

WebSphere Application Server



# Concepts, Planning, and Installation for Edge Components

*Version 8.5*



WebSphere Application Server



# Concepts, Planning, and Installation for Edge Components

*Version 8.5*

**Note**

Before using this information and the product it supports, be sure to read the general information under “Notices” on page 89.

**First edition (May 2012)**

This edition applies to:

WebSphere Edge Components, Version 8.5

and to all subsequent releases and modifications until otherwise indicated in new editions.

Order publications through your IBM representative or through the IBM branch office serving your locality.

© **Copyright IBM Corporation 2011.**

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

---

# Contents

<b>Figures</b> . . . . .	<b>v</b>
--------------------------	----------

<b>About this book</b> . . . . .	<b>vii</b>
Who should read this book . . . . .	vii
Accessibility . . . . .	vii
Conventions and terminology used in this book . . . . .	vii

---

## Part 1. Overview . . . . . 1

### Chapter 1. Introducing WebSphere Application Server Edge components . . . 3

Caching Proxy . . . . .	3
Load Balancer . . . . .	4
Dispatcher . . . . .	5
Content Based Routing . . . . .	5
Site Selector . . . . .	5
Cisco CSS Controller . . . . .	6
Nortel Alteon Controller . . . . .	6
Metric Server . . . . .	6

### Chapter 2. Edge Components and the WebSphere family . . . . . 7

Tivoli Access Manager . . . . .	7
WebSphere Portal Server . . . . .	7
WebSphere Site Analyzer . . . . .	7
WebSphere Transcoding Publisher . . . . .	8

### Chapter 3. More information on Application Server and Edge components . . . . . 9

---

## Part 2. Edge component concepts and discussions . . . . . 11

### Chapter 4. Caching . . . . . 13

Basic Caching Proxy configurations . . . . .	13
Reverse Caching Proxy (default configuration) . . . . .	13
Forward Caching Proxy . . . . .	14
Advanced caching . . . . .	17
Load-balanced Caching Proxy clusters . . . . .	17
Caching dynamic content . . . . .	18
Additional caching features . . . . .	18

### Chapter 5. Network performance . . . . . 21

Network hardware . . . . .	21
Memory considerations . . . . .	21
Hard disk considerations . . . . .	21
Network considerations . . . . .	22
CPU considerations . . . . .	22
Network architecture . . . . .	22
Web site popularity and proxy server load considerations . . . . .	22

Traffic type considerations . . . . .	23
---------------------------------------	----

### Chapter 6. Availability . . . . . 25

Load balancing . . . . .	25
Load balancing multiple content hosts . . . . .	25
Load balancing multiple reverse proxy servers . . . . .	26
Load Balancer with multiple forward proxy servers . . . . .	27
Failover support . . . . .	30

### Chapter 7. Content Based Routing . . . . . 33

---

## Part 3. Scenarios . . . . . 37

### Chapter 8. Business-to-consumer network . . . . . 39

Phase 1 . . . . .	39
Phase 2 . . . . .	40
Phase 3 . . . . .	41

### Chapter 9. Business to client-banking-solution . . . . . 43

### Chapter 10. Web portal network . . . . . 45

---

## Part 4. Installing Edge Components 47

### Chapter 11. Requirements for Edge Components . . . . . 49

### Chapter 12. Installing Edge Components with IBM Installation Manager . . . . . 51

Installing IBM Installation Manager . . . . .	55
Installing Edge Components with the IBM Installation Manager GUI . . . . .	58
Installing Edge Components silently with IBM Installation Manager . . . . .	61

### Chapter 13. Installing fix packs to Edge Components . . . . . 67

Installing fix packs using the Installation Manager GUI . . . . .	67
Installing fix packs with the command line . . . . .	67

### Chapter 14. Uninstalling fix packs to Edge Components with IBM Installation Manager . . . . . 71

**Chapter 15. Uninstalling Edge Components . . . . . 73**

Uninstalling Edge Components with the IBM  
Installation Manager GUI. . . . . 73  
Uninstalling Edge Components silently with IBM  
Installation Manager . . . . . 73  
Uninstalling previous versions of Edge Components 75

---

**Part 5. Building networks with Edge Components . . . . . 79**

**Chapter 16. Build a Caching Proxy network . . . . . 81**

Workflow . . . . . 81  
Review required computer systems and software. . 81  
Build Server 1 (Linux and UNIX systems) . . . . 82  
Build Server 1 (Windows system) . . . . . 82

Configure Server 1 . . . . . 82  
Test the Caching Proxy network . . . . . 83

**Chapter 17. Build a Load Balancer network . . . . . 85**

Workflow . . . . . 85  
Review required computer systems and software. . 85  
Configure the network. . . . . 86  
Configure the Dispatcher. . . . . 86  
    Configuring using the command line. . . . . 87  
    Configuring using the configuration wizard . . 87  
    Configuring using the graphical user interface  
    (GUI) . . . . . 88  
Test the Load Balancer network . . . . . 88

**Notices . . . . . 89**  
Trademarks . . . . . 90

---

## Figures

1. Caching Proxy acting as a reverse proxy	14	10. Using a primary and backup Load Balancer to make Web content highly available.	31
2. Caching Proxy acting as a forward proxy	15	11. Locating the backup Load Balancer on a content host	32
3. The Caching Proxy acting as a transparent forward proxy	16	12. Routing HTTP requests with CBR	34
4. Caching Proxy acting as proxy server for a load-balanced cluster	17	13. Load balancing HTTP requests routed with CBR	35
5. Load balancing multiple content hosts	26	14. Business to consumer network (Phase 1)	40
6. Load balancing multiple reverse proxy servers and content hosts	27	15. Business to consumer network (Phase 2)	41
7. Using Dispatcher to load balance multiple caching proxies.	28	16. Business to consumer network (Phase 3)	42
8. Using a primary and backup Dispatcher to provide highly available Internet access.	29	17. Business to consumer banking solution	44
9. Locating the backup Dispatcher on a Caching Proxy machine.	30	18. Web portal.	46
		19. Caching Proxy demonstration network	81
		20. Load Balancer demonstration network	85





---

## About this book

This book, *WebSphere® Application Server Concepts, Planning, and Installation for Edge Components*, serves as an introduction to the WebSphere Application Server Edge Components. It provides high-level product overviews, detailed functionality discussions for key components, edge-of-the-network scenarios, installation and initial configuration information, and demonstration networks.

---

## Who should read this book

*WebSphere Application Server Concepts, Planning, and Installation for Edge Components* is written for experienced network and system administrators who are familiar with their operating systems and with providing Internet services. Prior exposure to the WebSphere Application Server or to WebSphere Application Server Edge Components is not required.

---

## Accessibility

Accessibility features help a user who has a physical disability, such as restricted mobility or limited vision, to use software products successfully. These are the major accessibility features in WebSphere Application Server, Version 8.5:

- You can use screen-reader software and a digital speech synthesizer to hear what is displayed on the screen. You can also use voice recognition software, such as IBM® ViaVoice®, to enter data and to navigate the user interface.
- You can operate features by using the keyboard instead of the mouse.
- You can configure and administer Application Server features by using standard text editors or command-line interfaces instead of the graphical interfaces provided. For more information about the accessibility of particular features, refer to the documentation about those features.

---

## Conventions and terminology used in this book

This documentation uses the following typographical and keying conventions.

Table 1. Conventions used in this book

Convention	Meaning
<b>Bold</b>	When referring to graphical user interfaces (GUIs), bold face indicates menus, menu items, labels, buttons, icons, and folders. It also can be used to emphasize command names that otherwise might be confused with the surrounding text.
Monospace	Indicates text you must enter at a command prompt. Monospace also indicates screen text, code examples, and file excerpts.
<i>Italics</i>	Indicates variable values that you must provide (for example, you supply the name of a file for <i>fileName</i> ). Italics also indicates emphasis and the titles of books.
Ctrl- <i>x</i>	Where <i>x</i> is the name of a key, indicates a control-character sequence. For example, Ctrl-c means hold down the Ctrl key while you press the c key.
Return	Refers to the key labeled with the word Return, the word Enter, or the left arrow.
%	Represents the Linux and UNIX command-shell prompt for a command that does not require root privileges.
#	Represents the Linux and UNIX command-shell prompt for a command that requires root privileges.

Table 1. Conventions used in this book (continued)

<b>Convention</b>	<b>Meaning</b>
C:\	Represents the Windows command prompt.
Entering commands	When instructed to “enter” or “issue” a command, type the command and then press Return. For example, the instruction “Enter the <b>ls</b> command” means type <b>ls</b> at a command prompt and then press Return.
[ ]	Enclose optional items in syntax descriptions.
{ }	Enclose lists from which you must choose an item in syntax descriptions.
	Separates items in a list of choices enclosed in { }(braces) in syntax descriptions.
...	Ellipses in syntax descriptions indicate that you can repeat the preceding item one or more times. Ellipses in examples indicate that information was omitted from the example for the sake of brevity.

---

## Part 1. Overview

This part introduces the WebSphere Application Server Edge Components, Caching Proxy and Load Balancer, and discusses their integration with Application Server. It also defines the components of Caching Proxy and Load Balancer. In addition, this section introduces other related WebSphere family products.

This part contains the following chapters:

- Chapter 1, “Introducing WebSphere Application Server Edge components,” on page 3
- Chapter 2, “Edge Components and the WebSphere family,” on page 7
- Chapter 3, “More information on Application Server and Edge components,” on page 9



---

## Chapter 1. Introducing WebSphere Application Server Edge components

WebSphere is Internet infrastructure software that enables companies to develop, deploy, and integrate next-generation e-business applications such as those for business-to-business e-commerce. WebSphere middleware supports business applications from simple Web publishing through enterprise-scale transaction processing.

As the foundation of the WebSphere platform, WebSphere Application Server offers a comprehensive set of middleware that enables users to design, implement, deploy, and manage business applications. These applications could be a simple Web site storefront to a more complex tasks, like a complete revision of an organization's computing infrastructure.

Processor-intensive features, such as personalization, offer a competitive advantage to every e-business. However, habitually relegating these features to central servers can prevent valuable functions from scaling to Internet proportions. Consequently, with the constant addition of new Web applications, a business's Internet infrastructure must grow in scope and impact. In addition, reliability and security are extraordinarily important to an e-business. Even a minimal service disruption can result in a loss of business.

Edge Components (formerly Edge Server) are now a part of the WebSphere Application Server offering. Edge Components can be used in conjunction with WebSphere Application Server to control client access to Web servers and to enable business enterprises to provide better service to users who access Web-based content over the Internet or a corporate intranet. Using Edge Components can reduce Web server congestion, increase content availability, and improve Web server performance. As the name indicates, Edge Components usually run on machines that are close (in a network configuration sense) to the boundary between an enterprise's intranet and the Internet.

The WebSphere Application Server includes the Caching Proxy and Load Balancer Edge components.

---

### Caching Proxy

Caching Proxy reduces bandwidth use and improves a Web site's speed and reliability by providing a point-of-presence node for one or more back-end content servers. Caching Proxy can cache and serve static content and content dynamically generated by WebSphere Application Server.

Caching Proxy can be configured in the role of a reverse proxy server (default configuration) or a forward proxy server, providing either a point-of-presence for a network or an internal network server that is responsible for improving request and response time. For more information on reverse and forward configurations, see "Basic Caching Proxy configurations" on page 13.

The proxy server intercepts data requests from a client, retrieves the requested information from content-hosting machines, and delivers that content back to the client. Most commonly, the requests are for documents stored on Web server

machines (also called *origin servers* or *content hosts*) and delivered using the Hypertext Transfer Protocol (HTTP). However, you can configure the proxy server to handle other protocols, such as File Transfer Protocol (FTP) and Gopher.

The proxy server stores cacheable content in a local cache before delivering it to the requester. Examples of cacheable content include static Web pages and JavaServer Pages files that contain dynamically generated, but infrequently changing, information. Caching enables the proxy server to satisfy subsequent requests for the same content by delivering it directly from the local cache, which is much quicker than retrieving it again from the content host.

Plug-ins for Caching Proxy add functionality to the proxy server.

- The ICP plug-in enables the proxy server to query Internet Caching Protocol (ICP)-compliant caches in search of HTML pages and other cacheable resources.
- The Tivoli<sup>®</sup> Access Manager (formerly Policy Director) plug-in enables the proxy server to use Access Manager's integrated authorization or authentication services.
- The PAC-LDAP Authentication Module enables the proxy server to access an LDAP server when performing authorization or authentication routines.
- The WebSphere Transcoding Publisher plug-in allows the proxy server to cache multiple transcoded versions of content for mobile devices when used in conjunction with WebSphere Transcoding Publisher.

You can further extend the functions of Caching Proxy by writing custom plug-in modules to an application programming interface (API). The API is flexible, easy to use, and platform independent. The proxy performs a sequence of steps for each client request it processes. A plug-in application modifies or replaces a step within the request-processing workflow, such as client authentication or request filtering. The powerful Transmogriify interface, for example, provides access to HTTP data and enables substitution or transformation of URLs and Web content. Plug-ins can modify or replace designated processing steps, and you can invoke more than one plug-in for a particular step.

---

## Load Balancer

Load Balancer creates edge-of-network systems that direct network traffic flow, reducing congestion and balancing the load on various other services and systems. Load Balancer provides site selection, workload management, session affinity, and transparent failover.

Load Balancer is installed between the Internet and the enterprise's back-end servers, which can be content hosts or Caching Proxy machines. Load Balancer acts as the enterprise's single point-of-presence node on the Internet, even if the enterprise uses multiple back-end servers because of high demand or a large amount of content. You can also guarantee high availability by installing a backup Load Balancer to take over if the primary one fails temporarily.

Load Balancer intercepts data requests from clients and forwards each request to the server that is currently best able to fill the request. In other words, it balances the load of incoming requests among a defined set of machines that service the same type of requests. Load Balancer can distribute requests to many types of servers, including WebSphere Application Servers and Caching Proxy machines. Load balancing can be customized for a particular application or platform by using custom advisors. Special purpose advisors are available to obtain information for load balancing WebSphere Application Servers.

If the Content Based Routing component is installed together with the Caching Proxy, HTTP and HTTPS requests can even be distributed based on URLs or other administrator-determined characteristics, eliminating the need to store identical content on all back-end servers. The Dispatcher component can also provide the same function for HTTP requests.

Load balancing improves your Web site's availability and scalability by transparently clustering content servers, including HTTP servers, application servers, and proxy servers, which are surrogate content servers. Availability is achieved through parallelism, load balancing, and failover support. When a server is down, business is not interrupted. An infrastructure's scalability is greatly improved because back-end processing power can be added transparently.

Load Balancer includes the following components:

## Dispatcher

For all Internet services, such as HTTP, FTP, HTTPS, and Telnet, the Dispatcher component performs load balancing for servers within a local area network (LAN) or wide area network (WAN). For HTTP services, Dispatcher can perform load balancing of servers based on the URL content of the client request.

The Dispatcher component enables stable, efficient management of a large, scalable network of servers. With Dispatcher, you can link many individual servers into what appears to be a single virtual server. Your site thus appears as a single IP address to the world.

## Content Based Routing

For HTTP and HTTPS services, the Content Based Routing component performs load balancing for servers based on the content of the client request. The Content Based Routing component works in conjunction with the Application Server Caching Proxy component.

**IMPORTANT:** The Content Based Routing (CBR) component is available on all supported platforms, with the following exceptions:

- CBR is not available with Load Balancer installations on platforms running a 64-bit JVM.

Alternatively, for this type of installation, you can use the cbr forwarding method of Load Balancer's Dispatcher component to provide content-based routing of HTTP and HTTPS requests without the use of Caching Proxy. See the *WebSphere Application Server Load Balancer Administration Guide* for more information.

- CBR is not available with Load Balancer for IPv4 and IPv6 installations.

Load Balancer for IPv4 and IPv6 supports only the Dispatcher component's mac forwarding method. The nat and cbr forwarding methods are not supported.

## Site Selector

The Site Selector component enhances a load-balancing system by allowing it to act as the point-of-presence node for a network and load balance incoming requests by mapping DNS names to IP addresses. In conjunction with Metric Server, Site Selector can monitor the level of activity on a server, detect when a server has the least amount of network traffic, and detect a failed server.

This component is supported on all Edge component installations, with the following exception:

- This component is not available with Edge component installations of Load Balancer for IPv4 and IPv6.

## Cisco CSS Controller

**Note:** The Cisco CSS Controller component is shipped with Load Balancer for IPv4, but this component might not support newer hardware. Consult the prerequisites page for supported hardware: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

The Cisco CSS Controller component generates server-weighting metrics that are sent to a Cisco CSS switch for server selection, load optimization, and fault tolerance.

This component is supported on all Edge component installations, with the following exception:

- This component is not available with Edge component installations of Load Balancer for IPv4 and IPv6.

## Nortel Alteon Controller

**Note:** The Nortel Alteon Controller component is shipped with Load Balancer for IPv4, but this component might not support newer hardware. Consult the prerequisites page for supported hardware: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>

The Nortel Alteon Controller component generates server-weighting metrics that are sent to a Nortel Alteon switch for server selection, load optimization, and fault tolerance.

This component is supported on all Edge component installations, with the following exception:

- This component is not available with Edge component installations of Load Balancer for IPv4 and IPv6.

## Metric Server

The Metric Server component runs as a daemon on a load-balanced server and provides information about system loads to Load Balancer components.



---

## Chapter 2. Edge Components and the WebSphere family

The IBM WebSphere family is designed to help users realize the promise of e-business. It is a set of software products that helps users develop and manage high-performance Web sites and integrate Web sites with new or existing non-Web business information systems.

The WebSphere family consists of WebSphere Application Server, including the Edge Components, and other WebSphere family software that is tightly integrated with the WebSphere Application Server and enhances its performance. For an overview of WebSphere Application Server and its components, see Chapter 1, "Introducing WebSphere Application Server Edge components," on page 3.

---

### Tivoli Access Manager

Tivoli Access Manager (formerly Tivoli Policy Director) is available separately. It provides access control and centralized security for existing Web applications and offers one-time authentication capability with access to multiple Web resources. A Caching Proxy plug-in exploits Access Manager's security framework, enabling the proxy server to use Access Manager's integrated authorization or authentication services.

---

### WebSphere Portal Server

WebSphere Portal Server (available separately) offers a framework to meet the presentation, security, scalability, and availability issues associated with portals. Using Portal Server, companies can build their own custom portal Web site to serve the needs of employees, business partners, and customers. Users can sign on to the portal and receive personalized Web pages that provide access to the information, people, and applications they need. This personalized single point of access to all necessary resources reduces information overload, accelerates productivity, and increases Web site usage.

WebSphere Portal Server runs in a WebSphere Application Server cluster to achieve scalability and reliability. The Application Server Load Balancer component can also be used for additional load balancing and high availability.

---

### WebSphere Site Analyzer

WebSphere Site Analyzer (available separately) helps enterprises to anticipate capacity and performance problems. With Site Analyzer, Caching Proxy and Load Balancer logs and other manageability aids can be used to anticipate the demand for additional resources by monitoring, analyzing, and reporting your Web site usage. In addition, Site Analyzer manageability components assist users who install and upgrade Edge Components, manage and store configurations, operate Edge Components remotely, and view and report events.

---

## WebSphere Transcoding Publisher

WebSphere Transcoding Publisher (available separately) can convert a Web page for viewing on a mobile device, such as an Internet-capable phone, translate Web content to the user's preferred national language (by invoking WebSphere Translation Server), and convert markup languages. Transcoding Publisher enhances Caching Proxy's capabilities by allowing it to serve content for different devices and users. After accessing content from a Web server, Caching Proxy's Transmogrify interface can be configured to invoke Transcoding Publisher to transform the data and tag it for variant caching and possible reuse. At Caching Proxy's post-authentication interface, Transcoding Publisher then checks the proxy server for content matching the user and device requirements and, if a match is found, serves the content from the proxy server's cache.

---

## Chapter 3. More information on Application Server and Edge components

The following documentation specific to the WebSphere Application Server Edge Components is available in the Edge Components Information Center.

- *Programming Guide for Edge Components* GC31-6919-00
- *Caching Proxy Administration Guide* GC31-6920-00
- *Load Balancer Administration Guide* GC31-6921-00

Other WebSphere Application Server documentation is available from the WebSphere Application Server library page.

Technote support information on Edge Components is available from the WebSphere Application Server support page.

The following is a list of Web sites to obtain information on Edge Components or related information:

- IBM Web site home <http://www.ibm.com/>
- IBM WebSphere Application Server <http://www.ibm.com/software/webservers/appserv/>
- IBM WebSphere Application Server library Web site <http://www.ibm.com/software/webservers/appserv/was/library/>
- IBM WebSphere Application Server support Web site <http://www.ibm.com/software/webservers/appserv/support.html>



---

## Part 2. Edge component concepts and discussions

This part includes detailed discussions that highlight some of the functionality available with Edge components. See Chapter 1, “Introducing WebSphere Application Server Edge components,” on page 3 for an overview of the Application Server’s Caching Proxy component.

This part contains the following chapters:

Chapter 4, “Caching,” on page 13

Chapter 5, “Network performance,” on page 21

Chapter 6, “Availability,” on page 25

Chapter 7, “Content Based Routing,” on page 33



---

## Chapter 4. Caching

Caching Proxy's caching functionality helps to minimize network bandwidth utilization and ensure that end users receive faster, more reliable service. This is accomplished because the caching performed by the proxy server offloads back-end servers and peering links. Caching Proxy can cache static content and content dynamically generated by WebSphere Application Server. To provide enhanced caching, Caching Proxy also functions in conjunction with the Application Server Load Balancer component. See Chapter 1, "Introducing WebSphere Application Server Edge components," on page 3 for an introduction to these systems.

**IMPORTANT:** Caching Proxy is available on all Edge component installations, with the following exceptions:

- Caching Proxy is not available for Edge component installations that run on Itanium 2 or AMD Opteron 64-bit processors.
- Caching Proxy is not available for Edge component installations of Load Balancer for IPv4 and IPv6.

---

### Basic Caching Proxy configurations

Caching Proxy can be configured in the role of a reverse caching proxy server (default configuration) or a forward caching proxy server. When used by content hosts, the Caching Proxy is configured in the role of reverse caching proxy server, located between the Internet and the enterprises's content hosts. When used by Internet access providers, the Caching Proxy is configured in the role of a forward caching proxy server, located between a client and the Internet.

#### Reverse Caching Proxy (default configuration)

When using a reverse proxy configuration, Caching Proxy machines are located between the Internet and the enterprise's content hosts. Acting as a surrogate, the proxy server intercepts user requests arriving from the Internet, forwards them to the appropriate content host, caches the returned data, and delivers that data to the users across the Internet. Caching enables Caching Proxy to satisfy subsequent requests for the same content directly from the cache, which is much quicker than retrieving it again from the content host. Information can be cached depending on when it will expire, how large the cache should be and when the information should be updated. Faster download times for cache hits mean better quality of service for customers. Figure 1 on page 14 depicts this basic Caching Proxy functionality.

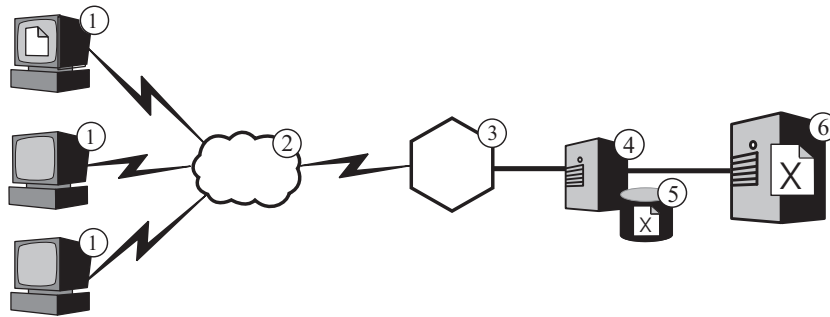


Figure 1. Caching Proxy acting as a reverse proxy. Legend:  
 1—Client 2—Internet 3—Router/Gateway 4—Caching Proxy  
 5—Cache 6—Content host

In this configuration, the proxy server (4) intercepts requests whose URLs include the content host's host name (6). When a client (1) requests file X, the request crosses the Internet (2) and enters the enterprise's internal network through its Internet gateway (3). The proxy server intercepts the request, generates a new request with its own IP address as the originating address, and sends the new request to the content host (6).

The content host returns file X to the proxy server rather than directly to the end user. If the file is cacheable, Caching Proxy stores a copy in its cache (5) before passing it to the end user. The most prominent example of cacheable content is static Web pages; however, Caching Proxy also provides the ability to cache and serve content dynamically generated by WebSphere Application Server.

## Forward Caching Proxy

Providing direct Internet access to end users can be very inefficient. Every user who fetches a given file from a Web server generates the same amount of traffic in your network and through your Internet gateway as the first user who fetched the file, even if the file has not changed. The solution is to install a forward Caching Proxy near the gateway.

When using a forward proxy configuration, Caching Proxy machines are located between the client and the Internet. Caching Proxy forwards a client's request to content hosts located across the Internet, caches the retrieved data, and delivers the retrieved data to the client.



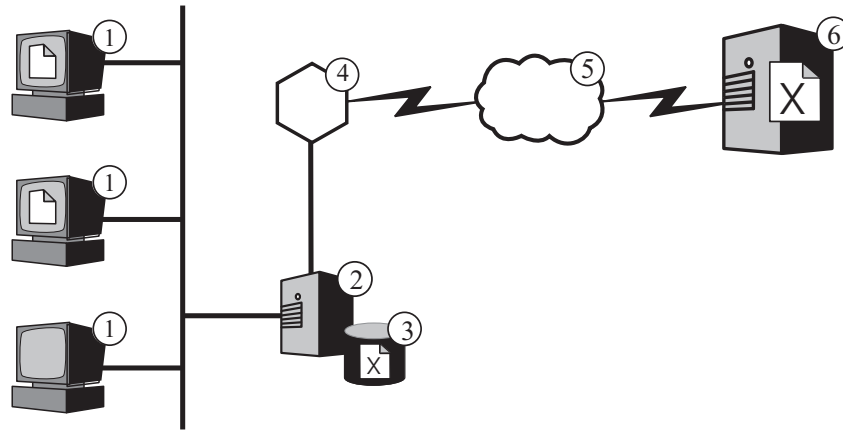


Figure 2. Caching Proxy acting as a forward proxy. Legend: 1—Client 2—Caching Proxy 3—Cache 4—Router/Gateway 5—Internet 6—Content host

Figure 2 depicts the forward Caching Proxy configuration. The clients' browser programs (on the machines marked 1) are configured to direct requests to the forward caching proxy (2), which is configured to intercept the requests. When an end user requests file X stored on the content host (6), the forward caching proxy intercepts the request, generates a new request with its own IP address as the originating address, and sends the new request out by means of the enterprise's router (4) across the Internet (5).

In this way the origin server returns file X to the forward caching proxy rather than directly to the end user. If the caching feature of the forward Caching Proxy is enabled, Caching Proxy determines whether file X is eligible for caching by checking settings in its return header, such as the expiration date and an indication whether the file was dynamically generated. If the file is cacheable, the Caching Proxy stores a copy in its cache (3) before passing it to the end user. By default, caching is enabled and the forward Caching Proxy uses a memory cache; however, you can configure other types of caching.

For the first request for file X, forward Caching Proxy does not improve the efficiency of access to the Internet very much. Indeed, the response time for the first user who accesses file X is probably slower than without the forward caching proxy, because it takes a bit more time for the forward Caching Proxy to process the original request packet and examine file X's header for cacheability information when it is received. Using the forward caching proxy yields benefits when other users subsequently request file X. The forward Caching Proxy checks that its cached copy of file X is still valid (has not expired), and if so it serves file X directly from the cache, without forwarding the request across the Internet to the content host.

Even when the forward Caching Proxy discovers that a requested file is expired, it does not necessarily have to refetch the file from the content host. Instead, it sends a special status checking message to the content host. If the content host indicates that the file has not changed, the forward caching proxy can still deliver the cached version to the requesting user.

Configuring the forward Caching Proxy in this way is termed forward proxy, because the Caching Proxy is acting on behalf of browsers, forwarding their requests to content hosts via the Internet. The benefits of forward proxy with caching are two-fold:

- If a file is cached, end users receive it much more quickly than when their requests must cross the Internet, because the forward caching proxy is on the local network. As more and more files are cached, the total response time that users experience for Internet requests continues to go down.
- There is no traffic generated outside the enterprise's local network. This effectively increases the capacity (available bandwidth) of the enterprise's gateway to the Internet by freeing it to handle requests for files that are not cached. It also reduces Internet access charges, which is especially important in environments where such charges are based on the number of packets.

Caching Proxy can proxy several network transfer protocols, including HTTP (Hypertext Transfer Protocol, FTP (File Transfer Protocol), and Gopher.

### Transparent forward Caching Proxy (Linux systems only)

A variation of the forward Caching Proxy is a transparent Caching Proxy. In this role, Caching Proxy performs the same function as a basic forward Caching Proxy, but it does so without the client being aware of its presence. The transparent Caching Proxy configuration is supported on Linux systems only.

In the configuration described in "Forward Caching Proxy" on page 14, each client browser is separately configured to direct requests to a certain forward Caching Proxy. Maintaining such a configuration can become inconvenient, especially for large numbers of client machines. The Caching Proxy supports several alternatives that simplify administration. One possibility is to configure the Caching Proxy for transparent proxy as depicted in Figure 3. As with regular forward Caching Proxy, the transparent Caching Proxy is installed on a machine near the gateway, but client browser programs are not configured to direct requests to a forward Caching Proxy. Clients are not aware that a proxy exists in the configuration. Instead, a router is configured to intercept client requests and direct them to the transparent Caching Proxy. When a client working on one of the machines, marked 1, requests file X stored on a content host (6), the router (2) passes the request to the Caching Proxy. Caching Proxy generates a new request with its own IP address as the originating address and sends the new request out by means of the router (2) across the Internet (5). When file X arrives, the Caching Proxy caches the file if appropriate (subject to the conditions described in "Forward Caching Proxy" on page 14) and passes the file to the requesting client.

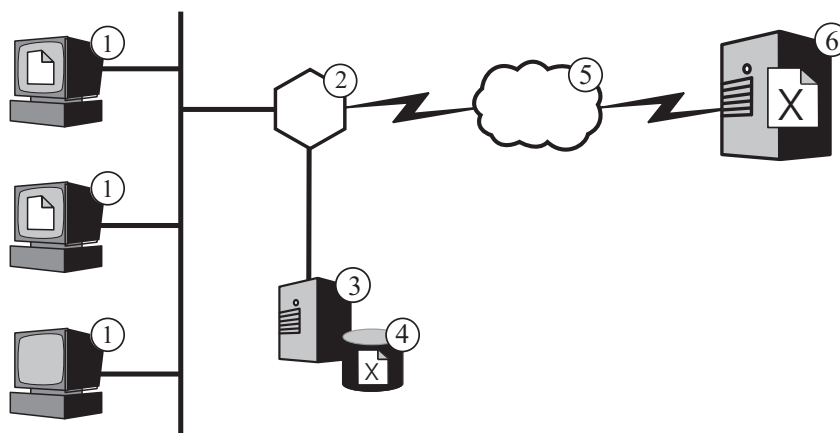


Figure 3. The Caching Proxy acting as a transparent forward proxy. Legend:  
 1—Client 2—Router/Gateway 3—Caching Proxy 4—Cache 5—Internet 6—Content host

For HTTP requests, another possible alternative to maintaining proxy configuration information on each browser is to use the automatic proxy configuration feature available in several browser programs, including Netscape Navigator version 2.0 and higher and Microsoft Internet Explorer version 4.0 and higher. In this case, you create one or more central proxy automatic configuration (PAC) files and configure browsers to refer to one of them rather than to local proxy configuration information. The browser automatically notices changes to the PAC and adjusts its proxy usage accordingly. This not only eliminates the need to maintain separate configuration information on each browser, but also makes it easy to reroute requests when a proxy server becomes unavailable.

A third alternative is to use the Web Proxy Auto Discovery (WPAD) mechanism available in some browser programs, such as Internet Explorer version 5.0 and higher. When you enable this feature on the browser, it automatically locates a WPAD-compliant proxy server in its network and directs its Web requests there. You do not need to maintain central proxy configuration files in this case. Caching Proxy is WPAD-compliant.

---

## Advanced caching

### Load-balanced Caching Proxy clusters

To provide more advanced caching functionality, use Caching Proxy as a reverse proxy in conjunction with the Load Balancer component. By integrating caching and load-balancing capabilities, you can create an efficient, highly manageable Web performance infrastructure.

Figure 4 depicts how you can combine Caching Proxy with Load Balancer to deliver Web content efficiently even in circumstances of high demand. In this configuration, the proxy server (4) is configured to intercept requests whose URLs include the host name for a cluster of content hosts (7) being load-balanced by Load Balancer (6).

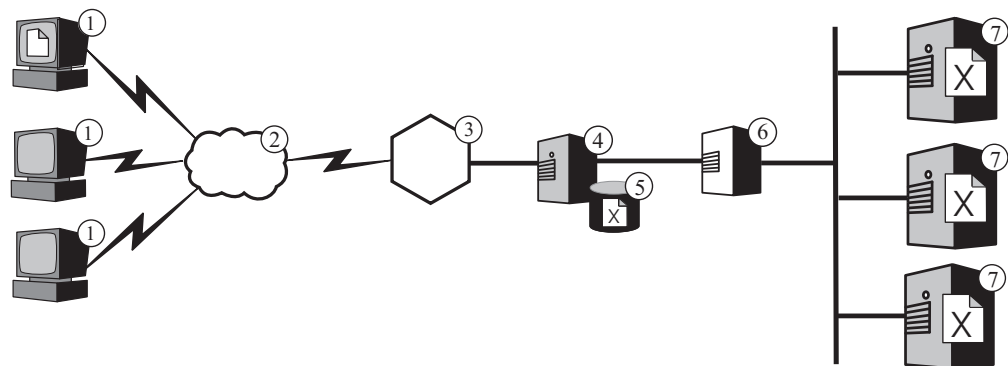


Figure 4. Caching Proxy acting as proxy server for a load-balanced cluster. Legend:  
1—Client 2—Internet 3—Router/Gateway 4—Caching Proxy 5—Cache 6—Load Balancer 7—Content host

When a client (1) requests file X, the request crosses the Internet (2) and enters the enterprise's internal network through its Internet gateway (3). The proxy server intercepts the request, generates a new request with its own IP address as the originating address, and sends the new request to Load Balancer at the cluster address. Load Balancer uses its load-balancing algorithm to determine which content host is currently best able to satisfy the request for file X. That content host

returns file X to the proxy server rather than via Load Balancer. The proxy server determines whether to cache it and delivers it to the end user in the same way as described previously.

## Caching dynamic content

Advanced caching functionality is also provided by Caching Proxy's Dynamic Caching plug-in. When used in conjunction with WebSphere Application Server, Caching Proxy has the ability to cache, serve, and invalidate dynamic content in the form of JavaServer Pages (JSP) and servlet responses generated by a WebSphere Application Server.

Generally, dynamic content with an indefinite expiration time must be marked "do not cache" because the standard time-based cache expiration logic does not ensure its timely removal. The Dynamic Caching plug-in's event-driven expiration logic enables content with an indefinite expiration time to be cached by the proxy server. Caching such content at the edge of the network relieves content hosts from repeatedly invoking an Application Server to satisfy requests from clients. This can offer the following benefits:

- Reduced workload on Web servers, WebSphere Application Servers, and back-end content hosts
- Faster response to users by eliminating network delays
- Reduced bandwidth usage due to fewer Internet traversals
- Better scalability of Web sites that serve dynamically generated content

Servlet response caching is ideal for dynamically produced Web pages that expire based on application logic or an event such as a message from a database. Although such a page's lifetime is finite, the time-to-live value cannot be set at the time of creation because the expiration trigger cannot be known in advance. When the time-to-live for such pages is set to zero, content hosts incur a high penalty when serving dynamic content.

The responsibility for synchronizing the dynamic cache of Caching Proxy and Application Server is shared by both systems. For example, a public Web page dynamically created by an application that gives the current weather forecast can be exported by Application Server and cached by Caching Proxy. Caching Proxy can then serve the application's execution results repeatedly to many different users until notified that the page is invalid. Content in Caching Proxy's servlet response cache is valid until the proxy server removes an entry because the cache is congested, the default timeout set by the ExternalCacheManager directive in Caching Proxy's configuration file expires, or Caching Proxy receives an Invalidate message directing it to purge the content from its cache. Invalidate messages originate at the WebSphere Application Server that owns the content and are propagated to each configured Caching Proxy.

**Note:** Dynamically generated private pages (such as a page showing the contents of a user's shopping cart) generally cannot and should not be cached by Caching Proxy. Caching Proxy can cache and serve private pages only when it is configured to perform authentication and authorization to ensure that the private pages are served only to their intended users.

## Additional caching features

Caching Proxy offers other key advanced caching features:

- The ability to use very large caches

- An option to automatically refresh the cache with the most-frequently accessed pages
- The possibility to cache even those pages where the header information says to fetch them every time
- Configurable daily garbage collection to improve server performance and ensure cache maintenance
- Remote Cache Access (RCA), a function that allows multiple Caching Proxy machines to share the same cache, thereby reducing the redundancy of cached content
- The ICP plug-in, which enables Caching Proxy to query Internet Caching Protocol (ICP)-compliant caches in search of HTML pages and other cacheable resources



---

## Chapter 5. Network performance

Network performance is affected by the introduction of Caching Proxy functionality. Use Caching Proxy alone or in conjunction with Load Balancer to improve the performance of your network. See Chapter 1, “Introducing WebSphere Application Server Edge components,” on page 3 for an introduction to these systems.

Caching Proxy's performance within your enterprise is only as good as the hardware on which it runs and the overall architecture of the system into which it is introduced. To optimize network performance, model hardware and overall network architecture to the characteristics of proxy servers.

Basic configuration and administration of the Caching Proxy software and tuning at the operating system level also greatly contribute to Caching Proxy performance. Many software configuration changes can be made to yield increased performance; these include, but are not limited to, adjusting the logging directives, mapping rules, plug-ins, timeout values, cache configuration values, and active thread values. Details on configuring Caching Proxy software is presented in the *Caching Proxy Administration Guide*.

Many operating system configuration changes can be made to yield increased performance as well; these include, but are not limited to, tuning TCP and ARP, increasing the file descriptor limits, synchronizing system clocks, tuning Network Interface Cards (NICs), and following common good practices when performing system administration tasks.

**IMPORTANT:** Caching Proxy is available on all Edge component installations, with the following exceptions:

- Caching Proxy is not available for Edge component installations that run on Itanium 2 or AMD Opteron 64-bit processors.
- Caching Proxy is not available for Edge component installations of Load Balancer for IPv4 and IPv6.

---

### Network hardware

This section discusses network hardware issues to consider when introducing Caching Proxy functionality into your network.

#### Memory considerations

A large amount of memory must be dedicated to the proxy server. Caching Proxy can consume 2 GB of virtual address space when a large memory-only cache is configured. Memory is also needed for the kernel, shared libraries, and network buffers. Therefore, it is possible to have a proxy server that consumes 3 or 4 GB of physical memory. Note that a memory-only cache is significantly faster than a raw disk cache, and this configuration change alone can be considered a performance improvement.

#### Hard disk considerations

It is important to have a large amount of disk space on the machine on which Caching Proxy is installed. This is especially true when disk caches are used.

Reading and writing to a hard disk is an intensive process for a computer. Although Caching Proxy's I/O procedures are efficient, the mechanical limitations of hard drives can limit performance when the Caching Proxy is configured to use a disk cache. The disk I/O bottleneck can be alleviated with practices such as using multiple hard disks for raw cache devices and log files and by using disk drives with fast seek times, rotational speeds, and transfer rates.

## Network considerations

Network requirements such as the speed, type, and number of NICs, and the speed of the network connection to the proxy server affect the performance of Caching Proxy. It is generally in the best interest of performance to use two NICs on a proxy server machine: one for incoming traffic and one for outgoing traffic. It is likely that a single NIC can reach its maximum limit by HTTP request and response traffic alone. Furthermore, NICs should be at least 100 MB, and they should always be configured for full-duplex operation. This is because automatic negotiation between routing and switching equipment can possibly cause errors and hinder throughput. Finally, the speed of the network connection is very important. For example, you cannot expect to service a high request load and achieve optimal throughput if the connection to the Caching Proxy machine is a saturated T1 carrier.

## CPU considerations

The central processing unit (CPU) of a Caching Proxy machine can possibly become a limiting factor. CPU power affects the amount of time it takes to process requests and the number of CPUs in the network affects scalability. It is important to match the CPU requirements of the proxy server to the environment, especially to model the peak request load that the proxy server will service.

---

## Network architecture

For overall performance, it is generally beneficial to scale the architecture and not just add individual pieces of hardware. No matter how much hardware you add to a single machine, that hardware still has a maximum level of performance.

The section discusses network architecture issues to take into consideration when introducing Caching Proxy functionality into your network.

## Web site popularity and proxy server load considerations

If your enterprise's Web site is popular, there can be greater demand for its content than a single proxy server can satisfy effectively, resulting in slow response times. To optimize network performance, consider including clustered, load-balanced Caching Proxy machines or using a shared cache architecture with Remote Cache Access (RCA) in your overall network architecture.

- **Load-balanced clusters**

One way to scale the architecture is to cluster proxy servers and use the Load Balancer component to balance the load among them. Clustering proxy servers is a beneficial design consideration not only for performance and scalability reasons but for redundancy and reliability reasons as well. A single proxy server represents a single point of failure; if it fails or becomes inaccessible because of a network failure, users cannot access your Web site.

- **Shared cache architecture**

Also consider a shared cache architecture with RCA. A shared cache architecture spreads the total virtual cache among multiple Caching Proxy servers that



usually use an intercache protocol like the Internet Cache Protocol (ICP) or the Cache Array Routing Protocol (CARP). RCA is designed to maximize clustered cache hit ratios by providing a large virtual cache.

Performance benefits result from using an RCA array of proxy servers as opposed to a single stand-alone Caching Proxy or even a cluster of stand alone Caching Proxy machines. For the most part, the performance benefits are caused by the increase in the total virtual cache size, which maximizes the cache hit ratio and minimizes cache inconsistency and latency. With RCA, only one copy of a particular document resides in the cache. With a cluster of proxy servers, the total cache size is increased, but multiple proxy servers are likely to fetch and cache the same information. The total cache hit ratio is therefore not increased.

RCA is commonly used in large enterprise content-hosting scenarios. However, RCA's usefulness is not limited to extremely large enterprise deployments. Consider using RCA if your network's load requires a cluster of cache servers and if the majority of requests are cache hits. Depending on your network setup, RCA does not always improve enterprise performance due to an increase in the number of TCP connections that a client uses when RCA is configured. This is because an RCA member is not only responsible for servicing URLs for which it has the highest score but it must also forward requests to other members or clusters if it gets a request for a URL for which it does not have the highest score. This means that any given member of an RCA array might have more open TCP connections than it would if it operated as a stand-alone server.

## Traffic type considerations

Major contributions to improved performance stem from Caching Proxy's caching capabilities. However, the cache of the proxy server can become a bottleneck if it is not properly configured. To determine the best cache configuration, a significant effort must be made to analyze traffic characteristics. The type, size, amount, and attributes of the content affect the performance of the proxy server in terms of the time it takes to retrieve documents from origin servers and the load on the server. When you understand the type of traffic that Caching Proxy is going to proxy or serve from its cache, then you can factor in those characteristics when configuring the proxy server. For example, knowing that 80% of the objects being cached are images (\*.gif or \*.jpg) and are approximately 200 KB in size can certainly help you tune caching parameters and determine the size of the cache. Additionally, understanding that most of the content is personalized dynamic pages that are not candidates for caching is also pertinent to tuning Caching Proxy.

Analyzing traffic characteristics enables you to determine whether using a memory or disk cache can optimize your cache's performance. Also, familiarity with your network's traffic characteristics enables you to determine whether improved performance can result from using the Caching Proxy's dynamic caching feature.

- **Memory versus disk cache**

Disk caches are appropriate for sites with large amounts of information to be cached. For example, if the site content is large (greater than 5 GB) and there is an 80 to 90% cache hit rate, then a disk cache is recommended. However, it is known that using a memory (RAM) cache is faster, and there are many scenarios when using a memory-only cache is feasible for large sites. For example, if Caching Proxy's cache hit rate is not as important or if a shared cache configuration is being used, then a memory cache is practical.

- **Caching dynamically generated content**

Caching Proxy can cache and invalidate dynamic content (JSP and servlet results) generated by the WebSphere Application Server dynamic cache,

providing a virtual extension of the Application Server cache into network-based caches. Enabling the caching of dynamically generated content is beneficial to network performance in an environment where there are many requests for dynamically produced public Web pages that expire based on application logic or an event such as a message from a database. The page's lifetime is finite, but an expiration trigger cannot be set in at the time of its creation; therefore, hosts without a dynamic caching and invalidation feature must designate such as page as having a time-to-live value of zero.

If such a dynamically generated page will be requested more than once during its lifetime by one or more users, then dynamic caching provides a valuable offload and reduces the workload on your network's content hosts. Using dynamic caching also improves network performance by providing faster response to users by eliminating network delays and reducing bandwidth usage due to fewer Internet traversals.

---

## Chapter 6. Availability

Functioning in conjunction with content hosts, such as WebSphere Application Server, or with the Application Server Caching Proxy component, the Application Server Load Balancer component enables you to enhance your network's availability and scalability. (See Chapter 1, "Introducing WebSphere Application Server Edge components," on page 3 for an introduction to these Edge Components.) Load Balancer is used by enterprise networks and is installed between the Internet and the enterprise's back-end servers. Load Balancer acts as the enterprise's single point-of-presence on the Internet, even if the enterprise uses multiple back-end servers because of high demand or a large amount of content.

Availability is achieved through load balancing and failover support.

**IMPORTANT:** Caching Proxy is available on all Edge component installations, with the following exceptions:

- Caching Proxy is not available for Edge component installations that run on Itanium 2 or AMD Opteron 64-bit processors.
- Caching Proxy is not available for Edge component installations of Load Balancer for IPv4 and IPv6.

---

### Load balancing

Load balancing improves your Web site's availability and scalability by transparently clustering proxy servers and application servers. An IT infrastructure's scalability is greatly improved because back-end processing power can be added transparently.

#### Load balancing multiple content hosts

You can satisfy high demand by duplicating content on multiple hosts, but then you need a way to balance the load among them. Domain Name Service (DNS) can provide basic round-robin load balancing, but there are several situations in which it does not perform well.

A more sophisticated solution for load balancing multiple content hosts is to use Load Balancer's Dispatcher component as depicted in Figure 5 on page 26. In this configuration, all of the content hosts (the machines marked 5) store the same content. They are defined to form a load-balanced *cluster*, and one of the network interfaces of the Load Balancer machine (4) is assigned a host name and IP address dedicated to the cluster. When an end user working on one of the machines marked 1 requests file X, the request crosses the Internet (2) and enters the enterprise's internal network through its Internet gateway (3). The Dispatcher intercepts the request because its URL is mapped to the Dispatcher's host name and IP address. The Dispatcher determines which of the content hosts in the cluster is currently best able to service the request, and forwards the request to that host, which, when the MAC forwarding method is configured, returns file X directly to the client (that is, file X does not pass through Load Balancer).

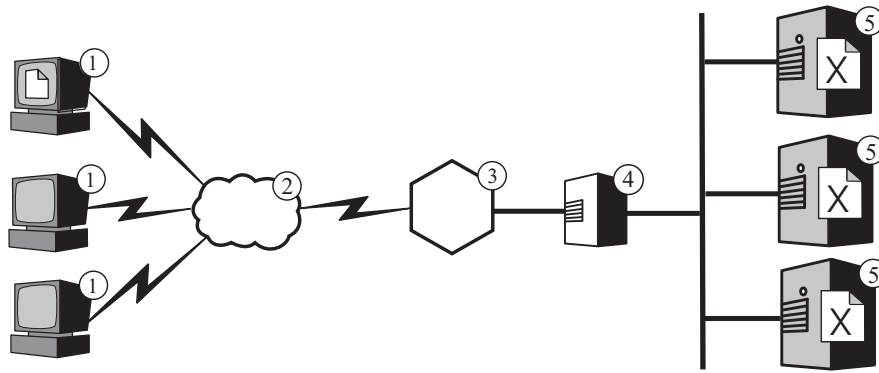


Figure 5. Load balancing multiple content hosts. Legend:  
 1—Client 2—Internet 3—Router/Gateway 4—Dispatcher 5—Content host

**Note:** The Dispatcher provides three forwarding methods:

- The MAC forwarding method is used to load balance incoming requests to the server. Responses are returned directly to the client.
- The NAT/NAPT forwarding method is used with servers located remotely. Incoming requests are load balanced by Dispatcher, which receives the responses from the server and returns them to the clients. (On Load Balancer for IPv4 and IPv6 installations, this forwarding method is not supported.)
- The content-based routing method (cbr) provides content-based routing of HTTP and HTTPS requests without the use of Caching Proxy. Content-based routing is performed for HTTP by using the "content" type rule and for HTTPS by using SSL session ID affinity. (On Load Balancer for IPv4 and IPv6 installations, this forwarding method is not supported.)

By default, the Dispatcher uses round-robin load balancing like DNS, but even so it addresses many of DNS's inadequacies. Unlike DNS, it tracks whether a content host is unavailable or inaccessible and does not continue to direct clients to an unavailable content host. Further, it considers the current load on the content hosts by tracking new, active, and finished connections. You can further optimize load balancing by activating Load Balancer's optional advisor and manager components, which track a content host's status even more accurately and incorporate the additional information into the load-balancing decision process. The manager enables you to assign different weights to the different factors used in the decision process, further customizing load balancing for your site.

## Load balancing multiple reverse proxy servers

Load Balancer's Dispatcher can also perform load balancing for multiple Caching Proxy machines. If your enterprise's Web site is popular, there can be greater demand for its contents than a single proxy server can satisfy effectively, potentially degrading the proxy server's performance.

You can have multiple Caching Proxy systems performing proxy functions for a single content host (similar to the configuration depicted in Figure 1 on page 14), but if your site is popular enough to need multiple proxy servers, then you probably also need multiple contents hosts whose loads are balanced by Load Balancer. Figure 6 on page 27 depicts this configuration. The Dispatcher marked 4 load balances a cluster of two proxy servers (5), and the Dispatcher marked 7 load balances a cluster of three content hosts (8).

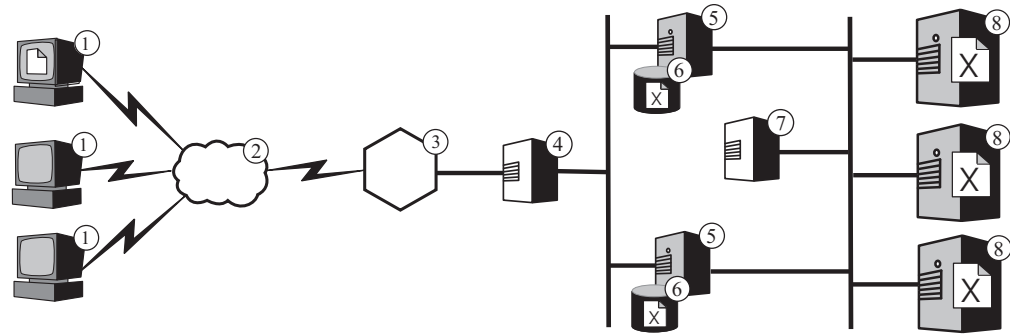


Figure 6. Load balancing multiple reverse proxy servers and content hosts.  
 1—Client 2—Internet 3—Router/Gateway 4—Dispatcher 5—Caching Proxy 6—Cache 7—Dispatcher 8—Content host

The cluster host name of the Dispatcher marked 4 is the host name that appears in URLs for the enterprise's Web site (that is, it is the name of the Web site as visible on the Internet). The cluster host name for the Dispatcher marked 7 is not visible on the Internet and so can be any value you wish. As an example, for the ABC Corporation an appropriate host name for the Dispatcher marked 4 is `www.abc.com`, whereas the Dispatcher marked 7 can be called something like `http-balancer.abc.com`.

Suppose that a browser on one of the client machines marked 1 needs to access file X stored on the content servers marked 8. The HTTP request crosses the Internet (2) and enters the enterprise's internal network at the gateway (3). The router directs the request to the Dispatcher marked 4, which passes it to the proxy server (5), which is currently best able to handle it according to the load-balancing algorithm. If the proxy server has file X in its cache (6), it returns it directly to the browser, bypassing the Dispatcher marked 4.

If the proxy server does not have a copy of file X in its cache, it creates a new request that has its own host name in the header's origin field and sends that to the Dispatcher marked 7. The Load Balancer determines which content host (8) is currently best able to satisfy the request, and directs the request there. The content host retrieves file X from storage and returns it directly to the proxy server, bypassing the Dispatcher marked 7. The proxy server caches file X if appropriate, and forwards it to the browser, bypassing the Dispatcher marked 4.

## Load Balancer with multiple forward proxy servers

If you provide Internet access to a large number of clients, they can generate more demand for Internet access than a single proxy can provide efficiently. As the Caching Proxy becomes overloaded with requests, clients can possibly experience worse response times than with direct Internet access. And if the Caching Proxy fails or becomes inaccessible because of a network failure, Internet access becomes impossible. The solution is to install multiple Caching Proxy machines and use the Load Balancer's Dispatcher to balance the load between them.

Without the Dispatcher, you can provide true transparent proxy with multiple Caching Proxy machines only if your routers can route the same type of traffic to more than one Caching Proxy; not all routers support this. It is possible to provide regular forward proxy service on multiple Caching Proxy machines without the Dispatcher, but you must explicitly configure the client browsers to use one of the Caching Proxy machines as their primary proxy. If that Caching Proxy fails or

becomes overloaded or inaccessible, end users can become unable to access the Internet. To avoid that situation, you can create a proxy automatic configuration (PAC) file (as described in “Transparent forward Caching Proxy (Linux systems only)” on page 16) that directs the browser to fail over to one or more secondary caching proxies. A PAC file does not address the need to balance the load among the Caching Proxy machines; however, if one Caching Proxy receives many more requests than another, its performance is likely to degrade, subjecting its browser clients to slower response times. For all clients to experience similar performance, you must configure an approximately equal number of browsers to use each Caching Proxy, and track the distribution manually so that you can keep the load even as you add or remove browsers.

Figure 7 depicts a network configuration in which the Dispatcher load balances a cluster of Caching Proxy machines. One of the Dispatcher machine’s network interfaces is configured to have the cluster’s dedicated host name and IP address. Client browsers are configured to direct their Internet requests to the cluster host name. When, for example, a browser on one of the client machines marked 1 needs to access file X on a content host (7), it directs its request to the cluster host name or address, where the Dispatcher (2) intercepts it and directs it to the appropriate Caching Proxy (3). The Caching Proxy creates a new request, passes it through the enterprise’s gateway (5) and across the Internet (6), and if appropriate stores the returned file in its cache (4) as described in more detail in “Forward Caching Proxy” on page 14

**Note:** The transparent proxy feature is available on Linux systems only.

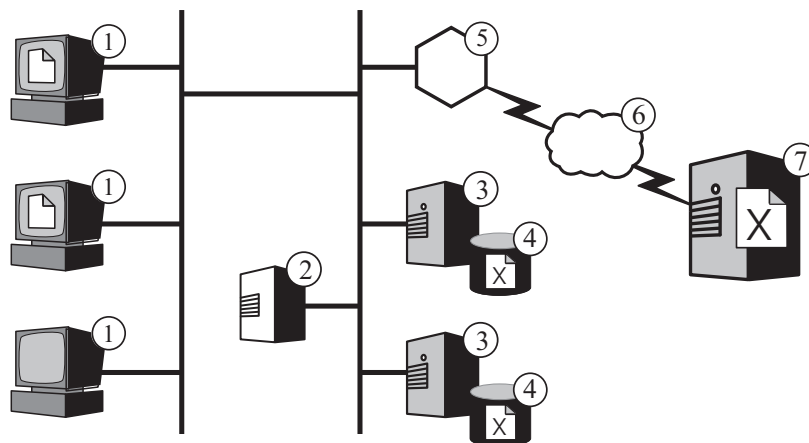


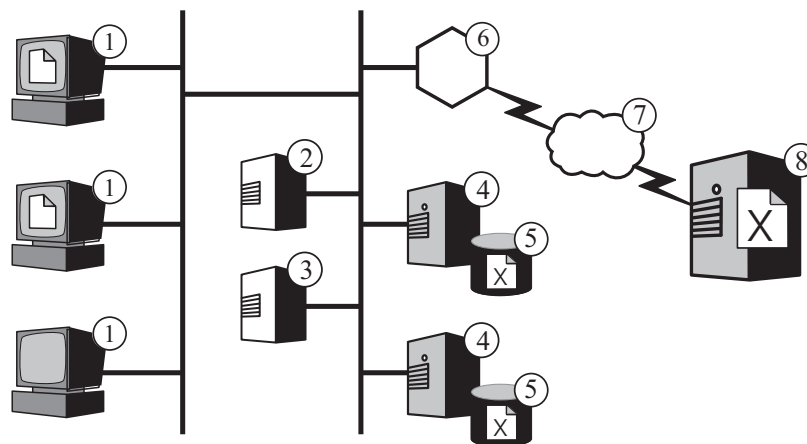
Figure 7. Using Dispatcher to load balance multiple caching proxies.. Legend:  
 1—Client 2—Dispatcher 3—Caching Proxy 4—Cache 5—Router/  
 Gateway 6—Internet 7—Content host

The Dispatcher detects when one of the Caching Proxy machines is unavailable and automatically routes requests to the other one. This allows you to shut down a Caching Proxy machine for maintenance without interrupting Internet access. The Dispatcher has numerous configuration options that you can use to control the factors it considers when making load-balancing decisions. You can also install auxiliary Dispatcher programs on the Caching Proxy machines to monitor their status and return the information to the Dispatcher. For details, see the *WebSphere Application Server Load Balancer Administration Guide*. Using multiple Caching Proxy's introduces a potential inefficiency, because more than one Caching Proxy can cache the same file if different clients request the file via different Caching Proxy machines. To eliminate the redundancy, you can configure remote cache access (RCA), which enables all of the proxies in a defined group to share the

contents of their caches with one another. The proxies in the RCA group all use the same algorithm to determine which Caching Proxy is responsible for a given URL. When a Caching Proxy intercepts a URL for which it is not responsible, it passes the request to the responsible Caching Proxy. The responsible Caching Proxy does the work necessary to satisfy the request, either retrieving it from its cache or forwarding the request to the relevant content host and caching the returned file if appropriate. The responsible Caching Proxy then passes the file to the original Caching Proxy, which delivers it to the requesting end user.

In the RCA group, if the Caching Proxy responsible for a given URL is failed, then the original Caching Proxy, which received the client request, will directly access the content host (or a backup Caching Proxy server, if it is defined). This implies that users can access files as long as at least one Caching Proxy in an RCA group is functioning correctly.

This configuration satisfies high demand for Internet access by using the Dispatcher to balance the load of requests across multiple Caching Proxy machines. One potential problem is that the Dispatcher is a single point of failure. If it fails or becomes inaccessible due to a network failure, browser clients cannot reach the Caching Proxy's or the Internet. The solution is to configure another Dispatcher to act as a backup for the primary Dispatcher, as depicted in Figure 8.



*Figure 8. Using a primary and backup Dispatcher to provide highly available Internet access..*  
 Legend: 1—Client 2—Primary Dispatcher 3—Backup Dispatcher 4—Caching Proxy 5—Cache 6—Router/Gateway 7—Internet 8—Content host

Here a browser running on one of the machines marked 1 normally directs its request for a file X to the primary Dispatcher (2), which routes the request to the Caching Proxy (4) selected on the basis of the Dispatcher's load-balancing criteria. The Caching Proxy creates a new request, routes it through the enterprise's gateway (6) across the Internet (7) to the content host (8), and stores the returned file X in its cache (5) if appropriate (for a more detailed description of this part of the process, see "Forward Caching Proxy" on page 14).

In this configuration, the backup Dispatcher (3) does not perform load balancing as long as the primary is operational. The primary and backup Dispatchers track each other's status by periodically exchanging messages called heartbeats. If the backup Dispatcher detects that the primary has failed, it automatically takes over the responsibility for load balancing by intercepting requests directed to the primary's host name and IP address. It is also possible to configure two Dispatchers for mutual high availability. In this case each actively performs load balancing for a



separate cluster of caching proxies, simultaneously acting as the backup for its colleague. For further discussion, see the *WebSphere Application Server Load Balancer Administration Guide*.

The Dispatcher does not generally consume many processing or memory resources, and other applications can run on the Dispatcher machine. If it is important to minimize equipment costs, it is even possible to run the backup Dispatcher on the same machine as Caching Proxy. Figure 9 depicts such a configuration, in which the backup Dispatcher runs on the same machine (3) as the Caching Proxy.

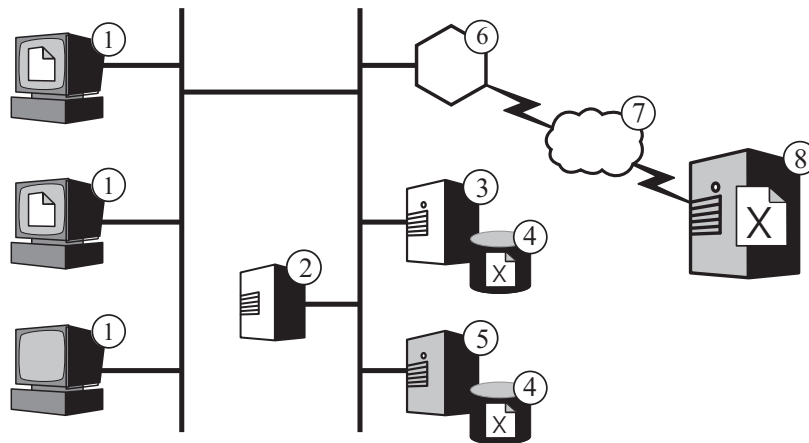


Figure 9. Locating the backup Dispatcher on a Caching Proxy machine.. Legend:  
1—Client 2—Primary Dispatcher 3—Backup Dispatcher and Caching Proxy  
4—Caching Proxy 5—Cache 6—Router/Gateway 7—Internet 8—Content host

## Failover support

Load Balancer acts as a single point-of-presence for your enterprise's content hosts. This is beneficial because you advertise the cluster host name and address in DNS, rather than the host name and address of each content host, which provides a level of protection against casual attacks and provides a unified feel for your enterprise's Web site. To further enhance Web site availability, configure another Load Balancer to act as a backup for the primary Load Balancer, as depicted in Figure 10 on page 31. If one Load Balancer fails or becomes inaccessible due to a network failure, end users can still reach the content hosts.



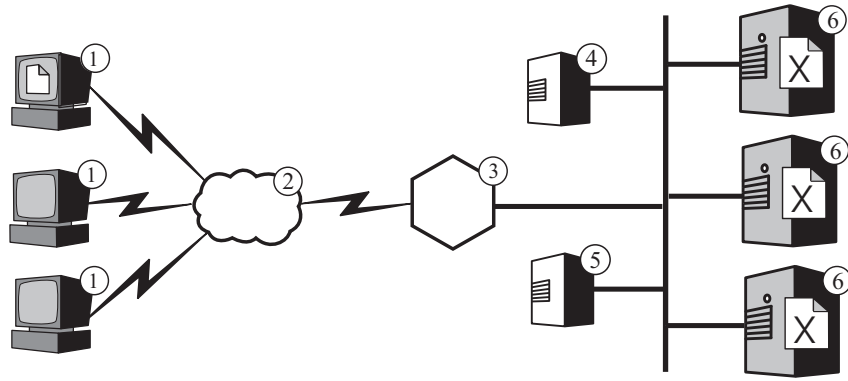


Figure 10. Using a primary and backup Load Balancer to make Web content highly available.  
 Legend: 1—Client 2—Internet 3—Router/Gateway 4—Primary Dispatcher 5—Backup Dispatcher 6—Content host

In the normal case, a browser running on one of the machines marked 1 directs its request for a file X to the cluster host name that is mapped to the primary Load Balancer (4). The Dispatcher routes the request to the content host (6) selected on the basis of the Dispatcher's load-balancing criteria. The content host sends file X directly to the browser, routing it through the enterprise's gateway (3) across the Internet (2) but bypassing Load Balancer.

The backup Dispatcher (5) does not perform load balancing as long as the primary one is operational. The primary and backup Dispatchers track each other's status by periodically exchanging messages called *heartbeats*. If the backup Dispatcher detects that the primary has failed, it automatically takes over the responsibility for load balancing by intercepting requests directed to the primary's cluster host name and IP address.

It is also possible to configure two Dispatchers for *mutual high availability*. In this case, each actively performs load balancing for a separate cluster of content hosts, simultaneously acting as the backup for its colleague. (On Load Balancer for IPv4 and IPv6 installations simple high availability is supported, but mutual high availability is not.)

The Dispatcher does not generally consume many processing or memory resources, and other applications can run on the Load Balancer machine. If it is vital to minimize equipment costs, it is even possible to run the backup Dispatcher on one of the machines in the cluster it is load balancing. Figure 11 on page 32 depicts such a configuration, in which the backup Dispatcher runs on one of the content hosts (5) in the cluster.

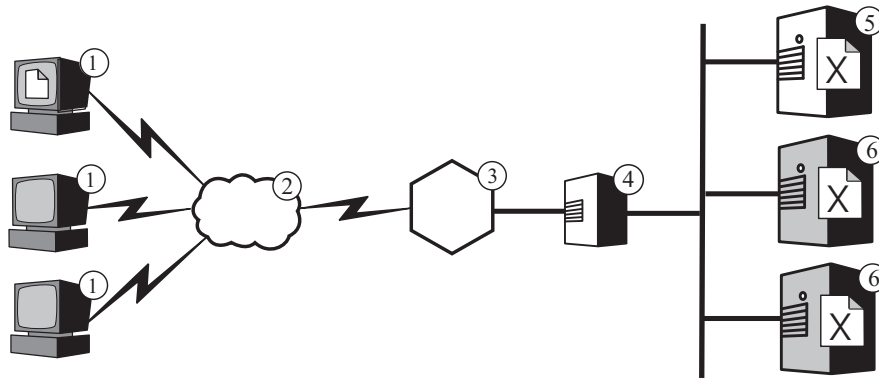


Figure 11. Locating the backup Load Balancer on a content host. Legend:  
 1—Client 2—Internet 3—Router/Gateway 4—Primary Dispatcher 5—Backup  
 Dispatcher and content host 6—Content host

---

## Chapter 7. Content Based Routing

IMPORTANT: The Content Based Routing (CBR) component is available on all supported platforms, with the following exceptions:

- CBR is not available with Load Balancer installations on platforms running a 64-bit JVM.

Alternatively, for this type of installation, you can use the cbr forwarding method of Load Balancer's Dispatcher component to provide content-based routing of HTTP and HTTPS requests without the use of Caching Proxy. See the *WebSphere Application Server Load Balancer Administration Guide* for more information.

- CBR is not available with Load Balancer for IPv4 and IPv6 installations.

Load Balancer for IPv4 and IPv6 supports only the Dispatcher component's mac forwarding method. The nat and cbr forwarding methods are not supported.

Functioning in conjunction with the Application Server Caching Proxy component, the Application Server Load Balancer component enables you to distribute requests to multiple back-end servers that host different content. (See Chapter 1, "Introducing WebSphere Application Server Edge components," on page 3 for an introduction to these Edge Components.)

If Load Balancer's Content Based Routing (CBR) component is installed together with Caching Proxy, HTTP requests can be distributed based on URL or other administrator-determined characteristics, eliminating the need to store identical content on all back-end servers.

Using CBR is especially appropriate if your Web servers need to perform several different functions or offer several types of services. For example, an online retailer's Web site must both display its catalog, a large portion of which is static, and accept orders, which means running an interactive application such as a Common Gateway Interface (CGI) script to accept item numbers and customer information. Often it is more efficient to have two different sets of machines perform the distinct functions, and to use CBR to route the different types of traffic to different machines. Similarly, an enterprise can use CBR to provide better service to paying customers than to casual visitors to its Web site, by routing the paid requests to more powerful Web servers.

CBR routes requests based on rules that you write. The most common type is the *content rule*, which directs requests based on the path name in the URL. For example, the ABC Corporation can write rules that direct requests for the URL `http://www.abc.com/catalog_index.html` to one cluster of servers and `http://www.abc.com/orders.html` to another cluster. There are also rules that route requests based on the IP address of the client who sent them or on other characteristics. For a discussion, see the *WebSphere Application Server Load Balancer Administration Guide* chapters about configuring CBR and about advanced Load Balancer and CBR functions. For syntax definitions for the rules, see the *WebSphere Application Server Load Balancer Administration Guide* appendix about CBR rule types.

Figure 12 on page 34 depicts a simple configuration in which Load Balancer's CBR component and Caching Proxy are installed together on the machine marked 4 and route requests to three content hosts (6, 7, and 8) that house different content.

When an end user working on one of the machines marked 1 requests file X, the request crosses the Internet (2) and enters the enterprise's internal network through its Internet gateway (3). The proxy server intercepts the request and passes it to the CBR component on the same machine, which parses the URL in the request and determines that content host 6 houses file X. The proxy server generates a new request for file X, and if its caching feature is enabled, determines whether the file is eligible for caching when host 6 returns it. If the file is cacheable, the proxy server stores a copy in its cache (5) before passing it to the end user. Routing for other files works in the same manner: requests for file Y go to content host 7, and requests for file Z go to content host 8.

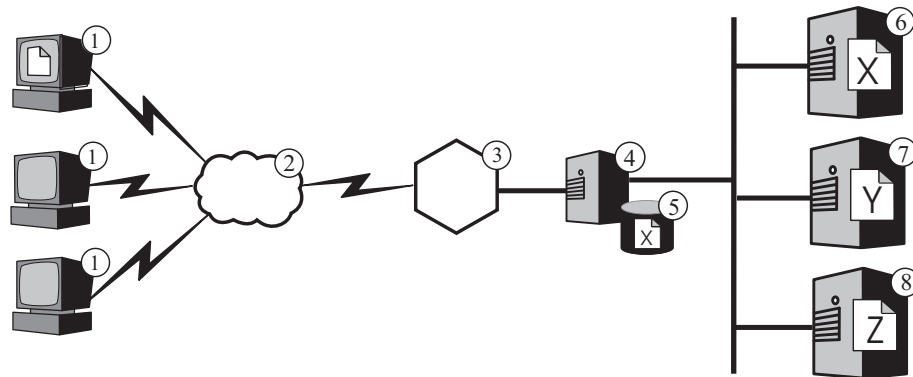


Figure 12. Routing HTTP requests with CBR. Legend: 1—Client 2—Internet 3—Router/Gateway 4—Caching Proxy and Load Balancer's CBR component 5—Cache 6, 7, 8—Content host

Figure 13 on page 35 depicts a more complex configuration which is possibly suitable for an online retailer. Load Balancer's CBR component and the proxy server are installed together on the machine marked 4 and route requests to two Load Balancer machines. The Load Balancer machine marked 6 load balances a cluster of content hosts (8) that house the mostly static content of the retailer's catalog, whereas the Load Balancer marked 7 load balances a cluster of Web servers that handle orders (9).

When an end user working on one of the machines marked 1 accesses the URL for the retailer's catalog, the request crosses the Internet (2) and enters the enterprise's internal network through its Internet gateway (3). The proxy server intercepts the request and passes it to the CBR component on the same machine, which parses the URL and determines that the Load Balancer machine marked 6 handles that URL. The proxy server creates a new access request and sends it to the Load Balancer, which determines which of the content hosts marked 8 is currently best able to service the request (based on criteria that you define). That content host passes the catalog content directly to the proxy server, bypassing Load Balancer. As in the preceding example, the proxy server determines whether the content is cacheable and stores it in its cache (5) if appropriate.

The end user places an order by accessing the retailer's ordering URL, presumably via a hyperlink in the catalog. The request travels the same path as the catalog access request, except that the CBR component on machine 4 routes it to the Load Balancer machine marked 7. Load Balancer forwards it to the most suitable of the Web servers marked 9, which replies directly to the proxy server. Because ordering information is generally dynamically generated, the proxy server probably does not cache it.

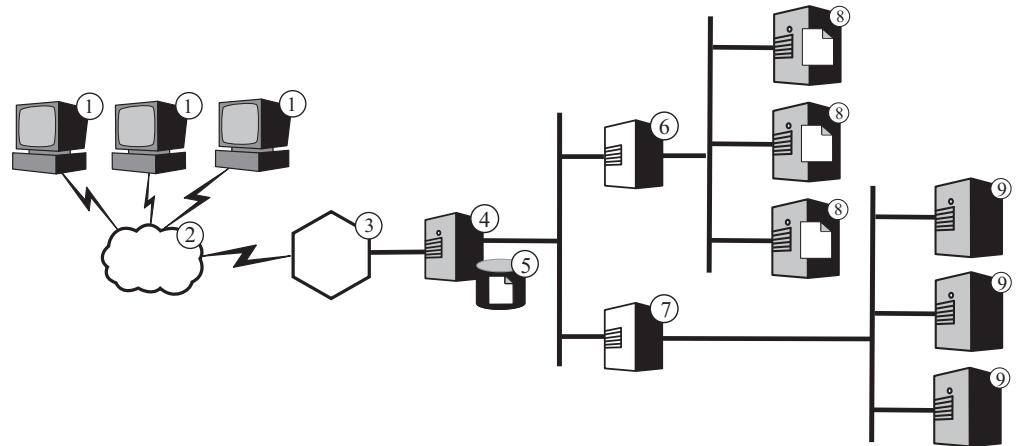


Figure 13. Load balancing HTTP requests routed with CBR. Legend:  
 1—Client 2—Internet 3—Router/Gateway 4—Caching Proxy and Load Balancer's  
 CBR component 5—Cache 6, 7—Load Balancer 8—Content host 9—Web Server

Load Balancer's CBR function supports *cookie affinity*. This means that the identity of the server that serviced an end user's first request is recorded in a special packet of data (a *cookie*) included in the server's response. When the end user accesses the same URL again within a period of time that you define, and the request includes the cookie, CBR routes the request to the original server rather than reapplying its standard rules. This generally improves response time if the server has stored information about the end user that it does not have to obtain again (such as a credit card number).



---

## Part 3. Scenarios

This part discusses business scenarios that use IBM WebSphere Application Server Edge components. These are architecturally sound and tested solutions that can provide excellent performance, availability, scalability, and reliability.

This part contains the following chapters:

Chapter 8, “Business-to-consumer network,” on page 39

Chapter 9, “Business to client-banking-solution,” on page 43

Chapter 10, “Web portal network,” on page 45





---

## Chapter 8. Business-to-consumer network

The basic electronic commerce Web site is a business-to-consumer network. In the first phase of Internet growth, businesses typically focus on simply creating a Web presence. Corporate information and product catalogs are converted to digital formats and made available on the Web site. Shopping can be available by providing e-mail addresses, telephone and fax numbers, and even automated forms. True online shopping, however, is not available. All transactions have an inherent latency because humans need to process the order.

In phase two, businesses eliminate this latency and streamline their sales operation by implementing secure shopping carts for direct online purchases. Synchronization with warehouse databases and integration with banking systems are crucial to completing these sales transactions. Product that is not available cannot be sold, and a customer's account cannot be charged for that item. Likewise, a product cannot be taken from inventory and shipped to a customer until a valid financial transaction occurs.

In the third phase, the corporate Web site evolves into a dynamic presentation site where the consumer begins to take on the aspects of a client and is provided with personalized content.

The following scenerio includes both Load Balancer and Caching Proxy.

**IMPORTANT:** Caching Proxy is available on all Edge component installations, with the following exceptions:

- Caching Proxy is not available for Edge component installations that run on Itanium 2 or AMD Opteron 64-bit processors.
- Caching Proxy is not available for Edge component installations of Load Balancer for IPv4 and IPv6.

---

### Phase 1

Figure 14 on page 40 shows a small commercial Web site designed to provide efficient catalog browsing. All client requests pass through the firewall to a Dispatcher that routes the requests to a cluster of proxy servers with active caches that act as surrogate servers to the Web servers. Metric servers are colocated with the proxy servers to provide load-balancing data to the Dispatcher. This arrangement reduces the network load on the Web servers and isolates them from direct contact with the Internet.

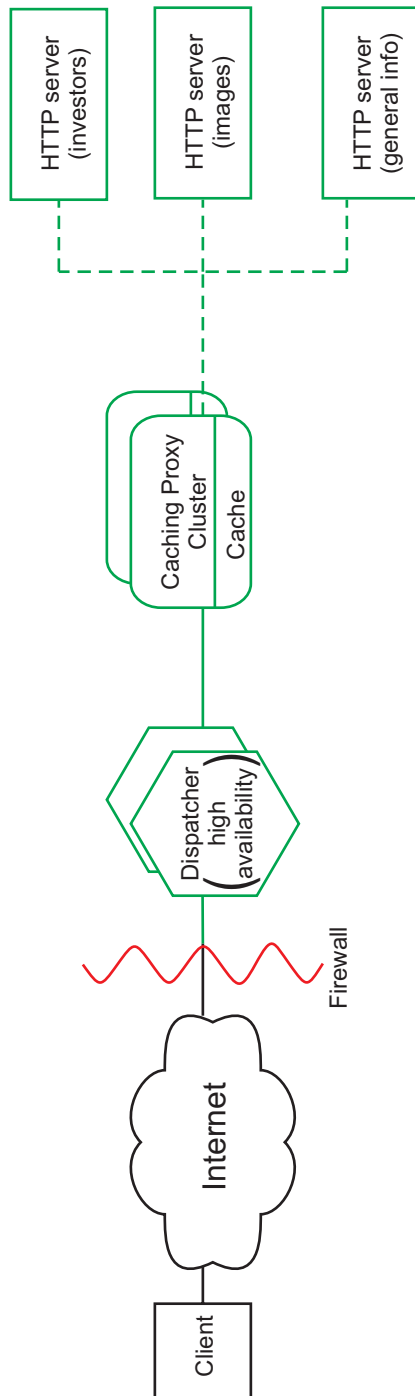


Figure 14. Business to consumer network (Phase 1)

## Phase 2

Figure 15 on page 41 shows a the second phase of evolution for a commercial Web site designed to provide efficient catalog browsing and fast, secure shopping carts for potential customers. All customer requests are routed to the appropriate branch of the network by a Dispatcher that separates requests based on Internet protocol. HTTP requests go to the static Web site; HTTPS requests go to the shopping

network. The primary, static Web site is still served by a cluster of proxy servers with active caches that acts as a surrogate for the Web servers. This part of the network mirrors the network in the first phase.

The electronic commerce portion of the Web site is also served by a cluster of proxy servers. However, the Caching Proxy nodes are enhanced with several plug-in modules. The SSL handshaking is offloaded to a cryptographic hardware card, and authentication is performed through the Access Manager (formerly Policy Director) plug-in. A Dynamic Caching plug-in reduces the workload on the WebSphere Application Server by storing common data. A plug-in on the application server invalidates objects in the Dynacache when necessary.

All shopping cart applications are tied into the customer database that was used to authenticate the user. This prevents the user from having to enter personal information into the system twice, once for authentication and once for shopping.

This network divides traffic according to client usage, removing the processor-intensive SSL authentication and electronic commerce shopping carts from the primary Web site. This dual-track Web site allows the network administrator to tune the various servers to provide excellent performance based on the role of the server within the network.

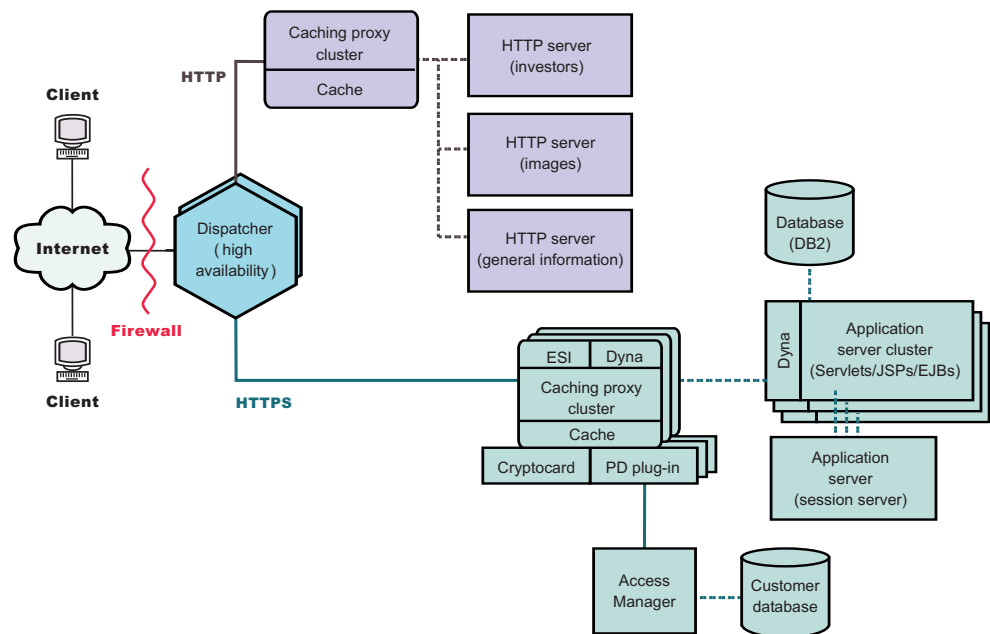


Figure 15. Business to consumer network (Phase 2)

### Phase 3

Figure 16 on page 42 shows the third phase of the evolution of a business-to-consumer network, with the static Web adopting a dynamic presentation method. The proxy server cluster has been enhanced to support the caching of dynamic Web content and assembly of page fragments written to comply with the Edge Side Includes (ESI) protocol. Rather than using server-side include mechanisms to build Web pages on the content servers and then propagating these client-specific, noncacheable, pages through the entire network,

ESI mechanisms permit pages to be assembled from cached content at the edge of the network, thereby reducing bandwidth consumption and decreasing response time.

ESI mechanisms are crucial in this third-phase scenario, where each client receives a personalized home page from the Web site. The building blocks of these pages are retrieved from a series of WebSphere Application Servers. Application servers containing sensitive business logic and ties to secure databases are isolated behind a firewall.

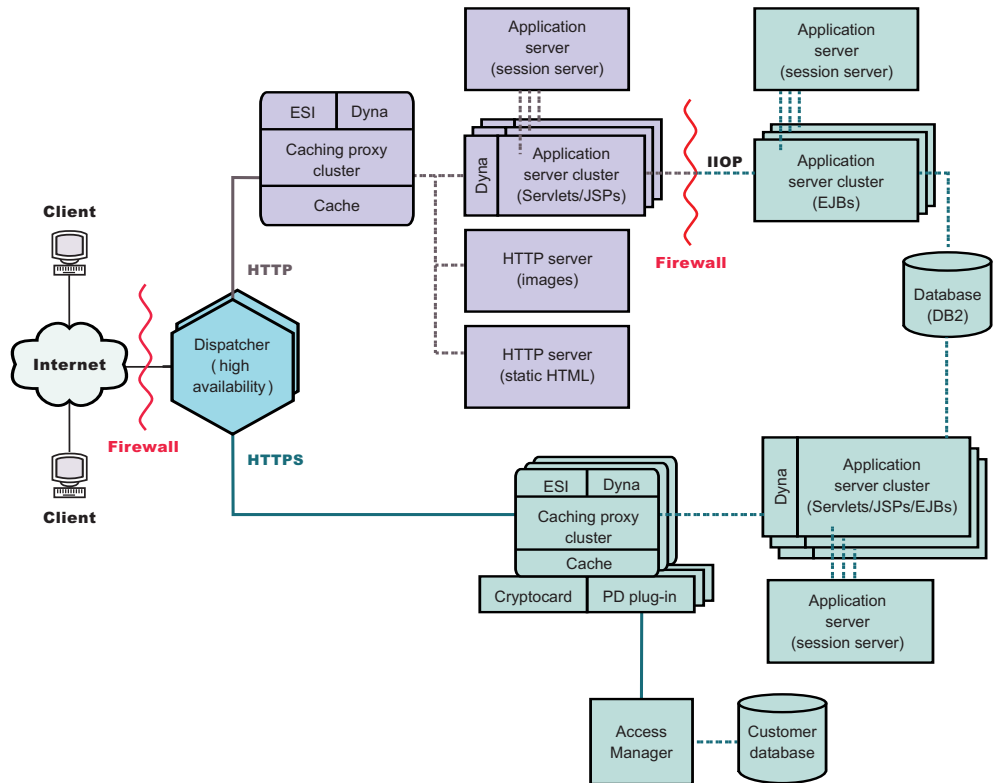


Figure 16. Business to consumer network (Phase 3)

---

## Chapter 9. Business to client-banking-solution

Figure 17 on page 44 shows an efficient online-banking solution that is similar to the business-to-consumer network described in Chapter 8, “Business-to-consumer network,” on page 39. All client requests pass through the firewall to a Dispatcher that separates traffic according to Internet protocol. HTTP requests pass to a cluster of proxy servers with active caches that act as surrogate servers for the Web servers. Metric servers are colocated with the proxy servers to provide load-balancing data to the Dispatcher. This arrangement reduces the network load on the Web servers and creates an additional buffer between them and the Internet.

HTTPS requests are passed to a secure network designed to provide clients with personal financial information and permit online-banking transactions. A cluster of enhanced proxy servers provides scalability to the site. These proxy servers support the caching of dynamic Web content and assembly of page fragments written to comply with the Edge Side Includes (ESI) protocol. A cryptographic hardware card manages SSL handshakes, which significantly reduces the processing required of the proxy server host, and an Access Manager (formerly Policy Director) directs client authentication.

A collection of application servers clusters distribute the processing of requests by separating the business logic, contained in EJB components, from these presentation layer, contained in servlets and JSP files. Each of these clusters is managed by a separate session server.

The following scenerio includes both Load Balancer and Caching Proxy.

**IMPORTANT:** Caching Proxy is available on all Edge component installations, with the following exceptions:

- Caching Proxy is not available for Edge component installations that run on Itanium 2 or AMD Opteron 64-bit processors.
- Caching Proxy is not available for Edge component installations of Load Balancer for IPv4 and IPv6.

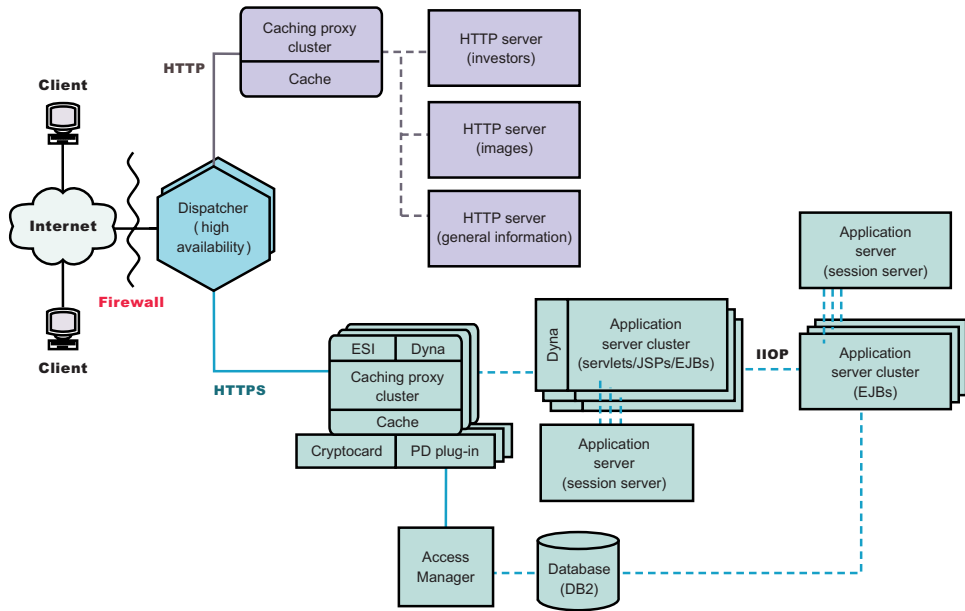


Figure 17. Business to consumer banking solution

---

## Chapter 10. Web portal network

Figure 18 on page 46 shows a Web portal network designed to support a heavy volume of traffic while providing each client with personalized content. To minimize the processing load on the various servers, no part of the network carries SSL traffic. Because the portal does not deliver sensitive data, security is not an important issue. It is important for the databases containing client IDs, passwords, and settings to remain moderately secure and uncorrupted, but this requirement does not impair the performance of the rest of the Web site.

All client requests pass through the firewall to a Dispatcher that balances the network load across a cluster of proxy servers with active caches that act as surrogate servers for the Web servers. Metric servers are collocated with the proxy servers to provide load-balancing data to the Dispatcher.

The actual dynamic Web site is a cluster of application servers that generate ESI fragments that are passed to the proxy servers for assembly. Because of the reduced security concerns, each application server performs all necessary functions for constructing the Web site. All application servers are identical. If one application server goes out of service, the session server can route requests to the other servers, providing high availability for the entire site. This configuration also allows for rapid escalation of the Web site if excessive traffic occurs, for example, the hosting of a special event by the portal. Additional proxy servers and application servers can quickly be configured into the site.

All static content, such as image files and boilerplate text is stored on separate Web servers, allowing it to be updated as necessary without risking corruption to the more complex application servers.

The following scenerio includes both Load Balancer and Caching Proxy.

**IMPORTANT:** Caching Proxy is available on all Edge component installations, with the following exceptions:

- Caching Proxy is not available for Edge component installations that run on Itanium 2 or AMD Opteron 64-bit processors.
- Caching Proxy is not available for Edge component installations of Load Balancer for IPv4 and IPv6.

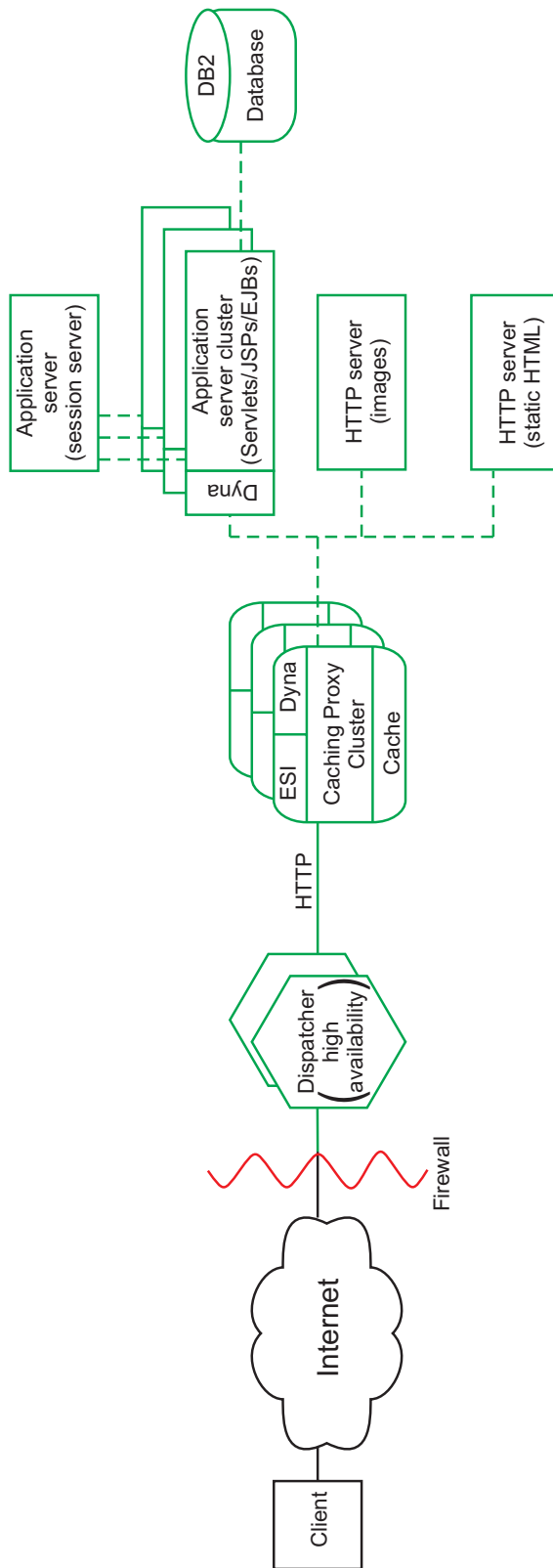


Figure 18. Web portal



---

## Part 4. Installing Edge Components

This part provides the procedures for installing Edge Components.

This part contains the following chapters:

- Chapter 11, “Requirements for Edge Components,” on page 49
- Chapter 12, “Installing Edge Components with IBM Installation Manager,” on page 51
  - “Installing IBM Installation Manager” on page 55
  - “Installing Edge Components with the IBM Installation Manager GUI” on page 58
  - “Installing Edge Components silently with IBM Installation Manager” on page 61
- Chapter 13, “Installing fix packs to Edge Components,” on page 67
- Chapter 14, “Uninstalling fix packs to Edge Components with IBM Installation Manager,” on page 71
- “Uninstalling Edge Components with the IBM Installation Manager GUI” on page 73
- “Uninstalling Edge Components silently with IBM Installation Manager” on page 73



---

## Chapter 11. Requirements for Edge Components

This chapter provides a link to hardware and software requirements for Edge Components and describes guidelines for using Web browsers with the Caching Proxy Configuration and Administration forms and with the Load Balancer online help.

For information on supported hardware and software requirements for WebSphere Application Server, Version 8.5 Edge Components, link to the following Web page: <http://www.ibm.com/support/docview.wss?rs=180&uid=swg27006921>.

SDK installation: The Java SDK automatically installs with Load Balancer on all platforms.



---

## Chapter 12. Installing Edge Components with IBM Installation Manager

IBM Installation Manager is a common installer for many IBM software products that you use to install this version of Edge Components.

**New Feature:** Edge Components Version 8.0 is the first full version of the product to be installed by Installation Manager rather than the programs that are based on InstallShield MultiPlatform (ISMP) or native packages that are used to install, update, and uninstall previous versions. Installation Manager is a single installation program that can use remote or local software flat-file repositories to install, modify, or update the products. It determines and shows available packages — including products, fix packs, interim fixes, etc. — checks prerequisites and interdependencies, and installs the selected packages. You also use Installation Manager to easily uninstall the packages that it installed.

### Overview of IBM Installation Manager

IBM Installation Manager is a general-purpose software installation and update tool that runs on a range of computer systems. Installation Manager can be invoked through a graphical user interface (GUI) or a command-line interface. You can also create response files in XML and use them to direct the performance of Installation Manager tasks in silent mode. For more information on using Installation Manager, read the Installation Manager Information Center.

### Packages and package groups

Each software product that can be installed with Installation Manager is referred to as a "package." An installed package has a product level and an installation location. A package group consists of all of the products that are installed at a single location.

### Installation Manager modes

IBM Installation Manager can be installed in one of the following three modes:

- In admin mode, the Installation Manager is installed from an administrator or a root ID and can be invoked by any administrator or root user. Load Balancer for IPv4 requires you to install and run the program in admin or root mode for the processes to run.
- In nonAdmin mode (also called "user mode"), only the user that installed Installation Manager can invoke it.

### How many Installation Managers do you need

You only need to run Installation Manager on those systems on which you install or update product code. You normally need only one Installation Manager on a system because one Installation Manager can keep track of any number of product installations.

### Getting the Installation Manager installation kit

IBM Installation Manager comes in the form of an installation kit, which contains a set of Installation Manager binaries and a flat-file repository for the Installation Manager product. The installation kit is only used for setup and maintenance of the Installation Manager.

### Installing Installation Manager

When the installation kit is available on your system, you can install

Installation Manager. Installation Manager consists of a set of binaries that are copied from the installation kit and a set of runtime data that describe the products that have been installed by this particular Installation Manager. Before installing Installation Manager, you must decide in which mode the Installation Manager will run as well as where the binaries and runtime data — called "agent data" or "appdata" — will reside. Then, you issue the Installation Manager `installc`, `userinstc`, or `groupinstc` command from the appropriate user ID to install Installation Manager.

### Accessing product repositories

All software materials that will be installed with IBM Installation Manager are stored in flat-file repositories. Each repository contains program objects and metadata for one or more packages—that is, software products at a particular level. Repositories can also contain product maintenance, such as fix packs and ifixes. Whenever you install a new product, you can choose from any of the available product levels in any accessible repository.

### Installing the product

After you have installed Installation Manager and have access to all necessary product repositories, you can use Installation Manager command-line commands or response files to perform the actual product installations. When you install a product, you provide the package name, optionally the product level to be installed, the product location, and any other optional properties. For example, some products have optional features that you can select at installation time or a list of optional supported language packs from which you can select.

### Working with installed products

You can use Installation Manager commands to list installed products and product levels. You can also obtain this information for installed copies of the products by issuing the `versionInfo` command from the product file system. You can use Installation Manager commands or response files to install a new product level, roll back to a previous level, or modify the product by adding or removing optional features or language packs.

### Using the IBM Packaging Utility

With the Packaging Utility, you can create and manage packages for installation repositories. You can copy multiple packages into one repository or copy multiple disks for one product into a repository. You can copy packages from Passport Advantage into a repository, for example. For more information on the Packaging Utility, go to the IBM Installation Manager Version 1.4 Information Center or the IBM Installation Manager Version 1.5 Information Center.

### Restrictions

- Do not use the same response files that are used with other versions of the product to install or uninstall Version 8.0 silently. Use response files that are based on Installation Manager to install, update, or uninstall Version 8.0 and later.
- **Solaris** The Installation Manager GUI is not supported on Solaris 10 x64 systems. Perform the following actions to install or uninstall the product on these systems:
  1. Use the Installation Manager GUI on a supported system to record a response file that will allow you to install or uninstall the product silently.
  2. Edit the recorded response file if necessary.

3. Use the response file to install or uninstall the product silently on your system.

- **Linux** For any Linux system that is enabled for Security Enhanced Linux (SELinux), such as Red Hat Enterprise Linux Version 5 or SUSE Linux Enterprise Server Version 11, you must identify the Java shared libraries in the Installation Manager 1.4.2 or later installation image to the system. Also, you must identify the Java shared libraries in the Installation Manager 1.4.2 or later installation after it has been installed. For example:

```
chcon -R -t texrel_shlib_t ${IM_Image}/jre_5.0.3.sr8a_20080811b/jre/bin
chcon -R -t texrel_shlib_t
    ${IM_Install_root}/eclipse/jre_5.0.3.sr8a_20080811b/jre/bin
```

- **Windows** If a non-administrator installs the product on a Windows Vista, Windows 7, or Windows Server 2008 operating system into the Program Files or Program Files (x86) directory with User Account Control (UAC) enabled, the product might not function correctly. UAC is an access-control mechanism that allows non-administrative users to install a software product into the Program Files or Program Files (x86) directory; but it then prohibits any write access to that directory after the installation has completed. The product requires write access in the main installation directory to function correctly. To resolve this issue, perform one of the following actions:
  - Install the product into a directory other than Program Files or Program Files (x86). For example:  
C:\IBM\Edge
  - Disable UAC.
- When you install the product using Installation Manager with local repositories, the installation takes a significantly longer amount of time if you use the repository.zip file directly without extracting it. Before you install the product using Installation Manager local repositories, extract the repository.zip file to a location on your local system and add that location to your Installation Manager preferences.

**Note:** Do not transfer the content of a repository in non-binary mode and do not convert any content on extraction.

**New Feature:** In addition to the GUI and silent methods described in this information, you can also use Installation Manager to manage installation using the Installation Manager imcl installation command. For information on using Installation Manager using this method, read the Installation Manager Information Center.

Perform the following tasks:

1. Ensure that any previous versions of the product are uninstalled. You must uninstall previous versions of Edge Components through the native installation program or system package tools.  
Read “Uninstalling previous versions of Edge Components” on page 75 for instructions.
2. Perform the tasks in “Installing IBM Installation Manager” on page 55, which includes adding the product repositories for the Edge Components products.
3. Install the product:
  - “Installing Edge Components with the IBM Installation Manager GUI” on page 58

- “Installing Edge Components silently with IBM Installation Manager” on page 61
4. To update your products, read Chapter 13, “Installing fix packs to Edge Components,” on page 67.

If you need to uninstall the product through IBM Installation Manager, perform one of the following tasks:

- “Uninstalling Edge Components with the IBM Installation Manager GUI” on page 73
- “Uninstalling Edge Components silently with IBM Installation Manager” on page 73

After installing the product, be aware of the following notes for logging, tracing, and troubleshooting:

#### Notes on logging and tracing

- An easy way to view the logs is to open Installation Manager and go to **File > View Log**. An individual log file can be opened by selecting it in the table and then clicking the Open log file icon.
- Logs are located in the logs directory of Installation Manager's application data location. For example:
  - **Windows Administrative installation:**  
C:\Documents and Settings\All Users\Application Data\IBM\Installation Manager
  - **Windows Non-administrative installation:**  
C:\Documents and Settings\user\_name\Application Data\IBM\Installation Manager
  - **AIX HP-UX Linux Solaris Administrative installation:**  
/opt/IBM/InstallationManager
  - **AIX HP-UX Linux Solaris Non-administrative installation:**  
user\_home/var/ibm/InstallationManager
- The main log files are time-stamped XML files in the logs directory, and they can be viewed using any standard web browser.
- The log.properties file in the logs directory specifies the level of logging or tracing that Installation Manager uses. To turn on tracing for the Edge component plug-ins, for example, create a log.properties file with the following content:
 

```
com.ibm.ws=DEBUG
com.ibm.cic.agent.core.Engine=DEBUG
global=DEBUG
```

Restart Installation Manager as necessary, and Installation Manager outputs traces for the Edge component plug-ins.

#### Notes on troubleshooting

- **HP-UX** By default, some HP-UX systems are configured to not use DNS to resolve host names. This could result in Installation Manager not being able to connect to an external repository. You can ping the repository, but nslookup does not return anything.  
Work with your system administrator to configure your machine to use DNS, or use the IP address of the repository.
- In some cases, you might need to bypass existing checking mechanisms in Installation Manager.



- On some network file systems, disk space might not be reported correctly at times, and you might need to bypass disk-space checking and proceed with your installation.

To disable disk-space checking, specify the following system property in the config.ini file in *IM\_install\_root/eclipse/configuration* and restart Installation Manager:

```
cic.override.disk.space=sizeunit
```

where *size* is a positive integer and unit is blank for bytes, k for kilo, m for megabytes, or g for gigabytes. For example:

```
cic.override.disk.space=120 (120 bytes)
cic.override.disk.space=130k (130 kilobytes)
cic.override.disk.space=140m (140 megabytes)
cic.override.disk.space=150g (150 gigabytes)
cic.override.disk.space=true
```

Installation Manager will report a disk-space size of Long.MAX\_VALUE. Instead of displaying a very large amount of available disk space, N/A is displayed.

- To bypass operating-system prerequisite checking, add the following line to the to the config.ini file in *IM\_install\_root/eclipse/configuration* directory:

```
disableOSPrereqChecking=true
```

Then restart Installation Manager.






If you need to use any of these bypass methods, contact IBM Support for assistance in developing a solution that does not involve bypassing the Installation Manager checking mechanisms.

- Installation Manager might display a warning message during the uninstallation process.

Uninstalling the product using Installation Manager requires that the data repositories remain valid and available.

- For more information on using Installation Manager, read the IBM Installation Manager Information Center. Read the release notes to learn more about the latest version of Installation Manager.

To access the release notes, complete the following task:

-  Click **Start > Programs > IBM Installation Manager > Release Notes**.
-     Go to the documentation subdirectory in the directory where Installation Manager is installed, and open the readme.html file.

#### Note on version and history information

The versionInfo and historyInfo commands return version and history information based on all of the installation, uninstallation, update, and roll back activities performed on the system.

---

## Installing IBM Installation Manager

Before you install any Edge Components products, you need to install IBM Installation Manager.

1. Obtain the necessary files from the physical media or the web. There are three basic options for obtaining and installing Installation Manager and the product.

- **Access the physical media, and use local installation:** You can access Installation Manager and the product repositories on the product media. Install Installation Manager on your system, and use it to install the product from the product repositories on the media.
  - **Download the files from the IBM Passport Advantage site, and use local installation:** Licensed customers can download Installation Manager as well as the necessary product repositories from the Passport Advantage site. You can then install Installation Manager on your system and use it to install the product from the repositories.
  - **Download a file from the Installation Manager website, and use web-based installation:** You can download and unpack a compressed file containing Installation Manager from the IBM Installation Manager website. You can then install Installation Manager on your local system and use it to install the product from the web-based repository.
2. Install the product through the GUI or “silently.”
- To install through the GUI:
    - a. Change to the location containing the Installation Manager installation files, and run one of the following commands:

#### Administrative installation

- **Windows**  
install.exe
- **AIX** **HP-UX** **Linux** **Solaris**  
./install

#### Non-administrative installation

- **Windows**  
userinst.exe
- **AIX** **HP-UX** **Linux** **Solaris**  
./userinst

- b. Make sure that the Installation Manager package is selected, and click **Next**.
  - c. Accept the terms in the license agreements, and click **Next**.  
The program creates the directory for your installation.
  - d. Click **Next**.
  - e. Review the summary information, and click **Install**.
    - If the installation is successful, the program displays a message indicating that installation is successful.
    - If the installation is not successful, click **View Log File** to troubleshoot the problem.
- To install silently:

#### Administrative installation

- **Windows**  
installc.exe -acceptLicense -log *log\_file\_path\_and\_name*
- **AIX** **HP-UX** **Linux** **Solaris**  
./installc -acceptLicense -log *log\_file\_path\_and\_name*

#### Non-administrative installation

- **Windows**  
userinstc.exe -acceptLicense -log *log\_file\_path\_and\_name*
- **AIX** **HP-UX** **Linux** **Solaris**

```
./userinstc -acceptLicense -log log_file_path_and_name
```

3. Add the product repository to your Installation Manager preferences. There are three ways to access repositories for your product:

- **Access the physical media, and use local installation**

You can access the product repositories on the product media. Use your existing Installation Manager to install the product from the product repositories on the media.

- **Download the files from the Passport Advantage site, and use local installation**

Licensed customers can download the necessary product repositories from the Passport Advantage site. You can then use your existing Installation Manager to install the product from the repositories.

- **Access the live repositories, and use web-based installation**

You can install Installation Manager on your local system and use it to install the product from the web-based repositories.

**Note:** This live repository is accessed by using Passport Advantage authentication. After you install Installation Manager, you can set the Passport Advantage preference to connect to the live repositories. To set Passport Advantage preferences, follow this procedure:

- a. Open Installation Manager.
- b. Open the Passport Advantage preferences page by selecting **File > Preferences > Passport Advantage**.
- c. Select **Connect to Passport Advantage** to connect to the Passport Advantage repository. The Password Required dialog box opens.
- d. Enter a user name and password for Passport Advantage.
- e. **Optional:** Select **Save password** to save the user name and password credentials. If you do not save the user name and password credentials, you are prompted for these credentials each time you access Passport Advantage.
- f. Click **OK** to close the Password Required dialog box.
- g. Click **OK** to close the Preferences window.

For more information on setting your Installation Manager preferences, see the Installation Manager Information Center.

To add the repository:

- a. Start Installation Manager.
- b. In the top menu, click **File > Preferences**.
- c. Select **Repositories**.
- d. Click **Add Repository**.
- e. Enter the path to the repository.config file in the location containing the repository files.
  - This is a sample path for a local repository:

**Windows**

```
C:\repositories\product_name\local-repositories
```

**AIX**

**HP-UX**

**Linux**

**Solaris**

```
/var/repositories/product_name/local-repositories
```

- This is the path for the remote repository:

```
http://www.ibm.com/software/repositorymanager/V8WASNDEdge
```

**Note:** Be aware of the following conditions:

- You must have Administrator or root access to install and use Load Balancer for IPv4 and Caching Proxy.
- Do not install Caching Proxy on 64-bit operating systems.
- **Windows** The default paths for Caching Proxy, Version 7 and earlier are different than the default paths for the current version. Update the local path in your configuration files to the new path after you migrate to the new version of Caching Proxy.
- **Windows** Load Balancer for IPv4 is not supported on Windows 7 or Windows Vista.
- **Windows** Before you install Load Balancer for IPv4 on Windows 2003 64-bit machines, you must install the required Microsoft hotfix. If you do not install the Microsoft hotfix prior to installing Load Balancer for IPv4, Load Balancer does not run correctly.

To correct this problem, you must install the required Microsoft hotfix for KB942589. After you apply the Microsoft hotfix, restart your system. You must restart your system before you install Load Balancer for IPv4 on Windows 2003 64-bit machines. You will need to restart your system again after completing the installation of Load Balancer for IPv4.

If you do not install the Microsoft hotfix before you install Load Balancer for IPv4 and you encounter this problem, you must uninstall Load Balancer for IPv4, install the Microsoft hotfix, and then reinstall Load Balancer for IPv4.

- f. Click **OK**.
  - g. Deselect any locations listed in the Repositories window that you will not be using.
  - h. Click **Apply**.
  - i. Click **OK**.
  - j. Click **File > Exit** to close Installation Manager.
4. Install the Edge Components product:
- “Installing Edge Components with the IBM Installation Manager GUI”
  - “Installing Edge Components silently with IBM Installation Manager” on page 61

---

## Installing Edge Components with the IBM Installation Manager GUI

Use the Installation Manager GUI to install Edge Components products.

**Note:** Be aware of the following conditions when you install this product:

- You must have Administrator or root access to install and use Load Balancer for IPv4 and Caching Proxy.
  - Do not install Caching Proxy on 64-bit operating systems.
  - **Windows** Load Balancer for IPv4 is not supported on Windows 7 or Windows Vista.
1. “Installing IBM Installation Manager” on page 55 if it is not installed. These instructions also include steps to add repository information for the Edge Components products.
  2. Start Installation Manager.

3. Click **Install**.

**Note:** If you are prompted to authenticate, use the Passport Advantage credentials that you registered with on the program website. Installation Manager searches its defined repositories for available packages.

4. Select the product you want to install. For example, select **IBM WebSphere Edge Components: Caching Proxy** or **IBM WebSphere Edge Components: Load Balancer for IPv4**.

If you already have the product installed on your system, a message displays indicating that the product is already installed.

5. Click **Next**.

**Note:** If you try to install a newer level of the product with a previous version of Installation Manager, Installation Manager might prompt you to update to the latest level of Installation Manager when it connects to the repository. Update to the newer version before you continue if you are prompted to do so. Read the Installation Manager Information Center for information about automatic updates.

6. Accept the terms in the license agreements, and click **Next**.

7. Specify the installation root directory for the product binaries, which are also referred to as the core product files or system files.

The panel also displays the shared resources directory and disk-space information.

**Note:** The first time that you install a package using Installation Manager, specify the shared resources directory. The shared resources directory is where installation artifacts are located that can be used by one or more package groups. Use your largest drive for this installation. You cannot change the directory location until after you uninstall all packages.

**Restrictions:**

- Installation manager requires that you install the product to an empty directory. The default installation paths are:

**Caching Proxy**

– **AIX** **HP-UX** **Linux** **Solaris**  
/opt/ibm/edge/

This cannot be modified.

– **Windows**  
C:\Program Files\IBM\edge\cachingproxy\

**Load Balancer for IPv4**

– **AIX** **HP-UX** **Linux** **Solaris**  
/opt/ibm/edge/lb/

This cannot be modified.

– **Windows**  
C:\Program Files\loadbalancer\IBM\edge\lb\

The beginning of this installation path can be changed, but be aware that Installation Manager requires you to install the product into an empty directory - in the default path it is

C:\Program Files\loadbalancer. The IBM\edge\lb portion of the path is hard coded and cannot be modified.

- Deleting the default target location and leaving an installation-directory field empty prevents you from continuing.
- Do not use symbolic links as the destination directory. Symbolic links are not supported.
- Do not use a semicolon in the directory name. The product cannot install properly if the target directory includes a semicolon.
- **Windows** A semicolon is the character used to construct the class path on Windows systems.
- **Windows** The maximum path length on the Windows Server 2008, Windows Vista, and Windows 7 operating systems is 60 characters.
- **Windows** Before you install Load Balancer for IPv4 on Windows 2003 64-bit machines, you must install the required Microsoft hotfix. If you do not install the Microsoft hotfix prior to installing Load Balancer for IPv4, Load Balancer does not run correctly.

To correct this problem, you must install the required Microsoft hotfix for KB942589. After you apply the Microsoft hotfix, restart your system. You must restart your system before you install Load Balancer for IPv4 on Windows 2003 64-bit machines. You will need to restart your system again after completing the installation of Load Balancer for IPv4.

If you do not install the Microsoft hotfix before you install Load Balancer for IPv4 and you encounter this problem, you must uninstall Load Balancer for IPv4, install the Microsoft hotfix, and then reinstall Load Balancer for IPv4.

8. Click **Next**.
  9. Select the languages for which translated content should be installed. English is always selected.
  10. Click **Next**.
  11. Select the features that you want to install.
  12. Click **Next**.
  13. Review the summary information, and click **Install**.
    - If the installation is successful, the program displays a message indicating that installation is successful.
- Note:** The program might also display important post-installation instructions as well.
- If the installation is not successful, click **View Log File** to troubleshoot the problem.
14. Click **Finish**.
  15. Click **File > Exit** to close Installation Manager.
  16. **Windows** **Optional:** If you are migrating from a previous version of Caching Proxy, update the configuration files with the new installation path. The default paths for Caching Proxy Version 7 and earlier are different than the default paths for the current version.

**Tip:** Installation Manager optionally can search for updates to itself whenever the Install Packages, Modify Packages, or Update Packages page is opened from the Start page as well as when clicking **Check for Other Versions and Extensions** on the Install Packages page. To enable this option, perform the following actions:

1. Start Installation Manager.

2. In the top menu, click **File > Preferences**.
3. Select **Updates**.
4. Select **Search for Installation Manager** updates.
5. Click **Apply**.
6. Click **OK**.

Do not enable this option if you do not have access to the service repository.

---

## Installing Edge Components silently with IBM Installation Manager

**Note:** Be aware of the following conditions when you install this product:

- You must have Administrator or root access to install and use Load Balancer for IPv4 and Caching Proxy.
  - Do not install Caching Proxy on 64-bit operating systems.
  - **Windows** Load Balancer for IPv4 is not supported on Windows 7 or Windows Vista.
1. “Installing IBM Installation Manager” on page 55 if it is not installed. These instructions also include steps to add repository information for the Edge Components products.
  2. **Optional:** Record a response file to install the product. If you already have a response file, proceed directly to Step Use the response files to install the product silently. on page 64.

On one of your systems, perform the following actions to record a response file that will install the product:

- a. From a command line, change to the eclipse subdirectory in the directory where you installed Installation Manager.
- b. Start Installation Manager from the command line using the `-record` option. For example:

- **Windows Administrator or non-administrator:**

```
IBMIM.exe -skipInstall "C:\temp\imRegistry"
-record C:\temp\install_response_file.xml
```

- **AIX HP-UX Linux Solaris Administrator:**

```
./IBMIM -skipInstall /var/temp/imRegistry
-record /var/temp/install_response_file.xml
```

- **AIX HP-UX Linux Solaris Non-administrator:**

```
./IBMIM -skipInstall user_home/var/temp/imRegistry
-record user_home/var/temp/install_response_file.xml
```

**Tip:** When you record a new response file, you can specify the `-skipInstall` parameter. Using this parameter has the following benefits:

- No files are actually installed, and this speeds up the recording.
- If you use a temporary data location with the `-skipInstall` parameter, Installation Manager writes the installation registry to the specified data location while recording. When you start Installation Manager again without the `-skipInstall` parameter, you then can use your response file to install against the real installation registry.
- The `-skipInstall` operation should not be used on the actual agent data location used by Installation Manager. This is unsupported. Use a clean writable location, and re-use that location for future recording sessions

For more information, read the Installation Manager Information Center.



- c. Ensure that you added the repository information for your product in Step 1.
- d. Click **Install**.

**Note:** If you are prompted to authenticate, use the Passport Advantage credentials that you registered with on the program website. Installation Manager searches its defined repositories for available packages.

Installation Manager searches its defined repositories for available packages.

- e. Select your product and the appropriate version. For example, select **IBM WebSphere Edge Components: Caching Proxy** or **IBM WebSphere Edge Components: Load Balancer for IPv4**.

If you already have the product installed on a WebSphere Application Server installation on your system, a message displays indicating that the product is already installed. To create another installation of the product in another location, click **Continue**.

- f. Click **Next**.
- g. Accept the terms in the license agreements, and click **Next**.
- h. Specify the installation root directory for the product binaries, which are also referred to as the core product files or system files.

The panel also displays the shared resources directory and disk-space information.

**Note:** The first time that you install a package using Installation Manager, specify the shared resources directory. The shared resources directory is where installation artifacts are located that can be used by one or more package groups. Use your largest drive for this installation. You cannot change the directory location until after you uninstall all packages.

**Restrictions:**

- Installation manager requires that you install the product to an empty directory. The default installation paths are:

**Caching Proxy**

– **AIX** **HP-UX** **Linux** **Solaris**  
/opt/ibm/edge/

This cannot be modified.

– **Windows**  
C:\Program Files\IBM\edge\cachingproxy\

**Load Balancer for IPv4**

– **AIX** **HP-UX** **Linux** **Solaris**  
/opt/ibm/edge/lb/

This cannot be modified.

– **Windows**  
C:\Program Files\loadbalancer\IBM\edge\lb\

The beginning of this installation path can be changed, but be aware that Installation Manager requires you to install the product into an empty directory - in the default path it is



**C:\Program Files\loadbalancer.** The **IBM\edge\lb** portion of the path is hard coded and cannot be modified.

- Deleting the default target location and leaving an installation-directory field empty prevents you from continuing.
- Do not use symbolic links as the destination directory. Symbolic links are not supported.
- Do not use a semicolon in the directory name. The product cannot install properly if the target directory includes a semicolon.
- **Windows** A semicolon is the character used to construct the class path on Windows systems.
- **Windows** The maximum path length on the Windows Server 2008, Windows Vista, and Windows 7 operating systems is 60 characters.
- **Windows** Before you install Load Balancer for IPv4 on Windows 2003 64-bit machines, you must install the required Microsoft hotfix. If you do not install the Microsoft hotfix prior to installing Load Balancer for IPv4, Load Balancer does not run correctly.

To correct this problem, you must install the required Microsoft hotfix for KB942589. After you apply the Microsoft hotfix, restart your system. You must restart your system before you install Load Balancer for IPv4 on Windows 2003 64-bit machines. You will need to restart your system again after completing the installation of Load Balancer for IPv4.

If you do not install the Microsoft hotfix before you install Load Balancer for IPv4 and you encounter this problem, you must uninstall Load Balancer for IPv4, install the Microsoft hotfix, and then reinstall Load Balancer for IPv4.

- Click **Next**.
- Select the languages for which translated content should be installed. English is always selected.
- Click **Next**.
- Select the features that you want to install.
- Click **Next**.
- Review the summary information, and click **Install**.
  - If the installation is successful, the program displays a message indicating that installation is successful.

**Note:** The program might also display important post-installation instructions as well.

- If the installation is not successful, click **View Log File** to troubleshoot the problem.
- Click **Finish**.
- Click **File > Exit** to close Installation Manager.
- Optional:** If you are using an authenticated remote repository, create a keyring file for silent installation.

**Note:** In a keyring file, you can store credentials for URLs that require authentication, such as your remote repositories.

- 1) From a command line, change to the eclipse subdirectory in the directory where you installed Installation Manager.
- 2) Start Installation Manager from the command line using the **-record** option.

For example:

- **Windows Administrator or non-administrator:**  
IBMIM.exe -skipInstall "C:\temp\imRegistry"  
-keyring C:\IM\im.keyring  
-record C:\temp\keyring\_response\_file.xml
- **AIX HP-UX Linux Solaris Administrator:**  
./IBMIM -skipInstall /var/temp/imRegistry  
-keyring /var/IM/im.keyring  
-record /var/temp/keyring\_response\_file.xml
- **AIX HP-UX Linux Solaris Non-administrator:**  
./IBMIM -skipInstall user\_home/var/temp/imRegistry  
-keyring user\_home/var/IM/im.keyring  
-record user\_home/var/temp/keyring\_response\_file.xml

3) When a window opens that requests your credentials for the authenticated remote repository, enter the correct credentials and save them.

4) Click **File > Exit** to close Installation Manager.

For more information, read the Installation Manager Information Center.

3. Use the response files to install the product silently.
- a. **Optional:** Use the response file to install the keyring file silently: Go to a command line on each of the systems on which you want to install the product, change to the eclipse/tools subdirectory in the directory where you installed Installation Manager, and install the keyring file silently.

For example:

- **Windows Administrator or non-administrator:**  
imcl.exe -acceptLicense  
input C:\temp\keyring\_response\_file.xml  
-log C:\temp\keyring\_log.xml
- **AIX HP-UX Linux Solaris Administrator:**  
./imcl -acceptLicense  
input /var/temp/keyring\_response\_file.xml  
-log /var/temp/keyring\_log.xml
- **AIX HP-UX Linux Solaris Non-administrator:**  
./imcl -acceptLicense  
input user\_home/var/temp/keyring\_response\_file.xml  
-log user\_home/var/temp/keyring\_log.xml

For more information on keyring files, read the Installation Manager Information Center.

- b. Use the response file to install the product silently. Go to a command line on each of the systems on which you want to install the product, change to the eclipse/tools subdirectory in the directory where you installed Installation Manager, and install the product silently.

For example:

- **Windows Administrator or non-administrator:**  
imcl.exe -acceptLicense  
input C:\temp\install\_response\_file.xml  
-log C:\temp\install\_log.xml  
-keyring C:\IM\im.keyring
- **AIX HP-UX Linux Solaris Administrator:**  
./imcl -acceptLicense  
input /var/temp/install\_response\_file.xml  
-log /var/temp/install\_log.xml  
-keyring /var/IM/im.keyring

- **AIX** **HP-UX** **Linux** **Solaris** **Non-administrator:**

```
./imcl -acceptLicense
input user_home/var/temp/install_response_file.xml
-log user_home/var/temp/install_log.xml
-keyring user_home/var/IM/im.keyring
```

#### Notes:

- The relevant terms and conditions, notices, and other information are provided in the license-agreement files in the `lfiles` or `product_name/lfiles` subdirectory of the installation image or repository for this product.
  - The program might write important post-installation instructions to standard output.
4. **Windows** **Optional:** If you are migrating from a previous version of Caching Proxy, update the configuration files with the new installation path. The default paths for Caching Proxy Version 7 and earlier are different than the default paths for the current version.

**Windows** **Example:** The following is a sample response file for installing Load Balancer for IPv4:

```
<?xml version="1.0" encoding="UTF-8"?>
<agent-input acceptLicense='true'>
<server>
<repository location='C:\repositories\edge80\local-repositories' />
</server>
<profile id='Load Balancer for IPV4 v8.0'
installLocation='c:\program files\loadbalancer\ibm\edge\lb'>
<data key='eclipseLocation' value=c:\program files\loadbalancer\ibm\edge\lb' />
<data key='user.import.profile' value='false' />
<data key='cic.selector.nl' value='cs,ja,es,zh_HK,en,zh_TW,pt_BR,it' />
</profile>
<install modify='false'>
<offering id='com.ibm.websphere.EDGEIPV4.v80' version='8.0.0.0'
profile='IBM WebSphere Edge Components: Load Balancer for IPV4 V8.0'
features='Dispatcher,Load Balancer,Metric Server' />
</install>
<preference name='com.ibm.cic.common.core.preferences.eclipseCache'
value='C:\Program Files\IBM\IMShared' />
<preference value='30' name='com.ibm.cic.common.core.preferences.connectTimeout' />
<preference value='30' name='com.ibm.cic.common.core.preferences.readTimeout' />
<preference value='0' name='com.ibm.cic.common.core.preferences.downloadAutoRetryCount' />
<preference value='false' name='offering.service.repositories.areUsed' />
<preference value='false' name='com.ibm.cic.common.core.preferences.ssl.nonsecureMode' />
<preference value='false'
name='com.ibm.cic.common.core.preferences.http.disablePreemptiveAuthentication' />
<preference value='true'
name='com.ibm.cic.common.core.preferences.preserveDownloadedArtifacts' />
<preference value='false' name='PassportAdvantageIsEnabled' />
<preference value='false'
name='com.ibm.cic.common.core.preferences.searchForUpdates' />
</agent-input>
```

#### Tips:

- To disable remote searches for updates in the response file, set the following preferences to false:

##### **offering.service.repositories.areUsed**

Used for searching remote repositories for updates to installed offerings.

##### **com.ibm.cic.common.core.preferences.searchForUpdates**

Used for searching for updates to Installation Manager.

For example:

```
<preference value='false' name='offering.service.repositories.areUsed' />  
<preference value='false' name='com.ibm.cic.common.core.preferences.searchForUpdates' />
```

- Installation Manager can save earlier versions of a package to roll back to if you experience issues later. When Installation Manager rolls back a package to a previous version, the current version of the files are uninstalled and the earlier versions are reinstalled.

If you choose not to save the files for rollback, you can prevent the files from being saved by changing the following preference in your response file:

```
<preference name='com.ibm.cic.common.core.preferences.preserveDownloadedArtifacts'  
  value='true' />
```

to this:

```
<preference name='com.ibm.cic.common.core.preferences.preserveDownloadedArtifacts'  
  value='false' />
```

---

## Chapter 13. Installing fix packs to Edge Components

Use one of these methods to update the product:

- “Installing fix packs using the Installation Manager GUI”
- “Installing fix packs with the command line”

---

### Installing fix packs using the Installation Manager GUI

You can use Installation Manager to update this product to a later version.

Make sure that your Installation Manager preferences are pointing to Web-based or local repositories that contain the appropriate updates for the product.

1. Stop all servers and applications on the product installation that is being updated.
2. Start Installation Manager.
3. Click **Update**.
4. Select the package group to update.
5. Click **Next**.

**Note:** If you are prompted to authenticate, use the Passport Advantage credentials that you registered with on the program website.

6. Select the version to which you want to update for your product.
7. Click **Next**.
8. Accept the terms in the license agreements, and click **Next**.
9. Review the summary information, and click **Update**.
  - If the installation is successful, the program displays a message indicating that the installation is successful.
  - If the installation is not successful, click **View Log File** to troubleshoot the problem.
10. Click **Finish**.
11. Click **File > Exit** to close Installation Manager.

---

### Installing fix packs with the command line

Contact the IBM Software Support Center for information about upgrades. The most current information is available from the IBM Software Support Center and Fix Central. IBM Installation Manager is used to apply product maintenance.

1. For a list of interim fixes and fix packs that are available and specific information about each fix, perform the following actions:
  - a. Go to Fix Central.
  - b. Select the product group and version to update.
  - c. Select your operating system as the platform, and click **Continue**.
  - d. Select **Browse for fixes**, and click **Continue**.
  - e. Click **More Information** under each fix to view information about the fix.
  - f. **Recommendation:** Make a note of the name of the fix pack that you would like to install.
2. Update the product with the fix pack using one of the following procedures:

- Access the live service repository that contains the fix pack, and use web-based updating.

Use Installation Manager on your local system to update the product with the fix pack from the live web-based service repositories. See Table 3 in WebSphere Application Server Version 8 product offerings for supported operating systems for a list of web-based service-repository URLs for the product offerings.

To install a fix pack from a service repository, perform the following actions:

- a. Log on to your system.
- b. If you do not already have an Installation Manager credentials file containing your IBM software user ID and password, create one that will allow you to access the repository.

**Note:** These are the credentials that you use to access protected IBM software websites. For information on creating a keyring file for Installation Manager, read Storing credentials in the Installation Manager Information Center.

When creating a keyring file, append `/repository.config` at the end of the URL location if the `imutils` command is unable to find the URL that is specified.

- c. Stop all servers and related applications on the installation that you are updating.
- d. Navigate to the `Installation_Manager_binaries/eclipse/tools` directory, where `Installation_Manager_binaries` is the installation root directory for the Installation Manager.
- e. Install the fix pack.

#### Windows

```
imcl.exe install offering_ID
  -acceptLicense
  -installationDirectory product_installation_location
  -repositories repository_URL
  -keyring keyring_file
```

#### AIX

#### HP-UX

#### Linux

#### Solaris

```
./imcl install offering_ID
  -acceptLicense
  -installationDirectory product_installation_location
  -repositories repository_URL
  -keyring keyring_file
```

where `offering_ID` is the offering ID that is listed in WebSphere Application Server Version 8 product offerings for supported operating systems.

**Tip:** If you are using Installation Manager Version 1.4.4 or later, you can also specify `none`, `recommended` or `all` with the `-installFixes` argument to indicate which interim fixes you want installed with the offering. If the offering version is specified, the `-installFixes` option defaults to `none`. If no offering version is specified, the `-installFixes` option defaults to `all`.

- f. **Optional:** List all installed packages to verify the installation:

#### Windows

```
imcl.exe listInstalledPackages -long
```

#### AIX

#### HP-UX

#### Linux

#### Solaris

```
./imcl listInstalledPackages -long
```

- g. Download the file that contains the fix pack from Fix Central, and use local updating.

You can download a compressed file that contains the fix pack from Fix Central. Each compressed fix-pack file contains an Installation Manager repository for the fix pack and usually has a .zip extension. After downloading the fix-pack file, use Installation Manager to update the product with the fix pack.

- a. Download the fix pack:
  - 1) Go to Fix Central.
  - 2) Select the product group and version to update.
  - 3) Select your operating system as the platform, and click **Continue**.
  - 4) Select **Browse for fixes**, and click **Continue**.
  - 5) Select the fix pack that you want to download, and click **Continue**.
  - 6) Select your download options, and click **Continue**.
  - 7) Click **I agree** to agree to the terms and conditions.
  - 8) Click **Download now** to download the fix pack.
  - 9) Transfer the compressed file in binary format to the system on which it will be installed.
  - 10) **Optional:** You might want to extract the compressed repository files to a directory on your system to improve installation performance.
- a. Install the fix pack:
  - 1) Log on to your system.
  - 2) Stop all servers and related applications on the installation that is being updated.
  - 3) Navigate to the *Installation\_Manager\_binaries/eclipse/tools* directory, where *Installation\_Manager\_binaries* is the installation root directory for the Installation Manager.
  - 4) Install the fix pack:

**Windows**

```
imcl.exe install offering_ID  
-acceptLicense  
-installationDirectory product_installation_location  
-repositories compressed_file_or_location_of_expanded_files
```

**AIX**

**HP-UX**

**Linux**

**Solaris**

```
./imcl install offering_ID  
-acceptLicense  
-installationDirectory product_installation_location  
-repositories compressed_file_or_location_of_expanded_files
```

where *offering\_ID* is the offering ID that is listed in WebSphere Application Server Version 8 product offerings for supported operating systems.

- 5) **Optional:** List all installed packages to verify the installation:

**Windows**

```
imcl.exe listInstalledPackages -long
```

**AIX**

**HP-UX**

**Linux**

**Solaris**

```
./imcl listInstalledPackages -long
```





---

## Chapter 14. Uninstalling fix packs to Edge Components with IBM Installation Manager

You can use Installation Manager to roll back updates to the product, and return to a previous version.

Make sure that your Installation Manager preferences are pointing to Web-based or local repositories that contain the appropriate earlier versions for the product.

1. Stop all servers and applications on the product installation that is being rolled back.
2. Start Installation Manager.
3. Click **Roll Back**.
4. Select the package group to roll back.
5. Click **Next**.

**Note:** If you are prompted to authenticate, use the Passport Advantage credentials that you registered with on the program website.

6. Select the version to which you want to roll back for your product.
7. Click **Next**.
8. Accept the terms in the license agreements, and click **Next**.
9. Review the summary information, and click **Update**.
  - If the roll back is successful, the program displays a message indicating that the roll back is successful.
  - If the roll back is not successful, click **View Log File** to troubleshoot the problem.
10. Click **Finish**.
11. Click **File > Exit** to close Installation Manager.



---

## Chapter 15. Uninstalling Edge Components

- “Uninstalling Edge Components with the IBM Installation Manager GUI”
- “Uninstalling Edge Components silently with IBM Installation Manager”
- “Uninstalling previous versions of Edge Components” on page 75

---

### Uninstalling Edge Components with the IBM Installation Manager GUI

Use the Installation Manager GUI to uninstall the product.

1. Uninstall the product.
  - a. Stop all instances and services for the Edge Components products.
  - b. Start Installation Manager.
  - c. Click **Uninstall**.
  - d. In the Uninstall Packages window, select the package and the appropriate version for the product to uninstall.

For example, select **IBM WebSphere Edge Components: Caching Proxy** or **IBM WebSphere Edge Components: Load Balancer for IPv4**.
  - e. Click **Next**.
  - f. Review the summary information.
  - g. Click **Uninstall**.
    - If the uninstallation is successful, the program displays a message that indicates success.
    - If the uninstallation is not successful, click **View log** to troubleshoot the problem.
  - h. Click **Finish**.
  - i. Click **File > Exit** to close Installation Manager.
2. **Optional:** Uninstall IBM Installation Manager. You must uninstall all of the packages that were installed by Installation Manager before you can uninstall IBM Installation Manager.

Read Uninstalling Installation Manager in the Installation Manager Information Center for information about performing this procedure.

---

### Uninstalling Edge Components silently with IBM Installation Manager

Use Installation Manager to uninstall this product silently.

1. **Optional:** Before you uninstall the product, you can perform or record the installation of Installation Manager and installation of the product to a temporary installation registry on one of your systems so that you can use this temporary registry to record the uninstallation without using the standard registry where Installation Manager is installed.

Read “Installing Edge Components with the IBM Installation Manager GUI” on page 58 or “Installing Edge Components silently with IBM Installation Manager” on page 61 for more information.
2. Stop all instances and services for the Edge Components products.
3. **Optional:** Record a response file to uninstall the product: On one of your systems, perform the following actions to record a response file that will uninstall the product:

- a. From a command line, change to the eclipse subdirectory in the directory in which you installed Installation Manager.
- b. Start Installation Manager from the command line using the `-record` option. For example:

- **Windows Administrator or non-administrator:**  

```
IBMIM.exe -skipInstall "C:\temp\imRegistry"
           -record C:\temp\uninstall_response_file.xml
```
- **AIX HP-UX Linux Solaris Administrator:**  

```
./IBMIM -skipInstall /var/temp/imRegistry
        -record /var/temp/uninstall_response_file.xml
```
- **AIX HP-UX Linux Solaris Non-administrator:**  

```
./IBMIM -skipInstall user_home/var/temp/imRegistry
        -record user_home/var/temp/uninstall_response_file.xml
```

**Tip:** If you choose to use the `-skipInstall` parameter with a temporary installation registry created as described in "Before you begin," Installation Manager uses the temporary installation registry while recording the response file. It is important to note that when you specify the `-skipInstall` parameter, no product packages are installed or uninstalled. All of the actions that you perform in Installation Manager simply update the installation data that is stored in the specified temporary registry. After the response file is generated, it can be used to uninstall the product, removing the product files and updating the standard installation registry.

The `-skipInstall` operation should not be used on the actual agent data location used by Installation Manager. This is unsupported. Use a clean writable location, and re-use that location for future recording sessions.

For more information, read the Installation Manager Information Center.

- c. Click **Uninstall**.
  - d. In the Uninstall Packages window, select the product and the appropriate version to uninstall. For example, select **IBM WebSphere Edge Components: Caching Proxy** or **IBM WebSphere Edge Components: Load Balancer for IPv4**.
  - e. Click **Next**.
  - f. Review the summary information.
  - g. Click **Uninstall**.
    - If the uninstallation is successful, the program displays a message that indicates success.
    - If the uninstallation is not successful, click **View log** to troubleshoot the problem.
  - h. Click **Finish**.
  - i. Click **File > Exit** to close Installation Manager.
4. Use the response file to uninstall the product silently. From a command line on each of the systems from which you want to uninstall the product, change to the `eclipse/tools` subdirectory in the directory where you installed Installation Manager and use the response file that you created to silently uninstall the product.

For example:

- **Windows Administrator or non-administrator:**  

```
imcl.exe
      input C:\temp\uninstall_response_file.xml
      -log C:\temp\uninstall_log.xml
```
- **AIX HP-UX Linux Solaris Administrator:**

```
./imcl
input /var/temp/uninstall_response_file.xml
-log /var/temp/uninstall_log.xml
```

- **AIX** **HP-UX** **Linux** **Solaris** **Non-administrator:**

```
./imcl
input user_home/var/temp/uninstall_response_file.xml
-log user_home/var/temp/uninstall_log.xml
```

Go to the Installation Manager Information Center for more information.

5. **Optional:** Uninstall IBM Installation Manager. You must uninstall all of the packages that were installed by Installation Manager before you can uninstall IBM Installation Manager.

Read Uninstalling Installation Manager in the Installation Manager Information Center for information about performing this procedure.

**Windows** **Example:** The following is a sample response file for uninstalling Load Balancer for IPv4:

```
<?xml version="1.0" encoding="UTF-8" ?>
<agent-input>
<server>
  <repository location='C:\repositories\edge80\local-repositories' />
</server>
<profile installLocation='c:\program files\loadbalancer\ibm\edge\lb'>
  <data key='eclipseLocation' value='c:\program files\loadbalancer\ibm\edge\lb' />
  <data key='cic.selector.nl'
    value='en, fr, it, zh, ru, zh_TW, de, ja, pl, es, cs, hu, ko, pt_BR' />
</profile>
<uninstall modify='false'>
  <offering id='com.ibm.websphere.EDGEIPV4.v80' version='8.0.0.0'
    profile='IBM WebSphere Edge Components: Load Balancer for IPV4 V8.0'
    features='Dispatcher,Load Balancer,Metric Server' />
</uninstall>
<preference name='com.ibm.cic.common.core.preferences.eclipseCache'
  value='C:\Program Files\IBM\IMShared' />
<preference value='30' name='com.ibm.cic.common.core.preferences.connectTimeout' />
<preference value='30' name='com.ibm.cic.common.core.preferences.readTimeout' />
<preference value='0' name='com.ibm.cic.common.core.preferences.downloadAutoRetryCount' />
<preference value='true' name='offering.service.repositories.areUsed' />
<preference value='false' name='com.ibm.cic.common.core.preferences.ssl.nonsecureMode' />
<preference value='false'
  name='com.ibm.cic.common.core.preferences.http.disablePreemptiveAuthentication' />
<preference value='true'
  name='com.ibm.cic.common.core.preferences.preserveDownloadedArtifacts' />
<preference value='false' name='PassportAdvantageIsEnabled' />
<preference value='false'
  name='com.ibm.cic.common.core.preferences.searchForUpdates' />
</agent-input>
```

## Uninstalling previous versions of Edge Components

If you installed a previous version of the Edge Components products, you need to use the native installers or system packaging tools to uninstall them. This section contains the following information:

- “Uninstalling Edge components for Windows”
- “Uninstalling Edge components for AIX” on page 76
- “Uninstalling Edge components for HP-UX” on page 76
- “Uninstalling Edge components for Linux” on page 76
- “Uninstalling Edge components for Solaris” on page 77

**Windows** **Uninstalling Edge Components for Windows**

1. To prevent the currently installed Load Balancer from starting, edit any start scripts that you created to temporarily suppress any commands that will start Load Balancer upon reboot.
2. Ensure that the Load Balancer service is set to Manual. For example:
  - a. Click **Start > Control Panel > Administrative Tools > Services**.
  - b. Right click the entry for Load Balancer, and select **Properties**.
  - c. Change the Startup type selection to Manual.
3. Restart your Windows machine.
4. Use the **Add/Remove Programs** feature in Control Panel to uninstall Caching Proxy or Load Balancer for IPv4.

#### **AIX** Uninstalling Edge Components for AIX

- To uninstall Caching Proxy, enter the following command:  
`installp -u wses`
- To uninstall Load Balancer for IPv4:
  1. Ensure that all executors and servers are stopped.
  2. enter the following command:  
`installp -u ibmlb`

For Version 6.0 or earlier, enter the following command:

`installp -u ibmnd`

#### **HP-UX** Uninstalling Edge Components for HP-UX

- To uninstall Caching Proxy, enter the following command:  
`swremove packagename`  
 Remove packages in the reverse order in which they were installed. These are the names and the recommended order for uninstalling the packages:
  1. WSES-DOC-en\_US
  2. WSES-CP
  3. WSES-cpmlang
  4. WSES-ADMIN
  5. WSES-ICU
  6. gsk7bas

The variable *lang* refers to the substitution of one of the following language specific codes: de\_DE, en\_US, es\_ES, fr\_FR, it\_IT, ja\_JP, ko\_KR, zh\_CN, zh\_TW. HP-UX does not support Portuguese Brazilian (pt\_BR).

To query the Caching Proxy packages that are still installed, use the command:

`swlist | grep WSES`

- To uninstall Load Balancer for IPv4:
  1. Ensure that all executors and servers are stopped.
  2. Enter the following command:  
`swremove ibmlb`

#### **Linux** Uninstalling Edge Components for Linux

- To uninstall Caching Proxy, enter the following command:  
`rpm -e packagename`

Remove the packages in the reverse order in which they were installed. These are the names and recommended order for uninstalling the packages:

1. WSES\_Doc\_en\_US-*release-version*.*hardw*.rpm
2. WSES\_CachingProxy-*release-version*.*hardw*.rpm
3. WSES\_CachingProxy\_msg\_lang3-*release-version*.*hardw*.rpm
4. WSES\_Admin\_Runtime-*release-version*.*hardw*.rpm
5. WSES\_ICU\_Runtime-*release-version*.*hardw*.rpm
6. gsk7bas.rpm

where:

- *release-version* is the current release, for example: 6.1.0-0
- *hardw* refers to the substitution of one of the following: i686, s390, ppc64.
- *lang* refers to the substitution of one of the following language specific codes: de\_DE, en\_US, es\_ES, fr\_FR, it\_IT, ja\_JP, ko\_KR, pt\_BR, zh\_CN, zh\_TW.

To query the Caching Proxy packages that are still installed, use the command:

```
rpm -qa |grep -i wses
```

- To uninstall Load Balancer for IPv4, enter the following command:

1. Ensure that all executors and servers are stopped.
2. Enter the following command:

```
rpm -e pkgname
```

When uninstalling the packages, reverse the order used for installation. These are the package names and the recommended order for uninstalling:

- a. *ibmlb-ms-release-version*.*hardw*.rpm (Metric Server)
- b. *ibmlb-doc-release-version*.*hardw*.rpm (Documentation)
- c. *ibmlb-component-release-version*.*hardw*.rpm (LB component)
- d. *ibmlb-lic-release-version*.*hardw*.rpm (License)
- e. *ibmlb-admin-release-version*.*hardw*.rpm (Administration)
- f. *ibmlb-base-release-version*.*hardw*.rpm (Base)

where:

- *release-version* is the current release, for example: 6.1.0-0
- *hardw* is one of the following values: i386, ppc64, ppc, s390, s390x, x86\_64.
- *component* is one of the following values: disp (Dispatcher component), cbr (CBR component), ss (Site Selector component), cco (Cisco CSS Controller), nal (Nortel Alteon Controller)

### **Solaris** Uninstalling Edge Components for Solaris

- To uninstall Caching Proxy, enter the following command:

```
pkgrm packagename
```

Remove the packages in the reverse order in which they were installed. These are the package names and the recommended order for uninstalling:

1. WSESdocen

2. WSEScp
3. WSEScplang
4. WSEAdmin
5. WSESicu
6. gsk7bas

To query the Caching Proxy packages that are still installed, use the command:

```
pkginfo | grep WSES
```

- To uninstall Load Balancer for IPv4:
  1. Ensure that all executors and servers are stopped.
  2. Enter the following command:  

```
pkgrm pkgname
```

When uninstalling the packages, reverse the order used for installation. These are the package names and the recommended order for uninstalling:

- a. ibmlbms (Metric Server)
- b. ibmlbdoc (Documentation)
- c. ibmlbna1 (Nortel Alteon Controller component)
- d. ibmlbcc0 (Cisco CSS Controller component)
- e. ibmlbss (Site Selector component)
- f. ibmlbcbr (CBR component)
- g. ibmlbdisp (Dispatcher component)
- h. ibmlblic (License)
- i. ibmlbadm (Administration)
- j. ibmlbbase (Base)



---

## Part 5. Building networks with Edge Components

This part provides procedures for building basic demonstration networks using Edge Components. These networks are not intended to be used in production environments. The process of initially configuring a network can clarify many edge-of-network concepts for administrators who are new to the product. For complete coverage of all component features and for in-depth configuration information, refer to the *Caching Proxy Administration Guide* and the *Load Balancer Administration Guide*.

The procedures permit any computer system supported by the component to be used at any node.

This part contains the following chapters:

Chapter 16, “Build a Caching Proxy network,” on page 81.

Chapter 17, “Build a Load Balancer network,” on page 85.



---

## Chapter 16. Build a Caching Proxy network

Figure 19 shows a basic proxy server network using three computer systems located at three network nodes. This network binds the proxy server to a dedicated content host (IBM HTTP Server), which is located on Server 2, and the proxy server serves the host. This is visually represented by the Internet being located between the workstation and Server 1.

**IMPORTANT:** Caching Proxy is available on all Edge component installations, with the following exceptions:

- Caching Proxy is not available for Edge component installations that run on Itanium 2 or AMD Opteron 64-bit processors.
- Caching Proxy is not available for Edge component installations of Load Balancer for IPv4 and IPv6.

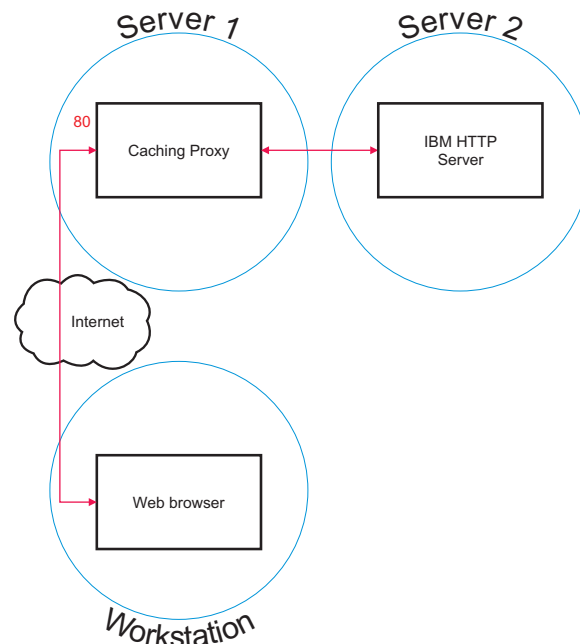


Figure 19. Caching Proxy demonstration network

---

### Workflow

To build a Caching Proxy network, perform these procedures in the following order:

1. Review required computer systems and software.
2. Build Server 1 (Linux and UNIX systems) or Build Server 1 (Windows system).
3. Configure Server 1.
4. Test the Caching Proxy network.

---

### Review required computer systems and software

The following computer systems and software components are needed:

- Computer system to act as Server 1. This system must have access to the Internet.
- Computer system to act as Server 2. An HTTP server must be installed on the content host.
- Computer system to act as the workstation. A Web browser must be installed.

---

## Build Server 1 (Linux and UNIX systems)

Install and configure the Caching Proxy as follows:

1. Ensure that the computer server meets all hardware and software requirements.
2. Log in as the superuser, typically root.
3. Install the Caching Proxy component.
4. Create an administrator identification and password for accessing the Configuration and Administration forms by entering the following command:  

```
# htadm -adduser /opt/ibm/edge/cp/server_root/protect/webadmin.passwd
```

When prompted, provide the **htadm** program with a user name, password, and real name for the administrator.
5. Continue with "Configure Server 1."

---

## Build Server 1 (Windows system)

Install and configure the Caching Proxy as follows:

1. Ensure that the Windows 2000 and Windows 2003 operating systems meet all hardware and software requirements.
2. Log in as a user with administrator privileges.
3. Install the Caching Proxy component.
4. Create an administrator identification and password for accessing the Configuration and Administration forms by entering the following command:  

```
cd "Program Files\IBM\edge\cp\server_root\protect"  
htadm -adduser webadmin.passwd"
```

When prompted, provide the **htadm** program with a user name, password, and real name for the administrator.
5. Continue with "Configure Server 1."

---

## Configure Server 1

From the workstation, do the following:

1. Start a Web browser.
2. In the **Address** field of your browser, enter `http://server_1`, where *server\_1* refers to the actual host name or IP address of the machine designated as Server 1.
3. Click **Configuration and Administration Forms**.
4. Enter your administrator name and password. The Configuration and Administration forms open in your browser.
5. Click **Server Configuration**—>**Request Processing**—>**Request Routing**.
6. Insert a new wildcard-mapping rule prior to the existing one by selecting the **Insert Before** radio button and the index value of the existing wildcard-mapping rule.
7. Select **Proxy** from the **Action** drop-down box.
8. Type `/*` in the **URL request template field**.

9. Type the host name for the site to which to redirect HTTP requests in the **Server IP Address or host name** field. Precede this value with `http://`.
10. Click **Submit**.
11. Create a mapping rule that allows access to the Configuration and Administration forms by selecting the **Insert Before** radio button and the index value of the mapping rule created in step 6.
12. Select **Pass** from the **Action** drop-down box.
13. Type `/pub/*` in the **URL request template field**.
14. Enter the location of the Configuration and Administration forms:
  -

**AIX**

**HP-UX** **Linux** **Solaris** Type `/opt/ibm/edge/cp/server_root/pub/en_US/*` in the **Server IP Address or host name** field.

- **Windows** Type `"C:\Program Files\IBM\edge\cp\server_root\pub\en_US\*"` in the **Server IP Address or host name** field.

15. Click **Submit**.
16. Click the **Restart Server** icon at the top of the configuration form.
17. Continue with "Test the Caching Proxy network."

---

## Test the Caching Proxy network

From the workstation, do the following:

1. Start a Web browser.
2. Enter `http://server_1` in the **Address** field of your browser. HTML pages from Server 2 will be proxied through Server 1 and delivered to the Web browser.
3. To access the Configuration and Administration forms, enter `http://server_1/pub/` into the **Address** field of your browser. The home page of the Configuration and Administration forms is displayed.



---

## Chapter 17. Build a Load Balancer network

Figure 20 shows a basic Load Balancer network with three locally attached workstations using the Dispatcher component's MAC forwarding method to load balance Web traffic between two Web servers. The configuration is similar when load balancing any other TCP or stateless UDP application traffic.

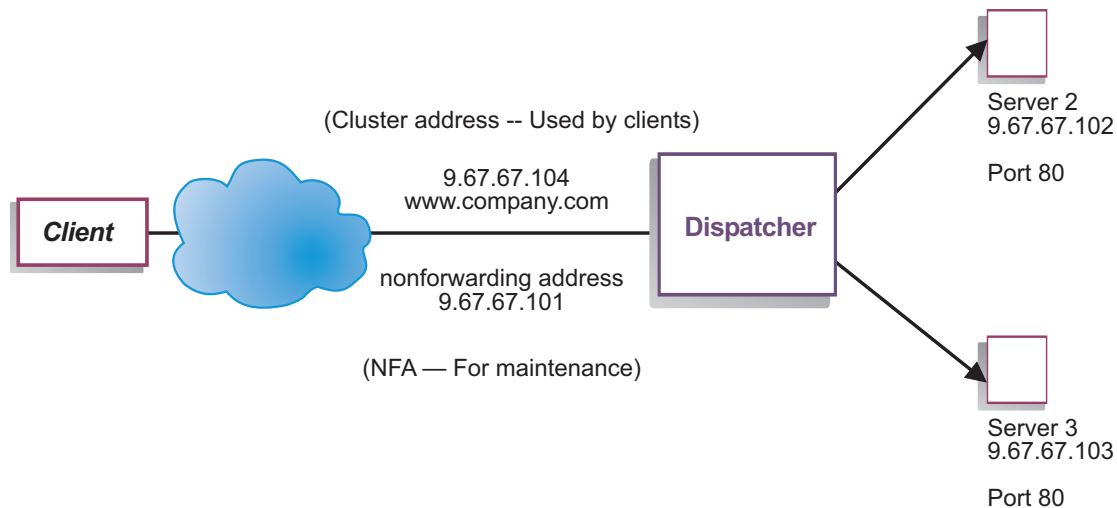


Figure 20. Load Balancer demonstration network

**Note:** This configuration can be completed using only two workstations with Dispatcher located on one of the Web server workstations. This represents a collocated configuration.

---

### Workflow

To build a Load Balancer network, perform these procedures in the following order:

1. Review required computer systems and software.
2. Configure the network.
3. Configure the Dispatcher.
4. Test the Load Balancer network.

---

### Review required computer systems and software

The following computer systems and software components are needed:

- Computer system to act as the Dispatcher. This system requires one actual IP address and one address to be load balanced.
- Two computer systems to act as Web servers. Each Web server requires one IP address.

---

## Configure the network

1. Set up your workstations so that they are on the same LAN segment. Ensure that network traffic among the three machines does not have to pass through any routers or bridges.
2. Configure the network adapters of the three workstations. For this example, assume you have the following network configuration:

Workstation	Name	IP Address
1	server1.company.com	9.67.67.101
2	server2.company.com	9.67.67.102
3	server3.company.com	9.67.67.103
Netmask = 255.255.255.0		

Each of the workstations contains only one standard Ethernet network interface card.

3. Ensure that server1.company.com can ping both server2.company.com and server3.company.com.
4. Ensure that server2.company.com and server3.company.com can ping server1.company.com.
5. Ensure that content is identical on the two Web servers (Server 2 and Server 3). This can be done by replicating data on both workstations, by using a shared file system such as NFS, AFS<sup>®</sup>, or DFS, or by any other means appropriate for your site.
6. Ensure that Web servers on server2.company.com and server3.company.com are operational. Use a Web browser to request pages directly from `http://server2.company.com` and `http://server3.company.com`.
7. Obtain another valid IP address for this LAN segment. This is the address you provide to clients who wish to access your site. For this example, the information is as follows:

Name= www.company.com  
IP=9.67.67.104

8. Configure the two Web server workstations to accept traffic for www.company.com.  
Add an alias for www.company.com to the **loopback** interface on server2.company.com and server3.company.com.
  - For AIX<sup>®</sup>:  
**ifconfig lo0 alias www.company.com netmask 255.255.255.0**
  - **Solaris** For Solaris 9:  
**ifconfig lo0:1 plumb www.company.com netmask 255.255.255.0**
9. Delete any extra route that might have been created as a result of aliasing the loopback interface.

You have now completed all configuration steps that are required on the two Web server workstations.

---

## Configure the Dispatcher

With Dispatcher, you can create a configuration by using the command line, the configuration wizard, or the graphical user interface (GUI).

**Note:** The parameter values must be typed in English characters. The only exceptions are parameter values for host names and file names.



## Configuring using the command line

If you are using the command line, follow these steps:

1. Start the dserver on Dispatcher:
  - **AIX** **HP-UX** **Linux** **Solaris** Run the following command as root user: `dserver`
  - **Windows** `dserver` runs as a service that starts automatically.
2. Start the executor function of Dispatcher:  
`dscontrol executor start`
3. Add the cluster address to the Dispatcher configuration:  
`dscontrol cluster add www.company.com`
4. Add the http protocol port to the Dispatcher configuration:  
`dscontrol port add www.company.com:80`
5. Add each of the Web servers to the Dispatcher configuration:  
`dscontrol server add www.company.com:80:server2.company.com`  
`dscontrol server add www.company.com:80:server3.company.com`
6. Configure the workstation to accept traffic for the cluster address:  
`dscontrol executor configure www.company.com`
7. Start the manager function of Dispatcher:  
`dscontrol manager start`  
Dispatcher now does load balancing based on server performance.
8. Start the advisor function of Dispatcher:  
`dscontrol advisor start http 80`  
Dispatcher now ensures that client requests are not sent to a failed Web server.

Your basic configuration with locally attached servers is now complete.

## Configuring using the configuration wizard

If you are using the configuration wizard, follow these steps:

1. Start the dserver on Dispatcher:
  - **AIX** **HP-UX** **Linux** **Solaris** Run the following command as root user:  
`dserver`
  - **Windows** `dserver` runs as a service that starts automatically.
2. Start the wizard function of Dispatcher, `dswizard`.

The wizard guides you step-by-step through the process of creating a basic configuration for the Dispatcher component. It asks questions about your network and guides you through the setup of a cluster for Dispatcher to load balance the traffic for a group of servers.

The configuration wizard contains the following panels:

- Introduction to the wizard
- What is going to happen
- Preparing for the setup
- Choosing a host to configure (if necessary)
- Defining a cluster
- Adding a port
- Adding a server

- Starting an advisor
- Server machine setup

## Configuring using the graphical user interface (GUI)

To start the GUI, follow these steps:

1. Ensure that the dsserver process is running:
  - **AIX** **HP-UX** **Linux** **Solaris** Run the following command as root:  
dsserver
  - **Windows** dsserver runs as a service that starts automatically.
2. Next, do one of the following:
  - **AIX** **HP-UX** **Linux** **Solaris** Type lbadmin.
  - **Windows** Click **Start > Programs > IBM WebSphere > Edge Components > IBM Load Balancer for IPv4 > Load Balancer for IPv4**.

---

## Test the Load Balancer network

1. From a Web browser, go to location <http://www.company.com> to verify that a page appears.
2. Reload the page in the Web browser.
3. Issue the following command: `dscontrol server report www.company.com:80:.`  
Verify that the total connections column of the two servers adds up to 2.

---

## Notices

### First edition (May 2012)

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Corporation  
Attn.: G71A/503  
P.O. box 12195  
3039 Cornwallis Rd.  
Research Triangle Park, N.C. 27709-2195  
U.S.A.

For license inquiries regarding double-byte (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

IBM World Trade Asia Corporation Licensing  
2-31 Roppongi 3-chome, Minato-ku  
Tokyo 106, Japan

**The following paragraph does not apply to the United Kingdom or any country where such provisions are inconsistent with local law:**

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OR CONDITIONS OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the document. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

IBM Corporation  
ATTN: Software Licensing  
11 Stanwix Street  
Pittsburgh, PA 15222-9183  
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM International Program License Agreement or any equivalent agreement between us.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

All statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples may include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

---

## Trademarks

The following terms are trademarks of IBM Corporation in the United States, other countries, or both:

- AFS
- AIX
- DFS
- IBM

- iSeries®
- RS/6000®
- SecureWay
- Tivoli
- ViaVoice
- WebSphere

Java and all Java-based trademarks are trademarks of Oracle in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel Inside (logos), MMX and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.







Printed in USA



Spine information:



WebSphere Application Server

**Concepts, Planning, and Installation for Edge  
Components**

Version 8.5