

IBM SPSS Analytic Server
Version 1

Installationsanweisungen

IBM

Inhaltsverzeichnis

IBM SPSS Analytic Server - Installationsanweisungen	1
Unterstützte Plattformen	1
Voraussetzungen	1
Installieren von Analytic Server mit dem Installationsprogramm	1
Manuelle Installation unter Hadoop	3
config.properties	4
Hinzufügen von JDBC-Treibern	5
WebSphere Liberty-Konfiguration	5
Basisregistrierung.	6
LDAP-Registrierungskonfiguration	6
Authentifizierung mithilfe eines Kerberos-Sicherheitsproviders	7

Aktivieren des Hochverfügbarkeitsmodus in Hadoop 2.0.	8
Aktivieren von HCatalog-Datenquellen	8
Aktivieren der Unterstützung für Essentials for R	10
Anmeldung neuer Benutzer	11
Aktualisieren von Analytic Server	11
Deinstallieren von Analytic Server-Server	12
Problembestimmung	12
Protokollierung	12
Versionsinformationen.	12
Log Collector.	13

IBM SPSS Analytic Server - Installationsanweisungen

Unterstützte Plattformen

Table 1. Umgebungen

Produkt	Release	Prozessor	Wortgröße
Red Hat Enterprise Linux	6.x	x64	64 Bit
Red Hat Enterprise Linux	5.x	x64	64 Bit
SuSE Linux Enterprise Server	11	x64	64 Bit
SuSE Linux Enterprise Server	10	x64	64 Bit

Voraussetzungen

- Eine der folgenden Verteilungen:
 - Apache Hadoop 0.20.2
 - Apache Hadoop 1.0.3
 - IBM BigInsights 2.0
 - Cloudera CDH3
 - Cloudera CDH4 MRv1
 - Hortonworks Data Platform 1.x
- Damit das Installationsprogramm die erforderlichen Dateien aus der Hadoop-Verteilung kopieren kann, müssen ssh und scp installiert sein und es muss möglich sein, eine Verbindung mit einer Public-Key-Authentifizierung herzustellen, ohne dass für die Anmeldung ein Kennwort benötigt wird. Andernfalls müssen Sie die Dateien manuell kopieren. So überprüfen Sie, ob Ihre SSH-Verbindung funktioniert:
 1. Melden Sie sich mit einer Public-Key-Authentifizierung am Hadoop-Namensknotenserver an.
 2. Führen Sie Folgendes aus und ersetzen Sie dabei {HADOOP-SERVER} durch Ihren Hadoop-Server.

```
export HADOOP_SERVER_USERNAME='whoami'  
export HADOOP_SERVER={HADOOP-SERVER}  
ssh -q -o "BatchMode=yes" -o "ConnectTimeout 5"  
    $HADOOP_SERVER_USERNAME@$HADOOP_SERVER "echo 2>&1"  
    && echo HOST_SSH_OK || echo HOST_SSH_NOK
```
 3. Wenn "HOST_SSH_OK" zurückgegeben wird, war die Prüfung erfolgreich. Andernfalls ist die Verbindung nicht ordnungsgemäß konfiguriert.

Installieren von Analytic Server mit dem Installationsprogramm

Das Installationsprogramm automatisiert eine Reihe von Schritten, die ansonsten manuell ausgeführt werden. Das Installationsprogramm muss Zugriff auf die HDFS-Umgebung haben. So installieren Sie Analytic Server:

1. Führen Sie die Installationsdatei `install.bin` aus.
 - a. Lesen Sie die Lizenzvereinbarung und akzeptieren Sie sie, um fortzufahren.
 - b. Wählen Sie einen Installationspfad aus.
 - c. Folgen Sie der Bedienungshilfe, um die Dateien auf Ihrem System zu installieren.

- d. Konfigurieren Sie Analytic Server-Optionen. Der Name des Administratorkontos (STANDARD: admin) ist der Name des Benutzers mit Superuser-Administratorrechten für die Konsole der Analytic Server-Anwendung.
- e. Keystore Password. In diesem Keystore wird der Schlüssel gespeichert, der zum Verschlüsseln von Kennwörtern für Datenbankdatenquellen verwendet wird, die in der Analytic Server-Datenbank gespeichert werden.

f. Option **Automatically Detect Hadoop Version:**

Yes Geben Sie den Hadoop-Namensknotenserver (STANDARD: localhost) an. Wenn Hadoop sich auf demselben Computer befindet, auf dem auch dieses Produkt installiert wird, verwenden Sie 'localhost' als Servernamen. Das Installationsprogramm stellt dann fest, welche Version von Hadoop installiert ist, und stellt die erforderlichen Dateien aus der Hadoop-Verteilung zusammen. Hinweis: ssh und scp müssen installiert sein und es muss möglich sein, eine Verbindung mit einer Public-Key-Authentifizierung herzustellen, ohne dass für die Anmeldung ein Kennwort benötigt wird.

No

- 1) Wählen Sie die Hadoop-Serverversion aus einer Liste aus.
- 2) Geben Sie die Position der erforderlichen Dateien aus der Hadoop-Verteilung an oder kopieren Sie die Dateien später manuell.
- 3) Wenn es sich bei Ihrer Verteilung um Cloudera CDH3 handelt, geben Sie die Position einer Oracle Java Runtime Environment 1.6 an oder konfigurieren Sie später manuell.

g. Konfigurieren Sie Hadoop-Parameter:

Hadoop user name

Der Name des UNIX-Benutzers, mit dem Analytic Server ausgeführt wird.

Password

Das Kennwort für den Hadoop-Benutzernamen.

Analytic Server root directory

In diesem Pfad definiert Analytic Server seinen Arbeitsbereich und speichert ausführbare Dateien, die an Hadoop-Knoten verteilt werden.

Namenode host

Der Name oder die IP-Adresse für den HDFS-Namensknotenhost.

Namenode port

Die Portnummer für den Namensknotenhost.

Job tracker

Der Name oder die IP-Adresse des Hadoop-Job-Tracker-Hosts.

Job tracker port number

Die Portnummer für den Job-Tracker.

2. Konfigurieren Sie WebSphere Liberty für die Registrierung von Benutzern. Weitere Informationen finden Sie in „WebSphere Liberty-Konfiguration“ auf Seite 5.
3. Starten Sie den Server, indem Sie {AS-Stammverzeichnis}/bin/start.sh ausführen. {AS-Stammverzeichnis} steht für den Ordner, in dem Analytic Server bereitgestellt wird, z. B. /opt/IBM/SPSS/AnalyticServer/1.0.
4. Öffnen Sie einen Web-Browser und geben Sie die Adresse http://localhost:8080 ein. Diese URL öffnet das Anmeldedialogfeld für die Analytic Server-Konsole. Melden Sie sich als der Superuser an, der oben definiert wurde. Details zur Nutzerverwaltung finden Sie im Handbuch *IBM® SPSS Analytic Server Verwaltung*.

Wenn das Installationsprogramm keine Verbindung zum Hadoop-Server herstellen kann oder keine gültige Hadoop-Version gefunden wird, führen Sie die manuellen Schritte aus.

Manuelle Installation unter Hadoop

1. Führen Sie `install.bin` aus. Hierdurch wird die Analytic Server-Datenbank erstellt und die Berechtigungsnachweise für Benutzer mit Verwaltungsaufgaben werden festgelegt.
 - a. {AS-Stammverzeichnis} steht für den Ordner, in dem Analytic Server bereitgestellt wird, z. B. `/opt/IBM/SPSS/AnalyticServer/1.0`.
 - b. {AS-Serverstammverzeichnis} steht für die Position der Konfigurations-, Protokoll- und Serverdateien, z. B. `/opt/IBM/SPSS/AnalyticServer/1.0/ae_wlpserver/usr/servers/aeserver`.
 - c. {AS-Ausgangsverzeichnis} steht für den HDFS-Ordner, der von Analytic Server als Stammordner verwendet wird.
2. Bearbeiten Sie die Konfigurationsdatei `{AS-Serverstammverzeichnis}/configuration/config.properties`. Weitere Informationen finden Sie in „`config.properties`“ auf Seite 4. Konfigurieren Sie die Hadoop-Einstellungen (**`hdfs.namenode.url`**, **`hdfs.user`**, **`hdfs.password`**) auf der Basis der spezifischen Hadoop-Konfiguration Ihres Systems.
3. Analytic Server wird auf IBM WebSphere Liberty aufgesetzt, das standardmäßig für die Ausführung an Port 8080 konfiguriert wird.
4. Kopieren Sie die richtigen Hadoop-JAR-Dateien in das Verzeichnis `{AS-Stammverzeichnis}/apps/AE_B00T.war/WEB-INF/lib`. Diese JAR-Dateien basieren auf Ihrer Installation und befinden sich auf dem Hadoop-Clusternamensknoten, auf dem Hadoop installiert ist, z. B. in den Ordnern `/usr/lib/hadoop` und `/usr/lib/hadoop/lib`.
 - Apache Hadoop 0.20.2
 - `hadoop-0.2*-core.jar`
 - Apache Hadoop 1.0.3
 - `commons-configuration-*.jar`
 - `commons-io-*.jar`
 - `commons-lang-*.jar`
 - `hadoop-core-1.*.jar`
 - `jackson-core-asl-*.jar`
 - `jackson-mapper-asl-*.jar`
 - `jersey-core-*.jar`
 - `jersey-server-*.jar`
 - Cloudera CDH3
 - `aspectjrt-1.6.5.jar`
 - `aspectjtools-1.6.5.jar`
 - `core-3.1.1.jar`
 - `guava-r09*.jar`
 - `hadoop-fairscheduler-0.20.2-*.jar`
 - `hadoop-core-0.20.2-*.jar`
 - `hadoop-tools-0.20.2-*.jar`
 - `jackson-core-asl-1.5.2.jar`
 - `jackson-mapper-asl-1.5.2.jar`
 - `jasper-compiler-5.5.12.jar`
 - `jasper-runtime-5.5.12.jar`

Für Cloudera CDH3 Aktualisierung 5 und höher sind die folgenden JAR-Dateien zusätzlich:

- `jersey-core-1.8.jar`
- `jersey-json-1.8.jar`
- `jersey-server-1.8.jar`
- Cloudera 4.0.1
 - `aspectjrt-1.6.5.jar`
 - `aspectjtools-1.6.5.jar`
 - `commons-io-2*.jar`
 - `guava-*.jar`
 - `hadoop-fairscheduler-2.*.jar`
 - `hadoop-core-2.*.jar`
 - `hadoop-tools-2.*.jar`
 - `jackson-core-asl-1*.jar`
 - `jackson-mapper-asl-1*.jar`
 - `jasper-compiler-5*.jar`
 - `jasper-runtime-5*.jar`
- Cloudera 4.1.*
 - `avro-*.jar`
 - `avro-compiler-*.jar`
 - `commons-cli-*.jar`
 - `commons-configuration-*.jar`
 - `commons-io-2*.jar`
 - `commons-lang-*.jar`
 - `guava-*.jar`
 - `hadoop-auth-*.jar`
 - `hadoop-common-2.*.jar`
 - `hadoop-core-2.*.jar`
 - `hadoop-hdfs-2.0.0-cdh4.1.2.jar`
 - `jackson-core-asl-*.jar`
 - `jackson-mapper-asl-*.jar`
 - `jersey-core-*.jar`
 - `jersey-server-*.jar`
 - `protobuf-java-*.jar`
 - `slf4j-api-*.jar`
 - `slf4j-log4j12-*.jar`

- IBM BigInsights 2.0

commons-configuration-*.jar	jackson-core-asl-*.jar
commons-lang-*.jar	jackson-mapper-asl-*.jar
commons-io-2*.jar	jersey-core-*.jar
hadoop-core-1.*.jar	jersey-server-*.jar
 - Hortonworks Data Platform 1.x

commons-configuration-*.jar	jackson-core-asl-*.jar
commons-lang-*.jar	jackson-mapper-asl-*.jar
commons-io-2*.jar	jersey-core-*.jar
hadoop-core-1.*.jar	jersey-server-*.jar
5. Cloudera CDH3 erfordert eine Oracle-JVM der Version 1.6. Aktualisieren Sie das Script setenv.sh {AE-Stammverzeichnis}/bin, indem Sie JAVA_HOME aktualisieren und den Standardpfad in den Pfad zu einer Oracle-JVM ändern.
 6. Wenn Sie Apache Hadoop 1.0.3 oder höher verwenden, ersetzen Sie commons-io-1.4.jar in {AS-Serverstammverzeichnis}/apps/AE_BOOT.war/WEB-INF/lib durch commons-io-2.x.jar.
 7. Wenn die Hadoop-Version 2.0 oder höher ist, kopieren Sie {AS-Serverstammverzeichnis}/cloudera/cloudera4/hadoopmapreduce_1.1-1.0.0.0.jar in {AS-Serverstammverzeichnis}/apps/AE_BOOT.war/WEB-INF/lib und entfernen Sie hadoopmapreduce-1.0.0.0.jar.
 8. Stellen Sie sicher, dass HDFS betriebsbereit ist und darauf zugegriffen werden kann. Kopieren Sie ausführbare Analytic Server-Dateien in HDFS und initialisieren Sie den Analytic Server-Arbeitsbereich, indem Sie {AS-Stammverzeichnis}/bin/hdfsUpdate.sh ausführen.
 9. Starten Sie den Server, indem Sie {AS-Stammverzeichnis}/bin/run.sh ausführen.
 10. Geben Sie http://localhost:8080 in Ihren Browser ein, um die Analytic Server-Konsole aufzurufen. Melden Sie sich mit dem Administratorbenutzernamen (admin.username) an, der in der Datei config.properties definiert ist. Details zur Nutzerverwaltung finden Sie im Handbuch *IBM SPSS Analytic Server Verwaltung*.
 11. Führen Sie zum Stoppen des Servers {AS-Stammverzeichnis}/bin/stop.sh aus.
 12. Das Serverprotokoll befindet sich im Verzeichnis {AS-Serverstammverzeichnis}/logs/.

config.properties

Die Datei config.properties enthält eine Reihe von Analytic Server-Einstellungen. Die folgenden Einstellungen entsprechen den Einstellungen, die durch das Installationsprogramm angegeben würden. Prüfen Sie bei der manuellen Installation des Produkts, ob die Standardeinstellungen übernommen werden können.

root.folder.path

Der Pfad, in dem die Analytic Server-Benutzerartefakte (Eingaben und Ausgaben) gespeichert werden.

admin.consumer=ibm

Der Standardadministratorkontonutzer. Nutzer stellen eine übergeordnete Einteilung der Rollen, Berechtigungen, Projekte und Datenquellen bereit.

admin.username=admin

Der Name des Standardadministratorkontos. Dieser Benutzername hat Basisadministratorrechte für die Analytic Server-Anwendung über die Konsole.

hdfs.namenode.url=hdfs://hadoop:54310/user/hdadmin

Die URL für den Hadoop-Namensknoten.

hdfs.user=hdadmin

Der Hadoop-Benutzername. Der Name des UNIX-Benutzers, mit dem Analytic Server ausgeführt wird.

hdfs.password=<KENNWORT EINGEBEN>

Das Kennwort für den Benutzer, der in der hdfs.user-Einstellung angegeben ist.

hdfs.classpath.folder=/user/hdpadmin/classpath

Das Analytic Server-Stammverzeichnis. In diesem Pfad definiert Analytic Server seinen Arbeitsbereich und speichert ausführbare Dateien, die an Hadoop-Knoten verteilt werden.

mapred.job.tracker=hadoop:54311

Name oder IP-Adresse und Portnummer des Hadoop-Job-Tracker-Hosts.

encryption.keystore.password

In diesem Keystore wird der Schlüssel gespeichert, der zum Verschlüsseln von Kennwörtern für Datenbankdatenquellen verwendet wird, die in der Analytic Server-Datenbank gespeichert werden.

Hinzufügen von JDBC-Treibern

Zur Unterstützung von Datenbankdatenquellen müssen Sie Analytic Server die JDBC-Treiber hinzufügen.

1. Stoppen Sie Analytic Server, indem Sie `{AS-Stammverzeichnis}/bin/stop.sh` ausführen.
2. Kopieren Sie die erforderlichen JAR-Dateien für JDBC-Treiber in das Verzeichnis `{AS-Stammverzeichnis}/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib`.
3. Aktualisieren Sie Analytic Server, indem Sie `{AS-Stammverzeichnis}/bin/hdfsUpdate.sh` ausführen.
4. Starten Sie Analytic Server, indem Sie `{AS-Stammverzeichnis}/bin/start.sh` ausführen.

Tabelle 2. Unterstützte Datenbanken

Datenbank	Unterstützte Versionen	JAR-Dateien für JDBC-Treiber	Anbieter
DB2 for Linux, UNIX, and Windows	9.5, 9.7, 10.0	db2jcc.jar	IBM
DB2 z/OS	10	db2jcc.jar, db2_license_cisuz.jar	IBM
Teradata	13.1, 14	tdgssconfig.jar, terajdbc4.jar	Teradata
SQL Server	2012, 2008 R2	sqljdbc4.jar	Microsoft
Netezza	6.x, 7	nzjdbc.jar	IBM
Oracle	12g, 11g R2	ojdbc6.jar, orai18n.jar	Oracle

WebSphere Liberty-Konfiguration

WebSphere Liberty Profile ist eine einfache Implementierung von IBM WebSphere. Analytic Server kann für die Authentifizierung von Benutzern die WebSphere-Anwendungssicherheit verwenden. Diese wird in der Datei `server.xml` für den Server eingerichtet, auf dem Analytic Server bereitgestellt wird. Damit die Anwendungssicherheit in Liberty aktiviert wird, muss das Feature `appSecurity-1.0` im Feature-Manager enthalten sein:

```
<featureManager onError="FAIL">
...
<feature>appSecurity-1.0</feature>
...
</featureManager>
```

Damit SSL auf einem Server aktiviert wird, muss die SSL-Funktion in der Datei `server.xml` angegeben werden:

```
<featureManager>
  <feature>ssl-1.0</feature>
</featureManager>
```

Detaillierte Informationen zur WebSphere-Sicherheit finden Sie unter ftp://ftp.software.ibm.com/software/webserver/appserv/library/v85/was85base_security.pdf.

Basisregistrierung

Mithilfe der Basisregistrierung kann der Administrator eine Datenbank mit Benutzern und Gruppen in der Datei {AS-Serverstammverzeichnis}/server.xml definieren. Kennwörter können mit dem Tool securityUtil codiert werden, um ihre Werte zu verschlüsseln. Dieses Tool befindet sich in {AS-Stammverzeichnis}/ae_wlpserver/bin.

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

```
<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>
```

LDAP-Registrierungskonfiguration

Die LDAP-Registrierung bietet für den Administrator eine Möglichkeit, Benutzer mit einem externen LDAP-Server wie Active Directory oder OpenLDAP zu authentifizieren. Im Folgenden finden Sie ein Beispiel für eine LDAP-Registrierung (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="(&uid=%v)(objectClass=inetOrgPerson)"
    groupFilter="(&cn=%v)(|(objectclass=organizationalUnit))"
    groupMemberIdMap="posixGroup:memberUid"/>
</ldapRegistry>
```

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

Authentifizierung mithilfe eines Kerberos-Sicherheitsproviders

Zum Konfigurieren von Kerberos benötigen Sie die folgenden Informationen von Ihrem Hadoop-Administrator:

1. Kerberos-Realm, z. B. ASSSO.COM
2. Hostname von Kerberos-Key-Distribution-Center (KDC), z. B. kdc.asso.com
3. Kerberos-Principal für Namensknoten, z. B. hdfs/namenode.asso.com@ASSSO.COM
4. Kerberos-Principal für MapReduce-Knoten, z. B. mapred/jobtracker.asso.com@ASSSO.COM

Anschließend müssen Sie die Datei `krb5.conf` unter `/etc/krb5.conf` konfigurieren. Beispiel:

```
[libdefaults]
default_realm = ASSSO.COM
default_tkt_enctypes = rc4-hmac des-cbc-md5
default_tgs_enctypes = rc4-hmac des-cbc-md5
dns_lookup_realm = false
dns_lookup_kdc = false
ticket_lifetime = 24h
forwardable = yes
```

```
[realms]
ASSSO.COM = {
    kdc = kdc.asso.com:88
    default_domain = asso.com
}
```

```
[dmain_realm]
.asso.com = ASSSO.COM
asso.com = ASSSO.COM
```

Ändern Sie die Datei `config.properties` nach der Erstellung der Datei `krb5.conf` wie folgt:

1. Im Abschnitt für die Analytic Server-Module:
 - Fügen Sie die Module `hdfsauth` und `kerberossecurityprovider` hinzu.
 - Entfernen Sie das Modul `wssecurityprovider`.
2. Fügen Sie die folgenden Konfigurationseigenschaften hinzu:

```
#Kerberos-Authentifizierungsparameter
hadoop.security.authentication=kerberos
dfs.namenode.kerberos.principal=hdfs/namenode.asso.com@ASSSO.COM
mapreduce.jobtracker.kerberos.principal=mapred/jobtracker.asso.com@ASSSO.COM
java.security.krb5.conf=/etc/krb5.conf
```

Dabei gilt Folgendes:

hadoop.security.authentication

Hadoop-Sicherheitsauthentifizierung. Geben Sie zur Aktivierung des Kerberos-Sicherheitsproviders `kerberos` an.

dfs.namenode.kerberos.principal

Kerberos-Principal, der für die Chiffrierschlüsseldatei verwendet wird, mit der der Namensknoten gestartet wird.

mapreduce.jobtracker.kerberos.principal

Kerberos-Principal, der für die Chiffrierschlüsseldatei verwendet wird, mit der der Job-Tracker gestartet wird.

java.security.krb5.conf

Position der Kerberos-Konfigurationsdatei.

3. Konfigurieren Sie das Liberty-LDAP-Benutzerrepository in der Datei `service.xml`. Informationen hierzu finden Sie in „WebSphere Liberty-Konfiguration“ auf Seite 5. Alle Benutzer, die im Liberty-Benutzerrepository angegeben sind, müssen mit Kerberos-Benutzerkonten abgeglichen werden, und dieselbe LDAP-Einstellung muss auf dem Kerberos-Server verwendet werden.

4. Standardmäßig verwendet Analytic Server das Verzeichnis `.temp` unter dem Ausgangsverzeichnis des Benutzers als temporäres Verzeichnis, aber wenn Sie das temporäre Verzeichnis an einer anderen Position konfigurieren wollen, befolgen Sie die nachfolgenden Anweisungen.

- a. Bearbeiten Sie `config.properties` und entfernen Sie das Kommentarzeichen bei folgender Konfigurationseinstellung.

```
#as.temp.folder=/.temp
```

Ändern Sie die Einstellung so, dass sie den absoluten Pfad zum temporären Verzeichnis angibt. Es werden keine Änderungen benötigt, wenn das Verzeichnis `/.temp` verwendet wird.

- b. Ändern Sie die Berechtigungen für diesen Ordner so, dass alle Benutzer auf ihn zugreifen können. Beispiel: `hadoop fs -chmod 777 /.temp`.

5. Erteilen Sie allen Kerberos-Benutzern Leseberechtigung für die Ordner `classpath` und `configuration`.

- a. Öffnen Sie die Datei `config.properties` und ermitteln Sie die Einstellungen für die Parameter **`hdfs.classpath.folder`** und **`component.framework.bin.path`**.

- b. Legen Sie die Leseberechtigungen wie folgt fest.

Wenn **`hdfs.classpath.folder`** auf `/user/hdpadding/classpath` und **`component.framework.bin.path`** auf `/user/hdpadding/configuration` gesetzt ist, führen Sie Folgendes aus:

```
hadoop fs -chmod -R 755 /user/hdpadding
hadoop fs -chmod -R 755 /user/hdpadding/classpath
hadoop fs -chmod -R 755 /user/hdpadding/configuration
```

Aktivieren des Hochverfügbarkeitsmodus in Hadoop 2.0

Analytic Server unterstützt die Ausführung benutzerdefinierter Hadoop-Modi wie des Hochverfügbarkeitsmodus (HA-Modus), indem die Konfiguration der Cluster-Client-API im Analytic Server-Konfigurationsordner bereitgestellt wird. Die Bereitstellung der Konfiguration der Hadoop-Cluster-Client-API für Analytic Server ist nur obligatorisch, wenn benutzerdefinierte Einstellungen wie die Hochverfügbarkeit vorhanden sind. Diese Einstellungen sind erforderlich, damit die Analytic Server-API die Anforderungen im Falle eines Fehlschlagens des Clusternamensknotens ausgleichen kann.

So konfigurieren Sie den benutzerdefinierten Modus nach einer erfolgreichen Analytic Server-Installation:

1. Fordern Sie die Konfigurationsdateien des Cluster-Clients an (`hdfs-site.xml` und `core-sites.xml`), die sich normalerweise auf einem Namensknotencomputer im Verzeichnis `/etc/hadoop/conf` befinden.
2. Kopieren Sie diese Dateien in das Verzeichnis `{AS-Stammverzeichnis}/ae_wlpserver/usr/servers/aeserver/configuration/hadoop-conf`.
3. Wenn der Hadoop-Cluster im Hochverfügbarkeitsmodus konfiguriert ist, stellen Sie sicher, dass die Analytic Server-Konfigurationseigenschaft `hdfs.namenode.url` in der Datei `config.properties` auf den HDFS-Servicenamen und `/user/Benutzername` verweist. Beispiel: `hdfs://nameservice1/user/hdpadding`.
4. Wenn Sie hohe Verfügbarkeit für den Job-Tracker-Service eingerichtet haben, aktualisieren Sie die Eigenschaft **`mapred.job.tracker`** in `config.properties` so, dass sie auf den HDFS-Servicenamen verweist.
5. Aktualisieren Sie das Hadoop-Dateisystem mit dem folgenden Befehl:

```
{AS-Stammverzeichnis}/bin/hdfsUpdate.sh
```
6. Starten Sie Analytic Server mit dem folgenden Befehl:

```
{AS-Stammverzeichnis}/bin/run.sh
```

Aktivieren von HCatalog-Datenquellen

So konfigurieren Sie Analytic Server für die Verwendung von HCatalog-Datenbanken nach einer erfolgreichen Analytic Server-Installation:

1. Fügen Sie der Liste der Analytic Server-Module in der Datei {AS-Stammverzeichnis}/ae_wlpserver/usr/servers/aeserver/configuration/config.properties den Eintrag hcataloginput hinzu. Beispiel:

```
ae.modules=securityprovidermanager,\
localClient,\
componentframework,\
...
hcataloginput
```

2. Fügen Sie der Datei config.properties die folgenden Zeilen hinzu oder entfernen Sie bei diesen Zeilen die Kommentarzeichen:

```
hive.metastore.local=false
hive.metastore.uris=thrift://Hostname:Portnummer
```

Dabei gilt Folgendes:

Hostname

Der Name des Computers, auf dem sich der Thrift-Server befindet

Portnummer

Die Portnummer, die im HCatalog-Installationsscript verwendet wird

3. Stellen Sie sicher, dass die folgenden Dateien im Verzeichnis {AS-Stammverzeichnis}/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib verfügbar sind, und kopieren Sie diese Dateien auch in das HDFS-Verzeichnis /user/{AE-Administrator}/classpath.

Die folgenden JAR-Dateien gelten für HCatalog 0.4.0 und Hive 0.9.0. Sie müssen die entsprechenden HCatalog- und Hive-JAR-Dateien sowie abhängige JAR-Dateien entsprechend für andere Versionen übernehmen.

- Die folgende Datei kann aus der HCatalog-Serverinstallation kopiert werden.

```
hcatalog-0.4.0.jar
```
- Die folgenden Dateien können aus der Hive-Serverinstallation kopiert werden.

```
hive-exec-0.9.0.jar
hive-metastore-0.9.0.jar
libfb303-0.7.0.jar
slf4j-api-1.6.1.jar
slf4j-log4j12-1.6.1.jar
```
- Die folgenden Dateien können aus der Hadoop 1.X-Installation oder aus <http://jackson.codehaus.org/> kopiert werden.

```
jackson-core-asl-1.8.8.jar
jackson-mapper-asl-1.8.8.jar
```

Die folgenden JAR-Dateien gelten für HCatalog 0.5.0.

- Die folgenden Dateien können aus <http://code.google.com/p/guava-libraries/wiki/Release13> kopiert werden.

```
guava-13.0.1.jar
```

4. Wenn Sie komplexe Hive-Typen verwenden möchten (Zuordnungen, Arrays, Strukturen), kopieren Sie die folgenden JAR-Dateien in Ihr Hive-Verzeichnis /usr/local/hive/hive-0.9.0/auxlib/:

```
hcatalog-0.4.0.jar
jackson-core-asl-1.8.8.jar
jackson-mapper-asl-1.8.8.jar
```

5. Die auf den komprimierten Sequenzdateien basierenden HCatalog-Datenquellen erfordern, dass native Hadoop-Bibliotheken auf dem Analytic Server-System verfügbar sind. Kopieren Sie die nativen Hadoop-Bibliotheken unter {HADOOP}/lib/native/Linux-amd64-64 in ein Verzeichnis auf dem Analytic Server-System und fügen Sie den Pfad zu diesem Verzeichnis der Variablen **LIB_PATH** in {AS-Stammverzeichnis}/bin/start.sh hinzu. Beispiel:

```
export LIB_PATH=$AE_BASE/ae_wlpserver/usr/servers/aeserver/configuration/lib_32:
$AE_BASE/ae_wlpserver/usr/servers/aeserver/configuration/lib_64:
<Verzeichnis_für_native_Hadoop-Bibliotheken>
```

Anmerkung: Wenn Ihre Analytic Server-Installation bereits über neuere Versionen dieser JAR-Dateien verfügt, dürfen Sie die älteren Versionen nicht kopieren. Wenn das Analytic Server-Installationsprogramm z. B. jackson-core-as1-1.8.0 bereits aus einer neueren Hadoop-Verteilung kopiert hat, dürfen Sie die JAR-Datei jackson-core-as1.1.7.3 nicht kopieren.

Aktivieren der Unterstützung für Essentials for R

Analytic Server unterstützt das Scoring von R-Modellen und die Ausführung von R-Skripts.

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

1. Installieren Sie die R-Engine auf dem Server, auf dem sich Analytic Server befindet, indem Sie die folgenden Schritte ausführen:

```
mkdir /home/hdpadmin/APPS/R
cd /home/hdpadmin/APPS/R
wget http://cran.r-project.org/src/base/R-2/R-2.15.2.tar.gz
tar -xvzf R-2.15.2.tar.gz
cd R-2.15.2
./configure --enable-R-shlib
make
```

2. Installieren Sie Essentials for R auf dem Server, auf dem sich Analytic Server befindet, indem Sie die Installationsdatei `install.bin` entsprechend den auf dem Bildschirm angezeigten Anweisungen ausführen. Das Installationsprogramm führt die folgenden Aktionen aus:

- a. Aktualisiert die R-Engine-Installation für Analytic Server und fügt das "R-Plug-in" hinzu.
- b. Aktualisiert das Verzeichnis `{AS-Stammverzeichnis}/ae_wlpserver/usr/servers/aeserver/configuration/ext_64/bin`, um dem Modul `pasw.rstats` eine native Bibliothek und eine Konfigurationsdatei hinzuzufügen.

3. Stellen Sie die R-Engine und die R-Komponente in Hadoop bereit.

- a. Wenn Analytic Server und alle Hadoop-Knoten dieselbe Version des Betriebssystems und dieselbe Prozessorarchitektur haben, gehen Sie folgendermaßen vor:

- 1) Erstellen Sie mithilfe des folgenden Skripts ein Archiv für die R-Engine:

```
#!/usr/bin/env bash
echo Creating R.zip...
cd /tmp
rm -r -f R
rm -f R.zip
mkdir R
cp -r $R_HOME/* ./R/
cp $(ldd ./R/bin/exec/R ./R/bin/Rscript|cut -d\ -f3|grep \.so\.|sort -u) ./R/lib/
cp -P /usr/lib64/libgfortran.so.3 ./R/lib
cp /usr/lib64/libgfortran.so.3.0.0 ./R/lib
rm -r ./R/doc ./R/src ./R/include ./R/tests
zip -r R.zip R
```

- 2) Kopieren Sie das Archiv `R.zip` in das Verzeichnis `{AS-Stammverzeichnis}/ae_wlpserver/usr/servers/aeserver/configuration/app_64`.

- b. Führen Sie `{AS-Stammverzeichnis}/bin/hdfsUpdate.sh` aus, um die Änderungen an HDFS weiterzugeben.
- c. Wenn die Versionen der Betriebssysteme für Analytic Server und Hadoop-Knoten unterschiedlich sind, installieren Sie die R-Engine und Essentials for R auf jedem Hadoop-Knoten und in einem Verzeichnis mit demselben Namen wie das, in dem die R-Engine in Analytic Server installiert ist. Wenn Sie Essentials for R auf dem Hadoop-Knoten installieren, müssen Sie nur die Position der R-Engine angeben und den Schritt zur Angabe der Position von `../ext_64/bin` überspringen.

Anmerkung: Die R-Engine muss an einer Position installiert sein, auf die alle Benutzer zugreifen können, da R auf dem Hadoop-Cluster als Benutzer ausgeführt wird, der sich vom Analytic Server-Benutzer unterscheidet.

Sie müssen Essentials for R auch auf dem Computer installieren, auf dem sich SPSS Modeler befindet. Weitere Informationen finden Sie in *IBM SPSS Modeler: Anweisungen zum Aktualisieren und Erweitern der Installation von IBM SPSS Analytic Server*.

Anmeldung neuer Benutzer

Weisen Sie die Benutzer an, zu `http://<Host>:<Port>/<Nutzer>` zu navigieren und Benutzernamen und Kennwort einzugeben, um sich bei der Analytic Server-Konsole anzumelden.

<Host>

Die Adresse des Analytic Server-Hosts

<Port>

Der Port, auf dem Analytic Server empfangsbereit ist

<Nutzer>

In einer Umgebung mit mehreren Nutzern der Nutzer, zu dem Sie gehören. In einer Umgebung mit einem einzigen Nutzer ist der Standardnutzer **ibm**.

Wenn auf IBM SPSS Analytic Catalyst zugegriffen werden soll, muss zu `http://<Host>:<Port>/catalyst.html` navigiert werden und zur Anmeldung Benutzernamen und Kennwörter müssen eingegeben werden.

Aktualisieren von Analytic Server

Wenn Sie mit dem Installationsprogramm eine Neuinstallation durchführen, können Sie die Analytic Server-Optionen, das Keystore-Kennwort und die Konfiguration der Hadoop-Parameter aktualisieren.

Aktualisieren der Hadoop-Serverversion

Gelegentlich kann es erforderlich sein, die Version des Hadoop-Servers, auf dem Analytic Server ausgeführt wird, zu aktualisieren, z. B. von Cloudera 3.x auf Cloudera 4.x. Wenn dies der Fall ist, muss Analytic Server mit den Client-JAR-Dateien aus der neuen Hadoop-Serverversion aktualisiert werden.

Anmerkung: Bei diesen Schritten wird angenommen, dass keine Konfigurationseigenschaften im Zusammenhang mit Hadoop wie Host/Port des Namensknotens, Host/Port des Job-Trackers oder Hadoop-Benutzername/-Kennwort geändert wurden.

1. Stoppen Sie Analytic Server mit dem folgenden Befehl.

```
{AS-Stammverzeichnis}/bin/stop.sh
```

2. Befolgen Sie die Anweisungen des Anbieters zur Aktualisierung des Hadoop-Servers und stellen Sie sicher, dass der Hadoop-Server aktiv ist.
3. Kopieren Sie die erforderlichen Client-JAR-Dateien in ein Verzeichnis, das für Analytic Server lokal ist; dieses Verzeichnis wird als {Hadoop-Quellenordner} bezeichnet. Die Liste erforderlicher JAR-Dateien hängt von der neuen Version des Hadoop-Servers ab und ist in „Manuelle Installation unter Hadoop“ auf Seite 3 aufgeführt.
4. Führen Sie das in {AS-Stammverzeichnis}/bin/hadoop_setup befindliche Shell-Script `install_hadoop_files` aus und geben Sie dabei {Hadoop-Quellenordner} als Argument an.
Beispiel: `./install_hadoop_files.sh /opt/hadoop_src`
 - a. Wählen Sie aus, ob Sie installierte Hadoop-Dateien entfernen möchten (empfohlen).
 - b. Wählen Sie die Hadoop-Version aus, die Sie verwenden möchten.
5. Führen Sie den folgenden Befehl aus, um das Hadoop-Dateisystem mit Analytic Server zu aktualisieren.

```
{AS-Stammverzeichnis}/bin/hdfsUpdate.sh
```
6. Führen Sie den folgenden Befehl aus, um Analytic Server erneut zu starten.

```
{AS-Stammverzeichnis}/bin/run.sh
```

Deinstallieren von Analytic Server-Server

1. Führen Sie `{AS-Stammverzeichnis}/_uninstall/uninstaller.bin` aus.

Anmerkung: Durch diese Operation werden keine Daten aus dem HDFS-Ordner für Analytic Server gelöscht. Wenn Sie den gesamten Analytic Server-Ausgangsordner entfernen wollen, entfernen Sie das HDFS-Verzeichnis, in dem die Analytic Server-Benutzerartefakte (Eingaben und Ausgaben) gespeichert werden. Dies ist das Verzeichnis, das für `root.folder.path` in `{AS-Serverstammverzeichnis}/configuration/config.properties` angegeben ist.

Anmerkung: Wenn das Deinstallationsprogramm wegen einer Java-Ausnahmebedingung `OutOfMemoryException` nicht gestartet werden kann, muss die maximale Größe des Heapspeichers erhöht werden. Die Eigenschaft `lax.nl.java.option.java.heap.size.max` in `{AS-Stammverzeichnis}/_uninstall/uninstaller.bin.lax` definiert die maximale Größe des Heapspeichers in Byte für das Installationsprogramm, das aufgerufen wird. Diese Anzahl wird immer in Byte angegeben, nicht in Kilobyte oder Megabyte, und ist mit dem VM-Parameter `-mx` oder `Xmx` gleichwertig. Der Standardwert ist 50331648 (48 MB). Es wird empfohlen, diesen Wert in 134217728 (128 MB) zu ändern und dann das Deinstallationsprogramm zu starten.

Problembestimmung

Analytic Server stellt einige hilfreiche Tools für die Problembestimmung bereit.

Protokollierung

Analytic Server erstellt Kundenprotokolldateien und `-tracedateien`, die für die Problemdiagnose hilfreich sind. Bei der Liberty-Standardinstallation finden Sie die Protokolldateien im Verzeichnis `{Liberty-Profilordner}/logs`.

Bei der Standardkonfiguration werden zwei Protokolldateien erstellt, in die im täglichen Wechsel geschrieben wird.

ae.log

Diese Datei enthält die übergeordnete Zusammenfassung von Informations- und Fehlernachrichten sowie Warnungen. Prüfen Sie zuerst diese Datei, wenn Serverfehler auftreten, die nicht mithilfe der Fehlernachricht, die an der Benutzeroberfläche angezeigt wird, behoben werden können.

ae_trace.log

Diese Datei enthält alle Einträge aus `ae.log`. Diese werden jedoch durch weitere Informationen ergänzt, die vor allem für den IBM Support und die Entwicklung bestimmt sind und zur Fehlerbehebung dienen sollen.

Analytic Server verwendet Apache LOG4J als Protokollierungseinrichtung. Mit LOG4J kann die Protokollierung dynamisch angepasst werden, indem die Konfigurationsdatei `{AS-Serverstammverzeichnis}/configuration/log4j.xml` bearbeitet wird. Dies kann erforderlich sein, wenn der Support Sie zur Erleichterung der Problemdiagnose dazu auffordert oder Sie diese Datei ändern möchten, um die Anzahl der Protokolldateien zu begrenzen. Änderungen an der Datei werden automatisch innerhalb weniger Sekunden erkannt, sodass Analytic Server nicht erneut gestartet werden muss.

Weitere Informationen zu `log4j` und zur Konfigurationsdatei finden Sie in der Dokumentation auf der offiziellen Apache-Website unter <http://logging.apache.org/log4j/>.

Versionsinformationen

Wenn Sie ermitteln wollen, welche Version von Analytic Server installiert ist, können Sie den Ordner `{AS-Stammverzeichnis}/properties/version` prüfen. Die folgenden Dateien enthalten Versionsinformationen.

analytic_engine.swtag

Enthält detaillierte Produktinformationen.

version.txt

Versions- und Buildnummer des installierten Produkts.

Log Collector

Wenn Probleme durch Prüfung der Protokolldateien nicht direkt behoben werden können, können Sie alle Protokolle bündeln und an den IBM Support senden. Es wird ein Dienstprogramm bereitgestellt, das die Erfassung aller erforderlichen Daten erleichtert.

Führen Sie die folgenden Befehle über eine Befehlshell aus:

```
cd <AS-Stammverzeichnis>/tools/support/logcollector  
run >sh ./logcollector.sh
```

Diese Befehle erstellen unter <AS-Stammverzeichnis>/tools/support/logcollector eine komprimierte Datei. Die komprimierte Datei enthält alle Protokolldateien und Produktversionsinformationen.



Gedruckt in Deutschland