

IBM SPSS Analytic Server
Version 1

Instructions d'installation



Table des matières

Avis aux lecteurs canadiens	v
Instructions d'installation d'IBM SPSS Analytic Server.	1
Plateformes prises en charge	1
Prérequis	1
Installation d'Analytic Server avec le programme d'installation	1
Installation manuelle sur Hadoop	2
config.properties	4
Ajout de pilotes JDBC	5
Configuration de WebSphere Liberty	5
Registre de base	6
Configuration du registre LDAP.	6
Authentification avec un fournisseur de sécurité Kerberos.	6
Activation du mode haute disponibilité dans Hadoop 2.0.	8
Activation des sources de données HCatalog	8
Activation de la prise en charge d'Essentials for R.	9
Initiation des utilisateurs	10
Mise à jour d'Analytic Server	11
Désinstallation du serveur Analytic Server	11
Identification des problèmes.	12
Consignation	12
Informations de version	12
Collecteur de journal	12

Avis aux lecteurs canadiens

Le présent document a été traduit en France. Voici les principales différences et particularités dont vous devez tenir compte.

Illustrations

Les illustrations sont fournies à titre d'exemple. Certaines peuvent contenir des données propres à la France.

Terminologie

La terminologie des titres IBM peut différer d'un pays à l'autre. Reportez-vous au tableau ci-dessous, au besoin.

IBM France	IBM Canada
ingénieur commercial	représentant
agence commerciale	succursale
ingénieur technico-commercial	informaticien
inspecteur	technicien du matériel

Claviers

Les lettres sont disposées différemment : le clavier français est de type AZERTY, et le clavier français-canadien de type QWERTY.

OS/2 et Windows - Paramètres canadiens

Au Canada, on utilise :

- les pages de codes 850 (multilingue) et 863 (français-canadien),
- le code pays 002,
- le code clavier CF.

Nomenclature

Les touches présentées dans le tableau d'équivalence suivant sont libellées différemment selon qu'il s'agit du clavier de la France, du clavier du Canada ou du clavier des États-Unis. Reportez-vous à ce tableau pour faire correspondre les touches françaises figurant dans le présent document aux touches de votre clavier.

France	Canada	Etats-Unis
 (Pos1)		Home
Fin	Fin	End
 (PgAr)		PgUp
 (PgAv)		PgDn
Inser	Inser	Ins
Suppr	Suppr	Del
Echap	Echap	Esc
Attn	Intrp	Break
Impr écran	ImpEc	PrtSc
Verr num	Num	Num Lock
Arrêt défil	Défil	Scroll Lock
 (Verr maj)	FixMaj	Caps Lock
AltGr	AltCar	Alt (à droite)

Brevets

Il est possible qu'IBM détienne des brevets ou qu'elle ait déposé des demandes de brevets portant sur certains sujets abordés dans ce document. Le fait qu'IBM vous fournisse le présent document ne signifie pas qu'elle vous accorde un permis d'utilisation de ces brevets. Vous pouvez envoyer, par écrit, vos demandes de renseignements relatives aux permis d'utilisation au directeur général des relations commerciales d'IBM, 3600 Steeles Avenue East, Markham, Ontario, L3R 9Z7.

Assistance téléphonique

Si vous avez besoin d'assistance ou si vous voulez commander du matériel, des logiciels et des publications IBM, contactez IBM direct au 1 800 465-1234.

Instructions d'installation d'IBM SPSS Analytic Server

Plateformes prises en charge

Tableau 1. Environnements

Produit	Edition	Processeur	Taille de mot
Red Hat Enterprise Linux	6.x	x64	64 bits
Red Hat Enterprise Linux	5.x	x64	64 bits
SuSE Linux Enterprise Server	11	x64	64 bits
SuSE Linux Enterprise Server	10	x64	64 bits

Prérequis

- L'une des distributions suivantes :
 - Apache Hadoop 0.20.2
 - Apache Hadoop 1.0.3
 - IBM BigInsights 2.0
 - Cloudera CDH3
 - Cloudera CDH4 MRv1
 - Hortonworks Data Platform 1.x
- Pour que le programme d'installation puisse copier les fichiers nécessaires à partir de la distribution Hadoop, vous devez avoir installé ssh et scp et être en mesure de vous connecter avec une authentification par clé publique sans nécessiter de mot de passe pour la connexion. Autrement, vous devrez copier manuellement les fichiers. Afin de vérifier si votre connexion ssh fonctionne :
 1. Connectez-vous au serveur du noeud Name Hadoop avec une authentification par clé publique.
 2. Exécutez le code ci-dessous, en remplaçant {SERVEUR_HADOOP} par votre serveur Hadoop.

```
export HADOOP_SERVER_USERNAME='whoami'  
export HADOOP_SERVER={SERVEUR_HADOOP}  
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. Si le système renvoie "HOST_SSH_OK", l'opération a réussi. Dans le cas contraire, la configuration est incorrecte.

Installation d'Analytic Server avec le programme d'installation

Le programme d'installation automatise un certain nombre d'étapes habituellement manuelles. Le programme d'installation doit avoir accès à l'environnement HDFS. Pour installer Analytic Server :

1. Exécutez le fichier du programme d'installation `install.bin`.
 - a. Consultez le contrat de licence et acceptez-le pour poursuivre l'installation.
 - b. Sélectionnez un chemin d'installation.
 - c. Suivez les invites pour installer les fichiers sur votre système.
 - d. Configurez les options d'Analytic Server. Le nom du compte administrateur (VALEUR PAR DEFAUT : admin) est le nom d'utilisateur qui possède les droits d'administrateur "superutilisateur" pour l'application Analytic Server via la console.

- e. Indiquez le mot de passe de fichier de clés. Ce fichier de clés est utilisé pour stocker la clé de chiffrement utilisée lors du chiffrement des mots de passe des sources de données de la base de données enregistrés dans la base de données d'Analytic Server.
- f. Détectez automatiquement la version Hadoop.

Oui Indiquez le serveur du noeud Name Hadoop (VALEUR PAR DEFALUT : localhost). Si Hadoop se trouve sur la machine sur laquelle ce produit est installé, utilisez localhost comme nom de serveur. Le programme d'installation détermine alors quelle version de Hadoop est installée et obtient les fichiers requis de la distribution Hadoop. Remarque : vous devez avoir installé ssh et scp et être en mesure de vous connecter avec une authentification de clé publique sans nécessiter de mot de passe pour la connexion.

Non

- 1) Sélectionnez la version du serveur Hadoop dans une liste.
 - 2) Indiquez l'emplacement des fichiers requis de la distribution Hadoop ou copiez manuellement les fichiers par la suite.
 - 3) Si la distribution est Cloudera CDH3, spécifiez l'emplacement d'Oracle Java Runtime Environment 1.6 ou configurez-le manuellement par la suite.
- g. Configurez les paramètres Hadoop.

Hadoop user name

Nom d'utilisateur UNIX avec lequel Analytic Server s'exécute.

Password

Mot de passe de l'utilisateur Hadoop.

Analytic Server root directory

Emplacement de configuration de l'espace de travail Analytic Server et d'enregistrement des fichiers exécutables distribués aux noeuds Hadoop.

Namenode host

Nom ou adresse IP de l'hôte du noeud Name HDFS.

Namenode port

Numéro de port de l'hôte du noeud Name.

Job tracker

Nom ou adresse IP de l'hôte du dispositif de suivi des travaux Hadoop.

Job tracker port number

Numéro de port du dispositif de suivi des travaux.

- 2. Configurez WebSphere Liberty pour enregistrer les utilisateurs. Pour plus d'informations, voir «Configuration de WebSphere Liberty», à la page 5.
- 3. Démarrez le serveur en exécutant {RACINE_AS}/bin/start.sh, où {RACINE_AS} désigne le dossier de déploiement d'Analytic Server, par exemple : /opt/IBM/SPSS/AnalyticServer/1.0.
- 4. Ouvrez un navigateur Web et entrez l'adresse http://localhost:8080. Cette adresse URL affiche la boîte de dialogue de connexion de la console Analytic Server. Connectez-vous comme le "superutilisateur" défini plus haut. Pour plus d'informations sur la gestion des titulaires, voir le *Guide d'administration d'IBM® SPSS Analytic Server*.

Si le programme d'installation ne parvient pas à se connecter au serveur Hadoop ou qu'une version d'Hadoop est introuvable, suivez la procédure manuelle.

Installation manuelle sur Hadoop

- 1. Exécutez install.bin. La base de données Analytic Server est créée et les données d'identification de l'utilisateur administrateur sont définies.

- a. {RACINE_AS} désigne le dossier de déploiement d'Analytic Server : par exemple, /opt/IBM/SPSS/AnalyticServer/1.0.
 - b. {RACINE_SERVEUR_AS} désigne l'emplacement des fichiers de configuration, des fichiers journaux et des fichiers serveur : par exemple, /opt/IBM/SPSS/AnalyticServer/1.0/ae_wlpserver/usr/servers/aeserver.
 - c. {DOSSIER_PRINCIPAL_AS} désigne le dossier sur HDFS utilisé par Analytic Server comme dossier principal.
2. Modifiez le fichier de configuration {RACINE_SERVEUR_AS}/configuration/config.properties. Voir «config.properties», à la page 4 pour plus d'informations. Configurez les paramètres Hadoop (**hdfs.namenode.url**, **hdfs.user**, **hdfs.password**) en fonction de la configuration Hadoop spécifique de votre système.
 3. Analytic Server s'exécute sur IBM WebSphere Liberty, qui est configuré par défaut pour s'exécuter sur le port 8080.
 4. Copiez les fichiers JAR d'Hadoop requis dans {RACINE_SERVEUR_AS}/apps/AE_BOOT.war/WEB-INF/lib. Ces fichiers JAR sont basés sur votre installation et se trouvent sur le noeud Name Hadoop du cluster d'installation d'Hadoop, par exemple, dans les dossiers /usr/lib/hadoop et /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

La mise à jour Cloudera CDH3 5 et version ultérieure implique d'ajouter des fichiers JAR supplémentaires

- 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 nécessite une machine virtuelle Oracle JVM 1.6. Mettez à jour le script setenv.sh dans le répertoire {AE_ROOT}/bin en mettant à jour JAVA_HOME, pour remplacer le chemin par défaut par une machine virtuelle Java (Oracle).
 6. Si vous utilisez Apache Hadoop 1.0.3 ou une version ultérieure, remplacez commons-io-1.4.jar par commons-io-2.x.jar dans {RACINE_SERVEUR_AS}/apps/AE_BOOT.war/WEB-INF/lib.
 7. Si la version Hadoop est la version 2.0 ou une version ultérieure, copiez {RACINE_AS}/cloudera/cloudera4/hadoopmapreduce_1.1-1.0.0.0.jar dans {RACINE_SERVEUR_AS}/apps/AE_BOOT.war/WEB-INF/lib et supprimez hadoopmapreduce-1.0.0.0.jar.
 8. Assurez-vous que HFDS est en cours d'exécution et accessible. Copiez les fichiers exécutables Analytic Server sur HDFS et initialisez l'espace de travail Analytic Server en exécutant {RACINE_AS}/bin/hdfsUpdate.sh.
 9. Démarrez le serveur en exécutant {RACINE_AS}/bin/run.sh
 10. Entrez http://localhost:8080 dans votre navigateur pour ouvrir la console Analytic Server. Connectez-vous avec le paramètre admin.username défini dans config.properties. Pour plus d'informations sur la gestion des titulaires, voir le *Guide d'administration d'IBM SPSS Analytic Server*.
 11. Pour arrêter le serveur, exécutez {RACINE_AS}/bin/stop.sh
 12. Le journal serveur se trouve sous : {RACINE_SERVEUR_AS}/logs/

config.properties

Le fichier config.properties contient différents paramètres Analytic Server. Les paramètres suivants correspondent aux paramètres à spécifier lors de l'exécution du programme d'installation. Vous devez confirmer que les paramètres par défaut sont acceptables lorsque vous effectuez une installation manuelle du produit.

root.folder.path

Emplacement d'enregistrement des artefacts utilisateur d'Analytic Server (entrées et sorties).

admin.consumer=ibm

Titulaire du compte administrateur par défaut. Les titulaires offrent une division de haut niveau des rôles, droits d'accès, projets et sources de données.

admin.username=admin

Nom du compte administrateur par défaut. Ce nom d'utilisateur dispose des droits d'administrateur pour l'application Analytic Server via la console.

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

URL du noeud Name Hadoop.

hdfs.user=hdadmin

Nom d'utilisateur Hadoop. Le nom d'utilisateur UNIX avec lequel Analytic Server s'exécute.

hdfs.password=<ENTER PASSWORD>

Mot de passe de l'utilisateur spécifié dans le paramètre hdfs.user.

hdfs.classpath.folder=/user/hdadmin/classpath

Répertoire racine d'Analytic Server. Emplacement de configuration de l'espace de travail Analytic Server et d'enregistrement des fichiers exécutables distribués aux noeuds Hadoop.

mapred.job.tracker=hadoop:54311

Nom ou adresse IP et numéro de port de l'hôte du dispositif de suivi des travaux Hadoop.

encryption.keystore.password

Ce fichier de clés est utilisé pour stocker la clé de chiffrement utilisée lors du chiffrement des mots de passe des sources de données de la base de données enregistrés dans la base de données d'Analytic Server.

Ajout de pilotes JDBC

Pour prendre en charge les sources de données de bases de données, vous devez ajouter les pilotes JDBC à Analytic Server.

1. Arrêtez Analytic Server en exécutant `{RACINE_AS}/bin/stop.sh`
2. Copiez les fichiers jar du pilote JDBC requis dans `{RACINE_AS}/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib`
3. Mettez à jour Analytic Server en exécutant `{RACINE_AS}/bin/hdfsUpdate.sh`
4. Démarrez Analytic Server en exécutant `{RACINE_AS}/bin/start.sh`

Tableau 2. Bases de données prises en charge

Base de données	Versions prises en charge	Fichiers jar du pilote JDBC	Fournisseur
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

Configuration de WebSphere Liberty

WebSphere Liberty Profile est une implémentation allégée d'IBM WebSphere. Analytic Server peut utiliser la sécurité d'application WebSphere pour authentifier les utilisateurs. Elle est configurée dans le fichier `server.xml` du serveur Analytic Server déployé. Pour activer la sécurité d'application dans Liberty, la fonction `appSecurity-1.0` doit être incluse dans le paramètre `featureManager` :

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

Pour activer SSL sur un serveur, vous devez inclure la fonction SSL dans le fichier `server.xml` :

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

Pour obtenir des informations détaillées sur la sécurité WebSphere, voir ftp://ftp.software.ibm.com/software/websphere/appserv/library/v85/was85base_security.pdf.

Registre de base

Le registre de base permet à l'administrateur de définir une base de données d'utilisateurs et de groupes au sein du fichier {RACINE_SERVEUR_AS}/server.xml. Les mots de passe peuvent être codés afin de masquer leur valeur avec l'outil securityUtil situé sous {RACINE_AS}/ae_wlpserver/bin.

Le registre de base est utile dans un environnement bac à sable, mais il n'est pas recommandé dans un environnement de production.

```
<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>
```

Configuration du registre LDAP

Le registre LDAP fournit à l'administrateur un moyen d'authentifier les utilisateurs avec un serveur LDAP externe tel qu'Active Directory ou OpenLDAP. Voici un exemple de registre LDAP pour 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>
```

Pour obtenir d'autres exemples de configurations, consultez le dossier des modèles {RACINE_AS}/ae_wlpserver/templates/config.

Authentification avec un fournisseur de sécurité Kerberos

Avant de pouvoir configurer Kerberos, vous devez obtenir les informations suivantes de votre administrateur Hadoop :

1. Domaine Kerberos, par exemple : ASSSO.COM
2. Nom d'hôte du centre de distribution de clés KDC) Kerberos, par exemple : kdc.assso.com
3. Principal Kerberos du noeud Name. Par exemple, hdfs/namenode.assso.com@ASSSO.COM.
4. Principal Kerberos du noeud MapReduce. Par exemple, mapred/jobtracker.assso.com@ASSSO.COM.

Vous devez ensuite configurer le fichier `krb5.conf` sous `/etc/krb5.conf`, par exemple :

```
[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
```

Une fois le fichier `krb5.conf` créé, modifiez le fichier `config.properties` comme suit :

1. Dans la section des modules Analytic Server :
 - Ajoutez les modules `hdfsauth` et `kerberossecurityprovider`.
 - Supprimez le module `wssecurityprovider`.

2. Ajoutez les propriétés de configuration suivantes :

```
#Kerberos authentication parameters
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

où
```

hadoop.security.authentication

Authentification de sécurité Hadoop. Entrez `kerberos` pour activer le fournisseur de sécurité Kerberos.

dfs.namenode.kerberos.principal

Principal Kerberos utilisé pour le fichier `keytab`, qui sert à démarrer le noeud Name.

mapreduce.jobtracker.kerberos.principal

Principal Kerberos utilisé pour le fichier `keytab`, qui sert à démarrer le programme de suivi des tâches.

java.security.krb5.conf

Emplacement du fichier de configuration Kerberos.

3. Configurez le référentiel d'utilisateurs Liberty LDAP dans `service.xml`. Voir «Configuration de WebSphere Liberty», à la page 5. Tous les utilisateurs spécifiés dans le référentiel d'utilisateurs Liberty doivent correspondre aux comptes utilisateurs Kerberos et le même paramètre LDAP doit être utilisé dans le serveur Kerberos.
4. Par défaut, Analytic Server utilise le répertoire `.temp` sous le répertoire principal de l'utilisateur comme répertoire temporaire, mais si vous souhaitez configurer le répertoire temporaire à un autre emplacement, suivez ces instructions.
 - a. Modifiez le fichier `config.properties` et supprimez la mise en commentaire du paramètre de configuration suivant.

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

En fonction des besoins, modifiez le paramètre sur le chemin absolu du répertoire temporaire. Aucune modification n'est nécessaire si le répertoire `/.temp` est utilisé.

- b. Modifiez les droits d'accès sur ce dossier de manière à permettre à tous les utilisateurs d'y accéder.
Par exemple, `hadoop fs -chmod 777 /.temp.`
- 5. Accordez les droits d'accès en lecture sur les dossiers `classpath` et `configuration` pour tous les utilisateurs Kerberos.
 - a. Ouvrez le fichier `config.properties` et notez les valeurs des paramètres **`hdfs.classpath.folder`** et **`component.framework.bin.path`**.
 - b. Définissez les droits d'accès en lecture, comme suit.
Si **`hdfs.classpath.folder`**=`/user/hdpadmin/classpath` et **`component.framework.bin.path`**=`/user/hdpadmin/configuration`, exécutez


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

Activation du mode haute disponibilité dans Hadoop 2.0

Analytic Server prend en charge l'exécution des modes Hadoop personnalisés, comme le mode haute disponibilité, en indiquant la configuration API client du cluster dans le dossier de configuration Analytic Server. Il est uniquement obligatoire de fournir la configuration API client du cluster Hadoop à Analytic Server en présence de paramètres personnalisés, comme haute disponibilité. Ces paramètres sont requis pour que l'API d'Analytic Server puisse répondre aux requêtes en cas d'échec d'un noeud Name de cluster.

Pour configurer le mode personnalisé suite à une installation réussie d'Analytic Server :

1. Obtenez les fichiers de configuration client du cluster (`hdfs-site.xml` et `core-sites.xml`) généralement sauvegardés sous `/etc/hadoop/conf` sur une machine noeud Name.
2. Copiez ces fichiers dans `{RACINE_AS}/ae_wlpserver/usr/servers/aeserver/configuration/hadoop-conf`.
3. Si le cluster Hadoop est configuré en mode haute disponibilité, assurez-vous que la propriété de configuration Analytic Server (dans le fichier `config.properties`) `hdfs.namenode.url` pointe vers le nom du service HDFS et `/user/Username`. Par exemple, `hdfs://nameservice1/user/hdpadmin`.
4. Si vous avez configuré le mode haute disponibilité pour le service de suivi des tâches, mettez à jour la propriété **`mapred.job.tracker`** dans `config.properties` afin de pointer vers le nom du service HDFS.
5. Mettez à jour le système de fichiers Hadoop en exécutant la commande suivante :


```
{RACINE_AS}/bin/hdfsUpdate.sh
```
6. Démarrez Analytic Server en exécutant la commande suivante :


```
{RACINE_AS}/bin/run.sh
```

Activation des sources de données HCatalog

Pour configurer Analytic Server pour une utilisation avec les bases de données HCatalog après une installation réussie d'Analytic Server :

1. Ajoutez une entrée `hcataloginput` à la liste des modules Analytic Server dans le fichier `{RACINE_AS}/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`. Par exemple :


```
ae.modules=securityprovidermanager,\
localClient,\
componentframework,\
...
hcataloginput
```
2. Supprimez la mise en commentaire ou ajoutez les lignes suivantes à `config.properties`

```
hive.metastore.local=false
hive.metastore.uris=thrift://nom_hôte:num_port
```

où

nom_hôte

Nom de la machine hébergeant le serveur Thrift

num_port

Numéro de port utilisé dans le script d'installation de HCatalog

3. Assurez-vous que les fichiers suivants sont disponibles dans le répertoire {RACINE_AS}/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib et copiez également ces fichiers dans le répertoire HDFS /user/{ae_admin}/classpath.

Les fichiers JAR ci-dessous concernent HCatalog 0.4.0 et Hive 0.9.0. Vous devez récolter les fichiers JAR HCatalog, Hive et dépendants correspondants pour les autres versions.

- Le fichier suivant peut être copié de l'installation du serveur HCatalog.
hcatalog-0.4.0.jar
- Les fichiers suivants peuvent être copiés de l'installation du serveur Hive.

```
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
```

- Les fichiers suivants peuvent être copiés de l'installation Hadoop 1.X ou de <http://jackson.codehaus.org/>.
jackson-core-asl-1.8.8.jar
jackson-mapper-asl-1.8.8.jar

Les fichiers JAR ci-dessous concernent HCatalog 0.5.0.

- Les fichiers ci-dessous peuvent être copiés à partir de <http://code.google.com/p/guava-libraries/wiki/Release13>.
guava-13.0.1.jar

4. Si vous envisagez d'utiliser des types complexes Hive (cartes, tableaux, structures), copiez ces fichiers JAR dans le répertoire /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. Les sources des données HCatalog basées sur les fichiers de séquence compressés impliquent que les bibliothèques natives Hadoop soient disponibles sur la machine Analytic Server. Copiez les bibliothèques natives Hadoop disponibles sous {HADOOP}/lib/native/Linux-amd64-64 dans un répertoire sur la machine Analytic Server et modifiez la variable **LIB_PATH** dans {RACINE_AS}/bin/start.sh pour inclure le chemin d'accès de ce répertoire. Par exemple :

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

Remarque : Si l'installation Analytic Server comporte des versions plus récentes de ces fichiers JAR, vous ne devez pas copier les anciennes versions. Par exemple, si le programme d'installation d'Analytic Server a déjà copié jackson-core-asl-1.8.0 à partir d'une distribution Hadoop plus récente, vous ne devez pas copier le fichier JAR jackson-core-asl-1.7.3.

Activation de la prise en charge d'Essentials for R

Analytic Server prend en charge l'évaluation de modèles R et l'exécution de scripts R.

Pour configurer la prise en charge de R suite à une installation réussie d'Analytic Server :

1. Installez R Engine sur le serveur hébergeant Analytic Server, comme suit.


```

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 -xzf R-2.15.2.tar.gz
cd R-2.15.2
./configure --enable-R-shlib
make

```

2. Installez Essentials for R sur le serveur hébergeant Analytic Server en exécutant le fichier du programme d'installation `install.bin` conformément aux instructions qui s'affichent à l'écran. Le programme d'installation :
 - a. Met à jour l'installation de R Engine sur Analytic Server, ajoutez le "R plug-in" et
 - b. Met à jour le répertoire `{RACINE_AS}/ae_wlpserver/usr/servers/aeserver/configuration/ext_64/bin` pour ajouter une bibliothèque native et un fichier de configuration au module `pasw.rstats`.
3. Déployez R Engine et R Component sur Hadoop.
 - a. Si Analytic Server et tous les noeuds Hadoop possèdent la même version de système d'exploitation et la même architecture de processeur :
 - 1) Créez une archive pour R Engine à l'aide du script suivant


```

#!/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) Copiez l'archive `R.zip` dans le répertoire `{RACINE_AS}/ae_wlpserver/usr/servers/aeserver/configuration/app_64`.
 - b. Exécutez `{RACINE_AS}/bin/hdfsUpdate.sh` pour propager les modifications sur HDFS.
 - c. Si les versions des systèmes d'exploitation d'Analytic Server et des noeuds Hadoop diffèrent, installez R Engine et Essentials for R sur chaque noeud Hadoop et dans un répertoire portant le même nom que le répertoire d'installation de R Engine sur Analytic Server. Lors de l'installation d'Essentials for R sur le noeud Hadoop, vous devez simplement spécifier l'emplacement de R Engine et ignorer l'étape de spécification de l'emplacement de `../ext_64/bin`.

Remarque : R Engine doit être installé à un emplacement accessible à tous les utilisateurs, car R est exécuté sur le cluster Hadoop comme un autre utilisateur que l'utilisateur Analytic Server.

Vous devez également installer Essentials for R sur la machine hébergeant SPSS Modeler. Pour plus d'informations, voir les *Instructions d'installation des mises à jour et extensions d'IBM SPSS Modeler pour IBM SPSS Analytic Server*.

Initiation des utilisateurs

Demandez aux utilisateurs d'accéder à `http://<hôte>:<port>/<titulaire>` et d'entrer leur nom et leur mot de passe pour se connecter à la console Analytic Server.

<hôte>

Adresse de l'hôte Analytic Server

<port>

Port sur lequel Analytic Server écoute

<titulaire>

Dans un environnement à service partagé, le titulaire auquel vous appartenez. Dans un environnement à service exclusif, le titulaire par défaut est **ibm**.

Pour accéder à IBM SPSS Analytic Catalyst, accédez à `http://<hôte>:<port>/catalyst.html` et entrez votre nom d'utilisateur et votre mot de passe pour vous connecter.

Mise à jour d'Analytic Server

Lorsque vous effectuez une réinstallation à l'aide du programme d'installation, vous avez la possibilité de mettre à jour les options Analytic Server, le mot de passe du fichier de clés et les paramètres de configuration Hadoop.

Mise à jour de la version du serveur Hadoop

Il peut être nécessaire, parfois, de mettre à jour la version du serveur Hadoop sur lequel Analytic Server est exécuté, par exemple, pour la mise à niveau de Cloudera 3.x vers Cloudera 4.x. Dans ce cas, il est nécessaire de mettre à jour Analytic Server avec les fichiers JAR du client à partir d'une nouvelle version du serveur Hadoop.

Remarque : Cette procédure considère qu'aucune propriété de configuration liée à Hadoop, comme Hôte/Port noeud Name, Hôte/Port du dispositif de suivi des tâches ou Nom d'utilisateur/Mot de passe Hadoop, n'a été modifiée.

1. Arrêtez Analytic Server en exécutant la commande suivante.
`{RACINE_AS}/bin/stop.sh`
2. Suivez les instructions du fournisseur pour la mise à niveau du serveur Hadoop et assurez-vous qu'il est en cours d'exécution.
3. Copiez les fichiers JAR du client nécessaires dans un répertoire local d'Analytic Server. Ce répertoire sera appelé `{DOSSIER_SRC_HADOOP}`. La liste des fichiers JAR nécessaires dépend de la nouvelle version du serveur Hadoop. Elle est décrite dans la section «Installation manuelle sur Hadoop», à la page 2.
4. Exécutez le script de shell `install_hadoop_files` qui se trouve sous `{RACINE_AS}/bin/hadoop_setup`, en indiquant l'argument `{DOSSIER_SRC_HADOOP}`. Par exemple : `./install_hadoop_files.sh /opt/hadoop_src`.
 - a. Choisissez si vous souhaitez supprimer les fichiers Hadoop installés (recommandé).
 - b. Sélectionnez la version Hadoop que vous souhaitez utiliser.
5. Mettez à jour le système de fichiers Hadoop avec Analytic Server en exécutant la commande suivante.
`{RACINE_AS}/bin/hdfsUpdate.sh`
6. Redémarrez Analytic Server en exécutant la commande suivante.
`{RACINE_AS}/bin/run.sh`

Désinstallation du serveur Analytic Server

1. Exécutez `{RACINE_AS}/_uninstall/uninstaller.bin`.

Remarque : Cette opération ne supprime pas les données du dossier Analytic Server sur HDFS. Pour supprimer le dossier de base complet d'Analytic Server, supprimez le répertoire HDFS dans lequel sont enregistrés les artefacts utilisateur d'Analytic Server (entrées et sorties). Il s'agit du répertoire spécifié comme `root.folder.path` dans `{RACINE_SERVEUR_AS}/configuration/config.properties`.

Remarque : Si le programme de désinstallation ne démarre pas en raison d'une exception Java `OutOfMemoryException`, la taille maximale de segment de mémoire doit être augmentée. La propriété `lax.nl.java.option.java.heap.size.max` dans `{RACINE_AS}/_uninstall/uninstaller.bin.lax` définit la taille maximale de segment de mémoire, en octets, pour le programme d'installation appelé. Ce

nombre est toujours spécifié en octets et non en kilo-octets ou en méga-octets. Il est analogue au paramètre VM -mx ou Xmx. La valeur par défaut est 50331648 (48 Mo). Il est suggéré de modifier ce nombre en 134217728 (128 Mo), puis de démarrer le programme de désinstallation.

Identification des problèmes

Analytic Server fournit plusieurs outils pratiques pour l'identification des problèmes.

Consignation

Analytic Server crée des fichiers journaux et des fichiers de trace client pouvant être utiles pour diagnostiquer les problèmes. Si vous avez choisi l'installation par défaut de Liberty, les fichiers journaux se trouvent dans le répertoire {LIBERTY_PROFILE_FOLDER}/logs.

La configuration de consignation par défaut génère deux fichiers journaux qui se renouvellent quotidiennement.

ae.log

Ce fichier contient le récapitulatif de haut niveau des messages informatifs d'avertissement et d'erreur. Vérifiez d'abord ce fichier lorsque des erreurs se produisent au niveau du serveur ne pouvant pas être résolues à l'aide du message d'erreur s'affichant sur l'interface utilisateur.

ae_trace.log

Ce fichier contient toutes les entrées du fichier ae.log, ainsi que des informations supplémentaires qui s'adressent essentiellement à l'équipe de support et de développement d'IBM à des fins de débogage.

Analytic Server utilise Apache LOG4J comme application de consignation sous-jacente. A l'aide de LOG4J, vous pouvez ajuster la consignation de manière dynamique en éditant le fichier de configuration {RACINE_SERVEUR_AS}/configuration/log4j.xml. Vous serez peut-être amené à le faire à la demande du support pour vous aider à diagnostiquer les problèmes ou pour limiter le nombre de fichiers journaux conservés. Les modifications apportées à ce fichier sont détectées automatiquement en quelques secondes de sorte qu'il n'est pas nécessaire de redémarrer Analytic Server.

Pour plus d'informations sur log4j et le fichier de configuration, voir la documentation sur le site officiel d'Apache à l'adresse <http://logging.apache.org/log4j/>.

Informations de version

Pour savoir quelle version d'Analytic Server est installée, reportez-vous au dossier {RACINE_AS}/properties/version. Les fichiers suivants contiennent des informations sur la version.

analytic_engine.swtag

Contient les infos produit détaillées.

version.txt

Version et numéro de compilation du produit installé.

Collecteur de journal

Lorsqu'il est impossible de résoudre des problèmes en consultant directement les fichiers journaux, vous pouvez regrouper tous les journaux et les envoyer au support IBM. Il existe un utilitaire permettant de simplifier la collecte de toutes les données nécessaires.

A l'aide d'un interpréteur de commandes, exécutez les commandes suivantes :

```
cd <RACINE_AS>/tools/support/logcollector
run >sh ./logcollector.sh
```

Ces commandes créent un fichier compressé sous <RACINE_AS>/tools/support/logcollector. Le fichier compressé contient tous les fichiers journaux et toutes les informations sur la version du produit.

