IBM SPSS Analytic Server
Version 3.2.1

Installation und Konfiguration

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen" auf Seite 77 gelesen werden.

Produktinformation

Diese Ausgabe bezieht sich auf Version 3, Release 2, Modifikation 1 von IBM SPSS Analytic Server und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Inhaltsverzeichnis

Kapitel 1. Voraussetzungen 1

Kapitel 2. Ambari-Installation und -Konfiguration. 3

Ambari-spezifische Voraussetzungen 3
Precheck- und Postcheck-Tools für Installation - Ambari 3
Installation in Ambari 6
Onlineinstallation. 6
Offlineinstallation 10
Installieren von Analytic Server in extern verwalteter MySQL-Umgebung 15
Konfiguration 16
Sicherheit 16
Aktivieren der Unterstützung für Essentials for R 22
Aktivieren relationaler Datenbankquellen . . . 24
Aktivieren von HCatalog-Datenquellen 25
Ändern der von Analytic Server verwendeten Ports 27
Analytic Server mit hoher Verfügbarkeit. . . . 27
Optimieren von JVM-Optionen für Small Data. . 29
Aktualisierung von Clientabhängigkeiten . . . 29
Konfigurieren von Apache Knox 29
Konfigurieren separater YARN-Warteschlangen für jeden IBM SPSS Analytic Server-Nutzer - HDP. 32
Migrieren von IBM SPSS Analytic Server unter Ambari 33
Deinstallation. 35
Deinstallation von Essentials for R. 36

Kapitel 3. Cloudera-Installation und -Konfiguration 37

Cloudera - Übersicht 37
Cloudera-spezifische Voraussetzungen 37
Kerberos-fähige Cloudera-Umgebungen 37
Konfigurieren von MySQL für Analytic Server. . 39
Precheck- und Postcheck-Tools für Installation - Cloudera 40
Installation in Cloudera 42
Konfigurieren von Cloudera. 48
Sicherheit 48
Aktivieren der Unterstützung für Essentials for R 53
Aktivieren relationaler Datenbankquellen . . . 54
Aktivieren von HCatalog-Datenquellen 55
Konfigurieren von Apache Impala. 57
Ändern der von Analytic Server verwendeten Ports 58
Analytic Server mit hoher Verfügbarkeit. . . . 58
Optimieren von JVM-Optionen für Small Data. . 60
Konfigurieren separater YARN-Warteschlangen für jeden IBM SPSS Analytic Server-Nutzer - Cloudera 60
Migration 61
Deinstallation von Analytic Server in Cloudera . . 62

Kapitel 4. Konfigurieren von IBM SPSS Modeler für die Verwendung mit IBM SPSS Analytic Server 65

Kapitel 5. Konfigurieren des Pushbacks für die benutzerdefinierte Funktion von Hive 67

Kapitel 6. Verwenden von SLM-Tags zum Überwachen der Lizenzierung . . 69

Kapitel 7. Fehlerbehebung 71

Bemerkungen. 77

Marken. 78

Kapitel 1. Voraussetzungen

Lesen Sie vor der Installation von Analytic Server die nachfolgenden Informationen.

Systemvoraussetzungen
Die aktuellen Informationen zu Systemanforderungen finden Sie in den Berichten mit den detaillierten Systemanforderungen auf der Site des IBM Technical Support unter http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html. Gehen Sie auf dieser Seite wie folgt vor:

1. Geben Sie SPSS Analytic Server als Produktnamen ein und klicken Sie auf **Search**.
2. Wählen Sie die gewünschte Version und den Berichtsumfang aus und klicken Sie dann auf **Submit**.

WebSocket-Datenverkehr
Sie müssen sicherstellen, dass der WebSocket-Datenverkehr zwischen Clients und Analytic Server nicht durch Firewalls, VPNs oder andere Methoden zum Blockieren von Ports blockiert werden. Der WebSocket-Port ist mit dem allgemeinen Analytic Server-Port identisch.

SuSE Linux (SLES) 12
Führen Sie die folgenden Aufgaben aus, bevor Sie Analytic Server unter SuSE Linux 12 installieren:

1. Laden Sie einen öffentlichen Schlüssel von der folgenden URL auf Ihren Host herunter:
   ```
   https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.1.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
   ```
2. Importieren Sie den öffentlichen Schlüssel, indem Sie den folgenden Befehl auf Ihrem Host ausführen:
   ```
   rpm --import IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
   ```

Power Systems
Stellen Sie sicher, dass die IBM Compiler XL C und XL F installiert und die Installationspfade in der PATH-Variablen aller Hosts im Cluster enthalten sind.

Weitere Informationen zum Erwerben einer Lizenz für diese Compiler finden Sie auf den folgenden Websites:

- XL C für Linux: http://www-03.ibm.com/software/products/en/xlcpp-linux
- XL Fortran für Linux: http://www-03.ibm.com/software/products/en/xlfortran-linux

Hortonworks Data Platform (HDP)
Vor der Installation von Analytic Server müssen Sie sicherstellen, dass in Ihrer Clusterumgebung mindestens ein HDP-Client bereitgestellt wurde. Da der Knoten, der Ambari Manager hostet, das Verzeichnis `/usr/hdp` erwartet, schlägt Analytic Server fehl, wenn kein HDP-Client vorhanden ist.

Hive/HCatalog
Wenn Sie NoSQL-Datenquellen verwenden wollen, konfigurieren Sie Hive und HCatalog für den Fernzugriff. Stellen Sie außerdem sicher, dass `hive-site.xml` eine Eigenschaft *hive.metastore.uris* im Format `thrift://<Hostname>:<Port>` enthält, die auf den aktiven Server für Thrift Hive Metastore verweist. Details finden Sie in der Dokumentation zur Hadoop-Verteilung.

Anmerkung: Der Analytic Server-Metaspeicher kann nicht auf demselben System wie der Hive-Metaspeicher installiert werden.

Wenn Sie Hive 2.1 verwenden wollen, müssen Sie Hive 2.1 aktivieren, indem Sie die Einstellung **Interactive Query** in der Ambari-Konsole aktivieren und während der Installation von Analytic Server `2.x` für die Eigenschaft `hive.version` angeben.

1. Öffnen Sie die Ambari-Konsole und fügen Sie im Abschnitt **Analytic Server Advanced analytics.cfg** die folgende Eigenschaft hinzu.

- Key: `hive.version`
- Value: Geben Sie die entsprechende Hive-Version ein (z. B. 2.x)

2. Speichern Sie die Konfiguration.

Anmerkung: Hive 2.1 wird unter HDP ab Version 2.5 mit Spark 2.x unterstützt.

Metadatenrepository

Standardmäßig installiert und verwendet Analytic Server eine MySQL-Datenbank. Alternativ können Sie Analytic Server für die Verwendung einer vorhandenen Db2-Installation konfigurieren. Unabhängig vom ausgewählten Typ der Datenbank muss sie eine UTF-8-Codierung haben.

MySQL

Der Standardzeichensatz für MySQL hängt von der Version und dem Betriebssystem ab. Verwenden Sie die folgenden Schritte, um festzustellen, ob Ihre Installation von MySQL auf UTF-8 gesetzt ist.

1. Bestimmen Sie die Version von MySQL.

   ```
   mysql -V
   ```

2. Führen Sie die folgende Abfrage über die MySQL-Befehlszeilenschnittstelle aus, um den Standardzeichensatz für MySQL zu bestimmen.

   ```
   mysql>show variables like 'char%';
   ```

   Wenn der Zeichensatz bereits auf UTF-8 gesetzt ist, sind keine weiteren Änderungen erforderlich.

3. Führen Sie die folgende Abfrage über die MySQL-Befehlszeilenschnittstelle aus, um die Standardsortierfolge für MySQL zu bestimmen.

   ```
   mysql>show variables like 'coll%';
   ```

   Wenn die Sortierfolge bereits auf UTF-8 gesetzt ist, sind keine weiteren Änderungen erforderlich.

4. Wenn der Standardzeichensatz oder die Standardsortierfolge nicht auf UTF-8 gesetzt ist, finden Sie in der Dokumentation zu MySQL Details zum Bearbeiten von /etc/my.cnf und zum Neustart des MySQL-Dämons, um den Zeichensatz in UTF-8 zu ändern.

Db2 Weitere Informationen zum Konfigurieren von Db2 finden Sie im Knowledge Center unter http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html.

Hochverfügbarkeitscluster

Lastausgleichsfunktion

Ihr Hochverfügbarkeitscluster sollte eine Lastausgleichsfunktion enthalten, die Sitzungsaffinität (auch als "permanente Sitzungen" bezeichnet) unterstützt. Analytic Server gibt Sitzungen mit dem Cookie "request-token" an. Dadurch wird eine Sitzung für die Dauer einer Benutzeranmeldung zur Verwendung in der anwendungsgesteuerten Sitzungsaffinität angegeben. Ziehen Sie die Dokumentation für Ihre spezielle Lastausgleichsfunktion zu Rate, die Details zur Unterstützung der Sitzungsaffinität enthält.

Analytic Server-Jobfehler

Wenn ein Analytic Server-Job fehlschlägt, da ein Cluster-Member fehlschlägt, wird der Job normalerweise auf einem anderen Cluster-Member fortgesetzt. Wenn der Job nicht fortgesetzt wird, stellen Sie sicher, dass der Hochverfügbarkeitscluster mindestens 4 Cluster-Member umfasst.

Kapitel 2. Ambari-Installation und -Konfiguration

Ambari-spezifische Voraussetzungen

Lesen Sie zusätzlich zu den Angaben zu allgemeinen Voraussetzungen die folgenden Informationen.

Services

Analytic Server ist als Ambari-Service installiert. Vor der Installation von Analytic Server müssen Sie sicherstellen, dass die folgenden Clients als Ambari-Services installiert sind:

- HDFS/HDFS_CLIENT
- MAPREDUCE2/MAPREDUCE2_CLIENT
- HIVE/HIVE_CLIENT
- SPARK/SPARK_CLIENT (wenn Spark 1.x verwendet wird)
- SPARK2/SPARK2_CLIENT (wenn Spark 2.x verwendet wird)
- HBASE/HBASE_CLIENT (wenn HBASE verwendet wird)
- YARN
- Zookeeper

Kennwortlose SSH

Konfigurieren Sie für den Rootbenutzer die kennwortlose SSH zwischen dem Analytic Metastore-Host und allen Hosts im Cluster.

Precheck- und Postcheck-Tools für Installation - Ambari

Übersicht über das Precheck-Tool

Das Precheck-Tool für die Analytic Server-Installation hilft bei der Reduzierung von Installationsproblemen und Laufzeitfehlern, indem es potenzielle Umgebungsprobleme vor der Analytic Server-Installation ermittelt.

Das Precheck-Tool prüft Folgendes:

- Betriebssystem- und Ambari-Versionen auf dem lokalen System
- ulimit-Betriebssystemeinstellungen auf dem lokalen System
- Verfügbarer Plattenspeicher auf dem lokalen System
- Hadoop-Version
- Ambari-Serviceverfügbarkeit (HDFS, HCatalog, Spark, Hive, MapReduce, Yarn, Zookeeper usw.)
- Bestimmte Analytic Server-Ambari-Einstellungen

Anmerkung: Das Precheck-Tool kann verwendet werden, nachdem die selbstextrahierende Analytic Server-Binärdatei ausgeführt wurde.

Übersicht über das Postcheck-Tool

Das Postcheck-Tool für die Analytic Server-Installation ermittelt Konfigurationsprobleme nach der Analytic Server-Installation, indem REST-API-Anforderungen zur Verarbeitung übergeben werden:

- Daten in HDFS
- Daten in Hive/HCatalog
- Komprimierte Daten (einschließlich deflate, bz2, snappy)
- Daten mit PySpark

- Daten, die native SPSS-Komponenten verwenden (einschließlich alm, tree, neuralnet, scoring, tascoring)
- Daten mit MapReduce
- Daten mit speicherinternem MapReduce

Speicherort und Voraussetzungen für das Tool

Führen Sie vor der Installation des Analytic Server-Service das Precheck-Tool auf allen Knoten aus, die Teil des Analytic Server-Service sein werden, um zu prüfen, ob Ihre Linux-Umgebung für die Installation von Analytic Server bereit ist.

Das Precheck-Tool wird automatisch als Teil der Installation aufgerufen. Das Tool prüft Analytic Metastore und jeden Analytic Server-Knoten, bevor die Installation auf jedem Host ausgeführt wird. Sie können das Precheck-Tool auch manuell auf dem Ambari-Server-Knoten aufrufen. Dadurch wird der Computer vor der Installation des Service validiert.

Nach dem Ausführen der selbstextrahierenden Analytic Server-Binärdatei befindet sich das Precheck-Tool in den folgenden Verzeichnissen:

- **HDP**

```
/opt/ibm/spss/analyticserver-ambari/3.2/ANALYTICSERVER/package/chktool/precheck.py

[root@servername chktool]# cd /opt/ibm/spss/analyticserver-ambari/3.2/ANALYTICSERVER/package/chktool
[root@servername chktool]# ls
checkers data lib logs postcheck.py precheck.py readme.txt
```

Nach der Installation von Analytic Server befindet sich das Postcheck-Tool im folgenden Verzeichnis:

- **HDP**

```
/opt/ibm/spss/analyticserver/3.2/tools/com.spss.ibm.checker.zip
```

Die Tools müssen als Root ausgeführt werden und erfordern Python 2.6.X (oder höher).

Wenn das Precheck-Tool Fehler meldet, müssen diese behoben werden, bevor Sie die Analytic Server-Installation fortsetzen.

Das Verzeichnis `chktool` ist nach der Ausführung der selbstextrahierenden Analytic Server-Binärdatei (Schritt 2 im Abschnitt „Installation in Ambari" auf Seite 6) verfügbar. Wenn Sie die Ausführung einer „Offlineinstallation" auf Seite 10 auswählen, ist das Verzeichnis `chktool` nach der Installation der Metadaten-RPM verfügbar.

Ausführen des Precheck-Tools

Automatisch

Das Precheck-Tool kann automatisch als Teil der Analytic Server-Installation aufgerufen werden, wenn Analytic Server über die Ambari-Konsole als Service installiert wird. Sie müssen den Benutzernamen und das Kennwort des Ambari-Serveradministrators manuell eingeben:

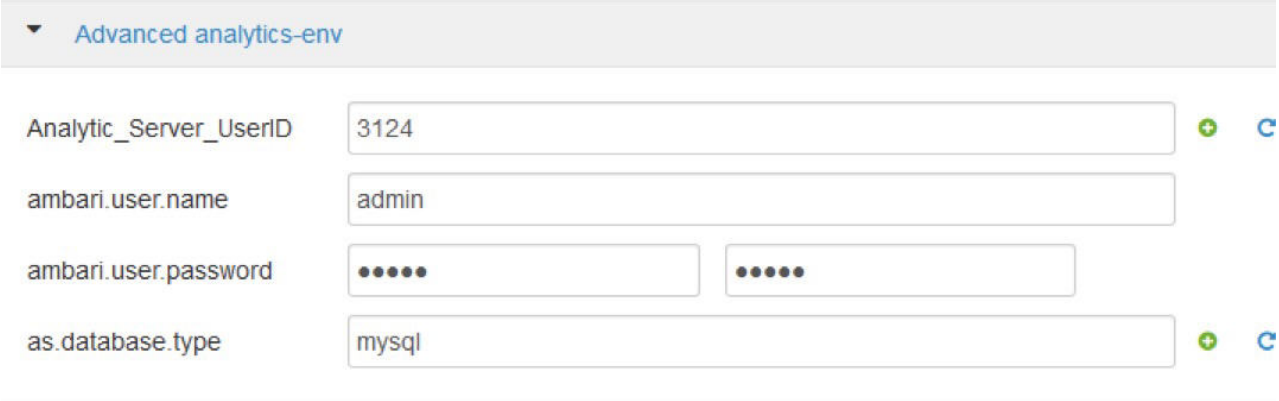

Abbildung 1. Erweiterte analytics-env-Einstellungen

Manuell

Sie können das Precheck-Tool manuell auf dem Ambari-Serverknoten aufrufen.

Das folgende Precheck-Beispiel prüft den Ambari-Cluster `MyCluster`, der auf `myambarihost.ibm.com:8080` mit aktiviertem SSL ausgeführt wird, und verwendet die Anmeldeberechtigungsnachweise `admin:admin`:

```
python ./precheck.py --target B --cluster MyCluster --username admin
--password admin --host myambarihost.ibm.com --port 8080 --as_host myashost.ibm.com --ssl
```

Hinweise:

- Der Wert `as_host` muss über die IP-Adresse oder einen vollständig qualifizierten Domänennamen bereitgestellt werden.
- Das Tool fordert zur Eingabe eines Kennworts auf, wenn das Kennwortargument ausgelassen wird.
- Der Befehl `precheck.py` enthält Syntaxhilfe, die mit dem Argument `--h` (`python ./precheck.py --help`) angezeigt werden kann.
- Das Argument `--cluster` ist optional. (Der aktuelle Cluster wird ermittelt, wenn `--cluster` nicht verwendet wird.)

Während das Precheck-Tool seine Prüfungen ausführt, wird der Status jeder Prüfung im Befehlsfenster angezeigt. Wenn ein Fehler auftritt, enthält die Protokolldatei detaillierte Informationen (die genaue Position der Protokolldatei wird im Befehlsfenster angegeben). Die Protokolldatei kann IBM Technical Support bereitgestellt werden, wenn mehr Unterstützung erforderlich ist.

Ausführen des Postcheck-Tools

Das Postcheck-Tool prüft, ob Analytic Server ordnungsgemäß ausgeführt wird und einfache Jobs verarbeiten kann. Das folgende Postcheck-Beispiel prüft eine Analytic Server-Instanz, die auf `myanalyticserverhost.ibm.com:9443` mit aktiviertem SSL ausgeführt wird, und verwendet die Anmeldeberechtigungsnachweise `admin:ibmspss`:

```
python ./postcheck.py --host myanalyticserverhost.ibm.com --port 9443
--username admin --password ibmspss --ssl
```

Wenn Knox mit Analytic Server verwendet wird, lautet der Befehl wie folgt:

```
python ./postcheck.py --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default
```

Führen Sie eine einzelne Prüfung mit dem folgenden Befehl durch:

```
python ./postcheck.py --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default --check AS_PYSPARK_BUILDMODEL
```

Hinweise:

- Das Tool fordert zur Eingabe eines Kennworts auf, wenn das Kennwortargument ausgelassen wird.
- Der Befehl `postcheck.py` enthält Syntaxhilfe, die mit dem Argument `--h` (`python ./postcheck.py --help`) angezeigt werden kann.

Während das Postcheck-Tool seine Prüfungen ausführt, wird der Status jeder Prüfung im Befehlsfenster angezeigt. Wenn ein Fehler auftritt, enthält die Protokolldatei detaillierte Informationen (die genaue Position der Protokolldatei wird im Befehlsfenster angegeben). Die Protokolldatei kann IBM Technical Support bereitgestellt werden, wenn mehr Unterstützung erforderlich ist.

Installation in Ambari

Der grundlegende Prozess ist, dass die Analytic Server-Dateien auf einem Host innerhalb des Ambari-Clusters installiert werden und Analytic Server dann als Ambari-Service hinzugefügt wird.

„Onlineinstallation"
Wählen Sie die Onlineinstallation aus, wenn Ihr Ambari-Server-Host und alle Knoten im Cluster auf https://ibm-open-platform.ibm.com zugreifen können.

„Offlineinstallation" auf Seite 10
Wählen Sie die Offlineinstallation aus, wenn Ihr Ambari-Server-Host keinen Internetzugriff hat.

Onlineinstallation

Wählen Sie die Onlineinstallation aus, wenn Ihr Ambari-Server-Host und alle Knoten im Cluster auf https://ibm-open-platform.ibm.com zugreifen können.

1. Navigieren Sie zur IBM Passport Advantage®-Website und laden Sie die selbstextrahierende Binärdatei, die Ihrem Stack, Ihrer Stackversion und Ihrer Hardwarearchitektur entspricht, auf den Ambari-Managerknoten herunter. Die verfügbaren Ambari-Binärdateien sind:

Tabelle 1. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM® SPSS Analytic Server 3.2.1 für Hortonworks Data Platform 2.5, 2.6, 3.0 und 3.1 Linux x86-64 (Englisch)	`spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.1 für Hortonworks Data Platform 2.6, 3.0 und 3.1 Linux on System p LE (Englisch)	`spss_as-3.2.1.0-hdp2.6-3.1-lppc64.bin`

2. Führen Sie die selbstextrahierende Binärdatei aus, folgen Sie den Anweisungen, um die Lizenz anzuzeigen und diese zu akzeptieren, wählen Sie die Onlineinstallation aus und wählen Sie den Installationsprozess für den Datenbanktyp aus, den Analytic Server verwendet. Ihnen werden die folgenden Optionen für den Datenbanktyp angeboten:
   - Neue MySQL-Instanz
   - Bereits vorhandene MySQL- oder Db2-Instanz
3. Führen Sie im Verzeichnis `/var/lib/ambari-server/resources/stacks/<Stackname>/<Stackversion>/services/ANALYTICSERVER/package/scripts` das Script `update_clientdeps.sh` mit den entsprechenden Argumenten aus (verwenden Sie z. B. das Argument `--help`).
4. Starten Sie Ihren Ambari-Server erneut.
   ```
   ambari-server restart
   ```
5. Melden Sie sich an Ihrem Ambari-Server an und installieren Sie Analytic Server als Service über die Ambari-Benutzerschnittstelle.

   Metadatenrepository

   Analytic Server verwendet standardmäßig MySQL, um Informationen zu Datenquellen, Projekten und Nutzern zu verfolgen. Während der Installation müssen Sie einen Benutzernamen

(**metadata.repository.user.name**) und ein Kennwort (**metadata.repository.password**) angeben, die in der JDBC-Verbindung zwischen Analytic Server und MySQL verwendet werden. Das Installationsprogramm erstellt den Benutzer in der MySQL-Datenbank. Dieser Benutzer ist spezifisch für die MySQL-Datenbank und muss kein vorhandener Linux- oder Hadoop-Benutzer sein.

Anmerkung: Wenn Sie während der Installation eine neue MySQL-Instanz installieren möchten, müssen Sie den Analytic Server-Metaspeicher auf einem System installieren, auf dem MySQL nicht installiert ist.

Führen Sie die folgenden Schritte aus, um das Metadatenrepository in Db2 zu ändern.

Anmerkung: Sie können das Metadatenrepository nach Abschluss der Installation nicht ändern.

a. Stellen Sie sicher, dass Db2 auf einem anderen Computer installiert ist. Weitere Informationen finden Sie im Abschnitt zum Metadatenrepository in Kapitel 1, „Voraussetzungen", auf Seite 1.
b. Navigieren Sie auf der Registerkarte "Ambari Services" zur Registerkarte "Configs" des Analytic Server-Service.
c. Öffnen Sie den Abschnitt **Advanced analytics-env**.
d. Ändern Sie den Wert von **as.database.type** von `mysql` in `db2`.
e. Öffnen Sie den Abschnitt **Advanced analytics-meta**.
f. Ändern Sie den Wert von **metadata.repository.driver** von `com.mysql.jdbc.Driver` in `com.ibm.db2.jcc.DB2Driver`.
g. Ändern Sie den Wert von **metadata.repository.url** in `jdbc:db2://{Db2-Host}:{Port}/{Datenbankname}:currentSchema={Schemaname};`. Dabei gilt Folgendes:
   - {Db2-Host} ist der Hostname des Servers, auf dem Db2 installiert ist.
   - {Port} ist der Port, an dem Db2 empfangsbereit ist.
   - {Schemaname} ist ein verfügbares, nicht verwendetes Schema.

   Wenn Sie sich nicht sicher sind, welche Werte eingegeben werden sollen, wenden Sie sich an Ihren Db2-Administrator.
h. Geben Sie in **metadata.repository.user.name** und **metadata.repository.password** gültige Db2-Berechtigungsnachweise an.
i. Klicken Sie auf **Save**.

LDAP-Konfiguration

Analytic Server verwendet einen LDAP-Server zum Speichern und Authentifizieren von Benutzern und Gruppen. Sie stellen die erforderlichen LDAP-Konfigurationsinformationen während der Installation von Analytic Server bereit.

Tabelle 2. LDAP-Konfigurationseinstellungen

LDAP-Einstellungen	Beschreibung
`as.ldap.type`	LDAP-Typ. Der Wert kann `ads`, `ad` oder `openladp` sein. • `ads` - Apache Directory Server (Standardeinstellung) • `ad` - Microsoft Active Directory • `openladp` - OpenLDAP
`as.ldap.host`	LDAP-Host
`as.ldap.port`	LDAP-Portnummer
`as.ldap.binddn`	LDAP-Bindungs-DN
`as.ldap.bindpassword`	Kennwort für LDAP-Bindungs-DN

Tabelle 2. LDAP-Konfigurationseinstellungen (Forts.)

LDAP-Einstellungen	Beschreibung
`as.ldap.basedn`	LDAP-Basis-DN
`as.ldap.filter`	LDAP-Benutzer- und -Gruppenfilterregel
`as.ldap.ssl.enabled`	Gibt an, ob SSL für die Kommunikation zwischen Analytic Server und LDAP verwendet werden soll. Der Wert kann `true` oder `false` sein.
`as.ldap.ssl.reference`	LDAP-SSL-Referenz-ID
`as.ldap.ssl.content`	LDAP-SSL-Konfiguration

- `as.ldap.type` ist standardmäßig auf `ads` gesetzt und die anderen zugehörigen Einstellungen enthalten Standardeinstellungen. Die Ausnahme ist, dass Sie ein Kennwort für die Einstellung `as.ldap.bindpassword` angeben müssen. Analytic Server verwendet die Konfigurationseinstellungen für die Installation einer ADS-Instanz (Apache Directory Server) und zum Ausführen der Serverinitialisierung. Das ADS-Standardprofil schließt den Benutzer `admin` mit dem Kennwort `admin` ein. Sie können die Benutzerverwaltung über die Analytic Server-Konsole durchführen oder Benutzer- und Gruppeninformationen über das Script `importUser.sh` im Ordner `<Analytic Server-Stammverzeichnis>/bin` importieren.
- Wenn Sie planen, einen externen LDAP-Server (z. B. Microsoft Active Directory oder OpenLDAP) zu verwenden, müssen Sie die Konfigurationseinstellungen den tatsächlichen LDAP-Werten entsprechend konfigurieren. Weitere Informationen finden Sie in LDAP-Benutzerregistrys in Liberty konfigurieren.
- Sie können die LDAP-Konfiguration ändern, nachdem Analytic Server installiert wurde (z. B. von Apache Directory Server auf OpenLDAP ändern). Wenn Sie allerdings ursprünglich mit Microsoft Active Directory oder OpenLDAP beginnen und später entscheiden, zu Apache Directory Server zu wechseln, installiert Analytic Server während der Installation keine Apache Directory Server-Instanz. Apache Directory Server wird nur installiert, wenn es während der Erstinstallation von Analytic Server ausgewählt wird.

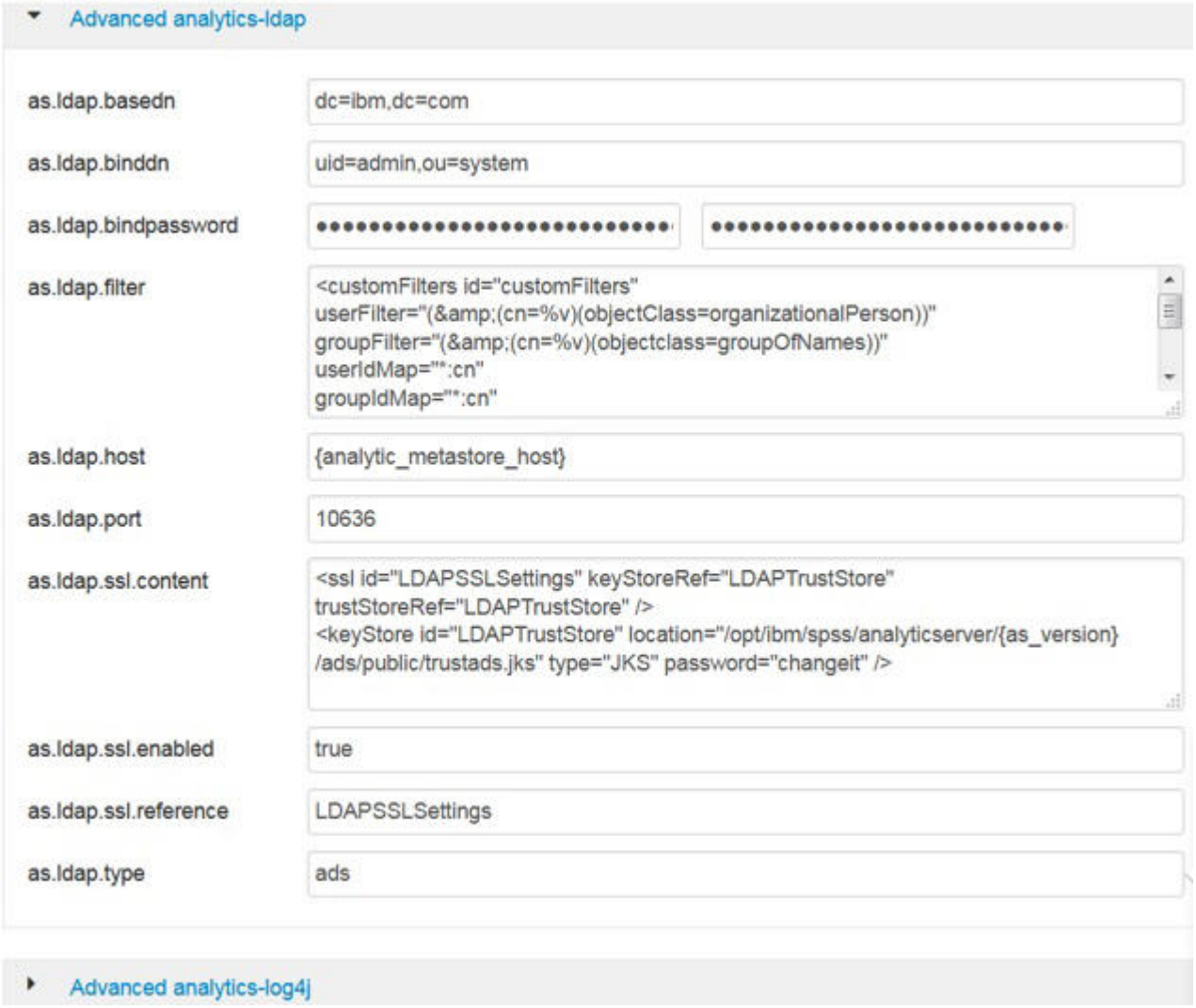

Abbildung 2. Beispiel für LDAP-Konfigurationseinstellungen

Konfigurationseinstellungen, die nach der Installation nicht geändert werden dürfen

Ändern Sie die folgenden Einstellungen nach der Installation nicht, da Analytic Server andernfalls nicht ausgeführt werden kann.

- Analytic_Server_User
- Analytic_Server_UserID
- as.database.type
- metadata.repository.driver
- distrib.fs.root

6. Nun haben Sie eine funktionierende Instanz von Analytic Server. Die weitere Konfiguration ist optional. Weitere Informationen zur Konfiguration und Verwaltung von Analytic Server finden Sie in „Konfiguration" auf Seite 16. Informationen zum Migrieren einer vorhandenen Konfiguration auf eine neue Installation finden Sie in „Migrieren von IBM SPSS Analytic Server unter Ambari" auf Seite 33.
7. Öffnen Sie einen Web-Browser und geben Sie die Adresse `http://<Host>:<Port>/analyticserver/admin/ibm` ein, wobei <Host> die Adresse des Analytic Server-Hosts und <Port> der Port ist, an dem Analytic Server empfangsbereit ist. Standardmäßig wird Port 9080 verwendet. Diese URL öffnet das Anmeldedialogfeld für die Analytic Server-Konsole. Melden Sie sich als Analytic Server-Administrator an. Standardmäßig ist die Benutzer-ID "admin" und das zugehörige Kennwort ist ebenfalls "admin".

Offlineinstallation

Eine Offlineinstallation von IBM SPSS Analytic Server kann automatisch oder manuell ausgeführt werden.

„Automatische Installation unter HDP"

Der automatische Installationsprozess verwendet die Ambari-REST-API und ist die bevorzugte Installationsmethode.

„Manuelle Installation unter HDP (RHEL, SLES)" auf Seite 11
Manuelle Installation von Analytic Server auf Hortonworks Data Platform

„Manuelle Installation unter HDP (Ubuntu)" auf Seite 13
Manuelle Installation von Analytic Server unter Ubuntu Linux

Automatische Installation unter HDP

Der automatische Installationsprozess verwendet die Ambari-REST-API und ist die bevorzugte Installationsmethode.

Wichtig:

- Die automatische Offlineinstallationsprozedur installiert eine eingebettete ADS-Instanz (Apache Directory Server). Wenn Sie einen LDAP-Server eines anderen Anbieters verwenden wollen, können Sie Ihre LDAP-Einstellungen konfigurieren, nachdem die Installation von IBM SPSS Analytic Server abgeschlossen ist.
- Die automatische Offlineinstallationsprozedur kann nur eine einzelne Analytic Server-Serviceinstanz installieren. Sie können weitere Instanzen hinzufügen, nachdem die Erstinstallation abgeschlossen ist.
- Die automatische Offlineinstallationsprozedur unterstützt keine Installation von Analytic Server auf einem Kerberos-fähigen Cluster.
- Die automatische Offlineinstallationsprozedur unterstützt keine Installation von Analytic Server unter HPD 3.0 oder 3.1.

Diese Einschränkungen gelten nicht für manuelle HDP- oder Ubuntu-Installationen.

1. Navigieren Sie zur IBM Passport Advantage®-Web-Site und laden Sie die selbstextrahierende Binärdatei auf einen Computer herunter, der auf https://ibm-open-platform.ibm.com zugreifen kann.

Tabelle 3. Selbstextrahierende Analytic Server-Binärdatei

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.2.1 für Hortonworks Data Platform 2.5, 2.6, 3.0 und 3.1 Linux x86-64 (Englisch)	`spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.1 für Hortonworks Data Platform 2.6, 3.0 und 3.1 Linux on System p LE (Englisch)	`spss_as-3.2.1.0-hdp2.6-3.1-lppc64.bin`

2. Führen Sie die ausführbare Binärdatei aus, die Sie in Schritt 1 heruntergeladen haben, und geben Sie eine Offlineinstallation an. Eine Offlineinstallation lädt die RPM- oder DEB-Dateien herunter, die später im Installationsprozess erforderlich sind, und sollte auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann. Die heruntergeladenen Dateien befinden sich im aktuellen Verzeichnis für ausführbare Binärdateien: `./IBM-SPSS-AnalyticServer`.
3. Kopieren Sie den gesamten Inhalt des Verzeichnisses für ausführbare Binärdateien (`./IBM-SPSS-AnalyticServer`) von dem Computer mit Internetzugriff auf den Ambari-Managerknoten (der sich hinter der Firewall befindet).
4. Prüfen Sie mit dem folgenden Befehl auf dem Ambari-Managerknoten, ob der Ambari-Server zurzeit aktiv ist:

```
ambari-server status
```

5. Installieren Sie das Tool, mit dem ein lokales Yum-Repository erstellt wird, auf dem Ambari-Managerknoten und auf allen anderen Knoten, auf denen Sie Analytic Server bereitstellen wollen.

   ```
   yum install createrepo (RHEL, CentOS)
   ```

   oder

   ```
   apt-get install dpkg-dev (Ubuntu)
   ```

6. Führen Sie die ausführbare Binärdatei `./IBM-SPSS-AnalyticServer/packages/spss_as-ambari-offlineinstall.bin` auf dem Ambari-Managerknoten aus. Während der Installation prüft die ausführbare Binärdatei, ob sich die erforderlichen Analytic Server-RPM/DEB-Dateien im Paketverzeichnis befinden. Die erforderlichen RPM-Dateien hängen von Ihrer Verteilung, Version und Architektur ab.

   HDP 2.5, 2.6, 3.0 und 3.1 (x86_64)
   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.x86_64.rpm

   HDP 2.6, 3.0 und 3.1 (PPC64LE)
   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.ppc64le.rpm

   HDP 2.5, 2.6, 3.0 und 3.1 (Ubuntu)
   IBM-SPSS-AnalyticServer-ambari-2.x_3.2.1.0_amd64.deb

   IBM-SPSS-AnalyticServer_1_amd64.deb

Während der Installation werden Sie zur Eingabe der Analytic Server-Version, des JDBC-Treibers, der Spark-Version, der Hive-Version usw. aufgefordert.

Manuelle Installation unter HDP (RHEL, SLES)

Der allgemeine Workflow für eine manuelle Offlineinstallation unter HDP (RHEL, SLES) ist folgender:

1. Navigieren Sie zur IBM Passport Advantage®-Web-Site und laden Sie die selbstextrahierende Binärdatei auf einen Computer herunter, der auf https://ibm-open-platform.ibm.com zugreifen kann.

Tabelle 4. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.2.1 für Hortonworks Data Platform 2.5, 2.6, 3.0 und 3.1 Linux x86-64 (Englisch)	`spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.1 für Hortonworks Data Platform 2.6, 3.0 und 3.1 Linux on System p LE (Englisch)	`spss_as-3.2.1.0-hdp2.6-3.1-lppc64.bin`

2. Führen Sie die ausführbare Binärdatei aus, die Sie in Schritt 1 heruntergeladen haben, und geben Sie eine Offlineinstallation an. Eine Offlineinstallation lädt die RPM-Dateien herunter, die später im Installationsprozess erforderlich sind, und sollte auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann. Die heruntergeladenen Dateien befinden sich im aktuellen Verzeichnis für ausführbare Binärdateien: `./IBM-SPSS-AnalyticServer`.
3. Kopieren Sie den gesamten Inhalt des Verzeichnisses für ausführbare Binärdateien (`./IBM-SPSS-AnalyticServer`) von dem Computer mit Internetzugriff in das Verzeichnis `<installierbares_AS-Ausgangsverzeichnis>` auf dem Ambari-Managerknoten (dieser befindet sich hinter der Firewall).
4. Prüfen Sie mit dem folgenden Befehl auf dem Ambari-Managerknoten, ob der Ambari-Server zurzeit aktiv ist:

   ```
   ambari-server status
   ```

5. Installieren Sie das Tool, mit dem Sie ein lokales yum-Repository erstellen können.

   ```
   yum install createrepo (RHEL, CentOS)
   ```

   oder

   ```
   zypper install createrepo (SLES)
   ```

6. Erstellen Sie ein Verzeichnis, das als Repository für die Analytic Server-RPM-Dateien verwendet wird. Siehe das folgende Beispiel.

   ```
   mkdir /home/root/repos/IBM-SPSS-AnalyticServer/x86_64
   ```

7. Kopieren Sie die erforderlichen Analytic Server-RPM-Dateien in das neue Verzeichnis. Die erforderlichen RPM-Dateien hängen von Ihrer Verteilung, Version und Architektur ab.

   HDP 2.5, 2.6, 3.0 und 3.1 (x86_64)

   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.x86_64.rpm

   HDP 2.6, 3.0 und 3.1 (PPC64LE)

   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.ppc64le.rpm

8. Erstellen Sie die Definition des lokalen Repositorys. Erstellen Sie z. B. eine Datei namens `IBM-SPSS-AnalyticServer-3.2.1.0.repo` mit dem folgenden Inhalt in `/etc/yum.repos.d/` (für RHEL, CentOS) oder `/etc/zypp/repos.d/` (für SLES).

   ```
   [IBM-SPSS-AnalyticServer]
   name=IBM-SPSS-AnalyticServer
   baseurl=file:///{Pfad zum lokalen Repository}
   enabled=1
   gpgcheck=0
   protect=1
   ```

9. Erstellen Sie das lokale `YUM`-Repository.

   ```
   createrepo /home/root/repos/IBM-SPSS-AnalyticServer/x86_64 (RHEL, CentOS, SLES)
   ```

10. Geben Sie im Befehlsfenster eines Rootbenutzers die folgenden Befehle ein: `cd` (zu `<installierbares_AS-Ausgangsverzeichnis>/IBM-SPSS-AnalyticServer`) und `run ./offLineInstall.sh`. Das Script liest auf Platte gespeicherte Antworten auf den zuvor ausgeführten Installationsbefehl für die ausführbare Binärdatei und setzt den entsprechenden Plattformbefehl ab (zur RPM-Installation).

    Anmerkung: Schritt 11 gilt nur bei einer extern verwalteten MySQL-Umgebung.

11. Führen Sie das Script `add_mysql_user.sh` auf dem Knoten/Host aus, auf dem die MySQL-Instanz, die als `AS_MetaStore` verwendet wird, installiert ist.

    a. Kopieren Sie das `add_mysql_user.sh`-Script aus `<INSTALLIERBARES_AS-AUSGANGSVERZEICHNIS>/IBM-SPSS-AnalyticServer` auf den Knoten/Host, auf dem die MySQL-Instanz als `AS_MetaStore` installiert ist.

       - Führen Sie das Script `add_mysql_user.sh` auf dem MySQL-Knoten/Host aus. Beispiel: `./add_mysql_user.sh -u as_user -p spss -d aedb`

       Hinweise:

       - Der Benutzername und das Kennwort müssen mit dem Datenbankbenutzernamen und -kennwort übereinstimmen, die für `AS_Metastore` in der Ambari-Konfigurationsanzeige eingegeben wurden.
       - Das Script `add_mysql_user.sh` kann manuell aktualisiert werden, um Befehle abzusetzen (bei Bedarf).
       - Verwenden Sie bei der Ausführung des Scripts `add_mysql_user.sh` für eine geschützte MySQL-Datenbank (Rootbenutzerzugriff) die Parameter `-r` und `-t` zum Übergeben von `dbuserid` und `dbuserid_password`. Das Script verwendet `dbuserid` und `dbuserid_password` zum Durchführen von MySQL-Operationen.

Anmerkung: Die Einstellung `metadata.repository.url` in der Anzeige **AS_Configuration** (**Advanced analytics-meta**) muss so geändert werden, dass sie auf den MySQL-Datenbankhost verweist. Än-

dern Sie z. B. die JDBC-Einstellung `mysql://{Analytic-Metaspeicher-Host}/aedb?createDatabaseIfNotExist=true` in `mysql://{MySQL-Datenbank}/aedb?createDatabaseIfNotExist=true`.

12. Fügen Sie Ihrer Ambari-Repository-Datei `repoinfo.xml`, die sich in der Regel im Verzeichnis `/var/lib/ambari-server/resources/stacks/$stackName/$stackVersion/repos/` befindet, die folgenden Zeilen hinzu, damit das lokale yum-Repository verwendet wird.

```
<os type="host_os">
   <repo>
        <baseurl>file:///{Pfad zum lokalen Repository}/</baseurl>
        <repoid>IBM-SPSS-AnalyticServer</repoid>
        <reponame>IBM-SPSS-AnalyticServer-3.2.1.0</reponame>
   </repo>
</os>
```

    Ein Beispiel für den `{Pfad zum lokalen Repository}` könnte wie folgt aussehen:

    `/home/root/repos/IBM-SPSS-AnalyticServer/x86_64/`

13. Wiederholen Sie die folgenden Schritte für jeden Ambari-Nicht-Server-Clusterknoten.
    a. Kopieren Sie den gesamten Inhalt des entsprechenden `<installierbaren_AS-Ausgangsverzeichnisses>` von dem Computer mit Internetzugriff auf den Ambari-Nicht-Server-Clusterknoten.
    b. Installieren Sie das Tool, mit dem Sie ein lokales yum-Repository erstellen können.

       `yum install createrepo (RHEL, CentOS)`

       oder

       `zypper install createrepo (SLES)`

    c. Erstellen Sie ein Verzeichnis, das als Repository für die Analytic Server-RPM-Dateien verwendet wird. Siehe das folgende Beispiel.

       `mkdir /home/root/repos/IBM-SPSS-AnalyticServer/x86_64`

    d. Kopieren Sie die erforderlichen Analytic Server-RPM-Dateien in das neue Verzeichnis. Die erforderlichen RPM-Dateien hängen von Ihrer Verteilung, Version und Architektur ab.

       HDP 2.5, 2.6, 3.0 und 3.1 (x86_64)
       IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

       IBM-SPSS-AnalyticServer-3.2.1.0-1.x86_64.rpm

       HDP 2.6, 3.0 und 3.1 (PPC64LE)
       IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

       IBM-SPSS-AnalyticServer-3.2.1.0-1.ppc64le.rpm

    e. Erstellen Sie die Definition des lokalen Repositorys. Erstellen Sie z. B. eine Datei namens `IBM-SPSS-AnalyticServer-3.2.1.0.repo` mit dem folgenden Inhalt in `/etc/yum.repos.d/` (für RHEL, CentOS) oder `/etc/zypp/repos.d/` (für SLES).

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer
baseurl=file:///{Pfad zum lokalen Repository}
enabled=1
gpgcheck=0
protect=1
```

    f. Erstellen Sie das lokale YUM-Repository.

       `createrepo /home/root/repos/IBM-SPSS-AnalyticServer/x86_64 (RHEL, CentOS, SLES)`

14. Setzen Sie den Vorgang mit Schritt 3 im Abschnitt „Onlineinstallation" auf Seite 6 fort.

Manuelle Installation unter HDP (Ubuntu)

Der allgemeine Workflow für eine manuelle Offlineinstallation unter HDP (Ubuntu) ist folgender:

1. Navigieren Sie zur IBM Passport Advantage®-Web-Site und laden Sie die entsprechende selbstextrahierende Binärdatei für Ubuntu auf einen Computer herunter, der auf https://ibm-open-platform.ibm.com zugreifen kann.

Tabelle 5. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.2.1 für Hortonworks Data Platform 2.5, 2.6, 3.0 und 3.1 Linux x86-64 (Englisch)	`spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin`

2. Führen Sie die ausführbare Binärdatei aus, die Sie in Schritt 1 heruntergeladen haben, und geben Sie eine Offlineinstallation an. Eine Offlineinstallation lädt die DEB-Dateien herunter, die später im Installationsprozess erforderlich sind, und sollte auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann. Die heruntergeladenen Dateien befinden sich im aktuellen Verzeichnis für ausführbare Binärdateien: `./IBM-SPSS-AnalyticServer`.
3. Kopieren Sie den gesamten Inhalt des Verzeichnisses für ausführbare Binärdateien (`./IBM-SPSS-AnalyticServer`) von dem Computer mit Internetzugriff in das Verzeichnis `<installierbares_AS-Ausgangsverzeichnis>` auf dem Ambari-Managerknoten (dieser befindet sich hinter der Firewall).
4. Prüfen Sie mit dem folgenden Befehl auf dem Ambari-Managerknoten, ob der Ambari-Server zurzeit aktiv ist:

   ```
   ambari-server status
   ```

5. Erstellen Sie ein Verzeichnis `<lokales_Repository>`, das als Repository für die Analytic Server-DEB-Dateien verwendet wird. Beispiel:

   ```
   mkdir –p /usr/local/mydebs
   ```

6. Kopieren Sie die erforderlichen Analytic Server-DEB-Dateien in das Verzeichnis `<lokales_Repository>`.
   - `IBM-SPSS-AnalyticServer-ambari-2.x_3.2.1.0_amd64.deb`
   - `IBM-SPSS-AnalyticServer_1_amd64.deb`
7. Erstellen Sie das lokale Repository.
   a. Installieren Sie das Tool, mit dem Sie ein lokales Repository erstellen können:

      ```
      apt-get install dpkg-dev
      ```

   b. Generieren Sie die Quellenpaketdatei:

      ```
      cd <lokales_Repository>
      dpkg-scanpackages . /dev/null | gzip -9c > Packages.gz
      ```

   c. Erstellen Sie die Komponente (main) und Architektur (z. B. `binary-i386`,`binary-amd64`) Ihres lokalen Repositorys:

      ```
      mkdir -p <lokales_Repository>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/
      mkdir -p <lokales_Repository>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/
      ```

   d. Kopieren Sie das Quellenpaket:

      ```
      cp -fr <lokales_Repository>/Packages.gz <lokales_Repository>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/Packages
      cp -fr <lokales_Repository>/Packages.gz <lokales_Repository>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/Packages
      ```

8. Erstellen Sie die Definition des lokalen Repositorys. Erstellen Sie z. B. eine Datei namens `IBM-SPSS-AnalyticServer-3.2.1.0.list` mit dem folgenden Inhalt in `/etc/apt/sources.list.d`.

   ```
   deb file:/usr/local/mydebs ./
   ```

9. Führen Sie den folgenden Befehl aus, um die Repository-Liste zu aktualisieren:

   ```
   apt-get update
   ```

10. Installieren Sie IBM SPSS Analytic Server 3.2.1, indem Sie den folgenden Befehl ausführen:

    ```
    apt-get install ./IBM-SPSS-AnalyticServer-ambari-2.x
    ```

    Anmerkung: Führen Sie den vorherigen Befehl nicht in Ihrem Verzeichnis `<lokales_Repository>` aus, um zu prüfen, ob Ihr lokales Repository korrekt eingerichtet ist. Wenn die Installation das Paket nicht finden kann, heißt das, dass Ihr lokales Repository nicht ordnungsgemäß eingerichtet ist. (In diesem Fall müssen Sie alle vorherigen Schritte prüfen.)
11. Wiederholen Sie die folgenden Schritte für jeden Ambari-Nicht-Server-Clusterknoten.
    a. Erstellen Sie ein Verzeichnis `<lokales_Repository>`, das als Repository für die Analytic Server-DEB-Dateien verwendet wird. Beispiel:

       ```
       mkdir –p /usr/local/mydebs
       ```

b. Kopieren Sie den gesamten Inhalt des Verzeichnisses <lokales_Repository> von dem Computer mit dem Ambari-Managerknoten in das Verzeichnis <lokales_Repository> auf dem Ambari-Nicht-Server-Clusterknoten. Das Verzeichnis sollte die folgenden Dateien enthalten:
   - <lokales_Repository>/IBM-SPSS-AnalyticServer-ambari-2.x_3.2.1.0_amd64.deb
   - <lokales_Repository>/IBM-SPSS-AnalyticServer_1_amd64.deb
   - <lokales_Repository>/Packages.gz
   - <lokales_Repository>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/Packages
   - <lokales_Repository>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/Packages

c. Erstellen Sie die Definition des lokalen Repositorys. Erstellen Sie z. B. eine Datei namens IBM-SPSS-AnalyticServer-3.2.1.0.list mit dem folgenden Inhalt in /etc/apt/sources.list.d.

```
deb file:/usr/local/mydebs ./
```

12. Setzen Sie den Vorgang mit Schritt 3 im Abschnitt „Onlineinstallation" auf Seite 6 fort.

Installieren von Analytic Server in extern verwalteter MySQL-Umgebung

Der Analytic Server-Installationsprozess unterscheidet sich von einer normalen Installation, wenn in einer extern verwalteten MySQL-Umgebung installiert wird.

In den folgenden Schritten wird der Prozess der Installation von Analytic Server in einer extern verwalteten MySQL-Umgebung erläutert.

1. Navigieren Sie zur IBM Passport Advantage®-Website und laden Sie die selbstextrahierende Binärdatei, die Ihrem Stack, Ihrer Stackversion und Ihrer Hardwarearchitektur entspricht, auf einen Host innerhalb des Ambari-Clusters herunter.
2. Führen Sie die selbstextrahierende Binärdatei aus und folgen Sie den Anweisungen, um (optional) die Lizenz anzuzeigen. Akzeptieren Sie diese.
   a. Wählen Sie die Option **Online** aus.
   b. Wählen Sie nach Aufforderung die Option **External MySQL Database** aus.
3. Kopieren Sie das add_mysql_user.sh-Script aus <INSTALLIERBARES_AS-AUSGANGSVERZEICHNIS>/IBM-SPSS-AnalyticServer auf den Knoten/Host, auf dem die MySQL-Instanz als AS_MetaStore installiert ist.
   - Führen Sie das Script add_mysql_user.sh auf dem MySQL-Knoten/Host aus. Beispiel:
     ```
     ./add_mysql_user.sh -u as_user -p spss -d aedb
     ```

   Hinweise:
   - Der Benutzername und das Kennwort müssen mit dem Datenbankbenutzernamen und -kennwort übereinstimmen, die für AS_Metastore in der Ambari-Konfigurationsanzeige eingegeben wurden.
   - Das Script add_mysql_user.sh kann manuell aktualisiert werden, um Befehle abzusetzen (bei Bedarf).
   - Verwenden Sie bei der Ausführung des Scripts add_mysql_user.sh für eine geschützte MySQL-Datenbank (Rootbenutzerzugriff) die Parameter -r und -t zum Übergeben von dbuserid und dbuserid_password. Das Script verwendet dbuserid und dbuserid_password zum Durchführen von MySQL-Operationen.
4. Starten Sie Ihren Ambari-Server erneut.
   ```
   ambari-server restart
   ```
5. Fügen Sie den Service AnalyticServer in der Ambari-Konsole als normal hinzu. (Geben Sie die gleichen Angaben für Datenbankbenutzername und -kennwort wie in Schritt 3 ein.)

   Anmerkung: Die Einstellung metadata.repository.url in der Anzeige **AS_Configuration** (**Advanced analytics-meta**) muss so geändert werden, dass sie auf den MySQL-Datenbankhost verweist. Ändern

Sie z. B. die JDBC-Einstellung `mysql://{Analytic-Metaspeicher-Host}/aedb?createDatabaseIfNotExist=true` in `mysql://{MySQL-Datenbank}/aedb?createDatabaseIfNotExist=true`.

Konfiguration

Nach der Installation können Sie Analytic Server optional über die Ambari-Benutzerschnittstelle konfigurieren und verwalten.

Anmerkung: Für Analytic Server-Dateipfade gelten die folgenden Konventionen:

- {AS-Stammverzeichnis} bezieht sich auf den Speicherort, an dem Analytic Server bereitgestellt wird, z. B. `/opt/IBM/SPSS/AnalyticServer/3.2`.
- {AS-Serverstammverzeichnis} bezieht sich auf den Speicherort der Konfigurations-, Protokoll- und Serverdateien, z. B. `/opt/IBM/SPSS/AnalyticServer/3.2/ae_wlpserver/usr/servers/aeserver`.
- {AS-Ausgangsverzeichnis} bezieht sich auf den HDFS-Speicherort, der von Analytic Server als Stammordner verwendet wird.

Sicherheit

Konfigurieren einer LDAP-Registry

Die LDAP-Registry ermöglicht Ihnen die Authentifizierung von Benutzern mit einem externen LDAP-Server wie beispielsweise Active Directory oder OpenLDAP.

Wichtig: Ein LDAP-Benutzer muss in Ambari als Analytic Server-Administrator angegeben werden.

Im Folgenden finden Sie ein Beispiel für eine LDAP-Registry (ldapRegistry) für OpenLDAP.

```
<ldapRegistry
    baseDN="ou=people,dc=aeldap,dc=org"
    ldapType="Custom"
    port="389"
    host="server"
    id="OpenLDAP"
    bindDN="cn=admin,dc=aeldap,dc=org"
    bindPassword="{xor}Dz4sLG5tbGs="
    searchTimeout="300000m"
    recursiveSearch="true">
    <customFilters
        id="customFilters"
        userFilter="(&amp;(uid=%v)(objectClass=inetOrgPerson))"
        groupFilter="(&amp;(cn=%v)(|(objectclass=organizationalUnit)))"
        groupMemberIdMap="posixGroup:memberUid"/>
</ldapRegistry>
```

Das folgende Beispiel stellt Analytic Server-Authentifizierung mit Active Directory bereit:

```
<ldapRegistry id="Microsoft Active Directory" realm="ibm"
  host="host"
  port="389"
  baseDN="cn=users,dc=adtest,dc=mycompany,dc=com"
  bindDN="cn=administrator,cn=users,dc=adtest,dc=mycompany,dc=com"
  bindPassword ="adminpassword"
  ldapType="Custom"
    <customFilters
    userFilter="(&amp;(sAMAccountName=%v)(objectcategory=user))"
    groupFilter="(&amp;(cn=%v)(objectcategory=group))"
    userIdMap="user:sAMAccountName"
    groupIdMap="*:cn"
    groupMemberIdMap="memberOf:member" />
</ldapRegistry>
```

Anmerkung: Oft ist es hilfreich, die LDAP-Konfiguration mit einem LDAP-Viewer eines anderen Anbieters zu prüfen.

Das folgende Beispiel stellt WebSphere Liberty-Profilauthentifizierung mit Active Directory bereit:

```
<ldapRegistry id="ldap" realm="SampleLdapADRealm"
    host="ldapserver.mycity.mycompany.com" port="389" ignoreCase="true"
    baseDN="cn=users,dc=adtest,dc=mycity,dc=mycompany,dc=com"
    bindDN="cn=testuser,cn=users,dc=adtest,dc=mycity,dc=mycompany,dc=com"
    bindPassword="testuserpwd"
    ldapType="Microsoft Active Directory"
    sslEnabled="true"
    sslRef="LDAPSSLSettings">
    <actived Filters
       userFilter="(&amp;(sAMAccountName=%v)(objectcategory=user))"
       groupFilter="(&amp;(cn=%v)(objectcategory=group))"
       userIdMap="user:sAMAccountName"
       groupIdMap="*:cn"
       groupMemberIdMap="memberOf:member" >
    </actived Filters>
</ldapRegistry>

<ssl id="LDAPSSLSettings" keyStoreRef="LDAPKeyStore" trustStoreRef="LDAPTrustStore" />

<keyStore id="LDAPKeyStore" location="${server.config.dir}/LdapSSLKeyStore.jks"
          type="JKS" password="{xor}CDo9Hgw=" />

<keyStore id="LDAPTrustStore" location="${server.config.dir}/LdapSSLTrustStore.jks"
          type="JKS" password="{xor}CDo9Hgw=" />
```

Hinweise:

- Unterstützung für LDAP in Analytic Server wird durch WebSphere Liberty gesteuert. Weitere Informationen finden Sie in LDAP-Benutzerregistrys in Liberty konfigurieren.
- Wenn LDAP mit SSL geschützt ist, befolgen Sie die Anweisungen im Abschnitt "Konfigurieren einer SSL-Verbindung von Analytic Server zu LDAP".

Konfigurieren einer SSL-Verbindung (Secure Socket Layer) von Analytic Server zu LDAP

Wenn Sie Apache Directory Server (`ads`) während der Installation von Analytic Server als LDAP-Option auswählen und die Standardkonfiguration verwenden, wird Apache Directory Server mit konfiguriertem und aktiviertem SSL installiert (Analytic Server verwendet automatisch SSL für die Kommunikation mit Apache Directory Server).

Konfigurieren Sie SSL mit den folgenden Schritten, wenn während der Installation von Analytic Server eine der anderen LDAP-Optionen ausgewählt wird (z. B. wenn ein externer LDAP-Server verwendet wird).

1. Melden Sie sich an allen Analytic Server-Computern als Analytic Server-Benutzer an und erstellen Sie ein allgemeines Verzeichnis für SSL-Zertifikate.

   Anmerkung: Der Analytic Server-Benutzer ist standardmäßig "as_user". Weitere Informationen finden Sie in der Ambari-Konsole auf der Registerkarte "Admin" unter **Service accounts**.

2. Kopieren Sie die Keystore- und Truststore-Dateien auf allen Analytic Server-Computern in dasselbe allgemeine Verzeichnis. Fügen Sie dem Truststore außerdem das Zertifikat einer Zertifizierungsstelle des LDAP-Clients hinzu. Es folgen einige Beispielanweisungen.

   ```
   mkdir /home/as_user/security
   cd /home/as_user/security
   openssl s_client -connect <LDAP-Hostname>:636 -showcerts > client.cert
   $JAVA_HOME/bin/keytool -import -file ./client.cert -alias ldapCA -keystore mytrust.jks
   password : changeit
   ```

Anmerkung: JAVA_HOME ist dieselbe Java-Ausführungsumgebung (JRE), die auch zum Starten von Analytic Server verwendet wird.

3. Kennwörter können mit dem Tool securityUtility codiert werden, um ihre Werte zu verschlüsseln. Dieses Tool befindet sich in `{AS-Stammverzeichnis}/ae_wlpserver/bin`. Es folgt ein Beispiel.

```
securityUtility encode changeit
        {xor}PDc+MTg6Nis=
```

4. Melden Sie sich an der Ambari-Konsole an und aktualisieren Sie die Analytic Server-Konfigurationseinstellung **ssl.keystore.config** mit den korrekten SSL-Konfigurationseinstellungen. Es folgt ein Beispiel.

```
<ssl id="defaultSSLConfig" keyStoreRef="defaultKeyStore" trustStoreRef="defaultTrustStore"
      clientAuthenticationSupported="true"/>
          <keyStore id="defaultKeyStore" location="/home/as_user/security/mykey.jks" type="JKS"
                    password="{xor}Ozo5PiozKxYdEgwPDAweDG1uDz4sLCg7"/>
          <keyStore id="defaultTrustStore" location="/home/as_user/security/mytrust.jks" type="JKS"
                    password="{xor}PDc+MTg6Nis="/>
```

Anmerkung: Verwenden Sie den absoluten Pfad zu den Keystore- und Truststore-Dateien.

5. Aktualisieren Sie die Konfigurationseinstellung **security.config** von Analytic Server mit den korrekten LDAP-Konfigurationseinstellungen. Setzen Sie beispielsweise im Element **ldapRegistry** das Attribut **sslEnabled** auf `true` und das Attribut **sslRef** auf `defaultSSLConfig`.

Konfigurieren von Kerberos

Analytic Server unterstützt Kerberos über Ambari.

Anmerkung: IBM SPSS Analytic Server unterstützt nicht Kerberos Single-Sign-On (SSO) bei Verwendung zusammen mit Apache Knox.

1. Sie können im Kerberos-Benutzerrepository für alle Benutzer, denen Sie Zugriff auf Analytic Server erteilen möchten, Konten erstellen.
2. Erstellen Sie dieselben Konten (aus dem vorherigen Schritt) auf dem LDAP-Server.
3. Erstellen Sie für jeden im vorherigen Schritt erstellten Benutzer auf jedem einzelnen Analytic Server-Knoten und Hadoop-Knoten ein Betriebssystembenutzerkonto.
   - Stellen Sie sicher, dass die Benutzer-ID für diese Benutzer auf allen Computern übereinstimmt. Dies können Sie prüfen, indem Sie sich mithilfe des Befehls "kinit" an jedem der Konten anmelden.
   - Stellen Sie sicher, dass die Benutzer-ID der YARN-Einstellung "Minimum user ID for submitting job" entspricht. Dies ist der Parameter **min.user.id** in container-executor.cfg. Wenn **min.user.id** beispielsweise auf 1000 gesetzt ist, muss die Benutzer-ID jedes erstellten Benutzerkontos größer-gleich 1000 sein.
4. Erstellen Sie in HDFS einen Benutzerausgangsordner für alle Principals in Analytic Server. Wenn Sie beispielsweise dem Analytic Server-System "testuser1" hinzufügen, erstellen Sie in HDFS einen Ausgangsordner wie `/user/testuser1` und stellen Sie sicher, dass "testuser1" über Lese- und Schreibberechtigungen für diesen Ordner verfügt.
5. [Optional] Wenn Sie HCatalog-Datenquellen verwenden wollen und Analytic Server auf einem anderen Computer als Hive-Metaspeicher installiert ist, müssen Sie in HDFS die Identität des Hive-Clients annehmen.
   a. Navigieren Sie in der Ambari-Konsole zur Registerkarte **Configs** des HDFS-Service.
   b. Bearbeiten Sie den Parameter **hadoop.proxyuser.hive.groups** so, dass er den Wert `*` hat oder eine Gruppe enthält, die alle Benutzer umfasst, die sich an Analytic Server anmelden können.
   c. Bearbeiten Sie den Parameter **hadoop.proxyuser.hive.hosts** so, dass er den Wert `*` hat oder die Liste der Hosts enthält, auf denen der Hive-Metaspeicher und alle Instanzen von Analytic Server als Service installiert sind.
   d. Starten Sie den HDFS-Service erneut.

Nachdem Sie diese Schritte ausgeführt haben und Analytic Server installiert ist, konfiguriert Analytic Server Kerberos automatisch im Hintergrund.

Konfigurieren von HAProxy für Kerberos-SSO (Single Sign On)

1. Konfigurieren und starten Sie HAProxy wie in der Dokumentation zu HAProxy unter http://www.haproxy.org/#docs beschrieben.
2. Erstellen Sie den Kerberos-Prinzipal (HTTP/<Proxy-Hostname>@<Realm>) und die Chiffrierschlüsseldatei für den HAProxy-Host, wobei <Proxy-Hostname> der vollständige Name des HAProxy-Hosts und <Realm> der Kerberos-Realm ist.
3. Kopieren Sie die Chiffrierschlüsseldatei als `/etc/security/keytabs/spnego_proxy.service.keytab` auf alle Analytic Server-Hosts.
4. Aktualisieren Sie die Berechtigungen für diese Datei auf allen Analytic Server-Hosts. Es folgt ein Beispiel.

   ```
   chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
   chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
   ```

5. Öffnen Sie die Amabri-Konsole und aktualisieren Sie die folgenden Eigenschaften im Analytic Server-Abschnitt 'Custom analytics.cfg'.

   ```
   web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
   web.authentication.kerberos.principal=HTTP/<vollständiger Name des Proxy-Computers>@<Realm>
   ```

6. Speichern Sie die Konfiguration und starten Sie alle Analytic Server-Services über die Amabri-Konsole erneut.

Benutzer können sich jetzt über die Option **Single sign on log in** auf dem Anmeldebildschirm von IBM SPSS Analytic Server bei Analytic Server anmelden.

Aktivieren des Kerberos-Identitätswechsels

Durch Identitätswechsel kann ein Thread in einem Sicherheitskontext ausgeführt werden, der sich vom Sicherheitskontext des Prozesses unterscheidet, der der Threadeigner ist. Beispielsweise können Hadoop-Jobs mithilfe von Identitätswechsel über einen anderen Benutzer als den Analytic Server-Standardbenutzer (`as_user`) ausgeführt werden. So aktivieren Sie den Kerberos-Identitätswechsel:

1. Fügen Sie HDFS (oder den Hive-Servicekonfigurationen) Konfigurationsattribute für Identitätswechsel bei Ausführung in einem Kerberos-aktivierten Cluster hinzu. Bei HDFS müssen der HDFS-Datei `core-site.xml` die folgenden Eigenschaften hinzugefügt werden:

   ```
   hadoop.proxyuser.<Analytic_Server-Service-Principal-Name> .hosts = *
   hadoop.proxyuser.<Analytic_Server-Service-Principal-Name> .groups = *
   ```

   Dabei ist `<Analytic_Server-Service-Principal-Name>` der Standardwert von `as_user`, der im Konfigurationsfeld `Analytic_Server_User` von Analytic Server angegeben ist.

   Die folgenden Eigenschaften müssen der HDFS-Datei `core-site.xml` hinzugefügt werden, wenn von HDFS über Hive/HCatalog auf Daten zugegriffen wird:

   ```
   hadoop.proxyuser.hive.hosts = *
   hadoop.proxyuser.hive.groups = *
   ```

2. Wenn Analytic Server für die Verwendung eines anderen Benutzernamens als `as_user` konfiguriert ist, müssen Sie die Eigenschaftsnamen ändern, um den Benutzernamen widerzuspiegeln (z. B. `hadoop-.proxyuser.xxxxx.hosts`, wobei `xxxxx` der konfigurierte Benutzername ist, der in der Analytic Server-Konfiguration angegeben ist).
3. Führen Sie den folgenden Befehl in einer Befehlsshell auf dem Analytic Server-Knoten aus:

   ```
   hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
   ```

Aktivieren mehrerer Realms

Beim Definieren mehrerer Realms ist die Einstellung **as.kdc.realms** erforderlich. Die Werte für **as.kdc.realms** befinden sich im Abschnitt "Advanced analytics.cfg" von Analytic Server der Ambari-Konsole.

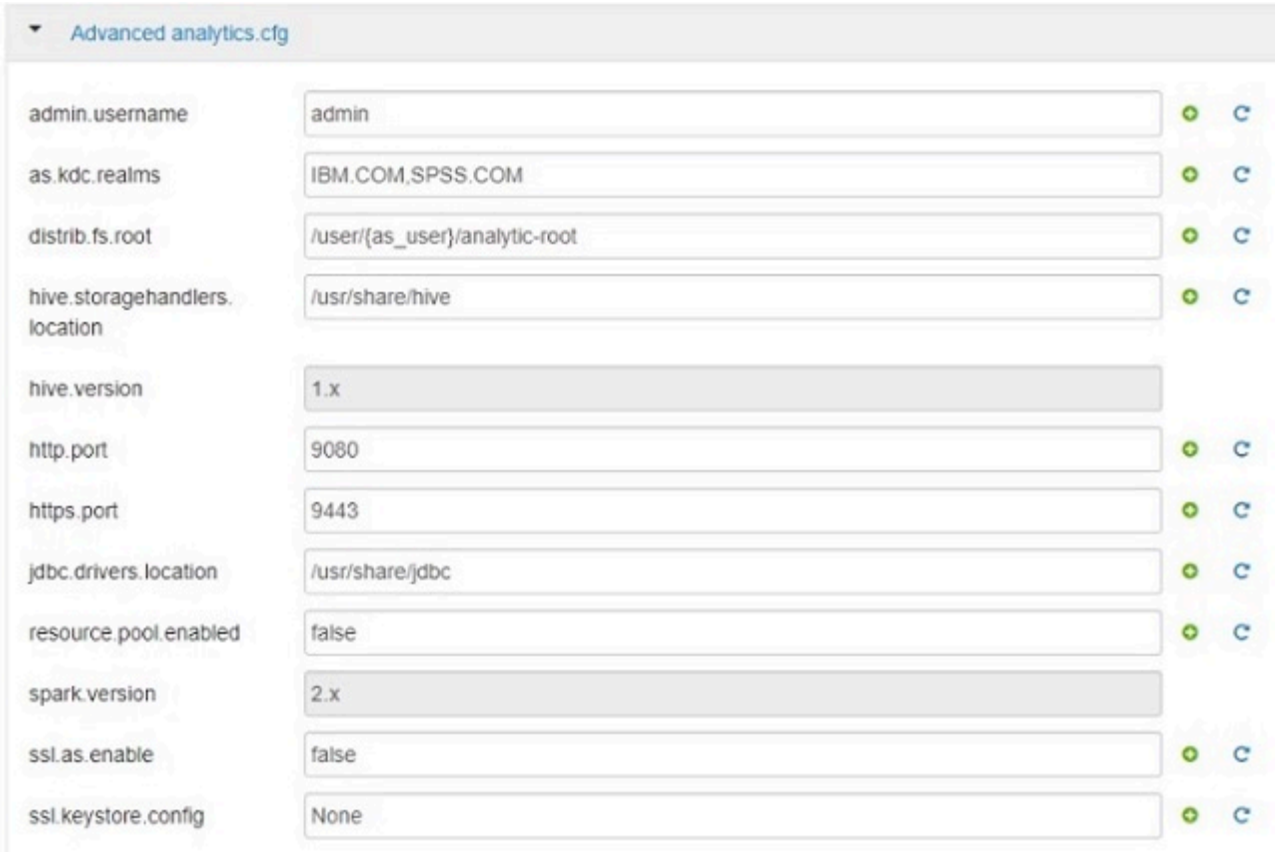

Abbildung 3. Einstellungen für Advanced analytics.cfg

Es werden mehrere Realmnamen unterstützt, wenn diese durch Kommas voneinander getrennt sind. Die angegebenen Kerberos-Realmnamen entsprechen Benutzernamen und sind Benutzernamen zugeordnet. Die Benutzernamen `UserOne@us.ibm.com` und `UserTwo@eu.ibm.com` würden beispielsweise den Realms `us.ibm.com, eu.ibm.com` entsprechen.

Vertrauensstellungen, die Kerberos-Realms übergreifen, müssen konfiguriert werden, wenn mehrere Realms als **Kerberos-Realmname** angegeben sind. Die Eingabe des Benutzernamens während der Anmeldeaufforderung der Analytic Server-Konsole erfolgt ohne das Suffix des Realmnamens. Infolgedessen wird Benutzern bei Verwendung mehrerer Realms die Dropdown-Liste **Realms** angezeigt, aus der diese den Realm auswählen können.

Anmerkung: Wenn nur ein Realm angegeben ist, wird Benutzern bei der Anmeldung bei Analytic Server die Dropdown-Liste **Realms** nicht angezeigt.

Inaktivieren von Kerberos

1. Sie können Kerberos in der Ambari-Konsole inaktivieren.
2. Stoppen Sie den Analytic Server-Service.
3. Entfernen Sie die folgenden Parameter aus **Custom analytics.cfg**.

```
default.security.provider
hdfs.keytab
```

```
hdfs.user
java.security.krb5.conf
as.db.connect.method
web.authentication.kerberos.keytab
web.authentication.kerberos.principal
```

4. Klicken Sie auf **Save** und starten Sie den Analytic Server-Service erneut.

Aktivieren von SSL-Verbindungen (Secure Socket Layer) zur Analytic Server-Konsole

Standardmäßig generiert Analytic Server selbst signierte Zertifikate, um SSL (Secure Socket Layer) zu aktivieren. Wenn Sie die selbst signierten Zertifikate akzeptieren, können Sie so über den sicheren Port auf die Analytic Server-Konsole zugreifen. Für einen sichereren HTTPS-Zugriff müssen Sie Zertifikate eines anderen Anbieters installieren.

Führen Sie die folgenden Schritte aus, um Zertifikate eines anderen Anbieters zu installieren.

1. Kopieren Sie auf allen Analytic Server-Knoten die Keystore- und Truststore-Zertifikate eines anderen Anbieters in dasselbe Verzeichnis, beispielsweise in `/home/as_user/security`.

   Anmerkung: Der Analytic Server-Benutzer muss über Lesezugriff auf dieses Verzeichnis verfügen.
2. Navigieren Sie auf der Registerkarte "Ambari Services" zur Registerkarte "Configs" des Analytic Server-Service.
3. Bearbeiten Sie den Parameter **ssl.keystore.config**.

   ```
   <ssl id="defaultSSLConfig"
        keyStoreRef="defaultKeyStore"
        trustStoreRef="defaultTrustStore"
        clientAuthenticationSupported="true"/>
   <keyStore id="defaultKeyStore"
             location="<KEYSTOREPOSITION>"
             type="<TYP>"
             password="<KENNWORT>"/>
   <keyStore id="defaultTrustStore"
             location="<TRUSTSTOREPOSITION>"
             type="<TYP>"
             password="<KENNWORT>"/>
   ```

   Ersetzen Sie Folgendes:

   - <KEYSTOREPOSITION> durch die absolute Position des Keystores. Beispiel: `/home/as_user/security/mykey.jks`
   - <TRUSTSTOREPOSITION> durch die absolute Position des Truststores. Beispiel: `/home/as_user/security/mytrust.jks`
   - <TYP> durch den Typ des Zertifikats. Beispiel: JKS, PKCS12 usw.
   - <KENNWORT> durch das verschlüsselte Kennwort im Base64-Verschlüsselungsformat. Für die Codierung können Sie das Tool securityUtility verwenden. Beispiel: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility encode <Kennwort>`

   Wenn Sie ein selbst signiertes Zertifikat generieren möchten, können Sie das Tool securityUtility verwenden. Beispiel: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility createSSLCertificate --server=myserver --password=mypassword --validity=365 --subject=CN=mycompany,O=myOrg,C=myCountry`.

   Anmerkung: Sie müssen einen entsprechenden Hostdomänennamen für den `CN`-Wert angeben.

   Weitere Informationen zu securityUtility und anderen SSL-Einstellungen finden Sie in der Dokumentation zum WebSphere Liberty-Profil.
4. Klicken Sie auf **Save** und starten Sie den Analytic Server-Service erneut.

Kommunizieren mit Apache Hive über SSL

Sie müssen die Datei `hive.properties` aktualisieren, um über eine SSL-Verbindung mit Apache Hive zu kommunizieren. Wenn die Hochverfügbarkeit in Ihrer Apache Hive-Umgebung aktiviert ist, können Sie alternativ die Hochverfügbarkeitsparameter auf der Analytic Server-Hauptseite für Datenquellen auswählen.

Aktualisieren der Datei 'hive.properties'

1. Öffnen Sie die Datei `hive.properties`. Die Datei befindet sich an der folgenden Position: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/database`
2. Suchen Sie die folgende Zeile:

   ```
   jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password}
   ```

3. Aktualisieren Sie die Zeile, indem Sie die folgenden **fett** angegebenen Informationen hinzufügen:

   ```
   jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password};
   ssl=true;sslTrustStore=PfadzurTruststore-Datei;trustStorePassword=xxxTruststore-Kennwort
   ```

4. Speichern Sie die Datei `hive.properties`.

Aktivieren der Unterstützung für Essentials for R

Analytic Server unterstützt das Scoren von R-Modellen und das Ausführen von R-Scripts.

So konfigurieren Sie die Unterstützung für R nach einer erfolgreichen Analytic Server-Installation:

1. Richten Sie die Serverumgebung für Essentials for R ein.

   RedHat Linux x86_64

   Führen Sie die folgenden Befehle aus:

   ```
   yum update
   yum install -y zlib zlib-devel
   yum install -y bzip2 bzip2-devel
   yum install -y xz xz-devel
   yum install -y pcre pcre-devel
   yum install -y libcurl libcurl-devel
   ```

   Ubuntu Linux

   Führen Sie die folgenden Befehle aus:

   ```
   apt-get update
   apt-get install -y zlib1g-dev
   apt-get install -y libreadline-dev
   apt-get install -y libxt-dev
   apt-get install -y bzip2
   apt-get install -y libbz2-dev
   apt-get install -y liblzma-dev
   apt-get install -y libpcre3 libpcre3-dev
   apt-get install -y libcurl4-openssl-dev
   apt-get install -y liblzma-dev
   apt-get install -y libpcre3 libpcre3-dev
   apt-get install -y libcurl4-openssl-dev
   ```

   SUSE Linux

   Die Installation von Essentials for R unter SUSE erfordert eine kompatible FORTRAN-Version. Diese ist in der Regel in den konfigurierten ZYPPER-Repositorys nicht verfügbar, sondern nur auf den SUSE-SDK-Datenträgern). Daher schlägt eine Ambari-Installation für Essentials for R auf einem SUSE-Server fehl, da FORTRAN nicht installiert werden kann. Führen Sie zum Einrichten unter SUSE die folgenden Schritte aus:

   a. Installieren Sie GCC-C++.

      ```
      zypper install gcc-c++
      ```

   b. Installieren Sie GCC-FORTRAN. Die erforderlichen RPM-Dateien können von den SUSE-SDK-Datenträgern kopiert werden und müssen in der folgenden Reihenfolge installiert werden:

      ```
      zypper install libquadmath0-4.7.2_20130108-0.19.3.x86_64.rpm
      zypper install libgfortran3-4.7.2_20130108-0.19.3.x86_64.rpm
      zypper install gcc43-fortran-4.3.4_20091019-0.37.30.x86_64.rpm
      zypper install gcc-fortran-4.3-62.200.2.x86_64.rpm
      ```

   c. Führen Sie den folgenden Befehl aus, um die Bibliotheken für Essentials for R zu installieren:

```
R_PREFIX=/opt/ibm/spss/R
cd $R_PREFIX
rm -fr $R_PREFIX/r_libs
mkdir -p $R_PREFIX/r_libs
cd $R_PREFIX/r_libs
wget https://zlib.net/fossils/zlib-1.2.11.tar.gz --no-check-certificate
tar zxvf zlib-1.2.11.tar.gz
cd zlib-1.2.11/
./configure
make && make install
cd $R_PREFIX/r_libs
wget http://www.bzip.org/1.0.6/bzip2-1.0.6.tar.gz
tar xzvf bzip2-1.0.6.tar.gz
cd bzip2-1.0.6
sed "s|^CC=gcc|CC=gcc -fPIC|" -i ./Makefile
make -f Makefile-libbz2_so
make clean
make
make install
cd $R_PREFIX/r_libs
wget https://tukaani.org/xz/xz-5.2.3.tar.gz
tar xzvf xz-5.2.3.tar.gz
cd xz-5.2.3
./configure
make -j3
make install
cd $R_PREFIX/r_libs
wget http://ftp.pcre.org/pub/pcre/pcre-8.38.tar.gz
tar xzvf pcre-8.38.tar.gz
cd pcre-8.38
./configure --enable-utf8
make
make install
cd $R_PREFIX/r_libs
wget https://www.openssl.org/source/openssl-1.0.2l.tar.gz --no-check-certificate
tar zxvf openssl-1.0.2l.tar.gz
cd openssl-1.0.2l/
./config shared
make
make install
echo '/usr/local/ssl/lib' >> /etc/ld.so.conf
ldconfig
cd $R_PREFIX/r_libs
wget --no-check-certificate https://curl.haxx.se/download/curl-7.50.1.tar.gz
tar xzvf curl-7.50.1.tar.gz
cd curl-7.50.1
./configure --with-ssl
make -j3
make install
cd $R_PREFIX/r_libs
wget ftp://rpmfind.net/linux/opensuse/distribution/12.3/repo/oss/suse/x86_64/libgomp1-4.7.2_20130108-2.1.6.x86_64.rpm --no-check-certificate
rpm -ivh libgomp1-4.7.2_20130108-2.1.6.x86_64.rpm
```

2. Laden Sie das selbstextrahierende Archiv (BIN) für den RPM oder DEB für IBM SPSS Modeler Essentials for R herunter. Essentials for R ist als Download verfügbar (https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp). Wählen Sie die Ihrem Stack, Ihrer Stackversion und Ihrer Hardwarearchitektur entsprechende Datei aus.
3. Führen Sie die sich selbst entpackende Binärdatei aus und folgen Sie den Anweisungen, um (optional) die Lizenz anzuzeigen, diese zu akzeptieren und die Online- oder Offlineinstallation auszuwählen.

   Onlineinstallation
   Wählen Sie die Onlineinstallation aus, wenn Ihr Ambari-Server-Host und alle Knoten im Cluster auf https://ibm-open-platform.ibm.com zugreifen können.

   Offlineinstallation
   Wählen Sie die Offlineinstallation aus, wenn Ihr Ambari-Server-Host keinen Internetzugriff hat. Die Offlineinstallation lädt die erforderlichen RPM-Dateien herunter und sollte auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann. Die RPM-Dateien können dann auf den Ambari-Server-Host kopiert werden.

   a. Kopieren Sie die für Essentials for R erforderlichen RPM- oder DEB-Dateien an einen beliebigen Speicherort auf dem Ambari-Server-Host. Die erforderlichen RPM/DEB-Dateien hängen wie nachfolgend aufgelistet von Ihrer Verteilung, Version und Architektur ab.

      HDP 2.5, 2.6, 3.0 und 3.1 (x86_64)
      IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.1.0.0-1.x86_64.rpm

      HDP 2.6, 3.0 und 3.1 (PPC64LE)
      IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.1.0.0-1.ppc64le.rpm

      HDP 2.5, 2.6, 3.0 und 3.1 (Ubuntu)
      IBM-SPSS-ModelerEssentialsR-ambari-3.2.1.0_3.2.1.0_amd64.deb

b. Installieren Sie den RPM oder DEB. Im folgenden Beispiel installiert der Befehl Essentials for R unter HDP 2.6 (x86_64).

```
rpm -i IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.1.0.0-1.x86_64.rpm
```

Im folgenden Beispiel installiert der Befehl Essentials for R unter HDP 2.5 (Ubuntu).

```
dpkg -i IBM-SPSS-ModelerEssentialsR-ambari-3.2.1.0_3.2.1.0_amd64.deb
```

4. Starten Sie Ihren Ambari-Server erneut.

```
ambari-server restart
```

5. Melden Sie sich an Ihrem Ambari-Server an und installieren Sie SPSS Essentials for R über die Ambari-Konsole als Service. SPSS Essentials for R muss auf jedem Host installiert werden, auf dem Analytic Server und der Analytic-Metaspeicher installiert sind.

   Anmerkung: Ambari versucht, gcc-c++ und gcc-gfortran (RHEL) sowie gcc-fortran (SUSE) vor der Installation von R zu installieren. Diese Pakete sind als Abhängigkeiten für die Ambari-Servicedefinition von R deklariert. Stellen Sie sicher, dass die Server, auf denen R installiert und ausgeführt werden soll, zum Herunterladen der RPMs für gcc-c++ und gcc-[g]fortran RPMs konfiguriert sind oder dass auf ihnen die GCC- und FORTRAN-Compiler installiert sind. Wenn die Installation von Essentials for R fehlschlägt, installieren Sie diese Pakete vor der Installation von Essentials for R manuell.

6. Aktualisieren Sie den Analytic Server-Service.
7. Führen Sie das Script `update_clientdeps` unter Beachtung der Anweisungen in „Aktualisierung von Clientabhängigkeiten" auf Seite 29 aus.
8. Sie müssen Essentials for R auch auf dem Computer installieren, der SPSS Modeler Server hostet. Details finden Sie in der Dokumentation zu SPSS Modeler.

Aktivieren relationaler Datenbankquellen

Wenn Sie die JDBC-Treiber in einem gemeinsam genutzten Verzeichnis in allen Analytic Server-Metaspeichern und auf allen Analytic Server-Hosts bereitstellen, kann Analytic Server relationale Datenbankquellen verwenden. Standardmäßig wird hierzu das Verzeichnis `/usr/share/jdbc` verwendet.

Führen Sie die folgenden Schritte aus, um das gemeinsam genutzte Verzeichnis zu ändern.

1. Navigieren Sie auf der Registerkarte "Ambari Services" zur Registerkarte "Configs" des Analytic Server-Service.
2. Öffnen Sie den Abschnitt **Advanced analytics.cfg**.
3. Geben Sie in **`jdbc.drivers.location`** den Pfad zum gemeinsam genutzten Verzeichnis mit den JDBC-Treibern an.
4. Klicken Sie auf **Save**.
5. Stoppen Sie den Analytic Server-Service.
6. Klicken Sie auf **Refresh**.
7. Starten Sie den Analytic Server-Service.

Tabelle 6. Unterstützte Datenbanken

Datenbank	Unterstützte Versionen	JAR-Dateien für JDBC-Treiber	Anbieter
Amazon Redshift	8.0.2 oder später	`RedshiftJDBC41-1.1.6.1006.jar` oder später	Amazon
BigSQL	4.1.0.0 oder später	db2jcc.jar	IBM
DashDB	Bluemix-Service	`db2jcc.jar`	IBM
Db2 for Linux, UNIX, and Windows	11.1, 10.5, 10.1, 9.7	`db2jcc.jar`	IBM
Db2 z/OS	11, 10	`db2jcc.jar`, `db2_license_cisuz.jar`	IBM

Tabelle 6. Unterstützte Datenbanken (Forts.)

Datenbank	Unterstützte Versionen	JAR-Dateien für JDBC-Treiber	Anbieter
Greenplum	5	postgresql.jar	Greenplum
Hive	1.2, 2.1	`hive-jdbc-*.jar`	Apache
MySQL	5.6, 5.7	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7, 6.x	`nzjdbc.jar`	IBM
Oracle	12c, 11g R2 (11.2)	`ojdbc6.jar`, `orai18n.jar`	Oracle
SQL Server	2014, 2012, 2008 R2	`sqljdbc4.jar`	Microsoft
Teradata	15, 15.1	`tdgssconfig.jar`, `terajdbc4.jar`	Teradata

Hinweise

- Wenn Sie vor der Installation von Analytic Server eine Redshift-Datenquelle erstellt haben, müssen Sie die folgenden Schritte ausführen, damit die Redshift-Datenquelle verwendet werden kann.
  1. Öffnen Sie die Redshift-Datenquelle in der Analytic Server-Konsole.
  2. Wählen Sie die `Redshift`-Datenbankdatenquelle aus.
  3. Geben Sie die Redshift-Serveradresse ein.
  4. Geben Sie den Datenbanknamen und den Benutzernamen ein. Das Kennwort sollte automatisch ausgefüllt werden.
  5. Wählen Sie die Datenbanktabelle aus.
- BigSQL ist die IBM SQL-Schnittstelle für die Apache Hadoop-Umgebung. BigSQL ist keine relationale Datenbank, aber Analytic Server unterstützt über JDBC Zugriff darauf. (Die JDBC-JAR-Datei ist dieselbe Datei, die für Db2 verwendet wird.)

  Eine gängige Verwendung von BigSQL mit Analytic Server ist der Zugriff auf BigSQL Hadoop/HBase-Tabellen über eine HCatalog-Datenquelle.

Aktivieren von HCatalog-Datenquellen

Analytic Server bietet über Hive/HCatalog Unterstützung für zahlreiche Datenquellen. Für einige Quellen sind Schritte zur manuellen Konfiguration erforderlich.

1. Erfassen Sie die für die Aktivierung der Datenquelle erforderlichen JAR-Dateien. Für die Aktivierung der Unterstützung für Apache HBase und Apache Accumulo sind keine zusätzlichen Schritte erforderlich. Bei anderen NoSQL-Datenquellen wenden Sie sich an den Datenbankanbieter, um den Speicherhandler und die entsprechenden JAR-Dateien zu erhalten. Informationen zu unterstützten HCatalog-Datenquellen finden Sie im Abschnitt "Verwenden von HCatalog-Datenquellen" im IBM SPSS Analytic Server 3.2.1 Benutzerhandbuch.
2. Fügen Sie diese JAR-Dateien zum Verzeichnis `{HIVE_HOME}/auxlib` und zum Verzeichnis `/usr/share/hive` in allen Analytic Server-Metaspeichern und auf allen Analytic Server-Knoten hinzu.
3. Starten Sie den Hive-Metaspeicherservice erneut.
4. Aktualisieren Sie den Analytic-Metaspeicherservice.
5. Starten Sie jede einzelne Instanz des Analytic Server-Service erneut.

Hinweise:

- Der Analytic Server-Metaspeicher kann nicht auf demselben System wie der Hive-Metaspeicher installiert werden.
- Wenn Sie über eine HCatalog-Datenquelle in Analytic Server auf HBase-Daten zugreifen, muss der zugreifende Benutzer über Leseberechtigung für die HBase-Tabellen verfügen.

– In Umgebungen, die kein Kerberos verwenden, greift Analytic Server mit `as_user` (`as_user` muss Leseberechtigung für HBase haben) auf HBase zu.
– In Kerberos-Umgebungen müssen `as_user` und der angemeldete Benutzer eine Leseberechtigung für HBase-Tabellen haben.

NoSQL-Datenbanken

Analytic Server unterstützt NoSQL-Datenbanken, für die ein Hive-Speicherhandler vom Anbieter verfügbar ist.

Für die Aktivierung der Unterstützung für Apache HBase und Apache Accumulo sind keine zusätzlichen Schritte erforderlich.

Bei anderen NoSQL-Datenbanken wenden Sie sich an den Datenbankanbieter, um den Speicherhandler und die entsprechenden JAR-Dateien zu erhalten.

Dateibasierte Hive-Tabellen

Analytic Server unterstützt dateibasierte Hive-Tabellen, für die ein integrierter oder angepasster Hive SerDe (Parallel-Seriell- und Seriell-Parallel-Umsetzer) verfügbar ist.

Der Hive XML SerDe für die Verarbeitung von XML-Dateien befindet sich im Maven Central Repository unter http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde.

Jobs für MapReduce Version 2

Verwenden Sie die Einstellung **`preferred.mapreduce`** im Analytic Server-Abschnitt **Custom analytic.cfg**, um zu steuern, wie MapReduce-Jobs verarbeitet werden:

Tabelle 7. Angepasste Eigenschaften in 'analytics.cfg'

Eigenschaft	Beschreibung
`preferred.mapreduce`	Steuert die Methode, in der MapReduce-Jobs ausgeführt werden. Gültige Werte sind unter anderem: • `spark` • `m3r` • `hadoop` Beispiel: `preferred.mapreduce=spark`

Apache Spark

Wenn Sie Spark (Version 1.5 oder höher) verwenden möchten, müssen Sie die Eigenschaft `spark.version` während der Installation von Analytic Server manuell hinzufügen.

1. Öffnen Sie die Amabri-Konsole und fügen Sie die folgende Eigenschaft im Analytic Server-Abschnitt **Advanced analytics.cfg** hinzu.
   • **Key**: `spark.version`
   • **Value**: Geben Sie die entsprechende Spark-Versionsnummer ein (z. B. `1.x`, `2.x` oder `None`).
2. Speichern Sie die Konfiguration.

Anmerkung: Sie können über eine angepasste Einstellung `analytics.cfg` erzwingen, dass HCatalog Spark nie verwendet.

1. Öffnen Sie die Amabri-Konsole und fügen Sie die folgende Eigenschaft im Analytic Server-Abschnitt **Custom analytic.cfg** hinzu.

- **Key**: `spark.hive.compatible`
- **Value**: `false`

Kerberos-fähige Umgebungen unter HDP 3.0 (oder höher)

Für Kerberos-fähige Umgebungen unter HDP 3.0 (oder höher) sind unter Umständen zusätzliche Sicherheitskonfigurationseinstellungen erforderlich. In HDFS werden im Verzeichnis `/warehouse/tablespace/managed/hive` FACLs des Dateisystems verwendet. Sie können den Bedarf zum Festlegen von FACLs im Hive-Metaspeicher ermitteln, wenn in der Datei `messages.log` oder `as_trace.log` folgende Ausnahmen auftreten:

```
Caused by: org.apache.hadoop.hive.ql.metadata.HiveException: java.security.AccessControlException:
Permission denied: user=xxxx, access=READ, inode="/warehouse/tablespace/managed/hive/hcat_primitives":hive:hadoop:drwxrwx---
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.check(FSPermissionChecker.java:399)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:261)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:193)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1850)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1834)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPathAccess(FSDirectory.java:1784)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkAccess(FSNamesystem.java:7767)
```

Im folgenden Beispiel wird der Befehl **`setfacl`** gezeigt, der umfassenden Zugriff (im vorliegenden Beispiel für alle Mitglieder der `hadoop`-Gruppe) auf das Hive-Verzeichnis `warehouse` bietet:

```
hadoop fs -setfacl -R -m group:hadoop:rwx /warehouse/tablespace/managed/hive/
```

Andere, restriktivere Variationen sollten verwendet werden, wenn eine differenziertere Zugriffssteuerung erforderlich ist.

Folgende Sites enthalten weitere Referenzinformationen.

https://docs.hortonworks.com/HDPDocuments/HDP3/HDP-3.1.0/hdfs-acls/content/acl_examples.html

https://docs.hortonworks.com/HDPDocuments/HDP3/HDP-3.1.0/securing-hive/content/hive_sba_permissions_model.html

Ändern der von Analytic Server verwendeten Ports

Analytic Server verwendet standardmäßig Port 9080 für HTTP und Port 9443 für HTTPS. Führen Sie die folgenden Schritte aus, um die Porteinstellungen zu ändern.

1. Navigieren Sie auf der Registerkarte "Ambari Services" zur Registerkarte "Configs" des Analytic Server-Service.
2. Öffnen Sie den Abschnitt **Advanced analytics.cfg**.
3. Geben Sie den gewünschten HTTP- und HTTPS-Port in **`http.port`** bzw. **`https.port`** ein.
4. Klicken Sie auf **Save**.
5. Starten Sie den Analytic Server-Service erneut.

Analytic Server mit hoher Verfügbarkeit

Sie können Hochverfügbarkeit für Analytic Server bereitstellen, indem Sie das Produkt als Service für mehrere Knoten in Ihrem Cluster hinzufügen.

1. Navigieren Sie in der Ambari-Konsole zur Registerkarte **Hosts**.
2. Wählen Sie einen Host aus, der Analytic Server noch nicht als Service ausführt.
3. Klicken Sie auf der Registerkarte "Summary" auf **Add** und wählen Sie Analytic Server aus.
4. Klicken Sie auf **Confirm Add**.

Unterstützung mehrerer Cluster

Die Mehrclusterfunktion ist eine Erweiterung der Hochverfügbarkeitsfunktion von IBM SPSS Analytic Server und ermöglicht eine bessere Isolation in Umgebungen mit mehreren Nutzern. Standardmäßig wird bei der Installation des Analytic Server-Service (in Ambari oder Cloudera Manager) ein einzelner Analyse-Server-Cluster definiert.

Die Clusterspezifikation definiert die Analytic Server-Clusterzugehörigkeit. Die Clusterspezifikation wird über XML-Inhalt geändert (im Feld `analytics-cluster` von Ambari für die Analytic Server-Konfiguration oder durch manuelles Bearbeiten der Cloudera Manager-Datei `configuration/analytics-cluster.xml`). Wenn Sie mehrere Analytic Server-Cluster konfigurieren, müssen den einzelnen Analytic Server-Clustern Anforderungen über die Lastausgleichsfunktionen der Cluster zugeführt werden.

Durch die Verwendung der Mehrclusterfunktion wird sichergestellt, dass die Arbeit für einen Nutzer sich nicht negativ auf die Arbeit im Cluster eines anderen Nutzers auswirkt. Bei Hochverfügbarkeitsjobs kommt es nur innerhalb des Analytic Server-Clusters zu einem Job-Failober, auf dem die Arbeit initialisiert wurde. Das folgende Beispiel zeigt eine XML-Spezifikation für mehrere Cluster:

Anmerkung: Analytic Server kann als hochverfügbar definiert werden, indem Sie das Produkt mehreren Knoten in Ihrem Cluster als Service hinzufügen.

```
<analayticServerClusterSpec>
    <cardinality>1+</cardinality>
    <cluster name="cluster1">
        <memberName>one.cluster</memberName>
        <memberName>two.cluster</memberName>
    </cluster>
    <cluster name="cluster2">
        <memberName>three.cluster</memberName>
        <memberName>four.cluster</memberName>
    </cluster>
</analayticServerClusterSpec>
```

Im vorherigen Beispiel sind zwei Lastausgleichsfunktionen erforderlich. Eine Lastausgleichsfunktion sendet Anforderungen an die Member von `cluster1` (`one.cluster` und `two.cluster`) und die andere sendet Anforderungen an die Member von `cluster2` (`three.cluster` und `four.cluster`).

Das folgende Beispiel stellt eine XML-Spezifikation für einen einzelnen Cluster bereit (Standardkonfiguration).

```
<analayticServerClusterSpec>
    <cardinality>1</cardinality>
    <cluster name="cluster1">
        <memberName>*</memberName>
    </cluster>
</analayticServerClusterSpec>
```

Im vorherigen Beispiel ist eine einzige Lastausgleichsfunktion erforderlich, um die Fälle zu bearbeiten, in denen mehrere Cluster-Member konfiguriert sind.

Hinweise

- Nur Singleton-Cluster unterstützen die Verwendung von Platzhalterzeichen im Element **`memberName`** (z. B. Clusterkardinalität = "1"). Gültige Werte für das Kardinalitätselement sind 1 und 1+.
- Der Membername (**`memberName`**) muss auf dieselbe Weise wie der Name des Hosts angegeben werden, dem die Analytic Server-Rolle zugewiesen ist.
- Alle Server in allen Clustern müssen erneut gestartet werden, nachdem die Änderungen der Clusterkonfiguration angewendet wurden.
- In Cloudera Manager müssen Sie die Datei `analytics-cluster.xml` auf allen Analytic Server-Knoten ändern und warten. Alle Knoten müssen gewartet werden, um sicherzustellen, dass sie denselben Inhalt haben.

Optimieren von JVM-Optionen für Small Data

Sie können JVM-Eigenschaften bearbeiten, um Ihr System für die Ausführung von Small Jobs (M3R) zu optimieren.

Rufen Sie in der Ambari-Konsole im Analytic Server-Service den Abschnitt "Advanced analytics-jvm-options" der Registerkarte "Configs" auf. Durch Ändern der folgenden Parameter wird die Größe des Heapspeichers für Jobs festgelegt, die auf dem Server ausgeführt werden, der Analytic Server hostet, also nicht Hadoop. Dies ist bei der Ausführung von Small Jobs (M3R) wichtig. Möglicherweise müssen Sie mit diesen Werten experimentieren, um Ihr System zu optimieren.

```
-Xms512M
-Xmx2048M
```

Aktualisierung von Clientabhängigkeiten

In diesem Abschnitt wird beschrieben, wie die Abhängigkeiten des Analytic Server-Service mit dem Script `update_clientdeps` aktualisiert werden.

1. Melden Sie sich am Ambari-Server-Host als Root an.
2. Wechseln Sie zum Verzeichnis `/var/lib/ambari-server/resources/stacks/<Stackname>/<Stackversion>/services/ANALYTICSERVER/package/scripts`; siehe das folgende Beispiel.

   ```
   cd "/var/lib/ambari-server/resources/stacks/HDP/2.6/services/ANALYTICSERVER/package/scripts"
   ```
3. Führen Sie das Script `update_clientdeps` mit den folgenden Argumenten aus.

   -u <Ambari-Benutzer>
   Benutzername des Ambari-Kontos.

   -p <Ambari-Kennwort>
   Kennwort für den Benutzer des Ambari-Kontos.

   -h <Ambari-Host>
   Hostname des Ambari-Servers.

   -x <Ambari-Port>
   Port, an dem Ambari empfangsbereit ist.

   Siehe das folgende Beispiel.

   ```
   ./update_clientdeps.sh -u admin -p admin -h host.domain -x 8080
   ```
4. Starten Sie den Ambari-Server mit dem folgenden Befehl erneut.

   ```
   ambari-server restart
   ```

Konfigurieren von Apache Knox

Der Apache Knox-Gateway ist ein System, das einen zentralen sicheren Zugriff auf Apache Hadoop-Services bereitstellt. Das System vereinfacht die Hadoop-Sicherheit sowohl für Benutzer (die auf die Clusterdaten zugreifen und Jobs ausführen) als auch für Bediener (die den Zugriff steuern und den Cluster verwalten). Der Gateway wird als Server (oder als Server-Cluster) ausgeführt, der mindestens einen Hadoop-Cluster bereitstellt.

Anmerkung: IBM SPSS Analytic Server unterstützt nicht Apache Knox bei Verwendung zusammen mit Kerberos Single-Sign-On.

Der Apache Knox-Gateway verbirgt effektiv die Details der Hadoop-Clustertopologie und ist in Enterprise LDAP und Kerberos integriert. Die folgenden Abschnitte enthalten Informationen zu den erforderlichen Konfigurationstasks für Apache Knox und Analytic Server.

Voraussetzungen

- Ein bekanntes Problem mit Apache Knox ist, dass die in HTTP-Cookies und -Headern enthaltenen Sicherheitsinformationen nicht weiter geleitet werden (weitere Informationen finden Sie in https://issue-

s.apache.org/jira/browse/KNOX-895). Das Problem wurde in Knox 0.14.0 (oder höher) behoben. Sie müssen eine aktualisierte Hortonworks-Verteilung abrufen, die Knox 0.14.0 (oder höher) einschließt, bevor Knox mit Analytic Server funktioniert. Wenden Sie sich an Ihren Hortonworks-Provider, wenn Sie weitere Informationen benötigen.

- Die Analytic Server-Knoten müssen über eine kennwortunabhängige SSH-Verbindung mit dem Knox-Server verbunden werden. Die kennwortunabhängige SSH-Verbindung verläuft von Analytic Server zu Knox (**Analytic Server** > **Knox**).
- Analytic Server muss nach der Installation des Knox-Service installiert werden.

In einigen Fällen führen nicht erwartete Probleme dazu, dass die Konfigurationsdateien nicht automatisch kopiert werden. In diesen Fällen müssen Sie die folgenden Konfigurationsdateien manuell kopieren:

- `com.ibm.spss.knox_0.6-3.2.1.0.jar`: Die Datei muss aus dem Analytic Server-Speicherort

  `<Analytic_Server-Installationspfad>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib`

  auf den Knox-Serverknoten kopiert werden:

  `/KnoxServicePath/ext`

  Beispiel: `/usr/iop/4.1.0.0/knox/ext`
- `rewrite.xml` und `service.xml`: Die Dateien müssen aus demAnalytic Server-Speicherort

  `<Analytic_Server-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/knox`

  auf den Knox-Serverknoten kopiert werden:

  `/KnoxServicePath/data/services`

  Beispiel: `/usr/iop/4.1.0.0/knox/data/services`

  Anmerkung: Es gibt zwei Gruppen mit den Dateien `rewrite.xml` und `service.xml` (eine Gruppe mit dem Datenverkehr für `http://rest` und eine Gruppe mit dem Datenverkehr für `ws://websocket`). Kopieren Sie alle Dateien `rewrite.xml` und `service.xml` für `analyticserver` und `analyticserver_ws` auf den Knox-Serverknoten.

Konfigurieren von Ambari

Der Analytic Server-Service muss über die Ambari-Benutzerschnittstelle konfiguriert werden:

1. Navigieren Sie in der Ambari-Benutzerschnittstelle zu **Knox** > **Configs** > **Advanced topology**. Die aktuellen Knox-Konfigurationseinstellungen werden im Inhaltsfenster angezeigt.
2. Fügen Sie dem Abschnitt **Advanced topology** in der Knox-Konfiguration die beiden folgenden Services hinzu:

```
<service>
    <role>ANALYTICSERVER</role>
    <url>http://{AS-Host}:{AS-Port}/analyticserver</url>
 </service>
<service>
    <role>ANALYTICSERVER_WS</role>
    <url>ws://{AS-Host}:{AS-Port}/analyticserver</url>
 </service>
```

   `{AS-Host}` und `{AS-Port}` müssen durch den entsprechenden Namen und die entsprechende Portnummer des Analytic Server-Servers ersetzt werden.

   - Die URL von `{AS-Host}` kann über die Ambari-Benutzerschnittstelle angezeigt werden (**SPSS Analytic Server** > **Summary** > **Analytic Server**).
   - Die Nummer von `{AS-Port}` kann über die Ambari-Benutzerschnittstelle angezeigt werden (**SPSS Analytic Server** > **Configs** > **Advanced analytics.cfg** > **http.port**).

   Anmerkung: Wenn Analytic Server auf mehreren Knoten bereitgestellt wird und die Lastausgleichsfunktion (LoadBalancer) verwendet wird, müssen `{AS-Host}` und `{AS-Port}` der URL und der Portnummer der Lastausgleichsfunktion entsprechen.

3. Starten Sie den Knox-Service erneut.

Wenn LDAP verwendet wird, übernimmt Knox standardmäßig die Werte des bereitgestellten LDAP-Demoservers. Sie können einen Enterprise LDAP-Server verwenden (beispielsweise Microsoft LDAP oder OpenLDAP).

Konfigurieren von Analytic Server

Wenn LDAP für Analytic Server verwendet werden soll, muss Analytic Server für die Verwendung des LDAP-Servers konfiguriert sein, der von Apache Knox verwendet wird. Die `<value>`-Werte für die folgenden Ambari-Einstellungen müssen aktualisiert werden, damit sie die entsprechenden Einstellungen des Knox-LDAP-Servers widerspiegeln.

- `main.ldapRealm.userDnTemplate`
- `main.ldapRealm.contextFactory.url`

Die Werte sind in der Ambari-Benutzerschnittstelle unter **Knox** > **Configs** > **Advanced topology** verfügbar. Beispiel:

```
<param>
   <name>main.ldapRealm.userDnTemplate</name>
   <value>uid={0},ou=people,dc=hadoop,dc=apache,dc=org</value>
</param>
<param>
   <name>main.ldapRealm.contextFactory.url</name>
   <value>ldap://{{Knox-Hostname}}:33389</value>
</param>
```

Starten Sie den Knox-Service nach der Aktualisierung der Knox-LDAP-Einstellungen erneut.

Wichtig: Das Analytic Server-Administratorkennwort muss mit dem Knox-Administratorkennwort identisch sein.

Konfigurieren von Apache Knox

1. Aktualisieren Sie die Knox-Datei `gateway.jks`:
   a. Stoppen Sie den Knox-Service auf dem Knox-Server.
   b. Löschen Sie die Datei `gateway.jks` aus dem Verzeichnis `/var/lib/knox/data-2.6.2.0-205/security/keystores`.
   c. Starten Sie den Knox-Service erneut.
2. Erstellen Sie auf dem Knox-Server das Unterverzeichnis `<Knox-Server>/data/service/analyticserver/3.2.1.0` und laden Sie anschließend die Dateien `service.xml` und `rewrite.xml` in das neue Verzeichnis hoch. Die beiden Dateien befinden sich in Analytic Server in `<Analytic_Server>/configuration/knox/analyticserver/` (z. B. `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/*.xml`)
3. Führen Sie das Script `./knoxcli.sh redeploy --cluster default` in `<Knox-Server>/bin` aus.
4. Laden Sie die Datei `com.ibm.spss.knoxservice_0.6-*.jar` in `<Know-Server>/ext` hoch. Die Datei befindet sich in Analytic Server in `<Analytic_Server>/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.1.0.jar` (z. B. `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.1.0.jar`).
5. Fügen Sie in der Ambari-Benutzerschnittstelle das folgende Element in **Knox** > **Configs** > **Advanced topology** hinzu:

```
<service>
   <role>ANALYTICSERVER</role>
   <url>http://{AS-Host}:{AS-Port}/analyticserver</url>
   <role>ANALYTICSERVER_WS</role>
   <url>ws://{AS-Host}:{AS-Port}/analyticserver</url>
</service>
```

Anmerkung: Die WebSocket-Funktion ist standardmäßig inaktiviert. Sie kann aktiviert werden, indem die Eigenschaft `gateway.websocket.feature.enabled` in der Datei `/conf/gateway-site.xml` in `true` geändert wird.

6. In der Ambari-Benutzerschnittstelle müssen Sie die Benutzer in **Knox** > **Configs** > **Advanced users-ldif** (z. B. `admin`, `qauser1`, `qauser2`) hinzufügen oder aktualisieren.
7. Starten Sie LDAP über **Knox** > **Service Actions** > **Start Demo LDAP** erneut.
8. Starten Sie den Knox-Service erneut.

URL-Struktur für die für Apache Knox aktivierte Analytic Server-Instanz

Die für Knox aktivierte URL der Analytic Server-Benutzerschnittstelle lautet `https://{Knox-Host}:{Knox-Port}/gateway/default/analyticserver/admin`.

- HTTPS-Protokoll - Benutzer müssen ein Zertifikat akzeptieren, um im Web-Browser fortfahren zu können.
- `Knox-Host` ist der Knox-Host.
- `Knox-Port` ist die Nummer des Knox-Ports.
- Der URI lautet `gateway/default/analyticserver`.

Konfigurieren separater YARN-Warteschlangen für jeden IBM SPSS Analytic Server-Nutzer - HDP

Die Konfiguration von YARN-Warteschlangen erfolgt durch die Verwendung von Spark-Techniken zur dynamischen Ressourcenzuordnung.

Hortonworks Data Platform 2.x

1. Navigieren Sie in der Ambari-Benutzerschnittstelle zur Registerkarte **SPSS Analytic Server service** > **Configs** > **Advanced analytics.cfg**.
2. Ändern Sie den Wert **resource.pool.enabled** in `true`.
3. Fügen Sie auf der Registerkarte **Custom analytics.cfg** die folgenden Eigenschaften hinzu:

```
config.folder.path=/etc/spark2/conf
resource.pool.mapping=tenant1:test,tenant2:production
resource.pool.default=default
spark.scheduler.mode=FAIR
spark.yarn.queue=default
```

Tabelle 8. Angepasste Eigenschaften in 'analytics.cfg'

Eigenschaft	Beschreibung
`config.folder.path`	Das Verzeichnis enthält die Datei `fairscheduler.xml`, die die Eigenschaftsinformationen für den Spark-Pool enthält. Die Datei ist erforderlich und muss manuell erstellt werden. Weitere Informationen finden Sie im Abschnitt **fairscheduler.xml - Beispiel**.
`resource.pool.mapping`	**Spark:** Ordnet die Nutzer den Pools zu, die in der Datei `fairscheduler.xml` definiert sind. Nutzerpaare müssen durch Kommas getrennt werden (Beispiel: `tenant1:test,tenant2:production`). Bevor Sie einen Pool angeben, stellen Sie sicher, dass der Pool in der Datei `fairscheduler.xml` angegeben wurde. **MapReduce:** Ordnet Nutzer der Warteschlange zu, die in YARN Queue Manager definiert wurde. Nutzerpaare müssen durch Kommas getrennt werden (Beispiel: tenant1:test,tenant2:production). Bevor Sie eine Warteschlange angeben, stellen Sie sicher, dass das System mit der Warteschlange konfiguriert wurde und dass der Zugriff zum Übergeben von Jobs an die Warteschlange zulässig ist. **Anmerkung:** Wenn Sie die Spark- und MapReduce-Jobs zusammen ausführen wollen, muss der Name der Nutzerzuordnungswerte in der Datei `fairscheduler.xml` und in YARN Queue Manager identisch sein.

Tabelle 8. Angepasste Eigenschaften in 'analytics.cfg' (Forts.)

Eigenschaft	Beschreibung
`resource.pool.default`	**Spark:** Definiert den Standardressourcenpool. Der Wert kann `default` oder ein Poolname sein, der in der Datei `fairscheduler.xml` definiert wurde. Verwenden Sie die Einstellung `default`, wenn Nutzer nicht (oder falsch) konfiguriert sind. **MapReduce:** Definiert die Standardwarteschlange, an die Jobs übergeben werden.
`spark.scheduler.mode=FAIR`	**Spark:** Aktiviert den Scheduler für den akzeptablen Modus. Die Eigenschaft sollte nicht geändert werden.
`spark.yarn.queue`	**Spark:** Der Name der YARN-Warteschlange, an die die Anwendung übergeben wird. In YARN Queue Manager können Sie einen angepassten YARN-Warteschlangennamen angeben.

4. Speichern Sie die Konfiguration und starten Sie den Analytic Server-Service erneut.

fairscheduler.xml - Beispiel

Die Datei `fairscheduler.xml` enthält die Eigenschaftsinformationen für den Spark-Pool. Die Datei ist erforderlich und muss manuell erstellt werden.

```
<?xml version="1.0"?>
<allocations>
  <pool name="production">
    <schedulingMode>FAIR</schedulingMode>
    <weight>1</weight>
    <minShare>2</minShare>
  </pool>
  <pool name="test">
    <schedulingMode>FIFO</schedulingMode>
    <weight>2</weight>
    <minShare>3</minShare>
  </pool>
</allocations>
```

Referenz

Weitere Informationen finden Sie auf den folgenden Sites:

- https://spark.apache.org/docs/latest/job-scheduling.html#dynamic-resource-allocation
- https://spark.apache.org/docs/latest/running-on-yarn.html

Migrieren von IBM SPSS Analytic Server unter Ambari

Analytic Server kann Daten und Konfigurationseinstellungen aus einer vorhandenen Analytic Server-Installation in eine neue Installation migrieren. Die Migration kann in derselben Clusterumgebung oder in einer neuen Clusterumgebung durchgeführt werden.

Migrieren von Analytic Server 3.1.2 zu 3.2.1 auf demselben Server-Cluster

Wenn Sie über eine vorhandene Installation von Analytic Server 3.1.2 verfügen, können Sie Ihre Konfigurationseinstellungen aus Version 3.1.2 zu Ihrer Installation der Version 3.2.1 auf demselben Server-Cluster migrieren.

1. Erfassen Sie die Konfigurationseinstellungen der alten Version von Analytic Server (Analytic Server 3.1.2).
   a. Erweitern Sie das Archiv `{AS_ROOT}\tools\unzip configcollector.zip` (dadurch wird ein neuer Ordner mit dem Namen `configcollector` erstellt).
   b. Führen Sie das Script `configcollector.sh` im Ordner `configcollector` aus. Kopieren Sie die resultierende komprimierte Datei `ASConfiguration_3.1.2.0.xxx.zip` (ZIP-Datei) in eine andere Ordnerposition (als Backup).

2. Sichern Sie das Analytic Server-Stammverzeichnis aus der Installation Ihrer alten Version von Analytic Server 3.1.2 an einer neuen Position.
   a. Wenn Sie sich nicht sicher sind, wo sich das Analytic Server-Stammverzeichnis befindet, führen Sie den Befehl **hhadoop fs -ls** aus. Der Pfad zum Analytic Server-Stammverzeichnis lautet ungefähr /user/as_user/analytic-root/analytic-workspace, wobei as_user die Benutzer-ID ist, die Eigner des Analytic Server-Stammverzeichnisses ist.
   b. Verwenden Sie die Befehle **hadoop fs -copyToLocal** und **hadoop fs -copyFromLocal**, um den Ordner analytic-workspace der alten Analytic Server-Version an die neue Position zu kopieren (Beispiel: /user/as_user/analytic-root/AS31Location).
3. Wenn Sie die eingebettete Apache Directory Server-Instanz verwenden, sichern Sie die aktuelle Benutzer-/Gruppenkonfiguration mit einem LDAP-Client-Tool eines anderen Anbieters. Importieren Sie die gesicherte Benutzer-/Gruppenkonfiguration nach der Installation von Analytic Server 3.2.1 in Apache Directory Server.

   Anmerkung: Dieser Schritt kann übersprungen werden, wenn Sie einen externen LDAP-Server verwenden.
4. Öffnen Sie die Ambari-Konsole und stoppen Sie **Analytic Server service**.
5. Deinstallieren Sie die alte Version von Analytic Server (Analytic Server 3.1.2) und installieren Sie anschließend Analytic Server 3.2.1. Installationsanweisungen finden Sie in Kapitel 2, „Ambari-Installation und -Konfiguration", auf Seite 3.
6. Öffnen Sie die Ambari-Konsole und stoppen Sie **Analytic Server service** (stellen Sie in Ambari sicher, dass **Analytic Metastore service** ausgeführt wird).
7. Kopieren Sie das gesicherte Analytic Server-Stammverzeichnis von Analytic Server 3.1.2 aus Schritt 2 an die neue Position der neuen Version von Analytic Server.
   a. Entfernen Sie analytic-workspace aus der neu installierten Version von Analytic Server.
   b. Kopieren Sie den gesicherten Analytic Server-Arbeitsbereichsordner von Analytic Server 3.1.2 (/user/as_user/analytic-root/AS31Location) an die Position der neuen Version (Beispiel: /user/as_user/analytic-root/analytic-workspace). Sie müssen sicherstellen, dass der Eigner des Analytic Server-Arbeitsbereichs als as_user definiert ist.
8. Löschen Sie den Zookeeper-Status. Führen Sie den folgenden Befehl im ZooKeeper-bin-Verzeichnis aus (z. B. /usr/hdp/current/zookeeper-client unter Hortonworks):
   ```
   ./zkCli.sh rmr /AnalyticServer
   ```
9. Kopieren Sie das Sicherungsarchiv ASConfiguration_3.1.2.0.xxx.zip aus Schritt 1 an die Position der neuen Version von Analytic Server (Beispiel: /opt/ibm/spss/analyticserver/3.2/).
10. Führen Sie das Migrationstool aus, indem Sie das Script **migrationtool.sh** ausführen und den Pfad der Archivdatei ASConfiguration_3.1.2.0.xxx.zip (die vom Konfigurationscollector erstellt wurde) übergeben. Beispiel:
    ```
    migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.1.2.0.xxx.zip
    ```
11. Führen Sie den folgenden Befehl in einer Befehlsshell auf dem Analytic Server-Knoten aus:
    ```
    hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
    ```
12. Starten Sie **Analytic Server service** in der Ambari-Konsole.

Migrieren von Analytic Server 3.1.2 zu 3.2.1 zu einem neuen Server-Cluster

Wenn Sie über eine vorhandene Installation von Analytic Server 3.1.2 verfügen, können Sie Ihre Konfigurationseinstellungen aus Version 3.1.2 zu Ihrer Installation der Version 3.2.1 auf einem neuen Server-Cluster migrieren.

1. Installieren Sie die neue Analytic Server-Version entsprechend den Anweisungen in „Installation in Ambari" auf Seite 6.
2. Kopieren Sie den Analytic Server-Arbeitsbereich von Ihrer alten Installation in die neue Installation.

a. Wenn Sie sich nicht sicher sind, wo sich der Analytic Server-Arbeitsbereich befindet, führen Sie den Befehl `hadoop fs -ls` aus. Der Pfad zum Analytic Server-Arbeitsbereich lautet ungefähr `/user/as_user/analytic-root/analytic-workspace`, wobei `as_user` die Benutzer-ID ist, die Eigner des Analytic Server-Arbeitsbereichs ist.
b. Entfernen Sie `analytic-workspace` auf dem neuen Server.
c. Verwenden Sie `hadoop fs -copyToLocal` und `hadoop fs -copyFromLocal`, um den Analytic Server-Arbeitsbereich des alten Servers in den Ordner `/user/as_user/analytic-root/analytic-workspace/` des neuen Servers zu kopieren (stellen Sie sicher, dass der Benutzer als `as_user` angegeben wird).

3. Wenn Sie die eingebettete Apache Directory Server-Instanz verwenden, sichern Sie die aktuelle Benutzer-/Gruppenkonfiguration mit einem LDAP-Client-Tool eines anderen Anbieters. Importieren Sie die gesicherte Benutzer-/Gruppenkonfiguration nach der Installation von Analytic Server 3.2.1 in Apache Directory Server.

   Anmerkung: Dieser Schritt kann übersprungen werden, wenn Sie einen externen LDAP-Server verwenden.
4. Öffnen Sie auf dem neuen Server die Ambari-Konsole und stoppen Sie den Analytic Server-Service (stellen Sie unter Ambari sicher, dass der Analytic Metastore-Service ausgeführt wird).
5. Erfassen Sie die Konfigurationseinstellungen der alten Installation.
   a. Kopieren Sie das Archiv `configcollector.zip` in Ihrer neuen Installation in `{AS-Stammverzeichnis}\tools` in Ihrer alten Installation.
   b. Extrahieren Sie die Kopie von `configcollector.zip`, wodurch in Ihrer alten Installation ein neues Unterverzeichnis `configcollector` erstellt wird.
   c. Führen Sie das Konfigurations-Collector-Tool in Ihrer alten Installation aus, indem Sie das Script **`configcollector`** im Verzeichnis `{AS-Stammverzeichnis}\tools\configcollector` aufrufen. Kopieren Sie die resultierende komprimierte Datei (ZIP-Datei) auf den Server, der Ihre neue Installation hostet.

   Wichtig: Das bereitgestellte Script **`configcollector`** ist möglicherweise nicht mit der aktuellen Version von Analytic Server kompatibel. Wenden Sie sich an einen IBM Technical Support-Mitarbeiter, wenn Probleme mit dem Script **`configcollector`** auftreten.
6. Löschen Sie den Zookeeper-Status. Führen Sie den folgenden Befehl im ZooKeeper-bin-Verzeichnis aus (z. B. `/usr/hdp/current/zookeeper-client` unter Hortonworks).

   ```
   ./zkCli.sh rmr /AnalyticServer
   ```
7. Führen Sie das Script **`migrationtool`** für das Migrationstool aus und übergeben Sie den Pfad der vom Konfigurationscollector erstellten komprimierten Datei als Argument. Es folgt ein Beispiel.

   ```
   migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.1.2.0.xxx.zip
   ```
8. Führen Sie den folgenden Befehl in einer Befehlsshell auf dem Analytic Server-Knoten aus:

   ```
   hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
   ```
9. Starten Sie den Analytic Server-Service in der Ambari-Konsole.

   Anmerkung: Wenn Sie R für die Verwendung mit der vorhandenen Analytic Server-Installation konfiguriert haben, befolgen Sie die Schritte zum Konfigurieren von R mit der neuen Analytic Server-Installation.

Deinstallation

Wichtig: Wenn Essentials for R installiert ist, müssen Sie zunächst das Script `remove_R.sh` ausführen. Wenn die Deinstallation von Essentials for R vor der Deinstallation von Analytic Server fehlschlägt, kann Essentials for R zu einem späteren Zeitpunkt nicht mehr deinstalliert werden. Bei der Deinstallation von Analytic Server wird das Script `remove_R.sh` entfernt. Informationen zur Deinstallation von Essentials for R finden Sie in „Deinstallation von Essentials for R" auf Seite 36.

1. Führen Sie auf dem Analytic Metastore-Host das Script `remove_as.sh` im Verzeichnis `{AS-Stammverzeichnis}/bin` mit den folgenden Parametern aus.

   u Erforderlich. Benutzer-ID des Ambari-Server-Administrators.

   p Erforderlich. Kennwort des Ambari-Server-Administrators.

   h Erforderlich. Name des Ambari-Server-Hosts.

   x Erforderlich. Ambari-Server-Port.

   l Optional. Aktiviert den sicheren Modus.

   Es folgen Beispiele.

   ```
   remove_as.sh -u admin -p admin -h one.cluster -x 8081
   ```

   Entfernt Analytic Server aus einem Cluster mit dem Ambari-Host one.cluster.

   ```
   remove_as.sh -u admin -p admin -h one.cluster -x 8081 -l
   ```

   Entfernt Analytic Server aus einem Cluster mit dem Ambari-Host one.cluster (sicherer Modus).

Anmerkung: Diese Operation entfernt den Analytic Server-Ordner aus HDFS.

Anmerkung: Durch diese Operation wird keines der Analytic Server zugeordneten Db2-Schemas entfernt. Informationen zum manuellen Entfernen von Schemas finden Sie in der Db2-Dokumentation.

Deinstallation von Essentials for R

1. Führen Sie auf dem Essentials for R-Host das Script `remove_R.sh` im Verzeichnis `{AS-Stammverzeichnis}/bin` mit den folgenden Parametern aus.

   u Erforderlich. Benutzer-ID des Ambari-Server-Administrators.

   p Erforderlich. Kennwort des Ambari-Server-Administrators.

   h Erforderlich. Name des Ambari-Server-Hosts.

   x Erforderlich. Ambari-Server-Port.

   l Optional. Aktiviert den sicheren Modus.

   Es folgen Beispiele.

   ```
   remove_R.sh -u admin -p admin -h one.cluster -x 8081
   ```

   Entfernt Essentials for R aus einem Cluster mit dem Ambari-Host one.cluster.

   ```
   remove_R.sh -u admin -p admin -h one.cluster -x 8081 -l
   ```

   Entfernt Essentials for R aus einem Cluster mit dem Ambari-Host one.cluster (sicherer Modus).
2. Entfernt das R-Serviceverzeichnis aus dem Ambari-Server-Service-Verzeichnis. Beispiel: Unter HDP 2.6 befindet sich das Verzeichnis `ESSENTIALR` im Verzeichnis `/var/lib/ambari-server/resources/stacks/HDP/2.6/services`.
3. Prüfen Sie in der Ambari-Konsole, dass der Essentials for R-Service nicht mehr vorhanden ist.

Kapitel 3. Cloudera-Installation und -Konfiguration

Cloudera - Übersicht

Cloudera ist eine Open-Source-Verteilung von Apache Hadoop. Cloudera Distribution Including Apache Hadoop (CDH) ist für auf Unternehmen abgestimmte Bereitstellungen dieser Technologie konzipiert.

Analytic Server kann auf der Plattform CDH ausgeführt werden. CDH enthält die zentralen Hauptelemente von Hadoop, die die zuverlässige, skalierbare, verteilte Datenverarbeitung großer Datasets (hauptsächlich MapReduce und HDFS) ermöglichen. Außerdem enthält es weitere unternehmensorientierte Komponenten, die Sicherheit, Hochverfügbarkeit und Integration in Hardware und andere Software bereitstellen.

Cloudera-spezifische Voraussetzungen

Lesen Sie zusätzlich zu den Angaben zu allgemeinen Voraussetzungen die folgenden Informationen.

Services

Stellen Sie sicher, dass die folgenden Instanzen auf jedem Analytic Server-Host installiert sind.

- HDFS: Gateway, Datenknoten oder Namensknoten
- Hive: Gateway, Hive-Metaspeicherserver oder HiveServer2
- YARN: Gateway, Ressourcenmanager oder Knotenmanager

Die folgenden Instanzen sind nur erforderlich, wenn die zugehörigen Funktionen verwendet werden.

- Accumulo: Gateway
- HBase: Gateway, Master oder Regionsserver
- Spark: Gateway
- Spark 2: Gateway

Metadatenrepository

Sie können Db2 und MySQL als Analytic Server-Metadatenrepository verwenden. Wenn Sie MySQL als Analytic Server-Metadatenrepository verwenden wollen, befolgen Sie die Anweisungen für „Konfigurieren von MySQL für Analytic Server" auf Seite 39.

Kerberos-fähige Cloudera-Umgebungen

Wenn Sie planen, Analytic Server in einer Kerberos-fähigen Cloudera-Umgebung zu installieren, müssen Sie sicherstellen, dass Kerberos ordnungsgemäß in einer Weise konfiguriert wurde, die mit Analytic Server kompatibel ist.

Die folgenden Abschnitte gelten für Cloudera-Umgebungen, in denen Kerberos bereits installiert ist. Sie müssen die folgenden Abschnitte befolgen, bevor Sie Analytic Server in Cloudera installieren. Es wird angenommen, dass Sie über grundlegende Kenntnisse der Kerberos-Authentifizierung verfügen, da die Abschnitte Kerberos-spezifische Terminologie (z. B. **`kinit`**, **`kadmin`** usw.) enthalten.

Anmerkung: Analytic Server überprüft die HDFS-Konfiguration für auf Kerberos-bezogene Werte, die für die Authentifizierung verwendet werden können.

Kerberos-Authentifizierung

Prüfen Sie vor der Installation von Analytic Server, ob die Kerberos-Authentifizierung auf allen Cloudera-Clustern konfiguriert ist. Weitere Informationen finden Sie in Authentifizierung in Cloudera Manager konfigurieren in der Cloudera-Produktdokumentation.

Anmerkung: Nachdem die Kerberos-Authentifizierung auf einem Cloudera-Clusterknoten konfiguriert wurde, müssen die Services **cloudera-scm-server** und **cloudera-scm-agent** erneut gestartet werden, bevor Analytic Server installiert wird. Der Service **cloudera-scm-agent** muss auf allen Clusterknoten erneut gestartet werden.

Erstellen des erforderlichen Kontos in Kerberos

1. Sie können im Kerberos-Benutzerrepository für alle Benutzer, denen Sie Zugriff auf Analytic Server erteilen möchten, Konten erstellen.
2. Erstellen Sie dieselben Konten (aus dem vorherigen Schritt) auf dem LDAP-Server.
3. Erstellen Sie für jeden im vorherigen Schritt erstellten Benutzer auf jedem einzelnen Analytic Server-Knoten und Hadoop-Knoten ein Betriebssystembenutzerkonto.
   - Stellen Sie sicher, dass die Benutzer-ID für diese Benutzer auf allen Computern übereinstimmt. Sie können dies testen, indem Sie sich mit dem Befehl `kinit` bei den einzelnen Konten anmelden.
   - Stellen Sie sicher, dass die Benutzer-ID der YARN-Einstellung **Minimum user ID for submitting job** entspricht. Dies ist die Einstellung **min.user.id** in `container-executor.cfg`. Wenn **min.user.id** beispielsweise auf 1000 gesetzt ist, muss die Benutzer-ID jedes erstellten Benutzerkontos größergleich 1000 sein.
4. Erstellen Sie in HDFS einen Benutzerausgangsordner für den Analytic Server-Administrator. Die Ordnerberechtigung muss auf `777` gesetzt werden, der Eigner muss als `admin` definiert werden und die Benutzergruppe muss auf `hdfs` gesetzt werden. Siehe das folgende Beispiel in Fettdruck:

```
[root@xxxxx configuration]# hadoop fs -ls /user

Found 9 items

drwxrwxrwx   - hdfs     supergroup      0 2017-07-26 03:41 /user/AE
drwxrwxrwx   - admin    hdfs            0 2017-06-08 01:33 /user/admin
drwxr-x--x   - as_user  hdfs            0 2017-06-06 01:00 /user/as_user
drwx------   - hdfs     supergroup      0 2017-07-31 00:17 /user/hdfs
drwxrwxrwx   - mapred   hadoop          0 2017-06-05 00:28 /user/history
drwxrwxr-t   - hive     hive            0 2017-06-05 00:30 /user/hive
drwxrwxr-x   - hue      hue             0 2017-06-05 00:30 /user/hue
drwxrwxr-x   - impala   impala          0 2017-07-19 00:52 /user/impala
drwxr-x--x   - spark    spark           0 2017-06-05 01:34 /user/spark
```

5. Wenn Sie HCatalog-Datenquellen verwenden wollen und Analytic Server auf einem anderen Computer als Hive-Metaspeicher installiert ist, müssen Sie in HDFS die Identität des Hive-Clients annehmen.
   a. Navigieren Sie in Cloudera Manager zur Registerkarte **Configuration** des HDFS-Service.

      Anmerkung: Die folgenden Einstellungen werden möglicherweise nicht auf der Registerkarte **Configuration** angezeigt, wenn sie nicht bereits festgelegt wurden. Führen Sie in diesem Fall eine Suche nach ihnen aus.

   b. Bearbeiten Sie die Einstellung **hadoop.proxyuser.hive.groups** so, dass sie den Wert `*` hat oder eine Gruppe enthält, die alle Benutzer umfasst, die sich an Analytic Server anmelden können.
   c. Bearbeiten Sie die Einstellung **hadoop.proxyuser.hive.hosts** so, dass sie den Wert `*` hat oder die Liste der Hosts enthält, auf denen der Hive-Metaspeicher und alle Instanzen von Analytic Server als Service installiert sind.
   d. Starten Sie den HDFS-Service erneut.

Nachdem Sie diese Schritte ausgeführt haben und Analytic Server installiert ist, konfiguriert Analytic Server Kerberos automatisch im Hintergrund.

Aktivieren des Kerberos-Identitätswechsels

Durch Identitätswechsel kann ein Thread in einem Sicherheitskontext ausgeführt werden, der sich vom Sicherheitskontext des Prozesses unterscheidet, der der Threadeigner ist. Beispielsweise können Hadoop-Jobs mithilfe von Identitätswechsel über einen anderen Benutzer als den Analytic Server-Standardbenutzer (`as_user`) ausgeführt werden. So aktivieren Sie den Kerberos-Identitätswechsel:

1. Öffnen Sie Cloudera Manager und fügen Sie im Bereich **Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml** (auf der Registerkarte **HDFS (Service-Wide)** > **Configuration**) die folgenden Eigenschaften hinzu oder aktualisieren diese.
   - **Name:** `hadoop.proxyuser.as_user.hosts`
   - **Value:** `*`
   - **Name:** `hadoop.proxyuser.as_user.groups`
   - **Value:** `*`

   Anmerkung: Die Einstellung **core-site.xml** gilt für die Hadoop-Konfiguration (nicht Analytic Server).
2. Führen Sie den folgenden Befehl in einer Befehlsshell auf dem Analytic Server-Knoten aus:
   ```
   hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
   ```

Konfigurieren von MySQL für Analytic Server

Zum Konfigurieren von IBM SPSS Analytic Server in Cloudera Manager ist die Installation und Konfiguration einer MySQL-Serverdatenbank erforderlich.

1. Führen Sie den folgenden Befehl in einem Befehlsfenster auf dem Knoten aus, auf dem die MySQL-Datenbank gespeichert ist:
   ```
   yum install mysql-server
   ```

   Anmerkung: Verwenden Sie für SuSE Linux `zypper install mysql`.
2. Führen Sie den folgenden Befehl in einem Befehlsfenster auf jedem Cloudera-Clusterknoten aus:
   ```
   yum install mysql-connector-java
   ```

   Anmerkung: Verwenden Sie für SuSE Linux `sudo zypper install mysql-connector-java`.
3. Legen Sie den Datenbanknamen, den Datenbankbenutzernamen und das Datenbankkennwort für Analytic Server fest, die Analytic Server beim Zugriff auf die MySQL-Datenbank verwendet, und notieren Sie sich diese Angaben.
4. Installieren Sie Analytic Server entsprechend den Anweisungen in „Installation in Cloudera" auf Seite 42.
5. Kopieren Sie das Script `/opt/cloudera/parcels/AnalyticServer/bin/add_mysql_user.sh` von einem der con Cloudera verwalteten Server auf den Knoten, auf dem die MySQL-Datenbank installiert ist. Führen Sie das Script mit den Ihrer Konfiguration entsprechenden Parametern aus. Beispiel:
   ```
   ./add_mysql_user.sh -u <Datenbankbenutzername> -p <Datenbankkennwort> -d
   <Datenbankname>
   ```

   Hinweise: Der Parameter `a -r <DB-Rootkennwort>` ist erforderlich, wenn die Datenbank im sicheren Modus (das Rootbenutzerkennwort ist festgelegt) ausgeführt wird.

   Die Parameter `-r <DB-Benutzerkennwort>` und `-t <DB-Benutzername>` sind erforderlich, wenn die Datenbank im sicheren Modus mit einem anderen Benutzernamen als `root` ausgeführt wird.

Precheck- und Postcheck-Tools für Installation - Cloudera

Speicherort und Voraussetzungen für das Tool

Führen Sie vor der Installation des Analytic Server-Service das Precheck-Tool auf allen Knoten aus, die Teil des Analytic Server-Service sein werden, um zu prüfen, ob Ihre Linux-Umgebung für die Installation von Analytic Server bereit ist.

Das Precheck-Tool wird automatisch als Teil der Installation aufgerufen. Das Tool prüft jeden Analytic Server-Knoten, bevor die Installation auf jedem Host ausgeführt wird. Sie können das Precheck-Tool auch manuell auf jedem Knoten aufrufen. Dadurch wird der Computer vor der Installation des Service validiert.

Nach dem Ausführen der selbstextrahierenden Analytic Server-Binärdatei befindet sich das Precheck-Tool in den folgenden Verzeichnissen:

- **Cloudera**

```
/opt/cloudera/parcels/AnalyticServer-3.2.1.0/tools/com.spss.ibm.checker.zip

[root@servername ~]# cd /opt/cloudera/parcels/AnalyticServer-3.2.1.0/tools/
[root@servername tools]# ls
com.spss.ibm.checker.zip configcollector.zip regex-files
```

Anmerkung: Das Precheck-Tool ist im Verzeichnis `tools` erst verfügbar, wenn Sie die ausführbare Binärdatei ausgeführt und anschließend verteilt (**Download** > **Distribute**) und Analytic Server auf der Seite "Parcels" von Cloudera Manager aktiviert haben.

Nach der Installation von Analytic Server befindet sich das Postcheck-Tool im folgenden Verzeichnis:

- **Cloudera**

```
/opt/cloudera/parcels/AnalyticServer-3.2.1.0/tools/com.spss.ibm.checker.zip
```

Die Tools müssen als Root ausgeführt werden und erfordern Python 2.6.X (oder höher).

Wenn das Precheck-Tool Fehler meldet, müssen diese behoben werden, bevor Sie die Analytic Server-Installation fortsetzen.

Ausführen des Precheck-Tools

Automatisch

Das Precheck-Tool kann automatisch als Teil der Analytic Server-Installation aufgerufen werden, wenn Analytic Server über die Cloudera Manager-Konsole als Service installiert wird. Sie müssen den Benutzernamen und das Kennwort des Cloudera Manager-Administrators manuell eingeben:

Abbildung 4. Cloudera Manager-Administratoreinstellungen

Manuell

Sie können das Precheck-Tool manuell auf jedem Clusterknoten aufrufen.

Das folgende Precheck-Beispiel prüft den Cloudera-Cluster `MyCluster`, der auf `myclouderahost.ibm.com:7180` ausgeführt wird, und verwendet die Anmeldeberechtigungsnachweise `admin:admin`:

```
python ./precheck.py --target C --cluster MyCluster --username admin
--password admin --host myclouderahost.ibm.com --port 7180 --as_host myashost.ibm.com
```

Hinweise:

- Der Wert `as_host` muss über die IP-Adresse oder einen vollständig qualifizierten Domänennamen bereitgestellt werden.
- Das Tool fordert zur Eingabe eines Kennworts auf, wenn das Kennwortargument ausgelassen wird.
- Der Befehl `precheck.py` enthält Syntaxhilfe, die mit dem Argument `--h` (`python ./precheck.py --help`) angezeigt werden kann.
- Das Argument `--cluster` ist optional. (Der aktuelle Cluster wird ermittelt, wenn `--cluster` nicht verwendet wird.)

Während das Precheck-Tool seine Prüfungen ausführt, wird der Status jeder Prüfung im Befehlsfenster angezeigt. Wenn ein Fehler auftritt, enthält die Protokolldatei detaillierte Informationen (die genaue Position der Protokolldatei wird im Befehlsfenster angegeben). Die Protokolldatei kann IBM Technical Support bereitgestellt werden, wenn mehr Unterstützung erforderlich ist.

Ausführen des Postcheck-Tools

Das Postcheck-Tool prüft, ob Analytic Server ordnungsgemäß ausgeführt wird und einfache Jobs verarbeiten kann. Das folgende Postcheck-Beispiel prüft eine Analytic Server-Instanz, die auf `myanalyticserverhost.ibm.com:9443` mit aktiviertem SSL ausgeführt wird, und verwendet die Anmeldeberechtigungsnachweise `admin:ibmspss`:

```
python ./postcheck.py --host myanalyticserverhost.ibm.com --port 9443
--username admin --password ibmspss --ssl
```

Wenn Knox mit Analytic Server verwendet wird, lautet der Befehl wie folgt:

```
python ./postcheck.py --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default
```

Führen Sie eine einzelne Prüfung mit dem folgenden Befehl durch:

```
python ./postcheck.py --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default --check AS_PYSPARK_BUILDMODEL
```

Hinweise:

- Das Tool fordert zur Eingabe eines Kennworts auf, wenn das Kennwortargument ausgelassen wird.
- Der Befehl `postcheck.py` enthält Syntaxhilfe, die mit dem Argument `--h` (`python ./postcheck.py --help`) angezeigt werden kann.

Während das Postcheck-Tool seine Prüfungen ausführt, wird der Status jeder Prüfung im Befehlsfenster angezeigt. Wenn ein Fehler auftritt, enthält die Protokolldatei detaillierte Informationen (die genaue Position der Protokolldatei wird im Befehlsfenster angegeben). Die Protokolldatei kann IBM Technical Support bereitgestellt werden, wenn mehr Unterstützung erforderlich ist.

Installation in Cloudera

In den folgenden Schritten wird der Prozess der manuellen Installation von IBM SPSS Analytic Server in Cloudera Manager erläutert.

Analytic Server 3.2.1

Onlineinstallation

1. Navigieren Sie zur IBM Passport Advantage®-Website und laden Sie die selbstextrahierende Binärdatei, die Ihrem Stack, Ihrer Stackversion und Ihrer Hardwarearchitektur entspricht, auf einen Host innerhalb des Cloudera-Clusters herunter. Die verfügbaren Cloudera-Binärdateien sind:

Tabelle 9. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.2.1 für Cloudera 5.11, 5.12, 5.13, 5.14, 5.15, 6.0 und 6.1 Ubuntu (Englisch)	`spss_as-3.2.1.0-cdh5.11-6.1-ubun.bin`
IBM SPSS Analytic Server 3.2.1 für Cloudera 5.11, 5.12, 5.13, 5.14, 5.15, 6.0 und 6.1 Linux x86-64 (Englisch)	`spss_as-3.2.1.0-cdh5.11-6.1-lx86.bin`

2. Führen Sie das selbstextrahierende Cloudera-Installationsprogramm `*.bin` auf dem Cloudera Manager-Master-Clusterknoten aus. Befolgen Sie die Eingabeaufforderungen bei der Installation, indem Sie die Lizenzvereinbarung akzeptieren und das CSD-Standardinstallationsverzeichnis beibehalten.

   Anmerkung: Sie müssen ein anderes CSD-Verzeichnis angeben, wenn die Standardposition geändert wurde.
3. Verwenden Sie den folgenden Befehl, um Cloudera Manager nach Abschluss der Installation erneut zu starten:

   ```
   service cloudera-scm-server restart
   ```
4. Öffnen Sie die Cloudera Manager-Schnittstelle (z. B. `http://${CM-Host}:7180/cmf/login` mit den Standardanmeldeberechtigungsnachweisen `admin/admin`), aktualisieren Sie **Remote Parcel Repository URLs** (in **Host** > **Parcels** > **Configuration**) und prüfen Sie, ob die URL korrekt ist. Beispiel:

   ```
   https://ibm-open-platform.ibm.com
   ```

   Anmerkung: Parcel Update Frequency und **Remote Parcel Repository URLs** können an Ihren Bedarf angepasst werden.

5. Nachdem Cloudera Manager die PARCEL-Dateien aktualisiert hat (Sie können die PARCEL-Dateien manuell aktualisieren, indem Sie auf **Check for New Parcels** klicken), sehen Sie, dass der Status der Analytic Server-PARCEL-Datei auf **Available Remotely** gesetzt ist.
6. Wählen Sie **Download** > **Distribute** > **Activate** aus. Der Status der Analytic Server-PARCEL-Datei wird in **Distributed, Activated** aktualisiert.
7. Fügen Sie Analytic Server in Cloudera Manager als Service hinzu und legen Sie die Position für Analytic Server fest. Im Assistenten zum Hinzufügen eines Service (**Add Service Wizard**) müssen Sie die folgenden Informationen angeben:

   Anmerkung: Der Assistent zum Hinzufügen eines Service (**Add Service Wizard**) zeigt während jeder Phase des Serviceerstellungsprozesses den Gesamtfortschritt an und gibt eine abschließende Bestätigungsnachricht aus, wenn der Service im Cluster erfolgreich erstellt und konfiguriert ist.
   - Hostname für Analytic Server-Metaspeicher
   - Datenbankname für Analytic Server-Metaspeicher
   - Benutzername für Analytic Server-Metaspeicher
   - Kennwort für Analytic Server-Metaspeicher

   MySQL als Metadatenrepository in Analytic Server
   - Analytic Server-Metaspeichertreiberklasse: `com.mysql.jdbc.Driver`
   - Analytic Server-Metaspeicherrepository-URL: `jdbc:mysql://${MySQL-Datenbank}/{Datenbankname}?createDatabaseIfNotExist=true`

     `{MySQL-Datenbank}` ist der Hostname des Servers, auf dem MySQL installiert ist.

   Db2 als Metadatenrepository in Analytic Server
   - Analytic Server-Metaspeichertreiberklasse: `com.ibm.db2.jcc.DB2Driver`
   - Analytic Server-Metaspeicherrepository-URL: `jdbc:db2://{Db2-Host}:{Port}/{Datenbankname}:currentSchema={Schemaname};`

     `{Db2-Host}` ist der Hostname des Servers, auf dem Db2 installiert ist.

     `{Port}` ist der Port, an dem Db2 empfangsbereit ist.

     `{Schemaname}` ist ein verfügbares, nicht verwendetes Schema.

     Wenden Sie sich an Ihren Db2-Administrator, wenn Sie sich nicht sicher sind, welche Werte eingegeben werden sollen.

   LDAP-Konfiguration

   Analytic Server verwendet einen LDAP-Server zum Speichern und Authentifizieren von Benutzern und Gruppen. Sie stellen die erforderlichen LDAP-Konfigurationsinformationen während der Installation von Analytic Server bereit.

Tabelle 10. LDAP-Konfigurationseinstellungen

LDAP-Einstellungen	Beschreibung
`as.ldap.type`	LDAP-Typ. Der Wert kann `ads`, `ad` oder `openladp` sein. • `ads` - Apache Directory Server (Standardeinstellung) • `ad` - Microsoft Active Directory • `openladp` - OpenLDAP
`as.ldap.host`	LDAP-Host
`as.ldap.port`	LDAP-Portnummer
`as.ldap.binddn`	LDAP-Bindungs-DN
`as.ldap.bindpassword`	Kennwort für LDAP-Bindungs-DN
`as.ldap.basedn`	LDAP-Basis-DN

Tabelle 10. LDAP-Konfigurationseinstellungen (Forts.)

LDAP-Einstellungen	Beschreibung
`as.ldap.filter`	LDAP-Benutzer- und -Gruppenfilterregel **Anmerkung:** Wenn dieser Wert vertikaler Striche (`|`) enthält, müssen die Zeichen mit umgekehrten Schrägstrichen als Escapezeichen entwertet werden (Beispiel: `\\|`).
`as.ldap.ssl.enabled`	Gibt an, ob SSL für die Kommunikation zwischen Analytic Server und LDAP verwendet werden soll. Der Wert kann `true` oder `false` sein.
`as.ldap.ssl.reference`	LDAP-SSL-Referenz-ID
`as.ldap.ssl.content`	LDAP-SSL-Konfiguration

- `as.ldap.type` ist standardmäßig auf `ads` gesetzt und die anderen zugehörigen Einstellungen enthalten Standardeinstellungen. Die Ausnahme ist, dass Sie ein Kennwort für die Einstellung `as.ldap.bindpassword` angeben müssen. Analytic Server verwendet die Konfigurationseinstellungen für die Installation einer ADS-Instanz (Apache Directory Server) und zum Ausführen der Serverinitialisierung. Das ADS-Standardprofil schließt den Benutzer `admin` mit dem Kennwort `admin` ein. Sie können die Benutzerverwaltung über die Analytic Server-Konsole durchführen oder Benutzer- und Gruppeninformationen über das Script `importUser.sh` im Ordner `<Analytic Server-Stammverzeichnis>/bin` importieren.
- Wenn Sie planen, einen externen LDAP-Server (z. B. Microsoft Active Directory oder OpenLDAP) zu verwenden, müssen Sie die Konfigurationseinstellungen den tatsächlichen LDAP-Werten entsprechend konfigurieren. Weitere Informationen finden Sie in LDAP-Benutzerregistrys in Liberty konfigurieren.
- Sie können die LDAP-Konfiguration ändern, nachdem Analytic Server installiert wurde (z. B. von Apache Directory Server auf OpenLDAP ändern). Wenn Sie allerdings ursprünglich mit Microsoft Active Directory oder OpenLDAP beginnen und später entscheiden, zu Apache Directory Server zu wechseln, installiert Analytic Server während der Installation keine Apache Directory Server-Instanz. Apache Directory Server wird nur installiert, wenn es während der Erstinstallation von Analytic Server ausgewählt wird.

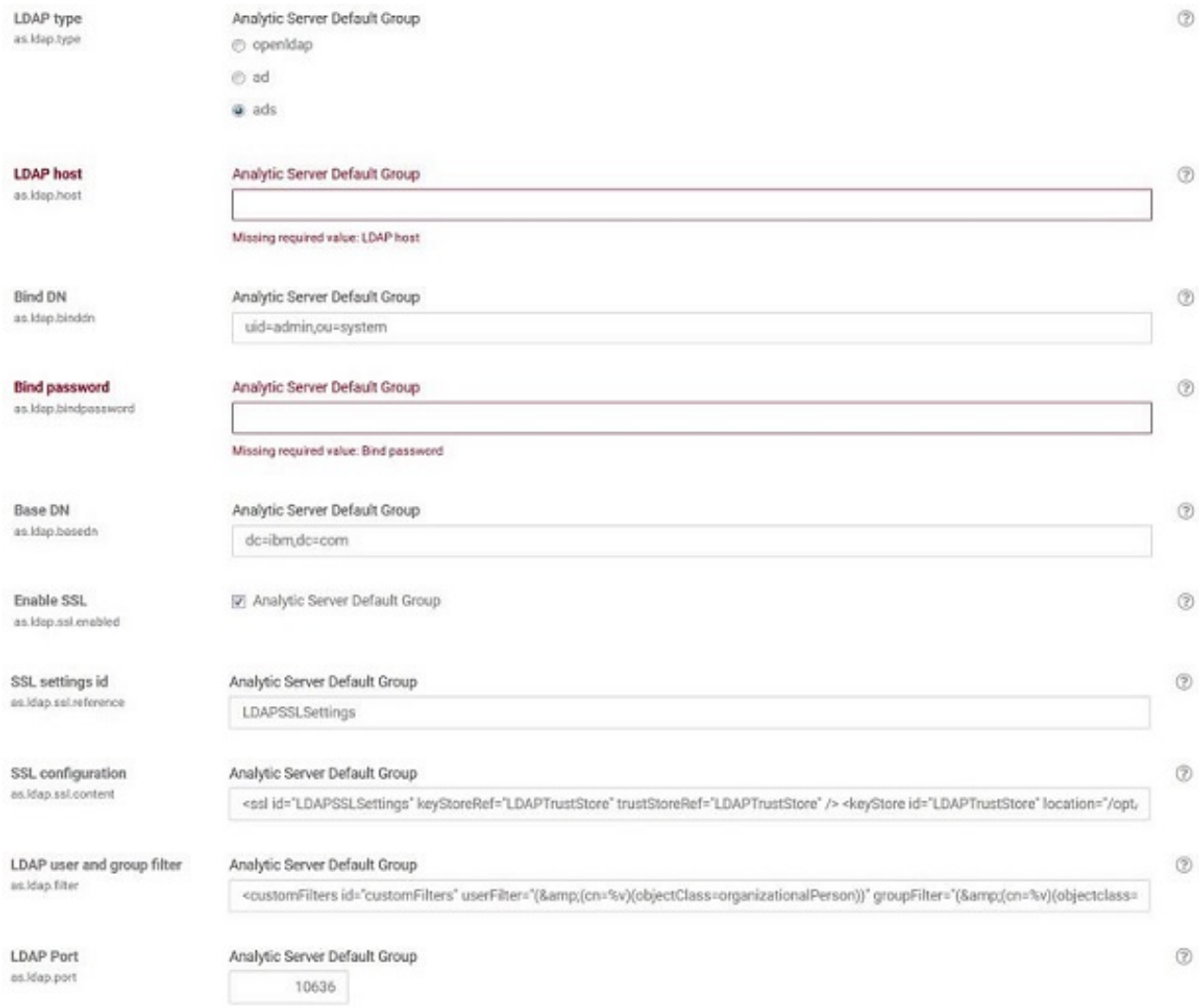

Abbildung 5. Beispiel für LDAP-Konfigurationseinstellungen

8. Wenn Sie Analytic Server in einer Kerberos-fähigen Cloudera-Umgebung installieren, müssen auch die folgenden Einstellungen unter "**Add Service Wizard**" konfiguriert werden:

   Anmerkung: Analytic Server überprüft die HDFS-Konfiguration für auf Kerberos-bezogene Werte, die für die Authentifizierung verwendet werden können.

   - Wählen Sie `Kerberos` für die Einstellung **Analytic Server security** aus, wenn Sie die Kerberos-Authentifizierung beim Anmelden an der Analytic Server-Konsole aktivieren wollen. Wenn **Kerberos** für die Einstellung **Analytic Server security** ausgewählt wird, verwendet die Analytic Server-Konsole standardmäßig den Kerberos-Anmeldemodus.
   - Wählen Sie `Kerberos` für die Einstellung **Analytic Server database data source connection method** aus, wenn Sie eine Verbindung zu Kerberos-fähigen Datenbanken herstellen wollen. Wenn **Kerberos** für die Einstellung **Analytic Server database source connection method** ausgewählt wird, verwendet die Analytic Server-Konsole den Kerberos-Modus, wenn eine Verbindung zu einer Datenbank hergestellt wird.
   - Die Einstellungen **Kerberos Realm Name** und **KDC host** sind erforderlich. Die Werte für den **Kerberos-Realmnamen** (**`as.kdc.realms`**) und den **KDC-Host** (**`kdcserver`**) befinden sich in der Datei `krb5.conf` auf dem KDC-Server (KDC = Kerberos Key Distribution Center).

     Es werden mehrere Realmnamen unterstützt, wenn diese durch Kommas voneinander getrennt sind. Die angegebenen Kerberos-Realmnamen entsprechen Benutzernamen und sind Benutzerna-

men zugeordnet. Die Benutzernamen `UserOne@us.ibm.com` und `UserTwo@eu.ibm.com` würden beispielsweise den Realms `us.ibm.com, eu.ibm.com` entsprechen.

Vertrauensstellungen, die Kerberos-Realms übergreifen, müssen konfiguriert werden, wenn mehrere Realms als **Kerberos-Realmname** angegeben sind. Die Eingabe des Benutzernamens während der Anmeldeaufforderung der Analytic Server-Konsole erfolgt ohne das Suffix des Realmnamens. Infolgedessen wird Benutzern bei Verwendung mehrerer Realms die Dropdown-Liste **Realms** angezeigt, aus der diese den Realm auswählen können.

Anmerkung: Wenn nur ein Realm angegeben ist, wird Benutzern bei der Anmeldung bei Analytic Server die Dropdown-Liste **Realms** nicht angezeigt.

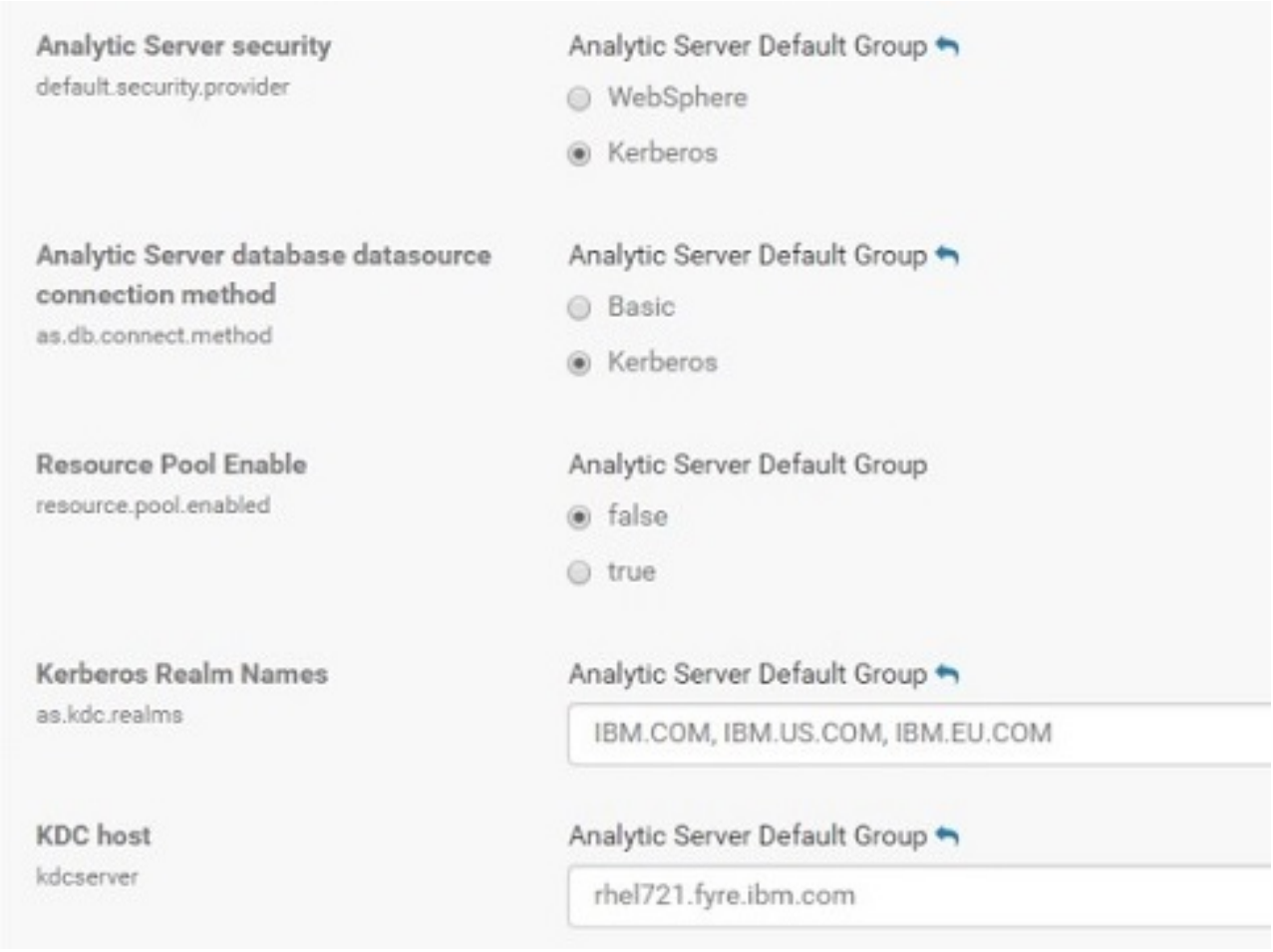

Abbildung 6. Kerberos-Beispieleinstellungen

Hinweise:

- Die Einstellungen **Analytic Server security** und **Analytic Server database data source connection method** gelten für die Authentifizierung auf dem IBM SPSS Modeler-Client und an der Analytic Server-Konsole.
- Wenn **Analytic Server database data source connection method** auf `Kerberos` gesetzt ist, müssen Sie sicherstellen, dass die Zieldatenbanken ebenfalls Kerberos-fähig sind.
- Mit den Einstellungen **Analytic Server security** und **Analytic Server database data source connection method** wird keine Kerberos-Authentifizierung auf dem Hadoop-Cluster konfiguriert. Weitere Informationen finden Sie im Abschnitt "Aktivieren des Kerberos-Identitätswechsels".
- Wenn die Kerberos-Authentifizierung bei der Anmeldung aktiviert werden soll, müssen Sie den IBM SPSS Modeler-Client als gültigen Kerberos-Client bereitstellen. Dazu verwenden Sie den Be-

fehl **addprinc** auf dem Kerberos-KDC-Server (Key Distribution Center). Weitere Informationen finden Sie in der IBM SPSS Modeler-Dokumentation.

Wenn Sie Analytic Server in einer Kerberos-fähigen Cloudera-Umgebung installieren, müssen Sie auch die erforderlichen Konten in Kerberos erstellen und den Kerberos-Identitätswechsel aktivieren. Weitere Informationen finden Sie in „Konfigurieren von Kerberos" auf Seite 49.

Anmerkung: Klicken Sie nach der erfolgreichen Installation von Analytic Server in der Liste "Actions" auf der Seite für Analytic Server-Services in Cloudera Manager nicht auf **Create Analytic Server Metastore**. Beim Erstellen eines Metaspeichers wird das vorhandene Metadatenrepository überschrieben.

Offlineinstallation

Die Schritte für die Offlineinstallation sind dieselben wie für die Onlineinstallation, mit dem Unterschied, dass Sie zuerst die Ihrem Betriebssystem entsprechenden der PARCEL-Dateien und Metadaten manuell herunterladen müssen.

RedHat Linux erfordert die folgenden Dateien:
- AnalyticServer-3.2.1.0-el7.parcel
- AnalyticServer-3.2.1.0-el7.parcel.sha
- manifest.json

SuSE Linux erfordert die folgenden Dateien:
- AnalyticServer-3.2.1.0-sles11.parcel
- AnalyticServer-3.2.1.0-sles11.parcel.sha
- manifest.json

  oder
- AnalyticServer-3.2.1.0-sles12.parcel
- AnalyticServer-3.2.1.0-sles12.parcel.sha

Ubuntu Linux 14.04 erfordert die folgenden Dateien:
- AnalyticServer-3.2.1.0-trusty.parcel
- AnalyticServer-3.2.1.0-trusty.parcel.sha

Ubuntu Linux 16.04 erfordert die folgenden Dateien:
- AnalyticServer-3.2.1.0-xenial.parcel
- AnalyticServer-3.2.1.0-xenial.parcel.sha

1. Laden Sie das selbstextrahierende Cloudera-Installationsprogramm (`*.bin`) auf den Cloudera Manager-Master-Clusterknoten herunter und führen Sie es aus. Befolgen Sie die Eingabeaufforderungen bei der Installation, indem Sie die Lizenzvereinbarung akzeptieren und das Standardinstallationsverzeichnis `CSD` beibehalten.

   Anmerkung: Sie müssen ein anderes `CSD`-Verzeichnis angeben, wenn es sich von der Standardposition unterscheidet.
2. Kopieren Sie die erforderlichen PARCEL- und Metadatendateien in Ihren lokalen Cloudera-Pfad `repo` auf dem Cloudera Manager-Master-Clusterknoten. Der Standardpfad ist `/opt/cloudera/parcel-repo` (der Pfad kann in der Cloudera Manager-Benutzerschnittstelle konfiguriert werden).
3. Verwenden Sie den folgenden Befehl, um Cloudera Manager erneut zu starten:

   ```
   service cloudera-scm-server restart
   ```

   Die Analytic Server-PARCEL-Datei wird als **downloaded** angezeigt, nachdem Cloudera Manager die PARCEL-Datei aktualisiert hat. Sie können auf **Check for New Parcels** klicken, um eine Aktualisierung zu erzwingen.

4. Klicken Sie auf **Distribute** > **Activate**.

   Die Analytic Server-PARCEL-Datei wird als **distributed** und **activated** angezeigt.
5. Fügen Sie Analytic Server in Cloudera Manager als Service hinzu. Weitere Informationen finden Sie in den Schritten 7 und 8 im Abschnitt "Onlineinstallation".

Konfigurieren von Cloudera

Nach der Installation können Sie Analytic Server optional über Cloudera Manager konfigurieren und verwalten.

Anmerkung: Für Analytic Server-Dateipfade gelten die folgenden Konventionen:

- {AS-Stammverzeichnis} bezieht sich auf den Speicherort, an dem Analytic Server bereitgestellt wird, z. B. `/opt/cloudera/parcels/AnalyticServer`.
- {AS-Serverstammverzeichnis} bezieht sich auf den Speicherort der Konfigurations-, Protokoll- und Serverdateien, z. B. `/opt/cloudera/parcels/AnalyticServer/ae_wlpserver/usr/servers/aeserver`.
- {AS-Ausgangsverzeichnis} bezieht sich auf den HDFS-Speicherort, der von Analytic Server als Stammordner verwendet wird, z. B. `/user/as_user/analytic-root`.

Sicherheit

Der Standardwert **`Nutzer-ID`** in der IBM SPSS Modeler-Datei `options.cfg` ist **`ibm`**. Sie können Nutzer in der Analytic Server-Konsole anzeigen. Details zur Nutzerverwaltung finden Sie im Handbuch *IBM SPSS Analytic Server Verwaltung*.

Konfigurieren einer LDAP-Registry

LDAP wird während der Installation von Analytic Server konfiguriert. Sie können nach der Installation von Analytic Server zu einer anderen LDAP-Servermethode wechseln.

Anmerkung: Unterstützung für LDAP in Analytic Server wird durch WebSphere Liberty gesteuert. Weitere Informationen finden Sie in LDAP-Benutzerregistrys in Liberty konfigurieren.

Konfigurieren einer SSL-Verbindung (Secure Socket Layer) von Analytic Server zu LDAP

1. Melden Sie sich an allen Analytic Server-Computern als Analytic Server-Benutzer an und erstellen Sie ein allgemeines Verzeichnis für SSL-Zertifikate.

   Anmerkung: In Cloudera ist der Analytic Server-Benutzer immer der as_user und dies kann nicht geändert werden.
2. Kopieren Sie die Keystore- und Truststore-Dateien auf allen Analytic Server-Computern in dasselbe allgemeine Verzeichnis. Fügen Sie dem Truststore außerdem das Zertifikat einer Zertifizierungsstelle des LDAP-Clients hinzu. Es folgen einige Beispielanweisungen.

   ```
   mkdir /home/as_user/security
   cd /home/as_user/security
   openssl s_client -connect <LDAP-Hostname>:636 -showcerts > client.cert
   $JAVA_HOME/bin/keytool -import -file ./client.cert -alias ldapCA -keystore mytrust.jks
   password : changeit
   ```

   Anmerkung: JAVA_HOME ist dieselbe Java-Ausführungsumgebung (JRE), die auch zum Starten von Analytic Server verwendet wird.
3. Kennwörter können mit dem Tool securityUtility codiert werden, um ihre Werte zu verschlüsseln. Dieses Tool befindet sich in `{AS-Stammverzeichnis}/ae_wlpserver/bin`. Es folgt ein Beispiel.

   ```
   securityUtility encode changeit
           {xor}PDc+MTg6Nis=
   ```
4. Melden Sie sich an Cloudera Manager an und aktualisieren Sie die Analytic Server-Konfigurationseinstellung **`ssl_cfg`** mit den korrekten SSL-Konfigurationseinstellungen. Es folgt ein Beispiel.

```
<ssl id="defaultSSLConfig" keyStoreRef="defaultKeyStore" trustStoreRef="defaultTrustStore"
      clientAuthenticationSupported="true"/>
          <keyStore id="defaultKeyStore" location="/home/as_user/security/mykey.jks" type="JKS"
                    password="{xor}Ozo5PiozKxYdEgwPDAweDG1uDz4sLCg7"/>
          <keyStore id="defaultTrustStore" location="/home/as_user/security/mytrust.jks" type="JKS"
                    password="{xor}PDc+MTg6Nis="/>
```

Anmerkung: Verwenden Sie den absoluten Pfad zu den Keystore- und Truststore-Dateien.

5. Aktualisieren Sie die Konfigurationseinstellung **security_cfg** von Analytic Server mit den korrekten LDAP-Konfigurationseinstellungen. Setzen Sie beispielsweise im Element **ldapRegistry** das Attribut **sslEnabled** auf `true` und das Attribut **sslRef** auf `defaultSSLConfig`.

Konfigurieren von Kerberos

Analytic Server unterstützt Kerberos unter Cloudera. In den folgenden Abschnitten finden Sie die Konfigurationseinstellungen, um sicherzustellen, dass Kerberos ordnungsgemäß in einer Weise konfiguriert wurde, die mit Analytic Server kompatibel ist.

Anmerkung: Analytic Server überprüft die HDFS-Konfiguration für auf Kerberos-bezogene Werte, die für die Authentifizierung verwendet werden können.

Analytic Server- und Kerberos-Einstellungen

Denken Sie an die folgenden Einstellungen, wenn Sie Analytic Server in einer Kerberos-fähigen Cloudera-Umgebung installieren.

- Wählen Sie `Kerberos` für die Einstellung **Analytic Server security** aus, wenn Sie die Kerberos-Authentifizierung beim Anmelden an der Analytic Server-Konsole aktivieren wollen. Wenn **Kerberos** für die Einstellung **Analytic Server security** ausgewählt wird, verwendet die Analytic Server-Konsole standardmäßig den Kerberos-Anmeldemodus.
- Wählen Sie `Kerberos` für die Einstellung **Analytic Server database data source connection method** aus, wenn Sie eine Verbindung zu Kerberos-fähigen Datenbanken herstellen wollen. Wenn **Kerberos** für die Einstellung **Analytic Server database source connection method** ausgewählt wird, verwendet die Analytic Server-Konsole den Kerberos-Modus, wenn eine Verbindung zu einer Datenbank hergestellt wird.
- Die Einstellungen **Kerberos Realm Name** und **KDC host** sind erforderlich. Die Werte für den **Kerberos-Realmnamen** (**as.kdc.realms**) und den **KDC-Host** (**kdcserver**) befinden sich in der Datei `krb5.conf` auf dem KDC-Server (KDC = Kerberos Key Distribution Center).

  Es werden mehrere Realmnamen unterstützt, wenn diese durch Kommas voneinander getrennt sind. Die angegebenen Kerberos-Realmnamen entsprechen Benutzernamen und sind Benutzernamen zugeordnet. Die Benutzernamen `UserOne@us.ibm.com` und `UserTwo@eu.ibm.com` würden beispielsweise den Realms `us.ibm.com, eu.ibm.com` entsprechen.

  Vertrauensstellungen, die Kerberos-Realms übergreifen, müssen konfiguriert werden, wenn mehrere Realms als **Kerberos-Realmname** angegeben sind. Die Eingabe des Benutzernamens während der Anmeldeaufforderung der Analytic Server-Konsole erfolgt ohne das Suffix des Realmnamens. Infolgedessen wird Benutzern bei Verwendung mehrerer Realms die Dropdown-Liste **Realms** angezeigt, aus der diese den Realm auswählen können.

  Anmerkung: Wenn nur ein Realm angegeben ist, wird Benutzern bei der Anmeldung bei Analytic Server die Dropdown-Liste **Realms** nicht angezeigt.

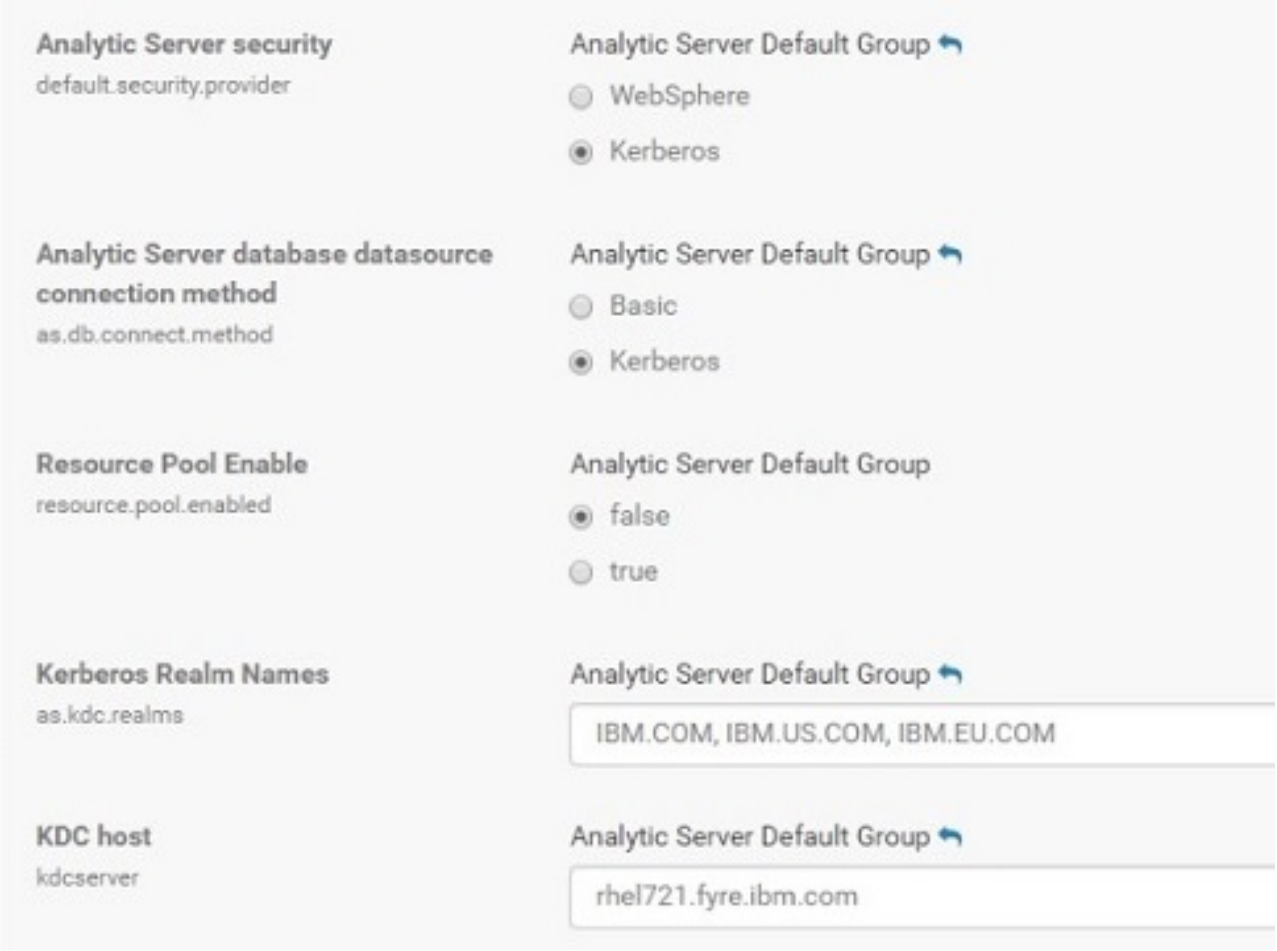

Abbildung 7. Kerberos-Beispieleinstellungen

Hinweise:

- Die Einstellungen **Analytic Server security** und **Analytic Server database data source connection method** gelten für die Authentifizierung auf dem IBM SPSS Modeler-Client und an der Analytic Server-Konsole.
- Wenn **Analytic Server database data source connection method** auf `Kerberos` gesetzt ist, müssen Sie sicherstellen, dass die Zieldatenbanken ebenfalls Kerberos-fähig sind.
- Mit den Einstellungen **Analytic Server security** und **Analytic Server database data source connection method** wird keine Kerberos-Authentifizierung auf dem Hadoop-Cluster konfiguriert. Weitere Informationen finden Sie im Abschnitt "Aktivieren des Kerberos-Identitätswechsels".
- Wenn die Kerberos-Authentifizierung bei der Anmeldung aktiviert werden soll, müssen Sie den IBM SPSS Modeler-Client als gültigen Kerberos-Client bereitstellen. Dazu verwenden Sie den Befehl **`addprinc`** auf dem Kerberos-KDC-Server (Key Distribution Center). Weitere Informationen finden Sie in der IBM SPSS Modeler-Dokumentation.

Erstellen des erforderlichen Kontos in Kerberos

1. Sie können im Kerberos-Benutzerrepository für alle Benutzer, denen Sie Zugriff auf Analytic Server erteilen möchten, Konten erstellen.
2. Erstellen Sie dieselben Konten (aus dem vorherigen Schritt) auf dem LDAP-Server.
3. Erstellen Sie für jeden im vorherigen Schritt erstellten Benutzer auf jedem einzelnen Analytic Server-Knoten und Hadoop-Knoten ein Betriebssystembenutzerkonto.

- Stellen Sie sicher, dass die Benutzer-ID für diese Benutzer auf allen Computern übereinstimmt. Sie können dies testen, indem Sie sich mit dem Befehl `kinit` bei den einzelnen Konten anmelden.
- Stellen Sie sicher, dass die Benutzer-ID der YARN-Einstellung **Minimum user ID for submitting job** entspricht. Dies ist die Einstellung **`min.user.id`** in `container-executor.cfg`. Wenn **`min.user.id`** beispielsweise auf 1000 gesetzt ist, muss die Benutzer-ID jedes erstellten Benutzerkontos größergleich 1000 sein.

4. Erstellen Sie in HDFS einen Benutzerausgangsordner für den Analytic Server-Administrator. Die Ordnerberechtigung muss auf 777 gesetzt werden, der Eigner muss als `admin` definiert werden und die Benutzergruppe muss auf `hdfs` gesetzt werden. Siehe das folgende Beispiel in Fettdruck:

```
[root@xxxxx configuration]# hadoop fs -ls /user

Found 9 items

drwxrwxrwx   - hdfs     supergroup      0 2017-07-26 03:41 /user/AE
drwxrwxrwx   - admin    hdfs            0 2017-06-08 01:33 /user/admin
drwxr-x--x   - as_user  hdfs            0 2017-06-06 01:00 /user/as_user
drwx------   - hdfs     supergroup      0 2017-07-31 00:17 /user/hdfs
drwxrwxrwx   - mapred   hadoop          0 2017-06-05 00:28 /user/history
drwxrwxr-t   - hive     hive            0 2017-06-05 00:30 /user/hive
drwxrwxr-x   - hue      hue             0 2017-06-05 00:30 /user/hue
drwxrwxr-x   - impala   impala          0 2017-07-19 00:52 /user/impala
drwxr-x--x   - spark    spark           0 2017-06-05 01:34 /user/spark
```

5. Wenn Sie HCatalog-Datenquellen verwenden wollen und Analytic Server auf einem anderen Computer als Hive-Metaspeicher installiert ist, müssen Sie in HDFS die Identität des Hive-Clients annehmen.
   a. Navigieren Sie in Cloudera Manager zur Registerkarte **Configuration** des HDFS-Service.

      Anmerkung: Die folgenden Einstellungen werden möglicherweise nicht auf der Registerkarte **Configuration** angezeigt, wenn sie nicht bereits festgelegt wurden. Führen Sie in diesem Fall eine Suche nach ihnen aus.
   b. Bearbeiten Sie die Einstellung **`hadoop.proxyuser.hive.groups`** so, dass sie den Wert `*` hat oder eine Gruppe enthält, die alle Benutzer umfasst, die sich an Analytic Server anmelden können.
   c. Bearbeiten Sie die Einstellung **`hadoop.proxyuser.hive.hosts`** so, dass sie den Wert `*` hat oder die Liste der Hosts enthält, auf denen der Hive-Metaspeicher und alle Instanzen von Analytic Server als Service installiert sind.
   d. Starten Sie den HDFS-Service erneut.

Nachdem Sie diese Schritte ausgeführt haben und Analytic Server installiert ist, konfiguriert Analytic Server Kerberos automatisch im Hintergrund.

Aktivieren des Kerberos-Identitätswechsels

Durch Identitätswechsel kann ein Thread in einem Sicherheitskontext ausgeführt werden, der sich vom Sicherheitskontext des Prozesses unterscheidet, der der Threadeigner ist. Beispielsweise können Hadoop-Jobs mithilfe von Identitätswechsel über einen anderen Benutzer als den Analytic Server-Standardbenutzer (`as_user`) ausgeführt werden. So aktivieren Sie den Kerberos-Identitätswechsel:

1. Öffnen Sie Cloudera Manager und fügen Sie im Bereich **Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml** (auf der Registerkarte **HDFS (Service-Wide)** > **Configuration**) die folgenden Eigenschaften hinzu oder aktualisieren diese.
   - **Name:** `hadoop.proxyuser.as_user.hosts`
   - **Value:** `*`
   - **Name:** `hadoop.proxyuser.as_user.groups`
   - **Value:** `*`

   Anmerkung: Die Einstellung **core-site.xml** gilt für die Hadoop-Konfiguration (nicht Analytic Server).
2. Führen Sie den folgenden Befehl in einer Befehlsshell auf dem Analytic Server-Knoten aus:

```
hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
```

Konfigurieren von HAProxy für Kerberos-SSO (Single Sign On)

1. Konfigurieren und starten Sie HAProxy wie in der Dokumentation zu HAProxy unter http://www.haproxy.org/#docs beschrieben.
2. Erstellen Sie den Kerberos-Prinzipal (HTTP/<Proxy-Hostname>@<Realm>) und die Chiffrierschlüsseldatei für den HAProxy-Host, wobei <Proxy-Hostname> der vollständige Name des HAProxy-Hosts und <Realm> der Kerberos-Realm ist.
3. Kopieren Sie die Chiffrierschlüsseldatei als `/etc/security/keytabs/spnego_proxy.service.keytab` auf alle Analytic Server-Hosts.
4. Aktualisieren Sie die Berechtigungen für diese Datei auf allen Analytic Server-Hosts. Es folgt ein Beispiel.

   ```
   chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
   chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
   ```

5. Öffnen Sie Cloudera Manager und fügen Sie die folgenden Eigenschaften im Analytic Server-Bereich **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties** hinzu oder aktualisieren Sie sie.

   ```
   web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
   web.authentication.kerberos.principal=HTTP/<vollständiger Name des Proxy-Computers>@<Realm>
   ```

6. Speichern Sie die Konfiguration und starten Sie alle Analytic Server-Services über Cloudera Manager erneut.
7. Weisen Sie Benutzer an, ihre Browser für die Verwendung von Kerberos zu konfigurieren.

Benutzer können sich jetzt über die Option **Single sign on log in** auf dem Anmeldebildschirm von IBM SPSS Analytic Server bei Analytic Server anmelden.

Inaktivieren von Kerberos

1. Inaktivieren Sie Kerberos in der Cloudera Manager-Konsole.
2. Stoppen Sie den Analytic Server-Service.
3. Entfernen Sie die folgenden Einstellungen aus dem Bereich **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**.

   ```
   default.security.provider
   hdfs.keytab
   hdfs.user
   java.security.krb5.conf
   as.db.connect.method
   web.authentication.kerberos.keytab
   web.authentication.kerberos.principal
   ```

4. Klicken Sie auf **Save Changes** und starten Sie den Analytic Server-Service erneut.

Aktivieren von SSL-Verbindungen (Secure Socket Layer) zur Analytic Server-Konsole

Standardmäßig generiert Analytic Server selbst signierte Zertifikate, um SSL (Secure Socket Layer) zu aktivieren. Wenn Sie die selbst signierten Zertifikate akzeptieren, können Sie so über den sicheren Port auf die Analytic Server-Konsole zugreifen. Für einen sichereren HTTPS-Zugriff müssen Sie Zertifikate eines anderen Anbieters installieren.

Führen Sie die folgenden Schritte aus, um Zertifikate eines anderen Anbieters zu installieren.

1. Kopieren Sie auf allen Analytic Server-Knoten die Keystore- und Truststore-Zertifikate eines anderen Anbieters in dasselbe Verzeichnis, beispielsweise in `/home/as_user/security`.

   Anmerkung: Der Analytic Server-Benutzer muss über Lesezugriff auf dieses Verzeichnis verfügen.
2. Navigieren Sie in Cloudera Manager zur Registerkarte "Configuration" des Analytic Server-Service.
3. Bearbeiten Sie den Parameter **`ssl_cfg`**.

```
<ssl id="defaultSSLConfig"
     keyStoreRef="defaultKeyStore"
     trustStoreRef="defaultTrustStore"
     clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
          location="<KEYSTOREPOSITION>"
          type="<TYP>"
          password="<KENNWORT>"/>
<keyStore id="defaultTrustStore"
          location="<TRUSTSTOREPOSITION>"
          type="<TYP>"
          password="<KENNWORT>"/>
```

Ersetzen Sie Folgendes:

- <KEYSTOREPOSITION> durch die absolute Position des Keystores. Beispiel: `/home/as_user/security/mykey.jks`
- <TRUSTSTOREPOSITION> durch die absolute Position des Truststores. Beispiel: `/home/as_user/security/mytrust.jks`
- <TYP> durch den Typ des Zertifikats. Beispiel: JKS, PKCS12 usw.
- <KENNWORT> durch das verschlüsselte Kennwort im Base64-Verschlüsselungsformat. Für die Verschlüsselung können Sie das Tool securityUtility verwenden. Beispiel: `{AS-Stammverzeichnis}/ae_wlpserver/bin/securityUtility encode <Kennwort>`

Wenn Sie ein selbst signiertes Zertifikat generieren wollen, können Sie das Tool securityUtility verwenden. Beispiel: `{AS-Stammverzeichnis}/ae_wlpserver/bin/securityUtility createSSLCertificate --server=myserver --password=mypassword --validity=365 --subject=CN=mycompany,O=myOrg,C=myCountry`. Weitere Informationen zu securityUtility und anderen SSL-Einstellungen finden Sie in der Dokumentation zum WebSphere Liberty-Profil.

Anmerkung: Sie müssen einen entsprechenden Hostdomänennamen für den `CN`-Wert angeben.

4. Klicken Sie auf **Save Changes** und starten Sie den Analytic Server-Service erneut.

Kommunizieren mit Apache Hive über SSL

Sie müssen die Datei `hive.properties` aktualisieren, um über eine SSL-Verbindung mit Apache Hive zu kommunizieren. Wenn die Hochverfügbarkeit in Ihrer Apache Hive-Umgebung aktiviert ist, können Sie alternativ die Hochverfügbarkeitsparameter auf der Analytic Server-Hauptseite für Datenquellen auswählen.

Aktualisieren der Datei 'hive.properties'

1. Öffnen Sie die Datei `hive.properties`. Die Datei befindet sich an der folgenden Position: `/opt/cloudera/parcels/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/database`
2. Suchen Sie die folgende Zeile:

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password}
```

3. Aktualisieren Sie die Zeile, indem Sie die folgenden **fett** angegebenen Informationen hinzufügen:

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password};
ssl=true;sslTrustStore=PfadzurTruststore-Datei;trustStorePassword=xxxTruststore-Kennwort
```

4. Speichern Sie die Datei `hive.properties`.

Aktivieren der Unterstützung für Essentials for R

Analytic Server unterstützt das Scoren von R-Modellen und das Ausführen von R-Scripts.

So installieren Sie Essentials for R nach einer erfolgreichen Installation von Analytic Server in Cloudera Manager:

1. Richten Sie die Serverumgebung für Essentials for R ein. Weitere Informationen finden Sie in Schritt 1 in „Aktivieren der Unterstützung für Essentials for R" auf Seite 22.

2. Laden Sie das sich selbst entpackende Archiv (BIN) für den RPM für IBM SPSS Modeler Essentials for R herunter. Essentials for R ist als Download verfügbar (https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp). Wählen Sie die Ihrem Stack, Ihrer Stackversion und Ihrer Hardwarearchitektur entsprechende Datei aus.
3. Führen Sie das sich selbst entpackende Archiv als Root- oder sudo-Benutzer auf dem Cloudera Manager-Server-Host aus. Die folgenden Pakete müssen installiert oder in den konfigurierten Repositorys verfügbar sein:
   - Red Hat Linux: gcc-gfortran, zip, gcc-c++
   - SUSE Linux: gcc-fortran, zip, gcc-c++
   - Ubuntu Linux: gcc-fortran, zip, gcc-c++
4. Das sich selbst entpackende Installationsprogramm führt die folgenden Aufgaben aus:
   a. Zeigt die erforderlichen Lizenzen an und fordert den Installationsverantwortlichen auf, sie zu akzeptieren.
   b. Fordert den Installationsverantwortlichen auf, die R-Quellenposition anzugeben oder mit der Standardposition fortzufahren. Standardmäßig wird R Version 3.3.2 installiert. So installieren Sie eine andere Version:
      - Onlineinstallation: Geben Sie die URL zum Archiv der erforderlichen R-Version an. Beispiel: https://cran.r-project.org/src/base/R-2/R-2.15.3.tar.gz für R 2.15.3.
      - Offlineinstallation: Laden Sie das Archiv der erforderlichen R-Version herunter und kopieren Sie es auf den Cloudera Manager-Server-Host. Benennen Sie das Archiv nicht um (standardmäßig heißt es `R-x.x.x.tar.gz`). Geben Sie die URL zu dem kopierten R-Archiv wie folgt an: `file://<R-Archivverzeichnis>/R-x.x.x.tar.gz`. Wenn das Archiv `R-2.15.3.tar.gz` heruntergeladen und dann in `/root` kopiert wurde, lautet die URL `file:///root/R-2.15.3.tar.gz`.

      Anmerkung: Weitere R-Versionen finden Sie unter https://cran.r-project.org/src/base/.
   c. Installiert die für R erforderlichen Pakete.
   d. Lädt R und das Plug-in Essentials for R herunter und installiert sie.
   e. Erstellt die PARCEL-Datei und die Datei `parcel.sha` und kopiert sie in `/opt/cloudera/parcel-repo`. Geben Sie den korrekten Speicherort ein, wenn der Speicherort geändert wurde.
5. Nachdem die Installation abgeschlossen wurde, verteilen und aktivieren Sie die PARCEL-Datei für **Essentials for R** in Cloudera Manager (klicken Sie auf **Check for New Parcels**, um die Liste der PARCEL-Dateien zu aktualisieren).
6. Wenn der Analytic Server-Service bereits installiert ist:
   a. Stoppen Sie den Service.
   b. Aktualisieren Sie die Analytic Server-Binärdateien.
   c. Starten Sie den Service, um die Installation von Essentials for R abzuschließen.
7. Wenn der Analytic Server-Service nicht installiert ist, fahren Sie mit dessen Installation fort.

Anmerkung: Für alle Analytic Server-Hosts müssen die entsprechenden Archivpakete (`zip` und `unzip`) installiert sein.

Aktivieren relationaler Datenbankquellen

Wenn Sie die JDBC-Treiber in einem gemeinsam genutzten Verzeichnis in allen Analytic Server-Metaspeichern und auf allen Analytic Server-Hosts bereitstellen, kann Analytic Server relationale Datenbankquellen verwenden. Standardmäßig wird hierzu das Verzeichnis `/usr/share/jdbc` verwendet.

Führen Sie die folgenden Schritte aus, um das gemeinsam genutzte Verzeichnis zu ändern.

1. Navigieren Sie in Cloudera Manager zur Registerkarte "Configuration" des Analytic Server-Service.
2. Geben Sie in **`jdbc.drivers.location`** den Pfad zum gemeinsam genutzten Verzeichnis mit den JDBC-Treibern an.

3. Klicken Sie auf **Save Changes**.
4. Wählen Sie aus dem Dropdown-Menü **Actions** die Option **Stop** aus, um den Analytic Server-Service zu stoppen.
5. Wählen Sie **Refresh Analytic Server Binaries** im Dropdown-Menü **Actions** aus.
6. Wählen Sie aus dem Dropdown-Menü **Actions** die Option **Start** aus, um den Analytic Server-Service zu starten.

Tabelle 11. Unterstützte Datenbanken

Datenbank	Unterstützte Versionen	JAR-Dateien für JDBC-Treiber	Anbieter
Amazon Redshift	8.0.2 oder später	`RedshiftJDBC41-1.1.6.1006.jar` oder später	Amazon
Apache Impala	JDBC 4 mit Version 2.5.5 oder später	`ImpalaJDBC4.jar`, `commons-codec-*.jar`, `commons-logging-*.jar`, `httpclient-*.jar`, `httpcore-*.jar`, `log4j-*.jar`, `libthrift-*.jar`, `libfb303-*.jar`, `slf4j-api-*.jar`, `ql.jar`, `zookeeper-*.jar`, `TCLIServiceClient.jar`	Apache
DashDB	Bluemix-Service	`db2jcc.jar`	IBM
Db2 for Linux, UNIX, and Windows	11.1, 10.5, 10.1, 9.7	`db2jcc.jar`	IBM
Db2 z/OS	11, 10	`db2jcc.jar`, `db2_license_cisuz.jar`	IBM
Greenplum	5.x	postgresql.jar	Greenplum
Hive	1.1	`hive-jdbc-*.jar`	Apache
MySQL	5.6, 5.7	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7, 6.x	`nzjdbc.jar`	IBM
Oracle	12c, 11g R2 (11.2)	`ojdbc6.jar`, `orai18n.jar`	Oracle
SQL Server	2014, 2012, 2008 R2	`sqljdbc4.jar`	Microsoft
Teradata	15, 15.1	`tdgssconfig.jar`, `terajdbc4.jar`	Teradata

Hinweise

- Wenn Sie vor der Installation von Analytic Server eine Redshift-Datenquelle erstellt haben, müssen Sie die folgenden Schritte ausführen, damit die Redshift-Datenquelle verwendet werden kann.
  1. Öffnen Sie die Redshift-Datenquelle in der Analytic Server-Konsole.
  2. Wählen Sie die `Redshift`-Datenbankdatenquelle aus.
  3. Geben Sie die Redshift-Serveradresse ein.
  4. Geben Sie den Datenbanknamen und den Benutzernamen ein. Das Kennwort sollte automatisch ausgefüllt werden.
  5. Wählen Sie die Datenbanktabelle aus.

Aktivieren von HCatalog-Datenquellen

Analytic Server bietet über Hive/HCatalog Unterstützung für zahlreiche Datenquellen. Für einige Quellen sind Schritte zur manuellen Konfiguration erforderlich.

1. Erfassen Sie die für die Aktivierung der Datenquelle erforderlichen JAR-Dateien. Details hierzu finden Sie in den folgenden Abschnitten.
2. Fügen Sie diese JAR-Dateien zum Verzeichnis `{HIVE_HOME}/auxlib` und zum Verzeichnis `/usr/share/hive` in allen Analytic Server-Metaspeichern und auf allen Analytic Server-Knoten hinzu.
3. Starten Sie den Hive-Metaspeicherservice erneut.
4. Starten Sie jede einzelne Instanz des Analytic Server-Service erneut.

Anmerkung:

Wenn Sie über eine HCatalog-Datenquelle in Analytic Server auf HBase-Daten zugreifen, muss der zugreifende Benutzer über Leseberechtigung für die HBase-Tabellen verfügen.

- In Umgebungen, die kein Kerberos verwenden, greift Analytic Server mit `as_user` (`as_user` muss Leseberechtigung für HBase haben) auf HBase zu.
- In Kerberos-Umgebungen müssen `as_user` und der angemeldete Benutzer eine Leseberechtigung für HBase-Tabellen haben.

NoSQL-Datenbanken

Analytic Server unterstützt NoSQL-Datenbanken, für die ein Hive-Speicherhandler vom Anbieter verfügbar ist.

Für die Aktivierung der Unterstützung für Apache HBase und Apache Accumulo sind keine zusätzlichen Schritte erforderlich.

Bei anderen NoSQL-Datenbanken wenden Sie sich an den Datenbankanbieter, um den Speicherhandler und die entsprechenden JAR-Dateien zu erhalten.

Dateibasierte Hive-Tabellen

Analytic Server unterstützt dateibasierte Hive-Tabellen, für die ein integrierter oder angepasster Hive SerDe (Parallel-Seriell- und Seriell-Parallel-Umsetzer) verfügbar ist.

Der Hive XML SerDe für die Verarbeitung von XML-Dateien befindet sich im Maven Central Repository unter http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde.

Jobs für MapReduce Version 2

Verwenden Sie die Einstellung **`preferred.mapreduce`** im Bereich **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**, um zu steuern, wie MapReduce-Jobs verarbeitet werden:

Tabelle 12. Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties

Eigenschaft	Beschreibung
`preferred.mapreduce`	Steuert die Methode, in der MapReduce-Jobs ausgeführt werden. Gültige Werte sind unter anderem: • `spark` • `m3r` • `hadoop` Beispiel: `preferred.mapreduce=spark`

Apache Spark

Wenn Sie Spark (Version 1.5 oder höher) verwenden möchten, müssen Sie die Spark-Version (`spark.version`) während der Installation von Analytic Server auswählen.

1. Öffnen Sie Cloudera Manager und wählen Sie die geeignete Spark-Version (`spark.version`, z. B. `None`, `1.x` oder `2.x`), im Bereich **Analytic Server Spark Version** aus.

   Anmerkung: Wenn Sie Spark 1.x verwenden, müssen Sie auch die folgende Zeile im Bereich **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties** hinzufügen.

   ```
   spark.extraListeners=org.apache.spark.JavaSparkListener
   ```

2. Speichern Sie die Konfiguration.

Konfigurieren von Apache Impala

Apache Impala wird unterstützt, wenn es unter Cloudera für eine Analytic Server-Datenbankdatenquelle oder eine HCatalog-Datenquelle ausgeführt wird (unabhängig davon, ob für Impala SSL aktiviert ist).

Erstellen einer Datenbankdatenquelle für Apache Impala-Daten

1. Klicken Sie auf der Analytic Server-Hauptseite **Data sources** auf **New**, um eine neue Datenquelle zu erstellen. Das Dialogfeld **New data source** wird angezeigt.
2. Geben Sie einen passenden Namen in das Feld **New data source** ein, wählen Sie `Database` als Wert für den Inhaltstyp aus und klicken Sie dann auf **OK**.
3. Öffnen Sie den Abschnitt **Database Selections** und geben Sie die folgenden Informationen ein.

   Database:
   Wählen Sie **Impala** im Dropdown-Menü aus.

   Server address:
   Geben Sie die URL des Servers ein, auf dem sich der Impala-Dämon befindet. Wenn Kerberos für Analytic Server aktiviert ist, ist ein vollständig qualifizierter Domänenname erforderlich.

   Server port:
   Geben Sie die Nummer des Ports ein, an dem die Impala-Datenbank empfangsbereit ist.

   Database name:
   Geben Sie den Namen der Datenbank ein, zu der Sie eine Verbindung herstellen wollen.

   Username:
   Geben Sie einen Benutzernamen mit der Berechtigung zum Anmelden an der Impala-Datenbank ein.

   Password:
   Geben Sie das zum Benutzernamen gehörige Kennwort ein.

   Table name:
   Geben Sie den Namen einer Tabelle aus der Datenbank ein, die Sie verwenden wollen. Klicken Sie auf **Select**, um eine Datei manuell auszuwählen.

   Maximum concurrent reads:
   Geben Sie den Grenzwert für die Anzahl paralleler Abfragen ein, die von Analytic Server zur Datenbank gesendet werden können, um aus der in der Datenquelle angegebenen Tabelle zu lesen.

4. Klicken Sie auf **Save**, nachdem Sie alle erforderlichen Informationen eingegeben haben.

Erstellen einer HCatalog-Datenquelle für Apache Impala-Daten

1. Klicken Sie auf der Analytic Server-Hauptseite **Data sources** auf **New**, um eine neue Datenquelle zu erstellen. Das Dialogfeld **New data source** wird angezeigt.

2. Geben Sie einen passenden Namen in das Feld **New data source** ein, wählen Sie `HCatalog` als Wert für den Inhaltstyp aus und klicken Sie dann auf **OK**.
3. Öffnen Sie den Abschnitt **Database Selections** und geben Sie die folgenden Informationen ein.

   Database:
   Wählen Sie **default** im Dropdown-Menü aus.

   Table name:
   Geben Sie den Namen einer Tabelle aus der Datenbank ein, die Sie verwenden wollen.

   HCatalog Schema
   Wählen Sie die Option **HCatalog Element** und anschließend die entsprechenden Optionen für die HCatalog-Feldzuordnungen aus.
4. Klicken Sie auf **Save**, nachdem Sie alle erforderlichen Informationen eingegeben haben.

Herstellen der Verbindung zu Apache Impala-Daten

1. Definieren Sie die folgenden Impala-SSL-Einstellungen in der Cloudera Manager-Konsole.

   Enable TLS/SSL for Impala (client_services_ssl_enabled)
   Wählen Sie die Option **Impala (Service-Wide)** aus.

   Impala TLS/SSL Server Certificate File (PEM Format) (ssl_server_certificate)
   Geben Sie den Speicherort und den Dateinamen des selbst signierten Zertifikats im PEM-Format ein (Beispiel: `/tmp/<Benutzername>/ssl/ll4200v21.crt`).

   Impala TLS/SSL Server Private Key File (PEM Format) (ssl_private_key)
   Geben Sie den Speicherort und den Dateinamen des privaten Schlüssels ein (Beispiel: `/tmp/<Benutzername>/ssl/ll4200v21.key`).
2. Importieren Sie auf dem Analytic Server-Host die Datei `*.crf` (wird zum Aktivieren von SSL für Impala verwendet) in eine `*.jks`-Datei. Dies kann eine Datei 'cacerts' sein (zum Beispiel `/etc/pki/java/cacerts`) oder eine andere, beliebige `*.jks`-Datei.
3. Aktualisieren Sie auf dem Analytic Server-Host die Impala-Konfigurationsdatei (`impala.properties`), indem Sie den folgenden `jdbcurl`-Schlüsselwert hinzufügen:

   ```
   SSL=1;AllowSelfSignedCerts=1;CAIssuedCertNamesMismatch=1;
   ```

   Anmerkung: Wenn eine `*.jks`-Datei (nicht 'cacerts') verwendet wird, müssen Sie zudem Folgendes angeben:

   ```
   SSLTrustStore=<Ihre_pks-Datei>;SSLTrustStorePwd=<Kennwort_für_pks-Datei>;
   ```
4. Starten Sie Analytic Server in der Cloudera Manager-Konsole erneut.

Ändern der von Analytic Server verwendeten Ports

Analytic Server verwendet standardmäßig Port 9080 für HTTP und Port 9443 für HTTPS. Führen Sie die folgenden Schritte aus, um die Porteinstellungen zu ändern.

1. Navigieren Sie in Cloudera Manager zur Registerkarte "Configuration" des Analytic Server-Service.
2. Geben Sie den gewünschten HTTP- und HTTPS-Port in den Parametern **`http.port`** bzw. **`https.port`** an.

   Anmerkung: Möglicherweise müssen Sie die Kategorie **Ports and Addresses** im Abschnitt **Filters** auswählen, damit diese Parameter angezeigt werden.
3. Klicken Sie auf **Save Changes**.
4. Starten Sie den Analytic Server-Service erneut.

Analytic Server mit hoher Verfügbarkeit

Sie können Hochverfügbarkeit für Analytic Server bereitstellen, indem Sie das Produkt als Service für mehrere Knoten in Ihrem Cluster hinzufügen.

1. Navigieren Sie in Cloudera Manager zur Registerkarte "Instances" des Analytic Server-Service.
2. Klicken Sie auf **Add Role Instances** und wählen Sie die Hosts aus, auf denen Analytic Server als Service hinzugefügt werden soll.

Unterstützung mehrerer Cluster

Die Mehrclusterfunktion ist eine Erweiterung der Hochverfügbarkeitsfunktion von IBM SPSS Analytic Server und ermöglicht eine bessere Isolation in Umgebungen mit mehreren Nutzern. Standardmäßig wird bei der Installation des Analytic Server-Service (in Ambari oder Cloudera Manager) ein einzelner Analyse-Server-Cluster definiert.

Die Clusterspezifikation definiert die Analytic Server-Clusterzugehörigkeit. Die Clusterspezifikation wird über XML-Inhalt geändert (im Feld `analytics-cluster` von Ambari für die Analytic Server-Konfiguration oder durch manuelles Bearbeiten der Cloudera Manager-Datei `configuration/analytics-cluster.xml`). Wenn Sie mehrere Analytic Server-Cluster konfigurieren, müssen den einzelnen Analytic Server-Clustern Anforderungen über die Lastausgleichsfunktionen der Cluster zugeführt werden.

Durch die Verwendung der Mehrclusterfunktion wird sichergestellt, dass die Arbeit für einen Nutzer sich nicht negativ auf die Arbeit im Cluster eines anderen Nutzers auswirkt. Bei Hochverfügbarkeitsjobs kommt es nur innerhalb des Analytic Server-Clusters zu einem Job-Failober, auf dem die Arbeit initialisiert wurde. Das folgende Beispiel zeigt eine XML-Spezifikation für mehrere Cluster:

Anmerkung: Analytic Server kann als hochverfügbar definiert werden, indem Sie das Produkt mehreren Knoten in Ihrem Cluster als Service hinzufügen.

```
<analayticServerClusterSpec>
    <cardinality>1+</cardinality>
    <cluster name="cluster1">
        <memberName>one.cluster</memberName>
        <memberName>two.cluster</memberName>
    </cluster>
    <cluster name="cluster2">
        <memberName>three.cluster</memberName>
        <memberName>four.cluster</memberName>
    </cluster>
</analayticServerClusterSpec>
```

Im vorherigen Beispiel sind zwei Lastausgleichsfunktionen erforderlich. Eine Lastausgleichsfunktion sendet Anforderungen an die Member von `cluster1` (`one.cluster` und `two.cluster`) und die andere sendet Anforderungen an die Member von `cluster2` (`three.cluster` und `four.cluster`).

Das folgende Beispiel stellt eine XML-Spezifikation für einen einzelnen Cluster bereit (Standardkonfiguration).

```
<analayticServerClusterSpec>
    <cardinality>1</cardinality>
    <cluster name="cluster1">
        <memberName>*</memberName>
    </cluster>
</analayticServerClusterSpec>
```

Im vorherigen Beispiel ist eine einzige Lastausgleichsfunktion erforderlich, um die Fälle zu bearbeiten, in denen mehrere Cluster-Member konfiguriert sind.

Hinweise

- Nur Singleton-Cluster unterstützen die Verwendung von Platzhalterzeichen im Element **`memberName`** (z. B. Clusterkardinalität = "1"). Gültige Werte für das Kardinalitätselement sind 1 und 1+.
- Der Membername (**`memberName`**) muss auf dieselbe Weise wie der Name des Hosts angegeben werden, dem die Analytic Server-Rolle zugewiesen ist.
- Alle Server in allen Clustern müssen erneut gestartet werden, nachdem die Änderungen der Clusterkonfiguration angewendet wurden.

- In Cloudera Manager müssen Sie die Datei `analytics-cluster.xml` auf allen Analytic Server-Knoten ändern und warten. Alle Knoten müssen gewartet werden, um sicherzustellen, dass sie denselben Inhalt haben.

Optimieren von JVM-Optionen für Small Data

Sie können JVM-Eigenschaften bearbeiten, um Ihr System für die Ausführung von Small Jobs (M3R) zu optimieren.

In Cloudera Manager befindet sich das Steuerelement **Jvm Options (jvm.options)** auf der Registerkarte "Configuration" im Analytic Server-Service. Durch Ändern der folgenden Parameter wird die Größe des Heapspeichers für Jobs festgelegt, die auf dem Server ausgeführt werden, der Analytic Server hostet, also nicht Hadoop. Dies ist bei der Ausführung von Small Jobs (M3R) wichtig. Möglicherweise müssen Sie mit diesen Werten experimentieren, um Ihr System zu optimieren.

```
-Xms512M
-Xmx2048M
```

Konfigurieren separater YARN-Warteschlangen für jeden IBM SPSS Analytic Server-Nutzer - Cloudera

Die Konfiguration von YARN-Warteschlangen erfolgt durch die Verwendung von Spark-Techniken zur dynamischen Ressourcenzuordnung.

Cloudera 5.x

Führen Sie die folgenden Schritte aus, wenn Sie den SPSS Analytic Server-Service einem vorhandenen Cluster hinzufügen.

1. Navigieren Sie in Cloudera Manager zu **SPSS Analytic Server Service** > **Configuration**.
2. Ändern Sie den Wert **Resource Pool Enable: resource.pool.enabled** in `true`.
3. Fügen Sie die folgenden Eigenschaften zu **Analytic Server Advanced Configuration Snippet (Safety Valve)** > **analyticserver-conf.config.properties** hinzu:

```
config.folder.path=/etc/spark2/conf
resource.pool.mapping=tenant1:test,tenant2:production
resource.pool.default=default
spark.scheduler.mode=FAIR
spark.yarn.queue=default
```

Tabelle 13. Einstellungen in 'analyticserver-conf.config.properties'

Eigenschaft	Beschreibung
`config.folder.path`	Das Verzeichnis enthält die Datei `fairscheduler.xml`, die die Eigenschaftsinformationen für den Spark-Pool enthält. Die Datei ist erforderlich und muss manuell erstellt werden. Weitere Informationen finden Sie im Abschnitt **fairscheduler.xml - Beispiel**.
`resource.pool.mapping`	**Spark:** Ordnet die Nutzer den Pools zu, die in der Datei `fairscheduler.xml` definiert sind. Nutzerpaare müssen durch Kommas getrennt werden (Beispiel: `tenant1:test,tenant2:production`). Bevor Sie einen Pool angeben, stellen Sie sicher, dass der Pool in der Datei `fairscheduler.xml` angegeben wurde. **MapReduce:** Ordnet Nutzer der Warteschlange zu, die in der Konfiguration des dynamische Ressourcenpools definiert wurde. Nutzerpaare müssen durch Kommas getrennt werden (Beispiel: tenant1:test,tenant2:production). Bevor Sie eine Warteschlange angeben, stellen Sie sicher, dass das System mit der Warteschlange konfiguriert wurde und dass der Zugriff zum Übergeben von Jobs an die Warteschlange zulässig ist. **Anmerkung:** Wenn Sie die Spark- und MapReduce-Jobs zusammen ausführen wollen, muss der Name der Nutzerzuordnungswerte in der Datei `fairscheduler.xml` und in der Konfiguration des dynamischen Ressourcenpools identisch sein.

Tabelle 13. Einstellungen in 'analyticserver-conf.config.properties' (Forts.)

Eigenschaft	Beschreibung
`resource.pool.default`	**Spark:** Definiert den Standardressourcenpool. Der Wert kann `default` oder ein Poolname sein, der in der Datei `fairscheduler.xml` definiert wurde. Verwenden Sie die Einstellung `default`, wenn Nutzer nicht (oder falsch) konfiguriert sind. **MapReduce:** Definiert die Standardwarteschlange, an die Jobs übergeben werden.
`spark.scheduler.mode=FAIR`	**Spark:** Aktiviert den Scheduler für den akzeptablen Modus. Die Eigenschaft sollte nicht geändert werden.
`spark.yarn.queue`	**Spark:** Der Name der YARN-Warteschlange, an die die Anwendung übergeben wird. In der Konfiguration des dynamischen Ressourcenpools können Sie einen angepassten YARN-Warteschlangennamen angeben.

4. Speichern Sie die Konfiguration und starten Sie den Analytic Server-Service erneut.

fairscheduler.xml - Beispiel

Die Datei `fairscheduler.xml` enthält die Eigenschaftsinformationen für den Spark-Pool. Die Datei ist erforderlich und muss manuell erstellt werden.

```
<?xml version="1.0"?>
<allocations>
  <pool name="production">
    <schedulingMode>FAIR</schedulingMode>
    <weight>1</weight>
    <minShare>2</minShare>
  </pool>
  <pool name="test">
    <schedulingMode>FIFO</schedulingMode>
    <weight>2</weight>
    <minShare>3</minShare>
  </pool>
</allocations>
```

Referenz

Weitere Informationen finden Sie auf den folgenden Sites:

- https://spark.apache.org/docs/latest/job-scheduling.html#dynamic-resource-allocation
- https://spark.apache.org/docs/latest/running-on-yarn.html

Migration

Analytic Server ermöglicht Ihnen das Migrieren von Daten und Konfigurationseinstellungen aus einer vorhandene Analytic Server-Installation in eine neue Installation.

Upgrade auf eine neue Version von Analytic Server

Wenn Sie über eine vorhandene Installation von Analytic Server 3.1.2 verfügen und eine neuere Version erworben haben, können Sie Ihre Konfigurationseinstellungen von Version 3.1.2 zu Ihrer neuen Installation migrieren.

Einschränkung: Installationen von 3.1.2 und neueren Versionen können nicht in demselben Hadoop-Cluster koexistieren. Wenn Sie Ihre neue Installation für die Verwendung desselben Hadoop-Clusters wie die Installation von 3.1.2 konfigurieren, funktioniert die Installation von 3.1.2 nicht mehr.

Migrationsschritte, 3.1.2 auf neuere Version

1. Führen Sie die Neuinstallation von Analytic Server entsprechend den Anweisungen in „Installation in Cloudera" auf Seite 42 durch.
2. Kopieren Sie den Analytic Server-Arbeitsbereich von Ihrer alten Installation in die neue Installation.

a. Wenn Sie sich nicht sicher sind, wo sich der Analytic Server-Arbeitsbereich befindet, führen Sie den Befehl `hadoop -fs ls` aus. Der Pfad zum Analytic Server-Arbeitsbereich hat das Format `/user/as_user/analytic-root/analytic-workspace`, wobei `as_user` die Benutzer-ID ist, die Eigner des Analytic Server-Arbeitsbereichs ist.

b. Melden Sie sich als `as_user` am Host der neuen Analytic Server-Installation an. Löschen Sie das Verzeichnis `/user/as_user/analytic-root/analytic-workspace`, falls es vorhanden ist.

c. Führen Sie das folgende Kopierscript aus:

```
hadoop distcp hftp://{Host des 3.1.2-Namensknotens}:50070/{Pfad zum Analytic Server-Arbeitsbereich Version 3.1.2}
hdfs://{Host des 3.2.1-Namensknotens}/user/as_user/analytic-root/analytic-workspace
```

3. Wenn Sie die eingebettete Apache Directory Server-Instanz verwenden, sichern Sie die aktuelle Benutzer-/Gruppenkonfiguration mit einem LDAP-Client-Tool eines anderen Anbieters. Importieren Sie die gesicherte Benutzer-/Gruppenkonfiguration nach der Installation von Analytic Server 3.2.1 in Apache Directory Server.

   Anmerkung: Dieser Schritt kann übersprungen werden, wenn Sie einen externen LDAP-Server verwenden.

4. Stoppen Sie den Analytic Server-Service in Cloudera Manager.
5. Erfassen Sie die Konfigurationseinstellungen der alten Installation.
   a. Kopieren Sie das Archiv `configcollector.zip` in Ihrer neuen Installation in `{AS-Stammverzeichnis}\tools` in Ihrer alten Installation.
   b. Extrahieren Sie die Kopie von `configcollector.zip`. Hierdurch wird ein neues Unterverzeichnis `configcollector` in Ihrer alten Installation erstellt.
   c. Führen Sie das Konfigurations-Collector-Tool in Ihrer alten Installation aus, indem Sie das Script **`configcollector`** im Verzeichnis `{AS-Stammverzeichnis}\tools\configcollector` aufrufen. Kopieren Sie die resultierende komprimierte Datei (ZIP-Datei) auf den Server, der Ihre neue Installation hostet.

   Wichtig: Das bereitgestellte Script **`configcollector`** ist möglicherweise nicht mit der aktuellen Version von Analytic Server kompatibel. Wenden Sie sich an einen IBM Technical Support-Mitarbeiter, wenn Probleme mit dem Script **`configcollector`** auftreten.
6. Löschen Sie den Zookeeper-Status. Führen Sie den folgenden Befehl im ZooKeeper-bin-Verzeichnis aus (z. B. `/opt/cloudera/parcels/CDH-5.4...../lib/zookeeper/bin` in Cloudera).

   ```
   ./zkCli.sh rmr /AnalyticServer
   ```

7. Führen Sie das Script **`migrationtool`** für das Migrationstool aus und übergeben Sie den Pfad der vom Konfigurationscollector erstellten komprimierten Datei als Argument. Es folgt ein Beispiel.

   ```
   migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.1.2.xxx.zip
   ```

8. Führen Sie den folgenden Befehl in einer Befehlsshell auf dem Analytic Server-Knoten aus:

   ```
   hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
   ```

9. Starten Sie den Analytic Server-Service in Cloudera Manager.

   Anmerkung: Wenn Sie R für die Verwendung mit der vorhandenen Analytic Server-Installation konfiguriert haben, müssen Sie die Schritte zum Konfigurieren von R mit der neuen Analytic Server-Installation befolgen.

Deinstallation von Analytic Server in Cloudera

Cloudera verarbeitet die meisten Schritte, die zum Deinstallieren des Service und der PARCEL-Datei von Analytic Server erforderlich sind, automatisch.

Die folgenden Schritte sind zum Löschen von Analytic Server aus der Cloudera-Umgebung erforderlich:

1. Stoppen Sie den Analytic Server-Service und löschen Sie ihn.
2. Verwenden Sie die Optionen **Deactivate**, **Remove From Hosts** und **Delete** für die PARCEL-Dateien von Analytic Server.

3. Löschen Sie das Analytic Server-Benutzerverzeichnis in HDFS. Die Standardposition ist `/user/as_user/analytic-root`.
4. Löschen Sie die Datenbank oder das Schema, die bzw. das von Analytic Server verwendet wird.
5. Bereinigen Sie alle Überreste des Analytic Server-Installationspakets. Dazu müssen Sie Folgendes löschen:
   - Ordner `csd`
   - Alle vorhandenen Dateien von Version 3.2.1, die sich in den Ordnern `parcels`, `parcel-cache` und `parcel-repo` befinden.

Kapitel 4. Konfigurieren von IBM SPSS Modeler für die Verwendung mit IBM SPSS Analytic Server

Sie müssen eine Reihe von Aktualisierungen an der SPSS Modeler Server-Installation vornehmen, um SPSS Modeler für die Verwendung mit Analytic Server zu aktivieren.

1. Konfigurieren Sie SPSS Modeler Server so, dass er einer Analytic Server-Installation zugeordnet ist.
   a. Bearbeiten Sie die Datei `options.cfg` im Unterverzeichnis `config` des Hauptserverinstallationsverzeichnisses und fügen Sie die folgenden Zeilen hinzu bzw. bearbeiten Sie sie:

```
as_ssl_enabled, {Y|N}
as_host, "{AS-Server}"
as_port, Port
as_context_root, "{Kontextstammverzeichnis}"
as_tenant, "{Nutzer}"
as_prompt_for_password, {Y|N}
as_kerberos_auth_mode, {Y|N}
as_kerberos_krb5_conf, {Konfigurationspfad}
as_kerberos_krb5_spn, {AS-SPN}
```

   as_ssl_enabled
   Geben Sie "Y" an, wenn bei Analytic Server sichere Kommunikation konfiguriert ist; geben Sie andernfalls "N" an.

   as_host
   Die IP-Adresse bzw. der Hostname des Servers, der als Host für Analytic Server fungiert.

   Anmerkung: Wenn SSL für Analytic Server aktiviert ist, müssen Sie eine entsprechende IP-Adresse bzw. einen entsprechenden Hostdomänennamen angeben.

   as_port
   Der Port, an dem Analytic Server empfangsbereit ist (standardmäßig 8080).

   as_context_root
   Das Analytic Server-Kontextstammverzeichnis (dies ist standardmäßig "analyticserver").

   as_tenant
   Der Nutzer, zu dem die SPSS Modeler Server-Installation gehört (der Standardnutzer ist ibm).

   as_prompt_for_password
   Geben Sie "N" an, wenn SPSS Modeler Server mit demselben Authentifizierungssystem für Benutzer und Kennwörter konfiguriert wurde wie Analytic Server, beispielsweise bei Verwendung der Kerberos-Authentifizierung. Geben Sie andernfalls **Y** an.

   Bei der Ausführung von SPSS Modeler im Stapelmodus fügen Sie `-analytic_server_username {AS-Benutzername} -analytic_server_password {AS-Kennwort}` dem Befehl `clemb` als Argumente hinzu.

   as_kerberos_auth_mode
   Geben Sie "Y" an, um Kerberos-SSO über SPSS Modeler zu aktivieren.

   as_kerberos_krb5_conf
   Geben Sie den Pfad zur Kerberos-Konfigurationsdatei an, die Analytic Server verwenden soll, z. B. `\etc\krb5.conf`.

   as_kerberos_krb5_spn
   Geben Sie den Kerberos-SPN von Analytic Server an, z. B. HTTP/ashost.mydomain.com@MYDOMAIN.COM.

   b. Starten Sie den SPSS Modeler Server-Service erneut.

Zum Herstellen einer Verbindung zu einer Analytic Server-Installation, bei der SSL/TLS aktiviert ist, müssen einige weitere Schritte ausgeführt werden, um Ihre SPSS Modeler Server-Installation und Clientinstallationen zu konfigurieren.

a. Navigieren Sie zu `http{s}://{Host}:{Port}/{Kontextstammverzeichnis}/admin/{Nutzer}` und melden Sie sich an der Analytic Server-Konsole an.
b. Laden Sie die Zertifizierungsdatei aus dem Browser herunter und speichern Sie sie in Ihrem Dateisystem.
c. Fügen Sie die Zertifizierungsdatei der Java-Ausführungsumgebung (JRE) sowohl der SPSS Modeler Server-Installation als auch der SPSS Modeler-Clientinstallation hinzu. Den zu aktualisierenden Speicherort finden Sie im Unterverzeichnis `/jre/lib/security/cacerts` des SPSS Modeler-Installationspfads.
   1) Stellen Sie sicher, dass die Datei `cacerts` nicht schreibgeschützt ist.
   2) Verwenden Sie das mit Modeler gelieferte Programm **keytool**, das sich im Unterverzeichnis `/jre/bin/keytool` des SPSS Modeler-Installationspfads befindet.

      Führen Sie den folgenden Befehl aus:

      ```
      keytool -import -alias <AS-Alias> -file <Zertifikatsdatei> -keystore "<cacerts-Datei>"
      ```

      Beachten Sie, dass <AS-Alias> ein Alias für die Datei `cacerts` ist. Sie können einen beliebigen Namen verwenden, solange er für die Datei `cacerts` eindeutig ist.

      Ein Beispielbefehl könnte wie folgt aussehen:

      ```
      keytool -import -alias MySSLCertAlias -file C:\Download\as.cer
              -keystore "c:\Programme\IBM\SPSS\Modeler\{ModelerVersion}\jre\lib\security\cacerts"
      ```
d. Starten Sie SPSS Modeler Server und den SPSS Modeler-Client erneut.

2. [Optional] Installieren Sie IBM SPSS Modeler - Essentials for R, wenn Sie vorhaben, ein Scoring für R-Modelle in Datenströmen mit Analytic Server-Datenquellen durchzuführen. IBM SPSS Modeler - Essentials for R ist als Download verfügbar (https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp).

Kapitel 5. Konfigurieren des Pushbacks für die benutzerdefinierte Funktion von Hive

Auf sämtlichen Pushback-fähigen IBM SPSS Analytic Server-Knoten wird nach Möglichkeit ein Pushback für die benutzerdefinierte Funktion von Hive durchgeführt. Nach der Registrierung der benutzerdefinierten Funktion von Hive bei HiveDB kann Analytic Server mithilfe der neuen benutzerdefinierten Funktionen ein Pushback durchführen.

Die Pushback-Funktion ist in der benutzerdefinierten Funktion von Hive standardmäßig inaktiviert und muss manuell über die Einstellung **udfmodule** in der Datei ASModules.xml aktiviert werden (ändern Sie hierzu den Wert **disabled** in **enabled**). Nach der Aktivierung der Einstellung müssen Sie Analytic Server neu starten und die benutzerdefinierte Funktion manuell bei Hive registrieren.

Die folgenden Beispiele veranschaulichen die Registrierung bzw. die Aufhebung der Registrierung einer benutzerdefinierten Funktion bei Hive in HDP- und Cloudera-Umgebungen.

Registrierung/Aufhebung der Registrierung einer benutzerdefinierten Funktion unter HDP

Registrierung einer benutzerdefinierten Funktion

```
su - hive -c 'hive -f /opt/ibm/spss/analyticserver/3.2/bin/udfRegister.sql'
```

Aufhebung der Registrierung einer benutzerdefinierten Funktion

```
su - hive -c 'hive -f /opt/ibm/spss/analyticserver/3.2/bin/udfUnregister.sql'
```

Registrierung/Aufhebung der Registrierung einer benutzerdefinierten Funktion in Cloudera

Registrierung einer benutzerdefinierten Funktion

```
sudo -u hive kinit -k -t hive.keytab hive/bosperf5-master.fyre.ibm.com@IBM.COM
sudo -u hive hive -f /opt/cloudera/parcels/AnalyticServer/bin/udfRegister.sql
```

Aufhebung der Registrierung einer benutzerdefinierten Funktion

```
sudo -u hive hive -f /opt/cloudera/parcels/AnalyticServer/bin/udfUnregister.sql
```

Kapitel 6. Verwenden von SLM-Tags zum Überwachen der Lizenzierung

SLM-Tags basieren auf dem Entwurf des Standards ISO/IEC 19770-4 für die Messung der Ressourcennutzung. SLM-Tags bieten eine standardisierte Möglichkeit für ein Produkt, die Nutzung von Lizenzmetriken (Ressourcen, die sich auf die Nutzung eines Software-Assets beziehen) aufzulisten. Nach der Aktivierung von SLM in einem Produkt wird eine Laufzeit-XML-Datei generiert, über die das Programm seine Lizenznutzung ausgeben kann.

Beim Start von Analytic Server werden `slmtag`-Dateien im Ordner `<AS-Installationspfad>/logs/slmtag` erstellt.

Da es zwei Lizenztypen gibt, werden laufend zwei unterschiedliche Metriken aufgezeichnet:

- Für die aktuelle Version von Analytic Server basiert die Lizenzierung auf der Gesamtzahl von Datenknoten im Hadoop-Cluster (auf Virtual Server basierend). Die Anzahl der Knoten wird im folgenden Abschnitt der `slmtag`-Datei aufgezeichnet:

```
<Type>VIRTUAL_SERVER</Type>
   <SubType>Number of Data Nodes in Hadoop</SubType>
   <Value>2</Value>
   ...
```

- Bei Analytic Server-Versionen vor 3.1 basierte die Lizenzierung auf der Größe des HDFS-Speichers im Hadoop-Cluster (auf RVU basiererend). Die Speichergröße (in Tegabyte) wird beispielsweise im folgenden Abschnitt der `slmtag`-Datei aufgezeichnet.

```
<Type>RESOURCE_VALUE_UNIT</Type>
   <SubType>HDFS storage (Unit: Tega byte)</SubType>
   <Value>0.21</Value>
```

Die SLM-Tagausgabe wird in einem Thread gestartet und wird über die Eigenschaften gesteuert, die in der Datei `SlmTagOutput.properties` definiert werden. Die Datei befindet sich im Ordner `<AS-Installationspfad>/configuration`.

Tabelle 14. SLM-Tageigenschaften

Eigenschaften	Beschreibung
license.metric.logger.output.enabled	Steuert die Generierung der SLM-Protokolldatei. Der Standardwert ist `False`.
license.metric.logger.output.dir	Der relative Pfad zu dem Verzeichnis, in dem die SLM-Tagdateien gespeichert werden. Das Standardverzeichnis ist `<AS-Installationspfad>/logs`.
license.metric.logger.output.SLMLogFrequency	Das Zeitintervall (Einheit: Millisekunden) für die Erfassung von SLM-Protokollen.
icense.metric.logger.file.size	Die maximale Größe der SML-Tagdatei in Byte.
license.metric.logger.file.number	Die maximale Anzahl von SLM-Tagdateien für eine Instanz einer Softwareidentität.

Kapitel 7. Fehlerbehebung

In diesem Abschnitt werden einige allgemeine Installations- und Konfigurationsprobleme sowie Wege zu deren Lösung beschrieben.

Allgemeine Probleme

Installation wird zwar mit Warnungen, aber erfolgreich abgeschlossen, Benutzer können jedoch keine Datenquellen erstellen. Es wird der folgende Fehler angezeigt: "Die Anforderung kann nicht abgeschlossen werden. Ursache: Berechtigung verweigert"

Wenn der Parameter **distrib.fs.root** auf ein Verzeichnis gesetzt wird, auf das der Analytic Server-Benutzer (standardmäßig as_user) keinen Zugriff hat, kommt es zu Fehlern. Stellen Sie sicher, dass der Analytic Server-Benutzer Lese-, Schreib- und Ausführungsberechtigung für das Verzeichnis **distrib.fs.root** hat.

Die Analytic Server-Leistung wird zunehmend schlechter.

Wenn die Analytic Server-Leistung die Erwartungen nicht erfüllt, entfernen Sie alle `*.war`-Dateien aus dem Knox-Servicebereitstellungspfad: `/<Knox-Servicepfad>/data/deployments`. Beispiel: `/usr/hdp/3.1.0.0-78/knox/data/deployments`.

Deinstallieren von Analytic Server oder Essentials for R unter Ambari

In einigen Fällen wird der Deinstallationsprozess beim Deinstallieren von Analytic Server oder Essentials for R unter Ambari blockiert. Wenn das Problem auftritt, müssen Sie die Ambari-Server-Prozess-ID manuell stoppen.

Probleme bei Installation von Analytic Server auf POWER System-Einheit, die OpenJDK verwendet

Wenn Analytic Server auf einer POWER System-Einheit verwendet wird, die OpenJDK verwendet, müssen Sie die folgenden Konfigurationsschritte manuell ausführen, um sicherzustellen, dass die Koordinatensystem-API wie erwartet funktioniert.

Anmerkung: Sie können die Konfigurationsanforderung ignorieren, wenn Sie die Koordinatensystem-API nicht verwenden.

1. Navigieren Sie in der Ambari-Konsole zu **Analytic Server service** > **Configs tab** > **Advanced analytics-jvm-options** und fügen Sie folgende Zeile zum Inhaltsbereich hinzu:

   ```
   -XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*
   ```

2. Navigieren Sie in der Ambari-Konsole zum Bereich **Custom analytics.cfg** und fügen Sie die folgenden 3 Konfigurationen hinzu:

   spark.executor.extraJavaOptions
   Setzen Sie den Wert auf `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`.

   spark.driver.extraJavaOptions
   Setzen Sie den Wert auf `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`.

   mapred.child.java.opts
   Setzen Sie den Wert auf `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`.

Fehler beim Installieren von Analytic Server unter SuSE Linux 12

Möglicherweise tritt der folgende Fehler auf, wenn Sie Analytic Server unter SuSE Linux 12 installieren:

```
Signature verification failed [4-Signatures public key is not available]
```

Das Problem kann gelöst werden, indem Sie die folgenden Aufgaben durchführen, bevor Sie Analytic Server unter SuSE Linux 12 installieren:

1. Laden Sie einen öffentlichen Schlüssel von der folgenden URL auf Ihren Host herunter:

   ```
   https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.1.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
   ```

2. Importieren Sie den öffentlichen Schlüssel, indem Sie den folgenden Befehl auf Ihrem Host ausführen:

   ```
   rpm --import IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
   ```

Probleme mit bestimmten Hadoop-Verteilungen

Aktualisierungsaktion für Analytic Server-Service ist unter Hortonworks 2.3 - 2.6 inaktiviert

Führen Sie die folgenden Schritte aus, um Analytic Server-Bibliotheken unter Hortonworks 2.3 - 2.6 manuell zu aktualisieren.

1. Melden Sie sich an dem Host, der den Analytic-Metaspeicher ausführt, als Analytic Server-Benutzer (standardmäßig as_user) an.

   Anmerkung: Sie können diesen Hostnamen über die Ambari-Konsole ermitteln.

2. Führen Sie das Script **refresh** im Verzeichnis {AS-Stammverzeichnis}/bin aus; Beispiel:

   ```
   cd /opt/ibm/spss/analyticserver/3.2/bin
   ./refresh
   ```

3. Starten Sie den Analytic Server-Service in der Ambari-Konsole erneut.

Von einer externen Site heruntergeladene Pakete lassen die Hashprüfung in Cloudera Manager fehlschlagen

Der Hashverifizierungsfehler wird in der Paketliste angezeigt. Das Problem kann behoben werden, indem Sie warten, bis der Downloadprozess abgeschlossen ist, und dann Cloudera über den Service cloudera-scm-server erneut starten. Der Fehler tritt nach dem Serviceneustart nicht auf.

HDFS-Supergruppeneigenschaften

Analytic Server protokolliert eine Ausnahmebedingung während des Starts, wenn as_user kein Member der folgenden HDFS-Gruppeneigenschaften ist: **dfs.permissions.supergroup/ dfs.permissions.superusergroup**. Beispiel:

```
[11/15/17 7:32:35:510 PST] 000000bf SystemOut
O 2017-11-15 07:32:35,510 | : |  |  |  |  | ERROR | slmtagoutput.SlmOuputAgent | SLM Logger => Error in performing callback function when calculating number
of nodes in kerberos environment: org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.AccessControlException): Access denied for user as_user.
Superuser privilege is required
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkSuperuserPrivilege(FSPermissionChecker.java:93)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkSuperuserPrivilege(FSNamesystem.java:6606)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.datanodeReport(FSNamesystem.java:5595)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getDatanodeReport(NameNodeRpcServer.java:928)
at org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.getDatanodeReport(AuthorizationProviderProxyClientProtocol.java:390)
at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getDatanodeReport(ClientNamenodeProtocolServerSideTranslatorPB.java:694)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:617)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1073)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2141)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2137)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:415)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1912)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2135)
```

Sie müssen as_user manuell der BS-Gruppe hinzufügen, die in den Konfigurationseigenschaften von hdfs-site definiert wurde: **dfs.permissions.supergroup/dfs.permissions.superusergroup**.

- Für Cloudera ist der Standardeigenschaftswert **supergroup** und muss in eine BS-Gruppe geändert werden, die tatsächlich vorhanden ist. Information zu den Supergruppeneinstellungen in Cloudera finden Sie in der Cloudera-Dokumentation.
- Für Ambari ist der Standardeigenschaftswert **hdfs**. Während einer Ambari-Installation fügt Analytic Server standardmäßig as_user zu den HDFS- und Hadoop-Gruppen hinzu.

Verwenden Sie unter Linux den Befehl **usermod**, um as_user zur HDFS-Gruppe **superusergroup** hinzuzufügen (wenn sie nicht bereits vorhanden ist).

Allgemeine Informationen zu HDFS-Berechtigungen finden Sie in HDFS Permissions Guide.

MapReduce-Jobs schlagen unter HDP 3.0 fehl

Bei MapReduce-Jobs unter HDP 3.0 tritt unter Umständen der folgende Fehler auf:

```
Unable to complete the request. Reason: java.lang.IllegalStateException: Job in state DEFINE instead of RUNNING (as_trace.log)
```

Der Fehlerstatus kann wie folgt aufgelöst werden:

1. Fügen Sie folgende Konfiguration zur Datei `Custom analytics.cfg` hinzu:
   ```
   exclude.mapreduce.jars=icu4j-
   ```
2. Starten Sie Analytic Server erneut.

Nach dem Neustart von Analytic Server werden die MapReduce-Jobs wie gewohnt ausgeführt.

Probleme mit dem Metadatenrepository

Operation CREATE USER schlägt bei Ausführung des Scripts add_mysql_user fehl

Bevor Sie das Script **`add_mysql_user`** ausführen, müssen Sie zuerst den Benutzer manuell entfernen, den Sie aus der MySQL-Datenbank hinzufügen wollen. Sie können die Benutzer über die MySQL Workbench-Benutzerschnittstelle oder über MySQL-Befehle entfernen. Beispiel:

```
mysql -u root -e "DROP USER '$AEDB_USERNAME_VALUE'@'localhost';"
mysql -u root -e "DROP USER '$AEDB_USERNAME_VALUE'@'$METASTORE_HOST';"
mysql -u root -e "DROP USER '$AEDB_USERNAME_VALUE'@'%';"
```

Ersetzen Sie in den oben genannten Befehlen $AEDB_USERNAME_VALUE durch den Benutzernamen, der entfernt werden soll, und $METASTORE_HOST durch den Namen des Hosts, auf dem die Datenbank installiert ist.

Probleme mit Apache Spark

Probleme mit Datenströmen, die in einem Spark-Prozess ausgeführt werden

SPSS Modeler-Datenströme können nicht abgeschlossen werden, wenn sie zur Ausführung in einem Spark-Prozess gezwungen werden. Die fehlschlagenden SPSS Modeler-Datenströme werden mit einem Analytic Server-Quellenknoten (HDFS-Datei) erstellt, der mit einem `Sortierknoten` verknüpft ist und so eingerichtet ist, dass Daten in eine andere Analytic Server-Datenquelle exportiert werden. Nach der Datenstromausführung gibt die Benutzerschnittstelle des Ressourcenmanagers an, dass die neue Anwendung ausgeführt wird, der Datenstrom wurde jedoch nie abgeschlossen und verbleibt im Ausführungsstatus. Es gibt in den Analytic Server-Protokollen, YARN-Protokollen oder Spark-Protokollen keine Nachrichten, die angeben, warum der Datenstrom nicht abgeschlossen wird.

Das Problem kann behoben werden, indem der angepassten Datei `analytics.cfg` in der Analytic Server-Konfiguration die Einstellung `spark.executor.memory` hinzugefügt wird. Wenn Sie den Speicherwert auf 4 GB festlegen, können die zuvor fehlgeschlagenen SPSS Modeler-Datenströme (in einer Umgebung mit einem einzelnen Knotencluster) in weniger als 2 Minuten abgeschlossen werden.

Fehler beim Ausführen von Spark-Jobs mit Cloudera 5.x und Spark 1.x

Möglicherweise tritt die folgende Ausnahmebedingung auf, wenn Sie Cloudera 5.x mit Spark 1.x verwenden:

```
org.apache.spark.SparkException: Exception when registering SparkListener
```

Die Ausnahmebedingung wird verursacht, weil `org.apache.spark.scheduler.SparkListener` nicht von `java.lang.ClassCastException: com.cloudera.spark.lineage.ClouderaNavigatorListener` umgesetzt werden kann.

Sie müssen die folgende Zeile im Bereich **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties** hinzufügen, um diese Ausnahmebedingung zu vermeiden.

```
spark.extraListeners=org.apache.spark.JavaSparkListener
```

Beim Ausführen von SparkML-Fällen tritt der Fehler "Exception during HdfsAuthcom.spss.utilities.i18n.LocException:Execution failed. Reason: com.spss.ae.filesystem.exception.FileSystemException: Unable to initialize the file system access." auf.

Der Fehler wird generiert, wenn Spark das Verzeichnis für das Herkunftsprotokoll nicht finden kann. Eine Problemumgehung ist das Umleiten von `spark.lineage.log.dir` nach `/ae_wlpserver/usr/servers/aeserver/logs/spark`.

Hochverfügbarkeitscluster

Analytic Server kann aufgrund von Änderungen der Abhängigkeiten keinen weiteren Hosts hinzugefügt werden

Führen Sie das Script `update_clientdeps` unter Beachtung der Anweisungen in „Aktualisierung von Clientabhängigkeiten” auf Seite 29 aus.

"Analytic Cluster Service hat unerwarteterweise Kontakt mit Zookeeper verloren, diese JVM wird beendet, um die Clusterintegrität zu bewahren."

Eine mögliche Ursache dafür kann sein, dass ein zu großes Datenvolumen in Zookeeper geschrieben wird. Falls die Zookeeper-Protokolle Ausnahmebedingungen wie die folgende enthalten:

```
java.io.IOException: Unreasonable length = 2054758
```

oder die Analytic Server-Protokolle Nachrichten wie die folgende enthalten:

```
Caused by: java.io.UTFDataFormatException: encoded string too long: 2054758 bytes
  at java.io.DataOutputStream.writeUTF(DataOutputStream.java:375)
```

1. Navigieren Sie in der Ambari-Konsole zur Registerkarte **Configs** für den Zookeeper-Service und fügen Sie env-template die folgende Zeile hinzu. Starten Sie den Zookeeper-Service anschließend erneut.

   ```
   export JVMFLAGS="-Xmx2048m -Djute.maxbuffer=2097152"
   ```

2. Navigieren Sie in der Ambari-Konsole zur Registerkarte "Configs" des Analytic Server-Service, fügen Sie unter "Advanced analytics-jvm-options" Folgendes hinzu und starten Sie anschließend den Analytic Cluster-Service erneut.

   ```
   -Djute.maxbuffer=2097152
   ```

Die Zahl, die für die Einstellung jute.maxbuffer angegeben wird, sollte größer als die in den Ausnahmebedingungsnachrichten angegebene Zahl sein.

Zookeeper-Transaktionsdaten können nicht mehr verwaltet werden

Setzen Sie den Parameter **`autopurge.purgeInterval`** in `zoo.cfg` auf 1, um das automatische Bereinigen des Zookeeper-Transaktionsprotokolls zu aktivieren.

Analysecluster-Service verliert Zookeeper-Kontakt

Prüfen und ändern Sie die Parameter **`tickTime`**, **`initLimit`** und **`syncLimit`** in `zoo.cfg`. Beispiel:

```
# The number of milliseconds of each tick
tickTime=2000
# The number of ticks that the initial
# synchronization phase can take
initLimit=30
# The number of ticks that can pass between
# sending a request and getting an acknowledgement
syncLimit=15
```

Details finden Sie in der Dokumentation zu Zookeeper unter https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html.

Analytic Server-Jobs werden nicht fortgesetzt

Es gibt eine allgemeine Situation, in der Analytic Server-Jobs nicht fortgesetzt werden.

- Wenn ein Analytic Server-Job fehlschlägt, da ein Cluster-Member fehlschlägt, wird der Job normalerweise auf einem anderen Cluster-Member fortgesetzt. Wenn der Job nicht fortgesetzt wird, stellen Sie sicher, dass der Hochverfügbarkeitscluster mindestens 4 Cluster-Member umfasst.

Bemerkungen

Die vorliegenden Informationen wurden für Produkte und Services entwickelt, die auf dem deutschen Markt angeboten werden. IBM stellt dieses Material möglicherweise auch in anderen Sprachen zur Verfügung. Für den Zugriff auf das Material in einer anderen Sprache kann eine Kopie des Produkts oder der Produktversion in der jeweiligen Sprache erforderlich sein.

Möglicherweise bietet IBM die in dieser Dokumentation beschriebenen Produkte, Services oder Funktionen in anderen Ländern nicht an. Informationen über die gegenwärtig im jeweiligen Land verfügbaren Produkte und Services sind beim zuständigen IBM Ansprechpartner erhältlich. Hinweise auf IBM Lizenzprogramme oder andere IBM Produkte bedeuten nicht, dass nur Programme, Produkte oder Services von IBM verwendet werden können. Anstelle der IBM Produkte, Programme oder Services können auch andere, ihnen äquivalente Produkte, Programme oder Services verwendet werden, solange diese keine gewerblichen oder anderen Schutzrechte von IBM verletzen. Die Verantwortung für den Betrieb von Produkten, Programmen und Services anderer Anbieter liegt beim Kunden.

Für in diesem Handbuch beschriebene Erzeugnisse und Verfahren kann es IBM Patente oder Patentanmeldungen geben. Mit der Auslieferung dieses Handbuchs ist keine Lizenzierung dieser Patente verbunden. Lizenzanforderungen sind schriftlich an folgende Adresse zu richten (Anfragen an diese Adresse müssen auf Englisch formuliert werden):

IBM Director of Licensing
IBM Corporation
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France

Diese Informationen können technische Ungenauigkeiten oder typografische Fehler enthalten. Die hier enthaltenen Informationen werden in regelmäßigen Zeitabständen aktualisiert und als Neuausgabe veröffentlicht. IBM kann ohne weitere Mitteilung jederzeit Verbesserungen und/oder Änderungen an den in dieser Veröffentlichung beschriebenen Produkten und/oder Programmen vornehmen.

Verweise in diesen Informationen auf Websites anderer Anbieter werden lediglich als Service für den Kunden bereitgestellt und stellen keinerlei Billigung des Inhalts dieser Websites dar. Das über diese Websites verfügbare Material ist nicht Bestandteil des Materials für dieses IBM Produkt. Die Verwendung dieser Websites geschieht auf eigene Verantwortung.

Werden an IBM Informationen eingesandt, können diese beliebig verwendet werden, ohne dass eine Verpflichtung gegenüber dem Einsender entsteht.

Lizenznehmer des Programms, die Informationen zu diesem Produkt wünschen mit der Zielsetzung: (i) den Austausch von Informationen zwischen unabhängig voneinander erstellten Programmen und anderen Programmen (einschließlich des vorliegenden Programms) sowie (ii) die gemeinsame Nutzung der ausgetauschten Informationen zu ermöglichen, wenden sich an folgende Adresse:

IBM Director of Licensing
IBM Corporation
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
USA

Die Bereitstellung dieser Informationen kann unter Umständen von bestimmten Bedingungen - in einigen Fällen auch von der Zahlung einer Gebühr - abhängig sein.

Die Lieferung des in diesem Dokument beschriebenen Lizenzprogramms sowie des zugehörigen Lizenzmaterials erfolgt auf der Basis der IBM Rahmenvereinbarung bzw. der Allgemeinen Geschäftsbedingungen von IBM, der IBM Internationalen Nutzungsbedingungen für Programmpakete oder einer äquivalenten Vereinbarung.

Die angeführten Leistungsdaten und Kundenbeispiele dienen nur zur Illustration. Die tatsächlichen Ergebnisse beim Leistungsverhalten sind abhängig von der jeweiligen Konfiguration und den Betriebsbedingungen.

Alle Informationen zu Produkten anderer Anbieter stammen von den Anbietern der aufgeführten Produkte, deren veröffentlichten Ankündigungen oder anderen allgemein verfügbaren Quellen. IBM hat diese Produkte nicht getestet und kann daher keine Aussagen zu Leistung, Kompatibilität oder anderen Merkmalen machen. Fragen zu den Leistungsmerkmalen von Produkten anderer Anbieter sind an den jeweiligen Anbieter zu richten.

Aussagen über Pläne und Absichten von IBM unterliegen Änderungen oder können zurückgenommen werden und repräsentieren nur die Ziele von IBM.

Alle von IBM angegebenen Preise sind empfohlene Richtpreise und können jederzeit ohne weitere Mitteilung geändert werden. Händlerpreise können u. U. von den hier genannten Preisen abweichen.

Diese Veröffentlichung dient nur zu Planungszwecken. Die in dieser Veröffentlichung enthaltenen Informationen können geändert werden, bevor die beschriebenen Produkte verfügbar sind.

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

COPYRIGHTLIZENZ:

Diese Veröffentlichung enthält Beispiele für Daten und Berichte des alltäglichen Geschäftsablaufs. Sie sollen nur die Funktionen des Lizenzprogramms illustrieren und können Namen von Personen, Firmen, Marken oder Produkten enthalten. Alle diese Namen sind frei erfunden; Ähnlichkeiten mit tatsächlichen Namen und Adressen sind rein zufällig.

Kopien oder Teile der Beispielprogramme bzw. daraus abgeleiteter Code müssen folgenden Copyrightvermerk beinhalten:

Marken

IBM, das IBM Logo und ibm.com sind Marken oder eingetragene Marken der IBM Corporation in den USA und/oder anderen Ländern. Weitere Produkt- und Servicenamen können Marken von IBM oder anderen Unternehmen sein. Eine aktuelle Liste der IBM Marken finden Sie auf der Webseite "Copyright and trademark information" unter www.ibm.com/legal/copytrade.shtml.

Adobe, das Adobe-Logo, PostScript und das PostScript-Logo sind Marken oder eingetragene Marken der Adobe Systems Incorporated in den USA und/oder anderen Ländern.

IT Infrastructure Library ist eine eingetragene Marke der Central Computer and Telecommunications Agency. Die Central Computer and Telecommunications Agency ist nunmehr in das Office of Government Commerce eingegliedert worden.

Intel, das Intel-Logo, Intel Inside, das Intel Inside-Logo, Intel Centrino, das Intel Centrino-Logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium und Pentium sind Marken oder eingetragene Marken der Intel Corporation oder ihrer Tochtergesellschaften in den USA oder anderen Ländern.

Linux ist eine eingetragene Marke von Linus Torvalds in den USA und/oder anderen Ländern.

Microsoft, Windows, Windows NT und das Windows-Logo sind Marken der Microsoft Corporation in den USA und/oder anderen Ländern.

ITIL ist eine eingetragene Marke, eine eingetragene Gemeinschaftsmarke des Cabinet Office (The Minister for the Cabinet Office) und eine eingetragene Marke, die beim U.S. Patent and Trademark Office eingetragen ist.

UNIX ist eine eingetragene Marke von The Open Group in den USA und anderen Ländern.

Cell Broadband Engine wird unter Lizenz verwendet und ist eine Marke der Sony Computer Entertainment, Inc. in den USA und/oder anderen Ländern.

Linear Tape-Open, LTO, das LTO-Logo, Ultrium und das Ultrium-Logo sind Marken von HP, der IBM Corporation und von Quantum in den USA und/oder anderen Ländern.

Gedruckt in Deutschland