapitel 13, „Planung für Site Selector“, auf Seite 127
- Kapitel 14, „Site Selector konfigurieren“, auf Seite 131

Kapitel 12. Schnellkonfiguration

Dieses Beispiel für einen schnellen Start zeigt das Erstellen einer Sitekonfiguration, bei der Site Selector den Datenverkehr ausgehend von dem in einer Client-Anfrage verwendeten Domännennamen auf die Server einer Gruppe verteilt.

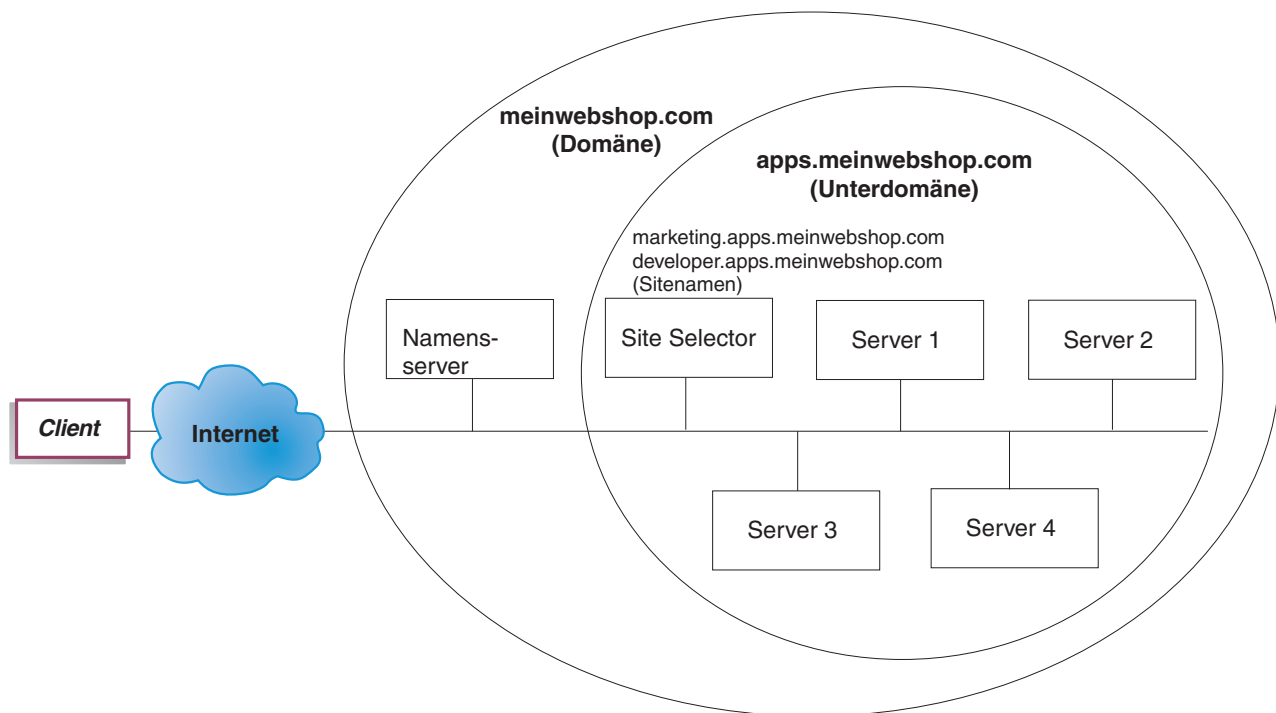


Abbildung 23. Einfache Site-Selector-Konfiguration

Voraussetzungen

Für dieses Beispiel für schnellen Start benötigen Sie Folgendes:

- Administratorzugriff auf den Namensserver Ihrer Site.
- Vier Server (Server 1, Server 2, Server 3, Server 4), die für das Netz konfiguriert sind, und einen zusätzlichen Server, auf dem Site Selector installiert ist.

Anmerkung: Falls Sie Site Selector mit einem der am Lastausgleich beteiligten Server verknüpfen, benötigen Sie nicht fünf, sondern nur vier Server. Die Verknüpfung (Kollokation) wirkt sich jedoch auf den Durchsatz der am Lastausgleich beteiligten Server aus.

Vorbereitungen

In diesem Beispiel für schnellen Start ist die Domäne der Firmensite **meinwebshop.com**. Site Selector ist für eine Unterdomäne innerhalb von **meinwebshop.com** verantwortlich. Sie müssen daher eine Unterdomäne von **meinwebshop.com** definieren. Beispiel: **apps.meinwebshop.com**. Site Selector ist kein vollständig implementierter DNS wie BIND und fungiert in einer DNS-Hierarchie als Blattknoten. Site Selector ist für die Unterdomäne **apps.meinwebshop.com** autoritativ.

Die Unterdomäne apps.meinwebshop.com umfasst die folgenden Sitenamen: marketing.apps.meinwebshop.com und developer.apps.meinwebshop.com.

1. Aktualisieren Sie den Domännennamensserver der Firmensite (vergleichen Sie hierzu Abb. 23 auf Seite 123). Erstellen Sie in der Datei named.data einen Namensservereintrag für die Unterdomäne (apps.meinwebshop.com) mit Site Selector als autoritativem Namensserver:
apps.meinwebshop.com. IN NS siteselector.meinwebshop.com
2. Vergewissern Sie sich, dass der vollständig qualifizierte Hostname oder die Site im aktuellen Domännennamenssystem nicht aufgelöst wird.
3. Installieren Sie auf den Servern (Server 1, Server 2, Server 3, Server 4), für die Site Selector den Lastausgleich durchführen soll, Metric Server. Weitere Informationen hierzu finden Sie im Abschnitt „Metric Server“ auf Seite 196.

Site Selector konfigurieren

Für Site Selector können Sie eine Konfiguration unter Verwendung der Befehlszeile, des Konfigurationsassistenten oder der grafischen Benutzerschnittstelle (GUI) erstellen. Dieses Beispiel für schnellen Start zeigt die Ausführung der Konfigurationsschritte in der Befehlszeile.

Anmerkung: Die Parameterwerte müssen mit Ausnahme der Parameterwerte für Hostnamen und Dateinamen in englischen Zeichen eingegeben werden.

Konfiguration von der Befehlszeile aus

Führen Sie an einer Eingabeaufforderung die folgenden Schritte aus:

1. Starten Sie den ssserver auf der Maschine, die das Hosting für Site Selector übernimmt. Setzen Sie an einer Eingabeaufforderung als Benutzer root oder Administrator den Befehl **ssserver** ab.

Anmerkung: Windows-Plattform: Starten Sie ssserver (IBM Site Selector) von der Anzeige "Dienste" aus, indem Sie nacheinander auf **Start** > **Einstellungen** (für Windows 2000) > **Systemsteuerung** > **Verwaltung** > **Dienste** klicken.

2. Starten Sie wie folgt den Namensserver für die Site-Selector-Konfiguration:
sscontrol nameserver start
3. Konfigurieren Sie für Site Selector die Sitenamen (marketing.apps.meinwebshop.com und developer.apps.meinwebshop.com):
sscontrol sitename add marketing.apps.meinwebshop.com
sscontrol sitename add developer.apps.meinwebshop.com
4. Fügen Sie die Server zur Site-Selector-Konfiguration hinzu. (Konfigurieren Sie Server 1 und Server 2 für den Sitenamen marketing.apps.meinwebshop.com. Konfigurieren Sie Server 3 und Server 4 für den Sitenamen developer.apps.meinwebshop.com):
sscontrol server add marketing.apps.meinwebshop.com:server1+server2
sscontrol server add developer.apps.meinwebshop.com:server3+server4
5. Starten Sie wie folgt die Manager-Funktion von Site Selector:
sscontrol manager start

6. Starten Sie die Advisor-Funktion von Site Selector (HTTP-Advisor-Funktion für marketing.apps.meinwebshop.com und FTP-Advisor-Funktion für developer.apps.meinwebshop):

sscontrol advisor start http marketing.apps.meinwebshop.com:80

sscontrol advisor start ftp developer.apps.meinwebshop.com:21

Site Selector stellt jetzt sicher, dass keine Client-Anforderungen an einen ausgefallenen Server gesendet werden.

7. Vergewissern Sie sich, dass auf allen am Lastausgleich beteiligten Servern Metric Server gestartet wurde.

Die Basiskonfiguration für Site Selector ist damit vollständig.

Konfiguration testen

Testen Sie wie folgt, ob die Konfiguration korrekt ist:

1. Versuchen Sie von einem Client aus, auf dem ein primärer DNS als verantwortlicher Namensserver für meinwebshop.com konfiguriert ist, mit ping einen der konfigurierten Sitenamen zu erreichen.
2. Stellen Sie eine Verbindung zur Anwendung her. Beispiel:
 - Öffnen Sie einen Browser und fordern Sie marketing.apps.meinwebshop.com an. Daraufhin sollte eine gültige Seite bereitgestellt werden.
 - Öffnen Sie einen FTP-Client zu developer.apps.meinwebshop.com und geben Sie einen gültigen Benutzernamen mit Kennwort ein.
3. Überprüfen Sie die Ergebnisse des folgenden Befehls:

sscontrol server status marketing.apps.meinwebshop.com:

sscontrol server status developer.apps.meinwebshop.com:

Der Eintrag unter "Summe Treffer" müsste für jeden Server den abgesetzten ping- und Anwendungsanforderungen entsprechen.

Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus

Informationen zur Verwendung der Site-Selector-GUI finden Sie im Abschnitt „GUI“ auf Seite 133 und in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Konfiguration mit dem Konfigurationsassistenten

Informationen zur Verwendung des Site-Selector-Assistenten können Sie dem Abschnitt „Konfigurationsassistent“ auf Seite 134 entnehmen.

Kapitel 13. Planung für Site Selector

In diesem Kapitel wird beschrieben, was die für die Planung des Netzes zuständige Person vor der Installation und Konfiguration der Komponente Site Selector berücksichtigen muss.

- In Kapitel 3, „Netz verwalten: Bestimmen der erforderlichen Features von Load Balancer“, auf Seite 19 finden Sie einen Überblick über die Features, die für die Verwaltung Ihres Netzes verfügbar sind.
- Informationen zum Konfigurieren der Lastausgleichparameter von Site Selector finden Sie in Kapitel 14, „Site Selector konfigurieren“, auf Seite 131.
- Informationen zum Konfigurieren von Load Balancer für erweiterte Funktionen finden Sie in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Dieses Kapitel enthält die folgenden Abschnitte:

- „Überlegungen bei der Planung“
- „Hinweise zu TTL“ auf Seite 129
- „Netzproximität verwenden“ auf Seite 130

Überlegungen bei der Planung

Site Selector verteilt zusammen mit einem Domänennamensserver die Last auf eine Gruppe von Servern. Dazu verwendet Site Selector erfasste Messwerte und Wertigkeiten. Sie können eine Sitekonfiguration erstellen, bei der die Last innerhalb einer Servergruppe auf der Grundlage des für eine Client-Anfrage verwendeten Domänennamens verteilt wird.

Einschränkungen: Site Selector unterstützt nur DNS-Abfragen des Typs A. Bei allen anderen Abfragetypen erscheint der Rückkehrcode NOTIMPL (nicht implementiert). Falls eine ganze Domäne an Site Selector delegiert wird, stellen Sie sicher, dass die Domäne nur Abfragen des Typs A empfängt.

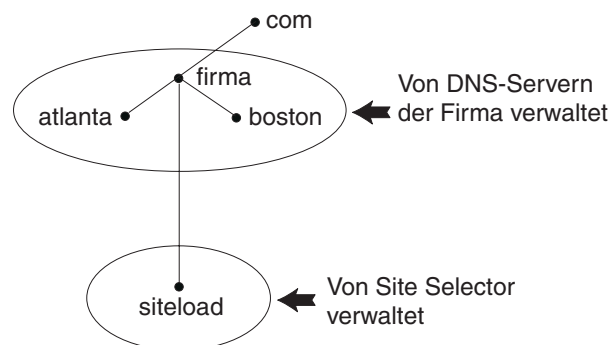


Abbildung 24. Beispiel für eine DNS-Umgebung

Wenn Sie innerhalb Ihrer DNS-Umgebung eine Unterdomäne für Site Selector einrichten, sollte Site Selector die Berechtigung für diese Unterdomäne haben. Beispiel (siehe Abb. 24 auf Seite 127): Ihre Firma hat die Berechtigung für die Domäne **firma.com** erhalten. Innerhalb der Firma gibt es mehrere Unterdomänen. Site Selector hätte in diesem Fall die Berechtigung für **siteload.firma.com** und die DNS-Server hätten weiterhin die Berechtigung für **atlanta.firma.com** und **boston.firma.com**.

Damit der Namensserver der Firma erkennt, dass Site Selector die Berechtigung für die Unterdomäne siteload hat, muss zur benannten Datendatei für den Server ein Namensservereintrag hinzugefügt werden. Auf AIX-Systemen würde ein solcher Namensservereintrag etwa wie folgt aussehen:

```
siteload.firma.com. IN NS siteselector.firma.com.
```

Hier ist **siteselector.firma.com** der Hostname der Site-Selector-Maschine. In allen anderen benannten Datendateien für DNS-Server sind äquivalente Einträge erforderlich.

Ein Client fordert die Auflösung eines Domänennamens bei einem Namensserver innerhalb seines Netzes an. Der Namensserver leitet die Anforderung an die Site-Selector-Maschine weiter. Site Selector löst den Domänennamen dann in die IP-Adresse eines der Server auf, die für den Sitenamen konfiguriert wurden. Anschließend gibt Site Selector die IP-Adresse des ausgewählten Servers an den Namensserver zurück. Der Namensserver liefert die IP-Adresse an den Client. (Site Selector arbeitet als nicht rekursiver Namensserver (Blattknotenserver) und meldet einen Fehler, wenn die Domänennamensanforderung nicht aufgelöst werden kann.)

Abb. 5 auf Seite 14 veranschaulicht eine Site, bei der Site Selector zusammen mit einem DNS-System die Last auf lokale und ferne Server verteilt.

Site Selector stellt die folgenden Funktionen bereit:

- Der **sss**server bearbeitet Anforderungen von der Befehlszeile an den Namensserver, den Manager und die Advisor-Funktionen.
- Die **Namensserver**-Funktion unterstützt die Verteilung eingehender Namensserveranforderungen. Eine DNS-Auflösung ist erst möglich, wenn Sie die Namensserverfunktion für Site Selector gestartet haben. Site Selector ist am Port 53 bereit, eingehende DNS-Anforderungen zu empfangen. Wenn der Name der anfragenden Site konfiguriert ist, gibt Site Selector eine Serveradresse (aus einer Gruppe von Serveradressen) zurück, die dem Sitenamen zugeordnet ist.
- Der **Manager** definiert Wertigkeiten, die vom Namensserver benutzt werden und auf folgenden Kriterien basieren:
 - von den Advisor-Funktionen bereitgestellte Rückmeldungen von den Servern
 - Rückmeldungen von einem Systemüberwachungsprogramm wie Metric Server.

Die Benutzung des Managers ist optional. Ohne den Manager wird der Lastausgleich nach einer gewichteten RoundRobin-Zeitplanung und ausgehend von den aktuellen Serverwertigkeiten durchgeführt. Es stehen keine Advisor-Funktionen zur Verfügung.

- **Metric Server** ist eine Load-Balancer-Komponente zur Systemüberwachung, die auf der Back-End-Servermaschine installiert wird. (Wenn Sie Load Balancer mit einer Servermaschine verknüpfen, für die ein Lastausgleich durchgeführt wird, müssen Sie Metric Server auf der Maschine mit Load Balancer installieren.)

Mit Metric Server kann Site Selector den Grad der Aktivität eines Servers überwachen, den Server mit der geringsten Auslastung feststellen und einen ausgefallenen Server erkennen. Die Last ist ein Maß für das Arbeitsaufkommen eines Servers. Der Administrator des Systems mit Site Selector steuert die Art der Lastmessung. Sie können Site Selector an die Anforderungen der eigenen Umgebung anpassen und dabei Faktoren wie die Zugriffshäufigkeit, die Gesamtzahl der Benutzer und die Zugriffsarten (beispielsweise kurze Abfragen, lange Abfragen, Transaktionen mit hoher CPU-Belastung) berücksichtigen.

Der Lastausgleich wird auf der Basis von Serverwertigkeiten vorgenommen. Für Site Selector gibt es vier Proportionen, die der Manager zur Ermittlung der Wertigkeiten verwendet:

- CPU
- Speicher
- Port
- System

Alle CPU- und Speicherwerte werden von Metric Server bereitgestellt. Wenn Sie mit Site Selector arbeiten, sollten Sie demzufolge auch Metric Server verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Metric Server“ auf Seite 196.

- Die **Advisor-Funktionen** richten Abfragen an die Server und analysieren die Ergebnisse nach Protokoll, bevor sie den Manager zum Festlegen der entsprechenden Wertigkeiten aufrufen. Es ist nicht immer sinnvoll, einige dieser Advisor-Funktionen in einer typischen Konfiguration zu verwenden. Sie können auch eigene Advisor-Funktionen schreiben. Die Benutzung der Advisor-Funktionen ist optional, wird jedoch empfohlen. Weitere Informationen hierzu finden Sie im Abschnitt „Advisor-Funktionen“ auf Seite 185.
- Zum Konfigurieren und Verwalten des Namensservers, der Advisor-Funktionen, von Metric Server und des Managers können Sie die Befehlszeile (**sscontrol**) oder die grafische Benutzerschnittstelle (**lbadmin**) verwenden.

Die vier wichtigsten Funktionen von Site Selector (Namensserver, Manager, Metric Server und Advisor-Funktionen) interagieren, um die eingehenden Anforderungen auf die Server zu verteilen und aufzulösen.

Hinweise zu TTL

Der DNS-gestützte Lastausgleich erfordert, dass Namensauflösungen zwischengespeichert werden können. Der TTL-Wert (Time To Live) bestimmt die Effizienz des DNS-gestützten Lastausgleichs. TTL legt fest, wie lange ein anderer Namensserver die aufgelöste Antwort zwischenspeichert. Bei einem kleinen TTL-Wert können geringfügige Änderungen der Server- oder Netzlast schneller realisiert werden. Wird die Zwischenspeicherung inaktiviert, müssen die Clients sich mit jeder Namensauflösungsanforderung an den maßgeblichen Namensserver wenden, was potenziell die Latenzzeit erhöht. Bei der Auswahl eines TTL-Wertes sollte sorgfältig abgewogen werden, welchen Einfluss das Inaktivieren der Zwischenspeicherung auf eine Umgebung hat. Es ist auch zu bedenken, dass der DNS-gestützte Lastausgleich potenziell von der Zwischenspeicherung von Namensauflösungen auf dem Client eingeschränkt werden kann.

TTL kann mit dem Befehl **sscontrol sitename [add | set]** konfiguriert werden. Weitere Informationen hierzu finden Sie im Abschnitt „sscontrol sitename — Sitennamen konfigurieren“ auf Seite 428.

Netzproximität verwenden

Die Netzproximität ist die Berechnung der Nähe jedes einzelnen Servers zum anfordernden Client. Zum Bestimmen der Netzproximität sendet der Agent Metric Server (der auf jedem Server mit Lastausgleich installiert sein muss) ein ping an die Client-IP-Adresse und meldet Site Selector die Antwortzeit. Site Selector bezieht die Proximitätsantwort in die Lastausgleichsentscheidung ein. Site Selector kombiniert den Wert der Netzproximitätsantwort mit der Wertigkeit vom Manager und ermittelt so die endgültige Wertigkeit für den Server.

Die Verwendung der Netzproximität mit Site Selector ist optional.

Site Selector stellt die folgenden Netzproximitätsoptionen bereit, die pro Sitenamen festgelegt werden können:

- Cache-Lebensdauer: Die Zeitperiode, während der eine Proximitätsantwort gültig und im Cache gespeichert ist.
- Prozentsatz für Proximität: Die Bedeutung der Proximitätsantwort, gemessen am Zustand des Servers (vom Manager vorgegebene Wertigkeit).
- Auf alle warten: Legt fest, ob vor der Beantwortung der Client-Anfrage auf alle Proximitätsantworten (ping-Antworten) der Server gewartet werden soll.

Ist dieser Wert auf **ja** gesetzt, sendet Metric Server ein ping an den Client, um die Zeit für die Proximitätsantwort zu ermitteln. Der Namensserver wartet auf die Antworten aller Metric-Server oder das Eintreten einer Zeitlimit-überschreitung. Anschließend erstellt der Namensserver für jeden Server aus der Zeit für die Proximitätsantwort und der vom Manager berechneten Wertigkeit eine kombinierte Wertigkeit. Site Selector teilt dem Client die Server-IP-Adresse mit der besten kombinierten Wertigkeit mit. (Es wird davon ausgegangen, dass die meisten Client-Namensserver ein Zeitlimit von 5 Sekunden haben. Site Selector versucht, vor Ablauf dieses Zeitlimits zu antworten.)

Ist dieser Wert auf **nein** gesetzt, erfolgt die Namensauflösung für den Client auf der Basis der aktuellen Wertigkeiten vom Manager. Anschließend sendet Metric Server ein ping an den Client, um die Zeit für die Proximitätsantwort zu ermitteln. Der Namensserver stellt die von Metric Server empfangene Antwortzeit in den Cache. Wenn der Client eine zweite Anforderung stellt, erstellt der Namensserver für jeden Server aus der aktuellen Wertigkeit vom Manager und dem zwischengespeicherten Wert der ping-Antwort eine kombinierte Wertigkeit. Site Selector gibt auf die zweite Anforderung des Clients die IP-Adresse des Servers mit der besten kombinierten Wertigkeit zurück.

Optionen für die Netzproximität können mit dem Befehl **sscontrol sitename [add | set]** gesetzt werden. Weitere Informationen hierzu finden Sie in Kapitel 28, „Befehlsreferenz für Site Selector“, auf Seite 405.

Kapitel 14. Site Selector konfigurieren

Lesen Sie vor Ausführung der in diesem Kapitel beschriebenen Schritte Kapitel 13, „Planung für Site Selector“, auf Seite 127. Dieses Kapitel erläutert das Erstellen einer Basiskonfiguration für die Komponente Site Selector von Load Balancer.

- Komplexere Konfigurationen für Load Balancer finden Sie in Kapitel 21, „Manager, Advisor-Funktionen und Metric Server für Dispatcher, CBR und Site Selector“, auf Seite 179 und in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Konfigurations-Tasks im Überblick

Anmerkung: Vergewissern Sie sich vor Ausführung der Konfigurationsschritte in dieser Tabelle, dass die Site Selector-Maschine und alle Servermaschinen mit dem Netz verbunden sind, gültige IP-Adressen haben und sich gegenseitig mit ping-Aufrufen erreichen können.

Tabelle 9. Konfigurations-Tasks für Site Selector

Task	Beschreibung	Referenzinformationen
Maschine mit Site Selector konfigurieren.	Stellen Sie fest, welche Voraussetzungen zu erfüllen sind.	„Maschine mit Site Selector konfigurieren“ auf Seite 134
Am Lastausgleich beteiligte Maschinen konfigurieren.	Definieren Sie Ihre Lastausgleichskonfiguration.	„Schritt 4. Am Lastausgleich beteiligte Servermaschinen definieren“ auf Seite 135

Konfigurationsmethoden

Es gibt im Wesentlichen vier Methoden für das Erstellen einer Basiskonfiguration für die Komponente Site Selector von Load Balancer:

- Befehlszeile
- Scripts
- grafische Benutzerschnittstelle (GUI)
- Konfigurationsassistent.

Befehlszeile

Dies ist die direkte Methode für die Konfiguration von Site Selector. Die Werte der Befehlsparameter müssen in englischen Zeichen eingegeben werden. Die einzige Ausnahme hiervon bilden Hostnamen (die z. B. in den Befehlen sitename und server verwendet werden) und Dateinamen.

Starten Sie Site Selector wie folgt von der Befehlszeile aus:

1. Setzen Sie an der Eingabeaufforderung den Befehl **sssserver** ab. Geben Sie zum Stoppen des Services **sssserver stop** ein.

Anmerkung: Klicken Sie auf Windows-Systemen nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf **IBM Site Selector** und wählen Sie **Starten** aus. Zum Stoppen des Services müssen Sie dieselben Schritte ausführen und **Beenden** auswählen.

2. Setzen Sie anschließend die gewünschten Steuerbefehle für Site Selector ab, um die Konfiguration einzurichten. Bei den in diesem Handbuch beschriebenen Prozeduren wird von der Verwendung der Befehlszeile ausgegangen. Der Befehl lautet **sscontrol**. Weitere Informationen zu Befehlen finden Sie in Kapitel 28, „Befehlsreferenz für Site Selector“, auf Seite 405.

Sie können eine Minimalversion der Parameter für den Befehl **sscontrol** eingeben. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **sscontrol he f** an Stelle von **sscontrol help file** eingeben.

Wenn Sie die Befehlszeilenschnittstelle starten möchten, setzen Sie den Befehl **sscontrol** ab, um die Eingabeaufforderung **sscontrol** aufzurufen.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie den Befehl **exit** oder **quit** absetzen.

Anmerkung: Auf der Windows-Plattform wird **dsserver** für die Komponente Dispatcher automatisch gestartet. Falls Sie nur mit Site Selector und nicht mit der Komponente Dispatcher arbeiten, können Sie den automatischen Start von **dsserver** wie folgt unterbinden:

1. Klicken Sie im Fenster "Dienste" von Windows mit der rechten Maustaste auf "IBM Dispatcher".
2. Wählen Sie den Menüeintrag "Eigenschaften" aus.
3. Wählen Sie im Feld **Starttyp** die Option "Manuell" aus.
4. Klicken Sie auf OK und schließen Sie das Fenster "Dienste".

Scripts

Die Befehle zum Konfigurieren von Site Selector können in eine Konfigurations-Script-Datei eingegeben und dann zusammen ausgeführt werden.

Anmerkung: Falls Sie den Inhalt einer Script-Datei (z. B. *meinScript*) schnell ausführen möchten, verwenden Sie einen der folgenden Befehle:

- Führen Sie zum Aktualisieren der derzeitigen Konfiguration die ausführbaren Befehle Ihrer Script-Datei mit folgendem Befehl aus:
sscontrol file appendload *meinScript*
- Führen Sie zum vollständigen Ersetzen der derzeitigen Konfiguration die ausführbaren Befehle Ihrer Script-Datei mit folgendem Befehl aus:
sscontrol file newload *meinScript*

Führen Sie den folgenden Befehl aus, um die aktuelle Konfiguration in einer Script-Datei (z. B. *sicherungsscript*) zu speichern:

sscontrol file save *sicherungsscript*

Dieser Befehl speichert die Script-Datei mit der Konfiguration im Verzeichnis **...ibm/edge/lb/servers/configurations/ss**.

GUI

Abb. 41 auf Seite 470 zeigt ein Beispiel für die GUI mit allgemeinen Anweisungen.

Gehen Sie zum Starten der GUI wie folgt vor:

1. Vergewissern Sie sich, dass ssserver aktiv ist. Setzen Sie an einer Eingabeaufforderung als Benutzer root oder Administrator den Befehl **ssserver** ab.
2. Führen Sie anschließend einen der folgenden Schritte aus:
 - Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Geben Sie **lbadmin** ein.
 - Auf Windows-Systemen: Klicken Sie nacheinander auf **Start > Programme > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**.

Zum Konfigurieren von Site Selector auf der GUI müssen Sie zunächst in der Baumstruktur **Site Selector** auswählen. Nachdem Sie eine Verbindung zu einem Host hergestellt haben, auf dem ssserver ausgeführt wird, können Sie Sitenamen mit Servern erstellen und den Manager sowie Advisor-Funktionen starten.

Von der GUI aus können Sie die gleichen Schritte wie mit dem Befehl **sscontrol** ausführen. Wenn Sie beispielsweise einen Sitenamen von der Befehlszeile aus definieren möchten, müssen Sie den Befehl **sscontrol sitename add Sitename** eingeben. Zum Definieren eines Sitenamens von der GUI aus müssen Sie mit der rechten Maustaste auf "Namensserver" klicken und in dem daraufhin angezeigten Popup-Menü mit der linken Maustaste auf **Sitenamen hinzufügen**. Geben Sie im Dialogfenster den Sitenamen ein und klicken Sie auf **OK**.

Bereits vorhandene Site-Selector-Konfigurationsdateien können Sie mit der im Popup-Menü **Host** angezeigten Option **Neue Konfiguration laden** (zum vollständigen Ersetzen der derzeitigen Konfiguration) oder **An aktuelle Konfiguration anfügen** (zum Aktualisieren der derzeitigen Konfiguration) laden. Sie sollten Ihre Site-Selector-Konfiguration von Zeit zu Zeit mit der Option **Konfigurationsdatei sichern unter** in einer Datei sichern. Diese Option ist ebenfalls im Popup-Menü **Host** enthalten. Das oben auf der GUI befindliche Menü **Datei** bietet Ihnen die Möglichkeit, die aktuellen Hostverbindungen in einer Datei zu speichern oder Verbindungen aus vorhandenen Dateien für alle Komponenten von Load Balancer wiederherzustellen.

Wenn Sie von der GUI aus einen Befehl ausführen möchten, gehen Sie wie folgt vor: Heben Sie in der GUI-Baumstruktur den Hostknoten hervor und wählen Sie im Popup-Menü "Host" **Befehl senden...** aus. Geben Sie im Befehlseingabefeld den gewünschten Befehl ein, z. B. **nameserver status**. In einem Fenster sehen Sie die Ergebnisse und die Historie der in der aktuellen Sitzung ausgeführten Befehle.

Sie können auf **Hilfe** zugreifen, indem Sie auf das Fragezeichen in der oberen rechten Ecke des Fensters von Load Balancer klicken.

- **Hilfe: Feldebene** — beschreibt jedes Feld und gibt die Standardwerte an.
- **Hilfe: Wie funktioniert das** — listet Tasks auf, die von dieser Anzeige aus ausgeführt werden können.
- **InfoCenter** — ermöglicht den zentralen Zugriff auf Produktinformationen.

Weitere Informationen zur Verwendung der GUI finden Sie in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Konfigurationsassistent

Führen Sie folgende Schritte aus, wenn Sie den Konfigurationsassistenten verwenden:

1. Starten Sie wie folgt den `ssserver` für Site Selector:
 - Führen Sie den folgenden Befehl als Benutzer `root` oder als Administrator aus:

`ssserver`

2. Starten Sie den Assistenten von Site Selector, **`sswizard`**.

Sie können den Assistenten von der Eingabeaufforderung aus starten, indem Sie den Befehl **`sswizard`** absetzen. Sie können den Konfigurationsassistenten aber auch auf der GUI unter der Komponente Site Selector auswählen.

Der Site-Selector-Assistent führt Sie schrittweise durch den Prozess zum Erstellen einer Basiskonfiguration für die Komponente Site Selector. Er stellt Ihnen Fragen zu Ihrem Netz und leitet Sie beim Konfigurieren eines Sitenamens an, mit dem Site Selector den Datenverkehr auf eine Gruppe von Servern verteilen kann.

Maschine mit Site Selector konfigurieren

Vor dem Konfigurieren der Maschine mit Site Selector müssen Sie (auf AIX-, HP-UX-, Linux- oder Solaris-Systemen) als Benutzer `root` oder (auf Windows-Systemen) als Administrator registriert sein.

Für eine Gruppe von Servern, die Sie konfigurieren, benötigen Sie einen nicht auflösbaren vollständig qualifizierten Hostnamen als Sitenamen. Der Sitename ist der Name, mit dem Clients auf Ihre Site zugreifen (z. B. `www.IhreFirma.com`). Site Selector verteilt mit dem DNS den Datenverkehr für die Site auf die Server der Gruppe.

Schritt 1. Serverfunktion starten

Geben Sie zum Starten der Site-Selector-Serverfunktion in der Befehlszeile **`ssserver`** ein.

Anmerkung: Eine Standardkonfigurationsdatei (`default.cfg`) wird beim Starten von `ssserver` automatisch geladen. Wenn Sie die Konfiguration in `default.cfg` sichern, werden alle in dieser Datei gesicherten Angaben beim nächsten Starten von `ssserver` automatisch geladen.

Schritt 2. Namensserver starten

Geben Sie zum Starten des Namensservers den Befehl **`sscontrol nameserver start`** ein.

Optional können Sie den Namensserver starten, indem Sie ihn mit dem Schlüsselwort `bindaddress` ausschließlich an die angegebene Adresse binden.

Schritt 3. Sitenamen definieren und Optionen für Sitenamen festlegen

Site Selector verteilt die an den Sitenamen gesendeten Anforderungen auf die entsprechenden Server, die für die Site konfiguriert sind.

Der Sitename ist ein nicht auflösbarer Hostname, den der Client anfordert. Der Sitename muss ein vollständig qualifizierter Domänenname sein (z. B. `www.dns-`

download.com). Wenn ein Client diesen Sitenamen anfordert, wird eine der dem Sitenamen zugeordneten Server-IP-Adressen zurückgegeben.

Setzen Sie zum Definieren eines Sitenamens den folgenden Befehl ab:

```
sscontrol sitename add Sitename
```

Wenn Sie Optionen für den Sitenamen festlegen möchten, setzen Sie den folgenden Befehl ab:

```
sscontrol sitename  
set Wert_der_Sitenamenoption
```

Weitere Informationen hierzu finden Sie in Kapitel 28, „Befehlsreferenz für Site Selector“, auf Seite 405.

Schritt 4. Am Lastausgleich beteiligte Servermaschinen definieren

Die Servermaschinen sind die Maschinen, auf denen die Anwendungen ausgeführt werden, deren Last verteilt werden soll. Für den *Server* wird der symbolische Name der Servermaschine oder deren Adresse in Schreibweise mit Trennzeichen angegeben. Setzen Sie den folgenden Befehl ab, um für den Sitenamen von Schritt 3 einen Server zu definieren:

```
sscontrol server add Sitename:Server
```

Für einen Sitenamen müssen Sie mehrere Server definieren, um einen Lastausgleich durchführen zu können.

Schritt 5. Manager-Funktion starten (optional)

Die Manager-Funktion ergänzt den Lastausgleich. Vergewissern Sie sich vor dem Starten der Manager-Funktion, dass auf allen am Lastausgleich beteiligten Maschinen Metric Server installiert ist.

Setzen Sie zum Starten des Managers den folgenden Befehl ab:

```
sscontrol manager start
```

Schritt 6. Advisor-Funktion starten (optional)

Die Advisor-Funktionen liefern dem Manager weitere Informationen über die Fähigkeit der am Lastausgleich beteiligten Servermaschinen, auf Anforderungen zu antworten. Advisor-Funktionen sind protokollspezifisch. Load Balancer stellt zahlreiche Advisor-Funktionen bereit. Wenn Sie beispielsweise die HTTP-Advisor-Funktion für einen bestimmten Sitenamen starten möchten, setzen Sie den folgenden Befehl ab:

```
sscontrol advisor start http Sitename:Port
```

Schritt 7. Systemmesswert definieren (optional)

Informationen zur Verwendung von Systemmesswerten und Metric Server finden Sie im Abschnitt „Metric Server“ auf Seite 196.

Schritt 8. Proportionen für den Sitenamen festlegen

Wenn Sie Advisor-Funktionen starten, können Sie die Wichtigkeit ändern, die in Entscheidungen für den Lastausgleich einfließenden Informationen von Advisor-Funktionen (Port-Informationen) beigemessen wird. Setzen Sie zum Festlegen der Proportionen für den Sitenamen den Befehl **sscontrol sitename set *Sitename* pro-**

portions ab. Weitere Informationen hierzu finden Sie im Abschnitt „Proportionale Bedeutung von Statusinformationen“ auf Seite 180.

Servermaschinen für Lastausgleich konfigurieren

Verwenden Sie Metric Server zusammen mit der Komponente Site Selector. Informationen zum Konfigurieren von Metric Server auf allen Maschinen, für die Site Selector einen Lastausgleich durchführt, finden Sie im Abschnitt „Metric Server“ auf Seite 196.

Teil 5. Cisco CSS Controller

Dieser Teil enthält Informationen zu einer schnellen Erstkonfiguration sowie zur Planung und beschreibt die Konfigurationsmethoden für die Komponente Cisco CSS Controller von Load Balancer. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 15, „Schnellkonfiguration“, auf Seite 139
- Kapitel 16, „Planung für Cisco CSS Controller“, auf Seite 143
- Kapitel 17, „Cisco CSS Controller konfigurieren“, auf Seite 149

Kapitel 15. Schnellkonfiguration

Dieses Beispiel für schnellen Start demonstriert das Erstellen einer Konfiguration mit der Komponente Cisco CSS Controller. Der Cisco CSS Controller stellt Serverwertigkeiten bereit, die dem Cisco CSS Switch helfen, für seine Lastausgleichsentscheidungen eine optimale Serverauswahl zu treffen.

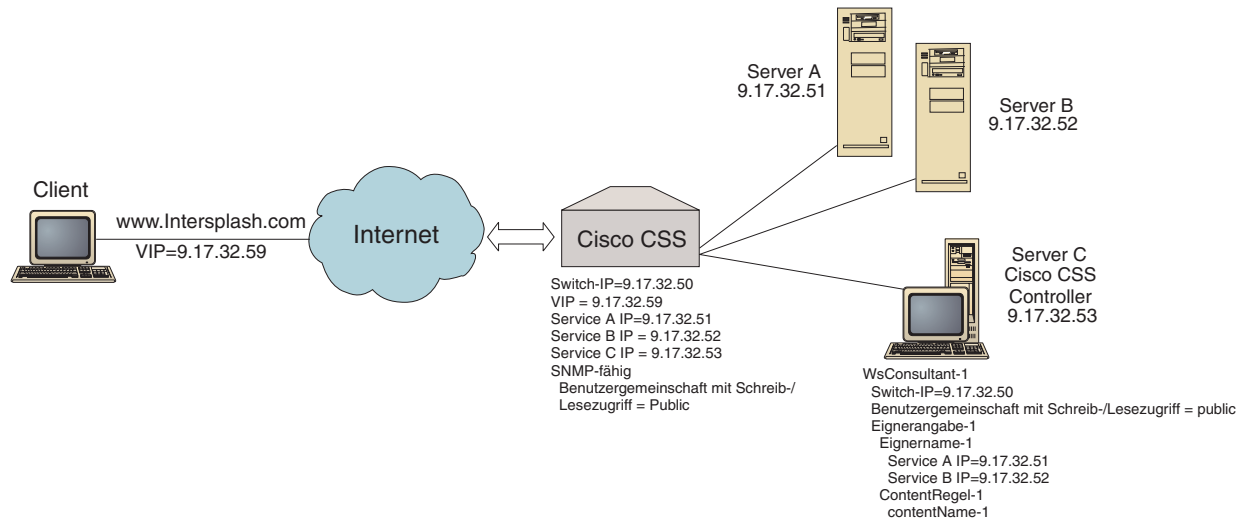


Abbildung 25. Einfache Konfiguration mit Cisco CSS Controller

Voraussetzungen

Für dieses Beispiel für schnellen Start benötigen Sie Folgendes:

- einen Cisco CSS Switch
- eine Servermaschine mit der Komponente Cisco CSS Controller
- zwei Webservermaschinen
- die folgenden fünf IP-Adressen:
 - eine IP-Adresse, die Sie Clients für den Zugriff auf Ihre Website www.Intersplashx.com (9.17.32.59) zur Verfügung stellen
 - eine IP-Adresse für eine Schnittstelle (Gateway) zum Cisco CSS Switch (9.17.32.50)
 - eine IP-Adresse für Server A (9.17.32.51)
 - eine IP-Adresse für Server B (9.17.32.52)
 - eine IP-Adresse für Server C mit Cisco CSS Controller (9.17.32.53)

Vorbereitungen

Vergewissern Sie sich vor dem Konfigurieren dieses Beispiels, dass die folgenden Schritte abgeschlossen sind:

- Stellen Sie sicher, dass der Cisco CSS Switch ordnungsgemäß konfiguriert ist. Informationen zur Konfiguration können Sie dem *Cisco Content Services Switch Getting Started Guide* entnehmen.

- Stellen Sie sicher, dass die Maschine mit Cisco CSS Controller den Cisco CSS Switch (9.17.32.50), den Server A (9.17.32.51) und den Server B (9.17.32.52) mit ping-Aufrufen erreichen kann.
- Vergewissern Sie sich, dass die Client-Maschine die VIP (9.17.32.59) mit ping erreichen kann.

Cisco CSS Controller konfigurieren

Für Cisco CSS Controller können Sie eine Konfiguration unter Verwendung der Befehlszeile oder der grafischen Benutzerschnittstelle (GUI) erstellen. Dieses Beispiel für schnellen Start zeigt die Ausführung der Konfigurationsschritte in der Befehlszeile.

Anmerkung: Die Parameterwerte müssen mit Ausnahme der Parameterwerte für Hostnamen und Dateinamen in englischen Zeichen eingegeben werden.

Konfiguration von der Befehlszeile aus

Führen Sie an einer Eingabeaufforderung die folgenden Schritte aus:

1. Starten Sie den ccoserver für Load Balancer. Setzen Sie an einer Eingabeaufforderung als Benutzer root oder Administrator den Befehl **ccoserver** ab.
2. Fügen Sie zur Konfiguration für Cisco CSS Controller einen Switch-Consultant hinzu. Geben Sie dazu die IP-Schnittstellenadresse des Cisco CSS Switch und den Namen der Benutzergemeinschaft mit Schreib-/Lesezugriff an. Diese Werte müssen mit den entsprechenden Attributen des Cisco CSS Switch übereinstimmen.

```
cococontrol consultant add SwConsultant-1 address 9.17.32.50 community public
```

Damit wird die Konnektivität zum Cisco CSS Switch überprüft und festgestellt, ob der Name der SNMP-Benutzergemeinschaft mit Schreib-/Lesezugriff funktioniert.

3. Fügen Sie zum Switch-Consultant Eignerangaben (Eignerangabe-1) hinzu. Geben Sie dazu den Eigernamen (Eigername-1) und die content-Regel (ContentRegel-1) an:

```
cococontrol ownercontent add SwConsultant-1:Eignerangabe-1 ownername Eigername-1 contentrule ContentRegel-1
```

Diese Werte müssen mit den entsprechenden Attributen des Cisco CSS Switch übereinstimmen.

Der Cisco CSS Controller kann jetzt über SNMP mit dem Switch kommunizieren und die notwendigen Konfigurationsdaten für den Switch abrufen. Nach diesem Schritt sollte der Cisco CSS Controller anzeigen, welche Services für den Cisco CSS Switch und die entsprechenden Eignerangaben konfiguriert sind.

4. Konfigurieren Sie die zu erfassenden Messwerte (aktive Verbindungen, Verbindungsrate, HTTP) und für jeden Messwert zu diesen Eignerangaben die proportionale Bedeutung:

```
cococontrol ownercontent metrics SwConsultant-1:Eignerangabe-1 activeconn 45 connrate 45 http 10
```

Dieser Befehl legt fest, welche Messwerte mit welcher Proportion von den Services zur Berechnung der Wertigkeit erfasst werden sollen. Die proportionalen Angaben aller Messwerte müssen in der Summe 100 ergeben.

5. Starten Sie den Switch-Consultant für den Cisco CSS Controller:

ccocontrol consultant start SwConsultant-1

Mit diesem Befehl werden alle Programme für Messwerterfassung gestartet, so dass mit dem Berechnen der Servicewertigkeit begonnen werden kann. Der Cisco CSS Controller teilt dem Cisco CSS Switch über SNMP die berechneten Servicewertigkeiten mit.

Die Basiskonfiguration für den Cisco CSS Controller ist damit vollständig.

Konfiguration testen

Testen Sie wie folgt, ob die Konfiguration korrekt ist:

1. Rufen Sie vom Client-Webbrowser aus die Adresse **http://www.Intersplashx.com** auf. Wenn eine Seite angezeigt wird, funktioniert die Konfiguration.
2. Laden Sie die Seite erneut im Webbrowser.
3. Überprüfen Sie die Ergebnisse des folgenden Befehls: **ccocontrol service report SwConsultant-1:Eignerangabe-1:Service-1**. Die Einträge der Spalte "Summe Verbindungen" für beide Webserver sollten addiert "2" ergeben.

Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus

Informationen zur Verwendung der GUI für den Cisco CSS Controller finden Sie im Abschnitt „GUI“ auf Seite 151 und in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Kapitel 16. Planung für Cisco CSS Controller

In diesem Kapitel wird beschrieben, was die für die Planung des Netzes zuständige Person vor der Installation und Konfiguration der Komponente Cisco CSS Controller berücksichtigen muss.

- Informationen zum Konfigurieren der Lastausgleichsparameter für die Komponente Cisco CSS Controller finden Sie in Kapitel 17, „Cisco CSS Controller konfigurieren“, auf Seite 149.
- Informationen zum Konfigurieren von Load Balancer für erweiterte Funktionen finden Sie in Kapitel 23, „Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller“, auf Seite 245.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Dieses Kapitel umfasst die folgenden Abschnitte:

- „Systemvoraussetzungen“
- „Überlegungen bei der Planung“
 - „Position des Consultant im Netz“ auf Seite 144
 - „Hohe Verfügbarkeit“ auf Seite 146
 - „Wertigkeiten berechnen“ auf Seite 147
 - „Fehlerbestimmung“ auf Seite 147

Systemvoraussetzungen

Informationen zu den Hardware- und Softwarevoraussetzungen finden Sie auf der folgenden Webseite: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Außerdem benötigen Sie Folgendes:

- Ein System für den Cisco CSS Controller.
- Einen installierten und konfigurierten Cisco CSS 11000 Series Content Services Switch.

Überlegungen bei der Planung

Der Cisco CSS Controller verwaltet eine Gruppe von Switch-Consultants. Jeder Consultant bestimmt Wertigkeiten für Services, deren Arbeitslast von nur einem Switch verteilt wird. Der Switch, für den der Consultant die Wertigkeiten bereitstellt, ist für die Inhaltsverteilung konfiguriert. Der Consultant sendet die berechneten Wertigkeiten mit dem Protokoll SNMP an den Switch. Wenn der Lastausgleichsalgorithmus mit einer RoundRobin-Methode gewichtet wird, verwendet der Switch die Wertigkeiten, um für die content-Regel des Lastausgleichs einen Service auszuwählen. Für die Bestimmung der Wertigkeiten verwendet der Consultant eine oder mehrere der folgenden Informationen:

- Verfügbarkeit und Antwortzeiten, die mit Hilfe von **Advisor-Funktionen** ermittelt werden, die mit den für den Service ausgeführten Anwendungen kommunizieren.

- Angaben zur Systembelastung, die durch Abrufen eines Messwerts von für den Service ausgeführten **Metric-Server-Agenten** ermittelt werden.
- Verbindungsdaten zum Service, die vom Switch abgerufen werden.
- Erreichbarkeitsdaten, die durch Senden von ping-Aufrufen an den Service ermittelt werden.

Eine Beschreibung der Inhaltsverteilung und ausführliche Informationen zum Konfigurieren des Switch können Sie dem *Cisco Content Services Switch Getting Started Guide* entnehmen.

Sie benötigen Folgendes, damit ein Consultant die zur Bestimmung der Servicewertigkeiten erforderlichen Informationen abrufen kann:

- IP-Konnektivität zwischen dem Consultant und den Services, deren Wertigkeit berechnet werden soll.
- IP-Konnektivität zwischen dem Consultant und dem Switch, der den Lastausgleich für die Server durchführt, deren Wertigkeit berechnet werden soll.
- Aktiviertes SNMP auf dem Switch. Aktivierter Schreib-/Lesezugriff.

Position des Consultant im Netz

Wie in Abb. 26 auf Seite 145 gezeigt, kann sich die Verbindung zwischen Consultant und Netz hinter den Switches befinden, für die der Consultant Wertigkeiten bereitstellen soll. Für die Konnektivität zwischen dem Controller, dem Switch und den Services müssen einige Parameter für den Switch und einige für den Controller konfiguriert werden.

Abb. 26 auf Seite 145:

- Die Verbindung zwischen Consultant und Netz befindet sich hinter den Switches, für die der Consultant Wertigkeiten bereitstellt.
- Das Netz besteht aus zwei virtuellen LANs.
- An den Schnittstellen zwischen Services und Netz sowie an der Schnittstelle zwischen Consultant und Netz muss die IP-Weiterleitung aktiviert sein, damit der Consultant mit den Services in beiden virtuellen LANs kommunizieren kann.
- Die IP-Adresse des Switch muss auf dem Consultant-System und den Service-Systemen als Standard-Gateway konfiguriert sein.

Ausführliche Informationen zum Konfigurieren von virtuellen LANs und des IP-Routing auf dem Switch können Sie dem *Cisco Content Services Switch Getting Started Guide* entnehmen.

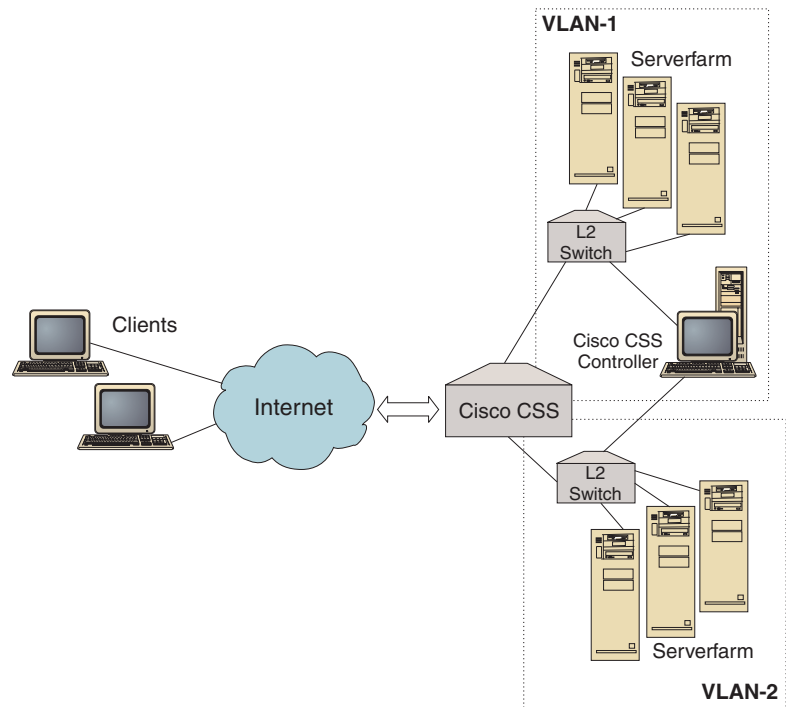


Abbildung 26. Beispiel für einen Consultant, der hinter den Switches mit dem Netz verbunden ist

Für die Verwaltung des Cisco CSS Controller können Sie eine der folgenden Schnittstellen nutzen:

- Browser
- GUI (fern oder lokal)
- Befehlszeile (fern oder lokal).

Abb. 27 auf Seite 146 für Fernverwaltung:

- Die Verbindung zwischen dem Consultant und dem Netz befindet sich hinter dem Switch, für den der Consultant Wertigkeiten bereitstellt.
- Die Benutzerschnittstelle befindet sich auf einem fernen System vor dem Switch.
- Der Switch muss so konfiguriert werden, dass das ferne System über den Switch mit dem Controllersystem kommunizieren kann.

Ausführliche Informationen hierzu können Sie dem *Cisco Content Services Switch Getting Started Guide* entnehmen.

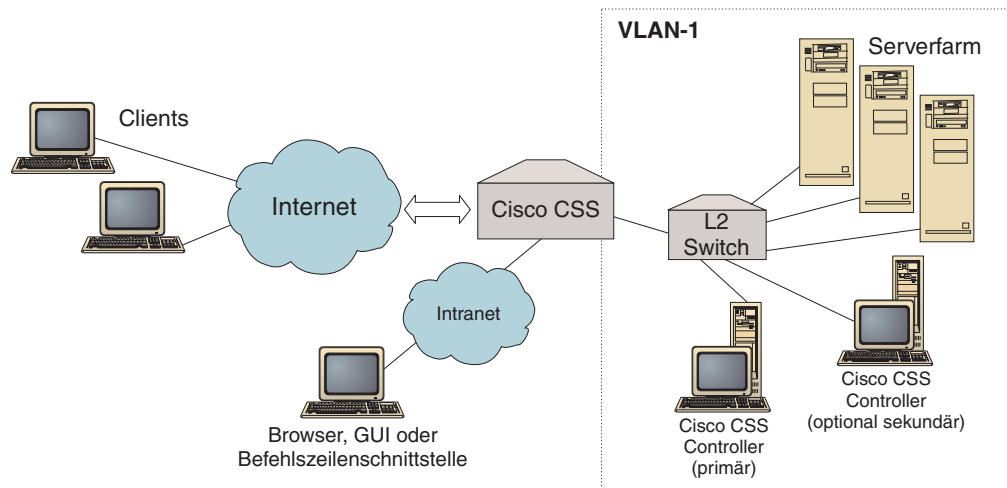


Abbildung 27. Beispiel für einen Consultant (optional mit einem Partner für hohe Verfügbarkeit) hinter dem Switch mit einer Benutzerschnittstelle vor dem Switch

Hohe Verfügbarkeit

Die hohe Controllerverfügbarkeit erweitert die Fähigkeiten von Load Balancer im Bereich der Fehlertoleranz. Die Entwicklung der hohen Controllerverfügbarkeit wurde im Hinblick auf die Verfügbarkeit für die Paketweiterleitung entwickelt und bedeutet die Verfügbarkeit zweier gleichzeitig aktiver Controller, von denen einer der primäre und der andere der sekundäre Controller ist.

Beide Controller werden mit identischen Switch-Daten konfiguriert und es ist immer nur ein Controller zur Zeit aktiv. Entsprechend der Logik der hohen Verfügbarkeit bedeutet dies, dass nur der aktive Controller neue Wertigkeiten für den Switch berechnet und aktualisiert.

Für die hohe Controllerverfügbarkeit kommuniziert der Controller mit seinem Partner über eine von Ihnen konfigurierte Adresse und einen von Ihnen konfigurierten Port unter Verwendung des UDP (User Datagram Protocol). Mit diesen Paketen tauschen die Controller die für hohe Verfügbarkeit wichtigen Daten (Erreichbarkeitsdaten) aus und stellen (mit Überwachungssignalen) fest, ob der Partner verfügbar ist. Wenn der Bereitschaftscontroller erkennt, dass der aktive Controller ausgefallen ist, übernimmt er die Aufgaben des aktiven Controllers. Damit wird der Bereitschaftscontroller zum aktiven Controller und beginnt, neue Wertigkeiten für den Switch zu berechnen und zu aktualisieren.

Neben der Verfügbarkeit des Partners können Sie für die hohe Verfügbarkeit Erreichbarkeitsziele konfigurieren. Bei der hohen Controllerverfügbarkeit werden Erreichbarkeitsdaten verwendet, um festzustellen, welcher der Controller aktiv und welcher der Bereitschaftscontroller ist. Der aktive Controller kann mehr Ziele mit ping erreichen und ist von seinem Partnercontroller aus erreichbar.

Weitere Informationen finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 245.

Wertigkeiten berechnen

Wenn der Consultant feststellt, dass ein Service nicht verfügbar ist, setzt er diesen Service auf dem Switch aus, damit der Switch den Service bei der Verteilung von Anforderungen nicht berücksichtigt. Sobald der Service wieder verfügbar ist, aktiviert der Consultant den Service auf dem Switch, so dass er bei Lastausgleichsanforderungen wieder Berücksichtigung findet.

Fehlerbestimmung

Der Cisco CSS Controller schreibt Einträge in die folgenden Protokolle:

- server.log
- consultant.log
- highavailability.log
- metriccollector.log
- binary.log

Diese Protokolle befinden sich in den folgenden Verzeichnissen:

- AIX-, HP-UX-, Linux- und Solaris-Systeme: ...ibm/edge/lb/servers/logs/cco/*Consultant-Name*
- Windows-Systeme: ...ibm\edge\lb\servers\logs\cco*Consultant-Name*

In jedem Protokoll können Sie die Protokollgröße und -stufe festlegen. Weitere Informationen hierzu finden Sie im Abschnitt „Protokolle von Load Balancer verwenden“ auf Seite 267.

Kapitel 17. Cisco CSS Controller konfigurieren

Lesen Sie vor Ausführung der in diesem Kapitel beschriebenen Schritte Kapitel 16, „Planung für Cisco CSS Controller“, auf Seite 143. Dieses Kapitel erläutert das Erstellen einer Basiskonfiguration für die Komponente Cisco CSS Controller von Load Balancer.

- Komplexere Konfigurationen finden Sie in Kapitel 23, „Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller“, auf Seite 245.
- Informationen zur Fernverwaltung mit Authentifizierung, zu Protokollen sowie zur Verwendung der Komponente Cisco CSS Controller finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Konfigurations-Tasks im Überblick

Vor Ausführung einer der in diesem Kapitel beschriebenen Konfigurationsmethoden müssen Sie die folgenden Schritte ausführen:

1. Vergewissern Sie sich, dass der Cisco CSS Switch und alle Servermaschinen richtig konfiguriert sind.
2. Konfigurieren Sie Cisco CSS Controller so, dass die Adresse des Cisco CSS Switch und der Name der SNMP-Benutzergemeinschaft mit den entsprechenden Attributen des Cisco CSS Switch übereinstimmen. Informationen zum Konfigurieren des Consultant finden Sie im Abschnitt „ccocontrol consultant — Consultant konfigurieren und steuern“ auf Seite 434.

Tabelle 10. Konfigurations-Tasks für Cisco CSS Controller

Task	Beschreibung	Referenzinformationen
Konfigurieren der Maschine mit Cisco CSS Controller	Ermitteln Sie die Voraussetzungen.	„Maschine mit Controller für Cisco CSS Switches konfigurieren“ auf Seite 152
Testen der Konfiguration	Überprüfen Sie, ob die Konfiguration funktioniert.	„Konfiguration testen“ auf Seite 154

Konfigurationsmethoden

Es gibt im Wesentlichen drei Methoden für das Erstellen einer Basiskonfiguration für die Load-Balancer-Komponente Cisco CSS Controller:

- Befehlszeile
- XML-Datei
- grafische Benutzerschnittstelle (GUI).

Befehlszeile

Dies ist die direkte Methode für die Konfiguration von Cisco CSS Controller. Bei den in diesem Handbuch beschriebenen Prozeduren wird von der Verwendung der Befehlszeile ausgegangen. Die Werte der Befehlsparameter müssen in englischen Zeichen eingegeben werden. Die einzige Ausnahme hiervon bilden Hostnamen (die z. B. im Befehl **consultant add** verwendet werden) und Dateinamen.

Starten Sie Cisco CSS Controller wie folgt von der Befehlszeile aus:

1. Setzen Sie an der Eingabeaufforderung den Befehl **ccoserver** ab. Geben Sie zum Stoppen des Servers **ccoserver stop** ein.

Anmerkungen:

- a. Klicken Sie auf Windows-Systemen nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf **IBM Cisco CSS Controller** und wählen Sie **Starten** aus. Zum Stoppen des Services müssen Sie dieselben Schritte ausführen und **Beenden** auswählen.
 - b. Auf Windows-Systemen können Sie **ccoserver** automatisch beim Booten starten. Gehen Sie dazu wie folgt vor:
 - 1) Klicken Sie nacheinander auf **Start > Einstellungen > Systemsteuerung > Verwaltung > Dienste**.
 - 2) Klicken Sie mit der rechten Maustaste auf **IBM Cisco CSS Controller** und wählen Sie **Eigenschaften** aus.
 - 3) Klicken Sie rechts neben **Starttyp** auf den Abwärtspfeil und wählen Sie **Automatisch** aus.
 - 4) Klicken Sie auf **OK**.
2. Setzen Sie anschließend die gewünschten Steuerbefehle für Cisco CSS Controller ab, um die Konfiguration einzurichten. Bei den in diesem Handbuch beschriebenen Prozeduren wird von der Verwendung der Befehlszeile ausgegangen. Der Befehl lautet **cococontrol**. Weitere Informationen zu Befehlen finden Sie in Kapitel 29, „Befehlsreferenz für Cisco CSS Controller“, auf Seite 433.

Sie können eine gekürzte Version der Parameter für den Befehl **cococontrol** eingeben. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **cococontrol he f** an Stelle von **cococontrol help file** angeben.

Wenn Sie die Befehlszeilenschnittstelle starten möchten, setzen Sie den Befehl **ccocontrol** ab, um die Eingabeaufforderung **cococontrol** aufzurufen.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie den Befehl **exit** oder **quit** absetzen.

Anmerkung: Auf Windows-Plattformen wird **dsserver** für die Komponente Dispatcher automatisch gestartet. Falls Sie nur mit Cisco CSS Controller und nicht mit der Komponente Dispatcher arbeiten, können Sie den automatischen Start von **dsserver** wie folgt unterbinden:

1. Klicken Sie im Fenster "Dienste" von Windows mit der rechten Maustaste auf **IBM Dispatcher**.
2. Wählen Sie den Eintrag **Eigenschaften** aus.
3. Wählen Sie im Feld **Starttyp** die Option **Manuell** aus.
4. Klicken Sie auf **OK** und schließen Sie das Fenster "Dienste".

XML

Die soeben definierte Konfiguration kann in einer XML-Datei gespeichert werden. So kann die Konfiguration später schnell erneut geladen werden.

Verwenden Sie zum Ausführen des Inhaltes einer XML-Datei (z. B. von **mein-script.xml**) einen der folgenden Befehle:

- Setzen Sie zum Speichern der aktuellen Konfiguration in einer XML-Datei den folgenden Befehl ab:
ccocontrol file save XML-Dateiname
- Setzen Sie zum Laden einer gespeicherten Konfiguration den folgenden Befehl ab:
ccocontrol file load XML-Dateiname
Verwenden Sie den Befehl **load** nur, wenn Sie zuvor den Befehl **file save** abgesetzt haben.

Die XML-Dateien werden im Verzeichnis **...ibm/edge/lb/servers/configurations/ccol** gespeichert.

GUI

Abb. 41 auf Seite 470 zeigt ein Beispiel für die grafische Benutzerschnittstelle (GUI) mit allgemeinen Anweisungen.

Gehen Sie zum Starten der GUI wie folgt vor:

1. Wenn **ccoserver** noch nicht aktiv ist, starten Sie den Dienst jetzt, indem Sie als **root** den folgenden Befehl absetzen:
ccoserver.
2. Führen Sie anschließend einen der folgenden Schritte aus:
 - Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Geben Sie **lbadmin** ein.
 - Auf Windows-Systemen: Klicken Sie nacheinander auf **Start > Programme > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer.**

Gehen Sie zum Konfigurieren von Cisco CSS Controller von der GUI aus wie folgt vor:

1. Klicken Sie in der Baumstruktur mit der rechten Maustaste auf "Cisco CSS Controller".
2. Stellen Sie eine Verbindung zu einem Host her.
3. Erstellen Sie einen oder mehrere Switch-Consultant(s) mit den gewünschten Eigenerangaben und den zugehörigen Messwerten.
4. Starten Sie den Consultant.

Von der GUI aus können Sie alle mit dem Befehl **ccocontrol** ausführbaren Schritte ausführen. Beispiel:

- Wenn Sie in der Befehlszeile einen Consultant definieren möchten, geben Sie **ccocontrol consultant add Consultant-ID address IP-Adresse community Name** ein.
- Wenn Sie einen Consultant von der GUI aus definieren möchten, klicken Sie mit der rechten Maustaste auf den Hostknoten. Klicken Sie dann auf **Switch-Consultant hinzufügen**. Geben Sie im Dialogfenster die Switch-Adresse und den Namen der Benutzergemeinschaft ein und klicken Sie auf "OK".
- Bei Auswahl des Eintrags **Konfiguration laden** im Popup-Menü "Host" können Sie eine bereits vorhandene Konfigurationsdatei für Cisco CSS Controller laden und an die aktive Konfiguration anhängen.
- Wählen Sie regelmäßig **Konfigurationsdatei sichern als** aus, um Ihre Konfiguration für Cisco CSS Controller in einer Datei zu speichern.
- Wenn Sie in der Menüleiste **Datei** auswählen, können Sie Ihre aktuellen Hostverbindungen in einer Datei speichern oder Verbindungen aus vorhandenen Dateien aller Load-Balancer-Komponenten wiederherstellen.

Gehen Sie wie folgt vor, um von der GUI aus einen Befehl auszuführen:

1. Klicken Sie mit der rechten Maustaste auf den Knoten **Host** und wählen Sie **Befehl senden...** aus.
2. Geben Sie im Befehlseingabefeld den gewünschten Befehl ein, z. B. **consultant report**.
3. Klicken Sie auf "Senden".

Im Fenster "Ergebnis" sehen Sie die Ergebnisse und die Historie der in der aktuellen Sitzung ausgeführten Befehle.

Falls Sie **Hilfe** benötigen, klicken Sie oben rechts im Load-Balancer-Fenster auf das Fragezeichen.

- **Hilfe: Feldebene** — beschreibt jedes Feld und gibt die Standardwerte an.
- **Hilfe: Wie funktioniert das** — listet Tasks auf, die von dieser Anzeige aus ausgeführt werden können.
- **InfoCenter** — ermöglicht den zentralen Zugriff auf Produktinformationen.

Weitere Informationen zur Verwendung der GUI finden Sie in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Maschine mit Controller für Cisco CSS Switches konfigurieren

Vor dem Konfigurieren der Maschine mit Cisco CSS Controller müssen Sie (auf AIX-, HP-UX-, Linux- oder Solaris-Systemen) als Benutzer root oder (auf Windows-Systemen) als Administrator registriert sein.

Der Consultant muss eine Verbindung zum Cisco CSS Switch als Cisco-CSS-Switch-Administrator herstellen können.

Wenn Sie den Consultant konfigurieren, müssen die Adresse und der Name der SNMP-Benutzergemeinschaft mit den entsprechenden Attributen des Cisco CSS Switch übereinstimmen.

Hilfe zu den in dieser Prozedur verwendeten Befehlen finden Sie in Kapitel 29, „Befehlsreferenz für Cisco CSS Controller“, auf Seite 433.

Schritt 1. Serverfunktion starten

Wenn ccoserver noch nicht aktiv ist, starten Sie den Dienst jetzt, indem Sie als root **ccoserver** eingeben.

Anmerkung: Klicken Sie auf Windows-Systemen nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf "IBM Cisco Controller" und wählen Sie "Starten" aus.

Schritt 2. Befehlszeilenschnittstelle aufrufen

Geben Sie **cococontrol** ein, um die Befehlszeilenschnittstelle aufzurufen.

Schritt 3. Consultant konfigurieren

Sie müssen die Switch-Adresse und einen Namen für die SNMP-Benutzergemeinschaft konfigurieren. Diese Werte müssen mit den entsprechenden Attributen des Cisco CSS Switch übereinstimmen.

Geben Sie Folgendes ein, um einen Consultant hinzuzufügen:

```
consultant add Switch-Consultant-ID address Switch-IP-Adresse  
community Name_der_Benutzergemeinschaft
```

Schritt 3. Eignerangaben konfigurieren

Eignerangaben sind die Darstellung einer content-Regel für einen Eigner, der für den Cisco CSS Switch definiert ist. Der Eigernamen und der Name der content-Regel müssen mit der entsprechenden Definition auf dem Switch übereinstimmen.

Geben Sie Folgendes ein, um Eignerangaben zu definieren:

```
ownercontent add Switch-Consultant-ID:ID_für_Eignerangaben ownername Eigernamen  
contentrule Name_der_content-Regel
```

Schritt 4. Konfiguration der Services prüfen

Wenn die Eignerangaben definiert sind, schließt der Consultant die Konfiguration ab, indem er die für den Switch konfigurierten Services abrufen. Vergleichen Sie die Konfiguration auf dem Switch mit der Konfiguration für den Consultant, um sicherzustellen, dass die Services übereinstimmen.

Schritt 5. Messwerte konfigurieren

Anhand von Messwerten werden die Wertigkeiten von Services und ihre proportionale Gewichtung (im Vergleich zu anderen Services) bestimmt. Sie können eine beliebige Kombination von Messwerten für Verbindungsdaten, für Advisor-Funktionen der Anwendung und für Metric Server verwenden. Die proportionalen Gewichtungen müssen in der Summe stets 100 ergeben.

Wenn Sie die Eignerangaben konfigurieren, werden die Standardmesswerte **activeconn** und **connrate** definiert. Falls Sie zusätzliche oder gänzlich andere Messwerte verwenden möchten, geben Sie Folgendes ein:

```
ownercontent metrics  
Switch-Consultant-ID:ID_für_Eignerangaben Messwert1 Proportion1  
Messwert2 Proportion2...MesswertN ProportionN
```

Schritt 6. Consultant starten

Geben Sie zum Starten des Consultant Folgendes ein:

```
consultant start Switch-Consultant-ID
```

Mit diesem Befehl wird die Erfassung von Messwerten gestartet, und die Berechnung von Wertigkeiten beginnt.

Schritt 7. Metric Server starten (optional)

Wenn Sie in Schritt 5 Systemmesswerte definiert haben, muss auf den Service-Maschinen Metric Server gestartet werden. Informationen zur Verwendung von Metric Server finden Sie im Abschnitt „Metric Server“ auf Seite 196.

Schritt 8. Hohe Verfügbarkeit konfigurieren (optional)

Geben Sie Folgendes ein, um die hohe Verfügbarkeit zu konfigurieren:

```
highavailability add  
address IP-Adresse partneraddress IP-Adresse port 80  
role primary
```

In einer Umgebung mit hoher Verfügbarkeit können Sie mehrere Switches konfigurieren. Der Cisco CSS Controller muss so konfiguriert werden, dass er für alle Switches und die entsprechenden Ausweicheinheiten Wertigkeiten bereitstellt, um auch bei der Übernahme der Aufgaben eines Switch durch einen anderen die ständige Verfügbarkeit von Wertigkeitsinformationen zu gewährleisten.

Ausführliche Informationen zur Verwendung und zum Konfigurieren der hohen Controllerverfügbarkeit finden Sie in Kapitel 23, „Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller“, auf Seite 245.

Konfiguration testen

Testen Sie wie folgt, ob die Konfiguration korrekt ist:

1. Setzen Sie die Protokollstufe des Consultant auf 4.
2. Trennen Sie einen Server für eine Minute vom Cisco CSS Switch oder fahren Sie den Anwendungsserver für eine Minute herunter.
3. Stellen Sie die Verbindung des Servers zum Switch wieder her oder führen Sie einen Neustart für den Anwendungsserver aus.
4. Setzen Sie die Protokollstufe des Consultant auf den gewünschten Wert (1) zurück.
5. Rufen Sie die Datei consultant.log in den folgenden Verzeichnissen auf und suchen Sie nach dem Eintrag **setServerWeights setting service**:
 - AIX-, HP-UX-, Linux- und Solaris-Systeme: ...ibm/edge/lb/servers/logs/cco/*Consultant-Name*
 - Windows-Systeme: ...ibm\edge\lb\servers\logs\cco*Consultant-Name*

Teil 6. Nortel Alteon Controller

Dieser Teil enthält Informationen zu einer schnellen Erstkonfiguration sowie zur Planung und beschreibt die Konfigurationsmethoden für die Komponente Nortel Alteon Controller von Load Balancer. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 18, „Schnellkonfiguration“, auf Seite 157
- Kapitel 19, „Planung für Nortel Alteon Controller“, auf Seite 161
- Kapitel 20, „Nortel Alteon Controller konfigurieren“, auf Seite 171

Kapitel 18. Schnellkonfiguration

Dieses Beispiel für schnellen Start demonstriert das Erstellen einer Konfiguration mit der Komponente Nortel Alteon Controller. Der Nortel Alteon Controller stellt Serverwertigkeiten für den Nortel Alteon Web Switch bereit. Anhand dieser Wertigkeiten werden Server für Services ausgewählt, für die der Switch einen Lastausgleich durchführt.

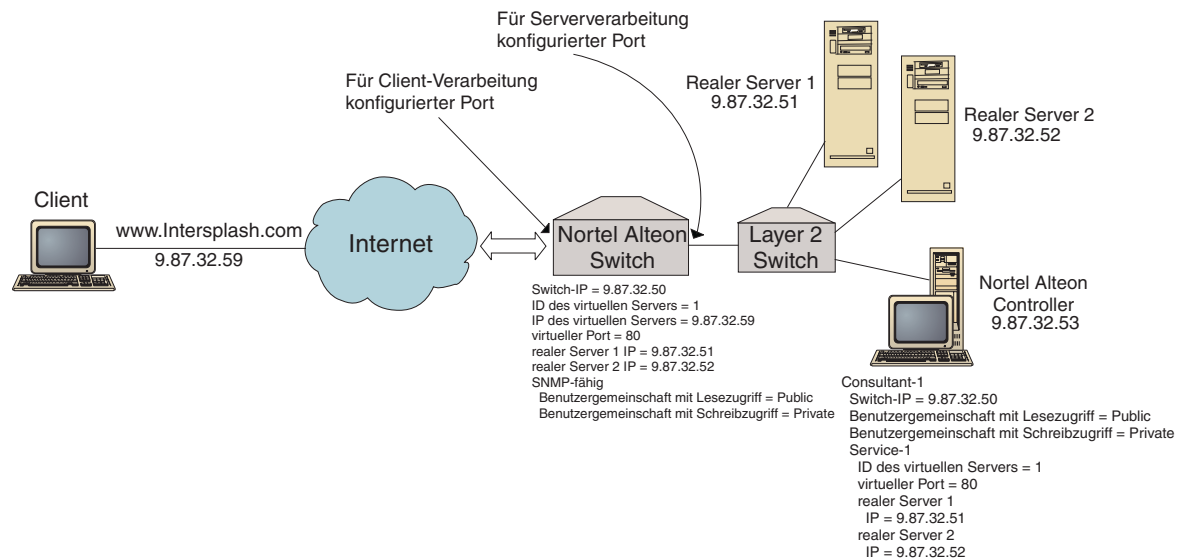


Abbildung 28. Einfache Konfiguration mit Nortel Alteon Controller

Voraussetzungen

Für dieses Beispiel für schnellen Start benötigen Sie Folgendes:

- Ein Nortel Alteon Web Switch mit Web OS Version 9.0 oder 10.0
- Eine Servermaschine mit der Komponente Nortel Alteon Controller
- Zwei Webservermaschinen
- Einen Layer 2 Switch, der mit einem Anschluss des Nortel Alteon Web Switch verbunden ist

Anmerkung: Wenn kein Layer 2 Switch verwendet wird, können die Maschine mit Nortel Alteon Controller und die Webservermaschinen direkt mit Anschlüssen des Nortel Alteon Web Switch verbunden werden.

- Die folgenden fünf IP-Adressen:
 - Eine IP-Adresse, die Sie Clients für den Zugriff auf Ihre Website www.Intersplashx.com (9.87.32.59) zur Verfügung stellen
 - eine IP-Adresse für eine Schnittstelle, die zum Nortel Alteon Web Switch (9.87.32.50) konfiguriert ist
 - eine IP-Adresse für den realen Server 1 (9.87.32.51)
 - eine IP-Adresse für den realen Server 2 (9.87.32.52)
 - eine IP-Adresse für den Nortel Alteon Controller (9.87.32.53)

Vorbereitungen

Vergewissern Sie sich vor dem Konfigurieren dieses Beispiels, dass die folgenden Schritte abgeschlossen sind:

- Stellen Sie sicher, dass der Nortel Alteon Web Switch ordnungsgemäß konfiguriert ist. (Ausführlichere Konfigurationsdaten können Sie dem "Nortel Alteon Web OS Application Guide" entnehmen):
 - Aktivieren Sie für den Switch den Serverlastausgleich der Ebene 4.
 - Konfigurieren Sie auf dem Nortel Alteon Web Switch eine IP-Schnittstelle (9.87.32.50).
 - Aktivieren Sie auf dem Nortel Alteon Web Switch SNMP.
 - Aktivieren Sie die Client-Verarbeitung des Serverlastausgleichs an dem Port des Nortel Alteon Web Switch, der die Client-Anforderungen empfängt.
 - Aktivieren Sie die Serververarbeitung des Lastausgleichs an dem Port des Nortel Alteon Web Switch, mit dem die Server verbunden sind.
 - Konfigurieren Sie für den realen Server 1, den realen Server 2 und den Nortel Alteon Controller als Standard-Gateway die Switch-IP-Schnittstelle (9.87.32.50).
 - Konfigurieren Sie den Nortel Alteon Web Switch mit dem realen Server 1 und dem realen Server 2.
 - Konfigurieren Sie den Nortel Alteon Web Switch mit einer Servergruppe, die die realen Server 1 und 2 umfasst. Weisen Sie dieser Gruppe die ID 1 zu.
 - Konfigurieren Sie den Nortel Alteon Web Switch mit einem virtuellen Server. Die IP-Adresse des virtuellen Servers lautet 9.87.32.59. Weisen Sie dem virtuellen Server die ID 1 zu.
 - Konfigurieren Sie den Nortel Alteon Web Switch mit einem Service, der den virtuellen Port 80 verwendet und von Gruppe 1 bedient wird.
- Vergewissern Sie sich, dass die Client-Maschine die IP-Adresse des virtuellen Servers 9.87.32.59 mit ping erreichen kann.
- Stellen Sie sicher, dass die Maschine mit Nortel Alteon Controller die IP-Schnittstelle des Nortel Alteon Web Switch (9.17.32.50), den realen Server 1 (9.87.32.51) und den realen Server 2 (9.87.32.52) mit ping-Aufrufen erreichen kann.

Nortel Alteon Controller konfigurieren

Für Nortel Alteon Controller können Sie eine Konfiguration unter Verwendung der Befehlszeile oder der grafischen Benutzerschnittstelle (GUI) erstellen. Dieses Beispiel für schnellen Start zeigt die Ausführung der Konfigurationsschritte in der Befehlszeile.

Anmerkung: Die Parameterwerte müssen mit Ausnahme der Parameterwerte für Hostnamen und Dateinamen in englischen Zeichen eingegeben werden.

Konfiguration von der Befehlszeile aus

Führen Sie an einer Eingabeaufforderung die folgenden Schritte aus:

1. Starten Sie auf der Maschine mit Nortel Alteon Controller den `nalserver`. Setzen Sie an einer Eingabeaufforderung als Benutzer `root` oder Administrator den Befehl `nalserver` ab.

2. Fügen Sie zur Konfiguration für Nortel Alteon Controller einen Consultant hinzu. Geben Sie dazu die IP-Schnittstellenadresse des Nortel Alteon Web Switch an. (Geben Sie die Benutzergemeinschaft mit Lesezugriff und die Benutzergemeinschaft mit Schreibzugriff nur an, wenn sie vom Standard (public, private) abweichen):

nalcontrol consultant add Consultant-1 address 9.87.32.50

Damit wird die Konnektivität zum Nortel Alteon Web Switch überprüft und festgestellt, ob die Namen der SNMP-Benutzergemeinschaften funktionieren.

3. Fügen Sie zum Consultant (Consultant-1) einen Service (Service-1) hinzu. Geben Sie dazu für den Service die ID des virtuellen Servers (1) und die Nummer des virtuellen Ports (80) an:

nalcontrol service add Consultant-1:Service-1 vsid 1 vport 80

Der Nortel Alteon Controller kommuniziert über SNMP mit dem Switch und ruft die notwendigen Konfigurationsdaten für den Switch ab. Nach diesem Schritt sollte der Nortel Alteon Controller anzeigen, welche Server auf dem Nortel Alteon Web Switch für diesen Service konfiguriert sind.

4. Konfigurieren Sie die Messwerte, die für die dem Service zugeordnete Servergruppe erfasst werden sollen:

nalcontrol service metrics Consultant-1:Service-1 http 40 activeconn 30 conn-rate 30

Dieser Befehl legt fest, welche Messwerte von den Servern erfasst und welche relative Bedeutung sie bei der Berechnung der Wertigkeit haben sollen.

5. Starten Sie die Consultant-Funktion für den Nortel Alteon Controller:

nalcontrol consultant start Consultant-1

Mit diesem Befehl werden alle Programme für Messwerterfassung gestartet, so dass mit dem Berechnen der Serverwertigkeit begonnen werden kann. Der Nortel Alteon Controller teilt dem Nortel Alteon Web Switch über SNMP die berechneten Serverwertigkeiten mit.

Die Basiskonfiguration für den Nortel Alteon Controller ist damit vollständig.

Konfiguration testen

Testen Sie wie folgt, ob die Konfiguration korrekt ist:

1. Rufen Sie vom Client-Webbrowser aus die Adresse **http://www.Intersplashx.com** auf. Wenn eine Seite angezeigt wird, funktioniert die Konfiguration.
2. Laden Sie die Seite erneut im Webbrowser.
3. Überprüfen Sie die Ergebnisse des folgenden Befehls: **nalcontrol service report Consultant-1:Service-1**. Die Einträge der Spalte "Summe Verbindungen" für beide Webserver sollten addiert "2" ergeben.

Konfiguration von der grafischen Benutzerschnittstelle (GUI) aus

Informationen zur Verwendung der GUI für den Nortel Alteon Controller finden Sie im Abschnitt „GUI“ auf Seite 173 und in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Kapitel 19. Planung für Nortel Alteon Controller

In diesem Kapitel wird beschrieben, was die für die Planung des Netzes zuständige Person vor der Installation und Konfiguration der Komponente Nortel Alteon Controller berücksichtigen muss.

- Informationen zum Konfigurieren der Lastausgleichsparameter für die Komponente Nortel Alteon Controller finden Sie in Kapitel 20, „Nortel Alteon Controller konfigurieren“, auf Seite 171.
- Informationen zum Konfigurieren von Advisor-Funktionen und Metric Servern finden Sie in Kapitel 23, „Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller“, auf Seite 245.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen von Load Balancer sowie zur Verwendung der Komponenten von Load Balancer finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Dieses Kapitel umfasst die folgenden Abschnitte:

- „Systemvoraussetzungen“
- „Überlegungen bei der Planung“
 - „Position des Consultant im Netz“ auf Seite 162
 - „Vom Controller festgelegte Serverattribute auf dem Switch“ auf Seite 165
 - „Ausweichserver konfigurieren“ auf Seite 165
 - „Gruppen konfigurieren“ auf Seite 166
 - „Hohe Verfügbarkeit“ auf Seite 167
 - „Optimierung“ auf Seite 169
 - „Fehlerbestimmung“ auf Seite 169

Systemvoraussetzungen

Informationen zu den Hardware- und Softwarevoraussetzungen finden Sie auf der folgenden Webseite: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Außerdem benötigen Sie Folgendes:

- Ein System für den Nortel Alteon Controller.
- Einen installierten und konfigurierten Nortel Alteon Web Switch. Die Web-Switch-Plattformen sind AD3, AD4, 180e, 184 und das Blade der Stufe 4/7 für den Passport 8600.

Überlegungen bei der Planung

Der Nortel Alteon Controller verwaltet eine Gruppe von Switch-Consultants. Jeder Consultant bestimmt Wertigkeiten für Server, deren Arbeitslast von nur einem Switch verteilt wird. Der Switch, für den der Consultant die Wertigkeiten bereitstellt, ist für den Serverlastausgleich konfiguriert. Der Consultant sendet die berechneten Wertigkeiten mit dem Protokoll SNMP an den Switch. Ausgehend von diesen Wertigkeiten wählt der Switch für den Service, dessen Last verteilt werden soll, einen Server aus. Für die Bestimmung der Wertigkeiten verwendet der Consultant eine oder mehrere der folgenden Informationen:

- Verfügbarkeit und Antwortzeiten, die mit Hilfe von **Advisor-Funktionen** ermittelt werden, die mit den auf den Servern ausgeführten Anwendungen kommunizieren.
- Angaben zur Systembelastung, die durch Abrufen eines Messwerts von auf den Servern ausgeführten **Metric-Server-Agenten** ermittelt werden.
- Verbindungsdaten zu den Servern, die vom Switch abgerufen werden.
- Erreichbarkeitsdaten, die durch Senden von ping-Aufrufen an die Server ermittelt werden.

Eine Beschreibung des Serverlastausgleichs und ausführliche Informationen zum Konfigurieren des Switch können Sie dem "Nortel Alteon Web OS Application Guide" entnehmen.

Sie benötigen Folgendes, damit ein Consultant die zur Bestimmung der Serverwertigkeiten erforderlichen Informationen abrufen kann:

- IP-Konnektivität zwischen dem Consultant und den Servern, deren Wertigkeit berechnet werden soll.
- IP-Konnektivität zwischen dem Consultant und dem Switch, der den Lastausgleich für die Server durchführt, deren Wertigkeit berechnet werden soll.
- Aktiviertes SNMP auf dem Switch. Aktivierter Schreib-/Lesezugriff.

Position des Consultant im Netz

Die Verbindung zwischen Consultant und Netz kann sich vor oder hinter den Switches befinden, für die der Consultant Wertigkeiten bereitstellen soll. Für die Konnektivität zwischen dem Controller, dem Switch und den Servern müssen einige Parameter für den Switch und einige für den Controller konfiguriert werden.

Abb. 29 auf Seite 163:

- Die Verbindung zwischen Consultant und Netz befindet sich hinter den Switches, für die der Consultant Wertigkeiten bereitstellt.
- Das Netz besteht aus zwei virtuellen LANs.
- An den Schnittstellen zwischen Servern und Netz sowie an der Schnittstelle zwischen Consultant und Netz muss die IP-Weiterleitung aktiviert sein, damit der Consultant mit den Servern in beiden virtuellen LANs kommunizieren kann.
- Die IP-Adresse des Switch muss auf dem Consultant-System und den Serversystemen als Standard-Gateway konfiguriert sein.

Ausführliche Informationen zum Konfigurieren von virtuellen LANs und des IP-Routing auf dem Switch können Sie dem "Nortel Alteon Web OS Application Guide" oder der Veröffentlichung "Command Reference" entnehmen.

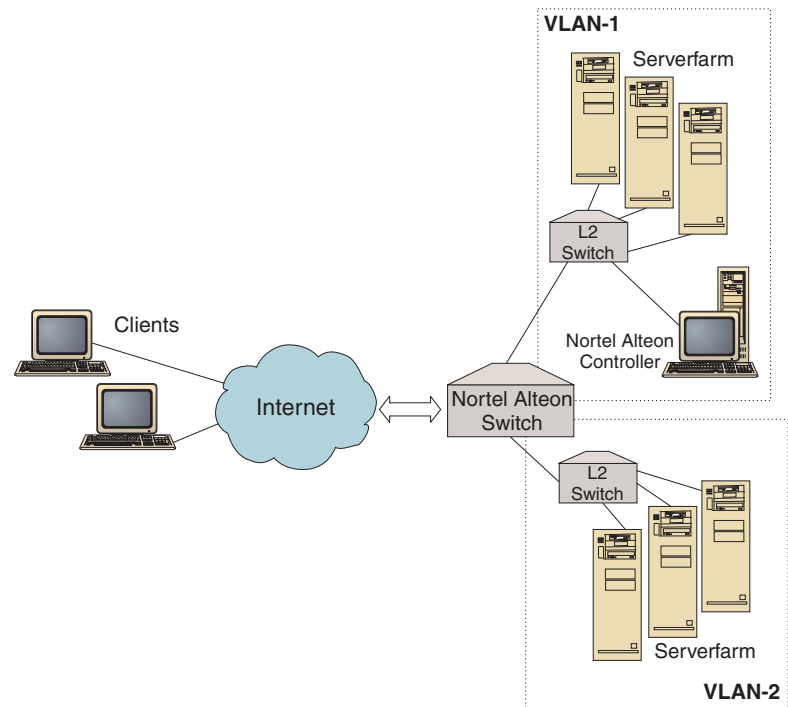


Abbildung 29. Beispiel für einen Consultant, der hinter dem Switch mit dem Netz verbunden ist

Abb. 30 auf Seite 164:

- Der Consultant ist über ein Intranet vor dem Switch mit dem Netz verbunden.
- Auf dem Switch muss der direkte Zugriffsmodus für den Serverlastausgleich aktiviert werden, damit der Consultant mit dem Switch und den Servern kommunizieren kann.
- Bei aktiviertem direktem Zugriffsmodus für den Lastausgleich kann jeder Client direkt Daten an jeden Server senden. Wenn Sie den direkten Serverzugriff auf den Consultant begrenzen möchten, können Sie auf dem Switch für den Lastausgleich *mnet* und *mmask* angeben. Ausführliche Informationen zum Konfigurieren des Serverlastausgleichs und zur direkten Serverinteraktion können Sie dem "Nortel Alteon Web OS Application Guide" oder der Veröffentlichung "Command Reference" entnehmen.

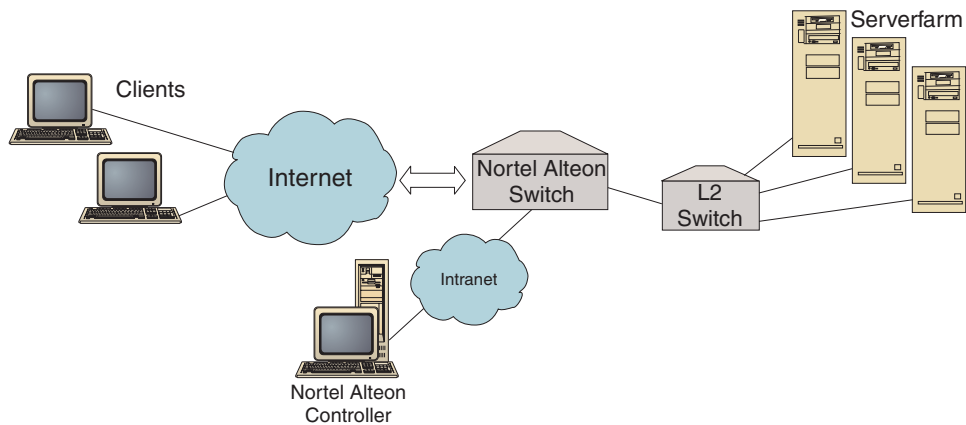


Abbildung 30. Beispiel für einen Consultant, der über ein Intranet vor dem Switch mit dem Netz verbunden ist.

Für die Verwaltung des Nortel Alteon Controller können Sie eine der folgenden Schnittstellen nutzen:

- Browser
- GUI
- ferne Befehlszeile

Abb. 31:

- Die Verbindung zwischen dem Consultant und dem Netz befindet sich hinter dem Switch, für den der Consultant Wertigkeiten bereitstellt.
- Die Benutzerschnittstelle befindet sich auf einem fernen System vor dem Switch.
- Das Netz muss so konfiguriert sein, dass die Benutzerschnittstelle mit dem Controller kommunizieren kann.

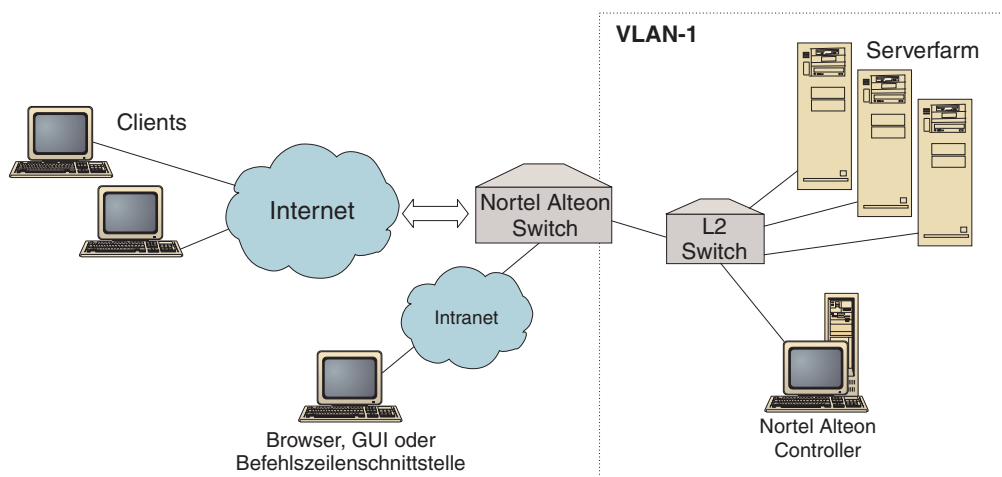


Abbildung 31. Beispiel für einen Consultant hinter dem Switch mit einer Benutzerschnittstelle vor dem Switch

Vom Controller festgelegte Serverattribute auf dem Switch

Wenn ein Consultant Wertigkeiten für Server berechnet, die einen Service bereitstellen, für den ein Switch den Lastausgleich durchführt, inaktiviert der Consultant die normale Überprüfung des Serverzustandes auf dem Switch, um unnötigen Datenverkehr zu den Servern zu vermeiden. Die Zustandsprüfung wird vom Consultant reaktiviert, sobald die Bereitstellung der Wertigkeiten für den Service abgeschlossen ist. Das Intervall für die Überprüfung des Serverzustandes entspricht der MIB-Variablen `slbNewCgRealServerPingInterval`.

Wenn der Consultant feststellt, dass ein Server nicht verfügbar ist, setzt er die maximale Verbindungsanzahl des Servers auf null, damit der Server bei der Verteilung von Anforderungen nicht berücksichtigt wird. Sobald der Server wieder verfügbar ist, wird der ursprüngliche Wert für die maximale Verbindungsanzahl wiederhergestellt. Die maximale Anzahl von Verbindungen des Servers entspricht der MIB-Variablen `slbNewCgRealServerMaxCons`.

Wird eine Wertigkeit für einen realen Server berechnet, wird diese für den Server festgelegt. Die Serverwertigkeit entspricht der MIB-Variablen `slbNewCgRealServerWeight`.

Ausweichserver konfigurieren

Der Switch lässt die Konfiguration von Servern zu, die als Ausweichserver fungieren. Wenn der Switch feststellt, dass ein Server mit verfügbarem Ausweichserver nicht erreichbar ist, kann er Anforderungen an den Ausweichserver senden. Berechnet der Consultant Wertigkeiten für einen Service mit verfügbarem Ausweichservice, werden Wertigkeiten für den Ausweichserver und den primären Server berechnet, damit Wertigkeiten für die Serverauswahl bereit stehen, wenn der Ausweichservice genutzt werden muss.

Die Wertigkeit eines Ausweichservers kann über der eines primären Servers liegen. Der Grund dafür ist, dass keine Anforderungen an den Ausweichserver weitergeleitet werden und dieser somit nicht belastet ist, solange der Switch nicht auf ihn zurückgreift.

Zur Vermeidung ungenutzter Serverressourcen ist es allgemein üblich, dass Server, die einem Service zugeordnet sind, als Ausweichserver für Server verwendet werden, die einem anderen Service zugeordnet sind. Bei der Implementierung einer solchen Konfiguration sollten Sie es vermeiden, dieselben realen Server mehreren gleichzeitig aktiven Services zuzuordnen. Andernfalls überschreibt der Consultant die Wertigkeit des Servers für jeden Service, an dem der Server beteiligt ist.

Jeder reale Server ist durch eine ganze Zahl, eine Wertigkeit und ein IP-Attribut gekennzeichnet. Zwei reale Server können dieselbe IP-Adresse haben. In einem solchen Fall sind zwei reale Server derselben physischen Servermaschine zugeordnet. Die als Ausweichserver vorgesehenen Server sollten nur für die Sicherung eines Services konfiguriert werden. Dient eine physische Maschine zur Sicherung von Servern, die mehreren Services zugeordnet sind, müssen Sie für jeden Service einmal konfiguriert werden und eine für jeden Service eindeutige Server-ID erhalten. Dadurch kann den Ausweichservern für jeden Service eine eindeutige Wertigkeit zugeordnet werden.

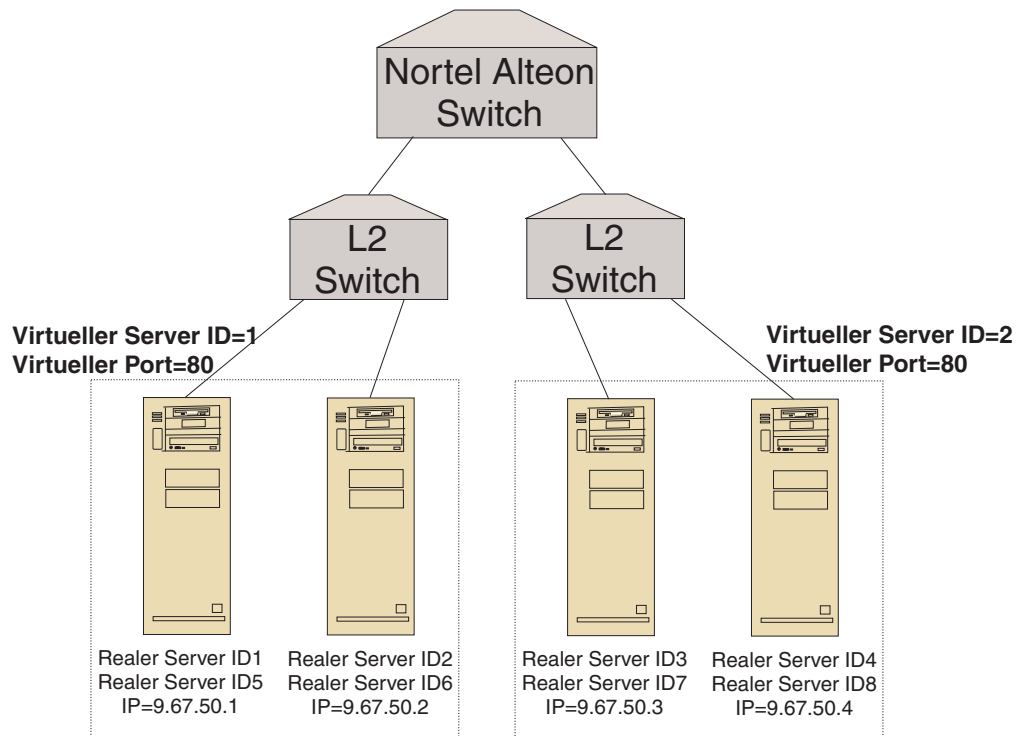


Abbildung 32. Beispiel für einen Consultant mit Ausweichservern

Gruppen konfigurieren

Die Server eines Switch können als Teil mehrerer Gruppen konfiguriert werden. Ebenso können die Gruppen eines Switch für mehrere Services konfiguriert werden.

Da ein Server für mehrere Services konfiguriert werden kann, wird die Wertigkeit für jeden Service berechnet, den der Server anbietet. Deshalb ist nicht sicher, dass die Wertigkeit korrekt ist, denn es ist unbekannt, auf welchen Service sie sich bezieht.

Bestimmt der Consultant Wertigkeiten für einen Service, für einen anderen jedoch nicht, kann der Fall eintreten, dass der Service, für den keine Wertigkeit berechnet wurde, keine Überprüfung des Serverzustandes übernimmt (weil diese inaktiviert ist). In einem solchen Fall kann der Switch keine ordnungsgemäße Lastverteilung für diesen Service sicherstellen.

In Kenntnis dieser Möglichkeiten müssen Sie gewährleisten, dass ein realer Server nicht mehreren am Lastausgleich beteiligten Services zugeordnet wird. Das bedeutet jedoch nicht, dass eine Servermaschine nicht Anfragen nach verschiedenen Services bedienen kann. Es bedeutet viel mehr, dass ein realer Server auf dem Switch für jeden Service, für den er Anfragen bearbeitet, mit einer eindeutigen Kennung konfiguriert werden muss.

Hohe Verfügbarkeit

Sowohl der Nortel Alteon Controller als auch der Nortel Alteon Web Switch bieten eine Funktion für hohe Verfügbarkeit an.

Sie können zwei Controller konfigurieren, die auf zwei verschiedenen Systemen in einer fehlertoleranten Konfiguration ausgeführt werden.

Zwei oder mehr Switches können sich gegenseitig sichern, wenn Sie sie als virtuellen IP-Schnittstellen-Router (VIR) oder virtuellen IP-Server-Router (VSR) konfigurieren.

Ein (vom Controller verwalteter) Consultant stellt nur Wertigkeiten für einen Switch bereit. Da es möglich ist, dass ein Ausweich-Switch die Aufgaben des Master-Switch übernimmt, müssen Sie den Controller mit dem Consultant für jeden Switch konfigurieren, der potenziell als Master-Switch eingesetzt werden könnte. Auf diese Weise ist sichergestellt, dass ein Master-Switch immer Wertigkeiten empfängt.

Wenn die Controller mit einem VIR verbunden sind, ist darüber hinaus ihre Kommunikation mit den Servern, den Switches und dem Ausweichcontroller auch bei Verlust der Konnektivität zu einem der Switches gewährleistet.

Informationen zur hohen Verfügbarkeit der Switches können Sie dem "Nortel Alteon Web OS Application Guide" entnehmen.

Die hohe Controllerverfügbarkeit erweitert die Fähigkeiten von Load Balancer im Bereich der Fehlertoleranz. Die Entwicklung der hohen Controllerverfügbarkeit wurde im Hinblick auf die Verfügbarkeit für die klassische Paketweiterleitung entwickelt und bedeutet die Verfügbarkeit zweier gleichzeitig aktiver Controller, von denen einer der primäre und der andere der sekundäre Controller ist.

Beide Controller werden mit identischen Switch-Daten konfiguriert. Ähnlich wie bei der klassischen hohen Verfügbarkeit ist immer nur ein Controller zur Zeit aktiv. Entsprechend der Logik der hohen Verfügbarkeit bedeutet dies, dass nur der aktive Controller neue Wertigkeiten für den Switch berechnet und aktualisiert.

Für die hohe Controllerverfügbarkeit kommuniziert der Controller mit seinem Partner über eine von Ihnen konfigurierte Adresse und einen von Ihnen konfigurierten Port unter Verwendung des UDP (User Datagram Protocol). Mit diesen Paketen tauschen die Controller die für hohe Verfügbarkeit wichtigen Daten (Erreichbarkeitsdaten) aus und stellen (mit Überwachungssignalen) fest, ob der Partner verfügbar ist. Wenn der Bereitschaftscontroller erkennt, dass der aktive Controller ausgefallen ist, übernimmt er die Aufgaben des aktiven Controllers. Damit wird der Bereitschaftscontroller zum aktiven Controller und beginnt, neue Wertigkeiten für den Switch zu berechnen und zu aktualisieren.

Neben der Verfügbarkeit des Partners können Sie für die hohe Verfügbarkeit Erreichbarkeitsziele konfigurieren. Wie bei der klassischen hohen Verfügbarkeit werden die Erreichbarkeitsdaten bei der hohen Controllerverfügbarkeit verwendet, um festzustellen, welcher der Controller aktiv und welcher der Bereitschaftscontroller ist. Der aktive Controller kann mehr Ziele mit ping erreichen und ist von seinem Partnercontroller aus erreichbar.

Weitere Informationen finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 245.

Abb. 33:

- Zwei Nortel Alteon Controller sind hinter den Switches mit dem Netz verbunden.
- Ein Controller ist der primäre Controller, der Serverwertigkeiten für den Switch bereitstellt. Der andere Controller ist der Ausweichcontroller.
- Die Controller müssen über TCP/IP kommunizieren, damit der Ausweichcontroller weiß, wann er die Aufgaben des primären Controllers übernehmen soll.
- Es werden zwei Nortel Alteon Web Switches konfiguriert, einer als VIR und einer als VSR.
- Der VIR bietet hohe Verfügbarkeit für Verbindungen zu den Servern.
- Der VSR bietet hohe Verfügbarkeit für den Zugriff auf die virtuellen Server, die auf den Switches konfiguriert sind.
- Einer der Switches muss der Master sein und der andere der Ausweich-Switch.
- Der primäre Controller stellt Wertigkeiten für beide Switches bereit.
- Der Ausweichcontroller sendet Überwachungssignale an den primären Controller, damit er weiß, wann er die Aufgaben des primären Controllers übernehmen soll.

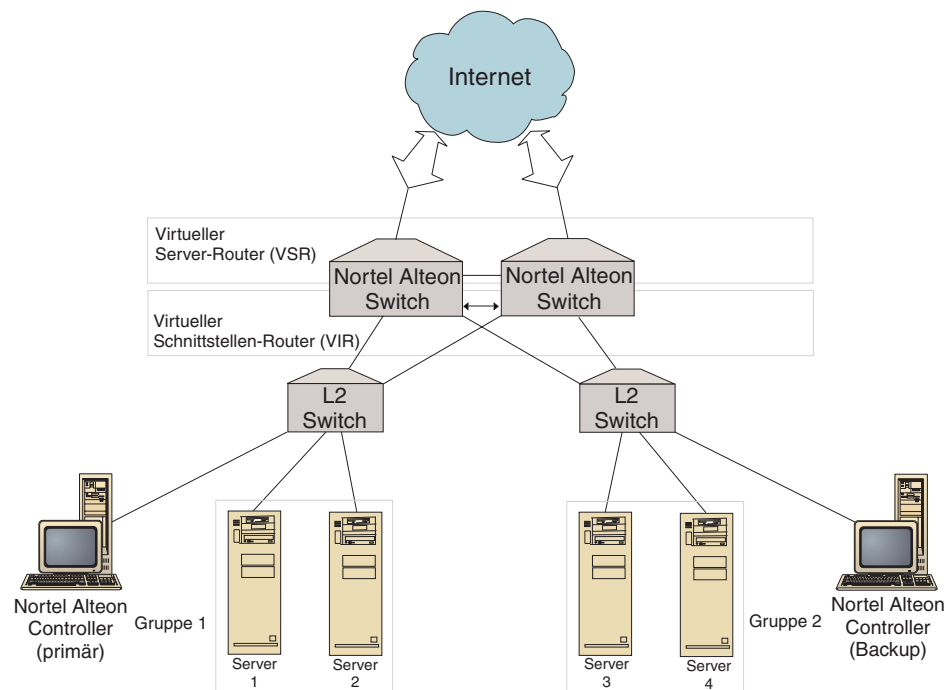


Abbildung 33. Beispiel für hohe Verfügbarkeit mit einem Nortel Alteon Controller und einem Nortel Alteon Web Switch

Optimierung

Sie können den Consultant mit einer Sensitivitätsschwelle konfigurieren, damit sich die Wertigkeiten nicht zu oft ändern. Die Sensitivitätsschwelle gibt an, in welchem Maße sich eine Wertigkeit ändern muss, damit die Änderung als relevant angesehen wird. Weitere Informationen hierzu finden Sie im Abschnitt „Sensitivitätsschwelle“ auf Seite 250.

Wenn der Switch einen zu großen Aufwand mit dem Aktualisieren der Wertigkeiten betreiben muss, können Sie die Inaktivitätszeit des Consultant erhöhen, um den Datenverkehr zwischen dem Controller, den Servern und dem Switch zu reduzieren. Dieser Zeitwert legt die Zeit der Inaktivität zwischen den Definitionszyklen für die Wertigkeit in Sekunden fest.

Wenn die Server zu viele Überwachungsanfragen vom Consultant bearbeiten müssen, können Sie die Inaktivitätszeit der Erfassungsprogramme für Messwerte ändern. Eine ausführliche Beschreibung dieses Prozesses finden Sie im Abschnitt „Ruhezeiten für Wertigkeitsberechnung“ auf Seite 249.

Fehlerbestimmung

Der Cisco CSS Controller schreibt Einträge in die folgenden Protokolle:

- server.log
- consultant.log
- highavailability.log
- metriccollector.log
- binary.log

Diese Protokolle befinden sich in den folgenden Verzeichnissen:

- AIX-, HP-UX-, Linux- und Solaris-Systeme: ...ibm/edge/lb/servers/logs/nal/*Consultant-Name*
- Windows-Systeme: ...ibm\edge\lb\servers\logs\nal*Consultant-Name*

In jedem Protokoll können Sie die Protokollgröße und -stufe festlegen. Weitere Informationen hierzu finden Sie im Abschnitt „Protokolle von Load Balancer verwenden“ auf Seite 267.

Kapitel 20. Nortel Alteon Controller konfigurieren

Lesen Sie vor Ausführung der in diesem Kapitel beschriebenen Schritte Kapitel 19, „Planung für Nortel Alteon Controller“, auf Seite 161. Dieses Kapitel erläutert das Erstellen einer Basiskonfiguration für die Load-Balancer-Komponente Nortel Alteon Controller.

- Komplexere Konfigurationen finden Sie in Kapitel 23, „Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller“, auf Seite 245.
- Informationen zur Fernverwaltung mit Authentifizierung, zu den Protokollen sowie zur Verwendung der Komponente Nortel Alteon Controller finden Sie in Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263.

Konfigurations-Tasks im Überblick

Vor Ausführung einer der in diesem Kapitel beschriebenen Konfigurationsmethoden müssen Sie gewährleisten, dass der Nortel Alteon Web Switch und alle Servermaschinen ordnungsgemäß konfiguriert sind.

Tabelle 11. Konfigurations-Tasks für Nortel Alteon Controller

Task	Beschreibung	Referenzinformationen
Nortel Alteon Web Switch und Server konfigurieren	Konfigurieren des Switch.	Hinweise zum Konfigurieren des Switch auf Seite 174
Maschine mit Nortel Alteon Controller konfigurieren	Konfigurieren des Controllers.	„Schritt 1. Serverfunktion starten“ auf Seite 174
Konfiguration testen	Überprüfung der Konfiguration.	„Konfiguration testen“ auf Seite 176

Konfigurationsmethoden

Es gibt im Wesentlichen drei Methoden für das Erstellen einer Basiskonfiguration für die Load-Balancer-Komponente Nortel Alteon Controller:

- Befehlszeile
- XML-Datei
- grafische Benutzerschnittstelle (GUI).

Befehlszeile

Dies ist die direkte Methode für die Konfiguration von Nortel Alteon Controller. Bei den in diesem Handbuch beschriebenen Prozeduren wird von der Verwendung der Befehlszeile ausgegangen.

Starten Sie Nortel Alteon Controller wie folgt von der Befehlszeile aus:

1. Setzen Sie an der Eingabeaufforderung den Befehl **nalserver** ab. Geben Sie zum Stoppen des Services **nalserver stop** ein.

Anmerkungen:

- a. Klicken Sie auf Windows-Systemen nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf "IBM Nortel Alteon Controller" und wäh-

len Sie "Starten" aus. Zum Stoppen des Services müssen Sie dieselben Schritte ausführen und "Beenden" auswählen.

- b. Auf Windows-Systemen können Sie nalservice automatisch beim Booten starten. Gehen Sie dazu wie folgt vor:
 - 1) Klicken Sie nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**.
 - 2) Klicken Sie mit der rechten Maustaste auf "IBM Nortel Alteon Controller" und wählen Sie "Eigenschaften" aus.
 - 3) Klicken Sie rechts neben "Starttyp" auf den Abwärtspfeil und wählen Sie "Automatisch" aus.
 - 4) Klicken Sie auf "OK".
2. Setzen Sie anschließend die gewünschten Steuerbefehle für Nortel Alteon Controller ab, um die Konfiguration einzurichten. Bei den in diesem Handbuch beschriebenen Prozeduren wird von der Verwendung der Befehlszeile ausgegangen. Der Befehl lautet **nalcontrol**. Weitere Informationen zu Befehlen finden Sie in Kapitel 30, „Befehlsreferenz für Nortel Alteon Controller“, auf Seite 451.

Für die Parameter des Befehls **nalcontrol** können Sie die abgekürzte Form verwenden. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **nalcontrol h e f** an Stelle von **nalcontrol help file** angeben.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie **exit** oder **quit** eingeben.

Anmerkungen:

1. Für alle Parameterwerte des Befehls müssen Sie die englischen Zeichen verwenden. Die einzige Ausnahme hiervon bilden Hostnamen (die in den server-Befehlen verwendet werden) und (die in Dateibefehlen verwendeten) Dateinamen.
2. Auf Windows-Systemen wird dsservice für die Komponente Dispatcher automatisch gestartet. Falls Sie nur mit Nortel Alteon Controller und nicht mit der Komponente Dispatcher arbeiten, können Sie den automatischen Start von ndservice wie folgt unterbinden:
 - a. Klicken Sie im Fenster "Dienste" von Windows mit der rechten Maustaste auf "IBM Dispatcher".
 - b. Wählen Sie den Menüeintrag "Eigenschaften" aus.
 - c. Wählen Sie im Feld **Starttyp** die Option "Manuell" aus.
 - d. Klicken Sie auf OK und schließen Sie das Fenster "Dienste".

XML

Die soeben definierte Konfiguration kann in einer XML-Datei gespeichert werden. So kann die Konfiguration später schnell erneut geladen werden.

Verwenden Sie zum Ausführen des Inhaltes einer XML-Datei (z. B. von **mein-script.xml**) die folgenden Befehle:

- Setzen Sie zum Speichern der aktuellen Konfiguration in einer XML-Datei den folgenden Befehl ab:

nalcontrol file save XML-Dateiname

Verwenden Sie den Befehl **load** nur, wenn Sie zuvor den Befehl **file save** abgesetzt haben.

- Setzen Sie zum Laden einer gespeicherten Konfiguration den folgenden Befehl ab:

nalcontrol file load *XML-Dateiname*

Verwenden Sie den Befehl **load** nur, wenn Sie zuvor den Befehl **file save** abgesetzt haben.

Die XML-Dateien werden im Verzeichnis **...ibm/edge/lb/servers/configurations/nal/** gespeichert.

GUI

Abb. 41 auf Seite 470 zeigt ein Beispiel für die grafische Benutzerschnittstelle (GUI).

Rufen Sie die GUI wie folgt auf:

1. Wenn **nalserver** noch nicht aktiv ist, starten Sie den Dienst jetzt, indem Sie als root **nalserver** eingeben.
2. Führen Sie anschließend einen der folgenden Schritte aus:
 - Auf AIX-, HP-UX-, Linux- oder Solaris-Systemen: Geben Sie **lbadm** ein.
 - Auf Windows-Systemen: Klicken Sie nacheinander auf **Start > Programme > IBM WebSphere > Edge Components > IBM Load Balancer > Load Balancer**.

Gehen Sie zum Konfigurieren von Nortel Alteon Controller von der GUI aus wie folgt vor:

1. Klicken Sie in der Baumstruktur mit der rechten Maustaste auf "Nortel Alteon Controller".
2. Stellen Sie eine Verbindung zu einem Host her.
3. Erstellen Sie einen oder mehrere Switch-Consultant(s) mit den gewünschten Services und den zugehörigen Messwerten.
4. Starten Sie den Consultant.

Von der GUI aus können Sie alle mit dem Befehl **nalcontrol** ausführbaren Schritte ausführen. Beispiel:

- Wenn Sie in der Befehlszeile ein Erreichbarkeitsziel definieren möchten, geben Sie **nalcontrol highavailability usereach** *Adresse* ein. Auf der GUI müssen Sie zum Definieren eines Erreichbarkeitsziels mit der rechten Maustaste auf "Hochverfügbarkeit" > "Erreichbarkeitsziel hinzufügen..." klicken. Geben Sie im Dialogfenster die Erreichbarkeitsadresse ein und klicken Sie auf "OK".
- Bei Auswahl des Eintrags **Konfiguration laden** im Popup-Menü "Host" können Sie die in einer Datei gespeicherte Konfiguration an die aktive Konfiguration anhängen. Wenn Sie eine *neue* Konfiguration laden möchten, müssen Sie vor dem Laden der Datei den Server beenden und neu starten.
- Klicken Sie regelmäßig mit der rechten Maustaste auf **Konfigurationsdatei sichern als**, um Ihre Konfiguration für Nortel Alteon Controller in einer Datei zu speichern.
- Wenn Sie in der Menüleiste **Datei** auswählen, können Sie Ihre aktuellen Hostverbindungen in einer Datei speichern oder Verbindungen aus vorhandenen Dateien aller Load-Balancer-Komponenten wiederherstellen.

Gehen Sie wie folgt vor, um von der GUI aus einen Befehl auszuführen:

1. Klicken Sie mit der rechten Maustaste auf den Knoten **Host** und wählen Sie **Befehl senden...** aus.
2. Geben Sie im Befehlseingabefeld den gewünschten Befehl ein, z. B. **consultant report**.
3. Klicken Sie auf "Senden".

Im Fenster "Ergebnis" sehen Sie die Ergebnisse und die Historie der in der aktuellen Sitzung ausgeführten Befehle.

Falls Sie Hilfe benötigen, klicken Sie oben rechts im Load-Balancer-Fenster auf das Fragezeichen.

- **Hilfe: Feldebene** — beschreibt jedes Feld und gibt die Standardwerte an.
- **Hilfe: Wie funktioniert das** — listet Tasks auf, die von dieser Anzeige aus ausgeführt werden können.
- **InfoCenter** — ermöglicht den zentralen Zugriff auf Produktinformationen.

Weitere Informationen zur Verwendung der GUI finden Sie in Anhang A, „Allgemeine Anweisungen zur GUI“, auf Seite 469.

Nortel Alteon Controller konfigurieren

Hilfe zu den in dieser Prozedur verwendeten Befehlen finden Sie in Kapitel 30, „Befehlsreferenz für Nortel Alteon Controller“, auf Seite 451.

Vor dem Konfigurieren der Maschine mit Nortel Alteon Controller müssen Sie die folgenden Schritte ausführen:

- Auf AIX-, HP-UX-, Linux- und Solaris-Systemen müssen Sie der Benutzer root und auf Windows-Systemen der Administrator sein.
- Nortel Alteon Controller muss IP-Konnektivität zu einem Nortel Alteon Web Switch und zu allen Servern, deren Wertigkeit berechnet werden soll, haben.
- Der Nortel Alteon Web Switch muss wie folgt konfiguriert werden:
 1. Aktivieren Sie für den Switch den Serverlastausgleich der Ebene 4.
 2. Konfigurieren Sie eine IP-Schnittstelle.
 3. Aktivieren Sie SNMP.
 4. Aktivieren Sie die Client-Verarbeitung des Serverlastausgleichs an dem Port, der die Client-Anforderungen empfängt.
 5. Aktivieren Sie die Serververarbeitung des Lastausgleichs an dem Port, mit dem die realen Server verbunden sind.
 6. Konfigurieren Sie reale Server als Webservermaschinen.
 7. Konfigurieren Sie eine reale Servergruppe mit realen Servern, auf denen der Anwendungsserver ausgeführt wird.
 8. Konfigurieren Sie einen virtuellen Server.
 9. Konfigurieren Sie einen Service an einem virtuellen Port und ordnen Sie ihn der realen Servergruppe zu, damit diese den Port bedient.

Schritt 1. Serverfunktion starten

Wenn nalserver noch nicht aktiv ist, starten Sie den Dienst jetzt, indem Sie als root **nalserver** eingeben.

Anmerkung: Klicken Sie auf Windows-Systemen nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf "IBM Nortel Alteon Controller" und wählen Sie "Starten" aus.

Schritt 2. Befehlszeilenschnittstelle aufrufen

Geben Sie **nalcontrol** ein, um die Befehlszeilenschnittstelle aufzurufen.

Schritt 3. Consultant für den Nortel Alteon Web Switch definieren

Geben Sie Folgendes ein, um einen Switch-Consultant hinzuzufügen:

```
consultant add Switch-Consultant-ID address Switch-IP-Adresse
```

Schritt 4. Service zum Switch-Consultant hinzufügen

Geben Sie Folgendes ein, um einen Service hinzuzufügen:

```
service add Switch-Consultant-ID:Service-ID  
vsid ID_des_virtuellen_Servers vport  
Nummer_des_virtuellen_Ports
```

Ein Service ist durch die ID eines virtuellen Servers (VSID) und die Nummer eines virtuellen Ports (VPORT) gekennzeichnet. Beide Werte sind einem virtuellen Server zugeordnet, der zuvor für den Switch konfiguriert wurde.

Schritt 5. Messwerte konfigurieren

Anhand der Messwerte wird die Wertigkeit der Server bestimmt. Jedem Messwert wird eine proportionale Bedeutung zugeordnet, um seine Wertigkeit im Verhältnis zu anderen Messwerten anzugeben. Sie können beliebige Kombinationen von Messwerten konfigurieren: Messwerte für Verbindungsdaten, für Advisor-Funktionen der Anwendung und für Metric Server. Die proportionalen Gewichtungen müssen in der Summe stets 100 ergeben.

Wenn Sie einen Service konfigurieren, werden die Standardmesswerte **activeconn** und **connrate** definiert. Falls Sie zusätzliche oder gänzlich andere Messwerte verwenden möchten, geben Sie Folgendes ein:

```
service metrics  
Switch-Consultant-ID:Service-ID Messwertname 50  
Messwertname2 50
```

Schritt 6. Consultant starten

Geben Sie zum Starten des Consultant Folgendes ein:

```
consultant start Switch-Consultant-ID
```

Mit diesem Befehl wird die Erfassung von Messwerten gestartet, und die Berechnung von Wertigkeiten beginnt.

Schritt 7. Hohe Verfügbarkeit konfigurieren (optional)

Geben Sie Folgendes ein, um die hohe Verfügbarkeit zu konfigurieren:

```
highavailability add  
address IP-Adresse partneraddress IP-Adresse port 80  
role primary
```

Ausführliche Informationen zur Verwendung und zum Konfigurieren der hohen Controllerverfügbarkeit finden Sie in Kapitel 23, „Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller“, auf Seite 245.

Schritt 8. Metric Server starten (optional)

Wenn Sie in Schritt 5 Systemmesswerte definiert haben, muss auf den Service-maschinen Metric Server gestartet werden. Informationen zur Verwendung von Metric Server finden Sie im Abschnitt „Metric Server“ auf Seite 255.

Schritt 9. Konfiguration für Nortel Alteon Controller aktualisieren

Wenn Sie die Konfiguration auf dem Nortel Alteon Web Switch ändern, können Sie die Controllerkonfiguration aktualisieren. Geben Sie Folgendes ein:

service refresh

Beenden Sie den Consultant vor dem Aktualisieren der Konfiguration. Starten Sie den Consultant neu, wenn der Befehl refresh die Konfiguration aktualisiert hat.

Konfiguration testen

Testen Sie wie folgt, ob die Konfiguration korrekt ist:

1. Setzen Sie die Protokollstufe des Consultant auf 4.
2. Trennen Sie einen Server für eine Minute vom Nortel Alteon Web Switch oder fahren Sie den Anwendungsserver für eine Minute herunter.
3. Stellen Sie die Verbindung des Servers zum Switch wieder her oder führen Sie einen Neustart für den Anwendungsserver aus.
4. Setzen Sie die Protokollstufe des Consultant auf den gewünschten Wert (1) zurück.
5. Rufen Sie die Datei consultant.log in den folgenden Verzeichnissen auf und suchen Sie nach dem Eintrag **setServerWeights setting service**. Dieser Eintrag weist darauf hin, dass versucht wurde, Wertigkeiten an den Switch zu senden.
 - AIX-, HP-UX-, Linux- und Solaris-Systeme: ...ibm/edge/lb/servers/logs/cco/*Consultant-Name*
 - Windows-Systeme: ...ibm\edge\lb\servers\logs\cco*Consultant-Name*
6. Sehen Sie sich die Serverwertigkeiten auf dem Switch an und prüfen Sie, ob diese mit den im Controllerbericht angezeigten Wertigkeiten übereinstimmen.

Teil 7. Funktionen und erweiterte Features von Load Balancer

Dieser Teil enthält Informationen zu den Funktionen und erweiterten Konfigurationsfeatures, die für Load Balancer verfügbar sind. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 21, „Manager, Advisor-Funktionen und Metric Server für Dispatcher, CBR und Site Selector“, auf Seite 179
- Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201
- Kapitel 23, „Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller“, auf Seite 245

Kapitel 21. Manager, Advisor-Funktionen und Metric Server für Dispatcher, CBR und Site Selector

In diesem Kapitel wird erklärt, wie die Lastausgleichparameter, die Manager, die Advisor-Funktionen und die Funktion Metric Server von Load Balancer konfiguriert werden.

Anmerkung: Falls Sie die Komponente Dispatcher *nicht* verwenden, ersetzen Sie beim Lesen dieses Kapitels dscontrol durch Folgendes:

- Für CBR: **cbrcontrol**
- Für Site Selector: **sscontrol** (lesen Sie hierzu die Informationen in Kapitel 28, „Befehlsreferenz für Site Selector“, auf Seite 405)

Wichtiger Hinweis: Falls Sie Load Balancer für IPV4 und IPV6 verwenden, sollten Sie sich vor der Beschäftigung mit dem Inhalt dieses Abschnitts in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81 über Einschränkungen und Konfigurationsunterschiede informieren.

Tabelle 12. Erweiterte Konfigurations-Tasks für Load Balancer

Task	Beschreibung	Referenzinformationen
Ändern der Einstellungen für den Lastausgleich (optional)	Sie können die folgenden Einstellungen für den Lastausgleich ändern: <ul style="list-style-type: none">• Die proportionale Gewichtung der Statusinformationen. Das Standardverhältnis ist 50-50-0-0. Wenn Sie den Standardwert verwenden, werden die Informationen von den Advisor-Funktionen, von Metric Server und von WLM nicht benutzt.• Wertigkeiten• Feste Wertigkeiten vom Manager• Manager-Intervalle• Sensitivitätsschwelle• Glättungsfaktor	„Lastausgleich mit Load Balancer optimieren“ auf Seite 180
Verwendung von Scripts, um einen Alert zu generieren oder Serverausfälle zu protokollieren, wenn der Manager Server als inaktiv oder aktiv markiert	Load Balancer stellt Benutzer-Exits bereit, die Scripts aktivieren, wenn der Manager Server als inaktiv oder aktiv markiert.	„Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden“ auf Seite 184
Advisor-Funktionen verwenden	Beschreibt die Advisor-Funktionen, die Berichte zu spezifischen Status Ihrer Server ausgeben.	„Advisor-Funktionen“ auf Seite 185
Option "Anforderung/Antwort (URL)" der HTTP- oder HTTPS-Advisor-Funktion verwenden	Definieren Sie eine eindeutige HTTP-URL-Zeichenfolge für einen spezifischen Dienst, der auf der Maschine abgefragt werden soll.	„Option 'Anforderung/Antwort (URL)' der HTTP- oder HTTPS-Advisor-Funktion konfigurieren“ auf Seite 191
Advisor-Funktion self verwenden	Stellt den Auslastungsstatus von Back-End-Servern für Load Balancer in einer Client/Server-WAN-Konfiguration bereit.	„Advisor-Funktion 'self' in einer Client/Server-WAN-Konfiguration“ auf Seite 192

Tabelle 12. Erweiterte Konfigurations-Tasks für Load Balancer (Forts.)

Task	Beschreibung	Referenzinformationen
Angepasste Advisor-Funktionen erstellen	Erklärt das Schreiben eigener angepasster Advisor-Funktionen.	„Kundenspezifische (anpassbare) Advisor-Funktion erstellen“ auf Seite 193
Verwendung des Agenten Metric Server	Metric Server stellt Informationen zur Systembelastung für Load Balancer bereit.	„Metric Server“ auf Seite 196
Verwendung der WLM-Advisor-Funktion (Workload Manager)	Die WLM-Advisor-Funktion stellt Informationen zur Systembelastung für Load Balancer bereit.	„Advisor-Funktion Workload Manager“ auf Seite 199

Lastausgleich mit Load Balancer optimieren

Die Managerfunktion von Load Balancer führt den Lastausgleich ausgehend von den folgenden Einstellungen durch:

- „Proportionale Bedeutung von Statusinformationen“
- „Wertigkeiten“ auf Seite 181
- „Manager-Intervalle“ auf Seite 183
- „Advisor-Intervalle“ auf Seite 187
- „Berichtszeitlimit für Advisor-Funktion“ auf Seite 188
- „Sensitivitätsschwelle“ auf Seite 183
- „Glättungsfaktor“ auf Seite 184

Zur Optimierung des Lastausgleichs für Ihr Netz können Sie diese Einstellungen ändern.

Proportionale Bedeutung von Statusinformationen

Der Manager kann in seine Gewichtungsentscheidung alle oder einige der nachfolgend genannten externen Faktoren einfließen lassen:

- *Aktive Verbindungen*: Die (vom Executor protokollierte) Anzahl aktiver Verbindungen auf jeder am Lastausgleich beteiligten Servermaschine. Diese Proportion gilt nicht für Site Selector.

Oder —

CPU: Prozentsatz der auf jeder am Lastausgleich beteiligten Servermaschine genutzten CPU (Vorgabe vom Agenten Metric Server). Für Site Selector wird diese Proportion an Stelle der Spalte für aktive Verbindungen angezeigt.

- *Neue Verbindungen*: Die (vom Executor protokollierte) Anzahl neuer Verbindungen auf jeder am Lastausgleich beteiligten Servermaschine. Diese Proportion gilt nicht für Site Selector.

Oder —

Speicher: Prozentsatz des auf jeder am Lastausgleich beteiligten Servermaschine genutzten Speichers (Vorgabe vom Metric Server Agent). Für Site Selector wird diese Proportion an Stelle der Spalte für neue Verbindungen angezeigt.

- *Port-spezifisch*: Vorgaben von den Advisor-Funktionen, die am Port empfangsbereit sind.
- *Systemmesswert*: Vorgabe von den Systemüberwachungs-Tools wie Metric Server oder WLM.

Der Manager erhält die beiden ersten Werte (aktive und neue Verbindungen) neben der aktuellen Wertigkeit jedes Servers und anderen für die Berechnungen erforder-

lichen Informationen vom Executor. Diese Werte basieren auf Informationen, die intern vom Executor generiert und gespeichert werden.

Anmerkung: Für Site Selector erhält der Manager die beiden ersten Werte (CPU und Speicher) von Metric Server.

Sie können die relative Bedeutung der vier Werte pro Cluster (oder Sitenamen) ändern. Die Proportionen sind vergleichbar mit Prozentsätzen. Die Summe der relativen Proportionen muss 100 % ergeben. Das Standardverhältnis ist 50/50/0/0, wobei die Advisor- und Systeminformationen ignoriert werden. In Ihrer Umgebung sollten Sie andere Proportionen ausprobieren, um die Kombination mit der besten Leistung zu finden.

Anmerkung: Wenn Sie eine Advisor-Funktion (mit Ausnahme von WLM) hinzufügen und die **Port-Proportion** null ist, erhöht der Manager diesen Wert auf 1. Da die Summe der relativen Proportionen immer 100 ist, muss der höchste Wert um 1 vermindert werden.

Wenn Sie die Advisor-Funktion WLM hinzufügen und die **Proportion der Systemmesswerte** null ist, erhöht der Manager diesen Wert auf 1. Da die Summe der relativen Proportionen immer 100 ist, muss der höchste Wert um 1 vermindert werden.

Die Anzahl aktiver Verbindungen hängt sowohl von der Anzahl der Clients als auch von der Zeit ab, die für die Nutzung der von den am Lastausgleich beteiligten Servermaschinen bereitgestellten Dienste erforderlich ist. Sind die Client-Verbindungen schnell (wie bei kleinen Webseiten, die mit HTTP GET bedient werden), ist die Anzahl aktiver Verbindungen ziemlich klein. Wenn die Client-Verbindungen langsamer sind (z. B. bei einer Datenbankabfrage), wird die Anzahl aktiver Verbindungen höher sein.

Sie sollten eine zu niedrige Proportionseinstellung für aktive und neue Verbindungen vermeiden. Wenn Sie diese beiden ersten Werte nicht jeweils auf mindestens 20 gesetzt haben, werden der Lastausgleich und die Glättungsfunktion inaktiviert.

Verwenden Sie zum Festlegen der proportionalen Bedeutung den Befehl **dscontrol cluster set Cluster proportions**. Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol cluster — Cluster konfigurieren“ auf Seite 357.

Wertigkeiten

Die Manager-Funktion legt Wertigkeiten ausgehend von internen Zählern des Executors, vom Feedback der Advisor-Funktionen und vom Feedback eines Systemüberwachungsprogramms wie Metric Server fest. Falls Sie Wertigkeiten bei Ausführung des Managers manuell festlegen möchten, geben Sie den Befehl **dscontrol server** mit der Option **fixedweight** an. Eine Beschreibung der Option **fixedweight** finden Sie im Abschnitt „Feste Wertigkeiten vom Manager“ auf Seite 182.

Wertigkeiten gelten für alle Server an einem Port. An einem bestimmten Port werden die Anfragen ausgehend von einem Vergleich der Wertigkeiten der einzelnen Server verteilt. Hat beispielsweise ein Server die Wertigkeit 10 und der andere Server die Wertigkeit 5, erhält der Server mit der Wertigkeit 10 doppelt so viele Anforderungen wie der Server mit der Wertigkeit 5.

Für die Wertigkeit, die ein Server haben kann, können Sie einen oberen Grenzwert angeben. Verwenden Sie dazu den Befehl **dscontrol port set Port weightbound**

Wertigkeit. Mit diesem Befehl wird die Differenz festgelegt, die für die einzelnen Server hinsichtlich der Anzahl der Anforderungen gelten soll. Wird die Wertigkeitsobergrenze auf 1 gesetzt, können alle Server die Wertigkeit 1 haben. Stillgelegte Server haben die Wertigkeit 0 und inaktive Server die Wertigkeit -1. Wenn Sie diese Zahl erhöhen, vergrößern sich die Unterschiede bei der Gewichtung von Servern. Bei einer Wertigkeitsobergrenze von 2 kann ein Server doppelt so viele Anforderungen wie ein anderer Server erhalten. Bei einer Wertigkeitsobergrenze von 10 kann ein Server zehn Mal so viele Anforderungen wie ein anderer Server erhalten. Der Standardwert für die Wertigkeitsobergrenze ist 20.

Stellt eine Advisor-Funktion fest, dass ein Server inaktiviert wurde, informiert er den Manager, der die Wertigkeit für den Server auf null setzt. Der Executor sendet in diesem Fall keine weiteren Verbindungen an diesen Server, solange die Wertigkeit bei null liegt. Falls es vor Änderung der Wertigkeit aktive Verbindungen zum Server gab, können diese normal beendet werden.

Wenn alle Server inaktiv sind, setzt der Manager die Wertigkeit auf die Hälfte der Gewichtungsgrenze.

Feste Wertigkeiten vom Manager

Ohne den Manager können Advisor-Funktionen nicht ausgeführt werden und nicht erkennen, ob ein Server inaktiv ist. Wenn Sie die Advisor-Funktionen ausführen möchten, der Manager jedoch *nicht* die von Ihnen für einen bestimmten Server festgelegte Wertigkeit aktualisieren soll, verwenden Sie den Befehl `dscontrol server` mit der Option **fixedweight**. Beispiel:

```
dscontrol server set Cluster:Port:Server fixedweight yes
```

Nachdem Sie **fixedweight** auf **yes** gesetzt haben, können Sie die Wertigkeit mit dem Befehl **dscontrol server set weight** auf den gewünschten Wert setzen. Der Wert für die Serverwertigkeit bleibt während der Ausführung des Managers unverändert erhalten, bis Sie einen weiteren Befehl `dscontrol server` absetzen, bei dem **fixedweight** auf **no** gesetzt ist. Weitere Informationen hierzu finden Sie im Abschnitt „`dscontrol server` — Server konfigurieren“ auf Seite 394.

TCP-Rücksetzanforderung an einen inaktiven Server senden (nur Komponente Dispatcher)

Wenn TCP **reset** aktiviert ist, sendet der Dispatcher eine TCP-Rücksetzanforderung an den Client, sofern dieser eine Verbindung zu einem Server mit der Wertigkeit 0 hat. Die Wertigkeit eines Servers kann 0 sein, weil sie mit dem Wert 0 konfiguriert wurde oder weil der Server von einer Advisor-Funktion als inaktiv markiert wurde. Eine TCP-Rücksetzanforderung bewirkt eine sofortige Beendigung der Verbindung. Dieses Feature ist hilfreich für lang andauernde Verbindungen, denn sie gibt dem Client die Möglichkeit, eine unterbrochene Verbindung schnell neu auszuhandeln. Verwenden Sie zum Aktivieren von TCP-Rücksetzanforderungen den Befehl **dscontrol port add | set Port reset yes**. Der Standardwert für **reset** ist **no**.

Anmerkung: Die TCP-Rücksetzanforderung kann auf alle Dispatcher-Weiterleitungsmethoden angewendet werden. Voraussetzung für die Verwendung des Feature für TCP-Rücksetzanforderungen ist jedoch, dass der Parameter **clientgateway** des Befehls **dscontrol executor** auf eine Router-Adresse gesetzt ist.

In Verbindung mit TCP-Rücksetzanforderungen kann es sinnvoll sein, das Feature **advisor retry** zu konfigurieren. Dieses Feature gibt einer Advisor-Funktion die Möglichkeit, einen erneuten Verbindungsversuch zu starten, bevor ein Server als

inaktiv markiert wird. Auf diese Weise wird verhindert, dass die Advisor-Funktion den Server vorschnell als inaktiv markiert, was zu Problemen beim Zurücksetzen der Verbindung führen könnte. Ein erster gescheiterter Verbindungsversuch der Advisor-Funktion muss nicht zwingend bedeuten, dass die vorhandenen Verbindungen ebenfalls unterbrochen sind. Weitere Informationen hierzu finden Sie im Abschnitt „Wiederholungsversuche der Advisor-Funktion“ auf Seite 188.

Manager-Intervalle

Um den Gesamtdurchsatz zu optimieren, wird die Interaktion von Manager und Executor in ihrer Häufigkeit eingeschränkt. Sie können dieses Intervall mit den Befehlen **dscontrol manager interval** und **dscontrol manager refresh** ändern. Mit dem Manager-Intervall wird angegeben, wie oft der Manager die Serverwertigkeiten aktualisiert, die der Executor für die Weiterleitung von Verbindungen benutzt. Ein zu niedriges Manager-Intervall kann sich negativ auf die Leistung auswirken, da der Manager den Executor permanent unterbricht. Ein zu hohes Manager-Intervall kann bedeuten, dass die Weiterleitung von Anforderungen durch den Executor nicht auf genauen, auf dem neuesten Stand befindlichen Informationen basiert.

Wollen Sie beispielsweise das Manager-Intervall auf 1 Sekunde setzen, geben Sie den folgenden Befehl ein:

```
dscontrol manager interval 1
```

Der Manager-Aktualisierungszyklus (Refresh) gibt an, wie oft der Manager Statusinformationen vom Executor anfordert. Der Aktualisierungszyklus basiert auf der Intervallzeit.

Wollen Sie beispielsweise den Manager-Aktualisierungszyklus auf 3 setzen, geben Sie den folgenden Befehl ein:

```
dscontrol manager refresh 3
```

In diesem Fall wartet der Manager 3 Intervalle ab, bevor er Statusinformationen vom Executor anfordert.

Sensitivitätsschwelle

Zur Optimierung des Lastausgleichs für Ihre Server stehen weitere Methoden zur Verfügung. Im Interesse einer hohen Übertragungsgeschwindigkeit werden die Wertigkeiten der Server nur aktualisiert, wenn sich signifikante Änderungen der Wertigkeit ergeben. Das permanente Aktualisieren der Wertigkeiten bei geringfügigen oder nicht vorhandenen Änderungen des Serverstatus würde zu einem unnötigen Systemaufwand führen. Wenn die prozentuale Änderung der Wertigkeit innerhalb der summierten Wertigkeit für alle Server an einem Port über der Sensitivitätsschwelle liegt, aktualisiert der Manager die vom Executor für die Verteilung der Verbindungen verwendeten Wertigkeiten. Nehmen wir beispielsweise an, die Gesamtwertigkeit ändert sich von 100 % auf 105 %. Die Änderung beträgt also 5 %. Beim standardmäßigen Sensitivitätsschwellenwert von 5 aktualisiert der Manager nicht die vom Executor verwendeten Wertigkeiten, da die prozentuale Änderung nicht **über** dem Schwellenwert liegt. Ändert sich die Gesamtwertigkeit jedoch von 100 % auf 106 %, aktualisiert der Manager die Wertigkeiten. Wollen Sie beispielsweise die Sensitivitätsschwelle des Managers auf einen anderen Wert als den Standardwert setzen (zum Beispiel 6), geben Sie den folgenden Befehl ein:

```
dscontrol manager sensitivity 6
```

In den meisten Fällen müssen Sie diesen Wert nicht ändern.

Glättungsfaktor

Der Manager berechnet die Serverwertigkeiten dynamisch. Daher kann eine aktualisierte Wertigkeit erheblich von der vorherigen Wertigkeit abweichen. In den meisten Fällen stellt dies kein Problem dar. Gelegentlich kann dies jedoch zu erheblichen Schwankungen bei der Verteilung von Anforderungen führen. So kann beispielsweise ein Server aufgrund seiner hohen Wertigkeit den größten Teil der Anforderungen erhalten. Der Manager stellt fest, dass der Server über eine hohe Anzahl von aktiven Verbindungen verfügt und sehr langsam antwortet. Der Manager verschiebt die Wertigkeit dann auf die freien Server, so dass dort derselbe Effekt auftritt und Ressourcen folglich ineffizient genutzt werden.

Um die Auswirkungen dieses Problems zu verringern, benutzt der Manager einen Glättungsfaktor. Der Glättungsfaktor begrenzt das Maß, in dem sich die Wertigkeit eines Servers ändern kann, und dämpft so die Änderung bei der Verteilung von Anforderungen. Ein höherer Glättungsfaktor führt zu einer weniger drastischen Änderung der Serverwertigkeiten. Ein geringerer Glättungsfaktor führt zu einer drastischeren Änderung der Serverwertigkeiten. Der Standardwert für den Glättungsfaktor ist 1,5. Bei einem Wert von 1,5 können Serverwertigkeiten sehr dynamisch sein. Bei einem Faktor von 4 oder 5 sind die Wertigkeiten stabiler. Wenn Sie den Glättungsfaktor beispielsweise auf 4 setzen möchten, geben Sie den folgenden Befehl ein:

```
dscontrol manager smoothing 4
```

In den meisten Fällen müssen Sie diesen Wert nicht ändern.

Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden

Load Balancer stellt Benutzer-Exits bereit, die Scripts aktivieren, die von Ihnen angepasst werden können. Sie können Scripts für die Ausführung automatisierter Aktionen erstellen. Eine solche Aktion wäre beispielsweise das Informieren eines Administrators über inaktive Server per Alert oder das Registrieren eines Ausfalls. Scripts, die Sie anpassen können, finden Sie im Installationsverzeichnis **...ibm/edge/lb/servers/samples**. Zum Ausführen der Dateien müssen Sie sie in das Verzeichnis **...ibm/edge/lb/servers/bin** verschieben und die Erweiterung "sample" löschen. Es stehen die folgenden Beispiel-Scripts bereit:

- **serverDown** — Ein Server wird vom Manager als inaktiv markiert.
- **serverUp** — Ein Server wird vom Manager als aktiv markiert.
- **managerAlert** — Alle Server für einen bestimmten Port werden als inaktiv markiert.
- **managerClear** — Mindestens ein Server ist aktiv, nachdem alle Server für einen bestimmten Port als inaktiv markiert wurden.

Wenn alle Server eines Clusters (vom Benutzer oder von den Advisor-Funktionen) als inaktiv markiert wurden, wird **managerAlert** gestartet (sofern konfiguriert), und Load Balancer versucht, den Datenverkehr mit einer RoundRobin-Technik an die Server weiterzuleiten. Das Script **serverDown** wird nicht gestartet, wenn festgestellt wird, dass der letzte Server des Clusters offline ist.

Load Balancer ist so konzipiert, dass die Weiterleitungsversuche fortgesetzt werden, falls ein Server wieder online geht und auf die Anforderung reagiert. Würde Load Balancer dagegen den gesamten Datenverkehr übergehen, würde der Client keine Antwort empfangen.

Stellt Load Balancer fest, dass der erste Server eines Clusters wieder online ist, wird das Script `managerClear` gestartet (sofern konfiguriert). Das Script `serverUp` (sofern konfiguriert) wird erst ausgeführt, wenn ein weiterer Server wieder online ist.

Hinweise zur Verwendung der Scripts **serverUp** und **serverDown**:

- Wenn Sie den Manager-Zyklus mit weniger als 25 % der Advisor-Zeit definieren, können falsche Berichte zur Aktivität bzw. Inaktivität von Servern erstellt werden. Standardmäßig wird der Manager alle zwei Sekunden ausgeführt, der Advisor jedoch alle sieben Sekunden. Der Manager erwartet demzufolge innerhalb von vier Zyklen neue Advisor-Informationen. Wenn Sie diese Einschränkung umgehen wollen (und den Manager-Zyklus mit mehr als 25 % der Advisor-Zeit definieren) geht der Durchsatz spürbar zurück, weil mehrere Advisor-Funktionen Informationen zu einem Server geben können.
- Fällt ein Server aus, wird das Script `serverDown` gestartet. Wenn Sie jedoch einen `serverUp`-Befehl absetzen, wird angenommen, dass der Server solange aktiv ist, bis der Manager neue Informationen vom Advisor-Zyklus erhalten hat. Ist der Server noch inaktiv, wird erneut das Script `serverDown` ausgeführt.

Advisor-Funktionen

Advisor-Funktionen sind Agenten von Load Balancer. Ihr Zweck ist es, den Zustand und die Belastung der Servermaschinen zu beurteilen. Dies erfolgt durch einen proaktiven Austausch mit den Servern, der dem von Clients vergleichbar ist. Advisor können als transportable Clients der Anwendungsserver betrachtet werden.

Das Produkt stellt mehrere protokollspezifische Advisor-Funktionen für die am häufigsten verwendeten Protokolle zur Verfügung. Es ist jedoch nicht sinnvoll, alle verfügbaren Advisor-Funktionen mit jeder Komponente von Load Balancer zu verwenden. (Die Telnet-Advisor-Funktion wird beispielsweise nicht mit der Komponente CBR verwendet.) Load Balancer unterstützt auch das Konzept der „angepassten Advisor-Funktion“, so dass Benutzer eigene Advisor-Funktionen schreiben können.

Einschränkung für die Verwendung bindungsspezifischer Serveranwendungen: Wenn Sie Advisor-Funktionen für einen bindungsspezifischen Server verwenden möchten, müssen Sie zwei Instanzen des Servers starten, eine Instanz, die an den Cluster:Port gebunden wird, und eine, die an den Server:Port gebunden wird. Mit dem Befehl `netstat -an` können Sie feststellen, ob der Server bindungsspezifisch ist. Achten Sie auf den Server:Port. Wenn der Server nicht bindungsspezifisch ist, gibt der Befehl `0.0.0.0:80` zurück. Andernfalls sehen Sie eine Adresse wie `192.168.15.103:80`.

Einschränkung auf HP-UX- und Solaris-Systemen für die Verwendung bindungsspezifischer Serveranwendungen: Wenn an Stelle des Befehls `ifconfig alias` der Befehl `arp publish` verwendet wird, *unterstützt* Load Balancer die Verwendung von Advisor-Funktionen beim Lastausgleich für Server mit bindungsspezifischen Serveranwendungen (einschließlich anderer Komponenten von Load Balancer wie CBR oder Site Selector), wenn die Bindung an die Cluster-IP-Adresse

erfolgt. Bei Verwendung von Advisor-Funktionen für eine bindungsspezifische Anwendung sollten Sie Load Balancer jedoch nicht mit der Serveranwendung auf derselben Maschine verknüpfen.

Anmerkung: Wenn Load Balancer auf einem Computer mit mehreren Netzwerkadaptoren ausgeführt wird und der Advisor-Datenverkehr über einen bestimmten Adapter fließen soll, können Sie für die Pakete eine bestimmte Quellen-IP-Adresse erzwingen. Falls Sie für die Advisor-Pakete eine bestimmte Quellenadresse erzwingen möchten, fügen Sie zur Zeile `java...SRV_XXXConfigServer...` des entsprechenden Start-Scripts für Load Balancer (dsserver, cbrserver oder ssserver) Folgendes hinzu:

```
-DLB_ADV_SRC_ADDR=IP-Adresse
```

Arbeitsweise der Advisor-Funktionen

Advisor-Funktionen öffnen regelmäßig eine TCP-Verbindung zu jedem Server und senden eine Anforderungsnachricht an den Server. Der Inhalt der Nachricht ist spezifisch für das Protokoll, das auf dem Server ausgeführt wird. Die HTTP-Advisor-Funktion sendet beispielsweise eine HTTP-Anfrage „HEAD“ an den Server.

Die Advisor-Funktionen warten dann auf den Empfang einer Antwort vom Server. Nach Empfang der Antwort beurteilt die Advisor-Funktion den Server. Um diesen „Lastwert“ zu ermitteln, messen die meisten Advisor-Funktionen die Zeit, bis der Server antwortet, und verwenden dann diesen Wert (in Millisekunden) als Lastwert.

Die Advisor-Funktionen übergeben dann den Lastwert an die Manager-Funktion, die ihn im Manager-Bericht in der Spalte „Port“ angibt. Der Manager addiert anschließend die Wertigkeiten für alle Quellen entsprechend ihren Proportionen und übergibt diese Werte an die Executor-Funktion. Der Executor benutzt diese Wertigkeiten dann für den Lastausgleich neuer ankommender Client-Verbindungen.

Stellt die Advisor-Funktion fest, dass ein Server aktiv ist und ordnungsgemäß arbeitet, meldet er einen positiven Lastwert ungleich null an den Manager. Stellt die Advisor-Funktion fest, dass ein Server inaktiv ist, gibt er den speziellen Lastwert -1 zurück. Der Manager und der Executor leiten daraufhin keine Verbindungen an diesen Server weiter, solange dieser inaktiv ist.

Anmerkung: Vor dem Senden der ersten Anforderungsnachricht sendet die Advisor-Funktion ein ping an den Server. Auf diese Weise steht schnell ein Status zur Verfügung, mit dem festgestellt werden kann, ob die Maschine online ist. Antwortet der Server auf das ping, werden keine weiteren ping-Signale gesendet. Wenn Sie die ping-Signale inaktivieren möchten, fügen Sie zur Start-Script-Datei von Load Balancer `-DLB_ADV_NB_PING` hinzu.

Advisor-Funktion starten und stoppen

Sie können eine Advisor-Funktion clusterübergreifend für einen bestimmten Port starten (Gruppen-Advisor-Funktion). Sie können aber auch an einem Port verschiedene Advisor-Funktionen für verschiedene Cluster ausführen (cluster-/sitespezifische Advisor-Funktion). Wenn Sie Load Balancer beispielsweise mit drei Clustern (*ClusterA*, *ClusterB*, *ClusterC*), jeweils mit Port 80, konfiguriert haben, können Sie folgende Schritte ausführen:

- Cluster-/sitespezifische Advisor-Funktion: Geben Sie zum Starten einer Advisor-Funktion am Port 80 für *ClusterA* wie folgt den Cluster und den Port an:

```
dscontrol advisor start
http ClusterA:80
```

Dieser Befehl startet die Advisor-Funktion HTTP am Port 80 für *ClusterA*. Die Advisor-Funktion HTTP wird für alle Port 80 von *ClusterA* zugeordneten Server ausgeführt.

- Gruppen-Advisor-Funktion: Geben Sie zum Starten einer angepassten Advisor-Funktion am Port 80 für alle anderen Cluster wie folgt den Port an:

```
dscontrol advisor start angepasster_Advisor 80
```

Dieser Befehl startet die Advisor-Funktion *angepasster_Advisor* am Port 80 von *ClusterB* und *ClusterC*. Die angepasste Advisor-Funktion wird für alle Port 80 von *ClusterB* und *ClusterC* zugeordneten Server ausgeführt. (Weitere Informationen zu angepassten Advisor-Funktionen finden Sie im Abschnitt „Kundenspezifische (anpassbare) Advisor-Funktion erstellen“ auf Seite 193.)

Anmerkung: Die Gruppen-Advisor-Funktion wird für alle Cluster/Sites ausgeführt, für die es derzeit keine cluster-/sitespezifische Advisor-Funktion gibt.

Wenn Sie das vorherige Konfigurationsbeispiel für die Gruppen-Advisor-Funktion verwenden, können Sie bei Bedarf die angepasste Advisor-Funktion *angepasster_Advisor* am Port 80 für einen oder beide Cluster (*ClusterB* und *ClusterC*) stoppen.

- Geben Sie zum Stoppen der angepassten Advisor-Funktion am Port 80 von *ClusterB* wie folgt Cluster und Port an:

```
dscontrol advisor stop angepasster_Advisor ClusterB:80
```

- Zum Stoppen der angepassten Advisor-Funktion am Port 80 von *ClusterB* und *ClusterC* müssen Sie wie folgt nur den Port angeben:

```
dscontrol advisor stop angepasster_Advisor 80
```

Advisor-Intervalle

Anmerkung: Die Advisor-Standardwerte funktionieren in den meisten Fällen effizient. Gehen Sie mit Vorsicht vor, wenn Sie andere Werte als die Standardwerte verwenden.

Das Advisor-Intervall legt fest, wie oft eine Advisor-Funktion den Status der Server an dem von ihr überwachten Port abfragt und die Ergebnisse dann an den Manager übergibt. Ein zu niedriges Advisor-Intervall kann sich negativ auf die Leistung auswirken, da der Advisor die Server permanent unterbricht. Ist das Advisor-Intervall zu hoch, basieren die Entscheidungen des Managers hinsichtlich der Gewichtung unter Umständen nicht auf exakten aktuellen Informationen.

Wenn Sie das Intervall der HTTP-Advisor-Funktion am Port 80 beispielsweise auf 3 Sekunden setzen möchten, geben Sie den folgenden Befehl ein:

```
dscontrol advisor interval http 80 3
```

Es ist nicht sinnvoll, ein Advisor-Intervall anzugeben, das kleiner als das Manager-Intervall ist. Das Standard-Advisor-Intervall liegt bei sieben Sekunden.

Berichtszeitlimit für Advisor-Funktion

Der Manager verwendet keine Informationen einer Advisor-Funktion, deren Zeitmarke älter als die für das Berichtszeitlimit der Advisor-Funktion festgelegte Zeit ist, um sicherzustellen, dass keine veralteten Informationen verwendet werden. Das Berichtszeitlimit der Advisor-Funktion muss größer als das Sendeaufrufintervall der Advisor-Funktion sein. Ist das Zeitlimit kleiner, ignoriert der Manager Berichte, die logisch betrachtet verwendet werden müssten. Für Berichte der Advisor-Funktion gilt standardmäßig kein Zeitlimit, so dass der Standardwert unlimited ist.

Wenn Sie das Berichtszeitlimit für die HTTP-Advisor-Funktion am Port 80 beispielsweise auf 30 Sekunden setzen möchten, geben Sie den folgenden Befehl ein:

```
dscontrol advisor timeout http 80 30
```

Weitere Informationen zum Festlegen des Berichtszeitlimits für die Advisor-Funktion finden Sie im Abschnitt „dscontrol advisor — Advisor-Funktion steuern“ auf Seite 351.

Serververbindungs- und -empfangszeitlimit der Advisor-Funktion

Für Load Balancer können Sie die Zeitlimits der Advisor-Funktion festlegen, innerhalb derer der Ausfall eines bestimmten Server- oder Service-Ports festgestellt werden soll. Die Zeitlimits für ausgefallene Server (connecttimeout und receivetimeout) bestimmen, wie lange eine Advisor-Funktion wartet, bis sie einen gescheiterten Sende- oder Empfangsvorgang meldet.

Für eine schnellstmögliche Erkennung ausgefallener Server müssen Sie das Verbindungs- und Empfangszeitlimit der Advisor-Funktion auf den kleinsten Wert (eine Sekunde) sowie die das Intervall Advisor-Funktion und Manager auf den kleinsten Wert (eine Sekunde) setzen.

Anmerkung: Falls es in Ihrer Umgebung ein mittleres bis hohes Datenverkehrsaufkommen gibt, so dass sich die Serverantwortzeit erhöht, sollten Sie die Werte connecttimeout und receivetimeout nicht zu niedrig festlegen. Andernfalls könnte die Advisor-Funktion einen ausgelasteten Server vorschnell als ausgefallenen Server markieren.

Wenn Sie connecttimeout und receivetimeout für die HTTP-Advisor-Funktion am Port 80 beispielsweise auf 9 Sekunden setzen möchten, geben Sie den folgenden Befehl ein:

```
dscontrol advisor connecttimeout http 80 9  
dscontrol advisor receivetimeout http 80 9
```

Der Standardwert für das Verbindungs- und Empfangszeitlimit liegt beim Dreifachen des Wertes, der für das Intervall der Advisor-Funktion angegeben wurde.

Wiederholungsversuche der Advisor-Funktion

Advisor-Funktionen können wiederholt versuchen, eine Verbindung herzustellen, bevor sie einen Server als inaktiv markieren. Die Advisor-Funktion markiert einen Server erst als inaktiv, wenn die Abfrage nach der festgelegten Anzahl Wiederholungen und einem weiteren Versuch nicht beantwortet wird. Der Wert für **retry** sollte nicht über 3 liegen. Mit dem folgenden Befehl wird retry für den LDAP-Advisor am Port 389 auf 2 gesetzt:

```
dscontrol advisor retry ldap 389 2
```

Liste der Advisor-Funktionen

- Die Advisor-Funktion **http** öffnet eine Verbindung, sendet standardmäßig eine HEAD-Anfrage, wartet auf eine Antwortverbindung und gibt die verstrichene Zeit als Arbeitslast zurück. Im Abschnitt „Option 'Anforderung/ Antwort (URL)' der HTTP- oder HTTPS-Advisor-Funktion konfigurieren“ auf Seite 191 können Sie nachlesen, wie Sie die Art der Anfrage ändern können, die von der HTTP-Advisor-Funktion gesendet wird.
- Die Advisor-Funktion **https** ist ein Heavyweight-Advisor für SSL-Verbindungen. Er stellt eine reine SSL-Socket-Verbindung zum Server her. Die HTTPS-Advisor-Funktion öffnet eine SSL-Verbindung, sendet eine HTTPS-Anfrage, wartet auf eine Antwort, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück. (Vergleichen Sie dazu die SSL-Advisor-Funktion, die ein Lightweight-Advisor für SSL-Verbindungen ist.)

Anmerkung: Die HTTPS-Advisor-Funktion ist nicht vom Serverschlüssel oder vom Inhalt des Serverzertifikats abhängig. Beide dürfen jedoch nicht abgelaufen sein.

- Die Advisor-Funktion **sip** öffnet eine Verbindung, sendet eine OPTIONS-Anfrage, wartet auf eine Antwort, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück. Der unterstützte SIP-Advisor funktioniert nur mit TCP und erfordert eine auf einem Server installierte Anwendung, die auf eine OPTIONS-Anfrage antwortet.
- Die Advisor-Funktion **ftp** öffnet eine Verbindung, sendet eine SYST-Anfrage, wartet auf eine Antwort, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **ldap** öffnet eine Verbindung, sendet eine anonyme BIND-Anfrage, wartet auf eine Antwort, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **telnet** öffnet eine Verbindung, wartet auf die erste Nachricht vom Server, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **nntp** öffnet eine Verbindung, wartet auf die erste Nachricht vom Server, sendet einen quit-Befehl, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **imap** öffnet eine Verbindung, wartet auf die erste Nachricht vom Server, sendet einen quit-Befehl, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **pop3** öffnet eine Verbindung, wartet auf die erste Nachricht vom Server, sendet einen quit-Befehl, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **smtp** öffnet eine Verbindung, wartet auf die erste Nachricht vom Server, sendet einen quit-Befehl, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **ssl** ist ein Lightweight-Advisor für SSL-Verbindungen. Er stellt keine reine SSL-Socket-Verbindung zum Server her. Die SSL-Advisor-Funktion öffnet eine Verbindung, sendet eine Anfrage SSL CLIENT_HELLO, wartet auf eine Antwort, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück. (Vergleichen Sie dazu die HTTPS-Advisor-Funktion, die ein Heavyweight-Advisor für SSL-Verbindungen ist.)

Anmerkung: Die SSL-Advisor-Funktion ist nicht von der Schlüsselverwaltung oder von Zertifikaten abhängig.

- Die Advisor-Funktion **ssl2http** wird für die unter Port 443 aufgelisteten Server gestartet und ausgeführt, öffnet jedoch einen Socket zum "mapport" für HTTP-Anforderungen. Wenden Sie die Advisor-Funktion **ssl2http** nur für CBR an, wenn vom Client zum Proxy das Protokoll SSL und vom Proxy zum Server das Protokoll HTTP verwendet wird. Weitere Informationen hierzu finden Sie im Abschnitt „Lastausgleich für SSL-Datenverkehr vom Client zum Proxy und HTTP-Datenverkehr vom Proxy zum Server“ auf Seite 108.
- Die Caching-Proxy-Advisor-Funktion öffnet eine Verbindung, sendet eine für Caching Proxy spezifische HTTP-GET-Anfrage und interpretiert die Antwort als eine Caching-Proxy-Arbeitslast.

Anmerkung: Wenn Sie die Caching-Proxy-Advisor-Funktion verwenden möchten, muss Caching Proxy auf allen Servern mit Lastausgleich aktiv sein. Auf der Maschine mit Load Balancer muss Caching Proxy nur installiert werden, wenn es sich um die Maschine handelt, für die gleichzeitig der Lastausgleich durchgeführt wird.

- Die Advisor-Funktion **dns** öffnet eine Verbindung, sendet eine Zeigeranfrage für DNS, wartet auf eine Antwort, beendet die Verbindung und gibt die verstrichene Zeit als Arbeitslast zurück.
- Die Advisor-Funktion **connect** tauscht keine protokollspezifischen Daten mit dem Server aus. Er misst nur die Zeit, die benötigt wird, um eine TCP-Verbindung zu dem Server zu öffnen und zu schließen. Dieser Advisor ist für Serveranwendungen nützlich, die TCP verwenden, jedoch mit einem Protokoll höherer Ebene, für das keine von IBM gelieferte oder anpassbare Advisor-Funktion verfügbar ist.
- Die Advisor-Funktion **ping** öffnet keine TCP-Verbindung zu Servern und meldet stattdessen, ob der Server auf ein ping antwortet. Die Advisor-Funktion **ping** kann für jeden Port verwendet werden, ist jedoch speziell für Konfigurationen mit einem Platzhalter-Port konzipiert, über den Datenverkehr mit verschiedenen Protokollen fließen kann. Er ist außerdem für Konfigurationen mit Servern nützlich, die andere als die TCP-Protokolle verwenden, z. B. UDP.
- Die Advisor-Funktion **reach** sendet ein ping an die zugehörigen Zielmaschinen. Dieser Advisor wurde für die Komponente Dispatcher für hohe Verfügbarkeit entwickelt, um die Erreichbarkeit der Erreichbarkeitsziele zu bestimmen. Die Ergebnisse werden an die Komponente für hohe Verfügbarkeit übergeben und erscheinen *nicht* im Manager-Bericht. Im Gegensatz zu anderen Advisor-Funktionen wird **reach** automatisch von der Manager-Funktion der Komponente Dispatcher gestartet.
- Die Advisor-Funktion **db2** arbeitet mit den DB2-Servern zusammen. Der Dispatcher verfügt über die Fähigkeit, den Status von DB2-Servern zu überprüfen, ohne dass Kunden eigene angepasste Advisor-Funktionen schreiben müssen. Die DB2-Advisor-Funktion kommuniziert nur mit dem Port für DB2-Verbindungen, nicht mit dem Port für Java-Verbindungen.
- Die Advisor-Funktion **self** sammelt Informationen zum Auslastungsstatus von Back-End-Servern. Sie können die Advisor-Funktion **self** anwenden, wenn Sie den Dispatcher in einer Client/Server-Konfiguration verwenden, so dass der Dispatcher Informationen von der Advisor-Funktion **self** an den übergeordneten Load Balancer liefert. Die Advisor-Funktion **self** misst insbesondere die Verbindungen pro Sekunde für die Back-End-Server des Dispatchers auf Executor-Ebene. Weitere Informationen hierzu finden Sie im Abschnitt „Advisor-Funktion 'self' in einer Client/Server-WAN-Konfiguration“ auf Seite 192.

- Die Advisor-Funktion **wlm** (Workload Manager) ist für Server auf OS/390-Großrechnern bestimmt, die die Komponente MVS Workload Manager (WLM) ausführen. Weitere Informationen hierzu finden Sie im Abschnitt „Advisor-Funktion Workload Manager“ auf Seite 199.
- Der Dispatcher bietet einem Kunden die Möglichkeit, eine *angepasste* (anpassbare) Advisor-Funktion zu schreiben. Damit werden persönliche Protokolle unterstützt (zusätzlich zu TCP), für die IBM keinen spezifischen Advisor entwickelt hat. Weitere Informationen hierzu finden Sie im Abschnitt „Kundenspezifische (anpassbare) Advisor-Funktion erstellen“ auf Seite 193.
- Die **WAS**-Advisor-Funktion (WebSphere Application Server) arbeitet mit den WebSphere Application Servern zusammen. Anpassbare Beispieldateien für die Advisor-Funktion finden Sie im Installationsverzeichnis. Der Abschnitt „WAS-Advisor-Funktion“ auf Seite 194 enthält weitere Informationen zu diesem Thema.

Option 'Anforderung/Antwort (URL)' der HTTP- oder HTTPS-Advisor-Funktion konfigurieren

Die URL-Option der HTTP- oder HTTPS-Advisor-Funktion ist für die Komponenten Dispatcher und CBR verfügbar.

Nach dem Starten einer HTTP- oder HTTPS-Advisor-Funktion können Sie für den Dienst, den Sie vom Server anfordern möchten, eine eindeutige Client-HTTP-URL-Zeichenfolge definieren. Mit dieser Zeichenfolge kann die Advisor-Funktion den Status einzelner Dienste auf einem Server bewerten. Dies können Sie erreichen, indem Sie logische Server, die dieselbe physische IP-Adresse haben, mit eindeutigen Servernamen definieren. Weitere Informationen hierzu finden Sie im Abschnitt „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58.

Für jeden unter dem HTTP-Port definierten logischen Server können Sie eine für den Dienst, den Sie vom Server anfordern möchten, eine eindeutige Client-HTTP-URL-Zeichenfolge angeben. Die HTTP- oder HTTPS-Advisor-Funktion verwendet die Zeichenfolge **advisorrequest**, um den Status der Server abzufragen. Der Standardwert ist HEAD / HTTP/1.0. Die Zeichenfolge **advisorresponse** ist die Antwort, nach der die Advisor-Funktion die HTTP-Antwort durchsucht. Die Advisor-Funktion vergleicht die Zeichenfolge **advisorresponse** mit der tatsächlich vom Server empfangenen Antwort. Der Standardwert ist null.

Wichtiger Hinweis: Wenn die HTTP-URL-Zeichenfolge ein Leerzeichen enthält, gilt Folgendes:

- Bei Absetzen des Befehls von der Shell-Eingabeaufforderung **dscontrol>>** müssen Sie die Zeichenfolge in Anführungszeichen setzen. Beispiel:

```
server set Cluster:Port:Server advisorrequest "head / http/1.0"
server set Cluster:Port:Server advisorresponse "HTTP 200 OK"
```
- Beim Absetzen des Befehls **dscontrol** an der Eingabeaufforderung des Betriebssystems müssen Sie dem Text die Zeichen **"\"** voranstellen und den Text mit den Zeichen **\""** beenden. Beispiel:

```
dscontrol server set Cluster:Port:Server
advisorrequest "\"head / http/1.0\""

dscontrol server set Cluster:Port:Server advisorresponse "\"HTTP 200 OK\""
```

Wenn Sie die Anforderung erstellen, die die HTTP- oder HTTPS-Advisor-Funktion an Back-End-Server sendet, um deren Funktionstüchtigkeit zu überprüfen, geben Sie nur den Anfang der HTTP-Anforderung ein. Load Balancer vervollständigt die Anforderung mit folgenden Angaben:

```
\r\nAccept:
*/*\r\nUser-Agent:IBM_Network_Dispatcher_HTTP_Advisor\r\n\r\n
```

Falls Sie weitere HTTP-Header-Felder hinzufügen möchten, bevor Load Balancer diese Zeichenfolge an das Ende der Anforderung anfügt, können Sie eine eigene \r\n-Zeichenfolge in die Anforderung aufnehmen. Nachfolgend sehen Sie eine Beispieleingabe für das Hinzufügen des HTTP-Header-Feldes "Host" zur Anforderung:

```
GET /pub/WWW/TheProject.html HTTP/1.0 \r\nHost: www.w3.org
```

Anmerkung: Nach dem Start einer HTTP- oder HTTPS-Advisor-Funktion für eine angegebene HTTP-Port-Nummer, wird für Server an diesem HTTP-Port der Abfrage-/Antwortwert der Advisor-Funktion aktiviert.

Weitere Informationen hierzu finden Sie im Abschnitt „dscontrol server — Server konfigurieren“ auf Seite 394.

Advisor-Funktion 'self' in einer Client/Server-WAN-Konfiguration

Die Advisor-Funktion self ist für die Komponente Dispatcher verfügbar.

Wenn Load Balancer in einer Client/Server-WAN-Konfiguration (Weitverkehrsnetz) installiert ist, stellt der Dispatcher eine self-Advisor-Funktion bereit, die Informationen zum Auslastungsstatus von Back-End-Servern sammelt.

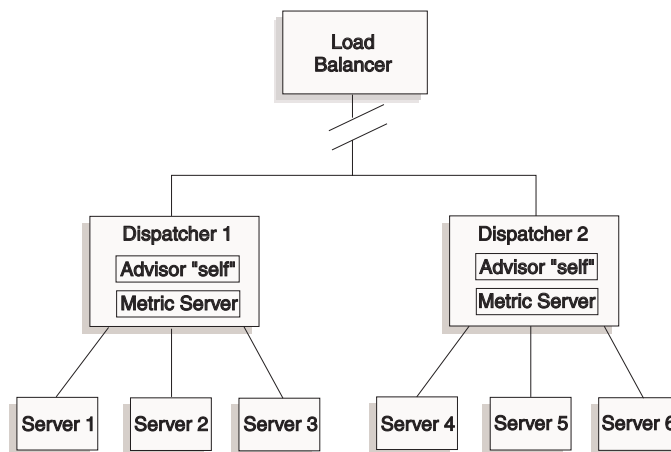


Abbildung 34. Beispiel für eine Client/Server-WAN-Konfiguration mit Advisor-Funktion self

In diesem Beispiel befinden sich die Advisor-Funktion self und Metric Server auf den beiden Dispatcher-Maschinen, deren Lastausgleich vom übergeordneten Load Balancer durchgeführt wird. Die Advisor-Funktion self misst insbesondere die Verbindungen pro Sekunde für die Back-End-Server des Dispatchers auf Executor-Ebene.

Der Selbst-Advisor schreibt die Ergebnisse in die Datei dsloadstat. Load Balancer stellt außerdem den externen Messwert dsload bereit. Bei Ausführung der Konfiguration des Agenten Metric Server auf den Dispatcher-Maschinen wird der externe

Messwert dsload aufgerufen. Das dsload-Script extrahiert eine Zeichenfolge aus der Datei dsloadstat und gibt sie an den Agenten Metric Server zurück. Anschließend geben die Metric-Server-Agenten (von den einzelnen Dispatcher-Maschinen) den Wert für den Auslastungsstatus an den übergeordneten Load Balancer zurück, damit dieser bestimmen kann, welcher Dispatcher Client-Anforderungen weiterleiten soll.

Die ausführbare Datei dsload befindet sich im Load-Balancer-Verzeichnis **...ibm/edge/lb/ms/script**.

Weitere Informationen zur Verwendung von Dispatcher in WAN-Konfigurationen finden Sie im Abschnitt „Dispatcher-WAN-Unterstützung konfigurieren“ auf Seite 229. Weitere Informationen zu Metric Server können Sie dem Abschnitt „Metric Server“ auf Seite 196 entnehmen.

Kundenspezifische (anpassbare) Advisor-Funktion erstellen

Die kundenspezifische (anpassbare) Advisor-Funktion ist ein kurzer Java-Code, den Sie als Klassendatei bereitstellen, die vom Basiscode aufgerufen wird. Der Basiscode gewährleistet alle Verwaltungsdienste wie das Starten und Stoppen einer Instanz der angepassten Advisor-Funktion, das Bereitstellen von Status und Berichten sowie das Aufzeichnen von Protokolldaten in einer Protokolldatei. Er übergibt auch die Ergebnisse an die Manager-Funktion. Der Basiscode führt regelmäßig einen Advisor-Zyklus aus, wobei alle Server in der Konfiguration individuell ausgewertet werden. Dieser beginnt mit dem Öffnen einer Verbindung zu einer Servermaschine. Wenn das Socket geöffnet wird, ruft der Basiscode die Methode (Funktion) „getLoad“ der angepassten Advisor-Funktion auf. Die angepasste Advisor-Funktion führt dann alle für die Auswertung des Serverstatus erforderlichen Schritte aus. Normalerweise sendet er eine benutzerdefinierte Nachricht an den Server und wartet dann auf eine Antwort. (Die angepasste Advisor-Funktion erhält Zugriff auf den geöffneten Socket.) Der Basiscode schließt dann den Socket zu dem Server und übergibt die Lastinformationen an den Manager.

Der Basiscode und die angepasste Advisor-Funktion können im normalen Modus oder im Ersetzungsmodus arbeiten. Die Auswahl der Betriebsart wird in der Datei der angepassten Advisor-Funktion als Parameter der Methode constructor angegeben.

Im normalen Modus tauscht die angepasste Advisor-Funktion Daten mit dem Server aus. Der Basiscode der Advisor-Funktion misst die Zeit für den Austausch und berechnet den Lastwert. Der Basiscode übergibt dann diesen Lastwert an den Manager. Die angepasste Advisor-Funktion muss nur null (bei Erfolg) oder -1 (bei einem Fehler) zurückgeben. Zur Angabe des normalen Modus wird die Markierung replace der Methode constructor auf false (falsch) gesetzt.

Im Ersetzungsmodus führt der Basiscode keine Zeitmessungen aus. Der Code der angepassten Advisor-Funktion führt alle für die funktionspezifischen Anforderungen erforderlichen Operationen aus und gibt dann einen tatsächlichen Lastwert zurück. Der Basiscode akzeptiert diesen Wert und übergibt ihn an den Manager. Um bestmögliche Ergebnisse zu erzielen, sollten Sie den Lastwert zwischen 10 und 1000 normalisieren, wobei 10 einen schnellen Server und 1000 einen langsamen Server angibt. Zur Angabe des Ersetzungsmodus muss die Markierung replace der Methode constructor auf true gesetzt werden.

Auf diese Weise können Sie eigene Advisor-Funktionen schreiben, die die benötigten präzisen Informationen über Server zur Verfügung stellen. Zu Load Balancer

wird ein Beispiel für eine angepasste Advisor-Funktion, **ADV_sample.java**, geliefert. Nach der Installation von Load Balancer finden Sie den Beispielcode im Installationsverzeichnis

...<Installationsverzeichnis>/servers/samples/CustomAdvisors.

Das *Installationsverzeichnis* ist standardmäßig folgendes Verzeichnis:

- AIX-, HP-UX-, Linux-, Solaris-Systeme: /opt/ibm/edge/lb
- Windows-Systeme: C:\Programme\IBM\edge\lb

Anmerkung: Wenn Sie eine angepasste Advisor-Funktion zum Dispatcher oder einer anderen Komponente von Load Balancer hinzufügen, müssen Sie **dsserver** (bzw. auf Windows-Systemen den Dienst) stoppen und dann erneut starten, um den Java-Prozess zu veranlassen, die neuen angepassten Advisor-Klassendateien zu lesen. Die angepassten Advisor-Klassendateien werden nur beim Systemstart geladen. Es ist nicht nötig, den Executor zu stoppen. Der Executor wird weiter ausgeführt, auch wenn dsserver oder der entsprechende Dienst gestoppt wurde.

Wenn die angepasste Advisor-Funktion auf zusätzliche Java-Klassen verweist, muss der Klassenpfad (classpath) in der Start-Script-Datei von Load Balancer (dsserver, cbrserver, ssserver) mit dem Verzeichnis dieser Klassen aktualisiert werden.

WAS-Advisor-Funktion

Das Installationsverzeichnis von Load Balancer enthält Beispieldateien für angepasste Advisor-Funktionen, insbesondere für die WAS-Advisor-Funktion (WebSphere Application Server).

- ADV_was.java ist die Datei, die kompiliert und auf der Load-Balancer-Maschine ausgeführt werden muss.
- Die Datei LBAdvisor.java.servlet muss kompiliert und auf der WAS-Maschine ausgeführt werden (nachdem sie in LBAdvisor.java umbenannt wurde).

Die Beispieldateien für die WAS-Advisor-Funktion befinden sich in demselben Verzeichnis wie die Datei ADV_sample.java.

Namenskonvention

Der Dateiname für Ihre angepasste Advisor-Funktion muss das Format „ADV_*meinAdvisor*.java“ haben. Er muss mit dem Präfix „ADV_“ in Großbuchstaben beginnen. Alle nachfolgenden Zeichen müssen Kleinbuchstaben sein.

Aufgrund von Java-Konventionen muss der Name der in der Datei definierten Klasse mit dem Namen der Datei übereinstimmen. Wenn Sie den Beispielcode kopieren, stellen Sie sicher, dass alle Exemplare von „ADV_sample“ in der Datei in den neuen Klassennamen geändert werden.

Kompilierung

Angepasste Advisor-Funktionen werden in der Sprache Java geschrieben. Verwenden Sie den Java-Compiler, der zusammen mit Load Balancer installiert wird. Während der Kompilierung wird auf die folgenden Dateien Bezug genommen:

- die Datei der angepassten Advisor-Funktion
- die Basisklassendatei ibmlb.jar im Installationsverzeichnis ...ibm/edge/lb/servers/lib.

Der Klassenpfad muss während der Kompilierung auf die Datei der angepassten Advisor-Funktion und die Datei mit den Basisklassen zeigen.

Beispielkompilierungsbefehl für Windows-Systeme:

```
Installationsverzeichnis/java/bin/javac -classpath  
Installationsverzeichnis\lb\servers\lib\ibmlb.jar ADV_fred.java
```

Für diesen Befehl gilt Folgendes:

- Ihre Advisor-Datei hat den Namen ADV_fred.java.
- Ihre Advisor-Datei ist im aktuellen Verzeichnis gespeichert.

Die Ausgabe der Kompilierung ist eine Klassendatei, zum Beispiel
ADV_fred.class

Kopieren Sie vor dem Starten der Advisor-Funktion die Klassendatei in das Installationsverzeichnis ...**ibm/edge/lb/servers/lib/CustomAdvisors**.

Anmerkung: Bei Bedarf können angepasste Advisor-Funktionen unter einem Betriebssystem kompiliert und unter einem anderen Betriebssystem ausgeführt werden. Sie können beispielsweise Ihre Advisor-Funktion auf Windows-Systemen kompilieren, die Klassendatei (im Binärformat) auf eine AIX-Maschine kopieren und die Advisor-Funktion dort ausführen.

Auf AIX-, HP-UX-, Linux- und Solaris-Systemen ist die Syntax ähnlich.

Ausführung

Bevor Sie die angepasste Advisor-Funktion ausführen, müssen Sie die Klassendatei in das richtige Unterverzeichnis des Installationsverzeichnisses kopieren:

```
...ibm/edge/lb/servers/lib/CustomAdvisors/ADV_fred.class
```

Konfigurieren Sie die Komponente, starten Sie die zugehörige Manager-Funktion und setzen Sie wie folgt den Befehl zum Starten der angepassten Advisor-Funktion ab:

```
dscontrol advisor start fred 123
```

Für diesen Befehl gilt Folgendes:

- Der Name Ihrer Advisor-Funktion ist fred, wie in ADV_fred.java.
- Der Port, an dem Ihre Advisor-Funktion ausgeführt wird, ist 123.

Wenn die angepasste Advisor-Funktion auf zusätzliche Java-Klassen verweist, muss der Klassenpfad (classpath) in der Start-Script-Datei von Load Balancer (dsserver, cbrserver, ssserver) mit dem Verzeichnis dieser Klassen aktualisiert werden.

Erforderliche Routinen

Eine angepasste Advisor-Funktion erweitert wie alle anderen Advisor-Funktionen den Advisor-Basiscode ADV_Base. Es ist der Advisor-Basiscode, der die meisten Funktionen ausführt. Dazu gehört das Zurückmelden von Belastungen an den Manager, die für den Wertigkeitsalgorithmus des Managers verwendet werden. Darüber hinaus stellt der Advisor-Basiscode Socket-Verbindungen her, schließt Sockets und stellt Sende- und Empfangsmethoden für die Advisor-Funktion bereit. Die Advisor-Funktion selbst wird nur zum Senden von Daten an den Port bzw. Empfangen von Daten vom Port des empfohlenen Servers verwendet. Die TCP-Methoden innerhalb des Advisor-Basiscodes sind zeitlich gesteuert, um die Last zu

berechnen. Mit einer Markierung der Methode constructor in ADV_base kann bei Bedarf die vorhandene Last mit der neuen, von der Advisor-Funktion zurückgegebenen Last überschrieben werden.

Anmerkung: Der Advisor-Basiscode stellt in angegebenen Intervallen die Last ausgehend von einem in der Methode constructor gesetzten Wert für den Wertigkeitsalgorithmus bereit. Wenn die eigentliche Advisor-Funktion noch keine gültige Last zurückgeben kann, verwendet der Advisor-Basiscode die vorherige Last.

Basisklassenmethoden sind:

- Eine Routine **constructor**. Die constructor-Routine ruft die constructor-Methode base class auf (siehe Advisor-Beispieldatei).
- Eine Methode **ADV_AdvisorInitialize**. Diese Methode stellt einen Hook für den Fall zur Verfügung, dass zusätzliche Schritte ausgeführt werden müssen, nachdem die Basisklasse ihre Initialisierung beendet hat.
- Eine Routine **getload**. Die Basis-Advisor-Klasse führt das Öffnen des Sockets aus. Daher muss getload nur die entsprechenden Sende- und Empfangsanforderungen absetzen, um den Advisor-Zyklus zu beenden.

Suchreihenfolge

Load Balancer durchsucht zunächst die Liste der eigenen Advisor-Funktionen. Wenn eine bestimmte Advisor-Funktion dort nicht aufgeführt ist, durchsucht Load Balancer die Kundenliste der angepassten Advisor-Funktionen.

Benennung und Pfad

- Die Klasse der angepassten Advisor-Funktion muss sich im Unterverzeichnis **...ibm/edge/lb/servers/lib/CustomAdvisors/** des Basisverzeichnis von Load Balancer befinden. Die Standardwerte für dieses Verzeichnis hängen vom verwendeten Betriebssystem ab:
 - AIX-, HP-UX-, Linux- und Solaris-Systeme:
/opt/ibm/edge/lb/servers/lib/CustomAdvisors/
 - Windows-Systeme:
C:\Programme\IBM\edge\lb\servers\lib\CustomAdvisors
- Es sind nur alphabetische Zeichen in Kleinschreibung zulässig. Ein Bediener muss somit bei der Eingabe von Befehlen in der Befehlszeile nicht auf die Groß-/Kleinschreibung achten. Der Name der Advisor-Datei muss das Präfix **ADV_** haben.

Beispiel-Advisor-Funktion

Die Programmliste für eine Beispiel-Advisor-Funktion finden Sie im Abschnitt „Beispiel-Advisor-Funktion“ auf Seite 487. Nach der Installation befindet sich diese Beispiel-Advisor-Funktion im Verzeichnis **...ibm/edge/lb/servers/samples/CustomAdvisors**.

Metric Server

Diese Funktion ist für alle Komponenten von Load Balancer verfügbar.

Metric Server gibt Load Balancer Informationen zur Serverauslastung. Diese Informationen werden in Form systemspezifischer Messwerte für den Serverzustand bereitgestellt. Der Manager von Load Balancer richtet Anfragen an den Agenten

Metric Server, der sich auf jedem der Server befindet, und legt anhand der Messwerte, die er von den Agenten erhalten hat, Wertigkeiten für den Lastausgleich fest. Die Ergebnisse werden auch in den Manager-Bericht gestellt.

Anmerkung: Wenn für jeden Server zwei oder mehr Messwerte ermittelt und in einen Systemauslastungswert normalisiert werden, kann es zu Rundungsfehlern kommen.

Informationen zur Verwendung (zum Starten und Stoppen) von Metric Server und zur Verwendung von Metric-Server-Protokollen finden Sie im Abschnitt „Metric Server verwenden“ auf Seite 282.

Ein Konfigurationsbeispiel ist in Abb. 5 auf Seite 14 dargestellt.

WLM-Einschränkung

Wie die Advisor-Funktion WLM gibt Metric Server Berichte zu kompletten Serversystemen aus und nicht zu einzelnen protokollspezifischen Serverdämonen. WLM und Metric Server stellen ihre Ergebnisse in die Spalte "System" des Manager-Berichts. Deshalb wird die gleichzeitige Ausführung der Advisor-Funktionen WLM und Metric Server nicht unterstützt.

Vorbedingungen

Der Agent Metric Server muss auf allen Servern installiert und ausgeführt werden, die am Lastausgleich teilnehmen.

Metric Server verwenden

Nachfolgend sind die Schritte aufgeführt, mit denen Sie Metric Server für den Dispatcher konfigurieren. Wenn Sie Metric Server für andere Komponenten von Load Balancer konfigurieren möchten, sind ähnliche Schritte auszuführen.

- Manager von Load Balancer (Load-Balancer-Seite)
 1. Starten Sie **dsserver**.
 2. Setzen Sie den Befehl **dscontrol manager start *manager.log Port*** ab.
Port ist hier der ausgewählte RMI-Port für alle Metric-Server-Agenten. Der in der Datei `metricserver.cmd` festgelegte Standard-RMI-Port ist 10004.
 3. Setzen Sie den Befehl **dscontrol metric add *Cluster:Systemmesswert*** ab.
Systemmesswert ist hier der Name des Scripts (auf dem Back-End-Server), das für jeden Server, der unter dem angegebenen Cluster (oder Sitenamen) in der Konfiguration enthalten ist, ausgeführt werden soll. Für den Kunden stehen die beiden Scripts **cpupload** und **memload** bereit. Sie können auch angepasste Scripts für Systemmesswerte erstellen. Das Script enthält einen Befehl, der einen numerischen Wert im Bereich von 0-100 oder, bei inaktivem Server, den Wert -1 zurückgeben sollte. Dieser numerische Wert sollte eine Lastmessung und keinen Verfügbarkeitswert darstellen.

Anmerkung: Für Site Selector werden `cpupload` und `memload` automatisch ausgeführt.

Einschränkung: Wenn der Name Ihres Scripts für Systemmesswerte auf der Windows-Plattform eine andere Erweiterung als `.exe` hat, müssen Sie den vollständigen Namen der Datei (z. B. "meinsystemscript.bat") angeben. Dies ergibt sich aus einer Java-Einschränkung.

4. Fügen Sie zur Konfiguration nur Server hinzu, die einen Metric-Server-Agenten enthalten, der für den in der Datei `metricserver.cmd` angegebenen Port ausgeführt wird. Der Port sollte mit dem im Befehl **manager start** angegebenen Port-Wert übereinstimmen.

Anmerkung: Gewährleisten Sie wie folgt die Sicherheit:

- Erstellen Sie auf der Load-Balancer-Maschine einen Schlüsselring. (Verwenden Sie dazu den Befehl **lbkeys create**.) Weitere Informationen zu `lbkeys` finden Sie im Abschnitt „Remote Method Invocation (RMI)“ auf Seite 264.
 - Kopieren Sie den resultierenden Schlüsselring für die Komponente, die ausgeführt wird, auf der Back-End-Servermaschine in das Verzeichnis `...ibm/edge/lb/admin/keys`. Stellen Sie sicher, dass die Berechtigungen für den Schlüsselring dem Benutzer `root` den Lesezugriff ermöglichen.
- Agent Metric Server (Seite der Servermaschine)
 1. Installieren Sie das Metric-Server-Paket aus dem Installationsverzeichnis von Load Balancer.
 2. Überprüfen Sie anhand des Scripts **metricserver** im Verzeichnis `/usr/bin`, ob der gewünschte RMI-Port verwendet wird. (Für Windows 2003 lautet das Verzeichnis `C:\WINDOWS\system32`.) Der Standard-RMI-Port ist 10004.

Anmerkung: Der für den RMI-Port angegebene Wert muss mit dem RMI-Port-Wert für Metric Server auf der Load-Balancer-Maschine übereinstimmen.

3. Die beiden folgenden Scripts werden dem Kunden bereits zur Verfügung gestellt: **cpuload** (gibt den Prozentsatz der verwendeten CPU im Bereich von 0-100 zurück) und **memload** (gibt den Prozentsatz des belegten Speichers im Bereich von 0-100 zurück). Diese Scripts befinden sich im Verzeichnis `...ibm/edge/lb/ms/script`.

Optional können Kunden ihre eigenen angepassten Script-Dateien für Messwerte schreiben, in denen definiert ist, welchen Befehl Metric Server auf den Servermaschinen absetzen soll. Vergewissern Sie sich, dass alle angepassten Scripts ausführbar sind und sich im Verzeichnis `...ibm/edge/lb/ms/script` befinden. Angepasste Scripts **müssen** einen numerischen Lastwert im Bereich von 0 bis 100 zurückgeben.

Anmerkung: Ein angepasstes Script für Messwerte muss ein gültiges Programm oder Script mit der Erweiterung `.bat` oder `.cmd` sein. Auf Linux- und UNIX-Systemen müssen Scripts mit der Shell-Deklaration beginnen, da sie sonst möglicherweise nicht richtig ausgeführt werden.

4. Starten Sie den Agenten durch Absetzen des Befehls **metricserver**.
5. Zum Stoppen des Agenten Metric Server müssen Sie den Befehl **metricserver stop** absetzen.

Wenn Metric Server für eine vom lokalen Host abweichende Adresse ausgeführt werden soll, müssen Sie die Datei `metricserver` auf der am Lastausgleich beteiligten Servermaschine editieren. Fügen Sie in der Datei `metricserver` nach dem Eintrag `"java"` Folgendes ein:

```
-Djava.rmi.server.hostname=andere_Adresse
```

Fügen Sie außerdem vor den Anweisungen `if` die folgende Zeile zur Datei `metricserver` hinzu: `hostname andere_Adresse`.

Auf der Windows-Plattform müssen Sie außerdem in Microsoft Stack auf der Metric-Server-Maschine den Aliasnamen für *andere_Adresse* angeben. Beispiel:

```
call netsh interface ip add address "Local Area Connection"  
    addr=9.37.51.28 mask=255.255.240.0
```

Wenn Sie Messwerte domänenübergreifend erfassen, müssen Sie `java.rmi.server.hostname` im Server-Script (`dsserver`, `cbrserver` usw.) auf den vollständig qualifizierten Domännennamen (FQDN) der Maschine setzen, die die Messwerte anfordert. Dies ist notwendig, weil `InetAddress.getLocalHost.getHostName()` nicht für jede Konfiguration und unter jedem Betriebssystem den FQDN zurückgibt.

Advisor-Funktion Workload Manager

WLM ist Code, der auf MVS-Großrechnern ausgeführt wird. Er kann abgefragt werden, um die Belastung auf der MVS-Maschine zu bestimmen.

Wurde MVS Workload Management auf Ihrem OS/390-System konfiguriert, kann der Dispatcher Kapazitätsinformationen von WLM akzeptieren und die Informationen für den Lastausgleich verwenden. Mit der Advisor-Funktion WLM öffnet der Dispatcher regelmäßig Verbindungen über den WLM-Port der einzelnen Server in der Dispatcher-Hosttabelle und akzeptiert die zurückgegebenen ganzzahligen Kapazitätswerte. Da diese ganzen Zahlen die noch verfügbare Kapazität darstellen und Dispatcher Werte erwarten, die die Belastung auf jeder Maschine angeben, werden die ganzzahligen Kapazitätswerte vom Advisor in Lastwerte umgekehrt und normalisiert. (Ein hoher ganzzahliger Kapazitätswert und ein niedriger Lastwert geben beide einen akzeptablen Zustand eines Servers an). Die daraus resultierenden Belastungen werden in die Spalte 'System' des Manager-Berichts gestellt.

Es gibt mehrere wichtige Unterschiede zwischen dem WLM-Advisor und anderen Advisor-Funktionen des Dispatchers:

1. Andere Advisor-Funktionen öffnen Verbindungen zu den Servern unter Verwendung des Ports, über den der normale Client-Datenverkehr fließt. Die WLM-Advisor-Funktion benutzt für das Öffnen von Verbindungen zu den Servern nicht den für normalen Datenverkehr verwendeten Port. Der WLM-Agent muss auf den einzelnen Servermaschinen so konfiguriert werden, dass er an dem Port empfangsbereit ist, an dem die WLM-Advisor-Funktion des Dispatchers gestartet wurde. Der Standard-WLM-Port ist 10007.
2. Andere Advisor-Funktionen bewerten nur die in der Konfiguration Cluster:Port:Server des Dispatchers definierten Server, deren Server-Port mit dem Port der Advisor-Funktion übereinstimmt. Die WLM-Advisor-Funktion wird (unabhängig von der Angabe Cluster:Port) für *alle* Server in der Dispatcher-Konfiguration ausgeführt. Daher dürfen Sie bei Verwendung der WLM-Advisor-Funktion nur WLM-Server definieren.
3. Andere Advisor-Funktionen stellen ihre Lastinformationen in die Spalte „Port“ des Manager-Berichts. Die Advisor-Funktion WLM stellt ihre Lastinformationen in die Spalte 'System' des Manager-Berichts.
4. Es ist möglich, protokollspezifische Advisor-Funktionen zusammen mit der Advisor-Funktion WLM zu verwenden. Die protokollspezifischen Advisor-Funktionen fragen die Server an den regulären Ports für Datenverkehr ab. Die WLM-Advisor-Funktion fragt die Systembelastung dagegen am WLM-Port ab.

Einschränkung für Metric Server

Der WLM-Agent gibt wie der Agent Metric Server Berichte zu kompletten Serversystemen aus und nicht zu einzelnen protokollspezifischen Serverdämonen. Metric Server und WLM stellen ihre Ergebnisse in die Spalte "System" des Manager-Berichts. Deshalb wird die gleichzeitige Ausführung der Advisor-Funktionen WLM und Metric Server nicht unterstützt.

Kapitel 22. Erweiterte Funktionen für Dispatcher, CBR und Site Selector

In diesem Kapitel wird erklärt, wie die Lastausgleichparameter konfiguriert werden und Load Balancer für die Verwendung der erweiterten Funktionen eingerichtet wird.

Anmerkung: Falls Sie die Komponente Dispatcher *nicht* verwenden, ersetzen Sie beim Lesen dieses Kapitels dscontrol durch Folgendes:

- Für CBR: **cbrcontrol**
- Für Site Selector: **sscontrol** (lesen Sie hierzu die Informationen in Kapitel 28, „Befehlsreferenz für Site Selector“, auf Seite 405)

Wichtiger Hinweis: Falls Sie Load Balancer für IPV4 und IPV6 verwenden, sollten Sie sich vor der Beschäftigung mit dem Inhalt dieses Kapitels in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81 über Einschränkungen und Konfigurationsunterschiede informieren.

Tabelle 13. Erweiterte Konfigurations-Tasks für Load Balancer

Task	Beschreibung	Referenzinformationen
Verknüpfung von Load Balancer mit einer am Lastausgleich beteiligten Maschine	Verknüpfte Load-Balancer-Maschine konfigurieren.	„Verknüpfte Server verwenden“ auf Seite 202
Konfigurieren der hohen Verfügbarkeit oder der gegenseitigen hohen Verfügbarkeit	Konfigurieren Sie eine zweite Dispatcher-Maschine, um eine Ausweichmaschine zu haben.	„Hohe Verfügbarkeit“ auf Seite 204
Konfigurieren des regelbasierten Lastausgleichs	Definieren Sie Bedingungen, unter denen eine Untergruppe Ihrer Server verwendet wird.	„Regelbasierten Lastausgleich konfigurieren“ auf Seite 212
Außerkräftsetzung der Port-Affinität für Server	Ermöglicht einem Server, die Einstellung für stickytime an seinem Port zu überschreiben.	„Port-Affinität außer Kraft setzen“ auf Seite 220
Verwendung der Affinitätsfunktion zum Konfigurieren einer Haltezeit für einen Cluster-Port	Client-Anforderungen werden immer an denselben Server gerichtet.	„Funktionsweise der Affinität für Load Balancer“ auf Seite 222
Verwendung der Port-übergreifenden Affinität, um die Affinität an allen Ports zu nutzen	Von verschiedenen Ports empfangene Client-Anforderungen werden an einen Server gerichtet.	„Port-übergreifende Affinität“ auf Seite 223
Verwendung der Affinitätsadressmaske zum Festlegen einer gemeinsamen IP-Teilnetzadresse	Aus einem Teilnetz empfangene Client-Anforderungen werden an denselben Server gerichtet.	„Affinitätsadressmaske (stickymask)“ auf Seite 223
Verwendung der aktiven Cookie-Affinität für den Serverlastausgleich mit CBR	Diese Regeloption ermöglicht die Bindung einer Sitzung an einen bestimmten Server.	„Aktive Cookie-Affinität“ auf Seite 225
Verwendung der passiven Cookie-Affinität für den Serverlastausgleich mit der inhaltsabhängige Weiterleitung des Dispatchers und mit CBR	Mit dieser Regeloption kann eine Sitzung ausgehend vom Cookie-Namen/Wert an einen bestimmten Server gebunden werden.	„Passive Cookie-Affinität“ auf Seite 227

Tabelle 13. Erweiterte Konfigurations-Tasks für Load Balancer (Forts.)

Task	Beschreibung	Referenzinformationen
Verwendung der URI-Affinität für den Lastausgleich bei Caching-Proxy-Servern mit Zwischenspeicherung spezifischer Inhalte auf jedem einzelnen Server	Mit dieser Regeloption kann eine Sitzung ausgehend vom URI an einen bestimmten Server gebunden werden.	„URI-Affinität“ auf Seite 228
Konfigurieren der Weitverkehrsunterstützung von Dispatcher	Konfigurieren Sie einen fernen Dispatcher für den Lastausgleich in einem WAN. Der Lastausgleich in einem WAN kann auch (ohne fernen Dispatcher) mit einer Serverplattform durchgeführt werden, die GRE unterstützt.	„Dispatcher-WAN-Unterstützung konfigurieren“ auf Seite 229
Verwendung expliziter Verbindungen	Vermeiden Sie es, den Dispatcher in Verbindungen zu umgehen.	„Explizite Verbindungen benutzen“ auf Seite 236
Verwendung eines privaten Netzes	Konfigurieren Sie den Dispatcher für den Lastausgleich bei Servern in einem privaten Netz.	„Konfiguration für ein privates Netz verwenden“ auf Seite 236
Zusammenfassung allgemeiner Serverkonfigurationen durch einen Platzhaltercluster	Adressen, die nicht explizit konfiguriert sind, verwenden den Platzhaltercluster für die Verteilung des Datenverkehrs.	„Platzhaltercluster zum Zusammenfassen von Serverkonfigurationen verwenden“ auf Seite 237
Verwendung eines Platzhalterclusters für den Lastausgleich bei Firewalls	Der gesamte Datenverkehr für Firewalls wird verteilt.	„Platzhaltercluster für den Lastausgleich von Firewalls verwenden“ auf Seite 238
Verwendung eines Platzhalterclusters mit Caching Proxy für transparente Weiterleitung	Der Dispatcher kann zum Aktivieren einer transparenten Weiterleitung verwendet werden.	„Platzhaltercluster mit Caching Proxy für transparente Weiterleitung verwenden“ auf Seite 239
Verwendung eines Platzhalterports für die Übertragung von Datenverkehr mit nicht konfiguriertem Port	Ermöglicht die Bearbeitung von Datenverkehr, der für keinen bestimmten Port konfiguriert ist.	„Platzhalter-Port für die Übertragung von Datenverkehr mit nicht konfiguriertem Port verwenden“ auf Seite 239
Verwendung der DoS-Erkennung für die Benachrichtigung von Administratoren über potenzielle Attacks (per Alert)	Der Dispatcher analysiert eingehende Anforderungen auf eine verdächtige Anzahl halboffener TCP-Verbindungen auf Servern.	„Erkennung von DoS-Attacken“ auf Seite 240
Binäre Protokollierung zur Analyse der Serverstatistik	Ermöglicht das Speichern von Serverinformationen in Binärdateien und das Abrufen dieser Informationen aus Binärdateien.	„Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 241
Verwendung einer verknüpften Client-Konfiguration	Load Balancer kann sich auf derselben Maschine wie ein Client befinden.	„Verknüpften Client verwenden“ auf Seite 243

Verknüpfte Server verwenden

Load Balancer kann sich auf derselben Maschine befinden wie ein Server, für dessen Anforderungen er einen Lastausgleich durchführt. Dies wird als das *Verknüpfen* eines Servers bezeichnet. Die Verknüpfung gilt für die Komponenten Dispatcher und Site Selector. Für CBR wird die Verknüpfung auch unterstützt. Dies gilt jedoch nur bei Verwendung bindungsspezifischer Webserver und Caching-Proxy-Server.

Anmerkung: In Zeiten hohen Datenverkehrs konkurriert ein verknüpfter Server mit Load Balancer um Ressourcen. Sind jedoch keine überlasteten Maschinen vorhanden, kann mit einem verknüpften Server die Gesamtzahl der Maschinen reduziert werden, die für das Einrichten eines Standortes mit Lastausgleich erforderlich sind.

Für Dispatcher

Linux-Systeme: Wenn Sie bei Ausführung der Komponente Dispatcher mit der Weiterleitungsmethode `mac` sowohl die Verknüpfung als auch die hohe Verfügbarkeit konfigurieren möchten, lesen Sie die Informationen im Abschnitt „Alternativen für die Festlegung eines Loopback-Aliasnamens unter Linux bei Verwendung der Load-Balancer-Weiterleitungsmethode `mac`“ auf Seite 78.

Windows-Systeme: Wenn Sie bei Ausführung der Komponente Dispatcher mit der Weiterleitungsmethode `mac` sowohl die Verknüpfung als auch die hohe Verfügbarkeit konfigurieren möchten, lesen Sie die Informationen im Abschnitt „Verknüpfung und hohe Verfügbarkeit konfigurieren (Windows-Systeme)“ auf Seite 212.

Solaris-Systeme: Sie können keine WAN-Advisor-Funktionen konfigurieren, wenn der Eingangspunkt-Dispatcher verknüpft ist. Lesen Sie hierzu die Informationen im Abschnitt „Ferne Advisor-Funktionen mit der Dispatcher-WAN-Unterstützung verwenden“ auf Seite 231.

In früheren Releases mussten die Adresse des verknüpften Servers und die NFA (Non-Forwarding Address) in der Konfiguration übereinstimmen. Diese Einschränkung wurde aufgehoben.

Für das Konfigurieren eines verknüpften Servers bietet der Befehl **`dscontrol server`** eine Option mit dem Namen **`collocated`** an, die auf *yes* oder *no* gesetzt werden kann. Die Standardeinstellung ist *no*. Die Adresse des Servers muss eine gültige IP-Adresse einer Netzschnittstellenkarte in der Maschine sein. Der Parameter `collocated` sollte nicht für Server gesetzt werden, die mittels der Dispatcher-Weiterleitungsmethode `nat` oder `cbr` verknüpft wurden.

Einen verknüpften Server können Sie auf eine der folgenden Arten konfigurieren:

- Wenn Sie die NFA als Adresse des verknüpften Servers verwenden: Legen Sie die NFA mit dem Befehl **`dscontrol executor set nfa IP-Adresse`** fest. Fügen Sie den Server, der die NFA verwendet, mit dem Befehl **`dscontrol server add Cluster:Port:Server`** hinzu.
- Wenn Sie eine andere Adresse als die NFA verwenden: Fügen Sie den Server mit der gewünschten IP-Adresse hinzu, indem Sie den Parameter `collocated` wie folgt auf *yes* setzen: **`dscontrol server add Cluster:Port:Server collocated yes`**.

Bei der Dispatcher-Weiterleitungsmethode `nat` oder `cbr` müssen Sie für die NFA eine nicht verwendete Adapteradresse konfigurieren (bzw. einen Aliasnamen festlegen). Der Server sollte so konfiguriert werden, dass er an dieser Adresse empfangsbereit ist. Konfigurieren Sie den Server mit der folgenden Befehlssyntax:

```
dscontrol server add Cluster:Port:neuer_Aliasname address neuer_Aliasname router  
Router-IP returnaddress Rückkehradresse
```

Wird der Server nicht wie angegeben konfiguriert, können Systemfehler auftreten und/oder Antworten vom Server ausbleiben.

Serververknüpfung mit Dispatcher-Weiterleitungsmethode nat konfigurieren

Wenn Sie einen verknüpften Server mit der Weiterleitungsmethode nat von Dispatcher konfigurieren, müssen Sie im Befehl `dscontrol server add` für den Router eine reale Router-Adresse und nicht die Server-IP-Adresse angeben.

Beim Konfigurieren der Dispatcher-Weiterleitungsmethode nat kann jetzt unter allen Betriebssystemen Unterstützung für die Verknüpfung konfiguriert werden. Dazu müssen auf der Dispatcher-Maschine die folgenden Schritte ausgeführt werden:

- **Auf AIX-Systemen** wird der verknüpfte Server wie jeder andere Server konfiguriert. Die Konfiguration muss nicht geändert werden.
- **Auf Linux-Systemen** wird der verknüpfte Server wie jeder andere Server konfiguriert. Die Konfiguration muss nicht geändert werden.
- **Auf Solaris- und HP-UX-Systemen** wird ganz normal mit `ifconfig` ein Aliasname für den Cluster erstellt. Für die Rückkehradresse darf jedoch kein Aliasname erstellt werden. Stattdessen muss für die Adresse der Befehl `arp publish` ausgeführt werden. Führen Sie dazu den folgenden Befehl aus:

```
arp -s hostname ether_addr pub
```

Verwenden Sie für `ether_addr` die lokale MAC-Adresse. Dies ermöglicht der lokalen Anwendung, Datenverkehr an die Rückkehradresse im Kernel zu senden.

- **Auf Windows-Plattformen** müssen der Cluster und die Rückkehradresse mit dem Befehl `dscontrol executor configure` konfiguriert werden. Sie dürfen nicht in die Windows-Netzwerkconfiguration aufgenommen werden. Für die lokale Anwendung müssen Sie in der Windows-Netzwerkconfiguration einen neuen IP-Aliasnamen zum lokalen Adapter hinzufügen. Verwenden Sie unter "TCP/IP-Einstellungen" die Option "Erweitert", um weitere IP-Adressen für einen Adapter hinzuzufügen. Diese zweite IP-Adresse wird in der Dispatcher-Konfiguration als Serverdefinition verwendet.

Für CBR

CBR unterstützt die Verknüpfung auf allen Plattformen, ohne zusätzliche Konfigurationsschritte zu erfordern. Die verwendeten Webserver und der verwendete Caching Proxy müssen jedoch bindungsspezifisch sein.

Für Site Selector

Site Selector unterstützt die Verknüpfung auf allen Plattformen, ohne zusätzliche Konfigurationsschritte zu erfordern.

Hohe Verfügbarkeit

Die hohe Verfügbarkeit (die mit dem Befehl `dscontrol highavailability` konfiguriert wird) ist für die Komponente Dispatcher verfügbar (jedoch nicht für die Komponenten CBR und Site Selector).

Um die Verfügbarkeit des Dispatchers zu verbessern, benutzt die Dispatcher-Funktion für hohe Verfügbarkeit die folgenden Mechanismen:

- Zwei Dispatcher, die mit denselben Clients und demselben Cluster von Servern sowie untereinander verbunden sind. Beide Dispatcher müssen auf derselben Plattform und unter demselben Betriebssystemtyp ausgeführt werden.

- Ein Mechanismus mit Überwachungssignalen zwischen den beiden Dispatcher-Maschinen, um einen Dispatcher-Ausfall zu erkennen. Für mindestens ein Überwachungssignale austauschendes Paar müssen die NFAs als Quellen- und Zieladresse definiert sein.

Nach Möglichkeit sollte mindestens eines der Paare die Überwachungssignale über ein anderes als das für den regulären Clusterdatenverkehr vorgesehene Teilnetz austauschen. Durch Abgrenzung des durch die Überwachungssignale verursachten Datenverkehrs können in Spitzenbelastungszeiten Fehler bei der Übernahme vermieden werden. Außerdem kann so die Zeit verkürzt werden, die nach einer Überbrückung für eine vollständige Wiederherstellung benötigt wird.

- Eine Liste mit Erreichbarkeitszielen. Beide Dispatcher-Maschinen müssen diese Adressen ansprechen können, damit ein normaler Lastausgleich des Datenverkehrs stattfinden kann. Weitere Informationen hierzu finden Sie im Abschnitt „Fehlererkennung mit Hilfe von Überwachungssignal und Erreichbarkeitsziel“ auf Seite 208.
- Synchronisation der Dispatcher-Informationen (d. h. der Verbindungstabellen, der Erreichbarkeitstabellen und anderer Informationen).
- Logik zur Auswahl des aktiven Dispatchers, der für einen bestimmten Cluster von Servern zuständig ist, und des Dispatchers in Bereitschaft, der permanent für diesen Cluster von Servern synchronisiert wird.
- Ein Mechanismus zur Ausführung der IP-Übernahme, wenn die Logik oder ein Bediener entscheidet, dass der aktive Dispatcher und der Dispatcher in Bereitschaft ihren Status tauschen sollen.

Anmerkung: Eine Abbildung und Beschreibung einer Konfiguration mit *gegenseitiger hoher Verfügbarkeit*, in der sich zwei Dispatcher-Maschinen, die zwei Clustergruppen gemeinsam benutzen, gegenseitig als Ausweichmaschine verwenden, finden Sie im Abschnitt „Gegenseitige hohe Verfügbarkeit“ auf Seite 61. Die gegenseitige hohe Verfügbarkeit ähnelt der einfachen hohen Verfügbarkeit, basiert jedoch speziell auf einer Clusteradresse und nicht auf einer Dispatcher-Maschine als Ganzes. Auf beiden Maschinen müssen die gemeinsam benutzten Clustergruppen identisch konfiguriert werden.

Hohe Verfügbarkeit konfigurieren

Die vollständige Syntax des Befehls **dscontrol highavailability** können Sie dem Abschnitt „dscontrol highavailability — Hohe Verfügbarkeit steuern“ auf Seite 369 entnehmen.

Im Abschnitt „Dispatcher-Maschine konfigurieren“ auf Seite 66 sind die meisten der nachfolgend aufgeführten Tasks genauer beschrieben.

1. Erstellen Sie auf beiden Dispatcher-Maschinen Alias-Script-Dateien. Weitere Informationen hierzu finden Sie im Abschnitt „Scripts verwenden“ auf Seite 209.
2. Starten Sie den Server auf beiden Dispatcher-Servermaschinen.
3. Starten Sie den Executor auf beiden Maschinen.
4. Vergewissern Sie sich, dass die NFA jeder Dispatcher-Maschine konfiguriert und eine für das Teilnetz der Dispatcher-Maschinen gültige IP-Adresse ist.
5. Fügen Sie auf beiden Maschinen Überwachungssignalinformationen hinzu:
`dscontrol highavailability heartbeat add Quellenadresse Zieladresse`

Anmerkung: *Quellenadresse* und *Zieladresse* sind die IP-Adressen (entweder DNS-Namen oder IP-Adressen) der Dispatcher-Maschinen. Die Werte auf den beiden Maschinen werden umgekehrt. Beispiel:

Primäre Maschine - highavailability heartbeat add 9.67.111.3 9.67.186.8
Partnermaschine - highavailability heartbeat add 9.67.186.8 9.67.111.3

Für mindestens ein Überwachungssignale austauschendes Paar müssen die NFAs als Quellen- und Zieladresse definiert sein.

Nach Möglichkeit sollte mindestens eines der Paare die Überwachungssignale über ein anderes als das für den regulären Clusterdatenverkehr vorgesehene Teilnetz austauschen. Durch Abgrenzung des durch die Überwachungssignale verursachten Datenverkehrs können in Spitzenbelastungszeiten Fehler bei der Übernahme vermieden werden. Außerdem kann so die Zeit verkürzt werden, die nach einer Überbrückung für eine vollständige Wiederherstellung benötigt wird.

Legt die Zeit in Sekunden fest, die der Executor als Zeitlimit für die Überwachungssignale für hohe Verfügbarkeit verwendet. Beispiel:

```
dscontrol executor set hatimeout 3
```

Der Standardwert sind 2 Sekunden.

6. Konfigurieren Sie auf beiden Maschinen über den Befehl **reach add** die Liste der IP-Adressen, die der Dispatcher erreichen muss, um einen vollständigen Service zu gewährleisten. Beispiel:

```
dscontrol highavailability reach add 9.67.125.18
```

Erreichbarkeitsziele werden empfohlen, sind aber nicht erforderlich. Weitere Informationen hierzu finden Sie im Abschnitt „Fehlererkennung mit Hilfe von Überwachungssignal und Erreichbarkeitsziel“ auf Seite 208.

7. Fügen Sie zu jeder Maschine die Sicherungsdaten hinzu:

- Für die **primäre** Maschine:

```
dscontrol highavailability backup add primary [auto | manual] Port
```

- Für die **Partnermaschine**:

```
dscontrol highavailability backup  
add backup [auto | manual] Port
```

- Bei der gegenseitigen hohen Verfügbarkeit übernimmt jede Dispatcher-Maschine **beide** Rollen (primäre Maschine und Partnermaschine):

```
dscontrol highavailability backup add both [auto | manual] Port
```

Anmerkung: Wählen Sie als *Port* einen nicht verwendeten Port Ihrer Maschinen aus. Die eingegebene Port-Nummer wird als ein Schlüssel verwendet, um sicherzustellen, dass das Paket vom richtigen Host empfangen wird.

8. Überprüfen Sie den Status der hohen Verfügbarkeit auf den beiden Maschinen:

```
dscontrol highavailability status
```

Die Maschinen sollten jeweils die korrekte Rolle (Partnermaschine und/oder primäre Maschine), die korrekten Status und die korrekten untergeordneten Status aufweisen. Die primäre Maschine sollte aktiv und synchronisiert sein. Die Ausweichmaschine sollte sich im Bereitschaftsmodus befinden und innerhalb kurzer Zeit synchronisiert werden. Die Strategien müssen übereinstimmen.

9. Konfigurieren Sie auf beiden Maschinen die Cluster-, Port- und Serverinformationen.

Anmerkung: Für die Konfiguration der gegenseitigen hohen Verfügbarkeit (Abb. 14 auf Seite 61) konfigurieren Sie beispielsweise die Clustergruppen, die von den beiden Dispatchern gemeinsam benutzt werden, wie folgt:

- Setzen Sie für Dispatcher 1 den folgenden Befehl ab:
`dscontrol cluster set ClusterA primaryhost NFADispatcher1`
`dscontrol cluster set ClusterB primaryhost NFADispatcher2`
- Setzen Sie für Dispatcher 2 den folgenden Befehl ab:
`dscontrol cluster set ClusterB primaryhost NFADispatcher2`
`dscontrol cluster set ClusterA primaryhost NFADispatcher1`

10. Starten Sie auf beiden Maschinen den Manager und die Advisor-Funktionen.

Anmerkungen:

1. Wollen Sie für die Paketweiterleitung eine Dispatcher-Maschine ohne Ausweichmaschine konfigurieren, setzen Sie beim Start keine Befehle für hohe Verfügbarkeit ab.
2. Wollen Sie eine für die hohe Verfügbarkeit erstellte Konfiguration mit zwei Dispatcher-Maschinen in eine Konfiguration mit nur einer Maschine ändern, beenden Sie den Executor auf einer der Maschinen und löschen Sie dann die Funktionen für hohe Verfügbarkeit (Überwachungssignale, Erreichbarkeit und Partnermaschine) auf der anderen Maschine.
3. In beiden oben geschilderten Fällen müssen Sie ggf. auf der Netzschnittstellenkarte Aliasnamen für Clusteradressen angeben.
4. Wenn zwei Dispatcher-Maschinen in einer Konfiguration mit hoher Verfügbarkeit synchronisiert werden, führen Sie zunächst alle dscontrol-Befehle (zum Aktualisieren der Konfiguration) auf der Bereitschaftsmaschine und dann auf der aktiven Maschine aus.
5. Wenn Sie zwei Dispatcher-Maschinen in einer Umgebung mit hoher Verfügbarkeit verwenden, können unerwartete Ergebnisse auftreten, wenn einer der Parameter für Executor, Cluster, Port oder Server (z. B. port stickytime) auf beiden Maschinen auf verschiedene Werte gesetzt ist.
6. Berücksichtigen Sie bei der gegenseitigen hohen Verfügbarkeit den Fall, bei dem einer der Dispatcher aktiv Pakete für seinen primären Cluster weiterleiten muss und außerdem das Weiterleiten von Paketen für den Partnercluster übernehmen muss. Stellen Sie sicher, dass damit die Kapazität für den Durchsatz auf dieser Maschine nicht überschritten wird.
7. Wenn Sie auf Linux-Systemen bei Verwendung der MAC-Port-Weiterleitungsmethode von Dispatcher gleichzeitig hohe Verfügbarkeit und Verknüpfung konfigurieren, lesen Sie die Informationen im Abschnitt „Alternativen für die Festlegung eines Loopback-Aliasnamens unter Linux bei Verwendung der Load-Balancer-Weiterleitungsmethode mac“ auf Seite 78.
8. Wenn Sie auf Windows-Systemen gleichzeitig hohe Verfügbarkeit und Verknüpfung konfigurieren, lesen Sie die Informationen im Abschnitt „Verknüpfung und hohe Verfügbarkeit konfigurieren (Windows-Systeme)“ auf Seite 212.
9. Die Konfiguration der hohen Verfügbarkeit kann Probleme wie die folgenden auslösen:
 - Übergangene Verbindungen nach der Übernahme
 - keine Synchronisation mit Partnermaschinen
 - versehentliche Übertragung von Anforderungen an die Ausweichpartnermaschine

Wie Sie die Auswirkungen solcher Probleme verringern können, ist im Abschnitt „Problem: Konfigurationstipps für hohe Verfügbarkeit“ auf Seite 321 beschrieben.

Fehlererkennung mit Hilfe von Überwachungssignal und Erreichbarkeitsziel

Neben den Basiskriterien der Fehlererkennung (durch Überwachungssignale erkannter Verlust der Konnektivität zwischen aktivem Dispatcher und Bereitschafts-Dispatcher) gibt es einen weiteren Fehlererkennungsmechanismus, der als *Erreichbarkeitskriterien* bezeichnet wird. Wenn Sie den Dispatcher konfigurieren, können Sie eine Liste von Hosts angeben, die für jeden der Dispatcher erreichbar sein sollten, damit die Dispatcher fehlerfrei arbeiten können. Die beiden Partner für hohe Verfügbarkeit tauschen ständig Überwachungssignale und Informationen darüber aus, wie viele Erreichbarkeitsziele jeder von ihnen mit einem ping-Signal erreichen kann. Wenn der Bereitschafts-Dispatcher mehr Erreichbarkeitsziele mit einem ping-Signal erreichen kann, übernimmt er die Aufgaben des primären Dispatchers.

Der aktive Dispatcher sendet jede halbe Sekunde ein Überwachungssignal, das normalerweise vom Bereitschafts-Dispatcher empfangen wird. Sollte der Bereitschafts-Dispatcher zwei Sekunden kein Überwachungssignal empfangen, beginnt er mit der Funktionsübernahme. Die Funktionsübernahme durch den Bereitschafts-Dispatcher findet jedoch nur statt, wenn alle Überwachungssignale ausfallen. Wenn zwei Überwachungssignalpaare konfiguriert sind, müssen demnach beide ausfallen. Konfigurieren Sie zur Stabilisierung einer Umgebung mit hoher Verfügbarkeit mehr als ein Überwachungssignalpaar, damit ständige Funktionsübernahmen vermieden werden.

Als Erreichbarkeitsziele sollten Sie mindestens einen Host für jedes Teilnetz auswählen, das die Dispatcher-Maschine verwendet. Die Hosts können Router, IP-Server oder andere Arten von Hosts sein. Die Erreichbarkeit von Hosts wird über ping-Aufrufe der Advisor-Funktion *reach* abgefragt. Es findet eine Funktionsübernahme statt, wenn keine Überwachungssignalnachrichten durchkommen oder die Erreichbarkeitskriterien vom Bereitschafts-Dispatcher eher erfüllt werden als vom primären Dispatcher. Damit die Entscheidung anhand aller verfügbaren Informationen getroffen wird, sendet der aktive Dispatcher regelmäßig Informationen über seine Erreichbarkeit an den Dispatcher in Bereitschaft. Der Dispatcher in Bereitschaft vergleicht dann diese Informationen mit seinen eigenen Erreichbarkeitsinformationen und entscheidet, ob eine Übernahme vorgenommen werden soll oder nicht.

Anmerkung: Wenn Sie das Erreichbarkeitsziel konfigurieren, müssen Sie die Advisor-Funktion *reach* starten. Die Advisor-Funktion *reach* wird automatisch gestartet, wenn Sie die Manager-Funktion starten. Weitere Informationen zur Advisor-Funktion *reach* finden Sie auf Seite 190.

Wiederherstellungsstrategie

Es werden zwei Dispatcher-Maschinen konfiguriert, die primäre Maschine und eine zweite Maschine, die so genannte *Partnermaschine*. Wird die primäre Maschine gestartet, leitet sie die gesamten Verbindungsdaten so lange an die Partnermaschine weiter, bis die beiden Maschinen synchronisiert sind. Die primäre Maschine wird *aktiv*, d. h., sie beginnt mit dem Lastausgleich. Die Partnermaschine überwacht in der Zwischenzeit den Status der primären Maschine und befindet sich in *Bereitschaft*.

Stellt die Partnermaschine an einem beliebigen Punkt fest, dass die primäre Maschine ausgefallen ist, *übernimmt* sie die Lastausgleichsfunktionen der primären Maschine und wird zur aktiven Maschine. Ist die primäre Maschine wieder betriebsbereit, gehen die Maschinen anhand der vom Benutzer konfigurierten *Wiederherstellungsstrategie* vor. Es gibt zwei Strategiearten:

Auto Die primäre Maschine nimmt das Weiterleiten von Paketen automatisch wieder auf, sobald sie wieder betriebsbereit ist.

Manual

Die Partnermaschine setzt das Weiterleiten von Paketen fort, auch wenn die primäre Maschine wieder betriebsbereit ist. Soll die primäre Maschine wieder in den Status der aktiven Maschine und die Partnermaschine wieder in den Bereitschaftsstatus zurückgesetzt werden, ist ein manueller Eingriff erforderlich.

Der Parameter für die Strategie muss für beide Maschinen auf denselben Wert gesetzt werden.

Bei der Strategie der manuellen Wiederherstellung können Sie über den Befehl **takeover** das Weiterleiten von Paketen durch eine bestimmte Maschine erzwingen. Die manuelle Wiederherstellung ist nützlich, wenn die andere Maschine gewartet wird. Die automatische Wiederherstellung ist für den normalen, nicht überwachten Betrieb konzipiert.

In einer Konfiguration mit gegenseitiger hoher Verfügbarkeit gibt es keinen Fehler auf Clusterbasis. Tritt ein Fehler bei einer Maschine auf, übernimmt die andere Maschine die Rolle für beide Cluster, auch wenn der Fehler nur einen Cluster betrifft.

Anmerkung: In einer Übernahme-situation können einige Verbindungsaktualisierungen verloren gehen. Vorhandene Verbindungen, die längere Zeit bestehen (z. B. telnet-Verbindungen), auf die zum Zeitpunkt der Übernahme zugegriffen wird, können dadurch beendet werden.

Scripts verwenden

Damit der Dispatcher Pakete weiterleiten kann, muss ein Aliasname für jede Clusteradresse auf einer Netzschnittstelleneinheit angegeben werden.

- In einer Standalone-Dispatcher-Konfiguration muss ein Aliasname für jede Clusteradresse auf einer Netzschnittstellenkarte (beispielsweise en0, tr0) definiert werden.
- In einer Konfiguration mit hoher Verfügbarkeit gilt Folgendes:
 - Auf der aktiven Maschine muss ein Aliasname für jede Clusteradresse auf einer Netzschnittstellenkarte (beispielsweise en0, tr0) definiert werden.
 - Auf der Bereitschaftsmaschine muss ein Aliasname für jede Clusteradresse auf einer Loopback-Einheit (z. B. lo0) angegeben werden, wenn Sie die Weiterleitungsmethode mac verwenden und mit verknüpften Servern arbeiten.
- In allen Maschinen, in denen der Executor beendet wurde, müssen alle Aliasnamen entfernt werden, um Konflikte mit einer anderen Maschine zu vermeiden, die möglicherweise gestartet wird.

Informationen zur Festlegung eines Aliasnamens auf der Netzschnittstellenkarte finden Sie im Abschnitt „Schritt 5. Aliasnamen auf der Netzschnittstellenkarte erstellen“ auf Seite 69.

Da die Dispatcher-Maschinen bei einem erkannten Fehler ihren Status tauschen, müssen die oben angegebenen Befehle automatisch abgesetzt werden. Dazu führt der Dispatcher vom Benutzer erstellte Scripts aus. Beispiel-Scripts finden Sie im Verzeichnis `...ibm/edge/lb/servers/samples`. Zum Ausführen *müssen* Sie diese in das Verzeichnis `...ibm/edge/lb/servers/bin` verschieben. Die Scripts werden automatisch nur ausgeführt, wenn dsserver aktiv ist.

Anmerkungen:

1. In einer Konfiguration mit gegenseitiger hoher Verfügbarkeit wird jedes "go"-Script vom Dispatcher mit einem Parameter aufgerufen, der die Adresse des primären Dispatchers angibt. Das Script muss diesen Parameter abfragen und die Befehle **executor configure** für die Clusteradressen ausführen, die diesem primären Dispatcher zugeordnet sind.
2. Wenn Sie für die Dispatcher-Weiterleitungsmethode `nat` die hohe Verfügbarkeit konfigurieren möchten, müssen Sie die Rückkehradressen zu den Script-Dateien hinzufügen.

Sie können die folgenden Beispiel-Scripts verwenden:

goActive

Das Script `goActive` wird ausgeführt, wenn ein Dispatcher in den aktiven Status wechselt und mit dem Weiterleiten von Paketen beginnt.

- Wenn Sie den Dispatcher in einer Konfiguration mit hoher Verfügbarkeit verwenden, müssen Sie dieses Script erstellen. Dieses Script löscht die Loopback-Aliasnamen und fügt Einheitenaliasnamen hinzu.
- Wird der Dispatcher in einer Standalone-Konfiguration ausgeführt, benötigen Sie dieses Script nicht.

goStandby

Das Script `goStandby` wird ausgeführt, wenn ein Dispatcher in den Bereitschaftsstatus wechselt, in dem der Status der aktiven Maschine überwacht wird, jedoch keine Pakete weitergeleitet werden.

- Wenn Sie den Dispatcher in einer Konfiguration mit hoher Verfügbarkeit verwenden, müssen Sie dieses Script erstellen. Dieses Script sollte Einheitenaliasnamen löschen und Loopback-Aliasnamen hinzufügen.
- Wird der Dispatcher in einer Standalone-Konfiguration ausgeführt, benötigen Sie dieses Script nicht.

goInOp

Das Script `goInOp` wird beim Stoppen eines Dispatcher-Executors ausgeführt.

- Wenn Sie den Dispatcher normalerweise in einer Konfiguration mit hoher Verfügbarkeit verwenden, können Sie dieses Script erstellen. Dieses Script löscht alle Einheiten- und Loopback-Aliasnamen.
- Wird der Dispatcher normalerweise in einer Standalone-Konfiguration ausgeführt, ist dieses Script optional. Sie können das Script erstellen und zum Löschen Einheitenaliasnamen benutzen oder Aliasnamen manuell löschen.

goIdle Das Script `goIdle` wird ausgeführt, wenn ein Dispatcher in den Freizustand und mit dem Weiterleiten von Paketen beginnt. Dieser Fall tritt ein, wenn die hohe Verfügbarkeit nicht hinzugefügt wurde, wie es in einer Standalone-Konfiguration der Fall ist. In einer Konfiguration mit hoher Verfügbarkeit geschieht dies auch vor dem Hinzufügen bzw. nach dem Entfernen der Merkmale für hohe Verfügbarkeit.

- Wenn Sie den Dispatcher normalerweise in einer Konfiguration mit hoher Verfügbarkeit verwenden, sollten Sie dieses Script *nicht* erstellen.
- Wird der Dispatcher normalerweise in einer Standalone-Konfiguration ausgeführt, ist dieses Script optional. Sie können dieses Script erstellen und zum Hinzufügen von Einheitenaliasnamen benutzen oder Aliasnamen manuell hinzufügen. Wenn Sie dieses Script nicht für Ihre Standalone-Konfiguration erstellen, müssen Sie den Befehl **dscontrol executor configure** verwenden oder bei jedem Start des Executors die Aliasnamen manuell konfigurieren.

highavailChange

Das Script highavailChange wird ausgeführt, wenn sich der Status der hohen Verfügbarkeit auf einer Dispatcher-Maschine so ändert, dass eines der go-Scripts aufgerufen wird. Der einzige an dieses Script übergebene Parameter ist der Name des gerade vom Dispatcher ausgeführten go-Scripts. Sie können dieses Script beispielsweise so schreiben, dass Informationen zu Statusänderungen verwendet werden, um Alerts an einen Administrator zu senden, oder derartige Ereignisse einfach erfasst werden.

Auf Windows-Systemen: Wenn Sie Ihre Konfiguration so eingerichtet haben, dass Site Selector den Lastausgleich für zwei Dispatcher-Maschinen in einer Umgebung mit hoher Verfügbarkeit durchführt, müssen Sie im Microsoft Stack einen Aliasnamen für die Messwertserver hinzufügen. Dieser Aliasname sollte auch zum Script goActive hinzugefügt werden. Beispiel:

```
call netsh interface ip add address "Local Area Connection"
    addr=9.37.51.28 mask=255.255.240.0
```

In den Scripts goStandby und goInOp muss der Aliasname entfernt werden. Beispiel:

```
call netsh interface ip delete address "Local Area Connection"
    addr=9.37.51.28
```

Wenn die Maschine mehrere NICs enthält, überprüfen Sie zunächst, welche Schnittstelle verwendet werden sollte. Setzen Sie dazu an der Eingabeaufforderung den Befehl netsh interface ip show address ab. Dieser Befehl gibt eine Liste der zur Zeit konfigurierten Schnittstellen zurück und versieht die Angabe "Local Area Connection" mit einer Nummer (z. B. "Local Area Connection 2"), so dass Sie bestimmen, welche Schnittstelle Sie verwenden sollten.

Unter Linux für S/390: Der Dispatcher setzt unaufgefordert ein ARP ab (gratuitous ARP), um IP-Adressen von einem Dispatcher auf einen anderen Dispatcher zu versetzen. Dieser Mechanismus ist somit an den Typ des zugrundeliegenden Netzes gebunden. Wenn Sie Linux für S/390 verwenden, kann der Dispatcher seine native Funktion für eine vollständig Funktionsübernahme für hohe Verfügbarkeit (mit Verschiebung der IP-Adressen) nur für Schnittstellen ausführen, die unaufgefordert ein ARP absetzen und die Adresse auf der lokalen Schnittstelle konfigurieren können. Auf Punkt-zu-Punkt-Schnittstellen wie IUCV und CTC sowie in bestimmten Konfigurationen von qeth/QDIO kann dieser Mechanismus nicht ordnungsgemäß funktionieren.

Für Schnittstellen und Konfigurationen, bei denen die native Dispatcher-Funktion der IP-Übernahme nicht ordnungsgemäß ausgeführt werden kann, kann der Kunde entsprechende Befehle in die go-Scripts aufnehmen, um die Adressen manuell zu versetzen. Auf diese Weise wird sichergestellt, dass auch diese Netztopologien von der hohen Verfügbarkeit profitieren können.

Verknüpfung und hohe Verfügbarkeit konfigurieren (Windows-Systeme)

Auf Windows-Servern können Sie die Load-Balancer-Features "Hohe Verfügbarkeit" und "Verknüpfung" zusammen konfigurieren. Hierfür sind allerdings einige zusätzliche Schritte erforderlich.

Wenn Sie auf Windows-Systemen die Verknüpfung zusammen mit der hohen Verfügbarkeit nutzen möchten, benötigen Sie eine zusätzliche IP-Adresse. Dies ist eine Art Pseudoadresse, die zum Loopback-Adapter des Windows-Systems hinzugefügt werden kann. Der Loopback-Adapter muss auf der primären Maschine und der Ausweichmaschine installiert sein. Die Schritte für die Installation der Loopback-Einheit auf Windows-Systemen sind im Abschnitt „Servermaschinen für Lastausgleich konfigurieren“ auf Seite 72 aufgelistet.

Wenn Sie in diesen Schritten aufgefordert werden, die Cluster-IP-Adresse zur Loopback-Einheit hinzuzufügen, müssen Sie an Stelle der Clusteradresse eine Pseudo-IP-Adresse hinzufügen. Dies liegt daran, dass die go*-Scripts für hohe Verfügbarkeit für Windows die Clusteradresse für die Loopback-Einheit löschen oder zur Loopback-Einheit hinzufügen müssen, je nachdem, ob die Load-Balancer-Maschine aktiv oder in Bereitschaft ist.

Die letzte konfigurierte IP-Adresse für die Loopback-Einheit kann auf Windows-Systemen nicht entfernt werden, weil die Loopback-Einheit nicht im DHCP-Modus funktioniert. Durch die Pseudoadresse kann Load Balancer seine Clusteradresse jederzeit entfernen. Die Pseudo-IP-Adresse wird für keine Art von Datenverkehr verwendet und kann sowohl auf der aktiven Maschine als auch auf der Ausweichmaschine verwendet werden.

Aktualisieren Sie die go*-Scripts von Load Balancer und verschieben Sie sie auf die aktive Maschine und die Ausweichmaschine. Starten Sie dann Dispatcher. Die Clusteradresse wird zu den entsprechenden Zeiten zur Netzschnittstelle und zur Loopback-Einheit hinzugefügt oder von der Netzschnittstelle und der Loopback-Einheit entfernt.

Regelbasierten Lastausgleich konfigurieren

Mit einem auf Regeln basierenden Lastausgleich kann genau abgestimmt werden, wann und warum Pakete an welche Server gesendet werden. Load Balancer überprüft alle hinzugefügten Regeln von der ersten Priorität bis zur letzten Priorität, stoppt bei der ersten Regel, die wahr ist, und verteilt dann die Last auf alle Server, die der Regel zugeordnet sind. Load Balancer verteilt die Last bereits ausgehend von Ziel und Port. Mit der Anwendung von Regeln haben Sie jedoch erweiterte Möglichkeiten für die Verteilung von Verbindungen.

In den meisten Fällen sollten Sie beim Konfigurieren von Regeln eine **immer gültige** Standardregel konfigurieren, um auch Anforderungen zu registrieren, die von den anderen Regeln höherer Priorität nicht erfasst werden. Diese Standardregel könnte beispielsweise die Antwort "Die Site ist derzeit leider nicht verfügbar, versuchen Sie es später erneut" sein, wenn alle anderen Server nicht für die Client-Anforderung verwendet werden können.

Sie sollten den regelbasierten Lastausgleich für Dispatcher und Site Selector verwenden, wenn Sie aus bestimmten Gründen nur einen Teil Ihrer Server nutzen möchten. Für die Komponente CBR *müssen* Sie in jedem Fall Regeln verwenden.

Es sind folgende Arten von Regeln verfügbar:

- Für Dispatcher:
 - Client-IP-Adresse
 - Client-Port
 - Uhrzeit
 - Diensttyp (TOS, Type of Service)
 - Verbindungen pro Sekunde
 - Summe der aktiven Verbindungen
 - Reservierte Bandbreite
 - Gemeinsame Bandbreite
 - Immer wahr
 - Inhalt einer Anforderung
- Für CBR:
 - Client-IP-Adresse
 - Uhrzeit
 - Verbindungen pro Sekunde
 - Summe der aktiven Verbindungen
 - Immer wahr
 - Inhalt einer Anforderung
- Für Site Selector:
 - Client-IP-Adresse
 - Uhrzeit
 - Metrik gesamt
 - Metrik Durchschnitt
 - Immer wahr

Erstellen Sie einen Plan der Logik, die von den Regeln befolgt werden soll, bevor der Konfiguration Regeln hinzugefügt werden.

Wie werden Regeln ausgewertet?

Jede Regel hat einen Namen, einen Typ und eine Priorität und kann neben einer Servergruppe auch einen Bereichsanfang und ein Bereichsende haben. Dem Regeltyp content für die Komponente CBR ist ein regulärer Ausdruck (pattern) für den Abgleich zugeordnet. (Beispiele und Szenarien für die Verwendung der content-Regel sowie eine gültige pattern-Syntax für die content-Regel finden Sie in Anhang B, „Syntax der content-Regel“, auf Seite 477.)

Regeln werden in der Reihenfolge ihrer Priorität ausgewertet. Eine Regel mit der Priorität 1 (kleinere Nummer) wird vor einer Regel mit der Priorität 2 (größere Nummer) ausgewertet. Die erste Regel, die erfüllt ist, wird verwendet. Sobald eine Regel erfüllt ist, werden keine weiteren Regeln ausgewertet.

Eine Regel ist erfüllt, wenn die beiden folgenden Bedingungen zutreffen:

1. Das Prädikat der Regel muss wahr sein. Das heißt, dass der Wert, der ausgewertet wird, zwischen Bereichsanfang und -ende liegen muss, oder der Inhalt mit dem regulären Ausdruck übereinstimmen muss, der für pattern in der content-Regel angegeben wurde. Für Regeln des Typs "true" stimmt das Prädikat unabhängig vom Bereichsanfang und -ende immer überein.

2. Sind der Regel Server zugeordnet, muss mindestens ein Server eine Wertigkeit von mehr 0 haben, damit Pakete an ihn weitergeleitet werden.

Sind einer Regel keine Server zugeordnet, muss für die Regel nur die erste Bedingung zutreffen, um erfüllt zu sein. In diesem Fall löscht der Dispatcher die Verbindungsanforderung. Site Selector gibt die Namensserveranforderung mit einem Fehler zurück und CBR veranlasst Caching Proxy, eine Fehlerseite auszugeben.

Wird keine der Regeln erfüllt, wählt Dispatcher aus allen für den Port verfügbaren Servern einen Server aus. Site Selector wählt aus allen für den Sitenamen verfügbaren Servern einen Server aus, und CBR veranlasst Caching Proxy, eine Fehlerseite auszugeben.

Auf der Client-IP-Adresse basierende Regeln verwenden

Dieser Regeltyp ist für die Komponenten Dispatcher, CBR und Site Selector verfügbar.

Sie können Regeln auf der Basis der Client-IP-Adresse verwenden, wenn die Herkunft das Kriterium für die Auswahl von Kunden und die Ressourcenzuordnung sein soll.

Stellen Sie sich vor, dass in Ihrem Netz in großem Umfang ein unbezahlter und deshalb unerwünschter Datenaustausch von Clients mit bestimmten IP-Adressen stattfindet. In diesem Fall könnten Sie mit dem Befehl **dscontrol rule** eine Regel erstellen. Beispiel:

```
dscontrol rule add 9.67.131.153:80:ni type ip  
beginrange 9.0.0.0 endrange 9.255.255.255
```

Diese ni-Regel blendet alle Verbindungen für unerwünschte Clients aus. Anschließend fügen Sie die Server zur Regel hinzu, die zugänglich sein sollen. Werden keine Server zur Regel hinzugefügt, werden Anforderungen von den Adressen 9.x.x.x von keinem Ihrer Server bedient.

Auf dem Client-Port basierende Regeln verwenden

Dieser Regeltyp ist nur für die Komponente Dispatcher verfügbar.

Wenn Ihre Clients eine Software verwenden, die für Anforderungen von TCP/IP einen bestimmten Port anfordert, möchten Sie vielleicht Regeln auf der Basis des Client-Ports verwenden.

Sie könnten beispielsweise eine Regel erstellen, die angibt, dass für alle Anforderungen mit einem Client-Port von 10002 eine Gruppe besonders schneller Server bereitgestellt wird, da bekannt ist, dass alle Client-Anforderungen mit diesem Port von einer besonders wichtigen Kundengruppe stammen.

Auf der Uhrzeit basierende Regeln verwenden

Dieser Regeltyp ist für die Komponenten Dispatcher, CBR und Site Selector verfügbar.

Möglicherweise sollen aus Gründen der Kapazitätsplanung Regeln verwendet werden, die auf der Uhrzeit basieren. Ist beispielsweise Ihre Website täglich zu bestimmten Zeiten besonders stark frequentiert, können Sie während der Spitzenzeit fünf zusätzliche Server hinzufügen.

Ein anderer Grund für die Verwendung einer Regel, die auf der Uhrzeit basiert, kann vorliegen, wenn Sie jede Nacht um Mitternacht einige der Server zur Wartung herunterfahren möchten. In diesem Fall können Sie eine Regel erstellen, mit der die Server während der benötigten Wartungszeit ausgeschlossen werden.

Auf der Serviceart basierende Regeln verwenden

Dieser Regeltyp ist nur für die Komponente Dispatcher verfügbar.

Möglicherweise sollen Regeln verwendet werden, die auf dem Inhalt des Felds "Type of Service" (TOS) im IP-Header basieren. Wird beispielsweise eine Client-Anforderung mit einem TOS-Wert empfangen, der einen normalen Service angibt, kann die Anforderung an eine Servergruppe weitergeleitet werden. Wird eine andere Client-Anforderung mit einem anderen TOS-Wert empfangen, der einen Service mit höherer Priorität angibt, kann die Anforderung an eine andere Servergruppe weitergeleitet werden.

Die TOS-Regel ermöglicht die vollständige Konfiguration jedes Bits im TOS-Byte unter Verwendung des Befehls **dscontrol rule**. Für signifikante Bits, die im TOS-Byte abgeglichen werden sollen, verwenden Sie 0 oder 1. Andernfalls wird der Wert x verwendet. Das folgende Beispiel zeigt das Hinzufügen einer TOS-Regel:

```
dscontrol rule add 9.67.131.153:80:tsr type service tos 0xx1010x
```

Regeln auf der Basis der Verbindungen pro Sekunde verwenden

Dieser Regeltyp ist für die Komponenten Dispatcher und CBR verfügbar.

Anmerkung: Der Manager muss aktiv sein, damit die folgenden Regeln angewendet werden können.

Vielleicht möchten Sie Regeln verwenden, die auf den Verbindungen pro Sekunde basieren, wenn einige Ihrer Server auch von anderen Anwendungen benutzt werden sollen. Sie können beispielsweise zwei Regeln erstellen:

1. Wenn Verbindungen pro Sekunde am Port 80 zwischen 0 und 2000, diese 2 Server verwenden
2. Wenn Verbindungen pro Sekunde am Port 80 über 2000, diese 10 Server verwenden

Möglicherweise verwenden Sie Telnet und möchten zwei Ihrer fünf Server für Telnet reservieren, es sei denn, die Verbindungen pro Sekunde überschreiten eine bestimmte Zahl. In diesem Fall würde der Dispatcher zu Spitzenzeiten die Last auf alle fünf Server verteilen.

Regelauswertungsoption upserversonrule in Verbindung mit Regeln des Typs connection setzen: Wenn Sie Regeln des Typs connection verwenden und die Option **upserversonrule** setzen, können Sie sicherstellen, dass die übrigen Server nicht überlastet werden, falls einige Server der Gruppe heruntergefahren sind. Weitere Informationen hierzu finden Sie im Abschnitt „Regeloption für Serverauswertung“ auf Seite 220.

Regeln auf der Basis der Summe aktiver Verbindungen verwenden

Dieser Regeltyp ist für die Komponenten Dispatcher und CBR verfügbar.

Anmerkung: Der Manager muss aktiv sein, damit die folgenden Regeln angewendet werden können.

Wenn Ihre Server überlastet sind und beginnen, Pakete zu verwerfen, möchten Sie vielleicht Regeln anwenden, die auf der Gesamtanzahl der an einem Port aktiven Verbindungen basieren. Von bestimmten Webservern werden weiterhin Verbindungen akzeptiert, auch wenn sie nicht über genügend Threads verfügen, um auf die Anforderung zu antworten. Von dem Client wird daraufhin eine Zeitlimitüberschreitung angefordert, und der Kunde, der Ihre Website aufruft, erhält keinen Service. Sie können Regeln verwenden, die auf den aktiven Verbindungen basieren, um die Kapazität innerhalb eines Pools mit Servern auszugleichen.

Sie wissen beispielsweise aus Erfahrung, dass Ihre Server den Service einstellen, nachdem sie 250 Verbindungen akzeptiert haben. Sie können eine Regel mit dem Befehl **dscontrol rule** oder mit dem Befehl **cbrcontrol rule** erstellen. Beispiel:

```
dscontrol rule add 130.40.52.153:80:pool2 type active
beginrange 250 endrange 500
```

oder

```
cbrcontrol rule add 130.40.52.153:80:pool2 type active
beginrange 250 endrange 500
```

Sie würden dann der Regel Ihre aktuellen Server plus einige zusätzliche Server hinzufügen, die andernfalls für eine andere Verarbeitung verwendet werden.

Regeln auf der Basis der reservierten und gemeinsam benutzten Bandbreite verwenden

Regeln für reservierte und gemeinsam genutzte Bandbreite sind nur für die Komponente Dispatcher verfügbar.

Für die Bandbreitenregeln errechnet der Dispatcher die Bandbreite als Geschwindigkeit, mit der die Daten von einer bestimmten Servergruppe an Clients geliefert werden. Der Dispatcher protokolliert die Kapazität auf Server-, Regel-, Port-, Cluster- und Executor-Ebene. Für jede dieser Ebenen gibt es ein Feld "Bytezähler", das übertragene Kilobytes pro Sekunde angibt. Der Dispatcher berechnet diese Geschwindigkeiten über einen Zeitraum von 60 Sekunden. Sie können diese Geschwindigkeitswerte auf der grafischen Benutzerschnittstelle (GUI) oder in der Ausgabe eines Befehlszeilenberichts anzeigen.

Regel "Reservierte Bandbreite"

Mit der Regel "Reservierte Bandbreite" können Sie die von einer Servergruppe gelieferte Datenmenge in Kilobytes pro Sekunde steuern. Durch Festlegen eines Schwellenwertes (Zuordnen eines bestimmten Bandbreitenbereichs) für jede Gruppe von Servern in der Konfiguration können Sie die von jeder Cluster-Port-Kombination genutzte Bandbreite steuern und gewährleisten.

Nachfolgend sehen Sie ein Beispiel für das Hinzufügen einer reservedbandwidth-Regel:

```
dscontrol rule add 9.67.131.153:80:rbw type reservedbandwidth
beginrange 0 endrange 300
```

Bereichsanfang und -ende werden in Kilobytes pro Sekunde angegeben.

Regel "Gemeinsame Bandbreite"

Vor dem Konfigurieren der Regel "Gemeinsame Bandbreite" müssen Sie die maximale Bandbreite (Kilobytes pro Sekunde) angeben, die auf Executor- oder Cluster-ebene gemeinsam genutzt werden kann, indem Sie den Befehl **dscontrol executor** oder **dscontrol cluster** mit der Option **sharedbandwidth** verwenden. Der Wert für **sharedbandwidth** sollte die insgesamt verfügbare Bandbreite (gesamte Netzkapazität) nicht überschreiten. Das Setzen der gemeinsam genutzten Bandbreite mit dem Befehl **dscontrol** gibt nur eine Obergrenze für die Regel an.

Nachfolgend sind Beispiele für die Befehlssyntax aufgeführt:

```
dscontrol executor set sharedbandwidth Größe  
dscontrol cluster [add | set] 9.12.32.9 sharedbandwidth Größe
```

Größe ist für **sharedbandwidth** ein ganzzahliger Wert (in Kilobytes pro Sekunde). Der Standardwert ist null. Ist der Wert gleich null, kann die Bandbreite nicht gemeinsam benutzt werden.

Bei gemeinsamer Nutzung von Bandbreite auf Clusterebene steht dem Cluster eine angegebene maximale Bandbreite zur Nutzung zur Verfügung. Solange die vom Cluster genutzte Bandbreite unter dem angegebenen Wert liegt, ist das Ergebnis der Auswertung für diese Regel **true**. Liegt die insgesamt genutzte Bandbreite über dem angegebenen Wert, ist das Ergebnis der Regelauswertung **false**.

Bei gemeinsamer Nutzung der Bandbreite auf Executor-Ebene kann die gesamte Dispatcher-Konfiguration eine maximale Bandbreite nutzen. Solange die auf Executor-Ebene genutzte Bandbreite unter dem angegebenen Wert liegt, ist das Ergebnis der Auswertung für diese Regel **true**. Liegt die insgesamt genutzte Bandbreite über der definierten Bandbreite, ist das Ergebnis der Regelauswertung **false**.

Nachfolgend einige Beispiele für das Hinzufügen oder Definieren einer **sharedbandwidth**-Regel:

```
dscontrol rule add 9.20.30.4:80:shbw type sharedbandwidth sharelevel Wert  
dscontrol rule set 9.20.34.11:80:shrule sharelevel Wert
```

Der *Wert* für **sharelevel** ist **executor** oder **cluster**. Der Parameter **sharelevel** ist für die Regel **sharedbandwidth** erforderlich.

Regeln für reservierte und gemeinsam genutzte Bandbreite verwenden

Dispatcher gibt Ihnen die Möglichkeit, mit der Regel *Reservierte Bandbreite* Gruppen von Servern in Ihrer Konfiguration eine angegebene Bandbreite zuzuordnen. Durch Angabe von Bereichsanfang und -ende können Sie steuern, wie viele Kilobytes eine Servergruppe an Clients senden kann. Wenn die Regelauswertung nicht mehr das Ergebnis **true** ergibt (das Bereichsende überschritten wird), wird die Regel mit der nächst niedrigeren Priorität ausgewertet. Falls die Regel mit der nächst niedrigeren Priorität eine immer gültige Regel (**always true**) ist, könnte ein Server ausgewählt werden, der Clients eine Antwort vom Typ "Site ausgelastet" sendet.

Beispiel: Gehen wir von einer Gruppe mit drei Servern am Port 2222 aus. Wenn die reservierte Bandbreite auf 300 gesetzt ist, werden über einen Zeitraum von 60 Sekunden maximal 300 Kilobytes pro Sekunden übertragen. Wird diese Geschwindigkeit überschritten, ist das Ergebnis der Regelauswertung nicht mehr **true**. Falls dies die einzige Regel ist, würde der Dispatcher einen der drei Server zur Bearbeitung der Anfrage auswählen. Wenn es eine Regel mit niedrigerer Priorität ("always true") gibt, könnte die Anfrage an einen anderen Server umgeleitet und mit "Site ausgelastet" beantwortet werden.

Die Regel für gemeinsam genutzte Bandbreite erweitert den Serverzugriff für Clients. Wenn diese Regel als Regel niedrigerer Priorität nach einer Regel für reservierte Bandbreite verwendet wird, kann ein Client auch dann noch auf einen Server zugreifen, wenn die reservierte Bandbreite überschritten wurde.

Beispiel: Durch Verwendung einer Regel für gemeinsam genutzte Bandbreite nach einer Regel für reservierte Bandbreite können Sie Clients kontrolliert den Zugriff auf die drei Server gewähren. Solange gemeinsam genutzte Bandbreite verfügbar ist, wird die Regel mit dem Ergebnis true ausgewertet und der Zugriff gewährleistet. Ist die gemeinsam genutzte Bandbreite erschöpft, ist die Regel nicht true und es wird die nächste Regel ausgewertet. Folgt eine immer gültige Regel ("always true"), kann die Anfrage bei Bedarf umgeleitet werden.

Wenn Sie die reservierte und die gemeinsam genutzte Bandbreite wie im obigen Beispiel beschrieben verwenden, können Sie den Zugriff auf die Server geregelter und flexibler gewähren (oder verweigern). Für Server an einem spezifischen Port können Sie die nutzbare Bandbreite einschränken, während andere Server zusätzliche Bandbreite (im Rahmen der Verfügbarkeit) nutzen können.

Anmerkung: Der Dispatcher verfolgt die Bandbreite, indem er den zu einem Server fließenden Client-Datenverkehr (z. B. Bestätigungspakete) misst. Falls der Dispatcher diesen Datenverkehr aus irgend einem Grund nicht "sehen" kann, kommt es bei der Anwendung der Bandbreitenregeln zu unvorhersehbaren Ergebnissen.

Regel 'Metrik gesamt'

Dieser Regeltyp ist nur für die Komponente Site Selector verfügbar.

Für die Regel "Metrik gesamt" wählen Sie einen Systemmesswert (cpuload, memload oder ein eigenes angepasstes Script für Systemmesswerte) aus. Site Selector vergleicht den Systemmesswert (der vom Agenten Metric Server auf jedem Server mit Lastausgleich zurückgegeben wird) mit dem von Ihnen für die Regel festgelegten Bereichsanfang und -ende. Die Regel wird erst angewendet, wenn der aktuelle Messwert für alle Server der Gruppe innerhalb des für die Regel festgelegten Bereichs liegt.

Anmerkung: Das von Ihnen gewählte Script für Systemmesswerte muss sich auf jedem der Server mit Lastausgleich befinden.

Nachfolgend sehen Sie ein Beispiel für das Hinzufügen einer Regel "Metrik gesamt" zu Ihrer Konfiguration:

```
sscontrol rule add dnsload.com:allrule1 type metrical1  
metricname cpuload beginrange 0 endrange 100
```

Regel 'Metrik Durchschnitt'

Dieser Regeltyp ist nur für die Komponente Site Selector verfügbar.

Für die Regel "Metrik Durchschnitt" wählen Sie einen Systemmesswert (cpuload, memload oder ein eigenes angepasstes Script für Systemmesswerte) aus. Site Selector vergleicht den Systemmesswert (der vom Agenten Metric Server auf jedem Server mit Lastausgleich zurückgegeben wird) mit dem von Ihnen für die Regel festgelegten Bereichsanfang und -ende. Die Regel wird erst angewendet, wenn der *Durchschnitt* der aktuellen Messwerte auf allen Servern der Gruppe innerhalb des für die Regel festgelegten Bereichs liegt.

Anmerkung: Das von Ihnen gewählte Script für Systemmesswerte muss sich auf jedem der Server mit Lastausgleich befinden.

Nachfolgend sehen Sie ein Beispiel für das Hinzufügen einer Regel "Metrik Durchschnitt" zu Ihrer Konfiguration:

```
sscontrol rule add dnsload.com:avgrule1 type metricavg  
metricname cpuload beginrange 0 endrange 100
```

Immer gültige Regeln verwenden

Dieser Regeltyp ist für die Komponenten Dispatcher, CBR und Site Selector verfügbar.

Es kann eine Regel erstellt werden, die "immer wahr" ist. Eine solche Regel wird immer ausgewählt, es sei denn, alle ihr zugeordneten Server sind inaktiv. Aus diesem Grund sollte sie eine niedrigere Priorität als andere Regeln haben.

Sie können sogar mehrere Regeln haben, die "immer wahr" sind. Jeder Regel kann eine Gruppe mit Servern zugeordnet sein. Die erste wahre Regel mit einem verfügbaren Server wird ausgewählt. Angenommen, Sie haben sechs Server. Zwei dieser Server sollen unter allen Umständen den Datenaustausch steuern, es sei denn, beide Server sind inaktiv. Sind die ersten beiden Server inaktiv, soll eine zweite Gruppe mit Servern den Datenaustausch steuern. Sind diese vier Server inaktiv, sollen die letzten zwei Server den Datenaustausch steuern. Sie könnten drei Regeln erstellen, die "immer wahr" sind. Die erste Gruppe mit Servern wird dann immer ausgewählt, wenn mindestens ein Server aktiv ist. Sind beide inaktiv, wird ein Server aus der zweiten Gruppe ausgewählt, usw.

Als weiteres Beispiel können Sie mit einer Regel, die "immer wahr" ist, sicherstellen, dass eingehende Clients keinen Service erhalten, wenn sie nicht den festgelegten Regeln entsprechen. Mit Hilfe des Befehls **dscontrol rule** würden Sie die folgende Regel erstellen:

```
dscontrol rule add 130.40.52.153:80:jamais type true priority 100
```

Wenn Sie anschließend keine Server zur Regel hinzufügen, werden die Client-Pakete ohne Antwort gelöscht.

Anmerkung: Beim Erstellen einer immer gültigen Regel müssen Sie keinen Bereichsanfang und kein Bereichsende festlegen.

Sie können mehrere Regeln definieren, die "immer wahr" sind, und dann durch Ändern der Prioritätsebene festlegen, welche Regel angewendet werden soll.

Auf dem Inhalt der Anforderung basierende Regeln verwenden

Dieser Regeltyp ist in den Komponenten CBR und Dispatcher (bei Verwendung der Dispatcher-Weiterleitungsmethode cbr) verfügbar.

Dieser Regeltyp wird verwendet, wenn Anforderungen an Gruppen von Servern gesendet werden sollen, die speziell für die Bearbeitung eines bestimmten Teils des Sitedatenverkehrs konfiguriert wurden. Beispielsweise wollen Sie eine Gruppe von Servern für die Bearbeitung aller *cgi-bin*-Anforderungen, eine andere Gruppe für die Bearbeitung aller Audiodatenstromanforderungen und eine dritte Gruppe für die Bearbeitung aller anderen Anforderungen verwenden. Sie würden eine Regel mit einem pattern-Wert hinzufügen, der mit dem Pfad zu Ihrem cgi-bin-Verzeichnis übereinstimmt, eine zweite Regel, die mit dem Dateityp Ihrer Audio-Streaming-

Dateien übereinstimmt, und eine dritte Regel, die immer wahr ist, um den restlichen Datenverkehr zu bearbeiten. Sie würden dann jeder Regel die entsprechenden Server hinzufügen.

Wichtiger Hinweis: Beispiele und Szenarien für die Verwendung der content-Regel sowie eine gültige pattern-Syntax für die content-Regel finden Sie in Anhang B, „Syntax der content-Regel“, auf Seite 477.

Port-Affinität außer Kraft setzen

Mit der Außerkraftsetzung der Port-Affinität können Sie die Affinität eines Ports für einen bestimmten Server außer Kraft setzen. Angenommen, Sie verwenden eine Regel, um die Anzahl der Verbindungen mit jedem Anwendungsserver zu begrenzen, und haben einen Überlaufserver mit einer Regel 'immer wahr', die "Bitte später erneut versuchen" für diese Anwendung angibt. Der Port hat einen stickytime-Wert von 25 Minuten. Sie möchten also nicht, dass der Client an diesen Server gebunden wird. Durch Außerkraftsetzung der Port-Affinität können Sie bewirken, dass der Überlaufserver die diesem Port normalerweise zugeordnete Affinität außer Kraft setzt. Fordert der Client das nächste Mal den Cluster an, erfolgt ein Lastausgleich auf der Basis des besten verfügbaren Anwendungsservers und nicht des Überlaufservers.

Der Abschnitt „dscontrol server — Server konfigurieren“ auf Seite 394 enthält ausführliche Informationen zur Befehlssyntax für das Außerkraftsetzen der Port-Affinität mit der Serveroption **sticky**.

Regeln zur Konfiguration hinzufügen

Zum Hinzufügen von Regeln können Sie den Befehl **dscontrol rule add** verwenden, die Beispielkonfigurationsdatei editieren oder die grafische Benutzerschnittstelle (GUI) benutzen. Sie können für jeden definierten Port eine oder mehrere Regel(n) hinzufügen.

Der Prozess besteht aus zwei Schritten: Hinzufügen der Regel und Definieren der Server, die verwendet werden sollen, wenn die Regel wahr ist. Beispielsweise möchte der Systemadministrator die Auslastung der Proxy-Server durch die einzelnen Unternehmensbereiche verfolgen. Er ordnet jedem Unternehmensbereich IP-Adressen zu. Er erstellt die erste Gruppe mit Regeln auf der Basis der Client-IP-Adressen, um zwischen den Lasten der einzelnen Unternehmensbereiche unterscheiden zu können.

```
dscontrol rule add 130.40.52.153:80:div1 type ip b 9.1.0.0 e 9.1.255.255
dscontrol rule add 130.40.52.153:80:div2 type ip b 9.2.0.0 e 9.2.255.255
dscontrol rule add 130.40.52.153:80:div3 type ip b 9.3.0.0 e 9.3.255.255
```

Anschließend fügt der Systemadministrator zu jeder Regel einen anderen Server hinzu und misst dann die Last auf jedem der Server, um dem Unternehmensbereich die verwendeten Services korrekt in Rechnung zu stellen. Beispiel:

```
dscontrol rule useserver 130.40.52.153:80:div1 207.72.33.45
dscontrol rule useserver 130.40.52.153:80:div2 207.72.33.63
dscontrol rule useserver 130.40.52.153:80:div3 207.72.33.47
```

Regeloption für Serverauswertung

Die Option für Serverauswertung ist nur für die Komponente Dispatcher verfügbar.

Der Befehl **dscontrol rule** bietet eine Serverauswertungsoption für Regeln an. Mit der Option *evaluate* können Sie die Regelbedingungen für alle Server an einem Port

oder für die in der Regel angegebenen Server auswerten. (In früheren Versionen von Load Balancer konnten nur die Regelbedingungen für alle Server an einem Port erfasst werden.)

Anmerkungen:

1. Die Option für Serverauswertung ist nur für Regeln gültig, die ihre Entscheidungen ausgehend von den Kenndaten der Server treffen. Dazu gehören die Regel "Summe Verbindungen (pro Sekunde)", die Regel "Aktive Verbindungen" und die Regel "Reservierte Bandbreite".
2. Regeln des Typs connection bieten die zusätzliche Auswertungsoption **upserverversionrule** zur Auswahl an. Weitere Informationen hierzu finden Sie im Abschnitt „Regeln auf der Basis der Verbindungen pro Sekunde verwenden“ auf Seite 215.

Nachfolgend einige Beispiele für das Hinzufügen oder Definieren der Auswertungsoption für eine Regel "Reservierte Bandbreite":

```
dscontrol rule add 9.22.21.3:80:rbweval type reservedbandwidth evaluate Ebene
dscontrol rule set 9.22.21.3:80:rbweval evaluate Ebene
```

Die *Ebene* für die Option evaluate kann auf port, rule oder upserverversionrule gesetzt werden. Der Standardwert ist port.

In der Regel angegebene Server auswerten

Mit der Option zum Erfassen der Regelbedingungen für die in der Regel definierten Server können Sie zwei Regeln mit den folgenden Kenndaten konfigurieren:

- Die erste auszuwertende Regel enthält alle Server, die den Inhalt der Website verwalten. Die Option evaluate wird auf *rule* gesetzt (damit die Regelbedingungen für die in der Regel definierten Server ausgewertet werden).
- Die zweite Regel ist eine immer gültige Regel, in der nur ein Server definiert ist, der mit einer Antwort des Typs "Site ausgelastet" reagiert.

Wenn der Datenverkehr den Schwellenwert für die in der ersten Regel angegebenen Server überschreitet, wird er an den in der zweiten Regel definierten Server ("Site ausgelastet") gesendet. Sinkt die Zahl der Datenübertragungen unter den Schwellenwert, der für die Server in der ersten Regel definiert ist, werden die nachfolgenden Datenübertragungen erneut an die Server in der ersten Regel gesendet.

Server am Port auswerten

Wenn Sie bei den für das vorherige Beispiel beschriebenen Regeln den Wert der Option evaluate für die erste Regel auf *port* setzen (damit die Regelbedingungen für alle Server am Port ausgewertet werden) und der Datenverkehr den Schwellenwert für diese Regel überschreitet, wird er an den der zweiten Regel zugeordneten Server ("Site ausgelastet") gesendet.

Die erste Regel misst den Datenverkehr aller Server (einschließlich des Verkehrs für den Server "Site ausgelastet") am Port, um festzustellen, ob der Schwellenwert überschritten wird. Geht die Überlastung der der ersten Regel zugeordneten Server zurück, kann der Datenverkehr entgegen der Absicht weiterhin an den Server "Site ausgelastet" gesendet werden, sofern der Datenverkehr am Port weiterhin den Schwellenwert für die erste Regel überschreitet.

Funktionsweise der Affinität für Load Balancer

Für die Komponenten Dispatcher und CBR: Wenn Sie den Port eines Clusters als sticky konfigurieren, aktivieren Sie die Affinitätsfunktion. Wird der Port eines Clusters als sticky konfiguriert, können nachfolgende Client-Anforderungen an denselben Server übertragen werden. Dies geschieht, indem für die Option **stickytime** auf Executor-, Cluster- oder Port-Ebene eine Haltezeit von einigen Sekunden angegeben wird. Sie können die Funktion inaktivieren, indem Sie stickytime auf null setzen.

Wird die Port-übergreifende Affinität aktiviert, müssen die Werte für stickytime der gemeinsam benutzten Ports identisch (und ungleich null) sein. Weitere Informationen hierzu finden Sie im Abschnitt „Port-übergreifende Affinität“ auf Seite 223.

Für die Komponente Site Selector: Wenn Sie einen Sitenamen als sticky konfigurieren, aktivieren Sie die Affinitätsfunktion. Bei einem als sticky konfigurierten Sitenamen kann der Client für mehrere Namensserviceanforderungen denselben Server verwenden. Geben Sie dazu als **stickytime** für den Sitenamen eine Haltezeit von einigen Sekunden an. Sie können die Funktion inaktivieren, indem Sie stickytime auf null setzen.

Die Haltezeit (stickytime) für einen Server ist das Intervall zwischen dem Schließen einer Verbindung und dem Öffnen einer neuen Verbindung. Innerhalb dieses Intervalls wird der Client an denselben Server wie bei der ersten Verbindung vermittelt. Nach Ablauf der Haltezeit kann der Client an einen anderen Server vermittelt werden. Die Haltezeit für einen Server wird mit den Befehlen `dscontrol executor, port` oder `cluster` konfiguriert. Wenn Sie einen Server mit dem Befehl `"server down"` (`dscontrol server down`) offline setzen und der stickytime-Wert für diesen Server ungleich null ist, werden vorhandene Clients weiter von diesem Server bedient, bis das stickytime-Intervall abläuft. Der Server wird erst nach Ablauf der Haltezeit heruntergefahren.

Verhalten bei inaktivierter Affinität

Wird bei inaktivierter Affinität eine neue TCP-Verbindung von einem Client empfangen, verwendet Load Balancer den zu diesem Zeitpunkt richtigen Server und leitet die Pakete an diesen Server weiter. Wird eine weitere Verbindung von demselben Client empfangen, behandelt Load Balancer diese Verbindung als eine neue Verbindung und wählt wieder den zu diesem Zeitpunkt richtigen Server aus.

Verhalten bei aktivierter Affinität

Bei Aktivierung der Affinität wird eine nachfolgende Anforderung von demselben Client an denselben Server gerichtet.

Nach einer gewissen Zeit hört der Client auf, Transaktionen zu senden, so dass der Affinitätseintrag entfernt wird. Jeder Affinitätseintrag bleibt nur für die für "stickytime" festlegte Zeit in Sekunden erhalten. Werden innerhalb der Haltezeit (stickytime) weitere Verbindungen empfangen, ist der Affinitätseintrag noch gültig, so dass die Anforderung an denselben Server weitergeleitet wird. Wenn eine Verbindung nicht innerhalb der Haltezeit empfangen wird, wird der Eintrag gelöscht. Für eine nach Ablauf der Haltezeit empfangene Verbindung wird ein neuer Server ausgewählt.

Ein Server wird mit dem Befehl `"server down"` (`dscontrol server down`) offline gesetzt, allerdings erst heruntergefahren, wenn das stickytime-Intervall abgelaufen ist.

Port-übergreifende Affinität

Die Port-übergreifende Affinität gilt nur für die Dispatcher-Weiterleitungsmethoden MAC und NAT/NATP.

Die Port-übergreifende Affinität ist die Ausdehnung der Haltefunktion auf mehrere Ports. Wird beispielsweise eine Client-Anforderung zuerst an einem Port und die nächste Anforderung an einem anderen Port empfangen, kann der Dispatcher die Client-Anforderungen bei Port-übergreifender Affinität an denselben Server senden. Die Ports müssen die folgenden Bedingungen erfüllen, um diese Funktion verwenden zu können:

- Sie müssen dieselbe Clusteradresse gemeinsam benutzen.
- Sie müssen dieselben Server gemeinsam benutzen.
- Sie müssen denselben Wert (ungleich null) für **stickytime** haben. Typ
- Sie müssen denselben Wert für **stickymask** haben. Typ

Mehrere Ports können eine Verbindung zu einem **crossport** herstellen. Wenn vom selben Client weitere Verbindungen an demselben Port oder einem gemeinsam benutzten Port ankommen, wird auf denselben Server zugegriffen. Nachfolgend sehen Sie eine Beispielkonfiguration für mehrere Ports mit einer Port-übergreifenden Affinität für Port 10:

```
dscontrol port set Cluster:20 crossport 10
dscontrol port set Cluster:30 crossport 10
dscontrol port set Cluster:40 crossport 10
```

Nachdem Sie die Port-übergreifende Affinität konfiguriert haben, können Sie den Wert für stickytime des Ports flexibel ändern. Sie sollten stickytime jedoch für alle gemeinsam benutzten Ports auf denselben Wert setzen, da andernfalls unerwartete Ergebnisse auftreten können.

Wenn Sie die Port-übergreifende Affinität aufheben möchten, setzen Sie den Wert für crossport auf seine eigene Port-Nummer zurück. Der Abschnitt „dscontrol port — Ports konfigurieren“ auf Seite 382 enthält ausführliche Informationen zur Befehlssyntax für die Option **crossport**.

Affinitätsadressmaske (stickymask)

Die Affinitätsadressmaske gilt nur für die Komponente Dispatcher.

Die Affinitätsadressmaske ist eine Erweiterung der Sticky-Funktion, mit der Clients auf der Basis gemeinsamer Teilnetzadressen zusammengefasst werden. Die Angabe von **stickymask** im Befehl **dscontrol port** ermöglicht Ihnen, die gemeinsamen höherwertigen Bits der 32-Bit-IP-Adresse zu maskieren. Wenn diese Funktion konfiguriert ist und eine Client-Anforderung zum ersten Mal eine Verbindung zu dem Port herstellt, werden alle nachfolgenden Anforderungen von Clients mit derselben Teilnetzadresse (repräsentiert vom maskierten Abschnitt der Adresse) an denselben Server übertragen.

Anmerkung: Wenn Sie stickymask aktivieren möchten, muss **stickytime** für den Port einen Wert ungleich null haben.

Wenn Sie beispielsweise alle eingehenden Client-Anforderungen mit derselben Netzadresse der Klasse A an einen Server übergeben möchten, setzen Sie den stickymask-Wert für den Port auf 8 (Bits). Sollen Client-Anforderungen mit derselben Netzadresse der Klasse B zusammengefasst werden, setzen Sie den Wert für sticky-

mask auf 16 (Bits). Sollen Client-Anforderungen mit derselben Netzadresse der Klasse C zusammengefasst werden, setzen Sie den Wert für stickymask auf 24 (Bits).

Die besten Ergebnisse werden erzielt, wenn Sie den Wert für stickymask beim erstmaligen Starten von Load Balancer definieren. Wird der Wert für stickymask dynamisch geändert, können unvorhersehbare Ergebnisse auftreten.

Interaktion mit Port-übergreifender Affinität: Wenn Sie die Port-übergreifende Affinität aktivieren, müssen die Werte für stickymask der gemeinsam benutzten Ports identisch sein. Weitere Informationen hierzu finden Sie im Abschnitt „Port-übergreifende Affinität“ auf Seite 223.

Um die Affinitätsadressmaske zu aktivieren, setzen Sie einen ähnlichen Befehl **dscontrol port** wie den folgenden ab:

```
dscontrol port set Cluster:Port stickytime 10 stickymask 8
```

Gültige Werte für stickymask sind 8, 16, 24 und 32. Der Wert 8 gibt an, dass die ersten 8 höherwertigen Bits der IP-Adresse (Netzadresse der Klasse A) maskiert werden. Der Wert 16 gibt an, dass die ersten 16 höherwertigen Bits der IP-Adresse (Netzadresse der Klasse B) maskiert werden. Der Wert 24 gibt an, dass die ersten 24 höherwertigen Bits der IP-Adresse (Netzadresse der Klasse C) maskiert werden. Wird der Wert 32 angegeben, wird die gesamte IP-Adresse maskiert, wodurch die Affinitätsadressmaskenfunktion inaktiviert wird. Der Standardwert für stickymask ist 32.

Der Abschnitt „dscontrol port — Ports konfigurieren“ auf Seite 382 enthält ausführliche Informationen zur Befehlssyntax für stickymask (Affinitätsadressmaskenfunktion).

Bearbeitung von Serververbindungen stilllegen

Die Stilllegung gilt für die Komponenten Dispatcher und CBR.

Wenn Sie aus bestimmten Gründen (Aktualisierungen, Upgrades, Wartung usw.) einen Server aus der Load-Balancer-Konfiguration entfernen müssen, können Sie den Befehl **dscontrol manager quiesce** verwenden. Mit dem Unterbefehl quiesce können vorhandene Verbindungen beendet werden (ohne weiter bedient zu werden). Nachfolgende neue Verbindungen vom Client zum stillgelegten Server werden nur weitergeleitet, wenn die Verbindung als gehaltene Verbindung (sticky) bezeichnet ist und die Haltezeit (stickytime) nicht abgelaufen ist. Alle anderen neuen Verbindungen zum Server werden vom Unterbefehl quiesce unterbunden.

Stilllegung gehaltener Verbindungen

Verwenden Sie die Option quiesce “now”, wenn Sie die Haltezeit definiert haben und vor Ablauf der Haltezeit neue Verbindungen an einen anderen als den stillgelegten Server gesendet werden sollen. Im folgenden Beispiel wird die Option now für die Stilllegung des Servers 9.40.25.67 verwendet:

```
dscontrol manager quiesce 9.40.25.67 now
```

Die Option now bestimmt wie folgt, was mit gehaltenen Verbindungen geschehen soll:

- Wenn Sie *nicht* die Option “now” angeben, können vorhandene Verbindungen beendet werden. Nachfolgende neue Verbindungen zum stillgelegten Server werden weitergeleitet, sofern sie von Clients mit vorhandenen und als gehaltene Verbindungen bezeichneten Verbindungen stammen und der stillgelegte Server

die neue Anforderung vor Ablauf der Haltezeit empfängt. (Falls Sie das Merkmal "Haltezeit" (Affinität) jedoch nicht aktiviert haben, kann der stillgelegte Server keine neuen Verbindungen empfangen.)

Auf diese Weise können Server schrittweise stillgelegt werden. Sie können einen Server beispielsweise nach und nach stilllegen und dann auf den Zeitpunkt des geringsten Datenverkehrsaufkommens warten (vielleicht am frühen Morgen), um den Server vollständig aus der Konfiguration zu entfernen.

- Durch Angabe von "now" legen Sie den Server so still, dass vorhandene Verbindungen beendet werden können, alle neuen Verbindungen, einschließlich der nachfolgenden Verbindungen von Clients mit bereits vorhandenen gehaltenen Verbindungen, jedoch unterbunden werden. Diese abrupte Methode der Serverstilllegung war in früheren Versionen von Load Balancer die einzig mögliche Methode.

Affinitätsoption der Regel ausgehend vom Inhalt der Client-Anfrage

Mit dem Befehl **dscontrol rule** können Sie die folgenden Arten der Affinität angeben:

- Aktive Cookie-Affinität — Aktiviert die Verteilung von Webdatenverkehr mit Affinität zu einem Server ausgehend von den von Load Balancer generierten Cookies.

Die aktive Cookie-Affinität gilt nur für die Komponente CBR.

- Passive Cookie-Affinität — Aktiviert die Verteilung von Webdatenverkehr mit Affinität zu einem Server ausgehend von den Identifizierungs-Cookies, die von den Servern generiert werden. Bei Verwendung der passiven Cookie-Affinität müssen Sie den Befehl **rule** mit dem Parameter **cookie** angeben.

Die passive Cookie-Affinität gilt für die Komponente CBR sowie für die Weiterleitungsmethode **cbr** der Komponente Dispatcher.

- URI-Affinität — Aktiviert den Lastausgleich für Webdatenverkehr auf Caching-Proxy-Servern mit effektiver Erhöhung der Kapazität des Cache.

Die URI-Affinität gilt für die Komponente CBR sowie für die Weiterleitungsmethode **cbr** der Komponente Dispatcher.

Der Standardwert für die Option **affinity** ist **none**. Die Option **stickytime** für den Port-Befehl (**port**) muss auf null gesetzt (inaktiviert) sein, damit die Option **affinity** des Regelbefehls (**rule**) auf die aktive oder passive Cookie-Affinität bzw. auf die URI-Affinität gesetzt werden kann. Ist für die Regel eine Affinität definiert, kann keine Haltezeit für den Port aktiviert werden.

Aktive Cookie-Affinität

Die aktive Cookie-Affinität gilt nur für die Komponente CBR.

Sie bietet eine Möglichkeit, Clients an einen bestimmten Server zu "binden". Diese Funktion wird aktiviert, indem der Wert **stickytime** einer Regel auf eine positive Zahl und die Affinität auf "activecookie" gesetzt wird. Dies kann beim Hinzufügen der Regel oder mit dem Befehl **rule set** geschehen. Ausführliche Informationen zur Befehlssyntax finden Sie im Abschnitt „dscontrol rule — Regeln konfigurieren“ auf Seite 388.

Wenn eine Regel für aktive Cookie-Affinität aktiviert wurde, wird der Lastausgleich für neue Client-Anforderungen mit Standard-CBR-Algorithmen durchgeführt. Aufeinanderfolgende Anforderungen eines Clients werden dabei an den zu Beginn ausgewählten Server gesendet. Der ausgewählte Server ist als Cookie in der

Antwort an den Client gespeichert. Solange die zukünftigen Anforderungen des Clients das Cookie enthalten und jede Anforderung innerhalb der Haltezeit empfangen wird, bleibt der Client an den anfänglichen Server gebunden.

Mit der aktiven Cookie-Affinität wird sichergestellt, dass die Arbeitslast eines Clients über einen bestimmten Zeitraum hinweg an denselben Server weitergeleitet wird. Dies wird erreicht, indem ein Cookie gesendet wird, das von dem Client-Browser gespeichert wird. Das Cookie enthält die für die Entscheidungsfindung verwendete Angabe Cluster:Port:Regel, den Server, an den die Arbeitslast weitergeleitet wurde, und eine Zeitmarke für das Zeitlimit, bei dessen Erreichung die Affinität ungültig wird. Das Cookie hat das folgende Format:

IBMCCR=Cluster:Port:Regel+Server-Zeit! Die Angaben *Cluster:Port:Regel* und *Server* sind codiert, so dass die CBR-Konfiguration nicht erkennbar ist.

Funktionsweise der aktiven Cookie-Affinität

Bei Erfüllung einer Regel mit gesetzter aktiver Cookie-Affinität wird das vom Client gesendete Cookie überprüft.

- Wenn ein Cookie mit der Kennung für die Cluster:Port:Regel-Kombination gefunden wird, die die Regel erfüllt, werden der Server, an den die Arbeitslast weitergeleitet werden soll, und die Zeitmarke für das Zeitlimit aus dem Cookie extrahiert.
- Wenn der Server noch zu der von der Regel verwendeten Gruppe gehört, seine Wertigkeit positiv ist oder der Server ein stillgelegter Server ist und die Zeitmarke für den Verfall einen späteren Zeitpunkt als die aktuelle Zeit angibt, wird der Server in dem Cookie für den Lastausgleich ausgewählt.
- Ist eine der oben aufgeführten Bedingungen nicht erfüllt, wird ein Server unter Verwendung des normalen Algorithmus ausgewählt.
- Nach Auswahl eines Servers (mit einer der beiden Methoden) wird ein neues Cookie erstellt, das IBMCCR, die Angabe Cluster:Port:Regel, ausgewählterServer und eine Zeitmarke enthält. Die Zeitmarke gibt die Uhrzeit an, zu der die Affinität ungültig wird. Die Angaben "Cluster:Port:Regel" und "ausgewählterServer" werden codiert, so dass keine Informationen zur CBR-Konfiguration erkennbar sind.
- Ein Parameter "expires" wird auch in das Cookie eingefügt. Dieser Parameter hat ein Format, das der Browser verstehen kann, und bewirkt, dass das Cookie sieben Tage nach Erreichen der Zeitmarke für den Verfall ungültig wird. Damit soll vermieden werden, dass die Cookie-Datenbank des Clients zu sehr anwächst.

Dieses neue Cookie wird dann in die Kopfzeilen eingefügt, die an den Client gesendet werden. Ist der Browser des Clients so konfiguriert, dass er Cookies akzeptiert, sendet er nachfolgende Anforderungen an diese Adresse zurück.

Jede Affinitätsinstanz im Cookie ist 65 Bytes lang und endet mit dem Ausrufezeichen. Ein 4096-Byte-Cookie kann demzufolge ungefähr 60 individuelle aktive Cookie-Regeln pro Domäne enthalten. Wenn das Cookie komplett gefüllt ist, werden alle abgelaufenen Affinitätsinstanzen gelöscht. Sollten noch alle Instanzen gültig sein, wird die älteste gelöscht und dafür die neue Instanz für die aktuelle Regel hinzugefügt.

Anmerkung: CBR ersetzt alle IBMCCR-Cookies im alten Format auf dem Proxy.

Die Option für aktive Cookie-Affinität für den Regelbefehl (rule) kann nur auf activecookie gesetzt werden, wenn die Haltezeit (stickytime) für den Port gleich null (inaktiviert) ist. Ist die aktive Cookie-Affinität für eine Regel aktiviert, kann keine Haltezeit für den Port aktiviert werden.

Aktive Cookie-Affinität aktivieren

Verwenden Sie zum Aktivieren der aktiven Cookie-Affinität für eine bestimmte Regel wie folgt den Befehl rule set:

```
rule set Cluster:Port:Regel stickytime 60
rule set Cluster:Port:Regel affinity activecookie
```

Grund für die Verwendung der aktiven Cookie-Affinität

Die Haltezeit wird in der Regel für CGIs oder Servlets verwendet, die den Client-Status auf dem Server speichern. Der Status wird durch eine Cookie-ID identifiziert (dies sind Server-Cookies). Der Client-Status ist nur auf dem ausgewählten Server gespeichert. Der Client benötigt also das Cookie von diesem Server, um diesen Status zwischen Anforderungen zu wahren.

Außerkraftsetzung der Verfallszeit für Cookie-Affinität

Die aktive Cookie-Affinität verfällt standardmäßig nach der Zeit, die sich aus der Summe der aktuellen Serverzeit, der Haltezeit (stickytime) und 24 Stunden ergibt. Wenn Ihre Clients (die Anfragen an Ihre CBR-Maschine senden) auf ihren Systemen eine falsche Zeit eingestellt haben (so dass sie z. B. der Serverzeit mehr als einen Tag voraus sind), ignorieren die Systeme dieser Clients die Cookies von CBR, weil sie davon ausgehen, dass die Cookies schon verfallen sind. Zum Einstellen einer längeren Verfallszeit müssen Sie das Script cbrserver modifizieren. Editieren Sie die Zeile javaw in der Script-Datei. Fügen Sie nach LB_SERVER_KEYS den folgenden Parameter ein: -DCOOKIEEXPIREINTERVAL=X. X steht hier für die Anzahl der Tage, die zur Verfallszeit addiert werden sollen.

Auf AIX-, Solaris- und Linux-Systemen finden Sie die Datei cbrserver im Verzeichnis /usr/bin.

Auf Windows-Systemen finden Sie die Datei cbrserver im Verzeichnis \winnt\system32.

Passive Cookie-Affinität

Die passive Cookie-Affinität gilt für die inhaltsabhängige Weiterleitung (cbr) durch die Komponente Dispatcher und die Komponente CBR. Informationen zum Konfigurieren der Dispatcher-Weiterleitungsmethode cbr finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Die passive Cookie-Affinität bietet eine Möglichkeit, Clients an einen bestimmten Server zu binden. Wenn Sie für eine Regel die Affinität auf passivecookie setzen, können Sie den Webdatenverkehr mit Affinität zu einem Server verteilen. Die Affinität basiert auf den von den Servern generierten Identifizierungs-Cookies. Die passive Cookie-Affinität wird auf Regelebene konfiguriert.

Wird eine Regel mit aktivierter passiver Cookie-Affinität erfüllt, wählt Load Balancer den Server ausgehend von dem im HTTP-Header der Client-Anforderung enthaltenen Cookie-Namen aus. Load Balancer vergleicht den Cookie-Namen aus dem HTTP-Header des Clients mit den für die einzelnen Server konfigurierten Cookie-Werten.

Findet Load Balancer einen Server, in dessen Cookie-Wert der Cookie-Name des Clients *enthalten* ist, wählt Load Balancer diesen Server für die Anforderung aus.

Anmerkung: Load Balancer bietet diese Flexibilität für den Fall, dass der Server einen Cookie-Wert generiert, der aus einem statischen Teil besteht, an den ein variabler Teil angehängt wird. Der Cookie-Wert des Servers könnte beispielsweise der Servername (statischer Wert) sein, an den eine Zeitmarke (variabler Wert) angehängt wird.

Wenn in der Client-Anforderung kein Cookie-Name gefunden wird oder dieser in keinem der Server-Cookie-Werte enthalten ist, wird der Server mit Hilfe der vorhandenen Serverauswahlmethoden oder der gewichteten RoundRobin-Methode ausgewählt.

Gehen Sie zum Konfigurieren der **passiven Cookie-Affinität** wie folgt vor:

- Für den Dispatcher müssen Sie zunächst die Weiterleitungsmethode cbr konfigurieren. (Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.) Für die Komponente CBR ist dieser Schritt nicht erforderlich.
- Setzen Sie den Parameter **affinity** des Befehls **dscontrol rule [add|set]** auf **passivecookie**. Der Parameter **cookieName** muss auf den Namen des Cookies gesetzt werden, nach dem Load Balancer im HTTP-Header der Client-Anforderung suchen soll.
- Legen Sie für jeden in der Regel definierten Server den Parameter **cookievalue** des Befehls **dscontrol server [add|set]** fest.

Die Option für passive Cookie-Affinität für den Regelbefehl (rule) kann nur auf **passivecookie** gesetzt werden, wenn die Haltezeit (stickytime) für den Port gleich null (inaktiviert) ist. Ist die passive Cookie-Affinität für eine Regel aktiviert, kann keine Haltezeit für den Port aktiviert werden.

URI-Affinität

Die URI-Affinität gilt für die Weiterleitungsmethode cbr von Dispatcher und die Komponente CBR. Informationen zum Konfigurieren der Weiterleitungsmethode cbr finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Bei Verwendung der URI-Affinität kann die Arbeitslast des Webdatenverkehrs so auf Caching-Proxy-Server verteilt werden, dass auf den einzelnen Servern unterschiedlicher Inhalt im Cache gespeichert werden kann. Auf diese Weise vergrößern Sie effektiv die Cache-Kapazität Ihrer Site, da eine redundante Zwischenspeicherung von Inhalten auf mehreren Maschinen vermieden wird. Konfigurieren Sie die URI-Affinität auf Regelebene. Wenn eine Regel mit aktivierter URI-Affinität erfüllt ist und die entsprechende Gruppe von Servern verfügbar und aktiv ist, leitet Load Balancer neue eingehende Client-Anforderungen mit demselben URI an einen Server weiter.

Normalerweise verteilt Load Balancer Anforderungen auf mehrere Server, die identische Inhalte bereitstellen. Wenn Sie Load Balancer mit einer Gruppe von Caching-Servern verwenden, wird häufig abgerufener Inhalt unter Umständen auf allen Servern zwischengespeichert. Daraus ergibt sich eine sehr hohe Client-Belastung, wenn auf mehreren Maschinen zwischengespeicherte identische Inhalte repliziert werden. Diese Vorgehensweise ist besonders für Websites mit großem Datenvolumen sinnvoll.

Wenn Ihre Website jedoch nur ein mittleres Client-Datenvolumen mit den verschiedensten Inhalten unterstützt und Sie einen großen, auf mehrere Server verteilten Cache bevorzugen, ist der Durchsatz Ihrer Site besser, wenn jeder Caching Server eindeutige Inhalte enthält und Load Balancer die Anforderungen nur an den Caching Server mit den entsprechenden Inhalten weiterleitet.

Bei Verwendung der URI-Affinität können Sie mit Load Balancer den zwischengespeicherten Inhalt auf einzelne Server verteilen und so eine redundante Zwischenspeicherung von Inhalten auf mehreren Maschinen vermeiden. Durch diese Erweiterung kann der Durchsatz von Serversites mit vielfältigen Inhalten, die Caching-Proxy-Server verwenden, verbessert werden. Identische Anforderungen werden an einen Server gesendet, so dass der Inhalt nur auf einem Server zwischengespeichert wird. Mit jeder zum Pool hinzugefügten Servermaschine vergrößert sich der effektive Cache.

Gehen Sie zum Konfigurieren der **URI-Affinität** wie folgt vor:

- Für den Dispatcher müssen Sie zunächst die Weiterleitungsmethode cbr konfigurieren. (Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.) Für die Komponente CBR ist dieser Schritt nicht erforderlich.
- Setzen Sie den Parameter **affinity** des Befehls **dscontrol rule [add|set]** oder **cbrcontrol rule [add|set]** auf uri.

Die Option für URI-Affinität für den Regelbefehl (rule) kann nur auf URI gesetzt werden, wenn die Haltezeit (stickytime) für den Port gleich null (inaktiviert) ist. Ist die URI-Affinität für eine Regel aktiviert, kann keine Haltezeit für den Port aktiviert werden.

Dispatcher-WAN-Unterstützung konfigurieren

Diese Funktion ist nur für die Komponente Dispatcher verfügbar.

Wenn Sie die WAN-Unterstützung und die Weiterleitungsmethode nat von Dispatcher nicht verwenden, erfordert die Dispatcher-Konfiguration, dass die Dispatcher-Maschine und die zugehörigen Server demselben LAN-Segment zugeordnet sind (siehe Abb. 35). Eine Client-Anfrage wird auf der Dispatcher-Maschine empfangen und an den Server gesendet. Der Server sendet die Antwort direkt zurück an den Client.

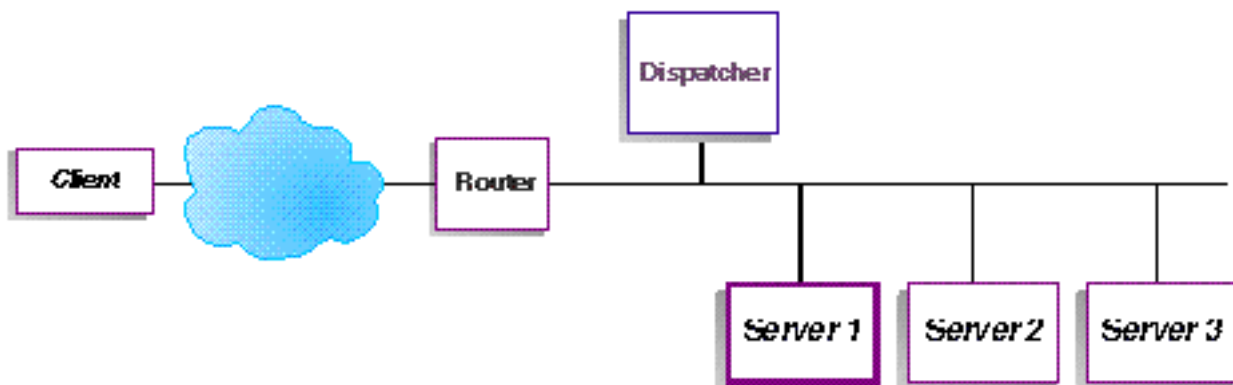


Abbildung 35. Beispiel einer Konfiguration mit einem LAN-Segment

Durch das WAN-Feature von Dispatcher werden Server an anderen Standorten, die als *ferne Server* bezeichnet werden, unterstützt (siehe Abb. 36). Wenn GRE am fernen Standort nicht unterstützt wird und Sie nicht die Dispatcher-Weiterleitungsmethode nat verwenden, muss der ferne Standort aus einer Dispatcher-Maschine (Dispatcher 2) und den lokal angeschlossenen Servern (ServerG, ServerH und ServerI) bestehen. Ein Client-Paket wird vom Internet an die erste Dispatcher-Maschine und anschließend von dieser Maschine an eine Dispatcher-Maschine an einem anderen geografischen Standort sowie an einen der dort lokal angeschlossenen Server gesendet.

Für WAN-Konfigurationen müssen alle Dispatcher-Maschinen (lokale und ferne) ein Betriebssystem desselben Typs ausführen und eine gemeinsame Plattform haben.

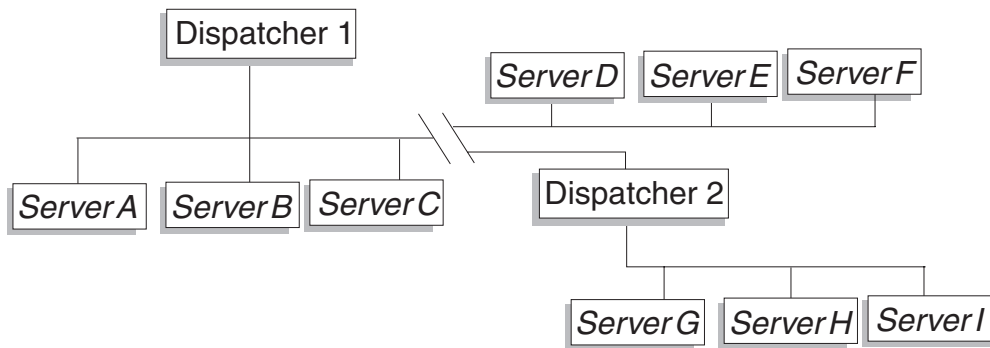


Abbildung 36. Beispiel einer Konfiguration mit lokalen und fernen Servern

Damit kann eine Clusteradresse weltweit alle Client-Anforderungen unterstützen und die Last auf Server auf der ganzen Welt verteilen.

An die Dispatcher-Maschine, die das Paket zunächst empfängt, können weiterhin lokale Server angeschlossen sein, und die Dispatcher-Maschine kann die Last auf ihre lokalen Server und auf die fernen Server verteilen.

Befehlssyntax

Führen Sie folgende Schritte aus, um die Weitverkehrsunterstützung zu konfigurieren:

1. Fügen Sie die Server hinzu. Wird einem Dispatcher ein Server hinzugefügt, müssen Sie definieren, ob es sich bei dem Server um einen lokalen Server oder einen fernen Server handelt (siehe oben). Soll ein Server hinzugefügt und der Server als lokaler Server definiert werden, geben Sie den Befehl **dscontrol server add** ohne Angabe eines Routers ein. Dieser Wert ist der Standardwert. Soll der Server als ferner Server definiert werden, müssen Sie den Router angeben, über den der Dispatcher das Paket senden muss, um den fernen Server zu erreichen. Der Server muss ein anderer Dispatcher sein, und die Adresse des Servers muss die NFA des Dispatchers sein. Wird beispielsweise in Abb. 37 auf Seite 233 *LB 2* als ferner Server unter *LB 1* hinzugefügt, müssen Sie *Router 1* als Router-Adresse definieren. Allgemeine Syntax:

```
dscontrol server add Cluster:Port:Server router Adresse
```

Weitere Informationen zum Schlüsselwort `router` finden Sie im Abschnitt „dscontrol server — Server konfigurieren“ auf Seite 394.

2. Konfigurieren Sie Aliasnamen. Auf der ersten Dispatcher-Maschine (auf der die Client-Anforderung aus dem Internet empfangen wird) muss für die Clusteradresse mit dem Befehl **executor configure** ein Aliasname definiert werden. (Für Linux- oder UNIX-Systeme können Sie den Befehl **executor configure** oder **ifconfig** verwenden.) Auf den fernen Dispatcher-Maschinen wird die Clusteradresse jedoch *nicht* als Aliasname auf einer Netzschnittstellenkarte definiert.

Ferne Advisor-Funktionen mit der Dispatcher-WAN-Unterstützung verwenden

Eingangspunkt-Dispatcher:

Eingangspunkt-Dispatcher, die auf der Plattform AIX, LINUX (mit GRE) oder Solaris ausgeführt werden, zeigen die Advisor-Last korrekt an. Andere Plattformen sind in Ermangelung eines Weitverkehrsnetzes auf einen RoundRobin-Lastausgleich bzw. auf die Dispatcher-Weiterleitungsmethode nat oder cbr angewiesen.

AIX-Systeme

- Es sind keine speziellen Konfigurationsschritte erforderlich.

HP-UX-Systeme

- Wenn Sie in einer WAN-Konfiguration einen Eingangspunkt-Dispatcher auf einer HP-UX-Plattform benutzen, gilt eine Einschränkung für die Verwendung ferner Advisor-Funktionen. Mit der Dispatcher-Weiterleitungsmethode mac werden sich Advisor-Funktionen unter HP-UX immer direkt an die Serveradresse und nicht an den Cluster wenden. Da sie den Cluster nicht adressieren, verteilt der ferne Dispatcher die Last der Advisor-Anfrage nicht auf die fernen Server. Bei Verwendung der Dispatcher-Weiterleitungsmethoden cbr und nat arbeiten die fernen Advisor-Funktionen jedoch ordnungsgemäß.

Linux-Systeme

- Wenn Sie in einer WAN-Konfiguration einen Eingangspunkt-Dispatcher auf einer Linux-Plattform benutzen, gilt eine Einschränkung für die Verwendung ferner Advisor-Funktionen. Mit der Dispatcher-Weiterleitungsmethode mac werden sich Advisor-Funktionen unter Linux immer direkt an die Serveradresse und nicht an den Cluster wenden. Da sie den Cluster nicht adressieren, verteilt der ferne Dispatcher die Last der Advisor-Anfrage nicht auf die fernen Server. Bei Verwendung der Dispatcher-Weiterleitungsmethoden cbr und nat arbeiten die fernen Advisor-Funktionen jedoch ordnungsgemäß.
- Wenn Sie Datenverkehr mit Generic Routing Encapsulation (GRE) an einen fernen Server senden und Ihre Konfiguration keinen fernen Dispatcher enthält, gelten bei Verwendung der Dispatcher-Weiterleitungsmethoden mac, nat und cbr auf einer Linux-Plattform keine Einschränkungen für Advisor-Funktionen. Weitere Informationen hierzu finden Sie im Abschnitt „Unterstützung für GRE (Generic Routing Encapsulation)“ auf Seite 235.

Solaris-Systeme

- Wenn Sie in einer WAN-Konfiguration einen Eingangspunkt-Dispatcher auf einer Solaris-Plattform verwenden, müssen Sie an Stelle der Konfigurationsmethoden ifconfig oder dscontrol executor die Konfigurationsmethode arp verwenden. Beispiel:

```
arp -s meine_Clusteradresse meine_MAC-Adresse pub
```

- Für die Solaris-Plattform gelten die folgenden Einschränkungen:
 - WAN-Advisor-Funktionen können nur mit der Clusterkonfigurationsmethode arp ausgeführt werden.
 - Advisor-Funktionen für bindungsspezifische Server können nur mit der Clusterkonfigurationsmethode arp ausgeführt werden.
 - Advisor-Funktionen für bindungsspezifische Server können nur mit der Clusterkonfigurationsmethode arp ausgeführt werden. Wenn Sie Advisor-Funktionen für bindungsspezifische Server verwenden, verknüpfen Sie Load Balancer nicht auf demselben Server mit der bindungsspezifischen Anwendung.

Windows-Systeme

- Wenn Sie in einer WAN-Konfiguration einen Eingangspunkt-Dispatcher auf einer Windows-Plattform benutzen, gilt eine Einschränkung für die Verwendung ferner Advisor-Funktionen. Mit der Dispatcher-Weiterleitungsmethode mac werden sich Advisor-Funktionen unter Windows immer direkt an die Serveradresse und nicht an den Cluster wenden. Da sie den Cluster nicht adressieren, verteilt der ferne Dispatcher die Last der Advisor-Anfrage nicht auf die fernen Server. Bei Verwendung der Dispatcher-Weiterleitungsmethoden cbr und nat arbeiten die fernen Advisor-Funktionen jedoch ordnungsgemäß.

Ferne Dispatcher: Führen Sie für jede ferne Clusteradresse die folgenden Konfigurationsschritte aus. Für eine Konfiguration mit hoher Verfügbarkeit am fernen Dispatcher-Standort müssen Sie diese Schritte auf beiden Maschinen ausführen.

AIX-Systeme

- Für Dispatcher ist es erforderlich, dass jeder Cluster auf der Schnittstelle mit der Netzmaske 255.255.255.255 konfiguriert sein, da sonst die Advisor-Funktionen nicht ordnungsgemäß arbeiten. Verwenden Sie für das Konfigurieren eines Clusters eines der folgenden Syntaxformate:
 - `ifconfig Schnittstellename alias Clusteradresse netmask 255.255.255.255`. Beispiel:
`ifconfig en0 alias 10.10.10.99 netmask 255.255.255.255`
 - `dscontrol executor configure Schnittstellenadresse Schnittstellename Netzmaske`. Beispiel:
`dscontrol executor configure 204.67.172.72 en0 255.255.255.255`

Anmerkung: Sie benötigen Advisor-Funktionen, die sowohl auf der lokalen als auch auf der fernen Dispatcher-Maschine ausgeführt werden.

HP-UX-Systeme, Linux-, Solaris- und Windows-Systeme

- Es sind keine zusätzlichen Konfigurationsschritte erforderlich.

Konfigurationsbeispiel

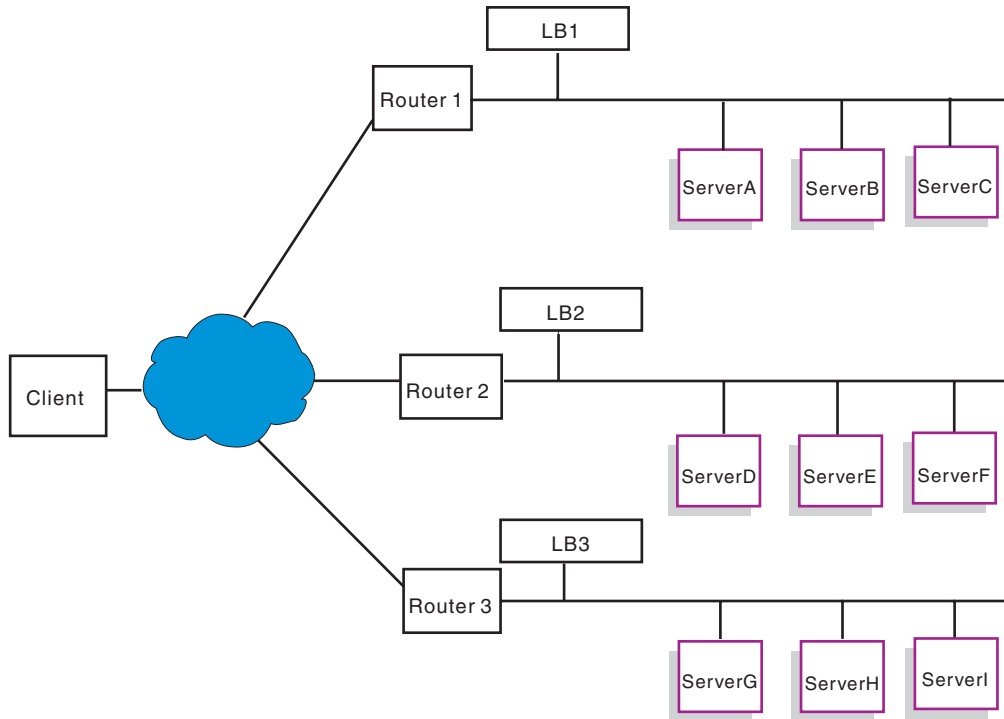


Abbildung 37. WAN-Beispielkonfiguration mit fernen Load-Balancer-Maschinen

Dieses Beispiel bezieht sich auf die in Abb. 37 gezeigte Konfiguration.

Nachfolgend wird beschrieben, wie die Dispatcher-Maschinen für die Unterstützung der Clusteradresse xebec am Port 80 konfiguriert werden. LB1 ist als „Eingangspunkt“-Load-Balancer definiert. Es wird eine Ethernet-Verbindung vorausgesetzt. Für LB1 sind fünf Server definiert, drei lokale (ServerA, ServerB, ServerC) und zwei ferne (LB2 und LB3). Für die fernen Load Balancer LB2 und LB3 sind jeweils drei lokale Server definiert.

Führen Sie an der Konsole des ersten Dispatchers (LB1) die folgenden Schritte aus:

1. Starten Sie den Executor.
dscontrol executor start
2. Definieren Sie die NFA der Dispatcher-Maschine.
dscontrol executor set nfa LB1
3. Definieren Sie den Cluster.
dscontrol cluster add xebec
4. Definieren Sie den Port.
dscontrol port add xebec:80
5. Definieren Sie die Server.
 - a. **dscontrol server add xebec:80:ServerA**
 - b. **dscontrol server add xebec:80:ServerB**
 - c. **dscontrol server add xebec:80:ServerC**
 - d. **dscontrol server add xebec:80:LB2 router Router1**
 - e. **dscontrol server add xebec:80:LB3 router Router1**

6. Konfigurieren Sie die Clusteradresse.
dscontrol executor configure xebec

An der Konsole des zweiten Dispatchers (LB2):

1. Starten Sie den Executor.
dscontrol executor start
2. Definieren Sie die NFA der Dispatcher-Maschine.
dscontrol executor set nfa LB2
3. Definieren Sie den Cluster.
dscontrol cluster add xebec
4. Definieren Sie den Port.
dscontrol port add xebec:80
5. Definieren Sie die Server.
 - a. **dscontrol server add xebec:80:ServerD**
 - b. **dscontrol server add xebec:80:ServerE**
 - c. **dscontrol server add xebec:80:ServerF**

An der Konsole des dritten Dispatchers (LB3):

1. Starten Sie den Executor.
dscontrol executor start
2. Definieren Sie die NFA der Dispatcher-Maschine.
dscontrol executor set nfa LB3
3. Definieren Sie den Cluster.
dscontrol cluster add xebec
4. Definieren Sie den Port.
dscontrol port add xebec:80
5. Definieren Sie die Server.
 - a. **dscontrol server add xebec:80:ServerG**
 - b. **dscontrol server add xebec:80:ServerH**
 - c. **dscontrol server add xebec:80:ServerI**

Anmerkungen

1. Geben Sie auf den Loopback-Einheiten aller Servern (A-I) einen Aliasnamen für die Clusteradresse an.
2. Cluster und Ports werden mit **dscontrol** auf allen beteiligten Dispatcher-Maschinen hinzugefügt. Dies gilt für den Dispatcher, der als Eingangspunkt definiert ist, und für alle fernen Dispatcher.
3. Der Abschnitt „Ferne Advisor-Funktionen mit der Dispatcher-WAN-Unterstützung verwenden“ auf Seite 231 enthält Informationen zur Verwendung ferner Advisor-Funktionen mit Weitverkehrsunterstützung.
4. Die Weitverkehrsunterstützung verbietet unendliche Routenschleifen. (Wenn eine Dispatcher-Maschine ein Paket von einem anderen Dispatcher empfängt, wird das Paket nicht an einen dritten Dispatcher weitergeleitet.) Mit der Weitverkehrsunterstützung wird nur eine Ebene von fernen Dispatchern unterstützt.
5. Die Weitverkehrsunterstützung impliziert Unterstützung für UDP und TCP.
6. Die Weitverkehrsunterstützung kann zusammen mit der Funktion für hohe Verfügbarkeit verwendet werden. Jedem Dispatcher kann eine benachbarte Bereitschaftsmaschine (im selben LAN-Segment) zugeordnet werden.

7. Der Manager und die Advisor-Funktionen können zusammen mit der Weitverkehrsunterstützung verwendet werden. In diesem Fall sollten sie auf allen beteiligten Dispatcher-Maschinen gestartet werden.
8. Load Balancer unterstützt WAN nur bei vergleichbaren Betriebssystemen.

Unterstützung für GRE (Generic Routing Encapsulation)

Generic Routing Encapsulation (GRE) ist ein Internet-Protokoll, das in RFC 1701 und RFC 1702 spezifiziert ist. Bei Verwendung von GRE kann Load Balancer Client-IP-Pakete in IP/GRE-Pakete integrieren und an Serverplattformen mit GRE-Unterstützung wie OS/390 weiterleiten. Mit der GRE-Unterstützung kann die Komponente Dispatcher die Arbeitslast von Paketen auf mehrere Serveradressen verteilen, die einer MAC-Adresse zugeordnet sind.

Load Balancer implementiert GRE im Rahmen der WAN-Funktion. Auf diese Weise stellt Load Balancer WAN-Lastausgleich direkt für alle Serversysteme zur Verfügung, die GRE-Pakete entpacken können. Load Balancer muss nicht auf einem fernen System installiert sein, wenn die fernen Server eingebundene GRE-Pakete unterstützen. Load Balancer integriert WAN-Pakete, deren GRE-Feld auf den Dezimalwert 3735928559 gesetzt ist.

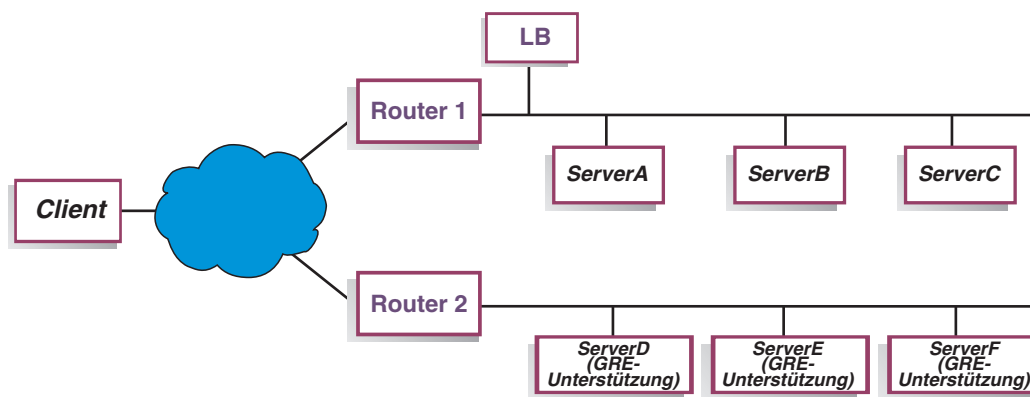


Abbildung 38. WAN-Beispielkonfiguration mit einer Serverplattform, die GRE unterstützt

Wenn Sie für dieses Beispiel (Abb. 38) einen fernen ServerD mit GRE-Unterstützung hinzufügen möchten, müssen Sie ihn in Ihrer Load-Balancer-Konfiguration definieren, wie Sie einen WAN-Server in der Hierarchie Cluster:Port:Server definieren würden:

```
dscontrol server add Cluster:Port:ServerD router Router1
```

GRE für WAN auf Linux-Systemen konfigurieren

Linux-Systeme haben native GRE-Fähigkeiten, so dass Load Balancer für Serverimages von Linux für S/390 einen Lastausgleich durchführen kann, wenn mehrere Serverimages eine MAC-Adresse gemeinsam benutzen. Der Eingangspunkt-Load-Balancer kann somit die Last direkt auf Linux-WAN-Server verteilen und ist nicht auf einen Load Balancer am fernen Standort angewiesen. Die Advisor-Funktionen des Eingangspunkt-Load-Balancer können ebenfalls direkt mit jedem der fernen Server zusammenarbeiten.

Erstellen Sie auf dem Eingangspunkt-Load-Balancer wie beschrieben eine Konfiguration für WAN.

Zum Konfigurieren der einzelnen Linux-Back-End-Server müssen Sie als root die folgenden Befehle absetzen. (Diese Befehle können zum Tool für Systemstart hinzugefügt werden, so dass die Änderungen für alle folgenden Bootvorgänge erhalten bleiben.)

```
# modprobe ip_gre
# ip tunnel add gre-nd mode gre ikey 3735928559
# ip link set gre-nd up
# ip addr add Clusteradresse dev gre-nd
```

Anmerkung: Der mit diesen Anweisungen konfigurierte Linux-Server *darf sich nicht* in demselben physischen Segment wie der Eingangspunkt-Load-Balancer befinden. Dies liegt daran, dass der Linux-Server auf Anfragen nach der Clusteradresse ("ARP who-has") antwortet, was zu einer Konkurrenzbedingung und damit zu einem möglichen "Kurzschluss" führt, so dass der gesamte an die Clusteradresse gerichtete Verkehr nur zum Gewinner der ARP-Konkurrenz übertragen wird.

Explizite Verbindungen benutzen

Normalerweise sind die Lastausgleichsfunktionen des Dispatchers unabhängig vom Inhalt der Sites, auf denen das Produkt benutzt wird. In einem bestimmten Bereich kann der Inhalt der Site jedoch von Bedeutung sein und können Entscheidungen über den Inhalt erhebliche Auswirkungen auf die Effektivität des Dispatchers haben. Dies ist der Bereich der Verbindungsadressierung.

Wenn Ihre Seiten Links enthalten, die auf einzelne Server für Ihre Site zeigen, zwingen Sie einen Client, auf eine bestimmte Maschine zuzugreifen und so die sonst wirksame Lastausgleichsfunktion zu umgehen. Benutzen Sie aus diesem Grund für alle Links Ihrer Seiten immer die Adresse des Dispatchers. Berücksichtigen Sie, dass die Art der verwendeten Adressierung nicht immer offensichtlich ist, wenn Ihre Site eine automatisierte Programmierung benutzt, bei der HTML dynamisch erstellt wird. Um den Lastausgleich zu optimieren, sollten Sie auf alle expliziten Adressierungen achten und sie, falls möglich, vermeiden.

Konfiguration für ein privates Netz verwenden

Sie können den Dispatcher und die TCP-Servermaschinen für ein privates Netz konfigurieren. Durch diese Konfiguration können Konkurrenzsituationen im öffentlichen oder externen Netz, die sich auf die Leistung auswirken, verringert werden.

Auf AIX-Systemen hat diese Konfiguration auch den Vorteil, dass der schnelle SP High Performance Switch genutzt werden kann, wenn der Dispatcher und die TCP-Servermaschinen auf Knoten in einem SP Frame ausgeführt werden.

Zum Einrichten eines privaten Netzes muss jede Maschine mindestens zwei LAN-Karten enthalten, von denen eine mit dem privaten Netz verbunden ist. Die zweite LAN-Karte muss für ein anderes Teilnetz konfiguriert werden, damit die Dispatcher-Maschine die Client-Anforderungen über das private Netz an die TCP-Servermaschinen sendet.

Windows-Systeme: Konfigurieren Sie die NFA (nicht für Weiterleitung bestimmte Adresse) mit dem Befehl `executor configure`.

Die über den Befehl **dscontrol server add** hinzugefügten Server müssen mit den Adressen des privaten Netzes hinzugefügt werden. Für das Beispiel in Abb. 39 müsste der Befehl für den Server "Apfel" wie folgt codiert sein:

dscontrol server add Clusteradresse:80:10.0.0.1

Er darf nicht wie folgt aussehen:

dscontrol server add Clusteradresse:80:9.67.131.18

Wenn Sie mit Site Selector Lastinformationen für den Dispatcher bereitstellen, muss Site Selector so konfiguriert werden, dass die Last an den privaten Adressen gemeldet wird.

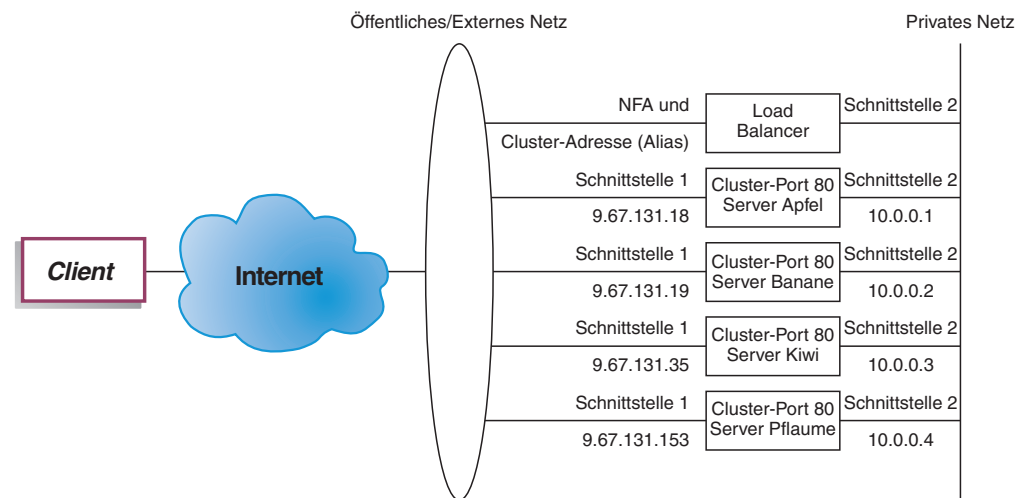


Abbildung 39. Beispiel für ein privates Netz mit dem Dispatcher

Die Verwendung der Konfiguration für ein privates Netz gilt nur für die Komponente Dispatcher.

Platzhaltercluster zum Zusammenfassen von Serverkonfigurationen verwenden

Die Verwendung eines Platzhalterclusters für die Zusammenfassung von Serverkonfigurationen gilt nur für die Komponente Dispatcher.

Das Wort "Platzhalter" bezieht sich auf die Fähigkeit des Clusters, mit mehreren IP-Adressen übereinzustimmen. Die Clusteradresse 0.0.0.0 wird verwendet, um einen Platzhaltercluster anzugeben.

Wenn Sie für viele Clusteradressen einen Lastausgleich durchführen müssen und die Port/Server-Konfigurationen für alle Cluster identisch sind, können Sie die Cluster in einer Platzhalterclusterkonfiguration zusammenfassen.

Sie müssen dennoch jede Clusteradresse explizit für einen der Netzwerkadapter Ihrer Dispatcher-Workstation konfigurieren. Sie sollten jedoch keine der Clusteradressen mit dem Befehl **dscontrol cluster add** zur Dispatcher-Konfiguration hinzufügen.

Fügen Sie nur den Platzhaltercluster (Adresse 0.0.0.0) hinzu und konfigurieren Sie die Ports und Server wie für den Lastausgleich erforderlich. Für den Datenverkehr an jede der auf dem Adapter konfigurierten Adressen erfolgt ein Lastausgleich unter Verwendung der Platzhalterclusterkonfiguration.

Ein Vorteil dieser Methode besteht darin, dass der Datenverkehr an alle Clusteradressen bei der Bestimmung des besten Servers berücksichtigt wird. Ist der Datenverkehr bei einem Cluster besonders hoch und hat der Cluster viele aktive Verbindungen auf einem der Server erstellt, findet für den Datenverkehr an andere Clusteradressen ein Lastausgleich unter Verwendung dieser Informationen statt.

Sie können den Platzhaltercluster mit tatsächlichen Clustern kombinieren, wenn Sie einige Clusteradressen mit eindeutiger Port/Server-Konfiguration und einige Clusteradressen mit gemeinsamer Konfigurationen haben. Die eindeutigen Konfigurationen müssen jeweils einer tatsächlichen Clusteradresse zugeordnet werden. Alle gemeinsamen Konfigurationen können dem Platzhaltercluster zugeordnet werden.

Platzhaltercluster für den Lastausgleich von Firewalls verwenden

Die Verwendung eines Platzhalterclusters für den Lastausgleich von Firewalls gilt nur für die Komponente Dispatcher. Die Clusteradresse 0.0.0.0 wird verwendet, um einen Platzhaltercluster anzugeben.

Der Platzhaltercluster kann für den Lastausgleich von Datenverkehr an Adressen verwendet werden, die nicht explizit für einen Netzwerkadapter der Dispatcher-Workstation konfiguriert sind. Dazu muss der Dispatcher mindestens den gesamten Datenverkehr sehen können, für den ein Lastausgleich erfolgen soll. Die Dispatcher-Workstation erkennt keinen Datenverkehr an Adressen, die nicht explizit für einen ihrer Netzwerkadapter konfiguriert wurden. Eine Ausnahme hiervon bilden Adressen, die für bestimmten Datenverkehr als Standardroute konfiguriert sind.

Wurde der Dispatcher als Standardroute konfiguriert, erfolgt der Lastausgleich für den TCP- oder UDP-Datenverkehr, der über die Dispatcher-Maschine transportiert wird, unter Verwendung der Platzhalterclusterkonfiguration.

Diese Methode kann für den Lastausgleich von Firewalls verwendet werden. Da Firewalls Pakete für jede Zieladresse und jeden Ziel-Port verarbeiten können, müssen Sie den Lastausgleich für den Datenverkehr unabhängig von der Zieladresse und dem Ziel-Port durchführen können.

Firewalls werden für die Bearbeitung des Datenverkehrs von nicht gesicherten Clients zu gesicherten Servern und die Antworten von den gesicherten Servern sowie den Datenverkehr von Clients auf der gesicherten Seite zu Servern auf der nicht gesicherten Seite mit den entsprechenden Antworten verwendet.

Sie müssen zwei Dispatcher-Maschinen konfigurieren, eine für die Verteilung des nicht gesicherten Datenverkehrs an die nicht gesicherten Firewall-Adressen und eine für die Verteilung des gesicherten Datenverkehrs an die gesicherten Firewall-Adressen. Da beide Dispatcher den Platzhaltercluster und den Platzhalter-Port mit verschiedenen Gruppen von Serveradressen verwenden müssen, ist es erforderlich, dass sich die beiden Dispatcher auf zwei separaten Workstations befinden.

Platzhaltercluster mit Caching Proxy für transparente Weiterleitung verwenden

Die Verwendung eines Platzhalterclusters mit Caching Proxy für transparente Weiterleitung ist nur für die Komponente Dispatcher. Die Clusteradresse 0.0.0.0 wird verwendet, um einen Platzhaltercluster anzugeben.

Bei Verwendung der Platzhalterclusterfunktion kann der Dispatcher eine transparente Proxy-Funktion für einen Caching-Proxy-Server aktivieren, der sich auf derselben Maschine wie der Dispatcher befindet. Dies ist nur eine AIX-Funktion, da zwischen der Komponente Dispatcher und der TCP-Komponente des Betriebssystems eine Kommunikation stattfinden muss.

Zum Aktivieren dieses Features müssen Sie Caching Proxy für den Empfang von Client-Anforderungen am Port 80 starten. Anschließend konfigurieren Sie einen Platzhaltercluster (0.0.0.0). Konfigurieren Sie im Platzhaltercluster den Port 80. Für Port 80 konfigurieren Sie die NFA der Dispatcher-Maschine als einzigen Server. Der gesamte Client-Datenverkehr an Adressen des Ports 80 wird nun an den Caching-Proxy-Server auf der Dispatcher-Workstation gesendet. Anschließend wird die Client-Anforderung wie üblich weitergeleitet. Die Antwort wird von Caching Proxy an den Client zurückgesendet. In diesem Modus führt die Komponente Dispatcher keinen Lastausgleich durch.

Platzhalter-Port für die Übertragung von Datenverkehr mit nicht konfiguriertem Port verwenden

Der Platzhalter-Port kann für Datenverkehr verwendet werden, der nicht für einen explizit konfigurierten Port bestimmt ist. Eine solche Verwendung wäre der Lastausgleich für Firewalls. Zum anderen kann mit einem Platzhalter-Port sichergestellt werden, dass an einen nicht konfigurierten Port gerichteter Datenverkehr entsprechend bearbeitet wird. Durch Definieren eines Platzhalter-Ports ohne Server stellen Sie sicher, dass alle Anforderungen an einen nicht konfigurierten Port gelöscht und nicht an das Betriebssystem zurückgesendet werden. Ein Platzhalter-Port wird mit der Port-Nummer 0 (null) angegeben. Beispiel:

```
dscontrol port add Cluster:0
```

Platzhalter-Port für FTP-Datenverkehr verwenden

Wenn Sie einen Cluster für passives FTP und den Platzhalter-Port konfigurieren, verwendet das passive FTP für Datenverbindungen standardmäßig den gesamten Bereich der nicht reservierten TCP-Ports. Für einen Client, der über einen Lastausgleichscluster eine Datenverbindung zu einem FTP-Steuer-Port aufgebaut hat, bedeutet dies, dass Load Balancer nachfolgende Steuerverbindungen zu diesem Cluster und Verbindungen zu Ports mit höheren Nummern als 1023 in diesem Cluster automatisch an denselben Server weiterleitet wie die FTP-Steuerverbindung.

Wenn für den Platzhalter-Port und den FTP-Port in einem Cluster nicht derselbe Server definiert ist, können Anwendungen an Ports mit hohen Nummern (höher als 1023) scheitern, wenn ein Client bereits eine FTP-Steuerverbindung hat. Das Definieren verschiedener Servergruppen für den FTP-Port und den Platzhalter-Port in einem Cluster wird daher nicht empfohlen. Falls ein solches Szenario gewünscht wird, muss in der Konfiguration von Load Balancer der Bereich passiver Ports für den FTP-Dämon festgelegt werden.

Erkennung von DoS-Attacken

Diese Funktion ist nur für die Komponente Dispatcher verfügbar.

Der Dispatcher ist in der Lage, potenzielle DoS-Attacken zu erkennen und Administratoren durch einen Alert zu benachrichtigen. Dazu analysiert der Dispatcher eingehende Anforderungen auf eine verdächtige Anzahl halboffener TCP-Verbindungen von Servern, die ein allgemeines Kennzeichen einfacher DoS-Attacken sind. Bei einer DoS-Attacke empfängt eine Site eine große Anzahl fabrizierter SYN-Pakete von einer Vielzahl von Quellen-IP-Adressen und Quellen-Port-Nummern. Folgepakete für diese TCP-Verbindungen werden jedoch nicht empfangen. Dies führt zu einer großen Anzahl halboffener TCP-Verbindungen auf den Servern, so dass diese mit der Zeit sehr langsam werden und keine neuen ankommenden Verbindungen mehr akzeptieren können.

Anmerkung: Erst wenn über den attackierten Cluster und Port wieder Datenverkehr empfangen wird, kann der Dispatcher feststellen, dass eine DoS-Attacke beendet ist. Solange der Datenverkehr nicht erneut fließt, kann der Dispatcher das Ende der Attacke nicht erkennen.

Load Balancer stellt Benutzer-Exits bereit, die Scripts aufrufen. Diese Scripts können so angepasst werden, dass der Administrator per Alert von einer möglichen DoS-Attacke informiert wird. Dispatcher stellt im Verzeichnis **...ibm/edge/lb/servers/samples** das folgende Beispiel-Script bereit:

- **halfOpenAlert** — es wurde eine mögliche DoS-Attacke (Denial of Service) festgestellt
- **halfOpenAlertDone** — die DoS-Attacke ist beendet

Zum Ausführen der Dateien müssen Sie sie in das Verzeichnis **...ibm/edge/lb/servers/bin** verschieben und die Erweiterung **.sample** löschen.

Zum Implementieren der Erkennung von DoS-Attacken müssen Sie wie folgt den Parameter **maxhalfopen** des Befehls **dscontrol port** setzen:

```
dscontrol port set 127.40.56.1:80 maxhalfopen 1000
```

Im obigen Beispiel vergleicht der Dispatcher die aktuelle Gesamtanzahl halboffener Verbindungen (für alle Server des Clusters 127.40.56.1 am Port 80) mit dem Schwellenwert 1000 (der vom Parameter **maxhalfopen** angegeben ist). Übersteigt die Anzahl der aktuellen halboffenen Verbindungen den Schwellenwert, wird ein Alert-Script (**halfOpenAlert**) aufgerufen. Fällt die Anzahl halboffener Verbindungen unter den Schwellenwert, wird ein anderes Alert-Script aufgerufen, um das Ende der Attacke anzuzeigen.

Bestimmen Sie wie folgt den Wert, den Sie für maxhalfopen definieren sollten: Wenn auf Ihrer Site ein mäßiger bis starker Datenverkehr zu verzeichnen ist, erstellen Sie in regelmäßigen Abständen (vielleicht alle 10 Minuten) mit **dscontrol port halfopenaddressreport Cluster:Port** einen Bericht zu halboffenen Verbindungen. Der Bericht gibt die aktuelle Gesamtanzahl der empfangenen halboffenen Verbindungen an. Setzen Sie **maxhalfopen** auf einen Wert, der 50 bis 200 % über der höchsten Anzahl halboffener Verbindungen liegt, die auf Ihrer Site aufgetreten sind.

Neben statistischen Daten generiert `halfopenaddressreport` für alle Client-Adressen (maximal 8000 Adresspaare), deren Serverzugriff halboffene Verbindungen zur Folge hatten, Einträge im Protokoll (`../ibm/edge/lb/servers/logs/dispatcher/halfOpen.log`).

Anmerkung: Es gibt SNMP-Alarmnachrichten, die den Scripts `halfOpenAlert` und `halfOpenAlertDone` entsprechen. Wenn der SNMP-Subagent konfiguriert und aktiv ist, werden unter den Bedingungen, die die Scripts aufrufen, die entsprechenden Alarmnachrichten gesendet. Weitere Informationen zum SNMP-Subagenten finden Sie im Abschnitt „Simple Network Management Protocol mit Dispatcher verwenden“ auf Seite 271.

Back-End-Server können Sie zusätzlich vor DoS-Attacken schützen, indem Sie Platzhaltercluster und -Ports konfigurieren. Fügen Sie unter jedem konfigurierten Cluster einen Platzhalter-Port ohne Server hinzu. Fügen Sie außerdem einen Platzhaltercluster mit einem Platzhalter-Port und ohne Server hinzu. Dies hat zur Folge, dass alle Pakete, die an einen Platzhaltercluster oder -Port gesendet werden, gelöscht werden. Informationen zu Platzhalterclustern und -Ports finden Sie in den Abschnitten „Platzhaltercluster zum Zusammenfassen von Serverkonfigurationen verwenden“ auf Seite 237 und „Platzhalter-Port für die Übertragung von Datenverkehr mit nicht konfiguriertem Port verwenden“ auf Seite 239.

Binäre Protokolle für die Analyse von Serverstatistiken verwenden

Anmerkung: Das Feature für binäre Protokollierung ist für die Komponenten Dispatcher und CBR verfügbar.

Mit dem Feature für binäre Protokollierung können Serverinformationen in Binärdateien gespeichert werden. Diese Dateien können dann verarbeitet werden, um die Serverinformationen zu analysieren, die über einen bestimmten Zeitraum gesammelt wurden.

Die folgenden Informationen werden für jeden in der Konfiguration definierten Server in dem binären Protokoll gespeichert:

- Clusteradresse
- Port-Nummer
- Server-ID
- Serveradresse
- Serverwertigkeit
- Summe Verbindungen für Server
- Aktive Verbindungen für Server
- Last am Server-Port
- Serversystembelastung

Einige dieser Informationen werden im Rahmen des Manager-Zyklus vom Executor abgerufen. Der Manager muss daher aktiv sein, damit die Informationen in den binären Protokollen aufgezeichnet werden können.

Verwenden Sie den Befehlssatz **`dscontrol binlog`**, um das binäre Protokollieren zu konfigurieren.

- `binlog start`
- `binlog stop`

- `binlog set interval <Sekunde>`
- `binlog set retention <Stunden>`
- `binlog status`

Mit der Option 'start' wird die Protokollierung von Serverinformationen in binären Protokollen im Protokollverzeichnis gestartet. Ein Protokoll wird zu Beginn jeder Stunde mit dem Datum und der Uhrzeit als Name der Datei erstellt.

Mit der Option 'stop' wird die Protokollierung von Serverinformationen in binären Protokollen gestoppt. Standardmäßig ist der Protokolldienst gestoppt.

Mit der Option 'set interval' wird gesteuert, wie oft Informationen in die Protokolle geschrieben werden. Der Manager sendet in jedem Manager-Intervall Serverdaten an den Protokollserver. Die Daten werden nur in die Protokolle geschrieben, wenn seit dem Schreiben des letzten Protokolleintrags die für das Protokollintervall angegebene Zeit in Sekunden verstrichen ist. Standardmäßig wird das Protokollintervall auf 60 Sekunden gesetzt. Zwischen den Einstellungen für das Manager-Intervall und das Protokollintervall gibt es eine gewisse Interaktion. Da dem Protokollserver Informationen nicht schneller zur Verfügung gestellt werden, als dies im Manager-Intervall (in Sekunden) angegeben ist, wird durch Angabe eines Protokollintervalls, das kleiner als das Manager-Intervall ist, das Protokollintervall de facto auf denselben Wert wie das Manager-Intervall gesetzt. Mit dieser Protokollierungstechnik können Sie Serverinformationen detaillierter erfassen. Sie können alle vom Manager festgestellten Änderungen der Serverinformationen für die Berechnung von Serverwertigkeiten erfassen. Dieser Informationsumfang ist jedoch wahrscheinlich nicht erforderlich, um die Serverauslastung und Trends zu analysieren. Werden Serverinformationen alle 60 Sekunden protokolliert, erhalten Sie Momentaufnahmen von Serverinformationen in Abhängigkeit vom zeitlichen Verlauf. Wird das Protokollintervall auf einen sehr niedrigen Wert gesetzt, kann dies zu großen Datenmengen führen.

Mit der Option 'set retention' wird gesteuert, wie lange Protokolldateien aufbewahrt werden. Protokolldateien, die älter als die angegebene Verweildauer (Stunden) sind, werden von dem Protokollserver gelöscht. Dies geschieht nur, wenn der Protokollserver von dem Manager aufgerufen wird, d. h., wird der Manager gestoppt, werden alte Protokolldateien nicht gelöscht.

Mit der Option 'status' werden die aktuellen Einstellungen des Protokolldienstes zurückgegeben. Diese Einstellungen geben an, ob der Service gestartet ist und welche Werte für das Intervall und die Verweildauer angegeben sind.

Im Verzeichnis `...ibm/edge/lb/servers/samples/BinaryLog` stehen ein Beispiel-Java-Programm und eine Beispielbefehlsdatei zur Verfügung. Dieses Beispiel zeigt, wie alle Informationen aus den Protokolldateien abgerufen und angezeigt werden können. Es kann für jede Art von Datenanalyse angepasst werden. Beispiel unter Verwendung des bereitgestellten Scripts und Programms für Dispatcher:

```
dslogreport 2001/05/01 8:00 2001/05/01 17:00
```

Dieser Befehl liefert einen Bericht mit den Serverdaten der Komponente Dispatcher vom 1. Mai 2001 in der Zeit von 8.00 Uhr bis 17.00 Uhr. (Verwenden Sie für CBR `cbrlogreport`.)

Verknüpften Client verwenden

Konfigurationen mit einem Client auf derselben Maschine wie Load Balancer werden nur auf Linux-Systemen unterstützt.

Konfigurationen mit verknüpftem Client funktionieren auf anderen Plattformen möglicherweise nicht ordnungsgemäß, weil Load Balancer die ankommenden Pakete unter den vielen unterstützten Betriebssystemen mit verschiedenen Techniken untersucht. In den meisten Fällen empfängt Load Balancer nur unter Linux Pakete von der lokalen Maschine. Unter anderen Betriebssystemen werden nur die Pakete aus dem Netz empfangen. Aus diesem Grund empfängt Load Balancer keine Anforderungen, die von der lokalen Maschine an die Clusteradresse gesendet werden, und kann diese nicht erfüllen.

Kapitel 23. Erweiterte Funktionen für Cisco CSS Controller und Nortel Alteon Controller

Dieses Kapitel enthält die folgenden Abschnitte:

- „Verknüpfung“
- „Hohe Verfügbarkeit“
- „Lastausgleich mit Load Balancer optimieren“ auf Seite 248
- „Advisor-Funktionen“ auf Seite 250
- „Metric Server“ auf Seite 255
- „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 258
- „Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden“ auf Seite 259

Anmerkung: In diesem Kapitel ist **xxxcontrol** durch **cococontrol** für Cisco CSS Controller bzw. **nalcontrol** für Nortel Alteon Controller zu ersetzen.

Verknüpfung

Cisco CSS Controller oder Nortel Alteon Controller kann sich auf derselben Maschine befinden wie ein Server, dessen Anforderungen verteilt werden. Dies wird als *Verknüpfen* eines Servers bezeichnet. Es sind keine zusätzlichen Konfigurationsschritte erforderlich.

Anmerkung: In Zeiten hohen Datenverkehrs konkurriert ein verknüpfter Server mit Load Balancer um Ressourcen. Sind jedoch keine überlasteten Maschinen vorhanden, kann mit einem verknüpften Server die Gesamtzahl der Maschinen reduziert werden, die für das Einrichten eines Standortes mit Lastausgleich erforderlich sind.

Hohe Verfügbarkeit

Die Funktion der hohen Verfügbarkeit ist jetzt für Cisco CSS Controller und Nortel Alteon Controller verfügbar.

Zur Verbesserung der Fehlertoleranz des Controllers stellt die Funktion für hohe Verfügbarkeit Folgendes bereit:

- Mechanismen zur Bestimmung der Partnercontroller durch Überwachungssignale. Zwischen den Adressen, die mit dem Befehl **xxxcontrol highavailability add** konfiguriert wurden, werden Überwachungssignale ausgetauscht. Sie können festlegen, in welchen Zeitabständen die Signale ausgetauscht werden und nach welcher Zeit ein Controller die Aufgaben seines Partners übernimmt.
- Eine Liste von Zielen, die für jeden Controller erreichbar sein müssen, damit die Wertigkeiten berechnet werden können und der Switch aktualisiert werden kann. Weitere Informationen hierzu finden Sie im Abschnitt „Ausfallerkennung“ auf Seite 247.
- Logik für die Auswahl des aktiven Controllers ausgehend von der Verfügbarkeit und den Erreichbarkeitsinformationen.
- Konfigurierbare Übernahmestrategie für die Übernahme der Aufgaben eines Controllers durch seinen Partner.

- Manueller Übernahmemechanismus für die Wartung aktiver Controller.
- Berichte mit den aktuellen Angaben zu Controllerrolle, -status, -synchronisation usw.

Konfiguration

Die vollständige Syntax für den Befehl **xxxcontrol highavailability** finden Sie in den Abschnitten „ccocontrol highavailability — Hohe Verfügbarkeit steuern“ auf Seite 441 und „nalcontrol highavailability — Hohe Verfügbarkeit steuern“ auf Seite 459.

Gehen Sie wie folgt vor, um die hohe Verfügbarkeit für den Controller zu konfigurieren:

1. Starten Sie auf beiden Controllermaschinen den Controllerserver.
2. Konfigurieren Sie alle Controller identisch.
3. Konfigurieren Sie für den lokalen Controller wie folgt die Rolle für hohe Verfügbarkeit, die Adresse und die Partneradresse:

```
xxxcontrol highavailability add address 10.10.10.10
partneraddress 10.10.10.20 port 143 role primary
```
4. Konfigurieren Sie für den Partnercontroller wie folgt die Rolle für hohe Verfügbarkeit, die Adresse und die Partneradresse:

```
xxxcontrol highavailability add address 10.10.10.20
partneraddress 10.10.10.10 port 143 role secondary
```

Die Parameter `address` und `partneraddress` sind auf der primären und der sekundären Maschine jeweils ausgetauscht.

5. Konfigurieren Sie optional auf dem lokalen und dem Partnercontroller Parameter für hohe Verfügbarkeit. Beispiel:

```
xxxcontrol highavailability set beatinterval 1000
```
6. Konfigurieren Sie bei Bedarf wie folgt Erreichbarkeitsziele auf dem lokalen und dem Partnercontroller:

```
xxxcontrol highavailability usereach 10.20.20.20
```

Auf dem lokalen und dem Partnercontroller muss dieselbe Anzahl von Erreichbarkeitszielen konfiguriert werden.

7. Starten Sie die Komponente mit hoher Verfügbarkeit und definieren Sie auf dem lokalen und dem Partnercontroller wie folgt eine Wiederherstellungsstrategie:

```
xxxcontrol highavailability start auto
```
8. Optional können Sie auf dem lokalen und Partnercontroller wie folgt Informationen zur hohen Verfügbarkeit anzeigen:

```
xxxcontrol highavailability report
```
9. Geben Sie optional auf dem Bereitschaftscontroller wie folgt `takeover` an, wenn dieser die Aufgaben des aktiven Controllers übernehmen soll:

```
xxxcontrol highavailability takeover
```

Dies ist nur für Wartungszwecke erforderlich.

Anmerkungen:

1. Wenn Sie einen Controller ohne hohe Verfügbarkeit konfigurieren möchten, setzen Sie keine Befehle zur hohen Verfügbarkeit ab.
2. Falls Sie eine Konfiguration mit hoher Verfügbarkeit, die zwei Controller umfasst, in eine Konfiguration mit nur einem Controller ändern möchten, stop-

pen Sie zunächst auf dem Bereitschaftscontroller die hohe Verfügbarkeit. Bei Bedarf können Sie dann auch auf dem aktiven Controller die hohe Verfügbarkeit stoppen.

3. In einer Konfiguration mit hoher Verfügbarkeit, die zwei Controller umfasst, kann es zu unerwarteten Ergebnissen kommen, wenn die Controllermerkmale (z. B. switch consultantid, switch address usw.) auf den Switches nicht übereinstimmen. Unerwartete Ergebnisse können sich auch einstellen, wenn die Controllermerkmale für hohe Verfügbarkeit nicht identisch sind, z. B. port, role, reach targets, beatinterval, takeoverinterval und recovery strategy.

Ausfallerkennung

Neben dem Verlust der Konnektivität zwischen aktivem Controller und Bereitschaftscontroller, der durch die Überwachungssignale festgestellt wird, kann jetzt auch die *Erreichbarkeit* erkannt werden.

Wenn Sie die hohe Verfügbarkeit für Controller konfigurieren, können Sie eine Liste von Hosts angeben, die für jeden der Controller erreichbar sein müssen. Für jedes Teilnetz, das Ihre Controllermaschine benutzt, muss mindestens ein Host angegeben sein. Die Hosts können Router, IP-Server oder andere Arten von Hosts sein.

Die Erreichbarkeit von Hosts wird über ping-Aufrufe der Advisor-Funktion reach abgefragt. Es findet eine Übernahme statt, wenn keine Überwachungssignalnachrichten durchkommen oder wenn die Erreichbarkeitskriterien eher vom Bereitschaftscontroller als vom primären Controller erfüllt werden. Damit die Entscheidung anhand aller verfügbaren Informationen getroffen wird, tauschen der aktive Controller und der Bereitschaftscontroller regelmäßig Informationen über ihre Erreichbarkeit aus. Die Controller vergleichen dann ihre Erreichbarkeitsdaten mit denen des Partners und entscheiden, welcher Controller der aktive Controller sein soll.

Wiederherstellungsstrategie

Die beiden Controllermaschinen werden nach ihrer Rolle als primärer und sekundärer Controller konfiguriert. Beim Systemstart tauschen die Controller Informationen aus, bis beide Maschinen synchronisiert sind. Anschließend wechselt der primäre Controller in den aktiven Status und beginnt mit der Berechnung von Wertigkeiten und dem Aktualisieren des Switch. Die sekundäre Maschine wechselt in den Bereitschaftsstatus und überwacht die Verfügbarkeit der primären Maschine.

Stellt die Bereitschaftsmaschine an einem beliebigen Punkt fest, dass die aktive Maschine ausgefallen ist, übernimmt sie die Lastausgleichsfunktionen der (ausgefallenen) aktiven Maschine und wird selbst zur aktiven Maschine. Ist die primäre Maschine wieder betriebsbereit, ermitteln die beiden Maschinen anhand der konfigurierten Wiederherstellungsstrategie, welcher Controller der aktive Controller sein soll.

Es gibt zwei Arten von Wiederherstellungsstrategien:

Automatische Wiederherstellung

Sobald der primäre Controller wieder betriebsbereit ist, wechselt er in den aktiven Status, berechnet und aktualisiert Wertigkeiten etc. Die sekundäre Maschine wechselt in den Bereitschaftsstatus, wenn die primäre Maschine wieder aktiv ist.

Manuelle Wiederherstellung

Der aktive sekundäre Controller bleibt aktiv, wenn der primäre Controller wieder betriebsbereit ist.

Der primäre Controller wechselt in den Bereitschaftsstatus und kann nur manuell in den aktiven Status versetzt werden.

Der Parameter für die Strategie muss für beide Maschinen auf denselben Wert gesetzt werden.

Beispiele

Beispiele zur hohen Verfügbarkeit für Cisco CSS Controller finden Sie im Abschnitt „Beispiele“ auf Seite 443.

Beispiele zur hohen Verfügbarkeit für Nortel Alteon Controller finden Sie im Abschnitt „Beispiele“ auf Seite 461.

Lastausgleich mit Load Balancer optimieren

Die Controllerfunktion von Load Balancer führt den Lastausgleich ausgehend von den folgenden Einstellungen durch:

- „Bedeutung von Messwerten“
- „Wertigkeiten“ auf Seite 249
- „Ruhezeiten für Wertigkeitsberechnung“ auf Seite 249
- „Advisor-Ruhezeiten“ auf Seite 251
- „Sensitivitätsschwelle“ auf Seite 250

Zur Optimierung des Lastausgleichs für Ihr Netz können Sie diese Einstellungen ändern.

Bedeutung von Messwerten

Der Controller kann in seine Gewichtungsentscheidung alle oder einige der nachfolgend genannten summierten Messwerte einfließen lassen:

- *Aktive Verbindungen*: Die vom Switch abgerufene Anzahl aktiver Verbindungen auf jeder am Lastausgleich beteiligten Servermaschine.
- *Verbindungsrate*: Die vom Switch abgerufene Anzahl neuer Verbindungen seit der letzten Anfrage auf jeder am Lastausgleich beteiligten Servermaschine.
- *CPU*: Prozentsatz der auf jeder am Lastausgleich beteiligten Servermaschine genutzten CPU (Vorgabe vom Agenten Metric Server).
- *Speicher*: Prozentsatz des auf jeder am Lastausgleich beteiligten Servermaschine genutzten Speichers (Vorgabe vom Agenten Metric Server).
- *Systemmesswert*: Vorgabe von den Systemüberwachungs-Tools wie Metric Server oder WLM.
- *Anwendungsspezifisch*: Vorgaben von den Advisor-Funktionen, die am Port empfangsbereit sind.

Die Standardmesswerte sind activeconn und connrate.

Sie können die relative Gewichtung der Messwerte ändern. Die Proportionen sind vergleichbar mit Prozentsätzen. Die Summe der relativen Proportionen muss 100 % ergeben. Standardmäßig werden die Messwerte "Aktive Verbindungen" und "Neue

Verbindungen" in der Gewichtung 50:50 verwendet. In Ihrer Umgebung sollten Sie andere Messwertproportionen ausprobieren, um die Kombination mit der besten Leistung zu finden.

Gehen Sie wie folgt vor, um die Proportionswerte festzulegen:

Für Cisco CSS Controller

```
cococontrol ownercontent metrics Messwertname1 Proportion1 Messwertname2 Proportion2
```

Für Nortel Alteon Controller

```
nalcontrol service metrics Messwertname1 Proportion1 Messwertname2 Proportion2
```

Wertigkeiten

Die Wertigkeiten werden ausgehend von Reaktionszeit und Verfügbarkeit der Anwendung, vom Feedback der Advisor-Funktionen und vom Feedback eines Systemüberwachungsprogramms wie Metric Server festgelegt. Falls Sie Wertigkeiten manuell festlegen möchten, geben Sie für den Server die Option **fixedweight** an. Eine Beschreibung der Option **fixedweight** finden Sie im Abschnitt „Feste Wertigkeiten vom Controller“.

Wertigkeiten gelten für alle Server, die einen Service anbieten. Für jeden einzelnen Service werden die Anforderungen entsprechend ihrer relativen Wertigkeit auf die Server verteilt. Hat beispielsweise ein Server die Wertigkeit 10 und der andere Server die Wertigkeit 5, erhält der Server mit der Wertigkeit 10 doppelt so viele Anforderungen wie der Server mit der Wertigkeit 5.

Stellt eine Advisor-Funktion fest, dass ein Server heruntergefahren wurde, wird seine Wertigkeit auf -1 gesetzt. Für Cisco CSS Controller und Nortel Alteon Controller: Der Switch wird informiert, dass der Server nicht verfügbar ist, und hört auf, dem Server Verbindungen zuzuordnen.

Feste Wertigkeiten vom Controller

Ohne den Controller können Advisor-Funktionen nicht ausgeführt werden und nicht erkennen, ob ein Server inaktiv ist. Wenn Sie die Advisor-Funktionen ausführen möchten, der Controller jedoch *nicht* die von Ihnen für einen bestimmten Server festgelegte Wertigkeit aktualisieren soll, verwenden Sie für Cisco CSS Controller den Befehl **cococontrol service** und für Nortel Alteon Controller den Befehl **nalcontrol server** mit der Option **fixedweight**.

Mit dem Befehl **fixedweight** können Sie die Wertigkeit auf den gewünschten Wert setzen. Der Wert für die Serverwertigkeit bleibt während der Ausführung des Controllers unverändert erhalten, bis Sie einen weiteren Befehl absetzen, bei dem **fixedweight** auf **no** gesetzt ist.

Ruhezeiten für Wertigkeitsberechnung

Zur Optimierung der Gesamtleistung können Sie festlegen, wie oft Messwerte zusammengestellt werden sollen.

Die Consultant-Ruhezeit gibt an, wie oft der Consultant die Serverwertigkeiten aktualisiert. Eine zu kurze Consultant-Ruhezeit kann sich negativ auf den Durchsatz auswirken, da der Consultant den Switch permanent unterbricht. Eine zu lange Consultant-Ruhezeit kann bedeuten, dass der Lastausgleich des Switch nicht auf genauen, auf dem neuesten Stand befindlichen Informationen basiert.

Eine Consultant-Ruhezeit von 1 Sekunde könnten Sie wie folgt festlegen:

```
xxxcontrol consultant set Consultant-ID sleeptime Intervall
```

Sensitivitätsschwelle

Zur Optimierung des Lastausgleichs für Ihre Server stehen weitere Methoden zur Verfügung. Im Interesse einer hohen Übertragungsgeschwindigkeit werden die Wertigkeiten der Server nur aktualisiert, wenn sich signifikante Änderungen der Wertigkeit ergeben. Das permanente Aktualisieren der Wertigkeiten bei geringfügigen oder nicht vorhandenen Änderungen des Serverstatus würde zu einem unnötigen Systemaufwand führen. Wenn die prozentuale Änderung der Wertigkeit innerhalb der summierten Wertigkeit für alle Server, die einen Service anbieten, über der Sensitivitätsschwelle liegt, werden die von Load Balancer für die Verteilung der Verbindungen verwendeten Wertigkeiten aktualisiert. Nehmen wir beispielsweise an, die Gesamtwertigkeit ändert sich von 100 % auf 105 %. Die Änderung beträgt also 5 %. Beim standardmäßigen Sensitivitätsschwellenwert von 5 werden die von Load Balancer verwendeten Wertigkeiten nicht aktualisiert, da die prozentuale Änderung nicht **über** dem Schwellenwert liegt. Ändert sich die Gesamtwertigkeit jedoch von 100 % auf 106 %, werden die Wertigkeiten aktualisiert. Wenn Sie die Consultant-Sensitivitätsschwelle auf einen anderen Wert als den Standardwert setzen möchten, geben Sie den folgenden Befehl ein:

```
xxxcontrol consultant set Consultant-ID sensitivity geänderterProzentsatz
```

In den meisten Fällen müssen Sie diesen Wert nicht ändern.

Advisor-Funktionen

Advisor-Funktionen sind Agenten von Load Balancer. Ihr Zweck ist es, den Zustand und die Belastung der Servermaschinen zu beurteilen. Dies erfolgt durch einen proaktiven Austausch mit den Servern, der dem von Clients vergleichbar ist. Advisor-Funktionen können als transportable Clients der Anwendungsserver betrachtet werden.

Anmerkung: Eine detaillierte Liste der Advisor-Funktionen finden Sie im Abschnitt „Liste der Advisor-Funktionen“ auf Seite 189.

Arbeitsweise der Advisor-Funktionen

Advisor-Funktionen öffnen regelmäßig eine TCP-Verbindung zu jedem Server und senden eine Anforderungsnachricht an den Server. Der Inhalt der Nachricht ist spezifisch für das Protokoll, das auf dem Server ausgeführt wird. Die HTTP-Advisor-Funktion sendet beispielsweise eine HTTP-Anfrage „HEAD“ an den Server.

Die Advisor-Funktionen warten dann auf den Empfang einer Antwort vom Server. Nach Empfang der Antwort beurteilt die Advisor-Funktion den Server. Um diesen *Lastwert* zu ermitteln, messen die meisten Advisor-Funktionen die Zeit, bis der Server antwortet, und verwenden dann diesen Wert (in Millisekunden) als Lastwert.

Die Advisor-Funktionen übergeben dann den Lastwert an die Consultant-Funktion, die ihn im Consultant-Bericht angibt. Der Consultant addiert anschließend die Wertigkeiten für alle Quellen entsprechend ihren Proportionen und sendet diese Werte an den Switch. Der Switch benutzt diese Wertigkeiten dann für den Lastausgleich neuer ankommender Client-Verbindungen.

Stellt die Advisor-Funktion fest, dass ein Server aktiv ist und ordnungsgemäß arbeitet, meldet er einen positiven Lastwert ungleich null an den Consultant. Stellt die Advisor-Funktion fest, dass ein Server inaktiv ist, gibt sie den speziellen Last-

wert -1 zurück, um dem Switch mitzuteilen, dass der Server heruntergefahren ist. Der Switch leitet daraufhin keine Verbindungen an diesen Server weiter, solange dieser inaktiv ist.

Advisor-Ruhezeiten

Anmerkung: Die Advisor-Standardwerte funktionieren in den meisten Fällen effizient. Gehen Sie mit Vorsicht vor, wenn Sie andere Werte als die Standardwerte verwenden.

Die Advisor-Ruhezeit legt fest, wie oft eine Advisor-Funktion den Status der Server an dem von ihr überwachten Port abfragt und die Ergebnisse dann an den Consultant übergibt. Eine zu kurze Advisor-Ruhezeit kann sich negativ auf den Durchsatz auswirken, da die Advisor-Funktion die Server permanent unterbricht. Eine zu lange Advisor-Ruhezeit kann bedeuten, dass die Gewichtungsentscheidungen des Consultant nicht auf genauen, auf dem neuesten Stand befindlichen Informationen basieren.

Wenn Sie das Intervall der HTTP-Advisor-Funktion beispielsweise auf 3 Sekunden setzen möchten, geben Sie den folgenden Befehl ein:

```
xxxcontrol metriccollector set Consultant-ID:HTTP sleeptime 3
```

Serververbindungs- und -empfangszeitlimit der Advisor-Funktion

Sie können festlegen, in welcher Zeit eine Advisor-Funktion feststellen soll, dass ein bestimmter Port auf einem Server oder für einen Service ausgefallen ist. Die Zeitlimits für ausgefallene Server connecttimeout und receivetimeout bestimmen, wie lange eine Advisor-Funktion wartet, bis sie einen gescheiterten Sende- oder Empfangsvorgang meldet.

Für eine schnellstmögliche Erkennung ausgefallener Server müssen Sie das Verbindungs- und Empfangszeitlimit der Advisor-Funktion auf den kleinsten Wert (eine Sekunde) sowie die Ruhezeit für Advisor-Funktion und Consultant auf den kleinsten Wert (eine Sekunde) setzen.

Anmerkung: Falls das Datenverkehrsaufkommen in Ihrer Umgebung mäßig bis hoch ist und die Reaktionszeit des Servers ansteigt, legen Sie keine zu niedrigen Werte für timeoutconnect und timeoutreceive fest. Andernfalls könnte die Advisor-Funktion einen ausgelasteten Server vorschnell als ausgefallenen Server markieren.

Wenn Sie timeoutconnect für die HTTP-Advisor-Funktion beispielsweise auf 9 Sekunden setzen möchten, geben Sie den folgenden Befehl ein:

```
xxxcontrol metriccollector set Consultant-ID:HTTP timeoutconnect 9
```

Der Standardwert für das Verbindungs- und Empfangszeitlimit liegt beim Dreifachen des Wertes, der für die Ruhezeit der Advisor-Funktion angegeben wurde.

Wiederholungsversuche der Advisor-Funktion

Advisor-Funktionen können wiederholt versuchen, eine Verbindung herzustellen, bevor sie einen Server als inaktiv markieren. Die Advisor-Funktion markiert einen Server erst als inaktiv, wenn die Abfrage nach der festgelegten Anzahl Wiederho-

lungen und einem weiteren Versuch nicht beantwortet wird. Ist keine Anzahl Wiederholungen festgelegt, wird standardmäßig von null Wiederholungen ausgegangen.

Für den Cisco CSS Controller können Sie den Wert **retry** mit dem Befehl **cococontrol ownercontent set** setzen. Weitere Informationen hierzu finden Sie im Abschnitt „cococontrol ownercontent — Eignernamen und content-Regel steuern“ auf Seite 446.

Für den Nortel Alteon Controller können Sie den Wert **retry** mit dem Befehl **nalcontrol service set** setzen. Weitere Informationen hierzu finden Sie im Abschnitt „nalcontrol service — Service konfigurieren“ auf Seite 466.

Kundenspezifische (anpassbare) Advisor-Funktion erstellen

Anmerkung: In diesem Abschnitt wird **Server** als generischer Begriff verwendet und bezeichnet für Cisco CSS Controller einen Service und für Nortel Alteon Controller einen Server.

Die kundenspezifische (anpassbare) Advisor-Funktion ist ein kurzer Java-Code, den Sie als Klassendatei bereitstellen, die vom Basiscode aufgerufen wird. Der Basiscode stellt alle Verwaltungsservices bereit. Dazu gehören unter anderem:

- Starten und Stoppen einer Instanz der angepassten Advisor-Funktion
- Bereitstellen von Status und Berichten
- Aufzeichnen von History-Daten in einer Protokolldatei.

Er übergibt auch die Ergebnisse an den Consultant. Der Basiscode führt regelmäßig einen Advisor-Zyklus aus, wobei alle Server in der Konfiguration individuell ausgewertet werden. Dieser beginnt mit dem Öffnen einer Verbindung zu einer Servermaschine. Wenn das Socket geöffnet wird, ruft der Basiscode die Methode (Funktion) `getLoad` der angepassten Advisor-Funktion auf. Die angepasste Advisor-Funktion führt dann alle für die Auswertung des Serverstatus erforderlichen Schritte aus. Normalerweise sendet sie eine benutzerdefinierte Nachricht an den Server und wartet dann auf eine Antwort. (Die angepasste Advisor-Funktion erhält Zugriff auf den geöffneten Socket.) Der Basiscode schließt dann den Socket zu dem Server und übergibt die Lastinformationen an den Consultant.

Der Basiscode und die angepasste Advisor-Funktion können im normalen Modus oder im Ersetzungsmodus arbeiten. Die Auswahl der Betriebsart wird in der Datei der angepassten Advisor-Funktion als Parameter der Methode `constructor` angegeben.

Im normalen Modus tauscht die angepasste Advisor-Funktion Daten mit dem Server aus. Der Basiscode der Advisor-Funktion misst die Zeit für den Austausch und berechnet den Lastwert. Der Basiscode übergibt dann diesen Lastwert an den Consultant. Die angepasste Advisor-Funktion muss nur null (bei Erfolg) oder -1 (bei einem Fehler) zurückgeben. Zur Angabe des normalen Modus wird die Markierung `replace` der Methode `constructor` auf `false` (falsch) gesetzt.

Im Ersetzungsmodus führt der Basiscode keine Zeitmessungen aus. Der Code der angepassten Advisor-Funktion führt alle für die funktionspezifischen Anforderungen erforderlichen Operationen aus und gibt dann einen tatsächlichen Lastwert zurück. Der Basiscode akzeptiert diesen Wert und übergibt ihn an den Consultant. Um bestmögliche Ergebnisse zu erzielen, sollten Sie den Lastwert zwischen 10 und 1000 normalisieren, wobei 10 einen schnellen Server und 1000 einen langsamen

Server angibt. Zur Angabe des Ersetzungsmodus muss die Markierung `replace` der Methode `constructor` auf `true` gesetzt werden.

Auf diese Weise können Sie eigene Advisor-Funktionen schreiben, die die benötigten präzisen Informationen über Server zur Verfügung stellen. Für die Controller wird ein Beispiel für eine angepasste Advisor-Funktion, **ADV_ctrlsample.java**, geliefert. Nach der Installation von Load Balancer finden Sie den Beispielcode im Installationsverzeichnis **...ibm/edge/lb/servers/samples/CustomAdvisors**.

Die Standardinstallationsverzeichnisse sind:

- AIX-, HP-UX-, Linux-, Solaris-Systeme: `/opt/ibm/edge/lb`
- Windows-Systeme: `C:\Programme\IBM\ibm\edge\lb`

Anmerkung: Wenn Sie eine angepasste Advisor-Funktion zum Cisco CSS Controller oder Nortel Alteon Controller hinzufügen, müssen Sie **ccoserver** oder **nalserver** stoppen und dann erneut starten (verwenden Sie auf Windows-Systemen "Dienste"), um den Java-Prozess zu veranlassen, die neuen angepassten Advisor-Klassendateien zu lesen. Die angepassten Advisor-Klassendateien werden nur beim Systemstart geladen.

Namenskonvention

Der Dateiname für Ihre angepasste Advisor-Funktion muss das Format **ADV_mein-Advisor.java** haben. Er muss mit dem Präfix **ADV_** in Großbuchstaben beginnen. Alle nachfolgenden Zeichen müssen Kleinbuchstaben sein.

Aufgrund von Java-Konventionen muss der Name der in der Datei definierten Klasse mit dem Namen der Datei übereinstimmen. Wenn Sie den Beispielcode kopieren, stellen Sie sicher, dass alle Exemplare von **ADV_ctrlsample** in der Datei in den neuen Klassennamen geändert werden.

Kompilierung

Angepasste Advisor-Funktionen werden in der Sprache Java geschrieben. Verwenden Sie den Java-Compiler, der zusammen mit Load Balancer installiert wird. Während der Kompilierung wird auf die folgenden Dateien Bezug genommen:

- die Datei der angepassten Advisor-Funktion
- die Basisklassendatei **ibmlb.jar** im Installationsverzeichnis **...ibm/edge/lb/servers/lib**.

Der Klassenpfad muss während der Kompilierung auf die Datei der angepassten Advisor-Funktion und die Datei mit den Basisklassen zeigen.

Ein Kompilierungsbefehl für die Windows-Plattform könnte wie folgt aussehen:

```
Installationsverzeichnis/java/bin/javac -classpath  
Installationsverzeichnis\lb\servers\lib\ibmlb.jar ADV_pam.java
```

Für diesen Befehl gilt Folgendes:

- Ihre Advisor-Datei hat den Namen **ADV_pam.java**.
- Ihre Advisor-Datei ist im aktuellen Verzeichnis gespeichert.

Die Ausgabe der Kompilierung ist eine Klassendatei, zum Beispiel:

ADV_pam.class

Kopieren Sie vor dem Starten der Advisor-Funktion die Klassendatei in das Installationsverzeichnis **...ibm/edge/lb/servers/lib/CustomAdvisors**.

Anmerkung: Bei Bedarf können angepasste Advisor-Funktionen unter einem Betriebssystem kompiliert und unter einem anderen Betriebssystem ausgeführt werden. Sie können beispielsweise Ihre Advisor-Funktion Auf Windows-Systemen kompilieren, die Klassendatei (im Binärformat) auf eine AIX-Maschine kopieren und die Advisor-Funktion dort ausführen.

Für AIX-, HP-UX-, Linux- und Solaris-Systeme ist die Syntax ähnlich.

Ausführung

Bevor Sie die angepasste Advisor-Funktion ausführen, müssen Sie die Klassendatei in das richtige Unterverzeichnis des Installationsverzeichnisses kopieren:

```
...ibm/edge/lb/servers/lib/CustomAdvisors/ADV_pam.class
```

Starten Sie den Consultant und setzen Sie dann zum Starten der angepassten Advisor-Funktion den folgenden Befehl ab:

Für Cisco CSS Controller

```
ccocontrol ownercontent metrics Consultant-ID:ID_für_Eignerangaben pam  
100
```

Für Nortel Alteon Controller

```
nalcontrol service metrics Consultant-ID:Service-ID pam 100
```

Für diesen Befehl gilt Folgendes:

- Der Name Ihrer Advisor-Funktion ist **pam** wie in **ADV_pam.java**.
- **100** ist die prozentuale Wertigkeit dieser Advisor-Funktion.

Erforderliche Routinen

Eine angepasste Advisor-Funktion erweitert wie alle anderen Advisor-Funktionen den Advisor-Basiscode **ADV_Base**. Es ist der Advisor-Basiscode, der die meisten Funktionen ausführt. Dazu gehört das Zurückmelden von Belastungen an den Consultant, die für den Wertigkeitsalgorithmus des Consultant verwendet werden. Darüber hinaus stellt der Advisor-Basiscode Socket-Verbindungen her, schließt Sockets und stellt Sende- und Empfangsmethoden für die Advisor-Funktion bereit. Die Advisor-Funktion selbst wird nur zum Senden von Daten an den Port bzw. Empfangen von Daten vom Port des empfohlenen Servers verwendet. Die TCP-Methoden innerhalb des Advisor-Basiscodes sind zeitlich gesteuert, um die Last zu berechnen. Mit einer Markierung der Methode **constructor** in **ADV_base** kann bei Bedarf die vorhandene Last mit der neuen, von der Advisor-Funktion zurückgegebenen Last überschrieben werden.

Anmerkung: Der Advisor-Basiscode stellt in angegebenen Intervallen die Last ausgehend von einem in der Methode **constructor** gesetzten Wert für den Wertigkeitsalgorithmus bereit. Wenn die eigentliche Advisor-Funktion noch keine gültige Last zurückgeben kann, verwendet der Advisor-Basiscode die vorherige Last.

Basisklassenmethoden sind:

- Eine Routine **constructor**. Die **constructor**-Routine ruft die **constructor**-Methode **base class** auf (siehe **Advisor-Beispieldatei**).

- Eine Methode **ADV_AdvisorInitialize**. Diese Methode stellt einen Hook für den Fall zur Verfügung, dass zusätzliche Schritte ausgeführt werden müssen, nachdem die Basisklasse ihre Initialisierung beendet hat.
- Eine Routine **getLoad**. Die Basis-Advisor-Klasse führt das Öffnen des Sockets aus. Daher muss getLoad nur die entsprechenden Sende- und Empfangsanforderungen absetzen, um den Advisor-Zyklus zu beenden.

Suchreihenfolge

Die Controller suchen zunächst in der bereitgestellten Liste nach nativen Advisor-Funktionen. Können Sie dort eine gegebene Advisor-Funktion nicht finden, durchsuchen Sie die Liste der angepassten Advisor-Funktionen.

Benennung und Pfad

- Die Klasse der angepassten Advisor-Funktion muss sich im Unterverzeichnis **...ibm/edge/lb/servers/lib/CustomAdvisors/** des Basisverzeichnis von Load Balancer befinden. Die Standardwerte für dieses Verzeichnis hängen vom verwendeten Betriebssystem ab:
 - AIX-, HP-UX-, Linux- oder Solaris-Systeme:
/opt/ibm/edge/lb/servers/lib/CustomAdvisors/
 - Windows-Systeme:
C:\Programme\IBM\edge\lb\servers\lib\CustomAdvisors
- Es sind nur alphabetische Zeichen in Kleinschreibung zulässig. Ein Bediener muss somit bei der Eingabe von Befehlen in der Befehlszeile nicht auf die Groß-/Kleinschreibung achten. Der Name der Advisor-Datei muss das Präfix **ADV_** haben.

Beispiel-Advisor-Funktion

Die Programmliste für eine Beispiel-Advisor-Funktion finden Sie im Abschnitt „Beispiel-Advisor-Funktion“ auf Seite 487. Nach der Installation befindet sich diese Beispiel-Advisor-Funktion im Verzeichnis **...ibm/edge/lb/servers/samples/CustomAdvisors**.

Metric Server

Metric Server gibt Load Balancer Informationen zur Serverauslastung. Diese Informationen werden in Form systemspezifischer Messwerte für den Serverzustand bereitgestellt. Der Consultant von Load Balancer richtet Anfragen an den Agenten Metric Server, der sich auf jedem der Server befindet, und legt anhand der Messwerte, die er von den Agenten erhalten hat, Wertigkeiten für den Lastausgleich fest. Die Ergebnisse werden für Cisco CSS Controller auch in den Servicebericht und für Nortel Alteon Controller in den Serverbericht gestellt.

Vorbedingungen

Der Agent Metric Server muss auf allen Servern installiert und ausgeführt werden, die am Lastausgleich teilnehmen.

Metric Server verwenden

Nachfolgend sind die Schritte aufgeführt, mit denen Sie Metric Server für die Controller konfigurieren.

- Controllerseite
 1. Starten Sie **ccoserver** oder **nalserver**.

2. Fügen Sie für Cisco CSS Controller einen Switch-Consultant und dann Eigenerangaben hinzu.
Fügen Sie für Nortel Alteon Controller einen Switch-Consultant und dann einen Service hinzu.
3. Geben Sie den Port an, an dem der Agent Metric Server empfangsbereit ist. Diese Angabe muss mit der Information in der Datei `metricserver.cmd` übereinstimmen. Der Standard-Port ist 10004. Verwenden Sie den folgenden Befehl:

Für Cisco CSS Controller

```
cococontrol service set Consultant-ID:ID_für_Eigenerangaben:Server-ID  
metricserverport Port-Nummer
```

Für Nortel Alteon Controller

```
nalcontrol server set Consultant-ID:Service-ID:Server-ID metricserver-  
port Port-Nummer
```

4. Setzen Sie wie folgt den Befehl für Systemmesswerte ab:

Für Cisco CSS Controller

```
cococontrol ownercontent metrics Consultant-ID:ID_für_Eigenerangaben  
Messwertname Bedeutung
```

Für Nortel Alteon Controller

```
nalcontrol service metrics Consultant-ID:Service-ID Messwertname  
Bedeutung
```

Messwertname steht hier für den Namen des Metric-Server-Scripts.

Das Script für Systemmesswerte befindet sich auf dem Back-End-Server und wird für jeden Server, der unter den aufgeführten Eigenerangaben oder dem angegebenen Service in der Konfiguration enthalten ist, ausgeführt. Die beiden Scripts **cpuload** und **memload** stehen für Sie bereit. Sie können aber auch angepasste Scripts für Systemmesswerte erstellen. Das Script enthält einen Befehl, der einen numerischen Wert zurückgeben muss. Dieser numerische Wert stellt eine Lastmessung und keinen Verfügbarkeitswert dar.

Einschränkung: Wenn der Name Ihres Scripts für Systemmesswerte auf Windows-Systemen eine andere Erweiterung als `.exe` hat, müssen Sie den vollständigen Namen der Datei (z. B. `meinsystemscript.bat`) angeben. Dies ist eine Java-Codeeinschränkung.

5. Setzen Sie wie folgt den Befehl für den Controller ab:

Für Cisco CSS Controller

```
cococontrol consultant start
```

Für Nortel Alteon Controller

```
nalcontrol consultant start
```

Anmerkung: Gewährleisten Sie wie folgt die Sicherheit:

- Erstellen Sie auf der Controllermaschine mit dem Befehl **lbkeys create** Schlüsseldateien. Weitere Informationen zu **lbkeys** finden Sie im Abschnitt „Remote Method Invocation (RMI)“ auf Seite 264.
- Kopieren Sie den resultierenden Schlüsselring auf der Servermaschine in das Verzeichnis `...ibm/edge/lb/admin/key`. Stellen Sie sicher, dass die Berechtigungen für den Schlüsselring dem Benutzer `root` den Lesezugriff ermöglichen.
- Agent Metric Server (Seite der Servermaschine)

1. Installieren Sie das Metric-Server-Paket aus dem Installationsverzeichnis von Load Balancer.
2. Überprüfen Sie anhand des Scripts **metricserver** im Verzeichnis **/usr/bin**, ob der gewünschte RMI-Port verwendet wird. (Auf Windows-Systemen lautet das Verzeichnis C:\WINNT\SYSTEM32.) Der Standard-RMI-Port ist 10004.

Anmerkung: Der für den RMI-Port angegebene Wert muss mit dem RMI-Port-Wert für Metric Server auf der Controllermaschine übereinstimmen.

3. Die beiden folgenden Scripts stehen zur Verfügung: **cpuload** (gibt den Prozentsatz der verwendeten CPU im Bereich von 0-100 zurück) und **memload** (gibt den Prozentsatz des belegten Speichers im Bereich von 0-100 zurück). Diese Scripts befinden sich im Verzeichnis **...ibm/edge/lb/ms/script**.

Optional können Sie Ihre eigenen angepassten Script-Dateien für Messwerte schreiben, in denen definiert ist, welchen Befehl Metric Server auf den Servermaschinen absetzen soll. Vergewissern Sie sich, dass alle angepassten Scripts ausführbar sind und sich im Verzeichnis **...ibm/edge/lb/ms/script** befinden. Angepasste Scripts **müssen** einen numerischen Lastwert zurückgeben.

Anmerkung: Ein angepasstes Script für Messwerte muss ein gültiges Programm oder Script mit der Erweiterung **.bat** oder **.cmd** sein. Auf Linux- und UNIX-Systemen müssen Scripts mit der Shell-Deklaration beginnen, da sie sonst möglicherweise nicht richtig ausgeführt werden.

4. Starten Sie den Agenten durch Absetzen des Befehls **metricserver**.
5. Geben Sie zum Stoppen des Agenten Metric Server **metricserver stop** ein.

Wenn Metric Server für eine vom lokalen Host abweichende Adresse ausgeführt werden soll, editieren Sie die Datei **metricserver** auf der am Lastausgleich beteiligten Servermaschine. Fügen Sie in der Datei **metricserver** nach dem Eintrag **java** Folgendes ein:

```
-Djava.rmi.server.hostname=andere_Adresse
```

Fügen Sie außerdem vor den Anweisungen if Folgendes zur Datei **metricserver** hinzu: **hostname *andere_Adresse***.

Auf Windows-Systemen: Geben Sie in Microsoft Stack einen Aliasnamen für *andere_Adresse* an. Informationen zum Angeben eines Aliasnamens für eine Adresse in Microsoft Stack finden Sie auf Seite 211.

Advisor-Funktion Workload Manager

WLM ist Code, der auf MVS-Großrechnern ausgeführt wird. Er kann abgefragt werden, um die Belastung auf der MVS-Maschine zu bestimmen.

Wurde MVS Workload Management auf Ihrem OS/390-System konfiguriert, können die Controller die Kapazitätsinformationen von WLM akzeptieren und die Informationen für den Lastausgleich verwenden. Mit der Advisor-Funktion WLM öffnen die Controller regelmäßig Verbindungen über den WLM-Port der einzelnen Server in der Consultant-Hosttabelle und akzeptiert die zurückgegebenen ganzzahligen Kapazitätswerte. Da diese ganzen Zahlen die noch verfügbare Kapazität darstellen und die Controller Werte erwarten, die die Belastung auf jeder Maschine angeben, werden die ganzzahligen Kapazitätswerte vom Advisor in Lastwerte

umgekehrt und normalisiert. (Ein hoher ganzzahliger Kapazitätswert und ein niedriger Lastwert geben beispielsweise beide einen akzeptablen Zustand eines Servers an). Es gibt mehrere wichtige Unterschiede zwischen dem WLM-Advisor und anderen Advisor-Funktionen des Controllers:

1. Andere Advisor-Funktionen öffnen Verbindungen zu den Servern unter Verwendung des Ports, über den der normale Client-Datenverkehr fließt. Die WLM-Advisor-Funktion benutzt für das Öffnen von Verbindungen zu den Servern nicht den für normalen Datenverkehr verwendeten Port. Der WLM-Agent muss auf den einzelnen Servermaschinen so konfiguriert werden, dass er an dem Port empfangsbereit ist, an dem die WLM-Advisor-Funktion des Controllers gestartet wurde. Der Standard-WLM-Port ist 10007.
2. Es ist möglich, protokollspezifische Advisor-Funktionen zusammen mit der Advisor-Funktion WLM zu verwenden. Die protokollspezifischen Advisor-Funktionen fragen die Server an den regulären Ports für Datenverkehr ab. Die WLM-Advisor-Funktion fragt die Systembelastung dagegen am WLM-Port ab.

Binäre Protokolle für die Analyse von Serverstatistiken verwenden

Mit dem Feature für binäre Protokollierung können Serverinformationen in Binärdateien gespeichert werden. Diese Dateien können dann verarbeitet werden, um die Serverinformationen zu analysieren, die über einen bestimmten Zeitraum gesammelt wurden.

Die folgenden Informationen werden für jeden in der Konfiguration definierten Server in dem binären Protokoll gespeichert:

- Übergeordnete Einheit (ID_für_Eignerangaben für Cisco CSS Controller; Service-ID für Nortel Alteon Controller)
- Server-ID
- Serveradresse
- Server-Port
- Serverwertigkeit
- Anzahl der für diesen Server konfigurierten Messgrößen
- Liste der Messwerte

Zum Protokollieren von Informationen in den binären Protokollen muss der Consultant aktiv sein.

Verwenden Sie zum Konfigurieren der binären Protokollierung die Befehlsgruppe **xxxcontrol consultant binarylog**.

- `binarylog start`
- `binarylog stop`
- `binarylog report`
- `binarylog set interval <Sekunden>`
- `binarylog set retention <Stunden>`

Mit der Option 'start' wird die Protokollierung von Serverinformationen in binären Protokollen im Protokollverzeichnis gestartet. Ein Protokoll wird zu Beginn jeder Stunde mit dem Datum und der Uhrzeit als Name der Datei erstellt.

Mit der Option 'stop' wird die Protokollierung von Serverinformationen in binären Protokollen gestoppt. Standardmäßig ist der Protokolldienst gestoppt.

Mit der Option 'set interval' wird gesteuert, wie oft Informationen in die Protokolle geschrieben werden. Der Consultant sendet in jedem Consultant-Intervall Serverdaten an den Protokollserver. Die Daten werden nur in die Protokolle geschrieben, wenn seit dem Schreiben des letzten Protokolleintrags die für das Protokollintervall angegebene Zeit in Sekunden verstrichen ist. Standardmäßig wird das Protokollierungsintervall auf 60 Sekunden gesetzt.

Zwischen den Einstellungen für das Consultant-Intervall und das Protokollierungsintervall gibt es eine gewisse Interaktion. Da dem Protokollserver Informationen nicht schneller zur Verfügung gestellt werden, als dies im Consultant-Intervall (in Sekunden) angegeben ist, wird durch Angabe eines Protokollierungsintervalls, das kleiner als das Consultant-Intervall ist, das Protokollierungsintervall de facto auf denselben Wert wie das Consultant-Intervall gesetzt.

Mit dieser Protokollierungstechnik können Sie Serverinformationen detaillierter erfassen. Sie können alle vom Consultant festgestellten Änderungen der Serverinformationen für die Berechnung von Serverwertigkeiten erfassen. Dieser Informationsumfang ist jedoch wahrscheinlich nicht erforderlich, um die Serverauslastung und Trends zu analysieren. Werden Serverinformationen alle 60 Sekunden protokolliert, erhalten Sie Momentaufnahmen von Serverinformationen in Abhängigkeit vom zeitlichen Verlauf. Wird das Protokollierungsintervall auf einen sehr niedrigen Wert gesetzt, kann dies zu großen Datenmengen führen.

Mit der Option 'set retention' wird gesteuert, wie lange Protokolldateien aufbewahrt werden. Protokolldateien, die älter als die angegebene Verweildauer (Stunden) sind, werden von dem Protokollserver gelöscht. Die geschieht nur, wenn der Protokollserver vom Consultant aufgerufen wird. Wenn Sie den Consultant stoppen, werden alte Protokolldateien demzufolge nicht gelöscht.

Im Verzeichnis **...ibm/edge/lb/servers/samples/BinaryLog** stehen ein Beispiel-Java-Programm und eine Beispielbefehlsdatei zur Verfügung. Dieses Beispiel zeigt, wie alle Informationen aus den Protokolldateien abgerufen und angezeigt werden können. Es kann für jede Art von Datenanalyse angepasst werden.

Beispiel für die Verwendung des bereitgestellten Scripts und Programms:

```
xxxlogreport 2002/05/01 8:00 2002/05/01 17:00
```

Dieser Befehl liefert einen Bericht mit den Serverdaten des Controllers vom 1. Mai 2002 in der Zeit von 8.00 bis 17.00 Uhr.

Scripts zum Generieren eines Alerts oder Protokollieren eines Serverausfalls verwenden

Load Balancer stellt Benutzer-Exits bereit, die Scripts aktivieren, die von Ihnen angepasst werden können. Sie können Scripts für die Ausführung automatisierter Aktionen erstellen. Eine solche Aktion wäre beispielsweise das Informieren eines Administrators über inaktive Server per Alert oder das Registrieren eines Ausfalls. Scripts, die Sie anpassen können, finden Sie im Installationsverzeichnis **...ibm/edge/lb/servers/samples**. Zum Ausführen müssen Sie die Dateien in das Verzeichnis **...ibm/edge/lb/servers/bin** kopieren und dann entsprechend den Anweisungen im Script umbenennen.

Die folgenden Beispiel-Scripts stehen zur Verfügung, bei denen die Angabe **xxx** durch **cco** für Cisco CSS Controller und **nal** für Nortel Alteon Controller zu ersetzen ist:

- **xxxserverdown** — Ein Server wird vom Controller als inaktiv markiert.
- **xxxserverUp** — Ein Server wird vom Controller als aktiv markiert.
- **xxxallserversdown** — Alle Server für einen bestimmten Service werden als inaktiv markiert.

Teil 8. Verwaltung von Load Balancer und Fehlerbehebung

Dieser Teil enthält Informationen zur Verwaltung von Load Balancer und zur Fehlerbehebung. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 24, „Betrieb und Verwaltung von Load Balancer“, auf Seite 263
- Kapitel 25, „Fehlerbehebung“, auf Seite 283

Kapitel 24. Betrieb und Verwaltung von Load Balancer

Anmerkung: Falls Sie die Komponente Dispatcher *nicht* verwenden, ersetzen Sie beim Lesen der allgemeinen Abschnitte dieses Kapitels, die sich nicht auf eine bestimmte Komponente beziehen, dscontrol und dsserver durch Folgendes:

- Für CBR: **cbrcontrol** und **cbrserver**
- Für Site Selector: **sscontrol** und **ssserver**
- Für Cisco CSS Controller: **ccocontrol** und **ccoserver**
- Für Nortel Alteon Controller: **nalcontrol** und **nalserver**

Wichtiger Hinweis: Falls Sie Load Balancer für IPV4 und IPV6 verwenden, sollten Sie sich vor der Beschäftigung mit dem Inhalt dieses Kapitels in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81 über Einschränkungen und Konfigurationsunterschiede informieren.

Das vorliegende Kapitel erläutert die Verwendung und Verwaltung von Load Balancer und ist in folgende Abschnitte untergliedert:

- „Fernverwaltung von Load Balancer“
 - „Remote Method Invocation (RMI)“ auf Seite 264
 - „Webgestützte Verwaltung“ auf Seite 265
- „Protokolle von Load Balancer verwenden“ auf Seite 267
 - „Für Dispatcher, CBR und Site Selector“ auf Seite 267
 - „Für Cisco CSS Controller und Nortel Alteon Controller“ auf Seite 269
- „Komponente Dispatcher verwenden“ auf Seite 270
 - „Simple Network Management Protocol mit Dispatcher verwenden“ auf Seite 271
- „Komponente Content Based Routing verwenden“ auf Seite 279
- „Site Selector verwenden“ auf Seite 280
- „Cisco CSS Controller verwenden“ auf Seite 280
- „Nortel Alteon Controller verwenden“ auf Seite 281

Fernverwaltung von Load Balancer

Load Balancer bietet zwei Möglichkeiten an, die Konfigurationsprogramme auf einer anderen Maschine als der Maschine mit Load Balancer auszuführen. Für die Kommunikation zwischen den Konfigurationsprogrammen (dscontrol, cbrcontrol, sscontrol, ccocontrol, nalcontrol) und dem Server (dsserver, cbrserver usw.) bestehen die beiden folgenden Möglichkeiten:

- Java Remote Method Invocation (RMI)
- webgestützte Verwaltung

Der Vorteil der Fernverwaltung mit RMI besteht darin, dass sie schneller als die webgestützte Verwaltung ist.

Die webgestützte Verwaltung hat den Vorteil, dass sie eine sichere Fernverwaltung mit Authentifizierung ist und auch bei Vorhandensein einer Firewall die Kommunikation mit der Load-Balancer-Maschine gewährleistet ist. Diese Verwaltungs-

methode erfordert außerdem *nicht* die Installation und Verwendung von Authentifizierungsschlüsseln (lbkeys) auf der fernen Client-Maschine, die mit der Load-Balancer-Maschine kommuniziert.

Remote Method Invocation (RMI)

Bei Verwendung von RMI lautet der Befehl zum Herstellen einer Verbindung zu einer Load-Balancer-Maschine für die Fernverwaltung **dscontrol host:ferner_Host**.

Stammt der RMI-Aufruf von einer anderen Maschine als der lokalen Maschine, muss eine Authentifizierung mit öffentlichem/privatem Schlüssel stattfinden, bevor der Konfigurationsbefehl akzeptiert wird.

Die Kommunikation zwischen den Steuerprogrammen, die auf derselben Maschine wie die Komponentenserver ausgeführt werden, wird nicht authentifiziert.

Verwenden Sie den folgenden Befehl, um öffentliche und private Schlüssel für die ferne Authentifizierung zu generieren:

lbkeys [create | delete]

Dieser Befehl muss auf derselben Maschine wie Load Balancer ausgeführt werden.

Bei Verwendung der Option **create** wird im Schlüsselverzeichnis des Servers (...**ibm/edge/lb/servers/key/**) ein privater Schlüssel erstellt. Im Verzeichnisverzeichnis für Schlüssel (...**ibm/edge/lb/admin/keys/**) der einzelnen Komponenten von Load Balancer werden öffentliche Schlüssel erstellt. Der Dateiname für den öffentlichen Schlüssel ist *Komponente-Serveradresse-RMI-Port*. Diese öffentlichen Schlüssel müssen anschließend zu den fernen Clients transportiert und in das Verzeichnisverzeichnis für Schlüssel gestellt werden.

Auf einer Load-Balancer-Maschine mit dem Hostnamen bzw. der Adresse 10.0.0.25, die für jede Komponente den Standard-RMI-Port verwendet, generiert der Befehl **lbkeys create** die folgenden Dateien:

- Privater Schlüssel: ...**ibm/edge/lb/servers/key/authorization.key**
- Öffentliche Schlüssel:
 - ...**ibm/edge/lb/admin/keys/dispatcher-10.0.0.25-10099.key**
 - ...**ibm/edge/lb/admin/keys/cbr-10.0.0.25-11099.key**
 - ...**ibm/edge/lb/admin/keys/ss-10.0.0.25-12099.key**
 - ...**ibm/edge/lb/admin/keys/cco-10.0.0.25-13099.key**
 - ...**ibm/edge/lb/admin/keys/na1-10.0.0.25-14099.key**

Die Verwaltungsdateigruppe wurde auf einer anderen Maschine installiert. Die Dateien der öffentlichen Schlüssel müssen auf der fernen Client-Maschine in das Verzeichnis ...**ibm/edge/lb/admin/keys** gestellt werden.

Jetzt ist der ferne Client berechtigt, Load Balancer auf der Maschine 10.0.0.25 zu konfigurieren.

Dieselben Schlüssel müssen Sie auf allen fernen Clients verwenden, die berechtigt sein sollen, Load Balancer auf der Maschine 10.0.0.25 zu konfigurieren.

Würde der Befehl **lbkeys create** erneut ausgeführt, hätte dies die Generierung einer neuen Gruppe von öffentlichen/privaten Schlüsseln zur Folge. Dies würde bedeuten, dass alle fernen Clients, die unter Verwendung der vorherigen Schlüssel

die Herstellung einer Verbindung versuchen, nicht berechtigt wären. Der neue Schlüssel müsste in das korrekte Verzeichnis auf den Clients gestellt werden, die erneut berechtigt werden sollen.

Mit dem Befehl **lbkeys delete** werden die privaten und öffentlichen Schlüssel von der Servermaschine gelöscht. Werden diese Schlüssel gelöscht, sind keine fernen Clients mehr berechtigt, eine Verbindung zu den Servern herzustellen.

Für lbkeys create und lbkeys delete gibt es die Option **force**. Die Option force unterdrückt die Eingabeaufforderungen, die von Ihnen eine Bestätigung für das Überschreiben oder Löschen der vorhandenen Schlüssel anfordern.

Wenn Sie eine RMI-Verbindung hergestellt haben, können Sie von einer Eingabeaufforderung aus über die Befehle dscontrol, cbrcontrol, sscontrol, ccocontrol, nalcontrol, dswizard, cbrwizard und sswizard mit den Konfigurationsprogrammen kommunizieren. Sie können Load Balancer auch auf der GUI konfigurieren. Geben Sie dazu an einer Eingabeaufforderung lbadmin ein.

Anmerkung: Aufgrund von Änderungen an den Sicherheitspaketen der Java-Version sind die Schlüssel, die Load Balancer für ältere Releases als Version 5.1.1 generiert hat, möglicherweise nicht mit den Schlüsseln für das aktuelle Release kompatibel. Wenn Sie ein neues Release installieren, müssen Sie die Schlüssel neu generieren.

Webgestützte Verwaltung

Voraussetzungen

Für die webgestützte Verwaltung ist auf der **Client-Maschine**, die die Fernverwaltung durchführt, Folgendes erforderlich:

- JRE ab Version 1.3.0
- Informationen zu unterstützten Browsern finden Sie auf der folgenden Webseite: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

Anmerkung: Wenn Sie Netscape verwenden, dürfen Sie nicht die Größe des Netscape-Browserfensters ändern, in dem die Load-Balancer-GUI angezeigt wird. Das heißt, Sie dürfen das Fenster nicht minimieren, maximieren, wiederherstellen usw. Da Netscape bei jeder Größenänderung des Browserfensters die Seite neu lädt, kommt es zu einer Trennung der Hostverbindung, die nach einer solchen Änderung demzufolge neu hergestellt werden muss.

Auf der **Hostmaschine**, auf die Sie zugreifen, um die webgestützte Fernverwaltung durchzuführen, ist Folgendes erforderlich:

- Caching Proxy Version 6
- Perl ab Version 5.5

Caching Proxy konfigurieren

- Für Caching Proxy benötigen Sie das Dienstprogramm IBM Key Management (iKeyman) oder ein anderes Dienstprogramm, um SSL-Serverzertifikate zu erstellen. (Informationen zum Erstellen der Zertifikate können Sie dem *Caching Proxy Administratorhandbuch* entnehmen.)
- Fügen Sie zum Abschnitt "Webgestützte Verwaltung von Load Balancer" der Caching-Proxy-Konfigurationsdatei (ibmpoxy.conf) die folgenden Anweisungen nach der Definition der Schutzdomänen und vor den Zuordnungsregeln hinzu:

Windows-Systeme —

```
Protect /lb-admin/lbwebaccess PROT-ADMIN
Exec /lb-admin/lbwebaccess C:\PROGRA~1\IBM\edge\lb\admin\lbwebaccess.pl
Pass /lb-admin/help/* C:\PROGRA~1\IBM\edge\lb\admin\help\*
Pass /lb-admin/*.jar C:\PROGRA~1\IBM\edge\lb\admin\lib\*.jar
Pass /lb-admin/* C:\PROGRA~1\IBM\edge\lb\admin\*
Pass /documentation/Sprache/* C:\PROGRA~1\IBM\edge\lb\documentation\Sprache\*
```

Sprache steht hier für das Sprachenunterverzeichnis (z. B. en_US).

Linux- und UNIX-Systeme —

```
Protect /lb-admin/lbwebaccess PROT-ADMIN
Exec /lb-admin/lbwebaccess /opt/ibm/edge/lb/admin/lbwebaccess.pl
Pass /lb-admin/help/* /opt/ibm/edge/lb/admin/help/*
Pass /lb-admin/*.jar /opt/ibm/edge/lb/admin/lib/*.jar
Pass /lb-admin/* /opt/ibm/edge/lb/admin/*
Pass /documentation/Sprache/* /opt/ibm/edge/lb/documentation/Sprache/*
```

Anmerkung: Auf HP-UX-Systemen geht das Script lbwebaccess.pl davon aus, dass sich die Perl-Binärdatei im Verzeichnis /usr/bin/ befindet. (Die erste Zeile des Scripts enthält die Angabe #!/usr/bin/perl.) Aktualisieren Sie diese Angabe mit dem Verzeichnis, das die Anwendung Perl enthält. Alternativ dazu können Sie eine symbolische Verbindung erstellen. Wenn Perl beispielsweise im Verzeichnis /opt/perl/bin/perl installiert ist, führen Sie den folgenden Befehl aus:

```
ln -s /opt/perl/bin/perl /usr/bin/perl
```

Zugriff auf die webgestützte Verwaltung und ihre Ausführung

Wenn Sie die webgestützte Verwaltung ausführen möchten, müssen Sie sie auf der Hostmaschine mit Load Balancer starten. Setzen Sie dazu an der Eingabeaufforderung der Hostmaschine den Befehl **lbwebaccess** ab.

Sie benötigen die Benutzer-ID und das Kennwort für die Hostmaschine, auf die Sie fern zugreifen. Diese Angaben stimmen mit der Benutzer-ID und dem Kennwort für die Verwaltung von Caching Proxy überein.

Greifen Sie im Webbrowser auf den folgenden URL des fernen Standortes zu, um die webgestützte Verwaltung von Load Balancer aufzurufen:

`http://Hostname/lb-admin/lbadmin.html`

Hostname ist der Name der Maschine, auf die Sie zugreifen, um mit Load Balancer zu kommunizieren.

Wenn die Webseite geladen ist, erscheint im Browserfenster die GUI von Load Balancer, so dass Sie die webgestützte Fernverwaltung durchführen können.

Von der Load-Balancer-GUI aus können Sie auch Steuerbefehle für die Konfiguration absetzen. Gehen Sie dazu wie folgt vor:

1. Heben Sie in der Baumstruktur der GUI den Hostknoten hervor.
2. Wählen Sie im Popup-Menü "Host" den Eintrag **Befehl senden...** aus.
3. Geben Sie im Befehlseingabefeld den gewünschten Befehl ein, z. B. **executor report**. In einem Fenster sehen Sie die Ergebnisse und die Historie der in der aktuellen Sitzung ausgeführten Befehle.

Konfiguration fern aktualisieren

Wenn bei der fernen webgestützten Verwaltung mehrere Administratoren die Load-Balancer-Konfiguration von verschiedenen Standorten aus aktualisieren, müssen Sie die Konfiguration aktualisieren, um (z. B.) den von einem anderen Administrator hinzugefügten (oder gelöschten) Cluster, Port oder Server zu sehen. Die GUI für die webgestützte Fernverwaltung bietet die Funktionen **Konfiguration aktualisieren** und **Alle Konfigurationen aktualisieren** an.

Gehen Sie zum Aktualisieren der Konfiguration auf der webgestützten GUI wie folgt vor:

- Für einen Host: Klicken Sie mit der rechten Maustaste in der Baumstruktur der GUI auf einen **Hostknoten** und wählen Sie **Konfiguration aktualisieren** aus.
- Für alle Hosts: Wählen Sie im Menü **Datei** aus und dann **Alle Konfigurationen aktualisieren**.

Protokolle von Load Balancer verwenden

Für Dispatcher, CBR und Site Selector

Load Balancer sendet Einträge an ein Serverprotokoll, ein Manager-Protokoll, an das Protokoll eines Metriküberwachungsprogramms (protokollbezogene Kommunikation mit Metric-Server-Agenten) und an das Protokoll jeder von Ihnen verwendeten Advisor-Funktion.

Anmerkung: Zusätzlich können Einträge in ein Subagentenprotokoll (SNMP) gestellt werden. Dies gilt allerdings nur für die Komponente Dispatcher.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode `cbr` der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (`cbr`)“ auf Seite 55.

Sie können die Protokollstufe festlegen, um den Umfang der Nachrichten zu definieren, die in das Protokoll geschrieben werden. Bei Stufe 0 werden Fehler protokolliert. Load Balancer protokolliert außerdem Header und Datensätze von Ereignissen, die nur einmal eintreten. (Beim Starten einer Advisor-Funktion wird beispielsweise eine Nachricht in das Manager-Protokoll geschrieben.) Bei Stufe 1 werden weitere Informationen aufgenommen. Bis Stufe 5 nimmt die Ausführlichkeit kontinuierlich zu. Bei Stufe 5 werden alle generierten Nachrichten aufgenommen, damit sie im Falle eines Fehlers für das Debugging verwendet werden können. Die Standardeinstellung für das Manager-Protokoll, das Protokoll der Advisor-Funktionen, das Serverprotokoll sowie das Protokoll der Subagenten ist 1.

Zudem können Sie die maximale Größe eines Protokolls festlegen. Wenn Sie eine maximale Größe für die Protokolldatei festlegen, findet ein Dateiumlauf statt. Hat die Datei die angegebene Größe erreicht, werden alle weiteren Einträge wieder an den Anfang der Datei geschrieben und die dort befindlichen Einträge überschrieben. Sie können für die Protokollgröße keinen Wert angeben, der kleiner als der aktuelle Wert für die Protokollgröße ist. Protokolleinträge werden mit einer Zeitmarke versehen, so dass Sie erkennen können, in welcher Reihenfolge sie geschrieben wurden.

Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen. Bei Stufe 0 ist es wahrscheinlich sicher, die Standardprotokollgröße von 1 MB zu verwenden. Ab Stufe 3 sollten Sie die Größe jedoch begrenzen. Bedenken Sie aber, dass bei einem zu kleinen Wert die Protokollierung nicht mehr sinnvoll ist.

- Konfigurieren Sie die Protokollstufe oder die maximale Größe eines Serverprotokolls mit dem Befehl **dscontrol set**. (Die Einstellungen für das Serverprotokoll können Sie mit dem Befehl **dscontrol logstatus** anzeigen.)
- Für ein Manager-Protokoll können Sie die Protokollstufe oder die maximale Größe mit dem Befehl **dscontrol manager** konfigurieren.
- Für das Protokoll der Metriküberwachung, in dem die Kommunikation mit den Metric-Server-Agenten protokolliert wird, können Sie die Protokollstufe oder die maximale Größe mit dem Befehl **dscontrol manager metric set** konfigurieren.
- Für das Protokoll einer Advisor-Funktion können Sie die Protokollstufe oder die maximale Größe mit dem Befehl **dscontrol advisor** konfigurieren.
- Konfigurieren Sie die Protokollstufe oder die maximale Größe eines Subagentenprotokolls mit dem Befehl **dscontrol subagent**. (Der SNMP-Subagent wird nur von der Komponente Dispatcher verwendet.)

Pfade für die Protokolldatei ändern

Die von Load Balancer generierten Protokolle werden standardmäßig im Unterverzeichnis logs des Installationsverzeichnisses von Load Balancer gespeichert. Wenn Sie diesen Pfad ändern möchten, setzen Sie die Variable *lb_logdir* im dsserver-Script entsprechend.

AIX-, HP-UX-, Linux- und Solaris-Systeme: Sie finden das dsserver-Script im Verzeichnis /usr/bin. In diesem Script ist die Variable *lb_logdir* auf das Standardverzeichnis gesetzt. Sie können diese Variable ändern, um Ihr Protokollverzeichnis anzugeben. Beispiel:

LB_LOGDIR=/pfad/zu/meinen/protokollen/

Windows-Systeme: Sie finden die dsserver-Datei im Windows-Systemverzeichnis (C:\WINNT\SYSTEM32 für Windows 2003). In der dsserver-Datei ist die Variable *lb_logdir* auf das Standardverzeichnis gesetzt. Sie können diese Variable ändern, um Ihr Protokollverzeichnis anzugeben. Beispiel:

set LB_LOGDIR=c:\pfad\zu\meinen\protokollen

Für alle Betriebssysteme ist sicherzustellen, dass sich rechts und links vom Gleichheitszeichen keine Leerzeichen befinden und dass der Pfad mit einem Schrägstrich endet ("/" bzw. "\").

Binäres Protokollieren

Anmerkung: Die binäre Protokollierung ist für die Komponente Site Selector nicht möglich.

Für die binäre Protokollierung von Load Balancer wird dasselbe Verzeichnis (log) wie für die übrigen Protokolldateien verwendet. Lesen Sie hierzu die Informationen im Abschnitt „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 241.

Für Cisco CSS Controller und Nortel Alteon Controller

Sie können die Protokollstufe festlegen, um den Umfang der Nachrichten zu definieren, die in das Protokoll geschrieben werden. Bei Stufe 0 werden Fehler protokolliert. Load Balancer protokolliert außerdem Header und Datensätze von Ereignissen, die nur einmal eintreten. (Beim Starten einer Advisor-Funktion wird beispielsweise eine Nachricht in das Consultant-Protokoll geschrieben.) Bei Stufe 1 werden weitere Informationen aufgenommen. Bis Stufe 5 nimmt die Ausführlichkeit kontinuierlich zu. Bei Stufe 5 werden alle generierten Nachrichten aufgenommen, damit sie im Falle eines Fehlers für das Debugging verwendet werden können. Der Standardwert für die Protokolle ist 1.

Zudem können Sie die maximale Größe eines Protokolls festlegen. Wenn Sie eine maximale Größe für die Protokolldatei festlegen, findet ein Dateiumlauf statt. Hat die Datei die angegebene Größe erreicht, werden alle weiteren Einträge wieder an den Anfang der Datei geschrieben und die dort befindlichen Einträge überschrieben. Sie können für die Protokollgröße keinen Wert angeben, der kleiner als der aktuelle Wert für die Protokollgröße ist. Protokolleinträge werden mit einer Zeitmarke versehen, so dass Sie erkennen können, in welcher Reihenfolge sie geschrieben wurden.

Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen. Bei Stufe 0 ist es wahrscheinlich sicher, die Standardprotokollgröße von 1 MB zu verwenden. Ab Stufe 3 sollten Sie die Größe jedoch begrenzen. Bedenken Sie aber, dass bei einem zu kleinen Wert die Protokollierung nicht mehr sinnvoll ist.

Controllerprotokolle

Für Cisco CSS Controller und Nortel Alteon Controller gibt es die folgenden Protokolle:

- Controllerprotokoll (Befehl **controller set**)
- Consultant-Protokoll (Befehl **consultant set**)
- Protokoll für hohe Verfügbarkeit (Befehl **highavailability set**)
- Protokoll der Messwerterfassung (Befehl **metriccollector set**)
- Binärprotokoll (Befehl **consultant binarylog**)

Das folgende Beispiel zeigt, wie Sie für das Protokoll der Metriküberwachung, in dem die Kommunikation mit den Metric-Server-Agenten protokolliert wird, die Protokollstufe und die maximale Größe konfigurieren können:

```
xxxcontrol metriccollector set Consultant-ID:Service-ID:Messwertname  
loglevel x logsize y
```

Pfade für die Protokolldatei ändern

Die von den Controllern generierten Protokolle werden standardmäßig im Unterverzeichnis logs des Installationsverzeichnisses der Controller gespeichert. Wenn Sie diesen Pfad ändern möchten, setzen Sie die Variable `xxx_logdir` im xxxserver-Script entsprechend.

AIX-, HP-UX-, Linux- und Solaris-Systeme: Sie finden das xxxserver-Script im Verzeichnis /usr/bin. In diesem Script ist die Variable `xxx_logdir` auf das Standardverzeichnis gesetzt. Sie können diese Variable ändern, um Ihr Protokollverzeichnis anzugeben. Beispiel:

```
xxx_LOGDIR=/pfad/zu/meinen/protokollen/
```

Windows-Systeme: Sie finden die xxxserver-Datei im Windows-Systemverzeichnis (in der Regel C:\WINNT\SYSTEM32). In der xxxserver-Datei ist die Variable *xxx_logdir* auf das Standardverzeichnis gesetzt. Sie können diese Variable ändern, um Ihr Protokollverzeichnis anzugeben. Beispiel:

```
set xxx_LOGDIR=c:\pfad\zu\meinen\protokollen\
```

Für alle Betriebssysteme ist sicherzustellen, dass sich rechts und links vom Gleichheitszeichen keine Leerzeichen befinden und dass der Pfad mit einem Schrägstrich endet ("/" bzw. "\").

Binäres Protokollieren

Für die binäre Protokollierung von Load Balancer wird dasselbe Verzeichnis (log) wie für die übrigen Protokolldateien verwendet. Lesen Sie hierzu die Informationen im Abschnitt „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 241.

Komponente Dispatcher verwenden

Die folgenden Abschnitte erläutern die Verwendung und Verwaltung der Komponente Dispatcher.

Dispatcher starten und stoppen

- Geben Sie zum Starten des Dispatchers in einer Befehlszeile **dsserver** ein.
- Geben Sie zum Stoppen des Dispatchers in einer Befehlszeile **dsserver stop** ein.

Inaktivitätszeitlimit verwenden

Load Balancer betrachtet Verbindungen als veraltet, wenn sie die durch das Inaktivitätszeitlimit angegebene Zeit (Sekunden) lang inaktiv waren. Wird das Inaktivitätszeitlimit überschritten, entfernt Load Balancer den Eintrag für diese Verbindung aus seinen Tabellen und löscht den nachfolgenden Datenverkehr für diese Verbindung.

Auf Port-Ebene können Sie das Inaktivitätszeitlimit beispielsweise mit dem Befehl **dscontrol port set staletimeout** angeben.

Das Inaktivitätszeitlimit kann auf Executor-, Cluster- und Port-Ebene gesetzt werden. Auf Executor- und Clusterebene liegt der Standardwert bei 300 Sekunden. Es wird bis hinunter zum Port gefiltert. Auf Port-Ebene ist der Standardwert vom jeweiligen Port abhängig. Einige herkömmliche Ports haben unterschiedliche Inaktivitätszeitlimits. Der Telnet-Port 23 hat beispielsweise ein Standardlimit von 259.200 Sekunden.

Dienste können auch eigene Inaktivitätszeitlimits haben. Für LDAP (Lightweight Directory Access Protocol) gibt es z. B. den Konfigurationsparameter *idletimeout*. Bei Überschreitung der von *idletimeout* angegebenen Zeit in Sekunden wird die Beendigung einer inaktiven Client-Verbindung erzwungen. Das Inaktivitätszeitlimit (*idletimeout*) kann auch auf 0 gesetzt werden, so dass Verbindungen nicht zwangsweise beendet werden können.

Wenn das Inaktivitätszeitlimit von Load Balancer kleiner als das des Dienstes ist, können Konnektivitätsprobleme auftreten. Im Falle von LDAP liegt das Inaktivitätslimit von Load Balancer (*staletimeout*) standardmäßig bei 300 Sekunden. Ist die Verbindung 300 Sekunden inaktiv, entfernt Load Balancer den Eintrag für die Ver-

bindung aus seinen Tabellen. Wenn das Inaktivitätszeitlimit (idletimeout) über 300 Sekunden liegt (oder auf 0 gesetzt ist), könnte der Client davon ausgehen, dass er weiterhin mit dem Server verbunden ist. Wenn der Client Pakete sendet, werden diese von Load Balancer gelöscht. Das hat zur Folge, dass LDAP blockiert, wenn eine Anfrage an den Server gesendet wird. Sie können dieses Problem vermeiden, indem Sie das Inaktivitätszeitlimit von LDAP (idletimeout) auf einen Wert ungleich null setzen, der genauso groß wie das Inaktivitätszeitlimit von Load Balancer (staletimeout) oder kleiner als dieses ist.

Bereinigung von Verbindungssätzen mit fintimeout und staletimeout steuern

Ein Client sendet ein FIN-Paket, nachdem er alle Pakete gesendet hat, um dem Server mitzuteilen, dass die Transaktion beendet ist. Wenn der Dispatcher das FIN-Paket erhält, kennzeichnet er die Transaktion nicht mehr als AKTIV, sondern als BEENDET. Wenn eine Transaktion als BEENDET gekennzeichnet ist, kann der für die Verbindung reservierte Speicher bereinigt werden.

Sie können den Durchsatz bei der Reservierung und Wiederverwendung von Verbindungssätzen verbessern, indem Sie mit dem Befehl **executor set fintimeout** steuern, wie lange der Dispatcher Verbindungen im Status BEENDET in den Dispatcher-Tabellen als aktiv und empfangsbereit speichern soll. Sobald eine Verbindung im Status BEENDET den mit **fintimeout** festgelegten Zeitraum überschreitet, wird sie aus den Dispatcher-Tabellen entfernt und kann wiederverwendet werden. Das Zeitlimit für die Beendigung inaktiver Verbindungen können Sie mit dem Befehl **dscontrol executor set fincount** ändern.

Mit dem Befehl **dscontrol executor set staletimeout** können Sie steuern, wie lange der Dispatcher Verbindungen im Status HERGESTELLT in den Dispatcher-Tabellen als aktiv und empfangsbereit speichern soll, wenn kein aktiver Datenverkehr festgestellt werden kann. Weitere Informationen hierzu finden Sie im Abschnitt „Inaktivitätszeitlimit verwenden“ auf Seite 270.

Berichte der GUI — Menüoption 'Überwachen'

Ausgehend von den Informationen des Executors können mehrere Diagramme angezeigt und an den Manager übergeben werden. (Die Menüoption "Überwachen" der GUI erfordert, dass die Manager-Funktion aktiviert ist):

- Verbindungen pro Sekunde je Server (mehrere Server können in demselben Diagramm angezeigt werden)
- Relative Wertigkeiten der Server an einem bestimmten Port
- Durchschnittliche Verbindungsdauer pro Server an einem bestimmten Port

Simple Network Management Protocol mit Dispatcher verwenden

Ein Netzverwaltungssystem ist ein Programm, das ständig ausgeführt und verwendet wird, um ein Netz zu überwachen, den Status eines Netzes wiederzugeben und ein Netz zu steuern. Simple Network Management Protocol (SNMP), ein häufig verwendetes Protokoll für die Kommunikation mit Einheiten in einem Netz, ist der aktuelle Netzverwaltungsstandard. Die Netzeinheiten verfügen normalerweise über einen *SNMP-Agenten* und einen oder mehrere Subagenten. Der SNMP-Agent kommuniziert mit der *Netzverwaltungsstation* oder antwortet auf SNMP-Befehlszeilenanforderungen. Der *SNMP-Subagent* ruft Daten ab und aktualisiert die Daten und übergibt diese Daten an den SNMP-Agenten, der sie an den Requester weiterleitet.

Der Dispatcher stellt eine SNMP *Management Information Base* (ibmNetDispatcherMIB) und einen SNMP-Subagenten zur Verfügung. Damit wird Ihnen die Verwendung jedes Netzverwaltungssystems ermöglicht, wie beispielsweise Tivoli NetView, Tivoli Distributed Monitoring oder HP OpenView, um den Zustand, die Leistung und die Aktivität des Dispatchers zu überwachen. Die MIB-Daten beschreiben den Dispatcher, der verwaltet wird, und geben den aktuellen Status des Dispatchers wieder. Die MIB wird im Unterverzeichnis `..lb/admin/MIB` installiert.

Anmerkung: Die MIB `ibmNetDispatcherMIB.02` wird nicht mit dem Tivoli-NetView-Programm `xnmloadmib2` geladen. Um den Fehler zu beheben, müssen Sie den Bereich NOTIFICATION-GROUP der MIB auf Kommentar setzen. Geben Sie dazu am Beginn der Zeile "indMibNotifications Group NOTIFICATION-GROUP" sowie der sechs darauf folgenden Zeilen "- " ein.

Das Netzverwaltungssystem verwendet SNMP-Befehle GET, um MIB-Werte auf anderen Maschinen zu überprüfen. Es kann dann den Benutzer benachrichtigen, wenn angegebene Schwellenwerte überschritten werden. Sie können anschließend die Leistung des Dispatchers beeinflussen, indem Sie Konfigurationsdaten für den Dispatcher so ändern, dass Dispatcher-Probleme im voraus bestimmt oder berichtigt werden, bevor sie den Ausfall des Dispatchers oder Webservers zur Folge haben.

SNMP - Befehle und Protokoll

Das System stellt normalerweise einen SNMP-Agenten für jede Netzverwaltungsstation zur Verfügung. Der Benutzer sendet einen Befehl GET an den SNMP-Agenten. Dieser SNMP-Agent sendet dann einen Befehl GET, um die angegebenen MIB-Variablenwerte von einem Subagenten abzurufen, der für diese MIB-Variablen zuständig ist.

Der Dispatcher stellt einen Subagenten zur Verfügung, der MIB-Daten aktualisiert und abrufen. Der Subagent antwortet mit den entsprechenden MIB-Daten, wenn der SNMP-Agent einen Befehl GET sendet. Der SNMP-Agent überträgt die Daten an die Netzverwaltungsstation. Die Netzverwaltungsstation kann Sie benachrichtigen, wenn angegebene Schwellenwerte überschritten werden.

Die Dispatcher-SNMP-Unterstützung beinhaltet einen SNMP-Subagenten, der DPI-Funktionalität verwendet (DPI = Distributed Program Interface). DPI ist eine Schnittstelle zwischen einem SNMP-Agenten und seinen Subagenten. Das Betriebssystem Windows verwendet den Windows-Erweiterungsagenten als Schnittstelle zwischen einem SNMP-Agenten und seinen Subagenten.

SNMP auf AIX-, HP-UX, Linux- und Solaris-Systemen aktivieren

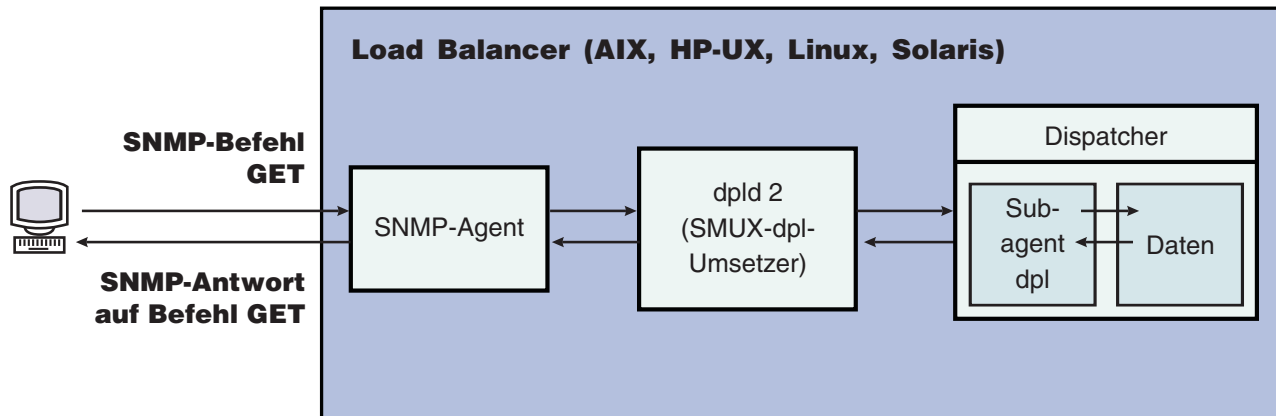


Abbildung 40. SNMP-Befehle für Linux- und UNIX-Systeme

AIX-Systeme stellen einen SNMP-Agenten bereit, der das SNMP-Multiplexer-Protokoll (SMUX) verwendet und die zusätzliche ausführbare Datei DPID2 anbietet, die als Umsetzer zwischen DPI und SMUX fungiert.

Für HP-UX-Systeme müssen Sie einen SMUX-fähigen SNMP-Agenten erwerben, da HP-UX keinen solchen bereitstellt. Load Balancer stellt DPID2 für HP-UX-Systeme bereit.

Linux-Systeme stellen einen SNMP-Agenten bereit, der SMUX verwendet. Im Lieferumfang der meisten Linux-Versionen (z. B. Red Hat) ist ein UCD-SNMP-Paket enthalten. UCD SNMP enthält ab Version 4.1 SMUX-fähige Agenten. Load Balancer stellt DPID2 für Linux-Systeme bereit.

Anmerkung: Für Systeme mit SuSE Linux müssen Sie einen SMUX-fähigen SNMP-Agenten erwerben, da dieses Betriebssystem keinen solchen bereitstellt.

Für Solaris-Systeme müssen Sie einen SMUX-fähigen SNMP-Agenten erwerben, da dieses Betriebssystem keinen solchen bereitstellt. Load Balancer stellt DPID2 für Solaris-Systeme im Verzeichnis `/opt/ibm/edge/lb/servers/samples/SNMP` bereit.

Den DPI-Agenten müssen Sie als Benutzer `root` ausführen. Bevor Sie den DPID2-Dämon ausführen, aktualisieren Sie die Dateien `/etc/snmpd.peers` und `/etc/snmpd.conf` wie folgt:

AIX- und Solaris-Systeme:

- Fügen Sie in der Datei `/etc/snmpd.peers` den folgenden Eintrag für `dpid` hinzu:
`"dpid2" 1.3.6.1.4.1.2.3.1.2.2.1.1.2 "dpid_password"`
- Fügen Sie in der Datei `/etc/snmpd.conf` den folgenden Eintrag für `dpid` hinzu:
`smux 1.3.6.1.4.1.2.3.1.2.2.1.1.2 dpid_password #dpid`

Für Linux-Systeme:

- Fügen Sie in der Datei `/etc/snmpd.peers` den folgenden Eintrag für `dpid` hinzu (ggf. müssen Sie die Datei neu erstellen):
`"dpid2" 1.3.6.1.4.1.2.3.1.2.2.1.1.2 "dpid_password"`

- Fügen Sie in der Datei `/etc/snmp/snmpd.conf` den folgenden Eintrag für `dpid` hinzu:

```
smuxpeer      .1.3.6.1.4.1.2.3.1.2.2.1.1.2      dpid_password
```

In der Datei `snmpd.conf` müssen Sie außerdem alle Zeilen auf Kommentar setzen, die mit den folgenden Wörter beginnen: `com2sec`, `group`, `view` oder `access`.

SNMP auf HP-UX-Systemen aktivieren

Führen Sie zum Installieren der SNMP-Unterstützung von HP-UX die folgenden Schritte aus:

1. Falls Sie keine Version von GNU SED installiert haben, laden Sie sie von der HP-Website <http://www.hp.com> herunter.
2. Laden Sie `ucd-snmp-4.2.4.tar.gz` von der Webseite http://sourceforge.net/project/showfiles.php?group_id=12694 herunter.
3. Vergewissern Sie sich, dass auf Ihrer Maschine `gcc` und `gmake` oder `make` installiert ist. Wenn das nicht der Fall ist, müssen Sie die Tools installieren.
4. Entpacken Sie die Datei `ucd-snmp-4.2.4.tar.gz`. Entpacken Sie dann alle Quellendateien in dem Verzeichnis.
5. Führen Sie im Verzeichnis mit den Quellendateien die folgenden Schritte aus:
 - a. `run ./configure --with-mib-modules=smux`
 - b. `make`
 - c. Führen Sie die beiden nächsten Befehle als Benutzer `root` aus:
 - 1) `umask 022`
 - 2) `make install`
 - d. `export SNMPCONFPATH=/etc/snmp`
 - e. `start /usr/local/sbin/snmpd -s` (Damit wird der SNMP-Agent gestartet.)
 - f. `start dpid2` (Damit wird der DPI-Umsetzer gestartet.)
 - g. `dscontrol subagent start` (Damit wird der Dispatcher-Subagent gestartet.)

SNMP auf Systemen mit SuSE Linux aktivieren

Um das Load-Balancer-SNMP auf Systemen mit SuSE Linux verwenden zu können, müssen Sie die folgenden Schritte ausführen:

1. Entfernen Sie vom SuSE-Rechner das Paket `ucd-snmp.rpm`.
2. Laden Sie `ucd-snmp-4.2.4.tar.gz` herunter (Adresse: http://sourceforge.net/project/showfiles.php?group_id=12694).
3. Vergewissern Sie sich, dass auf Ihrem SuSE-Rechner `gcc` und `gmake` oder `make` installiert ist. (Wenn das nicht der Fall ist, müssen Sie die Tools installieren.)
4. Entpacken Sie die Datei `ucd-snmp-4.2.4.tar.gz`. Entpacken Sie dann alle Quellendateien in dem Verzeichnis.
5. Führen Sie im Verzeichnis mit den Quellendateien die folgenden Schritte aus:
 - a. `run ./configure --with-mib-modules=smux`
 - b. `make`
 - c. Führen Sie die beiden nächsten Befehle als Benutzer `root` aus:
 - 1) `umask 022 #`
 - 2) `make install`
 - d. `export SNMPCONFPATH=/etc/snmp`
 - e. `start /usr/local/sbin/snmpd -s`
 - f. `start dpid2`

Aktualisieren Sie wie folgt snmpd (sofern der Dämon bereits aktiv ist), damit die Datei snmpd.conf neu gelesen wird:

```
refresh -s snmpd
```

Starten Sie den DPID-SMUX-Peer:

```
dpid2
```

Die Dämonen müssen in der folgenden Reihenfolge gestartet werden:

1. SNMP-Agent
2. DPI-Umsetzer
3. Dispatcher-Subagent

SNMP auf Solaris-Systemen aktivieren

Führen Sie zum Installieren der Solaris-SNMP-Unterstützung die folgenden Schritte aus:

1. Beenden Sie den aktiven Solaris-SNMP-Dämon (snmpdx und snmpXdmid) mit dem Befehl kill.
2. Benennen Sie die folgenden Dateien um:

```
/etc/rc3.d/S76snmpdx in /etc/rc3.d/K76snmpdx
```

```
/etc/rc3.d/S77dmi in /etc/rc3.d/K77dmi
```
3. Laden Sie von der Adresse <http://www.sunfreeware.com/> die folgenden Pakete herunter:
 - libgcc-3.0.3-sol8-sparc-local (SMClibgcc)
 - openssl-0.9.6c-sol8-sparc-local (SMCssl)
 - popt-1.6.3-sol8-sparc-local (SMCpopt)
4. Installieren Sie die heruntergeladenen Pakete mit pkgadd.
5. Laden Sie die Datei ucd-snmp-4.2.3-solaris8.tar.gz herunter (Adresse: http://sourceforge.net/project/showfiles.php?group_id=12694).
6. Entpacken Sie ucd-snmp-4.2.3-solaris8.tar.gz im Stammverzeichnis (/).
7. Setzen Sie die folgenden Befehle ab:

```
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH: /usr/local/lib:/usr/local/ssl/lib:/usr/lib
```

```
export PATH=/usr/local/sbin:/usr/local/bin:$PATH
```

```
export SNMPCONFPATH =/etc/snmp
```

```
export MIBDIRS=/usr/local/share/snmp/mibs
```

```
cp /opt/ibm/edge/lb/servers/samples/SNMP/dpid2 /usr/local/sbin/dpid2
```
8. Erstellen Sie die Datei /etc/snmpd.peers, sofern sie noch nicht vorhanden ist. Fügen Sie in die Datei snmpd.peers Folgendes ein:

```
"dpid2" 1.3.6.1.4.1.2.3.1.2.2.1.1.2      "dpid_password"
```
9. Erstellen Sie die Datei /etc/snmp/snmpd.conf, sofern sie noch nicht vorhanden ist. Fügen Sie in die Datei snmpd.conf Folgendes ein:

```
smuxpeer      1.3.6.1.4.1.2.3.1.2.2.1.1.2      dpid_password
```
10. Starten Sie /usr/local/sbin/snmpd.
11. Starten Sie /usr/local/sbin/dpid2.

Anmerkungen:

1. Die folgenden Pakete liegen im Paketformat vor.
 - libgcc-3.0.3-sol8-sparc-local (SMClibgcc)

- openssl-0.9.6c-sol8-sparc-local (SMCssl)
- pop-1.6.3-sol8-sparc-local (SMCpop)

Auf der Website <http://sunfreeware.com/> haben die Namen die Erweiterung .gz und können deshalb nicht wie ZIP- oder TAR-Dateien entpackt werden. Verwenden Sie stattdessen pkgadd *Paketname*.

2. Achten Sie beim Hinzufügen des Eintrags smuxpeer zur Datei /etc/snmp/snmpd.conf darauf, dass Sie zur Zeichenfolge **dpid_password** kein Leerzeichen hinzufügen.
3. Das SNMP-Feature von Load Balancer wurde mit der smux-fähigen ucd-snmp-Version 4.2.3 getestet. Künftige Releases von ucd-snmp mit smux sollten mit einer ähnlichen Konfiguration funktionieren.

SNMP unter Windows aktivieren

Führen Sie zum Installieren der Windows-SNMP-Unterstützung die folgenden Schritte aus:

1. Klicken Sie auf "Start" > "Einstellungen" (Windows 2000) > "Systemsteuerung" > "Software".
2. Klicken Sie auf **Windows-Komponenten hinzufügen/entfernen**.
3. Klicken Sie im Assistenten für Windows-Komponenten auf **Verwaltungs- und Überwachungsprogramme**. Markieren Sie jedoch nicht das zugehörige Markierungsfeld bzw. löschen Sie die Markierung nicht, sollte sie vorhanden sein. Klicken Sie dann auf **Details**.
4. Markieren Sie den Eintrag **SNMP (Simple Network Management Protocol)** und klicken Sie auf "OK".
5. Klicken Sie auf "Weiter".

Namen einer Benutzergemeinschaft für SNMP angeben

Verwenden Sie bei aktivem Executor den Befehl **dscontrol subagent start** [*Name_der_Benutzergemeinschaft*], um den Namen der Benutzergemeinschaft zu definieren, der zwischen dem Windows-Erweiterungsagenten und dem SNMP-Agenten verwendet wird.

Wichtiger Hinweis: Unter Windows 2003 reagiert SNMP standardmäßig nicht auf angegebene Namen von Benutzergemeinschaften. Der SNMP-Subagent antwortet demzufolge nicht auf SNMP-Anforderungen. Um sicherzustellen, dass der SNMP-Subagent auf den Namen der Benutzergemeinschaft reagiert, müssen Sie die SNMP-Servicemerkmale mit dem entsprechenden Namen der Benutzergemeinschaft und den entsprechenden Zielhosts konfigurieren. Konfigurieren Sie die SNMP-Sicherheitsmerkmale wie folgt:

1. Öffnen Sie den Ordner "Computerverwaltung".
2. Klicken Sie in der Baumstruktur auf **Dienste**.
3. Klicken Sie im Detailfenster auf **SNMP-Dienst**.
4. Klicken Sie im Menü "Vorgang" auf **Eigenschaften**.
5. Klicken Sie auf der Registerseite "Sicherheit" unter "Angenommene Communitynamen" auf **Hinzufügen**.
6. Wählen Sie unter "Communityrechte" eine Berechtigungsstufe aus, mit der dieser Host SNMP-Anforderungen von der ausgewählten Community verarbeiten soll (das Minimum ist der **Lesezugriff**).
7. Geben Sie im Feld "Communityname" den Namen einer Benutzergemeinschaft ein. Der Name muss unter Beachtung der Groß-/Kleinschreibung mit dem für den Load-Balancer-Subagenten bereitgestellten Namen übereinstimmen (Standardcommunityname: public). Klicken Sie dann auf **Hinzufügen**.

8. Geben Sie an, ob SNMP-Pakete von einem Host akzeptiert werden sollen. Wählen Sie dazu eine der folgenden Optionen aus:
 - Wenn SNMP-Anforderungen von jedem Host im Netz unabhängig von der Identität akzeptiert werden sollen, klicken Sie auf **SNMP-Pakete von jedem Host annehmen**. (Wurde diese Option ausgewählt, muss eine Person oder Organisation ausgehend durch eine Authentifizierung mittels Kennwort oder Zertifikat überprüft werden.)
 - Wenn Sie die zu akzeptierenden SNMP-Pakete einschränken möchten, klicken Sie auf **SNMP-Pakete von diesen Hosts annehmen**. Klicken Sie dann auf **Hinzufügen**. Geben Sie für jeden Eintrag den Hostnamen und die IP-Adresse oder IPX-Adresse ein und klicken Sie nach jedem Eintrag auf **Hinzufügen**.
9. Starten Sie den SNMP-Dienst neu, damit die Änderung in Kraft tritt.

Nachrichten

Die Kommunikation von SNMP erfolgt über das Senden und Empfangen von *Nachrichten*, die von verwalteten Einheiten gesendet werden, um Ausnahmsbedingungen oder das Auftreten besonderer Ereignisse, wie beispielsweise das Erreichen eines Schwellenwerts, zu melden.

Der Subagent verwendet die folgenden Alarmnachrichten:

- indHighAvailStatus
- indSrvrGoneDown
- indDOSAttack
- indDOSAttackDone

Die Nachricht **indHighAvailStatus** gibt an, dass sich der Wert der Statusvariablen (hasState) für hohe Verfügbarkeit geändert hat. Die gültigen Werte von hasState sind:

-Ruhend

Diese Maschine führt einen Lastausgleich durch und versucht nicht, Kontakt mit der Dispatcher-Partnermaschine aufzunehmen.

-Empfangsbereit

Die Funktion für hohe Verfügbarkeit wurde gerade gestartet und der Dispatcher ist für die Partnermaschine empfangsbereit.

-Aktiv Diese Maschine führt einen Lastausgleich durch.

-Bereitschaft

Diese Maschine überwacht die aktive Maschine.

-Vorwegnehmen

Diese Maschine befindet sich während des Wechsels von primärer Maschine zu Partnermaschine in einem Übergangszustand.

-Auswählen

Der Dispatcher wählt aus, welche die primäre Maschine und welche die Partnermaschine ist.

-Executor nicht aktiv

Der Executor ist nicht aktiv.

Die Alarmnachricht **indSrvrGoneDown** gibt an, dass die Wertigkeit des vom Abschnitt csID (Cluster-ID), psNum (Port-Nummer) und ssID (Server-ID) der Objektkennung angegebenen Servers gleich null ist. Die letzte bekannte Anzahl

aktiver Verbindungen für den Server wird in der Nachricht gesendet. Diese Alarmnachricht gibt an, dass der angegebene Server inaktiviert ist, soweit Dispatcher dies feststellen konnte.

Die Alarmnachricht **indDOSAttack** gibt an, dass der Wert für numhalfopen (die Anzahl halboffener Verbindungen, die nur SYN-Pakete enthalten) an dem vom Abschnitt csID (Cluster-ID) und psNum (Port-Nummer) der Objektkennung angegebenen Port den Schwellenwert maxhalfopen überschritten hat. Die Anzahl der für den Port konfigurierten Server wird in der Alarmnachricht gesendet. Diese Alarmnachricht zeigt an, dass bei Load Balancer möglicherweise eine DoS-Attacke aufgetreten ist.

Die Alarmnachricht **indDOSAttackDone** gibt an, dass der Wert für numhalfopen (die Anzahl halboffener Verbindungen, die nur SYN-Pakete enthalten) an dem vom Abschnitt csID und psNum der Objektkennung angegebenen Port unter den Schwellenwert maxhalfopen gefallen ist. Die Anzahl der für den Port konfigurierten Server wird in der Alarmnachricht gesendet. Wenn Load Balancer nach dem Senden einer indDOSAttack-Alarmnachricht feststellt, dass die mögliche DoS-Attacke vorüber ist, wird diese Alarmnachricht gesendet.

Auf Linux- und UNIX-Systemen kann es sich aufgrund einer Einschränkung in der SMUX-API bei der Unternehmenskennung, die in Nachrichten von dem ibmNetDispatcher-Subagenten gemeldet wird, um die Unternehmenskennung von dpid2 und nicht um die Unternehmenskennung von ibmNetDispatcher, 1.3.6.1.4.1.2.6.144, handeln. Die SNMP-Verwaltungsdienstprogramme können jedoch die Quelle der Nachricht bestimmen, da die Daten eine Objektkennung aus der ibmNetDispatcher-MIB enthalten.

SNMP-Unterstützung mit dem Befehl dscontrol aktivieren und inaktivieren

Mit dem Befehl **dscontrol subagent start** wird die SNMP-Unterstützung aktiviert. Mit dem Befehl **dscontrol subagent stop** wird die SNMP-Unterstützung inaktiviert.

Weitere Informationen zum Befehl dscontrol finden Sie im Abschnitt „dscontrol subagent — SNMP-Subagenten konfigurieren“ auf Seite 402.

Gesamten Datenverkehr zur Sicherheit der Load-Balancer-Maschine mit ipchains oder iptables zurückweisen (Linux-Systeme)

In den Linux-Kernel ist das Firewall-Tool ipchains integriert. Wenn Load Balancer und ipchains gleichzeitig ausgeführt werden, sieht Load Balancer die Pakete zuerst. Erst danach werden sie von ipchains gesehen. Deshalb kann ipchains verwendet werden, um die Sicherheit einer Linux-Maschine mit Load Balancer zu erhöhen. Bei einer solchen Maschine könnte es sich beispielsweise um einen Rechner mit Load Balancer handeln, der einen Lastausgleich für Firewalls durchführt.

Wenn ipchains oder iptables für eine vollständige Einschränkung konfiguriert ist (so dass kein ein- oder ausgehender Datenverkehr zulässig ist), arbeitet die Paketweiterleitungsfunktion von Load Balancer normal weiter.

ipchains und iptables *können nicht* zum Filtern von eingehendem Datenverkehr verwendet werden, für den noch kein Lastausgleich durchgeführt wurde.

Ein gewisses Maß an Datenverkehr muss erlaubt sein, da Load Balancer sonst nicht fehlerfrei arbeiten kann. Nachfolgend sind einige Beispiele für eine solche Kommunikation aufgelistet:

- Die Advisor-Funktionen auf der Maschine mit Load Balancer und auf den Back-End-Servern kommunizieren miteinander.
- Load Balancer sendet ping-Aufrufe an Back-End-Server, Erreichbarkeitsziele und Load-Balancer-Partnermaschinen für hohe Verfügbarkeit.
- Die Benutzerschnittstellen (grafische Benutzerschnittstelle, Befehlszeile und Assistenten) verwenden RMI.
- Die Back-End-Server müssen auf die ping-Aufrufe der Load-Balancer-Maschine reagieren.

Eine angemessene ipchains-Strategie für die Load-Balancer-Maschinen wäre, den gesamten Datenverkehr mit Ausnahme des Verkehrs von oder zu den Back-End-Servern, den Partnermaschinen für hohe Verfügbarkeit, allen Erreichbarkeitszielen oder Konfigurationshosts zu unterbinden.

Sie sollten iptables nicht aktivieren, wenn Sie Load Balancer mit einem Linux-Kernel der Version 2.4.10.x ausführen. Eine Aktivierung unter diesem Linux-Kernel kann im Laufe der Zeit zur Beeinträchtigung des Durchsatzes führen.

Wenn Sie iptables inaktivieren möchten, listen Sie die Module auf (`lsmod`), um festzustellen, welche Module `ip_tables` und `ip_conntrack` verwenden. Entfernen Sie sie anschließend mit den Befehlen `rmmod ip_tables` und `rmmod ip_conntrack`. Nach einem Warmstart der Maschine werden diese Module wieder hinzugefügt. Sie müssen diesen Schritt deshalb nach jedem Warmstart wiederholen.

Weitere Informationen hierzu finden Sie im Abschnitt „Problem: Mögliche Störung der Paketweiterleitung durch iptables unter Linux“ auf Seite 327.

Komponente Content Based Routing verwenden

Die folgenden Abschnitte erläutern die Verwendung und Verwaltung der Komponente CBR von Load Balancer.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode `cbr` der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (`cbr`)“ auf Seite 55.

CBR starten und stoppen

- Geben Sie zum Starten von CBR in einer Befehlszeile **`cbrserver`** ein.
- Geben Sie zum Stoppen von CBR in einer Befehlszeile **`cbrserver stop`** ein.

CBR und Caching Proxy kooperieren über die API des Caching-Proxy-Plug-in bei der Bearbeitung von HTTP- und HTTPS-Anfragen (SSL). CBR kann erst mit dem Lastausgleich für die Server beginnen, wenn Caching Proxy auf derselben Maschine ausgeführt wird. Konfigurieren Sie CBR und Caching Proxy wie im Abschnitt „CBR-Konfigurationsbeispiel“ auf Seite 119 beschrieben.

CBR steuern

Nachdem Sie CBR gestartet haben, können Sie die Komponente mit einer der folgenden Methoden steuern:

- Konfigurieren Sie CBR mit dem Befehl **cbrcontrol**. Die vollständige Syntax dieses Befehls ist in Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349 beschrieben. Einige Verwendungsbeispiele sind an dieser Stelle aufgeführt.
- Konfigurieren Sie CBR auf der grafischen Benutzerschnittstelle (GUI). Geben Sie in der Befehlszeile **ladmin** ein, um die GUI zu öffnen. Weitere Informationen zum Konfigurieren von CBR auf der GUI finden Sie im Abschnitt „GUI“ auf Seite 112.

CBR-Protokolle verwenden

Die von CBR verwendeten Protokolle ähneln den Protokollen, die im Dispatcher verwendet werden. Weitere Informationen hierzu finden Sie im Abschnitt „Protokolle von Load Balancer verwenden“ auf Seite 267.

Anmerkung:

In früheren Releases konnten Sie den Protokollverzeichnispfad für CBR in der Caching-Proxy-Konfigurationsdatei ändern. Jetzt können Sie den Verzeichnispfad, in dem das Protokoll gespeichert wird, in der cbrserver-Datei ändern. Lesen Sie hierzu die Informationen im Abschnitt „Pfade für die Protokolldatei ändern“ auf Seite 269.

Site Selector verwenden

Site Selector starten und stoppen

- Geben Sie zum Starten von Site Selector in einer Befehlszeile **sssserver** ein.
- Geben Sie zum Stoppen von Site Selector in einer Befehlszeile **sssserver stop** ein.

Site Selector steuern

Nachdem Sie Site Selector gestartet haben, können Sie die Komponente mit einer der folgenden Methoden steuern:

- Konfigurieren Sie Site Selector mit dem Befehl **sscontrol**. Die vollständige Syntax dieses Befehls ist in Kapitel 28, „Befehlsreferenz für Site Selector“, auf Seite 405 beschrieben. Einige Verwendungsbeispiele sind an dieser Stelle aufgeführt.
- Konfigurieren Sie Site Selector auf der grafischen Benutzerschnittstelle (GUI). Geben Sie in der Befehlszeile **ladmin** ein, um die GUI zu öffnen. Weitere Informationen zum Konfigurieren von Site Selector auf der GUI finden Sie im Abschnitt „GUI“ auf Seite 133.

Protokolle von Site Selector verwenden

Die von Site Selector verwendeten Protokolle ähneln den Protokollen des Dispatchers. Weitere Informationen hierzu finden Sie im Abschnitt „Protokolle von Load Balancer verwenden“ auf Seite 267.

Cisco CSS Controller verwenden

Cisco CSS Controller starten und stoppen

1. Geben Sie zum Starten von Cisco CSS Controller in einer Befehlszeile **ccoserver** ein.

2. Geben Sie zum Stoppen von Cisco CSS Controller in einer Befehlszeile **ccoserver stop** ein.

Cisco CSS Controller steuern

Nachdem Sie Cisco CSS Controller gestartet haben, können Sie die Komponente mit einer der folgenden Methoden steuern:

- Konfigurieren Sie Cisco CSS Controller mit dem Befehl **ccocontrol**. Die vollständige Syntax dieses Befehls ist in Kapitel 29, „Befehlsreferenz für Cisco CSS Controller“, auf Seite 433 beschrieben. Einige Verwendungsbeispiele sind an dieser Stelle aufgeführt.
- Konfigurieren Sie Cisco CSS Controller auf der grafischen Benutzerschnittstelle (GUI). Geben Sie in der Befehlszeile **ladmin** ein, um die GUI zu öffnen. Weitere Informationen zum Konfigurieren von Cisco CSS Controller auf der GUI finden Sie im Abschnitt „GUI“ auf Seite 151.

Protokolle von Cisco CSS Controller verwenden

Die von Cisco CSS Controller verwendeten Protokolle ähneln den Protokollen des Dispatchers. Weitere Informationen hierzu finden Sie im Abschnitt „Protokolle von Load Balancer verwenden“ auf Seite 267.

Nortel Alteon Controller verwenden

Nortel Alteon Controller starten und stoppen

1. Geben Sie zum Starten von Nortel Alteon Controller in einer Befehlszeile **nalserver** ein.
2. Geben Sie zum Stoppen von Nortel Alteon Controller in einer Befehlszeile **nalserver stop** ein.

Nortel Alteon Controller steuern

Nachdem Sie Nortel Alteon Controller gestartet haben, können Sie die Komponente mit einer der folgenden Methoden steuern:

- Konfigurieren Sie Nortel Alteon Controller mit dem Befehl **nalcontrol**. Die vollständige Syntax dieses Befehls ist in Kapitel 30, „Befehlsreferenz für Nortel Alteon Controller“, auf Seite 451 beschrieben. Einige Verwendungsbeispiele sind an dieser Stelle aufgeführt.
- Konfigurieren Sie Nortel Alteon Controller auf der grafischen Benutzerschnittstelle (GUI). Geben Sie in der Befehlszeile **ladmin** ein, um die GUI zu öffnen. Weitere Informationen zum Konfigurieren von Nortel Alteon Controller auf der GUI finden Sie im Abschnitt „GUI“ auf Seite 173.

Protokolle von Nortel Alteon Controller verwenden

Die von Nortel Alteon Controller verwendeten Protokolle ähneln den Protokollen des Dispatchers. Weitere Informationen hierzu finden Sie im Abschnitt „Protokolle von Load Balancer verwenden“ auf Seite 267.

Metric Server verwenden

Metric Server starten und stoppen

Metric Server stellt Informationen zur Serverauslastung für Load Balancer bereit. Metric Server befindet sich auf jedem Server, der in den Lastausgleich einbezogen ist.

Linux- und UNIX-Systeme:

- Geben Sie auf jeder Maschine mit Metric Server in einer Befehlszeile **metricserver start** ein, um Metric Server zu starten.
- Geben Sie auf jeder Maschine mit Metric Server in einer Befehlszeile **metricserver stop** ein, um Metric Server zu stoppen.

Windows-Systeme:

Klicken Sie nacheinander auf **Start > Einstellungen** (für Windows 2000) > **Systemsteuerung > Verwaltung > Dienste**. Klicken Sie mit der rechten Maustaste auf "IBM Metric Server" und wählen Sie "Starten" aus. Zum Stoppen des Services müssen Sie dieselben Schritte ausführen und "Beenden" auswählen.

Protokolle von Metric Server verwenden

Ändern Sie die Protokollstufe im Start-Script für Metric Server. Sie können eine Protokollstufe von 0 bis 5 angeben. Die Stufen sind denen für die Load-Balancer-Protokolle vergleichbar. Daraufhin wird im Verzeichnis **...ms/logs** ein Agentenprotokoll erstellt.

Kapitel 25. Fehlerbehebung

Anhand der Informationen in diesem Kapitel können Fehler erkannt und behoben werden, die sich auf Load Balancer beziehen.

- Bevor Sie den IBM Kundendienst anrufen, lesen Sie den Abschnitt „Informationen zur Fehlerbehebung abrufen“.
- Durchsuchen Sie die „Fehlerbehebungstabellen“ auf Seite 287 nach dem aufgetretenen Symptom.

Informationen zur Fehlerbehebung abrufen

Stellen Sie wie in diesem Abschnitt beschrieben die vom IBM Kundendienst geforderten Daten zusammen. Die Informationen sind wie folgt thematisch geordnet:

- „Basisinformationen (immer erforderlich)“
- „Probleme mit der hohen Verfügbarkeit“ auf Seite 284
- „Advisor-Fehler“ auf Seite 285
- „Fehler beim Content Based Routing“ auf Seite 286
- „Cluster nicht erreichbar“ auf Seite 286
- „Alle Versuche sind gescheitert“ auf Seite 287
- „Upgrades“ auf Seite 287
- „Nützliche Links“ auf Seite 287

Basisinformationen (immer erforderlich)

Für die Komponente Dispatcher gibt es ein Fehlerbestimmungs-Tool, das automatisch betriebssystemspezifische Daten und komponentenspezifische Konfigurationsdateien erfasst. Geben Sie zum Ausführen dieses Tools im entsprechenden Verzeichnis **lbpd** ein:

Für Linux- und UNIX-Systeme: `/opt/ibm/edge/lb/servers/bin/`

Für Windows-Systeme: `C:\Programme\IBM\edge\lb\servers\bin`

Dieses Fehlerbestimmungs-Tool packt die Daten wie folgt in Dateien:

Für Linux- und UNIX-Systeme: `/opt/ibm/edge/lb/lbpmr.tar.Z`

Für Windows-Systeme: `C:\Programme\IBM\edge\lb\lbpmr.zip`

Anmerkung: Sie benötigen ein ZIP-Dienstprogramm für Windows-Systeme, das in einer Befehlszeile ausgeführt werden kann.

Bevor Sie den IBM Kundendienst anrufen, sollten Sie die folgenden Informationen zur Hand haben.

- Für Dispatcher die vom oben erwähnten Fehlerbestimmungs-Tool generierte `lbpmr`-Datei.
- In einer Umgebung mit hoher Verfügbarkeit die Konfigurationsdateien von beiden Load-Balancer-Maschinen. Verwenden Sie für alle Betriebssysteme das Script, mit dem Sie die Konfiguration laden, oder setzen Sie den folgenden Befehl ab:

```
dscontrol file save primary.cfg
```

Dieser Befehl stellt die Konfigurationsdatei in das Verzeichnis `.../ibm/edge/lb/servers/configuration/Komponentel`.

- Das Betriebssystem und die von Ihnen verwendete Version.
- Die Version von Load Balancer.
 - Wenn Load Balancer aktiv ist, setzen Sie die folgenden Befehle ab:
 - Für Dispatcher: `dscontrol executor report`
 - Für CBR: `cbrcontrol executor status`
 - Überprüfen Sie für Site Selector den Anfang der Datei `server.log` im Verzeichnis `.../ibm/edge/lb/servers/logs/ss/`.
 - Für Cisco CSS Controller und Nortel Alteon Controller: `xxxcontrol controller report`
 - Setzen Sie die folgenden Befehle ab, um sicherzustellen, dass Load Balancer installiert ist, und um die aktuelle Version von Load Balancer abzurufen:
 - Auf AIX-Systemen: `ls|pp -l|grep ibmlb`
 - Auf HP-UX-Systemen: `swlist | grep ibmlb`
 - Auf Linux-Systemen: `rpm -qa | grep ibmlb`
 - Auf Solaris-Systemen: `pkginfo | grep ibm`

Auf Windows-Systemen können Sie wie folgt überprüfen, ob Load Balancer installiert ist: Wählen Sie 'Start' > 'Einstellungen' > 'Systemsteuerung' > 'Software' aus.
- Setzen Sie den folgenden Befehl ab, um die aktuelle Java-Version abzurufen:


```
java -fullversion
```
- Wird Token-Ring oder Ethernet verwendet?
- Setzen Sie einen der folgenden Befehle ab, um Informationen zur Protokollstatistik und zur TCP/IP-Verbindung zu erhalten:


```
AIX-, HP-UX-, Linux- und Solaris-Systeme: netstat -ni
```

```
Windows-Systeme: ipconfig /all
```

Dies ist für alle Server und für Load Balancer erforderlich.
- Setzen Sie einen der folgenden Befehle ab, um Informationen zur Routentabelle zu erhalten:


```
AIX-, HP-UX-, Linux- und Solaris-Systeme: netstat -nr
```

```
Windows-Systeme: route print
```

Dies ist für alle Server und für Load Balancer erforderlich.

Probleme mit der hohen Verfügbarkeit

Stellen Sie bei Problemen in einer Umgebung mit hoher Verfügbarkeit die folgenden erforderlichen Informationen zusammen:

- Setzen Sie `hamon.log` auf die Protokollstufe 5: `dscontrol set loglevel 5`.
- Setzen Sie `reach.log` auf die Protokollstufe 5: `dscontrol manager reach set loglevel 5`.
- Stellen Sie die Scripts aus den folgenden Verzeichnissen zusammen:


```
AIX-, HP-UX-, Linux- und Solaris-Systeme: /opt/ibm/edge/lb/servers/bin
```

```
Windows-Systeme: C:\Programme\ibm\edge\lb\servers\bin
```

Die Script-Namen lauten wie folgt:

```
goActive
goStandby
goIdle (sofern vorhanden)
goInOp (sofern vorhanden)
```

Fügen Sie die Konfigurationsdateien hinzu. Lesen Sie hierzu die Informationen im Abschnitt „Basisinformationen (immer erforderlich)“ auf Seite 283.

Advisor-Fehler

Stellen Sie bei Advisor-Fehlern (z. B., wenn Advisor-Funktionen Server fälschlicherweise als inaktiv markieren) die folgenden erforderlichen Informationen zusammen.

- Setzen Sie das Advisor-Protokoll auf die Protokollstufe 5:

```
dscontrol advisor loglevel http 80 5
```

oder

```
dscontrol advisor loglevel Advisor-Name Port Protokollstufe
```

oder

```
dscontrol advisor loglevel Advisor-Name Cluster:Port Protokollstufe
```

oder

```
nalcontrol metriccollector set Consultant-ID:Service-ID:Messwertname loglevel Wert
```

Dieser Befehl erstellt ein Protokoll mit dem Namen `ADV_Advisor-Name.log`, z. B. `ADV_http.log`. Dieses Protokoll finden Sie in den folgenden Verzeichnissen:

AIX, HP-UX, Linux und Solaris: `/opt/ibm/edge/lb/servers/logs/Komponente`

Windows-Plattformen: `C:\Programme\ibm\edge\lb\servers\logs\Komponente`

Komponente steht hier für Folgendes:

dispatcher = Dispatcher

cbr = Content Based Routing

cco = Cisco CSS Controller

nal = Nortel Alteon Controller

ss = Site Selector

Anmerkung: Wenn Sie eigene Advisor-Funktionen schreiben, sollten Sie mit `ADV-LOG(Protokollstufe,Nachricht)` überprüfen, ob die Advisor-Funktion fehlerfrei ausgeführt werden kann.

Der Aufruf von `ADVLOG` gibt Anweisungen in der Protokolldatei der Advisor-Funktionen aus, wenn die Stufe niedriger als die den Advisor-Funktionen zugeordnete Protokollstufe ist. Bei der Protokollstufe 0 wird die Anweisung immer geschrieben. Sie können `ADVLOG` nicht vom Konstruktor aus verwenden. Die Protokolldatei wird erst unmittelbar nach Beendigung des Konstruktors für die benutzerdefinierte Advisor-Funktion erstellt, weil der Name der Protokolldatei von Angaben abhängt, die im Konstruktor definiert werden.

Sie können Ihre eigene Advisor-Funktion aber auch auf andere Weise testen und dabei die obige Einschränkung umgehen. Mit den Anweisungen `System.out.println(Nachricht)` können Sie Nachrichten in einem Fenster ausgeben. Editieren Sie das Script `dserver` und ändern Sie `javaw` in `java`, damit die `print`-Anweisungen im Fenster erscheinen. Die Nachrichten können nur angezeigt werden, wenn das zum Starten von `dserver` verwendete Fenster geöffnet bleibt. Auf Windows-Plattformen müssen Sie den als Dienst ausgeführten Dispatcher stoppen und ihn manuell in einem Fenster starten, um die Nachrichten anzuzeigen.

Weitere Informationen zu ADVLOG finden Sie im *Programming Guide for Edge Components*.

Fehler beim Content Based Routing

Stellen Sie bei Problemen mit dem Content Based Routing die folgenden erforderlichen Informationen zusammen:

- Setzen Sie diesen Befehl ab, um die Version zu erfragen: `cbrcontrol executor status`.
- Stellen Sie die folgenden Dateien zusammen:
 - `ibmproxy.conf` aus den folgenden Verzeichnissen:
 - Linux- und UNIX-Systeme: `/etc/`
 - Windows-Systeme: `C:\Programme\IBM\edge\cp\etc\en_US\`
 - CBR-Konfigurationsdatei aus den folgenden Verzeichnissen:
 - Linux- und UNIX-Systeme: `/opt/ibm/edge/lb/servers/configurations/cbr`
 - Windows-Systeme: `C:\Programme\IBM\edge\lb\servers\configurations\cbr`
 - Vergewissern Sie sich, dass in `ibmproxy.conf` die richtigen Einträge enthalten sind. Lesen Sie hierzu die Informationen unter „Schritt 1. Caching Proxy für die Verwendung von CBR konfigurieren“ auf Seite 114.

Cluster nicht erreichbar

Wenn Sie den Cluster nicht erreichen können, wurde möglicherweise auf einer der Load-Balancer-Maschinen oder auf beiden kein Aliasname für den Cluster definiert. Sie können wie folgt feststellen, welche Maschine Eigner des Clusters ist:

1. Dasselbe Teilnetz, aber *nicht* auf einer Load-Balancer-Maschine:

```
ping Cluster
arp -a
```

Falls Sie die Dispatcher-Weiterleitungsmethode `nat` oder `cbr` verwenden, rufen Sie mit `ping` auch die Rückkehradresse ab.

2. Sehen Sie sich die Ausgabe von `arp` an und vergleichen Sie die (16-stellige hexadezimale) MAC-Adresse mit den Ausgaben von `netstat -ni`, um festzustellen, welche Maschine der physische Eigner des Clusters ist.
3. Verwenden Sie die folgenden Befehle, um die Ausgaben beider Maschinen auszuwerten und festzustellen, ob auf beiden die Clusteradresse definiert ist.

```
AIX- und HP-UX-Systeme: netstat -ni
Linux- und Solaris-Systeme: ifconfig -a
Windows-Systeme: ipconfig /all
```

Wenn Sie von `ping` keine Antwort erhalten und nicht mit ULB arbeiten, wurde möglicherweise auf den Schnittstellen beider Maschinen (z. B. `en0`, `tr0` usw.) kein Aliasname für die Cluster-IP-Adresse definiert.

Anmerkung: Wenn Sie auf Linux-Systemen mit installiertem Load Balancer für IPV4 und IPV6 keine Antwort auf das `ping`-Signal erhalten, ist dies nur ein Hinweis darauf, dass der Back-End-Server nicht verfügbar ist. Der `arp`-Eintrag sollte dennoch aktualisiert werden. Falls `arping` verfügbar ist, kann es alternativ verwendet werden.

Alle Versuche sind gescheitert

Wenn alle Versuche, Routing-Fehler zu beheben, gescheitert sind, setzen Sie den folgenden Befehl ab, um einen Trace für den Datenverkehr im Netz durchzuführen:

- Auf einem AIX-System mit Load Balancer:

```
iptrace -a  
-s fragliche_Client-IP-Adresse -d Cluster-IP-Adresse -b iptrace.trc
```

Führen Sie den Trace durch, reproduzieren Sie den Fehler und beenden Sie dann den Prozess mit kill.

- HP-UX-Systeme:

```
tcpdump -i lan0 host Cluster und host Client
```

Unter Umständen müssen Sie tcpdump von einer der GNU-Softwarearchivsites zu HP-UX herunterladen.

- Auf Linux-Systemen:

```
tcpdump -i eth0 host Cluster und host Client
```

Führen Sie den Trace durch, reproduzieren Sie den Fehler und beenden Sie dann den Prozess mit kill.

- Unter Solaris:

```
snoop -v Client-IP-Adresse Ziel-IP-Adresse > snooptrace.out
```

- Auf Windows-Systemen benötigen Sie einen Sniffer. Verwenden Sie dieselben Eingaben wie für einen Filter.

Sie können auch die Protokollstufe verschiedener Protokolle (z. B. des Manager- oder des Advisor-Protokolls) erhöhen und die Protokollausgaben auswerten.

Upgrades

Überprüfen Sie, ob Upgrades verfügbar sind, da ein vorliegender Fehler möglicherweise in einem neuen Servicerelease oder Patch-Code bereits behoben ist. Eine Liste der behobenen Fehler für Edge Components finden Sie auf der Support-Seite der Website zu WebSphere Application Server unter der folgenden Adresse: <http://www.ibm.com/software/webservers/appserv/was/support/>. Klicken Sie auf der Support-Seite auf den Link zur Downloadsite für Fehlerberichtigungen.

Java-Code

Die richtige Java-Version wird im Rahmen der Installation von Load Balancer installiert.

Nützliche Links

Im Abschnitt „Referenzliteratur“ auf Seite xvii finden Sie Links zu Webseiten für Unterstützung und Bibliotheken. Die Webseite für Unterstützung enthält einen Link zu technischen Hinweisen, die Ihnen bei Fragen weiterhelfen können.

Fehlerbehebungstabellen

In den genannten Tabellen finden Sie Informationen zu folgenden Themen:

- Fehlerbehebung für Dispatcher — Tabelle 14 auf Seite 288
- Fehlerbehebung für CBR — Tabelle 15 auf Seite 294
- Fehlerbehebung für Site Selector — Tabelle 16 auf Seite 295
- Fehlerbehebung für Cisco CSS Controller — Tabelle 17 auf Seite 297

- Fehlerbehebung für Nortel Alteon Controller — Tabelle 18 auf Seite 298
- Fehlerbehebung für Metric Server — Tabelle 19 auf Seite 299

Tabelle 14. Tabelle zur Fehlerbehebung für Dispatcher

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Dispatcher wird nicht korrekt ausgeführt.	In Konflikt stehende Port-Nummern	„Port-Nummern für Dispatcher überprüfen“ auf Seite 301
Ein auf der Dispatcher-Maschine konfigurierter Server antwortet nicht auf Lastausgleichsanforderungen.	Falsche oder einen Konflikt verursachende Adresse	„Problem: Dispatcher und Server antworten nicht“ auf Seite 305
Kein Service für Verbindungen von Client-Maschinen oder Zeitlimitüberschreitung bei Verbindungen	<ul style="list-style-type: none"> • Falsche Konfiguration für Weiterleitung • Auf der NIC wurde kein Aliasname für die Clusteradresse angegeben. • Auf der Loopback-Einheit des Servers wurde kein Aliasname für die Clusteradresse festgelegt. • Zusätzliche Route nicht gelöscht • Port nicht für jeden Cluster definiert 	„Problem: Dispatcher-Anforderungen werden nicht verteilt“ auf Seite 305
Client-Maschinen erhalten keinen Service oder überschreiten das Zeitlimit.	Funktion für hohe Verfügbarkeit arbeitet nicht	„Problem: Die Dispatcher-Funktion für hohe Verfügbarkeit kann nicht ausgeführt werden“ auf Seite 305
Es kann kein Überwachungssignal hinzugefügt werden (Windows-Plattform).	Die Quellenadresse ist auf keinem Adapter konfiguriert.	„Problem: Es kann kein Überwachungssignal hinzugefügt werden (Windows-Plattform)“ auf Seite 306
Server verarbeitet keine Anforderungen (Windows-Plattform).	Es wurde eine zusätzliche Route in der Routentabelle erstellt.	„Problem: Zusätzliche Routen (Windows 2000)“ auf Seite 306
Advisor-Funktionen arbeiten nicht korrekt mit der Weitverkehrsunterstützung.	Advisor-Funktionen werden auf fernen Maschinen nicht ausgeführt.	„Problem: Advisor-Funktionen arbeiten nicht korrekt“ auf Seite 306
Dispatcher, Microsoft IIS und SSL arbeiten nicht oder setzen die Arbeit nicht fort.	Protokollübergreifend können keine verschlüsselten Daten gesendet werden.	„Problem: Dispatcher, Microsoft IIS und SSL funktionieren nicht (Windows-Plattform)“ auf Seite 306
Verbindung zur fernen Maschine zurückgewiesen	Es wird noch eine ältere Version der Schlüssel verwendet.	„Problem: Dispatcher-Verbindung zu einer fernen Maschine“ auf Seite 307

Tabelle 14. Tabelle zur Fehlerbehebung für Dispatcher (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Der Befehl dscontrol oder lbadmin scheitert mit der Nachricht 'Server antwortet nicht' oder 'Zugriff auf RMI-Server nicht möglich'.	<ol style="list-style-type: none"> 1. Befehle können nicht ausgeführt werden, weil der Stack SOCKSifiziert ist oder dsserver nicht gestartet wurde. 2. Die RMI-Ports sind nicht richtig definiert. 3. Für die Hostdatei ist der falsche lokale Host angegeben. 	„Problem: Der Befehl dscontrol oder lbadmin scheitert“ auf Seite 307
Wenn Netscape als Standardbrowser zum Anzeigen der Onlinehilfe verwendet wird, erscheint die Fehlernachricht "Datei ... nicht gefunden" (Windows-Plattform).	Falsche Einstellung für die HTML-Dateizuordnung	„Problem: Fehlernachricht 'Datei nicht gefunden...' beim Anzeigen der Onlinehilfe (Windows-Plattform)" auf Seite 308
Die grafische Benutzerschnittstelle wird nicht richtig gestartet.	Unzureichender Paging-Bereich	„Problem: Die grafische Benutzerschnittstelle (GUI) wird nicht richtig gestartet" auf Seite 308
Fehler bei der Ausführung von Dispatcher mit installiertem Caching Proxy	Caching-Proxy-Datei-abhängigkeit	„Problem: Fehler bei der Ausführung von Dispatcher mit installiertem Caching Proxy" auf Seite 308
Die grafische Benutzerschnittstelle wird nicht richtig angezeigt.	Die Auflösung ist nicht korrekt.	„Problem: Die grafische Benutzerschnittstelle (GUI) wird nicht richtig angezeigt" auf Seite 308
Die Hilfetextanzeigen werden manchmal von anderen Fenstern verdeckt.	Java-Einschränkung	„Problem: Auf der Windows-Plattform sind die Hilfenfenster manchmal von anderen offenen Fenstern verdeckt" auf Seite 308
Load Balancer kann Rahmen nicht verarbeiten und weiterleiten.	Für jede NIC ist eine eindeutige MAC-Adresse erforderlich.	„Problem: Load Balancer kann Rahmen nicht verarbeiten und weiterleiten" auf Seite 309
Die Anzeige ist blau.	Es ist keine Netzwerkkarte installiert/konfiguriert.	„Problem: Beim Starten des Executors von Load Balancer erscheint eine blaue Anzeige" auf Seite 309
Automatische Pfaderkennung verhindert Datenrückfluss.	Auf der Loopback-Einheit wird ein Aliasname für den Cluster verwendet.	„Problem: Automatische Pfaderkennung verhindert Datenrückfluss mit Load Balancer" auf Seite 309
Keine hohe Verfügbarkeit im Weitverkehrsmodus von Load Balancer	Der ferne Dispatcher muss auf dem lokalen Dispatcher als Server eines Clusters definiert werden.	„Problem: Keine hohe Verfügbarkeit im Weitverkehrsmodus von Load Balancer" auf Seite 310

Tabelle 14. Tabelle zur Fehlerbehebung für Dispatcher (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Die GUI blockiert oder verhält sich nicht erwartungsgemäß, wenn versucht wird, eine große Konfigurationsdatei zu laden.	Java kann nicht auf so viel Speicher zugreifen, wie für die Bearbeitung einer so großen Änderung der GUI erforderlich ist.	„Problem: Beim Laden einer großen Konfigurationsdatei blockiert die GUI oder verhält sich nicht erwartungsgemäß“ auf Seite 311
IP-Adressen werden über die Fernverbindung nicht richtig aufgelöst.	Wenn Sie einen fernen Client mit einer gesicherten Socks-Implementierung verwenden, ist nicht sichergestellt, dass vollständig qualifizierte Domänen- oder Hostnamen in die richtige IP-Adresse aufgelöst werden.	„Problem: IP-Adressen werden über die Fernverbindung nicht richtig aufgelöst“ auf Seite 312
Auf der koreanischen Schnittstelle von Load Balancer werden unter AIX und Linux überlappende oder unpassende Schriftarten angezeigt.	Ändern Sie die Standardschriftarten.	„Problem: Auf der koreanischen Schnittstelle von Load Balancer werden bei AIX- und Linux-Systemen überlappende oder unpassende Schriftarten angezeigt“ auf Seite 312
Wenn Sie auf Windows-Systemen einen Aliasnamen auf dem MS Loopback-Adapter definiert haben und bestimmte Befehle absetzen (z. B. hostname), reagiert das Betriebssystem falsch und gibt die Aliasadresse zurück.	In der Liste der Netzwerkverbindungen darf der neu hinzugefügte Aliasname nicht oberhalb der lokalen Adresse aufgeführt sein.	„Problem: Auf Windows-Systemen wird beim Absetzen von Befehlen wie hostname an Stelle der lokalen Adresse die Aliasadresse zurückgegeben“ auf Seite 312
Bei Verwendung einer Matrox-AGP-Videokarte auf einer Windows-Plattform kommt es zu unerwartetem GUI-Verhalten.	Der Fehler tritt auf, wenn Matrox-AGP-Videokarten während der Ausführung der Load-Balancer-GUI verwendet werden.	„Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten“ auf Seite 313
Bei Ausführung von <code>rmmod ibmlb</code> auf Linux-Systemen kommt es zu einem unerwarteten Verhalten (z. B. zu einer Blockierung des Systems).	Der Fehler tritt auf, wenn das Kernel-Modul von Load Balancer (<code>ibmlb</code>) manuell entfernt wird.	„Problem: Unerwartetes Verhalten bei Ausführung von <code>'rmmod ibmlb'</code> (Linux-Systeme)“ auf Seite 313
Bei Ausführung von Befehlen auf der Dispatcher-Maschine sind die Antwortzeiten sehr lang.	Lange Antwortzeiten können auf eine Überlastung der Maschine durch ein hohes Client-Datenverkehrsaufkommen zurückzuführen sein.	„Problem: Lange Antwortzeiten beim Ausführen von Befehlen auf der Dispatcher-Maschine“ auf Seite 313
Bei Verwendung der Dispatcher-Weiterleitungsmethode <code>mac</code> registriert die SSL- oder HTTPS-Advisor-Funktion keine Serverbelastungen.	Dieser Fehler tritt auf, wenn die SSL-Serveranwendung nicht mit der Cluster-IP-Adresse konfiguriert ist.	„Problem: Bei Verwendung der Weiterleitungsmethode <code>mac</code> registriert die Advisor-Funktion SSL oder HTTPS keine Serverlast“ auf Seite 314

Tabelle 14. Tabelle zur Fehlerbehebung für Dispatcher (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Bei der fernen Webverwaltung mit Netscape wird die Verbindung zum Host getrennt.	Die Verbindung zum Host wird getrennt, wenn Sie die Größe des Browserfensters ändern.	„Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung“ auf Seite 314
Bei aktiviertem Socket-Pooling wird der Webserver an 0.0.0.0 gebunden.	Konfigurieren Sie den Microsoft IIS bindingspezifisch.	„Problem: Bei aktiviertem Socket-Pooling wird der Webserver an 0.0.0.0 gebunden“ auf Seite 314
Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1).	Ändern Sie die Schriftartmerkmale des Fensters mit der Eingabeaufforderung.	„Problem: Auf Windows-Systemen erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)“ auf Seite 315
Auf der HP-UX-Plattform wird die folgende Nachricht angezeigt: java.lang.OutOfMemoryError unable to create new native thread	Einige HP-UX-Installationen lassen standardmäßig 64 Threads pro Prozess zu. Dies ist unzureichend.	„Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads“ auf Seite 315
Unter Windows markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv.	Das Feature Task Offload ist nicht inaktiviert oder Sie müssen unter Umständen ICMP aktivieren.	„Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv“ auf Seite 316
Auf der Windows-Plattform tritt ein Problem bei der Auflösung von IP-Adressen in Hostnamen auf, wenn für einen Adapter mehrere Adressen konfiguriert sind.	Die IP-Adresse, die als Hostname verwendet werden soll, muss in der Registrierungsdatenbank als erste Adresse angegeben sein.	„Problem: Auflösung von IP-Adressen in Hostnamen auf Windows-Systemen, wenn für einen Adapter mehrere Adressen konfiguriert sind“ auf Seite 316
Auf der Windows-Plattform können die Advisor-Funktionen in einer Konfiguration für hohe Verfügbarkeit nach einem Netzwerkausfall nicht ausgeführt werden.	Wenn das System einen Netzwerkausfall erkennt, löscht es seinen ARP-Cache (Address Resolution Protocol).	„Problem: Advisor-Funktionen können auf Windows-Systemen in einer Konfiguration für hohe Verfügbarkeit nach einem Netzwerkausfall nicht ausgeführt werden“ auf Seite 317
Auf Linux-Systemen ist der Befehl "IP address add" nicht mit mehreren Clusteraliasnamen auf der Loopback-Einheit kompatibel.	Wenn Sie auf der Loopback-Einheit Aliasnamen für mehrere Adressen festlegen möchten, sollten Sie den Befehl ifconfig und nicht ip address add verwenden.	„Problem: Beim Festlegen von Aliasnamen für mehrere Cluster auf der Loopback-Einheit von Linux-Systemen nicht den Befehl 'IP address add' verwenden“ auf Seite 318
Bei dem Versuch, einen Server hinzuzufügen, erscheint die Fehlermeldung: "Router-Adresse nicht angegeben oder nicht gültig für die Port-Methode".	Stellen Sie anhand der Informationen in der Prüfliste fest, welches Problem bei Hinzufügen eines Servers aufgetreten ist.	„Problem: Fehlermeldung 'Router-Adresse nicht angegeben oder nicht gültig für Port-Methode' erscheint“ auf Seite 318

Tabelle 14. Tabelle zur Fehlerbehebung für Dispatcher (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Auf Solaris-Systemen werden die Load-Balancer-Prozesse beendet, wenn Sie das Fenster mit der Terminalsitzung verlassen, in dem die Prozesse gestartet wurden.	Verwenden Sie den Befehl nohup , um zu verhindern, dass die gestarteten Prozesse beim Verlassen der Terminalsitzung ein Stoppsignal empfangen.	„Problem: Auf Solaris-Systemen werden Load-Balancer-Prozesse beim Verlassen des Terminalfensters, in dem die Prozesse gestartet wurden, beendet“ auf Seite 319
Beim Laden von Load-Balancer-Konfigurationen geht der Durchsatz zurück.	Die Verzögerung kann auf DNS-Aufrufe (Domain Name System) zurückzuführen sein, die abgesetzt werden, um die Serveradresse aufzulösen und zu prüfen.	„Problem: Verzögerung beim Laden einer Load-Balancer-Konfiguration“ auf Seite 319
Auf Windows-Systemen erscheint die Fehlermeldung, dass ein IP-Adressenkonflikt mit einem anderen System im Netz vorliegt.	Wenn die hohe Verfügbarkeit konfiguriert ist, können für einen kurzen Zeitraum Clusteradressen auf beiden Maschinen konfiguriert sein, was zu dieser Fehlermeldung führt.	„Problem: Fehlermeldung zu einem IP-Adressenkonflikt auf Windows-Systemen“ auf Seite 319
In einer Konfiguration mit hoher Verfügbarkeit sind sowohl primäre Maschinen als auch Ausweichmaschinen aktiv.	Dieses Problem kann auftreten, wenn die go-Scripts weder auf der primären Maschine noch auf der Ausweichmaschine ausgeführt werden.	„Problem: In einer Konfiguration mit hoher Verfügbarkeit sind sowohl primäre Maschinen als auch Ausweichmaschinen aktiv“ auf Seite 320
Client-Anforderungen scheitern, wenn der Dispatcher versucht, Antworten in Form großer Seiten zurückzugeben.	Client-Anforderungen, die zu Antworten in Form großer Seiten führen, überschreiten das Zeitlimit, wenn die maximale Übertragungseinheit (MTU, Maximum Transmit Unit) auf der Dispatcher-Maschine bei Verwendung der Weiterleitungsmethode nat oder cbr nicht ordnungsgemäß gesetzt ist.	„Problem: Client-Anforderungen scheitern, wenn die Rückgabe von Antworten in Form großer Seiten versucht wird“ auf Seite 320
Auf Windows-Systemen wird beim Absetzen eines Befehls dscontrol oder lbadm der Fehler "Der Server antwortet nicht" angezeigt.	Dieser Fehler tritt auf, wenn es auf einem Windows-System mehr als eine IP-Adresse gibt und die Hostdatei nicht angibt, welche Adresse dem Hostnamen zugeordnet werden soll.	„Auf Windows-Systemen wird beim Absetzen eines Befehls dscontrol oder lbadm der Fehler 'Der Server antwortet nicht' angezeigt“ auf Seite 321
Dispatcher-Maschinen mit hoher Verfügbarkeit können unter Linux auf S/390 mit qeth-Treibern möglicherweise nicht synchronisiert werden.	Wenn Sie unter Linux für S/390 mit dem qeth-Netzreiter die hohe Verfügbarkeit verwenden, können der aktive Dispatcher und der Bereitschafts-Dispatcher unter Umständen nicht synchronisiert werden.	„Problem: Dispatcher-Maschinen mit hoher Verfügbarkeit können unter Linux auf S/390 mit qeth-Treibern möglicherweise nicht synchronisiert werden“ auf Seite 321

Tabelle 14. Tabelle zur Fehlerbehebung für Dispatcher (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Tipps für das Konfigurieren der hohen Verfügbarkeit für Load Balancer	Die Tipps werden Ihnen helfen, die Auswirkungen von Problemen mit der hohen Verfügbarkeit zu verringern. Es können unter anderem folgende Probleme auftreten: <ul style="list-style-type: none"> • Übergangene Verbindungen nach der Übernahme • keine Synchronisation mit Partnermaschinen • versehentliche Übertragung von Anforderungen an die Ausweichpartnermaschine 	„Problem: Konfigurationstipps für hohe Verfügbarkeit“ auf Seite 321
Einschränkungen für die Konfiguration der Dispatcher-Weiterleitungsmethode mac auf zSeries- und S/390-Plattformen	Unter Linux gibt es Einschränkungen, wenn Sie zSeries- oder S/390-Server mit OSA-Karten (Open System Adapter) verwenden. Strategien zur Lösung von Problemen stehen zur Verfügung.	„Problem: Einschränkungen für die Dispatcher-Konfiguration unter Linux auf zSeries- oder S/390-Servern mit OSA-Karten“ auf Seite 323
Speicherverlust unter einigen Red-Hat-Linux-Versionen bei Ausführung von Load Balancer mit Manager und Advisor-Funktionen	Die JVM-Versionen des IBM Java SDK und die im Lieferumfang einiger Linux-Distributionen (z. B. Red Hat Enterprise Linux 3.0) enthaltene NPTL (Native POSIX Thread Library) können einen Speicherverlust verursachen.	„Problem: Speicherverlust unter einigen Linux-Versionen bei Ausführung von Dispatcher mit Manager und Advisor-Funktionen“ auf Seite 325
Hinweis auf weitergeleitete Pakete (steigende Paketanzahl) im Dispatcher-Bericht unter SUSE Linux Enterprise Server 9, obwohl die Pakete den Back-End-Server nie wirklich erreichen	Das NAT-Modul von iptables wird geladen. In dieser Version von iptables gibt es möglicherweise einen (noch nicht nachgewiesenen) Fehler, der bei der Interaktion mit Dispatcher ein seltsames Verhalten auslöst.	„Problem: Unter SUSE Linux Enterprise Server 9 leitet Dispatcher Pakete weiter, die den Back-End-Server jedoch nicht erreichen“ auf Seite 325
Mögliche Probleme bei der Funktionsübernahme auf Windows-Systemen, wenn das Dispatcher-Feature für hohe Verfügbarkeit verwendet wird	Wenn das goScript, das die Cluster-IP-Adresse auf der aktiven Maschine konfiguriert, vor dem goScript ausgeführt wird, das die Konfiguration der Cluster-IP-Adresse auf der Ausweichmaschine aufhebt, kann es zu Problemen kommen.	„Problem: Fehlernachricht zu einem IP-Adressenkonflikt auf Windows-Systemen während der Funktionsübernahme bei hoher Verfügbarkeit“ auf Seite 326
Mögliche Störung der Paketweiterleitung durch iptables auf Linux-Systemen	Unter Linux kann iptables den Lastausgleich für den Datenverkehr stören und muss auf der Load-Balancer-Maschine inaktiviert werden.	„Problem: Mögliche Störung der Paketweiterleitung durch iptables unter Linux“ auf Seite 327

Tabelle 14. Tabelle zur Fehlerbehebung für Dispatcher (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Nachricht "Server kann nicht hinzugefügt werden" auf Solaris-Systemen bei dem Versuch, einen IPV6-Server auf der Dispatcher-Maschine zu konfigurieren	Dieser Fehler kann durch die Art und Weise ausgelöst werden, mit der das Betriebssystem Solaris die ping-Anforderung für eine IPV6-Adresse handhabt.	„Problem: Auf Solaris-Systemen kann kein IPV6-Server zur Load-Balancer-Konfiguration hinzugefügt werden“ auf Seite 327
Warnung zu Java-Dateigruppen bei der Installation von Servicekorrekturen oder bei einer Installation mit den nativen System-Tools für Paketinstallation	Die Produktinstallation umfasst mehrere Pakete, die nicht auf derselben Maschine installiert werden müssen. Jedes dieser Pakete installiert daher eine Java-Dateigruppe. Werden die Pakete auf derselben Maschine installiert, erscheint eine Warnung, dass eine andere Dateigruppe auch Eigner der Java-Dateigruppe ist.	„Java-Warnung bei Installation von Servicekorrekturen“ auf Seite 328
Upgrade der Java-Dateigruppe von Load-Balancer-Installationen	Falls ein Problem mit der Java-Dateigruppe auftritt, melden Sie es dem IBM Service. Sie erhalten dann ein Upgrade für die Java-Dateigruppe, die für die Load-Balancer-Installation bereitgestellt wurde.	„Upgrade der Java-Dateigruppe einer Load-Balancer-Installation“ auf Seite 328

Tabelle 15. Tabelle zur Fehlerbehebung für CBR

Fehler	Mögliche Ursache	Siehe Abschnitt ...
CBR wird nicht korrekt ausgeführt.	In Konflikt stehende Port-Nummern	„Port-Nummern für CBR überprüfen“ auf Seite 302
Der Befehl cbrcontrol oder lbadm in scheitert mit der Nachricht 'Server antwortet nicht' oder 'Zugriff auf RMI-Server nicht möglich'.	Befehle können nicht ausgeführt werden, weil der Stack SOCKSifiziert ist oder cbrserver nicht gestartet wurde.	„Problem: Der Befehl cbrcontrol oder lbadm in scheitert“ auf Seite 328
Die Last von Anforderungen wird nicht verteilt.	Caching Proxy wurde vor dem Executor gestartet.	„Problem: Anforderungen werden nicht verteilt“ auf Seite 329
Unter Solaris scheitert der Befehl cbrcontrol executor start mit der Nachricht 'Fehler: Executor wurde nicht gestartet'.	Unter Umständen müssen die IPC-Standardwerte geändert werden, um den Befehl ordnungsgemäß ausführen zu können. Es ist auch möglich, dass die Verknüpfung mit der Bibliothek nicht korrekt angegeben ist.	„Problem: Auf Solaris-Systemen scheitert der Befehl cbrcontrol executor start“ auf Seite 329
URL-Regel arbeitet nicht.	Syntax- oder Konfigurationsfehler	„Problem: Syntax- oder Konfigurationsfehler“ auf Seite 329

Tabelle 15. Tabelle zur Fehlerbehebung für CBR (Forts.)

Bei Verwendung einer Matrox-AGP-Videokarte in Windows-Systemen kommt es zu unerwartetem GUI-Verhalten.	Der Fehler tritt auf, wenn Matrox-AGP-Videokarten während der Ausführung der Load-Balancer-GUI verwendet werden.	„Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten“ auf Seite 330
Die GUI blockiert oder verhält sich nicht erwartungsgemäß, wenn versucht wird, eine große Konfigurationsdatei zu laden.	Java kann nicht auf so viel Speicher zugreifen, wie für die Bearbeitung einer so großen Änderung der GUI erforderlich ist.	„Problem: Beim Laden einer großen Konfigurationsdatei blockiert die GUI oder verhält sich nicht erwartungsgemäß“ auf Seite 311
Bei der fernen Webverwaltung mit Netscape wird die Verbindung zum Host getrennt.	Die Verbindung zum Host wird getrennt, wenn Sie die Größe des Browserfensters ändern.	„Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung“ auf Seite 330
Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1).	Ändern Sie die Schriftartmerkmale des Fensters mit der Eingabeaufforderung.	„Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)“ auf Seite 330
Auf der HP-UX-Plattform wird die folgende Nachricht angezeigt: java.lang.OutOfMemoryError unable to create new native thread	Einige HP-UX-Installationen lassen standardmäßig 64 Threads pro Prozess zu. Dies ist unzureichend.	„Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads“ auf Seite 330
Unter Windows markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv.	Das Feature Task Offload ist nicht inaktiviert oder Sie müssen unter Umständen ICMP aktivieren.	„Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv“ auf Seite 330
Auf der Windows-Plattform tritt ein Problem bei der Auflösung von IP-Adressen in Hostnamen auf, wenn für einen Adapter mehrere Adressen konfiguriert sind.	Die IP-Adresse, die als Hostname verwendet werden soll, muss in der Registrierungsdatenbank als erste Adresse angegeben sein.	„Problem: Auflösung von IP-Adressen in Hostnamen auf Windows-Systemen, wenn für einen Adapter mehrere Adressen konfiguriert sind“ auf Seite 331
Auf Solaris-Systemen werden die Load-Balancer-Prozesse beendet, wenn Sie das Fenster mit der Terminalsitzung verlassen, in dem die Prozesse gestartet wurden.	Verwenden Sie den Befehl nohup , um zu verhindern, dass die gestarteten Prozesse beim Verlassen der Terminalsitzung ein Stoppsignal empfangen.	„Problem: Auf Solaris-Systemen werden Load-Balancer-Prozesse beim Verlassen des Terminalfensters, in dem die Prozesse gestartet wurden, beendet“ auf Seite 319

Tabelle 16. Tabelle zur Fehlerbehebung für Site Selector

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Site Selector wird nicht korrekt ausgeführt.	Konflikt verursachende Port-Nummer	„Port-Nummern für Site Selector überprüfen“ auf Seite 302

Tabelle 16. Tabelle zur Fehlerbehebung für Site Selector (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Site Selector gewichtet vom Solaris-Client eingehende Anforderungen nicht nach der RoundRobin-Methode.	Solaris-Systeme führen einen Namensservice-Cache-Dämon aus.	„Problem: Site Selector verteilt den Datenverkehr von Solaris-Clients nicht nach der RoundRobin-Methode“ auf Seite 331
Der Befehl sscontrol oder lbadm in scheitert mit der Nachricht ‘Server antwortet nicht’ oder ‘Zugriff auf RMI-Server nicht möglich’.	Befehle können nicht ausgeführt werden, weil der Stack SOCKSifiziert ist oder sserver nicht gestartet wurde.	„Problem: Der Befehl sscontrol oder lbadm in scheitert“ auf Seite 331
„sserver“ kann auf der Windows-Plattform nicht gestartet werden.	Auf Windows-Systemen muss der Hostname nicht im DNS enthalten sein.	„Problem: sserver wird auf der Windows-Plattform nicht gestartet“ auf Seite 332
Für eine Maschine mit duplizierten Routen wird der Lastausgleich nicht richtig durchgeführt. Die Namensauflösung scheint nicht zu funktionieren.	Eine Site-Selector-Maschine enthält mehrere Adapter, die mit demselben Teilnetz verbunden sind.	„Problem: Site Selector führt bei duplizierten Routen den Lastausgleich nicht korrekt durch“ auf Seite 332
Bei Verwendung einer Matrox-AGP-Videokarte auf einer Windows-Plattform kommt es zu unerwartetem GUI-Verhalten.	Der Fehler tritt auf, wenn Matrox-AGP-Videokarten während der Ausführung der Load-Balancer-GUI verwendet werden.	„Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten“ auf Seite 333
Die GUI blockiert oder verhält sich nicht erwartungsgemäß, wenn versucht wird, eine große Konfigurationsdatei zu laden.	Java kann nicht auf so viel Speicher zugreifen, wie für die Bearbeitung einer so großen Änderung der GUI erforderlich ist.	„Problem: Beim Laden einer großen Konfigurationsdatei blockiert die GUI oder verhält sich nicht erwartungsgemäß“ auf Seite 311
Bei der fernen Webverwaltung mit Netscape wird die Verbindung zum Host getrennt.	Die Verbindung zum Host wird getrennt, wenn Sie die Größe des Browserfensters ändern.	„Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung“ auf Seite 333
Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1).	Ändern Sie die Schriftartmerkmale des Fensters mit der Eingabeaufforderung.	„Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)“ auf Seite 333
Auf der HP-UX-Plattform wird die folgende Nachricht angezeigt: java.lang.OutOfMemoryError unable to create new native thread	Einige HP-UX-Installationen lassen standardmäßig 64 Threads pro Prozess zu. Dies ist unzureichend.	„Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads“ auf Seite 333

Tabelle 16. Tabelle zur Fehlerbehebung für Site Selector (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Unter Windows markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv.	Das Feature Task Offload ist nicht inaktiviert oder Sie müssen unter Umständen ICMP aktivieren.	„Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv“ auf Seite 333
Auf Solaris-Systemen werden die Load-Balancer-Prozesse beendet, wenn Sie das Fenster mit der Terminalsitzung verlassen, in dem die Prozesse gestartet wurden.	Verwenden Sie den Befehl nohup , um zu verhindern, dass die gestarteten Prozesse beim Verlassen der Terminalsitzung ein Stoppsignal empfangen.	„Problem: Auf Solaris-Systemen werden Load-Balancer-Prozesse beim Verlassen des Terminalfensters, in dem die Prozesse gestartet wurden, beendet“ auf Seite 319

Tabelle 17. Tabelle zur Fehlerbehebung für Controller für Cisco CSS Switches

Fehler	Mögliche Ursache	Siehe Abschnitt ...
ccoserver wird nicht gestartet	In Konflikt stehende Port-Nummern	„Port-Nummern für Cisco CSS Controller überprüfen“ auf Seite 303
Der Befehl ccocontrol oder lbadm in scheitert mit der Nachricht ‘Server antwortet nicht’ oder ‘Zugriff auf RMI-Server nicht möglich’.	Befehle können nicht ausgeführt werden, weil der Stack SOCKSifiziert ist oder ccoserver nicht gestartet wurde.	„Problem: Der Befehl ccocontrol oder lbadm in scheitert“ auf Seite 334
Empfangener Fehler: Für Port 13099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden.	Abgelaufene Produktlizenz	„Problem: Für Port 13099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden“ auf Seite 334
Bei Verwendung einer Matrox-AGP-Videokarte auf einer Windows-Plattform kommt es zu unerwartetem GUI-Verhalten.	Der Fehler tritt auf, wenn Matrox-AGP-Videokarten während der Ausführung der Load-Balancer-GUI verwendet werden.	„Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten“ auf Seite 335
Empfang eines Verbindungsfehlers, wenn ein Consultant hinzugefügt wird.	Die Konfigurationseinstellungen für den Switch oder den Controller stimmen nicht.	„Problem: Beim Hinzufügen eines Consultant wird ein Verbindungsfehler empfangen“ auf Seite 335
Die Wertigkeiten werden auf dem Switch nicht aktualisiert.	Die Kommunikation zwischen Controller und Switch ist nicht möglich oder unterbrochen.	„Problem: Auf dem Switch werden die Wertigkeiten nicht aktualisiert“ auf Seite 335
Der Aktualisierungsbefehl (refresh) aktualisiert die Consultant-Konfiguration nicht.	Die Kommunikation zwischen Switch und Controller ist nicht möglich oder unterbrochen.	„Problem: Befehl refresh aktualisiert die Consultant-Konfiguration nicht“ auf Seite 336
Die GUI blockiert oder verhält sich nicht erwartungsgemäß, wenn versucht wird, eine große Konfigurationsdatei zu laden.	Java kann nicht auf so viel Speicher zugreifen, wie für die Bearbeitung einer so großen Änderung der GUI erforderlich ist.	„Problem: Beim Laden einer großen Konfigurationsdatei blockiert die GUI oder verhält sich nicht erwartungsgemäß“ auf Seite 311

Tabelle 17. Tabelle zur Fehlerbehebung für Controller für Cisco CSS Switches (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Bei der fernen Webverwaltung mit Netscape wird die Verbindung zum Host getrennt.	Die Verbindung zum Host wird getrennt, wenn Sie die Größe des Browserfensters ändern.	„Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung“ auf Seite 336
Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1).	Ändern Sie die Schriftartmerkmale des Fensters mit der Eingabeaufforderung.	„Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)“ auf Seite 336
Auf der HP-UX-Plattform wird die folgende Nachricht angezeigt: java.lang.OutOfMemoryError unable to create new native thread	Einige HP-UX-Installationen lassen standardmäßig 64 Threads pro Prozess zu. Dies ist unzureichend.	„Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads“ auf Seite 336
Auf Solaris-Systemen werden die Load-Balancer-Prozesse beendet, wenn Sie das Fenster mit der Terminalsitzung verlassen, in dem die Prozesse gestartet wurden.	Verwenden Sie den Befehl nohup , um zu verhindern, dass die gestarteten Prozesse beim Verlassen der Terminalsitzung ein Stoppsignal empfangen.	„Problem: Auf Solaris-Systemen werden Load-Balancer-Prozesse beim Verlassen des Terminalfensters, in dem die Prozesse gestartet wurden, beendet“ auf Seite 319

Tabelle 18. Tabelle zur Fehlerbehebung für Nortel Alteon Controller

Fehler	Mögliche Ursache	Siehe Abschnitt ...
nalserver wird nicht gestartet	In Konflikt stehende Port-Nummern	„Port-Nummern für Nortel Alteon Controller überprüfen“ auf Seite 304
Der Befehl nalcontrol oder lbadmin scheitert mit der Nachricht ‘Server antwortet nicht’ oder ‘Zugriff auf RMI-Server nicht möglich’.	Befehle können nicht ausgeführt werden, weil der Stack SOCKSifiziert ist oder nalserver nicht gestartet wurde.	„Problem: Der Befehl nalcontrol oder lbadmin scheitert“ auf Seite 337
Empfangener Fehler: Für Port 14099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden.	Abgelaufene Produktlizenz	„Problem: Für Port 14099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden“ auf Seite 337
Bei Verwendung einer Matrox-AGP-Videokarte auf einer Windows-Plattform kommt es zu unerwartetem GUI-Verhalten.	Der Fehler tritt auf, wenn Matrox-AGP-Videokarten während der Ausführung der Load-Balancer-GUI verwendet werden.	„Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten“ auf Seite 337
Die GUI blockiert oder verhält sich nicht erwartungsgemäß, wenn versucht wird, eine große Konfigurationsdatei zu laden.	Java kann nicht auf so viel Speicher zugreifen, wie für die Bearbeitung einer so großen Änderung der GUI erforderlich ist.	„Problem: Beim Laden einer großen Konfigurationsdatei blockiert die GUI oder verhält sich nicht erwartungsgemäß“ auf Seite 311

Tabelle 18. Tabelle zur Fehlerbehebung für Nortel Alteon Controller (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Bei der fernen Webverwaltung mit Netscape wird die Verbindung zum Host getrennt.	Die Verbindung zum Host wird getrennt, wenn Sie die Größe des Browserfensters ändern.	„Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung“ auf Seite 338
Empfang eines Verbindungsfehlers, wenn ein Consultant hinzugefügt wird.	Die Konfigurationseinstellungen für den Switch oder den Controller stimmen nicht.	„Problem: Beim Hinzufügen eines Consultant wird ein Verbindungsfehler empfangen“ auf Seite 338
Die Wertigkeiten werden auf dem Switch nicht aktualisiert.	Die Kommunikation zwischen Controller und Switch ist nicht möglich oder unterbrochen.	„Problem: Auf dem Switch werden die Wertigkeiten nicht aktualisiert“ auf Seite 338
Der Aktualisierungsbefehl (refresh) aktualisiert die Consultant-Konfiguration nicht.	Die Kommunikation zwischen Switch und Controller ist nicht möglich oder unterbrochen.	„Problem: Befehl refresh aktualisiert die Consultant-Konfiguration nicht“ auf Seite 338
Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1).	Ändern Sie die Schriftartmerkmale des Fensters mit der Eingabeaufforderung.	„Problem: Auf Windows-Systemen erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)“ auf Seite 339
Auf der HP-UX-Plattform wird die folgende Nachricht angezeigt: java.lang.OutOfMemoryError unable to create new native thread	Einige HP-UX-Installationen lassen standardmäßig 64 Threads pro Prozess zu. Dies ist unzureichend.	„Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads“ auf Seite 339
Auf Solaris-Systemen werden die Load-Balancer-Prozesse beendet, wenn Sie das Fenster mit der Terminalsitzung verlassen, in dem die Prozesse gestartet wurden.	Verwenden Sie den Befehl nohup , um zu verhindern, dass die gestarteten Prozesse beim Verlassen der Terminalsitzung ein Stoppsignal empfangen.	„Problem: Auf Solaris-Systemen werden Load-Balancer-Prozesse beim Verlassen des Terminalfensters, in dem die Prozesse gestartet wurden, beendet“ auf Seite 319

Tabelle 19. Tabelle zur Fehlerbehebung für Metric Server

Fehler	Mögliche Ursache	Siehe Abschnitt ...
IOException für Metric Server auf der Windows-Plattform bei Ausführung von benutzerdefinierten Messwertdateien mit der Erweiterung .bat oder .cmd	Es ist ein vollständiger Messwertname erforderlich.	„Problem: IOException für Metric Server auf der Windows-Plattform bei Ausführung von benutzerdefinierten Messwertdateien mit der Erweiterung .bat oder .cmd“ auf Seite 339

Tabelle 19. Tabelle zur Fehlerbehebung für Metric Server (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Metric Server meldet die Lastinformationen nicht an die Load-Balancer-Maschine.	<p>Mögliche Ursachen sind unter anderem:</p> <ul style="list-style-type: none"> • Auf der Metric-Server-Maschine gibt es keine Schlüsselringe. • Der Hostname der Metric-Server-Maschine ist nicht im lokalen Namensserver registriert. • Die Datei /etc/hosts legt fest, dass der lokale Hostname in die Loopback-Adresse 127.0.0.1 aufgelöst wird. 	„Problem: Metric Server meldet die Last nicht an die Load-Balancer-Maschine“ auf Seite 339
Beim Übertragen von Schlüsselringen zum Server enthält das Metric-Server-Protokoll den Eintrag "Für den Zugriff auf den Agenten ist eine Kennung erforderlich".	Der Schlüsselring ist beschädigt und kann deshalb nicht autorisiert werden.	„Problem: Metric-Server-Protokoll meldet, dass für den Zugriff auf den Agenten eine Kennung erforderlich ist“ auf Seite 340
Wenn Metric Server auf einem AIX-Multiprozessor-system (AIX 4.3.3 oder AIX 5.1) mit starker Belastung ausgeführt wird, kann die Ausgabe des Befehls ps -vg beschädigt werden.	Dieses bekannte AIX-Problem wird von APAR IY33804 korrigiert.	„Problem: Bei Ausführung von Metric Server auf AIX-Systemen kann die Ausgabe des Befehls ps -vg beschädigt werden“ auf Seite 340
Konfigurieren von Metric Server in einer zweistufigen Konfiguration mit Site Selector zur Lastverteilung auf Dispatcher mit hoher Verfügbarkeit	Metric Server (zweite Stufe) ist nicht für Empfangsbereitschaft an einer neuen IP-Adresse konfiguriert.	„Problem: Konfigurieren von Metric Server in einer zweistufigen Konfiguration mit Site Selector für die Lastverteilung auf Dispatcher mit hoher Verfügbarkeit“ auf Seite 340
Scripts (metricserver, cpuload, memload), die auf Solaris-Maschinen mit mehreren CPUs ausgeführt werden, erzeugen unerwünschte Konsolnachrichten.	Dieses Verhalten ist auf die Verwendung des VMSTAT-Systembefehls zum Abrufen von CPU- und Speicherstatistiken vom Kernel zurückzuführen.	„Problem: Unerwünschte Konsolnachrichten bei Ausführung von Scripts auf Solaris-Maschinen mit mehreren CPUs“ auf Seite 342
Auf Solaris-Systemen werden die Load-Balancer-Prozesse beendet, wenn Sie das Fenster mit der Terminalsitzung verlassen, in dem die Prozesse gestartet wurden.	Verwenden Sie den Befehl nohup , um zu verhindern, dass die gestarteten Prozesse beim Verlassen der Terminalsitzung ein Stoppsignal empfangen.	„Problem: Auf Solaris-Systemen werden Load-Balancer-Prozesse beim Verlassen des Terminalfensters, in dem die Prozesse gestartet wurden, beendet“ auf Seite 319

Tabelle 19. Tabelle zur Fehlerbehebung für Metric Server (Forts.)

Fehler	Mögliche Ursache	Siehe Abschnitt ...
Bei Ausführung von Load Balancer für IPV6 ist auf Linux-Systemen kein Abruf der Werte von Metric Server möglich.	Auf Linux-Plattformen gibt es eine Inkompatibilität bei der Auswahl der Quellen-IPV6-Adresse. Die Metriküberwachung versucht deshalb, über die falsche Quellen-IP-Adresse mit Metric Server zu kommunizieren.	„Bei Ausführung von Load Balancer für IPV6 ist auf Linux-Systemen kein Abruf der Werte von Metric Server möglich.“ auf Seite 342
Messwert gibt nach dem Start von Metric Server -1 zurück.	Dieses Problem kann dadurch entstehen, dass die Schlüsseldateien bei der Übertragung zum Client ihre Integrität verlieren.	„Problem: Nach dem Start von Metric Server wird der Wert -1 zurückgegeben“ auf Seite 343

Port-Nummern für Dispatcher überprüfen

Falls beim Ausführen des Dispatchers Probleme auftreten, verwendet unter Umständen eine Ihrer Anwendungen eine Port-Nummer, die normalerweise vom Dispatcher benutzt wird. Der Dispatcher-Server benutzt die folgenden Port-Nummern:

- 10099 zum Empfangen der Befehle von dscontrol
- 10004 zum Senden von Messwertabfragen an Metric Server
- 10199 für den RMI-Server-Port

Wenn eine andere Anwendung eine der Port-Nummern des Dispatchers verwendet, können Sie die Port-Nummern des Dispatchers *oder* die Port-Nummer der Anwendung ändern.

Gehen Sie zum Ändern der Port-Nummern des Dispatchers wie folgt vor:

- Den zum Empfang von Befehlen verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable `LB_RMIPORT` am Anfang der Datei `dsserver` auf den Port, an dem der Dispatcher Befehle empfangen soll.
- Den zum Empfang der Messwerte von Metric Server verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable `RMI_PORT` in der Datei `metricserver` auf den Port, über den Dispatcher mit Metric Server kommunizieren soll.
 - Geben Sie beim Starten des Managers das Argument `metric_port` an. Eine Beschreibung der Befehlssyntax für **dscontrol manager start** finden Sie im Abschnitt „dscontrol manager — Manager steuern“ auf Seite 375.

Gehen Sie zum Ändern der RMI-Port-Nummer der Anwendung wie folgt vor:

- Den von der Anwendung verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable `LB_RMISERVERPORT` in der Datei `dsserver` auf den Port, den die Anwendung benutzen soll. (Der Standardwert für den von der Anwendung verwendeten RMI-Port ist 10199.)

Anmerkung: Auf der Windows-Plattform befinden sich die Dateien `dsserver` und `metricserver` im Verzeichnis `C:\winnt\system32`. Auf anderen Plattformen sind diese Dateien im Verzeichnis `/usr/bin/` enthalten.

Port-Nummern für CBR überprüfen

Wenn beim Ausführen von CBR Fehler auftreten, verwendet unter Umständen eine Ihrer Anwendungen eine Port-Nummer, die normalerweise von CBR benutzt wird. CBR benutzt die folgenden Port-Nummern:

- 11099 zum Empfangen der Befehle von cbrcontrol
- 10004 zum Senden von Messwertabfragen an Metric Server
- 11199 für den RMI-Server-Port

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode cbr der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Wenn eine andere Anwendung eine der Port-Nummern für CBR verwendet, können Sie die Port-Nummern für CBR *oder* die Port-Nummer der Anwendung ändern.

Gehen Sie zum Ändern der Port-Nummern für CBR wie folgt vor:

- Den zum Empfang von Befehlen verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable LB_RMIPORT am Anfang der Datei cbrserver auf den Port, an dem CBR Befehle empfangen soll.
- Den zum Empfang der Messwerte von Metric Server verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable RMI_PORT in der Datei metricserver auf den Port, über den CBR mit Metric Server kommunizieren soll.
 - Geben Sie beim Starten des Managers das Argument metric_port an. Eine Beschreibung der Befehlssyntax für **manager start** finden Sie im Abschnitt „dscontrol manager — Manager steuern“ auf Seite 375.

Gehen Sie zum Ändern der RMI-Port-Nummer der Anwendung wie folgt vor:

- Den von der Anwendung verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable LB_RMISERVERPORT am Anfang der Datei cbrserver auf den Port, den die Anwendung benutzen soll. (Der Standardwert für den von der Anwendung verwendeten RMI-Port ist 11199.)

Anmerkung: Auf der Windows-Plattform befinden sich die Dateien cbrserver und metricserver im Verzeichnis C:\winnt\system32. Auf anderen Plattformen sind diese Dateien im Verzeichnis /usr/bin/ enthalten.

Port-Nummern für Site Selector überprüfen

Wenn bei Ausführung der Komponente Site Selector Fehler auftreten, verwendet unter Umständen eine Ihrer Anwendungen eine Port-Nummer, die normalerweise von Site Selector benutzt wird. Site Selector benutzt die folgenden Port-Nummern:

- 12099 zum Empfangen der Befehle von sscontrol
- 10004 zum Senden von Messwertabfragen an Metric Server
- 12199 für den RMI-Server-Port

Wenn eine andere Anwendung eine der Port-Nummern für Site Selector verwendet, können Sie die Port-Nummern für Site Selector *oder* die Port-Nummer der Anwendung ändern.

Gehen Sie zum Ändern der Port-Nummern für Site Selector wie folgt vor:

- Den zum Empfang von Befehlen verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable LB_RMIPORT am Anfang der Datei ssserver auf den Port, an dem Site Selector Befehle empfangen soll.
- Den zum Empfang der Messwerte von Metric Server verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable RMI_PORT in der Datei metricserver auf den Port, über den Site Selector mit Metric Server kommunizieren soll.
 - Geben Sie beim Starten des Managers das Argument metric_port an. Eine Beschreibung der Befehlssyntax für **manager start** finden Sie im Abschnitt „sscontrol manager — Manager steuern“ auf Seite 415.

Gehen Sie zum Ändern der RMI-Port-Nummer der Anwendung wie folgt vor:

- Den von der Anwendung verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable LB_RMISERVERPORT am Anfang der Datei ssserver auf den Port, den die Anwendung benutzen soll. (Der Standardwert für den von der Anwendung verwendeten RMI-Port ist 12199.)

Anmerkung: Auf der Windows-Plattform befinden sich die Dateien ssserver und metricserver im Verzeichnis C:\winnt\system32. Auf anderen Plattformen sind diese Dateien im Verzeichnis /usr/bin/ enthalten.

Port-Nummern für Cisco CSS Controller überprüfen

Wenn bei Ausführung der Komponente Cisco CSS Controller Fehler auftreten, verwendet unter Umständen eine andere Anwendung eine Port-Nummer, die normalerweise vom ccoserver der Komponente benutzt wird. Cisco CSS Controller benutzt die folgenden Port-Nummern:

- 13099 zum Empfangen der Befehle von ccocontrol
- 10004 zum Senden von Messwertabfragen an Metric Server
- 13199 für den RMI-Server-Port

Wenn eine andere Anwendung eine der Port-Nummern des Cisco CSS Controller verwendet, können Sie die Port-Nummern des Cisco CSS Controller *oder* die Port-Nummer der Anwendung ändern.

Gehen Sie zum Ändern der Port-Nummern des Cisco CSS Controller wie folgt vor:

- Zum Ändern des für den Empfang der Befehle von ccocontrol verwendeten Ports müssen Sie die Variable CCO_RMIPORT in der Datei ccoserver ändern. Ersetzen Sie den Wert 13099 durch die Nummer des Ports, an dem Cisco CSS Controller ccocontrol-Befehle empfangen soll.
- Den zum Empfang der Messwerte von Metric Server verwendeten Port können Sie wie folgt ändern:
 1. Ändern Sie die Variable RMI_PORT in der Datei metricserver. Ersetzen Sie den Wert 10004 durch die Nummer des Ports, über den Cisco CSS Controller mit Metric Server kommunizieren soll.
 2. Geben Sie beim Starten des Consultant das Argument metric_port an.

Gehen Sie zum Ändern der RMI-Port-Nummer der Anwendung wie folgt vor:

- Den von der Anwendung verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable CCO_RMISERVERPORT am Anfang der Datei ccoserver auf den Port, den die Anwendung benutzen soll. (Der Standardwert für den von der Anwendung verwendeten RMI-Port ist 13199.)

Anmerkung: Auf der Windows-Plattform befinden sich die Dateien ccoserver und metricserver im Verzeichnis C:\winnt\system32. Auf anderen Plattformen sind diese Dateien im Verzeichnis /usr/bin enthalten.

Port-Nummern für Nortel Alteon Controller überprüfen

Wenn bei Ausführung der Komponente Nortel Alteon Controller Fehler auftreten, verwendet unter Umständen eine andere Anwendung eine Port-Nummer, die normalerweise vom nalserver der Komponente benutzt wird. Nortel Alteon Controller benutzt die folgenden Port-Nummern:

- 14099 zum Empfang der Befehle von nalcontrol
- 10004 zum Senden von Messwertabfragen an Metric Server
- 14199 für den RMI-Server-Port

Wenn eine andere Anwendung eine der Port-Nummern des Nortel Alteon Controller verwendet, können Sie die Port-Nummern des Nortel Alteon Controller *oder* die Port-Nummern der Anwendung ändern.

Gehen Sie zum Ändern der Port-Nummern des Nortel Alteon Controller wie folgt vor:

- Zum Ändern des für den Empfang der Befehle von nalcontrol verwendeten Ports müssen Sie die Variable NAL_RMIPORT in der Datei nalserver ändern. Ersetzen Sie den Wert 14099 durch die Nummer des Ports, an dem Nortel Alteon Controller nalcontrol-Befehle empfangen soll.
- Den zum Empfang der Messwerte von Metric Server verwendeten Port können Sie wie folgt ändern:
 1. Ändern Sie die Variable RMI_PORT in der Datei metricserver. Ersetzen Sie den Wert 10004 durch die Nummer des Ports, über den Nortel Alteon Controller mit Metric Server kommunizieren soll.
 2. Geben Sie beim Starten des Consultant das Argument metric_port an.

Gehen Sie zum Ändern der RMI-Port-Nummer der Anwendung wie folgt vor:

- Den von der Anwendung verwendeten Port können Sie wie folgt ändern:
 - Setzen Sie die Variable NAL_RMISERVERPORT am Anfang der Datei nalserver auf den Port, den die Anwendung benutzen soll. (Der Standardwert für den von der Anwendung verwendeten RMI-Port ist 14199.)

Anmerkung: Auf der Windows-Plattform befinden sich die Dateien nalserver und metricserver im Verzeichnis C:\winnt\system32. Auf anderen Plattformen sind diese Dateien im Verzeichnis /usr/bin enthalten.

Allgemeine Probleme lösen — Dispatcher

Problem: Dispatcher wird nicht ausgeführt

Dieses Problem kann auftreten, wenn eine andere Anwendung einen der Ports benutzt, die normalerweise vom Dispatcher verwendet werden. Weitere Informationen enthält der Abschnitt „Port-Nummern für Dispatcher überprüfen“ auf Seite 301.

Problem: Dispatcher und Server antworten nicht

Dieses Problem tritt auf, wenn eine andere als die angegebene Adresse verwendet wird. Stellen Sie bei Verknüpfung von Dispatcher und Server sicher, dass die in der Konfiguration verwendete Serveradresse die NFA ist oder als verknüpft konfiguriert ist. Überprüfen Sie außerdem die Adresse der Hostdatei.

Problem: Dispatcher-Anforderungen werden nicht verteilt

Symptome für dieses Problem sind: Verbindungen von Client-Maschinen werden nicht bedient oder Verbindungen überschreiten ein Zeitlimit. Überprüfen Sie Folgendes, um diesen Fehler zu bestimmen:

1. Haben Sie die NFA, Cluster, Ports und Server für die Weiterleitung konfiguriert? Überprüfen Sie die Konfigurationsdatei.
2. Wurde auf der Netzschnittstellenkarte ein Aliasname für die Clusteradresse angegeben? Auf Linux- und UNIX-Systemen können Sie dies mit `netstat -ni` überprüfen.
3. Ist der Aliasname der Loopback-Einheit jedes Servers auf die Clusteradresse gesetzt? Auf Linux- und UNIX-Systemen können Sie dies mit `netstat -ni`überprüfen.
4. Wurde die zusätzliche Route gelöscht? Auf Linux- und UNIX-Systemen können Sie dies mit `netstat -nr`überprüfen.
5. Benutzen Sie den Befehl `dscontrol cluster status`, um die Informationen für die einzelnen definierten Cluster zu überprüfen. Vergewissern Sie sich, dass Sie für jeden Cluster einen Port definiert haben.
6. Stellen Sie mit dem Befehl `dscontrol server report ::` sicher, dass Ihre Server nicht inaktiv sind und nicht die Wertigkeit null haben.

Beachten Sie für Windows und andere Plattformen auch die Informationen im Abschnitt „Servermaschinen für Lastausgleich konfigurieren“ auf Seite 72.

Problem: Die Dispatcher-Funktion für hohe Verfügbarkeit kann nicht ausgeführt werden

Dieses Problem tritt auf, wenn eine Dispatcher-Umgebung mit hoher Verfügbarkeit konfiguriert ist und Verbindungen von Client-Maschinen nicht bedient werden oder Zeitlimits überschreiten. Überprüfen Sie Folgendes, um den Fehler zu korrigieren oder zu bestimmen:

- Überprüfen Sie, ob Sie die Scripts `goActive`, `goStandby` und `goInOp` erstellt haben, und stellen Sie sie in das Unterverzeichnis `bin` des Installationsverzeichnis von Dispatcher. Weitere Informationen zu diesen Scripts finden Sie im Abschnitt „Scripts verwenden“ auf Seite 209.
- Vergewissern Sie sich bei **AIX**-, **HP-UX**-, **Linux**- und **Solaris**-Systemen, dass in den Scripts `goActive`, `goStandby` und `goInOp` die Option `execute permission` gesetzt ist.

- Bei Windows-Systemen müssen Sie die NFA mit dem Befehl **executor configure** konfigurieren.

Mit den folgenden Schritten kann effektiv getestet werden, ob Scripts für hohe Verfügbarkeit fehlerfrei funktionieren:

1. Erstellen Sie einen Bericht, indem Sie auf der Maschine `netstat -an` und `ifconfig -a` absetzen.
2. Führen Sie das Script `goActive` aus.
3. Führen Sie das Script `goStandby` aus.
4. Erstellen Sie erneut einen Bericht, indem Sie die Befehle `netstat -an` und `ifconfig -a` absetzen.

Die beiden Berichte sind identisch, wenn die Scripts richtig konfiguriert sind.

Problem: Es kann kein Überwachungssignal hinzugefügt werden (Windows-Plattform)

Dieser Fehler tritt auf der Windows-Plattform auf, wenn die Quellenadresse auf keinem Adapter konfiguriert ist. Überprüfen Sie Folgendes, um den Fehler zu korrigieren oder zu bestimmen:

- Sie müssen die NFA mit der Token-Ring- oder Ethernet-Schnittstelle konfigurieren und einen der folgenden Befehle absetzen:
`dscontrol executor configure <IP-Adresse>`

Problem: Zusätzliche Routen (Windows 2000)

Nach dem Konfigurieren von Servermaschinen stellen Sie unter Umständen fest, dass Sie unbeabsichtigt eine oder mehrere zusätzliche Route(n) erstellt haben. Werden diese zusätzlichen Routen nicht entfernt, kann der Dispatcher nicht ordnungsgemäß arbeiten. Informationen zum Feststellen und Löschen zusätzlicher Routen finden Sie im Abschnitt „Servermaschinen für Lastausgleich konfigurieren“ auf Seite 72.

Problem: Advisor-Funktionen arbeiten nicht korrekt

Wenn Sie die Weitverkehrsunterstützung verwenden und die Advisor-Funktionen nicht ordnungsgemäß zu arbeiten scheinen, müssen Sie sicherstellen, dass sie sowohl auf den lokalen als auch auf den fernen Dispatchern gestartet wurden.

Vor der Advisor-Anforderung wurde ein ICMP ping an die Server abgesetzt. Falls es zwischen Load Balancer und den Servern eine Firewall gibt, stellen Sie sicher, dass pings über die Firewall unterstützt werden. Stellt eine solche Konfiguration ein Sicherheitsrisiko für Netz dar, modifizieren Sie die Anweisung `java` in `dsserver`, um alle pings an die Server zu inaktivieren. Fügen Sie dazu wie folgt das Merkmal `java` hinzu:

```
LB_ADV_NO_PING="true"
java -DLB_ADV_NO_PING="true"
```

Lesen Sie hierzu die Informationen im Abschnitt „Ferne Advisor-Funktionen mit der Dispatcher-WAN-Unterstützung verwenden“ auf Seite 231.

Problem: Dispatcher, Microsoft IIS und SSL funktionieren nicht (Windows-Plattform)

Wenn Sie Dispatcher, Microsoft IIS und SSL verwenden und diese Komponenten nicht zusammenarbeiten, kann dies auf ein Problem mit der Aktivierung der SSL-

Sicherheit zurückzuführen sein. Weitere Informationen dazu, wie Sie ein Schlüsselpaar generieren, ein Zertifikat erhalten und ein Verzeichnis so konfigurieren, dass es SSL erfordert, finden Sie in der Dokumentation *Microsoft Information and Peer Web Services*.

Problem: Dispatcher-Verbindung zu einer fernen Maschine

Der Dispatcher verwendet Schlüssel, die Ihnen ermöglichen, eine Verbindung zu einer fernen Maschine herzustellen und die Maschine zu konfigurieren. Die Schlüssel geben einen RMI-Port für die Verbindung an. Sie können den RMI-Port aus Sicherheitsgründen oder bei Konflikten ändern. Wird der RMI-Port geändert, ändert sich auch der Dateiname des Schlüssels. Wenn Ihr Schlüsselverzeichnis für eine ferne Maschine mehrere Schlüssel enthält, die verschiedene RMI-Ports angeben, verwendet die Befehlszeile nur den ersten gefundenen Schlüssel. Ist dies der falsche Schlüssel, wird die Verbindung zurückgewiesen. Die Verbindung wird erst hergestellt, wenn der falsche Schlüssel gelöscht wurde.

Problem: Der Befehl dscontrol oder lbadmin scheitert

1. Der Befehl dscontrol gibt die Nachricht **Fehler: Server antwortet nicht** zurück, oder der Befehl lbadmin gibt die Nachricht **Fehler: Zugriff auf RMI-Server nicht möglich** zurück. Diese Fehler können auftreten, wenn der Stack Ihrer Maschine SOCKSifiziert ist. Lösen Sie dieses Problem, indem Sie die Datei socks.cnf so editieren, dass sie die folgenden Zeilen enthält:

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Die Verwaltungskonsolen für Load-Balancer-Schnittstellen (Befehlszeile, grafische Benutzerschnittstelle und Assistenten) kommunizieren per RMI (Remote Method Invocation) mit dsserver. Für die Standardkommunikation werden drei Ports verwendet, die im Start-Skript für dsserver wie folgt definiert sind:
 - 10099 zum Empfangen der Befehle von dscontrol
 - 10004 zum Senden von Messwertabfragen an Metric Server
 - 10199 für den RMI-Server-Port

Diese Definition kann Fehler verursachen, wenn eine der Verwaltungskonsolen auf derselben Maschine als Firewall oder über eine Firewall ausgeführt wird. Wird beispielsweise Load Balancer auf derselben Maschine als Firewall ausgeführt, können beim Absetzen von dscontrol-Befehlen Fehler wie der folgende angezeigt werden: **Fehler: Server antwortet nicht**.

Sie können diesen Fehler vermeiden, indem Sie die dsserver-Script-Datei editieren und den von RMI für die Firewall (oder eine andere Anwendung) verwendeten Port festlegen. Ändern Sie die Zeile LB_RMISERVERPORT=10199 in LB_RMISERVERPORT=*Ihr_Port*. *Ihr_Port* ist ein anderer Port.

Starten Sie anschließend erneut dsserver und öffnen Sie den Datenverkehr für die Ports 10099, 10004, 10199 und 10100 oder für den Port, den Sie für die Hostadresse, an der die Verwaltungskonsolle ausgeführt wird, ausgewählt haben.

3. Derartige Fehler können auch auftreten, wenn Sie dsserver noch nicht gestartet haben.
4. Falls es in der Maschine mehrere Adapter gibt, müssen Sie festlegen, welchen Adapter dsserver verwenden soll. Fügen Sie dazu Folgendes zum dsserver-Script hinzu: `java.rmi.server.hostname=<Hostname oder IP-Adresse>`
Beispiel: `java -Djava.rmi.server.hostname="10.1.1.1"`

Problem: Fehlernachricht 'Datei nicht gefunden...' beim Anzeigen der Onlinehilfe (Windows-Plattform)

Wenn Sie auf Windows-Plattformen Netscape als Standardbrowser verwenden, kann bei diesem Fehler die folgende Nachricht erscheinen: "Netscape kann die Datei '<Dateiname>.html' (oder eine ihrer Komponenten) nicht finden. Stellen Sie sicher, dass Pfad- und Dateiname stimmen und alle erforderlichen Bibliotheken verfügbar sind".

Das Problem beruht auf einer falschen Einstellung für die HTML-Dateizuordnung. Das Problem kann wie folgt gelöst werden:

1. Klicken Sie auf **Arbeitsplatz**, klicken Sie auf **Extras**, wählen Sie **Ordneroptionen** aus und klicken Sie auf die Registerkarte **Dateitypen**.
2. Wählen Sie "Netscape Hypertext Document" aus.
3. Klicken Sie auf den Knopf **Erweitert**, wählen Sie **open** aus und klicken Sie auf den Knopf **Bearbeiten**.
4. Geben Sie im Feld **Anwendung**: (nicht im Feld "Anwendung für diesen Vorgang:") *NSShell* ein und klicken Sie auf **OK**.

Problem: Die grafische Benutzerschnittstelle (GUI) wird nicht richtig gestartet

Die grafische Benutzerschnittstelle lbadmin erfordert für eine einwandfreie Funktion einen ausreichenden Paging-Bereich. Wenn der Paging-Bereich nicht ausreicht, wird die GUI möglicherweise nicht vollständig gestartet. Überprüfen Sie in einem solchen Fall den Paging-Bereich und vergrößern Sie ihn gegebenenfalls.

Problem: Fehler bei der Ausführung von Dispatcher mit installiertem Caching Proxy

Wenn Sie Load Balancer deinstallieren, um eine andere Version zu installieren, und bei dem Versuch, die Komponente Dispatcher zu starten, eine Fehlernachricht empfangen, überprüfen Sie, ob Caching Proxy installiert ist. Caching Proxy ist von einer der Dispatcher-Dateien abhängig. Diese Datei wird nur bei der Deinstallation von Caching Proxy deinstalliert.

Sie können dieses Problem wie folgt vermeiden:

1. Deinstallieren Sie Caching Proxy.
2. Deinstallieren Sie Load Balancer.
3. Installieren Sie Load Balancer und Caching Proxy erneut.

Problem: Die grafische Benutzerschnittstelle (GUI) wird nicht richtig angezeigt

Falls die GUI von Load Balancer nicht richtig angezeigt wird, überprüfen Sie die Auflösung für den Desktop des Betriebssystems. Die GUI wird am besten bei einer Auflösung von 1024 x 768 Bildpunkten angezeigt.

Problem: Auf der Windows-Plattform sind die Hilfefenster manchmal von anderen offenen Fenstern verdeckt

Wenn Sie auf der Windows-Plattform zum ersten Mal Hilfefenster öffnen, werden diese manchmal hinter vorhandene Fenster gestellt. Klicken Sie in diesem Fall auf das Fenster, um es wieder in den Vordergrund zu stellen.

Problem: Load Balancer kann Rahmen nicht verarbeiten und weiterleiten

Unter Solaris haben alle Netzwerkadapter standardmäßig dieselbe MAC-Adresse. Wenn sich jeder Adapter in einem anderen IP-Teilnetz befindet, verursacht dies keine Probleme. In einer Switch-Umgebung, in der mehrere NICs mit derselben MAC-Adresse und derselben IP-Teilnetzadresse mit einem Switch kommunizieren, sendet der Switch den gesamten für die eine MAC-Adresse (und beide IP-Adressen) bestimmten Datenverkehr über eine Leitung. Nur der Adapter, der als letzter einen Rahmen über die Leitung gesendet hat, sieht die für beide Adapter bestimmten IP-Pakete. Solaris löscht unter Umständen Pakete für eine gültige IP-Adresse, die an der "falschen" Schnittstelle ankommen.

Wenn in `ibmlb.conf` nicht alle Netzschnittstellen für Load Balancer konfiguriert sind und die nicht in `ibmlb.conf` definierte NIC einen Rahmen empfängt, kann Load Balancer den Rahmen nicht verarbeiten und weiterleiten.

Sie können dieses Problem vermeiden, indem Sie die Standardeinstellung überschreiben und für jede Schnittstelle eine eindeutige MAC-Adresse definieren. Verwenden Sie dafür den folgenden Befehl:

```
ifconfig Schnittstelle ether MAC-Adresse
```

Beispiel:

```
ifconfig eri0 ether 01:02:03:04:05:06
```

Problem: Beim Starten des Executors von Load Balancer erscheint eine blaue Anzeige

Auf der Windows-Plattform müssen Sie vor dem Starten des Executors eine Netzwerkkarte installiert und konfiguriert haben.

Problem: Automatische Pfaderkennung verhindert Datenrückfluss mit Load Balancer

Unter AIX gibt es einen Netzwerkparameter für die automatische Erkennung der MTU, die auf einem Pfad transportiert werden kann. Stellt das Betriebssystem während einer Transaktion mit einem Client fest, dass es für ausgehende Pakete eine kleinere MTU (größte zu übertragende Einheit) verwenden muss, veranlasst die automatische Erkennung der MTU für einen Pfad AIX, eine Route zu erstellen, um sich diese Daten merken zu können. Die neue Route ist für diese spezielle Client-IP-Adresse bestimmt und zeichnet die für das Erreichen der Adresse erforderliche MTU auf.

Nachdem die Route erstellt wurde, könnte auf den Servern ein Problem auftreten, weil auf der Loopback-Einheit ein Aliasname für den Cluster verwendet wird. Wenn die Gateway-Adresse für die Route in das Teilnetz des Clusters / der Netzmaske fällt, erstellen AIX-Systeme die Route zur Loopback-Adresse. Der Grund hierfür ist, dass dies die letzte Schnittstelle war, deren Aliasname dieses Teilnetz enthielt.

Wenn der Cluster beispielsweise 9.37.54.69 ist, die Netzmaske 255.255.255.0 lautet und das angestrebte Gateway 9.37.54.1 ist, verwenden AIX-Systeme die Loopback-Adresse für die Route. Dadurch können die Antworten des Servers die Maschine nicht verlassen und der Client wartet bis zur Überschreitung des Zeitlimits. Normalerweise sieht der Client eine Antwort vom Cluster. Danach wird die Route erstellt und der Client empfängt keine weiteren Antworten.

Für dieses Problem gibt es zwei mögliche Lösungen.

1. Inaktivieren Sie die automatische Erkennung der MTU für einen Pfad, so dass das AIX-System nicht dynamisch Routen hinzufügt. Verwenden Sie dazu die folgenden Befehle:
no -a Listet die AIX-Einstellungen für den Netzbetrieb auf.
no -o option=value
Legt die TCP-Parameter auf AIX-Systemen fest.
2. Geben Sie auf der Loopback-Einheit den Aliasnamen für die Cluster-IP-Adresse mit der Netzmaske 255.255.255.255 an. Dies bedeutet, dass das über den Aliasnamen erreichbare Teilnetz nur die Cluster-IP-Adresse ist. Wenn AIX-Systeme die dynamischen Routen erstellen, stimmt die IP-Adresse des Ziel-Gateway nicht mit diesem Teilnetz überein, weshalb eine Route erstellt wird, die die korrekte Netzschnittstelle verwendet. Löschen Sie anschließend die neue lo0-Route, die während der Festlegung der Aliasnamen erstellt wurde. Suchen Sie dazu die Route zur Loopback-Adresse, deren Netzziel die Cluster-IP-Adresse ist, und löschen Sie sie. Dieser Schritt muss immer ausgeführt werden, wenn für den Cluster ein Aliasname erstellt wird.

Anmerkungen:

1. Bis AIX 4.3.2 ist die automatische Erkennung der MTU für einen Pfad standardmäßig inaktiviert. Ab AIX Version 4.3.3 ist sie jedoch standardmäßig aktiviert.
2. Die folgenden Befehle schalten die automatische Erkennung der MTU für einen Pfad aus und müssen bei jedem Booten des Systems ausgeführt werden. Fügen Sie diese Befehle zur Datei /etc/rc.net hinzu.
 - -o udp_pmtu_discover=0
 - -o tcp_pmtu_discover=0

Problem: Keine hohe Verfügbarkeit im Weitverkehrsmodus von Load Balancer

Wenn Sie Load Balancer für ein WAN konfigurieren, müssen Sie den fernen Dispatcher auf Ihrem lokalen Dispatcher als Server in einem Cluster definieren. In der Regel werden Sie die NFA des fernen Dispatchers als Zieladresse des fernen Servers verwenden. Wenn Sie anschließend die Funktion für hohe Verfügbarkeit auf dem fernen Dispatcher konfigurieren, kann diese nicht ausgeführt werden. Der Grund hierfür ist, dass der lokale Dispatcher immer auf die primäre Maschine des fernen Standorts zeigt, wenn Sie für den Zugriff auf den fernen Server die NFA verwenden.

Sie können dieses Problem wie folgt umgehen:

1. Definieren Sie auf dem fernen Dispatcher einen zusätzlichen Cluster. Für diesen Cluster müssen Sie keine Ports oder Server definieren.
2. Fügen Sie diese Clusteradresse zu Ihren Scripts goActive und goStandby hinzu.
3. Definieren Sie diesen Cluster auf Ihrem lokalen Dispatcher als Server und nicht als NFA des fernen primären Dispatchers.

Wenn der ferne primäre Dispatcher aktiviert wird, verwendet er auf seinem Adapter diese Adresse als Aliasnamen, so dass er Datenverkehr akzeptieren kann. Tritt ein Fehler auf, wird die Adresse auf die Ausweichmaschine versetzt. Der weitere Datenverkehr für diese Adresse wird dann von der Ausweichmaschine akzeptiert.

Problem: Beim Laden einer großen Konfigurationsdatei blockiert die GUI oder verhält sich nicht erwartungsgemäß

Wenn Sie mit lbadmin oder der Webverwaltung (lbwaccess) eine große Konfigurationsdatei (im Schnitt mit mehr als 200 **add**-Befehlen) laden, kann die GUI blockieren oder ein unerwartetes Verhalten zeigen. Ein solches Verhalten wäre beispielsweise ein extrem langsames Reagieren auf Anzeigeänderungen.

Dieses Problem tritt auf, weil Java nicht auf so viel Speicher zugreifen kann, wie für die Bearbeitung einer so großen Konfiguration erforderlich ist.

Die Laufzeitumgebung bietet eine Option an, mit der der Java zur Verfügung stehende Speicherzuordnungspool vergrößert werden kann.

Die Option ist `-Xmxn`, bei der n die maximale Größe des Speicherzuordnungspools in Bytes angibt. Der Wert n muss ein Vielfaches von 1024 und größer als 2 MB sein. Nach dem Wert n können Sie k bzw. K für Kilobytes oder m bzw. M für Megabytes angeben. Zwei Beispiele für gültige Angaben sind `-Xmx128M` und `-Xmx81920k`. Der Standardwert ist `"64M"`. Für Solaris 8 gilt ein Maximalwert von `"4000M"`.

Zum Hinzufügen dieser Option müssten Sie die lbadmin-Script-Datei editieren und wie folgt `"javaw"` in `"javaw -Xmxn"` ändern. (Bei AIX-Systemen müssen Sie `"java"` in `"java -Xmxn"` ändern):

- **AIX-Systeme**

```
javaw -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **HP-UX-Systeme**

```
java -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **Linux-Systeme**

```
javaw -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **Solaris-Systeme**

```
java -Xmx256m -cp $LB_CLASSPATH $LB_INSTALL_PATH $LB_CLIENT_KEYS  
com.ibm.internet.nd.framework.FWK_Main 1>/dev/null 2>&1 &
```

- **Windows-Systeme:**

```
START javaw -Xmx256m -cp %LB_CLASSPATH% %LB_INSTALL_PATH%  
%LB_CLIENT_KEYS% com.ibm.internet.nd.framework.FWK_Main
```

Für n wird kein bestimmter Wert empfohlen. Er sollte jedoch über dem Standardwert für die Option liegen. Ein guter Ausgangswert wäre das Zweifache des Standardwertes.

Problem: lbadmin trennt nach dem Aktualisieren der Konfiguration die Verbindung zum Server

Wenn das Verwaltungsprogramm von Load Balancer (lbadmin) nach dem Aktualisieren der Konfiguration die Verbindung zum Server trennt, überprüfen Sie auf dem Server, den Sie konfigurieren möchten, die Version von dsserver. Stellen Sie sicher, dass diese mit der Version von lbadmin oder dscontrol identisch ist.

Problem: IP-Adressen werden über die Fernverbindung nicht richtig aufgelöst

Wenn Sie einen fernen Client mit einer gesicherten Socks-Implementierung verwenden, ist nicht sichergestellt, dass vollständig qualifizierte Domänen- oder Hostnamen in die richtige IP-Adresse im IP-Adressenformat aufgelöst werden. Möglicherweise fügt die Socks-Implementierung spezifische Socks-bezogene Daten zur DNS-Auflösung hinzu.

Wenn die IP-Adressen über die Fernverbindung nicht richtig aufgelöst werden, geben Sie die IP-Adressen im IP-Adressenformat an.

Problem: Auf der koreanischen Schnittstelle von Load Balancer werden bei AIX- und Linux-Systemen überlappende oder unpassende Schriftarten angezeigt

Überlappende und unpassende Schriftarten auf der koreanischen Load-Balancer-Schnittstelle können Sie wie folgt korrigieren:

Auf AIX-Systemen

1. Stoppen Sie alle Java-Prozesse auf dem AIX-System.
2. Öffnen Sie die Datei `font.properties.ko` in einem Editor. Die Datei befindet sich im Verzeichnis `Ausgangsverzeichnis/jre/lib`. `Ausgangsverzeichnis` steht hier für das Java-Ausgangsverzeichnis.
3. Suchen Sie nach dieser Zeichenfolge:
`-Monotype-TimesNewRomanWT-medium-r-normal
--*-%d-75-75-*--ksc5601.1987-0`
4. Ersetzen Sie alle Vorkommen dieser Zeichenfolge durch Folgendes:
`-Monotype-SansMonoWT-medium-r-normal
--*-%d-75-75-*--ksc5601.1987-0`
5. Sichern Sie die Datei.

Auf Linux-Systemen

1. Stoppen Sie alle Java-Prozesse auf dem System.
2. Öffnen Sie die Datei `font.properties.ko` in einem Editor. Die Datei befindet sich im Verzeichnis `Ausgangsverzeichnis/jre/lib`. `Ausgangsverzeichnis` steht hier für das Java-Ausgangsverzeichnis.
3. Suchen Sie nach dieser Zeichenfolge (ohne Leerzeichen):
`-monotype-timesnewromanwt-medium-r-normal--*-%d-75-75-p-*--microsoft-symbol`
4. Ersetzen Sie alle Vorkommen dieser Zeichenfolge durch Folgendes:
`-monotype-sansmonowt-medium-r-normal--*-%d-75-75-p-*--microsoft-symbol`
5. Sichern Sie die Datei.

Problem: Auf Windows-Systemen wird beim Absetzen von Befehlen wie `hostname` an Stelle der lokalen Adresse die Aliasadresse zurückgegeben

Wenn Sie auf Windows-Systemen einen Aliasnamen auf dem MS Loopback-Adapter definiert haben und bestimmte Befehle absetzen (z. B. `hostname`), reagiert das Betriebssystem falsch und gibt an Stelle der lokalen Adresse die Aliasadresse zurück. Dieser Fehler tritt nicht mehr auf, wenn der neu hinzugefügte Aliasname in der Liste der Netzwerkverbindungen unterhalb der lokalen Adresse aufgeführt ist. Dadurch ist sichergestellt, dass vor der Loopback-Aliasadresse auf die lokale Adresse zugegriffen wird.

Gehen Sie wie folgt vor, um die Liste der Netzwerkverbindungen zu überprüfen:

1. Klicken Sie nacheinander auf **Start > Einstellungen > Netzwerk- und DFÜ-Verbindungen**.
2. Wählen Sie im Menü **Erweitert** den Eintrag **Erweiterte Einstellungen...** aus.
3. Vergewissern Sie sich, dass unter **Verbindungen** als erstes **LAN-Verbindung** aufgeführt ist.
4. Verwenden Sie ggf. auf der rechten Seite die Knöpfe zum Sortieren, um Einträge in der Liste nach oben oder unten zu verschieben.

Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten

Wenn Sie auf der Windows-Plattform eine Matrox-AGP-Karte verwenden, kann es auf der Load-Balancer-GUI zu unerwartetem Verhalten kommen. Beim Klicken mit der Maus kann ein Block etwa von der Größe des Mauszeigers beschädigt werden und zur Umkehrung von Hervorhebungen oder zur Verschiebung von Abbildungen führen. Bei älteren Matrox-Karten wurde dieses Verhalten nicht beobachtet. Für Matrox-AGP-Karten gibt es keine bekannte Korrektur.

Problem: Unerwartetes Verhalten bei Ausführung von 'rmmod ibmlb' (Linux-Systeme)

Wenn dsserver beim manuellen Entfernen des Kernel-Moduls von Load Balancer noch aktiv ist, kann es auf Linux-Systemen zu einem unerwarteten Verhalten kommen. Das System könnte beispielsweise blockieren oder es könnten javacores geschrieben werden. Wenn Sie das Kernel-Modul manuell entfernen möchten, stoppen Sie vorher den dsserver.

Falls dsserver stop nicht funktioniert, stoppen Sie den Java-Prozess mit SRV_KND-ConfigServer. Stoppen Sie den Prozess. Stellen Sie dazu mit dem Befehl `ps -ef|grep SRV_KNDConfigServer` die Prozesskennung des Dienstes fest. Beenden Sie dann den Prozess mit dem Befehl `kill Prozess-ID`.

Sie können den Befehl `rmmod ibmlb` sicher ausführen, um das Load-Balancer-Modul aus dem Kernel zu entfernen.

Problem: Lange Antwortzeiten beim Ausführen von Befehlen auf der Dispatcher-Maschine

Wenn Sie die Komponente Dispatcher für den Lastausgleich einsetzen, kann der Computer mit Client-Datenverkehr überlastet werden. Das Kernel-Modul von Load Balancer hat die höchste Priorität. Wenn dieses Modul ständig mit der Bearbeitung von Client-Paketen beschäftigt ist, kann der Rest des Systems möglicherweise nicht mehr reagieren. Die Ausführung von Befehlen im Benutzerbereich kann dann sehr lange dauern. Unter Umständen werden die Befehle gar nicht vollständig ausgeführt.

In einer solchen Situation sollten Sie Ihre Konfiguration neu strukturieren, um eine Überlastung der Load-Balancer-Maschine zu vermeiden. Mögliche Alternativen wären die Verteilung der Last auf mehrere Load-Balancer-Maschinen oder das Ersetzen der Maschine durch einen leistungstärkeren und schnelleren Computer. Wenn Sie untersuchen möchten, ob die langen Antwortzeiten auf ein hohes Client-Datenverkehrsaufkommen zurückzuführen ist, überlegen Sie, ob dieser Zustand auftritt, wenn viel Client-Datenverkehr generiert wird. Dieselben Symptome können durch schlecht konfigurierte Systeme hervorgerufen werden, die Routen-

schleifen produzieren. Stellen Sie vor dem Ändern der Load-Balancer-Konfiguration fest, ob die Symptome durch eine hohe Client-Last verursacht werden.

Problem: Bei Verwendung der Weiterleitungsmethode mac registriert die Advisor-Funktion SSL oder HTTPS keine Serverlast

Bei Verwendung der Weiterleitungsmethode mac sendet Load Balancer Pakete an die Server und verwendet dabei die Clusteradresse, für die auf der Loopback-Einheit ein Aliasname definiert ist. Einige Serveranwendungen (z. B. SSL) erfordern, dass Konfigurationsdaten (wie Zertifikate) auf den IP-Adressen basieren. Die IP-Adresse muss die Clusteradresse sein, die an der Loopback-Adresse definiert ist. Nur in diesem Fall stimmt sie mit dem Inhalt eingehender Pakete überein. Wird beim Konfigurieren der Serveranwendung nicht die IP-Adresse des Clusters verwendet, kann die Client-Anforderung nicht ordnungsgemäß zum Server weitergeleitet werden.

Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung

Wenn Sie Load Balancer mit der fernen Webverwaltung konfigurieren, dürfen Sie nicht die Größe des Netscape-Browserfensters ändern, in dem die Load-Balancer-GUI angezeigt wird. Das heißt, Sie dürfen das Fenster nicht minimieren, maximieren, wiederherstellen usw. Da Netscape bei jeder Größenänderung des Browserfensters die Seite neu lädt, kommt es zu einer Trennung der Hostverbindung, die nach einer solchen Änderung demzufolge neu hergestellt werden muss. Wenn Sie die ferne Webverwaltung auf einer Windows-Plattform ausführen, verwenden Sie den Internet Explorer.

Problem: Bei aktiviertem Socket-Pooling wird der Webserver an 0.0.0.0 gebunden

Wenn Sie Microsoft IIS Version 5.0 auf Back-End-Servern mit Windows ausführen, müssen Sie Microsoft IIS bindungsspezifisch konfigurieren. Andernfalls ist das Socket-Pooling standardmäßig aktiviert, so dass der Webserver nicht an die virtuellen IP-Adressen gebunden wird, die als mehrere Identitäten der Site konfiguriert wurden, sondern an 0.0.0.0, und somit den gesamten Datenverkehr empfangen kann. Wenn eine Anwendung auf dem lokalen Host bei aktiviertem Socket-Pooling inaktiviert wird, können die Advisor-Funktionen dies auf AIX- oder Windows-ND-Servern erkennen. Wird jedoch eine Anwendung auf einem virtuellen Host inaktiviert, während der lokale Host aktiv bleibt, erkennen die Advisor-Funktionen den Ausfall der Anwendung nicht, so dass Microsoft IIS weiterhin auf den gesamten Datenverkehr reagiert, auch auf den der inaktivierten Anwendung.

Wenn Sie feststellen möchten, ob das Socket-Pooling aktiviert ist und der Webserver an 0.0.0.0 gebunden wird, setzen Sie den folgenden Befehl ab:

```
netstat -an
```

Anweisungen für das bindungsspezifische Konfigurieren von Microsoft IIS (Inaktivieren des Socket-Pooling) finden Sie auf der Microsoft-Website für Produktunterstützung. Bei Anzeige der folgenden Informationen können Sie auch den jeweils genannten URL aufrufen:

IIS5: Hardware Load Balance Does Not Detect a Stopped Web Site (Q300509)

<http://support.microsoft.com/default.aspx?scid=kb;en-us;Q300509>

Problem: Auf Windows-Systemen erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)

Unter Windows können im Fenster mit der Eingabeaufforderung einige nationale Sonderzeichen der Zeichensatzfamilie Latin-1 beschädigt angezeigt werden. Der Buchstabe "a" mit Tilde kann beispielsweise als Pi-Symbol erscheinen. Zum Korrigieren dieses Fehlers müssen Sie die Schriftartmerkmale für das Fenster mit der Eingabeaufforderung ändern. Gehen Sie zum Ändern der Schriftart wie folgt vor:

1. Klicken Sie oben links in der Ecke des Fensters mit der Eingabeaufforderung auf das Symbol.
2. Wählen Sie "Eigenschaften" aus und klicken Sie auf das Register "Schriftart".
3. Die Standardeinstellung für "Schriftart" ist "Rasterschriftarten". Setzen Sie die Schriftart auf "Lucida Console" und klicken Sie auf OK.

Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads

Einige Installationen von HP-UX 11i sind so vorkonfiguriert, dass nur 64 Threads pro Prozess zulässig sind. Manche Load-Balancer-Konfigurationen erfordern jedoch mehr Threads. Setzen Sie die Anzahl der Threads pro Prozess auf HP-UX-Systemen auf mindestens 256. Verwenden Sie zum Erhöhen dieses Wertes das Dienstprogramm "sam" und definieren Sie den Kernel-Parameter `max_thread_proc` neu. Bei einer erwarteten starken Auslastung, müssen Sie `max_thread_proc` möglicherweise auf einen noch höheren Wert als 256 setzen.

Gehen Sie zum Erhöhen von `max_thread_proc` wie folgt vor:

1. Geben Sie in der Befehlszeile Folgendes ein: `sam`
2. Wählen Sie **Kernel Configuration > Configurable Parameters** aus.
3. Wählen Sie in der Schiebeleiste **max_thread_proc** aus.
4. Drücken Sie die Leertaste, um **max_thread_proc** hervorzuheben.
5. Drücken Sie die Tabulatortaste. Drücken Sie dann so oft die Taste mit dem Rechtspfeil, bis Sie die Option **Actions** ausgewählt haben.
6. Drücken Sie die Eingabetaste, um das Menü **Actions** aufzurufen. Geben Sie dann **M** ein, um die Option "Modify Configurable Parameter" auszuwählen. (Falls diese Option nicht angezeigt wird, heben Sie **max_thread_proc** hervor.)
7. Drücken Sie die Tabulatortaste, bis Sie das Feld **Formula/Value** ausgewählt haben.
8. Geben Sie 256 oder einen höheren Wert ein.
9. Klicken Sie auf **OK**.
10. Drücken Sie die Tabulatortaste. Wählen Sie dann **Actions** aus.
11. Geben Sie **K** ein, um die Funktion "Process New Kernel" auszuwählen.
12. Wählen Sie **Yes** aus.
13. Führen Sie einen Warmstart des Systems durch.

Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv

Wenn Sie Ihren Adapter in einer Load-Balancer-Maschine konfigurieren, müssen Sie für eine fehlerfreie Ausführung der Advisor-Funktion sicherstellen, dass die beiden folgenden Einstellungen richtig definiert sind:

- Inaktivieren Sie das Feature Task Offload, das von den meisten 3Com-Adapterkarten verwendet wird.
 - Gehen Sie zum Inaktivieren von Task Offload wie folgt vor: Wählen Sie "Start" > "Einstellungen" > "Systemsteuerung" > "Netzwerk- und DFÜ-Verbindungen" aus. Wählen Sie dann den Adapter aus.
 - Klicken Sie im angezeigten Dialogfenster auf "Eigenschaften".
 - Klicken Sie auf "Konfigurieren" und wählen Sie das Register "Erweiterte Einstellungen" aus.
 - Wählen Sie im Teilfenster "Eigenschaft" das Merkmal "Task Offload" aus. Wählen Sie dann im Feld "Wert" den Eintrag "Disable" aus.
- Aktivieren Sie Protokoll 1 (ICMP) für IP-Protokolle, wenn Sie TCP/IP-Filter aktivieren. Wenn ICMP nicht aktiviert ist, scheitert der ping-Test zum Back-End-Server. Überprüfen Sie wie folgt, ob ICMP aktiviert ist:
 - Wählen Sie "Start" > "Einstellungen" > "Systemsteuerung" > "Netzwerk- und DFÜ-Verbindungen" aus. Wählen Sie dann den Adapter aus.
 - Klicken Sie im angezeigten Dialogfenster auf "Eigenschaften".
 - Wählen Sie im Komponententeilfenster "Internetprotokoll (TCP/IP)" aus und klicken Sie auf "Eigenschaften".
 - Klicken Sie auf "Erweitert" und wählen Sie das Register "Optionen" aus.
 - Wählen Sie im Teilfenster "Optionale Einstellungen" den Eintrag "TCP/IP-Filter" aus und klicken Sie auf "Eigenschaften".
 - Wenn Sie **TCP/IP-Filter aktivieren** ausgewählt und für IP-Protokolle die Option **Nur zulassen** markiert haben, müssen Sie das IP-Protokoll 1 hinzufügen. Dieses Protokoll muss zusätzlich zu den von Ihnen für TCP- und UDP-Ports aktivierten Protokollen hinzugefügt werden.

Problem: Auflösung von IP-Adressen in Hostnamen auf Windows-Systemen, wenn für einen Adapter mehrere Adressen konfiguriert sind

Wenn Sie auf der Windows-Plattform einen Adapter mit mehreren IP-Adressen konfigurieren, muss die IP-Adresse, die mit dem Hostnamen verbunden werden soll, als erste Adresse in der Registrierungsdatenbank konfiguriert werden.

Da Load Balancer in vielen Instanzen (z. B. lbkeys create) von InetAddress.getLocalHost() abhängig ist, können mehrere IP-Adressen, die per Aliasname an einen Adapter gebunden sind, zu Fehlern führen. Zur Vermeidung dieses Problems sollten die IP-Adressen, in die der Hostname aufgelöst werden soll, als erste Adressen in der Registrierungsdatenbank aufgelistet sein. Beispiel:

1. Starten Sie Regedit.
2. Modifizieren Sie wie folgt die folgenden Wertnamen:
 - HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet001 -> Services -> {IhreSchnittstellenadresse} -> Parameters -> Tcpip -> IPAddress
 - Versetzen Sie die IP-Adresse, in die der Hostname aufgelöst werden soll, an die erste Stelle.

- HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet001 -> Services -> Tcpip -> Parameters -> Interfaces -> {IhreSchnittstellenadresse} -> IPAddress
 - Versetzen Sie die IP-Adresse, in die der Hostname aufgelöst werden soll, an die erste Stelle.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet002 -> Services -> {IhreSchnittstellenadresse} -> Parameters -> Tcpip -> IPAddress
 - Versetzen Sie die IP-Adresse, in die der Hostname aufgelöst werden soll, an die erste Stelle.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> ControlSet002 -> Services -> Tcpip -> Parameters -> Interfaces -> {IhreSchnittstellenadresse} -> IPAddress
 - Versetzen Sie die IP-Adresse, in die der Hostname aufgelöst werden soll, an die erste Stelle.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> CurrentControlSet -> Services -> {IhreSchnittstellenadresse} -> Parameters -> Tcpip -> IPAddress
 - Versetzen Sie die IP-Adresse, in die der Hostname aufgelöst werden soll, an die erste Stelle.
 - HKEY_LOCAL_MACHINE -> SYSTEM -> CurrentControlSet -> Services -> Tcpip -> Parameters -> Interfaces -> {IhreSchnittstellenadresse} -> IPAddress
 - Versetzen Sie die IP-Adresse, in die der Hostname aufgelöst werden soll, an die erste Stelle.
3. Führen Sie einen Warmstart durch.
 4. Überprüfen Sie, ob der Hostname in die richtige IP-Adresse aufgelöst wird. Verwenden Sie zur Prüfung beispielsweise ping *IhrHostname*.

Problem: Advisor-Funktionen können auf Windows-Systemen in einer Konfiguration für hohe Verfügbarkeit nach einem Netzwerkausfall nicht ausgeführt werden

Wenn das Betriebssystem Windows einen Netzwerkausfall erkennt, löscht es standardmäßig seinen ARP-Cache (Address Resolution Protocol) einschließlich aller statischen Einträge. Ist das Netzwerk wieder verfügbar, wird der ARP-Cache durch die im Netzwerk gesendeten ARP-Anforderungen wieder gefüllt.

In einer Konfiguration für hohe Verfügbarkeit übernehmen beide Server primäre Operationen, wenn einer der Server oder beide von einem Verlust der Netzkonnektivität betroffen ist/sind. Wird dann die ARP-Anforderung zur erneuten Füllung des ARP-Cache gesendet, antworten beide Server. Deshalb markiert der ARP-Cache den Eintrag als ungültig. Die Advisor-Funktionen können aus diesem Grund keinen Socket zu den Ausweichservern erstellen.

Dieses Problem lässt sich dadurch lösen, dass das Betriebssystem Windows daran gehindert wird, den ARP-Cache bei einem Konnektivitätsverlust zu löschen. Microsoft hat einen Artikel veröffentlicht, in dem die Ausführung dieser Task erklärt wird. Sie finden diesen Artikel auf der Website von Microsoft unter "Knowledge Base". Der Artikel hat die Nummer 239924 und die Webadresse <http://support.microsoft.com/default.aspx?scid=kb;de;239924>.

Hier finden Sie eine Zusammenfassung der im Artikel von Microsoft beschriebenen Schritte, die das System daran hindern sollen, den ARP-Cache zu löschen:

1. Öffnen Sie die Registrierungsdatenbank mit dem Registrierungseditor (regedit oder regedit32).
2. Sehen Sie sich den folgenden Schlüssel in der Registrierungsdatenbank an:

HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters

3. Fügen Sie den folgenden Wert zur Registrierungsdatenbank hinzu: Name: DisableDHCPMediaSense Typ: REG_DWORD.
4. Setzen Sie den Wert des hinzugefügten Schlüssels auf 1.
5. Führen Sie einen Warmstart der Maschine durch, um die Änderung in Kraft zu setzen.

Anmerkung: Diese Änderung wirkt sich unabhängig von der DHCP-Einstellung auf den ARP-Cache aus.

Problem: Beim Festlegen von Aliasnamen für mehrere Cluster auf der Loopback-Einheit von Linux-Systemen nicht den Befehl 'IP address add' verwenden

Bei Verwendung von Servern mit dem Linux-Kernel 2.4.x und der Dispatcher-Weiterleitungsmethode mac sind mehrere Punkte zu beachten. Wenn der Server mit dem Befehl **ip address add** eine Clusteradresse auf der Loopback-Einheit konfiguriert hat, kann nur für eine Clusteradresse ein Aliasname festgelegt werden.

Wenn Sie auf der Loopback-Einheit Aliasnamen für mehrere Cluster festlegen möchten, verwenden den Befehl **ifconfig**. Beispiel:

```
ifconfig lo:Num Clusteradresse netmask 255.255.255.255 up
```

Zusätzlich ist zu beachten, dass die Konfigurationsmethode ifconfig für Schnittstellen nicht vollständig mit der Konfigurationsmethode ip für Schnittstellen kompatibel ist. Deshalb sollte für einen Standort eine Methode ausgewählt und dann ausschließlich angewendet werden.

Problem: Fehlermeldung 'Router-Adresse nicht angegeben oder nicht gültig für Port-Methode' erscheint

Wenn Sie Server zu Ihrer Dispatcher-Konfiguration hinzufügen, kann die Fehlermeldung "Fehler: Router-Adresse nicht angegeben oder nicht gültig für Port-Methode" angezeigt werden.

Bestimmen Sie den Fehler anhand der Prüfliste.

- Stellen Sie sicher, dass die letzte Wartungsstufe angewendet wurde.
- Vergewissern Sie sich, dass Sie (mit Ausnahme von Solaris-Plattformen) eine IBM Version von Java verwenden.
- Stellen Sie sicher, dass Sie auf Windows-Systemen nicht die Verwendung von DHCP konfiguriert haben.
- Wenn die Weiterleitungsmethode mac ist (Standardeinstellung), müssen sich der Server, der Cluster und mindestens eine unterstützte NIC innerhalb eines Teilnetzes befinden. Sie können beispielsweise nicht den Cluster 10.1.1.1 und den Server 130.2.3.4 definieren, weil beide nicht in demselben Teilnetz enthalten sind.

Anmerkung: Ist die Weiterleitungsmethode nat oder cbr, müssen sich die Server nicht in demselben Teilnetz wie der Cluster befinden.

- Wenn alle in demselben Teilnetz enthalten sind und Sie für den Cluster einen Aliasnamen festgelegt haben, müssen Sie den Clusteraliasnamen auf einer NIC definieren, die Datenverkehr in dieses Teilnetz weiterleitet. Wenn en0 für 13.2.3.4 definiert ist und en1 für 9.1.2.3 und die Clusterdefinition 9.5.7.3 lautet, müssen Sie den Cluster beispielsweise auf en1 konfigurieren. Die Standardschnittstelle ist en0.

- Vergewissern Sie sich auf Linux-Plattformen, dass der richtige Kernel geladen wurde. Schauen Sie sich dazu die Datei `loadoutput.log` im Verzeichnis `/usr/lpp/ibm/internet/nd/logs/dispatcher` an. Prüfen Sie, ob die Datei Fehlernachrichten enthält.

Der Standardwert für den Parameter `router` ist 0. Dieser Wert gibt an, dass der Server lokal ist. Wenn Sie die Router-Adresse des Servers auf einen Wert ungleich 0 setzen, gibt dieser Wert an, dass es sich um einen fernen Server in einem anderen Teilnetz handelt. Weitere Informationen zum Parameter `router` des Befehls `server add` finden Sie im Abschnitt „dscontrol server — Server konfigurieren“ auf Seite 394.

Befindet sich der Server, den Sie hinzufügen, in einem anderen Teilnetz, sollte der Parameter `router` auf die Adresse des Routers gesetzt werden, der im lokalen Teilnetz für die Kommunikation mit dem fernen Server verwendet wird.

Problem: Auf Solaris-Systemen werden Load-Balancer-Prozesse beim Verlassen des Terminalfensters, in dem die Prozesse gestartet wurden, beendet

Wenn Sie auf Solaris-Systemen Load-Balancer-Skripts (wie `dsserver` oder `lbadmin`) in einem Terminalfenster starten, endet der Load-Balancer-Prozess beim Verlassen dieses Fensters.

Sie können dieses Problem lösen, indem Sie die Load-Balancer-Skripts mit dem Befehl **nohup** starten. Beispiel: **nohup dsserver**. Dieser Befehl verhindert, dass die in der Terminalsitzung gestarteten Prozesse beim Beenden des Terminals ein Stoppsignal empfangen, so dass die Prozesse nach Beendigung der Terminalsitzung weiter ausgeführt werden können. Geben Sie den Befehl **nohup** vor allen Load-Balancer-Skripts an, die nach dem Ende einer Terminalsitzung weiter verarbeitet werden sollen.

Problem: Verzögerung beim Laden einer Load-Balancer-Konfiguration

Das Laden einer Load-Balancer-Konfiguration kann aufgrund von DNS-Aufrufen (Domain Name System), die abgesetzt werden, um die Serveradresse aufzulösen und zu prüfen, sehr lange dauern.

Falls der DNS der Load-Balancer-Maschine falsch konfiguriert ist oder DNS generell viel Zeit benötigt, verlangsamt sich das Laden der Konfiguration, weil Java DNS-Anforderungen im Netz sendet.

Sie können dieses Problem umgehen, indem Sie die Serveradressen und Hostnamen zu Ihrer lokalen Datei `/etc/hosts` hinzufügen.

Problem: Fehlernachricht zu einem IP-Adressenkonflikt auf Windows-Systemen

Wenn die hohe Verfügbarkeit konfiguriert ist, können für einen kurzen Zeitraum Clusteradressen auf beiden Maschinen konfiguriert sein. Dies führt zu der Fehlermeldung, dass ein IP-Adressenkonflikt mit einem anderem System im Netz vorliegt. Sie können diese Nachricht ignorieren. Es ist möglich, dass eine Clusteradresse für kurze Zeit gleichzeitig auf beiden Maschinen mit hoher Verfügbarkeit konfiguriert ist. Dies gilt insbesondere während des Starts einer der Maschinen oder bei Einleitung einer Aufgabenübernahme.

Überprüfen Sie, ob die Clusteradressen in den go*-Scripts korrekt konfiguriert bzw. dekonfiguriert sind. Wenn Sie eine Konfigurationsdatei aufgerufen und go*-Scripts installiert haben, vergewissern Sie sich, dass Ihre Konfigurationsdatei für Clusteradressen keinen Befehl "executor configure" enthält, da dies zu Konflikten mit den Befehlen configure und unconfigure in den go*-Scripts führt.

Weitere Informationen zu go*-Scripts beim Konfigurieren der hohen Verfügbarkeit finden Sie im Abschnitt „Scripts verwenden“ auf Seite 209.

Problem: In einer Konfiguration mit hoher Verfügbarkeit sind sowohl primäre Maschinen als auch Ausweichmaschinen aktiv

Dieses Problem kann auftreten, wenn die go-Scripts weder auf der primären Maschine noch auf der Ausweichmaschine ausgeführt werden. Die go-Scripts können nicht ausgeführt werden, wenn dsserver nicht auf beiden Maschinen gestartet wurde. Überprüfen Sie beide Maschinen und vergewissern Sie sich, dass dsserver aktiv ist.

Problem: Client-Anforderungen scheitern, wenn die Rückgabe von Antworten in Form großer Seiten versucht wird

Client-Anforderungen, die zu Antworten in Form großer Seiten führen, überschreiten das Zeitlimit, wenn die maximale Übertragungseinheit (MTU, Maximum Transmit Unit) auf der Dispatcher-Maschine nicht ordnungsgemäß gesetzt ist. Dieser Fehler kann bei den Weiterleitungsmethoden cbr und nat der Dispatcher-Komponente auftreten, wenn der Dispatcher den MTU-Wert nicht aushandelt, sondern den Standardwert für die MTU verwendet.

Die MTU wird unter jedem Betriebssystem ausgehend vom Typ des Übertragungsmediums (z. B. Ethernet oder Token-Ring) festgelegt. Für Router des lokalen Segments kann eine kleinere MTU definiert sein, wenn sie eine Verbindung zu einem Übertragungsmedium eines anderen Typs herstellen. Bei normalem TCP-Datenverkehr wird die MTU beim Aufbau der Verbindung ausgehandelt. Für die Übertragung der Daten zwischen den Maschinen wird dann die kleinste MTU verwendet.

Der Dispatcher unterstützt das Aushandeln der MTU nicht für die Weiterleitungsmethoden cbr und nat, weil er an TCP-Verbindungen aktiv als Endpunkt beteiligt ist. Für die Weiterleitungsmethoden cbr und nat verwendet der Dispatcher standardmäßig den MTU-Wert 1500. Dieser Wert ist die typische MTU-Größe für Standard-Ethernet und muss daher von den meisten Kunden nicht angepasst werden.

Wenn Sie mit der Dispatcher-Weiterleitungsmethode cbr oder nat arbeiten und zum lokalen Segment einen Router mit einer kleineren MTU verwenden, müssen Sie die MTU auf der Dispatcher-Maschine auf diesen niedrigeren MTU-Wert setzen.

Verwenden Sie zur Lösung dieses Problems den folgenden Befehl, um die maximale Segmentgröße (mss) zu setzen: `dscontrol executor set mss neuer_Wert`

Beispiel:

```
dscontrol executor set mss 1400
```

Der Standardwert für mss ist 1460.

Die Einstellung `mss` gilt nicht für die Dispatcher-Weiterleitungsmethode `mac` oder für eine andere Komponente von Load Balancer. Sie wird nur auf die Komponente Dispatcher angewendet.

Auf Windows-Systemen wird beim Absetzen eines Befehls `dscontrol` oder `lbadm` der Fehler 'Der Server antwortet nicht' angezeigt

Wenn es auf einem Windows-System mehrere IP-Adressen gibt und die Datei `hosts` nicht angibt, welche Adresse dem Hostnamen zugeordnet werden soll, wählt das Betriebssystem die kleinste Adresse aus und ordnet sie dem Hostnamen zu.

Sie können dieses Problem lösen, indem Sie die Datei `c:\Windows\system32\drivers\etc\hosts` aktualisieren. Geben Sie den Hostnamen Ihrer Maschine an und die IP-Adresse, die Sie dem Hostnamen zuordnen möchten.

Wichtiger Hinweis: Die IP-Adresse darf keine Clusteradresse sein.

Problem: Dispatcher-Maschinen mit hoher Verfügbarkeit können unter Linux auf S/390 mit `qeth`-Treibern möglicherweise nicht synchronisiert werden

Wenn Sie unter Linux für S/390 mit dem `qeth`-Netztreiber die hohe Verfügbarkeit verwenden, können der aktive Dispatcher und der Bereitschafts-Dispatcher unter Umständen nicht synchronisiert werden. Dieses Problem könnte auf den Linux-Kernel 2.6 beschränkt sein.

Sollte dieses Problem auftreten, können Sie wie folgt vorgehen:

Definieren Sie zwischen den Images des aktiven Dispatchers und des Bereitschafts-Dispatchers ein CTC-Netz (Channel-to-Channel) und fügen Sie zwischen den IP-Adressen der beiden CTC-Endpunkte ein Überwachungssignal hinzu.

Problem: Konfigurationstipps für hohe Verfügbarkeit

Wenn Sie die Load-Balancer-Funktion für hohe Verfügbarkeit verwenden, kann eine Partnermaschine den Lastausgleich übernehmen, falls der primäre Partner ausfällt oder heruntergefahren wird. Für die Aufrechterhaltung der Verbindungen zwischen den Partnern mit hoher Verfügbarkeit werden zwischen den beiden Maschinen Verbindungssätze übergeben. Wenn der Ausweichpartner die Lastausgleichsfunktion übernimmt, wird die Cluster-IP-Adresse von der Ausweichmaschine entfernt und zur neuen primären Maschine hinzugefügt. Diese Übernahmeoperation kann durch diverse Aspekte der Ablaufsteuerung und der Konfiguration beeinflusst werden.

Mit den in diesem Abschnitt aufgeführten Tipps können Sie die Auswirkungen von Problemen mindern, die sich aus der Konfiguration der hohen Verfügbarkeit ergeben. Es kann sich unter anderem um folgende Probleme handeln:

- Übergangene Verbindungen nach der Übernahme
- keine Synchronisation mit Partnermaschinen
- versehentliche Übertragung von Anforderungen an die Ausweichpartnermaschine

Die folgenden Tipps sind hilfreich für eine erfolgreiche Konfiguration der hohen Verfügbarkeit auf Ihren Load-Balancer-Maschinen.

- Die Position der Befehle für hohe Verfügbarkeit in Ihren Script-Dateien kann den entscheidenden Unterschied ausmachen.

Beispielbefehle für hohe Verfügbarkeit:

```
dscontrol highavailability heartbeat add ...
dscontrol highavailability backup add ...
dscontrol highavailability reach add ...
```

In den meisten Fällen müssen Sie die Definitionen für hohe Verfügbarkeit an das Ende der Datei stellen. Die Anweisungen cluster, port und server müssen vor den highavailability-Anweisungen stehen. Dies ist notwendig, weil die hohe Verfügbarkeit nach den Cluster-, Port- und Serverdefinitionen sucht, wenn bei der Synchronisation ein Verbindungssatz empfangen wird.

Sind Cluster, Port und Server nicht vorhanden, wird der Verbindungssatz gelöscht. Wenn eine Übernahme stattfinden soll und der Verbindungssatz nicht auf der Partnermaschine repliziert wurde, scheitert die Verbindung.

Eine Ausnahme von dieser Regel bilden verknüpfte Server, die mit der Weiterleitungsmethode mac konfiguriert sind. Bei diesen Servern müssen die highavailability-Anweisungen vor den Anweisungen für verknüpfte Server stehen. Falls die highavailability-Anweisungen nicht vor den Anweisungen für verknüpfte Server stehen, kann Load Balancer eine Anforderung an den verknüpften Server empfangen, interpretiert diese jedoch wie eine ankommende Anforderung für den Cluster und unterzieht sie dem Lastausgleich. Auf diese Weise können die Pakete in einer Schleife im Netz gesendet werden und unnötigen Datenverkehr verursachen. Wenn die highavailability-Anweisungen vor dem verknüpften Server angegeben sind, weiß Load Balancer, dass ankommender Datenverkehr nur weitergeleitet werden muss, wenn er den Status AKTIV hat.

- Unter den Betriebssystemen z/OS und OS/390 steuert der Hypervisor die Schnittstelle und multiplext die reale Schnittstelle für die Gastsysteme. Der Hypervisor erlaubt immer nur jeweils einer Gastmaschine, sich für eine IP-Adresse zu registrieren, und zeigt ein Updatefenster an. Wenn die Cluster-IP-Adresse von der Ausweichmaschine entfernt wird, müssen Sie demzufolge unter Umständen für das Hinzufügen der Cluster-IP-Adresse zur Primärmaschine eine Verzögerung einrichten. Andernfalls scheitert das Hinzufügen der IP-Adresse und ankommende Verbindungen werden nicht verarbeitet.

Korrigieren Sie dieses Verhalten, indem Sie zum goActive-Script eine sleep-Verzögerung hinzufügen. Die jeweilige Verzögerungszeit ist vom Deployment abhängig. Beginnen Sie zunächst mit einer Verzögerungszeit von 10.

- Partner mit hoher Verfügbarkeit müssen sich in demselben Teilnetz befinden und sich gegenseitig ping-Signale senden können.

Die Maschinen versuchen standardmäßig, jede halbe Sekunde miteinander zu kommunizieren. Nach vier fehlgeschlagenen Versuchen wird ein Ausfall festgestellt. Bei einer ausgelasteten Maschine kann dies eine Funktionsübernahme zur Folge haben, obwohl das System noch ordnungsgemäß arbeitet. Setzen Sie den folgenden Befehl ab, um die Anzahl der Versuche bis zur Feststellung des Ausfalls zu erhöhen:

```
dscontrol executor set hatimeout <Wert>
```

- Während der Synchronisation der Partner werden alle Verbindungssätze von der aktiven Maschine an die Ausweichmaschine gesendet. Die Synchronisation muss innerhalb des Standardzeitlimits von 50 Sekunden abgeschlossen sein.

Alte Verbindungen dürfen daher nicht längere Zeit im Speicher verbleiben. Probleme hat es insbesondere mit LDAP-Ports und großen staletimeout-Zeiträumen (länger als ein Tag) gegeben. Wird ein langer staletimeout-Zeitraum festge-

legt, bleiben alte Verbindungen im Speicher. Dies hat zur Folge, dass bei der Synchronisation mehr Verbindungssätze übergeben werden und die Speicherlastung auf beiden Maschinen steigt.

Falls die Synchronisation bei einem vernünftigen `staletimeout`-Zeitraum scheitert, können Sie das Synchronisationszeitlimit mit folgendem Befehl erhöhen:

```
e xm 33 5 neues_Zeitlimit
```

Dieser Befehl wird nicht gespeichert, wenn Sie die Konfigurationsdatei speichern. Falls Sie diese Einstellung weiter verwenden möchten, müssen Sie sie nach jedem Systemabschluss manuell zur Konfigurationsdatei hinzufügen.

Das Zeitlimit wird in halben Sekunden gespeichert. Der Standardwert für `neues_Zeitlimit` liegt daher bei 100 (50 Sekunden).

- Wenn eine Partnermaschine die Auslastung übernimmt, setzt sie unaufgefordert eine ARP-Antwort ab (gratuitous ARP), um den Maschinen im selben Teilnetz die neue Hardwareadresse mitzuteilen, die der Cluster-IP-Adresse zugeordnet ist. Sie müssen sicherstellen, dass Ihre Router unaufgeforderte ARP-Nachrichten berücksichtigen und ihren Cache aktualisieren. Andernfalls werden die Anforderungen an den inaktiven Partner gesendet.

Anmerkung: Weitere Informationen zum Konfigurieren der hohen Verfügbarkeit finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 204.

Problem: Einschränkungen für die Dispatcher-Konfiguration unter Linux auf zSeries- oder S/390-Servern mit OSA-Karten

Wenn Sie die Weiterleitungsmethode `mac` verwenden, müssen sich die Server der Load-Balancer-Konfiguration generell und unabhängig von der Plattform in einem Netzsegment befinden. Aktive Netzeinheiten wie Router, Brücken und Firewalls behindern Load Balancer. Dies liegt daran, dass Load Balancer wie ein spezialisierter Router funktioniert, der nur die Header der Verbindungsschicht für seinen nächsten und letzten Hop modifiziert. Jede Netztopologie, bei der der nächste Hop nicht der letzte Hop ist, ist für Load Balancer nicht gültig.

Anmerkung: Tunnel wie CTC (Channel-To-Channel) oder IUCV (Inter-User Communication Vehicle) werden oft unterstützt. Die Weiterleitung von Load Balancer muss jedoch durch den Tunnel direkt zum Zielort erfolgen. Es darf sich nicht um einen Tunnel zwischen Netzen handeln.

Für zSeries- und S/390-Server, die die OSA-Karte gemeinsam nutzen, gibt es eine Einschränkung, weil dieser Adapter anders als die meisten Netzkarten funktioniert. Die OSA-Karte hat eine eigene virtuelle Implementierung der Verbindungsschicht, die nichts mit dem Ethernet zu tun hat, das sich den Linux- und z/OS-Hosts präsentiert. Die Verbindung über eine OSA-Karte sieht tatsächlich wie eine Verbindung zwischen Ethernet-Hosts (und nicht wie eine Verbindung zu OSA-Hosts) aus, und die Hosts, die eine solche Karte verwenden, behandeln sie wie eine Ethernet-Karte.

Die OSA-Karte führt auch einige Funktionen mit direktem Bezug zur IP-Schicht aus. Eine ihrer Funktionen ist beispielsweise die Beantwortung von ARP-Anforderungen (Address Resolution Protocol). Außerdem leitet die gemeinsam genutzte OSA-Karte IP-Pakete ausgehend von der Ziel-IP-Adresse weiter und nicht wie ein Switch der Schicht 2 ausgehend von der Ethernet-Adresse. Eigentlich ist die OSA-Karte ein Netzsegment mit einer Brücke zu sich selbst.

Wird Load Balancer auf einem S/390- oder zSeries-Host mit Linux ausgeführt, können Pakete an Hosts mit derselben OSA-Karte oder an Hosts im Ethernet weitergeleitet werden. Alle Hosts, die dieselbe OSA-Karte gemeinsam benutzen, befinden sich effektiv im selben Segment.

Die Natur der OSA-Brücke versetzt Load Balancer in die Lage, Pakete aus dem Bereich der gemeinsam benutzten OSA-Karte *hinaus* weiterzuleiten. Der Brücke ist der OSA-Port bekannt, der Eigner der Cluster-IP-Adresse ist. Außerdem kennt die Brücke die MAC-Adresse von Hosts, die direkt mit dem Ethernet-Segment verbunden sind. Deshalb kann Load Balancer über eine OSA-Brücke Datenpakete mit der Weiterleitungsmethode `mac` weiterleiten.

In den Bereich einer gemeinsam genutzten OSA-Karte hinein kann Load Balancer jedoch keine Daten weiterleiten. Dies gilt auch für Load Balancer auf einem S/390 mit Linux, wenn der Back-End-Server eine andere OSA-Karte als Load Balancer verwendet. Die OSA-Karte für den Back-End-Server macht die OSA-MAC-Adresse für die Server-IP-Adresse zugänglich. Wenn jedoch ein Paket mit der Ethernet-Zieladresse der Server-OSA-Karte und der IP-Adresse des Clusters ankommt, weiß die OSA-Karte des Servers nicht, welche Hosts das Paket ggf. empfangen sollten. Die Prinzipien, die zwischen OSA und Ethernet eine Weiterleitung mit der Methode `mac` bei einer gemeinsam benutzten OSA-Karte möglich machen, gelten nicht, wenn versucht wird, Daten in den Bereich einer gemeinsam genutzten OSA-Karte hinein weiterzuleiten.

Problemumgehung:

In Load-Balancer-Konfigurationen mit zSeries- oder S/390-Servern, die OSA-Karten enthalten, gibt es zwei Möglichkeiten, das oben beschriebene Problem zu umgehen.

1. Verwendung von Plattformfeatures

Falls die Server der Load-Balancer-Konfiguration denselben zSeries- oder S/390-Plattformtyp haben, können Sie zwischen Load Balancer und den einzelnen Servern Punkt-zu-Punkt-Verbindungen (CTC oder IUCV) definieren. Konfigurieren Sie die Endpunkte mit privaten IP-Adressen. Die Punkt-zu-Punkt-Verbindung wird nur für den Datenverkehr zwischen Load Balancer und dem jeweiligen Server verwendet. Fügen Sie dann die Server mit der IP-Adresse des Serverendpunkts für den Tunnel hinzu. Bei dieser Konfiguration fließt der Clusterdatenverkehr über die OSA-Karte von Load Balancer und wird über die Punkt-zu-Punkt-Verbindung weitergeleitet. Der Server antwortet dann über seine eigene Standardroute. Die Antwort wird über die OSA-Karte des Servers abgesetzt. Dabei muss es sich nicht um dieselbe Karte handeln.

2. Verwendung des GRE-Features von Load Balancer

Anmerkung: Anmerkung: Das GRE-Feature ist nicht in der Doppelprotokollumgebung von Load Balancer für IPV4 und IPV6 verfügbar.

Falls die Server der Load-Balancer-Konfiguration nicht denselben zSeries- oder S/390-Plattformtyp haben oder keine Punkt-zu-Punkt-Verbindung zwischen Load Balancer und den einzelnen Servern definiert werden kann, sollten Sie das Load-Balancer-Feature Generic Routing Encapsulation (GRE) verwenden. GRE ist ein Protokoll, mit dem Load Balancer Daten über Router weiterleiten kann.

Wenn Sie GRE verwenden, wird das vom Client an die Cluster-IP-Adresse von Load Balancer gekapselt empfangen und an den Server gesendet. Auf dem Server wird das ursprüngliche Paket von Client an die Cluster-IP-Adresse entkapselt, und der Server antwortet dem Client direkt. Der Vorteil der Verwendung von GRE besteht darin, dass Load Balancer nur den Datenverkehr vom Client

zum Server sieht, nicht aber den Datenverkehr vom Server zum Client. Der Nachteil ist, dass GRE durch den Kapselungsaufwand die maximale Segmentgröße der TCP-Verbindung herabsetzt.

Wenn Sie die Load-Balancer-Weiterleitung mit GRE konfigurieren möchten, fügen Sie die Server mit dem folgenden Befehl hinzu:

```
dscontrol server add Clusteradresse:Port:Back-End-Server router Back-End-Server
```

Die Angabe "router Back-End-Server" ist gültig, wenn sich Load Balancer und der Back-End-Server in demselben IP-Teilnetz befinden. Geben Sie andernfalls die gültige IP-Adresse des nächsten Hop als Router an.

Wenn Sie Linux-Systeme für natives GRE konfigurieren möchten, setzen Sie für jeden Back-End-Server die folgenden Befehle ab:

```
modprobe ip_gre
ip tunnel add gre1b0 mode gre ikey 3735928559
ip link set gre1b0 up
ip addr add Clusteradresse dev gre1b0
```

Anmerkung: Definieren Sie auf der Loopback-Einheit der Back-End-Server nicht die Clusteradresse. Wenn Sie z/OS-Back-End-Server verwenden, müssen Sie die Server mit z/OS-Befehlen für GRE konfigurieren.

Problem: Speicherverlust unter einigen Linux-Versionen bei Ausführung von Dispatcher mit Manager und Advisor-Funktionen

Wenn Sie Load Balancer konfiguriert mit Manager und Advisor-Funktionen ausführen, kann es bei einigen Red-Hat-Linux-Versionen zu großen Speicherverlusten kommen. Der Java-Speicherverlust nimmt noch zu, wenn Sie für die Advisor-Funktion ein kleines Zeitintervall konfigurieren.

Die JVM-Versionen des IBM Java SDK und die im Lieferumfang einiger Linux-Distributionen (z. B. Red Hat Enterprise Linux 3.0) enthaltene NPTL (Native POSIX Thread Library) können den Speicherverlust verursachen.

Die aktuellsten Informationen zu Linux-Systemen und zu dem mitgelieferten IBM Java SDK finden Sie unter <http://www.ibm.com/developerworks/java/jdk/linux/tested.html>.

Zur Fehlerbestimmung können Sie den Befehl `vmstat` oder `ps` verwenden, mit dem Speicherverluste festgestellt werden können.

Zum Korrigieren des Speicherverlustes müssen Sie den folgenden Befehl zum Inaktivieren der NPTL absetzen, bevor die Load-Balancer-Maschine aktiv ist:

```
export LD_ASSUME_KERNEL=2.4.10
```

Problem: Unter SUSE Linux Enterprise Server 9 leitet Dispatcher Pakete weiter, die den Back-End-Server jedoch nicht erreichen

Wenn Sie unter Suse Linux Enterprise Server 9 die Weiterleitungsmethode `mac` verwenden, kann der Dispatcher-Bericht angeben, dass das Paket weitergeleitet wurde (steigende Paketanzahl), obwohl das Paket den Back-End-Server nie erreicht.

Wenn dieses Problem auftritt, lässt sich unter Umständen Folgendes beobachten:

- Auf der Dispatcher-Maschine wird die folgende Nachricht angezeigt:
ip_finish_output2: No header cache and no neighbour!
- Auf der Client-Maschine wird die folgende Nachricht angezeigt:
ICMP Destination unreachable: Fragmentation Needed

Dieses Problem kann wegen des geladenen NAT-Moduls von iptables auftreten. Unter SLES 9 gibt es in dieser Version von iptables möglicherweise einen (noch nicht nachgewiesenen) Fehler, der bei der Interaktion mit Dispatcher ein seltsames Verhalten auslöst.

Lösung:

Entladen Sie das NAT-Modul und das Connection-Tracking-Modul von iptables.

Beispiel:

```
# lsmod | grep ip
iptables_filter      3072  0
iptables_nat         22060  0
ip_conntrack         32560  1 iptable_nat
ip_tables            17280  2
iptables_filter,iptable_nat
ip6                  236800  19
# rmmod iptable_nat
# rmmod ip_conntrack
```

Entfernen Sie die Module in der Reihenfolge, in der sie verwendet werden. Ein Modul kann nur dann entfernt werden, wenn die Referenzanzahl (letzte Spalte der lsmod-Ausgabe) null ist. Falls Sie in iptables Regeln konfiguriert haben, müssen Sie diese entfernen. Beispiel: iptables -t nat -F.

Das Modul iptable_nat verwendet ip_conntrack. Sie müssen demzufolge erst das Modul iptable_nat und dann das Modul ip_conntrack entfernen.

Anmerkung: Sobald Sie versuchen, die für eine Tabelle konfigurierten Regeln aufzulisten, wird das entsprechende Modul geladen. Beispiel: iptables -t nat -L. Stellen Sie sicher, dass solche Befehle nicht verwendet werden, nachdem Sie die Module entfernt haben.

Problem: Fehlermeldung zu einem IP-Adressenkonflikt auf Windows-Systemen während der Funktionsübernahme bei hoher Verfügbarkeit

Wenn Sie auf Windows-Systemen das Load-Balancer-Feature für hohe Verfügbarkeit ausführen, werden goScripts verwendet, um bei einer Funktionsübernahme die Cluster-IP-Adresse des aktiven Load Balancer zu konfigurieren und die Cluster-IP-Adresse des Ausweichsystems aus der Konfiguration zu entfernen. Wenn das goScript, das die Cluster-IP-Adresse auf der aktiven Maschine konfiguriert, vor dem goScript ausgeführt wird, das die Cluster-IP-Adresse der Ausweichmaschine aus der Konfiguration entfernt, kann es zu Problemen kommen. Unter Umständen erscheint ein Dialogfenster mit einer Nachricht, dass das System einen IP-Adressenkonflikt festgestellt hat. Wenn Sie den Befehl ipconfig /all ausführen, kann auch eine IP-Adresse 0.0.0.0 auf der Maschine angezeigt werden.

Lösung:

Setzen Sie den folgenden Befehl ab, um die Cluster-IP-Adresse der primären Maschine manuell aus der Konfiguration zu entfernen:

```
dscontrol executor unconfigure Cluster-IP-Adresse
```

Dieser Befehl entfernt die Adresse 0.0.0.0 aus dem Windows-IP-Stack.

Nachdem der Partner für hohe Verfügbarkeit die Cluster-IP-Adresse abgegeben hat, setzen Sie den folgenden Befehl ab, um die Cluster-IP-Adresse erneut hinzuzufügen:

```
dscontrol executor configure Cluster-IP-Adresse
```

Überprüfen Sie nach dem Absetzen dieses Befehls erneut die Cluster-IP-Adressen im Windows-IP-Stack. Setzen Sie dazu folgenden Befehl ab:

```
ipconfig /all
```

Problem: Mögliche Störung der Paketweiterleitung durch iptables unter Linux

Unter Linux kann iptables den Lastausgleich für den Datenverkehr stören und muss auf der Dispatcher-Maschine inaktiviert werden.

Setzen Sie den folgenden Befehl ab, um festzustellen, ob iptables geladen werden:

```
lsmod | grep ip_tables
```

Die Ausgabe dieses Befehls kann so ähnlich wie die folgende aussehen:

```
ip_tables          22400    3
iptables_mangle,iptable_nat,iptable_filter
```

Setzen Sie für jede in der Ausgabe aufgelistete iptable den folgenden Befehl ab, um die Regeln für die Tabellen anzuzeigen:

```
iptables -t <Kurzname> -L
```

Beispiel:

```
iptables -t mangle -L
iptables -t nat -L
iptables -t filter -L
```

Wenn die Tabelle iptable_nat geladen ist, muss sie entladen werden. Da iptable_nat von iptable_conntrack abhängig ist, muss iptable_conntrack ebenfalls entfernt werden. Setzen Sie zum Entladen dieser beiden iptables den folgenden Befehl ab:

```
rmmod iptable_nat iptable_conntrack
```

Problem: Auf Solaris-Systemen kann kein IPV6-Server zur Load-Balancer-Konfiguration hinzugefügt werden

Wenn Sie auf Solaris-Systemen versuchen, einen IPV6-Server in einer Installation von Load Balancer für IPV4 und IPV6 zu konfigurieren, erscheint die Nachricht Server kann nicht hinzugefügt werden. Dieser Fehler kann durch die Art und Weise ausgelöst werden, mit der das Betriebssystem Solaris die ping-Anforderung für eine IPV6-Adresse handhabt. Wenn Sie auf Solaris-Systemen einen Server zur Konfiguration hinzufügen, versucht Load Balancer, den Server mit ping zu erreichen, um die MAC-Adresse des Servers abzurufen. Die Solaris-Maschine könnte eine konfigurierte Clusteradresse als Quellenadresse der ping-Anfrage auswählen,

anstatt die NFA der Maschine zu verwenden. Falls die Clusteradresse auf der Loopback-Einheit des Servers konfiguriert ist, wird die ping-Antwort nicht von der Load-Balancer-Maschine empfangen, so dass der Server nicht zur Konfiguration hinzugefügt wird.

Sie müssen daher auf der Load-Balancer-Maschine eine andere IPV6-Adresse konfigurieren, bevor oder nachdem Sie die IPV6-Clusteradresse konfigurieren. Für diese Adresse darf kein Aliasname auf der Loopback-Einheit des Back-End-Servers festgelegt sein, den Sie zur Load-Balancer-Konfiguration hinzufügen möchten. Fügen Sie dann den Server zur Load-Balancer-Konfiguration hinzu.

Java-Warnung bei Installation von Servicekorrekturen

Bei der Produktinstallation von Load Balancer wird auch eine Java-Dateigruppe installiert. Die Produktinstallation umfasst mehrere Pakete, die nicht auf derselben Maschine installiert werden müssen. Dazu gehören unter anderem das Paket mit Metric Server, das Administrationspaket und das Basispaket. Alle diese Codepakete, die auf gesonderten Maschinen installiert werden können, erfordern eine Java-Dateigruppe. Deshalb wird zusammen mit jedem dieser Pakete eine Java-Dateigruppe installiert. Werden die Pakete auf derselben Maschine installiert, ist jede dieser Dateigruppen Eigner der Java-Dateigruppe. Wenn Sie die zweite und dritte Java-Dateigruppe installieren, empfangen Sie eine Warnung, dass noch eine weitere Dateigruppe Eigner der Java-Dateigruppe ist.

Wenn Sie Code mit den nativen Installationsmethoden installieren (unter AIX beispielsweise mit `installp`), sollten Sie diese Warnung ignorieren.

Upgrade der Java-Dateigruppe einer Load-Balancer-Installation

Während der Installation von Load Balancer wird auch eine Java-Dateigruppe installiert. Load Balancer ist die einzige Anwendung, die die zusammen mit dem Produkt installierte Java-Version verwendet. Sie sollten diese Version der Java-Dateigruppe nicht selbst aktualisieren. Falls ein Problem auftritt, das ein Upgrade der Java-Dateigruppe erforderlich macht, sollten Sie es dem IBM Service melden, damit für die im Lieferumfang von Load Balancer enthaltene Java-Dateigruppe eine offizielle Fixversion mit einem Upgrade bereitgestellt werden kann.

Allgemeine Probleme lösen — CBR

Problem: CBR wird nicht ausgeführt

Dieses Problem kann auftreten, wenn eine andere Anwendung einen der Ports benutzt, die von CBR verwendet werden. Weitere Informationen enthält der Abschnitt „Port-Nummern für CBR überprüfen“ auf Seite 302.

Problem: Der Befehl `cbrcontrol` oder `lbadmin` scheitert

1. Der Befehl `cbrcontrol` gibt die Nachricht **Fehler: Server antwortet nicht** zurück, oder der Befehl `lbadmin` gibt die Nachricht **Fehler: Zugriff auf RMI-Server nicht möglich** zurück. Diese Fehler können auftreten, wenn der Stack Ihrer Maschine SOCKSifiziert ist. Lösen Sie dieses Problem, indem Sie die Datei `socks.cnf` so editieren, dass sie die folgenden Zeilen enthält:

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Die Verwaltungskonsolen für Load-Balancer-Schnittstellen (Befehlszeile, grafische Benutzerschnittstelle und Assistenten) kommunizieren per RMI (Remote

Method Invocation) mit `cbrserver`. Für die Standardkommunikation werden drei Ports verwendet, die im Start-Script für `cbrserver` wie folgt definiert sind:

- 11099 zum Empfangen der Befehle von `cbrcontrol`
- 10004 zum Senden von Messwertabfragen an Metric Server
- 11199 für den RMI-Server-Port

Diese Definition kann Fehler verursachen, wenn eine der Verwaltungskonsolen auf derselben Maschine als Firewall oder über eine Firewall ausgeführt wird. Wird beispielsweise Load Balancer auf derselben Maschine als Firewall ausgeführt, können beim Absetzen von `cbrcontrol`-Befehlen Fehler wie der folgende angezeigt werden: **Fehler: Server antwortet nicht.**

Sie können diesen Fehler vermeiden, indem Sie die `cbrserver`-Script-Datei editieren und den von RMI für die Firewall (oder eine andere Anwendung) verwendeten Port festlegen. Ändern Sie die Zeile `LB_RMISERVERPORT=11199` in `LB_RMISERVERPORT=Ihr_Port`. *Ihr_Port* ist ein anderer Port.

Starten Sie anschließend erneut `cbrserver` und öffnen Sie den Datenverkehr für die Ports 11099, 10004, 11199 und 11100 oder für den Port, den Sie für die Hostadresse, an der die Verwaltungskonsole ausgeführt wird, ausgewählt haben.

3. Derartige Fehler können auch auftreten, wenn Sie **`cbrserver`** noch nicht gestartet haben.

Problem: Anforderungen werden nicht verteilt

Anforderungen werden nicht verteilt, obwohl Caching Proxy und CBR gestartet wurden. Dieser Fehler kann auftreten, wenn Sie Caching Proxy vor dem Executor starten. Ist dies der Fall, enthält das Protokoll `stderr` für Caching Proxy die Fehlermeldung `"ndServerInit: Keine Verbindung zum Executor möglich"`. Vermeiden Sie dieses Problem, indem Sie den Executor vor Caching Proxy starten.

Problem: Auf Solaris-Systemen scheitert der Befehl `cbrcontrol executor start`

Auf Solaris-Systemen gibt der Befehl **`cbrcontrol executor start`** die Nachricht "Fehler: Executor wurde nicht gestartet" zurück. Dieser Fehler tritt auf, wenn Sie die prozessübergreifende Kommunikation (IPC, Inter-Process Communication) für das System nicht so konfigurieren, dass die maximale Größe eines gemeinsam benutzten Speichersegments und die Anzahl der gemeinsam benutzten Semaphor-IDs über dem Standardwert des Betriebssystems liegen. Wenn Sie das gemeinsam benutzte Speichersegment vergrößern und die Anzahl der gemeinsam benutzten Semaphor-IDs erhöhen möchten, müssen Sie die Datei `/etc/system` editieren. Weitere Informationen zum Konfigurieren dieser Datei finden Sie auf Seite 114.

Problem: Syntax- oder Konfigurationsfehler

Wenn der URL nicht funktioniert, kann dies an einem Syntax- oder Konfigurationsfehler liegen. Überprüfen Sie bei diesem Problem Folgendes:

- Stellen Sie sicher, dass die Regel korrekt konfiguriert ist. Ausführliche Informationen hierzu finden Sie in Anhang B, „Syntax der content-Regel“, auf Seite 477.
- Setzen Sie für diese Regel einen Befehl **`cbrcontrol rule report`** ab und überprüfen Sie, ob die Spalte 'Anzahl Ausführungen' entsprechend der Anzahl der Anforderungen erhöht wurde. Wurde der Wert korrekt erhöht, überprüfen Sie erneut die Serverkonfiguration.
- Wird die Regel nicht ausgeführt, fügen Sie eine Regel 'Immer wahr' hinzu. Setzen Sie für die immer gültige Regel einen Befehl **`cbrcontrol rule report`** ab, um zu prüfen, ob sie angewendet wird.

Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten

Wenn Sie auf der Windows-Plattform eine Matrox-AGP-Karte verwenden, kann es auf der Load-Balancer-GUI zu unerwartetem Verhalten kommen. Beim Klicken mit der Maus kann ein Block etwa von der Größe des Mauszeigers beschädigt werden und zur Umkehrung von Hervorhebungen oder zur Verschiebung von Abbildungen führen. Bei älteren Matrox-Karten wurde dieses Verhalten nicht beobachtet. Für Matrox-AGP-Karten gibt es keine bekannte Korrektur.

Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung

Wenn Sie Load Balancer mit der fernen Webverwaltung konfigurieren, dürfen Sie nicht die Größe des Netscape-Browserfensters ändern, in dem die Load-Balancer-GUI angezeigt wird. Das heißt, Sie dürfen das Fenster nicht minimieren, maximieren, wiederherstellen usw. Da Netscape bei jeder Größenänderung des Browserfensters die Seite neu lädt, kommt es zu einer Trennung der Hostverbindung, die nach einer solchen Änderung demzufolge neu hergestellt werden muss. Wenn Sie die ferne Webverwaltung auf einer Windows-Plattform ausführen, verwenden Sie den Internet Explorer.

Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)

Unter Windows können im Fenster mit der Eingabeaufforderung einige nationale Sonderzeichen der Zeichensatzfamilie Latin-1 beschädigt angezeigt werden. Der Buchstabe "a" mit Tilde kann beispielsweise als Pi-Symbol erscheinen. Zum Korrigieren dieses Fehlers müssen Sie die Schriftartmerkmale für das Fenster mit der Eingabeaufforderung ändern. Gehen Sie zum Ändern der Schriftart wie folgt vor:

1. Klicken Sie oben links in der Ecke des Fensters mit der Eingabeaufforderung auf das Symbol.
2. Wählen Sie "Eigenschaften" aus und klicken Sie auf das Register "Schriftart".
3. Die Standardeinstellung für "Schriftart" ist "Rasterschriftarten". Setzen Sie die Schriftart auf "Lucida Console" und klicken Sie auf OK.

Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads

Einige Installationen von HP-UX 11i sind so vorkonfiguriert, dass nur 64 Threads pro Prozess zulässig sind. Manche Load-Balancer-Konfigurationen erfordern jedoch mehr Threads. Setzen Sie die Anzahl der Threads pro Prozess auf HP-UX-Systemen auf mindestens 256. Verwenden Sie zum Erhöhen dieses Wertes das Dienstprogramm "sam" und definieren Sie den Kernel-Parameter `max_thread_proc` neu. Bei einer erwarteten starken Auslastung, müssen Sie `max_thread_proc` möglicherweise auf einen noch höheren Wert als 256 setzen.

Führen Sie zum Erhöhen von `max_thread_proc` die Schritte auf Seite 315 aus.

Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv

Wenn Sie Ihren Adapter in einer Load-Balancer-Maschine konfigurieren, müssen Sie für eine fehlerfreie Ausführung der Advisor-Funktion sicherstellen, dass die beiden folgenden Einstellungen richtig definiert sind:

- Inaktivieren Sie das Feature Task Offload, das von den meisten 3Com-Adapterkarten verwendet wird.
- Aktivieren Sie Protokoll 1 (ICMP) für IP-Protokolle, wenn Sie TCP/IP-Filter aktivieren. Wenn ICMP nicht aktiviert ist, scheitert der ping-Test zum Back-End-Server.

Anweisungen für das Konfigurieren dieser Einstellung finden Sie auf Seite 316.

Problem: Auflösung von IP-Adressen in Hostnamen auf Windows-Systemen, wenn für einen Adapter mehrere Adressen konfiguriert sind

Wenn Sie auf der Windows-Plattform einen Adapter mit mehreren IP-Adressen konfigurieren, muss die IP-Adresse, die mit dem Hostnamen verbunden werden soll, als erste Adresse in der Registrierungsdatenbank konfiguriert werden.

Da Load Balancer in vielen Instanzen (z. B. lbkeys create) von InetAddress.getLocalHost() abhängig ist, können mehrere IP-Adressen, die per Aliasname an einen Adapter gebunden sind, zu Fehlern führen. Zur Vermeidung dieses Problems sollten die IP-Adressen, in die der Hostname aufgelöst werden soll, als erste Adressen in der Registrierungsdatenbank aufgelistet sein.

Die Schritte für das Konfigurieren des Hostnamens als ersten Eintrag in der Registrierungsdatenbank sind auf Seite 316 beschrieben.

Allgemeine Fehler beheben — Site Selector

Problem: Site Selector wird nicht ausgeführt

Dieser Fehler kann auftreten, wenn eine andere Anwendung einen der von Site Selector verwendeten Ports benutzt. Weitere Informationen enthält der Abschnitt „Port-Nummern für Site Selector überprüfen“ auf Seite 302.

Problem: Site Selector verteilt den Datenverkehr von Solaris-Clients nicht nach der RoundRobin-Methode

Symptom: Site Selector gewichtet von Solaris-Clients eingehende Anforderungen nicht nach der RoundRobin-Methode.

Mögliche Ursache: Solaris-Systeme führen einen Namensservice-Cache-Dämon aus. Wenn dieser Dämon aktiv ist, wird die nächste Anfrage aus diesem Cache beantwortet, ohne dass Site Selector abgefragt wird.

Lösung: Inaktivieren Sie den Namensserver-Cache-Dämon auf der Solaris-Maschine.

Problem: Der Befehl sscontrol oder lbadmin scheitert

1. Der Befehl sscontrol gibt die Nachricht **Fehler: Server antwortet nicht** zurück, oder der Befehl lbadmin gibt die Nachricht **Fehler: Zugriff auf RMI-Server nicht möglich** zurück. Diese Fehler können auftreten, wenn der Stack Ihrer Maschine SOCKSifiziert ist. Lösen Sie dieses Problem, indem Sie die Datei socks.cnf so editieren, dass sie die folgenden Zeilen enthält:

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Die Verwaltungskonsolen für Load-Balancer-Schnittstellen (Befehlszeile, grafische Benutzerschnittstelle und Assistenten) kommunizieren per RMI (Remote Method Invocation) mit sserver. Für die Standardkommunikation werden drei Ports verwendet, die im Start-Script für sserver wie folgt definiert sind:

- 12099 zum Empfangen der Befehle von sscontrol
- 10004 zum Senden von Messwertabfragen an Metric Server
- 12199 für den RMI-Server-Port
- 53 für das Senden und Empfangen von DNS-Datenverkehr

Diese Definition kann Fehler verursachen, wenn eine der Verwaltungskonsolen auf derselben Maschine als Firewall oder über eine Firewall ausgeführt wird. Wird beispielsweise Load Balancer auf derselben Maschine als Firewall ausgeführt, können beim Absetzen von sscontrol-Befehlen Fehler wie der folgende angezeigt werden: **Fehler: Server antwortet nicht.**

Sie können diesen Fehler vermeiden, indem Sie die sserver-Script-Datei editieren und den von RMI für die Firewall (oder eine andere Anwendung) verwendeten Port festlegen. Ändern Sie die Zeile `LB_RMISERVERPORT=10199` in `LB_RMISERVERPORT=Ihr_Port`. *Ihr_Port* ist ein anderer Port.

Starten Sie anschließend erneut sserver und öffnen Sie den Datenverkehr für die Ports 12099, 10004, 12199 und 12100 oder für den Port, den Sie für die Hostadresse, an der die Verwaltungskonsole ausgeführt wird, ausgewählt haben.

3. Derartige Fehler können auch auftreten, wenn Sie **ssserver** noch nicht gestartet haben.

Problem: sserver wird auf der Windows-Plattform nicht gestartet

Site Selector muss an einem DNS teilhaben können. Alle zur Konfiguration gehörenden Maschinen sollten ebenfalls an diesem System teilhaben. Auf Windows-Systemen muss nicht immer der konfigurierte Hostname im DNS enthalten sein. Site Selector wird nur ordnungsgemäß gestartet, wenn der Hostname der Komponente im DNS definiert ist.

Prüfen Sie, ob dieser Host im DNS definiert ist. Editieren Sie die Datei `ssserver.cmd` und löschen Sie das "w" des Eintrags "javaw". Auf diese Weise erhalten Sie weitere Fehlerinformationen.

Problem: Site Selector führt bei duplizierten Routen den Lastausgleich nicht korrekt durch

Der Namensserver von Site Selector wird an keine Adresse der Maschine gebunden. Er beantwortet alle Anfragen, die an gültige IP-Adressen auf der Maschine gerichtet sind. Site Selector verlässt sich darauf, dass das Betriebssystem die Antwort an den Client zurückgibt. Wenn die Site-Selector-Maschine mehrere Adapter enthält und eine beliebige Anzahl dieser Adapter mit demselben Teilnetz verbunden sind, sendet das Betriebssystem die Antwort an den Client unter Umständen nicht von der Adresse, an die der Client seine Anfrage gesendet hat. Einige Client-Anwendungen akzeptieren nur Antworten, die sie von der Adresse empfangen, an die sie die Anfrage gesendet haben. Das erweckt den Anschein, als würde die Namensauflösung nicht funktionieren.

Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten

Wenn Sie auf der Windows-Plattform eine Matrox-AGP-Karte verwenden, kann es auf der Load-Balancer-GUI zu unerwartetem Verhalten kommen. Beim Klicken mit der Maus kann ein Block etwa von der Größe des Mauszeigers beschädigt werden und zur Umkehrung von Hervorhebungen oder zur Verschiebung von Abbildungen führen. Bei älteren Matrox-Karten wurde dieses Verhalten nicht beobachtet. Für Matrox-AGP-Karten gibt es keine bekannte Korrektur.

Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung

Wenn Sie Load Balancer mit der fernen Webverwaltung konfigurieren, dürfen Sie nicht die Größe des Netscape-Browserfensters ändern, in dem die Load-Balancer-GUI angezeigt wird. Das heißt, Sie dürfen das Fenster nicht minimieren, maximieren, wiederherstellen usw. Da Netscape bei jeder Größenänderung des Browserfensters die Seite neu lädt, kommt es zu einer Trennung der Hostverbindung, die nach einer solchen Änderung demzufolge neu hergestellt werden muss. Wenn Sie die ferne Webverwaltung auf einer Windows-Plattform ausführen, verwenden Sie den Internet Explorer.

Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)

Unter Windows können im Fenster mit der Eingabeaufforderung einige nationale Sonderzeichen der Zeichensatzfamilie Latin-1 beschädigt angezeigt werden. Der Buchstabe "a" mit Tilde kann beispielsweise als Pi-Symbol erscheinen. Zum Korrigieren dieses Fehlers müssen Sie die Schriftartmerkmale für das Fenster mit der Eingabeaufforderung ändern. Gehen Sie zum Ändern der Schriftart wie folgt vor:

1. Klicken Sie oben links in der Ecke des Fensters mit der Eingabeaufforderung auf das Symbol.
2. Wählen Sie "Eigenschaften" aus und klicken Sie auf das Register "Schriftart".
3. Die Standardeinstellung für "Schriftart" ist "Rasterschriftarten". Setzen Sie die Schriftart auf "Lucida Console" und klicken Sie auf OK.

Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads

Einige Installationen von HP-UX 11i sind so vorkonfiguriert, dass nur 64 Threads pro Prozess zulässig sind. Manche Load-Balancer-Konfigurationen erfordern jedoch mehr Threads. Setzen Sie die Anzahl der Threads pro Prozess auf HP-UX-Systemen auf mindestens 256. Verwenden Sie zum Erhöhen dieses Wertes das Dienstprogramm "sam" und definieren Sie den Kernel-Parameter `max_thread_proc` neu. Bei einer erwarteten starken Auslastung, müssen Sie `max_thread_proc` möglicherweise auf einen noch höheren Wert als 256 setzen.

Führen Sie zum Erhöhen von `max_thread_proc` die Schritte auf Seite 315 aus.

Problem: Auf Windows-Systemen markieren Advisor-Funktionen und Erreichbarkeitsziele alle Server als inaktiv

Wenn Sie Ihren Adapter in einer Load-Balancer-Maschine konfigurieren, müssen Sie für eine fehlerfreie Ausführung der Advisor-Funktion sicherstellen, dass die beiden folgenden Einstellungen richtig definiert sind:

- Inaktivieren Sie das Feature Task Offload, das von den meisten 3Com-Adapterkarten verwendet wird.
- Aktivieren Sie Protokoll 1 (ICMP) für IP-Protokolle, wenn Sie TCP/IP-Filter aktivieren. Wenn ICMP nicht aktiviert ist, scheitert der ping-Test zum Back-End-Server.

Anweisungen für das Konfigurieren dieser Einstellung finden Sie auf Seite 316.

Allgemeine Probleme lösen — Cisco CSS Controller

Problem: ccoserver wird nicht gestartet

Dieser Fehler kann auftreten, wenn eine andere Anwendung einen Port verwendet, der vom ccoserver des Cisco CSS Controller verwendet wird. Weitere Informationen hierzu finden Sie im Abschnitt „Port-Nummern für Cisco CSS Controller überprüfen“ auf Seite 303.

Problem: Der Befehl ccocontrol oder lbadm in scheitert

1. Der Befehl ccocontrol gibt die Nachricht **Fehler: Server antwortet nicht** zurück, oder der Befehl lbadm in gibt die Nachricht **Fehler: Zugriff auf RMI-Server nicht möglich** zurück. Diese Fehler können auftreten, wenn der Stack Ihrer Maschine SOCKSifiziert ist. Lösen Sie dieses Problem, indem Sie die Datei socks.cnf so editieren, dass sie die folgenden Zeilen enthält:

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Die Verwaltungskonsolen für Load-Balancer-Schnittstellen (Befehlszeile und grafische Benutzerschnittstelle) kommunizieren per RMI (Remote Method Invocation) mit ccoserver. Für die Standardkommunikation werden drei Ports verwendet, die im Start-Script für ccoserver wie folgt definiert sind:
 - 13099 zum Empfangen der Befehle von ccocontrol
 - 10004 zum Senden von Messwertabfragen an Metric Server
 - 13199 für den RMI-Server-Port

Diese Definition kann Fehler verursachen, wenn eine der Verwaltungskonsolen auf derselben Maschine als Firewall oder über eine Firewall ausgeführt wird. Wird beispielsweise Load Balancer auf derselben Maschine als Firewall ausgeführt, können beim Absetzen von ccocontrol-Befehlen Fehler wie der folgende angezeigt werden: **Fehler: Server antwortet nicht**.

Sie können diesen Fehler vermeiden, indem Sie die ccoserver-Script-Datei editieren und den von RMI für die Firewall (oder eine andere Anwendung) verwendeten Port festlegen. Ändern Sie die Zeile `CCO_RMISERVERPORT=14199` in `CCO_RMISERVERPORT=Ihr_Port`. *Ihr_Port* ist ein anderer Port.

Starten Sie anschließend erneut ccoserver und öffnen Sie den Datenverkehr für die Ports 13099, 10004, 13199 und 13100 oder für den Port, den Sie für die Hostadresse, an der die Verwaltungskonsole ausgeführt wird, ausgewählt haben.

3. Derartige Fehler können auch auftreten, wenn Sie **ccoserver** noch nicht gestartet haben.

Problem: Für Port 13099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden

Dieses Problem kann auftreten, wenn eine gültige Produktlizenz fehlt. Wenn Sie versuchen, ccoserver zu starten, empfangen Sie die folgende Nachricht:

Die Lizenz ist abgelaufen. IBM Ansprechpartner oder autorisierten IBM Händler kontaktieren.

Sie können dieses Problem wie folgt lösen:

1. Falls Sie bereits versucht haben, ccoserver zu starten, geben Sie **ccoserver stop** ein.
2. Kopieren Sie Ihre gültige Lizenz in das Verzeichnis **...ibm/edge/lb/servers/conf**.
3. Geben Sie **ccoserver** ein, um den Server zu starten.

Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten

Wenn Sie auf der Windows-Plattform eine Matrox-AGP-Karte verwenden, kann es auf der Load-Balancer-GUI zu unerwartetem Verhalten kommen. Beim Klicken mit der Maus kann ein Block etwa von der Größe des Mauszeigers beschädigt werden und zur Umkehrung von Hervorhebungen oder zur Verschiebung von Abbildungen führen. Bei älteren Matrox-Karten wurde dieses Verhalten nicht beobachtet. Für Matrox-AGP-Karten gibt es keine bekannte Korrektur.

Problem: Beim Hinzufügen eines Consultant wird ein Verbindungsfehler empfangen

Beim Hinzufügen eines Consultant kann es aufgrund falscher Konfigurationseinstellungen zu einem Verbindungsfehler kommen. Beheben Sie diesen Fehler wie folgt:

- Vergewissern Sie sich, dass die angegebene Adresse oder Benutzergemeinschaft genau mit dem für den Switch konfigurierten Wert übereinstimmt.
- Stellen Sie sicher, dass die Konnektivität zwischen dem Controller und dem Switch verfügbar ist.
- Vergewissern Sie sich, dass die Benutzergemeinschaft für den Switch eine Schreib-/Leseberechtigung hat. Der Controller versucht, die Variable `ApSvcLoadEnable` (SNMP) zu aktivieren, wenn er die Verbindung testet, um den Schreibzugriff zu überprüfen.

Problem: Auf dem Switch werden die Wertigkeiten nicht aktualisiert

Beheben Sie diesen Fehler wie folgt:

- Wenn Sie den Messwert "Aktive Verbindungen" oder "Verbindungsrate" verwenden, setzen Sie `ccocontrol service SW-ID:ID_für_Eignerangaben:Service-E/A report` ab. Überprüfen Sie, ob sich die Messwerte entsprechend dem Datendurchfluss auf dem Switch ändern.
- Erhöhen Sie die Protokollstufe des Consultant-Protokolls und suchen Sie nach SNMP-TimeOut-Einträgen. Falls es zu Zeitlimitüberschreitungen gekommen ist, bieten sich unter anderem die folgenden Möglichkeiten an:
 - Verringern Sie die Arbeitslast auf dem Switch.
 - Verringern Sie die Netzverzögerung zwischen Switch und Controller.
- Stoppen Sie den Consultant und starten Sie ihn erneut.

Problem: Befehl refresh aktualisiert die Consultant-Konfiguration nicht

Erhöhen Sie die Protokollstufe für den Consultant und wiederholen Sie den Befehl. Sollte er erneut scheitern, suchen Sie im Protokoll nach SNMP-Zeitlimitüberschreitungen oder SNMP-Übertragungsfehlern.

Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung

Wenn Sie Load Balancer mit der fernen Webverwaltung konfigurieren, dürfen Sie nicht die Größe des Netscape-Browserfensters ändern, in dem die Load-Balancer-GUI angezeigt wird. Das heißt, Sie dürfen das Fenster nicht minimieren, maximieren, wiederherstellen usw. Da Netscape bei jeder Größenänderung des Browserfensters die Seite neu lädt, kommt es zu einer Trennung der Hostverbindung, die nach einer solchen Änderung demzufolge neu hergestellt werden muss. Wenn Sie die ferne Webverwaltung auf einer Windows-Plattform ausführen, verwenden Sie den Internet Explorer.

Problem: Auf der Windows-Plattform erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)

Unter Windows können im Fenster mit der Eingabeaufforderung einige nationale Sonderzeichen der Zeichensatzfamilie Latin-1 beschädigt angezeigt werden. Der Buchstabe "a" mit Tilde kann beispielsweise als Pi-Symbol erscheinen. Zum Korrigieren dieses Fehlers müssen Sie die Schriftartmerkmale für das Fenster mit der Eingabeaufforderung ändern. Gehen Sie zum Ändern der Schriftart wie folgt vor:

1. Klicken Sie oben links in der Ecke des Fensters mit der Eingabeaufforderung auf das Symbol.
2. Wählen Sie "Eigenschaften" aus und klicken Sie auf das Register "Schriftart".
3. Die Standardeinstellung für "Schriftart" ist "Rasterschriftarten". Setzen Sie die Schriftart auf "Lucida Console" und klicken Sie auf OK.

Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads

Einige Installationen von HP-UX 11i sind so vorkonfiguriert, dass nur 64 Threads pro Prozess zulässig sind. Manche Load-Balancer-Konfigurationen erfordern jedoch mehr Threads. Setzen Sie die Anzahl der Threads pro Prozess auf HP-UX-Systemen auf mindestens 256. Verwenden Sie zum Erhöhen dieses Wertes das Dienstprogramm "sam" und definieren Sie den Kernel-Parameter `max_thread_proc` neu. Bei einer erwarteten starken Auslastung, müssen Sie `max_thread_proc` möglicherweise auf einen noch höheren Wert als 256 setzen.

Führen Sie zum Erhöhen von `max_thread_proc` die Schritte auf Seite 315 aus.

Allgemeine Probleme lösen — Nortel Alteon Controller

Problem: nalserver wird nicht gestartet

Dieser Fehler kann auftreten, wenn eine andere Anwendung einen Port verwendet, der vom nalserver des Nortel Alteon Controller verwendet wird. Weitere Informationen hierzu finden Sie im Abschnitt „Port-Nummern für Nortel Alteon Controller überprüfen“ auf Seite 304.

Problem: Der Befehl `nalcontrol` oder `lbadmin` scheitert

1. Der Befehl `nalcontrol` gibt die Nachricht **Fehler: Server antwortet nicht** zurück, oder der Befehl `lbadmin` gibt die Nachricht **Fehler: Zugriff auf RMI-Server nicht möglich** zurück. Diese Fehler können auftreten, wenn der Stack Ihrer Maschine SOCKSifiziert ist. Lösen Sie dieses Problem, indem Sie die Datei `socks.cnf` so editieren, dass sie die folgenden Zeilen enthält:

```
EXCLUDE-MODULE java
EXCLUDE-MODULE javaw
```

2. Die Verwaltungskonsolen für Load-Balancer-Schnittstellen (Befehlszeile und grafische Benutzerschnittstelle) kommunizieren per RMI (Remote Method Invocation) mit `nalserver`. Für die Standardkommunikation werden drei Ports verwendet, die im Start-Script für `nalserver` wie folgt definiert sind:
 - 14099 zum Empfang der Befehle von `nalcontrol`
 - 10004 zum Senden von Messwertabfragen an Metric Server
 - 14199 für den RMI-Server-Port

Diese Definition kann Fehler verursachen, wenn eine der Verwaltungskonsolen auf derselben Maschine als Firewall oder über eine Firewall ausgeführt wird. Wird beispielsweise Load Balancer auf derselben Maschine als Firewall ausgeführt, können beim Absetzen von `nalcontrol`-Befehlen Fehler wie der folgende angezeigt werden: **Fehler: Server antwortet nicht**.

Sie können diesen Fehler vermeiden, indem Sie die `nalserver`-Script-Datei editieren und den von RMI für die Firewall (oder eine andere Anwendung) verwendeten Port festlegen. Ändern Sie die Zeile `NAL_RMISERVERPORT=14199` in `NAL_RMISERVERPORT=Ihr_Port`. *Ihr_Port* ist ein anderer Port.

Starten Sie anschließend erneut `nalserver` und öffnen Sie den Datenverkehr für die Ports 14099, 10004, 14199 und 14100 oder für den Port, den Sie für die Hostadresse, an der die Verwaltungskonsole ausgeführt wird, ausgewählt haben.

3. Derartige Fehler können auch auftreten, wenn Sie `nalserver` noch nicht gestartet haben.

Problem: Für Port 14099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden

Dieses Problem kann auftreten, wenn eine gültige Produktlizenz fehlt. Wenn Sie versuchen, `nalserver` zu starten, empfangen Sie die folgende Nachricht:

```
Die Lizenz ist abgelaufen. IBM Ansprechpartner
oder autorisierten IBM Händler kontaktieren.
```

Sie können dieses Problem wie folgt lösen:

1. Falls Sie bereits versucht haben, `nalserver` zu starten, geben Sie `nalserver stop` ein.
2. Kopieren Sie Ihre gültige Lizenz in das Verzeichnis `...ibm/edge/lb/servers/conf`.
3. Geben Sie `nalserver` ein, um den Server zu starten.

Problem: Unerwartetes GUI-Verhalten auf der Windows-Plattform bei Verwendung von Matrox-AGP-Videokarten

Wenn Sie auf der Windows-Plattform eine Matrox-AGP-Karte verwenden, kann es auf der Load-Balancer-GUI zu unerwartetem Verhalten kommen. Beim Klicken mit der Maus kann ein Block etwa von der Größe des Mauszeigers beschädigt werden und zur Umkehrung von Hervorhebungen oder zur Verschiebung von Abbildungen führen. Bei älteren Matrox-Karten wurde dieses Verhalten nicht beobachtet. Für Matrox-AGP-Karten gibt es keine bekannte Korrektur.

Problem: Trennen der Hostverbindung bei Änderung des Netscape-Browserfensters in der Webverwaltung

Wenn Sie Load Balancer mit der fernen Webverwaltung konfigurieren, dürfen Sie nicht die Größe des Netscape-Browserfensters ändern, in dem die Load-Balancer-GUI angezeigt wird. Das heißt, Sie dürfen das Fenster nicht minimieren, maximieren, wiederherstellen usw. Da Netscape bei jeder Größenänderung des Browserfensters die Seite neu lädt, kommt es zu einer Trennung der Hostverbindung, die nach einer solchen Änderung demzufolge neu hergestellt werden muss. Wenn Sie die ferne Webverwaltung auf einer Windows-Plattform ausführen, verwenden Sie den Internet Explorer.

Problem: Beim Hinzufügen eines Consultant wird ein Verbindungsfehler empfangen

Beim Hinzufügen eines Consultant kann es aufgrund falscher Konfigurationseinstellungen zu einem Verbindungsfehler kommen. Beheben Sie diesen Fehler wie folgt:

- Vergewissern Sie sich, dass die angegebene Adresse oder Benutzergemeinschaft genau mit dem für den Switch konfigurierten Wert übereinstimmt.
- Stellen Sie sicher, dass die Konnektivität zwischen dem Controller und dem Switch verfügbar ist.
- Vergewissern Sie sich, dass die Benutzergemeinschaft für den Switch eine Schreib-/Leseberechtigung hat. Der Controller versucht, die Variable `ApSvcLoadEnable` (SNMP) zu aktivieren, wenn er die Verbindung testet, um den Schreibzugriff zu überprüfen.

Problem: Auf dem Switch werden die Wertigkeiten nicht aktualisiert

Beheben Sie diesen Fehler wie folgt:

- Wenn Sie den Messwert "Aktive Verbindungen" oder "Verbindungsrate" verwenden, setzen Sie `ccocontrol service SW-ID:ID_für_Eignerangaben:Service-E/A report` ab. Überprüfen Sie, ob sich die Messwerte entsprechend dem Datendurchfluss auf dem Switch ändern.
- Erhöhen Sie die Protokollstufe des Consultant-Protokolls und suchen Sie nach SNMP-TimeOut-Einträgen. Falls es zu Zeitlimitüberschreitungen gekommen ist, bieten sich unter anderem die folgenden Möglichkeiten an:
 - Verringern Sie die Arbeitslast auf dem Switch.
 - Verringern Sie die Netzverzögerung zwischen Switch und Controller.
- Stoppen Sie den Consultant und starten Sie ihn erneut.

Problem: Befehl refresh aktualisiert die Consultant-Konfiguration nicht

Erhöhen Sie die Protokollstufe für den Consultant und wiederholen Sie den Befehl. Sollte er erneut scheitern, suchen Sie im Protokoll nach SNMP-Zeitlimitüberschreitungen oder SNMP-Übertragungsfehlern.

Problem: Auf Windows-Systemen erscheint die Eingabeaufforderung mit beschädigten nationalen Sonderzeichen (Latin-1)

Unter Windows können im Fenster mit der Eingabeaufforderung einige nationale Sonderzeichen der Zeichensatzfamilie Latin-1 beschädigt angezeigt werden. Der Buchstabe "a" mit Tilde kann beispielsweise als Pi-Symbol erscheinen. Zum Korrigieren dieses Fehlers müssen Sie die Schriftartmerkmale für das Fenster mit der Eingabeaufforderung ändern. Gehen Sie zum Ändern der Schriftart wie folgt vor:

1. Klicken Sie oben links in der Ecke des Fensters mit der Eingabeaufforderung auf das Symbol.
2. Wählen Sie "Eigenschaften" aus und klicken Sie auf das Register "Schriftart".
3. Die Standardeinstellung für "Schriftart" ist "Rasterschriftarten". Setzen Sie die Schriftart auf "Lucida Console" und klicken Sie auf OK.

Problem: Java-Fehler unter HP-UX wegen unzureichender Speicherkapazität/Threads

Einige Installationen von HP-UX 11i sind so vorkonfiguriert, dass nur 64 Threads pro Prozess zulässig sind. Manche Load-Balancer-Konfigurationen erfordern jedoch mehr Threads. Setzen Sie die Anzahl der Threads pro Prozess auf HP-UX-Systemen auf mindestens 256. Verwenden Sie zum Erhöhen dieses Wertes das Dienstprogramm "sam" und definieren Sie den Kernel-Parameter `max_thread_proc` neu. Bei einer erwarteten starken Auslastung, müssen Sie `max_thread_proc` möglicherweise auf einen noch höheren Wert als 256 setzen.

Führen Sie zum Erhöhen von `max_thread_proc` die Schritte auf Seite 315 aus.

Allgemeine Fehler beheben — Metric Server

Problem: IOException für Metric Server auf der Windows-Plattform bei Ausführung von benutzerdefinierten Messwertdateien mit der Erweiterung .bat oder .cmd

Für Metric Server auf der Windows-Plattform müssen Sie für benutzerdefinierte Messwerte den vollständigen Namen angeben. An Stelle von **benutzermesswert** müssten Sie beispielsweise **benutzermesswert.bat** angeben. Der Name **benutzermesswert** ist in der Befehlszeile gültig, funktioniert jedoch nicht bei Ausführung in einer Laufzeitumgebung. Wenn Sie nicht den vollständigen Namen des Messwerts verwenden, empfangen Sie eine Metric Server IOException. Setzen Sie in der metricserver-Befehlsdatei die Variable `LOG_LEVEL` auf den Wert 3 und überprüfen Sie die Protokollausgabe. In diesem Beispiel sieht die Ausnahmebedingung wie folgt aus:

```
... java.io.IOException: CreateProcess: usermetric error=2
```

Problem: Metric Server meldet die Last nicht an die Load-Balancer-Maschine

Dafür, dass Metric Server keine Lastinformationen an Load Balancer meldet, kann es mehrere Gründe geben. Überprüfen Sie Folgendes, um die Ursache zu ermitteln:

- Vergewissern Sie sich, dass die Schlüsselringe zu Metric Server übertragen wurden.
- Prüfen Sie, ob der Hostname der Metric-Server-Maschine im lokalen Namensserver registriert ist.

Sie können dieses Problem auch lösen, indem Sie im Script `metricserver` den Hostnamen im Java-Merkmal `java.rmi.server.hostname` angeben.

- Führen Sie einen Neustart mit einer höheren Protokollstufe durch und sehen Sie sich die Fehlernachrichten an.
- Erhöhen Sie auf der Load-Balancer-Maschine die Protokollstufe für das Protokoll der Metriküberwachung. Verwenden Sie dazu den Befehl **`dscontrol manager metric set`**. Suchen Sie in der Datei `MetricMonitor.log` nach Fehlern.

Problem: Metric-Server-Protokoll meldet, dass für den Zugriff auf den Agenten eine Kennung erforderlich ist

Das Metric-Server-Protokoll enthält diese Fehlernachricht, nachdem Schlüsselringe zum Server übertragen wurden.

Dieser Fehler wird registriert, wenn der Schlüsselring aufgrund einer Beschädigung des Schlüsselpaares nicht autorisiert werden kann. Versuchen Sie wie folgt, diesen Fehler zu beheben:

- Senden Sie den Schlüsselring erneut mit FTP und verwenden Sie die binäre Übertragungsmethode.
- Erstellen Sie einen neuen Schlüssel und verteilen Sie diesen.

Problem: Bei Ausführung von Metric Server auf AIX-Systemen kann die Ausgabe des Befehls `ps -vg` beschädigt werden

Wenn Metric Server auf einer AIX-Multiprozessorplattform (4.3.3, 5.1 32-Bit oder 64-Bit) mit starker Belastung ausgeführt wird, kann die Ausgabe des Befehls `ps -vg` beschädigt sein. Beispiel:

```
55742 - A 88:19 42 18014398509449680 6396 32768 22 36 2.8 1.0 java -Xms
```

Das Feld `SIZE` und/oder `RSS` des Befehls `ps` kann die Verwendung einer zu großen Speicherkapazität anzeigen.

Dies ist ein bekannter AIX-Kernel-Fehler, der mit APAR IY33804 korrigiert werden kann. Sie können die Korrektur von der AIX-Support-Site <http://techsupport.services.ibm.com/server/fixes> herunterladen oder sich an Ihr lokales AIX-Support-Team wenden.

Problem: Konfigurieren von Metric Server in einer zweistufigen Konfiguration mit Site Selector für die Lastverteilung auf Dispatcher mit hoher Verfügbarkeit

Wenn Site Selector (erste Stufe) in einer zweistufigen Load-Balancer-Konfiguration die Last auf ein Paar von Dispatcher-Partnern mit hoher Verfügbarkeit (zweite Stufe) verteilt, sind für die Komponente Metric Server bestimmte Konfigurationsschritte erforderlich. Sie müssen Metric Server für die Empfangsbereitschaft an einer neuen IP-Adresse konfigurieren, die speziell für Metric Server bestimmt ist. Auf den beiden Dispatcher-Maschinen mit hoher Verfügbarkeit ist Metric Server nur auf dem aktiven Dispatcher aktiv.

Führen Sie die folgenden Schritte aus, um diese Konfiguration korrekt zu definieren:

- Konfigurieren Sie Metric Server für die Empfangsbereitschaft an der neuen lokalen IP-Adresse. Metric Server sollte nicht an der lokalen NFA antworten. Weitere Konfigurationsinformationen finden Sie im Abschnitt „Metric Server“ auf Seite 196.

- Da Site Selector nur mit dem aktiven Dispatcher kommunizieren sollte, müssen Sie Metric Server in den go-Scripts für hohe Verfügbarkeit starten und stoppen. Zum korrekten Starten oder Stoppen von Metric Server müssen Sie für die neue IP-Adresse, die für Metric Server spezifisch ist, einen Aliasnamen auf der Maschine festlegen. Modifizieren Sie die go-Scripts (ähnlich wie für die Verschiebung von Clusteradressen), so dass das Script goActive die IP-Adresse für Metric Server vom Loopback-Adapter zu einem physischen Adapter verschiebt und das Script goStandby die Adresse vom physischen Adapter zu einem Loopback-Adapter verschiebt. Nach dem Versetzen der IP-Adresse muss das Script goActive den Befehl `metricserver` ausführen, um Metric Server zu starten. Das Script goStandby muss **`metricserver stop`** ausführen, um zu verhindern, dass Metric Server im Bereitschaftsmodus mit Site Selector kommuniziert.
- Informationen zum Versetzen der für Metric Server spezifischen IP-Adresse auf der Windows-Plattform finden Sie im Abschnitt „Scripts verwenden“ auf Seite 209.
- Zu den Änderungen am Script goStandby gehören unter anderem betriebssystemspezifische Anweisungen wie die folgenden:
 - **HP-UX-, Linux- und Solaris-Systeme:** Fügen Sie in den Abschnitt des Scripts goStandby, in dem die Clusteradresse zur Loopback-Einheit verschoben wird, Befehle zum Verschieben der für Metric Server spezifischen IP-Adresse zur Loopback-Einheit ein. Fügen Sie als nächstes den Befehl **`metricserver stop`** ein, um die Reaktion von Metric Server auf Site Selector zu stoppen.
 - **AIX-Systeme:** Fügen Sie in den Abschnitt des Scripts goStandby, in dem die Clusteradresse zur Loopback-Einheit verschoben wird, Befehle zum Verschieben der für Metric Server spezifischen IP-Adresse zur Loopback-Einheit ein. Fügen Sie als nächstes eine Route für die Kommunikation mit dem Aliasnamen auf der Loopback-Einheit ein. Führen Sie den Befehl **`route add IP-Adresse_für_Metric_Server 127.0.0.1`** aus. Fügen Sie dann den Befehl **`metricserver stop`** aus, um zu verhindern, dass Metric Server weiterhin auf Site Selector reagiert. Nachdem Metric Server gestoppt wurde, müssen Sie die Loopback-Route entfernen. Fügen Sie zur Vermeidung späterer Fehler **`route delete IP-Adresse_für_Metric_Server`** ein.

Beispiel:

```
ifconfig en0 delete 9.27.23.61
ifconfig lo0 alias 9.27.23.61 netmask 255.255.255.0
route add 9.27.23.61 127.0.0.1
metricserver stop
# Inaktivität von maximal 60 Sekunden oder bis zum Stoppen von Metric Server
let loopcount=0
while [[ "$loopcount" -lt "60" && 'ps -ef |grep AgentStop|
      grep -c -v gr ep' -eq "1" ]]
do
  sleep 1
  let loopcount=$loopcount+1
done
route delete 9.27.23.61
```

- **Windows-Systeme:** Zunächst muss der Loopback-Adapter für Metric Server (der im folgenden Beispiel als Verbindung 2 im lokalen Netz bezeichnet wird) mit einer IP-Adresse in Ihrer Maschine installiert werden. Verwenden Sie eine nicht benutzte Adresse für private Netze, z. B. 10.1.1.1. Nehmen Sie nach dem Konfigurieren der Loopback-Adresse die Änderungen an den go-Scripts vor. In das Script goStandby wird der netsh-Befehl zum Versetzen der IP-Adresse für Metric Server auf den Loopback-Adapter für Metric Server aufgenommen. Anschließend wird der Befehl **`metricserver stop`** ausgeführt.

Beispiel:

```
call netsh interface ip delete address "LAN-Verbindung" addr=9.27.23.61
call netsh interface ip add address "LAN-Verbindung 2" addr=9.27.2.3.61
mask = 255.255.255.0
sleep 3
metricserver stop
```

Problem: Unerwünschte Konsolnachrichten bei Ausführung von Scripts auf Solaris-Maschinen mit mehreren CPUs

Die Scripts `metricserver`, `cpuload` und `memload` können bei Ausführung auf Solaris-Maschinen mit mehreren CPUs unerwünschte Konsolnachrichten erzeugen. Dieses Verhalten ist auf die Verwendung des `VMSTAT`-Systembefehls zum Abrufen von CPU- und Speicherstatistiken vom Kernel zurückzuführen. Einige von `VMS-TAT` zurückgegebene Nachrichten geben an, dass sich der Status des Kernels geändert hat. Die Scripts können diese Nachrichten nicht bearbeiten, was zu unnötigen Konsolnachrichten von der Shell führt.

Beispiele für solche Konsolnachrichten:

```
/opt/ibm/edge/lb/ms/script/memload[29]: TOTAL=: syntax error
/opt/ibm/edge/lb/ms/script/memload[31]: LOAD=4*100/0: divide by zero
/opt/ibm/edge/lb/ms/script/memload[29]: TOTAL=659664+: more tokens expected
```

Sie können diese Nachrichten ignorieren.

Bei Ausführung von Load Balancer für IPV6 ist auf Linux-Systemen kein Abruf der Werte von Metric Server möglich.

Auf Linux-Plattformen gibt es eine Inkompatibilität bei der Auswahl der Quellen-IPV6-Adresse. Die Metriküberwachung versucht deshalb, über die falsche Quellen-IP-Adresse mit Metric Server zu kommunizieren.

Auf Linux-Systemen wird als IPV6-Quellenadresse für eine bestimmte Route standardmäßig die letzte konfigurierte Adresse ausgewählt, die mit dem Netzabschnitt der Route übereinstimmt.

Falls die letzte konfigurierte Schnittstelle ein IPV6-Cluster ist und diese Schnittstelle mit dem Netzabschnitt einer Route in der Routentabelle übereinstimmt, wird die Schnittstelle als Standardquellen-IP-Adresse für diese Route verwendet. Wird diese Route zwischen Load Balancer und Metric Server genutzt, ist zwischen den beiden Knoten keine Kommunikation möglich.

Die Kommunikation ist unmöglich, weil der Load-Balancer-Knoten versucht, unter Verwendung der Clusteradresse als Quellen-IP-Adresse mit Metric Server zu kommunizieren. Wenn der Cluster auf der Loopback-Einheit des Metric-Server-Knotens konfiguriert ist, wird die Antwort von Metric Server an die Loopback-Einheit gesendet, so dass keine Kommunikation erfolgen kann.

Lösung:

Mit dem folgenden Befehl können Sie feststellen, welche Adresse der Linux-Knoten für eine bestimmte Route verwendet und welche Schnittstelle für die RMI-Kommunikation zwischen der Metriküberwachung und Metric Server genutzt wird:

```
ip -6 route get Ihre_IPV6-Route
```

Nehmen wir an, Sie setzen diesen Befehl ab:

```
ip -6 route get fec0::/64
```

Der Befehl gibt Folgendes zurück:

```
fec0:: via fec0:: dev eth0 src fec0::4 metric 0 cache mtu 1500 advmss 1383
```

Falls fec0::4 eine Clusteradresse ist, muss eine andere Schnittstelle zur Einheit hinzugefügt werden, damit der Cluster nicht als Standardquelle verwendet wird. Sie können auch eine der vorherigen Nicht-Clusterschnittstellen entfernen und erneut hinzufügen.

Beispiel:

```
ip -6 addr add fec0::5/64 dev eth0
```

Problem: Nach dem Start von Metric Server wird der Wert -1 zurückgegeben

Dieses Problem kann auftreten, wenn die Integrität der Schlüsseldateien bei der Übertragung zum Client verloren geht.

Wenn Sie Ihre Schlüsseldateien von der Load-Balancer-Maschine mit FTP zum Back-End-Server übertragen, müssen Sie den Befehl put oder get im Binärmodus verwenden, um die Schlüsseldateien auf den FTP-Server zu stellen oder vom FTP-Server abzurufen.

Teil 9. Befehlsreferenz

Dieser Teil enthält Referenzinformationen zu den Befehlen aller Komponenten von Load Balancer. Zu diesem Teil gehören die folgenden Kapitel:

- Kapitel 26, „Syntaxdiagramm lesen“, auf Seite 347
- Kapitel 27, „Befehlsreferenz für Dispatcher und CBR“, auf Seite 349
- Kapitel 28, „Befehlsreferenz für Site Selector“, auf Seite 405
- Kapitel 29, „Befehlsreferenz für Cisco CSS Controller“, auf Seite 433
- Kapitel 30, „Befehlsreferenz für Nortel Alteon Controller“, auf Seite 451

Kapitel 26. Syntaxdiagramm lesen

Im Syntaxdiagramm wird gezeigt, wie ein Befehl angegeben wird, damit das Betriebssystem die Eingabe korrekt interpretieren kann. Lesen Sie das Syntaxdiagramm von links nach rechts und von oben nach unten entlang der horizontalen Linie (Hauptpfad).

Symbole und Interpunktion

In den Syntaxdiagrammen werden die folgenden Symbole benutzt:

Symbol	Beschreibung
--------	--------------

»	Markiert den Anfang der Befehlssyntax.
---	--

«	Markiert das Ende der Befehlssyntax.
---	--------------------------------------

Alle im Syntaxdiagramm aufgeführten Interpunktionszeichen, beispielsweise Doppelpunkte, Fragezeichen und Minuszeichen, müssen wie gezeigt übernommen werden.

Parameter

In den Syntaxdiagrammen werden die folgenden Arten von Parametern benutzt:

Parameter	Beschreibung
-----------	--------------

Erforderlich	Erforderliche Parameter werden im Hauptpfad gezeigt.
---------------------	--

Optional	Optionale Parameter werden unter dem Hauptpfad gezeigt.
-----------------	---

Parameter werden als Schlüsselwörter oder Variablen klassifiziert. Schlüsselwörter werden in Kleinbuchstaben gezeigt und können in Kleinbuchstaben eingegeben werden. Ein Befehlsname ist beispielsweise ein Schlüsselwort. Variablen stehen in Kursivschrift und stellen Namen oder Werte dar, die von Ihnen zur Verfügung gestellt werden müssen.

Beispiele für die Syntax

In dem folgenden Beispiel ist der Befehl **user** ein Schlüsselwort. Die erforderliche Variable ist die *Benutzer-ID* und die optionale Variable das *Kennwort*. Ersetzen Sie die Variablen durch Ihre eigenen Werte.

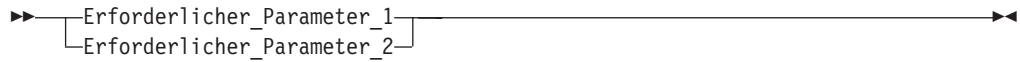
»—user—*Benutzer-ID*—*Kennwort*—«

Erforderliche Schlüsselwörter: Erforderliche Schlüsselwörter und Variablen stehen im Hauptpfad.

»—Erforderliches_Schlüsselwort—«

Erforderliche Schlüsselwörter und Werte **müssen** eingegeben werden.

Sich gegenseitig ausschließende Parameter aus einer Gruppe auswählen: Sind mehrere Schlüsselwörter oder Variablen aufgeführt, die sich gegenseitig ausschließen und aus denen ein Schlüsselwort oder eine Variable ausgewählt werden muss, sind sie vertikal in alphanumerischer Anordnung aufgeführt.



Optionale Werte: Optionale Schlüsselwörter und Variablen stehen unter dem Hauptpfad.



Sie können auswählen, ob sie optionale Schlüsselwörter oder Variablen angeben wollen oder nicht.

Optionale Schlüsselwörter oder Parameter aus einer Gruppe auswählen: Sind mehrere Schlüsselwörter oder Variablen aufgeführt, die sich gegenseitig ausschließen und aus denen ein Schlüsselwort oder eine Variable ausgewählt werden kann, sind sie vertikal in alphanumerischer Anordnung unter dem Hauptpfad aufgeführt.



Variablen: Wörter in Kursivschrift sind *Variablen*. Erscheint eine Variable in der Syntax, muss sie wie im Text angegeben durch einen ihrer erlaubten Namen oder Werte ersetzt werden.



Zeichen, die keine alphanumerischen Zeichen sind: Enthält ein Diagramm ein Zeichen, das kein alphanumerisches Zeichen ist (beispielsweise einen Doppelpunkt, ein Anführungszeichen oder ein Minuszeichen), müssen Sie dieses Zeichen als Teil der Syntax angeben. Im folgenden Beispiel müssen Sie den Cluster und den Port im Format *Cluster:Port* angeben.



Kapitel 27. Befehlsreferenz für Dispatcher und CBR

Dieses Kapitel beschreibt die Verwendung der **dscontrol**-Befehle von Dispatcher. Sie können dieses Kapitel auch als Befehlsreferenz für CBR verwenden.

Frühere Versionen dieses Produkts liefen unter dem Namen Network Dispatcher. In diesen Versionen war der Dispatcher-Steuerbefehl **ndcontrol**. Jetzt ist der Dispatcher-Steuerbefehl **dscontrol**. Stellen Sie sicher, dass alle bisherigen Script-Dateien so aktualisiert werden, dass **dscontrol** (und nicht **ndcontrol**) zum Konfigurieren von Dispatcher verwendet wird.

CBR verwendet eine Untergruppe der in dieser Befehlsreferenz aufgeführten Dispatcher-Befehle. Wenn Sie diese Syntaxdiagramme für **CBR** verwenden, ersetzen Sie **dscontrol** durch **cbrcontrol**. Weitere Informationen hierzu finden Sie im Abschnitt „Konfigurationsunterschiede bei CBR und Dispatcher“ auf Seite 350.

Wichtiger Hinweis: Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, steht nur die Dispatcher-Komponente zur Verfügung. Bei einer solchen Installation verwendet der Dispatcher eine Untergruppe der in dieser Befehlsreferenz aufgeführten Dispatcher-Befehle. Wenn Sie die Syntaxdiagramme verwenden, ersetzen Sie den Doppelpunkt (:) als Begrenzer im Befehl **dscontrol** durch **at (@)**. Weitere Informationen zur Installation von Load Balancer für IPV4 und IPV6 finden Sie in den Abschnitten „Unterschiede in der Befehlssyntax“ auf Seite 92 und „Unterstützte **dscontrol**-Befehle“ auf Seite 93.

Nachfolgend sind die in diesem Kapitel beschriebenen Befehle aufgelistet:

- „**dscontrol advisor** — Advisor-Funktion steuern“ auf Seite 351
- „**dscontrol binlog** — Binäre Protokolldatei steuern“ auf Seite 356
- „**dscontrol cluster** — Cluster konfigurieren“ auf Seite 357
- „**dscontrol executor** — Executor steuern“ auf Seite 361
- „**dscontrol file** — Konfigurationsdateien verwalten“ auf Seite 366
- „**dscontrol help** — Hilfetext für diesen Befehl anzeigen oder drucken“ auf Seite 368
- „**dscontrol highavailability** — Hohe Verfügbarkeit steuern“ auf Seite 369
- „**dscontrol host** — Ferne Maschine konfigurieren“ auf Seite 373
- „**dscontrol logstatus** — Protokolleinstellungen des Servers anzeigen“ auf Seite 374
- „**dscontrol manager** — Manager steuern“ auf Seite 375
- „**dscontrol metric** — Systemmesswerte konfigurieren“ auf Seite 381
- „**dscontrol port** — Ports konfigurieren“ auf Seite 382
- „**dscontrol rule** — Regeln konfigurieren“ auf Seite 388
- „**dscontrol server** — Server konfigurieren“ auf Seite 394
- „**dscontrol set** — Serverprotokoll konfigurieren“ auf Seite 400
- „**dscontrol status** — Aktivitätsanzeige für Manager und Advisor-Funktionen“ auf Seite 401
- „**dscontrol subagent** — SNMP-Subagenten konfigurieren“ auf Seite 402

Sie können eine Minimalversion der Parameter für den Befehl **dscontrol** eingeben. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben.

Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **dscontrol he f** an Stelle von **dscontrol help file** angeben.

Wenn Sie die Befehlszeilenschnittstelle starten möchten, setzen Sie den Befehl **dscontrol** ab, um die Eingabeaufforderung **dscontrol** aufzurufen.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie den Befehl **exit** oder **quit** absetzen.

Die Werte der Befehlsparameter müssen in englischen Zeichen eingegeben werden. Die einzige Ausnahme hiervon bilden Hostnamen (die in den Befehlen **cluster**, **server** und **highavailability** verwendet werden) und (die in Dateibefehlen verwendeten) Dateinamen.

Konfigurationsunterschiede bei CBR und Dispatcher

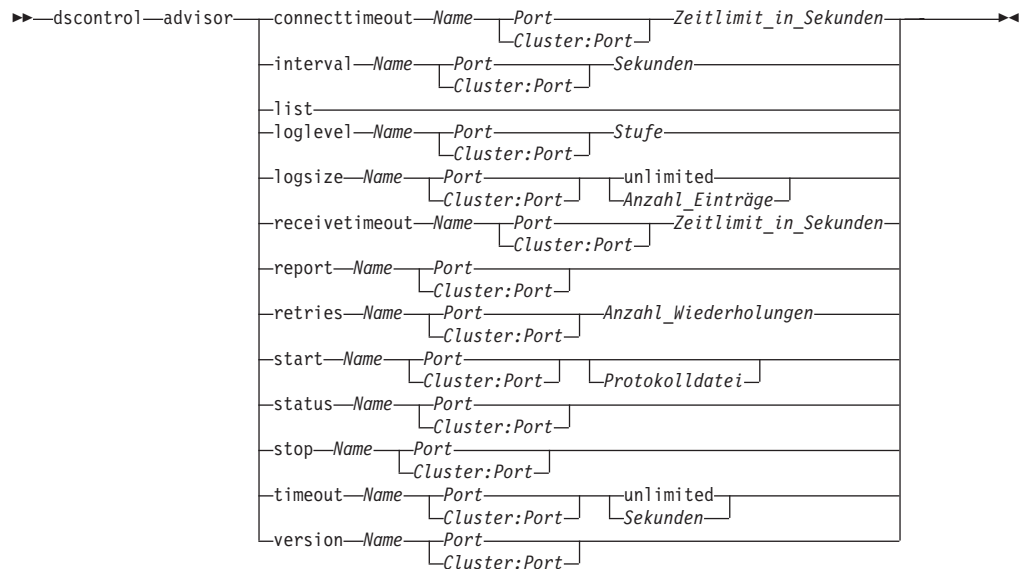
Die Befehlszeilenschnittstelle von CBR umfasst einen Teil der Befehlszeilenschnittstelle von Dispatcher. Verwenden Sie für CBR an Stelle des Befehls **dscontrol** den Befehl **cbrcontrol**, um die Komponente zu konfigurieren.

Anmerkung: Die Komponente CBR (Content Based Routing) ist auf allen unterstützten Plattformen, die keine 64-Bit-JVM ausführen, verfügbar. Alternativ können Sie für CBR ohne Caching Proxy die Weiterleitungsmethode **cbr** der Dispatcher-Komponente von Load Balancer verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Nachfolgend sind einige der Befehle aufgelistet, die in CBR *ignoriert* werden.

1. **highavailability**
2. **subagent**
3. **executor**
 - **report**
 - **set nfa <Wert>**
 - **set fintimeout <Wert>**
 - **set hatimeout <Wert>**
 - **set hasynctimeout <Wert>**
 - **set porttype <Wert>**
4. **Cluster**
 - **report {c}**
 - **set {c} porttype**
5. **port**
 - **add {c:p} porttype**
 - **add {c:p} protocol**
 - **set {c:p} porttype**
6. **rule add {c:p:r} type port**
7. **server**
 - **add {c:p:s} router**
 - **set {c:p:s} router**

dscontrol advisor — Advisor-Funktion steuern



connecttimeout

Definiert, wie lange eine Advisor-Funktion wartet, bevor sie meldet, dass zu einem bestimmten Port eines Servers (einem Service) keine Verbindung hergestellt werden kann. Weitere Informationen hierzu finden Sie im Abschnitt „Serververbindungs- und -empfangszeitlimit der Advisor-Funktion“ auf Seite 188.

Name

Der Name der Advisor-Funktion. Zu den gültigen Werten gehören **connect**, **db2**, **dns**, **ftp**, **http**, **https**, **cachingproxy**, **imap**, **ldap**, **nntp**, **ping**, **pop3**, **self**, **sip**, **smtp**, **ssl**, **ssl2http**, **telnet** und **wlm**.

Weitere Informationen zu den von Load Balancer bereitgestellten Advisor-Funktionen finden Sie im Abschnitt „Liste der Advisor-Funktionen“ auf Seite 189.

Die Namen angepasster Advisor-Funktionen haben das Format **xxxx**, wobei **ADV_xxxx** der Name der Klasse ist, die die angepasste Advisor-Funktion implementiert. Weitere Informationen hierzu finden Sie im Abschnitt „Kundenspezifische (anpassbare) Advisor-Funktion erstellen“ auf Seite 193.

Port

Die Nummer des Ports, der von der Advisor-Funktion überwacht wird.

Cluster:Port

Der Wert **"Cluster"** ist in den advisor-Befehlen optional, der Wert **"Port"** jedoch erforderlich. Wenn kein Wert für **"Cluster"** angegeben ist, wird die Advisor-Funktion an dem Port für alle Cluster gestartet. Wenn Sie einen Cluster angeben, wird die Advisor-Funktion an dem Port gestartet, jedoch nur für den von Ihnen genannten Cluster. Weitere Informationen hierzu finden Sie im Abschnitt „Advisor-Funktion starten und stoppen“ auf Seite 186.

Der Cluster kann als Adresse im IP-Adressenformat oder als symbolischer Name angegeben werden. Der Port wird als Nummer des Ports angegeben, der von der Advisor-Funktion überwacht wird.

Zeitlimit_in_Sekunden

Eine positive ganze Zahl, die das Zeitlimit in Sekunden angibt, nach dessen Ablauf die Advisor-Funktion meldet, dass zu einem Server keine Verbindung hergestellt werden kann. Dieser Wert liegt standardmäßig beim Dreifachen des für das Advisor-Intervall angegebenen Wertes.

interval

Legt fest, wie oft der Advisor Informationen von den Servern abfragt.

Sekunden

Eine positive ganze Zahl, die die Zeit zwischen den an die Server gerichteten Statusabfragen in Sekunden angibt. Der Standardwert ist 7.

list

Zeigt eine Liste der Advisor an, die derzeit Informationen an den Manager liefern.

loglevel

Legt die Protokollstufe für ein Advisor-Protokoll fest.

Stufe

Die Nummer der Stufe (0 bis 5). Der Standardwert ist 1. Je größer die Zahl ist, desto mehr Informationen werden in das Advisor-Protokoll geschrieben. Die folgenden Werte sind gültig: 0 für keine Einträge, 1 für eine minimale Protokollierung, 2 für eine Basisprotokollierung, 3 für eine normale, 4 für eine erweiterte und 5 für eine ausführliche Protokollierung.

logsize

Legt die maximale Größe eines Advisor-Protokolls fest. Wenn Sie eine maximale Größe für die Protokolldatei festlegen, findet ein Dateiumlauf statt. Hat die Datei die angegebene Größe erreicht, werden alle weiteren Einträge wieder an den Anfang der Datei geschrieben und die dort befindlichen Einträge überschrieben. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Anzahl_Einträge

Die maximale Größe der Advisor-Protokolldatei in Bytes. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** (unbegrenzt) angeben. Möglicherweise erreicht die Protokolldatei nicht genau die maximale Größe, bevor der Dateiumlauf stattfindet, da die Größe der Protokolleinträge variiert. Der Standardwert ist 1 MB.

receivetimeout

Definiert, wie lange eine Advisor-Funktion wartet, bevor sie meldet, dass von einem bestimmten Port eines Servers (einem Service) keine Daten empfangen werden können. Weitere Informationen hierzu finden Sie im Abschnitt „Serververbindungs- und -empfangszeitlimit der Advisor-Funktion“ auf Seite 188.

Zeitlimit_in_Sekunden

Eine positive ganze Zahl, die das Zeitlimit in Sekunden angibt, nach dessen Ablauf die Advisor-Funktion meldet, dass von einem Server keine Daten empfangen werden können. Dieser Wert liegt standardmäßig beim Dreifachen des für das Advisor-Intervall angegebenen Wertes.

report

Anzeigen eines Berichts zum Advisor-Status.

retry

Der Parameter `retry` legt die Wiederholungsversuche einer Advisor-Funktion fest, bevor diese einen Server als inaktiv markiert.

Anzahl_Wiederholungen

Eine ganze Zahl größer als oder gleich null. Dieser Wert sollte nicht größer als 3 sein. Wenn das Schlüsselwort für Wiederholungen nicht konfiguriert ist, wird standardmäßig von null Wiederholungsversuchen ausgegangen.

start

Den Advisor starten. Für alle Protokolle stehen Advisor zur Verfügung. Die Standard-Ports sind:

Advisor-Name	Protokoll	Port
cachingproxy	HTTP (über Caching Proxy)	80
connect	ICMP	12345
db2	privat	50000
dns	DNS	53
ftp	FTP	21
http	HTTP	80
https	SSL	443
imap	IMAP	143
ldap	LDAP	389
nntp	NNTP	119
ping	PING	0
pop3	POP3	110
self	privat	12345
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
ssl2http	SSL	443
telnet	Telnet	23
WLM	privat	10007

Anmerkung: Die FTP-Advisor-Funktion darf nur für den FTP-Steuer-Port (21) ausgeführt werden. Starten Sie eine FTP-Advisor-Funktion nicht für den FTP-Daten-Port (20).

Protokolldatei

Der Name der Datei, in die die Verwaltungsdaten geschrieben werden. Jeder Eintrag im Protokoll wird mit einer Zeitmarke versehen.

Die Standarddatei ist *Advisor-Name_Port.log*, z. B. **http_80.log**. Wollen Sie das Verzeichnis ändern, in dem die Protokolldateien gespeichert werden, lesen Sie die Informationen im Abschnitt „Pfade für die Protokolldatei ändern“ auf Seite 269. Die Standardprotokolldateien für cluster- oder sitespezifische Advisor-Funktionen werden mit der Clusteradresse erstellt, z. B.

http_127.40.50.1_80.log.

status

Zeigt den aktuellen Status aller Werte in einem Advisor an, die global gesetzt werden können. Zudem werden die Standardwerte dieser Werte angezeigt.

stop

Den Advisor stoppen.

timeout

Legt die Anzahl von Sekunden fest, in denen der Manager von dem Advisor erhaltene Informationen als gültig ansieht. Stellt der Manager fest, dass die Advisor-Informationen älter als dieses Zeitlimit sind, verwendet der Manager diese Informationen nicht zum Bestimmen Wertigkeiten für die Server am Port, die von der Advisor-Funktion überwacht werden. Dieses Zeitlimit gilt nicht, wenn die Advisor-Funktion den Manager darüber informiert hat, dass ein bestimmter Server inaktiv ist. Der Manager verwendet diese Information über den Server auch nach Überschreitung des Informationszeitlimits für die Advisor-Funktion weiter.

Sekunden

Eine positive Zahl, die die Anzahl von Sekunden darstellt, oder das Wort **unlimited** (unbegrenzt). Der Standardwert ist unlimited.

version

Zeigt die aktuelle Advisor-Version an.

Beispiele

- Starten der Advisor-Funktion http am Port 80 für Cluster 127.40.50.1:
dscontrol advisor start http 127.40.50.1:80
- Starten der Advisor-Funktion http am Port 88 für alle Cluster:
dscontrol advisor start http 88
- Stoppen der Advisor-Funktion http am Port 80 für Cluster 127.40.50.1:
dscontrol
advisor stop http 127.40.50.1:80
- Festlegen der Zeit (30 Sekunden), die eine HTTP-Advisor-Funktion für Port 80 wartet, bevor sie meldet, dass zu einem Server keine Verbindung hergestellt werden kann:
dscontrol advisor connecttimeout http 80 30
- Festlegen der Zeit (20 Sekunden), die eine HTTP-Advisor-Funktion für Port 80 des Clusters 127.40.50.1 wartet, bevor sie meldet, dass zu einem Server keine Verbindung hergestellt werden kann:
dscontrol advisor connecttimeout http 127.40.50.1:80 20
- Festlegen des Intervalls für die FTP-Advisor-Funktion (für Port 21) auf 6 Sekunden:
dscontrol advisor interval ftp 21 6
- Geben Sie den folgenden Befehl ein, um eine Liste der Advisor anzuzeigen, die derzeit Informationen an den Manager liefern:
dscontrol advisor list

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

ADVISOR	CLUSTER:PORT	ZEITLIMIT
http	127.40.50.1:80	unlimited
ftp	21	unlimited

- Ändern der Protokollstufe für das Advisor-Protokoll auf 0, um einen höheren Durchsatz zu erreichen:
dscontrol advisor loglevel http 80 0
- Ändern der Protokollgröße für die Advisor-Funktion ftp am Port 21 auf 5000 Bytes:
dscontrol advisor logsize ftp 21 5000
- Festlegen der Zeit (60 Sekunden), die eine HTTP-Advisor-Funktion (für Port 80) wartet, bevor sie meldet, dass von einem Server keine Daten empfangen werden können:
dscontrol advisor receivetimeout http 80 60
- Anzeigen eines Berichts zum Status der Advisor-Funktion ftp (für Port 21):
dscontrol advisor report ftp 21

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Advisor-Bericht:

Advisor-Name Ftp
Port-Nummer 21

Clusteradresse 9.67.131.18
Serveradresse 9.67.129.230
Last 8

Clusteradresse 9.67.131.18
Serveradresse 9.67.131.215
Last -1

- Anzeigen des aktuellen Status der Werte, die der Advisor-Funktion http für Port 80 zugeordnet sind:
dscontrol advisor status http 80

Dieser Befehl erzeugt eine Ausgabe, die ungefähr wie folgt aussieht:

Advisor-Status:

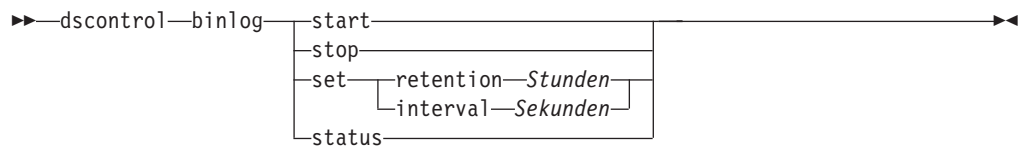
Intervall (Sekunden) 7
Zeitlimit (Sekunden) Unlimited
Zeitlimit für Verbindung (Sekunden) 21
Zeitlimit für Empfang (Sekunden) 21
Advisor-Protokolldateiname Http_80.log
Protokollstufe 1
Maximale Managerprotokollgröße (Bytes)... Unlimited
Anzahl Wiederholungen 0

- Festlegen des Zeitlimits für Informationen der Advisor-Funktion ftp am Port 21 auf 5 Sekunden:
dscontrol advisor timeout ftp 21 5
- Anzeigen der aktuellen Versionsnummer der Advisor-Funktion ssl für Port 443:
dscontrol advisor version ssl 443

Dieser Befehl erzeugt eine Ausgabe, die ungefähr wie folgt aussieht:

Version: 04.00.00.00 - 07/12/2001-10:09:56-EDT

dscontrol binlog — Binäre Protokolldatei steuern



start

Binäres Protokoll starten.

stop

Binäres Protokoll stoppen.

set

Legt Felder für die binäre Protokollierung fest. Weitere Informationen zum Festlegen von Feldern für die binäre Protokollierung finden Sie im Abschnitt „Binäre Protokolle für die Analyse von Serverstatistiken verwenden“ auf Seite 241.

retention

Die Anzahl der Stunden, die binäre Protokolldateien aufbewahrt werden. Der Standardwert für retention ist 24.

Stunden

Die Anzahl der Stunden.

interval

Die Anzahl der Sekunden zwischen dem Protokollieren von Einträgen. Der Standardwert für interval ist 60.

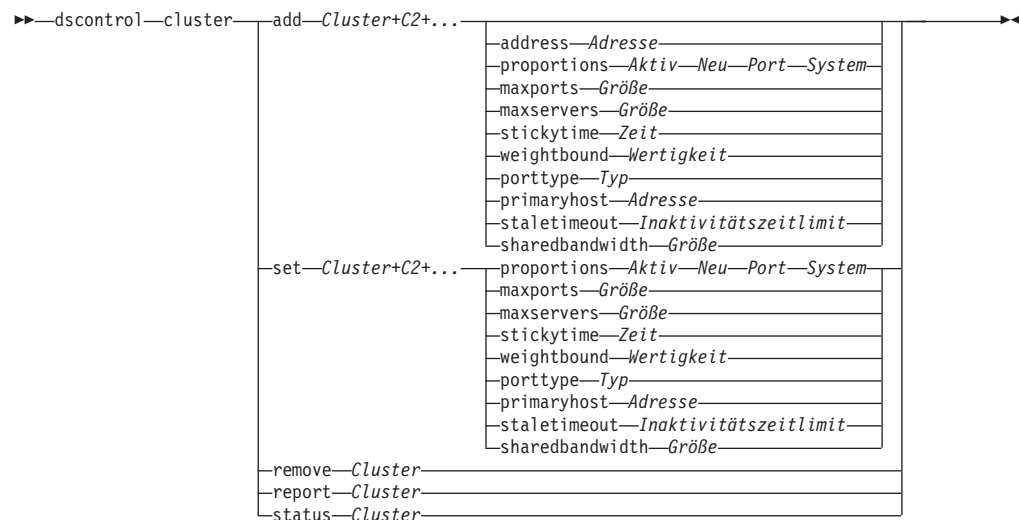
Sekunden

Die Anzahl der Sekunden.

status

Zeigt die Verweildauer und das Intervall des binären Protokolls.

dscontrol cluster — Cluster konfigurieren



add

Diesen Cluster hinzufügen. Sie müssen mindestens 1 Cluster definieren.

Cluster

Der Name oder die Adresse des Clusters, zu dem die Clients eine Verbindung herstellen. Der Clusterwert ist ein symbolischer Name oder eine IP-Adresse im IP-Adressenformat. Mit der Clusteradresse 0.0.0.0 kann ein Platzhaltercluster angegeben werden. Weitere Informationen hierzu finden Sie im Abschnitt „Platzhaltercluster zum Zusammenfassen von Serverkonfigurationen verwenden“ auf Seite 237.

Generell können Sie einen Doppelpunkt (:) als Platzhalter verwenden. Die einzige Ausnahme hiervon bildet der Befehl `dscontrol cluster add`. Der Befehl `dscontrol cluster set : weightbound 80` bewirkt beispielsweise, dass für alle Cluster eine Wertigkeit von 80 festgelegt wird.

Anmerkung: Zusätzliche Cluster werden durch ein Pluszeichen (+) getrennt.

address

Die eindeutige IP-Adresse der TCP-Maschine als Hostname oder im IP-Adressenformat. Falls der Clusterwert nicht aufgelöst werden kann, müssen Sie diese Adresse der physischen Maschine angeben.

Anmerkung: Die Adresse gilt nur für die Komponente Dispatcher.

Adresse

Wert für die Adresse des Clusters.

proportions

Legt auf Clusterebene die proportionale Bedeutung von aktiven Verbindungen (*Aktiv*), von neuen Verbindungen (*Neu*), von Informationen der Advisor-Funktionen (*Port*) und von Informationen eines Systemüberwachungsprogramms wie Metric Server (*System*), anhand derer der Manager Serverwertigkeiten festlegt. Alle diese Werte, die nachfolgend beschrieben werden, werden als Prozentsatz der Summe angegeben und müssen daher immer 100 ergeben. Weitere Informationen hierzu finden Sie im Abschnitt „Proportionale Bedeutung von Statusinformationen“ auf Seite 180.

Aktiv

Eine Zahl von 0 bis 100 für die proportionale Gewichtung der aktiven Verbindungen. Der Standardwert ist 50.

Neu

Eine Zahl von 0 bis 100 für die proportionale Gewichtung von neuen Verbindungen. Der Standardwert ist 50.

Port

Eine Zahl von 0 bis 100 für die proportionale Gewichtung der Informationen von Advisor-Funktionen. Der Standardwert ist 0.

Anmerkung: Wenn eine Advisor-Funktion gestartet wird und die Port-Proportion 0 ist, setzt Load Balancer diesen Wert automatisch auf 1, damit der Manager die Informationen der Advisor-Funktion als Vorgabe für die Berechnung der Serverwertigkeit verwendet.

System

Eine Zahl von 0-100, die die proportionale Gewichtung der Systemmesswerte von einem Programm wie Metric Server. Der Standardwert ist 0.

maxports

Die maximale Port-Anzahl. Der Standardwert für maxports ist 8.

Größe

Die zulässige Port-Anzahl.

maxservers

Die standardmäßige Höchstzahl von Servern pro Port. Dieser Wert kann für einzelne Ports mit **port maxservers** überschrieben werden. Der Standardwert für maxservers ist 32.

Größe

Die zulässige Anzahl von Servern für einen Port.

stickytime

Die Standardhaltezeit für Ports, die erstellt werden sollen. Dieser Wert kann für einzelne Ports mit dem Befehl **port stickytime** außer Kraft gesetzt werden. Der Standardwert für stickytime ist 0.

Anmerkung: Für die Dispatcher-Weiterleitungsmethode cbr gilt: Wenn Sie stickytime festlegen (einen Wert ungleich null angeben), ist port stickytime aktiviert, sofern der Port SSL (und nicht HTTP) verwendet. Wenn für zu erstellende Ports der Wert für stickytime ungleich null ist, wird beim Hinzufügen eines neuen SSL-Ports die Affinität der SSL-IDs für den Port aktiviert. Sie können die Affinität der SSL-IDs für den Port inaktivieren, indem Sie stickytime für den Port explizit auf 0 setzen.

Zeit

Der Wert für stickytime in Sekunden.

weightbound

Die standardmäßige Wertigkeitsgrenze für Ports. Dieser Wert kann für einzelne Ports mit dem Befehl **port weightbound** außer Kraft gesetzt werden. Der Standardwert für weightbound ist 20.

Wertigkeit

Die Wertigkeitsgrenze.

porttype

Der Standard-Port-Typ. Dieser Wert kann für einzelne Ports mit **port porttype** überschrieben werden.

Typ

Gültige Werte sind **tcp**, **udp** und **both**.

primaryhost

Die NFA dieser Dispatcher-Maschine oder die NFA-Adresse der Dispatcher-Partnermaschine. In einer Konfiguration mit gegenseitiger hoher Verfügbarkeit ist ein Cluster entweder der primären Maschine oder der Ausweichmaschine zugeordnet.

Wird der primäre Host (primaryhost) eines Clusters geändert, nachdem die primäre Maschine und die Partnermaschine gestartet wurden, und ist die gegenseitige hohe Verfügbarkeit aktiv, müssen Sie auch den neuen primären Host zur Übernahme zwingen. Außerdem müssen Sie die Scripts aktualisieren und den Cluster manuell aus der Konfiguration entfernen und dann richtig konfigurieren. Weitere Informationen hierzu finden Sie im Abschnitt „Gegenseitige hohe Verfügbarkeit“ auf Seite 61.

Adresse

Der Wert für die Adresse des primären Hosts. Der Standardwert ist die NFA-Adresse dieser Maschine.

staletimeout

Die Zeit der Inaktivität einer Verbindung in Sekunden, bevor die Verbindung entfernt wird. Der Standardwert für FTP ist 900 und für Telnet 32.000.000. Für alle anderen Protokolle liegt der Standardwert bei 300. Dieser Wert kann für einzelne Ports mit dem Befehl **port staletimeout** außer Kraft gesetzt werden. Weitere Informationen hierzu finden Sie im Abschnitt „Inaktivitätszeitlimit verwenden“ auf Seite 270.

Inaktivitätszeitlimit

Der Wert für staletimeout.

sharedbandwidth

Die maximale Bandbreite (in Kilobytes pro Sekunde), die auf Clusterebene gemeinsam genutzt werden kann. Weitere Informationen zur gemeinsam genutzten Bandbreite finden Sie in den Abschnitten „Regeln auf der Basis der reservierten und gemeinsam benutzten Bandbreite verwenden“ auf Seite 216 und „Regel "Gemeinsame Bandbreite"“ auf Seite 217.

Anmerkung: "shared bandwidth" gilt für die Komponente Dispatcher.

Größe

Der Parameter **sharedbandwidth** muss einen ganzzahligen Wert haben. Der Standardwert ist null. Bei einem Wert von null kann keine Bandbreite auf Clusterebene gemeinsam genutzt werden.

set

Die Merkmale des Clusters festlegen.

remove

Diesen Cluster entfernen.

report

Die internen Felder des Clusters anzeigen.

Anmerkung: "report" gilt für die Komponente Dispatcher.

status

Den aktuellen Status eines bestimmten Clusters anzeigen.

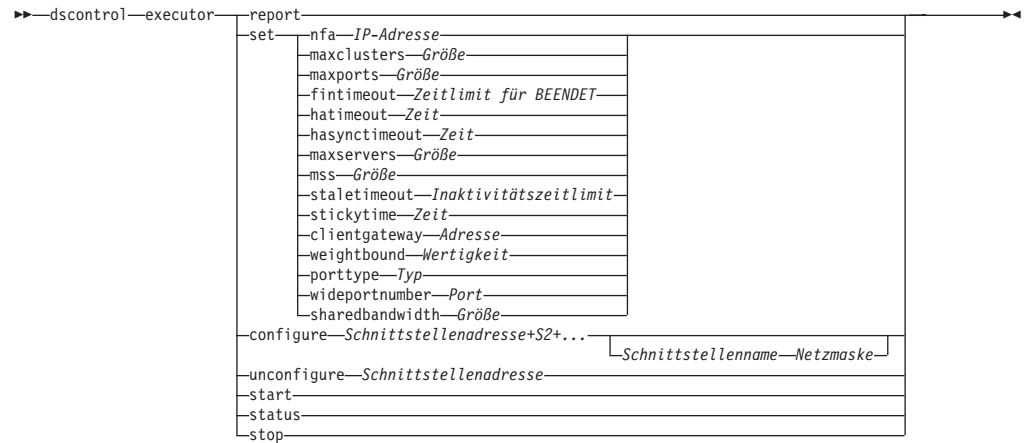
Beispiele

- Geben Sie den folgenden Befehl ein, um die Clusteradresse 130.40.52.153 hinzuzufügen:
`dscontrol cluster add 130.40.52.153`
- Geben Sie den folgenden Befehl ein, um die Clusteradresse 130.40.52.153 zu entfernen:
`dscontrol cluster remove 130.40.52.153`
- Festlegen der relativen Bedeutung von Vorgaben (Aktiv, Neu, Port, System), die vom Manager für Server des Clusters 9.6.54.12 empfangen werden:
`dscontrol cluster set 9.6.54.12 proportions 60 35 5 0`
- Geben Sie den folgenden Befehl ein, um einen Platzhaltercluster hinzuzufügen:
`dscontrol cluster add 0.0.0.0`
- Geben Sie den folgenden Befehl ein, um in einer Konfiguration mit gegenseitiger hoher Verfügbarkeit die Clusteradresse 9.6.54.12 mit der NFA der Partnermaschine (9.65.70.19) als primären Host zu definieren:
`dscontrol cluster set 9.6.54.12 primaryhost 9.65.70.19`
- Geben Sie den folgenden Befehl ein, um den Status für Clusteradresse 9.67.131.167 anzuzeigen:
`dscontrol cluster status 9.67.131.167`

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
Clusterstatus:
-----
Cluster ..... 9.67.131.167
Adresse ..... 9.67.131.167
Anzahl Ziel-Ports ..... 3
Standardhaltezeit ..... 0
Strd.-Zeitlimit für Inaktivität ..... 30
Strd.-Gewichtungsgrenze für Port ..... 20
Max. Anzahl Ports ..... 8
Strd.-Port-Protokoll ..... tcp/udp
Strd. max. Anzahl Server ..... 32
Proportion für aktive Verbindungen ..... 0.5
Proportion für neue Verbindungen ..... 0.5
Port-spezifische Proportion ..... 0
Proportion für Systemmetrik ..... 0
Gemeinsame Bandbreite (KBytes) ..... 0
Adresse des primären Hosts ..... 9.67.131.167
```

dscontrol executor — Executor steuern



report

Zeigt eine statistische Momentaufnahme an, z. B. die Gesamtanzahl der empfangenen, gelöschten oder mit Fehlern weitergeleiteten Pakete usw.

Anmerkung: "report" gilt für die Komponente Dispatcher.

set

Die Felder des Executors festlegen.

nfa

NFA definieren. Alle an diese Adresse gesendeten Pakete werden von der Dispatcher-Maschine nicht weitergeleitet.

Anmerkung: NFA gilt für die Komponente Dispatcher.

IP-Adresse

Die Internet-Protocol-Adresse als symbolischer Name oder in Schreibweise mit Trennzeichen.

maxclusters

Die maximale Anzahl Cluster, die konfiguriert werden können. Der Standardwert für maxclusters ist 100.

Größe

Die maximale Anzahl Cluster, die konfiguriert werden können.

maxports

Der Standardwert für maxports für Cluster, die erstellt werden sollen. Dieser Wert kann mit dem Befehl **cluster set** oder **cluster add** überschrieben werden. Der Standardwert für maxports ist 8.

Größe

Die Port-Anzahl.

fintimeout

Die Anzahl Sekunden, die eine Verbindung im Speicher verbleiben soll, nachdem die Verbindung in den Status BEENDET gesetzt wurde. Der Standardwert für fintimeout ist 60.

Zeitlimit für BEENDET

Der Wert für fintimeout.

Anmerkung: `fintimeout` gilt für die Komponente Dispatcher.

hatimeout

Die Zeit in Sekunden, die der Executor als Zeitlimit für die Überwachungssignale für hohe Verfügbarkeit verwendet. Der Standardwert ist 2.

Anmerkung: Der `hatimeout`-Wert gilt nur für die Komponente Dispatcher.

Zeit

Der Wert für `hatimeout`.

hasynctimeout

Die Zeit in Sekunden, die der Executor als Zeitlimit für die Replikation der Verbindungssätze zwischen der primären Maschine und der Ausweichmaschine verwendet. Der Standardwert ist 50.

Mit diesem Zeitgeber wird sichergestellt, dass die primäre Maschine und die Ausweichmaschine versuchen, ihre Daten zu synchronisieren. Wenn jedoch zu viele Verbindungen vorhanden sind und die aktive Maschine weiterhin einer signifikanten Belastung durch ankommenden Datenverkehr ausgesetzt ist, wird die Synchronisation möglicherweise nicht vor Ablauf des Zeitgebers beendet. Load Balancer würde in einem solchen Fall endlos versuchen, eine Synchronisation durchzuführen, ohne dass die Daten der beiden Maschinen tatsächlich synchronisiert werden. Setzen Sie `hasynctimeout` in dieser Situation auf einen höheren Wert als den Standardwert, damit die Maschinen genug Zeit haben, Informationen über vorhandene Verbindungen auszutauschen. Zum Setzen dieses Zeitgebers müssen Sie den Befehl `hasynctimeout` nach dem Befehl `dscontrol executor start`, aber vor den Befehlen für hohe Verfügbarkeit (`dscontrol highavailability`) absetzen.

Anmerkung: Der `hasynctimeout`-Wert gilt nur für die Komponente Dispatcher.

Zeit

Der Wert für `hasynctimeout`.

maxservers

Die standardmäßig geltende maximale Anzahl von Servern pro Port. Dieser Wert kann mit dem Befehl **cluster** oder **port** überschrieben werden. Der Standardwert für `maxservers` ist 32.

mss

Die maximale Bytezahl im Datensegment der TCP/UDP-Verbindung. Die Summe aus der Bytezahl im Datensegment und dem Header muss kleiner als die Bytezahl in der maximalen Übertragungseinheit (MTU) sein. Der Standardwert für `mss` ist 1460.

Anmerkung: Die maximale Segmentgröße ist nur für die Weiterleitungsmethoden `nat` und `cbr` der Komponente Dispatcher von Relevanz.

Größe

Die Anzahl Server.

staletimeout

Die Zeit der Inaktivität einer Verbindung in Sekunden, bevor die Verbindung entfernt wird. Der Standardwert für FTP ist 900 und für Telnet 32.000.000. Der Standardwert für alle anderen Ports ist 300. Dieser Wert kann mit dem Befehl **cluster** oder **port** überschrieben werden. Weitere Informationen hierzu finden Sie im Abschnitt „Inaktivitätszeitlimit verwenden“ auf Seite 270.

Inaktivitätszeitlimit

Der Wert für `staletimeout`.

stickytime

Die Standardhaltezeit für Ports für alle künftigen Cluster. Dieser Wert kann mit dem Befehl **cluster** oder **port** überschrieben werden. Der Standardwert für `stickytime` ist 0.

Zeit

Der Wert für `stickytime` in Sekunden.

clientgateway

Der Parameter `clientgateway` ist eine IP-Adresse, die für NAT/NAPT oder inhaltsabhängige Weiterleitung durch die Komponente Dispatcher verwendet wird. Er gibt die Router-Adresse an, über die der Antwortdatenverkehr von Load Balancer zu den Clients weitergeleitet wird. Der Parameter `clientgateway` muss auf einen Wert ungleich null gesetzt werden, bevor mit der Weiterleitungsmethode NAT/NAPT oder inhaltsabhängige Weiterleitung durch die Komponente Dispatcher ein neuer Port hinzugefügt wird. Weitere Informationen hierzu finden Sie in „Dispatcher-Weiterleitungsmethode nat“ auf Seite 53 und „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Anmerkung: Der Parameter `clientgateway` gilt nur für die Komponente Dispatcher.

Adresse

Die für den Parameter `clientgateway` angegebene Adresse ist ein symbolischer Name oder eine Adresse in Schreibweise mit Trennzeichen. Der Standardwert ist 0.0.0.0.

weightbound

Die standardmäßige Port-Wertigkeitsgrenze für alle künftigen Ports. Dieser Wert kann mit dem Befehl **cluster** oder **port** überschrieben werden. Der Standardwert für `weightbound` ist 20.

Wertigkeit

Der Wert für `weightbound`.

porttype

Der für alle künftigen Ports gültige Standardwert für Port-Typ. Dieser Wert kann mit dem Befehl **cluster** oder **port** überschrieben werden.

Anmerkung: „`porttype`“ gilt für die Komponente Dispatcher.

Typ

Gültige Werte sind **tcp**, **udp** und **both**.

wideportnumber

Ein nicht verwendeter TCP-Port auf jeder Dispatcher-Maschine. Die *wideportnumber* muss für alle Dispatcher-Maschinen identisch sein. Der Standardwert für `wideportnumber` ist 0. Dieser Wert gibt an, dass die Weitverkehrsunterstützung nicht verwendet wird.

Anmerkung: „`wideportnumber`“ gilt für die Komponente Dispatcher.

Port

Der Wert für **wideportnumber**.

sharedbandwidth

Die maximale Bandbreite (in Kilobytes pro Sekunde), die auf Executor-Ebene gemeinsam genutzt werden kann. Weitere Informationen zur gemeinsam genutzten Bandbreite finden Sie in den Abschnitten „Regeln auf der Basis der reservierten und gemeinsam benutzten Bandbreite verwenden“ auf Seite 216 und „Regel "Gemeinsame Bandbreite"“ auf Seite 217.

Anmerkung: "shared bandwidth" gilt für die Komponente Dispatcher.

Größe

Der Parameter **sharedbandwidth** muss einen ganzzahligen Wert haben. Der Standardwert ist null. Bei einem Wert von null kann keine Bandbreite auf Executor-Ebene gemeinsam genutzt werden.

configure

Konfigurieren einer Adresse (z. B. einer Clusteradresse, einer Rückkehradresse oder einer Adresse für die Überwachung der hohen Verfügbarkeit) für die Netzschnittstellenkarte der Dispatcher-Maschine. Dieser Schritt wird auch als das Konfigurieren eines Aliasnamens für die Dispatcher-Maschine bezeichnet.

Anmerkung: "configure" gilt für die Komponente Dispatcher.

Schnittstellenadresse

Die Adresse ist ein symbolischer Name oder eine Adresse im IP-Adressenformat.

Anmerkung: Weitere Schnittstellenadressen werden durch ein Pluszeichen (+) getrennt angegeben.

Schnittstellenname Netzmaske

Diese Angabe ist nur erforderlich, wenn die Adresse mit keinem Teilnetz für vorhandene Adressen übereinstimmt. Der *Schnittstellenname* kann ein Wert wie en0, eth1, eri0 sein. Die *Netzmaske* ist die 32-Bit-Maske, mit der die Teilnetzadressbits im Hostabschnitt einer IP-Adresse identifiziert werden.

unconfigure

Löscht die Aliasadresse von der Netzschnittstellenkarte.

Anmerkung: "unconfigure" gilt für die Komponente Dispatcher.

start

Den Executor starten.

status

Anzeigen des aktuellen Status für die im Executor definierbaren Werte und ihrer Standardeinstellungen.

stop

Stoppen des Executors.

Anmerkung: Der Parameter stop gilt für den Dispatcher und für CBR.

Beispiele

- Geben Sie den folgenden Befehl ein, um die internen Zähler für den Dispatcher anzuzeigen:

```
dscontrol executor status
```

```
Executor-Status:
```

```
-----
```

```
NFA ..... 9.67.131.151
Client-Gateway-Adresse ..... 0.0.0.0
Zeitlimit beend. inakt. Verbindungen .... 60
Port-Nr. für Weitverkehrsnetz ..... 0
Gemeinsame Bandbreite (KBytes) ..... 0
Strd. max. Ports pro Cluster ..... 8
Max. Anzahl Cluster ..... 100
Strd. max. Anzahl Server pro Port ..... 32
Strd.-Zeitlimit für Inaktivität ..... 300
Standardhaltezeit ..... 0
Strd.-Gewichtungsgrenze ..... 20
Standard-Port-Typ ..... tcp/udp
```

- Definieren der NFA von 130.40.52.167:

```
dscontrol executor set nfa 130.40.52.167
```
- Geben Sie den folgenden Befehl ein, um die maximale Anzahl von Clustern festzulegen:

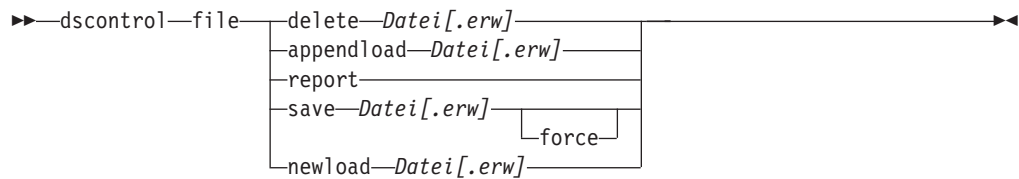
```
dscontrol executor set maxclusters 4096
```

- Starten des Executors:

```
dscontrol executor start
```
- Stoppen des Executors:

```
dscontrol executor stop
```

dscontrol file — Konfigurationsdateien verwalten



delete

Die Datei löschen.

Datei[.erw]

Eine Konfigurationsdatei mit dscontrol-Befehlen.

Die Dateierweiterung (.erw) kann vom Benutzer festgelegt oder übergangen werden.

appendload

Zum Aktualisieren der momentanen Konfiguration führt der Befehl appendload die Befehle in Ihrer Script-Datei aus.

report

Bericht über die verfügbare(n) Datei(en).

save

Sichern der aktuellen Konfiguration für Load Balancer in der Datei.

Anmerkung: Dateien werden in den nachfolgend genannten Verzeichnissen gespeichert und aus diesen geladen. Für *Komponente* gilt der Wert dispatcher oder cbr.

- Linux- und UNIX-Systeme: **/opt/ibm/edge/lb/servers/configurations/Komponente**
- Windows-Plattform: **C:\Programme\ibm\edge\lb\servers\configurations\Komponente**

force

Wenn Sie Ihre Datei in einer vorhandenen Datei mit demselben Namen speichern möchten, verwenden Sie **force**, um die vorhandene Datei vor dem Speichern der neuen Datei zu löschen. Bei Nichtverwendung der Option force wird die vorhandene Datei nicht überschrieben.

newload

Laden einer neuen Konfigurationsdatei in Load Balancer und ausführen derselben. Die neue Konfigurationsdatei ersetzt die aktuelle Konfiguration.

Beispiele

- Geben Sie den folgenden Befehl ein, um eine Datei zu löschen:
`dscontrol file delete Datei3`

Datei (Datei3) wurde gelöscht.
- Geben Sie den folgenden Befehl ein, um eine neue Konfigurationsdatei zu laden, die die aktuelle Konfiguration ersetzt:
`dscontrol file newload Datei1.sv`

Datei (Datei1.sv) wurde in den Dispatcher geladen.

- Geben Sie den folgenden Befehl ein, um eine Konfigurationsdatei an die aktuelle Konfiguration anzuhängen und zu laden:

```
dscontrol file appendload Datei2.sv
```

Datei (Datei2.sv) wurde an die aktuelle Konfiguration angehängt und geladen.

- Geben Sie den folgenden Befehl ein, um einen Bericht über Ihre Dateien anzuzeigen (die Dateien, die zuvor gesichert wurden):

```
dscontrol file report
```

DATEIBERICHT:

Datei1.save

Datei2.sv

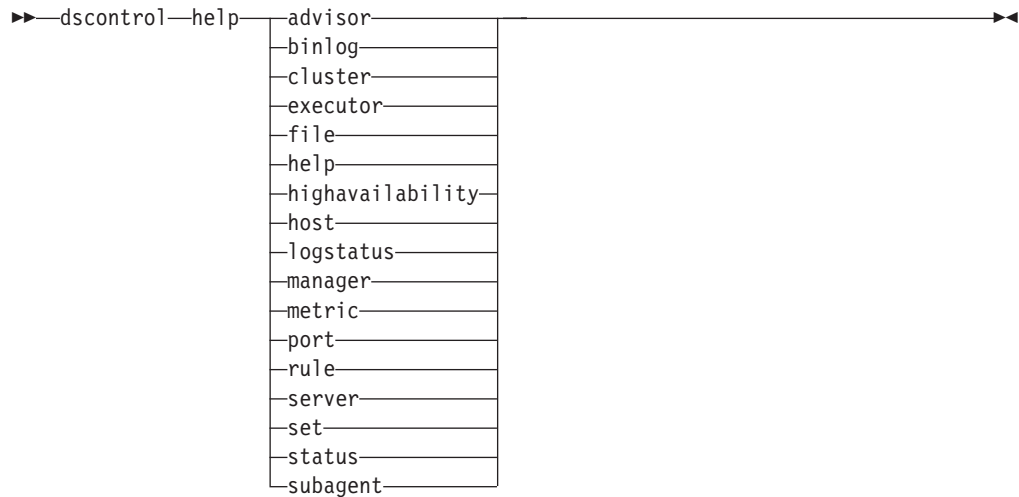
Datei3

- Geben Sie den folgenden Befehl ein, um die Konfiguration in der Datei Datei3 zu sichern:

```
dscontrol file save Datei3
```

Die Konfiguration wurde in Datei (Datei3) gesichert.

dscontrol help — Hilfetext für diesen Befehl anzeigen oder drucken



Beispiele

- Hilfetext zum Befehl dscontrol abrufen:

```
dscontrol help
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

ARGUMENTE BEFEHL HELP:

Verwendung: help <Hilfeoption>

Beispiel: help cluster

help	- vollständigen Hilfetext drucken
advisor	- Hilfe zum Befehl advisor
cluster	- Hilfe zum Befehl cluster
executor	- Hilfe zum Befehl executor
file	- Hilfe zum Befehl file
host	- Hilfe zum Befehl host
binlog	- Hilfe zum Befehl binlog
manager	- Hilfe zum Befehl manager
metric	- Hilfe zum Befehl metric
port	- Hilfe zum Befehl port
rule	- Hilfe zum Befehl rule
server	- Hilfe zum Befehl server
set	- Hilfe zum Befehl set
status	- Hilfe zum Befehl status
logstatus	- Hilfe zum Befehl logstatus
subagent	- Hilfe zum Befehl subagent
highavailability	- Hilfe zum Befehl highavailability

Parameter innerhalb von <> sind Variablen.

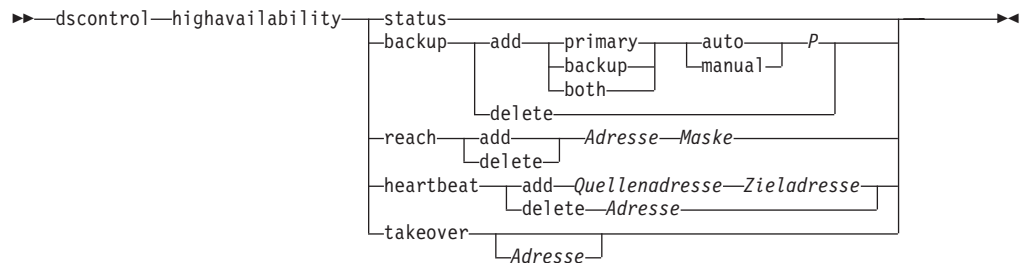
- Manchmal enthält der Hilfetext Optionen für die Variablen, die durch das Zeichen | voneinander getrennt sind:

```
fintimeout <Clusteradresse>|all <Zeit>
```

- Zeitlimit für beendete inaktive Verbindungen ändern
(Verwenden Sie 'all', um alle Cluster zu ändern)

dscontrol highavailability — Hohe Verfügbarkeit steuern

Anmerkung: Das Syntaxdiagramm zu dscontrol highavailability gilt nur für die Komponente Dispatcher.



status

Einen Bericht über die hohe Verfügbarkeit zurückgeben. Maschinen können eine von drei Statusbedingungen oder einen von drei Status haben:

Aktiv Eine bestimmte Maschine (primäre Maschine und/oder Partnermaschine) leitet Pakete weiter.

Bereitschaft

Eine bestimmte Maschine (primäre Maschine und/oder Partnermaschine) leitet keine Pakete weiter. Sie überwacht den Status eines **aktiven** Dispatchers.

Ruhend

Eine bestimmte Maschine leitet Pakete weiter und versucht nicht, Kontakt mit der Dispatcher-Partnermaschine aufzunehmen.

Darüber hinaus gibt das Schlüsselwort **status** Informationen über verschiedene untergeordnete Status zurück:

Synchronisiert

Eine bestimmte Maschine hat Kontakt mit einer anderen Dispatcher-Maschine aufgenommen.

Andere untergeordnete Status

Diese Maschine versucht, Kontakt zu ihrer Dispatcher-Partnermaschine aufzunehmen, die Kontaktaufnahme ist aber bisher nicht gelungen.

backup

Informationen über die primäre Maschine oder die Partnermaschine angeben.

add

Definiert die Funktionen der hohen Verfügbarkeit für diese Maschine und führt sie aus.

primary

Identifiziert die Dispatcher-Maschine, die die Rolle als *primäre Maschine* einnimmt.

backup

Identifiziert die Dispatcher-Maschine, die die Rolle als *Partnermaschine* einnimmt.

both

Identifiziert die Dispatcher-Maschine, die die Rolle als *primäre Maschine und Partnermaschine* einnimmt. Hierbei handelt es sich um die Funktion für gegen-

seitige hohe Verfügbarkeit, bei der die Rollen als primäre Maschine und als Partnermaschine auf der Basis einer Clustergruppe zugeordnet werden. Weitere Informationen hierzu finden Sie im Abschnitt „Gegenseitige hohe Verfügbarkeit“ auf Seite 61.

auto

Gibt eine *automatische* Wiederanlaufstrategie an, bei der die primäre Maschine das Weiterleiten von Paketen übernimmt, sobald sie wieder betriebsbereit ist.

manual

Gibt eine *manuelle* Wiederanlaufstrategie an, bei der die primäre Maschine das Weiterleiten von Paketen erst dann wieder übernimmt, wenn der Administrator den Befehl **takeover** ausgibt.

P[ort]

Ein auf beiden Maschinen nicht verwendeter TCP-Port für die Nachrichten zu den Dispatcher-Überwachungssignalen. Der *Port* muss für die primäre und die Ausweichmaschine identisch sein.

delete

Entfernt diese Maschine aus der hohen Verfügbarkeit, diese Maschine kann daher nicht mehr als Partnermaschine oder als primäre Maschine benutzt werden.

reach

Hinzufügen oder Löschen der Zieladresse für den primären Dispatcher und den Ausweich-Dispatcher. Die Advisor-Funktion reach sendet *pings* vom primären und Ausweich-Dispatcher, um die Erreichbarkeit ihrer Ziele festzustellen.

Anmerkung: Wenn Sie das Ziel für reach konfigurieren, müssen Sie auch die Advisor-Funktion reach starten. Die Advisor-Funktion reach wird automatisch von der Manager-Funktion gestartet.

add

Fügt eine Zieladresse zur Advisor-Funktion reach hinzu.

delete

Löscht eine Zieladresse aus der Advisor-Funktion reach.

Adresse

Die IP-Adresse (im IP-Adressenformat oder als symbolischer Name) des Zielknotens.

Maske

Eine Teilnetzmaske.

heartbeat

Definiert eine Übertragungssitzung zwischen der primären Dispatcher-Maschine und der Ausweichmaschine.

add

Teilt dem Quellen-Dispatcher die Adresse seines Partners (Zieladresse) mit.

Quellenadresse

Quellenadresse. Die Adresse (IP-Adresse oder symbolischer Name) dieser Dispatcher-Maschine.

Zieladresse

Zieladresse. Die Adresse (IP-Adresse oder symbolischer Name) der anderen Dispatcher-Maschine.

Anmerkung: Die Quellen- und die Zieladresse müssen für mindestens ein Überwachungssignalpaar die NFAs der Maschinen sein.

delete

Entfernt das Adressenpaar aus den Informationen zum Überwachungssignal. Sie können die Ziel- oder die Quellenadresse des Überwachungssignalpaares angeben.

Adresse

Die Adresse (IP-Adresse oder symbolischer Name); entweder die Ziel- oder die Quellenadresse.

takeover

Konfiguration mit einfacher Hochverfügbarkeit (Rolle der Dispatcher-Maschinen lautet entweder *primary* oder *backup*):

- Takeover weist einen Dispatcher in Bereitschaft an, aktiv zu werden und mit dem Weiterleiten von Paketen zu beginnen. Damit wird der gegenwärtig aktive Dispatcher in Bereitschaft versetzt. Der Befehl takeover muss auf der Maschine in Bereitschaft ausgegeben werden. Er wird nur ausgeführt, wenn die Strategie **manual** lautet. Der untergeordnete Status muss *synchronisiert* lauten.

Konfiguration mit gegenseitiger hoher Verfügbarkeit (Rolle jeder Dispatcher-Maschine lautet *beide*):

- Die Dispatcher-Maschine mit der Funktion für gegenseitige hohe Verfügbarkeit enthält zwei Cluster, die denen ihres Partners entsprechen. Einer der Cluster wird als primärer Cluster (Partnercluster der Partnermaschine) und der andere als Partnercluster (primärer Cluster der Partnermaschine) betrachtet. Takeover weist die Dispatcher-Maschine an, mit dem Weiterleiten von Paketen für den oder die Cluster der anderen Maschine zu beginnen. Der Befehl takeover kann nur ausgegeben werden, wenn der oder die Cluster der Dispatcher-Maschine den Status *Bereitschaft* haben und der untergeordnete Status *synchronisiert* lautet. Damit werden die gegenwärtig aktiven Cluster der Partnermaschine in den Bereitschaftsstatus geändert. Der Befehl takeover wird nur ausgeführt, wenn die Strategie **manual** lautet. Weitere Informationen hierzu finden Sie im Abschnitt „Gegenseitige hohe Verfügbarkeit“ auf Seite 61.

Anmerkungen:

1. Beachten Sie, dass sich die *Rollen* der Maschinen (*primary*, *backup*, *both*) nicht ändern. Es ändert sich lediglich der relative *Status* (*Aktiv* oder *Bereitschaft*).
2. Es gibt drei mögliche takeover-Scripts: *goActive*, *goStandby* und *goInOp*. Lesen Sie hierzu die Informationen im Abschnitt „Scripts verwenden“ auf Seite 209.

Adresse

Der Wert für die Übernahmeadresse ist optional. Er sollte nur verwendet werden, wenn die Maschine die Rolle als *primäre Maschine und Partnermaschine* einnimmt (Konfiguration mit gegenseitiger hoher Verfügbarkeit). Die angegebene Adresse ist die NFA der Dispatcher-Maschine, die normalerweise den Datenverkehr dieses Clusters weiterleitet. Erfolgt eine Übernahme beider Cluster, geben Sie die eigene NFA-Adresse des Dispatchers an.

Beispiele

- Geben Sie den folgenden Befehl ein, um den Hochverfügbarkeitsstatus einer Maschine zu überprüfen:
`dscontrol highavailability status`
Ausgabe:
Hochverfügbarkeitsstatus:

Rolle primary
Wiederanlaufstrategie manual
Status Aktiv
Untergeordneter Status Synchronisiert
Primärer Host 9.67.131.151
Port 12345
Bevorzugtes Ziel 9.67.134.223

Signalstatus:

Anzahl 1
Quelle/Ziel 9.67.131.151/9.67.134.223

Erreichbark.-Status:

Anzahl 1
Adresse 9.67.131.1 erreichbar
- Hinzufügen der Sicherungsinformationen zur primären Maschine unter Verwendung der automatischen Wiederherstellungsstrategie und von Port 80:
`dscontrol highavailability backup add primary auto 80`
- Geben Sie den folgenden Befehl ein, um eine Adresse hinzuzufügen, die der Dispatcher erreichen muss:
`dscontrol highavailability reach add 9.67.125.18`
- Geben Sie die folgenden Befehle ein, um Überwachungssignalinformationen für die primäre Maschine und Partnermaschine hinzuzufügen:
Primäre Maschine - `highavailability heartbeat add 9.67.111.3 9.67.186.8`
Partnermaschine - `highavailability heartbeat add 9.67.186.8 9.67.111.3`
- Geben Sie den folgenden Befehl ein, wenn der Dispatcher in Bereitschaft angewiesen werden soll, aktiv zu werden, und die aktive Maschine in Bereitschaft versetzt werden soll:
`dscontrol highavailability takeover`

dscontrol host — Ferne Maschine konfigurieren

►►—dscontrol—host:—*ferner_Host*—◄◄

ferner_Host

Der Name der fernen Maschine mit Load Balancer, die konfiguriert wird. Stellen Sie bei der Eingabe dieses Befehls sicher, dass sich zwischen **host:** und *ferner_Host* kein Leerzeichen befindet. Beispiel:

dscontrol host:*ferner_Host*

Nachdem dieser Befehl in der Eingabeaufforderung ausgegeben wurde, geben Sie einen beliebigen gültigen Befehl dscontrol ein, der für die ferne Load-Balancer-Maschine abgesetzt werden soll.

dscontrol logstatus — Protokolleinstellungen des Servers anzeigen

►►—dscontrol—logstatus—◄◄

logstatus

Zeigt die Einstellungen des Serverprotokolls (Name der Protokolldatei, Protokollstufe und -größe) an.

Beispiele

Geben Sie den folgenden Befehl ein, um den Protokollstatus anzuzeigen:

```
dscontrol logstatus
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

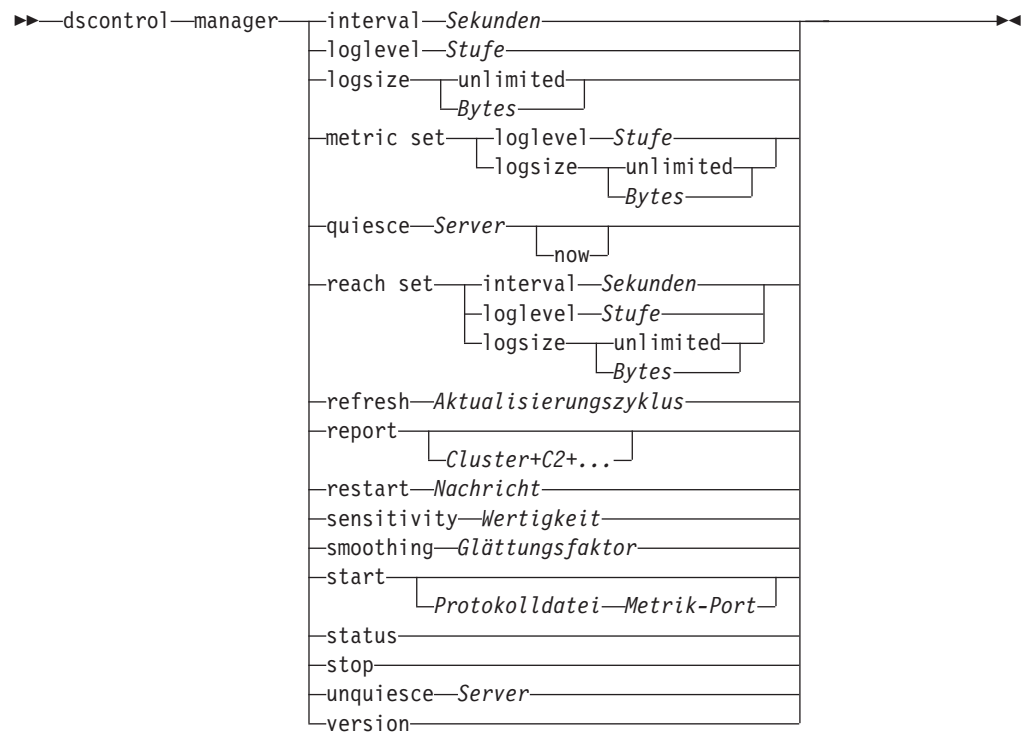
Protokollstatus von Dispatcher:

Name der Protokolldatei C:\PROGRA~1\IBM\edge\lb\servers
 \logs\dispatcher\server.log

Protokollstufe 1

Maximale Protokollgröße (Bytes) ... 1048576

dscontrol manager — Manager steuern



interval

Legt fest, wie oft der Manager die Wertigkeit der Server für den Executor aktualisiert. Dabei werden die Kriterien aktualisiert, die der Executor für die Weiterleitung von Client-Anforderungen verwendet.

Sekunden

Eine positive Zahl, die in Sekunden darstellt, wie oft der Manager Wertigkeiten für den Executor aktualisiert. Der Standardwert ist 2.

loglevel

Legt die Protokollstufe für das Protokoll des Managers fest.

Stufe

Die Nummer der Stufe (0 bis 5). Je größer die Zahl, desto mehr Informationen werden in das Manager-Protokoll geschrieben. Der Standardwert ist 1. Die folgenden Werte sind gültig: 0 für keine Einträge, 1 für eine minimale Protokollierung, 2 für eine Basisprotokollierung, 3 für eine normale, 4 für eine erweiterte und 5 für eine ausführliche Protokollierung.

logsize

Legt die maximale Größe des Protokolls des Managers fest. Wenn Sie eine maximale Größe für die Protokolldatei festlegen, findet ein Dateiumlauf statt. Hat die Datei die angegebene Größe erreicht, werden alle weiteren Einträge wieder an den Anfang der Datei geschrieben und die dort befindlichen Einträge überschrieben. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge werden mit einer Zeitmarke versehen, damit Sie erkennen können, in welcher Reihenfolge die Einträge geschrieben wurden. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Bytes

Die maximale Größe in Byte für die Protokolldatei des Managers. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** (unbegrenzt) angeben. Möglicherweise erreicht die Protokolldatei nicht genau die maximale Größe, bevor der Dateiumlauf stattfindet, da die Größe der Protokolleinträge variiert. Der Standardwert ist 1 MB.

metric set

Legt die Werte **loglevel** und **logsize** für das Metriküberwachungsprotokoll fest. Der Wert für loglevel ist die Stufe für die Protokollierung der Metriküberwachung (0 - Keine, 1 - Minimal, 2 - Grundlegend, 3 - Mäßig, 4 - Erweitert oder 5 - Ausführlich). Die Standardprotokollstufe ist 1. Der Wert für logsize ist die Datenmenge (in Bytes), die maximal in der Protokolldatei für Metriküberwachung erfasst wird. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder unlimited (unbegrenzt) angeben. Die Standardprotokollgröße ist 1 MB.

quiesce

Es werden keine weiteren Verbindungen an einen Server gesendet. Hiervon ausgenommen sind nur nachfolgende neue Verbindungen vom Client zum stillgelegten Server, sofern diese als sticky markiert sind und die Haltezeit (stickytime) nicht abgelaufen ist. Der Manager setzt die Wertigkeit für diesen Server an jedem Port, für den er definiert ist, auf 0. Diesen Befehl verwenden, wenn auf einem Server eine schnelle Wartung erfolgen soll und der Server anschließend wieder aktiviert werden soll. Wenn Sie einen stillgelegten Server aus der Konfiguration löschen und ihn der Konfiguration anschließend wieder hinzufügen, behält er nicht seinen Status, den er vor der Stilllegung hatte. Weitere Informationen hierzu finden Sie im Abschnitt „Bearbeitung von Serververbindungen stilllegen“ auf Seite 224.

Server

Die IP-Adresse des Servers als symbolischer Name oder in Schreibweise mit Trennzeichen.

Wenn Sie die Serverpartitionierung verwenden, geben Sie den eindeutigen Namen des logischen Servers an. Weitere Informationen hierzu finden Sie im Abschnitt „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58.

now

Verwenden Sie quiesce “now” nur, wenn Sie die Haltezeit definiert haben und vor Ablauf der Haltezeit neue Verbindungen an einen anderen als den stillgelegten Server gesendet werden sollen. Weitere Informationen hierzu finden Sie im Abschnitt „Bearbeitung von Serververbindungen stilllegen“ auf Seite 224.

reach set

Legt das Intervall, die Protokollstufe und die Protokollgröße für die Advisor-Funktion reach fest.

refresh

Legt die Anzahl von Intervallen fest, nach denen der Manager die Informationen über neue und aktive Verbindungen für den Executor aktualisiert.

Aktualisierungszyklus

Eine positive Zahl, die die Anzahl von Intervallen darstellt. Der Standardwert ist 2.

report

Zeigt eine statistische Momentaufnahme an.

Cluster

Die Adresse des Clusters, die im Bericht angezeigt werden soll. Die Adresse kann ein symbolischer Name oder eine Adresse im IP-Adressenformat sein. Der Standardwert ist ein Manager-Bericht, in dem alle Cluster angezeigt werden.

Anmerkung: Zusätzliche Cluster werden durch ein Pluszeichen (+) getrennt.

restart

Startet alle Server (die nicht inaktiv sind) mit der Standardwertigkeit (1/2 der maximalen Wertigkeit).

Nachricht

Eine Nachricht, die in die Protokolldatei des Managers gestellt werden soll.

sensitivity

Legt die Mindestsensitivität für die Aktualisierung von Wertigkeiten fest. Diese Einstellung definiert, wann der Manager seine Serverwertigkeit ausgehend von externen Informationen ändern sollte.

Wertigkeit

Eine Zahl von 1 bis 100, die als prozentuale Wertigkeit verwendet werden soll. Der Standardwert 5 bewirkt eine Mindestsensitivität von 5 %.

smoothing

Festlegen eines Faktors, der Wertigkeitsabweichungen während des Lastausgleichs glättet. Ein höherer Glättungsfaktor führt zu einer weniger drastischen Änderung von Serverwertigkeiten bei Änderungen an den Netzbedingungen. Ein geringerer Glättungsfaktor führt zu einer drastischeren Änderung von Serverwertigkeiten.

Glättungsfaktor

Eine positive Gleitkommazahl. Der Standardwert ist 1,5.

start

Den Manager starten.

Protokolldatei

Der Name der Datei, in der die Daten des Managers protokolliert werden. Jeder Eintrag im Protokoll wird mit einer Zeitmarke versehen.

Die Standarddatei ist im Verzeichnis **logs** installiert. Lesen Sie hierzu die Informationen in Anhang C, „Beispielkonfigurationsdateien“, auf Seite 481. Wollen Sie das Verzeichnis ändern, in dem die Protokolldateien gespeichert werden, lesen Sie die Informationen im Abschnitt „Pfade für die Protokolldatei ändern“ auf Seite 269.

Metrik-Port

Der von Metric Server für Meldungen zur Systembelastung verwendete Port. Wenn Sie einen Metrik-Port angeben, müssen Sie auch einen Protokoll-dateinamen angeben. Der Standard-Metrik-Port ist 10004.

status

Zeigt den aktuellen Status aller Werte in dem Manager an, die global gesetzt werden können. Zudem werden die Standardwerte dieser Werte angezeigt.

stop

Den Manager stoppen.

unquiesce

Festlegung, dass der Manager einem zuvor stillgelegten Server an jedem Port, für den er definiert ist, eine Wertigkeit größer als null zuordnen kann.

Server

Die IP-Adresse des Servers als symbolischer Name oder in Schreibweise mit Trennzeichen.

version

Zeigt die aktuelle Version des Managers an.

Beispiele

- Geben Sie den folgenden Befehl ein, um das Aktualisierungsintervall für den Manager auf 5 Sekunden zu setzen:
`dscontrol manager interval 5`
- Geben Sie den folgenden Befehl ein, um die Stufe der Protokollierung zwecks Verbesserung der Leistung auf 0 zu setzen:
`dscontrol manager loglevel 0`
- Geben Sie den folgenden Befehl ein, um die Größe des Protokolls des Managers auf 1.000.000 Byte zu setzen:
`dscontrol manager logsize 1000000`
- Festlegung, dass keine weiteren Verbindungen an den Server 130.40.52.153 gesendet werden sollen:
`dscontrol manager quiesce 130.40.52.153`
- Setzen der Anzahl Aktualisierungsintervalle auf 3, bevor die Wertigkeiten aktualisiert werden:
`dscontrol manager refresh 3`
- Geben Sie den folgenden Befehl ein, um eine statistische Momentaufnahme des Managers abzurufen:
`dscontrol manager report`

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

SERVER		IP-ADRESSE		STATUS	
mach14.dmz.com		10.6.21.14		AKTIV	
mach15.dmz.com		10.6.21.15		AKTIV	

LEGENDE ZUM MANAGERBERICHT	
AKTV	Aktive Verbindungen
NEUV	Neue Verbindungen
SYS	Systemmetrik
JETZT	Aktuelle Gewichtung
NEU	Neue Gewichtung
GWT	Gewichtung
VERB	Verbindungen

www.dmz.com	GEWICHTUNG		AKTIV	NEU	PORT	SYS
10.6.21.100	JETZT NEU		49 %	50 %	1 %	0 %
PORT: 21						

mach14.dmz.com	10	10	0	0	-1	0
mach15.dmz.com	10	10	0	0	-1	0

www.dmz.com	GEWICHTUNG		AKTIV	NEU	PORT	SYS
10.6.21.100	JETZT NEU		49 %	50 %	1 %	0 %
PORT: 80						

mach14.dmz.com	10	10	0	0	23	0
mach15.dmz.com	9	9	0	0	30	0

ADVISOR	CLUSTER:PORT	ZEITLIMIT
http	80	unlimited
ftp	21	unlimited

- Neustart aller Server mit Standardwertigkeit und Schreiben einer Nachricht in die Manager-Protokolldatei:

`dscontrol manager restart` Neustart des Managers für Codeaktualisierung

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

320-14:04:54 Neustart des Managers für Codeaktualisierung

- Setzen der Sensitivität für Wertigkeitsänderungen auf 10:

`dscontrol manager sensitivity 10`

- Geben Sie den folgenden Befehl ein, um den Glättungsfaktor auf 2,0 zu setzen:

`dscontrol manager smoothing 2.0`

- Geben Sie den folgenden Befehl ein, um den Manager zu starten und die Protokolldatei `ndmgr.log` anzugeben (Pfade können nicht angegeben werden):

`dscontrol manager start ndmgr.log`

- Geben Sie den folgenden Befehl ein, um den aktuellen Status der Werte anzuzeigen, die dem Manager zugeordnet sind:

`dscontrol manager status`

Dieser Befehl erzeugt eine Ausgabe, die ungefähr wie folgt aussieht:

Manager-Status:

=====

```
Metrik-Port ..... 10004
Name der Managerprotokolldatei ..... manager.log
Managerprotokollstufe ..... 1
Max. Managerprotokollgröße (Bytes) ..... unlimited
Sensitivitätsstufe ..... 0,05
Glättungsfaktor ..... 1,5
Aktualisierungsintervall (Sekunden) ..... 2
Gewichtungsaktualisierungszyklus ..... 2
Erreichbarkeit - Protokollstufe ..... 1
Erreichbarkeit - Max. Protokollgröße (Bytes) ..... unlimited
Erreichbarkeit - Aktualisierungsintervall (Sekunden) .... 7
Name der Protokolldatei für Metriküberwachung ..... MetricMonitor.log
Protokollstufe für Metriküberwachung ..... 1
Maximale Protokollgröße für Metriküberwachung (Bytes) ... 1048576
```

- Stoppen des Managers:

```
dscontrol manager stop
```

- Festlegung, dass keine neuen Verbindungen an einen Server mit der Adresse 130.40.52.153 gesendet werden sollen (Anmerkung: Verwenden Sie zum Stilllegen des Servers nur die Option "now", wenn Sie die Haltezeit festgelegt haben und neue Verbindungen bis zum Ablauf der Haltezeit an einen anderen Server gesendet werden sollen):

```
dscontrol manager quiesce 130.40.52.153 now
```

- Festlegung, dass keine neuen Verbindungen an einen Server mit der Adresse 130.40.52.153 gesendet werden sollen (Anmerkung: Wenn Sie die Haltezeit definiert haben, werden nachfolgende neue Verbindungen vom Client bis zum Ablauf der Haltezeit an diesen Server gesendet):

```
dscontrol manager quiesce 130.40.52.153
```

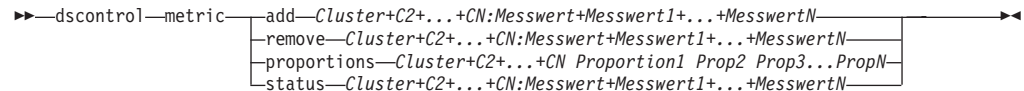
- Festlegung, dass der Manager einem zuvor stillgelegten Server im Cluster 130.40.52.153 eine Wertigkeit größer als 0 zuordnen kann:

```
dscontrol manager unquiesce 130.40.52.153
```

- Geben Sie den folgenden Befehl ein, um die aktuelle Versionsnummer des Managers aufzurufen:

```
dscontrol manager version
```

dscontrol metric — Systemmesswerte konfigurieren



add

Hinzufügen des angegebenen Messwerts.

Cluster

Die Adresse, zu der die Clients eine Verbindung herstellen. Die Adresse kann der Hostname der Maschine oder die IP-Adresse im IP-Adressenformat sein. Zusätzliche Cluster werden durch ein Pluszeichen (+) getrennt.

Messwert

Name des Systemmesswerts. Es muss sich um den Namen einer ausführbaren Datei oder Script-Datei im Verzeichnis des Messwertservers handeln.

remove

Entfernen des angegebenen Messwerts.

proportions

Festlegen der Proportionen für alle diesem Objekt zugeordneten Messwerte.

status

Anzeigen des aktuellen Messwerts.

Beispiele

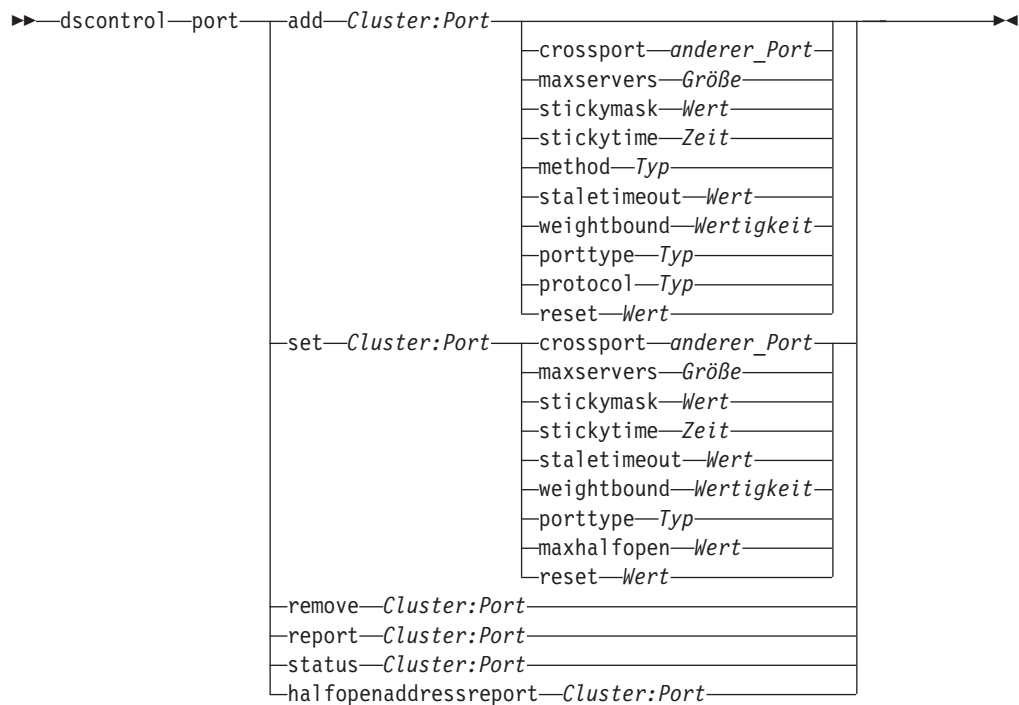
- Hinzufügen eines Systemmesswerts:
`dscontrol metric add Site1:Messwert1`
- Festlegen der Proportionen für einen Sitenamen mit zwei Systemmesswerten:
`dscontrol metric proportions Site1 0 100`
- Anzeigen des aktuellen Status der zugeordneten Messwerte:
`dscontrol metric status Site1:Messwert1`

Dieser Befehl erzeugt eine Ausgabe, die ungefähr wie folgt aussieht:

Metrikstatus:

```
Cluster ..... 10.10.10.20
Metrikname ..... Messwert1
Metrikproportion ..... 50
  Server ..... plm3
  Metrikdaten ..... -1
```


dscontrol port — Ports konfigurieren



add

Hinzufügen eines Ports zu einem Cluster. Sie müssen einen Port zu einem Cluster hinzufügen, bevor Sie Server zu diesem Port hinzufügen können. Sind keine Ports für einen Cluster vorhanden, werden alle Client-Anforderungen lokal verarbeitet. Mit diesem Befehl können Sie mehrere Ports auf einmal hinzufügen.

Cluster

Die Adresse des Clusters als symbolischer Name oder im IP-Adressenformat. Sie können einen Doppelpunkt (:) als Platzhalter verwenden. Der Befehl `dscontrol port add :80` bewirkt beispielsweise, dass Port 80 zu allen Clustern hinzugefügt wird.

Anmerkung: Zusätzliche Cluster werden durch ein Pluszeichen (+) getrennt.

Port

Nummer des Ports. Mit der Port-Nummer 0 (null) kann ein Platzhalter-Port angegeben werden.

Anmerkung: Zusätzliche Ports werden durch ein Pluszeichen (+) getrennt.

crossport

Mit `crossport` können Sie die Affinität (Merkmal *sticky*) auf mehrere Ports ausdehnen, so dass künftige Anforderungen von Clients, deren Anforderungen an verschiedenen Ports empfangen werden, dennoch an einen Server gesendet werden können. Geben Sie als Wert für `crossport` die Nummer des *anderen* Ports an, der in die Port-übergreifende Affinität einbezogen werden soll. Die Ports müssen die folgenden Bedingungen erfüllen, um diese Funktion verwenden zu können:

- Sie müssen dieselbe Clusteradresse gemeinsam benutzen.
- Sie müssen dieselben Server gemeinsam benutzen.

- Sie müssen denselben Wert (ungleich null) für stickytime haben.
- Sie müssen denselben Wert für stickymask haben.

Wenn Sie die Port-übergreifende Affinität aufheben möchten, setzen Sie den Wert für crossport auf seine eigene Port-Nummer zurück. Weitere Informationen zur Port-übergreifenden Affinität finden Sie im Abschnitt „Port-übergreifende Affinität“ auf Seite 223.

Anmerkung: "crossport" gilt nur für die Dispatcher-Weiterleitungsmethoden MAC und NAT/NATP.

anderer_Port

Der Wert von crossport. Der Standardwert entspricht der eigenen Port-Nummer.

maxservers

Die maximale Anzahl von Servern. Der Standardwert für maxservers ist 32.

Größe

Der Wert für maxservers.

stickymask

Mit der Affinitätsadressmaske werden eingehende Client-Anforderungen auf der Basis gemeinsamer Teilnetzadressen zusammengefasst. Wenn eine Client-Anforderung eine Verbindung zu dem Port hergestellt hat, werden alle nachfolgenden Anforderungen von Clients mit derselben (durch den maskierten Abschnitt der IP-Adresse angegebenen) Teilnetzadresse an denselben Server übertragen. Wenn Sie stickymask aktivieren möchten, muss port stickytime ein Wert ungleich null sein. Weitere Informationen hierzu finden Sie im Abschnitt „Affinitätsadressmaske (stickymask)“ auf Seite 223.

Anmerkung: Das Schlüsselwort stickymask gilt nur für die Komponente Dispatcher.

Wert

Der Wert für stickymask ist die Anzahl der höherwertigen Bits der 32-Bit-IP-Adresse, die maskiert werden sollen. Gültige Werte sind: 8, 16, 24 und 32. Der Standardwert ist 32. Damit wird die Funktion der Affinitätsadressmaske inaktiviert.

stickytime

Das Intervall zwischen dem Schließen einer Verbindung und dem Öffnen einer neuen Verbindung. Innerhalb dieses Intervalls wird der Client an denselben Server wie bei der ersten Verbindung vermittelt. Nach Ablauf der Haltezeit kann der Client an einen anderen Server vermittelt werden.

Für die Komponente Dispatcher:

- Für die Dispatcher-Weiterleitungsmethode cbr
 - Sie können stickytime nur für einen SSL-Port festlegen (auf einen Wert ungleich null setzen) und nicht für einen HTTP-Port, weil dadurch die SSL-ID-Affinität aktiviert wird.
 - Wenn Sie die Haltezeit (stickytime) für den Port setzen, muss der Affinitätstyp der Regel auf den Standard (none) gesetzt sein. Die regelbasierte Affinität (passive Cookie-Affinität und URI-Affinität) kann nicht gleichzeitig festgelegt werden, wenn die Haltezeit für den Port gesetzt ist.
- Für die Dispatcher-Weiterleitungsmethoden mac und nat

- Wenn Sie die Haltezeit (stickytime) für den Port (auf einen Wert ungleich null) setzen, können Sie in der Regel keinen Affinitätstyp festlegen. Die regelbasierte Affinität kann nicht gleichzeitig festgelegt werden, wenn die Haltezeit für den Port gesetzt ist.
- Das Festlegen von port stickytime aktiviert die Affinität der IP-Adressen.

Für die Komponente CBR: Wenn Sie die Haltezeit (stickytime) für den Port auf einen Wert ungleich null setzen, muss der Affinitätstyp der Regel auf den Standard (none) gesetzt sein. Die regelbasierte Affinität (passive Cookie-Affinität, URI-Affinität, aktive Cookie-Affinität) kann nicht gleichzeitig festgelegt werden, wenn die Haltezeit für den Port gesetzt ist.

Zeit

Die Zeit in Sekunden für die Weiterleitung der Anforderungen eines Clients an denselben Server. Null gibt an, dass die Anforderungen eines Clients nicht an denselben Server weitergeleitet werden.

method

Weiterleitungsmethode. Es gibt die Weiterleitungsmethoden mac und nat und die inhaltsabhängige Weiterleitung (cbr). Die Weiterleitungsmethode nat oder die inhaltsabhängige Weiterleitung (cbr) können Sie *erst* hinzufügen, nachdem Sie mit dem Parameter clientgateway des Befehls dscontrol executor eine IP-Adresse ungleich null angegeben haben. Weitere Informationen hierzu finden Sie in „Dispatcher-Weiterleitungsmethode nat“ auf Seite 53 und „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Anmerkungen:

1. Die Methode gilt nur für die Komponente Dispatcher.
2. Wenn sich der Back-End-Server im selben Teilnetz wie die Rückkehradresse befindet und Sie die Weiterleitungsmethode cbr oder nat verwenden, müssen Sie als Router-Adresse die Adresse des Back-End-Servers definieren.
3. Fügen Sie eine mac-Weiterleitungsmethode hinzu, werden Sie aufgefordert, für den Parameter protocol entweder HTTP oder SSL anzugeben.

Typ

Der Typ der Weiterleitungsmethode. Mögliche Werte sind mac, nat oder cbr. Der Standardwert ist mac.

staletimeout

Die Zeit der Inaktivität einer Verbindung in Sekunden, bevor die Verbindung entfernt wird. Für die Komponente Dispatcher lautet der Standardwert 900 für Port 21 (FTP) und 32.000.000 für Port 23 (Telnet). Für alle anderen Dispatcher-Ports und alle CBR-Ports ist der Standardwert 300. Das Inaktivitätszeitlimit kann auch auf Executor- oder Clusterebene gesetzt werden. Weitere Informationen hierzu finden Sie im Abschnitt „Inaktivitätszeitlimit verwenden“ auf Seite 270.

Wert

Der Wert für **staletimeout** in Sekunden.

weightbound

Legt die maximale Wertigkeit von Servern an diesem Port fest. Mit diesem Wert wird die Differenz festgelegt, die hinsichtlich der Anzahl der Anforderungen, die der Executor den einzelnen Servern zuordnet, gelten soll. Der Standardwert ist 20.

Wertigkeit

Eine Zahl von 1 bis 100 für die maximale Wertigkeitsgrenze.

porttype

Der Port-Typ.

Anmerkung: Der Parameter porttype gilt nur für Dispatcher.

Typ

Gültige Werte sind **tcp**, **udp** und **both**. Der Standardwert ist both (tcp/udp).

protocol

Der Protokolltyp. Für die Komponente Dispatcher ist dies ein erforderlicher Parameter, sofern Sie für den Port eine cbr-Methode angeben. Wenn Sie für einen Port den Protokolltyp **SSL** auswählen, sollten Sie außerdem eine Haltezeit (stickytime) ungleich null angeben, um die SSL-ID-Affinität zu aktivieren. Bei Auswahl des Protokolls **HTTP** können Sie mit content-Regeln die Serveraffinität konfigurieren. Weitere Informationen hierzu finden Sie im Abschnitt „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55.

Anmerkung: "protocol" gilt nur für die Weiterleitungsmethode cbr von Dispatcher.

Typ

Gültige Werte sind **HTTP** und **SSL**.

maxhalfopen

Der Schwellenwert für das Maximum halboffener Verbindungen. Verwenden Sie diesen Parameter, um mögliche DoS-Attacken festzustellen, die eine große Anzahl halboffener TCP-Verbindungen auf Servern nach sich ziehen.

Ein positiver Wert gibt an, dass überprüft wird, ob die aktuelle Anzahl halboffener Verbindungen den Schwellenwert überschreitet. Wenn der aktuelle Wert über dem Schwellenwert liegt, wird ein Alert-Script aufgerufen. Weitere Informationen finden Sie im Abschnitt „Erkennung von DoS-Attacken“ auf Seite 240.

Anmerkung: Der Parameter maxhalfopen gilt nur für Dispatcher.

Wert

Der Wert für maxhalfopen. Der Standardwert ist null (es findet keine Überprüfung statt).

reset

Mit reset können Sie angeben, ob Load Balancer an inaktive Server am Port eine TCP-Rücksetzanforderung senden soll. Eine TCP-Rücksetzanforderung bewirkt eine sofortige Beendigung der Verbindung. Weitere Informationen hierzu finden Sie im Abschnitt „TCP-Rücksetzanforderung an einen inaktiven Server senden (nur Komponente Dispatcher)“ auf Seite 182.

Anmerkung: Das Schlüsselwort reset gilt nur für die Komponente Dispatcher. Wenn Sie das Schlüsselwort reset verwenden möchten, muss der Parameter clientgateway des Befehls dscontrol executor auf eine Router-Adresse gesetzt sein.

Wert

Gültige Werte für reset sind yes und no. Der Standardwert ist no (es wird keine TCP-Rücksetzanforderung an inaktive Server gesendet). Wenn reset auf yes gesetzt ist, wird eine TCP-Rücksetzanforderung an inaktive Server gesendet.

set

Festlegen der Felder eines Ports.

remove

Entfernen dieses Ports.

report

Bericht zu diesem Port.

status

Anzeigen des Status für den Server an diesem Port. Wenn Sie den Status für alle Ports sehen möchten, geben Sie diesen Befehl ohne *Port* an. Vergessen Sie jedoch nicht den Doppelpunkt.

Sekunden

Die Zeit in Sekunden, nach der halboffene Verbindungen zurückgesetzt werden.

halfopenaddressreport

Generiert für alle Client-Adressen (bis zu 8000 Adresspaare), deren Serverzugriff halboffene Verbindungen zur Folge hatten, Einträge im Protokoll (halfOpen.log). Außerdem werden statistische Daten an die Befehlszeile zurückgegeben. Dazu gehören unter anderem die Gesamtzahl, die größte Anzahl und die durchschnittliche Anzahl halboffener Verbindungen sowie die durchschnittliche Dauer halboffener Verbindungen (in Sekunden). Weitere Informationen finden Sie im Abschnitt „Erkennung von DoS-Attacken“ auf Seite 240.

Beispiele

- Geben Sie den folgenden Befehl ein, um die Ports 80 und 23 zur Clusteradresse 130.40.52.153 hinzuzufügen:
`dscontrol port add 130.40.52.153:80+23`
- Geben Sie den folgenden Befehl ein, um einen Platzhalter-Port zur Clusteradresse 130.40.52.153 hinzuzufügen:
`dscontrol port set 130.40.52.153:0`
- Festlegen der maximalen Wertigkeit 10 für Port 80 an der Clusteradresse 130.40.52.153:
`dscontrol port set 130.40.52.153:80 weightbound 10`
- Geben Sie den folgenden Befehl ein, um den Wert für stickytime für die Ports 80 und 23 der Clusteradresse 130.40.52.153 auf 60 Sekunden zu setzen:
`dscontrol port set 130.40.52.153:80+23 stickytime 60`
- Geben Sie den folgenden Befehl ein, um die Port-übergreifende Affinität von Port 80 zu Port 23 für die Clusteradresse 130.40.52.153 zu setzen:
`dscontrol port set 130.40.52.153:80 crossport 23`
- Entfernen von Port 23 aus dem Cluster mit der Adresse 130.40.52.153:
`dscontrol port remove 130.40.52.153:23`
- Abrufen des Status für Port 80 an der Clusteradresse 9.67.131.153:
`dscontrol port status 9.67.131.153:80`

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Port-Status:

```
Port-Nummer ..... 80
Cluster ..... 9.67.131.153
Zeitlimit für Inaktivität ..... 300
Gewichtungsgrenze ..... 20
Max. Anzahl Server ..... 32
Haltezeit ..... 0
Port-Typ ..... tcp/udp
Port-übergreifende Affinität ..... 80
StickyBits der Maske ..... 32
Max. Anz. halb geöffneter Verb. ... 0
TCP-Rücksetzanforderungen senden ... no
```

- Abrufen des Berichts für Port 80 an der Clusteradresse 9.62.130.157:

```
dscontrol port report 9.62.130.157:80
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Port-Bericht:

```
Clusteradresse ..... 9.62.130.157
Port-Nummer ..... 80
Anzahl Server ..... 5
Max. Servergewichtung ..... 10
Summe aktiver Verbindungen ..... 55
Verbindungen pro Sekunde ..... 12
KBytes pro Sekunde ..... 298
Anzahl halbgeöffneter Verbindungen ... 0
Gesendete TCP-Rücksetzanforderungen .. 0
Weiterleitungsmethode ..... MAC-gestützte Weiterleitung
```

- Abrufen des Berichts zu Adressen mit halboffenen Verbindungen für Port 80 an der Clusteradresse 9.67.127.121:

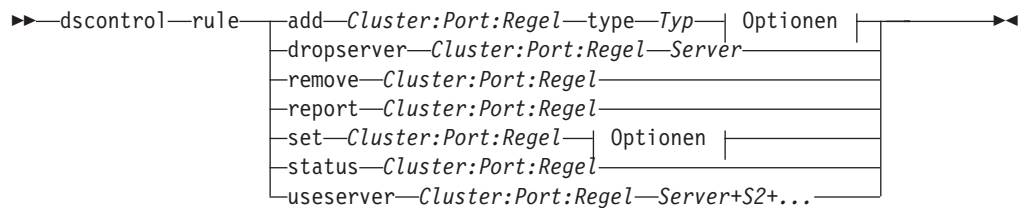
```
dscontrol port halfopenaddressreport 9.67.127.121:80
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Bericht zu halboffenen Verbindungen wurde erfolgreich erstellt.

```
Adressenbericht zu halb geöffneten Verbindungen für Cluster:Port = 9.67.127.121:80
Summe Adressen mit halb geöffneten Verbindungen ..... 0
Gesamtanzahl halb geöffneten Verbindungen ..... 0
Größte Anzahl halb geöffneten Verbindungen ..... 0
Durchschnittliche Zahl halb geöffneten Verbindungen ... 0
Durchschnittl. Zeit für halb geöffnete Verb. (Sek.) ... 0
Summe empfangener halb geöffneten Verbindungen ..... 0
```

dscontrol rule — Regeln konfigurieren



Optionen:

beginrange	Niedrig	endrange	Hoch
priority	Stufe		
pattern	Muster		
tos	Wert		
stickytime	Zeit		
affinity	Affinitätstyp		
cookie name	Wert		
evaluate	Stufe		
sharelevel	Stufe		

add

Diese Regel zu einem Port hinzufügen.

Cluster

Die Adresse des Clusters als symbolischer Name oder im IP-Adressenformat. Sie können einen Doppelpunkt (:) als Platzhalter verwenden. Der Befehl `dscontrol rule add :80:RegelA type Typ` bewirkt beispielsweise, dass RegelA für alle Cluster zu Port 80 hinzugefügt wird.

Anmerkung: Zusätzliche Cluster werden durch ein Pluszeichen (+) getrennt.

Port

Nummer des Ports. Sie können einen Doppelpunkt (:) als Platzhalter verwenden. Der Befehl `dscontrol rule add ClusterA::RegelA type Typ` bewirkt beispielsweise, dass RegelA zu allen Ports für ClusterA hinzugefügt wird.

Anmerkung: Zusätzliche Ports werden durch ein Pluszeichen (+) getrennt.

Regel

Der für die Regel ausgewählte Name. Dieser Name kann eine beliebige Kombination aus alphanumerischen Zeichen, Unterstreichungszeichen, Silbentrennungsstrichen und Punkten sein. Der Name kann 1 bis 20 Zeichen lang sein und darf keine Leerzeichen enthalten.

Anmerkung: Zusätzliche Regeln werden durch ein Pluszeichen (+) voneinander getrennt.

type

Der Regeltyp.

Typ

Die Auswahlmöglichkeiten für *Typ* sind:

ip Die Regel basiert auf der Client-IP-Adresse.

time Die Regel basiert auf der Uhrzeit.

connection

Die Regel basiert auf der Anzahl der Verbindungen pro Sekunde für den Port. Diese Regel kann nur verwendet werden, wenn der Manager aktiv ist.

active Die Regel basiert auf der Gesamtzahl der aktiven Verbindungen für den Port. Diese Regel kann nur verwendet werden, wenn der Manager aktiv ist.

port Die Regel basiert auf dem Client-Port.

Anmerkung: "port" gilt für die Komponente Dispatcher.

service

Diese Regel basiert auf dem Feld für die Service-Art (Type of Service = TOS) im IP-Header.

Anmerkung: Service gilt nur für die Komponente Dispatcher.

reservedbandwidth

Diese Regel basiert auf der Bandbreite (in Kilobytes pro Sekunde), die von einer Servergruppe bereitgestellt wird. Weitere Informationen finden Sie unter „Regeln auf der Basis der reservierten und gemeinsam benutzten Bandbreite verwenden“ auf Seite 216 und „Regel "Reservierte Bandbreite"“ auf Seite 216.

Anmerkung: Der Parameter reservedbandwidth gilt nur für die Komponente Dispatcher.

sharedbandwidth

Diese Regel basiert auf der Bandbreite (in Kilobytes pro Sekunde), die auf Executor- oder Clusterebene gemeinsam genutzt wird. Weitere Informationen finden Sie unter „Regeln auf der Basis der reservierten und gemeinsam benutzten Bandbreite verwenden“ auf Seite 216 und „Regel "Gemeinsame Bandbreite"“ auf Seite 217.

Anmerkung: Der Parameter sharedbandwidth gilt nur für die Komponente Dispatcher.

true Diese Regel ist immer wahr. Sie kann mit der Anweisung ELSE in der Programmierlogik verglichen werden.

content

Diese Regel beschreibt einen regulären Ausdruck, der mit den URLs verglichen wird, die vom Client angefordert werden. Dies gilt für Dispatcher und CBR.

beginrange

Der untere Wert des Bereichs, mit dem bestimmt wird, ob die Regel wahr ist oder nicht.

Niedrig

Hängt von der Art der Regel ab. Die Art des Werts und der Standardwert werden nachfolgend nach der Art der Regel aufgelistet:

ip Die Adresse des Clients als symbolischer Name oder im IP-Adressenformat. Der Standardwert ist 0.0.0.0.

Zeit Eine ganze Zahl. Der Standardwert ist 0. Er stellt Mitternacht dar.

Verbindung

Eine ganze Zahl. Der Standardwert ist 0.

aktiv Eine ganze Zahl. Der Standardwert ist 0.

Port Eine ganze Zahl. Der Standardwert ist 0.

Reservierte Bandbreite

Eine ganze Zahl (Kilobytes pro Sekunde). Der Standardwert ist 0.

endrange

Der obere Wert des Bereichs, mit dem bestimmt wird, ob die Regel wahr ist oder nicht.

Hoch

Hängt von der Art der Regel ab. Die Art des Werts und der Standardwert werden nachfolgend nach der Art der Regel aufgelistet:

ip Die Adresse des Clients als symbolischer Name oder im IP-Adressenformat. Der Standardwert ist 255.255.255.254.

Zeit Eine ganze Zahl. Der Standardwert ist 24. Er stellt Mitternacht dar.

Anmerkung: Beim Definieren des Bereichsbeginns (*beginrange*) und Bereichsendes (*endrange*) der Zeitintervalle ist darauf zu achten, dass jeder Wert eine ganze Zahl sein muss, die nur den Stundenteil der Uhrzeit darstellt. Es werden keine Teilwerte einer Stunde angegeben. Soll beispielsweise die Stunde von 3 Uhr bis 4 Uhr angegeben werden, geben Sie für *beginrange* 3 und für *endrange* ebenfalls 3 an. Damit werden alle Minuten von 3 Uhr bis 3 Uhr 59 angegeben. Wird für *beginrange* 3 und für *endrange* 4 angegeben, wird die zweistündige Periode von 3 Uhr bis 4 Uhr 59 angegeben.

Verbindung

Eine ganze Zahl. Der Standardwert ist 2 hoch 32 minus 1.

aktiv Eine ganze Zahl. Der Standardwert ist 2 hoch 32 minus 1.

Port Eine ganze Zahl. Der Standardwert ist 65535.

Reservierte Bandbreite

Eine ganze Zahl (Kilobytes pro Sekunde). Der Standardwert ist 2 hoch 32 minus 1.

Priorität

Die Reihenfolge, in der die Regeln überprüft werden.

Stufe

Eine ganze Zahl. Wird die Priorität der ersten hinzugefügten Regel nicht angegeben, wird sie vom Dispatcher standardmäßig auf 1 gesetzt. Wird eine nachfolgende Regel hinzugefügt, wird ihre Priorität standardmäßig als 10 + der derzeit niedrigsten Priorität aller vorhandenen Regeln berechnet. Beispiel: Sie haben eine Regel mit der Priorität 30. Sie fügen eine neue Regel hinzu und setzen ihre Priorität auf 25 (die, wie Sie wissen, eine *höhere* Priorität als 30 ist). Anschließend fügen Sie eine dritte Regel ohne Angabe einer Priorität hinzu. Die Priorität der dritten Regel wird wie folgt berechnet: 40 (30 + 10).

pattern

Gibt das Muster an, das für eine Regel der des Typs content verwendet werden soll.

Muster

Das zu verwendende Muster. Weitere Informationen zu gültigen Werten finden Sie in Anhang B, „Syntax der content-Regel“, auf Seite 477.

tos

Gibt den Wert für "Type of Service" (TOS) an, der für die Regel der Art **Service** verwendet wird.

Anmerkung: TOS gilt nur für die Komponente Dispatcher.

Wert

Die aus 8 Zeichen bestehende Zeichenfolge, die für den tos-Wert verwendet werden soll. Gültige Zeichen sind: 0 (binäre Null), 1 (binäre Eins) und x (beliebig). Beispiel: 0xx1010x. Weitere Informationen hierzu finden Sie im Abschnitt „Auf der Serviceart basierende Regeln verwenden“ auf Seite 215.

stickytime

Gibt die für eine Regel zu verwendende Haltezeit an. Wenn der Parameter affinity des Befehls rule auf activecookie gesetzt wird, muss stickytime auf einen Wert ungleich null gesetzt werden, um diesen Affinitätstyp zu aktivieren. Die Haltezeit für die Regel gilt nicht für Regeln mit dem Affinitätstyp passivecookie oder uri.

Weitere Informationen hierzu finden Sie im Abschnitt „Aktive Cookie-Affinität“ auf Seite 225.

Anmerkung: Der Parameter stickytime für Regeln gilt nur für die Komponente CBR.

Zeit

Zeit in Sekunden.

affinity

Gibt den für eine Regel zu verwendenden Affinitätstyp an: Aktive Cookie-Affinität, passive Cookie-Affinität, URI-Affinität oder Keine.

Der Affinitätstyp activecookie aktiviert eine Lastverteilung des Webdatenverkehrs mit Affinität an einen Server. Die Affinität basiert auf Cookies, die von Load Balancer generiert werden.

Der Affinitätstyp passivecookie aktiviert die Verteilung von Webdatenverkehr mit Affinität zu einem Server ausgehend von den Identifizierungs-Cookies, die von den Servern generiert werden. Für die passive Cookie-Affinität müssen Sie den Parameter cookiename verwenden.

Der Affinitätstyp "URI" aktiviert den Lastausgleich für Webdatenverkehr auf Caching-Proxy-Servern mit effektiver Vergrößerung des Cache.

Weitere Informationen hierzu finden Sie in den Abschnitten „Aktive Cookie-Affinität“ auf Seite 225, „Passive Cookie-Affinität“ auf Seite 227 und „URI-Affinität“ auf Seite 228.

Anmerkung: Die Affinität gilt für Regeln, die mit der Dispatcher-Weiterleitungsmethode cbr konfiguriert wurden, und für die Komponente CBR.

Affinitätstyp

Mögliche Werte für den Affinitätstyp sind: Keine (Standardwert), Aktives Cookie, Passives Cookie oder URI.

cookiename

Ein vom Administrator willkürlich festgelegter Name, der als Kennung für Load Balancer verwendet wird. Nach diesem Namen muss Load Balancer die HTTP-Header-Anforderung durchsuchen. Der Cookie-Name dient neben dem Cookie-Wert als Kennung für Load Balancer, so dass Load Balancer nachfol-

gende Anforderungen einer Website immer an dieselbe Servermaschine senden kann. Der Cookie-Name kann nur für die Affinität "Passives Cookie" angewendet werden.

Weitere Informationen finden Sie im Abschnitt „Passive Cookie-Affinität“ auf Seite 227.

Anmerkung: Der Cookie-Name gilt für Regeln, die mit der Dispatcher-Weiterleitungsmethode cbr konfiguriert wurden, und für die Komponente CBR.

Wert

Wert des Cookie-Namens.

evaluate

Diese Option ist nur für die Komponente Dispatcher verfügbar. Sie gibt an, ob die Regelbedingungen für alle Server an einem Port oder für alle Server in einer Regel ausgewertet werden sollen. Diese Option ist nur für Regeln gültig, die Entscheidungen ausgehend von den Kenndaten der Server treffen. Dazu gehören die Regeln "Aktive Verbindungen" und "Reservierte Bandbreite". Weitere Informationen hierzu finden Sie im Abschnitt „Regeloption für Serverauswertung“ auf Seite 220.

Für Regeln des Typs connection können Sie auch eine Auswertungsoption (upserversonrule) angeben. Durch Angabe von upserversonrule können Sie sicherstellen, dass die übrigen Server, die dieser Regel unterliegen, nicht überlastet werden, wenn einige Server der Gruppe inaktiv sind.

Stufe

Mögliche Werte sind port, rule oder upserversonrule. Der Standardwert ist port. upserversonrule ist nur für den Regeltyp connection verfügbar.

sharelevel

Dieser Parameter gilt nur für die Regel "Gemeinsam genutzte Bandbreite". Er gibt an, ob die gemeinsame Nutzung von Bandbreite auf Cluster- oder Executor-Ebene stattfindet. Bei gemeinsamer Nutzung von Bandbreite auf Cluster-Ebene steht innerhalb eines Clusters Port-übergreifend eine maximale Bandbreite zur gemeinsamen Nutzung zur Verfügung. Bei gemeinsamer Nutzung von Bandbreite auf Executor-Ebene steht für Cluster innerhalb der gesamten Dispatcher-Konfiguration eine maximale Bandbreite zur gemeinsamen Nutzung zur Verfügung. Weitere Informationen hierzu finden Sie im Abschnitt „Regel "Gemeinsame Bandbreite"“ auf Seite 217.

Stufe

Mögliche Werte sind executor oder cluster.

dropserver

Einen Server aus einem Regelsatz entfernen.

Server

Die IP-Adresse der TCP-Servermaschine als symbolischer Name oder im IP-Adressenformat.

Wenn Sie die Serverpartitionierung verwenden, geben Sie den eindeutigen Namen des logischen Servers an. Weitere Informationen hierzu finden Sie im Abschnitt „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58.

Anmerkung: Zusätzliche Server werden durch ein Pluszeichen (+) voneinander getrennt.

remove

Eine oder mehrere Regeln, die durch Pluszeichen voneinander getrennt sind, entfernen.

report

Die internen Werte einer oder mehrerer Regeln anzeigen.

set

Werte für diese Regel festlegen.

status

Die einstellbaren Werte einer oder mehrerer Regeln anzeigen.

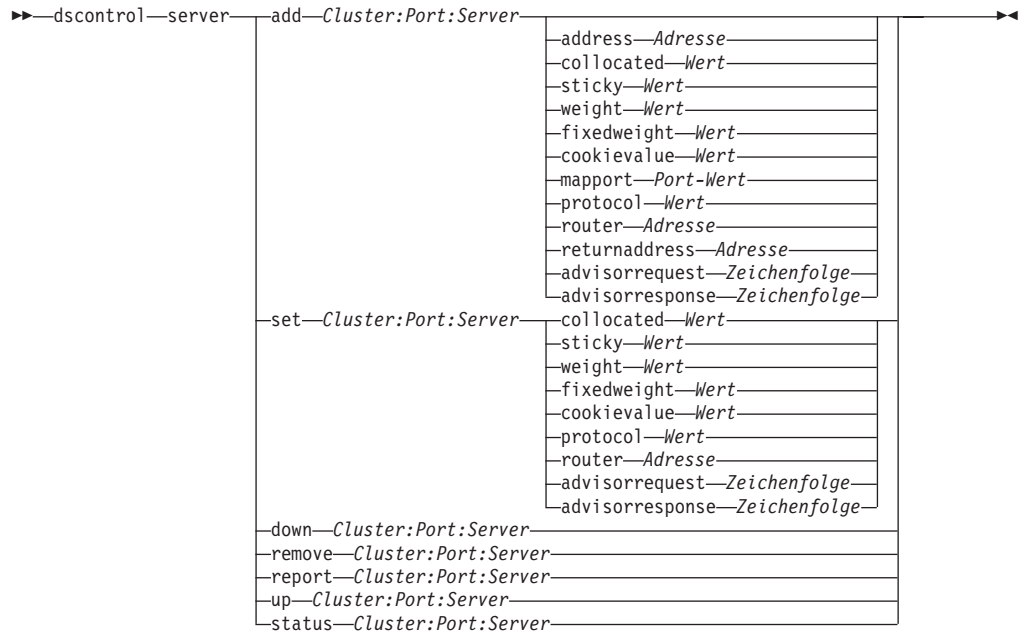
useserver

Server in einen Regelsatz einfügen.

Beispiele

- Hinzufügen einer immer gültigen Regel ohne Angabe von Bereichsanfang oder -ende:
`dscontrol rule add 9.37.67.100:80:trule type true priority 100`
- Erstellen einer Regel, die den Zugriff auf einen Bereich von IP-Adressen unterbindet, der in diesem Fall mit "9:" beginnt:
`dscontrol rule add 9.37.131.153:80:ni type ip b 9.0.0.0 e 9.255.255.255`
- Soll eine Regel erstellt werden, die die Verwendung eines bestimmten Servers in der Zeit von 11 Uhr bis 15 Uhr angibt, den folgenden Befehl eingeben:
`dscontrol rule add cluster1:80:timerule type time beginrange 11 endrange 14`
`dscontrol rule useserver Cluster1:80:timerule Server05`
- Soll eine Regel erstellt werden, die auf dem Inhalt des TOS-Bytefelds im IP-Header basiert, den folgenden Befehl eingeben:
`dscontrol rule add 9.67.131.153:80:tosrule type service tos 0xx1001x`
- Erstellen einer Regel ausgehend von der reservierten Bandbreite, die einer Gruppe von Servern (die innerhalb der Regel ausgewertet wird) zugeordnet wird, um Daten mit einer Geschwindigkeit von 100 Kilobytes pro Sekunde zu liefern:
`dscontrol rule add 9.67.131.153:80:rbwrule type reservedbandwidth`
`beginrange 0 endrange 100 evaluate rule`
- Erstellen einer Regel ausgehend von der gemeinsam genutzten Bandbreite, bei der es sich um verfügbar gemachte Bandbreite auf Clusterebene handelt, die nicht genutzt wurde (Anmerkung: Zuerst müssen Sie mit dem Befehl `dscontrol cluster` die maximale Bandbreite in Kilobytes pro Sekunde angeben, die auf Clusterebene gemeinsam genutzt werden kann):
`dscontrol cluster set 9.67.131.153 sharedbandwidth 200`
`dscontrol rule add 9.67.131.153:80:shbwrule type sharedbandwidth`
`sharelevel cluster`

dscontrol server — Server konfigurieren



add

Diesen Server hinzufügen.

Cluster

Die Adresse des Clusters als symbolischer Name oder im IP-Adressenformat. Sie können einen Doppelpunkt (:) als Platzhalter verwenden. Der Befehl `dscontrol server add :80:ServerA` bewirkt beispielsweise, dass ServerA für alle Cluster zu Port 80 hinzugefügt wird.

Anmerkung: Zusätzliche Cluster werden durch ein Pluszeichen (+) getrennt.

Port

Nummer des Ports. Sie können einen Doppelpunkt (:) als Platzhalter verwenden. Der Befehl `dscontrol server add ::ServerA` bewirkt beispielsweise, dass ServerA für alle Cluster zu allen Ports hinzugefügt wird.

Anmerkung: Zusätzliche Ports werden durch ein Pluszeichen (+) getrennt.

Server

Der **Server** ist die eindeutige IP-Adresse der TCP-Servermaschine als symbolischer Name oder im IP-Adressenformat.

Wenn Sie einen eindeutigen Namen verwenden, der nicht in eine IP-Adresse aufgelöst wird, müssen Sie den Befehl **dscontrol server add** mit dem Serverparameter **address** verwenden. Weitere Informationen hierzu finden Sie im Abschnitt „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58.

Anmerkung: Zusätzliche Server werden durch ein Pluszeichen (+) voneinander getrennt.

address

Die eindeutige IP-Adresse der TCP-Servermaschine als Hostname oder im IP-Adressenformat. Falls der Servername nicht aufgelöst werden kann, müssen

Sie die Adresse der physischen Servermaschine angeben. Weitere Informationen hierzu finden Sie im Abschnitt „Serverpartitionierung - Konfigurieren logischer Server für einen physischen Server (IP-Adresse)“ auf Seite 58.

Adresse

Wert für die Adresse des Servers.

collocated

Mit dem Parameter `collocated` können Sie angeben, ob der Dispatcher auf einer der Servermaschinen installiert ist, für die er den Lastausgleich durchführt.

Anmerkung: Der Parameter `collocated` ist gültig, wenn die Dispatcher-Weiterleitungsmethode `mac`, `nat` oder `cbr` verwendet wird. Site Selector und CBR können auf allen Plattformen verknüpft werden, erfordern jedoch nicht dieses Schlüsselwort. Weitere Informationen hierzu finden Sie im Abschnitt „Verknüpfte Server verwenden“ auf Seite 202.

Wert

Wert für `collocated`: ja oder nein. Der Standardwert ist Nein.

sticky

Ermöglicht einem Server, die Einstellung für `stickytime` an seinem Port zu überschreiben. Bei Verwendung des Standardwertes „yes“ bleibt die normale Affinität wie für den Port definiert erhalten. Bei Verwendung des Wertes „no“ wird der Client beim Absetzen der nächsten Anforderung an diesem Port *nicht* wieder an diesen Server verwiesen. Dies gilt unabhängig von der Einstellung für `stickytime` des Ports. Dies ist in bestimmten Situationen nützlich, wenn Regeln verwendet werden. Weitere Informationen hierzu finden Sie im Abschnitt „Port-Affinität außer Kraft setzen“ auf Seite 220.

Wert

Wert für `sticky`: ja oder nein. Der Standardwert ist Ja.

weight

Eine Zahl von 0 bis 100 (die den angegebenen Wert für die Wertigkeitsgrenze des Ports nicht überschreiten darf) zur Angabe der Gewichtung dieses Servers. Wird die Wertigkeit auf 0 gesetzt, werden keine neuen Anforderungen an den Server gesendet, die derzeit aktiven Verbindungen zu diesem Server werden jedoch nicht beendet. Der Standardwert entspricht der Hälfte der für den Port angegebenen Wertigkeitsgrenze. Ist der Manager aktiv, wird diese Einstellung schnell überschrieben.

Wert

Wertigkeit des Servers.

fixedweight

Mit der Option `fixedweight` können Sie angeben, ob der Manager die Serverwertigkeit ändern soll. Wird der Wert für `fixedweight` auf `yes` gesetzt, kann der Manager die Serverwertigkeit nicht ändern. Weitere Informationen hierzu finden Sie im Abschnitt „Feste Wertigkeiten vom Manager“ auf Seite 182.

Wert

Wert für `fixedweight`: ja oder nein. Der Standardwert ist Nein.

cookievalue

Der Parameter `cookievalue` ist ein zufälliger Wert, der die Serverseite des aus Cookie-Namen und Cookie-Wert bestehenden Paares. Der Cookie-Wert dient neben dem Cookie-Namen als Kennung, mit der Load Balancer nachfolgende Client-Anforderungen an nur einen Server senden kann. Weitere Informationen hierzu finden Sie im Abschnitt „Passive Cookie-Affinität“ auf Seite 227.

Anmerkung: Der Parameter cookievalue ist für Dispatcher (bei Verwendung der Weiterleitungsmethode cbr) und CBR gültig.

Wert

Ein zufälliger Wert. Standardmäßig ist kein Cookie-Wert angegeben.

mapport

Zuordnen der Nummer des Ziel-Ports für die Client-Anforderung (die für den Dispatcher angegeben ist) zur Nummer des Server-Ports, an dem der Dispatcher den Lastausgleich für die Client-Anforderungen durchführt. Erlaubt Load Balancer, die Anforderung eines Clients an einem Port zu empfangen und sie an einen anderen Port auf der Servermaschine zu übertragen. Mit mapport können Sie den Lastausgleich für eine Client-Anforderung auf einem Server durchführen, auf dem mehrere Serverdämonen aktiv sind.

Anmerkung: Der Parameter mapport gilt für Dispatcher (bei Verwendung der Weiterleitungsmethode nat oder cbr) und für CBR. Weitere Informationen zum Dispatcher finden Sie in den Abschnitten „Dispatcher-Weiterleitungsmethode nat“ auf Seite 53 und „Inhaltsabhängige Weiterleitung durch die Komponente Dispatcher (cbr)“ auf Seite 55. Weitere Informationen zu CBR können Sie dem Abschnitt „Lastausgleich für SSL-Datenverkehr vom Client zum Proxy und HTTP-Datenverkehr vom Proxy zum Server“ auf Seite 108 entnehmen.

protocol

Die gültigen Werte für protocol sind HTTP und HTTPS. Der Standardwert ist HTTP.

Anmerkung: Der Parameter protocol gilt nur für die Komponente CBR.

Port-Wert

Nummer des zugeordneten Ports. Der Standardwert ist die Nummer des Ziel-Ports für die Client-Anforderung.

router

Wenn Sie ein Weitverkehrsnetz konfigurieren, ist dies die Adresse des Routers zum fernen Server. Der Standardwert ist 0. Er gibt einen lokalen Server an. Wurde die Router-Adresse eines Servers einmal auf einen anderen Wert als null gesetzt (gibt einen fernen Server an), kann die Adresse nicht mehr auf null zurückgesetzt werden, um einen lokalen Server anzugeben. Der Server muss stattdessen entfernt und dann erneut ohne Angabe einer Router-Adresse hinzugefügt werden. Genauso kann ein als lokal definierter Server (Router-Adresse = 0) nicht als ferner Server definiert werden, indem die Router-Adresse geändert wird. Der Server muss entfernt und erneut hinzugefügt werden. Weitere Informationen hierzu finden Sie im Abschnitt „Dispatcher-WAN-Unterstützung konfigurieren“ auf Seite 229.

Anmerkung: Der Parameter router gilt nur für Dispatcher. Wenn Sie die Weiterleitungsmethode nat oder cbr verwenden, müssen Sie beim Hinzufügen eines Servers zur Konfiguration die Router-Adresse angeben.

Adresse

Wert der Adresse des Routers.

returnaddress

Eine eindeutige IP-Adresse oder ein eindeutiger Hostname. Diese Adresse wird auf der Dispatcher-Maschine konfiguriert und wird vom Dispatcher bei der

Lastverteilung der Client-Anforderung an den Server als Quellenadresse verwendet. Auf diese Weise wird sichergestellt, dass der Server das Paket an die Dispatcher-Maschine zurückgibt und es nicht direkt an den Client sendet. (Der Dispatcher leitet das IP-Paket dann an den Client weiter.) Sie müssen den Wert für die Rückkehradresse beim Hinzufügen des Servers angeben. Die Rückkehradresse kann nur geändert werden, wenn Sie den Server entfernen und dann erneut hinzufügen. Die Rückkehradresse darf nicht mit dem Wert für Cluster, Server oder NFA übereinstimmen.

Anmerkung: Der Parameter `returnaddress` gilt nur für Dispatcher. Wenn Sie die Weiterleitungsmethode `nat` oder `cbr` verwenden, müssen Sie beim Hinzufügen eines Servers zur Konfiguration die Rückkehradresse angeben.

Adresse

Wert der Rückkehradresse.

advisorrequest

Die HTTP- oder HTTPS-Advisor-Funktion verwendet die Zeichenfolge `advisorrequest`, um den Status der Server abzufragen. Dieser Wert ist nur für die Server gültig, die mit der HTTP- oder HTTPS-Advisor-Funktion zusammenarbeiten. Zum Aktivieren dieses Wertes müssen Sie die HTTP- oder HTTPS-Advisor-Funktion starten. Weitere Informationen hierzu finden Sie im Abschnitt „Option ‘Anforderung/Antwort (URL)’ der HTTP- oder HTTPS-Advisor-Funktion konfigurieren“ auf Seite 191.

Anmerkung: Der Parameter `advisorrequest` gilt für die Komponenten Dispatcher und CBR.

Zeichenfolge

Wert der Zeichenfolge, die von der HTTP- oder HTTPS-Advisor-Funktion verwendet wird. Der Standardwert ist `HEAD / HTTP/1.0`.

Anmerkung: Wenn die Zeichenfolge ein Leerzeichen enthält, gilt Folgendes:

- Bei Absetzen des Befehls von der Shell-Eingabeaufforderung **dscontrol**>> müssen Sie die Zeichenfolge in Anführungszeichen setzen. Beispiel: **server set Cluster:Port:Server advisorrequest "head / http/1.0"**
- Beim Absetzen des Befehls **dscontrol** an der Eingabeaufforderung des Betriebssystems müssen Sie dem Text die Zeichen `"\"` voranstellen und den Text mit den Zeichen `\"` beenden. Beispiel: **dscontrol server set Cluster:Port:Server advisorrequest "\"head / http/1.0\""**

advisorresponse

Die Antwortzeichenfolge der Advisor-Funktion, nach der die HTTP- oder HTTPS-Advisor-Funktion die HTTP-Antwort durchsucht. Dieser Wert ist nur für die Server gültig, die mit der HTTP- oder HTTPS-Advisor-Funktion zusammenarbeiten. Zum Aktivieren dieses Wertes müssen Sie die HTTP- oder HTTPS-Advisor-Funktion starten. Weitere Informationen hierzu finden Sie im Abschnitt „Option ‘Anforderung/Antwort (URL)’ der HTTP- oder HTTPS-Advisor-Funktion konfigurieren“ auf Seite 191.

Anmerkung: Der Parameter `advisorresponse` gilt für die Komponenten Dispatcher und CBR.

Zeichenfolge

Wert der Zeichenfolge, die von der HTTP- oder HTTPS-Advisor-Funktion verwendet wird. Der Standardwert ist null.

Anmerkung: Wenn die Zeichenfolge ein Leerzeichen enthält, gilt Folgendes:

- Bei Absetzen des Befehls von der Shell-Eingabeaufforderung **dscontrol>>** müssen Sie die Zeichenfolge in Anführungszeichen setzen.
- Beim Absetzen des Befehls **dscontrol** an der Eingabeaufforderung des Betriebssystems müssen Sie dem Text die Zeichen "\" voranstellen und den Text mit den Zeichen \"\" beenden.

down

Diesen Server als inaktiv markieren. Durch diesen Befehl werden alle aktiven Verbindungen zu diesem Server unterbrochen, und es wird verhindert, dass weitere Verbindungen oder Pakete an diesen Server gesendet werden.

Wenn Sie einen Server mit dem Befehl "server down" offline setzen und der stickytime-Wert für diesen Server ungleich null ist, werden vorhandene Clients weiter von diesem Server bedient, bis das stickytime-Intervall abläuft. Der Server wird erst nach Ablauf des stickytime-Intervalls heruntergefahren.

remove

Diesen Server entfernen.

report

Bericht über diesen Server erstellen. Der Bericht enthält zu jedem Server die folgenden Angaben: Anzahl gleichzeitiger Verbindungen pro Sekunde, innerhalb einer Sekunde übertragene Kilobytes, Summe der Verbindungen, Summe der aktiven Verbindungen, Anzahl der Verbindungen mit Beendigungsstatus (FIN) und Anzahl beendeter Verbindungen.

set

Werte für diesen Server festlegen.

status

Status der Server anzeigen.

up Diesen Server als aktiv markieren. Der Dispatcher sendet jetzt neue Verbindungen zu diesem Server.

Beispiele

- Hinzufügen des Servers mit der Adresse 27.65.89.42 zum Port 80 an der Clusteradresse 130.40.52.153:
`dscontrol server add 130.40.52.153:80:27.65.89.42`
- Setzen des Servers an der Adresse 27.65.89.42 auf nonsticky (Merkmal zur Außerkraftsetzung der Port-Affinität):
`dscontrol server set 130.40.52.153:80:27.65.89.42 sticky no`
- Geben Sie den folgenden Befehl ein, um den Server an 27.65.89.42 als "ausgefallen" zu kennzeichnen:
`dscontrol server down 130.40.52.153:80:27.65.89.42`
- Geben Sie den folgenden Befehl ein, um den Server 27.65.89.42 von allen Ports aller Cluster zu entfernen:
`dscontrol server remove ::27.65.89.42`

- Geben Sie den folgenden Befehl ein, um den Server an 27.65.89.42 als "zusammengeführt" zu definieren (Server befindet sich auf derselben Maschine wie Load Balancer):
dscontrol server set 130.40.52.153:80:27.65.89.42 collocated yes
- Festlegen der Wertigkeit 10 für Server 27.65.89.42 am Port 80 der Clusteradresse 130.40.52.153:
dscontrol server set 130.40.52.153:80:27.65.89.42 weight 10
- Geben Sie den folgenden Befehl ein, um den Server an 27.65.89.42 als "aktiv" zu kennzeichnen:
dscontrol server up 130.40.52.153:80:27.65.89.42
- Geben Sie den folgenden Befehl ein, um einen fernen Server hinzuzufügen:
dscontrol server add 130.40.52.153:80:130.60.70.1 router 130.140.150.0
- Zielsetzung, dass die HTTP-Advisor-Funktion eine HTTP-URL-Anforderung HEAD / HTTP/1.0 für Server 27.65.89.42 am HTTP-Port 80 abfragt:
dscontrol server set 130.40.52.153:80:27.65.89.42
advisorrequest "\"HEAD / HTTP/1.0\""
- Anzeigen des Status für Server 9.67.143.154 am Port 80:
dscontrol server status 9.67.131.167:80:9.67.143.154

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Serverstatus:

```

-----
Server ..... 9.67.143.154
Port-Nummer ..... 80
Cluster ..... 9.67.131.167
Clusteradresse ..... 9.67.131.167
Stillgelegt ..... N
Server aktiv ..... J
Gewichtung ..... 10
Feste Gewichtung ..... N
Haltemodus für Regel ..... J
Ferner Server ..... N
Netz-Router-Adresse ..... 0.0.0.0
Zusammengelegt ..... N
Advisor-Anforderung ..... HEAD / HTTP/1.0
Advisor-Antwort .....
Cookie-Wert ..... nicht anwendbar
Klon-ID ..... nicht anwendbar

```

dscontrol set — Serverprotokoll konfigurieren



loglevel

Die Stufe für die Protokollierung von dserver-Aktivitäten.

Stufe

Der Standardwert für **loglevel** ist 0. Der gültige Bereich ist 0 bis 5. Die folgenden Werte sind gültig: 0 für keine Einträge, 1 für eine minimale Protokollierung, 2 für eine Basisprotokollierung, 3 für eine normale, 4 für eine erweiterte und 5 für eine ausführliche Protokollierung.

logsize

Die maximale Anzahl von Byte, die in der Protokolldatei protokolliert werden können.

Größe

Der Standardwert für logsize ist 1 MB.

dscontrol status — Aktivitätsanzeige für Manager und Advisor-Funktionen

►►—dscontrol—status—◀◀

Beispiele

- Geben Sie den folgenden Befehl ein, um festzustellen, ob der Manager und die Advisor aktiv sind:

```
dscontrol status
```

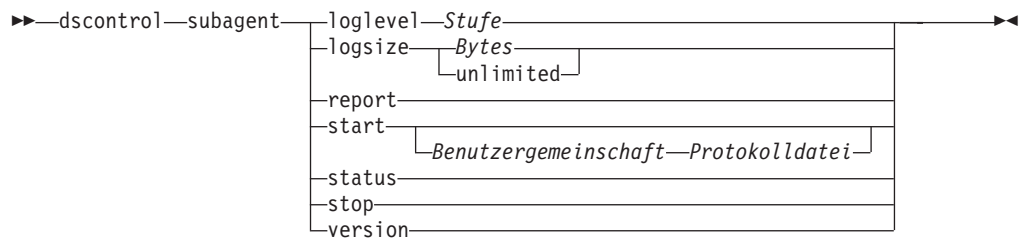
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Executor wurde gestartet.
Manager wurde gestartet.

ADVISOR	CLUSTER:PORT	ZEITLIMIT
reach	0	unlimited
http	80	unlimited
ftp	21	unlimited

dscontrol subagent — SNMP-Subagent konfigurieren

Anmerkung: Die Befehlssyntax für dscontrol subagent gilt für die Komponente Dispatcher.



loglevel

Die Stufe, auf der der Subagent seine Aktivitäten in einer Datei protokolliert.

Stufe

Die Nummer der Stufe (0 bis 5). Je größer die Zahl, desto mehr Informationen werden in das Manager-Protokoll geschrieben. Der Standardwert ist 1. Die folgenden Werte sind gültig: 0 für keine Einträge, 1 für eine minimale Protokollierung, 2 für eine Basisprotokollierung, 3 für eine normale, 4 für eine erweiterte und 5 für eine ausführliche Protokollierung.

logsize

Festlegen der Maximalen Größe des Subagentenprotokolls in Bytes. Der Standardwert ist 1 MB. Wenn Sie eine maximale Größe für die Protokolldatei festlegen, findet ein Dateiumlauf statt. Hat die Datei die angegebene Größe erreicht, werden alle weiteren Einträge wieder an den Anfang der Datei geschrieben und die dort befindlichen Einträge überschrieben. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge werden mit einer Zeitmarke versehen, damit Sie erkennen können, in welcher Reihenfolge die Einträge geschrieben wurden. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Bytes

Die maximale Größe in Byte für die Protokolldatei des Subagenten. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** (unbegrenzt) angeben. Möglicherweise erreicht die Protokolldatei nicht genau die maximale Größe, bevor der Dateiumlauf stattfindet, da die Größe der Protokolleinträge variiert. Der Standardwert ist unlimited.

report

Zeigt eine statistische Momentaufnahme an.

start

Den Subagenten starten.

Benutzergemeinschaft

Der Name der SNMP-Benutzergemeinschaft, den Sie als Sicherheitskennwort verwenden können. Der Standardwert ist public.

Windows-Plattform: Es wird der Name der Benutzergemeinschaft für das Betriebssystem verwendet.

Protokolldatei

Der Name der Datei, in der die Daten des SNMP-Subagenten protokolliert werden. Jeder Eintrag im Protokoll wird mit einer Zeitmarke versehen. Der Standardwert ist `subagent.log`. Die Standarddatei ist im Verzeichnis **logs** installiert. Lesen Sie hierzu die Informationen in Anhang C, „Beispielkonfigurationsdateien“, auf Seite 481. Wollen Sie das Verzeichnis ändern, in dem die Protokolldateien gespeichert werden, lesen Sie die Informationen im Abschnitt „Pfade für die Protokolldatei ändern“ auf Seite 269.

status

Zeigt den aktuellen Status aller Werte in dem SNMP-Subagenten an, die global gesetzt werden können. Zudem werden die Standardwerte dieser Werte angezeigt.

version

Zeigt die aktuelle Version des Subagenten an.

Beispiele

- Geben Sie den folgenden Befehl ein, um den Subagenten mit dem Benutzer-gemeinschaftsnamen `bigguy` zu starten:
`dscontrol subagent start bigguy bigguy.log`

Kapitel 28. Befehlsreferenz für Site Selector

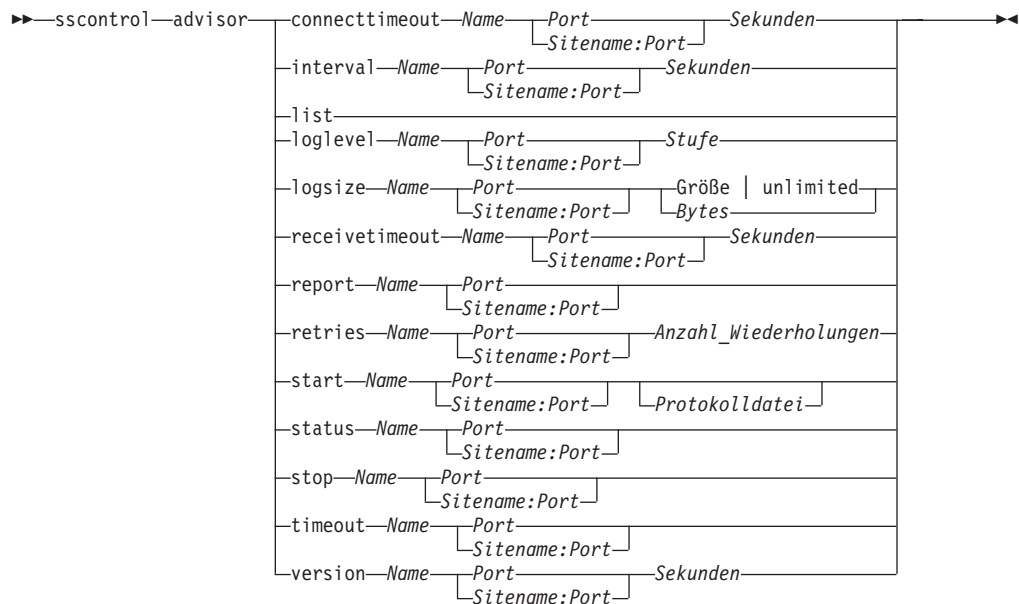
Dieses Kapitel beschreibt die Verwendung der folgenden **sscontrol**-Befehle für Site Selector:

- „sscontrol advisor — Advisor-Funktion steuern“ auf Seite 406
- „sscontrol file — Konfigurationsdateien verwalten“ auf Seite 411
- „sscontrol help — Hilfetext für diesen Befehl anzeigen oder drucken“ auf Seite 413
- „sscontrol logstatus — Protokolleinstellungen des Servers anzeigen“ auf Seite 414
- „sscontrol manager — Manager steuern“ auf Seite 415
- „sscontrol metric — Systemmesswerte konfigurieren“ auf Seite 420
- „sscontrol nameserver — Namensserver steuern“ auf Seite 421
- „sscontrol rule — Regeln konfigurieren“ auf Seite 422
- „sscontrol server — Server konfigurieren“ auf Seite 425
- „sscontrol set — Serverprotokoll konfigurieren“ auf Seite 427
- „sscontrol sitename — Sitenamen konfigurieren“ auf Seite 428
- „sscontrol status — Aktivitätsanzeige für Manager und Advisor-Funktionen“ auf Seite 431

Sie können eine Minimalversion der Parameter für den Befehl **sscontrol** eingeben. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **sscontrol he f** an Stelle von **sscontrol help file** eingeben.

Anmerkung: Die Werte der Befehlsparameter müssen in englischen Zeichen eingegeben werden. Die einzige Ausnahme hiervon bilden Hostnamen (die in den Befehlen **cluster** und **server** verwendet werden) und (die in Dateibefehlen verwendeten) Dateinamen.

sscontrol advisor — Advisor-Funktion steuern



connecttimeout

Definiert, wie lange eine Advisor-Funktion wartet, bevor sie meldet, dass zu einem Server keine Verbindung hergestellt werden kann. Weitere Informationen hierzu finden Sie im Abschnitt „Serververbindungs- und -empfangszeitlimit der Advisor-Funktion“ auf Seite 188.

Name

Der Name der Advisor-Funktion. Mögliche Werte sind **http**, **https**, **ftp**, **sip**, **ssl**, **smtp**, **imap**, **pop3**, **ldap**, **nntp**, **telnet**, **connect**, **ping**, **WLM** und **WTE**. Die Namen angepasster Advisor-Funktionen haben das Format **xxxx**, wobei **ADV_xxxx** der Name der Klasse ist, die die angepasste Advisor-Funktion implementiert.

Port

Die Nummer des Ports, der von der Advisor-Funktion überwacht wird.

Sekunden

Eine positive ganze Zahl, die die Zeit in Sekunden angibt, nach deren Ablauf die Advisor-Funktion meldet, dass zu einem Server keine Verbindung hergestellt werden kann. Dieser Wert liegt standardmäßig beim Dreifachen des für das Advisor-Intervall angegebenen Wertes.

interval

Legt fest, wie oft der Advisor Informationen von den Servern abfragt.

Sekunden

Eine positive ganze Zahl, die die Zeit zwischen den an die Server gerichteten Statusabfragen in Sekunden angibt. Der Standardwert ist 7.

list

Zeigt eine Liste der Advisor an, die derzeit Informationen an den Manager liefern.

loglevel

Legt die Protokollstufe für ein Advisor-Protokoll fest.

Stufe

Die Nummer der Stufe (0 bis 5). Der Standardwert ist 1. Je größer die Zahl ist, desto mehr Informationen werden in das Advisor-Protokoll geschrieben. Gültige Werte:

- 0 entspricht keiner Protokollierung
- 1 entspricht einer minimalen Protokollierung
- 2 entspricht einer Basisprotokollierung
- 3 entspricht einer normalen Protokollierung
- 4 entspricht einer erweiterten Protokollierung
- 5 entspricht einer ausführlichen Protokollierung.

logsize

Legt die maximale Größe eines Advisor-Protokolls fest. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, werden bei Erreichen der Größe die vorherigen Protokolleinträge überschrieben. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe | unlimited

Die maximale Größe der Advisor-Protokolldatei in Bytes. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder **unlimited** (unbegrenzt) angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen. Der Standardwert ist 1 MB.

receivetimeout

Definiert, wie lange eine Advisor-Funktion wartet, bevor sie meldet, dass von einem Server keine Daten empfangen werden können. Weitere Informationen hierzu finden Sie im Abschnitt „Serververbindungs- und -empfangszeitlimit der Advisor-Funktion“ auf Seite 188.

Sekunden

Eine positive ganze Zahl, die die Zeit in Sekunden angibt, nach deren Ablauf die Advisor-Funktion meldet, dass von einem Server keine Daten empfangen werden können. Dieser Wert liegt standardmäßig beim Dreifachen des für das Advisor-Intervall angegebenen Wertes.

report

Anzeigen eines Berichts zum Advisor-Status.

retries

Legt die Wiederholungsversuche einer Advisor-Funktion fest, bevor diese einen Server als inaktiv markiert.

Anzahl_Wiederholungen

Eine ganze Zahl größer als oder gleich null. Dieser Wert sollte nicht größer als 3 sein. Wenn das Schlüsselwort für Wiederholungen nicht konfiguriert ist, wird standardmäßig von null Wiederholungsversuchen ausgegangen.

start

Den Advisor starten. Für alle Protokolle stehen Advisor zur Verfügung. Die Standard-Ports sind:

Advisor-Name	Protokoll	Port
connect	nicht anwendbar	benutzerdefiniert
db2	privat	50000
ftp	FTP	21
http	HTTP	80
https	SSL	443
imap	IMAP	143
ldap	LDAP	389
nntp	NNTP	119
PING	PING	nicht anwendbar
pop3	POP3	110
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
telnet	Telnet	23

Name

Der Name der Advisor-Funktion.

Sitename:Port

Der Wert "Sitename" ist in den advisor-Befehlen optional, der Wert "Port" jedoch erforderlich. Wird der Wert "Sitename" nicht angegeben, wird die Advisor-Funktion für alle verfügbaren konfigurierten Sitenamen ausgeführt. Bei Angabe eines Sitenamens wird die Advisor-Funktion nur für den von Ihnen angegebenen Sitenamen ausgeführt. Zusätzliche Sitenamen werden durch ein Pluszeichen (+) voneinander getrennt.

Protokolldatei

Der Name der Datei, in die die Verwaltungsdaten geschrieben werden. Jeder Eintrag des Protokolls ist mit einer Zeitmarke versehen.

Die Standarddatei ist *Advisor-Name_Port.log*, z. B. **http_80.log**. Informationen zum Ändern des Verzeichnisses, in dem Protokolldateien gespeichert werden, finden Sie im Abschnitt „Pfade für die Protokolldatei ändern“ auf Seite 269.

Sie können nur eine Advisor-Funktion pro Sitenamen starten.

status

Anzeigen des aktuellen Status sowie der Standardeinstellungen für alle globalen Werte einer Advisor-Funktion.

stop

Den Advisor stoppen.

timeout

Legt die Zeit in Sekunden fest, in der der Manager von der Advisor-Funktion erhaltene Informationen als gültig ansieht. Stellt der Manager fest, dass die Advisor-Informationen älter als dieses Zeitlimit sind, verwendet der Manager diese Informationen nicht zum Bestimmen Wertigkeiten für die Server am Port, die von der Advisor-Funktion überwacht werden. Dieses Zeitlimit gilt nicht, wenn die Advisor-Funktion den Manager darüber informiert hat, dass ein bestimmter Server inaktiv ist. Der Manager verwendet diese Informationen über den Server auch, nachdem die Advisor-Informationen das Zeitlimit überschritten haben.

Sekunden

Eine positive Zahl, die die Sekunden angibt, oder **unlimited**. Der Standardwert ist unlimited.

version

Zeigt die aktuelle Advisor-Version an.

Beispiele

- Festlegen der Zeit (30 Sekunden), die eine HTTP-Advisor-Funktion (für Port 80) wartet, bevor sie meldet, dass zu einem Server keine Verbindung hergestellt werden kann:

```
sscontrol advisor connecttimeout http 80 30
```

- Festlegen des Intervalls für die FTP-Advisor-Funktion (für Port 21) auf 6 Sekunden:

```
sscontrol advisor interval ftp 21 6
```

- Geben Sie den folgenden Befehl ein, um eine Liste der Advisor anzuzeigen, die derzeit Informationen an den Manager liefern:

```
sscontrol advisor list
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

ADVISOR	SITENAME:PORT	ZEITLIMIT

http	80	unlimited
ftp	21	unlimited

- Ändern der Protokollstufe für das Protokoll der Advisor-Funktion http für den Sitenamen "meineSite" in 0, um einen höheren Durchsatz zu erreichen:

```
sscontrol advisor loglevel http meineSite:80 0
```

- Ändern der Protokollgröße der Advisor-Funktion ftp für den Sitenamen "meineSite" in 5000 Bytes:

```
sscontrol advisor  
logsize ftp meineSite:21 5000
```

- Festlegen der Zeit (60 Sekunden), die eine HTTP-Advisor-Funktion (für Port 80) wartet, bevor sie meldet, dass von einem Server keine Daten empfangen werden können:

```
sscontrol advisor receivetimeout http 80 60
```

- Anzeigen eines Berichts zum Status der Advisor-Funktion ftp (für Port 21):

```
sscontrol  
advisor report ftp 21
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Advisor-Bericht:

Advisor-Name http
Port-Nummer 80

Sitename meineSite
Serveradresse 9.67.129.230
Last 8

- Geben Sie den folgenden Befehl ein, um den Advisor mit der Datei ftpadv.log zu starten:

```
sscontrol advisor start ftp 21 ftpadv.log
```

- Anzeigen des aktuellen Status der Werte, die der Advisor-Funktion http zugeordnet sind:

```
sscontrol advisor status http 80
```

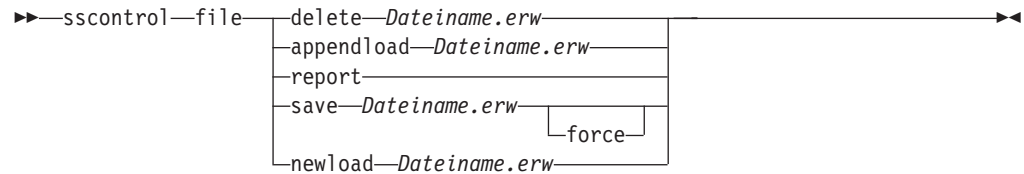
Dieser Befehl erzeugt eine Ausgabe, die ungefähr wie folgt aussieht:

Advisor-Status:

```
Intervall (Sekunden) ..... 7
Zeitlimit (Sekunden) ..... Unlimited
Zeitlimit für Verbindung (Sekunden) ..... 21
Zeitlimit für Empfang (Sekunden) ..... 21
Advisor-Protokolldateiname ..... Http_80.log
Protokollstufe ..... 1
Maximale Managerprotokollgröße (Bytes)... Unlimited
Anzahl Wiederholungen ..... 0
```

- Stoppen der Advisor-Funktion http am Port 80:
sscontrol advisor stop http 80
- Festlegen eines Zeitlimits für Advisor-Informationen von 5 Sekunden:
sscontrol advisor timeout ftp 21 5
- Ermitteln der aktuellen Versionsnummer für die Advisor-Funktion ssl:
sscontrol advisor version ssl 443

sscontrol file — Konfigurationsdateien verwalten



delete

Die Datei löschen.

Dateiname.erw

Eine Konfigurationsdatei.

Die Dateierweiterung (.*erw*) ist optional und kann beliebig gewählt werden.

appendload

Hinzufügen einer Konfigurationsdatei zur aktuellen Konfiguration und Laden der Datei in Site Selector.

report

Bericht über die verfügbare(n) Datei(en).

save

Sichern der aktuellen Konfiguration für Site Selector in der Datei.

Anmerkung: Dateien werden in den nachfolgend genannten Verzeichnissen gespeichert und aus diesen geladen:

- Linux- und UNIX-Systeme: **/opt/ibm/edge/lb/servers/configurations/ss**
- Windows-Systeme: **C:\Programme\ibm\edge\lb\servers\configurations\Komponente**

force

Wenn Sie Ihre Datei in einer vorhandenen Datei mit demselben Namen speichern möchten, verwenden Sie **force**, um die vorhandene Datei vor dem Speichern der neuen Datei zu löschen. Bei Nichtverwendung der Option **force** wird die vorhandene Datei nicht überschrieben.

newload

Laden einer neuen Konfigurationsdatei in Site Selector. Die neue Konfigurationsdatei ersetzt die aktuelle Konfiguration.

Beispiele

- Geben Sie den folgenden Befehl ein, um eine Datei zu löschen:
`sscontrol file delete Datei3`

Datei (Datei3) wurde gelöscht.
- Geben Sie den folgenden Befehl ein, um eine neue Konfigurationsdatei zu laden, die die aktuelle Konfiguration ersetzt:
`sscontrol file newload Datei1.sv`

Datei (Datei1.sv) wurde in den Dispatcher geladen.

- Geben Sie den folgenden Befehl ein, um eine Konfigurationsdatei an die aktuelle Konfiguration anzuhängen und zu laden:

```
sscontrol file appendload Datei2.sv
```

Datei (Datei2.sv) wurde an die aktuelle Konfiguration angehängt und geladen.

- Geben Sie den folgenden Befehl ein, um einen Bericht über Ihre Dateien anzuzeigen (die Dateien, die zuvor gesichert wurden):

```
sscontrol file report
```

DATEIBERICHT:

Datei1.save

Datei2.sv

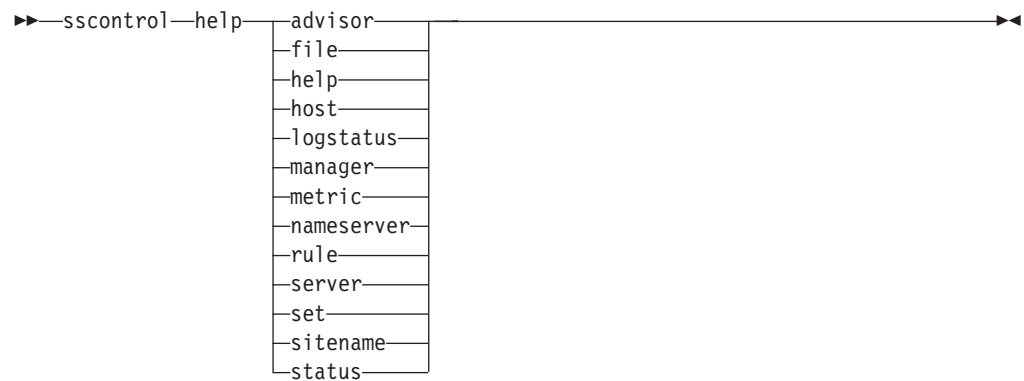
Datei3

- Geben Sie den folgenden Befehl ein, um die Konfiguration in der Datei Datei3 zu sichern:

```
sscontrol file save Datei3
```

Die Konfiguration wurde in Datei (Datei3) gesichert.

sscontrol help — Hilfetext für diesen Befehl anzeigen oder drucken



Beispiele

- Hilfetext zum Befehl sscontrol abrufen:

```
sscontrol help
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
ARGUMENTE BEFEHL HELP:
```

```
-----
```

```
Verwendung: help <Hilfeoption>
```

```
Beispiel: help name
```

help	- vollständigen Hilfetext drucken
advisor	- Hilfe zum Befehl advisor
file	- Hilfe zum Befehl file
host	- Hilfe zum Befehl host
manager	- Hilfe zum Befehl manager
metric	- Hilfe zum Befehl metric
sitename	- Hilfe zum Befehl sitename
nameserver	- Hilfe zum Befehl nameserver
rule	- Hilfe zum Befehl rule
server	- Hilfe zum Befehl server
set	- Hilfe zum Befehl set
status	- Hilfe zum Befehl status
logstatus	- Hilfe zum Befehl logstatus

Parameter innerhalb von < > sind Variablen.

- Manchmal enthält der Hilfetext Optionen für die Variablen, die durch das Zeichen | voneinander getrennt sind:

```
logsize <Bytezahl | unlimited>
```

```
-Maximale Anzahl Bytes für Protokolldatei festlegen
```

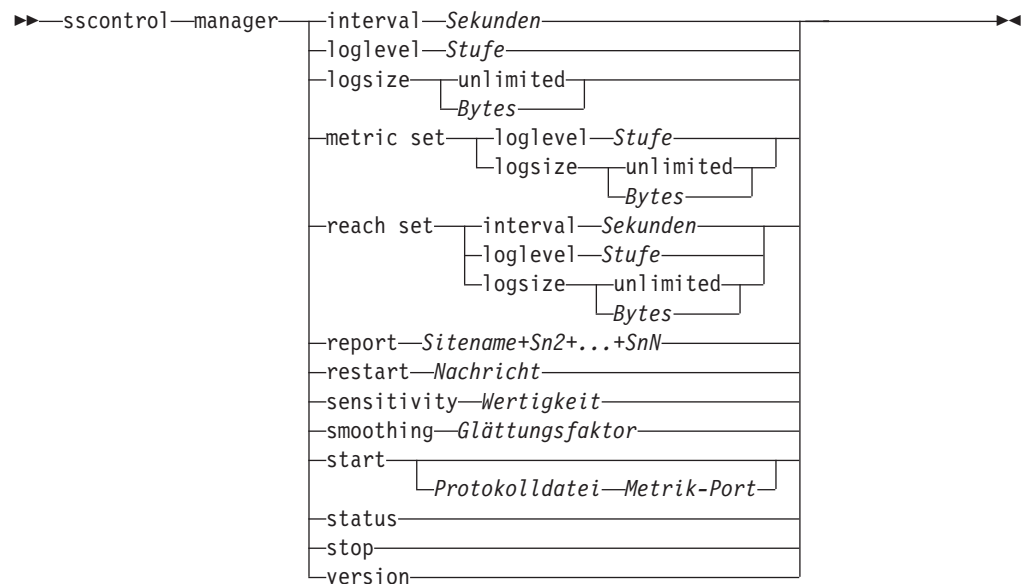
sscontrol logstatus — Protokolleinstellungen des Servers anzeigen

►►—sscontrol—logstatus—◄◄

logstatus

Zeigt die Einstellungen des Serverprotokolls (Name der Protokolldatei, Protokollstufe und -größe) an.

sscontrol manager — Manager steuern



interval

Legt fest, wie oft der Manager die Wertigkeit der Server aktualisiert.

Sekunden

Eine positive Zahl, die in Sekunden darstellt, wie oft der Manager Wertigkeiten aktualisiert. Der Standardwert ist 2.

loglevel

Legt die Protokollstufe für das Protokoll des Managers fest.

Stufe

Die Nummer der Stufe (0 bis 5). Je größer die Zahl, desto mehr Informationen werden in das Manager-Protokoll geschrieben. Der Standardwert ist 1. Gültige Werte:

- 0 entspricht keiner Protokollierung
- 1 entspricht einer minimalen Protokollierung
- 2 entspricht einer Basisprotokollierung
- 3 entspricht einer normalen Protokollierung
- 4 entspricht einer erweiterten Protokollierung
- 5 entspricht einer ausführlichen Protokollierung.

logsize

Legt die maximale Größe des Protokolls des Managers fest. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Bytes

Die maximale Größe in Byte für die Protokolldatei des Managers. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder **unlimited** (unbegrenzt) angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen. Der Standardwert ist 1 MB.

metric set

Legt die Werte **loglevel** und **logsize** für das Metriküberwachungsprotokoll fest. Der Wert für loglevel ist die Stufe für die Protokollierung der Metriküberwachung (0 - Keine, 1 - Minimal, 2 - Grundlegend, 3 - Mäßig, 4 - Erweitert oder 5 - Ausführlich). Die Standardprotokollstufe ist 1. Der Wert für logsize ist die Datenmenge (in Bytes), die maximal in der Protokolldatei für Metriküberwachung erfasst wird. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder unlimited (unbegrenzt) angeben. Die Standardprotokollgröße ist 1.

reach set

Legt das Intervall, die Protokollstufe und die Protokollgröße für die Advisor-Funktion reach fest.

report

Zeigt eine statistische Momentaufnahme an.

Sitename

Der Sitename, der im Bericht angezeigt werden soll. Dies ist ein nicht auflösbarer Hostname, den der Client abfragt. Der Sitename muss ein vollständig qualifizierter Domänenname sein.

Anmerkung: Zusätzliche Sitenamen werden durch ein Pluszeichen (+) voneinander getrennt.

restart

Startet alle Server (die nicht inaktiv sind) mit der Standardwertigkeit (1/2 der maximalen Wertigkeit).

Nachricht

Eine Nachricht, die in die Protokolldatei des Managers gestellt werden soll.

sensitivity

Legt die Mindestsensitivität für die Aktualisierung von Wertigkeiten fest. Diese Einstellung definiert, wann der Manager seine Serverwertigkeit ausgehend von externen Informationen ändern sollte.

Wertigkeit

Eine Zahl von 0 bis 100, die als prozentuale Wertigkeit verwendet wird. Der Standardwert 5 bewirkt eine Mindestsensitivität von 5 %.

smoothing

Festlegen eines Faktors, der Wertigkeitsabweichungen während des Lastausgleichs glättet. Ein höherer Glättungsfaktor führt zu einer weniger drastischen Änderung von Serverwertigkeiten bei veränderten Netzwerkbedingungen. Ein geringerer Glättungsfaktor führt zu einer drastischeren Änderung der Serverwertigkeiten.

Glättungsfaktor

Eine positive Gleitkommazahl. Der Standardwert ist 1,5.

start

Den Manager starten.

Protokolldatei

Der Name der Datei, in der die Daten des Managers protokolliert werden. Jeder Eintrag des Protokolls ist mit einer Zeitmarke versehen.

Die Standarddatei ist im Verzeichnis **logs** installiert. Lesen Sie hierzu die Informationen in Anhang C, „Beispielkonfigurationsdateien“, auf Seite 481. Wollen Sie das Verzeichnis ändern, in dem die Protokolldateien gespeichert werden, lesen Sie die Informationen im Abschnitt „Pfade für die Protokolldatei ändern“ auf Seite 269.

Metrik-Port

Der Port, an dem Metric Server Systembelastungen meldet. Wenn Sie einen Metrik-Port angeben, müssen Sie auch einen Protokolldateinamen angeben. Der Standard-Metrik-Port ist 10004.

status

Anzeigen des aktuellen Status sowie der Standardeinstellungen für alle globalen Werte des Managers.

stop

Den Manager stoppen.

version

Zeigt die aktuelle Version des Managers an.

Beispiele

- Geben Sie den folgenden Befehl ein, um das Aktualisierungsintervall für den Manager auf 5 Sekunden zu setzen:
`sscontrol manager interval 5`
- Geben Sie den folgenden Befehl ein, um die Stufe der Protokollierung zwecks Verbesserung der Leistung auf 0 zu setzen:
`sscontrol manager loglevel 0`
- Geben Sie den folgenden Befehl ein, um die Größe des Protokolls des Managers auf 1.000.000 Byte zu setzen:
`sscontrol manager logsize 1000000`
- Geben Sie den folgenden Befehl ein, um eine statistische Momentaufnahme des Managers abzurufen:
`sscontrol manager report`

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

SERVER	STATUS
9.67.129.221	AKTIV
9.67.129.213	AKTIV
9.67.134.223	AKTIV

LEGENDE ZUM MANAGERBERICHT

CPU	CPU-Last
MEM	Speicherlast
SYS	Systemmetrik
JETZT	Aktuelle Gewichtung
NEU	Neue Gewichtung
GWT	Gewichtung

meineSite		GWT		CPU 49 %		MEM 50 %		PORT 1 %		SYS 0 %	
		JETZT NEU		GWT LAST		GWT LAST		GWT LAST		GWT LAST	
9.37.56.180		10	10	-99	-1	-99	-1	-99	-1	0	0
SUMMEN:		10	10	-1		-1		-1		0	

ADVISOR	SITENAME:PORT	ZEITLIMIT
http	80	unlimited

- Neustart aller Server mit Standardwertigkeit und Schreiben einer Nachricht in die Manager-Protokolldatei:

`sscontrol manager restart` Neustart des Managers für Codeaktualisierung

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

320-14:04:54 Neustart des Managers für Codeaktualisierung

- Setzen der Sensitivität für Wertigkeitsänderungen auf 10:
`sscontrol manager sensitivity 10`
- Geben Sie den folgenden Befehl ein, um den Glättungsfaktor auf 2,0 zu setzen:
`sscontrol manager smoothing 2.0`
- Geben Sie den folgenden Befehl ein, um den Manager zu starten und die Protokolldatei `ndmgr.log` anzugeben (Pfade können nicht angegeben werden):
`sscontrol manager start ndmgr.log`
- Geben Sie den folgenden Befehl ein, um den aktuellen Status der Werte anzuzeigen, die dem Manager zugeordnet sind:
`sscontrol manager status`

Dieser Befehl erzeugt eine Ausgabe, die ungefähr wie folgt aussieht:

Manager-Status:

=====

```
Metrik-Port ..... 10004
Name der Managerprotokolldatei ..... manager.log
Managerprotokollstufe ..... 1
Max. Managerprotokollgröße (Bytes) ..... unlimited
Sensitivitätsstufe ..... 5
Glättungsfaktor ..... 1,5
Aktualisierungsintervall (Sekunden) ..... 2
Gewichtungsaktualisierungszyklus ..... 2
Erreichbarkeit - Protokollstufe ..... 1
Erreichbarkeit - Max. Protokollgröße (Bytes) ..... unlimited
Erreichbarkeit - Aktualisierungsintervall (Sekunden) ... 7
```

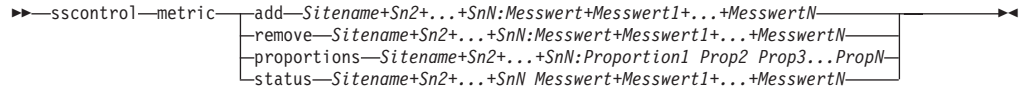
- Stoppen des Managers:

```
sscontrol manager stop
```

- Geben Sie den folgenden Befehl ein, um die aktuelle Versionsnummer des Managers aufzurufen:

```
sscontrol manager version
```

sscontrol metric — Systemmesswerte konfigurieren



add

Hinzufügen des angegebenen Messwerts.

Sitename

Der konfigurierte Sitename. Zusätzliche Sitenamen werden durch ein Pluszeichen (+) voneinander getrennt.

Messwert

Name des Systemmesswerts. Es muss sich um den Namen einer ausführbaren Datei oder Script-Datei im Verzeichnis des Messwertservers handeln.

remove

Entfernen des angegebenen Messwerts.

proportions

Der Parameter proportions gibt die Wichtigkeit an, die jedem Messwert verglichen mit anderen zugeordnet wird, wenn die Messwerte kombiniert werden, um die Systembelastung eines Servers zu ermitteln.

status

Anzeigen der aktuellen Serverwerte für diesen Messwert.

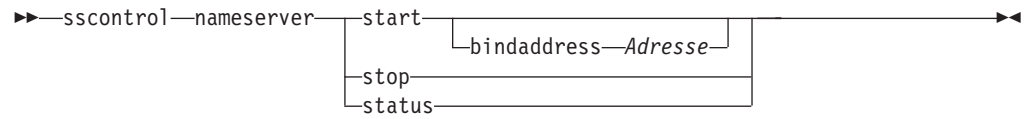
Beispiele

- Hinzufügen eines Systemmesswerts:
`sscontrol metric add Site1:Messwert1`
- Festlegen der Proportionen für einen Sitenamen mit zwei Systemmesswerten:
`sscontrol metric proportions Site1 0 100`
- Anzeigen des aktuellen Status der zugeordneten Messwerte:
`sscontrol metric status Site1:Messwert1`

Dieser Befehl erzeugt eine Ausgabe, die ungefähr wie folgt aussieht:

```
Metrikstatus:
-----
Sitename ..... Site1
Metrikname ..... Messwert1
Metrikproportion ..... 50
  Server ..... 9.37.56.100
  Metrikdaten .... -1
```

sscontrol nameserver — Namensserver steuern



start

Starten des Namensservers.

bindaddress

Startet den Namensserver, der an die angegebene Adresse gebunden ist. Der Namensserver antwortet nur auf Anfragen, die an diese Adresse gerichtet sind.

Adresse

Eine auf der Maschine mit Site Selector konfigurierte Adresse (IP-Adresse oder symbolischer Name).

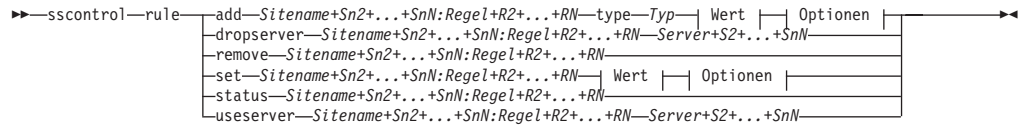
stop

Stoppt den Namensserver.

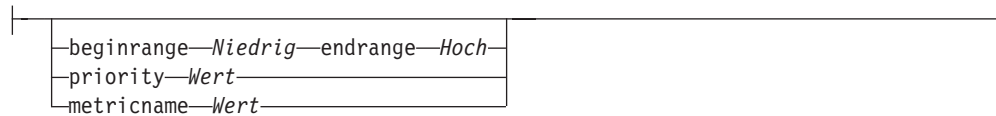
status

Zeigt den Status des Namensservers an.

sscontrol rule — Regeln konfigurieren



Optionen:



add

Diese Regel zu einem Sitenamen hinzufügen.

Sitename

Ein nicht auflösbarer Hostname, den der Client abfragt. Der Sitename muss ein vollständig qualifizierter Domänenname sein. Zusätzliche Sitenamen werden durch ein Pluszeichen (+) voneinander getrennt.

Regel

Der für die Regel ausgewählte Name. Dieser Name kann eine beliebige Kombination aus alphanumerischen Zeichen, Unterstreichungszeichen, Silbentrennungsstrichen und Punkten sein. Der Name kann 1 bis 20 Zeichen lang sein und darf keine Leerzeichen enthalten.

Anmerkung: Zusätzliche Regeln werden durch ein Pluszeichen (+) voneinander getrennt.

type

Der Regeltyp.

Typ

Die Auswahlmöglichkeiten für *Typ* sind:

ip Die Regel basiert auf der Client-IP-Adresse.

metrical

Die Regel basiert auf dem aktuellen Messwert für alle Server der Gruppe.

metricavg

Die Regel basiert auf dem Durchschnitt der aktuellen Messwerte für alle Server der Gruppe.

time Die Regel basiert auf der Uhrzeit.

true Diese Regel ist immer wahr. Sie kann mit der Anweisung ELSE in der Programmierlogik verglichen werden.

beginrange

Der untere Wert des Bereichs, mit dem bestimmt wird, ob die Regel wahr ist oder nicht.

Niedrig

Hängt von der Art der Regel ab. Die Art des Werts und der Standardwert werden nachfolgend nach der Art der Regel aufgelistet:

ip Die Adresse des Clients als symbolischer Name oder im IP-Adressenformat. Der Standardwert ist 0.0.0.0.

Zeit Eine ganze Zahl. Der Standardwert ist 0. Er stellt Mitternacht dar.

metricall
Eine ganze Zahl. Der Standardwert ist 100.

metricavg
Eine ganze Zahl. Der Standardwert ist 100.

endrange

Der obere Wert des Bereichs, mit dem bestimmt wird, ob die Regel wahr ist oder nicht.

Hoch

Hängt von der Art der Regel ab. Die Art des Werts und der Standardwert werden nachfolgend nach der Art der Regel aufgelistet:

ip Die Adresse des Clients als symbolischer Name oder im IP-Adressenformat. Der Standardwert ist 255.255.255.254.

Zeit Eine ganze Zahl. Der Standardwert ist 24. Er stellt Mitternacht dar.

Anmerkung: Beim Definieren des Bereichsanfangs (*beginrange*) und Bereichsendes (*endrange*) der Zeitintervalle ist darauf zu achten, dass jeder Wert eine ganze Zahl sein muss, die nur den Stundenteil der Uhrzeit darstellt. Es werden keine Teilwerte einer Stunde angegeben. Soll beispielsweise die Stunde von 3 Uhr bis 4 Uhr angegeben werden, geben Sie für *beginrange* 3 und für *endrange* ebenfalls 3 an. Damit werden alle Minuten von 3 Uhr bis 3 Uhr 59 angegeben. Wird *beginrange* auf 3 und *endrange* auf 4 gesetzt, ergibt sich ein fast zweistündiger Zeitraum von 3.00 Uhr bis 4.59 Uhr.

metricall
Eine ganze Zahl. Der Standardwert ist 2 hoch 32 minus 1.

metricavg
Eine ganze Zahl. Der Standardwert ist 2 hoch 32 minus 1.

priority

Die Reihenfolge, in der die Regeln überprüft werden.

Stufe

Eine ganze Zahl. Wenn Sie für die erste Regel, die Sie hinzufügen, keine Priorität angeben, setzt Site Selector die Priorität standardmäßig auf 1. Wird eine weitere Regel hinzugefügt, wird deren Priorität standardmäßig mit der folgenden Formel errechnet: 10 + derzeit niedrigste Priorität für alle vorhandenen Regeln. Beispiel: Sie haben eine Regel mit der Priorität 30. Sie fügen nun eine neue Regel hinzu und setzen deren Priorität auf 25 (also auf eine *höhere* Priorität als 30). Anschließend fügen Sie eine dritte Regel ohne Angabe einer Priorität hinzu. Als +Priorität der dritten Regel wird 40 ermittelt (30 + 10).

metricname

Name des für eine Regel gemessenen Messwerts.

dropserver

Einen Server aus einem Regelsatz entfernen.

Server

Die IP-Adresse der TCP-Servermaschine als symbolischer Name oder im IP-Adressenformat.

Anmerkung: Zusätzliche Sitenamen werden durch ein Pluszeichen (+) voneinander getrennt.

remove

Eine oder mehrere Regeln, die durch Pluszeichen voneinander getrennt sind, entfernen.

set

Werte für diese Regel festlegen.

status

Anzeigen aller Werte für eine oder mehrere Regel(n).

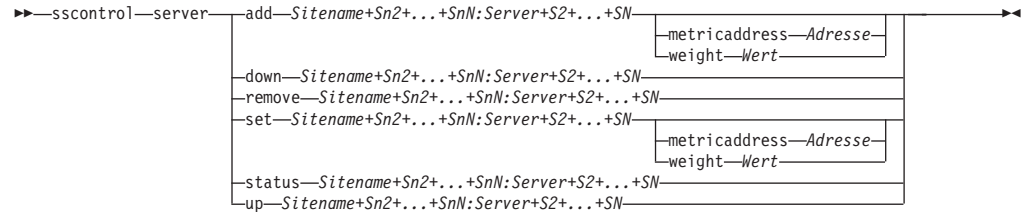
useserver

Server in einen Regelsatz einfügen.

Beispiele

- Hinzufügen einer immer gültigen Regel ohne Angabe von Bereichsanfang oder -ende:
`sscontrol rule add Sitename:Regelname type true priority 100`
- Erstellen einer Regel, die den Zugriff auf einen Bereich von IP-Adressen unterbindet, der in diesem Fall mit "9" beginnt:
`sscontrol rule add Sitename:Regelname type ip b 9.0.0.0 e 9.255.255.255`
- Soll eine Regel erstellt werden, die die Verwendung eines bestimmten Servers in der Zeit von 11 Uhr bis 15 Uhr angibt, den folgenden Befehl eingeben:
`sscontrol rule add Sitename:Regelname type time beginrange 11 endrange 14`
`sscontrol rule useserver Sitename:Regelname Server05`

sscontrol server — Server konfigurieren



add

Diesen Server hinzufügen.

Sitename

Ein nicht auflösbarer Hostname, den der Client abfragt. Der Sitename muss ein vollständig qualifizierter Domänenname sein. Zusätzliche Sitenamen werden durch ein Pluszeichen (+) voneinander getrennt.

Server

Die IP-Adresse der TCP-Servermaschine als symbolischer Name oder im IP-Adressenformat.

Anmerkung: Zusätzliche Server werden durch ein Pluszeichen (+) voneinander getrennt.

metricaddress

Die Adresse des Metric-Servers.

Adresse

Die Adresse des Servers als symbolischer Name oder im IP-Adressenformat.

weight

Eine Zahl von 0 bis 100 (die den Wert für die Wertigkeitsobergrenze des angegebenen Sitenamens nicht überschreiten darf) zur Angabe der Gewichtung dieses Servers. Wird die Wertigkeit auf 0 gesetzt, werden keine neuen Anforderungen an den Server gesendet. Der Standardwert entspricht der Hälfte der Wertigkeitsobergrenze für den angegebenen Sitenamen. Ist der Manager aktiv, wird diese Einstellung schnell überschrieben.

Wert

Die Wertigkeit des Servers.

down

Diesen Server als inaktiv markieren. Dieser Befehl verhindert, dass weitere Anforderungen an diesen Server übergeben werden.

remove

Diesen Server entfernen.

set

Werte für diesen Server festlegen.

status

Status der Server anzeigen.

up Diesen Server als aktiv markieren. Site Selector übergibt neue Anforderungen an diesen Server.

Beispiele

- Hinzufügen des Servers 27.65.89.42 zum Sitenamen Site1:
`sscontrol
server add Site1:27.65.89.42`
- Geben Sie den folgenden Befehl ein, um den Server an 27.65.89.42 als "ausgefallen" zu kennzeichnen:
`sscontrol server down Site1:27.65.89.42`
- Entfernen des Servers 27.65.89.42 für alle Sitenamen:
`sscontrol server remove :27.65.89.42`
- Geben Sie den folgenden Befehl ein, um den Server an 27.65.89.42 als "aktiv" zu kennzeichnen:
`sscontrol server up Site1:27.65.89.42`

sscontrol set — Serverprotokoll konfigurieren



loglevel

Die Stufe für die Protokollierung von sserver-Aktivitäten.

Stufe

Der Standardwert für **loglevel** ist 0. Gültige Werte sind:

- 0 entspricht keiner Protokollierung
- 1 entspricht einer minimalen Protokollierung
- 2 entspricht einer Basisprotokollierung
- 3 entspricht einer normalen Protokollierung
- 4 entspricht einer erweiterten Protokollierung
- 5 entspricht einer ausführlichen Protokollierung.

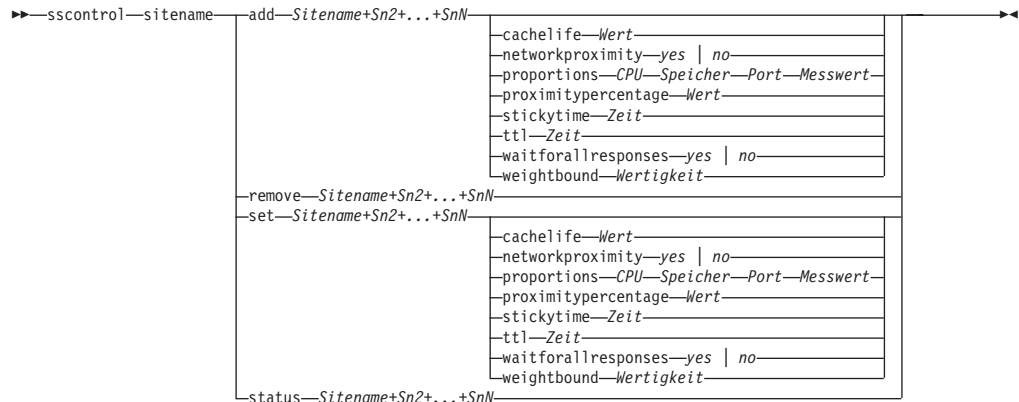
logsize

Die maximale Anzahl von Byte, die in der Protokolldatei protokolliert werden können.

Größe

Der Standardwert für logsize ist 1 MB.

sscontrol sitename — Sitenamen konfigurieren



add

Hinzufügen eines neuen Sitenamens.

Sitename

Ein nicht auflösbarer Hostname, der vom Client angefragt wird. Zusätzliche Sitenamen werden durch ein Pluszeichen (+) voneinander getrennt.

cachelife

Die Zeitperiode, während der eine Proximitätsantwort gültig und im Cache gespeichert ist. Der Standardwert ist 1800. Weitere Informationen hierzu finden Sie im Abschnitt „Netzproximität verwenden“ auf Seite 130.

Wert

Eine positive Zahl, die angibt, wie viele Sekunden eine Proximitätsantwort gültig ist und im Cache gespeichert wird.

networkproximity

Legt die Netzproximität des Servers zum anfordernden Client fest. Verwenden Sie diese Proximitätsantwort bei der Lastausgleichsentscheidung. Die Proximität ist "Ein" oder "Aus". Weitere Informationen hierzu finden Sie im Abschnitt „Netzproximität verwenden“ auf Seite 130.

Wert

Zur Auswahl stehen yes oder no. Der Standardwert no bedeutet, dass die Netzproximität inaktiviert ist.

proportions

Festlegen der proportionalen Bedeutung von CPU, Speicher, Port (Advisor-Informationen) und Systemmesswerten für den Metric Server. Anhand dieser Proportionen legt der Manager die Serverwertigkeiten fest. Jeder dieser Werte wird als Prozentsatz der Summe angegeben, die immer bei 100 liegt.

CPU Prozentsatz der auf jeder am Lastausgleich beteiligten Servermaschine genutzten CPU (Vorgabe vom Agenten Metric Server).

Speicher

Prozentsatz des auf jeder am Lastausgleich beteiligten Servermaschine genutzten Speichers (Vorgabe vom Agenten Metric Server).

Port Vorgaben von den am Port empfangsbereiten Advisor-Funktionen.

Messwert

Vorgaben von Metric Server.

proximitypercentage

Legt die Bedeutung der Proximitätsantwort, gemessen am Status des Servers (Wertigkeit vom Manager), fest. Weitere Informationen hierzu finden Sie im Abschnitt „Netzproximität verwenden“ auf Seite 130.

Wert

Der Standardwert ist 50.

stickytime

Das Intervall, in dem ein Client dieselbe Server-ID empfängt, die zuvor auf die erste Anfrage zurückgegeben wurde. Der Standardwert für stickytime ist 0 und bedeutet, dass für den Sitenamen keine Haltezeit gilt.

Zeit

Eine positive Zahl ungleich null, die die Zeit in Sekunden angibt, in der der Client dieselbe Server-ID empfängt, die zuvor auf die erste Anfrage zurückgegeben wurde.

ttl Legt die Lebensdauer fest. Dieser Parameter gibt an, wie lange ein anderer Namensserver die aufgelöste Antwort zwischenspeichert. Der Standardwert ist 5.

Wert

Eine positive Zahl, die angibt, wie viele Sekunden der Namensserver die aufgelöste Antwort zwischenspeichert.

waitforallresponses

Legt fest, ob vor der Beantwortung der Client-Anfrage auf alle Proximitätsantworten der Server gewartet werden soll. Weitere Informationen hierzu finden Sie im Abschnitt „Netzproximität verwenden“ auf Seite 130.

Wert

Zur Auswahl stehen yes oder no. Der Standardwert ist yes.

weightbound

Eine Zahl, die die maximale Wertigkeit angibt, die für Server dieser Site festgelegt werden kann. Der für den Sitenamen festgelegte weightbound-Wert kann mit **server weight** für einzelne Server überschrieben werden. Der Standardwert für sitename weightbound ist 20.

Wertigkeit

Die Wertigkeitsgrenze.

set

Festlegen der Merkmale für den Sitenamen.

remove

Entfernen dieses Sitenamens.

status

Anzeigen des aktuellen Status für einen bestimmten Sitenamen.

Beispiele

- Hinzufügen eines Sitenamens:
`sscontrol sitename add 130.40.52.153`
- Aktivieren der Netzproximität:
`sscontrol sitename set meineSite networkproximity yes`
- Festlegen einer Caching-Zeit von 1900000 Sekunden:
`sscontrol sitename set meineSite cachelife 1900000`

- Festlegen eines Proximitätsprozentsatzes von 45:
sscontrol sitename set meineSeite proximitypercentage 45
- Festlegung für einen Sitenamen, dass vor einer Reaktion nicht auf alle Antworten gewartet werden soll:
sscontrol
sitename set meineSite waitforallresponses no
- Festlegen einer Lebensdauer von 7 Sekunden:
sscontrol sitename set meineSite ttl 7
- Festlegen der proportionalen Bedeutung von cpuload, memload, Port und Systemmesswert:
sscontrol sitename
set meineSite proportions 50 48 1 1
- Entfernen eines Sitenamens:
sscontrol sitename remove 130.40.52.153
- Anzeigen des Status für den Sitenamen meineSite:
sscontrol sitename status meineSite

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Status des Sitenamens:

```
Sitename ..... meineSite
Gewichtungsgrenze ..... 20
TTL ..... 5
Haltezeit ..... 0
Anzahl Server ..... 1
Proportion für CpuLoad ..... 49
Proportion für MemLoad ..... 50
Proportion für Port ..... 1
Proportion für Systemmetrik ..... 0
Advisor am Port ..... 80
Verwendete Proximität ..... N
```

sscontrol status — Aktivitätsanzeige für Manager und Advisor-Funktionen

►►—sscontrol—status—◄◄

Beispiele

- Geben Sie den folgenden Befehl ein, um festzustellen, welche Funktionen ausgeführt werden:

```
sscontrol status
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
Namensserver wurde gestartet.  
Manager wurde gestartet.
```

	ADVISOR		SITENAME:PORT		ZEITLIMIT	
	http		80		unlimited	

Kapitel 29. Befehlsreferenz für Cisco CSS Controller

Dieses Kapitel beschreibt die Verwendung der folgenden **ccocontrol**-Befehle für Cisco CSS Controller:

- „ccocontrol consultant — Consultant konfigurieren und steuern“ auf Seite 434
- „ccocontrol controller — Controller steuern“ auf Seite 437
- „ccocontrol file — Konfigurationsdateien verwalten“ auf Seite 439
- „ccocontrol help — Hilfetext für diesen Befehl anzeigen oder drucken“ auf Seite 440
- „ccocontrol highavailability — Hohe Verfügbarkeit steuern“ auf Seite 441
- „ccocontrol metriccollector — Messwerterfassung konfigurieren“ auf Seite 444
- „ccocontrol ownercontent — Eignernamen und content-Regel steuern“ auf Seite 446
- „ccocontrol service — Service konfigurieren“ auf Seite 449

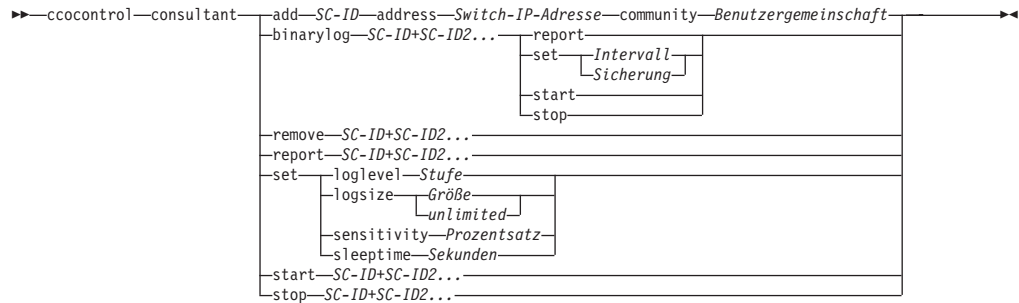
Für die Parameter des Befehls **ccocontrol** können Sie die abgekürzte Form verwenden. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **ccocontrol he f** an Stelle von **ccocontrol help file** angeben.

Geben Sie zum Aufrufen der **ccocontrol**-Eingabeaufforderung **ccocontrol** ein.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie **exit** oder **quit** eingeben.

Anmerkung: Für alle Parameterwerte des Befehls müssen Sie die englischen Zeichen verwenden. Die einzige Ausnahme hiervon bilden Hostnamen (die in den server-Befehlen verwendet werden) und (die in Dateibefehlen verwendeten) Dateinamen.

cococontrol consultant — Consultant konfigurieren und steuern



add

Fügt einen Switch-Consultant hinzu.

SC-ID (Switch-Consultant-ID)

Eine benutzerdefinierte Zeichenfolge, die sich auf den Consultant bezieht.

address

Die IP-Adresse des Cisco CSS Switch, für den der Consultant Wertigkeiten bereitstellt.

Switch-IP-Adresse

Die IP-Adresse des Switch.

community

Der in SNMP für das Anfordern und Festlegen der Kommunikation mit dem Cisco CSS Switch verwendete Name.

Benutzergemeinschaft

Der Name der Benutzergemeinschaft mit Schreib-/Lesezugriff für den Cisco CSS Switch.

binarylog

Steuert die binäre Protokollierung für einen Consultant.

report

Berichte mit den Kenndaten der binären Protokollierung.

set

Legt fest, wie oft (alle wie viel Sekunden) Daten in die binären Protokolle geschrieben werden. Die binäre Protokollierung ermöglicht das Speichern von Informationen zu jedem in der Konfiguration definierten Service in binären Protokolldateien. Die Daten werden nur in die Protokolle geschrieben, wenn seit dem Schreiben des letzten Protokolleintrags die für das Protokollintervall angegebene Zeit in Sekunden verstrichen ist. Das Standardintervall für binäre Protokollierung ist 60.

Intervall

Gibt in Sekunden die Zeit zwischen den Einträgen im binären Protokoll an.

Sicherung

Legt die Zeit in Stunden fest, die binäre Protokolldateien aufbewahrt werden.

start

Startet die binäre Protokollierung.

stop

Stoppt die binäre Protokollierung.

remove

Entfernt einen Switch-Consultant.

report

Berichte mit den Kenndaten von Switch-Consultants.

set

Legt die Kenndaten von Switch-Consultants fest.

loglevel

Legt die Protokollstufe fest, auf der der Switch-Consultant Aktivitäten protokolliert. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe von 0 bis 5. Die Standardeinstellung ist 1. Gültige Werte sind:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe

Die maximale Anzahl Bytes, die im Consultant-Protokoll protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

sensitivity

Gibt an, in welchem Maße sich eine Wertigkeit ändern muss, damit die Änderung als relevant angesehen wird. Der Unterschied zwischen der neuen und der alten Wertigkeit muss größer als der hier festgelegte Prozentsatz sein, bevor die Wertigkeit geändert wird. Der gültige Bereich liegt zwischen 0 und 100. Der Standardwert ist 5.

Prozentsatz

Eine ganze Zahl von 0 bis 100 zur Angabe der Sensitivität.

sleeptime

Legt die Zeit der Inaktivität zwischen den Definitionszyklen für die Wertigkeit in Sekunden fest. Der Standardwert ist 7.

Sekunden

Eine ganze Zahl, die die Inaktivität in Sekunden angibt. Der gültige Bereich beginnt mit 0 und endet mit 2.147.460.

start

Startet das Erfassen von Messwerten und das Definieren von Wertigkeiten.

stop

Stoppt das Erfassen von Messwerten und das Definieren von Wertigkeiten.

Beispiele

- Hinzufügen eines Switch-Consultant mit der Switch-ID SC1, der IP-Adresse 9.37.50.17 und der Benutzergemeinschaft Comm1:

```
cococontrol consultant add SC1 address 9.37.50.17 community Comm1
```

- Starten der binären Protokollierung:

```
cococontrol consultant binarylog SC1 start
```

- Anzeigen eines Berichts mit den Kenndaten des Switch-Consultant SC1:

```
cococontrol consultant report SC1
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Die Verbindung zwischen Consultant SC1 und Switch 9.37.50.1:Comm1 wurde hergestellt

Der Consultant wurde gestartet.

Ruhezeit = 7

Sensitivität = 5

Protokollstufe = 5

Protokollgröße = 1.048.576

Eignerangaben:

Eignerangaben EA1

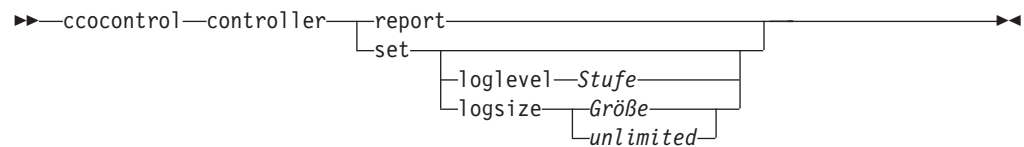
- Festlegen einer Inaktivitätszeit von 10 Sekunden zwischen den Definitionszyklen für die Wertigkeit für Switch-ID SC1:

```
cococontrol consultant set SC1 sleeptime 10
```

- Starten der Erfassung von Messwerten und des Definierens von Wertigkeiten für die Consultant-ID SC1:

```
cococontrol consultant start SC1
```

cococontrol controller — Controller steuern



report

Anzeigen der Kenndaten des Controllers. Dieser Bericht enthält auch die Versionsnummer.

set

Festlegen der Kenndaten des Controllers.

loglevel

Legt die Protokollstufe fest, auf der der Controller Aktivitäten protokolliert. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe von 0 bis 5. Die Standardeinstellung ist 1. Gültige Werte sind:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe | unlimited

Die maximale Anzahl Bytes, die im Consultant-Protokoll protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

Beispiele

- Anzeigen eines Berichts zum Controller:

```
cococontrol controller report
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Controller-Bericht:

Version Version: 05.00.00.00 - 03/21/2002-09:49:57-EST

Protokollstufe 1

Protokollgröße 1048576

Konfigurationsdatei . . . config1.xml

Consultants:

Consultant consult1 gestartet

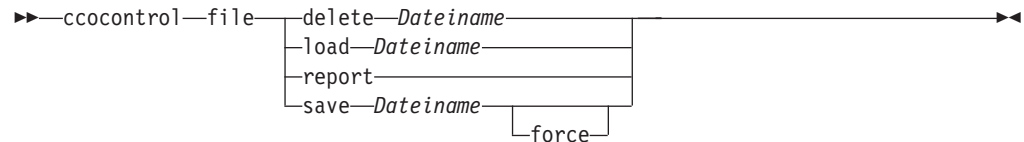
- Festlegen der Protokollstufe null zur Verbesserung des Durchsatzes:

```
cococontrol set loglevel 0
```

- Festlegen einer Controllerprotokollgröße von 1.000.000:

```
cococontrol controller set logsize 1000000
```

cococontrol file — Konfigurationsdateien verwalten



delete

Löschen der angegebenen Konfigurationsdatei.

Dateiname

Eine Konfigurationsdatei. Die Dateierweiterung muss .xml lauten. Wenn diese Erweiterung nicht angegeben wird, wird sie vorausgesetzt.

load

Laden der in der angegebenen Datei gespeicherten Konfiguration.

Anmerkung: Beim Laden einer Datei wird die darin gespeicherte Konfiguration an die aktive Konfiguration angehängt. Wenn Sie eine *neue* Konfiguration laden möchten, müssen Sie vor dem Laden der Datei den Server beenden und neu starten.

report

Auflisten der Konfigurationsdateien.

save

Sichern der aktuellen Konfiguration in der angegebenen Datei.

Anmerkung: Dateien werden in den nachfolgend genannten Verzeichnissen gespeichert und aus diesen geladen:

- AIX-Systeme: `/opt/ibm/edge/lb/servers/configurations/cco`
- Linux-Systeme: `/opt/ibm/edge/lb/servers/configurations/cco`
- Solaris-Systeme: `/opt/ibm/edge/lb/servers/configurations/cco`
- Windows-Systeme:

Installationsverzeichnis (Standard): `C:\Programme\ibm\edge\lb\servers\configurations\cco`

force

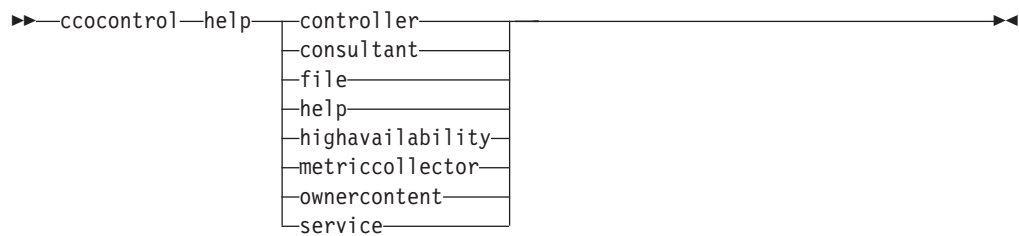
Speichern in einer vorhandenen Datei.

Beispiele

- Löschen der Datei Datei1:
`cococontrol file delete Datei1`
- Anhängen der Konfiguration in der Datei an die aktuelle Konfiguration:
`cococontrol file load Konfig2`
- Anzeigen eines Berichts zu den bisher gespeicherten Dateien:
`cococontrol file report`
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:
DATEIBERICHT:

Datei1.xml
Datei2.xml
Datei3.xml
- Speichern der Konfiguration in der Datei Konfig2.xml:
`cococontrol file save Konfig2`

cococontrol help — Hilfetext für diesen Befehl anzeigen oder drucken



Beispiele

- Hilfetext zum Befehl `ccocontrol` abrufen:

```
cococontrol help
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Die folgenden Befehle sind verfügbar:

controller	- für den Controller
consultant	- für Switch-Consultants
file	- für Konfigurationsdateien
help	- für Hilfetexte
highavailability	- für Hochverfügbarkeit
metriccollector	- für die Metrik-Erfassungsprogramme
ownerContent	- für Eignerangaben
service	- für Services

- In den Syntaxdiagrammen der Onlinehilfe werden die folgenden Symbole benutzt:

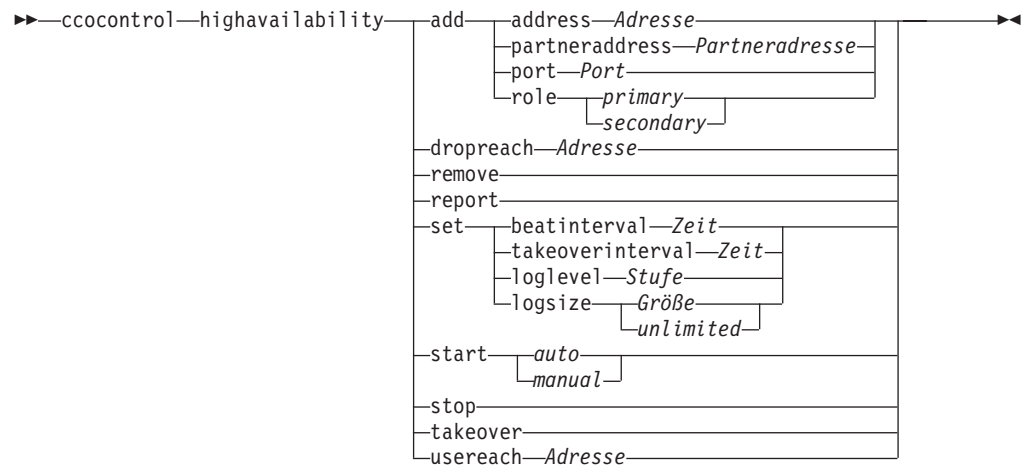
< > Parameter oder Zeichenfolgen sind in spitze Klammern gesetzt.

[] Optionale Elemente sind in eckige Klammern gesetzt.

Ein vertikaler Strich trennt alternative Angaben in eckigen und spitzen Klammern.

: Ein Doppelpunkt trennt Namen voneinander. Beispiel:
Consultant1:Eignerangaben1.

cococontrol highavailability — Hohe Verfügbarkeit steuern



add

Konfigurieren eines Knotens mit hoher Verfügbarkeit, des zugehörigen Partners und von Erreichbarkeitszielen.

address

Die Adresse, von der Überwachungssignale empfangen werden sollen.

Adresse

Die IP-Adresse des Knotens mit hoher Verfügbarkeit.

partneraddress

Die Adresse, an die Überwachungssignale gesendet werden sollen. Dies ist die IP-Adresse oder der Hostname, die bzw. der für den Partnerknoten konfiguriert wurde. Diese Adresse wird verwendet, um mit der Partnermaschine für hohe Verfügbarkeit zu kommunizieren.

Partneradresse

Die IP-Adresse der Partnermaschine.

port

Der für die Kommunikation mit dem Partner verwendete Port. Der Standardwert ist 12345.

Port

Die Port-Nummer.

role

Die Rolle für hohe Verfügbarkeit.

primary | secondary

Die primäre oder sekundäre Rolle.

dropreach

Entfernen dieses Erreichbarkeitsziels aus den Kriterien für hohe Verfügbarkeit.

Adresse

Die IP-Adresse des Erreichbarkeitsziels.

remove

Entfernen des Knotens, des Partners und des Erreichbarkeitsziels aus der Konfiguration für hohe Verfügbarkeit. Vor Verwendung dieses Befehls muss die hohe Verfügbarkeit inaktiviert werden.

report

Anzeigen von Informationen zur hohen Verfügbarkeit.

set

Festlegen der Kenndaten für die hohe Verfügbarkeit.

beatinterval

Legt fest, in welchem Abstand (in Millisekunden) Überwachungssignale an den Partner gesendet werden. Der Standardwert ist 500.

Zeit

Eine positive ganze Zahl, die das Intervall für die Signale in Millisekunden angibt.

takeoverinterval

Legt die Zeit in Millisekunden fest, nach der eine Übernahme erfolgt. (Während dieser Zeit werden keine Überwachungssignale empfangen.) Der Standardwert ist 2000.

Zeit

Eine positive ganze Zahl, die das Übernahmeintervall in Millisekunden angibt.

loglevel

Legt die Stufe fest, auf der Aktivitäten protokolliert werden. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe von 0 bis 5. Die Standardeinstellung ist 1. Gültige Werte sind:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei für hohe Verfügbarkeit protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe | unlimited

Die maximale Anzahl Bytes, die im Protokoll für hohe Verfügbarkeit protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

start

Aktiviert die hohe Verfügbarkeit. Vor Verwendung dieses Befehls müssen Sie einen Knoten mit hoher Verfügbarkeit, den zugehörigen Partner und ein Erreichbarkeitsziel konfiguriert haben.

auto | *manual*

Legt fest, ob die hohe Verfügbarkeit mit automatischer oder manueller Wiederherstellung aktiviert werden soll.

stop

Inaktiviert die hohe Verfügbarkeit.

takeover

Der Knoten mit hoher Verfügbarkeit gibt die Steuerung ab.

usereach

Die Adresse des Erreichbarkeitsziels, an dem die hohe Verfügbarkeit aktiviert wird. Fügen Sie ein Erreichbarkeitsziel hinzu, dass mit ping erreicht werden kann, damit die Partner für hohe Verfügbarkeit feststellen können, ob ihre Zieladressen tatsächlich erreichbar sind.

Adresse

Die IP-Adresse des Erreichbarkeitsziels.

Beispiele

- Hinzufügen eines Knotens mit hoher Verfügbarkeit, dessen IP-Adresse 9.37.50.17 lautet, der am Port 12345 die primäre Rolle hat und dessen Partner die Adresse 9.37.50.14 hat:

```
cococontrol highavailability add
address 9.37.50.17 role primary port 12345 partneraddress 9.37.50.14
```

- Hinzufügen eines Erreichbarkeitsziels mit der Adresse 9.37.50.9:

```
cococontrol highavailability usereach 9.37.50.9
```

- Entfernen des Erreichbarkeitsziels mit der Adresse 9.37.50.9:

```
cococontrol highavailability dropreach 9.37.50.9
```

- Aktivieren der hohen Verfügbarkeit mit manueller Wiederherstellung:

```
cococontrol highavailability start manual
```

- Abrufen einer statistischen Momentaufnahme der hohen Verfügbarkeit:

```
cococontrol highavailability report
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

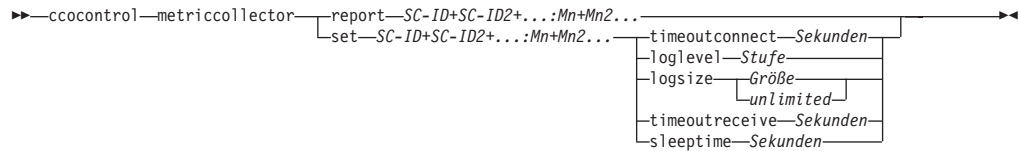
Hochverfügbarkeitsstatus:

```
Knoten . . . . . primär
Knotenadresse . . . . . 9.37.50.17
Port . . . . . 12345
Partneradresse . . . . . 9.37.50.14
Wiederherstellungsstrategie . . . . . manuell
Intervall für Überwachungssignal . . . . . 500
Übernahmeintervall . . . . . 2000
Status . . . . . inaktiv
Teilstatus . . . . . nicht synchronisiert
```

Erreichbarkeitsstatus: Knoten/Partner

Keine Erreichbarkeitsziele konfiguriert.

cococontrol metriccollector — Messwerterfassung konfigurieren



report

Anzeigen der Kenndaten eines Erfassungsprogramms für Messwerte.

SC-ID (Switch-Consultant-ID)

Eine benutzerdefinierte Zeichenfolge, die sich auf den Consultant bezieht.

Mn (Messwertname)

Ein Name, der den vorgegebenen oder benutzerdefinierten Messwert bezeichnet.

set

Festlegen der Kenndaten eines Erfassungsprogramms für Messwerte.

timeoutconnect

Festlegen der Zeit, nach deren Ablauf ein Erfassungsprogramm für Messwerte meldet, dass eine Verbindung unterbrochen ist.

Sekunden

Eine positive ganze Zahl, die die Zeit in Sekunden angibt, nach deren Ablauf das Erfassungsprogramm für Messwerte meldet, dass zu einem Service keine Verbindung hergestellt werden kann.

loglevel

Legt die Protokollstufe fest, auf der der angegebene Consultant Aktivitäten protokolliert. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe. Der Standardwert ist 1. Je größer die Zahl ist, desto mehr Informationen werden in das Consultant-Protokoll geschrieben. Gültige Werte:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe | unlimited

Die maximale Anzahl Bytes, die im Consultant-Protokoll protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

timeoutreceive

Legt die Zeit fest, nach deren Ablauf der Consultant meldet, dass von einem Service keine Daten empfangen werden können.

Sekunden

Eine positive ganze Zahl, die die Zeit in Sekunden angibt, nach deren Ablauf der Consultant meldet, dass von einem Service keine Daten empfangen werden können.

sleeptime

Legt für das Erfassungsprogramm für Messwerte die Zeit der Inaktivität zwischen den Erfassungszyklen in Sekunden fest.

Eine positive ganze Zahl, die die Inaktivitätszeit in Sekunden angibt.

Beispiele

- Anzeigen eines Berichts mit den Kenndaten eines Erfassungsprogramms für Messwerte:

```
cococontrol metriccollector report SC1:http
```

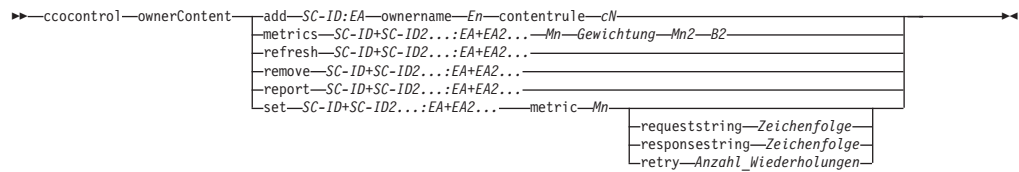
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
MetricCollector SC1:http
  Erfasste Metrik ..... http
  Protokollstufe ..... 5
  Protokollgröße ..... 1048576
  Ruhezeit in Sekunden ..... 7
  Verbindungszeitlimit in Sekunden ... 21
  Empfangszeitlimit in Sekunden ..... 21
```

- Festlegen eines Verbindungszeitlimits von 15 Sekunden, einer unbegrenzten Protokollgröße für den Switch-Consultant SC1 und des Messwerts http:

```
cococontrol metriccollector set SC1:http timeoutconnect 15 logsize unlimited
```


cococontrol ownercontent — Eigernamen und content-Regel steuern



add

Fügt Eignerangaben zum angegebenen Consultant hinzu.

SC-ID (Switch-Consultant-ID)

Eine benutzerdefinierte Zeichenfolge, die den Consultant bezeichnet.

EA (Name der Eignerangaben)

Eine benutzerdefinierte Zeichenfolge, die für den Switch den Eigernamen und die content-Regel repräsentiert.

ownername

Der auf dem Switch für die Eignerkonfiguration definierte Name.

En (Eigername)

Eine eindeutige Zeichenfolge ohne Leerzeichen. Der Eigername muss mit dem auf dem Cisco-Switch angegebenen übereinstimmen.

contentrule

Der auf dem Switch für die content-Regel des Eigners definierte Name.

cN (content-Name)

Eine eindeutige Zeichenfolge ohne Leerzeichen. Der content-Name muss mit dem auf dem Cisco-Switch angegebenen übereinstimmen.

metrics

Gibt die Messwerte an, die für die Berechnung der Wertigkeiten herangezogen werden sollen, sowie die Gewichtung der einzelnen Messwerte. Die Gewichtung wird als Prozentsatz ausgedrückt. Die Summe aller Gewichtungen muss 100 ergeben. Die Messwerte können eine beliebige Kombinationen aus Messwerten für Verbindungsdaten, für Advisor-Funktionen der Anwendung und für Metric Server sein. Die Standardmesswerte sind activeconn (aktive Verbindungen) und connrate (Verbindungsrate) mit einer Gewichtung von jeweils 50 %.

Mn (Messwertname)

Name des Erfassungsprogramms für Messwerte, das die Messwerte für die Bestimmung der Serverwertigkeit erfasst.

Nachfolgend sehen Sie eine Liste gültiger Messwertnamen mit den zugehörigen Ports.

Advisor-Name	Protokoll	Port
connect	ICMP	12345
db2	privat	50000
dns	DNS	53
ftp	FTP	21
http	HTTP	80
https	SSL	443

Advisor-Name	Protokoll	Port
cachingproxy	HTTP (über Caching Proxy)	80
imap	IMAP	143
ldap	LDAP	389
nntp	NNTP	119
ping	PING	0
pop3	POP3	110
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
telnet	Telnet	23
wlm	privat	10007
activeconn	nicht anwendbar	nicht anwendbar
connrate	nicht anwendbar	nicht anwendbar
cpuload	nicht anwendbar	nicht anwendbar
memload	nicht anwendbar	nicht anwendbar

Gewichtung

Eine Zahl von 0 bis 100, die die Bedeutung dieses Messwerts für die Berechnung der Serverwertigkeiten angibt.

refresh

Aktualisiert die konfigurierten Services mit der Konfiguration vom Cisco CSS Switch.

remove

Entfernen von Eignerangaben.

report

Ausgeben eines Berichts mit den Kenndaten der Eignerangaben.

set

Festlegen der Kenndaten von Eignerangaben.

metric

Festlegen der Kenndaten eines Messwerts.

Mn

Der Name des gewünschten Messwerts.

requeststring

Festlegen einer Anfragezeichenfolge für den angegebenen Messwert. Diese Zeichenfolge repräsentiert die Anfrage, die ein Erfassungsprogramm zum Abrufen von Messdaten sendet.

Zeichenfolge

Die vom Erfassungsprogramm für Messwerte an den Server gesendete Anfragezeichenfolge.

responsestring

Festlegen einer Antwortzeichenfolge für den angegebenen Messwert. Die angegebene Antwortzeichenfolge vergleicht das Erfassungsprogramm für Messwerte mit den Antworten, die es von den Servern empfängt, und ermittelt daraufhin die Serververfügbarkeit.

Zeichenfolge

Die Antwortzeichenfolge, mit der das Erfassungsprogramm für Messwerte vom Server empfangene Antworten vergleicht.

retry

Der Parameter retry legt die Wiederholungsversuche fest, bevor ein Server als inaktiv markiert wird.

Anzahl_Wiederholungen

Eine ganze Zahl größer als oder gleich null. Dieser Wert sollte nicht größer als 3 sein. Wenn das Schlüsselwort retry nicht konfiguriert ist, wird standardmäßig von null Wiederholungsversuchen ausgegangen.

Beispiele

- Hinzufügen der Eignerangaben EA1 (mit dem Eigernamen Eigner1 und dem content-Namen content1) zum Switch-Consultant mit der ID SC1:

```
cococontrol ownerContent add SC1:EA1 ownername Eigner1 contentrule content1
```

- Angeben einer Proportion von 50:50 für die Messwerte activeconn und http:

```
cococontrol ownerContent metrics SC1:EA1 activeconn 50 http 50
```

- Anzeigen eines Berichtes mit den Kerndaten des Eigners:

```
cococontrol ownerContent report SC1:EA1
```

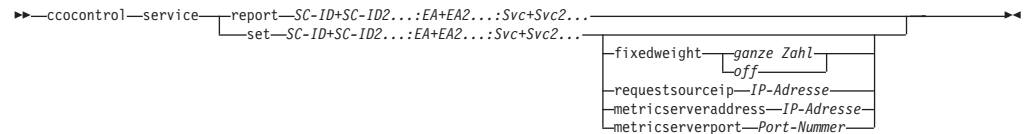
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
ownerContent SC1:EA1
  Gewichtungsgrenze = 10
  Metrik activeconn hat die Proportion 25
    Antwortzeichenfolge... nicht anwendbar
    Anforderungszeichenfolge... nicht anwendbar
  Metrik http hat die Proportion 50
    Antwortzeichenfolge... nicht anwendbar
    Anforderungszeichenfolge... nicht anwendbar
  Metrik connrate hat die Proportion 25
    Antwortzeichenfolge... nicht anwendbar
    Anforderungszeichenfolge... nicht anwendbar
  Enthält Service t3
  Enthält Service t2
  Enthält Service t1
```

- Festlegen einer Anfragezeichenfolge für http:

```
cococontrol ownerContent set SC1:EA1 metric http requeststring getCookie
```

cococontrol service — Service konfigurieren



report

Anzeigen der Kenndaten von Services.

SC-ID (Switch-Consultant-ID)

Eine benutzerdefinierte Zeichenfolge, die den Consultant bezeichnet.

EA (Name der Eignerangaben)

Eine benutzerdefinierte Zeichenfolge, die für den Switch den Eigernamen und die content-Regel repräsentiert.

Svc (Service)

Eine benutzerdefinierte Zeichenfolge auf dem Switch, die den Service repräsentiert.

set

Festlegen der Kenndaten von Services.

fixedweight

Definieren einer festen Wertigkeit für diesen Service. Der Standardwert ist off.

ganze Zahl | off

Eine positive ganze Zahl von 0 bis 10, die die feste Wertigkeit für diesen Service angibt, oder das Wort **off**, um anzugeben, dass es keine feste Wertigkeit gibt.

requestsourceip

Festlegen der Adresse, von der aus Anwendungsanfragen an den Service zu richten sind.

IP-Adresse

Die IP-Adresse, von der aus der Kontakt zum Service hergestellt werden soll, als symbolischer Name oder im IP-Adressenformat.

metricserveraddress

Festlegen der Adresse, unter der der Service für Anfragen an Metric Server erreichbar ist.

IP-Adresse

Die IP-Adresse von Metric Server als symbolischer Name oder im IP-Adressenformat.

metricserverport

Festlegen des Ports, von dem aus der Kontakt zum Metric Server hergestellt werden soll.

Port-Nummer

Die Nummer des Ports, von dem aus der Kontakt zum Metric Server hergestellt wird.

Beispiele

- Anzeigen eines Berichts zum Service t1 für den Consultant SC1:

```
ccocontrol service report SC1:EA1:t1
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
Service SC1:EA1:t1 hat die Gewichtung 10
Feste Gewichtung ist off
IP der Anforderungsquelle ..... 9.27.24.156
Port der Anwendung ..... 80
Adresse des MetricServer ..... 1.0.0.1
Port des MetricServer ..... 10004
  Metrik activeconn hat den Wert -99
  Metrik http hat den Wert -99
  Metrik connrate hat den Wert -99
```

- Festlegen einer Metric-Server-Adresse für den Service t2:

```
ccocontrol service set SC1:EA1:t2 metricserveraddress 9.37.50.17
```

Kapitel 30. Befehlsreferenz für Nortel Alteon Controller

Dieses Kapitel beschreibt die Verwendung der folgenden **nalcontrol**-Befehle für Nortel Alteon Controller:

- „**nalcontrol consultant** — Consultant konfigurieren und steuern“ auf Seite 452
- „**nalcontrol controller** — Controller steuern“ auf Seite 455
- „**nalcontrol file** — Konfigurationsdateien verwalten“ auf Seite 457
- „**nalcontrol help** — Hilfetext für diesen Befehl anzeigen oder drucken“ auf Seite 458
- „**nalcontrol highavailability** — Hohe Verfügbarkeit steuern“ auf Seite 459
- „**nalcontrol metriccollector** — Messwerterfassung konfigurieren“ auf Seite 462
- „**nalcontrol service** — Service konfigurieren“ auf Seite 466
- „**nalcontrol server** — Server konfigurieren“ auf Seite 464

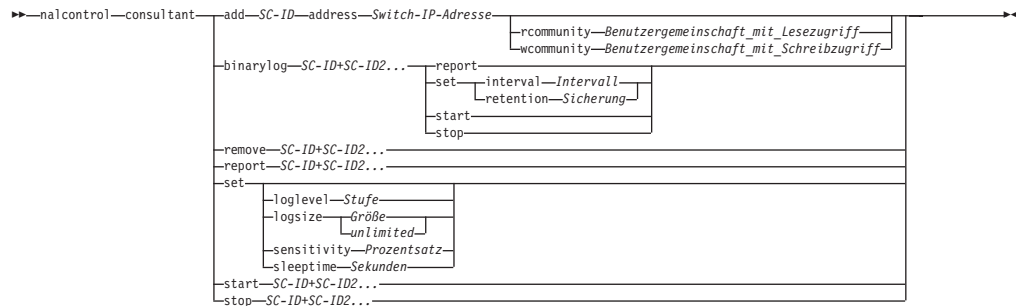
Für die Parameter des Befehls **nalcontrol** können Sie die abgekürzte Form verwenden. Sie müssen nur die eindeutigen Buchstaben der Parameter eingeben. Beispiel: Wenn Sie Hilfe für den Befehl zum Speichern von Dateien aufrufen möchten, können Sie **nalcontrol he f** an Stelle von **nalcontrol help file** angeben.

Geben Sie zum Aufrufen der **nalcontrol**-Eingabeaufforderung **nalcontrol** ein.

Sie können die Befehlszeilenschnittstelle verlassen, indem Sie **exit** oder **quit** eingeben.

Anmerkung: Für alle Parameterwerte des Befehls müssen Sie die englischen Zeichen verwenden. Die einzige Ausnahme hiervon bilden Hostnamen (die in den **server**-Befehlen verwendet werden) und (die in Dateibefehlen verwendeten) Dateinamen.

nalcontrol consultant — Consultant konfigurieren und steuern



add

Fügt einen Switch-Consultant hinzu.

SC-ID

Eine benutzerdefinierte Zeichenfolge, die sich auf den Consultant bezieht.

address

Die IP-Adresse des Nortel Alteon Web Switch, für den der Consultant Wertigkeiten bereitstellt.

Switch-IP-Adresse

Die IP-Adresse des Switch.

rcommunity

Der für SNMP-Kommunikation mit dem Nortel Alteon Web Switch verwendete Name der Benutzergemeinschaft mit Lesezugriff. Der Standardwert ist public.

Benutzergemeinschaft_mit_Lesezugriff

Die Zeichenfolge für den Namen der Benutzergemeinschaft mit Lesezugriff, wie sie für den Nortel Alteon Web Switch konfiguriert ist. Der Standardwert ist public.

wcommunity

Der für SNMP-Kommunikation verwendete Name der Benutzergemeinschaft mit Schreibzugriff.

Benutzergemeinschaft_mit_Schreibzugriff

Die Zeichenfolge für den Namen der Benutzergemeinschaft mit Schreibzugriff, wie sie für den Nortel Alteon Web Switch konfiguriert ist. Der Standardwert ist private.

binarylog

Steuert die binäre Protokollierung für einen Consultant.

report

Berichte mit den Kenndaten der binären Protokollierung.

set

Legt fest, wie oft (alle wie viel Sekunden) Daten in die binären Protokolle geschrieben werden. Die binäre Protokollierung ermöglicht das Speichern von Informationen zu jedem in der Konfiguration definierten Service in binären Protokolldateien. Die Daten werden nur in die Protokolle geschrieben, wenn seit dem Schreiben des letzten Protokolleintrags die für das Protokollintervall angegebene Zeit in Sekunden verstrichen ist. Das Standardintervall für binäre Protokollierung ist 60.

interval

Gibt in Sekunden die Zeit zwischen den Einträgen im binären Protokoll an.

retention

Legt die Zeit in Stunden fest, die binäre Protokolldateien aufbewahrt werden.

start

Startet die binäre Protokollierung.

stop

Stoppt die binäre Protokollierung.

remove

Entfernt einen Switch-Consultant.

report

Berichte mit den Kenndaten von Switch-Consultants.

set

Legt die Kenndaten von Switch-Consultants fest.

loglevel

Legt die Protokollstufe fest, auf der der Switch-Consultant Aktivitäten protokolliert. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe von 0 bis 5. Die Standardeinstellung ist 1. Gültige Werte sind:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe

Die maximale Anzahl Bytes, die im Consultant-Protokoll protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

sensitivity

Gibt an, in welchem Maße sich eine Wertigkeit ändern muss, damit die Änderung als relevant angesehen wird. Der Unterschied zwischen der neuen und der alten Wertigkeit muss größer als der hier festgelegte Prozentsatz sein, bevor die Wertigkeit geändert wird. Der gültige Bereich liegt zwischen 0 und 100. Der Standardwert ist 5.

Prozentsatz

Eine ganze Zahl von 0 bis 100 zur Angabe der Sensitivität.

sleeptime

Legt die Zeit der Inaktivität zwischen den Definitionszyklen für die Wertigkeit in Sekunden fest. Der Standardwert ist 7.

Sekunden

Eine ganze Zahl, die die Inaktivität in Sekunden angibt. Der gültige Bereich beginnt mit 0 und endet mit 2.147.460.

start

Startet das Erfassen von Messwerten und das Definieren von Wertigkeiten.

stop

Stoppt das Erfassen von Messwerten und das Definieren von Wertigkeiten.

Beispiele

- Hinzufügen eines Switch-Consultant mit der Switch-ID SC1 und der IP-Adresse 9.37.50.17:

```
nalcontrol consultant add SC1 address 9.37.50.17
```

- Starten der binären Protokollierung:

```
nalcontrol consultant binarylog SC1 start
```

- Anzeigen eines Berichts mit den Kenndaten des Switch-Consultant SC1:

```
nalcontrol consultant report SC1
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Consultant-ID: SC1 IP-Adresse des Switch: 9.37.50.1

Community mit Lesezugriff: public

Community mit Schreibzugriff: private

Der Consultant wurde gestartet.

Ruhezeit = 7

Sensitivität = 5

Protokollstufe = 5

Protokollgröße = 1.048.576

Services:

Service Svc1

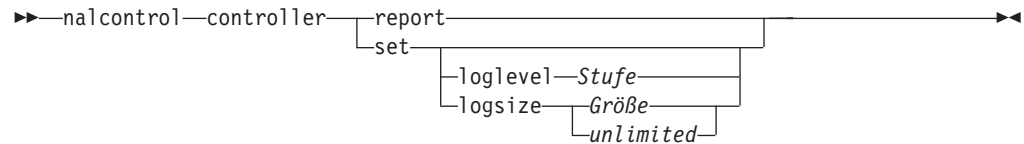
- Festlegen einer Inaktivitätszeit von 10 Sekunden zwischen den Definitionszyklen für die Wertigkeit für Switch-ID SC1:

```
nalcontrol consultant set SC1 sleeptime 10
```

- Starten der Erfassung von Messwerten und des Definierens von Wertigkeiten für die Consultant-ID SC1:

```
nalcontrol consultant start sc1
```

nalcontrol controller — Controller steuern



report

Anzeigen der Kenndaten des Controllers. Dieser Bericht enthält auch die Versionsnummer.

set

Festlegen der Kenndaten des Controllers.

loglevel

Legt die Protokollstufe fest, auf der der Controller Aktivitäten protokolliert. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe von 0 bis 5. Die Standardeinstellung ist 1. Gültige Werte sind:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe | unlimited

Die maximale Anzahl Bytes, die im Consultant-Protokoll protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

Beispiele

- Anzeigen eines Berichts zum Controller:

```
nalcontrol controller report
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Controller-Bericht:

Version Version: 05.00.00.00 - 03/21/2002-09:49:57-EST

Protokollstufe 1

Protokollgröße 1048576

Konfigurationsdatei . . . config1.xml

Consultants:

Consultant consult1 gestartet

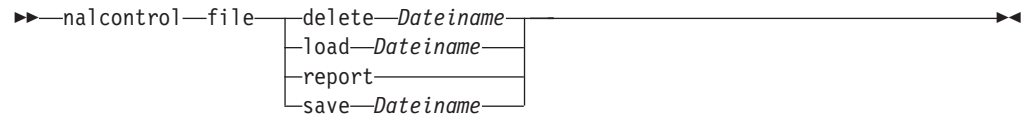
- Festlegen der Protokollstufe null zur Verbesserung des Durchsatzes:

```
nalcontrol set loglevel 0
```

- Festlegen einer Controllerprotokollgröße von 1.000.000:

```
nalcontrol controller set logsize 1000000
```

nalcontrol file — Konfigurationsdateien verwalten



delete

Löschen der angegebenen Konfigurationsdatei.

Dateiname

Eine Konfigurationsdatei. Die Dateierweiterung muss .xml lauten. Wenn diese Erweiterung nicht angegeben wird, wird sie vorausgesetzt.

load

Laden der in der angegebenen Datei gespeicherten Konfiguration.

Anmerkung: Beim Laden einer Datei wird die darin gespeicherte Konfiguration an die aktive Konfiguration angehängt. Wenn Sie eine *neue* Konfiguration laden möchten, müssen Sie vor dem Laden der Datei den Server beenden und neu starten.

report

Auflisten der Konfigurationsdateien.

save

Sichern der aktuellen Konfiguration in der angegebenen Datei.

Anmerkung: Dateien werden in den nachfolgend genannten Verzeichnissen gespeichert und aus diesen geladen:

- AIX-Systeme: `/opt/ibm/edge/lb/servers/configurations/nal`
- Linux-Systeme: `/opt/ibm/edge/lb/servers/configurations/nal`
- Solaris-Systeme: `/opt/ibm/edge/lb/servers/configurations/nal`
- Windows-Systeme:

Allgemeiner Installationsverzeichnispfad — `C:\Programme\ibm\edge\lb\servers\configurations\nal`

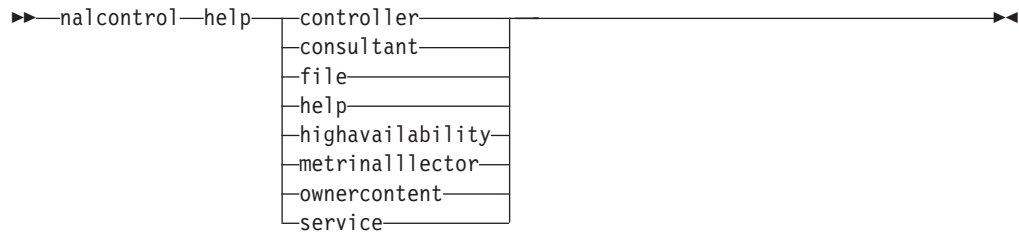
Interner Installationsverzeichnispfad — `C:\Programme\ibm\lb\servers\configurations\nal`

Beispiele

- Löschen der Datei Date1:
`nalcontrol file delete Date1`
- Laden einer neuen Konfigurationsdatei, die die aktuelle Konfiguration ersetzen soll:
`nalcontrol file load Konfig2`
- Anzeigen eines Berichts zu den bisher gespeicherten Dateien:
`nalcontrol file report`
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:
DATEIBERICHT:

Date1.xml
Datei2.xml
Datei3.xml
- Speichern der Konfiguration in der Datei Konfig2:
`nalcontrol file save Konfig2`

nalcontrol help — Hilfetext für diesen Befehl anzeigen oder drucken



Beispiele

- Hilfetext zum Befehl `nalcontrol` abrufen:

```
nalcontrol help
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Die folgenden Befehle sind verfügbar:

<code>controller</code>	- für den Controller
<code>consultant</code>	- für Switch-Consultants
<code>file</code>	- für Konfigurationsdateien
<code>help</code>	- für Hilfetexte
<code>highavailability</code>	- für Hochverfügbarkeit
<code>metriccollector</code>	- für die Metrik-Erfassungsprogramme
<code>server</code>	- für Server
<code>service</code>	- für Services

- In den Syntaxdiagrammen der Onlinehilfe werden die folgenden Symbole benutzt:

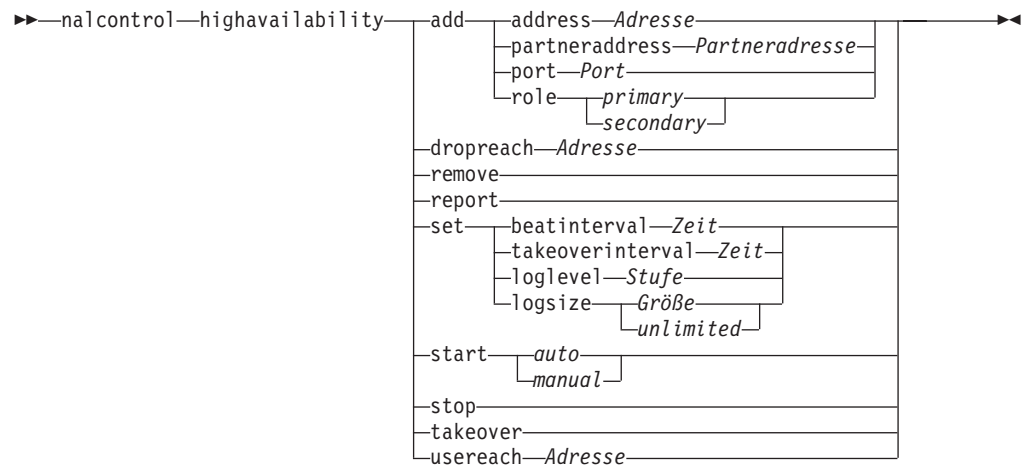
`< >` Parameter oder Zeichenfolgen sind in spitze Klammern gesetzt.

`[]` Optionale Elemente sind in eckige Klammern gesetzt.

`|` Ein vertikaler Strich trennt alternative Angaben in eckigen und spitzen Klammern.

`:` Ein Doppelpunkt trennt Namen voneinander. Beispiel:
Consultant1:Service1.

nalcontrol highavailability — Hohe Verfügbarkeit steuern



add

Konfigurieren eines Knotens mit hoher Verfügbarkeit, des zugehörigen Partners und von Erreichbarkeitszielen.

address

Die Adresse, von der Überwachungssignale empfangen werden sollen.

Adresse

Die IP-Adresse des Knotens mit hoher Verfügbarkeit.

partneraddress

Die Adresse, an die Überwachungssignale gesendet werden sollen. Dies ist die IP-Adresse oder der Hostname, die bzw. der für den Partnerknoten konfiguriert wurde. Diese Adresse wird verwendet, um mit der Partnermaschine für hohe Verfügbarkeit zu kommunizieren.

Partneradresse

Die IP-Adresse der Partnermaschine.

port

Der für die Kommunikation mit dem Partner verwendete Port. Der Standardwert ist 12345.

Port

Die Port-Nummer.

role

Die Rolle für hohe Verfügbarkeit.

primary | secondary

Die primäre oder sekundäre Rolle.

dropreach

Entfernen dieses Erreichbarkeitsziels aus den Kriterien für hohe Verfügbarkeit.

Adresse

Die IP-Adresse des Erreichbarkeitsziels.

remove

Entfernen des Knotens, des Partners und des Erreichbarkeitsziels aus der Konfiguration für hohe Verfügbarkeit. Vor Verwendung dieses Befehls muss die hohe Verfügbarkeit inaktiviert werden.

report

Anzeigen von Informationen zur hohen Verfügbarkeit.

set

Festlegen der Kenndaten für die hohe Verfügbarkeit.

beatinterval

Legt fest, in welchem Abstand (in Millisekunden) Überwachungssignale an den Partner gesendet werden. Der Standardwert ist 500.

Zeit

Eine positive ganze Zahl, die das Intervall für die Signale in Millisekunden angibt.

takeoverinterval

Legt die Zeit in Millisekunden fest, nach der eine Übernahme erfolgt. (Während dieser Zeit werden keine Überwachungssignale empfangen.) Der Standardwert ist 2000.

Zeit

Eine positive ganze Zahl, die das Übernahmeintervall in Millisekunden angibt.

loglevel

Legt die Stufe fest, auf der Aktivitäten protokolliert werden. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe von 0 bis 5. Die Standardeinstellung ist 1. Gültige Werte sind:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei für hohe Verfügbarkeit protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe | unlimited

Die maximale Anzahl Bytes, die im Protokoll für hohe Verfügbarkeit protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

start

Aktiviert die hohe Verfügbarkeit. Vor Verwendung dieses Befehls müssen Sie einen Knoten mit hoher Verfügbarkeit, den zugehörigen Partner und ein Erreichbarkeitsziel konfiguriert haben.

auto | *manual*

Legt fest, ob die hohe Verfügbarkeit mit automatischer oder manueller Wiederherstellung aktiviert werden soll.

stop

Inaktiviert die hohe Verfügbarkeit.

takeover

Der Knoten mit hoher Verfügbarkeit gibt die Steuerung ab.

usereach

Die Adresse des Erreichbarkeitsziels, an dem die hohe Verfügbarkeit aktiviert wird. Fügen Sie ein Erreichbarkeitsziel hinzu, dass mit ping erreicht werden kann, damit die Partner für hohe Verfügbarkeit feststellen können, ob ihre Zieladressen tatsächlich erreichbar sind.

Adresse

Die IP-Adresse des Erreichbarkeitsziels.

Beispiele

- Hinzufügen eines Knotens mit hoher Verfügbarkeit, dessen IP-Adresse 9.37.50.17 lautet, der am Port 12345 die primäre Rolle hat und dessen Partner die Adresse 9.37.50.14 hat:

```
nalcontrol highavailability add
  address 9.37.50.17 role primary port 12345 partneraddress 9.37.50.14
```

- Hinzufügen eines Erreichbarkeitsziels mit der Adresse 9.37.50.9:

```
nalcontrol highavailability usereach 9.37.50.9
```

- Entfernen des Erreichbarkeitsziels mit der Adresse 9.37.50.9:

```
nalcontrol highavailability dropreach 9.37.50.9
```

- Aktivieren der hohen Verfügbarkeit mit manueller Wiederherstellung:

```
nalcontrol highavailability start manual
```

- Abrufen einer statistischen Momentaufnahme der hohen Verfügbarkeit:

```
nalcontrol highavailability report
```

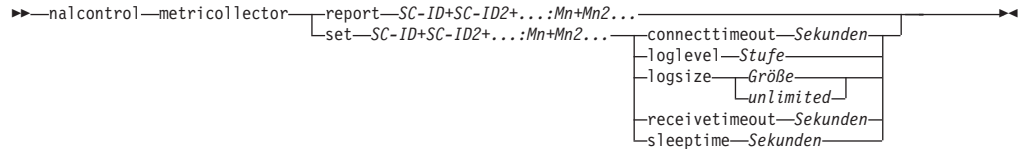
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

Hochverfügbarkeitsstatus:

```
Knoten . . . . . primär
Knotenadresse . . . . . 9.37.50.17
Port . . . . . 12345
Partneradresse . . . . . 9.37.50.14
Wiederherstellungsstrategie . . . . . manuell
Intervall für Überwachungssignal . . . . . 500
Übernahmeintervall . . . . . 2000
Gestartet . . . . . N
Status . . . . . inaktiv
Teilstatus . . . . . nicht synchronisiert
```

Erreichbarkeitsstatus: Knoten/Partner

nalcontrol metriccollector — Messwerterfassung konfigurieren



report

Anzeigen der Kenndaten eines Erfassungsprogramms für Messwerte.

SC-ID (Switch-Consultant-ID)

Eine benutzerdefinierte Zeichenfolge, die sich auf den Consultant bezieht.

Mn (Messwertname)

Ein Name, der den vorgegebenen oder benutzerdefinierten Messwert bezeichnet.

set

Festlegen der Kenndaten eines Erfassungsprogramms für Messwerte.

connecttimeout

Festlegen der Zeit, nach deren Ablauf ein Erfassungsprogramm für Messwerte meldet, dass eine Verbindung unterbrochen ist.

Sekunden

Eine positive ganze Zahl, die die Zeit in Sekunden angibt, nach deren Ablauf das Erfassungsprogramm für Messwerte meldet, dass zu einem Service keine Verbindung hergestellt werden kann.

loglevel

Legt die Protokollstufe fest, auf der der angegebene Consultant Aktivitäten protokolliert. Der Standardwert ist 1.

Stufe

Die Nummer der Stufe. Der Standardwert ist 1. Je größer die Zahl ist, desto mehr Informationen werden in das Consultant-Protokoll geschrieben. Gültige Werte:

- 0 = Keine
- 1 = Minimal
- 2 = Grundlegend
- 3 = Mäßig
- 4 = Erweitert
- 5 = Ausführlich

logsize

Legt die maximale Anzahl Bytes fest, die in der Protokolldatei protokolliert werden. Der Standardwert ist 1048576. Wenn Sie für die Protokolldatei eine maximale Größe festlegen, wird die Protokollierung bei Erreichen der Größe am Anfang der Datei fortgesetzt, so dass die vorherigen Protokolleinträge überschrieben werden. Die Protokollgröße kann nicht auf einen geringeren Wert als die aktuelle Dateigröße gesetzt werden. Protokolleinträge sind mit einer Zeitmarke versehen, damit Sie die Reihenfolge, in der sie geschrieben wurden, erkennen können. Je höher Sie die Protokollstufe setzen, desto vorsichtiger müssen Sie die Protokollgröße auswählen, da die Protokolldatei sehr schnell voll ist, wenn Sie eine Protokollierung auf einer höheren Stufe wählen.

Größe | unlimited

Die maximale Anzahl Bytes, die im Consultant-Protokoll protokolliert werden. Sie können entweder eine positive Zahl, die größer als 0 sein muss, oder das Wort **unlimited** angeben. Da die Protokolleinträge unterschiedlich lang sind, wird mit dem Überschreiben von Einträgen unter Umständen schon vor Erreichen der exakten maximalen Größe begonnen.

receivetimeout

Legt die Zeit fest, nach deren Ablauf der Consultant meldet, dass von einem Service keine Daten empfangen werden können.

Sekunden

Eine positive ganze Zahl, die die Zeit in Sekunden angibt, nach deren Ablauf der Consultant meldet, dass von einem Service keine Daten empfangen werden können.

sleeptime

Legt für das Erfassungsprogramm für Messwerte die Zeit der Inaktivität zwischen den Erfassungszyklen in Sekunden fest.

Sekunden

Eine positive ganze Zahl, die die Inaktivitätszeit in Sekunden angibt.

Beispiele

- Anzeigen eines Berichts mit den Kenndaten eines Erfassungsprogramms für Messwerte:

```
nalcontrol metrinallector report sc1:http
```

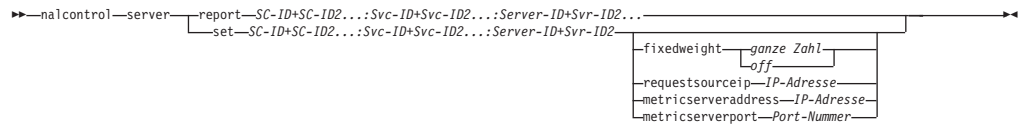
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
Metrinallector sc1:http
  Erfasste Metrik ..... http
  Protokollstufe ..... 5
  Protokollgröße ..... 1048576
  Ruhezeit in Sekunden ..... 7
  Verbindungszeitlimit in Sekunden ... 21
  Empfangszeitlimit in Sekunden ..... 21
```

- Festlegen eines Verbindungszeitlimits von 15 Sekunden, einer unbegrenzten Protokollgröße für den Switch-Consultant SC1 und des Messwerts http:

```
nalcontrol metrinallector set SC1:http connecttimeout 15 logsize unlimited
```

nalcontrol server — Server konfigurieren



report

Anzeigen der Kenndaten von Servern.

SC-ID

Eine benutzerdefinierte Zeichenfolge, die den Consultant bezeichnet.

Svc-ID

Eine benutzerdefinierte Zeichenfolge für die Kennung des virtuellen Services und die Nummer des virtuellen Ports auf dem Switch.

Server-ID

Eine ganze Zahl, die den Server auf dem Switch repräsentiert.

set

Festlegen der Kenndaten von Servern.

fixedweight

Definieren einer festen Wertigkeit für diesen Server. Der Standardwert ist off. Der Maximalwert für fixedweight ist 48.

ganze Zahl | off

Eine positive ganze Zahl, die die feste Wertigkeit für diesen Server angibt, oder das Wort **off**, um anzugeben, dass es keine feste Wertigkeit gibt.

requestsourceip

Festlegen der Adresse, von der aus Anwendungsanfragen an den Server zu richten sind.

IP-Adresse

Die IP-Adresse, von der aus der Kontakt zum Server hergestellt werden soll, als symbolischer Name oder im IP-Adressenformat.

metricserveraddress

Festlegen der Adresse, von der aus Metric-Server-Anfragen an den Server zu richten sind.

IP-Adresse

Die IP-Adresse von Metric Server als symbolischer Name oder im IP-Adressenformat.

metricserverport

Festlegen des Ports, von dem aus der Kontakt zum Metric Server hergestellt werden soll.

Port-Nummer

Die Nummer des Ports, von dem aus der Kontakt zum Metric Server hergestellt wird.

Beispiele

- Anzeigen eines Berichts zu Server 1 für den Consultant SC1:

```
nalcontrol server report SC1:Svc1:1
```

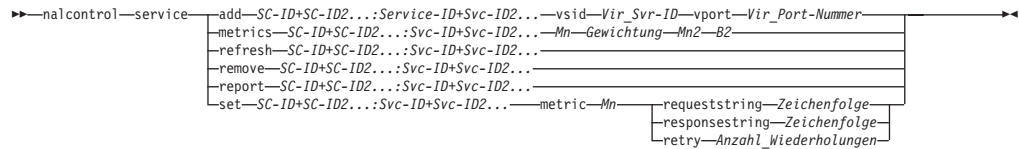
Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
Server SC1:Svc1:1 hat die Gewichtung -99
  Feste Gewichtung ist off
  IP der Anforderungsquelle ..... 9.27.24.156
  Port der Anwendung ..... 99
  Adresse des MetricServer ..... 9.99.99.98
  Port des MetricServer ..... 10004
    Metrik activeconn hat den Wert -99
    Metrik connrate hat den Wert -99
```

- Festlegen einer Metric-Server-Adresse für den Service 2:

```
nalcontrol server set SC1:Svc1:2 metricserveraddress 9.37.50.17
```

nalcontrol service — Service konfigurieren



add

Fügt einen Service zum angegebenen Consultant hinzu.

SC-ID (Switch-Consultant-ID)

Eine benutzerdefinierte Zeichenfolge, die sich auf den Consultant bezieht.

Svc-ID (Service-ID)

Eine benutzerdefinierte Zeichenfolge, die den Service bezeichnet.

vsid

Das Schlüsselwort, das den virtuellen Service bezeichnet.

Vir_Svr-ID (ID des virtuellen Servers)

Die Zahl, die auf dem Switch den virtuellen Server repräsentiert.

vport

Das Schlüsselwort für den virtuellen Port.

Vir_Port-Nummer (Nummer des virtuellen Ports)

Die derzeit auf dem Switch konfigurierte Port-Nummer für den Service.

metrics

Gibt die Messwerte an, die für die Berechnung der Wertigkeiten herangezogen werden sollen, sowie die Gewichtung der einzelnen Messwerte. Die Gewichtung wird als Prozentsatz ausgedrückt. Die Summe aller Gewichtungen muss 100 ergeben. Die Messwerte können eine beliebige Kombinationen aus Messwerten für Verbindungsdaten, für Advisor-Funktionen der Anwendung und für Metric Server sein. Die Standardmesswerte sind activeconn (aktive Verbindungen) und connrate (Verbindungsrate) mit einer Gewichtung von jeweils 50 %.

Mn (Messwertname)

Name des Erfassungsprogramms für Messwerte, das die Messwerte für die Bestimmung der Serverwertigkeit erfasst.

Nachfolgend sehen Sie eine Liste gültiger Messwertnamen mit den zugehörigen Ports.

Advisor-Name	Protokoll	Port
connect	ICMP	12345
db2	privat	50000
dns	DNS	53
ftp	FTP	21
http	HTTP	80
https	SSL	443
cachingproxy	HTTP (über Caching Proxy)	80
imap	IMAP	143
ldap	LDAP	389

Advisor-Name	Protokoll	Port
nntp	NNTP	119
ping	PING	0
pop3	POP3	110
sip	SIP	5060
smtp	SMTP	25
ssl	SSL	443
telnet	Telnet	23
WLM	privat	10007
activeconn	nicht anwendbar	nicht anwendbar
connrate	nicht anwendbar	nicht anwendbar
cpuload	nicht anwendbar	nicht anwendbar
memload	nicht anwendbar	nicht anwendbar

Gewichtung

Eine Zahl von 0 bis 100, die die Bedeutung dieses Messwerts für die Berechnung der Serverwertigkeiten angibt.

refresh

Aktualisieren eines Services mit Informationen vom Nortel Alteon Web Switch.

remove

Entfernen eines Services.

report

Ausgeben eines Berichts mit den Kenndaten eines Services.

set

Festlegen der Kenndaten eines Services.

metric

Festlegen der Kenndaten eines konfigurierten Messwerts.

Mn (Messwertname)

Der Name des gewünschten Messwerts.

requeststring

Festlegen einer Anfragezeichenfolge für den angegebenen Messwert. Diese Zeichenfolge repräsentiert die Anfrage, die ein Erfassungsprogramm zum Abrufen von Messdaten sendet.

Zeichenfolge

Die vom Erfassungsprogramm für Messwerte an den Server gesendete Anfragezeichenfolge.

responsestring

Festlegen einer Antwortzeichenfolge für den angegebenen Messwert. Die angegebene Antwortzeichenfolge vergleicht das Erfassungsprogramm für Messwerte mit den Antworten, die es von den Servern empfängt, und ermittelt daraufhin die Serververfügbarkeit.

Zeichenfolge

Die Antwortzeichenfolge, mit der das Erfassungsprogramm für Messwerte vom Server empfangene Antworten vergleicht.

retry

Der Parameter `retry` legt die Wiederholungsversuche fest, bevor ein Server als inaktiv markiert wird.

Anzahl_Wiederholungen

Eine ganze Zahl größer als oder gleich null. Dieser Wert sollte nicht größer als 3 sein. Wenn das Schlüsselwort für Wiederholungen nicht konfiguriert ist, wird standardmäßig von null Wiederholungsversuchen ausgegangen.

Beispiele

- Hinzufügen des Services `Svc1` (ID des virtuellen Servers: 1, Nummer des virtuellen Ports: 80) zum Switch-Consultant mit der ID `SC1`:

```
nalcontrol service add SC1:Svc1 vsid 1 vport 80
```

- Angeben einer Proportion von 50:50 für die Messwerte `activeconn` und `http`:

```
nalcontrol service metrics SC1:Svc1 activeconn 50 http 50
```

- Anzeigen eines Berichtes mit den Kenndaten des Eigners:

```
nalcontrol service report SC1:Svc1
```

Die Ausgabe dieses Befehls sieht etwa wie folgt aus:

```
Service SC1:Svc1
Gewichtungsgrenze = 48
Metrik activeconn hat die Proportion 50
Metrik connrate hat die Proportion 50
Enthält Server 4
Enthält Server 3
Enthält Server 2
Enthält Server 1
```

- Festlegen einer Anfragezeichenfolge für `http`:

```
nalcontrol service set SC1:Svc1 metric http requeststring getLastErrorCode
```

Anhang A. Allgemeine Anweisungen zur GUI

Auf der grafischen Benutzerschnittstelle (GUI) von Load Balancer erscheint auf der linken Seite der Anzeige eine Baumstruktur mit Load Balancer als Ausgangsebene und Dispatcher, Content Based Routing (CBR), Site Selector; Cisco CSS Controller sowie Nortel Alteon Controller als Komponenten.

Wenn Sie Load Balancer für IPV4 und IPV6 verwenden, steht nur die Dispatcher-Komponente zur Verfügung. Weitere Informationen hierzu finden Sie in Kapitel 8, „Deployment des Dispatchers von Load Balancer für IPV4 und IPV6“, auf Seite 81.

Beispiele der GUI von Load Balancer, die sich auf die verschiedenen Komponenten beziehen, zeigen die folgenden Abbildungen:

- Abb. 41 auf Seite 470 für Dispatcher
- Abb. 42 auf Seite 471 für CBR
- Abb. 43 auf Seite 472 für Site Selector
- Abb. 44 auf Seite 473 für Cisco CSS Controller
- Abb. 45 auf Seite 474 für Nortel Alteon Controller

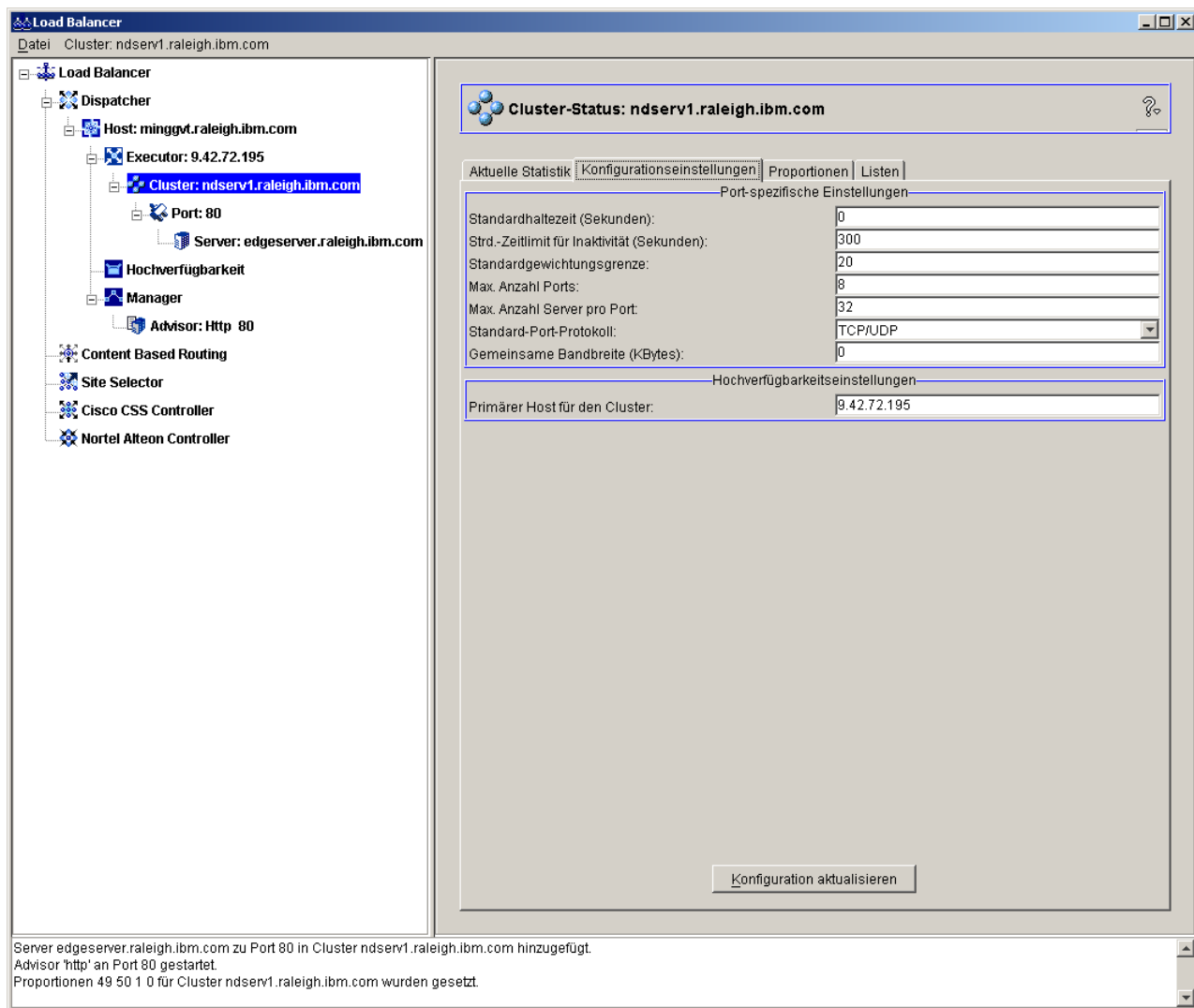


Abbildung 41. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Dispatcher

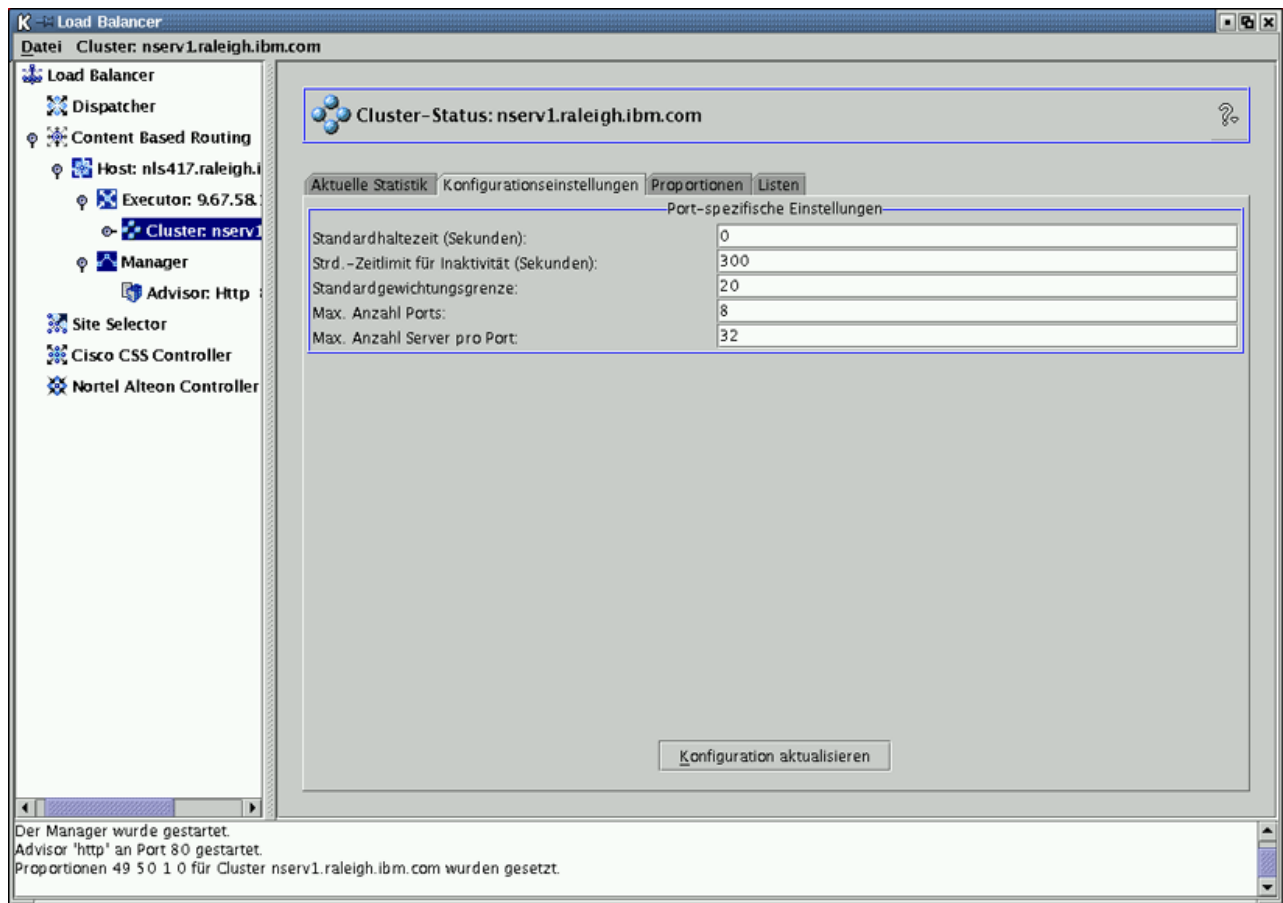


Abbildung 42. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente CBR

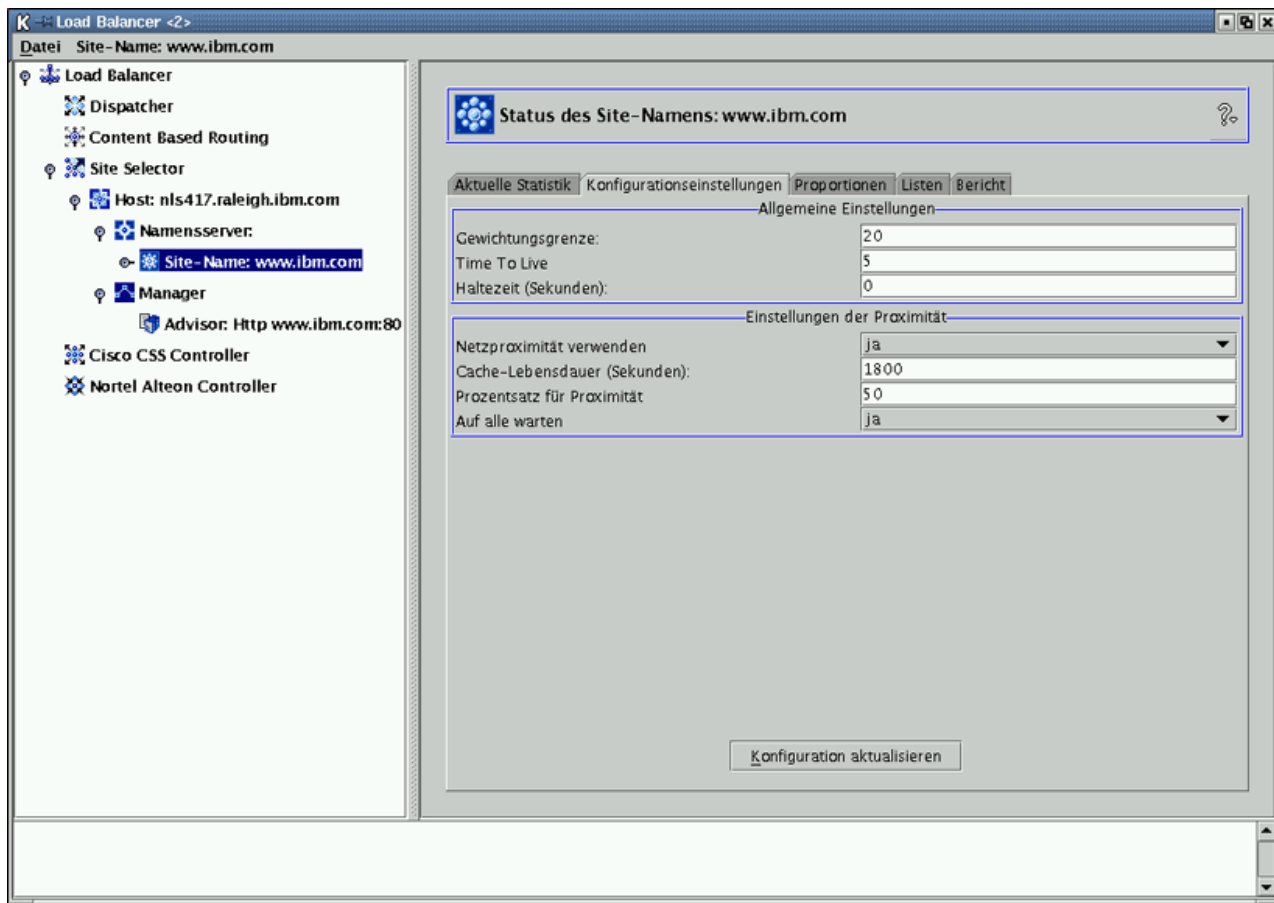


Abbildung 43. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Site Selector

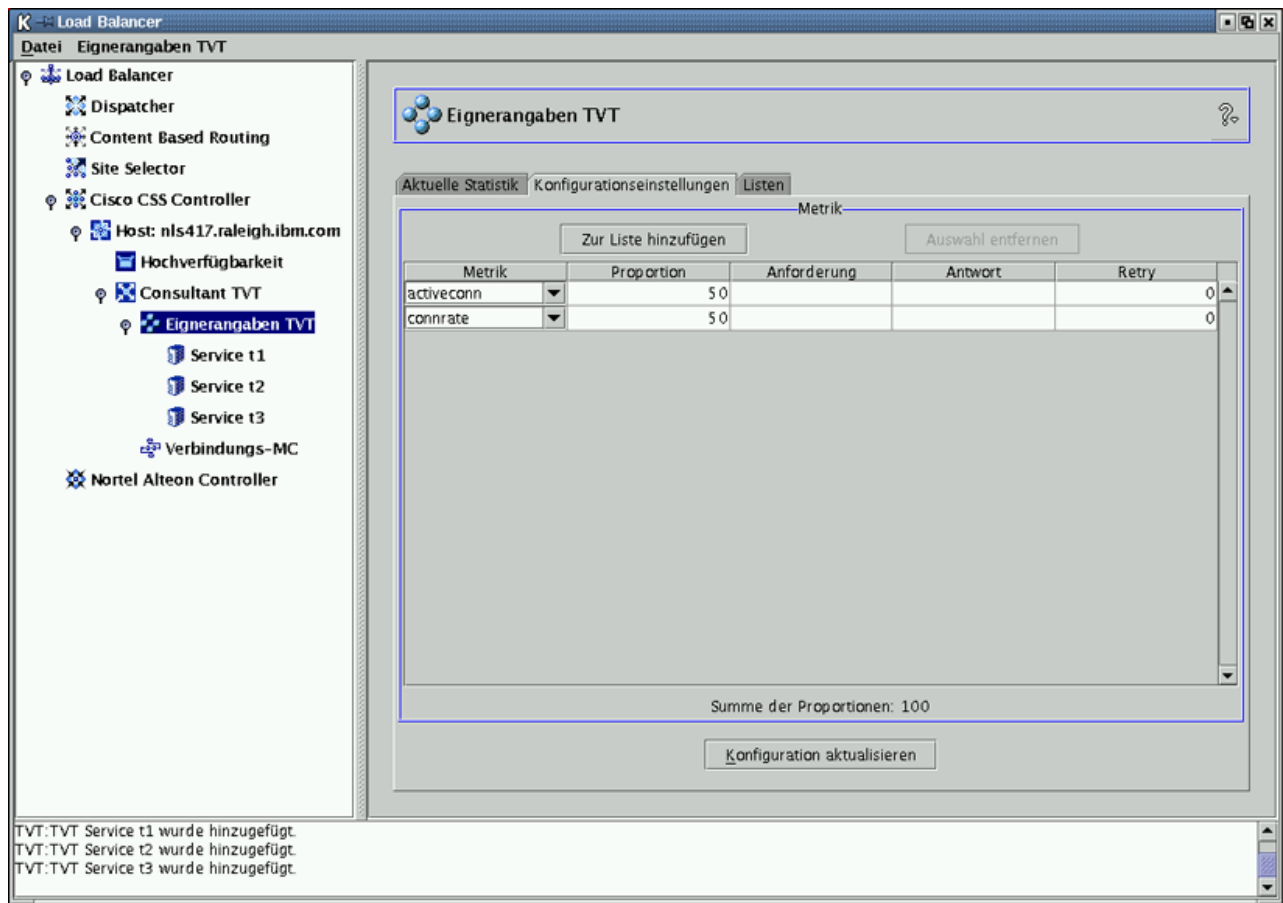


Abbildung 44. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Cisco CSS Controller

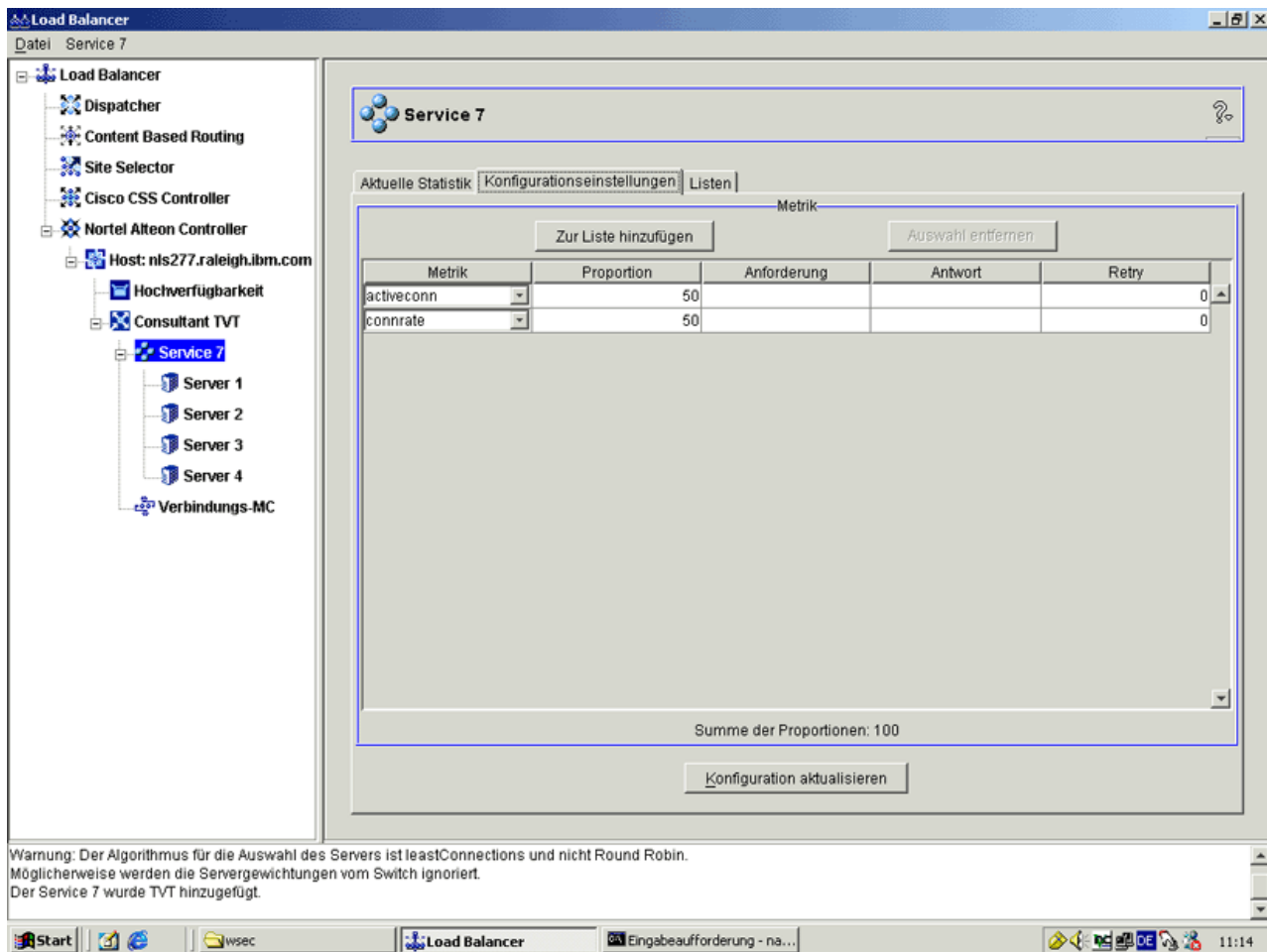


Abbildung 45. GUI mit der erweiterten Anzeige der Baumstruktur für die Komponente Nortel Alteon Controller

Alle Komponenten können auf der GUI konfiguriert werden. Sie können Elemente in der Baumstruktur auswählen, indem Sie mit der ersten Maustaste (normalerweise der linken Maustaste) darauf klicken. Zum Aufrufen von Popup-Menüs müssen Sie die zweite Maustaste (normalerweise die rechte Maustaste) drücken. Auf die Popup-Menüs für die Baumstrukturelemente kann auch über die Menüleiste zugegriffen werden, die sich oben in der Anzeige befindet.

Durch Klicken auf das Plus- oder Minuszeichen können Sie die Elemente der Baumstruktur ein- bzw. ausblenden.

Wenn Sie von der GUI aus einen Befehl ausführen möchten, gehen Sie wie folgt vor: Heben Sie in der GUI-Baumstruktur den Hostknoten hervor und wählen Sie im Popup-Menü "Host" **Befehl senden...** aus. Geben Sie im Befehlseingabefeld den gewünschten Befehl ein, z. B. **executor report**. In einem Fenster sehen Sie die Ergebnisse und die Historie der in der aktuellen Sitzung ausgeführten Befehle.

Auf der rechten Seite der Anzeige erscheinen Registerseiten mit Statusanzeigen für das derzeit ausgewählte Element.

- Die Registerseite **Aktuelle Statistik** stellt statistische Daten zum Element bereit. Diese Registerseite wird nicht für alle Elemente der Baumstruktur angezeigt.
- Durch Klicken auf den Knopf **Statistik aktualisieren** können Sie die aktuellen statistischen Daten aufrufen. Sollte kein Knopf "Statistik aktualisieren" vorhanden sein, wird die Statistik automatisch aktualisiert und ist somit immer auf dem neuesten Stand.
- Die Registerseite **Konfigurationseinstellungen** stellt Konfigurationsparameter bereit, die wie in den Kapiteln zur Konfiguration der einzelnen Komponenten beschrieben definiert werden können. Diese Registerseite wird nicht für alle Elemente der Baumstruktur angezeigt.
- Mit dem Knopf **Konfiguration aktualisieren** werden die letzten Änderungen auf die gegenwärtig aktive Konfiguration angewendet.
- Die Registerseite **Proportionen** enthält Proportionsparameter (oder Wertigkeitsparameter), die wie in Kapitel 22, „Erweiterte Funktionen für Dispatcher, CBR und Site Selector“, auf Seite 201 beschrieben konfiguriert werden können. Diese Registerseite wird nicht für alle Elemente der Baumstruktur angezeigt.
- Die Registerseite **Listen** stellt zusätzliche Details zum ausgewählten Baumstrukturelement zur Verfügung. Diese Registerseite wird nicht für alle Elemente der Baumstruktur angezeigt.
- Mit dem Knopf **Entfernen** können Sie die hervorgehobenen Einträge aus der Liste löschen.
- Die Registerseite **Bericht** zeigt den Manager-Bericht zum Element. Diese Registerseite wird nicht für alle Elemente der Baumstruktur angezeigt.
- Durch Klicken auf den Knopf **Bericht aktualisieren** können Sie die neuesten Daten des Manager-Berichts aufrufen.

Falls Sie **Hilfe** benötigen, klicken Sie oben rechts im Load-Balancer-Fenster auf das Fragezeichen (?).

- **Hilfe: Feldebene** — beschreibt jedes Feld und gibt die Standardwerte an.
- **Hilfe: Wie funktioniert das** — listet Tasks auf, die in der aktuellen Anzeige ausgeführt werden können.
- **InfoCenter** — bietet Zugang zu Produktinformationen. Dazu gehören eine Übersicht über neue Funktionen, ein Link zur Website für das Produkt, ein Index der Onlinehilfedateien sowie ein Glossar.

Anhang B. Syntax der content-Regel

Dieser Anhang beschreibt die Syntax der content-Regel (des Musters) für die Komponente CBR und die Weiterleitungsmethode cbr der Komponente Dispatcher sowie Szenarien und Beispiele für ihre Verwendung.

Syntax der content-Regel

Diese Angaben gelten nur, wenn Sie als Regeltyp content ausgewählt haben.

Beachten Sie bei der Syntax des gewünschten Musters die folgenden Einschränkungen:

- Geben Sie keine Leerzeichen ein.
- Den folgenden Zeichen gelten als Sonderzeichen, sofern ihnen kein umgekehrter Schrägstrich (\) vorangestellt wird:
 - * Platzhalterzeichen (entspricht 0 bis x beliebigen Zeichen)
 - (linke runde Klammer für logische Gruppierung
 -) rechte runde Klammer für logische Gruppierung
 - & logisches UND
 - | logisches ODER
 - ! logisches NICHT

Reservierte Schlüsselwörter

Auf reservierte Schlüsselwörter folgt immer ein Gleichheitszeichen „=“.

Methode

HTTP-Methode in der Anforderung, z. B. GET, POST usw.

URI Pfad der URL-Anforderung (Groß-/Kleinschreibung muss beachtet werden)

Version

Spezifische Version der Anforderung, entweder HTTP/1.0 oder HTTP/1.1

Host Wert vom Host: Header (Groß-/Kleinschreibung kann ignoriert werden)

Anmerkung: In HTTP/1.0-Protokollen optional.

<Schlüssel>

Ein gültiger HTTP-Header-Name, nach dem Dispatcher suchen kann. Beispiele für HTTP-Header sind User-Agent, Connection, Referer usw.

Ein Browser, der `http://www.firma.com/pfad/webseite.htm` aufruft, kann folgende Werte ergeben:

```
Method=GET
URI=/pfad/webseite.htm
Version=HTTP/1.1
Host=www.firma.com
Connection=Keep-Alive
Referer=http://www.firma.com/pfad/externwebseite.htm
```


Anmerkung: Die Shell des Betriebssystems interpretiert Sonderzeichen wie "&" unter Umständen und konvertiert sie in alternativen Text, bevor sie von **cbrcontrol** ausgewertet werden.

Der folgende Befehl ist beispielsweise nur gültig, wenn die Eingabeaufforderung **cbrcontrol>>** verwendet wird.

```
rule add 10.1.203.4:80:cbr_prod_rule_ek type content
  pattern uri=/nipoek/*
```

Wenn dieser Befehl mit Sonderzeichen an der Eingabeaufforderung des Betriebssystems (oder in der Konfigurationsdatei) funktionieren soll, müssen Sie den pattern-Wert wie folgt in Anführungszeichen (" ") setzen:

```
cbrcontrol rule add 10.1.203.4:80:cbr_prod_rule_ek type content
  pattern "uri=/nipoek/*"
```

Fehlen die Anführungszeichen, könnte beim Speichern der Regel in CBR ein Teil des Musters abgeschnitten werden. An der Eingabeaufforderung **cbrcontrol>>** wird die Verwendung von Anführungszeichen nicht unterstützt.

Nachfolgend finden Sie eine Zusammenstellung möglicher Szenarien und Beispiele für die Mustersyntax.

Szenario 1:

Zur Konfiguration für einen Clusternamen gehört eine Gruppe von Webservern für standardmäßigen HTML-Inhalt, eine weitere Gruppe von Webservern mit WebSphere Application Server für Servlet-Anforderungen, eine dritte Gruppe von Lotus-Notes-Servern für NSF-Dateien usw. Der Zugriff auf die Client-Daten ist erforderlich, um zwischen den angeforderten Seiten unterscheiden zu können. Die Daten müssen außerdem an die jeweils geeigneten Server gesendet werden. Die Erkennungsregeln für das content-Muster ermöglichen die für diese Tasks notwendige Trennung. Es wird eine Reihe von Regeln konfiguriert, die die nötige Trennung der Anforderungen automatisch vornehmen. Mit den folgenden Befehlen können Sie die genannte Trennung in drei Gruppen erreichen:

```
>>rule add Cluster1:80:servlets type content pattern uri=*/servlet/* priority 1
>>rule uses Cluster1:80:servlets Server1+Server2

>>rule add Cluster1:80:notes type content pattern uri=*.nsf* priority 2
>>rule uses Cluster1:80:notes Server3+Server4

>>rule add Cluster1:80:regular type true priority 3
>>rule uses Cluster1:80:regular Server5+Server6
```

Wenn Load Balancer eine Anforderung für eine NSF-Datei empfängt, wird zuerst die servlets-Regel angewendet, bei der jedoch keine Übereinstimmung gefunden wird. Anschließend wird die Anforderung mit der notes-Regel abgeglichen und eine Übereinstimmung festgestellt. Die Client-Daten werden auf Server3 und Server4 verteilt.

Szenario 2

Ein weiteres allgemeines Szenario ist die Steuerung unterschiedlicher interner Gruppen durch die Hauptwebsite. Beispiel: `www.firma.com/software` bezieht verschiedene Server und Inhalte aus dem Bereich `www.firma.com/hardware` ein. Da alle Anforderungen vom Root-Cluster `www.firma.com` ausgehen, sind content-Regeln erforderlich, um die URI-Unterscheidung für den Lastausgleich vorzunehmen. Die Regel für dieses Szenario würde etwa wie folgt aussehen:

```
>>rule add Cluster1:80:Bereich1 type content pattern uri=/software/* priority 1
>>rule uses Cluster1:80:Bereich1 Server1+Server2
>>rule add Cluster1:80:Bereich2 type content pattern uri=/hardware/* priority 2
>>rule uses Cluster1:80:Bereich2 Server3+Server4
```

Szenario 3

Bei bestimmten Kombinationen ist die Reihenfolge wichtig, in der die Regeln durchsucht werden. Im Szenario 2 wurden die Clients beispielsweise ausgehend von einem Verzeichnis in ihrem Anforderungspfad aufgeteilt. Der Zielverzeichnispfad kann jedoch auf verschiedenen Ebenen dieses Pfades vorhanden sein und dort jeweils zu anderen Ergebnissen führen. So ist das Ziel `www.firma.com/pcs/fixes/software` beispielsweise von `www.firma.com/mainframe/fixes/software` verschieden. Die Regeln müssen dieser Möglichkeit Rechnung tragen und so konfiguriert werden, dass die nicht zu vielen Szenarien gleichzeitig gerecht werden. Die Platzhaltersuche „`uri=*/software/*`“ wäre in diesem Falle zu breit angelegt. Alternative Regeln könnten wie folgt strukturiert sein:

Mit einer kombinierten Suche kann hier eine Eingrenzung vorgenommen werden:

```
>>rule add Cluster1:80:pcs type content pattern (uri=/pcs/*)&(uri=*/software/*)
>>rule uses Cluster 1:80:pcs Server1
```

In Fällen, wo keine Kombinationen anwendbar sind, ist die Reihenfolge wichtig:

```
>>rule add Cluster1:80:pc1 type content pattern uri=/pcs/*
>>rule uses Cluster1:80:pc1 Server2
```

Die zweite Regel wird angewendet, wenn „`pcs`“ im Verzeichnis nicht an erster Stelle, sondern an einer späteren Stelle erscheint.

```
>>rule add Cluster1:80:pc2 type content pattern uri=*/pcs/*
>>rule uses Cluster1:80:pc2 Server3
```

In fast allen Fällen sollten Sie die Regeln durch eine Standardregel **always true** ergänzen, die für alles gelten, was nicht unter die übrigen Regeln fällt. In Szenarien, in denen alle Server für einen bestimmten Client nicht in Frage kommen, könnte dies auch ein Server mit der Antwort „Die Site ist derzeit nicht verfügbar, versuchen Sie es später erneut“ sein.

```
>>rule add Cluster1:80:sorry type true priority 100
>>rule uses Cluster1:80:sorry Server5
```

Anhang C. Beispielkonfigurationsdateien

Dieser Anhang enthält Beispielkonfigurationsdateien für die Komponente Dispatcher von Load Balancer.

Wichtiger Hinweis: Wenn Sie mit Load Balancer für IPV4 und IPV6 arbeiten, müssen Sie bei Verwendung der Syntaxdiagramme den Doppelpunkt (:) als Begrenzer in dscontrol-Befehlen durch at (@) ersetzen.

Beispielkonfigurationsdateien für Load Balancer

Die Beispieldateien finden Sie im Verzeichnis ...ibm/edge/lb/servers/samples/.

Dispatcher-Konfigurationsdatei — AIX-, Linux- und Solaris-Systeme

```
#!/bin/bash
#
# configuration.sample - Beispielkonfigurationsdatei für die
# Komponente Dispatcher
#
#
# Dieses Script muss vom Benutzer root ausgeführt werden.
#
# iam=`wer bin ich`

# if [ "$iam" != "root" ]if [ "$iam" != "root" ]
# then
#   echo "Zur Ausführung dieses Scripts müssen Sie als root angemeldet sein"
#   exit 2
# fi
#
# Starten Sie zunächst den Server
#
# dsserver start
# sleep 5

#
# Starten Sie dann den Executor.
#
# dscontrol executor start

#
# Der Dispatcher kann vor dem Löschen der Dispatcher-Software
# jederzeit mit den Befehlen dscontrol executor stop und
# dsserver stop zum Stoppen von Executor und Server entfernt
# werden.
#
# Der nächste Konfigurationsschritt für den Dispatcher ist das
# Festlegen der NFA und der Clusteradresse(n).
#
# Die NFA wird für den Fernzugriff auf die Dispatcher-Maschine
# zu Verwaltungs- oder Konfigurationszwecken verwendet. Diese
# Adresse ist erforderlich, da der Dispatcher Pakete an die
# Clusteradresse(n) weiterleitet.
#
# Die CLUSTER-Adresse ist der Hostname (oder die IP-Adresse)
# zu dem (bzw. zu der) ferne Clients eine Verbindung herstellen.
#
# Hostnamen und IP-Adressen sind an jeder Stelle dieser
```

```

# Datei beliebig gegeneinander austauschbar.
#

# NFA=Hostname.Domäne.Name
# CLUSTER=www.IhreFirma.com
# echo "NFA wird geladen"
# dscontrol executor set nfa $NFA
#
# Der nächste Konfigurationsschritt für den Dispatcher ist
# das Erstellen eines Clusters. Der Dispatcher leitet an die
# Clusteradresse gesendete Anforderungen an die entsprechenden,
# für diesen Cluster definierten Servermaschinen weiter. Mit
# Dispatcher können Sie mehrere Clusteradressen konfigurieren
# und bedienen.
# Verwenden Sie für CLUSTER2, CLUSTER3 usw. eine ähnliche Konfiguration.
#

# echo "Erste CLUSTER-Adresse wird geladen"
# dscontrol cluster add $CLUSTER
#
# Jetzt müssen die Ports definiert werden, die dieser Cluster
# verwendet. Alle vom Dispatcher an einem definierten Port
# empfangenen Anforderungen werden an den entsprechenden Port
# einer der Servermaschinen weitergeleitet.
#

# echo "Ports für CLUSTER $CLUSTER werden erstellt"

# dscontrol port add $CLUSTER:20+21+80

#
# Der letzte Schritt ist das Hinzufügen der einzelnen Servermaschinen
# zu den Ports dieses Clusters.
# Auch hier können Sie entweder den Hostnamen oder die IP-Adresse der
# Servermaschinen verwenden.
#

# SERVER1=Servername1.Domäne.Name
# SERVER2=Servername2.Domäne.Name
# SERVER3=Servername3.Domäne.Name

# echo "Servermaschinen werden hinzugefügt"
# dscontrol server add $CLUSTER:20+21+80:
# $SERVER1+$SERVER2+$SERVER3

#
# Jetzt werden die Lastausgleichskomponenten von Dispatcher
# gestartet. Die Hauptkomponente ist der Manager. Die
# sekundären Komponenten sind die Advisor-Funktionen. Sind
# Manager und Advisor-Funktionen nicht aktiv, sendet der
# Dispatcher Anforderungen in einem RoundRobin-Format.
# Sobald der Manager gestartet ist, werden Wertigkeitsentscheidungen
# auf der Grundlage der Anzahl neuer und aktiver Verbindungen
# getroffen und eingehende Anforderungen an den am besten geeigneten
# Server gesendet. Die Advisor-Funktionen geben dem Manager
# Einblick in die Fähigkeit eines Servers, Anforderungen zu bedienen,
# und können feststellen, ob ein Server aktiv ist. Erkennt eine
# Advisor-Funktion, dass ein Server inaktiv ist, wird dieser
# entsprechend markiert (sofern die Manager-Proportionen
# auf das Einbeziehen von Advisor-Eingaben gesetzt sind) und es
# werden keine weiteren Anforderungen an den Server weitergeleitet.

# Der letzte Schritt beim Konfigurieren der Lastausgleichskomponenten
# ist das Festlegen der Manager-Proportionen. Der Manager aktualisiert
# die Wertigkeit der Server ausgehend von vier verschiedenen Ansätzen:
# 1. Anzahl der aktiven Verbindungen für jeden Server.
# 2. Anzahl der neuen Verbindungen zu jedem Server.

```

```

# 3. Eingaben von den Advisor-Funktionen.
# 4. Eingaben von der Advisor-Funktion auf Systemebene.
# Diese Proportionen müssen in der Summe 100 ergeben. Sie können
# die Manager-Proportionen beispielsweise wie folgt festlegen:
# dscontrol manager proportions 48 48 4 0
# Damit fließen die aktiven und neuen Verbindungen mit jeweils 48 %
# in die Gewichtungseinscheidung ein. Die Advisor-Funktionen fließen
# zu 4 % ein und die Systemeingaben werden nicht berücksichtigt.
#
# ANMERKUNG. Standardmäßig sind die Manager-Proportionen auf 50 50 0 0 gesetzt.
#

# echo "Manager wird gestartet..."
# dscontrol manager start
# echo "FTP-Advisor-Funktion wird an Port 21 gestartet ..."
# dscontrol advisor start ftp 21
# echo "HTTP-Advisor-Funktion wird an Port 80 gestartet ..."
# dscontrol advisor start http 80
# echo "Telnet-Advisor-Funktion wird an Port 23 gestartet ..."
# dscontrol advisor start telnet 23
# echo "SMTP-Advisor-Funktion wird an Port 25 gestartet ..."
# dscontrol advisor start smtp 25
# echo "POP3-Advisor-Funktion wird an Port 110 gestartet ..."
# dscontrol advisor start pop3 110
# echo "NNTP-Advisor-Funktion wird an Port 119 gestartet ..."
# dscontrol advisor start nntp 119
# echo "SSL-Advisor-Funktion wird an Port 443 gestartet ..."
# dscontrol advisor start ssl 443
#

# echo "Manager-Proportionen werden festgelegt..."
# dscontrol manager proportions 58 40 2 0
#
# Der letzte Konfigurationsschritt für die Dispatcher-Maschine
# ist das Festlegen eines Aliasnamens auf der Netzschnittstellenkarte (NIC).
#
# ANMERKUNG: Verwenden Sie diesen Befehl NICHT in einer Umgebung mit hoher
# Verfügbarkeit. Die NIC und die Loopback-Adresse werden von den
# go*-Scripts konfiguriert.
# dscontrol executor configure $CLUSTER

# Wenn die Clusteradresse sich auf einer von der NFA abweichenden
# NIC oder in einem abweichenden Teilnetz befindet, verwenden Sie für
# den Befehl cluster configure das folgende Format:
# dscontrol executor configure $CLUSTER tr0 0xfffff800
# tr0 ist hier die NIC (tr1 die zweite Token-Ring-Karte, en0
# die erste Ethernet-Karte) und 0xfffff800 ist eine für
# Ihre Site gültige Teilnetzmaske.
#

#
# Die folgenden Befehle aktivieren die Standardwerte.
# Verwenden Sie diese Befehle als Ausgangspunkt für Änderungen der Standardwerte.
# dscontrol manager loglevel 1
# dscontrol manager logsize 1048576
# dscontrol manager sensitivity 5
# dscontrol manager interval 2
# dscontrol manager refresh 2
#
# dscontrol advisor interval ftp 21 5
# dscontrol advisor loglevel ftp 21 1
# dscontrol advisor logsize ftp 21 1048576
# dscontrol advisor timeout ftp 21 unlimited
# dscontrol advisor interval telnet 23 5
# dscontrol advisor loglevel telnet 23 1
# dscontrol advisor logsize telnet 23 1048576
# dscontrol advisor timeout telnet 23 unlimited

```

```
# dscontrol advisor interval smtp 25 5
# dscontrol advisor loglevel smtp 25 1
# dscontrol advisor logsize smtp 25 1048576
# dscontrol advisor timeout smtp 25 unlimited
# dscontrol advisor interval http 80 5
# dscontrol advisor loglevel http 80 1
# dscontrol advisor logsize http 80 1048576
# dscontrol advisor timeout http 80 unlimited
# dscontrol advisor interval pop3 110 5
# dscontrol advisor loglevel pop3 110 1
# dscontrol advisor logsize pop3 110 1048576
# dscontrol advisor timeout pop3 110 unlimited
# dscontrol advisor interval nntp 119 5
# dscontrol advisor loglevel nntp 119 1
# dscontrol advisor logsize nntp 119 1048576
# dscontrol advisor timeout nntp 119 unlimited
# dscontrol advisor interval ssl 443 5
# dscontrol advisor loglevel ssl 443 1
# dscontrol advisor logsize ssl 443 1048576
# dscontrol advisor timeout ssl 443 unlimited
#
```

Dispatcher-Konfigurationsdatei — Windows-Systeme

Die folgende Konfigurationsdatei ist die Load-Balancer-Beispielkonfigurationsdatei **configuration.cmd.sample**, die mit Windows verwendet wird.

```
@echo off
rem configuration.cmd.sample - Beispielkonfigurationsdatei für die
rem Komponente Dispatcher.
rem

rem dsserver muss im Fenster "Dienste" gestartet werden.

rem

rem
rem Starten Sie dann den Executor.
rem
rem call dscontrol executor start

rem

rem Der nächste Konfigurationsschritt für den Dispatcher
rem ist das Festlegen der NFA und der Clusteradresse(n).
rem

rem Die NFA wird für den Fernzugriff auf die Dispatcher-Maschine
rem zu Verwaltungs- oder Konfigurationszwecken verwendet. Diese
rem Adresse ist erforderlich, da der Dispatcher Pakete an die
rem Clusteradresse(n) weiterleitet.
rem
rem Die CLUSTER-Adresse ist der Hostname (oder die IP-Adresse)
rem zu dem (bzw. zu der) ferne Clients eine Verbindung herstellen.
rem

rem Hostnamen und IP-Adressen sind an jeder Stelle dieser
rem Datei beliebig gegeneinander austauschbar.
rem NFA=[Non-Forwarding Address]
rem CLUSTER=[Clustername]
rem

rem set NFA=Hostname.Domäne.Name
rem set CLUSTER=www.IhreFirma.com

rem echo "NFA wird geladen"
rem call dscontrol executor set nfa %NFA%
```

```

rem
rem Mit den folgenden Befehlen werden die Standardwerte festgelegt.
rem Verwenden Sie diese Befehle zum Ändern der Standardwerte.

rem  call dscontrol executor set fintimeout 30
rem
rem Der nächste Konfigurationsschritt für den Dispatcher ist
rem das Erstellen eines Clusters. Der Dispatcher leitet an die
rem Clusteradresse gesendete Anforderungen an die entsprechenden,
rem für diesen Cluster definierten Servermaschinen weiter. Mit
rem Dispatcher können Sie mehrere Clusteradressen konfigurieren
rem und bedienen.
rem Verwenden Sie für CLUSTER2, CLUSTER3 usw. eine ähnliche Konfiguration.
rem

rem echo "Erste CLUSTER-Adresse wird geladen"
rem call dscontrol cluster add %CLUSTER%
rem
rem Jetzt müssen die Ports definiert werden, die dieser Cluster
rem verwendet. Alle vom Dispatcher an einem definierten Port
rem empfangenen Anforderungen werden an den entsprechenden Port
rem einer der Servermaschinen weitergeleitet.
rem

rem echo "Ports für CLUSTER %CLUSTER% werden erstellt"
rem call dscontrol port add %CLUSTER%:20+21+80
rem
rem Der letzte Schritt ist das Hinzufügen der einzelnen Servermaschinen
rem zu den Ports dieses Clusters. Auch hier können Sie entweder den
rem Hostnamen oder die IP-Adresse der Servermaschinen verwenden.
rem

rem set SERVER1=Servername1.Domäne.Name
rem set SERVER2=Servername2.Domäne.Name
rem set SERVER3=Servername3.Domäne.Name

rem echo "Servermaschinen werden hinzugefügt"
rem call dscontrol server add %CLUSTER%:20+21+80:
rem %SERVER1%+%SERVER2%+%SERVER3%

rem
rem Jetzt werden die Lastausgleichskomponenten von Dispatcher
rem gestartet. Die Hauptkomponente ist der Manager. Die
rem sekundären Komponenten sind die Advisor-Funktionen. Sind
rem Manager und Advisor-Funktionen nicht aktiv, sendet der
rem Dispatcher Anforderungen in einem RoundRobin-Format.
rem Sobald der Manager gestartet ist, werden Wertigkeitsentscheidungen
rem auf der Grundlage der Anzahl neuer und aktiver Verbindungen
rem getroffen und eingehende Anforderungen an den am besten geeigneten
rem Server gesendet. Die Advisor-Funktionen geben dem Manager
rem Einblick in die Fähigkeit eines Servers, Anforderungen zu bedienen,
rem und können feststellen, ob ein Server aktiv ist. Erkennt eine
rem Advisor-Funktion, dass ein Server inaktiv ist, wird dieser
rem entsprechend markiert (sofern die Manager-Proportionen
rem auf das Einbeziehen von Advisor-Eingaben gesetzt sind) und es
rem werden keine weiteren Anforderungen an den Server weitergeleitet.
rem Der letzte Schritt beim Konfigurieren der Lastausgleichskomponenten
rem ist das Festlegen der Manager-Proportionen. Der Manager aktualisiert
rem die Wertigkeit der Server ausgehend von vier verschiedenen Ansätzen:
rem 1. Anzahl der aktiven Verbindungen für jeden Server.
rem 2. Anzahl der neuen Verbindungen zu jedem Server.
rem 3. Eingaben von den Advisor-Funktionen.
rem 4. Eingaben von der Advisor-Funktion auf Systemebene.
rem
rem Diese Proportionen müssen in der Summe 100 ergeben. Sie können
rem die Manager-Proportionen beispielsweise wie folgt festlegen:
rem      dscontrol cluster set <Cluster> proportions 48 48 4 0

```



```

rem Damit fließen die aktiven und neuen Verbindungen mit jeweils 48 %
rem in die Gewichtungseinscheidung ein. Die Advisor-Funktionen fließen
rem zu 4 % ein und die Systemeingaben werden nicht berücksichtigt.
rem
rem ANMERKUNG. Standardmäßig sind die Manager-Proportionen auf
rem 50 50 0 0 gesetzt.
rem echo "Manager wird gestartet..."
rem call dscontrol manager start
rem echo "FTP-Advisor-Funktion wird an Port 21 gestartet..."
rem call dscontrol advisor start ftp 21
rem echo "HTTP-Advisor-Funktion wird an Port 80 gestartet..."
rem call dscontrol advisor start http 80
rem echo "Telnet-Advisor-Funktion wird an Port 23 gestartet..."
rem call dscontrol advisor start telnet 23
rem echo "SMTP-Advisor-Funktion wird an Port 25 gestartet..."
rem call dscontrol advisor start smtp 25
rem echo "POP3-Advisor-Funktion wird an Port 110 gestartet..."
rem call dscontrol advisor start pop3 110
rem echo "NNTP-Advisor-Funktion wird an Port 119 gestartet..."
rem call dscontrol advisor start nntp 119
rem echo "SSL-Advisor-Funktion wird an Port 443 gestartet..."
rem call dscontrol advisor start ssl 443
rem

rem echo "Clusterproportionen werden festgelegt..."
rem call dscontrol cluster set %CLUSTER% proportions 58 40 2 0
rem
rem Der letzte Konfigurationsschritt für die Dispatcher-Maschine
rem ist das Festlegen eines Aliasnamens auf der Netzschnittstellenkarte (NIC).
rem
rem ANMERKUNG: Verwenden Sie diesen Befehl NICHT in einer Umgebung mit hoher
rem Verfügbarkeit. Die NIC und die Loopback-Adresse werden von den
rem go*-Scripts konfiguriert.
rem
rem dscontrol executor configure %CLUSTER%

rem Wenn die Clusteradresse sich auf einer von der NFA abweichenden
rem NIC oder in einem abweichenden Teilnetz befindet, verwenden Sie für
rem den Befehl cluster configure das folgende Format:
rem dscontrol executor configure %CLUSTER% tr0 0xfffff800
rem tr0 ist hier die NIC (tr1 die zweite Token-Ring-Karte, en0
rem die erste Ethernet-Karte) und 0xfffff800 ist eine für
rem Ihre Site gültige Teilnetzmaske.
rem

rem
rem Mit den folgenden Befehlen werden die Standardwerte festgelegt.
rem Verwenden Sie diese Befehle als Ausgangspunkt zum Ändern der Standardwerte.
rem call dscontrol manager loglevel 1
rem call dscontrol manager logsize 1048576
rem call dscontrol manager sensitivity 5
rem call dscontrol manager interval 2
rem call dscontrol manager refresh 2
rem
rem call dscontrol advisor interval ftp 21 5
rem call dscontrol advisor loglevel ftp 21 1
rem call dscontrol advisor logsize ftp 21 1048576
rem call dscontrol advisor timeout ftp 21 unlimited
rem call dscontrol advisor interval telnet 23 5
rem call dscontrol advisor loglevel telnet 23 1
rem call dscontrol advisor logsize telnet 23 1048576
rem call dscontrol advisor timeout telnet 23 unlimited
rem call dscontrol advisor interval smtp 25 5
rem call dscontrol advisor loglevel smtp 25 1
rem call dscontrol advisor logsize smtp 25 1048576
rem call dscontrol advisor timeout smtp 25 unlimited
rem call dscontrol advisor interval http 80 5

```

```

rem call dscontrol advisor loglevel http 80 1
rem call dscontrol advisor logsize http 80 1048576
rem call dscontrol advisor timeout http 80 unlimited
rem call dscontrol advisor interval pop3 110 5
rem call dscontrol advisor loglevel pop3 110 1
rem call dscontrol advisor logsize pop3 110 1048576
rem call dscontrol advisor timeout pop3 110 unlimited
rem call dscontrol advisor interval nntp 119 5
rem call dscontrol advisor loglevel nntp 119 1
rem call dscontrol advisor logsize nntp 119 1048576
rem call dscontrol advisor timeout nntp 119 unlimited
rem call dscontrol advisor interval ssl 443 5
rem call dscontrol advisor loglevel ssl 443 1
rem call dscontrol advisor logsize ssl 443 1048576
rem call dscontrol advisor timeout ssl 443 unlimited
rem

```

Beispiel-Advisor-Funktion

Nachfolgend ist die Advisor-Beispieldatei **ADV_sample** wiedergegeben.

```

/**
 * ADV_sample: HTTP-Advisor-Funktion von Load Balancer
 *
 *
 * Diese Klasse definiert eine angepasste Beispiel-Advisor-Funktion für Load Balancer.
 * Diese angepasste Advisor-Funktion erweitert wie alle anderen Advisor-Funktionen
 * den Advisor-Basiscode ADV_Base. Es ist der Advisor-Basiscode, der die meisten
 * Advisor-Funktionen ausführt. Dazu gehört das Zurückmelden von Belastungen an Load
 * Balancer, die für den Wertigkeitsalgorithmus von Load Balancer verwendet werden.
 * Darüber hinaus stellt der Advisor-Basiscode Socket-Verbindungen her, schließt
 * Sockets und stellt Sende- und Empfangsmethoden für die Advisor-Funktion bereit.
 * Die Advisor-Funktion selbst wird nur zum Senden von Daten an den Port bzw.
 * Empfangen von Daten vom Port des empfohlenen Servers verwendet. Die TCP-Methoden
 * im Advisor-Basiscode sind zeitlich gesteuert, um die Last zu berechnen. Mit einer
 * Markierung der Methode constructor in ADV_base kann bei Bedarf die vorhandene Last
 * mit der neuen, von der Advisor-Funktion zurückgegebenen Last überschrieben werden.
 *
 * Anmerkung: Der Advisor-Basiscode stellt in angegebenen Intervallen
 * die Last ausgehend von einem in der Methode constructor gesetzten
 * Wert für den Wertigkeitsalgorithmus bereit. Ist die eigentliche
 * Advisor-Funktion noch nicht abgeschlossen, so dass sie keinen gültigen
 * Lastwert zurückgeben kann, verwendet der Advisor-Basiscode die
 * bisherige Last.
 *
 *
 * NAMEN
 *
 * Es gilt die folgende Namenskonvention:
 *
 * - Die Datei muss sich in den folgenden Load-Balancer-Verzeichnissen befinden:
 *
 *    lb/servers/lib/CustomAdvisors/ (Windows: lb\servers\lib\CustomAdvisors)
 *
 * - Der Name der Advisor-Funktion muss das Präfix ADV_ haben. Zum Starten
 *   der Advisor-Funktion genügt jedoch der Name. Die Advisor-Funktion
 *   ADV_sample kann beispielsweise mit sample gestartet werden.
 *
 * - Der Name der Advisor-Funktion muss in Kleinbuchstaben angegeben werden.
 *
 * Unter Beachtung dieser Regeln wird auf dieses Beispiel wie folgt verwiesen:
 *
 *    <Basisverzeichnis>/lib/CustomAdvisors/ADV_sample.class
 *
 *
 * Advisor-Funktionen müssen, wie für Load Balancer generell gültig,
 * mit der erforderlichen Java-Version kompiliert werden. Um den Zugriff auf
 * die Load-Balancer-Klassen zu gewährleisten, müssen Sie sicherstellen, dass die

```

```

* Datei ibmlb.jar (aus dem Unterverzeichnis lib des Basisverzeichnisses) im CLASSPATH
* des Systems enthalten ist.
*
* Von ADV_Base bereitgestellte Methoden:
*
* - ADV_Base (Constructor):
*
*   - Parameter
*     - String sName = Name der Advisor-Funktion
*     - String sVersion = Version der Advisor-Funktion
*     - int iDefaultPort = Standard-Port-Nummer für die Advisor-Funktion
*     - int iInterval = Intervall für die Ausführung der Advisor-Funktion
*                       auf den Servern
*     - String sDefaultName = Nicht verwendet; muss als "" übergeben werden.
*     - boolean replace = True - Den vom Advisor-Basiscode berechneten Lastwert
*                               ersetzen
*                               False - Zu dem vom Advisor-Basiscode berechneten Lastwert
*                               addieren
*   - Rückgabe
*     - constructor-Methoden haben keine Rückgabewerte.
*
* Da der Advisor-Basiscode auf Threads basiert, stehen verschiedene andere
* Methoden für Advisor-Funktionen zur Verfügung. Auf diese kann mit dem von
* getLoad() übergebenen Parameter CALLER verwiesen werden.
*
* Es handelt sich um die folgenden Methoden:
*
* - send - Informationspaket über die eingerichtete Socket-Verbindung
*         an den Server am angegebenen Port senden.
*   - Parameter
*     - String sDataString - Daten werden in Form einer Zeichenfolge
*                           gesendet
*   - Rückgabe
*     - int RC - Null gibt unabhängig vom erfolgreichen/gescheiterten Senden
*               der Daten an, dass die Daten gesendet wurden. Eine negative
*               ganze Zahl zeigt einen Fehler an.
*
* - receive - Empfang von Informationen von der Socket-Verbindung.
*   - Parameter
*     - StringBuffer sbDataBuffer - Die während des Aufrufs von receive
*                                   empfangenen Daten
*   - Rückgabe
*     - int RC - Null gibt unabhängig vom erfolgreichen/gescheiterten Empfang
*               der Daten an, dass die Daten gesendet wurden. Eine negative
*               ganze Zahl zeigt einen Fehler an.
*
* Falls die vom Advisor-Basiscode bereitgestellte Funktionalität nicht
* ausreicht, können Sie die gewünschte Funktion innerhalb des Advisors
* erstellen. Die vom Advisor-Basiscode bereitgestellten Methoden werden
* dann ignoriert.
*
* Eine wichtige Frage hinsichtlich der zurückgegebenen Last ist, ob
* sie auf die vom Advisor-Basiscode generierte Last angewendet oder
* diese ersetzen soll. Es gibt gültige Instanzen für beide Situationen.
*
* Dieses sehr einfache Beispiel entspricht im Wesentlichen der
* HTTP-Advisor-Funktion von Load Balancer.
* Es wird eine Sende-anforderung (HTTP HEAD) abgesetzt. Beim Empfang einer
* Antwort wird die Methode getLoad beendet und der Advisor-Basiscode
* angewiesen, die Ablaufsteuerung der Anforderung zu stoppen. Die Methode
* ist damit abgeschlossen. Die zurückgegebenen Informationen werden keiner
* Syntaxanalyse unterzogen. Die Last basiert auf der für das Senden und
* Empfangen benötigten Zeit.
*/

package CustomAdvisors;
import com.ibm.internet.nd.advisors.*;

```

```

public class ADV_sample extends ADV_Base implements ADV_MethodInterface
{
    String COPYRIGHT = "(C) Copyright IBM Corporation 1997, alle Rechte vorbehalten.\n";
    static final String ADV_NAME = "Sample";
    static final int ADV_DEF_ADV_ON_PORT = 80;
    static final int ADV_DEF_INTERVAL = 7;

    // Anmerkung: Die meisten Serverprotokolle erfordern am Ende von Nachrichten
    // eine Zeilenschaltung ("\r") und einen Zeilenvorschub ("\n"). Sollte dies
    // für Sie zutreffen, nehmen Sie sie an dieser Stelle in Ihre Zeichenfolge
    // auf.
    static final String ADV_SEND_REQUEST =
        "HEAD / HTTP/1.0\r\nAccept: */*\r\nUser-Agent: " +
        "IBM_Load_Balancer_HTTP_Advisor\r\n\r\n";

    /**
     * Constructor.
     *
     * Parameter: Keine. An die constructor-Methode für ADV_Base müssen
     * jedoch mehrere Parameter übergeben werden.
     */
    public ADV_sample()
    {
        super( ADV_NAME,
              "2.0.0.0-03.27.98",
              ADV_DEF_ADV_ON_PORT,
              ADV_DEF_INTERVAL,
              "", // not used
              false);
        super.setAdvisor( this );
    }

    /**
     * ADV_AdvisorInitialize
     *
     * Eine Advisor-spezifische Initialisierung, die nach dem Start der
     * Advisor-Funktion stattfinden muss. Diese Methode wird nur einmal aufgerufen
     * und in der Regel nicht verwendet.
     */
    public void ADV_AdvisorInitialize()
    {
        return;
    }

    /**
     * getLoad()
     *
     * Diese Methode wird vom Advisor-Basiscode aufgerufen, um die Operation der
     * Advisor-Funktion auf der Grundlage protokollspezifischer Details zu beenden.
     * In diesem Beispiel sind nur eine Sende- und eine Empfangsoperation
     * notwendig. Wenn eine komplexere Logik erforderlich ist, können mehrere
     * Sende- und Empfangsoperationen ausgeführt werden.
     * Es könnte beispielsweise eine Antwort empfangen werden. Die sich aus der
     * Syntexanalyse dieser Antwort ergebenden Informationen könnten eine
     * weitere Sende- und Empfangsoperation nach sich ziehen.
     *
     * Parameter:
     *
     * - iConnectTime - Derzeitige Last entsprechend der Zeit, die für das Herstellen
     * der Verbindung zum Server über den angegebenen Port benötigt
     * wurde.
     *
     * - caller - Verweis auf die Advisor-Basisklasse, wo die von Load

```

```

*           Balancer bereitgestellten Methoden einfache TCP-Anforderungen
*           wie Sende- und Empfangsaufrufe durchführen sollen.
*
* Ergebnisse:
*
* - Last: Ein in Millisekunden angegebener Wert, der entsprechend der
* Markierung replace der constructor-Methode zur vorhandenen Last
* addiert wird oder die vorhandene Last ersetzt.
*
* Je größer die Last ist, desto länger benötigt der Server für die
* Antwort. Um so geringer wird auch die Wertigkeit im Load Balancer
* ausfallen.
*
* Wenn der Wert negativ ist, wird von einem Fehler ausgegangen. Ein Fehler
* von einer Advisor-Funktion zeigt an, dass der Server, den die
* Advisor-Funktion zu erreichen versucht, nicht zugänglich und inaktiv ist.
* Load Balancer versucht nicht, einen inaktiven Server am Lastausgleich zu
* beteiligen. Der Server wird erst wieder in den Lastausgleich einbezogen,
* wenn ein positiver Wert empfangen wird.
*
*/
public int getLoad(int iConnectTime, ADV_Thread caller)
{
    int iRc;
    int iLoad = ADV_HOST_INACCESSIBLE; // -1

    // TCP-Anforderung senden
    iRc = caller.send(ADV_SEND_REQUEST);
    if (iRc >= 0)
    {
        // Empfang ausführen
        StringBuffer sbReceiveData = new StringBuffer("");
        iRc = caller.receive(sbReceiveData);

        /**
         * Im normalen Advisor-Modus (Markierung replace ist auf false gesetzt),
         * wird der Lastwert 0 oder 1 zurückgegeben, um anzugeben, ob der Server
         * aktiv oder inaktiv ist.
         * Bei erfolgreichem Empfang wird als Lastwert null zurückgegeben, um
         * anzuzeigen, dass der von der Basis-Advisor-Funktion ermittelte
         * Lastwert verwendet werden soll.
         *
         * Andernfalls (Markierung replace ist auf true gesetzt) müssen Sie
         * den gewünschten Lastwert zurückgeben.
         */

        if (iRc >= 0)
        {
            iLoad = 0;
        }
    }
    return iLoad;
}

} // Ende von ADV_sample

```

Anhang D. Beispiel für eine Client/Server-Konfiguration mit hoher Verfügbarkeit unter Verwendung von Dispatcher, CBR und Caching Proxy

Dieser Anhang beschreibt das Einrichten einer Client/Server-Konfiguration mit hoher Verfügbarkeit, die das Leistungsspektrum der Komponenten von Load Balancer (Dispatcher und CBR) mit dem von Caching Proxy verbindet.

Servermaschine einrichten

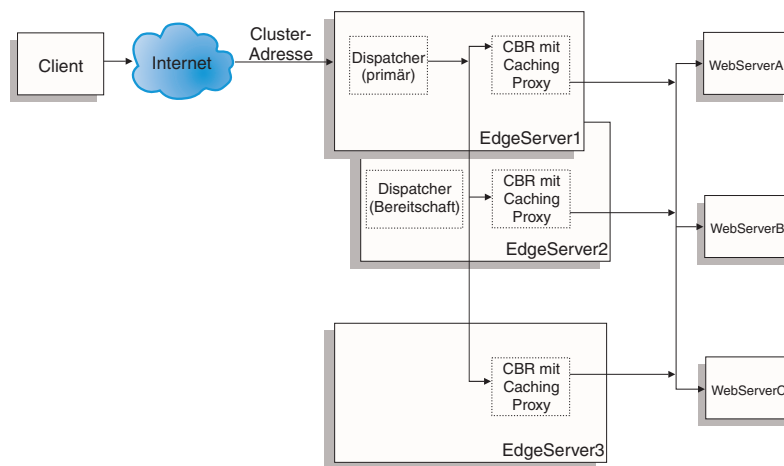


Abbildung 46. Beispiel für eine Client/Server-Konfiguration mit hoher Verfügbarkeit unter Verwendung von Dispatcher, CBR und Caching Proxy

Die Servermaschine in Abb. 46 ist wie folgt konfiguriert:

- EdgeServer1: primäre Dispatcher-Maschine (hohe Verfügbarkeit), die mit CBR und Caching Proxy verknüpft ist und die Last auf Webserver verteilt.
- EdgeServer2: Bereitschafts-Dispatcher-Maschine (hohe Verfügbarkeit), die mit CBR und Caching Proxy verknüpft ist.
- EdgeServer3: Maschine mit CBR und Caching Proxy.
- WebServerA, WebServerB, WebServerC: Ausweichwebserver.

Abb. 46 zeigt eine grundlegende Darstellung mehrerer Server (EdgeServer1, EdgeServer2, EdgeServer3), die die Last auf mehrere Back-End-Webserver verteilen. Die Komponente CBR verwendet Caching Proxy für eine vom Inhalt des URL abhängige Weiterleitung von Anforderungen an die Back-End-Webserver. Die Komponente Dispatcher verteilt die Last der CBR-Komponenten auf alle Edge-Server. Die Dispatcher-Funktion für hohe Verfügbarkeit stellt sicher, dass Anforderungen an die Back-End-Server auch dann möglich sind, wenn die primäre Maschine mit hoher Verfügbarkeit (EdgeServer1) ausfällt.

Basisrichtlinien für die Konfiguration:

- Konfigurieren Sie Caching Proxy auf allen Edge-Servern identisch. Zur Verbesserung der Zugriffsmöglichkeiten auf die Webseiten der Back-End-Server sollten Sie Caching Proxy für das Speicher-Caching konfigurieren. So können die Edge-Server häufiger angeforderte Webseiten zwischenspeichern. Weitere Informationen zum Konfigurieren von Caching Proxy finden Sie im *Caching Proxy Administratorhandbuch*.
- Definieren Sie für die Load-Balancer-Komponenten CBR und Dispatcher identische Clusteradressen und Ports.
- Konfigurieren Sie die Komponente CBR auf allen Edge-Servern gleich. Verwenden Sie an den Ports, die Sie für den Cluster definieren möchten, die Webserver A, B und C. Weitere Informationen zum Konfigurieren von CBR finden Sie in Kapitel 11, „Content Based Routing konfigurieren“, auf Seite 109.
- Konfigurieren Sie die Komponente Dispatcher auf den Edge-Servern 1 und 2 identisch. Definieren Sie an den Ports, die als Cluster-Ports definiert werden sollen, an denen Dispatcher einen Lastausgleich durchführt, alle Edge-Server als zu verwendende Server. Weitere Informationen zum Konfigurieren von Dispatcher finden Sie in Kapitel 7, „Dispatcher konfigurieren“, auf Seite 63.
- Konfigurieren Sie den Edge-Server 1 als primäre Maschine mit hoher Verfügbarkeit und den Edge-Server 2 als Bereitschaftsmaschine (Sicherung) mit hoher Verfügbarkeit. Weitere Informationen hierzu finden Sie im Abschnitt „Hohe Verfügbarkeit“ auf Seite 204.

Anmerkung:

1. Wenn Sie vermeiden möchten, dass die Back-End-Serveradressen auf einem Client im URL angezeigt werden, müssen Sie für jede Back-End-Serveradresse die Anweisung ReversePass in der Konfigurationsdatei von Caching Proxy setzen.
2. Sie können sicherstellen, dass das Webspeicher-Caching effizient genutzt wird, indem Sie in der Konfigurationsdatei von Caching Proxy die Anweisung "Caching" auf "ON" und den Wert für die Anweisung "CacheMemory" auf die erforderliche Größe setzen.
3. Beispielzeilen zu den obigen Anmerkungen 1-2:

```
Caching          ON
CacheMemory      128000 K
ReversePass /* http://websrvA.firma.com/* http://www.firma.com/*
```

4. Vergessen Sie nicht, auf der Netzschnittstellenkarte für EdgeServer1 und auf den Loopback-Einheiten der übrigen Edge-Server einen Aliasnamen für die Clusteradresse festzulegen.
5. Wenn Sie die Edge-Server auf einer Linux-Plattform verwenden, müssen Sie einen Patch-Code für den Linux-Kernel installieren oder eine alternative Methode nutzen, um auf der Loopback-Einheit einen Aliasnamen festzulegen. Weitere Informationen hierzu finden Sie im Abschnitt „Alternativen für die Festlegung eines Loopback-Aliasnamens unter Linux bei Verwendung der Load-Balancer-Weiterleitungsmethode mac“ auf Seite 78.
6. Wenn Sie CBR mit content-Regeln verwenden, dürfen Sie nicht die Port-Affinität (stickytime) anwenden, da andernfalls die content-Regeln beim Verarbeiten von Anforderungen an die Back-End-Webserver nicht erfüllt werden.

Beispielkonfigurationsdateien:

Die folgenden Beispielkonfigurationsdateien ähneln den Dateien, die beim Einrichten einer Edge-Components-Konfiguration, wie sie in Abb. 46 auf Seite 491 dargestellt ist, erstellt werden. Die Beispielkonfigurationsdateien sind Dateien für die Load-Balancer-Komponenten Dispatcher und CBR. In der Beispielkonfiguration wird für jede Edge-Server-Maschine ein Ethernet-Adapter verwendet und alle Adressen befinden sich innerhalb eines privaten Teilnetzes. In den Beispielkonfigurationsdateien sind für die angegebenen Maschinen die folgenden IP-Adressen angegeben:

- EdgeServer1 (primärer Edge-Server mit hoher Verfügbarkeit): 192.168.1.10
- EdgeServer2 (Ausschweich-Edge-Server mit hoher Verfügbarkeit): 192.168.1.20
- EdgeServer3 (Edge-Server für Web-Caching): 192.168.1.30
- Clusteradresse der Website: 192.168.1.11
- WebServerA-C (Back-End-Webserver): 192.168.1.71, 192.168.1.72 und 192.168.1.73

Beispielkonfigurationsdatei für die Komponente Dispatcher auf dem primären Edge-Server mit hoher Verfügbarkeit:

```
dscontrol executor start
```

```
dscontrol cluster add 192.168.1.11 primaryhost 192.168.1.10
```

```
dscontrol port add 192.168.1.11:80
```

```
dscontrol server add 192.168.1.11:80:edgeserver1 address 192.168.1.10
```

```
dscontrol server add 192.168.1.11:80:edgeserver2 address 192.168.1.20
```

```
dscontrol server add 192.168.1.11:80:edgeserver3 address 192.168.1.30
```

```
dscontrol manager start manager.log 10004
```

```
dscontrol highavailability heartbeat add 192.168.1.10 192.168.1.20
```

```
dscontrol highavailability backup add primary auto 4567
```


Beispielkonfigurationsdatei für die Komponente CBR auf den Edge-Servern:

```
cbrcontrol set loglevel 1
cbrcontrol executor start

cbrcontrol cluster add 192.168.1.11

cbrcontrol port add 192.168.1.11:80

cbrcontrol server add 192.168.1.11:80:webserverA address 192.168.1.71
cbrcontrol server add 192.168.1.11:80:webserverB address 192.168.1.72
cbrcontrol server add 192.168.1.11:80:webserverC address 192.168.1.73

cbrcontrol rule add 192.168.1.11:80:webA_Regel type content
    pattern (URI=*WSA*)|(URI=*wsA*) priority 21
cbrcontrol rule useserver 192.168.1.11:80:webA_Regel webserverA

cbrcontrol rule add 192.168.1.11:80:webB_Regel type content
    pattern (URI=/WS_B*) priority 22
cbrcontrol rule useserver 192.168.1.11:80:webB_Regel webserverB

cbrcontrol rule add 192.168.1.11:80:webC_Regel type content
    pattern URI=*webC* priority 23
cbrcontrol rule useserver 192.168.1.21:80:webC_Regel webserverC
```

Anhang E. Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden.

Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, dass nur Programme, Produkte oder Dienstleistungen von IBM verwendet werden können. An Stelle der Produkte, Programme oder Dienstleistungen können auch andere ihnen äquivalente Produkte, Programme oder Dienstleistungen verwendet werden, solange diese keine gewerblichen oder andere Schutzrechte der IBM verletzen. Die Verantwortung für den Betrieb von Fremdprodukten, Fremdprogrammen und Fremddienstleistungen liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieser Veröffentlichung ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

IBM Director of Licensing
IBM Europe, Middle East & Africa
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France

Trotz sorgfältiger Bearbeitung können technische Ungenauigkeiten oder Druckfehler in dieser Veröffentlichung nicht ausgeschlossen werden. Die Angaben in diesem Handbuch werden in regelmäßigen Zeitabständen aktualisiert. Die Änderungen werden in Überarbeitungen oder in Technical News Letters (TNLs) bekannt gegeben. IBM kann ohne weitere Mitteilung jederzeit Verbesserungen und/oder Änderungen an den in dieser Veröffentlichung beschriebenen Produkten und/oder Programmen vornehmen.

Verweise in diesen Informationen auf Websites anderer Anbieter dienen lediglich als Benutzerinformationen und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt; die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängigen, erstellten Programmen und anderen Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

IBM Corporation
Attn.: G71A./503.
P.O. Box 12195
3039 Cornwallis Rd.
Research Triangle Park, N.C. 27709-2195
USA

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des im Handbuch aufgeführten Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt im Rahmen der Allgemeinen Geschäftsbedingungen der IBM oder einer äquivalenten Vereinbarung.

Alle in diesem Dokument enthaltenen Leistungsdaten stammen aus einer gesteuerten Umgebung. Die Ergebnisse, die in anderen Betriebsumgebungen erzielt werden, können daher erheblich von den hier erzielten Ergebnissen abweichen. Einige Daten stammen möglicherweise von Systemen, deren Entwicklung noch nicht abgeschlossen ist. Eine Garantie, dass diese Daten auch in allgemein verfügbaren Systemen erzielt werden, kann nicht gegeben werden. Darüber hinaus wurden einige Daten unter Umständen durch Extrapolation berechnet. Die tatsächlichen Ergebnisse können abweichen. Benutzer dieses Dokuments sollten die entsprechenden Daten in ihrer spezifischen Umgebung prüfen.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen machen. Fragen zum Leistungsspektrum von Produkten anderer Hersteller sind an die Hersteller dieser Produkte zu richten.

Die oben genannten Erklärungen bezüglich der Produktstrategien und Absichtserklärungen von IBM stellen die gegenwärtige Absicht der IBM dar, unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele der IBM.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufes. Sie sollen nur die Funktionen des Lizenzprogrammes illustrieren; sie können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

Marken

Folgende Namen sind in gewissen Ländern Marken der IBM Corporation:

AFS
AIX
DFS
IBM
iSeries
NetView
OS/2
Redbooks
RS/6000
SecureWay
ViaVoice
WebSphere
zSeries

Java und alle Java-basierten Marken sind in gewissen Ländern Marken von Sun Microsystems, Inc.

Microsoft, Windows, Windows NT und das Windows-Logo sind in gewissen Ländern Marken der Microsoft Corporation.

Intel, Intel Inside (Logos), MMX und Pentium sind in gewissen Ländern Marken der Intel Corporation.

UNIX ist in gewissen Ländern eine eingetragene Marke von The Open Group.

Linux ist in gewissen Ländern eine Marke von Linus Torvalds.

Andere Namen von Unternehmen, Produkten und Dienstleistungen können Marken oder Dienstleistungsmarken anderer Unternehmen sein.

Glossar

A

ACK. Ein Steuerungsbit (zur Bestätigung), das keinen Platz in der Folge beansprucht. Es zeigt an, dass das Bestätigungsfeld dieses Segments die nächste Folgennummer angibt, die der Absender dieses Segments erwartet, und somit den Empfang aller vorherigen Folgennummern bestätigt.

Adresse. Der eindeutige Code, der jeder Einheit oder Workstation zugeordnet wird, die mit einem Netz verbunden ist. Eine Standard-IPv4-Adresse ist ein 32-Bit-Adressfeld mit zwei Abschnitten. Der erste Abschnitt ist die Netzadresse, der zweite die Hostnummer. Eine IPv6-Adresse ist ein 128-Bit-Adressfeld, das eine deutlich höhere Anzahl von Adressen als IPv4 unterstützt. Eine IPv6-Adresse unterstützt außerdem weitere Features wie die Multicast- und Anycast-Adressierung.

Advisor-Funktion. Die Advisor-Funktionen sind Bestandteil von Load Balancer. Advisor-Funktionen erfassen und analysieren Rückmeldungen von einzelnen Servern und informieren die Manager-Funktion.

Agent. (1) In der Systemverwaltung ein Benutzer, der für eine bestimmte Interaktion die Rolle eines Agenten übernommen hat. (2) Eine Definitionseinheit, die verwaltete Objekte repräsentiert. Dies geschieht durch (a) die Ausgabe von Mitteilungen über die Objekte und (b) die Bearbeitung von Manager-Anforderungen für Verwaltungsoperationen zum Ändern oder Abfragen der Objekte.

Aliasname. Ein zusätzlicher Name, der einem Server zugeordnet wird. Der Aliasname macht den Server vom Namen seiner Hostmaschine unabhängig. Der Aliasname muss im Domänennamensserver definiert sein.

Aliasname der Loopback-Adresse. Eine der Loopback-Schnittstelle zugeordnete alternative IP-Adresse. Die alternative Adresse hat den nützlichen Nebeneffekt, dass sie sich nicht auf einer realen Schnittstelle befindet.

Als aktiv markieren. Einem Server das Empfangen neuer Verbindungen erlauben.

Als inaktiv markieren. Alle aktiven Verbindungen zu einem Server werden unterbrochen und das Senden neuer Verbindungen oder Pakete an diesen Server wird unterbunden.

API. Anwendungsprogrammierschnittstelle. Die Schnittstelle (Anrufvereinbarungen), durch die ein Anwendungsprogramm auf Dienste des Betriebssystems und andere Dienste zugreift. Eine API ist auf Quellencodesebene definiert und bietet eine Abstraktionsstufe zwischen der Anwendung und dem Kernel (oder anderen privilegierten Dienstprogrammen), um die Portierbarkeit des Codes sicherzustellen.

Assistent. Ein Dialog innerhalb einer Anwendung, der einen Benutzer schrittweise bei der Ausführung einer bestimmten Task anleitet.

Ausweichmaschine. Bei der Funktion für hohe Verfügbarkeit von Dispatcher die Partnermaschine der primären Maschine. Sie überwacht den Status der primären Maschine und übernimmt ggf. deren Aufgaben. Siehe auch "Hohe Verfügbarkeit" und "Primäre Maschine".

B

Bandbreite. Die Differenz zwischen der höchsten und der niedrigsten Frequenz eines Übertragungskanal. Die Datenmenge, die pro Sekunde über eine bestimmte Kommunikationsverbindung gesendet werden kann.

Bereichsanfang. Bei regelbasierten Lastausgleich der niedrigste Wert, der für eine Regel angegeben wird. Der Standardwert hängt vom Regeltyp ab.

Bereichsende. Beim regelbasierten Lastausgleich der höchste für eine Regel angegebene Wert. Die Standardeinstellung ist vom Regeltyp abhängig.

Binäre Protokollierung. Erlaubt das Speichern von Serverdaten in Binärdateien, die anschließend verarbeitet werden, um die zeitabhängig gesammelten Serverdaten zu analysieren.

C

Caching Proxy. Ein Caching-Proxy-Server, der durch sehr effiziente Caching-Schemata die Antwortzeit für Endbenutzer verkürzen hilft. Flexible PICS-Filter unterstützen Netzadministratoren bei der Steuerung des Zugriffs auf webbasierte Informationen an einem zentralen Standort.

CBR. Content Based Routing. Eine Komponente von Load Balancer. CBR verteilt zusammen mit Caching Proxy eingehende Client-Anforderungen ausgehend vom Inhalt der Webseite und unter Verwendung bestimmter Regeltypen auf HTTP- oder HTTPS-Server.

cbrcontrol. Die Schnittstelle zur Komponente Content Based Routing von Load Balancer.

cbrserver. Bearbeitet beim Content Based Routing die Anfragen von der Befehlszeile an Executor, Manager und Advisor-Funktionen.

ccocontrol. Die Schnittstelle zwischen Cisco CSS Controller und dem Cisco CSS Switch.

ccoserver. Bearbeitet im Cisco CSS Controller die Anforderungen von der Befehlszeile an die Consultants.

CGI. Common Gateway Interface. Ein Standard für den Austausch von Informationen zwischen einem Webserver und einem externen Programm. Das externe Programm kann in einer beliebigen vom Betriebssystem unterstützten Sprache geschrieben sein und führt Tasks aus, die der Server normalerweise nicht ausführt, z. B. die Formularverarbeitung.

CGI-Script. Ein CGI-Programm, das in einer Script-basierten Sprache wie Perl oder REXX geschrieben ist und mit der Common Gateway Interface Tasks ausführt, die der Server in der Regel nicht ausführt, z. B. die Formularverarbeitung.

Cisco CSS Controller. Eine Komponente von IBM Load Balancer. Cisco CSS Controller stellt mit der Load-Balancer-Technologie Echtzeitdaten zum Lastausgleich für den Cisco Content Services Switch bereit.

Cisco CSS Switch. Switches der Cisco CSS 11000 Series, die zur Weiterleitung von Paketen und Inhalten verwendet werden.

Client. Ein Datenverarbeitungssystem oder -prozess, das bzw. der einen Dienst von einem anderen Datenverarbeitungssystem oder -prozess anfordert. Eine Workstation oder ein Personal Computer, die bzw. der HTML-Dokumente von einem Lotus Domino Go Webserver anfordert, ist beispielsweise ein Client dieses Servers.

Cluster. Im Kontext der Komponente Dispatcher eine Gruppe von TCP- oder UDP-Servern, die für denselben Zweck verwendet werden und mit einem Hostnamen identifiziert werden.

Clusteradresse. Im Kontext der Komponente Dispatcher die Adresse, zu der Clients eine Verbindung herstellen.

Clusterserver. Ein Server, den der Dispatcher mit anderen Servern zu einem virtuellen Server zusammenfasst. Load Balancer verteilt den TCP- oder UDP-Datenverkehr auf diese Clusterserver.

Consultant. Erfasst Servermesswerte von den am Lastausgleich beteiligten Servern und sendet Daten zur Serverwertigkeit an den Switch, der den Lastausgleich durchführt.

Controller. Eine Gruppe von Consultants.

D

Dämon. (DAEMon, Disk And Execution Monitor) Ein Programm, das nicht explizit beteiligt ist, sondern ruht und darauf wartet, dass bestimmte Bedingungen erfüllt sind. Der Verursacher der Bedingungen muss nichts von dem wartenden Dämon wissen (obwohl ein Programm häufig eine Aktion aus genau dem Grund ausführt, weil es weiß, dass damit implizit ein Dämon aufgerufen wird).

Dienst. (1) Eine von Knoten bereitgestellte Funktion wie HTTP, FTP oder Telnet. (2) Für Nortel Alteon Controller ist ein Dienst die Funktion oder Information, die ein Endbenutzer von einer Site anfordert. Er ist durch eine virtuelle IP-Adresse und eine virtuelle Port-Nummer für eine Endbenutzeranforderung gekennzeichnet. Auf dem Switch ist er durch eine virtuelle Server-ID gekennzeichnet. Diese ID besteht aus einer ganzen Zahl und einer virtuellen Port-

Nummer oder einem Servicenamen. (3) Für Cisco CSS Consultant ist ein Dienst eine Zieladresse, unter der ein Inhaltselement physisch gespeichert ist. Diese Adresse könnte ein lokaler oder ferner Server (mit Port) sein.

Dispatcher. Eine Komponente von Load Balancer, die den TCP- oder UDP-Datenverkehr effizient auf Gruppen einzeln verbundener Server verteilt. Die Dispatcher-Maschine ist der Server, der den Dispatcher-Code ausführt.

Domänennamensserver. DNS. Ein vielseitig einsetzbarer, verteilter und replizierter Datenabfragedienst, der hauptsächlich im Internet für die Umsetzung von Hostnamen in Internet-Adressen verwendet wird. Bezeichnet außerdem die Darstellung des Hostnamens im Internet, obwohl ein solcher Name eigentlich ein vollständig qualifizierter Domänenname ist. Der DNS kann in der Weise konfiguriert werden, dass er basierend auf den Domänen im gesuchten Namen eine Folge von Namensservern verwendet, bis er eine Übereinstimmung findet.

dscontrol. Die Schnittstelle zur Komponente Dispatcher von Load Balancer.

dsserver. Bearbeitet in der Komponente Dispatcher die Anforderungen von der Befehlszeile an Executor, Manager und Advisor-Funktionen.

E

Eignerangaben. Repräsentiert den Eigernamen und die content-Regel des Eigners. Beide Werte werden für den Cisco CSS Switch definiert.

Erfassungsprogramm für Messwerte. Ist Bestandteil des Consultant und erfasst Messwerte.

Ethernet. Ein Standardtyp eines lokalen Netzes (LAN). Dieser Standard erlaubt mehreren Stationen den beliebigen Zugriff auf das Übertragungsmedium ohne Koordination, verhindert durch Trägerprüfung und Verzögerung Konkurrenzsituationen und beseitigt Konkurrenzsituationen durch Kollisionserkennung und Übertragung. Die von Ethernet verwendeten Softwareprotokolle variieren, umfassen aber TCP/IP.

Executor. Eine von mehreren Load-Balancer-Funktionen. Der Executor leitet Anforderungen an die TCP- oder UDP-Server weiter, überwacht die Anzahl neuer, aktiver und beendeter Verbindungen und führt für beendete oder zurückgesetzte Verbindungen eine Garbage Collection durch. Der Executor liefert die neuen und aktiven Verbindungen an die Manager-Funktion.

F

FIN. Ein Steuerungsbit (finis), das eine Folgenummer belegt. Damit wird angezeigt, dass der Sender keine weiteren Daten oder Steuerzeichen sendet, die einen Platz in der Folge beanspruchen.

FIN-Status. Der Status einer Transaktion, die beendet wurde. Ist eine Transaktion im FIN-Status, kann der Garbage Collector von Load Balancer den für diese Verbindung reservierten Speicher bereinigen.

Firewall. Ein Computer, der ein privates Netz (z. B. ein Unternehmen) mit einem öffentlichen Netz (z. B. dem Internet) verbindet. Er enthält Programme, die den Zugriff zwischen zwei Netzen einschränken.

FTP (File Transfer Protocol). Ein Anwendungsprotokoll, das von Computern in einem Netz für die Übertragung von Dateien verwendet wird. FTP erfordert für den Zugriff auf Dateien eines fernen Hostsystems eine Benutzer-ID und manchmal auch ein Kennwort.

G

Gateway. Eine Funktionseinheit, die zwei Computernetze mit unterschiedlichen Architekturen verbindet.

Gegenseitige hohe Verfügbarkeit. Bei gegenseitiger hoher Verfügbarkeit können zwei Dispatcher-Maschinen primäre Maschinen sein und gleichzeitig als Ausweichmaschine der jeweils anderen Dispatcher-Maschine verwendet werden. Siehe auch "Ausweichmaschine", "Hohe Verfügbarkeit", "Primäre Maschine".

GRE. Generic Routing Encapsulation. Ein Protokoll, das die Übertragung eines beliebigen Netzprotokolls A über ein beliebiges anderes Protokoll B ermöglicht, indem es die Pakete von A in GRE-Paketen kapselt, die dann in den Paketen von B enthalten sind.

H

Haltezeit. Das Intervall zwischen dem Schließen einer Verbindung und dem Öffnen einer neuen Verbindung. Innerhalb dieses Intervalls wird der Client an denselben Server wie bei der ersten Verbindung vermittelt. Nach Ablauf der Haltezeit kann der Client an einen anderen Server vermittelt werden.

Hohe Verfügbarkeit. Eine Funktion von Load Balancer, die die Übernahme der Aufgaben eines Load Balancers durch einen anderen ermöglicht, sollte der erste Load Balancer ausfallen.

Host. Ein mit einem Netz verbundener Computer, der ein Eingangspunkt für dieses Netz bildet. Ein Host kann ein Client, ein Server oder beides gleichzeitig sein.

Hostname. Der einem Host zugeordnete symbolische Name. Hostnamen werden über einen Domänennamensserver in IP-Adressen aufgelöst.

HTML (Hypertext Markup Language). Die Sprache, die zum Erstellen von Hypertext-Dokumenten benutzt wird. Hypertext-Dokumente enthalten Links zu anderen Dokumenten mit zusätzlichen Informationen zum hervorgehobenen Begriff oder Thema. HTML steuert beispielsweise das Format von Text und die Position von Eingabefeldern in Formularen sowie die navigierbaren Links.

HTTP (Hypertext Transfer Protocol). Das Protokoll, das zum Übertragen und Anzeigen von Hypertext-Dokumenten verwendet wird.

HTTPS (Hypertext Transfer Protocol, Secure). Das Protokoll, das zum Übertragen und Anzeigen von Hypertext-Dokumenten mit SSL verwendet wird.

I

ICMP. Internet Control Message Protocol. Ein Nachrichtensteuerungs- und Fehlermeldungsprotokoll zwischen einem Hostserver und einem Gateway zum Internet.

IMAP. Internet Message Access Protocol. Ein Protokoll, mit dem ein Client auf E-Mail-Nachrichten auf einem Server zugreifen und diese bearbeiten kann. Es ermöglicht für ferne ferne Nachrichtenordner (Mailboxes) dieselbe Art der Bearbeitung wie für lokale Mailboxes.

Internet. Der weltweite Verbund von Netzen, die die Internet-Protokollgruppe verwenden und öffentlich zugänglich sind.

Intranet. Ein sicheres privates Netz, das Internet-Standards und -Anwendungen (wie Webbrowser) in die vorhandene Computerinfrastruktur für den Netzbetrieb integriert.

IP. Internet Protocol. Ein verbindungsloses Protokoll zur Weiterleitung von Daten über ein Netz oder miteinander verbundene Netze. IP agiert als Vermittler zwischen den höheren Protokollschichten und der Bitübertragungsschicht.

IP-Adresse. Internet-Protocol-Adresse. Eine eindeutige Adresse, die die tatsächliche Position jeder Einheit oder Workstation in einem Netz angibt. Diese Adresse wird auch als Internet-Adresse bezeichnet.

IPSEC. Internet Protocol Security. Ein sich entwickelnder Standard für die Sicherheit der Vermittlungs- oder Paketebene bei der Netzkommunikation.

K

Kollokation. Installation von Load Balancer auf der Maschine, für die der Lastausgleich durchgeführt wird.

L

LAN. Local Area Network (lokales Netz). Ein Computernetz mit Einheiten, die innerhalb eines begrenzten geographischen Bereichs verbunden sind, um miteinander zu kommunizieren, und mit einem größeren Netz verbunden werden können.

Loopback-Schnittstelle. Eine Schnittstelle, die nicht erforderliche Übertragungsfunktionen umgeht, wenn die Informationen an eine Definitionseinheit innerhalb desselben Systems adressiert sind.

M

MAC-Adresse. Die Hardwareadresse (Media Access Control) einer Einheit, die mit einem gemeinsam benutzten Netzwerkdatenträger verbunden ist.

Manager. Eine von mehreren Load-Balancer-Funktionen. Der Manager legt ausgehend von internen Zählern des Executors und Rückmeldungen der Advisor-Funktionen Wertigkeiten fest. Der Executor verwendet die Wertigkeiten dann für den Lastausgleich.

Metric Server. Früher bekannt als Server Monitor Agent (SMA). Metric Server stellt systemspezifische Messwerte für den Manager von Load Balancer bereit.

Metrik. Ein Prozess oder Befehl, der einen numerischen Wert zurückgibt, der beim Lastausgleich im Netz verwendet werden kann, z. B. die Anzahl der derzeit angemeldeten Benutzer.

Metrikadresse. Die Adresse, zu der Metric Server eine Verbindung herstellt.

MIB. (1) Management Information Base. Eine Gruppe von Objekten, auf die mit einem Netzwerkverwaltungsprotokoll zugegriffen werden kann. (2) Eine Definition für Verwaltungsinformationen, die die von einem Host oder Gateway verfügbaren Informationen und die zulässigen Operationen angibt.

N

nalcontrol. Die Schnittstelle zur Load-Balancer-Komponente Nortel Alteon Controller.

nalserver. Bearbeitet im Nortel Alteon Controller die Anforderungen von der Befehlszeile an die Consultants.

Netzmaske. In IPV4 eine 32-Bit-Maske, mit der die Teilnetzadressbits im Hostabschnitt einer IP-Adresse identifiziert werden.

Netzproximität. Die Proximität zweier vernetzter Einheiten wie Client und Server, die Site Selector durch Messung der durchschnittlichen Umlaufzeit ermittelt.

Netzverwaltungsstation. In SNMP (Simple Network Management Protocol) eine Station, die Verwaltungsanwendungsprogramme ausführt, mit denen Netzelemente überwacht und gesteuert werden.

Netzwerk. Hardware und Software umfassendes Datenübertragungssystem. Netzwerke werden häufig nach ihrer räumlichen Ausdehnung in lokale Netze (LAN), Hochgeschwindigkeitsnetze (MAN) und Weitverkehrsnetze (WAN) unterteilt. Es gibt aber auch die Unterteilung nach verwendetem Protokoll.

NFA (Non-Forwarding Address). Die für Verwaltungs- und Konfigurationszwecke verwendete primäre IP-Adresse der Load-Balancer-Maschine.

NIC. Network Interface Card. Eine Adapterschaltkarte, die in einem Computer installiert ist, um eine physische Verbindung zu einem Netz zu ermöglichen.

NNTP. Network News Transfer Protocol. Ein TCP/IP-Protokoll zur Übertragung von Nachrichten.

Nortel Alteon Controller. Eine Komponente von IBM Load Balancer. Nortel Alteon Controller stellt mit der Load-Balancer-Technologie Echtzeitdaten zum Lastausgleich für den Nortel Alteon Web Switch bereit.

Nortel Alteon Web Switch. Der Nortel Alteon ACE Director Series Switch und der Nortel Alteon 180 Series Switch aus dem Alteon Web Switching Portfolio für die Weiterleitung von Paketen und Inhalten.

P

Paket. Die Dateneinheit, die im Internet oder einem anderen paketvermittelten Netz zwischen Ursprung und Ziel weitergeleitet wird.

PICS. Platform for Internet Content Selection. PICS-fähige Clients ermöglichen den Benutzern, selbst zu bestimmen, welche Bewertungsdienste sie verwenden möchten und welche Bewertungen der einzelnen Dienste akzeptabel bzw. inakzeptabel sind.

Ping. Ein Befehl, der ICMP-Echoanforderungspakete (ICMP = Internet Control Message Protocol) an einen Host, ein Gateway oder einen Router sendet und eine Antwort erwartet.

POP3. Post Office Protocol 3. Ein Protokoll, das zum Austausch von Netzpost und zum Zugriff auf Mailboxes benutzt wird.

port. Eine Nummer, die eine abstrakte Übertragungseinheit bezeichnet. Webserver verwenden standardmäßig Port 80.

Port-übergreifende Affinität. Die Port-übergreifende Affinität ist eine Affinität (Haltefunktion), die sich über mehrere Ports erstreckt. Siehe auch "Haltezeit".

Port-Umsetzung für Netzadressen. NAT (Network Address Port Translation), auch als Port-Zuordnung bekannt. Mit NAT können Sie auf einem physischen Server mehrere Serverdämonen konfigurieren, die an verschiedenen Port-Nummern empfangsbereit sind.

Primäre Maschine. Bei der Funktion für hohe Verfügbarkeit der Komponente Dispatcher die Maschine, die aktiv Pakete weiterleitet. Die zugehörige Partner- oder Ausweichmaschine überwacht den Status der primären Maschine und übernimmt ggf. deren Aufgaben. Siehe auch "Ausweichmaschine" und "Hohe Verfügbarkeit".

Priorität. Bei dem auf Regeln basierenden Lastausgleich das Maß an Bedeutung, das einer bestimmten Regel beigegeben wird. Der Dispatcher wertet die Regeln beginnend bei der ersten Prioritätsebene bis hin zur letzten Prioritätsebene aus.

Privates Netz. Ein separates Netz, in dem der Dispatcher aus Gründen des Durchsatzes mit Clusterservern kommuniziert.

Protokoll. Die Regeln, die den Betrieb von Funktionseinheiten eines DFV-Systems steuern, wenn eine Kommunikation stattfinden soll. Protokolle können Details der unteren Ebene zu Schnittstellen zwischen Maschinen festlegen, wie beispielsweise die Reihenfolge, in der die Bits eines Byte gesendet werden. Sie können auch Austauschprozesse der höheren Ebene zwischen Anwendungsprogrammen festlegen, z. B. die Dateiübertragung.

Q

Quality of Service (QoS). Leistungsmerkmale eines Netzdienstes wie Durchsatz, Transitverzögerung und Priorität. Bei einigen Protokollen können Pakete oder Datenströme QoS-Anforderungen enthalten.

Quellenadresse. Bei der Funktion für hohe Verfügbarkeit des Dispatchers die Adresse der Partnermaschine, die Überwachungssignale sendet.

R

reach. Eine Advisor-Funktion des Dispatchers, die ping-Aufrufe an eine bestimmte Zieladresse absetzt und meldet, ob die Zieladresse antwortet.

reach-Adresse. Bei der Funktion für hohe Verfügbarkeit von Dispatcher die Zieladresse, an die die Advisor-Funktion ping-Aufrufe absetzen soll, um festzustellen, ob die Zieladresse antwortet.

Regel. Beim regelbasierten Lastausgleich ein Mechanismus zum Gruppieren von Servern, der die Auswahl eines Servers ausgehend von anderen Informationen als der Zieladresse und dem Port ermöglicht.

Regeltyp. Beim regelbasierten Lastausgleich ein Anzeiger für die Informationen, die ausgewertet werden müssen, um zu bestimmen, ob eine Regel erfüllt wird.

RMI. Remote Method Invocation. Teil der Bibliothek der Programmiersprache Java, der einem Java-Programm, das auf einem Computer ausgeführt wird, ermöglicht, auf Objekte und Methoden eines auf einem anderen Computer ausgeführten Java-Programms zuzugreifen.

Root. Die uneingeschränkte Berechtigung zum Zugriff auf und Ändern von beliebige(n) Teilen des Betriebssystems AIX, Red Hat Linux oder Solaris. Diese Berechtigung wird normalerweise dem Benutzer erteilt, der das System verwaltet.

Route. Der Pfad für den Datenaustausch im Netz von der Ursprungsadresse zur Zieladresse.

Router. Eine Einheit, die Pakete zwischen Netzen weiterleitet. Die Weiterleitungsentscheidung wird ausgehend von Informationen der Vermittlungsschicht und von Routentabellen, die häufig von Routing-Produkten erstellt werden, getroffen.

RPM. Red Hat Package Manager.

Rückkehradresse. Eine eindeutige IP-Adresse oder ein eindeutiger Hostname. Die Rückkehradresse wird auf der Dispatcher-Maschine konfiguriert und vom Dispatcher bei der Verteilung der Client-Anforderungen auf die Server als Quellenadresse verwendet.

S

Schreibweise mit Trennzeichen. Die syntaktische Darstellung eines 32-Bit-Integers, das aus vier 8-Bit-Zahlen besteht, die in Dezimalschreibweise angegeben werden und durch Punkte voneinander getrennt sind. Diese Schreibweise dient zur Darstellung von IPV4-Adressen.

Server. Ein Computer, der gemeinsam genutzte Dienste für andere Computer über ein Netz bereitstellt, z. B. ein Dateiserver, ein Druckserver oder ein Postserver.

Serveradresse. Der eindeutige Code, der jedem Computer zugeordnet wird, der gemeinsam genutzte Dienste für andere Computer über ein Netz bereitstellt, z. B. einem Dateiserver, einem Druckserver oder einem Postserver. Die Serveradresse kann die IP-Adresse oder der Hostname sein.

Servermaschine. Ein Server, den der Dispatcher mit anderen Servern zu einem virtuellen Server zusammenfasst. Der Dispatcher verteilt den Datenverkehr auf die Servermaschinen. Synonym für Clusterserver.

Shell. Die Software, die Befehlszeilen von der Workstation eines Benutzers akzeptiert und verarbeitet. Die bash-Shell ist eine von mehreren verfügbaren UNIX-Shells.

Sitename. Ein Sitename ist ein nicht auflösbarer Hostname, den der Client anfordert. Beispiel: Eine Website hat drei Server (1.2.3.4, 1.2.3.5 und 1.2.3.6), die für den Sitenamen *www.dnsload.com* konfiguriert sind. Wenn ein Client diesen Sitenamen anfordert, wird eine der drei Server-IP-Adressen als Auflösung zurückgegeben. Der Sitename muss ein vollständig qualifizierter Domänenname wie *dnsload.com* sein. Ein nicht qualifizierter Name, z. B. *dnsload*, ist als Sitename ungültig.

Site Selector. Eine DNS-gestützte Lastausgleichskomponente von Load Balancer. Site Selector verteilt die Last auf Server innerhalb eines Weitverkehrsnetzes (WAN) und verwendet dafür Messungen und Wertigkeiten, die von der auf diesen Servern aktiven Komponente Metric Server erfasst werden.

Skalierbar. Im Kontext des Leistungsspektrums eines Systems die schnelle Anpassung an Schwankungen bei der Auslastung. Ein skalierbares System kann beispielsweise gut an größere oder kleinere Netze angepasst werden und Tasks unterschiedlicher Komplexität ausführen.

SMTP. Simple Mail Transfer Protocol. In der Internet-Protokollgruppe ein Anwendungsprotokoll zum Übertragen von Post zwischen Benutzern in der Internet-Umgebung. SMTP gibt die Postaustauschfolgen und das Nachrichtenformat an. SMTP setzt voraus, dass TCP (Transmission Control Protocol) das zugrundeliegende Protokoll ist.

SNMP. Simple Network Management Protocol. Das in STD 15, RFC 1157, definierte Internet-Standardprotokoll, das für die Verwaltung von Knoten in einem IP-Netz entwickelt wurde. SNMP ist nicht auf TCP/IP beschränkt. Es kann zum Verwalten und Überwachen aller Arten von Einrichtungen verwendet werden, einschließlich Computer, Router, Vernetzungs-Hubs, Toaster und Jukeboxes.

SPARC. Scalable Processor Architecture.

sscontrol. Die Schnittstelle zur Komponente Site Selector von Load Balancer.

SSL. Secure Socket Layer. Ein bekanntes Sicherheitsschema, das von der Netscape Communications Corporation in Zusammenarbeit mit RSA Data Security, Inc. entwickelt wurde und dem Client ermöglicht, den Server zu authentifizieren und alle Daten und Anforderungen zu verschlüsseln. Der URL eines mit SSL gesicherten Servers beginnt mit https und nicht mit HTTP.

ssserver. Bearbeitet für die Komponente Site Selector die Anforderungen von der Befehlszeile an Sitenamen, Manager und Advisor-Funktionen.

Standardeinstellung. Ein Wert, ein Attribut oder eine Option, die verwendet werden, wenn keine explizite Angabe vorliegt.

Stilllegen. Das Beenden eines Prozesses mit vollständigem normalem Abschluss laufender Operationen.

strategy. Bei der Funktion für hohe Verfügbarkeit des Dispatchers ein Schlüsselwort, das angibt, wie eine ausgefallene Maschine wiederhergestellt werden soll.

SYN. Ein Steuerungsbit im eingehenden Segment, das eine Folgenummer belegt und bei der Initialisierung einer Verbindung angibt, wo die Folgenummernvergabe beginnt.

T

TCP. Transmission Control Protocol. Ein im Internet verwendetes Übertragungsprotokoll. TCP ermöglicht einen zuverlässigen Austausch von Informationen zwischen Hosts. TCP verwendet IP als zugrundeliegendes Protokoll.

TCP/IP . Transmission Control Protocol/Internet Protocol. Eine Protokollgruppe, die die Übertragung zwischen Netzen unabhängig von den in den einzelnen Netzen verwendeten Übertragungstechnologien ermöglicht.

TCP-Servermaschine. Ein Server, den Load Balancer mit anderen Servern zu einem virtuellen Server zusammenfasst. Load Balancer verteilt den TCP-Datenverkehr auf die TCP-Servermaschinen. Synonym für Clusterserver.

Teilnetzmaske. In IPV4 eine 32-Bit-Maske, mit der die Teilnetzadressbits im Hostabschnitt einer IP-Adresse identifiziert werden.

Telnet. Terminalemulationsprotokoll. Ein TCP/IP-Anwendungsprotokoll für Fernverbindungsdienste. Mit Telnet kann ein Benutzer so auf einen fernen Host zugreifen, als wäre seine Workstation direkt mit diesem fernen Host verbunden.

TOS. Type of Service (Diensttyp). Ein 1-Byte-Feld im IP-Header des SYN-Pakets.

TTL. DNS TTL (Time To Live) ist die Zeit in Sekunden, die ein Client die Namensauflösungsantwort zwischenspeichern kann.

U

Überwachungssignal. Ein einfaches Paket, das zwischen zwei Load Balancer-Maschinen im Modus für hohe Verfügbarkeit übertragen wird und vom Bereitschafts-Load-Balancer zur Überwachung des Zustandes des aktiven Load Balancers verwendet wird.

UDP. User Datagram Protocol. In der Internet-Protokollgruppe ein Protokoll für einen unzuverlässigen, verbindungslosen Datagrammdienst. Mit UDP kann ein Anwendungsprogramm auf einer Maschine oder ein Prozess ein Datenpaket an ein Anwendungsprogramm auf einer anderen Maschine oder einen anderen Prozess senden. UDP benutzt das Internet Protocol (IP) zum Senden von Datenpaketen.

Umsetzer für Netzadressen. NAT (Network Address Translator), virtuelles LAN. Eine Hardwareeinheit, die zur Zeit in Entwicklung ist und zur Erweiterung der vorhandenen Internet-Adressen verwendet werden soll. Sie erlaubt duplizierte IP-Adressen innerhalb einer Firma und eindeutige Adressen außerhalb der Firma.

URI. Universal Resource Identifier. Die codierte Adresse für jede Ressource im Web, z. B. ein HTML-Dokument, ein Bild, ein Videoclip, ein Programm usw.

URL. Uniform Resource Locator. Eine standardisierte Angabe der Position eines Objektes, in der Regel einer Webseite im Internet. URLs sind das im World Wide Web verwendete Adressenformat. In HTML-Dokumenten gibt der URL das Ziel eines Hyperlink an, bei dem es sich häufig um ein anderes HTML-Dokument handelt (das eventuell auf einem anderen Computer gespeichert ist).

V

Verknüpfung mehrerer Adressen. Die Verknüpfung mehrerer Adressen ermöglicht dem Kunden, in der Konfiguration für den verknüpften Server eine andere Adresse als die NFA anzugeben. Siehe auch "Kollokation".

Verwalteter Knoten. In der Internet-Kommunikation eine Workstation, ein Server oder ein Router mit einem Netzwerkverwaltungsagenten. Im Internet Protocol (IP) enthält der verwaltete Knoten normalerweise einen SNMP-Agenten (SNMP = Simple Network Management Protocol).

Vollständig qualifizierter Domänenname. Der vollständige Name eines Systems, bestehend aus dem lokalen Hostnamen und dem Domänennamen einschließlich einer Domäne der höchsten Ebene. Wenn "venera" ein Hostname ist, wäre "venera.isi.edu" beispielsweise ein vollständig qualifizierter Domänenname. Anhand eines vollständig qualifizierten Domänennamens sollte für jeden Host im Internet eine eindeutige Internet-Adresse bestimmt werden können. Dieser Prozess wird als "Namensauflösung" bezeichnet und verwendet das Domänennamensystem (DNS).

VPN. Virtuelles privates Netz. Ein Netz, das aus einem oder mehreren gesicherten IP-Tunnel(n) besteht, die zwei oder mehr Netze verbinden.

W

WAN. Wide Area Network (Weitverkehrsnetz). Ein Netz, das Übertragungsdienste für ein geografisches Gebiet bereitstellt, das größer als das von einem lokalen Netz oder einem Hochgeschwindigkeitsnetz versorgte Gebiet ist. Ein WAN kann öffentliche Übertragungseinrichtungen verwenden oder zur Verfügung stellen.

WAP. Wireless Application Protocol. Ein offener internationaler Standard für Anwendungen, die festnetz-unabhängige Kommunikation verwenden, z. B. Internet-Zugriff über ein Handy.

WAS. WebSphere Application Server.

Web. Das Netz von HTTP-Servern, das Programme und Dateien enthält, von denen viele Hypertext-Dokumente mit Links zu anderen Dokumenten auf HTTP-Servern sind. Das Web wird auch als World Wide Web bezeichnet.

WLM. Workload Manager. Eine zum Dispatcher gehörige Advisor-Funktion. WLM ist für Server auf OS/390-Großrechnern bestimmt, die die Komponente MVS Workload Manager (WLM) ausführen.

Z

Zeitlimit. Das Zeitintervall, das für die Ausführung einer Operation zugeteilt wurde.

Zieladresse. Die Adresse der Partnermaschine für hohe Verfügbarkeit, an die Überwachungssignale und Antworten gesendet werden.

Index

A

Abrufen von Informationen 283
Advisor-Funktionen
 Anforderung/Antwort der HTTP-Advisor-Funktion 191
 angepasstes Beispiel 487
 Beispielkonfigurationsdatei 487
 cbrcontrol 351
 Controller 250
 anpassen 252
 Ruhezeit 251
 schnelle Ausfallerkennung 251
 Serverempfangszeitlimit 251
 Serververbindungszeitlimit 251
 Wiederholungen für Server 251
 dscontrol 351
 Einschränkung unter Solaris 185
 Hinweise zu IPV6 84
 Komponente CBR
 Advisor-Funktion ssl2http 190
 Komponente Dispatcher 185
 Advisor-Funktion self 190, 192
 anpassen 193
 Bericht 355
 Berichtszeitlimit 188, 354
 Caching-Proxy-Advisor-Funktion 190
 Intervall 187, 354
 Liste 189, 354
 Name 351
 Port 358
 schnelle Ausfallerkennung 188
 Serverempfangszeitlimit 188, 352, 355
 Serververbindungszeitlimit 188, 351, 354
 starten 71, 354
 starten/stoppen 186
 Statusbericht 355
 stoppen 354
 Version 355
 Wiederholungen für Server 182, 188, 353
 Liste 353
 Site Selector
 Berichtszeitlimit 408, 410
 Intervall 406, 409
 Liste 406, 407, 409
 Name 406
 Port 351, 406
 Protokollstufe 406
 schnelle Ausfallerkennung 188
 Serverempfangszeitlimit 188, 407, 409
 Serververbindungszeitlimit 188, 406, 409
 starten 407, 409
 Statusbericht 407, 409
 stoppen 408, 410
 Version 409, 410

Advisor-Funktionen (*Forts.*)
 Site Selector (*Forts.*)
 Wiederholungen für Server 188, 407
 sscontrol 406, 413
 URL-Option der HTTP-Advisor-Funktion 191
Advisor-Funktionen, Komponente von Load Balancer
 starten 71
Affinität (Bindung)
 Affinitätsadressmaske 223
 aktive Cookie-Affinität 225, 391
 Funktionsweise 222
 passive Cookie-Affinität 225, 227, 391
 Port-Affinität außer Kraft setzen 220
 Port-übergreifende Affinität 223, 224, 382
 quiesce now 224, 376, 380
 Regeloption 225
 SSL-ID (Weiterleitungsmethode cbr) 56
 sticky (Außerkräftsetzen der Port-Affinität) 220, 395
 stickymask 223, 383
 stickytime 56, 222, 223, 383, 391
 URI-Affinität 225, 228, 391
Affinitätsadressmaske 223, 383
AIX
 installieren 32
 Voraussetzungen 31
Aktive Cookie-Affinität 225, 391
Aktualisieren der Konfiguration von einem fernen Standort aus 266
Alerts
 Controller 259
 Dispatcher, CBR, Site Selector 184
Aliasname
 für NIC 69, 116
 Loopback-Einheit 72
Allgemeiner Schlüssel
 für ferne Authentifizierung 264
Anzeige
 globale Werte und ihre Standardeinstellungen
 für den Manager 379, 417, 418
 für einen Advisor 355, 408, 410
 interne Zähler 365
 Liste
 Advisor, die Messungen durchführen 354, 409
 Statistik 378, 416, 417
 Status
 ein Cluster oder alle Cluster 360
 Server für einen Port 386
 Statusbericht für eine Advisor-Funktion 355, 407, 409
Versionsnummer
 von Advisor 355, 409, 410
 von Manager 380, 417, 419

Assistent für Konfiguration
 CBR 113
 Dispatcher 66
 Site Selector 134
Auflösung, GUI 308
Ausweichmaschine, hohe Verfügbarkeit konfigurieren 205

B

Befehle
 cbrcontrol
 advisor 351
 binlog 356
 cluster 357
 executor 361
 file 366
 help 368
 host 373
 logstatus 374
 manager 375
 metric 381
 port 382
 rule 388
 server 394
 set 400
 status 401
 ccocontrol
 consultant 434, 437
 Eingabeaufforderung 433
 file 439
 help 440
 host 446
 metric 444
 Server konfigurieren 449
 Cisco CSS Controller 433
 dscontrol
 advisor 351
 binlog 356
 cluster 357
 Eingabeaufforderung 350
 executor 361
 file 366
 help 368
 hohe Verfügbarkeit steuern 369, 459
 host 373
 logstatus 374
 manager 375
 metric 381
 port 382
 rule 388
 server 394
 set 400
 SNMP-Subagenten konfigurieren 402
 status 401
 zum Definieren der NFA 69, 365
 zum Definieren eines Ports 70
 zur Definition eines Servers 71

- Befehle (*Forts.*)
 - dscontrol (*Forts.*)
 - zur Steuerung der Advisor-Funktion 71
 - zur Steuerung des Managers 71
 - ifconfig 70, 232
 - Aliasnamen auf der Loopback-Einheit angeben 73
 - nalcontrol
 - consultant 452, 455
 - Eingabeaufforderung 451
 - file 457
 - help 458
 - host 466
 - metric 462
 - Server konfigurieren 464
 - ndcontrol
 - hohe Verfügbarkeit steuern 441
 - netstat
 - zur Überprüfung der IP-Adressen und Aliasnamen 76
 - Nortel Alteon Controller 451
 - route
 - zum Löschen einer zusätzlichen Route 76, 77
 - Site Selector 405
 - sscontrol
 - advisor 406
 - file 411
 - help 413
 - logstatus 414
 - manager 415
 - metric 420
 - nameserver 421
 - rule 422
 - server 425
 - set 427
 - sitename 428
 - status 431
 - Befehlsreferenzen
 - lesen 347
 - Befehlszeile
 - Befehl senden (GUI) 474
 - Konfigurationsbeispiel
 - CBR 100
 - Cisco CSS Controller 140
 - Dispatcher 47
 - Nortel Alteon Controller 158
 - Site Selector 124
 - Beispiel für einen schnellen Start 45
 - CBR 99
 - Cisco CSS Controller 139
 - Nortel Alteon Controller 157
 - Site Selector 123
 - Beispiele
 - lokale Server verwalten 10, 11, 13, 14, 15
 - schneller Start 45
 - CBR 99
 - Cisco CSS Controller 139
 - Nortel Alteon Controller 157
 - Site Selector 123
 - Beispielkonfigurationsdateien 481
 - Advisor-Funktion 487
 - Komponente Dispatcher (AIX) 481
 - Komponente Dispatcher (Windows) 484
 - Bemerkungen 495
 - Benutzer-Exit, Scripts 184, 259
 - ccoallserversdown 260
 - ccoserverdown 260
 - ccoserverup 260
 - Denial of Service (DoS) erkennen 240
 - managerAlert 184
 - managerClear 184
 - nalallserversdown 260
 - naloserverup 260
 - nalserverdown 260
 - serverDown 184
 - serverUp 184
 - Benutzerbereich, Unterstützung 82
 - Bericht
 - Cisco CSS Controller 437
 - Nortel Alteon Controller 455
 - Binäre Protokollierung für Serverstatistik 241, 268, 270
 - Controller 258
 - Bindung (Affinität)
 - Affinitätsadressmaske 223
 - aktive Cookie-Affinität 225, 391
 - Funktionsweise 222
 - passive Cookie-Affinität 225, 227, 391
 - Port-Affinität außer Kraft setzen 220
 - Port-übergreifende Affinität 223, 224, 382
 - quiesce now 224, 376, 380
 - sticky (Außerkräftsetzen der Port-Affinität) 220, 395
 - stickymask 223, 383
 - stickytime 56, 222, 223, 383, 391
 - URI-Affinität 225, 391
 - Bindungsspezifische Server 70, 71, 185, 232
 - binlog
 - binäres Protokoll für Statistik 356
 - cbrcontrol 356
 - dscontrol 356
- ## C
- Caching Proxy 107
 - für CBR konfigurieren 114
 - Caching-Proxy-Advisor-Funktion 190
 - CBR
 - Advisor-Funktionen und Erreichbarkeitsziele markieren alle Server als inaktiv (Windows) 330
 - Aliasname für NIC 116
 - Anforderungen werden nicht verteilt 329
 - Befehl ifconfig 116
 - bei Webverwaltung wird Hostverbindung getrennt 330
 - Beispiel für einen schnellen Start 99
 - beschädigte nationale Sonderzeichen (Latin-1) unter Windows 330
 - cbrcontrol scheitert 328
 - cbrcontrol scheitert unter Solaris 329
 - Einstellungen für Lastausgleich 180
 - Wiederholungsversuche der Advisor-Funktion für Server 188
 - erforderliche Funktionen bestimmen 23
 - CBR (*Forts.*)
 - Fehler bei der Auflösung von IP-Adressen in Hostnamen (Windows) 331
 - Java-Fehler wegen Speicherkapazität/Threads (HP-UX) 330
 - Konfiguration
 - CBR-Maschine konfigurieren 113
 - Tasks im Überblick 109
 - lbadmin scheitert 328
 - mit Caching Proxy
 - Advisor-Funktion ssl2http 108
 - konfigurieren 119
 - Schlüsselwort mapport 108
 - SSL-Verbindungen 107
 - Übersicht 106
 - mit der Komponente Dispatcher 55
 - Planung 105
 - starten und stoppen 279
 - Syntax- oder Konfigurationsfehler 329
 - Tabellen zur Fehlerbehebung 294
 - unerwartetes GUI-Verhalten bei Verwendung von Matrox-AGP-Karten 330
 - wird nicht ausgeführt 328
 - cbr, Weiterleitungsmethode 55, 57
 - stickytime 56
 - cbrcontrol, Befehl
 - advisor 351
 - binlog 356
 - cluster 357
 - executor 361
 - file 366
 - help 368
 - host 373
 - logstatus 374
 - manager 375
 - metric 381
 - port 382
 - rule 388
 - server 394
 - set 400
 - status 401
 - cbrserver
 - starten 100
 - ccocontrol, Befehl
 - consultant 434, 437
 - Eingabeaufforderung 433
 - file 439
 - help 440
 - host 446
 - metric 444
 - server 449
 - ccoserver
 - starten 140
 - wird nicht gestartet 303, 304, 334
 - Cisco CSS Controller
 - Advisor-Funktion Workload Manager 257
 - Advisor-Funktionen 250
 - Alerts 259
 - Befehl refresh aktualisiert die Konfiguration nicht 336
 - Befehle 433
 - bei Webverwaltung wird Hostverbindung getrennt 336

Cisco CSS Controller (*Forts.*)
 Beispiel für einen schnellen Start 139
 Bericht
 Controller 437
 binäre Protokollierung für Server-
 statistik 258
 ccocontrol scheitert 334
 Consultant-Verbindungsfehler 335
 Einstellungen für Lastausgleich 248
 erforderliche Funktionen bestim-
 men 27
 für Port 13099 kann kein Eintrag in
 der Registrierungsdatenbank erstellt
 werden 334
 Hardware- und Softwarevorausset-
 zungen 143
 hohe Verfügbarkeit 245
 Java-Fehler wegen Speicherkapazität/
 Threads (HP-UX) 336
 Konfiguration
 Beispiel 15
 CSS-Maschine konfigurieren 152
 Tasks im Überblick 149
 lbadmian scheitert 334
 Metric Server 255
 Planung 143
 starten 280
 starten und stoppen 280
 Tabellen zur Fehlerbehebung 297
 unerwartetes GUI-Verhalten bei Ver-
 wendung von Matrox-AGP-Kar-
 ten 335
 verknüpfen 245
 verwenden 280
 Wertigkeiten werden auf dem Switch
 nicht aktualisiert 335
 wird nicht gestartet 334
 Cisco CSS Controller, Komponente
 beschädigte nationale Sonderzeichen
 (Latin-1) unter Windows 336
 Cluster
 Adresse konfigurieren 69
 Anzeige
 Status dieses Clusters 360
 cbrcontrol 357
 definieren 69, 360
 dscontrol 357
 entfernen 360, 429
 hinzufügen 360
 Platzhalter 69
 Proportionen 357
 Proportionen festlegen 72
 Clusterspezifisch
 Proportionen 428
 collocated, Schlüsselwort 203, 399
 Consultant
 ccocontrol 434, 437
 Cisco CSS Controller
 Bericht 434
 binäres Protokoll 434
 hinzufügen 434
 nalcontrol 452, 455
 Nortel Alteon Controller
 Bericht 452
 binäres Protokoll 452
 hinzufügen 452
 starten 153, 175

Content Based Routing 5
 Einstellungen für Lastausgleich 180
 Konfiguration
 CBR-Maschine konfigurieren 113
 Tasks im Überblick 109
 mit der Komponente Dispatcher 55
 Planung 105
 Tabellen zur Fehlerbehebung 294
 verwenden 279
 content-Regel 55, 219
 Controller
 Cisco CSS Controller
 Bericht 437
 definieren 437
 Protokollgröße 435, 437
 Protokollstufe 435, 437
 Einstellungen für Lastausgleich
 Advisor-Ruhezeiten 251
 Bedeutung von Messwerten 248
 Ruhezeiten 249
 Sensitivitätsschwelle 250
 Serverzeitlimit der Advisor-Funkti-
 on 251
 Wertigkeiten 249
 Wiederholungsversuche der Advi-
 sor-Funktion für Server 251
 feste Wertigkeit 249
 Kundenspezifische (anpassbare) Advi-
 sor-Funktion 252
 Nortel Alteon Controller
 Bericht 455
 definieren 455
 Protokollgröße 453, 455
 Protokollstufe 453, 455

D

Datei
 cbrcontrol 111, 366
 ccocontrol 439
 dscontrol 64, 366
 nalcontrol 457
 sscontrol 132, 411
 DB2-Advisor-Funktion 190
 default.cfg 68, 115, 134
 Definieren
 Cluster 360
 NFA 69, 365
 Port für einen Cluster 70, 386
 Server für einen Port 71, 398, 426
 Deinstallieren
 unter AIX 32
 unter HP-UX 36
 unter Linux 37
 unter Solaris 39
 unter Windows 41
 Dienst
 Konfiguration 175
 Dispatcher
 erforderliche Funktionen bestim-
 men 19
 Konfiguration
 Back-End-Server konfigurieren 72
 Dispatcher, Komponente
 Advisor-Funktionen arbeiten nicht
 korrekt 306

Dispatcher, Komponente (*Forts.*)
 Advisor-Funktionen und
 Erreichbarkeitsziele markieren alle
 Server als inaktiv (Windows) 316
 Advisor-Funktionen werden in Konfi-
 guration für hohe Verfügbarkeit
 nach Netzwerkausfall nicht ausge-
 führt (Windows) 317
 an Stelle der lokalen Adresse wird die
 Aliasadresse zurückgegeben 312
 automatische Pfaderkennung verhin-
 dert Datenrückfluss mit Load Balan-
 cer 309
 Back-End-Server empfängt unter
 Linux nicht von Dispatcher weiter-
 geleitete Pakete 325
 bei Webverwaltung wird Hostverbin-
 dung getrennt 314
 beschädigte nationale Sonderzeichen
 (Latin-1) unter Windows 315
 blaue Anzeige beim Starten des Exe-
 cutors 309
 Client-Anforderungen scheitern, wenn
 die Rückgabe von Antworten in
 Form großer Seiten versucht
 wird 320
 Content Based Routing 55
 Dispatcher mit hoher Verfügbarkeit
 kann unter Linux nicht synchroni-
 siert werden 321
 dscontrol scheitert 307
 Einschränkungen unter Linux bei Ver-
 wendung von zSeries- oder S/390-
 Servern 323
 Einstellungen für Lastausgleich 180
 Advisor-Intervalle 187
 Berichtszeitlimit für Advisor-Funkti-
 on 188
 Glättungsfaktor 184
 Manager-Intervalle 183
 proportionale Bedeutung von
 Statusinformationen 180
 Sensitivitätsschwelle 183
 Serverzeitlimit der Advisor-Funkti-
 on 188
 Wertigkeiten 181
 Wiederholungsversuche der Advi-
 sor-Funktion für Server 182, 188
 Fehler "Der Server antwortet nicht"
 unter Windows 321
 Fehler bei der Auflösung von IP-
 Adressen in Hostnamen (Win-
 dows) 316
 Fehler bei installiertem Caching Pro-
 xy 308
 GUI startet nicht richtig 308
 GUI wird nicht richtig angezeigt 308
 Hilfefenster kann nicht geöffnet wer-
 den 308
 Hilfefenster sind nicht zu sehen 308
 hohe Verfügbarkeit funktioniert
 nicht 305
 in Konfiguration mit hoher Verfügbar-
 keit sind primäre und Ausweich-
 maschinen aktiv 320
 inaktive Server zurücksetzen 385
 inaktiven Server zurücksetzen 182

Dispatcher, Komponente (Forts.)

- IP-Adresse wird über Fernverbindung nicht aufgelöst 312
- IP-Adressenkonflikt bei Verwendung der hohen Verfügbarkeit 319
- IPV6-Unterstützung 81
- Java-Fehler wegen Speicherkapazität/Threads (HP-UX) 315
- Java-Warnung bei Installation von Servicekorrekturen 328
- kein Lastausgleich für Anfragen 305
- keine hohe Verfügbarkeit im Verkehrsmodus von Load Balancer 310
- Konfiguration
 - Maschine mit Load Balancer konfigurieren 66
 - privates Netz konfigurieren 236
 - Tasks im Überblick 63
- Konfigurationstipps für hohe Verfügbarkeit 321
- lange Antwortzeiten 313
- lbadm scheitert 307
- lbadm trennt nach dem Aktualisieren der Konfiguration die Verbindung zum Server 311
- Load-Balancer-Prozesse enden (Solaris) 319
- MAC-Weiterleitung 53
- MS IIS und SSL funktionieren nicht 306
- NAT/NAPT 53
- nicht Befehl 'IP address add' zum Festlegen des Aliasnamens auf der Loopback-Einheit verwenden (Linux) 318
- Planung 51
- Problem mit Funktionsübernahme bei hoher Verfügbarkeit auf Windows-Systemen 326
- Rahmen kann nicht weitergeleitet werden 309
- Router-Adresse nicht angegeben oder nicht gültig für Port-Methode 318
- Server antwortet nicht 305
- Serverlast wird nicht registriert 314
- Speicherverlust unter Linux bei Verwendung mit Manager und Advisor-Funktionen 325
- starten 270
- Tabellen zur Fehlerbehebung 288
- Überwachungssignal kann nicht hinzugefügt werden 306
- unerwartetes GUI-Verhalten bei Verwendung von Matrox-AGP-Karten 313
- unerwartetes Verhalten bei Ausführung von rmmmod ibmlb 313
- unerwartetes Verhalten beim Laden einer großen Konfigurationsdatei 311
- unpassende koreanische Schriftarten unter AIX und Linux 312
- unter Solaris können keine IPV6-Server zur Konfiguration hinzugefügt werden 327

Dispatcher, Komponente (Forts.)

- Upgrade der Java-Dateigruppe für die Installation 328
- Verbindung zu einer fernen Maschine 307
- verwenden 270
- Verzögerung beim Laden einer Load-Balancer-Konfiguration 319
- Webserver wird an 0.0.0.0 gebunden 314
- wird nicht ausgeführt 305
- zusätzliche Routen (Windows) 306
- DoS-Attacke erkennen 240
 - halfopenaddressreport 386
 - maxhalfopen 385
- DPID2 273
- dscontrol, Befehl
 - advisor 71, 351
 - Befehlsparameter abkürzen 350
 - binlog 356
 - cluster 357
 - Eingabeaufforderung 350
 - executor 69, 361
 - file 366
 - help 368
 - highavailability 369, 459
 - host 373
 - logstatus 374
 - manager 71, 375
 - metric 381
 - port 70, 382
 - rule 388
 - server 71, 394
 - set 400
 - status 401
 - subagent 402
- dsserver
 - starten 47

E

- Einstellungen, Anzeige aller globalen Werte
 - für den Manager 379, 417, 418
 - für einen Advisor 355, 408, 410
- Entfernen
 - Cluster 360, 429
 - Port von einem Cluster 386
 - Server von einem Port 398, 425, 426
 - zusätzliche Route 77
- Ethernet-NIC
 - ibmlb.conf
 - für Solaris konfigurieren 67
- Executor
 - cbrcontrol 361
 - dscontrol 361
 - starten 365
 - stoppen 365
- Explizite Verbindung 236

F

- Fehlerbehebung 283
 - Advisor-Funktionen arbeiten nicht korrekt 306

Fehlerbehebung (Forts.)

- Advisor-Funktionen und Erreichbarkeitsziele markieren alle Server als inaktiv (Windows) 316, 330, 333
- Advisor-Funktionen werden in Konfiguration für hohe Verfügbarkeit nach Netzwerkausfall nicht ausgeführt (Windows) 317
- allgemeine Probleme und Lösungen 305, 307, 328, 331, 334, 336, 339
- an Stelle der lokalen Adresse wird die Aliasadresse zurückgegeben 312
- Anforderungen werden nicht verteilt 329
- Ausgabe des Befehls ps -vg unter AIX beschädigt 340
- automatische Pfaderkennung verhindert Datenrückfluss mit Load Balancer 309
- Back-End-Server empfängt unter Linux nicht von Dispatcher weitergeleitete Pakete 325
- Befehl cbrcontrol oder lbadm scheitert 328
- Befehl ccocontrol oder lbadm scheitert 334
- Befehl dscontrol oder lbadm scheitert 307
- Befehl nalcontrol oder lbadm scheitert 337
- Befehl refresh aktualisiert die Konfiguration nicht 336, 338
- Befehl sscontrol oder lbadm scheitert 331
- bei Webverwaltung wird Hostverbindung getrennt 314, 330, 333, 336, 338
- beschädigte nationale Sonderzeichen (Latin-1) unter Windows 315, 330, 333, 336, 339
- blaue Anzeige beim Starten des Executors von Load Balancer 309
- CBR wird nicht ausgeführt 328
- cbrcontrol scheitert unter Solaris 329
- ccoserver wird nicht gestartet 334
- Client-Anforderungen scheitern, wenn die Rückgabe von Antworten in Form großer Seiten versucht wird 320
- Consultant-Verbindungsfehler 335, 338
- Dispatcher, Microsoft IIS und SSL arbeiten nicht 306
- Dispatcher-Anforderungen werden nicht weitergeleitet 305
- Dispatcher-Funktion für hohe Verfügbarkeit arbeitet nicht 305
- Dispatcher mit hoher Verfügbarkeit kann unter Linux nicht synchronisiert werden 321
- Dispatcher und Server antworten nicht 305
- Dispatcher wird nicht ausgeführt 305
- Einschränkungen unter Linux bei Verwendung von zSeries- oder S/390-Servern 323

Fehlerbehebung (Forts.)

- Fehler "Der Server antwortet nicht" unter Windows 321
- Fehler bei der Auflösung von IP-Adressen in Hostnamen (Windows) 316, 331
- Fehler bei der Ausführung von Dispatcher mit installiertem Caching Proxy 308
- Fehlernachricht bei dem Versuch, Onlinehilfetexte anzuzeigen 308
- für Port 13099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden 334
- für Port 14099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden 337
- GUI startet nicht richtig 308
- GUI wird nicht richtig angezeigt 308
- Hilfenfenster sind nicht zu sehen 308
- Hinzufügen des Überwachungssignals nicht möglich 306
- in Konfiguration mit hoher Verfügbarkeit sind primäre und Ausweichmaschinen aktiv 320
- IOException für Metric Server unter Windows 339
- IP-Adresse wird über Fernverbindung nicht aufgelöst 312
- IP-Adressenkonflikt bei Verwendung der hohen Verfügbarkeit 319
- Java-Fehler wegen Speicherkapazität/Threads (HP-UX) 315, 330, 333, 336, 339
- Java-Warnung bei Installation von Servicekorrekturen 328
- kein Abruf der Werte von Metric Server auf Linux-Systemen möglich 342
- keine hohe Verfügbarkeit im Weitverkehrsmodus von Load Balancer 310
- Konfigurationstipps für hohe Verfügbarkeit 321
- lange Antwortzeiten 313
- lbadmin trennt nach dem Aktualisieren der Konfiguration die Verbindung zum Server 311
- Load Balancer kann Rahmen nicht verarbeiten und weiterleiten. 309
- Load-Balancer-Prozesse enden (Solaris) 319
- Metric Server in einer zweistufigen Konfiguration konfigurieren 340
- Metric Server meldet keine Last 339
- Metric-Server-Protokoll meldet, dass für den Zugriff auf den Agenten eine Kennung erforderlich ist 340
- nalserver wird nicht gestartet 336
- nicht Befehl 'IP address add' zum Festlegen des Aliasnamens auf der Loopback-Einheit verwenden (Linux) 318
- Problem mit Funktionsübernahme bei hoher Verfügbarkeit auf Windows-Systemen 326

Fehlerbehebung (Forts.)

- Router-Adresse nicht angegeben oder nicht gültig für Port-Methode 318
 - Rückgabe von -1 nach dem Start von Metric Server 343
 - Serverlast wird nicht registriert 314
 - Site Selector führt den Lastausgleich nicht korrekt durch 332
 - Site Selector wendet keine RoundRobin-Methode an (Solaris) 331
 - Site Selector wird nicht ausgeführt 331
 - Speicherverlust unter Linux bei Verwendung mit Manager und Advisor-Funktionen 325
 - ssserver wird unter Windows nicht gestartet 332
 - Syntax- oder Konfigurationsfehler 329
 - unerwartetes GUI-Verhalten bei Verwendung von Matrox-AGP-Karten 313, 330, 333, 335, 337
 - unerwartetes Verhalten bei Ausführung von rmmmod ibmlb 313
 - unerwartetes Verhalten beim Laden einer großen Konfigurationsdatei 311
 - unerwünschte Konsolnachrichten von Scripts unter Solaris 342
 - unpassende koreanische Schriftarten unter AIX und Linux 312
 - unter Solaris können keine IPV6-Server zur Konfiguration hinzugefügt werden 327
 - Upgrade der Java-Dateigruppe für die Installation 328
 - Verzögerung beim Laden einer Load-Balancer-Konfiguration 319
 - vom Dispatcher verwendete Port-Nummern 301
 - von CBR verwendete Port-Nummern 302
 - von Cisco CSS Controller verwendete Port-Nummern 303
 - von Nortel Alteon Controller verwendete Port-Nummern 304
 - von Site Selector verwendete Port-Nummern 302
 - Webserver wird an 0.0.0.0 gebunden 314
 - Wertigkeiten werden auf dem Switch nicht aktualisiert 335, 338
 - zusätzliche Routen 306
- ### Fehlerbehebung, Tabellen
- CBR 294
 - Cisco CSS Controller 297
 - Komponente Dispatcher 288
 - Metric Server 299
 - Nortel Alteon Controller 298
 - Site Selector 295
- ### Fernverwaltung
- 35, 39, 40, 41
 - RMI 263, 264
 - webgestützte Verwaltung 263, 265
- ### Fernverwaltung (webgestützt)
- aktualisieren 266

Festlegen

- Anzahl der Abfragen des Executors durch den Manager 183, 378
 - cbrcontrol 400
 - Clusteradresse 70
 - dscontrol 400
 - Glättungsfaktor 184, 379, 416, 418
 - Intervallzeit
 - für Advisor zur Abfrage der Server 354, 409
 - für Manager zur Aktualisierung des Executors 183, 378, 415, 417
 - maximale Größe des Protokolls
 - für Advisor-Funktion 267, 355, 407, 409
 - für den Manager 378, 415, 417
 - maximale Wertigkeit
 - für Server an einem bestimmten Port 181, 386
 - Name der Protokolldatei 408
 - für den Manager 417
 - NFA 66
 - proportionale Bedeutung beim Lastausgleich 360
 - Protokollstufe
 - für Advisor-Funktion 267, 355, 409
 - für den Manager 415
 - Sensitivität für Aktualisierung von Wertigkeiten 183, 379, 416, 418
 - sscontrol 427
 - Wertigkeit für einen Server 378, 380, 399, 425
- ### Firewall (Einschränkung) 41
- ftp-Advisor 351, 406

G

- Gegenseitige hohe Verfügbarkeit 61, 205, 207
 - primärer Host 359, 360
 - Scripts 210
 - Übernahme 209
- Glättungsfaktor einstellen 184, 379, 416, 418
- goActive 210
- goldle 210
- goInOp 210
- goStandby 210
- Grafische Benutzerschnittstelle (GUI)
 - allgemeine Anweisungen 469
 - CBR 112
 - Cisco CSS Controller 151
 - Dispatcher 65
 - Nortel Alteon Controller 173
 - Site Selector 133
- GRE (Generic Routing Encapsulation)
 - Linux 235
 - OS/390 235
 - WAN-Unterstützung 235
- GUI
 - allgemeine Anweisungen 469
 - Auflösung 308
 - CBR 112
 - Cisco CSS Controller 151
 - Dispatcher 65
 - Nortel Alteon Controller 173

GUI (*Forts.*)
Site Selector 133

H

Hardwarevoraussetzungen
Cisco CSS Controller 143
Nortel Alteon Controller 161
highavailChange 211
Hilfe
cbrcontrol 368
ccocontrol 440
dscontrol 368
nalcontrol 458
Hinzufügen
Cisco CSS Controller 434
Cluster 360
Nortel Alteon Controller 452
Port für einen Cluster 70, 386
Server für einen Port 71, 398, 426
Hohe Verfügbarkeit 5, 6, 60, 204
Cisco CSS Controller 245
dscontrol 369, 459
gegenseitig 61, 207, 359, 360, 371
Hinweise zu IPV6 85
Konfiguration 154, 175, 176
konfigurieren 205
Linux für S/390 211
ndcontrol 441
Nortel Alteon Controller 245
primärer Host 359, 360
Scripts 209
goActive 210
goIdle 210
goInOp 210
goStandby 210
highavailChange 211
Weiterleitungsmethode nat 210
Host
cbrcontrol 373
ccocontrol 446
dscontrol 373
nalcontrol 466
HP-UX
Befehl arp publish 70
installieren 35
Voraussetzungen 35
http-Advisor 351, 406

I

IBM Firewall (Einschränkung) 41
ibmlb.conf
für Solaris konfigurieren 67
ibmproxy 108, 114
ifconfig, Befehl 70, 73, 116, 232
Inaktivitätszeitlimit 270, 359, 362, 384
Informationen abrufen 283
Installation planen 3, 9, 51, 127
Installieren
Load Balancer 31
unter AIX 32
unter HP-UX 35
unter Linux 37
unter Solaris 39
unter Windows 40, 41

Intervall, Einstellung
Advisor fragt Server ab 354, 409
Manager aktualisiert Wertigkeiten für
den Executor 183, 378, 415, 417
Manager fragt Executor ab 183, 378
IPV6-Unterstützung 81
Advisor-Funktionen verwenden 84
Aliasnamen auf NIC festlegen 88
autoconf6 (AIX) 88
dsconfig 88
dscontrol-Befehle 93
hohe Verfügbarkeit 85
ifconfig 88
IP-Adresse 88
IPV6-Pakete aktivieren 88
Konfigurationshinweise 83
Metric Server 87
modprobe (Linux) 88
nicht unterstützte Features 84
Plattformunterstützung 82
Unterschiede in der Befehlssyntax 92
verbindungslokale Adresse 84
Verknüpfung 86

K

Kernel-unabhängige Unterstützung 82
Kollokation von Load Balancer und Server 395, 399
Konfiguration
Beispieldateien 481
cbrwizard 113
Cisco CSS Controller 149
Consultant starten 153, 175
Content Based Routing 109
Dienst 175
dswizard 66
erweiterte Tasks 179, 201
hohe Verfügbarkeit 154, 175, 176
Komponente Dispatcher 63
Messwerte 153, 175
Methoden
Assistent (CBR) 113
Assistent (Dispatcher) 66
Assistent (Site Selector) 134
Befehlszeile (CBR) 110
Befehlszeile (Cisco CSS Controller) 149
Befehlszeile (Dispatcher) 64
Befehlszeile (Nortel Alteon Controller) 171
Befehlszeile (Site Selector) 131
GUI (CBR) 112
GUI (Cisco CSS Controller) 151
GUI (Dispatcher) 65
GUI (Nortel Alteon Controller) 173
GUI (Site Selector) 133
Scripts (CBR) 111
Scripts (Cisco CSS Controller) 150
Scripts (Dispatcher) 64
Scripts (Nortel Alteon Controller) 172
Scripts (Site Selector) 132
Nortel Alteon Controller 171
prüfen 77
Site Selector 131

Konfiguration (*Forts.*)

sswizard 134
Switch-Consultant definieren 175
testen 154, 176
Kundenspezifische (anpassbare) Advisor-Funktion 193, 252
Beispiel 487

L

Lastausgleichseinstellungen (optimieren) 180, 248
lbkeys 198, 256, 264
lbwebaccess 266
Linux
hohe Verfügbarkeit auf S/390 211
installieren 37
Voraussetzungen 37
Load Balancer
Beispiel für einen schnellen Start 45
CBR 99
Cisco CSS Controller 139
Nortel Alteon Controller 157
Site Selector 123
Betrieb und Verwaltung 263, 280, 281
erweiterte Konfigurations-Tasks 179, 201
Fehlerbehebung 283
Funktionen 3, 9
installieren 31
IPV6-Unterstützung 81
konfigurieren
CBR 109
Cisco CSS Controller 149
Komponente Dispatcher 66, 113, 134
Nortel Alteon Controller 171
Site Selector 131
Überlegungen bei der Planung 51, 127
Übersicht 3, 9
Vorteile 5
Load Balancer für IPV4 und IPV6 81
Advisor-Funktionen verwenden 84
Aliasnamen auf der Loopback-Einheit festlegen 88
autoconf6 (AIX) 88
dsconfig 88
dscontrol-Befehle 93
hohe Verfügbarkeit 85
ifconfig 88
IP-Adresse 88
IPV6-Pakete aktivieren 88
Konfigurationshinweise 83
Metric Server 87
modprobe (Linux) 88
nicht unterstützte Features 84
Plattformunterstützung 82
Unterschiede in der Befehlssyntax 92
verbindungslokale Adresse 84
Verknüpfung 86
Loopback-Einheit
Aliasname 72
Alternativen für Festlegung eines Aliasnamens unter Linux 78
Löschen
Cluster 360, 429

Löschen (*Forts.*)

- Port von einem Cluster 386
- Server von einem Port 398, 425, 426
- zusätzliche Route 77

M

mac, Weiterleitungsmethode 53

Manager

- cbrcontrol 375
- dscontrol 375
- feste Wertigkeit 182
- Proportionen 180
- sscontrol 415
- starten 71, 379, 416, 418
- stoppen 380, 417, 419
- Version 380, 417, 419

Marken 497

Maximale Wertigkeit einstellen

- für Server an einem bestimmten Port 181, 386

Mehrere Adressen verknüpfen 71

Messwerte

- cbrcontrol 381
- cococontrol 444
- dscontrol 381
- Konfiguration 153, 175
- nalcontrol 462
- sscontrol 420

Metric Server

- Ausgabe des Befehls `ps -vg` unter AIX beschädigt 340
- Hinweise zu IPV6 87
- in einer zweistufigen Konfiguration konfigurieren 340
- IOException für Metric Server unter Windows 339
- kein Abruf der Werte auf Linux-Systemen möglich 342
- meldet keine Last 339
- Metric-Server-Protokoll meldet, dass für den Zugriff auf den Agenten eine Kennung erforderlich ist 340
- Rückgabe von -1 nach dem Start von Metric Server 343
- starten und stoppen 282
- Tabellen zur Fehlerbehebung 299
- Übersicht 196, 255
- unerwünschte Konsolnachrichten von Scripts unter Solaris 342
- verwenden 282

Migration 31

N

nalcontrol, Befehl

- consultant 452, 455
- Eingabeaufforderung 451
- file 457
- help 458
- host 466
- metric 462
- server 464

nalserver

- starten 158
- wird nicht gestartet 336

Namensserver

- sscontrol 421

nat, Serververknüpfung 204

nat, Weiterleitungsmethode 53, 57

- Scripts für hohe Verfügbarkeit 210

ndcontrol, Befehl

- highavailability 441

netstat, Befehl 76

Netzadressenkonvertierung (NAT) 53

Netzproximität 130

Neue Features, Version 6.1

- Kernel-unabhängige Unterstützung 7
- SIP-Advisor 7
- Unterstützung für Benutzerbereich 7
- Unterstützung für den Browser Firefox 7
- Unterstützung für HP-UX Itanium 64-Bit 7
- Unterstützung für Linux zSeries 64-Bit 7
- verknüpfte Client-Konfiguration 7

Neue Verbindungen, Bedeutung von Proportionen festlegen 180, 358

Neustart aller Server mit Standardwertigkeit 379, 416, 418

NFA

- definieren 69
- festlegen 365

NIC

- Aliasname 69
- Ethernet (für Solaris) 67
- zuordnen (für Windows) 70

Nortel Alteon Consultant

- erforderliche Funktionen bestimmen 28

Nortel Alteon Controller

- Advisor-Funktion Workload Manager 257
- Advisor-Funktionen 250
- Alerts 259
- Befehl `refresh` aktualisiert die Konfiguration nicht 338
- Befehle 451
- bei Webverwaltung wird Hostverbindung getrennt 338
- Beispiel für einen schnellen Start 157
- Bericht
- Controller 455
- beschädigte nationale Sonderzeichen (Latin-1) unter Windows 339
- binäre Protokollierung für Serverstatistik 258
- Consultant-Verbindungsfehler 338
- Einstellungen für Lastausgleich 248
- für Port 14099 kann kein Eintrag in der Registrierungsdatenbank erstellt werden 337
- Hardware- und Softwarevoraussetzungen 161
- hohe Verfügbarkeit 245
- Java-Fehler wegen Speicherkapazität/Threads (HP-UX) 339
- Konfiguration
- Maschine mit Nortel Alteon Controller konfigurieren 174
- Tasks im Überblick 171
- lbadmin scheitert 337

Nortel Alteon Controller (*Forts.*)

- Metric Server 255
- nalcontrol scheitert 337
- Planung 161
- starten und stoppen 281
- Tabellen zur Fehlerbehebung 298
- unerwartetes GUI-Verhalten bei Verwendung von Matrox-AGP-Karten 337
- verknüpfen 245
- verwenden 281
- Wertigkeiten werden auf dem Switch nicht aktualisiert 338
- wird nicht gestartet 336

O

OS/390

- GRE-Unterstützung 235

P

Passive Cookie-Affinität 225, 227, 391

Planung

- CBR 105
- Cisco CSS Controller 143
- Komponente Dispatcher 51
- Nortel Alteon Controller 161
- Site Selector 127

Platzhalter-Port 70, 386

- Advisor-Funktion `ping` 190
- für FTP-Datenverkehr verwenden 239
- für Übertragung von Datenverkehr mit nicht konfiguriertem Port 239

Platzhaltercluster 69, 360

- für den Lastausgleich von Firewalls 238
- mit Caching Proxy für transparente Weiterleitung 239
- zum Zusammenfassen von Serverkonfigurationen 237

Port

- cbrcontrol 382
- dscontrol 382
- Port-Affinität außer Kraft setzen Server 220, 395, 398

Port-übergreifende Affinität 223, 382

Port-Umsetzung für Netzadressen (NAPT) 53

Ports

- Anzeige
- Status von Servern an diesem Port 386
- entfernen 386
- für Advisor 351, 406
- für Cluster definieren 70, 386
- hinzufügen 386
- maximale Wertigkeit festlegen 181, 386
- Platzhalter 70
- Primärer Host 207, 360
- Privater Schlüssel
- für ferne Authentifizierung 264
- Privates Netz, Benutzung mit Dispatcher 236

- Produktkomponenten 51
- Proportionale Bedeutung für Lastausgleich festlegen 180, 360
- Protokoll
 - binär für Serverstatistik 241
 - CBR-Protokolle verwenden 280
 - Dateinamen festlegen
 - für Advisor-Funktion 408
 - für den Manager 417
 - Größe einstellen
 - für Advisor-Funktion 267, 355, 407, 409
 - für den Consultant 269
 - für den Manager 267, 378, 415, 417
 - für den Server 267, 269
 - für den Subagenten 267, 269
 - Stufe einstellen
 - für Advisor 267, 355, 409
 - für den Consultant 269
 - für den Manager 267, 415
 - für den Server 267, 269
 - für den Subagenten 267
 - von Cisco CSS Controller verwenden 281
 - von Load Balancer verwenden 267
 - von Metric Server verwenden 282
 - von Site Selector verwenden 280
- Protokollstatus
 - cbrcontrol 374
 - dscontrol 374
 - sscontrol 414
- Proximitätsoptionen 130

R

- Regel
 - cbrcontrol 388
 - dscontrol 388
 - sscontrol 422
- Regelbasierter Lastausgleich 212
 - aktive Verbindungen pro Port 215, 389
 - Auswertungsoption 220
 - Client-IP-Adresse 214, 388, 393, 422, 424
 - Client-Port 214, 389
 - Diensttyp (TOS) 215, 389, 393
 - gemeinsame genutzte Bandbreite 216, 217, 389, 393
 - immer gültig 219, 389, 393, 422, 424
 - Inhalt der Anforderung 55, 219, 389
 - metricall 422
 - metricavg 422
 - Metrik Durchschnitt 218
 - Metrik gesamt 218
 - Option für Serverauswertung 220
 - Regelauswahl für die einzelnen Komponenten 213
 - reservierte Bandbreite 216, 389, 393
 - Uhrzeit 214, 388, 393, 422, 424
 - Verbindungen pro Sekunde 215, 389
- RMI (Remote Method Invocation) 35, 39, 40, 41, 263, 264
- route, Befehl 76, 77
- Routen, zusätzliche 76

S

- Schlüssel
 - lbkeys 197, 255, 264
- Scripts 209
 - Benutzer-Exit 184, 259
 - ccoserverdown 260
 - goActive 210
 - goldle 210
 - goInOp 210
 - goStandby 210
 - highavailChange 211
- Secure Socket Layer 71
- Sensitivität für Aktualisierung der Wertigkeit einstellen 183, 379, 416, 418
- Sensitivitätsschwelle 250
- Server
 - Adresse 394
 - Advisor-Anforderung 397
 - Advisor-Antwort 397
 - alle mit Standardwertigkeit neu starten 379, 416, 418
 - als aktiv markieren 399, 425, 426
 - als inaktiv markieren 398, 425, 426
 - cbrcontrol 394
 - ccocontrol 449
 - Cookie-Wert 395
 - dscontrol 394
 - entfernen 398, 425, 426
 - feste Wertigkeit 395
 - für einen Port definieren 71, 398, 426
 - hinzufügen 398, 426
 - inaktiven Server zurücksetzen 182
 - logisch 58
 - mapport 108
 - mit nat verknüpft 204
 - nalcontrol 464
 - nonsticky (Außerkräftsetzen der Port-Affinität) 395, 398
 - Partitionierung 58
 - physisch 58
 - Protokoll 396
 - Router 396
 - Rückkehradresse 396
 - sscontrol 425
 - stillelegen 224, 376, 378, 380
 - verknüpft 395, 399
 - Wertigkeit 395
 - Wertigkeit festlegen 399, 425
 - wieder in Betrieb nehmen 380
 - zugeordneter Port 396
- Server als 'aktiv' markieren 399, 425, 426
- Server als 'inaktiv' markieren 398, 425, 426
- Server markieren als
 - aktiv 399, 425, 426
 - inaktiv 398, 425, 426
- Sicherung der hohen Verfügbarkeit 60, 369, 441, 459
- Simple Network Management Protocol (SNMP) 271
- Site Selector
 - Advisor-Funktionen und Erreichbarkeitsziele markieren alle Server als inaktiv (Windows) 333
 - Befehle 405

- Site Selector (*Forts.*)
 - bei Webverwaltung wird Hostverbindung getrennt 333
 - Beispiel für einen schnellen Start 123
 - beschädigte nationale Sonderzeichen (Latin-1) unter Windows 333
 - Einstellungen für Lastausgleich 180
 - Serverzeitlimit der Advisor-Funktion 188
 - Wiederholungsversuche der Advisor-Funktion für Server 188
 - erforderliche Funktionen bestimmen 26
 - Java-Fehler wegen Speicherkapazität/Threads (HP-UX) 333
 - kein korrekter Lastausgleich bei duplizierten Routen 332
 - Konfiguration
 - Maschine konfigurieren 134
 - Tasks im Überblick 131
 - Konfigurationsbeispiel 14
 - Lastausgleich für HA-Dispatcher 211
 - lbadm scheitert 331
 - Planung 127
 - sscontrol scheitert 331
 - ssserver wird unter Windows nicht gestartet 332
 - starten und stoppen 280
 - Tabellen zur Fehlerbehebung 295
 - Übersicht 13
 - unerwartetes GUI-Verhalten bei Verwendung von Matrox-AGP-Karten 333
 - verteilt Datenverkehr von Solaris-Clients nicht nach der RoundRobin-Methode 331
 - verwenden 280
 - wird nicht ausgeführt 331
- Sitename
 - sscontrol 428
- SNMP 267, 271
- Softwarevoraussetzungen
 - Cisco CSS Controller 143
 - Nortel Alteon Controller 161
- Solaris
 - Befehl arp publish 70
 - Dispatcher-Maschine konfigurieren 67
 - installieren 39
 - Voraussetzungen 39
- sscontrol, Befehl
 - advisor 406
 - file 411
 - help 413
 - logstatus 414
 - manager 415
 - metric 420
 - nameserver 421
 - rule 422
 - server 425
 - set 427
 - sitename 428
 - status 431
- SSL 71
- SSL-Verbindungen
 - Fehler beim Aktivieren 306
 - für CBR 107, 108

SSL-Verbindungen (*Forts.*)
 HTTPS-Advisor-Funktion 189
 ibmproxy konfigurieren 108
 SSL-Advisor-Funktion 189
 ssl2http, Advisor-Funktion 108, 190
 ssserver
 starten 124
 Starten
 Advisor-Funktion 71, 354, 407, 409
 CBR 100
 Cisco CSS Controller 140, 280
 Dispatcher 47
 Executor 68, 365
 Manager 71, 379, 416, 418
 Metric Server 282
 Nortel Alteon Controller 158, 281
 Server 68, 69
 Site Selector 124, 280
 Starten und stoppen
 CBR 279
 Dispatcher 270
 Statistische Momentaufnahme anzeigen 378, 416, 417
 Status
 cbrcontrol 401
 dscontrol 401
 Statusanzeige
 Server für einen bestimmten Port 386
 Stilllegen eines Servers 224, 376, 378, 380
 Stoppen
 Advisor-Funktion 354, 408, 410
 Cisco CSS Controller 280
 Executor 365
 Manager 380, 417, 419
 Nortel Alteon Controller 281
 Subagenten 267, 271
 dscontrol 402
 Switch-Consultant
 definieren 175
 Syntaxdiagramme
 Beispiele 347
 Interpunktion 347
 lesen 347
 Parameter 347
 Symbole 347
 Systemmesswerte
 konfigurieren 381, 420, 444, 462
 proportionale Bedeutung festlegen 181, 248, 357, 358

T

Testen
 Konfiguration 154, 176

U

Überprüfen
 zusätzliche Route 76
 Übersicht
 Konfiguration der Komponente Dispatcher 63
 Konfiguration von CBR 109

Übersicht (*Forts.*)
 Konfiguration von Cisco CSS Controller 149
 Konfiguration von Nortel Alteon Controller 171
 Konfiguration von Site Selector 131
 Überwachen, Menüoption 271
 URI-Affinität 225, 228, 391

V

Verbindungen, Bedeutung von Proportionen festlegen 181, 360
 Verbindungszeitlimit
 Site Selector 406
 Verknüpfung
 Cisco CSS Controller 245
 Hinweise zu IPV6 86
 Nortel Alteon Controller 245
 Verknüpfung mit nat 204
 Verknüpfung von Load Balancer und Client 243
 Verknüpfung von Load Balancer und Server 66, 71, 202, 232
 Version anzeigen
 Advisor-Funktion 355, 409, 410
 Manager 380, 417, 419
 Verwaltung von Load Balancer 263
 Verwendung von Load Balancer 263
 Voraussetzungen
 AIX 31
 HP-UX 35
 Linux 37
 Solaris 39
 Windows 40

W

WAN-Unterstützung 229
 Konfigurationsbeispiel 233
 Linux 235
 mit fernem Dispatcher 230
 mit fernem Advisor-Funktionen 231
 mit GRE 235
 WAS (WebSphere Application Server)
 Advisor-Funktion 191, 194
 WAS-Advisor-Funktion 191, 194
 Webgestützte Verwaltung 263, 265
 aktualisieren 266
 Weiterleitungsmethode
 cbr 55, 57
 mac 53, 54
 mac, nat oder cbr 56, 384
 nat 53, 57
 Wertigkeit
 Controller 249
 festlegen
 für einen Server 399, 425
 Grenzwert für alle Server an einem Port 181, 386
 Festlegung durch den Manager 182
 Windows
 Befehl executor configure 70
 Dispatcher-Maschine konfigurieren 68
 installieren 40

Windows (*Forts.*)
 Voraussetzungen 40
 Workload Manager (WLM), Advisor-Funktion 199, 257

Z

Zugriffsmöglichkeit xvii
 Zuordnungsdatei adressieren
 Beispiel 237
 Zusätzliche Routen 76, 77



GC12-3680-00

