IBM SPSS Analytic Server
Version 3.1.0

Installation und Konfiguration

Hinweis

Vor Verwendung dieser Informationen und des darin beschriebenen Produkts sollten die Informationen unter „Bemerkungen" auf Seite 71 gelesen werden.

Produktinformation

Diese Ausgabe bezieht sich auf Version 3, Release 1, Modifikation 0 von IBM SPSS Analytic Server und alle nachfolgenden Releases und Modifikationen, bis dieser Hinweis in einer Neuausgabe geändert wird.

Diese Veröffentlichung ist eine Übersetzung des Handbuchs
IBM SPSS Analytic Server, Version 3.1.0, Installation and Configuration Guide,
herausgegeben von International Business Machines Corporation, USA

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

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

Änderung des Textes bleibt vorbehalten.

Herausgegeben von:
TSC Germany
Kst. 2877
Mai 2017

Inhaltsverzeichnis

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**.

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

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.

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.

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 HDFS, YARN, MapReduce2, Hive und Zookeeper als Ambari-Services hinzugefügt wurden.

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 IBM SPSS Analytic Server-Installation

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 nach der Analytic Server-Installation ausgeführt werden.

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, cmx)

  Anmerkung: cmx wird nur in BigInsights unterstützt.
- 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

Die Precheck- und Postcheck-Tools sind in den folgenden Verzeichnissen angeordnet:

- **BigInsights**

  `/var/lib/ambari-server/resources/stacks/BigInsights/4.X/services/ANALYTICSERVER/package/chktool`
- **HDP**

  `/var/lib/ambari-server/resources/stacks/HDP/2.X/services/ANALYTICSERVER/package/chktool`
- **Cloudera**

  Die Tools sind in der Datei `AnalyticServer-*.jar` gepackt (diese Datei befindet sich in: `/opt/cloudera/csd`).

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

Vor der Installation von Analytic Server sollte das Precheck-Tool auf allen Ambari-Knoten ausgeführt werden, die den Analytic Server-Service hosten. Soll das Tool auf einem anderen Knoten ausgeführt werden, muss das gesamte Verzeichnis `chktool` auf den Knoten kopiert werden.

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 5) verfügbar. Wenn Sie die Ausführung einer „Offlineinstallation" auf Seite 7 auswählen, ist das Verzeichnis `chktool` nach der Installation der Metadaten-RPM verfügbar.

Ausführen des Precheck-Tools

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 Ambari –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` schließt Verwendungshilfe ein, 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` schließt Verwendungshilfe ein, 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. Es folgen detailliertere Schritte.

Wichtig: Analytic Server unterstützt nicht die Installation in einer Umgebung, in der der Ambari-Server als Benutzer ohne Rootberechtigung ausgeführt wird.

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. Die verfügbaren Ambari-Binärdateien sind:

Tabelle 1. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM® SPSS Analytic Server 3.1 für BigInsights 4.1 und 4.3 Linux on System p LE (Englisch)	`spss_as-3.1-bi4.1-4.3-lppc64le_en.bin`
IBM SPSS Analytic Server 3.1 für BigInsights 4.1, 4.2 und 4.3 Linux x86-64 (Englisch)	`spss_as-3.1-bi4.1-4.2-4.3-lx86_en.bin`
IBM SPSS Analytic Server 3.1 für Hortonworks Data Platform 2.3 und 2.4 Ubuntu (Englisch)	`spss_as-3.1-hdp2.3-2.4-ubun_en.bin`
IBM SPSS Analytic Server 3.1 für Hortonworks Data Platform 2.4, 2.5 und 2.6 Linux x86-64 (Englisch)	`spss_as-3.1-hdp2.4-2.6-lx86_en.bin`
IBM SPSS Analytic Server 3.1 für Hortonworks Data Platform 2.6 Linux on System p LE (Englisch)	`spss_as-3.1-hdp2.6-lppc64le_en.bin`

2. Führen Sie die selbstextrahierende 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.

   [Nur GPFS (Spectrum Scale)] Laden Sie die Datei https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms/IBM-SPSS-AnalyticServer-3.1.0.0.repo (x86 und ppc64le) oder https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/Ubuntu/IBM-SPSS-AnalyticServer.list (Ubuntu) herunter und verschieben Sie sie auf jedem

Knoten, auf dem Sie Analytic Server Metastore als Service hinzufügen, in den Ordner `/etc/yum.repos.d` (RHEL, CentOS), `/etc/zypp/repos.d` (SLES) oder `/etc/apt/sources.list.d` (Ubuntu).

Offlineinstallation
Wählen Sie die Offlineinstallation aus, wenn Ihr Ambari-Server-Host keinen Internetzugriff hat. Details finden Sie in „Offlineinstallation" auf Seite 7.

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. Wenn Sie eine Offlineinstallation durchführen, prüfen Sie, ob die Änderungen in den Dateien `repoinfo.xml` und `IBM-SPSS-AnalyticServer-3.1.0.0.repo` noch vorhanden sind. Wenden Sie sich an Ihren IBM Support-Mitarbeiter, wenn die Änderungen in `repoinfo.xml` und `IBM-SPSS-AnalyticServer-3.1.0.0.repo` nicht mehr vorhanden sind.
6. 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.

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**.

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

7. 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 11. Informationen zum Migrieren einer vorhandenen Konfiguration auf eine neue Installation finden Sie in „Upgrade und Migration" auf Seite 25.
8. 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. Der Standardwert ist **9080**. 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

Der allgemeine Workflow für eine Offlineinstallation sieht wie folgt aus:

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. Die verfügbaren Ambari-Binärdateien sind:

Tabelle 2. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.1 für BigInsights 4.1 und 4.3 Linux on System p LE (Englisch)	`spss_as-3.1-bi4.1-4.3-lppc64le_en.bin`
IBM SPSS Analytic Server 3.1 für BigInsights 4.1, 4.2 und 4.3 Linux x86-64 (Englisch)	`spss_as-3.1-bi4.1-4.2-4.3-lx86_en.bin`
IBM SPSS Analytic Server 3.1 für Hortonworks Data Platform 2.3 und 2.4 Ubuntu (Englisch)	`spss_as-3.1-hdp2.3-2.4-ubun_en.bin`
IBM SPSS Analytic Server 3.1 für Hortonworks Data Platform 2.4, 2.5 und 2.6 Linux x86-64 (Englisch)	`spss_as-3.1-hdp2.4-2.6-lx86_en.bin`
IBM SPSS Analytic Server 3.1 für Hortonworks Data Platform 2.6 Linux on System p LE (Englisch)	`spss_as-3.1-hdp2.6-lppc64le_en.bin`

2. Führen Sie die ausführbare Binärdatei aus und geben Sie eine Offlineinstallation an. Eine Offlineinstallation lädt die erforderlichen RPM- oder DEB-Dateien herunter und sollte auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann. Die ausführbare Binärdatei befindet sich in den verfügbaren Ambari-Verteilerverzeichnissen des `<installierbaren_AS-Ausgangsverzeichnisses>` .
3. Kopieren Sie den gesamten Inhalt des entsprechenden `<installierbaren_AS-Ausgangsverzeichnisses>` von dem Computer mit Internetzugriff auf den Ambari-Managerknoten (hinter der Firewall).
4. Setzen Sie die Anweisungsschritte für Ihre jeweilige Verteilung (BigInsights/HDP oder Ubuntu) auf den entsprechenden Ambari-Managerknoten (hinter der Firewall) fort.

BigInsights- und HDP-Anweisungen - 3.1.0

Wichtig: Analytic Server unterstützt nicht die Installation in einer Umgebung, in der der Ambari-Server als Benutzer ohne Rootberechtigung ausgeführt wird.

1. Prüfen Sie mit dem folgenden Befehl, ob der Ambari-Server zurzeit aktiv ist:

   ```
   ambari-server status
   ```

   Fahren Sie den Ambari-Serverknoten herunter (wenn er zurzeit aktiv ist):

   ```
   ambari-server stop
   ```

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

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

   oder

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

3. 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
   ```

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

   BigInsights 4.1, 4.2 und 4.3 (x86_64)
   SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm

   BigInsights 4.1 und 4.3 (PPC64LE)
   SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.1.0.0-1.ppc64le.rpm

   HDP 2.3, 2.4 und 2.5 (x86_64)
   SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm

5. Erstellen Sie die Definition des lokalen Repositorys. Erstellen Sie z. B. eine Datei namens IBM-SPSS-AnalyticServer-3.1.0.0.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
   ```

6. Erstellen Sie das lokale YUM-Repository.

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

7. 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: Die Schritte 8 und 9 gelten nur bei einer extern verwalteten MySQL-Umgebung.

8. 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 Script add_mysql_user.sh aus /opt/AS_Installable/IBM-SPSS-AnalyticServer auf den Knoten/Host, auf dem die MySQL-Instanz, die als AS_MetaStore verwendet wird, installiert ist. Beispiel: /opt/AS_InstallTools.
      - 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. Ändern Sie z. B. die JDBC-Einstellung mysql://{Analytic-Metaspeicher-Host}/aedb?createDatabaseIfNotExist=true in mysql://{MySQL-Datenbank}/aedb?createDatabaseIfNotExist=true.

9. 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-3.1.0.0</reponame>
   </repo>
</os>
```

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

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

10. 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.

       BigInsights 4.1, 4.2 und 4.3 (x86_64)
       SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

       IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm

       BigInsights 4.1 und 4.3 (PPC64LE)
       SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

       IBM-SPSS-AnalyticServer-3.1.0.0-1.ppc64le.rpm

       HDP 2.3, 2.4 und 2.5 (x86_64)
       SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

       IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm

e. Erstellen Sie die Definition des lokalen Repositorys. Erstellen Sie z. B. eine Datei namens `IBM-SPSS-AnalyticServer-3.1.0.0.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)
```

11. Setzen Sie den Vorgang mit Schritt 5 im Thema „Installation in Ambari" auf Seite 5 fort.

Ubuntu-Anweisungen - 3.1.0

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

Tabelle 3. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.1.0 für Hortonworks Data Platform 2.3 und 2.5 Ubuntu (Englisch)	`spss_as-3.1.0-hdp2.3-2.5-ubun_en.bin`

2. Führen Sie die ausführbare Binärdatei auf einem Computer mit Internetzugriff aus und geben Sie eine Offlineinstallation an. Eine Offlineinstallation lädt die erforderlichen DEB-Dateien herunter und sollte auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann. Die ausführbare Binärdatei befindet sich im verfügbaren Ambari-Verteilerverzeichnis `<installierbares_AS-Ausgangsverzeichnisses>`.
3. Die erforderlichen Analytic Server-DEB-Dateien befinden sich im folgenden Verzeichnis:

```
IBM-SPSS-AnalyticServer/packages
```

4. Installieren Sie Analytic Server 3.1.0 mit den folgenden Befehlen:

```
dpkg -i ./IBM-SPSS-AnalyticServer-ambari-HDP-2.5_3.1.0.0_amd64.deb (oder IBM-SPSS-AnalyticServer-ambari-HDP-2.3_3.1.0.0_amd64.deb)
dpkg -i ./IBM-SPSS-AnalyticServer_1_amd64.deb
```

5. Starten Sie Ihren Ambari-Server erneut.

```
ambari-server restart
```

6. Melden Sie sich an Ihrem Ambari-Server an und installieren Sie Analytic Server als Service über die Ambari-Benutzerschnittstelle.

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 Script `add_mysql_user.sh` aus `/opt/AS_Installable/IBM-SPSS-AnalyticServer` auf den Knoten/Host, auf dem die MySQL-Instanz, die als `AS_MetaStore` verwendet wird, installiert ist. Beispiel: `/opt/AS_InstallTools`.
   - 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.
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/{Version}`.
- {AS-Serverstammverzeichnis} bezieht sich auf den Speicherort der Konfigurations-, Protokoll- und Serverdateien, z. B. `/opt/IBM/SPSS/AnalyticServer/{Version}/ae_wlpserver/usr/servers/aeserver`.
- {AS-Ausgangsverzeichnis} bezieht sich auf den HDFS-Speicherort, der von Analytic Server als Stammordner verwendet wird.

Sicherheit

Der Parameter **`security.config`** definiert die Registrierung von Benutzern und Gruppen, die dem Analytic Server-System als Principals hinzugefügt werden können.

Standardmäßig ist eine Basisregistrierung mit einem Benutzer `admin` und dem Kennwort `admin` definiert. Sie können die Registrierung ändern, indem Sie den Parameter **`security.config`** bearbeiten oder Kerberos konfigurieren. Der Parameter **`security.config`** befindet sich im Abschnitt **Advanced analytics.cfg** der Registerkarte **Configs** des Analytic Server-Service.

Anmerkung: Wenn Sie den Parameter **`security.config`** bearbeiten, um die Registrierung zu ändern, müssen Sie dem Analytic Server-System alle neuen Benutzer als Principals hinzufügen. Details zur Nutzerverwaltung finden Sie im Handbuch *IBM SPSS Analytic Server Verwaltung*.

Vornehmen von Änderungen an der Basisregistrierung

Mithilfe der Basisregistrierung können Sie im Parameter **security.config** eine Datenbank mit Benutzern und Gruppen definieren.

Die Standardbasisregistrierung könnte wie folgt aussehen:

```
<basicRegistry id="basic" realm="ibm">
   <user name="admin" password="admin"/>
</basicRegistry>
```

Es folgt ein Beispiel für eine geänderte Basisregistrierung.

```
<basicRegistry id="basic" realm="ibm">
  <user name="user1" password="{xor}Dz4sLG5tbGs="/>
  <user name="user2" password="Pass"/>
  <user name="user3" password="Pass"/>
  <user name="user4" password="Pass"/>
  <user name="admin" password="{xor}KzosKw=="/>
  <group name="Development">
    <member name="user1"/>
    <member name="user2"/>
  </group>
  <group name="QA">
    <member name="user3"/>
    <member name="user4"/>
  </group>
  <group name="ADMIN">
    <member name="user1"/>
    <member name="admin"/>
  </group>
</basicRegistry>
```

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

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

Anmerkung: Details zum Tool securityUtility finden Sie unter http://www-01.ibm.com/support/knowledgecenter/SSD28V_8.5.5/com.ibm.websphere.wlp.core.doc/ae/rwlp_command_securityutil.html.

Anmerkung: Die Basisregistrierung ist in einer Sandboxumgebung hilfreich, sie wird jedoch für eine Produktionsumgebung nicht empfohlen.

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

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** in **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.

Anmerkung: Wenn die Analytic Server-Installation eine Basisregistrierung verwendet, muss sie die Kerberos-Benutzerkonten enthalten, wobei "-" als Kennwort verwendet wird. Es folgt ein Beispiel.

```
<basicRegistry id="basic" realm="ibm">
  <user name="admin" password="-"/>
  <user name="user1" password="-"/>
  <user name="user2" password="-"/>
  <group name="group1">
     <member name="admin"/>
     <member name="user1"/>
     <member name="user2"/>
  </group>
  <group name="group2">
     <member name="admin"/>
     <member name="user1"/>
  </group>
</basicRegistry>
```

2. 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.

3. 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.
4. [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 Kerberos-SSO an 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. Im Fall von 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 anderen Benutzernamen widerzuspiegeln (z. B. `hadoop.proxyuser.xxxxx.hosts`, wobei `xxxxx` der konfigurierte Benutzername ist, der in der Analytic Server-Konfiguration angegeben ist).

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/3.0/ae_wlpserver/bin/securityUtility encode <Kennwort>`

Wenn Sie ein selbst signiertes Zertifikat generieren wollen, können Sie das Tool securityUtility verwenden. Beispiel: `/opt/ibm/spss/analyticserver/3.0/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.

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

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. 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.
2. Führen Sie die selbstextrahierende 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.

   [Nur GPFS (Spectrum Scale)] Laden Sie die Datei https://ibm-open-platform.ibm.com/repos/IBM-SPSS-ModelerEssentialsR/3.1.0.0/x86_64/IBM-SPSS-AnalyticServer-3.1.0.0.repo (x86), https://ibm-open-platform.ibm.com/repos/IBM-SPSS-ModelerEssentialsR/3.1.0.0/ppc64le/IBM-SPSS-AnalyticServer-3.1.0.0.repo (ppc64le) oder https://ibm-open-platform.ibm.com/repos/IBM-SPSS-ModelerEssentialsR/3.1.0.0/Ubuntu/IBM-SPSS-ModelerEssentialsR-3.1.0.0.list (Ubuntu) herunter und verschieben Sie sie auf jedem Knoten, auf dem Sie Analytic Server Metastore als Service hinzufügen, in den Ordner `/etc/yum.repos.d` (RHEL, CentOS), `/etc/zypp/repos.d` (SLES) oder `/etc/apt/sources.list.d` (Ubuntu).

   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.

BigInsights 4.1 und 4.2 (x86_64)
IBM-SPSS-ModelerEssentialsR-ambari-2.1-BI-4.1-8.4.1.0-1.x86_64.rpm

BigInsights 4.1 (PPC64LE)
IBM-SPSS-ModelerEssentialsR-ambari-2.1-BI-4.1-8.4.1.0-1.ppc64le.rpm

HDP 2.3 und 2.4 (x86_64)
IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.3-8.4.1.0-1.x86_64.rpm

HDP 2.4 (Ubuntu)
IBM-SPSS-ModelerEssentialsR-ambari-3.1.0.0_3.1.0.0_amd64.deb

b. Installieren Sie den RPM oder DEB. Im folgenden Beispiel installiert der Befehl Essentials for R unter BigInsights 4.2.

```
rpm -i IBM-SPSS-ModelerEssentialsR-ambari-2.1-BI-4.1-8.4.1.0-1.x86_64.rpm
```

Im folgenden Beispiel installiert der Befehl Essentials for R unter HDP 2.4 (Ubuntu).

```
dpkg -i IBM-SPSS-ModelerEssentialsR-ambari-3.1.0.0_3.1.0.0_amd64.deb
```

3. Starten Sie Ihren Ambari-Server erneut.

```
ambari-server restart
```

4. 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 Analytic Metastore 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.

5. Aktualisieren Sie den Analytic Server-Service.
6. Führen Sie das Script `update_clientdeps` unter Beachtung der Anweisungen in „Aktualisierung von Clientabhängigkeiten" auf Seite 21 aus.
7. 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 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 4. 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	10.5, 10.1, 9.7	`db2jcc.jar`	IBM
DB2 z/OS	11, 10	`db2jcc.jar`, `db2_license_cisuz.jar`	IBM
Greenplum	5, 4.2.x	postgresql.jar	Greenplum
Hive	1.1, 1.2	`hive-jdbc-*.jar`	Apache
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
Sybase IQ	16.x, 15.4, 15.2	`jconnect70.jar`	Sybase
Teradata	14, 14.1, 15	`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 die gleiche Datei, die für DB2 verwendet wird.)

  BigSQL ist ein BigInsights-Mehrwertdienst. Daher ist seine Version mit der BigInsights-Version identisch. 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. Details hierzu finden Sie in den folgenden Abschnitten.
2. Fügen Sie diese JAR-Dateien dem Verzeichnis `{HIVE-Ausgangsverzeichnis}/auxlib` und dem Verzeichnis `/usr/share/hive` auf allen Analytic Server-Knoten hinzu.
3. Starten Sie den Hive-Metaspeicherservice erneut.
4. Aktualisieren Sie den Analytic Metastore-Service.
5. Starten Sie jede einzelne Instanz des Analytic Server-Service erneut.

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.

Apache Spark

Wenn Sie Spark (Version 1.5 oder höher) mit einer HCatalog-Eingabedatenquelle verwenden wollen, müssen Sie die Eigenschaft `spark.version` der angepassten Datei `analytics.cfg` 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 und starten Sie alle Analytic Server-Services über die Amabri-Konsole erneut.

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`

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**.

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 den Abschnitt **Advanced analytics-jvm-options** der Registerkarte **Configs** für den Analytic Server-Service 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/2.4/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

- 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_0.6-3.1.0.0.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:

  `/KnoxServicePath/ext`

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

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

  auf den Knox-Serverknoten:

  `/KnoxServicePath/data/services`

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

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 der Knox-Konfiguration den folgenden Service hinzu:

   ```
   <service>
       <role>ANALYTICSERVER</role>
       <url>http://{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 können in der Ambari-Benutzerschnittstelle durch Auswahl von **Knox** > **Configs** > **Advanced topology** angezeigt werden. 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. Erstellen Sie auf dem Knox-Server das Unterverzeichnis <Knox-Server>/data/service/analyticserver/3.1 und laden Sie dann die Dateien service.xml und rewrite.xml in das neue Verzeichnis hoch. Die beiden Dateien befinden sich auf Analytic Server in <Analytic_Server>/configuration/knox/analyticserver/3.1 (z. B. /opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/3.1/*.xml).
2. Führen Sie das Script ./knoxcli.sh redeploy --cluster default in <Knox-Server>/bin aus.
3. Laden Sie die Datei com.ibm.spss.knoxservice_0.6-*.jar in <Knox-Server>/ext hoch. Die Datei befindet sich auf Analytic Server in <Analytic_Server>/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.1.0.0.jar (z. B. /opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.1.0.0.jar).
4. 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>
   </service>
   ```
5. In der Ambari-Benutzerschnittstelle müssen Sie die Benutzer in **Knox** > **Configs** > **Advanced users-ldif** (z. B. admin, qauser1, qauser2) hinzufügen oder aktualisieren.
6. Starten Sie LDAP über **Knox** > **Service Actions** > **Start Demo LDAP** erneut.
7. Starten Sie den Knox-Service erneut.

Installieren von Apache Knox auf Hortonworks Data Platform (HDP)

Die folgenden Schritte legen die Installation von Apache Knox in einem HDP-Cluster dar.

1. Prüfen Sie, ob auf dem HDP-Cluster ein Knox-Benutzer vorhanden ist. Wenn kein Knox-Benutzer vorhanden ist, müssen Sie einen erstellen.
2. Laden Sie Apache Knox in einen Ordner unter /home/knox herunter und extrahieren Sie Apache Knox.
3. Wechseln Sie in HDP zum Knox-Benutzer und gehen Sie zum Ordner knox. Der Knox-Benutzer muss über permission(RWX) für alle knox-Unterordner verfügen.

4. Konfigurieren Sie Apache Knox für Analytic Server. Weitere Informationen finden Sie im Abschnitt **Konfigurieren von Apache Knox**.
   a. Erstellen Sie eine Ordnerhierarchie `analyticserver/3.0` unter `{Knox}/data/services`.
   b. Kopieren Sie die Dateien `rewrite.xml` und `service.xml` aus dem Analytic Server-Speicherort

      `/opt/ibm/spss/analyticserver/3.0/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/3.1`

      auf den Knox-Serverknoten:

      `{knox}/data/services/analyticserver/3.1`
   c. Kopieren Sie die *.jar-Knox-Datei vom Analytic Server-Host:

      `/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-*.jar`

      in das Knox-Verzeichnis`ext`:

      `{Knox}/ext`
   d. Aktualisieren Sie die Datei `default.xml` in `{Knox}/conf/topologies`, sodass sie mit dem folgenden Beispiel übereinstimmt:

      Anmerkung: Sie müssen die Datei erstellen, wenn sie nicht vorhanden ist.

```
<topology>
   <gateway>
      <provider>
         <role>authentication</role>
         <name>ShiroProvider</name>
         <enabled>true</enabled>
         <param>
            <name>sessionTimeout</name>
            <value>30</value>
         </param>
         <param>
            <name>main.ldapRealm</name>
            <value>org.apache.hadoop.gateway.shirorealm.KnoxLdapRealm</value>
         </param>
         <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://localhost:33389</value>
         </param>
         <param>
            <name>main.ldapRealm.contextFactory.authenticationMechanism</name>
            <value>simple</value>
         </param>
         <param>
            <name>urls./**</name>
            <value>authcBasic</value>
         </param>
      </provider>
      <provider>
         <role>identity-assertion</role>
         <name>Default</name>
         <enabled>true</enabled>
      </provider>
      <provider>
         <role>authorization</role>
         <name>AclsAuthz</name>
         <enabled>true</enabled>
      </provider>

   </gateway>

   <!--other service-->
   <service>
      <role>ANALYTICSERVER</role>
      <!--replace the {AS-host}nas {AS-port} with real value-->
      <url>http://{AS-host}:{AS-port}/analyticserver</url>
   </service>

</topology>
```

5. Führen Sie `{Knox}/bin/knoxcli.sh` aus.
6. Führen Sie `{Knox}/bin/ldap.sh start` aus.

Anmerkung: Das Script verwendet den Port 33389. Stellen Sie sicher, dass der Port zurzeit nicht belegt ist.

7. Führen Sie `{Knox}/bin/gateway.sh start` aus.

   Anmerkung: Das Script verwendet den Port 8443. Stellen Sie sicher, dass der Port zurzeit nicht belegt ist.

8. Prüfen Sie die Installation.
   a. Führen Sie den Befehl `curl` für Analytic Server über die Knox-URL aus:

      ```
      curl -ikvu {Benutzername}:{Kennwort} https://{Knox-Host}:8443/gateway/default/analyticserver/admin
      ```

Fehlerbehebung

Problem: Nach der Installation funktioniert Analytic Server nicht in Knox.

Lösung: Stoppen Sie Knox, entfernen Sie alle Dateien unter `{Knox}/data/deployments/*` und starten Sie Knox dann erneut.

Problem: Anmeldung an Analytic Server über Knox nicht möglich.

Lösung: Prüfen Sie die Benutzer in `{Knox}/conf/users.ldif`. Aktualisieren Sie vorhandene Benutzer oder fügen Sie neue Analytic Server-Benutzer hinzu. Die Principals und Berechtigungsnachweise des Knox-Benutzers müssen mit den Analytic Server-Benutzern übereinstimmen.

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`.

Upgrade und Migration

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

Durchführen eines Upgrades von Version 3.0.1 auf 3.1.0 - BigInsights und Hortonworks

Wenn Sie über eine vorhandene Installation von Analytic Server 3.0.1 verfügen, können Sie für diese Installation ein Upgrade auf Version 3.1.0 durchführen.

1. Stoppen Sie den Analytic Server-Service in der Ambari-Konsole.
2. Führen Sie abhängig von Ihrem Installationstyp die folgenden Schritte aus.

   Online-Upgrade

   a. Stellen Sie sicher, dass Ihr Ambari-Server-Host und alle Knoten im Cluster auf https://ibm-open-platform.ibm.com zugreifen können.
   b. Laden Sie die Datei `IBM-SPSS-AnalyticServer-3.1.0.0.repo` von https://ibm-open-platform.ibm-.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms/IBM-SPSS-AnalyticServer-3.1.0.0.repo (x86 und ppc64le) auf jeden Analytic Server-Host herunter und verschieben Sie sie in den Ordner `/etc/yum.repos.d` (RHEL oder CentOS) oder `/etc/zypp/repos.d` (SLES).

   Offline-Upgrade

a. Das Offline-Upgrade lädt die erforderlichen RPM-Dateien herunter und sollte auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann.
b. Erstellen Sie ein neues Verzeichnis, das als Repository für die Analytic Server-RPM-Dateien verwendet wird. Siehe das folgende Beispiel:

```
mkdir /home/root/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms
```

c. Kopieren Sie die erforderlichen Analytic Server-RPM-Dateien in dieses Verzeichnis. Die erforderlichen RPM-Dateien hängen von Ihrer Verteilung, Version und Architektur ab. Für BigInsights 4.2 werden die erforderlichen Dateien im Folgenden gezeigt.

Tabelle 5. BigInsights 4.2 RPMs

BigInsights 4.2 (x86_64)
IBM-SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

d. Erstellen Sie die Definition des lokalen Repositorys. Erstellen Sie beispielsweise eine Datei `analyticserver.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-3.1.0.0
baseurl=file:///{Pfad zum lokalen Repository}
enabled=1
gpgcheck=0
protect=1
```

e. Erstellen Sie das lokale YUM-Repository. Siehe das folgende Beispiel:

```
createrepo /home/root/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms
```

3. Löschen Sie die Ambari-Metadaten aus Ihrem lokalen Cache. Wenn Sie z. B. den Cache unter RHEL oder CentOS löschen wollen, führen Sie den folgenden Befehl aus:

```
sudo yum clean all
```

Anmerkung: Der Befehl `yum` kann nicht ausgeführt werden, wenn zwei Analytic Server-Repositorys aufgelistet sind. Daher müssen die ursprünglichen `*.repo`-Dateien, die zu Analytic Server gehören, umbenannt oder entfernt werden. Unter SLES ändert sich der Befehl wie folgt:

```
sudo zypper refresh
```

4. Führen Sie auf jedem Analytic Server-Host ein Upgrade für die RPMs durch. Wenn Sie z. B. ein Upgrade unter RHEL oder CentOS durchführen wollen, führen Sie die folgenden Befehle aus:

```
chown -R as_user:hadoop /opt/ibm/spss/analyticserver/3.0
sudo yum upgrade IBM-SPSS-AnalyticServer
```

Unter SLES ändert sich der Befehl wie folgt:

```
sudo zypper up IBM-SPSS-AnalyticServer
```

5. Aktualisieren Sie den Stack.

BigInsights

a. Starten und stoppen Sie anschließend den Analytic Server-Service in der Ambari-Konsole.
b. Führen Sie die angepasste Aktualisierungsaktion aus.

Hortonworks

Navigieren Sie zu einem Ihrer Analytic Server-Knoten und führen Sie den folgenden Befehl aus:

```
sudo -u as_user /opt/ibm/spss/analyticserver/3.1/bin/refresh.sh
```

6. Nur Offlineinstallation. 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-3.1.0.0</reponame>
    </repo>
</os>
```

7. Löschen Sie den Zookeeper-Status. Führen Sie den folgenden Befehl im Zookeper-bin-Verzeichnis aus (z. B. `/usr/iop/current/zookeeper-server/bin`):

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

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

Migration auf eine neue Version von Analytic Server

Wenn Sie über eine vorhandene Installation von Analytic Server 2.0 oder 2.1 verfügen und Version 3.1.0 erworben haben, können Sie Ihre Konfigurationseinstellungen von Version 2.0/2.1 auf Ihre Installation von Version 3.1.0 migrieren.

Einschränkungen:

- Wenn eine Version vor Version 2.0 installiert ist, müssen Sie die frühere Version zuerst auf Version 2.0/2.1 und dann von Version 2.0/2.1 auf 3.1.0 migrieren.
- Installationen von 2.0/2.1 und 3.1.0 können nicht auf demselben Hadoop-Cluster koexistieren. Wenn Sie Ihre Installation von 3.1.0 für die Verwendung desselben Hadoop-Clusters wie die Installation von 2.0/2.1 konfigurieren, funktioniert die Installation von 2.0/2.1 nicht mehr.

Migrationsschritte (2.0/2.1 auf 3.1.0)

1. Führen Sie die Neuinstallation von Analytic Server entsprechend den Anweisungen in „Installation in Ambari" auf Seite 5 durch.
2. Kopieren Sie das Analytic Server-Stammverzeichnis von Ihrer alten Installation in die neue Installation.
   a. Wenn Sie sich nicht sicher sind, wo sich das Analytic Server-Stammverzeichnis befindet, führen Sie den Befehl `hadoop -fs ls` aus. Der Pfad zum Analytic Server-Stammverzeichnis hat das Format `/user/aeuser/analytic-root`, wobei aeuser die Benutzer-ID ist, die Eigner des Analytic Server-Stammverzeichnisses ist.
   b. Ändern Sie das Eigentumsrecht von aeuser in as_user:

      ```
      hadoop dfs -chown -R {as_user:{Gruppe}} {Pfad zu 2.0/2.1-AS-Stammverzeichnis}
      ```

      Anmerkung: Wenn Sie die vorhandene Analytic Server-Installation nach der Migration verwenden wollen, erstellen Sie eine Kopie des Analytic Server-Stammverzeichnisses in HDFS und ändern dann das Eigentumsrecht für die Kopie des Verzeichnisses.
   c. Melden Sie sich als as_user am Host der neuen Analytic Server-Installation an. Löschen Sie das Verzeichnis `/user/as_user/analytic-root`, falls es vorhanden ist.
   d. Führen Sie das folgende Kopierscript aus:

      ```
      hadoop distcp hftp://{Host des 2.0/2.1-Namensknotens}:50070/{Pfad zu 2.0/2.1-AS-Stammverzeichnis}
      hdfs://{Host des 3.1.0-Namensknotens}/user/as_user/analytic-root
      ```

3. Stoppen Sie den Analytic Server-Service in der Ambari-Konsole.
4. Stellen Sie 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`. 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.

6. Löschen Sie den Zookeeper-Status. Führen Sie den folgenden Befehl im Zookeper-bin-Verzeichnis aus (z. B. `/usr/hdp/current/zookeeper-client` unter Hortonworks oder `/usr/iop/current/zookeeper-server` unter BigInsights).

   ```
   ./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/3.1/ASConfiguration_2.1.0.0.xxx.zip
   ```

8. 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, müssen Sie die Schritte zum Konfigurieren von R mit der neuen Analytic Server-Installation befolgen.

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".

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: In BigInsights 4.2 befindet sich das Verzeichnis `ESSENTIALR` im Verzeichnis `/var/lib/ambari-server/resources/stacks/BigInsights/4.2/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

Metadatenrepository

Wenn Sie MySQL als Metadatenrepository von Analytic Server verwenden wollen, befolgen Sie die Anweisungen für „Konfigurieren von MySQL für Analytic Server”.

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 32.

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.

Installation in Cloudera

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

Analytic Server 3.1.0

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 6. Selbstextrahierende Analytic Server-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.1 für Cloudera 5.8, 5.9 und 5.10 Ubuntu (Englisch)	spss_as-3.1-cdh5.8-5.10-ubun_en.bin
IBM SPSS Analytic Server 3.1 für Cloudera 5.8, 5.9 und 5.10 Linux x86-64 (Englisch)	spss_as-3.1-cdh5.8-5.10-lx86-en.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. Starten Sie Cloudera Manager nach Abschluss der Installation erneut.
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** 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. Konfigurieren Sie MySQL für Analytic Server.
8. 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:

- Hostname für Analytic Server-Metaspeicher
- Datenbankname für Analytic Server-Metaspeicher
- Benutzername für Analytic Server-Metaspeicher
- Kennwort für Analytic Server-Metaspeicher

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.

Anmerkung: Klicken Sie nach der erfolgreichen Installation von Analytic Server nicht auf **Create Analytic Server Metastore** in der Liste **Actions** der Seite für Analytic Server-Services in Cloudera Manager. 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-3.1.0.0-el6.parcel
- AnalyticServer-3.1.0.0-el6.parcel.sha
- manifest.json

  oder
- AnalyticServer-3.1.0.0-el7.parcel
- AnalyticServer-3.1.0.0-el7.parcel.sha

SuSE Linux erfordert die folgenden Dateien:
- AnalyticServer-3.1.0.0-sles11.parcel
- AnalyticServer-3.1.0.0-sles11.parcel.sha
- manifest.json

Ubuntu Linux erfordert die folgenden Dateien:
- AnalyticServer-3.1.0.0-trusty.parcel
- AnalyticServer-3.1.0.0-trusty.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).

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

   Die Analytic Server-PARCEL-Datei wird als **distributed** und **activated** angezeigt.

Durchführen eines Upgrades auf Analytic Server 3.1.0 in Cloudera

Wenn Sie über eine vorhandene Installation von Analytic Server 3.0/3.0.1 verfügen, können Sie für diese Installation ein Upgrade auf Version 3.1.0 durchführen.

1. Stoppen und löschen Sie den Analytic Server-Service in Cloudera Manager.
2. Inaktivieren Sie in Cloudera Manager die Vorgängerversion von Analytic Server.
3. Anweisungen zum Installieren von Analytic Server 3.1.0 finden Sie im Abschnitt "Online" oder "Offline" in „Upgrade und Migration" auf Seite 25.

   4. Nachdem der Analytic Server-Service installiert und in Cloudera Manager hinzugefügt wurde, führen Sie **Refresh Analytic Server Binaries** aus. Analytic Server 3.1.0 ist jetzt einsatzbereit.

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 Parameter **`security_cfg`** definiert die Registrierung von Benutzern und Gruppen, die dem Analytic Server-System als Principals hinzugefügt werden können.

Standardmäßig ist eine Basisregistrierung mit einem Benutzer `admin` und dem Kennwort `admin` definiert. Sie können die Registrierung ändern, indem Sie den Parameter **`security_cfg`** bearbeiten oder Kerberos als Sicherheitsprovider konfigurieren. Der Parameter **`security_cfg`** befindet sich im Abschnitt **Analytic Server Advanced Configuration Snippet** der Registerkarte **Configuration** des Analytic Server-Service.

Anmerkung: Wenn Sie den Parameter **`security_cfg`** bearbeiten, um die Registrierung zu ändern, müssen Sie dem Analytic Server-System alle neuen Benutzer als Principals hinzufügen. Details zur Nutzerverwaltung finden Sie im Handbuch *IBM SPSS Analytic Server Verwaltung*.

Vornehmen von Änderungen an der Basisregistrierung

Mithilfe der Basisregistrierung können Sie im Parameter **`security_cfg`** eine Datenbank mit Benutzern und Gruppen definieren.

Die Standardbasisregistrierung könnte wie folgt aussehen:

```
<basicRegistry id="basic" realm="ibm">
   <user name="admin" password="admin"/>
</basicRegistry>
```

Es folgt ein Beispiel für eine geänderte Basisregistrierung.

```
<basicRegistry id="basic" realm="ibm">
  <user name="user1" password="{xor}Dz4sLG5tbGs="/>
  <user name="user2" password="Pass"/>
  <user name="user3" password="Pass"/>
  <user name="user4" password="Pass"/>
  <user name="admin" password="{xor}KzosKw=="/>
  <group name="Development">
    <member name="user1"/>
```

```
    <member name="user2"/>
  </group>
  <group name="QA">
    <member name="user3"/>
    <member name="user4"/>
  </group>
  <group name="ADMIN">
    <member name="user1"/>
    <member name="admin"/>
  </group>
</basicRegistry>
```

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

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

Anmerkung: Details zum Tool securityUtility finden Sie unter http://www-01.ibm.com/support/knowledgecenter/SSD28V_8.5.5/com.ibm.websphere.wlp.core.doc/ae/rwlp_command_securityutil.html.

Anmerkung: Die Basisregistrierung ist in einer Sandboxumgebung hilfreich, sie wird jedoch für eine Produktionsumgebung nicht empfohlen.

Konfigurieren einer LDAP-Registry

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

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>
```

Weitere Beispiele für Konfigurationen finden Sie im Vorlagenordner {AS-Stammverzeichnis}/ae_wlpserver/templates/config.

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 in Cloudera.

1. Sie können im Kerberos-Benutzerrepository für alle Benutzer, denen Sie Zugriff auf Analytic Server erteilen möchten, Konten erstellen.

   Anmerkung: Wenn die Analytic Server-Installation eine Basisregistrierung verwendet, muss sie die Kerberos-Benutzerkonten enthalten, wobei "-" als Kennwort verwendet wird. Es folgt ein Beispiel.

```
<basicRegistry id="basic" realm="ibm">
  <user name="admin" password="-"/>
  <user name="user1" password="-"/>
  <user name="user2" password="-"/>
  <group name="group1">
     <member name="admin"/>
     <member name="user1"/>
     <member name="user2"/>
  </group>
  <group name="group2">
     <member name="admin"/>
     <member name="user1"/>
  </group>
</basicRegistry>
```

2. 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.

3. 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.
4. 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 Parameter 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 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 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 Kerberos-SSO an 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. Öffnen Sie Cloudera Manager und fügen Sie die folgenden Eigenschaften im Analytic Server-Bereich **Analytic Server Advanced Configuration Snippet (Safety Valve) for core-site.xml** hinzu oder aktualisieren Sie sie.

```
- hadoop.proxyuser.as_user.hosts = *
- hadoop.proxyuser.as_user.groups = *
```

2. Wenn Analytic Server für die Verwendung eines anderen Benutzernamens als `as_user` konfiguriert ist, müssen Sie die Eigenschaftsnamen ändern, um den anderen Benutzernamen widerzuspiegeln (z. B. `hadoop.proxyuser.xxxxx.hosts`, wobei `xxxxx` der konfigurierte Benutzername ist, der in der Analytic Server-Konfiguration angegeben ist).

Anmerkung: Die Eigenschaften werden in Ambari (basierend auf Werten in der Analytic Server-Konfiguration) automatisch hinzugefügt.

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 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.

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

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. Laden Sie das selbstextrahierende 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.
2. Führen Sie das selbstextrahierende 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++
3. Das selbstextrahierende 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 3.1.0 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.
4. 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).
5. 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.

6. 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 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 **Stop** im Dropdown-Menü **Actions** aus, um den Analytic Server-Service zu stoppen.
5. Wählen Sie **Refresh Analytic Server Binaries** im Dropdown-Menü **Actions** aus.
6. Wählen Sie **Start** im Dropdown-Menü **Actions** aus, um den Analytic Server-Service zu starten.

Tabelle 7. 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
DashDB	Bluemix-Service	`db2jcc.jar`	IBM
DB2 for Linux, UNIX, and Windows	10.5, 10.1, 9.7	`db2jcc.jar`	IBM
DB2 z/OS	11, 10	`db2jcc.jar`, `db2_license_cisuz.jar`	IBM
Greenplum	5, 4.2.x	postgresql.jar	Greenplum
Hive	1.1, 1.2	`hive-jdbc-*.jar`	Apache
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
Sybase IQ	16.x, 15.4, 15.2	`jconnect70.jar`	Sybase
Teradata	14, 14.1, 15	`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 dem Verzeichnis `{HIVE-Ausgangsverzeichnis}/auxlib` und dem Verzeichnis `/usr/share/hive` auf allen Analytic Server-Knoten hinzu.
3. Starten Sie den Hive-Metaspeicherservice erneut.
4. Starten Sie jede einzelne Instanz des Analytic Server-Service erneut.

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.

Apache Spark

Wenn Sie Spark (Version 1.5 oder höher) mit einer HCatalog-Eingabedatenquelle verwenden wollen, müssen Sie die Eigenschaft `spark.version=X.X.0` (z. B. `spark.version=2.0.0` manuell hinzufügen.

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

   ```
   spark.version=2.0.0
   ```

2. Speichern Sie die Konfiguration und starten Sie alle Analytic Server-Services über Cloudera Manager erneut.

Konfigurieren von Apache Impala

Apache Impala wird unterstützt, wenn es in 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 Analytic Server-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.

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
```

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 2.0/2.1 verfügen und eine neuere Version erworben haben, können Sie Ihre Konfigurationseinstellungen von Version 2.0/2.1 auf Ihre neue Installation migrieren.

Einschränkung: Wenn eine Version vor Version 2.0 installiert ist, müssen Sie die frühere Version zuerst auf Version 2.0/2.1 und dann von Version 2.0/2.1 auf die neuere Version migrieren.

Einschränkung: Installationen von 2.0/2.1 und neueren Versionen können nicht auf demselben Hadoop-Cluster koexistieren. Wenn Sie Ihre neue Installation für die Verwendung desselben Hadoop-Clusters wie die Installation von 2.0/2.1 konfigurieren, funktioniert die Installation von 2.0/2.1 nicht mehr.

Migrationsschritte (2.1 auf neuere Version)

1. Führen Sie die Neuinstallation von Analytic Server entsprechend den Anweisungen in „Installation in Cloudera" auf Seite 32 durch.
2. Kopieren Sie das Analytic Server-Stammverzeichnis von Ihrer alten Installation in die neue Installation.
   a. Wenn Sie sich nicht sicher sind, wo sich das Analytic Server-Stammverzeichnis befindet, führen Sie den Befehl `hadoop -fs ls` aus. Der Pfad zum Analytic Server-Stammverzeichnis hat das Format `/user/aeuser/analytic-root`, wobei aeuser die Benutzer-ID ist, die Eigner des Analytic Server-Stammverzeichnisses ist.
   b. Ändern Sie das Eigentumsrecht von aeuser in as_user:

      ```
      hadoop dfs -chown -R {as_user:{Gruppe}} {Pfad zu 2.1-AS-Stammverzeichnis}
      ```

      Anmerkung: Wenn Sie die vorhandene Analytic Server-Installation nach der Migration verwenden wollen, erstellen Sie eine Kopie des Analytic Server-Stammverzeichnisses in HDFS und ändern dann das Eigentumsrecht für die Kopie des Verzeichnisses.
   c. Melden Sie sich als as_user am Host der neuen Analytic Server-Installation an. Löschen Sie das Verzeichnis `/user/as_user/analytic-root`, falls es vorhanden ist.
   d. Führen Sie das folgende Kopierscript aus:

      ```
      hadoop distcp hftp://{Host des 2.1-Namensknotens}:50070/{Pfad zu 2.1-AS-Stammverzeichnis}
      hdfs://{Host des 3.1-Namensknotens}/user/as_user/analytic-root
      ```
3. Stoppen Sie den Analytic Server-Service in Cloudera Manager.
4. 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.
5. 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/3.1/ASConfiguration_2.1.0.0.xxx.zip
   ```
6. Löschen Sie den Zookeeper-Status. Führen Sie den folgenden Befehl im Zookeper-bin-Verzeichnis aus (z. B. `/opt/cloudera/parcels/CDH-5.4...../lib/zookeeper/bin` in Cloudera).

   ```
   ./zkCli.sh rmr /AnalyticServer
   ```
7. 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. Inaktivieren Sie die Analytic Server-PARCEL-Dateien und entfernen Sie sie von den Hosts, indem Sie **Deactivate** und **Remove From Hosts** auswählen.
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.

Kapitel 4. MapR-Installation und -Konfiguration

MapR - Übersicht

MapR ist eine vollständige Verteilung für Apache Hadoop, die mehr als ein Dutzend Projekte aus dem Hadoop-Ökosystem in einem Paket enthält, um ein umfangreiches Set mit Big Data-Funktionalität bereitzustellen.

Auf das MapR-Dateisystem kann nicht von außerhalb des Server-Clusters zugegriffen werden. Deshalb muss IBM SPSS Analytic Server in den MapR-Clusterknoten bereitgestellt werden. In diesem Bereitstellungsszenario muss Analytic Server von einem Benutzer ausgeführt werden, der eine Zugriffsberechtigung für das MapR-Dateisystem hat und Jobs an YARN übergibt, um die Bereitstellung für Analytic Server (als `<as_user>`) zu ermöglichen.

Installieren von Analytic Server in MapR

In den folgenden Schritten wird der Prozess der manuellen Installation von IBM SPSS Analytic Server in einem MapR-Cluster beschrieben.

Installieren von Analytic Server 3.1.0 unter MapR 5.0 oder 5.1

1. Navigieren Sie zur IBM Passport Advantage®-Web-Site und laden Sie die selbstextrahierende MapR-Binärdatei herunter.

Tabelle 8. Selbstextrahierende MapR-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.1 für MapR 5.0 und 5.1 Linux x86-64 (Englisch)	`spss_as-3.1.0-mapr5.0-5.1-lx86_en.bin`

2. Führen Sie das Analytic Server-Installationsprogramm als Root- oder sudo-Benutzer aus. Befolgen Sie während der Installation die Aufforderungen zum Akzeptieren der Lizenz und wählen Sie aus, ob Analytic Server online oder offline installiert werden soll.
   a. Wählen sie die Onlineoption aus, wenn der Server, der Analytic Server hostet, eine Internetverbindung zu `https://ibm-open-platform.ibm.com` hat. Das Installationsprogramm installiert Analytic Server automatisch.
   b. Wählen sie die Offlineoption aus, wenn der Server, der Analytic Server hostet, keine Internetverbindung zu `https://ibm-open-platform.ibm.com` hat. Führen Sie das Installationsprogramm auf einem anderen Server aus, der auf die URL zugreifen kann, und wählen Sie die Offlineinstallation von Analytic Server aus. Das Installationsprogramm lädt das RPM- oder DEB-Paket automatisch herunter.
3. Suchen Sie die RPM- oder DEB-Datei für Analytic Server und führen Sie sie aus:
   - RedHat oder SuSe Linux:

     ```
     rpm -ivh IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm
     ```

   - Ubuntu Linux:

     ```
     dpkg -i IBM-SPSS-AnalyticServer_1_amd64.deb
     ```

   Sowohl im Online- als auch im Offlineinstallationsmodus wird Analytic Server im Verzeichnis `/opt/ibm/spss/analyticserver/3.1` (als `<AS-Installationspfad>`) installiert.
4. Ändern Sie alle Dateien im Installationspfad in den Benutzer, der Analytic Server ausführt:

   ```
   chown -R <as_user> <AS-Installationspfad>
   ```

   Wechseln Sie den Benutzer zu `<as_user>`; in allen nachfolgenden Schritten wird `<as_user>` verwendet.

5. Konfigurieren Sie die HTTP-Eigenschaft. Erstellen Sie eine Datei mit dem Namen `http_endpoint.xml` im Pfad `<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver` und fügen Sie der Datei die folgenden Zeilen hinzu:

```
<server>
    <httpEndpoint host="*" id="defaultHttpEndpoint" httpPort="<HTTP-Port>" httpsPort="<HTTPS-Port>" onError="FAIL"/>
</server>
```

   `<HTTP-Port>` und `<HTTPS-Port>` sind die Ports, die von Analytic Server über die Protokolle HTTP und HTTPS verwendet werden. Ersetzen Sie sie durch verfügbare Ports.

6. Fügen Sie Benutzer und Gruppen hinzu. Erstellen Sie eine Datei mit dem Namen `security_cfg.xml` im Pfad `<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver` und fügen Sie der Datei die folgenden Zeilen hinzu:

```
<server>
    <basicRegistry id="basic" realm="ibm">
        <user name="admin" password="test"/>
    </basicRegistry>
</server>
```

   Standardmäßig enthält die XML-Datei nur den Benutzer admin. Sie müssen weitere Benutzer und Gruppen in der Einstellung `<basicRegistry>` manuell hinzufügen oder die Einstellung in `ldapRegistry` ändern.

7. Richten Sie die Metadatendatenbank ein. Analytic Server unterstützt die DB2- und MySQL-Datenbanken.

   a. Konfigurieren Sie die Datenbankbenutzer. Wird die MySQL-Datenbank verwendet, führen Sie das folgende SQL-Script in der MySQL-Shell aus:

```
DROP DATABASE IF EXISTS <DB-Name>;
CREATE DATABASE <DB-Name> DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8_bin;
CREATE USER '<DB-Benutzername>'@'%' IDENTIFIED BY '<DB-Kennwort>';
CREATE USER '<DB-Benutzername>'@'localhost' IDENTIFIED BY '<DB-Kennwort>';
GRANT ALL PRIVILEGES ON *.* TO '<DB-Benutzername>'@'%';
GRANT ALL PRIVILEGES ON *.* TO '<DB-Benutzername>'@'localhost';
```

   b. Verschlüsseln Sie das Kennwort. Die Kennwörter der Datenbankbenutzer müssen verschlüsselt werden, bevor sie an Analytic Server übergeben werden können. Führen Sie den folgenden Befehl aus:

```
java -Duser.language=en -cp <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/*
com.spss.ae.encryption.provider.EncryptKeystorePassword <DB-Kennwort>
```

      Anmerkung: Wird der Befehl direkt in einer Linux-Shell ausgeführt, muss das Zeichen `*` möglicherweise mit einem Escapezeichen (`\*`) versehen werden.

      Die Befehlsausgabe sieht wie folgt aus: `The encrypted password is '<verschlüsseltes_DB-Kennwort>'`. Notieren Sie das verschlüsselte Datenbankkennwort.

   c. Löschen Sie die Datei `<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`, falls sie vorhanden ist, und erstellen Sie eine neue Datei mit demselben Namen. Ändern Sie die folgenden Eigenschaften, wenn die DB2-Datenbank verwendet wird:

```
jndi.aedb=jdbc/aeds
jndi.aedb.url=jdbc:db2://<DB-Host>:<DB-Port>/<DB-Name>:currentSchema=<DB-Schemaname>;
jndi.aedb.driver=com.ibm.db2.jcc.DB2Driver
jndi.aedb.username=<DB-Benutzername>
jndi.aedb.password=<verschlüsseltes_DB-Kennwort>
```

      Ist das Schema `<DB-Schemaname>` nicht vorhanden, muss der Benutzer `<DB-Benutzername>` eine implizite Berechtigung zum Erstellen des Schemas haben. Ändern Sie die folgenden Eigenschaften, wenn die MySQL-Datenbank verwendet wird:

```
jndi.aedb=jdbc/aeds
jndi.aedb.url=jdbc:mysql://<DB-Host>:<DB-Port>/<DB-Name>?createDatabaseIfNotExist=true
jndi.aedb.driver=com.mysql.jdbc.Driver
jndi.aedb.username=<DB-Benutzername>
jndi.aedb.password=<verschlüsseltes_DB-Kennwort>
```

   d. Der MySQL-JDBC-Treiber muss installiert sein, wenn die MySQL-Datenbank verwendet wird. Führen Sie den folgenden Befehl aus:

```
yum install mysql-connector-java
```

   e. Führen Sie den folgenden Befehl aus, um die erforderlichen Tabellen zu erstellen:

```
cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/sql/<DB-Typ>
java -Xmx128m -Xms128m -cp <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/*:/usr/share/java/*
com.spss.ae.dbscript.ScriptRunner ../../configuration/config.properties schema.sql true
```

Abhängig von der verwendeten Datenbank ist der <DB-Typ> entweder db2 oder mysql.

Anmerkung: Wird MySQL mit der MYISAM-Engine verwendet, meldet der zweite Befehl die folgenden Fehlernachrichten, die ignoriert werden können:

```
Error executing: set global innodb_large_prefix=ON
java.sql.SQLException: Unknown system variable 'innodb_large_prefix'
Error executing: set global innodb_file_format=BARRACUDA
java.sql.SQLException: Unknown system variable 'innodb_file_format'
Error executing: set global innodb_file_format_max=BARRACUDA
java.sql.SQLException: Unknown system variable 'innodb_file_format_max'
Error executing: set global innodb_file_per_table=TRUE
java.sql.SQLException: Variable 'innodb_file_per_table' is a read only variable
```

8. Führen Sie den folgenden Befehl aus, um die cf-Bibliothek zu entpacken.

```
cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration
unzip cf.zip
```

9. Konfigurieren Sie den Klassenpfad für JAAS-Anmeldemodule, indem Sie eine Datei mit dem Namen private_library.xml im Pfad <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver erstellen und die folgenden Informationen in die Datei eingeben:

```
<server>
 <library id="maprLib">
  <fileset dir="${wlp.install.dir}/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib" includes="*.jar"/>
  <fileset dir="/usr/share/java" includes="*.jar"/>
  <folder dir="/opt/mapr/hadoop/hadoop-2.7.0/etc/hadoop"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/common" includes="*.jar"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/common/lib" includes="*.jar"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/hdfs" includes="*.jar"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/hdfs/lib" includes="*.jar"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/yarn" includes="*.jar"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/yarn/lib" includes="*.jar"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/mapreduce" includes="*.jar"/>
  <fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/mapreduce/lib" includes="*.jar"/>
 </library>
 <jaasLoginModule id="maprLoginModule1" className="org.apache.hadoop.security.login.GenericOSLoginModule"
  controlFlag="REQUIRED" libraryRef="maprLib"></jaasLoginModule>
 <jaasLoginModule id="maprLoginModule2" className="org.apache.hadoop.security.login.HadoopLoginModule"
  controlFlag="REQUIRED" libraryRef="maprLib"></jaasLoginModule>
 <jaasLoginContextEntry id="hadoop_simple" name="hadoop_simple" loginModuleRef="maprLoginModule1,maprLoginModule2" />
 <application context-root="/analyticserver" id="AS_BOOT" location="AE_BOOT.war" name="AS_BOOT" type="war">
   <classloader commonLibraryRef="maprLib"></classloader>
 </application>
 <application id="help" location="help.war" name="help" type="war" context-root="/analyticserver/help"/>
</server>
```

Anmerkung: Das vorherige Beispiel gilt für die Konfiguration des Anmeldemoduls hadoop_simple. Die Konfiguration muss geändert werden, wenn MapR andere Anmeldemodule verwendet.

10. Prüfen Sie, ob die Datei ASModules.xml im Pfad <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/ vorhanden ist. Ist die Datei nicht vorhanden, benennen Sie die Datei ASModules.xml.template (im selben Pfad) in ASModules.xml um.
11. Konfigurieren Sie die Clusterinformationen, indem Sie die folgenden Eigenschaften in der Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties hinzufügen.

```
ae.cluster.zookeeper.connect.string=
ae.cluster.member.name=
ae.cluster.collective.name=mapr_5.1
```

Die Eigenschaft ae.cluster.zookeeper.connect.string ist die durch Kommas getrennte Zookeeper-Knotenliste. Die Eigenschaft kann denselben Zookeeper-Cluster wie MapR verwenden. ae.cluster.member.name ist der Hostname des Knotens, der Analytic Server hostet.

Das folgende Beispiel veranschaulicht das Format von ae.cluster.zookeeper.connect.string:

```
ae.cluster.zookeeper.connect.string=<Zookeeper-Host1>:<Zookeeper-Port1>,<Zookeeper-Host2>:<Zookeeper-Port2>,<Zookeeper-Host3>:
<Zookeeper-Port3>...
```

Wenn Analytic Server den gleichen Zookeeper-Cluster mit MapR teilt, muss der Wert von ae.cluster.zookeeper.connect.string der gleiche sein wie für die Eigenschaft zookeeper.servers in der MapR-Datei warden.conf (Dateistandardposition: /opt/mapr/conf).

12. Öffnen Sie die Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/server.env und fügen Sie ihr die folgenden Zeilen hinzu:

```
JAVA_HOME=<Java-Ausgangsverzeichnis>

PATH=<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<Java-Ausgangsverzeichnis>/jre/lib/amd64:/usr/sbin:/
usr/bin:/sbin:/bin

IBM_SPSS_AS_NATIVE_PATH=<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64

LD_LIBRARY_PATH=<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:
<Java-Ausgangsverzeichnis>/jre/lib/amd64:/opt/mapr/hadoop/hadoop-2.7.0/lib/native
```

    Ersetzen Sie <AS-Installationspfad> und <Java-Ausgangsverzeichnis> durch den tatsächlichen Installationspfad und den Java-Ausgangspfad.

13. Bearbeiten Sie das Analytic Server-Stammverzeichnis, indem Sie die Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties öffnen und die folgende Zeile hinzufügen:

```
distrib.fs.root=<AS-Stammverzeichnis>
```

    <AS-Stammverzeichnis> ist ein Pfad im MapR-Dateisystem, der die erforderlichen fernen Dateien von Analytic Server enthält. Der empfohlene Pfad ist /user/<as_user>/analytic-root.

14. Legen Sie den Benutzer mit Administratorberechtigung fest, indem Sie die Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties öffnen und die folgende Zeile hinzufügen:

```
admin.username=admin
```

    Der Wert muss ein Benutzername eines Analytic Server-Administrators sein, der als einer der Benutzer in der Datei security_cfg.xml konfiguriert ist.

15. Laden Sie Analytic Server-Abhängigkeiten in das MapR-Dateisystem hoch, indem Sie die folgende Zeile in Zeile 69 in der Datei <AS-Installationspfad>/bin/hdfsUpdate.sh hinzufügen:

```
JAVA_CLASS_PATH='hadoop classpath':$JAVA_CLASS_PATH
```

    Führen Sie die folgenden Befehle aus, um das <AS-Stammverzeichnis> zu erstellen:

```
cd <AS-Installationspfad>/bin
./hdfsUpdate.sh
```

    <as_user> muss eine Schreibberechtigung für das übergeordnete Verzeichnis <AS-Stammverzeichnis> haben.

16. Starten und stoppen Sie Analytic Server.
    a. Führen Sie den folgenden Befehl aus, um Analytic Server zu starten:

```
cd <AS-Installationspfad>/ae_wlpserver/bin
./server start aeserver
```

    b. Führen Sie den folgenden Befehl aus, um Analytic Server zu stoppen:

```
cd <AS-Installationspfad>/ae_wlpserver/bin
./server stop aeserver
```

Installieren von Analytic Server 3.1.0 unter MapR 5.2

1. Navigieren Sie zur IBM Passport Advantage®-Web-Site und laden Sie die selbstextrahierende MapR-Binärdatei herunter.

Tabelle 9. Selbstextrahierende MapR-Binärdateien

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.1.0.0 für MapR 5.0, 5.1 und 5.2 Linux x86-64 (Englisch)	spss_as3.1.0.0-mapr5-5.2-lx86_en.bin

2. Die verbleibenden Schritte für die Installation von Analytic Server sind mehr oder weniger die gleichen wie bei der Installation von Analytic Server 3.1.0 unter MapR 5.0 oder 5.1. Die Informationen zum „Aktivieren von Apache HBase" auf Seite 52 und „Aktivieren von Apache Spark" auf Seite 53 unterscheiden sich jedoch zwischen MapR 5.1 und 5.2. Informationen zum Installieren unter MapR 5.2 finden Sie in diesen Abschnitten.

Konfigurieren von MapR

Nach der Installation können Sie optional MapR-Funktionen für Analytic Server konfigurieren und verwalten.

Aktivieren von Datenbank-Pushback

Datenbank-Pushback bezeichnet das Lesen von Daten aus einer Datenbank und das direkte Verarbeiten der Daten.

IBM SPSS Analytic Server unterstützt Pushback für die folgenden Datenbanken:

- DashDB
- DB2
- DB2 for Z
- Hive
- MySQL
- Netezza
- Oracle
- PostgreSQL
- Redshift
- SQL Server
- Sybase IQ
- Teradata

Führen Sie die folgenden Schritte aus, um Datenbank-Pushback zu aktivieren.

1. Kopieren Sie die entsprechenden JDBC-Treiber-JAR-Dateien in das Verzeichnis `<AS-Installationspfad>/jdbc`.
2. Öffnen Sie die Datei `<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/private_library.xml`, suchen Sie die Tagbibliothek mit der ID `maprLib` und fügen Sie im Tag die folgende Zeile hinzu:

   ```
   <fileset dir="<AS-Installationspfad>/jdbc" includes="*.jar"/>
   ```

3. Führen Sie die folgenden Befehle aus:

   ```
   cd <AS-Installationspfad>/jdbc
   hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
   ```

4. Starten Sie Analytic Server erneut.

Aktivieren von Apache Hive

Apache Hive ist eine Data-Warehouse-Infrastruktur, die auf Hadoop aufbaut und die Zusammenfassung, Abfrage und Analyse von Daten ermöglicht.

Anmerkung: Hive muss für die Verwendung von MySQL als Metaspeicher konfiguriert sein. Die Datei `hive-site.xml`, die sich auf dem Knoten befindet, der IBM SPSS Analytic Server hostet, muss identisch mit der Datei auf dem Knoten sein, auf dem der Hive-Metaspeicher ausgeführt wird.

So aktivieren Sie die Unterstützung für Apache Hive nach einer erfolgreichen Installation von MapR:

1. Laden Sie die Hive- und hcatalog-Abhängigkeiten in das MapR-Dateisystem hoch, indem Sie die folgenden Befehle ausführen:

   ```
   cd /opt/mapr/hive/hive-1.2/lib
   hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
   cd /opt/mapr/hive/hive-1.2/hcatalog/share/hcatalog
   hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
   ```

   `<AS-Stammverzeichnis>` ist der in „Installieren von Analytic Server in MapR" auf Seite 47 definierte Pfad des Analytic Server-Stammverzeichnisses.

2. Öffnen Sie die Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/private_library.xml, suchen Sie die Tagbibliothek mit der ID maprLib und fügen Sie im Tag die folgenden Zeilen hinzu:

```
<fileset dir="/opt/mapr/hive/hive-1.2/lib" includes="*.jar"/>
<fileset dir="/opt/mapr/hive/hive-1.2/hcatalog/share/hcatalog" includes="*.jar"/>
```

3. Führen Sie die folgenden Befehle aus, um Hive- und hcatalog-Konfigurationsdateilinks zu erstellen:

```
mkdir <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/hive-conf
ln -s /opt/mapr/hive/hive-1.2/conf/* <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/hive-conf
```

4. Fügen Sie der Datei private_library.xml die folgende Zeile hinzu, wenn sich zusätzliche JAR-Dateien im Verzeichnis auxlib für Hive befinden:

```
<fileset dir="/opt/mapr/hive/hive-1.2/auxlib" includes="*.jar"/>
```

   Führen Sie die folgenden Befehle aus, nachdem Sie die vorherige Zeile hinzugefügt haben:

```
cd /opt/mapr/hive/hive-1.2/auxlib
hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
```

5. Starten Sie Analytic Server erneut.

Ausführen von Hive im HTTP-Modus

Hive wird standardmäßig im Binärmodus (TCP-Modus) ausgeführt. Sie müssen die folgenden Hive-Konfigurationseigenschaften (vor allem die Eigenschaft hive.server2.transport.mode) aktualisieren, um Hive im HTTP-Modus auszuführen.

Anmerkung: Weitere Informationen zu jeder Eigenschaft finden Sie in Hive-Konfigurationseigenschaften.

Tabelle 10. Hive-Eigenschaften für HTTP-Modus

Eigenschaftsname	Standardwert	Beschreibung
hive.server2.transport.mode	binary	Der Servertransportmodus. Der Wert kann binary oder http sein. Setzen Sie ihn auf http, um den HTTP-Transportmodus zu aktivieren.
hive.server2.thrift.http.port	10001	Die Portnummer im HTTP-Modus.
hive.server2.thrift.http.path	cliservice	Die Pfadkomponente des URL-Endpunkts im HTTP-Modus.
hive.server2.thrift.http.min.worker.threads	5	Die minimale Anzahl Worker-Threads im Server-Pool im HTTP-Modus.
hive.server2.thrift.http.max.worker.threads	500	Die maximale Anzahl Worker-Threads im Server-Pool im HTTP-Modus.

Anmerkung: Hive muss nach der Aktualisierung der Eigenschaften erneut gestartet werden.

Aktivieren von Apache HBase

Apache HBase ist eine in Java geschriebene, nicht relationale verteilte Open-Source-Datenbank. Apache HBase wurde als Teil des Projekts Apache Hadoop von Apache Software Foundation entwickelt und setzt auf HDFS (Hadoop Distributed Filesystem) auf.

So aktivieren Sie die Unterstützung für Apache HBase nach einer erfolgreichen Installation von MapR:

IBM SPSS Analytic Server 3.1.0 on MapR 5.0/5.1

1. Laden Sie die HBase-Abhängigkeiten in das MapR-Dateisystem hoch und führen Sie die folgenden Befehle aus:

```
cd /opt/mapr/hbase/hbase-0.98.12/lib
hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
```

   <AS-Stammverzeichnis> ist der in „Installieren von Analytic Server in MapR" auf Seite 47 definierte Pfad des Analytic Server-Stammverzeichnisses.

2. Öffnen Sie die Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/private_library.xml, suchen Sie die Tagbibliothek mit der ID maprLib und fügen Sie im Tag die folgende Zeile hinzu:

   ```
   <fileset dir="/opt/mapr/hbase/hbase-0.98.12/lib" includes="*.jar"/>
   ```

3. Führen Sie die folgenden Befehle aus, um HBase- und hcatalog-Konfigurationsdateilinks zu erstellen:

   ```
   mkdir <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/hbase-conf
   ln -s /opt/mapr/hbase/hbase-0.98.12/conf/* <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/
   configuration/hbase-conf
   ```

4. Starten Sie IBM SPSS Analytic Server erneut.

IBM SPSS Analytic Server 3.1.0 on MapR 5.2

1. Laden Sie die HBase-Abhängigkeiten in das MapR-Dateisystem hoch, indem Sie die folgenden Befehle ausführen:

   ```
   cd /opt/mapr/hbase/hbase-1.1.1/lib
   hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
   ```

   <AS-Stammverzeichnis> ist der Pfad, der in Schritt 12 in „Installieren von Analytic Server in MapR" auf Seite 47 festgelegt wurde.

2. Öffnen Sie <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/private_library.xml und suchen Sie die Tagbibliothek mit der ID maprLib. Fügen Sie dem Tag die folgende Zeile hinzu:

   ```
   <fileset dir="/opt/mapr/hbase/hbase-1.1.1/lib" includes="*.jar"/>
   ```

3. Führen Sie die folgenden Befehle aus, um Links für die Hive- und HCatalog-Konfigurationsdateien zu erstellen:

   ```
   mkdir <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/hbase-conf
   ln -s /opt/mapr/hbase/hbase-1.1.1/conf/* <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/
   configuration/hbase-conf
   ```

4. Fügen Sie <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties die folgende Zeile hinzu:

   ```
   spark.executor.extraClassPath=/opt/mapr/hbase/hbase-1.1.1/lib/*
   ```

5. Starten Sie Analytic Server erneut.

Aktivieren von Apache Spark

Apache Spark ist ein offener Standard für flexible speicherinterne Datenverarbeitung für Stapel- und Echtzeitanalysen sowie erweiterte Analysen.

So aktivieren Sie die Unterstützung für Apache Spark nach einer erfolgreichen Installation von MapR:

IBM SPSS Analytic Server 3.1.0 on MapR 5.0/5.1

1. Kopieren Sie die Datei spark-assembly-1.4.1-hadoop2.5.1-mapr-1501.jar aus /opt/mapr/spark/spark-1.4.1/lib in das Verzeichnis <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/spark/.
2. Laden Sie die Spark-Abhängigkeiten in das MapR-Dateisystem hoch und führen Sie die folgenden Befehle aus:

   ```
   cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/spark/
   hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
   ```

   <AS-Stammverzeichnis> ist der in „Installieren von Analytic Server in MapR" auf Seite 47 definierte Pfad des Analytic Server-Stammverzeichnisses.

3. Öffnen Sie die Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/private_library.xml, suchen Sie die Tagbibliothek mit der ID maprLib und fügen Sie im Tag die folgende Zeile hinzu:

   ```
   <fileset dir="/opt/mapr/spark/spark-1.4.1/lib" includes="spark-assembly-*.jar"/>
   ```

4. Führen Sie die folgenden Befehle aus, um Spark-Konfigurationsdateilinks zu erstellen:

```
mkdir <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf
ln -s /opt/mapr/spark/spark-1.4.1/conf/* <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf
```

5. Fügen Sie in der Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/server.env die folgende Zeile hinzu:

   ```
   SPARK_HOME=/opt/mapr/spark/spark-1.4.1
   ```

6. Fügen Sie in der Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties die folgende Zeile hinzu:

   ```
   spark.executor.extraLibraryPath=/opt/mapr/hadoop/hadoop-2.7.0/lib/native
   ```

7. Starten Sie IBM SPSS Analytic Server erneut.
8. Zum Aktivieren der PySpark-Funktion fügen Sie in der Datei yarn-env.sh die folgende Zeile hinzu und starten Sie Ressourcen- und Knotenmanager erneut:

   ```
   export SPARK_HOME=/opt/mapr/spark/spark-1.4.1
   ```

IBM SPSS Analytic Server 3.1.0 on MapR 5.2

Die Schritte sind je nach Spark-Version unterschiedlich.

Spark 1.x

1. Kopieren Sie die Datei 'spark-assembly-1.4.1-hadoop2.5.1-mapr-1501.jar' von:

   ```
   /opt/mapr/spark/spark-1.4.1/lib
   ```

   in

   ```
   <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/spark/
   ```

2. Öffnen Sie die folgende Datei:

   ```
   <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/private_library.xml
   ```

   Lokalisieren Sie die Tagbibliothek mit der ID maprLib. Fügen Sie dem Tag die folgende Zeile hinzu:

   ```
   <fileset dir="/opt/mapr/spark/spark-1.4.1/lib" includes="spark-assembly-*.jar"/>
   ```

3. Löschen Sie die folgende Datei:

   ```
   <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/spark/com.ibm.spss.
   sparkmapreduce_2-3.1.0.0.jar
   ```

Spark 2.x

1. Löschen Sie die folgende Datei:

   ```
   <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/spark/com.ibm.spss.sparkmapreduce
   -3.1.0.0.jar
   ```

2. Öffnen Sie die folgende Datei:

   ```
   <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/private_library.xml
   ```

   Lokalisieren Sie die Tagbibliothek mit der ID maprLib. Fügen Sie dem Tag die folgenden Zeilen hinzu:

   ```
   <fileset dir="/opt/mapr/spark/spark-2.0.1/jars" includes="*.jar"/>
   <fileset dir="/opt/mapr/spark/spark-2.0.1/scala/lib" includes="*.jar"/>
   <fileset dir="<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/
   spark" includes="*.jar"/>
   ```

3. Fügen Sie der folgenden Datei die Zeile spark.version=2.0 hinzu:

   ```
   <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/
   configuration/config.properties
   ```

Die folgenden Schritte gelten für Spark 1.x und 2.x und müssen nach den vorherigen Schritten für Spark 1.x oder 2.x ausgeführt werden.

Anmerkung: Alle Referenzen auf <Spark-Version> müssen durch die tatsächliche Spark-Version ersetzt werden (z. B. 1.4.1 oder 2.0.1).

1. Laden Sie die Spark-Abhängigkeiten in das MapR-Dateisystem hoch, indem Sie die folgenden Befehle ausführen:

   ```
   cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/spark/
   hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
   ```

   <AS-Stammverzeichnis> ist der Pfad, der in Schritt 12 in „Installieren von Analytic Server in MapR" auf Seite 47 festgelegt wurde.
2. Führen Sie die folgenden Befehle aus, um Links für die Spark-Konfigurationsdatei zu erstellen:

   ```
   mkdir <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf
   ln -s /opt/mapr/spark/spark-<Spark-Version>/conf/*
   <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf
   ```
3. Fügen Sie der Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/server.env die folgende Zeile hinzu:

   ```
   SPARK_HOME=/opt/mapr/spark/spark-<Spark-Version>
   ```
4. Fügen Sie in der Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties die folgende Zeile hinzu:

   ```
   spark.executor.extraLibraryPath=/opt/mapr/hadoop/hadoop-2.7.0/lib/native
   ```
5. Starten Sie Analytic Server erneut.
6. Wenn Sie die PySpark-Funktion aktivieren müssen, fügen Sie der Datei yarn-env.sh die folgende Zeile hinzu:

   ```
   export SPARK_HOME=/opt/mapr/spark/spark-<Spark-Version>
   ```

   Starten Sie Ressourcen- und Knotenmanager erneut.

Aktivieren von Funktionsflags

Funktionsflags stellen die Fähigkeit zum Aktivieren und Inaktivieren bestimmter Anwendungsfeatures bereit.

So aktivieren Sie die Unterstützung für Funktionsflags nach einer erfolgreichen Installation von MapR:

1. Fügen Sie in der Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties die folgende Zeile hinzu:

   ```
   load.feature.flags.on.msg=true
   ```
2. Starten Sie IBM SPSS Analytic Server erneut.

Aktivieren von R

R ist eine Sprache und Umgebung für statistische Datenverarbeitung und Grafiken.

So aktivieren Sie die Unterstützung für R nach einer erfolgreichen Installation von MapR:

Anmerkung: Das folgende Paket muss installiert sein, bevor Sie das Installationsprogramm auf allen Clusterknoten ausführen können, die Node Manager und IBM SPSS Analytic Server hosten:

```
gcc-gfortran
libgfortran
gcc-c++
```

1. Führen Sie das Installationsprogramm spss_er-8.4.0.0-mapr5-lx86_64_en.bin auf allen Clusterknoten aus, die Node Manager und Analytic Server hosten. Der Benutzer, der das Installationsprogramm ausführt, muss Schreibberechtigung für die Installationspfade von R und Analytic Server haben.
2. Befolgen Sie die Installationsanweisungen, indem Sie die Lizenzvereinbarung akzeptieren und die erforderlichen Informationen eingeben. Wenn Analytic Server auf dem Installationsserver installiert ist, wählen Sie Ja aus, wenn Sie dazu aufgefordert werden, und geben Sie den <AS-Installationspfad> ein. Wenn Analytic Server nicht auf dem Installationsserver installiert ist, wählen Sie Nein aus, wenn Sie dazu aufgefordert werden.

3. Wenn Analytic Server installiert ist, wird Essentials for R automatisch im Installationspfad von Analytic Server installiert.
   - Wenn Analytic Server nicht installiert ist, wird Essentials for R im Pfad <Installationsprogrammpfad>/IBM_SPSS_ModelerEssentialsR/linux installiert.
   - Wenn Analytic Server später installiert wird, verwenden Sie den folgenden Befehl, um Essentials for R in den Konfigurationspfad von Analytic Server, in dem Analytic Server installiert wird.

     ```
     cp -r <Installationsprogrammpfad>/IBM_SPSS_ModelerEssentialsR/linux <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/
     configuration
     ```

4. Löschen Sie die Datei cf.zip im Pfad <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration und generieren Sie eine neue Datei mit dem folgenden Befehl:

   ```
   cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration
   zip -r cf.zip linux
   ```

5. Führen Sie die folgenden Befehle aus:

   ```
   cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration
   hadoop fs -rm <AS-Stammverzeichnis>/cluster1/configuration/cf.zip
   hadoop fs -put cf.zip <AS-Stammverzeichnis>/cluster1/configuration/
   ```

6. Starten Sie Analytic Server erneut.

Aktivieren von LZO

LZO ist eine verlustfreie Datenkomprimierungsbibliothek, bei der die Geschwindigkeit wichtiger als das Komprimierungsverhältnis ist. MapR muss manuell konfiguriert werden, damit es Unterstützung für LZO bereitstellt.

Auf der Site https://github.com/twitter/hadoop-lzo finden Sie Anweisungen zur Installation und Konfiguration von LZO.

In den folgenden Schritten wird der Prozess des Imports einer LZO-Bibliothek in MapR beschrieben.

1. Kopieren Sie die Datei hadoop-lzo-<Version>.jar in den Hadoop-Klassenpfad. Der empfohlene Pfad ist /opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/mapreduce/lib.
2. Kopieren Sie die nativen Dateien libgplcompression.so und liblzo2.so.2 in das Verzeichnis /opt/mapr/hadoop/hadoop-2.7.0/lib/native und fügen Sie der Datei core-site.xml die folgenden Eigenschaften hinzu:

   ```
   <property>
      <name>io.compression.codecs</name>
      <value>org.apache.hadoop.io.compress.GzipCodec,org.apache.hadoop.io.compress.DefaultCodec,com.hadoop.compression.lzo.
       LzoCodec,com.hadoop.compression.lzo.LzopCodec,org.apache.hadoop.io.compress.BZip2Codec</value>
   </property>
   <property>
       <name>io.compression.codec.lzo.class</name>
       <value>com.hadoop.compression.lzo.LzoCodec</value>
   </property>
   ```

3. Öffnen Sie die Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/server.env und fügen Sie <Pfad_native_LZO-Bibliothek> zum Parameter LD_LIBRARY_PATH hinzu. <Pfad_native_LZO-Bibliothek> ist der Ordner, der die native Hadoop-LZO-Bibliothek enthält.

   ```
   LD_LIBRARY_PATH=<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<Java-Ausgangsverzeichnis>/
   jre/lib/amd64:
   /opt/mapr/hadoop/hadoop-2.7.0/lib/native:<Pfad_native_LZO-Bibliothek>
   ```

4. Starten Sie IBM SPSS Analytic Server erneut.

Einrichten eines IBM SPSS Analytic Server-Clusters für MapR

Führen Sie die folgenden Schritte aus, um eine IBM SPSS Analytic Server-Clusterumgebung für MapR-Unterstützung einzurichten.

1. Fügen Sie in der Datei <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties die folgende Zeile hinzu.

   ```
   enable.resume=true
   ```

2. Kopieren Sie den Installationspfad auf die anderen Clusterknoten und ändern Sie die Eigenschaft ae.cluster.member.name in der Datei config.properties in den korrekten Hostnamen.

3. Starten Sie alle Clusterknoten.

Deinstallation von MapR

In den folgenden Schritten wird der Prozess der Deinstallation von MapR erläutert:

1. Stoppen Sie IBM SPSS Analytic Server.
2. Löschen Sie die Metadatendatenbank.
   a. Führen Sie die folgenden Befehle aus:

      ```
      cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/sql/<DB-Typ>
      java -Xmx128m -Xms128m -cp <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/*:/usr/
      share/java/* com.spss.ae.dbscript.ScriptRunner ../../configuration/config.properties drop.sql true
      ```

   b. Führen Sie die folgende SQL-Anweisung aus, um die Datenbank zu löschen:

      ```
      drop database <DB-Name>
      ```

3. Deinstallieren Sie das RPM-Paket:

   ```
   rpm -e IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64
   ```

4. Löschen Sie den Installationspfad:

   ```
   rm -r <AS-Installationspfad>
   ```

5. Löschen Sie das AS-Stammverzeichnis:

   ```
   hadoop fs -rm -r <AS-Stammverzeichnis>
   ```

6. Löschen Sie die Zookeeper-Daten:

   ```
   /opt/mapr/zookeeper/zookeeper-3.4.5/bin/zkCli.sh -server <Zookeeper-Host>:<Zookeeper-Port>
   rmr /AnalyticServer
   ```

Migration von IBM SPSS Analytic Server in MapR

IBM SPSS Analytic Server kann in MapR migriert werden.

Führen Sie die folgenden Schritte aus, um IBM SPSS Analytic Server 2.0 oder 2.1 auf Version 3.1.0 in MapR zu migrieren.

1. Installieren Sie Analytic Server 3.1.0 in einem MapR-Cluster, indem Sie die Installationsanweisungen in „Installieren von Analytic Server in MapR" auf Seite 47 befolgen.
2. Kopieren Sie das Analytic Server-Stammverzeichnis.

   Anmerkung: Dieser Schritt kann ignoriert werden, wenn das Analytic Server-Stammverzeichnis nicht geändert wurde.

   - Führen Sie den folgenden Befehl auf einem der Datenknoten aus, wenn sich die Analytic Server-Stammverzeichnisse für die Analytic Server-Versionen 2.0/2.1 und 3.1.0 in demselben MapR-Cluster befinden:

     ```
     hadoop fs -cp <altes_AS-Stammverzeichnis>/analytic-workspace/* <neues_AS-Stammverzeichnis>/analytic-workspace
     ```

   - Die installierten WEBHDFS- oder NFS-Services bestimmen, wann sich die Analytic Server-Stammverzeichnisse für Analytic Server Version 2.0/2.1 und 3.1.0 in verschiedenen MapR-Clustern befinden. WEBHDFS oder NFS sind zum Kopieren der Analytic Server-Stammverzeichnisdaten erforderlich, da außerhalb des Clusters nicht direkt auf das MapR-Dateisystem zugriffen werden kann.
     a. Führen Sie den folgenden Befehl auf einem der neuen Clusterknoten von Analytic Server 3.1.0 aus, wenn der alte Cluster von Analytic Server 2.0/2.1 den WEBHDFS-Service enthält:

        ```
        hadoop distcp webhdfs://<WEBHDFS-Server>:<WEBHDFS-Port>/<altes_AS-Stammverzeichnis>/
        analytic-workspace/* maprfs://<neues_AS-Stammverzeichnis>/
        analytic-workspace
        ```

     b. Führen Sie den folgenden Befehl auf einem der alten Clusterknoten von Analytic Server 2.0/2.1 aus, wenn der alte Cluster von Analytic Server 3.1.0 den WEBHDFS-Service enthält:

        ```
        hadoop distcp maprfs://<altes_AS-Stammverzeichnis>/
        analytic-workspace/*
        webhdfs://<WEBHDFS-Server>:<WEBHDFS-Port>/<neues_AS-Stammverzeichnis>/
        analytic-workspace
        ```

c. Führen Sie den folgenden Befehl auf einem der alten Clusterknoten von Analytic Server 2.0/2.1 aus, wenn der alte Cluster NFS enthält und NFS auch an einem der neuen Clusterknoten von Analytic Server 3.1.0 angehängt ist:

```
hadoop distcp file:///<Mountpfad>/<altes_AS-Stammverzeichnis>/analytic-workspace/* maprfs://<neues_AS-Stammverzeichnis>/
analytic-workspace
```

d. Führen Sie den folgenden Befehl auf einem der neuen Clusterknoten von Analytic Server 3.1.0 aus, wenn der neue Cluster NFS enthält und NFS auch an einem der alten Clusterknoten von Analytic Server 2.0/2.1 angehängt ist:

```
hadoop discp maprfs://<altes_AS-Stammverzeichnis>/analytic-workspace/* file:///<Mountpfad>/<neues_AS-Stammverzeichnis>/
analytic-workspace
```

Informationen zum Migrieren von Daten zwischen verschiedenen MapR-Clustern finden Sie auf der MapR-Site zur Datenmigration.

3. Führen Sie die folgenden Befehle aus, um den Eigner und die Berechtigungen für das neue Analytic Server-Stammverzeichnis zu ändern:

```
hadoop fs -chown -R <as_user> <AS-Stammverzeichnis>
hadoop fs -chmod -R 755 <>
```

4. Stoppen Sie Analytic Server 3.1.0, aber stellen Sie sicher, dass die Metadatendatenbank noch ausgeführt wird.
5. Erfassen Sie die Konfigurationseinstellungen der alten Clusterinstallation von Analytic Server 2.0/2.1.
   a. Kopieren Sie das Archiv `configcollector.zip` aus der neuen Clusterinstallation von Analytic Server 3.1.0 in das Verzeichnis `<alter_AS-Installationspfad>/tools` der alten Clusterinstallation von Analytic Server 2.0/2.1.
   b. Extrahieren Sie den Inhalt von `configcollector.zip` in der alten Clusterinstallation von Analytic Server 2.0/2.1. Ein neues Unterverzeichnis `configcollector` wird in der alten Clusterinstallation von Analytic Server 2.0/2.1 erstellt.
   c. Führen Sie das Konfigurations-Collector-Tool in der alten Clusterinstallation von Analytic Server 2.0/2.1 aus, indem Sie das Script `configcollector` im Verzeichnis `<alter_AS-Installationspfad>/tools/configcollector` ausführen. Kopieren Sie die resultierende komprimierte Datei (ZIP-Datei) in die neue Clusterinstallation von Analytic Server 3.1.0.
6. Führen Sie das Migrationstool für den neuen Cluster von Analytic Server 3.1.0 aus, indem Sie das Script `migrationtool` ausführen und den Pfad der vom Konfigurationscollector erstellten komprimierten Datei als Argument übergeben. Beispiel:

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.1/ASConfiguration_2.1.0.0.xxx.zip
```

7. Starten Sie Analytic Server 3.1.0.

MapR - Fehlerbehebung

In diesem Abschnitt werden einige allgemeine Probleme bei der MapR-Installation und -Konfiguration sowie Wege zu deren Lösung beschrieben.

Probleme mit dem Script hdfsUpdate.sh

Das Script `hdfsUpdate.sh` muss nur einmal ausgeführt werden, da es alle Dateien im AS-Stammverzeichnis löscht, bevor es neue Dateien hochlädt. Wird das Script mehrmals ausgeführt, müssen Sie die Abhängigkeiten für Datenbank-Pushback, Hive, HBase und Spark erneut hochladen. Führen Sie den folgenden Befehl aus, um die erforderlichen Abhängigkeiten erneut hochzuladen:

```
cd <AS-Installationspfad>/jdbc

hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath

cd /opt/mapr/hive/hive-1.2/lib
hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
cd /opt/mapr/hive/hive-1.2/hcatalog/share/hcatalog
hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath

cd /opt/mapr/hbase/hbase-0.98.12/lib
hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
```

```
cd <AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver/modules/spark/

hadoop fs -put *.jar <AS-Stammverzeichnis>/cluster1/classpath
```

Ein Konflikt zwischen MapR- und Spark-Versionen führte zu einer fehlgeschlagenen Spark-Jobausführung

Zwischen MapR und Spark (1.6.1) tritt ein Klassenkonfliktproblem auf, wenn die MapR-Version 5.1 oder höher ist. Der Konflikt führt zu einer fehlgeschlagenen Spark-Jobausführung. Sie können das Problem beheben, indem Sie die Datei `private_library.xml` in `<AS-Installationspfad>/ae_wlpserver/usr/servers/aeserver` ändern. Das folgende Beispiel gibt die erforderliche Änderung an:

```
......
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/common/lib" includes="*.jar" excludes="jackson-databind-*.jar" />
......
```

Kapitel 5. Huawei FusionInsight HD-Installation und -Konfiguration

FusionInsight HD - Übersicht

Huawei FusionInsight HD stellt eine umfassende Big Data-Softwareplattform für Stapel- und Echtzeitanalysen unter Verwendung der quelloffenen Hadoop- und Spark-Technologien bereit. Das System nutzt HDFS, HBase, MapReduce und YARN/Zookeeper für Hadoop-Clustering sowie Apache Spark für schnellere Echtzeitanalysen und interaktive Abfragen.

Analytic Server kann auf der Plattform FusionInsight HD ausgeführt werden. FusionInsight 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.

Installation in Huawei FusionInsight HD

In den folgenden Schritten wird der Prozess der manuellen Installation von IBM SPSS Analytic Server in Huawei FusionInsight HD erläutert.

Analytic Server 3.1.0

1. Navigieren Sie zur IBM Passport Advantage®-Website und laden Sie die folgende selbstextrahierende Binärdatei auf einen Host innerhalb des FusionInsight HD-Clusters herunter.

Tabelle 11. Selbstextrahierende Analytic Server-Binärdatei

Beschreibung	Name der Binärdatei
IBM SPSS Analytic Server 3.1 für FusionInsight HD 2.6 Linux x86-64 (Englisch)	`spss_as-3.1-fhd2.6-lx86_en.bin`

2. Führen Sie das selbstextrahierende Installationsprogramm `*.bin` auf dem FusionInsight Manager-Master-Clusterknoten aus. Befolgen Sie die Eingabeaufforderungen bei der Installation, indem Sie die Lizenzvereinbarung akzeptieren und das Standardinstallationsverzeichnis beibehalten. Das Installationsprogramm lädt die erforderlichen RPM-Dateien herunter und muss auf einem Computer ausgeführt werden, der auf https://ibm-open-platform.ibm.com zugreifen kann. Die ausführbare Binärdatei befindet sich im verfügbaren FusionInsight HD-Verteilerverzeichnis `<installierbares_AS-Ausgangsverzeichnisses>`.
3. Installieren Sie Analytic Server 3.1.0 mit dem folgenden Befehl:

```
# yum install -y IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm
```

4. Melden Sie sich mit `omm` an und erstellen Sie die Datei `analyticserver.keytab`:

```
# su omm
# source /opt/huawei/Bigdata/om-0.0.1/meta-0.0.1-SNAPSHOT/kerberos/scripts/component_env
# kadmin -p kadmin/admin
```

   Das `kadmin`-Standardkennwort lautet `Admin@123`. Sie müssen das Kennwort bei der ersten Verwendung ändern. Ersetzen Sie in den folgenden Befehlen `_HOST` durch den Namen Ihres Hosts.

```
kadmin > addprinc -randkey omm/_HOST@HADOOP.COM
kadmin > ktadd -k  /opt/ibm/spss/analyticserver/3.1/analyticserver.keytab HTTP/_HOST@HADOOP.COM
kadmin > ktadd -k /opt/ibm/AnalyticServer/analyticserver.keytab omm/_HOST@HADOOP.COM
```

5. Installieren Sie MYSQL und erstellen Sie die Datenbank `aedb` manuell. Beispiel:

```
# cd /etc/yum.repos.d
# wget http://dev.mysql.com/get/mysql57-community-release-el7-9.noarch.rpm
# yum -y install mysql57-community-release-el7-9.noarch.rpm
# yum repolist all | grep mysql
```

```
# yum -y install mysql-community-server
# yum install -y mysql-connector-java
# systemctl enable mysqld.service
# systemctl start mysqld.service
```

Rufen Sie das MYSQL-Rootbenutzerkennwort ab:

```
# grep 'temporäres Kennwort' /var/log/mysqld.log
# mysql -uroot -p
# MySQL> set global validate_password_policy=0;
# MySQL> DROP DATABASE IF EXISTS aedb;
# MySQL> CREATE DATABASE aedb  DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8_bin;
# MySQL> CREATE USER 'aeuser'@'%' IDENTIFIED BY 'Pass1234';
# MySQL> CREATE USER 'aeuser'@'localhost' IDENTIFIED BY 'Pass1234';
# MySQL> GRANT ALL PRIVILEGES ON *.* TO 'aeuser'@'%';
# MySQL>  GRANT ALL PRIVILEGES ON *.* TO 'aeuser'@'localhost';
```

6. Erstellen Sie ein Script `install_as.sh` in `/opt` und führen Sie es mit dem Benutzer `omm` aus:

```
# chown -R omm:wheel /opt/ibm/*

# su omm
# /opt/install_as.sh

install_as.sh
```

`install_as.sh` enthält das folgende Script:

```
cd /opt/ibm/spss/analyticserver/3.1
mkdir hadoop
mkdir zookeeper
mkdir spark-client

cd /opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/
echo "JAVA_HOME=/opt/huawei/Bigdata/jdk/jre" > server.env
echo "PATH=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:/opt/huawei/Bigdata/jdk/
echo jre/lib/amd64:/usr/sbin:/usr/bin:/sbin:/bin" >> server.env
echo "IBM_SPSS_AS_NATIVE_PATH=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/
echo lib_64" >> server.env
echo "LD_LIBRARY_PATH=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:/opt/huawei/
echo Bigdata/jdk/jre/lib/amd64=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/native" >> server.env
echo "SPARK_HOME=/opt/ibm/spss/analyticserver/3.1/spark-client" >> server.env

echo "<server description=\"new server\">" > server.xml
echo "<!-- Enable features -->" >> server.xml
echo "  <featureManager>" >> server.xml
echo "    <feature>servlet-3.1</feature>" >> server.xml
echo "    <feature>jsp-2.3</feature>" >> server.xml
echo "    <feature>jdbc-4.0</feature>" >> server.xml
echo "    <feature>jndi-1.0</feature>" >> server.xml
echo "    <feature>localConnector-1.0</feature>" >> server.xml
echo "    <feature>jaxrs-2.0</feature>" >> server.xml
echo "    <feature>json-1.0</feature>" >> server.xml
echo "    <feature>appSecurity-2.0</feature>" >> server.xml
echo "    <feature>ldapRegistry-3.0</feature>" >> server.xml
echo "    <feature>restConnector-1.0</feature>" >> server.xml
echo "    <feature>monitor-1.0</feature>" >> server.xml
echo "    <feature>ssl-1.0</feature>" >> server.xml
echo "  </featureManager>" >> server.xml
echo "  <applicationManager startTimeout=\"120s\" />" >> server.xml
echo "  <executor name=\"LargeThreadPool\" id=\"default\" coreThreads=\"100\" keepAlive=\"60s\" stealPolicy=\"STRICT\
echo " rejectedWorkPolicy=\"CALLER_RUNS\" />" >> server.xml
echo "  <webContainer deferServletLoad=\"false\" disallowAllFileServing=\"false\" fileServingEnabled=\"true\" trusted=\"false\
echo " directoryBrowsingEnabled=\"false\" asyncTimeoutDefault=\"300000\"/>" >> server.xml
echo "  <classloading useJarUrls=\"true\"/>" >> server.xml
echo "  <applicationMonitor updateTrigger=\"mbean\" />" >> server.xml
echo "  <mimeTypes>" >> server.xml
echo "    <type>svg=image/svg+xml</type>" >> server.xml
echo "  </mimeTypes>" >> server.xml
echo "  <variable name=\"AE_DATABASE\" value=\"\${wlp.install.dir}/usr/servers/aeserver/aedb\" />" >> server.xml
echo "  <administrator-role>" >> server.xml
echo "    <user>admin</user>" >> server.xml
echo "  </administrator-role>" >> server.xml
echo "  <include optional=\"true\" location=\"\${server.config.dir}/private_library.xml\"/>" >> server.xml
echo "  <include optional=\"true\" location=\"\${server.config.dir}/http_endpoint.xml\"/>" >> server.xml
echo "  <include optional=\"true\" location=\"\${server.config.dir}/security_cfg.xml\"/>" >> server.xml
echo "  <include optional=\"true\" location=\"\${server.config.dir}/ssl_cfg.xml\"/>" >> server.xml
echo "  <include optional=\"true\" location=\"\${server.config.dir}/configuration/key.xml\" />" >> server.xml
echo "</server>" >> server.xml

touch http_endpoint.xml
echo "<server>" > http_endpoint.xml
echo "  <httpEndpoint host=\"*\" id=\"defaultHttpEndpoint\" httpPort=\"9080\" httpsPort=\"9443\" onError=\"FAIL\
echo "/>" >> http_endpoint.xml
echo "</server>" >> http_endpoint.xml

touch private_library.xml
echo "<server>" > private_library.xml
echo "  <application context-root=\"/analyticserver\" id=\"AS_BOOT\" location=\"AE_BOOT.war\" name=\"AS_BOOT\" type=\"war\">
```

```
echo " >> private_library.xml
echo "    <classloader>" >> private_library.xml
echo "      <privateLibrary>" >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib\" includes=\"*.jar\"/>
echo " >> private_library.xml
echo "        <fileset dir=\"/usr/share/java\" includes=\"*.jar\"/>" >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/../lib\" includes=\"*.jar\"/>" >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/../spark-client/lib\" includes=\"spark-assembly-*.jar\"/>" >> private_library.xml
echo "        <folder  dir=\"\${wlp.install.dir}/usr/servers/aeserver/configuration/hadoop-conf\"/>" >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/../jdbc\" includes=\"postgresql-*.jar\"/>
echo " >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/../jdbc\" includes=\"*.jar\"/
echo >" >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/../hive\" includes=\"*.jar\"/>" >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/../zookeeper\" includes=\"*.jar\"/>" >> private_library.xml
echo "        <fileset dir=\"\${wlp.install.dir}/../hadoop\" includes=\"*.jar\"/>" >> private_library.xml
echo "      </privateLibrary>" >> private_library.xml
echo "    </classloader>" >> private_library.xml
echo "  </application>" >> private_library.xml
echo "  <application id=\"help\" location=\"help.war\" name=\"help\" type=\"war\" context-root=\"/
echo analyticserver/help\"/> " >> private_library.xml
echo "</server>" >> private_library.xml

touch security_cfg.xml
echo "<server>" > security_cfg.xml
echo "  <basicRegistry id=\"basic\" realm=\"ibm\">" >> security_cfg.xml
echo "    <user name=\"admin\" password=\"admin\"/>" >> security_cfg.xml
echo "  </basicRegistry>" >> security_cfg.xml
echo "</server>" >> security_cfg.xml

touch jaas.conf
echo "Client {" > jaas.conf
echo "com.sun.security.auth.module.Krb5LoginModule required" >> jaas.conf
echo "keyTab=\"/opt/ibm/spss/analyticserver/3.1/analyticserver.keytab\"" >> jaas.conf
echo "principal=\"omm/huawei-1@HADOOP.COM\"" >> jaas.conf
echo "useKeyTab=true" >> jaas.conf
echo "useTicketCache=true" >> jaas.conf
echo "storeKey=true" >> jaas.conf
echo "debug=true;" >> jaas.conf
echo "};" >> jaas.conf

cd /opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration
echo "spark.version=1.x" > config.properties
echo "http.port=9080" >> config.properties
echo "https.port=9443" >> config.properties
echo "ae.cluster.zookeeper.connect.string=172.16.155.123:24002,172.16.155.212:24002,172.16.186.208:24002"
echo >> config.properties
echo "ae.cluster.member.name=huawei-1" >> config.properties
echo "ae.cluster.collective.name=Test_01" >> config.properties
echo "jndi.aedb=jdbc/aeds" >> config.properties
echo "jndi.aedb.url=jdbc:mysql://huawei-1/aedb?createDatabaseIfNotExist=true" >> config.properties
echo "jndi.aedb.username=aeuser" >> config.properties
echo "jndi.aedb.driver=com.mysql.jdbc.Driver" >> config.properties
echo "distrib.fs.root=/user/as_user/analytic-root" >> config.properties
echo "admin.username=admin" >> config.properties
echo "enable.resume=true" >> config.properties
echo "load.feature.flags.on.msg=true" >> config.properties
echo "jndi.aedb.password=FEFFUy9FQ0IvUEtDUzVQYWRkaW5nAGk3bIuya2BzXYeXyFcOrxo=" >> config.properties
echo "ae.kerberos.principal=omm/huawei-1@HADOOP.COM" >> config.properties
echo "hdfs.user=omm/huawei-1@HADOOP.COM" >> config.properties
echo "web.authentication.kerberos.principal=HTTP/huawei-1@HADOOP.COM" >> config.properties
echo "java.security.krb5.conf=/home/omm/kerberos/var/krb5kdc/krb5.conf" >> config.properties
echo "web.authentication.kerberos.keytab=/opt/ibm/spss/analyticserver/3.1/analyticserver.keytab" >> config.properties
echo "hdfs.keytab=/opt/ibm/spss/analyticserver/3.1/analyticserver.keytab" >> config.properties
echo "ae.db.connect.method=Kerberos" >> config.properties
echo "kdcrealm=HADOOP.COM" >> config.properties
echo "kdcserver=172.16.155.212:21732" >> config.properties
echo "encryption.keystore.password=FEFFUy9FQ0IvUEtDUzVQYWRkaW5nAMDJul7PVsvdIyLlzjeS8ws=" >> config.properties
echo "encryption.keystore.base64=zs7OzgAAAAIAAAABAAAAAwA6Y29tLnNwc3MuYWUuZW5jcnlwdGlvbi5wcm92aWRlci5lbmNyeXB0aW9ucHJvdmlkZXJpbXBsLm
FlcwAAAUTg2AhyrO0ABXNyABlqYXZheC5jcnlwdG8uU2VhbGVkT2JqZWN0PjY9psO3VHACAARbAA1lbmNvZGVkUGFyYW1zdAAC
W0JbABBlbmNyeXB0ZWRDb250ZW50cQB+AAFMAAlwYXJhbXNBbGd0ABJMamF2YS9sYW5nL1N0cmluZztMAAdzZWFsQWxncQB
+AAJ4cHVyAAJbQqzzF/gGCFTgAgAAeHAAAAAPMA0ECEnr6ybTxO1mAgEUdXEAfgAEAAAAcGbNRpiJe0xkAuiMpWPjhzFuWCD2
OeK7YZ4pwutRbgEcx4ul3SfPDAQcMZDTH+Ze03p8p1m7Kb/yY7SK6xvaaFYvCC9IWNgU6pkz/FXswnVgb1G/Jsve7mYEX+
8R2FUC+t2CEuzioKdTChUZsnzz0xB0AANQQkV0ABZQQkVXaXRoTUQ1QW5kVHJpcGxlREVTqmaA1K/MuEHB/yIaqSe9NgA2JsY=
" >> config.properties
echo "jdbc.drivers.location=/usr/share/jdbc" >> config.properties
echo "default.security.provider=Kerberos" >> config.properties
echo "load.feature.flags.on.msg=true" >> config.properties
echo "spark.serializer=org.apache.spark.serializer.JavaSerializer" >> config.properties
echo "spark.executor.extraLibraryPath=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/
configuration/linux/lib_64" >> config.properties
echo "zookeeper.server.jaas.config=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/
configuration/jaas.conf" >> config.properties

chmod 644 *.xml
```

```
unzip cf.zip

mkdir hadoop-conf
```

7. Laden Sie den SPARK- und Zookeeper-Client von FusionInsight herunter, extrahieren Sie den Inhalt und kopieren Sie die Spark-Konfigurationsdateien in den Ordner `/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/hadoop-conf`.
8. Kopieren Sie die nativen Hadoop-Dateien in den folgenden Ordner: `/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64`
9. Fügen Sie der Datei `/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/jvm.option` die folgende Zeile hinzu:

   ```
   -Dconfig.folder.path=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration
   ```
10. Führen Sie das Script `hdfsUpdate.sh` (`/opt/ibm/spss/analyticserver/3.1/bin/hdfsUpdate.sh`) aus.
11. Führen Sie das Script `start.sh` (`/opt/ibm/spss/analyticserver/3.1/bin/start.sh`) aus, um den Analytic Server-Service zu starten. Führen Sie das Script `stop.sh` (`/opt/ibm/spss/analyticserver/3.1/bin/stop.sh`) aus, um den Analytic Server-Service zu stoppen.

Kapitel 6. 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; andernfalls geben Sie **N** an.

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

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

   as_context_root
   Das Analytic Server-Kontextstammverzeichnis (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 ist 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 SPSS Modeler-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 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/iop/4.1.0.0/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 mit bestimmten Hadoop-Verteilungen

Aktualisierungsaktion für Analytic Server-Service unter Hortonworks 2.3 inaktiviert

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

1. Melden Sie sich an dem Host, der Analytic Metastore 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/3.0/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.

LZO-Kompressor funktioniert unter BigInsights 4.2.X nicht

Sie müssen **yum install lzo* hadoop-lzo*** ausführen, um die LZO-Komponenten in BigInsights 4.2.X zu installieren. Da sich das Installationsverhalten in der BigInsights 4.2-Umgebung geändert hat, werden die LZO-Komponenten jetzt unter dem Verzeichnis `/usr/lib/hadoop-lzo/lib` installiert. (Dieses Verzeichnis schließt nicht die Hadoop-Ausführungsumgebung ein.) Konfigurieren Sie die LZO-Komponenten in BigInsights 4.2.X mithilfe der folgenden Schritte manuell.

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

   Anmerkung: Sie können den Hostnamen in der Ambari-Konsole feststellen.

2. Kopieren Sie die Hadoop-LZO-JAR-Datei in das Analytic Server-Bibliotheksverzeichnis. Beispiel:

   ```
   cp /usr/lib/hadoop-lzo/lib/hadoop-lzo*.jar {AS-Stammverzeichnis}/ae_wlpserver/usr/servers/
   aeserver/apps/AE_BOOT.war/WEB-INF/lib
   ```

3. Kopieren Sie die nativen LZO-Bibliotheken in das Analytic Server-Verzeichnis der nativen Bibliotheken. Beispiel:

   ```
   cp /usr/lib/hadoop-lzo/lib/native/* {AS-Stammverzeichnis}/ae_wlpserver/usr/servers/aeserver/
   configuration/linux/lib_64
   ```

4. Führen Sie das Script **refresh** aus, das sich im Verzeichnis `{AS-Stammverzeichnis}/bin` befindet. Beispiel:

   ```
   cd /opt/ibm/spss/analyticserver/3.0/bin
   ./refresh
   ```

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

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 IBM SPSS Modeler-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.

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 21 aus.

java.net.SocketTimeoutException: Lesezeitlimit überschritten

Ändern Sie die Umgebungsvariable für Zeitlimits in Liberty ND wie folgt:

```
export LIBERTYND_READ_TIMEOUT=<Millisekunden>
```

Dabei gibt <Millisekunden> die Anzahl der Sekunden für das JMX-Lesezeitlimit an.

java.io.IOException: CWWKX7202E: Der Zeitlimitwert von 60 (Sekunden) für den Befehl zum Starten des Servers wurde überschritten

Fügen Sie der Datei server.xml des Controller-Servers Folgendes hinzu:

```
<!-- Zeitlimit für Starten/Stoppen des Servers zur Berücksichtigung langsamer Hardware erhöhen -->
<serverCommands startServerTimeout="120" stopServerTimeout="120"/>
```

java.lang.OutOfMemoryError: Größe des Java-Heapspeichers

Fügen Sie der Datei `jvm.options` auf jedem Member des HA-Clusters die folgenden Zeilen hinzu:

```
-Xms512M
-Xmx2048M
```

"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** für den Analytic Server-Service, fügen Sie **Advanced analytics-jvm-options** Folgendes hinzu und starten Sie dann 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.

Gelegentlich blockieren Analytic Server-Server beim Herunterfahren des Servers
Beenden Sie den Server manuell.

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 Europe, Middle East & Africa
Tour Descartes
2, avenue Gambetta
92066 Paris La Defense
France

Trotz sorgfältiger Bearbeitung können technische Ungenauigkeiten oder Druckfehler in dieser Veröffentlichung nicht ausgeschlossen werden. Die 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
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
USA

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

Die Lieferung des im Dokument aufgeführten 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 unter Umständen 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 und jede Ähnlichkeit mit tatsächlichen Personen oder Unternehmen ist 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 und jede Ähnlichkeit mit tatsächlichen Personen oder Unternehmen ist rein zufällig.

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

© (Name Ihrer Firma) (Jahr). Teile des vorliegenden Codes wurden aus Beispielprogrammen der IBM Corp. abgeleitet.

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.