

IBM SPSS Analytic Server
version 3.1.0

Guide d'installation et de configuration

IBM

Important

Avant d'utiliser le présent document et le produit associé, prenez connaissance des informations générales figurant à la section «Remarques», à la page 65.

Cette édition s'applique à la version 3, édition 1, modification 0 de IBM SPSS Analytic Server, et à toutes les éditions et modifications ultérieures sauf mention contraire dans les nouvelles éditions.

LE PRESENT DOCUMENT EST LIVRE EN L'ETAT SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES INFORMATIONS EN CAS DE CONTREFACON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE.

Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. Les informations qui y sont fournies sont susceptibles d'être modifiées avant que les produits décrits ne deviennent eux-mêmes disponibles. En outre, il peut contenir des informations ou des références concernant certains produits, logiciels ou services non annoncés dans ce pays. Cela ne signifie cependant pas qu'ils y seront annoncés.

Pour plus de détails, pour toute demande d'ordre technique, ou pour obtenir des exemplaires de documents IBM, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial.

Vous pouvez également consulter les serveurs Internet suivants :

- <http://www.fr.ibm.com> (serveur IBM en France)
- <http://www.ibm.com/ca/fr> (serveur IBM au Canada)
- <http://www.ibm.com> (serveur IBM aux Etats-Unis)

*Compagnie IBM France
Direction Qualité
17, avenue de l'Europe
92275 Bois-Colombes Cedex*

Table des matières

Avis aux lecteurs canadiens	v	Configuration d'Apache Impala	39
Chapitre 1. Prérequis	1	Modification des ports utilisés par Analytic Server	41
Chapitre 2. Installation et configuration d'Ambari	3	Haute disponibilité d'Analytic Server	41
Conditions requises propres à Ambari	3	Optimisation des options JVM pour les petits volumes de données (Small Data)	41
Outils precheck et postcheck d'IBM SPSS Analytic Server	3	Migration	41
Installation dans Ambari	5	Désinstallation d'Analytic Server dans Cloudera	42
Installation hors ligne	7	Chapitre 4. Installation et configuration de MapR	43
Installation d'Analytic Server vis à vis d'un environnement MySQL géré de l'extérieur	10	Présentation de MapR	43
Configuration	11	Installation d'Analytic Server dans MapR	43
Sécurité	11	Configuration de MapR	46
Activation de la prise en charge d'Essentials for R	17	Activation de la répercussion de base de données	46
Activation des sources de base de données relationnelle	18	Activation d'Apache Hive	47
Activation des sources de données HCatalog	19	Activation d'Apache HBase	48
Modification des ports utilisés par Analytic Server	20	Activation d'Apache Spark	49
Haute disponibilité d'Analytic Server	20	Activation d'indicateurs de fonction	51
Optimisation des options JVM pour les petits volumes de données (Small Data)	20	Activation de R	51
Mise à jour des dépendances de client	21	Activation de LZO	52
Configuration d'Apache Knox	21	Configuration d'un cluster IBM SPSS Analytic Server pour MapR	52
Mise à niveau et migration	25	Désinstallation de MapR	52
Désinstallation	27	Migration d'IBM SPSS Analytic Server dans MapR	53
Désinstallation d'Essentials for R	28	Traitement des incidents de MapR	54
Chapitre 3. Installation et configuration de Cloudera	29	Chapitre 5. Installation et configuration de Huawei FusionInsight HD	55
Présentation de Cloudera	29	Présentation de FusionInsight HD	55
Conditions requises propres à Cloudera	29	Installation dans Huawei FusionInsight HD	55
Configuration de MySQL pour Analytic Server	29	Chapitre 6. Configuration d'IBM SPSS Modeler pour son utilisation avec IBM SPSS Analytic Server	59
Installation dans Cloudera	30	Chapitre 7. Traitement des incidents	61
Configuration de Cloudera	32	Remarques	65
Sécurité	32	Marques	67
Activation de la prise en charge d'Essentials for R	37		
Activation des sources de base de données relationnelle	37		
Activation des sources de données HCatalog	38		

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
 (Post)		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.

Chapitre 1. Prérequis

Avant d'installer Analytic Server, consultez les informations suivantes.

Configuration système requise

Pour les informations les plus récentes sur la configuration système requise, reportez-vous au document Detailed system sur le site du Support technique IBM : <http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html>. Sur cette page :

1. Entrez SPSS Analytic Server comme nom de produit et cliquez sur **Search**.
2. Sélectionnez la version voulue et la portée du rapport, puis cliquez sur **Submit**.

Systèmes Power

Vérifiez que les compilateurs IBM XLC et XLF sont installés et inclus dans la variable PATH sur tous les hôtes dans le cluster.

Pour plus d'informations sur l'obtention d'une licence pour ces compilateurs, visitez les sites Web suivants :

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

Hive/HCatalog

Si vous comptez utiliser des sources de données NoSQL, configurez Hive et HCatalog pour accès distant. Vérifiez également que le fichier `hive-site.xml` contient une propriété `hive.metastore.uris` sous la forme `thrift://<nom_hôte>:<port>` qui pointe vers le serveur Thrift Hive Metastore actif. Pour plus d'informations, reportez-vous à la documentation de votre distribution Hadoop.

Référentiel de métadonnées

Par défaut, Analytic Server installe et utilise une base de données MySQL. Vous avez toutefois la possibilité de configurer Analytic Server afin d'utiliser une installation DB2 existante. Quel que soit le type de base de données choisi, celle-ci doit utiliser le codage UTF-8.

MySQL

Le jeu de caractères par défaut pour MySQL est fonction de la version et du système d'exploitation. Procédez comme suit pour déterminer si votre installation MySQL est configurée pour utiliser UTF-8.

1. Déterminez la version de MySQL via la commande :

```
mysql -V
```
2. Déterminez le jeu de caractères par défaut pour MySQL en exécutant la requête suivante depuis l'interface de ligne de commande MySQL :

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

Si le jeu de caractères correspond déjà à UTF-8, aucune autre modification n'est requise.

3. Déterminez l'interclassement par défaut pour MySQL en exécutant la requête suivante depuis l'interface de ligne de commande MySQL :

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

Si l'interclassement correspond déjà à UTF-8, aucune autre modification n'est requise.

4. Si le jeu de caractères ou l'interclassement par défaut n'est pas configuré pour utiliser UTF-8, reportez-vous à la documentation MySQL pour plus d'informations sur les modifications à apporter à `/etc/my.cnf` et redémarrez le démon MySQL pour appliquer le jeu de caractères UTF-8.

DB2 Pour plus d'informations sur la configuration de DB2, accédez au Knowledge Center : http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html.

Clusters à haute disponibilité

Équilibreur de charge

Votre cluster à haute disponibilité doit disposer d'un équilibreur de charge prenant en charge l'affinité de session. Analytic Server identifie les sessions avec le cookie "request-token". Celui-ci identifie une session pour la durée de connexion d'un utilisateur et son utilisation dans une affinité de session contrôlée par l'application. Consultez la documentation de votre équilibreur de charge spécifique pour plus d'informations sur sa prise en charge de l'affinité de session.

Chapitre 2. Installation et configuration d'Ambari

Conditions requises propres à Ambari

En plus des conditions requises générales, prenez connaissance des informations ci-après.

Services

Analytic Server est installé en tant que service Ambari. Avant d'installer Analytic Server, vous devez vous assurer que HDFS, YARN, MapReduce2, Hive et Zookeeper ont été ajoutés en tant que services Ambari.

Connexion SSH sans mot de passe

Configurez une connexion SSH sans mot de passe pour l'utilisateur root entre l'hôte Analytic Metastore et tous les hôtes dans le cluster.

Outils precheck et postcheck d'IBM SPSS Analytic Server

Présentation de l'outil precheck

L'outil de vérification préalable de l'installation d'Analytic Server (outil precheck) permet de limiter les problèmes d'installation et les erreurs d'exécution en identifiant les problèmes d'environnement potentiels avant l'installation d'Analytic Server.

L'outil precheck vérifie :

- Les versions du système d'exploitation et d'Ambari sur le système local
- Les paramètres ulimit du système d'exploitation sur le système local
- L'espace disque disponible sur le système local
- La version de Hadoop
- La disponibilité du service Ambari (HDFS, HCatalog, Spark, Hive, MapReduce, Yarn, Zookeeper, etc.)
- Les paramètres Ambari propres à Analytic Server

Remarque : L'outil precheck peut être exécuté après l'installation d'Analytic Server.

Présentation de l'outil postcheck

L'outil de vérification postérieure à l'installation d'Analytic Server (outil postcheck) identifie les problèmes de configuration à l'issue de l'installation d'Analytic Server en soumettant des demandes API REST pour traiter :

- Des données du système de fichiers HDFS
- Des données dans Hive/HCatalog
- Des données compressées (y compris deflate, bz2, snappy, cmx)

Remarque : cmx est pris en charge uniquement dans BigInsights.

- Des données avec PySpark
- Des données qui utilisent des composants SPSS natifs (y compris alm, tree, neuralnet, scoring, tascoreing)
- Des données avec MapReduce
- Des données avec MapReduce en mémoire

Emplacement des outils et prérequis

Les outils precheck et postcheck se trouvent dans les répertoires suivants :

- **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**

Les outils sont conditionnés dans le fichier `AnalyticServer-*.jar` (qui se trouve dans `/opt/cloudera/csd`).

Vous devez exécuter les outils en tant que superutilisateur et disposer de Python 2.6.X (ou suivante).

Avant d'installer Analytic Server, vous devez exécuter l'outil precheck sur tous les noeuds Ambari qui hébergent le service Analytic Server. L'exécution de l'outil sur un noeud différent requiert la copie de l'ensemble du répertoire `chktool` sur le noeud concerné.

Si l'outil precheck signale des échecs, vous devez les résoudre avant de passer à l'installation d'Analytic Server.

Le répertoire `chktool` est disponible après l'exécution du fichier binaire autoextractible d'Analytic Server (étape 2 de la section «Installation dans Ambari», à la page 5). Si vous décidez d'exécuter un «Installation hors ligne», à la page 7, le répertoire `chktool` est disponible après l'installation du fichier RPM de métadonnées.

Exécution de l'outil precheck

L'exemple suivant vérifie au préalable le cluster Ambari MyCluster qui s'exécute sur `myambarihost.ibm.com:8080` avec SSL activé et utilise les données d'identification et de connexion `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
```

Remarques :

- La valeur `as_host` doit correspondre à l'adresse IP ou par un nom de domaine complet.
- L'outil vous invite à saisir un mot de passe lorsque l'argument `-password` est omis.
- La commande `precheck.py` utilise l'aide sur la syntaxe, qui s'affiche avec l'argument `-h` (python `./precheck.py -help`).
- L'argument `-cluster` est facultatif. (Le cluster en cours est identifié lorsque l'argument `-cluster` n'est pas utilisé.)

Lorsque l'outil precheck effectue ses vérifications préalables, le statut de chaque vérification s'affiche dans la fenêtre de commande. En cas d'échec, des informations détaillées sont disponibles dans le fichier `journal`. (L'emplacement exact du fichier `journal` est indiqué dans la fenêtre de commande.) Le fichier `journal` peut être transmis au service de support technique IBM lorsqu'une assistance supplémentaire est nécessaire.

Exécution de l'outil postcheck

L'outil postcheck vérifie qu'Analytic Server s'exécute correctement et peut traiter des travaux simples. L'exemple postcheck suivant vérifie une instance Analytic Server qui s'exécute sur `myanalyticserverhost.ibm.com:9443` avec SSL activé, et utilise les données d'identification et de connexion `admin:ibmspss` :

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

Lorsque Knox est utilisé avec Analytic Server, la commande prend la forme suivante :

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

Pour effectuer une vérification unique, utilisez la commande suivante :

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

Remarques :

- L'outil vous invite à saisir un mot de passe lorsque l'argument `-password` est omis.
- La commande `postcheck.py` inclut l'aide sur la syntaxe, qui s'affiche avec l'argument `-h` (python `./postcheck.py -help`).

Lorsque l'outil `postcheck` effectue ses vérifications, le statut de chaque vérification s'affiche dans la fenêtre de commande. En cas d'échec, des informations détaillées sont disponibles dans le fichier journal. (L'emplacement exact du fichier journal est indiqué dans la fenêtre de commande.) Le fichier journal peut être transmis au service de support technique IBM lorsqu'une assistance supplémentaire est nécessaire.

Installation dans Ambari

Le processus de base consiste à installer les fichiers Analytic Server sur un hôte dans le cluster Ambari, puis à ajouter Analytic Server en tant que service Ambari. Les étapes sont détaillées ci-après.

Important : Analytic Server ne prend pas en charge l'installation dans un environnement dans lequel le serveur Ambari est exécuté par un utilisateur non superutilisateur.

1. Accédez au [site Web IBM Passport Advantage®](#) et téléchargez le fichier binaire autoextractible approprié pour votre pile, votre version de pile et votre architecture matérielle sur un hôte dans le cluster Ambari. Les fichiers binaires Ambari disponibles sont :

Tableau 1. Fichiers binaires autoextractibles Analytic Server

Description	Nom du fichier binaire
IBM® SPSS Analytic Server 3.1 pour BigInsights 4.1 et 4.3 Linux on System p LE (Anglais)	spss_as-3.1-bi4.1-4.3-lppc64le_en.bin
IBM SPSS Analytic Server 3.1 pour BigInsights 4.1, 4.2 et 4.3 Linux x86-64 (Anglais)	spss_as-3.1-bi4.1-4.2-4.3-1x86_en.bin
IBM SPSS Analytic Server 3.1 pour Hortonworks Data Platform 2.3 et 2.4 Ubuntu (Anglais)	spss_as-3.1-hdp2.3-2.4-ubun_en.bin
IBM SPSS Analytic Server 3.1 pour Hortonworks Data Platform 2.4, 2.5 et 2.6 Linux x86-64 (Anglais)	spss_as-3.1-hdp2.4-2.6-1x86_en.bin
IBM SPSS Analytic Server 3.1 pour Hortonworks Data Platform 2.6 Linux on System p LE (Anglais)	spss_as-3.1-hdp2.6-lppc64le_en.bin

2. Exécutez le fichier binaire autoextractible et suivez les instructions pour (si vous le souhaitez) afficher la licence, l'accepter et sélectionner une installation en ligne ou hors ligne.

Installation en ligne

Sélectionnez l'installation en ligne si l'hôte du serveur Ambari et tous les noeuds du cluster peuvent accéder à <https://ibm-open-platform.ibm.com>.

[GPFS (Spectrum Scale) uniquement] Téléchargez le fichier <https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms/IBM-SPSS-AnalyticServer-3.1.0.0.repo> (x86 et ppc64le) ou <https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/Ubuntu/IBM-SPSS-AnalyticServer.list> (Ubuntu) et placez-le dans le dossier `/etc/yum.repos.d`

(RHEL, CentOS), /etc/zypp/repos.d (SLES), ou /etc/apt/sources.list.d (Ubuntu) sur chaque noeud sur lequel vous ajoutez Analytic Server Metastore en tant que service.

Installation hors ligne

Sélectionnez l'installation hors ligne si votre serveur Ambari n'a pas accès à Internet. Pour plus d'informations, voir «Installation hors ligne», à la page 7.

3. Depuis le répertoire `var/lib/ambari-server/resources/stacks/<nom-pile>/<version-pile>/services/ANALYTICSERVER/package/scripts`, exécutez le script `update_clientdeps.sh` avec les arguments appropriés (utilisez l'argument `--help` pour des exemples).
4. Redémarrez votre serveur Ambari.
`ambari-server restart`
5. Si vous procédez à une installation hors ligne, vérifiez que les modifications de `reponfo.xml` et `IBM-SPSS-AnalyticServer-3.1.0.0.repo` ont été conservées. Contactez votre interlocuteur du support IBM si vous constatez que les modifications de `reponfo.xml` et `IBM-SPSS-AnalyticServer-3.1.0.0.repo` n'ont pas été conservées.
6. Connectez-vous à votre serveur Ambari et installez Analytic Server en tant que service via l'interface utilisateur Ambari.

Référentiel de métadonnées

Analytic Server utilise MySQL par défaut pour assurer le suivi des informations sur les sources de données, les projets et les titulaires. Au cours de l'installation, vous devez indiquer un nom d'utilisateur (**metadata.repository.user.name**) et un mot de passe **metadata.repository.password** utilisés pour la connexion JDBC entre Analytic Server et MySQL. Le programme d'installation crée l'utilisateur dans la base de données MySQL, mais cet utilisateur est propre à la base de données MySQL et n'a pas besoin d'être un utilisateur Linux ou Hadoop existant.

Pour remplacer le référentiel de métadonnées par DB2, procédez comme suit.

Remarque : Le référentiel de métadonnées ne peut plus être modifié une fois l'installation terminée.

- a. Vérifiez que DB2 est installé sur une autre machine. Pour plus d'informations, consultez la section relative au référentiel de métadonnées à la rubrique Chapitre 1, «Prérequis», à la page 1.
 - b. Dans l'onglet Ambari Services, naviguez jusqu'à l'onglet Configs du service Analytic Server.
 - c. Ouvrez la section **Advanced analytics-env**.
 - d. Remplacez la valeur de **as.database.type**, `mysql`, par `db2`.
 - e. Ouvrez la section **Advanced analytics-meta**.
 - f. Remplacez la valeur de **metadata.repository.driver**, `com.mysql.jdbc.Driver`, par `com.ibm.db2.jcc.DB2Driver`.
 - g. Remplacez la valeur de **metadata.repository.url** par `jdbc:db2://{DB2_HOST}:{PORT}/{DBName}:currentSchema={SchemaName};`, où
 - {DB2_HOST} est le nom d'hôte du serveur sur lequel est installé DB2
 - {PORT} est le port sur lequel DB2 écoute.
 - {SchemaName} est un schéma disponible, non utilisé.
- Si vous n'êtes pas sûr des valeurs à entrer, demandez à votre administrateur DB2.
- h. Entrez des données d'identification DB2 valides dans **metadata.repository.user.name** et **metadata.repository.password**.
 - i. Cliquez sur **Save**.

Paramètres de configuration à ne pas changer après l'installation

Ne modifiez pas les paramètres suivants après l'installation, faute de quoi Analytic Server ne fonctionnera pas.

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

7. Vous disposez maintenant d'une instance fonctionnelle d'Analytic Server. Les autres opérations de configuration sont facultatives. Pour plus d'informations sur la configuration et l'administration d'Analytic Server, consultez la rubrique : «Configuration», à la page 11. Pour plus d'informations sur la migration d'une configuration existante vers une nouvelle installation, reportez-vous à la rubrique : «Mise à niveau et migration», à la page 25.
8. Ouvrez un navigateur Web et entrez l'adresse `http://<hôte>:<port>/analyticserver/admin/ibm`, où <hôte> est l'adresse de l'hôte Analytic Server, et <port> est le port sur lequel écoute Analytic Server. Par défaut, la valeur est 9080. Cette adresse URL affiche la boîte de dialogue de connexion de la console Analytic Server. Connectez-vous comme administrateur Analytic Server. Par défaut, l'ID utilisateur est admin et le mot de passe est admin.

Installation hors ligne

Le flux de travaux général d'une installation hors ligne est le suivant :

1. Accédez au [site Web IBM Passport Advantage®](#) et téléchargez le fichier binaire autoextractible approprié pour votre pile, votre version de pile et votre architecture matérielle sur un hôte dans le cluster Ambari. Les fichiers binaires Ambari disponibles sont :

Tableau 2. Fichiers binaires autoextractibles Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.1 pour BigInsights 4.1 et 4.3 Linux on System p LE (Anglais)	spss_as-3.1-bi4.1-4.3-1ppc64le_en.bin
IBM SPSS Analytic Server 3.1 pour BigInsights 4.1, 4.2 et 4.3 Linux x86-64 (Anglais)	spss_as-3.1-bi4.1-4.2-4.3-1x86_en.bin
IBM SPSS Analytic Server 3.1 pour Hortonworks Data Platform 2.3 et 2.4 Ubuntu (Anglais)	spss_as-3.1-hdp2.3-2.4-ubun_en.bin
IBM SPSS Analytic Server 3.1 pour Hortonworks Data Platform 2.4, 2.5 et 2.6 Linux x86-64 (Anglais)	spss_as-3.1-hdp2.4-2.6-1x86_en.bin
IBM SPSS Analytic Server 3.1 pour Hortonworks Data Platform 2.6 Linux on System p LE (Anglais)	spss_as-3.1-hdp2.6-1ppc64le_en.bin

2. Exécutez le fichier binaire exécutable et indiquez une installation hors ligne. Une installation hors ligne télécharge les fichiers RPM ou DEB nécessaires et doit être exécutée sur un système qui peut accéder à <https://ibm-open-platform.ibm.com>. Le fichier binaire exécutable est situé dans les répertoires de distribution Ambari <AS_INSTALLABLE_HOME> disponibles.
3. Copiez l'ensemble du contenu du répertoire <AS_INSTALLABLE_HOME> approprié du système doté d'un accès à Internet sur le noeud du gestionnaire Ambari (derrière le pare-feu).
4. Continuez en suivant les instructions pour votre distribution (BigInsights/HDP ou Ubuntu) sur les noeuds des gestionnaires Ambari appropriés (derrière le pare-feu).

Instructions BigInsights et HDP - 3.1.0

Important : Analytic Server ne prend pas en charge l'installation dans un environnement dans lequel le serveur Ambari est exécuté par un utilisateur non superutilisateur.

1. Utilisez la commande suivante pour vérifier si le serveur Ambari est démarré :
`ambari-server status`
 Arrêtez le noeud de serveur Ambari (s'il est démarré) :
`ambari-server stop`
2. Installez l'outil qui crée un référentiel yum local.
`yum install createrepo (RHEL, CentOS)`
 ou
`zypper install createrepo (SLES)`
3. Créez un répertoire utilisé comme référentiel pour les fichiers RPM d'Analytic Server. Voir l'exemple ci-après.
`mkdir /home/root/repos/IBM-SPSS-AnalyticServer/x86_64`
4. Copiez les fichiers RPM d'Analytic Server nécessaires dans le nouveau répertoire. Les fichiers RPM dont vous avez besoin varient en fonction de votre distribution, de votre version et de votre architecture.

BigInsights 4.1, 4.2 et 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 et 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 et 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. Créez la définition du référentiel local. Par exemple, créez un fichier nommé `IBM-SPSS-AnalyticServer-3.1.0.0.repo` dans `/etc/yum.repos.d/` (pour RHEL, CentOS) ou `/etc/zypp/repos.d/` (pour SLES) avec le contenu suivant :

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer
baseurl=file:///{"chemin du référentiel local"}
enabled=1
gpgcheck=0
protect=1
```
6. Créez le référentiel yum local.
`createrepo /home/root/repos/IBM-SPSS-AnalyticServer/x86_64 (RHEL, CentOS, SLES)`

7. Depuis une fenêtre de commande de l'utilisateur root, exécutez la commande `cd` en pointant sur le répertoire `<AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer` et exécutez la commande `run ./offlineInstall.sh`. Le script lit les réponses conservées adressées à la commande d'installation du fichier exécutable binaire qui a été précédemment lancée et émet la commande pour la plateforme appropriée (pour installer le fichier rpm).

Remarque : Les étapes 8 et 9 ne s'appliquent que si vous utilisez un environnement MySQL géré en externe.

8. Exécutez le script `add_mysql_user.sh` sur le noeud/l'hôte sur lequel l'instance MySQL, qui sera utilisée comme valeur `AS_MetaStore`, est installée.
 - a. Copiez le script `add_mysql_user.sh` depuis `/opt/AS_Installable/IBM-SPSS-AnalyticServer` vers le noeud/hôte sur lequel l'instance MySQL qui sera utilisée comme `AS_MetaStore` est installée. Par exemple, `/opt/AS_InstallTools`.
 - Exécutez le script `add_mysql_user.sh` sur le noeud/hôte MySQL. Par exemple, `./add_mysql_user.sh -u as_user -p spss -d aedb`

Remarques :

- Le nom d'utilisateur et le mot de passe doivent correspondre au nom d'utilisateur de la base de données et au mot de passe saisi pour AS_Metastore sur l'écran de configuration Ambari.
- Vous pouvez, si vous le désirez, mettre à jour manuellement le script `add_mysql_user.sh` pour émettre des commandes.
- Lors de l'exécution du script `add_mysql_user.sh` vis à vis d'une base de données MySQL sécurisée (accès utilisateur root), utilisez les paramètres `-r` et `-t` pour transmettre les valeurs de `dbuserid` et `dbuserid_password`. Le script utilise `dbuserid` et `dbuserid_password` pour effectuer des opérations MySQL.

Remarque : Le paramètre `metadata.repository.url` sur l'écran **AS_Configuration (Advanced analytics-meta)** doit être modifié afin de pointer sur l'hôte de base de données MySQL. Par exemple, remplacez le paramètre `JDBC mysql://{analytic_metastore_host}/aedb?createDatabaseIfNotExist=true` par `mysql://{BD_MYSQL}/aedb?createDatabaseIfNotExist=true`

9. Mettez à jour le fichier référentiel Ambari `repoinfo.xml`, généralement stocké dans `/var/lib/ambari-server/resources/stacks/$stackName/$stackVersion/repos/`, pour utiliser le référentiel yum local en ajoutant les lignes suivantes :

```
<os type="host_os">
  <repo>
    <baseurl>file:///path to local repository/</baseurl>
    <repoid>IBM-SPSS-AnalyticServer</repoid>
    <reponame>IBM-SPSS-AnalyticServer-3.1.0.0</reponame>
  </repo>
</os>
```

Voici un exemple de {chemin du référentiel local} :

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

10. Répétez la procédure suivante sur tous les noeuds de cluster Ambari qui ne sont pas un serveur.
- a. Copiez l'ensemble du contenu du répertoire `<AS_INSTALLABLE_HOME>` approprié du système doté d'un accès à Internet sur le noeud de cluster non-serveur Ambari.
 - b. Installez l'outil qui crée un référentiel yum local.
`yum install createrepo (RHEL, CentOS)`
ou
`zypper install createrepo (SLES)`
 - c. Créez un répertoire utilisé comme référentiel pour les fichiers RPM d'Analytic Server. Voir l'exemple ci-après.
`mkdir /home/root/repos/IBM-SPSS-AnalyticServer/x86_64`
 - d. Copiez les fichiers RPM d'Analytic Server nécessaires dans le nouveau répertoire. Les fichiers RPM dont vous avez besoin varient en fonction de votre distribution, de votre version et de votre architecture.

BigInsights 4.1, 4.2 et 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 et 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 et 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. Créez la définition du référentiel local. Par exemple, créez un fichier nommé IBM-SPSS-AnalyticServer-3.1.0.0.repo dans /etc/yum.repos.d/ (pour RHEL, CentOS) ou /etc/zypp/repos.d/ (pour SLES) avec le contenu suivant :

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer
baseurl=file:///{chemin du référentiel local}
enabled=1
gpgcheck=0
protect=1
```

- f. Créez le référentiel yum local.

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

11. Passez à l'étape 5 de la rubrique «Installation dans Ambari», à la page 5.

Instructions pour Ubuntu - 3.1.0

1. Accédez au [site Web d'IBM Passport Advantage](#) et téléchargez le fichier binaire autoextractible Ubuntu approprié pour votre pilote, votre version et votre architecture matérielle sur un hôte du cluster Ambari. Les fichiers binaires disponibles sont :

Tableau 3. Fichiers binaires autoextractibles Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.1.0 pour Hortonworks Data Platform 2.3 et 2.5 Ubuntu (Anglais)	spss_as-3.1.0-hdp2.3-2.5-ubun_en.bin

2. Sur un système doté d'un accès à Internet, lancez le fichier binaire exécutable et indiquez une installation hors ligne. Une installation hors ligne télécharge les fichiers DEB nécessaires et doit être exécutée sur un ordinateur qui peut accéder à <https://ibm-open-platform.ibm.com>. Le fichier binaire exécutable se trouve dans le répertoire de distribution Ambari <AS_INSTALLABLE_HOME> disponible.
3. Les fichiers DEB Analytic Server se trouvent dans le répertoire suivant :
IBM-SPSS-AnalyticServer/packages
4. Utilisez la commande suivante pour installer Analytic Server 3.1.0 :
dpkg -i ./IBM-SPSS-AnalyticServer-ambari-HDP-2.5_3.1.0.0_amd64.deb (ou IBM-SPSS-AnalyticServer-ambari-HDP-2.3_3.1.0.0_amd64.deb)
dpkg -i ./IBM-SPSS-AnalyticServer_1_amd64.deb
5. Redémarrez votre serveur Ambari.
ambari-server restart
6. Connectez-vous à votre serveur Ambari et installez Analytic Server en tant que service via l'interface utilisateur Ambari.

Installation d'Analytic Server vis à vis d'un environnement MySQL géré de l'extérieur

La procédure d'installation d'Analytic Server diffère d'une installation normale si elle concerne un environnement MySQL géré de l'extérieur.

Les étapes ci-après décrivent la procédure d'installation d'Analytic Server vis à vis d'un environnement MySQL géré de l'extérieur.

1. Accédez au [site Web IBM Passport Advantage®](#) et téléchargez le fichier binaire autoextractible approprié pour votre pile, votre version de pile et votre architecture matérielle sur un hôte dans le cluster Ambari.
2. Exécutez le fichier binaire autoextractible et suivez les instructions pour (si vous le souhaitez) afficher la licence et l'accepter.
 - a. Sélectionnez l'option en ligne.
 - b. A l'invite, sélectionnez l'option **Base de données MySQL externe**.

3. Copiez le script `add_mysql_user.sh` depuis `/opt/AS_Installable/IBM-SPSS-AnalyticServer` vers le noeud/hôte sur lequel l'instance MySQL qui sera utilisée comme `AS_MetaStore` est installée. Par exemple, `/opt/AS_InstallTools`.
 - Exécutez le script `add_mysql_user.sh` sur le noeud/hôte MySQL. Par exemple, `./add_mysql_user.sh -u as_user -p spss -d aedb`

Remarques :

- Le nom d'utilisateur et le mot de passe doivent correspondre au nom d'utilisateur de la base de données et au mot de passe saisis pour `AS_Metastore` sur l'écran de configuration Ambari.
 - Vous pouvez, si vous le désirez, mettre à jour manuellement le script `add_mysql_user.sh` pour émettre des commandes.
 - Lors de l'exécution du script `add_mysql_user.sh` vis à vis d'une base de données MySQL sécurisée (accès utilisateur root), utilisez les paramètres `-r` et `-t` pour transmettre les valeurs de `dbuserid` et `dbuserid_password`. Le script utilise `dbuserid` et `dbuserid_password` pour effectuer des opérations MySQL.
4. Redémarrez votre serveur Ambari.
 5. Depuis la console Ambari, ajoutez le service `AnalyticServer` en tant que service normal (entrez le même nom d'utilisateur et mot de passe qu'à l'étape 3).

Remarque : Le paramètre `metadata.repository.url` sur l'écran **AS_Configuration (Advanced analytics-meta)** doit être modifié afin de pointer sur l'hôte de base de données MySQL. Par exemple, remplacez le paramètre `JDBC mysql://{analytic_metastore_host}/aedb?createDatabaseIfNotExist=true` par `mysql://{BD_MySQL}/aedb?createDatabaseIfNotExist=true`

Configuration

Après l'installation, vous pouvez configurer et administrer Analytic Server à l'aide de l'interface utilisateur d'Ambari.

Remarque : Les conventions suivantes sont utilisées pour les chemins de fichier Analytic Server.

- `{RACINE_AS}` se réfère à l'emplacement dans lequel Analytic Server est déployée ; par exemple, `/opt/IBM/SPSS/AnalyticServer/{version}`.
- `{RACINE_SERVEUR_AS}` se réfère à l'emplacement des fichiers de configuration, des journaux et du serveur. Par exemple, `/opt/IBM/SPSS/AnalyticServer/{version}/ae_wlpserver/usr/servers/aeserver`.
- `{DOSSIER_PRINCIPAL_AS}` désigne l'emplacement sur HDFS utilisé par Analytic Server en tant que dossier principal.

Sécurité

Le paramètre **security.config** définit le registre des utilisateurs et des groupes qui peuvent être définis comme principaux dans le système Analytic Server.

Par défaut, un registre de base est défini avec un seul utilisateur, `admin`, dont le mot de passe est `admin`. Vous pouvez modifier le registre en éditant **security.config** ou en configurant Kerberos. Le paramètre **security.config** se trouve dans la section **Advanced analytics.cfg** de l'onglet Configs du service Analytic Server.

Remarque : Si vous éditez le paramètre **security.config** pour modifier le registre, vous devez définir les nouveaux utilisateurs comme principaux dans le système Analytic Server. Voir le manuel *IBM SPSS Analytic Server - Guide d'administration* pour plus d'informations sur la gestion des titulaires.

Modification du registre de base

Le registre de base vous permet de définir une base de données d'utilisateurs et de groupes dans le paramètre **security.config**.

Le registre de base par défaut ressemble à l'exemple qui suit.

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

Exemple de registre de base modifié :

```
<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>
```

Les mots de passe peuvent être codés afin de brouiller leur valeur via l'outil securityUtility, situé sous le répertoire {RACINE_AS}/ae_wlpserver/bin.

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

Remarque : Voir http://www-01.ibm.com/support/knowledgecenter/SSD28V_8.5.5/com.ibm.websphere.wlp.core.doc/ae/rwlp_command_securityutil.html pour plus de détails sur l'outil securityUtility.

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

Configuration d'un registre LDAP

Le registre LDAP vous permet d'authentifier les utilisateurs via un serveur LDAP externe tel que Active Directory ou OpenLDAP.

Important : Un utilisateur LDAP doit être désigné comme administrateur Analytic Server dans Ambari.

Voici un exemple de registre LDAP pour OpenLDAP.

```
<ldapRegistry
  baseDN="ou=people,dc=aelldap,dc=org"
  ldapType="Custom"
  port="389"
  host="server"
  id="OpenLDAP"
  bindDN="cn=admin,dc=aelldap,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>
```

L'exemple suivant fournit une authentification d'Analytic Server dans Active Directory :

```
<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="(&(sAMAccountName=%v)(objectcategory=user))"
    groupFilter="(&(cn=%v)(objectcategory=group))"
    userIdMap="user:sAMAccountName"
    groupIdMap="*:cn"
    groupMemberIdMap="memberOf:member" />
</ldapRegistry>
```

Remarque : Il est souvent utile d'utiliser un outil d'afficheur LDAP tiers pour vérifier la configuration LDAP.

L'exemple suivant fournit une authentification du profil WebSphere Liberty dans Active Directory :

```
<ldapRegistry id="ldap" realm="SampleLdapADRealm"
  host="ldapserverserver.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">
  <activatedFilters
    userFilter="(&(sAMAccountName=%v)(objectcategory=user))"
    groupFilter="(&(cn=%v)(objectcategory=group))"
    userIdMap="user:sAMAccountName"
    groupIdMap="*:cn"
    groupMemberIdMap="memberOf:member" >
  </activatedFilters>
</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=" />
```

Remarques :

- La prise en charge de LDAP dans Analytic Server est régie par WebSphere Liberty. Pour plus d'informations, voir Configuration de registres utilisateur LDAP dans Liberty.
- Une fois LDAP sécurisé avec SSL, suivez les instructions de la section "Configuration d'une connexion SSL (Secure Socket Layer) entre Analytic Server et LDAP".

Configuration d'une connexion SSL (Secure Socket Layer) entre Analytic Server et LDAP

1. Connectez-vous à toutes les machines Analytic Server en tant qu'utilisateur Analytic Server et créez un répertoire commun pour les certificats SSL.

Remarque : Par défaut, `as_user` est l'utilisateur Analytic Server (voir **Service accounts** dans l'onglet Admin de la console Ambari).

2. Copiez le magasin de clés et le magasin de clés de confiance dans le répertoire commun de toutes les machines Analytic Server. Ajoutez également le certificat de l'autorité de certification du client LDAP au magasin de clés de confiance. Par exemple :

```

mkdir /home/as_user/security
cd /home/as_user/security
openssl s_client -connect <nom d'hôte ldap>:636 -showcerts > client.cert
$JAVA_HOME/bin/keytool -import -file ./client.cert -alias ldapCA -keystore mytrust.jks
password : changeit

```

Remarque : JAVA_HOME est l'environnement d'exécution Java utilisé au démarrage d'Analytic Server.

3. Les mots de passe peuvent être codés afin de brouiller leur valeur via l'outil securityUtility, situé sous le répertoire {RACINE_AS}/ae_wlpserver/bin. Par exemple :

```

securityUtility encode changeit
{xor}PDC+MTg6Nis=

```

4. Connectez-vous à la console Ambari et définissez la valeur adéquate pour SSL dans le paramètre de configuration **ssl.keystore.config** d'Analytic Server. Par exemple :

```

<ssl id="defaultSSLConfig" keyStoreRef="defaultKeyStore" trustStoreRef="defaultTrustStore"
  clientAuthenticationSupported="true"/>
  <keyStore id="defaultKeyStore" location="/home/as_user/security/mykey.jks" type="JKS"
    password="{xor}0zo5PiozKxYdEgwPDaWeDG1uDz4sLCg7"/>
  <keyStore id="defaultTrustStore" location="/home/as_user/security/mytrust.jks" type="JKS"
    password="{xor}PDC+MTg6Nis="/>

```

Remarque : Utilisez le chemin absolu pour les fichiers du magasin de clés et du magasin de clés de confiance.

5. Définissez la valeur adéquate pour LDAP dans le paramètre de configuration **security.config** d'Analytic Server. Par exemple, dans l'élément **ldapRegistry**, définissez l'attribut **sslEnabled** sur true et l'attribut **sslRef** sur defaultSSLConfig.

Configuration de Kerberos

Analytic Server prend en charge Kerberos avec Ambari.

Remarque : IBM SPSS Analytic Server ne prend pas en charge le mécanisme de connexion unique (SSO) Kerberos lorsqu'il est utilisé en conjonction avec Apache Knox.

1. Créez des comptes dans le référentiel utilisateur de Kerberos pour tous les utilisateurs auxquels vous souhaitez donner accès à Analytic Server.

Remarque : Si l'installation Analytic Server utilise un registre de base, celui-ci doit contenir les comptes utilisateur Kerberos, avec le mot de passe "_". Par exemple :

```

<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. Créez un compte utilisateur de système d'exploitation pour chaque utilisateur créé à l'étape précédente sur chaque noeud Analytic Server et sur chaque noeud Hadoop.
 - Assurez-vous que l'ID de ces utilisateurs correspond sur toutes les machines. Vous pouvez tester cette condition en utilisant la commande kinit pour vous connecter à chaque compte.
 - Vérifiez que l'ID utilisateur se conforme au paramètre Yarn "Minimum user ID for submitting job" (ID utilisateur minimum pour soumission de travail). Il s'agit du paramètre **min.user.id** défini dans

container-executor.cfg. Par exemple, si **min.user.id** est défini à 1000, chaque compte utilisateur créé doit avoir un ID utilisateur supérieur ou égal à 1000.

3. Créez un dossier de base utilisateur sur HDFS pour tous les principaux d'Analytic Server. Par exemple, si vous ajoutez testuser1 au système Analytic Server, vous devez créer un dossier de base /user/testuser1 sur HDFS et autoriser testuser1 à y accéder en lecture et en écriture.
4. [Facultatif] Si vous prévoyez d'utiliser des sources de données HCatalog et si Analytic Server est installé sur une machine différente de celle de Hive Metastore, vous devez simuler les droits d'accès du client Hive sur HDFS.
 - a. Accédez à l'onglet Configs du service HDFS dans la console Ambari.
 - b. Editez le paramètre **hadoop.proxyuser.hive.groups** et entrez la valeur *, ou un groupe contenant tous les utilisateurs autorisés à se connecter à Analytic Server.
 - c. Editez le paramètre **hadoop.proxyuser.hive.hosts** et entrez la valeur *, ou la liste des hôtes sur lesquels Hive Metastore et les instances d'Analytic Server sont installés en tant que services.
 - d. Redémarrez le service HDFS.

Lorsque ces étapes ont été réalisées et qu'Analytic Server est installé, ce dernier configure Kerberos silencieusement et automatiquement.

Configuration de HAProxy pour mécanisme de connexion unique (SSO) à l'aide de Kerberos

1. Configurez et lancez HAProxy en suivant le manuel de la version correspondante dans la documentation HAProxy : <http://www.haproxy.org/#docs>
2. Créez le principal Kerberos (HTTP/<nom_d'hôte_proxy>@<domaine>) et le fichier de clés pour l'hôte HAProxy, où <nom_d'hôte_proxy> correspond au nom complet de l'hôte HAProxy et <domaine> au domaine Kerberos.
3. Copiez le fichier de clés sur chaque hôte Analytic Server en tant que /etc/security/keytabs/spnego_proxy.service.keytab
4. Mettez à jour les autorisations d'accès à ce fichier sur chaque hôte Analytic Server. Par exemple :

```
chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
```
5. Ouvrez la console Ambari et mettez à jour les propriétés suivantes dans la section 'Custom analytics.cfg' d'Analytic Server.

```
web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
web.authentication.kerberos.principal=HTTP/<nom_complet_machine_proxy>@<domaine>
```
6. Enregistrez la configuration et redémarrez tous les services Analytic Server depuis la console Ambari.

Les utilisateurs peuvent à présent se connecter à Analytic Server en utilisant le mécanisme de connexion unique (SSO) Kerberos.

Activation de l'emprunt d'identité Kerberos

L'emprunt d'identité permet de lancer une unité d'exécution dans un contexte de sécurité différent de celui du processus propriétaire de l'unité d'exécution. Par exemple, l'emprunt d'identité permet aux travaux Hadoop de s'exécuter avec des noms d'utilisateur différents du nom d'utilisateur Analytic Server standard (as_user). Pour activer l'emprunt d'identité Kerberos, procédez comme suit :

1. Ajoutez des attributs de configuration de l'emprunt d'identité au système de fichiers HDFS (ou les configurations du service Hive) lors d'une exécution dans un cluster activé pour Kerberos. Dans le cas du système de fichiers HDFS, vous devez ajouter les propriétés suivantes au fichier core-site.xml HDFS :

```
hadoop.proxyuser.<nom_principal_service_analytic_server> .hosts = *
hadoop.proxyuser.<nom_principal_service_analytic_server> .groups = *
```

où <nom_principal_service_analytic_server> est la valeur par défaut as_user indiquée dans la zone Analytic_Server_User de la configuration Analytic Server.

Vous devez également ajouter les propriétés suivantes dans le fichier core-site.xml HDFS via Hive/HCatalog:

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

2. Si Analytic Server est configuré pour utiliser un autre nom d'utilisateur que le nom d'utilisateur as_user, vous devez modifier les noms de propriété pour prendre en compte l'autre nom d'utilisateur (par exemple hadoop.proxyuser.xxxxx.hosts, où xxxxx est le nom d'utilisateur configuré qui est indiqué dans la configuration Analytic Server).

Désactivation de Kerberos

1. Désactivez Kerberos dans la console Ambari.
2. Arrêtez le service Analytic Server.
3. Supprimez les paramètres suivants du fichier analytics.cfg personnalisé.

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

4. Cliquez sur **Save** et redémarrez le service Analytic Server.

Activation des connexions SSL (Secure Socket Layer) à la console Analytic Server

Par défaut, Analytic Server génère des certificats autosignés pour SSL (Secure Socket Layer), ce qui vous permet d'accéder à Analytic Server par le port sécurisé en acceptant ces certificats. Pour protéger davantage l'accès HTTPS, vous devez installer des certificats tiers.

Pour installer des certificats tiers, procédez comme suit.

1. Copiez le magasin de clés et les certificats du magasin de clés de confiance du fournisseur tiers dans le même répertoire sur tous les noeuds Analytic Server. Exemple : /home/as_user/security.

Remarque : L'utilisateur Analytic Server doit avoir accès en lecture à ce répertoire.

2. Dans l'onglet Ambari Services, naviguez jusqu'à l'onglet Configs du service Analytic Server.
3. Editez le paramètre **ssl.keystore.config**.

```
<ssl id="defaultSSLConfig"
  keyStoreRef="defaultKeyStore"
  trustStoreRef="defaultTrustStore"
  clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
  location="<KEYSTORE-LOCATION>"
  type="<TYPE>"
  password="<PASSWORD>"/>
<keyStore id="defaultTrustStore"
  location="<TRUSTSTORE-LOCATION>"
  type="<TYPE>"
  password="<PASSWORD>"/>
```

Remplacez

- <KEYSTORE-LOCATION> par le chemin absolu du magasin de clés, par exemple /home/as_user/security/mykey.jks
- <TRUSTSTORE-LOCATION> par le chemin absolu du magasin de clés de confiance, par exemple /home/as_user/security/mytrust.jks
- <TYPE> par le type du certificat, par exemple JKS, PKCS12, etc.

- <PASSWORD> par le mot de passe chiffré en base 64. Pour l'encodage, vous pouvez utiliser l'utilitaire securityUtility ; par exemple : /opt/ibm/spss/analyticserver/3.0/ae_wlpserver/bin/securityUtility encode <mot_de_passe>

Si vous désirez générer un certificat autosigné; vous pouvez utiliser l'utilitaire securityUtility ; par exemple : /opt/ibm/spss/analyticserver/3.0/ae_wlpserver/bin/securityUtility createSSLCertificate --server=myserver --password=myspassword --validity=365 --subject=CN=mycompany,O=myOrg,C=myCountry.

Pour plus d'informations sur securityUtility et les autres paramètres SSL, reportez-vous à la documentation WebSphere Liberty Profile.

4. Cliquez sur **Save** et redémarrez le service Analytic Server.

Activation de la prise en charge d'Essentials for R

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

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

1. Téléchargez l'archive autoextractible (BIN) correspondant au fichier RPM ou DEB d'IBM SPSS Modeler Essentials for R. Vous pouvez télécharger Essentials for R depuis le site (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>). Sélectionnez le fichier spécifique à votre pile, à sa version et à l'architecture matérielle.
2. Exécutez le fichier binaire autoextractible et suivez les instructions pour (si vous le souhaitez) afficher la licence, l'accepter et sélectionner une installation en ligne ou hors ligne.

Installation en ligne

Sélectionnez l'installation en ligne si l'hôte du serveur Ambari et tous les noeuds du cluster peuvent accéder à <https://ibm-open-platform.ibm.com>.

[*GPFS (Spectrum Scale) uniquement*] Téléchargez le fichier 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) ou <https://ibm-open-platform.ibm.com/repos/IBM-SPSS-ModelerEssentialsR/3.1.0.0/Ubuntu/IBM-SPSS-ModelerEssentialsR-3.1.0.0.list> (Ubuntu) et transférez-le dans le dossier /etc/yum.repos.d (RHEL, CentOS), /etc/zypp/repos.d (SLES) ou /etc/apt/sources.list.d (Ubuntu) sur chaque noeud où vous ajoutez Analytic Server Metastore en tant que service.

Installation hors ligne

Sélectionnez l'installation hors ligne si votre serveur Ambari n'a pas accès à Internet. L'installation hors ligne télécharge les fichiers RPM nécessaires et doit être exécutée sur un système qui peut accéder à <https://ibm-open-platform.ibm.com>. Les fichiers RPM peuvent ensuite être copiés sur l'hôte du serveur Ambari.

- a. Copiez les fichiers RPM ou DEB Essentials for R dans un emplacement de l'hôte du serveur Ambari. Les fichiers RPM/DEB nécessaires varient en fonction de votre distribution, de votre version et de votre architecture, comme indiqué ci-après.

BigInsights 4.1 et 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 et 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. Installez le fichier RPM ou DEB. Dans l'exemple ci-dessous, la commande installe Essentials for R sur BigInsights 4.2.

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

Dans l'exemple ci-dessous, la commande installe Essentials for R sur HDP 2.4 (Ubuntu).

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

3. Redémarrez votre serveur Ambari.

```
ambari-server restart
```

4. Connectez-vous à votre serveur Ambari et installez SPSS Essentials for R en tant que service via l'interface utilisateur Ambari. SPSS Essentials for R doit être installé sur chaque hôte sur lequel Analytic Server et Analytic Metastore sont installés.

Remarque : Ambari va tenter d'installer gcc-c++, gcc-gfortran (RHEL) et gcc-fortran (SUSE) avant d'installer R. Ces packages sont déclarés en tant que dépendances dans la définition de service Ambari de R. Assurez-vous que les serveurs sur lesquels R doit être installé et exécuté sont configurés pour le téléchargement des packages RPM gcc-c++ et gcc-[g]fortran ou possèdent des compilateurs GCC et FORTRAN. Si l'installation d'Essentials for R échoue, installez ces packages manuellement avant d'installer Essentials for R.

5. Actualisez le service Analytic Server.
6. Exécutez le script `update_clientdeps` en suivant les instructions figurant dans «*Mise à jour des dépendances de client*», à la page 21.
7. Vous devez également installer Essentials for R sur la machine qui héberge SPSS Modeler Server. Reportez-vous à la documentation SPSS Modeler pour plus de détails.

Activation des sources de base de données relationnelle

Analytic Server peut utiliser des sources de base de données relationnelle si vous rendez disponibles les pilotes JDBC dans un répertoire partagé sur chaque hôte Analytic Server. Par défaut, ce répertoire est `/usr/share/jdbc`.

Pour utiliser un autre répertoire partagé, procédez comme suit.

1. Dans l'onglet Ambari Services, naviguez jusqu'à l'onglet Configs du service Analytic Server.
2. Ouvrez la section **Advanced analytics.cfg**.
3. Entrez le chemin du répertoire partagé des pilotes JDBC dans **jdbc.drivers.location**.
4. Cliquez sur **Save**.
5. Arrêtez le service Analytic Server.
6. Cliquez sur **Refresh**.
7. Démarrez le service Analytic Server.

Tableau 4. Bases de données prises en charge

Base de données	Versions prises en charge	Fichiers jar du pilote JDBC	Fournisseur
Amazon Redshift	8.0.2 ou version ultérieure	RedshiftJDBC41-1.1.6.1006.jar ou version ultérieure	Amazon
BigSQL	4.1.0.0 ou suivante	db2jcc.jar	IBM
DashDB	Service Bluemix	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

Tableau 4. Bases de données prises en charge (suite)

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

Remarques

- Si vous avez créé une source de données Redshift avant d'installer Analytic Server, vous devez effectuer les opérations ci-dessous pour pouvoir utiliser cette source de données.
 1. Dans la console Analytic Server, ouvrez la source de données Redshift.
 2. Sélectionnez la source de données de base de données Redshift.
 3. Entrez l'adresse du serveur Redshift.
 4. Entrez le nom de la base de données et le nom d'utilisateur. Le mot de passe devrait être renseigné automatiquement.
 5. Sélectionnez la table de base de données.
- BigSQL est l'interface SQL IBM de l'environnement Apache Hadoop. BigSQL n'est pas une base de données relationnelle mais Analytic Server permet d'y accéder via JDBC. (Le fichier jar JDBC est identique à celui utilisé pour DB2.)
BigSQL est un service supplémentaire de BigInsights. A ce titre, sa version est identique à la version BigInsights. Une utilisation courante de BigSQL avec Analytic Server est l'accès aux tables BigSQL Hadoop/HBase via une source de données HCatalog.

Activation des sources de données HCatalog

Analytic Server prend en charge différentes sources de données par l'intermédiaire de Hive et de HCatalog. Certaines nécessitent des opérations de configuration manuelles.

1. Collectez les fichiers JAR nécessaires pour activer la source de données. Voir les sections ci-dessous pour plus de détails.
2. Ajoutez ces fichiers JAR au répertoire {DOSSIER_PRINCIPAL_HIVE}/auxlib et au répertoire /usr/share/hive sur chaque noeud Analytic Server.
3. Redémarrez le service Hive Metastore.
4. Actualisez le service Analytic Metastore.
5. Redémarrez toutes les instances du service Analytic Server.

Bases de données NoSQL

Analytic Server prend en charge toutes les bases de données NoSQL pour lesquelles un gestionnaire d'espace de stockage Hive est disponible chez le fournisseur.

La prise en charge d'Apache HBase et d'Apache Accumulo ne demande aucune opération particulière.

Pour les autres bases de données NoSQL, procurez-vous le gestionnaire d'espace de stockage et les fichiers JAR associés auprès du fournisseur de la base.

Tables Hive sous forme de fichiers

Analytic Server prend en charge toutes les tables Hive sous forme de fichiers pour lesquelles un sérialiseur-désérialiseur (SerDe) Hive intégré ou personnalisé est disponible.

Le sérialiseur-désérialiseur XML Hive pour le traitement des fichiers XML est stocké dans le référentiel Maven Central à l'adresse <http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde>.

Apache Spark

Si vous désirez utiliser Spark (version 1.5 ou ultérieure) avec une source de données d'entrée HCatalog, vous devez ajouter manuellement la propriété `spark.version` au fichier `analytics.cfg` personnalisé.

1. Ouvrez la console Amabri et ajoutez la propriété suivante dans la section Analytic Server **Advanced analytics.cfg**.
 - **Key** : `spark.version`
 - **Value** : Entrez le numéro de version Spark approprié (par exemple 1.x, 2.x ou None).
2. Enregistrez la configuration et redémarrez tous les services Analytic Server depuis la console Ambari.

Remarque : Vous pouvez contraindre HCatalog de ne jamais utiliser Spark en utilisant un paramètre de la section Custom `analytics.cfg`.

1. Ouvrez la console Amabri et ajoutez la propriété suivante dans la section Analytic Server **Custom analytic.cfg**.
 - **Key** : `spark.hive.compatible`
 - **Value** : `false`

Modification des ports utilisés par Analytic Server

Analytic Server utilise par défaut le port 9080 pour HTTP et 9443 pour HTTPS. Pour modifier les paramètres de port, procédez comme suit.

1. Dans l'onglet Ambari Services, naviguez jusqu'à l'onglet Configs du service Analytic Server.
2. Ouvrez la section **Advanced analytics.cfg**.
3. Définissez les ports HTTP et HTTPS souhaités dans **http.port** et **https.port** respectivement.
4. Cliquez sur **Save**.
5. Redémarrez le service Analytic Server.

Haute disponibilité d'Analytic Server

Vous pouvez garantir la haute disponibilité d'Analytic Server en le définissant en tant que service à plusieurs noeuds de votre cluster.

1. Dans la console Ambari, naviguez jusqu'à l'onglet Hosts.
2. Sélectionnez un hôte sur lequel Analytic Server ne s'exécute pas encore en tant que service.
3. Sur l'onglet Summary, cliquez sur **Add**, et sélectionnez Analytic Server.
4. Cliquez sur **Confirm Add**.

Optimisation des options JVM pour les petits volumes de données (Small Data)

Vous pouvez éditer les propriétés JVM pour optimiser votre système en cas d'exécution de petits travaux (M3R).

Dans la console Ambari, affichez la section `Advanced analytics-jvm-options` de l'onglet Configs dans le service Analytic Server. La modification des paramètres ci-après définit la taille de segment de mémoire des travaux s'exécutant sur le serveur hébergeant Analytic Server (pas le serveur Hadoop). Cette option est importante pour l'exécution de petits travaux (M3R) et vous devrez éventuellement tester différentes valeurs afin d'optimiser votre système.

```
-Xms512M  
-Xmx2048M
```

Mise à jour des dépendances de client

Cette section explique comment mettre à jour les dépendances du service Analytic Server avec le script `update_clientdeps`.

1. Connectez-vous à l'hôte du serveur Ambari en tant que root.
2. Placez-vous dans le répertoire `/var/lib/ambari-server/resources/stacks/<nom_pile>/<version_pile>/services/ANALYTICSERVER/package/scripts`. Exemple :

```
cd
"/var/lib/ambari-server/resources/stacks/HDP/2.4/services/ANALYTICSERVER/package/scripts"
```

3. Exécutez le script `update_clientdeps` avec les arguments suivants :

```
-u <utilisateur_ambari>
    Nom d'utilisateur du compte Ambari

-p <mot_de_passe_ambari>
    Mot de passe de l'utilisateur du compte Ambari.

-h <hôte_ambari>
    Nom d'hôte du serveur Ambari.

-x <port_ambari>
    Port sur lequel Ambari est à l'écoute.
```

Voir l'exemple ci-après.

```
./update_clientdeps.sh -u admin -p admin -h host.domain -x 8080
```

4. Redémarrez le serveur Ambari à l'aide de la commande suivante :
- ```
ambari-server restart
```

## Configuration d'Apache Knox

Apache Knox Gateway est un système qui fournit un point d'accès sécurisé unique aux services Apache Hadoop. Le système simplifie la sécurité Hadoop des utilisateurs (ayant accès aux données de cluster et exécutant les travaux) et des opérateurs (dont le rôle est de contrôler l'accès et de gérer le cluster). Gateway s'exécute comme un serveur (ou cluster de serveurs) d'un ou de plusieurs clusters Hadoop.

**Remarque :** IBM SPSS Analytic Server ne prend pas en charge Apache Knox lorsqu'il est utilisé en conjonction avec le mécanisme de connexion unique (SSO) Kerberos.

Apache Knox Gateway masque efficacement les détails de topologie de cluster Hadoop et s'intègre à Enterprise LDAP et Kerberos. Les sections suivantes fournissent des informations sur les tâches de configuration requises pour Apache Knox et pour Analytic Server.

### Prérequis

- Les noeuds Analytic Server doivent se connecter au serveur Knox avec une connexion SSH sans mot de passe. La connexion SSH sans mot de passe circule dans le sens Analytic Server vers Knox (**Analytic Server > Knox**).
- Analytic Server doit être installé après que le service Knox a été installé.

Dans certains cas, des problèmes inattendus entraînent que les fichiers de configuration ne sont pas copiés automatiquement. Vous devez alors copier manuellement les fichiers de configuration suivants :

- `com.ibm.spss.knox_0.6-3.1.0.0.jar` : Le fichier doit être copié à partir de l'emplacement d'Analytic Server :  
`<Chemin_installation_Analytic_Server>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib`  
vers le noeud de serveur Knox :  
`/Chemin_service_Knox`

Par exemple : /usr/iop/4.1.0.0/knox/ext

- `rewrite.xml` et `service.xml` : les fichiers doivent être copiés depuis l'emplacement Analytic Server suivant :

<Chemin\_installation\_Analytic\_Server>/ae\_wlpserver/usr/servers/aeserver/configuration/knox  
vers le noeud de serveur Knox :

/Chemin\_service\_Knox/data/services

Par exemple : /usr/iop/4.1.0.0/knox/data/services

## Configuration d'Ambari

Le service Analytic Server doit être configuré dans l'interface utilisateur d'Ambari :

1. Dans l'interface utilisateur d'Ambari, allez à **Knox > Configs > Advanced topology**. Les paramètres de configuration Knox s'affichent dans la fenêtre **content**.
2. Ajoutez l'élément <service> suivant à la configuration Knox :

```
<service>
 <role>ANALYTICSERVER</role>
 <url>http://{analyticserver-host}:{analyticserver-port}/analyticserver</url>
</service>
```

{analyticserver-host} et {analyticserver-port} doivent être remplacés par le nom de serveur Analytic Server et le numéro de port appropriés :

- L'URL de l'élément {analyticserver-host} se trouve dans l'interface utilisateur d'Ambari (**SPSS Analytic Server > Summary > Analytic Server**).
- Le numéro de l'élément {analyticserver-port} se trouve dans l'interface utilisateur d'Ambari (**SPSS Analytic Server > Configs > Advanced analytics.cfg > http.port**).

**Remarque :** Lorsque Analytic Server est déployé sur plusieurs noeuds, et que l'équilibreur de charge est utilisé, les valeurs de {analyticserver-host} et de {analyticserver-port} doivent correspondre à l'URL et au numéro de port de l'équilibreur de charge.

3. Redémarrez le service Knox.

Lorsque LDAP est utilisé, Knox utilise par défaut la version de démonstration LDAP fournie. Vous pouvez la remplacer par un serveur LDAP d'entreprise (tel que Microsoft LDAP ou OpenLDAP).

## Configuration du produit Analytic Server

Pour utiliser LDAP pour Analytic Server, Analytic Server doit être configuré de façon à utiliser le même serveur LDAP qu'Apache Knox. Les entrées <value> pour les paramètres Ambari suivants doivent être mises à jour afin de refléter les paramètres de serveur Knox LDAP appropriés :

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

Les valeurs sont disponibles dans l'interface utilisateur d'Ambari via le menu : **Knox > Configs > Advanced topology**. Exemple :

```
<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_host_name}:33389</value>
</param>
```

Redémarrez le service Knox après avoir mis à jour les paramètres LDAP de Knox.

**Important :** Le mot de passe de l'administrateur d'Analytic Server doit être identique à celui de l'administrateur de Knox.

## Configuration d'Apache Knox

1. Sur le serveur Knox, créez le sous-répertoire `<serveur_knox>/data/service/analyticserver/3.1`, puis téléchargez les fichiers `service.xml` et `rewrite.xml` dans le nouveau répertoire. Les deux fichiers se trouvent sur le système Analytic Server, dans `<analytic_server>/configuration/knox/analyticserver/3.1` (par exemple, `/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/3.1/*.xml`)
2. Depuis le répertoire `<serveur_knox>/bin`, exécutez le script `./knoxcli.sh redeploy --cluster default`
3. Importez le fichier `com.ibm.spss.knoxservice_0.6-*.jar` dans `<knox_server>/ext`. Le fichier se trouve sur le système Analytic Server, dans `<analytic_server>/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.1.0.0.jar` (par exemple, `/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. Depuis l'interface utilisateur Ambari, ajoutez l'élément suivant dans **Knox > Configs > Advanced topology**:

```
<service>
 <role>ANALYTICSERVER</role>
 <url>http://{AS-Host}:{AS-port}/analyticserver</url>
</service>
```
5. Depuis l'interface utilisateur Ambari, ajoutez ou mettez à jour les utilisateurs dans **Knox > Configs > Advanced users-ldif** (par exemple, `admin`, `qauser1`, `qauser2`).
6. Redémarrez LDAP depuis **Knox > Service Actions > Start Demo LDAP**.
7. Redémarrez le service Knox.

## Installation d'Apache Knox sur Hortonworks Data Platform (HDP)

Les étapes suivantes décrivent le processus d'installation d'Apache HDP dans une grappe HDP.

1. Vérifiez si un utilisateur Knox existe sur la grappe HDP. Si un utilisateur Knox n'existe pas, vous devez en créer un.
2. Téléchargez et décompressez Apache Knox dans un dossier sous `/home/knox`.
3. Dans HDP, basculez sur l'utilisateur Knox et accédez au dossier `knox`. L'utilisateur Knox doit disposer de l'autorisation `permission(RWX)` sur tous les sous-dossiers `knox`.
4. Configurez Apache Knox pour Analytic Server. Pour plus d'informations, reportez-vous à la section **Configuration d'Apache Knox**.
  - a. Créez une hiérarchie de dossiers `analyticserver/3.0` sous `{knox}/data/services`.
  - b. Copiez les fichiers `rewrite.xml` and `service.xml` depuis l'emplacement Analytic Server :  
`/opt/ibm/spss/analyticserver/3.0/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/3.1`  
vers le noeud de serveur Knox :  
`{knox}/data/services/analyticserver/3.1`
  - c. Copiez le fichier Knox `*.jar` depuis l'hôte Analytic Server :  
`/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`  
vers le répertoire Knox `ext` :  
`{knox}/ext`
  - d. Mettez à jour le fichier `default.xml` dans `{knox}/conf/topologies` pour se conformer à l'exemple suivant :

**Remarque :** Vous devez créer le fichier s'il n'existe pas.

```

<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.shiro.realm.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>AcIsAuthz</name>
 <enabled>true</enabled>
 </provider>
 </gateway>

 <!--autre service-->
 <service>
 <role>ANALYTICSERVER</role>
 <!--remplacez {AS-host}nas {AS-port} par la valeur réelle-->
 <url>http://{AS-host}:{AS-port}/analyticserver</url>
 </service>
</topology>

```

5. Exécutez `{knox}/bin/knoxcli.sh`.

6. Exécutez `{knox}/bin/ldap.sh start`.

**Remarque :** Le script utilise le port 33389. Vérifiez que le port n'est pas actuellement utilisé.

7. Exécutez `{knox}/bin/gateway.sh start`.

**Remarque :** Le script utilise le port 8443. Vérifiez que le port n'est pas actuellement utilisé.

8. Vérifiez l'installation.

a. Exécutez la commande `curl` par rapport au serveur Analytic Server dans l'URL Knox :

```

curl -ikvu
{nom_utilisateur}:{mot_de_passe} https://{hôte_knox}:8443/gateway/default/analyticserver/admin

```

## Traitement des incidents

**Problème :** Analytic Server ne fonctionne pas dans Knox après l'installation.

**Solution :** arrêtez Knox, supprimez tous les fichiers sous `{knox}/data/deployments/*`, puis redémarrez Knox.

**Problème :** impossible de se connecter à Analytic Server via Knox.

**Solution :** vérifiez les utilisateurs dans `{knox}/conf/users.ldif`. Mettez à jour les utilisateurs existants ou ajoutez de nouveaux utilisateurs Analytic Server. Les utilisateurs Knox et leurs données d'identification

doivent correspondre à ceux des utilisateurs Analytic Server.

## Structure de l'URL d'Analytic Server avec Apache Knox activé

L'URL d'interface utilisateur Analytic Server activée pour Knox est `https://{hôte_knox}:{port_knox}/gateway/default/analyticserver/admin`

- https protocol - les utilisateurs doivent accepter un certificat pour poursuivre via ce navigateur Web.
- knox-host correspond à l'hôte Knox.
- knox-port correspond au numéro de port Knox.
- L'URI est `gateway/default/analyticserver`.

---

## Mise à niveau et migration

Analytic Server vous permet de mettre à niveau ou de migrer les données et les paramètres de configuration depuis une installation Analytic Server existante vers une nouvelle installation.

### Mise à niveau de la version 3.0.1 vers la version 3.1.0 - BigInsights et Hortonworks

Si vous disposez d'une installation d'Analytic Server 3.0.1 existante, vous pouvez la mettre à niveau vers la version 3.1.0.

1. Dans la console Ambari, arrêtez le service Analytic Server.
2. Selon le type de votre installation, procédez comme suit :

#### Mise à niveau en ligne

- a. Vérifiez que l'hôte du serveur Ambari et tous les noeuds du cluster peuvent accéder à <https://ibm-open-platform.ibm.com>.
- b. Téléchargez le fichier `IBM-SPSS-AnalyticServer-3.1.0.0.repo` depuis <https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms/IBM-SPSS-AnalyticServer-3.1.0.0.repo> (x86 et ppc64le) sur chaque hôte Analytic Server et placez-le dans le dossier `/etc/yum.repos.d` (RHEL ou CentOS) ou `/etc/zypp/repos.d` (SLES).

#### Mise à niveau hors ligne

- a. La mise à niveau hors ligne télécharge les fichiers RPM requis et doit être réalisée sur une machine pouvant accéder à <https://ibm-open-platform.ibm.com>.
- b. Créez un répertoire qui servira de référentiel pour les fichiers RPM d'Analytic Server. Reportez-vous à l'exemple suivant :  

```
mkdir /home/root/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms
```
- c. Copiez les fichiers RPM nécessaires d'Analytic Server dans ce répertoire. Les fichiers RPM dont vous avez besoin varient en fonction de votre distribution, de votre version et de votre architecture. Pour BigInsights 4,2, les fichiers requis sont décrits ci-après.

Tableau 5. Fichiers RPM BigInsights 4.2

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

- d. Créez la définition du référentiel local. Par exemple, créez un fichier nommé `analyticserver.repo` dans `/etc/yum.repos.d/` (pour RHEL, CentOS) ou `/etc/zypp/repos.d/` (pour SLES) avec le contenu suivant.

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer-3.1.0.0
baseurl=file:///chemin du référentiel local
enabled=1
gpgcheck=0
protect=1
```

e. Créez le référentiel yum local. Reportez-vous à l'exemple suivant :

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

3. Effacez les métadonnées Ambari de votre cache local. Par exemple, pour vider le cache sous RHEL ou CentOS, exécutez la commande suivante :

```
sudo yum clean all
```

**Remarque :** La commande yum ne fonctionne pas lorsque deux référentiels Analytic Server sont listés. Par conséquent, les fichiers \*.repo originaux associés à Analytic Server doivent être renommés ou supprimés. Sous SLES, la commande doit être remplacée par celle-ci :

```
sudo zypper refresh
```

4. Sur chaque hôte Analytic Server, mettez à niveau les packages RPM. Par exemple, pour effectuer une mise à niveau sous RHEL ou CentOS, exécutez la commande suivante :

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

Sous SLES, la commande doit être remplacée par celle-ci :

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

5. Actualisez la pile.

### BigInsights

a. Dans la console Ambari, démarrez et arrêtez le service Analytic Server.

b. Exécutez l'action **Actualiser** personnalisée

### Hortonworks

Accédez à l'un de vos noeuds Analytic Server et exécutez la commande suivante :

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

6. Installation hors ligne uniquement. Mettez à jour le fichier référentiel Ambari repoinfo.xml, qui se trouve généralement dans /var/lib/ambari-server/resources/stacks/\$stackName/\$stackVersion/repos/, pour utiliser le référentiel yum local en ajoutant les lignes suivantes :

```
<os type="host_os">
 <repo>
 <baseurl>file:///{{chemin du référentiel local}}/</baseurl>
 <repoid>IBM-SPSS-AnalyticServer</repoid>
 <reponame>IBM-SPSS-AnalyticServer-3.1.0.0</reponame>
 </repo>
</os>
```

7. Effacez l'état de Zookeeper. Exécutez la commande suivante dans le répertoire bin de Zookeeper (par exemple, /usr/iop/current/zookeeper-server/bin):

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

8. Dans la console Ambari, démarrez le service Analytic Server.

### Migration vers une nouvelle version d'Analytic Server

Si vous disposez d'une installation existante d'Analytic Server 2.0 ou 2.1 et que vous avez acheté la version 3.1.0, vous pouvez ensuite migrer vos paramètres de configuration de la version 2.0/2.1 dans l'installation de la version 3.1.0.

### Restrictions :

- Si une version antérieure à la version 2.0 est installée, vous devez d'abord effectuer la migration de la version antérieure vers la version 2.0/2.1, puis de la version 2.0/2.1 vers la version 3.1.0.
- Les installations 2.0/2.1 et 3.1.0 ne peuvent pas coexister sur le même cluster Hadoop. Si vous configurez votre installation 3.1.0 pour utiliser le même cluster Hadoop que votre installation 2.0/2.1, l'installation 2.0/2.1 ne fonctionnera plus.



## Etapes de migration de la version 2.0/2.1 vers la version 3.1.0

1. Installez la nouvelle installation d'Analytic Server en suivant les instructions figurant dans «Installation dans Ambari», à la page 5.
2. Copiez la racine d'analyse de l'ancienne installation dans la nouvelle.
  - a. Si vous n'êtes pas certain de l'emplacement de la racine d'analyse, exécutez la commande `hadoop -fs ls`. Le chemin de la racine d'analyse aura le format `/user/aeuser/analytic-root`, où `aeuser` est l'ID utilisateur auquel appartient la racine d'analyse.
  - b. Remplacez le propriétaire de la racine d'analyse (`aeuser`) par `as_user`.

```
hadoop dfs -chown -R {as_user:{groupe}} {chemin racine d'analyse 2.0/2.1}
```

**Remarque :** Si vous prévoyez d'utiliser l'installation Analytic Server existante après la migration, déposez une copie du répertoire `analytic-root` dans HDFS, puis changez le propriétaire de la copie.

  - c. Connectez-vous à l'hôte de la nouvelle installation Analytic Server en tant que `as_user`. Supprimez le répertoire `/user/as_user/analytic-root` s'il existe.
  - d. Exécutez le script de copie suivant :

```
hadoop distcp hftp://{hôte namenode 2.0/2.1}:50070/{chemin racine d'analyse 2.0/2.1} hdfs://{hôte namenode 3.1.0}/user/as_user/analytic-root
```
3. Dans la console Ambari, arrêtez le service Analytic Server.
4. Vérifiez que le service Analytic Metastore est actif.
5. Collectez les paramètres de configuration de l'ancienne installation.
  - a. Copiez l'archive `configcollector.zip` de la nouvelle installation dans le répertoire `{AS_ROOT}\tools` de l'ancienne installation.
  - b. Décompressez la copie de `configcollector.zip`. Cette opération crée un sous-répertoire `configcollector` dans l'ancienne installation.
  - c. Exécutez l'outil de collecte de la configuration dans votre ancienne installation en lançant le script **configcollector** situé sous `{RACINE_AS}\tools\configcollector`. Copiez le fichier compressé (ZIP) généré sur le serveur qui héberge la nouvelle installation.
6. Effacez l'état de Zookeeper. Depuis le répertoire `bin` de Zookeeper (par exemple, `/usr/hdp/current/zookeeper-client` sur Hortonworks ou `/usr/iop/current/zookeeper-server` sur BigInsights), exécutez la commande suivante.

```
./zkCli.sh rmr /AnalyticServer
```
7. Lancez l'outil de migration en exécutant le script **migrationtool** et en transmettant sous forme d'argument le chemin du fichier compressé créé par le collecteur de configuration. Exemple :

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.1/ASConfiguration_2.1.0.0.xxx.zip
```
8. Dans la console Ambari, démarrez le service Analytic Server.

**Remarque :** Si vous avez configuré R pour son utilisation avec l'installation Analytic Server existante, vous devez suivre les étapes permettant de le configurer avec la nouvelle installation Analytic Server.

---

## Désinstallation

**Important :** Lorsque Essentials for R est installé, vous devez d'abord exécuter le script `remove_R.sh`. Si vous ne désinstallez pas Essentials for R, avant de désinstaller Analytic Server, il sera impossible de le faire ultérieurement. Le script `remove_R.sh` est supprimé lorsque Analytic Server est désinstallé. Pour obtenir des informations sur la désinstallation d'Essentials for R, voir «Désinstallation d'Essentials for R», à la page 28.

1. Sur l'hôte Analytic Metastore, lancez le script `remove_as.sh` dans le répertoire `{RACINE_AS}/bin` avec les paramètres suivants :
  - u** Requis. ID utilisateur de l'administrateur du serveur Ambari.
  - p** Requis. Mot de passe de l'administrateur du serveur Ambari.

- h** Requis. Nom d'hôte du serveur Ambari.
- x** Requis. Port du serveur Ambari.
- l** Facultatif. Active le mode sécurisé.

Exemples :

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

Supprime Analytic Server d'un cluster sur l'hôte Ambari one.cluster.

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

Supprime Analytic Server d'un cluster sur l'hôte Ambari one.cluster, en mode sécurisé.

**Remarque :** Cette opération supprime le dossier Analytic Server sur le système HDFS.

**Remarque :** Cette opération ne supprime pas les schémas DB2 associés à Analytic Server. Voir la documentation de DB2 pour des informations sur la suppression manuelle des schémas.

## Désinstallation d'Essentials for R

1. Sur l'hôte Essentials for R, exécutez le script `remove_R.sh` dans le répertoire `{AS_ROOT}/bin` avec les paramètres suivants :

- u** Requis. ID utilisateur de l'administrateur du serveur Ambari.
- p** Requis. Mot de passe de l'administrateur du serveur Ambari.
- h** Requis. Nom d'hôte du serveur Ambari.
- x** Requis. Port du serveur Ambari.
- l** Facultatif. Active le mode sécurisé.

Exemples :

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

Supprime Essentials for R d'un cluster avec l'hôte Ambari one.cluster.

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

Supprime Essentials for R d'un cluster avec l'hôte Ambari one.cluster en mode sécurisé.

2. Supprimez le répertoire des services R du répertoire des services du serveur Ambari. Par exemple, dans BigInsights 4.2, le répertoire ESSENTIALR est situé sous `/var/lib/ambari-server/resources/stacks/BigInsights/4.2/services`.
3. Dans la console Ambari, vérifiez que le service Essentials for R n'existe plus.

---

## Chapitre 3. Installation et configuration de Cloudera

---

### Présentation de Cloudera

Cloudera est une distribution Apache Hadoop open source. Cloudera Distribution Including Apache Hadoop (CDH) cible les déploiements de cette technologie pour les entreprises.

Analytic Server peut s'exécuter sur la plateforme CDH. CDH contient les éléments de base principaux pour Hadoop, qui permettent une informatique répartie fiable et évolutive pour les ensembles de données volumineux (principalement MapReduce et HDFS), ainsi que d'autres composants orientés entreprise qui assurent la sécurité, la haute disponibilité et l'intégration au matériel et à d'autres logiciels.

---

### Conditions requises propres à Cloudera

En plus des conditions requises générales, prenez connaissance des informations ci-après.

#### Services

Assurez-vous que les instances ci-dessous sont installées sur chaque hôte Analytic Server.

- HDFS : Gateway, DataNode ou NameNode
- Hive : Gateway, Hive Metastore Server ou HiveServer2
- Yarn : Gateway, ResourceManager ou NodeManager

Les instances suivantes ne sont requises que lorsque leurs fonctions sont utilisées.

- Accumulo : Gateway
- HBase : Gateway, Master ou RegionServer

#### Référentiel de métadonnées

Si vous prévoyez d'utiliser MySQL comme référentiel de métadonnées d'Analytic Server, suivez les instructions de «Configuration de MySQL pour Analytic Server».

### Configuration de MySQL pour Analytic Server

La configuration d'IBM SPSS Analytic Server dans Cloudera Manager requiert l'installation et la configuration d'une base de données de serveur MySQL.

1. Exécutez la commande suivante depuis une fenêtre de commande sur le noeud sur lequel la base de données MySQL est stockée :

```
yum install mysql-server
```

**Remarque :** Utilisez `zypper install mysql` sous SuSE Linux.

2. Exécutez la commande suivante depuis une fenêtre de commande sur chaque noeud de cluster Cloudera :

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

**Remarque :** Utilisez `sudo zypper install mysql-connector-java` pour SuSE Linux.

3. Choisissez le nom de la base de données Analytic Server, le nom d'utilisateur de la base de données et le mot de passe de la base de données qu'Analytic Server doit utiliser pour accéder à la base de données MySQL, et prenez-en note.
4. Installez Analytic Server selon les instructions figurant dans «Installation dans Cloudera», à la page 30.
5. Copiez le script `/opt/cloudera/parcels/AnalyticServer/bin/add_mysql_user.sh` depuis l'un des serveurs gérés par Cloudera sur le noeud sur lequel la base de données MySQL est installée. Exécutez le script avec les paramètres appropriés à votre configuration. Exemple :

```
./add_mysql_user.sh -u <nom_utilisateur_base_de_données> -p <mot_de_passe_base_de_données> -d <nom_base_de_données>
```

**Remarques :** Le paramètre a -r <dbRootPassword> est requis lorsque la base de données s'exécute en mode sécurisé (le mot de passe de l'utilisateur root est défini).

Les paramètres -r <mot\_de\_passe\_utilisateur\_base\_de\_données> et -t <nom\_utilisateur\_base\_de\_données> sont requis lorsque la base de données s'exécute en mode sécurisé avec un nom d'utilisateur autre que root.

---

## Installation dans Cloudera

Les étapes ci-après décrivent le processus d'installation manuelle d'IBM SPSS Analytic Server dans Cloudera Manager.

### Analytic Server 3.1.0

#### Installation en ligne

1. Accédez au site Web d'IBM Passport Advantage et téléchargez le fichier binaire autoextractible approprié pour votre pile, version de pile et architecture matérielle sur un hôte dans le cluster Cloudera. Les fichiers binaires disponibles sont :

Tableau 6. Fichiers binaires autoextractibles Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.1 pour Cloudera 5.8, 5.9 et 5.10 Ubuntu (Anglais)	spss_as-3.1-cdh5.8-5.10-ubun_en.bin
IBM SPSS Analytic Server 3.1 pour Cloudera 5.8, 5.9 et 5.10 Linux x86-64 (Anglais)	spss_as-3.1-cdh5.8-5.10-1x86-en.bin

2. Exécutez le programme d'installation autoextractible \*.bin Cloudera sur le noeud de cluster maître Cloudera. Suivez les invites d'installation en acceptant le contrat de licence et en conservant le répertoire d'installation CSD par défaut.

**Remarque :** Vous devez spécifier un répertoire CSD différent si l'emplacement par défaut a été modifié.

3. Redémarrez Cloudera Manager une fois l'installation terminée.
4. Ouvrez l'interface de Cloudera Manager (par exemple [http://\\${CM\\_HOST}:7180/cm/login](http://${CM_HOST}:7180/cm/login) avec les données d'identification et de connexion par défaut admin/admin), actualisez **Remote Parcel Repository URLs** et vérifiez que l'adresse URL est correcte. Exemple :

<https://ibm-open-platform.ibm.com>

**Remarque :** Les informations dans **Parcel Update Frequency** et **Remote Parcel Repository URLs** peuvent être mises à jour pour répondre à vos besoins.

5. Une fois que Cloudera Manager a actualisé les fichiers parcel (vous pouvez les actualiser manuellement en cliquant sur **Check for New Parcels**), le statut du fichier parcel **AnalyticServer** indique **Available Remotely**.
6. Sélectionnez **Download > Distribute > Activate**. Le statut du fichier parcel **AnalyticServer** est mis à jour et devient **Distributed, Activated**.
7. Configurez MySQL pour Analytic Server.
8. Dans Cloudera Manager, ajoutez Analytic Server en tant que service et choisissez son emplacement. Vous devez fournir les informations suivantes dans l'assistant Add Service Wizard :
  - Le nom d'hôte du métamagasin Analytic Server
  - Nom de base de données du métamagasin Analytic Server

- Le nom de l'utilisateur du métamagasin d'Analytic Server
- Le mot de passe du métamagasin Analytic Server

L'assistant Add Service Wizard affiche la progression générale au cours de chaque phase du processus de création de service ainsi qu'un message de confirmation final une fois que le service a été installé et configuré correctement dans le cluster.

**Remarque :** Une fois Analytic Server installé, ne cliquez pas sur **Create Analytic Server Metastore** dans la liste Actions de la page des services Analytic Server dans Cloudera Manager. La création d'un métamagasin écrase le référentiel de métadonnées existant.

## Installation hors ligne

Les étapes d'installation hors ligne sont identiques aux étapes d'installation en ligne, si ce n'est que vous devez télécharger manuellement les fichiers parcel et les métadonnées correspondant à votre système d'exploitation.

RedHat Linux requiert les fichiers suivants :

- AnalyticServer-3.1.0.0-el6.parcel
- AnalyticServer-3.1.0.0-el6.parcel.sha
- manifest.json
- ou
- AnalyticServer-3.1.0.0-el7.parcel
- AnalyticServer-3.1.0.0-el7.parcel.sha

SuSE Linux requiert les fichiers suivants :

- AnalyticServer-3.1.0.0-sles11.parcel
- AnalyticServer-3.1.0.0-sles11.parcel.sha
- manifest.json

Ubuntu Linux requiert les fichiers suivants :

- AnalyticServer-3.1.0.0-trusty.parcel
- AnalyticServer-3.1.0.0-trusty.parcel.sha

1. Téléchargez et exécutez le programme d'installation \*.bin autoextractible de Cloudera sur le noeud de cluster maître Cloudera Manager. Suivez les invites d'installation en acceptant le contrat de licence et en conservant le répertoire d'installation CSD par défaut.

**Remarque :** Vous devez spécifier un répertoire CSD différent s'il ne réside pas sous l'emplacement par défaut.

2. Copiez les fichiers parcel et les fichiers de métadonnées requis vers votre chemin Cloudera local repo sur le noeud de cluster maître Cloudera Manager. Le chemin par défaut est /opt/cloudera/parcel-repo (il peut être configuré dans l'interface utilisateur de Cloudera Manager).

Le fichier parcel **AnalyticServer** est signalé comme téléchargé (**downloaded**) après son actualisation par Cloudera Manager. Vous pouvez cliquer sur **Check for New Parcels** pour forcer l'actualisation.

3. Cliquez sur **Distribute > Activate**.

Le fichier parcel **AnalyticServer** est signalé comme ayant été distribué et activé.

## Mise à niveau d'Analytic Server 3.1.0 sur Cloudera

Si vous disposez d'une installation d'Analytic Server 3.0/3.0.1 existante, vous pouvez la mettre à niveau vers la version 3.1.0.

1. Depuis Cloudera Manager, arrêtez, puis supprimez le service Analytic Server.

2. Depuis Cloudera Manager, désactivez la version antérieure d'Analytic Server.
3. Reportez-vous aux sections "en ligne" ou "hors ligne" de la rubrique «Mise à niveau et migration», à la page 25 pour connaître les instructions d'installation d'Analytic Server 3.1.0.
  4. Une fois le service Analytic Server installé et ajouté dans Cloudera Manager, exécutez **Refresh Analytic Server Binaries**. Analytic Server 3.1.0 est maintenant opérationnel.

---

## Configuration de Cloudera

Après l'installation, si vous le souhaitez, vous pouvez configurer et administrer Analytic Server via Cloudera Manager.

**Remarque :** Les conventions suivantes sont utilisées pour les chemins de fichier Analytic Server.

- {AS\_ROOT} référence l'emplacement dans lequel Analytic Server est déployé, par exemple /opt/cloudera/parcels/AnalyticServer.
- {AS\_SERVER\_ROOT} référence l'emplacement des fichiers de configuration, journal et serveur, par exemple /opt/cloudera/parcels/AnalyticServer/ae\_wlpserver/usr/servers/aeserver.
- {AS\_HOME} référence l'emplacement dans le système de fichiers HDFS qui est utilisé par Analytic Server comme dossier racine, par exemple /user/as\_user/analytic-root.

## Sécurité

Le paramètre **security\_cfg** définit le registre d'utilisateurs et de groupes pouvant être ajoutés en tant que principaux au système Analytic Server.

Par défaut, un registre de base est défini avec un seul utilisateur, `admin`, dont le mot de passe est `admin`. Vous pouvez changer le registre en éditant **security\_cfg** ou en configurant Kerberos comme fournisseur de sécurité. Le paramètre **security\_cfg** se trouve dans la section **Analytic Server Advanced Configuration Snippet** de l'onglet Configuration du service Analytic Server.

**Remarque :** Si vous éditez le paramètre **security\_cfg** pour modifier le registre, vous devez ajouter les nouveaux utilisateurs en tant que principaux au système Analytic Server. Voir le manuel *IBM SPSS Analytic Server - Guide d'administration* pour plus d'informations sur la gestion des titulaires.

## Modification du registre de base

Le registre de base vous permet de définir une base de données d'utilisateurs et de groupes dans le paramètre **security\_cfg**.

Le registre de base par défaut ressemble à l'exemple qui suit.

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

Exemple de registre de base modifié :

```
<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>

```

Les mots de passe peuvent être codés afin de brouiller leur valeur via l'outil `securityUtility`, situé sous le répertoire `{RACINE_AS}/ae_wlpserver/bin`.

```

securityUtility encode changeit
 {xor}PDC+MTg6Nis=

```

**Remarque :** Voir [http://www-01.ibm.com/support/knowledgecenter/SSD28V\\_8.5.5/com.ibm.websphere.wlp.core.doc/ae/rwlp\\_command\\_securityutil.html](http://www-01.ibm.com/support/knowledgecenter/SSD28V_8.5.5/com.ibm.websphere.wlp.core.doc/ae/rwlp_command_securityutil.html) pour plus de détails sur l'outil `securityUtility`.

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

## Configuration d'un registre LDAP

Le registre LDAP vous permet d'authentifier les utilisateurs via un serveur LDAP externe tel que 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 d'autres exemples de configuration, consultez le dossier des modèles `{RACINE_AS}/ae_wlpserver/templates/config`.

**Remarque :** La prise en charge de LDAP dans Analytic Server est régie par WebSphere Liberty. Pour plus d'informations, voir Configuration de registres utilisateur LDAP dans Liberty.

## Configuration d'une connexion SSL (Secure Socket Layer) entre Analytic Server et LDAP

1. Connectez-vous à toutes les machines Analytic Server en tant qu'utilisateur Analytic Server et créez un répertoire commun pour les certificats SSL.

**Remarque :** Dans Cloudera, l'utilisateur Analytic Server est toujours `as_user` et il ne peut pas être changé.

2. Copiez le magasin de clés et le magasin de clés de confiance dans le répertoire commun de toutes les machines Analytic Server. Ajoutez également le certificat de l'autorité de certification du client LDAP au magasin de clés de confiance. Par exemple :

```

mkdir /home/as_user/security
cd /home/as_user/security
openssl s_client -connect <nom d'hôte ldap>:636 -showcerts > client.cert
$JAVA_HOME/bin/keytool -import -file ./client.cert -alias ldapCA -keystore mytrust.jks
password : changeit

```

**Remarque :** JAVA\_HOME est l'environnement d'exécution Java utilisé au démarrage d'Analytic Server.

3. Les mots de passe peuvent être codés afin de brouiller leur valeur via l'outil securityUtility, situé sous le répertoire {RACINE\_AS}/ae\_wlpsserver/bin. Par exemple :

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

4. Connectez-vous à Cloudera Manager et mettez à jour le paramètre de configuration d'Analytic Server **ssl\_cfg** avec les paramètres de configuration SSL corrects. Exemple :

```
<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="/>
```

**Remarque :** Utilisez le chemin absolu pour les fichiers du magasin de clés et du magasin de clés de confiance.

5. Mettez à jour le paramètre de configuration d'Analytic Server **security\_cfg** avec les paramètres de configuration LDAP corrects. Par exemple, dans l'élément **ldapRegistry**, définissez l'attribut **sslEnabled** sur true et l'attribut **sslRef** sur defaultSSLConfig.

## Configuration de Kerberos

Analytic Server prend en charge Kerberos dans Cloudera.

1. Créez des comptes dans le référentiel utilisateur de Kerberos pour tous les utilisateurs auxquels vous souhaitez donner accès à Analytic Server.

**Remarque :** Si l'installation Analytic Server utilise un registre de base, celui-ci doit contenir les comptes utilisateur Kerberos, avec le mot de passe "\_". Par exemple :

```
<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. Créez un compte utilisateur de système d'exploitation pour chaque utilisateur créé à l'étape précédente sur chaque noeud Analytic Server et sur chaque noeud Hadoop.
  - Assurez-vous que l'ID de ces utilisateurs correspond sur toutes les machines. Vous pouvez tester cette condition en utilisant la commande kinit pour vous connecter à chaque compte.
  - Vérifiez que l'ID utilisateur se conforme au paramètre Yarn "Minimum user ID for submitting job" (ID utilisateur minimum pour soumission de travail). Il s'agit du paramètre **min.user.id** défini dans container-executor.cfg. Par exemple, si **min.user.id** est défini à 1000, chaque compte utilisateur créé doit avoir un ID utilisateur supérieur ou égal à 1000.
3. Créez un dossier de base utilisateur sur HDFS pour tous les principaux d'Analytic Server. Par exemple, si vous ajoutez testuser1 au système Analytic Server, vous devez créer un dossier de base /user/testuser1 sur HDFS et autoriser testuser1 à y accéder en lecture et en écriture.
4. Si vous prévoyez d'utiliser des sources de données HCatalog et si Analytic Server est installé sur une machine différente de celle de Hive Metastore, vous devez simuler les droits d'accès du client Hive sur HDFS.



- a. Accédez à l'onglet Configuration du service HDFS dans Cloudera Manager.

**Remarque :** Il se peut que les paramètres suivants n'apparaissent pas dans l'onglet Configuration s'ils n'ont pas encore été définis. Dans ce cas, lancez une recherche pour les localiser.

- b. Editez le paramètre **hadoop.proxyuser.hive.groups** et entrez la valeur \*, ou un groupe contenant tous les utilisateurs autorisés à se connecter à Analytic Server.
- c. Editez le paramètre **hadoop.proxyuser.hive.hosts** et entrez la valeur \*, ou la liste des hôtes sur lesquels Hive Metastore et les instances d'Analytic Server sont installés en tant que services.
- d. Redémarrez le service HDFS.

Lorsque ces étapes ont été réalisées et qu'Analytic Server est installé, ce dernier configure Kerberos silencieusement et automatiquement.

### Configuration de HAProxy pour mécanisme de connexion unique (SSO) à l'aide de Kerberos

1. Configurez et lancez HAProxy en suivant le manuel de la version correspondante dans la documentation HAProxy : <http://www.haproxy.org/#docs>
2. Créez le principal Kerberos (HTTP/<nom\_d'hôte\_proxy>@<domaine>) et le fichier de clés pour l'hôte HAProxy, où <nom\_d'hôte\_proxy> correspond au nom complet de l'hôte HAProxy et <domaine> au domaine Kerberos.
3. Copiez le fichier de clés sur chaque hôte Analytic Server en tant que `/etc/security/keytabs/spnego_proxy.service.keytab`
4. Mettez à jour les autorisations d'accès à ce fichier sur chaque hôte Analytic Server. Par exemple :  

```
chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
```
5. Ouvrez Cloudera Manager et ajoutez ou mettez à jour les propriétés suivantes dans la zone Analytic Server **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**.  

```
web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
web.authentication.kerberos.principal=HTTP/<nom_complet_machine_proxy>@<domaine>
```
6. Sauvegardez la configuration et redémarrez tous les services Analytic Server depuis Cloudera Manager.
7. Demandez aux utilisateurs de configurer leur navigateur pour l'utilisation de Kerberos.

Les utilisateurs peuvent à présent se connecter à Analytic Server en utilisant le mécanisme de connexion unique (SSO) Kerberos.

### Activation de l'emprunt d'identité Kerberos

L'emprunt d'identité permet de lancer une unité d'exécution dans un contexte de sécurité différent de celui du processus propriétaire de l'unité d'exécution. Par exemple, l'emprunt d'identité permet aux travaux Hadoop de s'exécuter avec des noms d'utilisateur différents du nom d'utilisateur Analytic Server standard (as\_user). Pour activer l'emprunt d'identité Kerberos, procédez comme suit :

1. Ouvrez Cloudera Manager et ajoutez ou mettez à jour les propriétés suivantes dans la section Analytic Server **Analytic Server Advanced Configuration Snippet (Safety Valve) for core-site.xml**.
  - `hadoop.proxyuser.as_user.hosts = *`
  - `hadoop.proxyuser.as_user.groups = *`
2. Si Analytic Server est configuré pour utiliser un autre nom d'utilisateur que le nom d'utilisateur as\_user, vous devez modifier les noms de propriété pour prendre en compte l'autre nom d'utilisateur (par exemple `hadoop.proxyuser.xxxxx.hosts`, où xxxxx est le nom d'utilisateur configuré qui est indiqué dans la configuration Analytic Server).

**Remarque :** Les propriétés sont automatiquement ajoutées dans Ambari (en fonction des valeurs de la configuration d'Analytic Server).

### Désactivation de Kerberos

1. Désactivez Kerberos dans la console Ambari.
2. Arrêtez le service Analytic Server.
3. Supprimez les paramètres suivants de la zone **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. Cliquez sur **Save Changes** et redémarrez le service Analytic Server.

### Activation des connexions SSL (Secure Socket Layer) à la console Analytic Server

Par défaut, Analytic Server génère des certificats autosignés pour SSL (Secure Socket Layer), ce qui vous permet d'accéder à Analytic Server par le port sécurisé en acceptant ces certificats. Pour protéger davantage l'accès HTTPS, vous devez installer des certificats tiers.

Pour installer des certificats tiers, procédez comme suit.

1. Copiez le magasin de clés et les certificats du magasin de clés de confiance du fournisseur tiers dans le même répertoire sur tous les noeuds Analytic Server. Exemple : `/home/as_user/security`.

**Remarque :** L'utilisateur Analytic Server doit disposer de l'accès en lecture à ce répertoire.

2. Dans Cloudera Manager, accédez à l'onglet Configuration du service Analytic Server.
3. Editez le paramètre **ssl\_cfg**.

```
<ssl id="defaultSSLConfig"
 keyStoreRef="defaultKeyStore"
 trustStoreRef="defaultTrustStore"
 clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
 location="<KEYSTORE-LOCATION>"
 type="<TYPE>"
 password="<PASSWORD>"/>
<keyStore id="defaultTrustStore"
 location="<TRUSTSTORE-LOCATION>"
 type="<TYPE>"
 password="<PASSWORD>"/>
```

Remplacez

- `<KEYSTORE-LOCATION>` par le chemin absolu du magasin de clés, par exemple `/home/as_user/security/mykey.jks`
- `<TRUSTSTORE-LOCATION>` par le chemin absolu du magasin de clés de confiance, par exemple `/home/as_user/security/mytrust.jks`
- `<TYPE>` par le type du certificat, par exemple JKS, PKCS12, etc.
- `<PASSWORD>` par le mot de passe chiffré en base 64. Pour le codage, vous pouvez utiliser `securityUtility`. Exemple : `{AS_ROOT}/ae_wlpserver/bin/securityUtility encode <mot_de_passe>`

Si vous voulez générer un certificat autosigné, vous pouvez utiliser `securityUtility`. Exemple : `{AS_ROOT}/ae_wlpserver/bin/securityUtility createSSLCertificate --server=myserver --password=myspassword --validity=365 --subject=CN=mycompany,O=myOrg,C=myCountry`. Pour plus d'informations sur `securityUtility` et sur les autres paramètres SSL, reportez-vous à la documentation WebSphere Liberty Profile.

4. Cliquez sur **Save Changes** et redémarrez le service Analytic Server.

## Activation de la prise en charge d'Essentials for R

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

Pour installer Essentials for R après une installation réussie d'Analytic Server dans Cloudera Manager :

1. Téléchargez l'archive autoextractible (bin) du gestionnaire de packages RPM contenant IBM SPSS Modeler Essentials for R. Essentials for R peut être téléchargé (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>). Sélectionnez le fichier spécifique à votre pile, à sa version et à l'architecture matérielle.
2. Exécutez l'archive autoextractible en tant qu'utilisateur root ou sudo sur l'hôte de serveur Cloudera Manager. Les packages suivants doivent être installés ou disponibles depuis les référentiels configurés :
  - Red Hat Linux : gcc-gfortran, zip, gcc-c++
  - SUSE Linux : gcc-fortran, zip, gcc-c++
  - Ubuntu Linux : gcc-fortran, zip, gcc-c++
3. Le programme d'installation autoextractible effectue les tâches suivantes :
  - a. Il affiche les licences requises et demande au programme d'installation de les accepter.
  - b. Il invite le programme d'installation à entrer l'emplacement de la source R ou à continuer avec l'emplacement par défaut. La version de R par défaut qui est installée est 3.1.0. Pour installer une autre version :
    - Installation en ligne : indiquez l'adresse URL de l'archive de version R requise. Par exemple, <https://cran.r-project.org/src/base/R-2/R-2.15.3.tar.gz> pour R 2.15.3.
    - Installation hors ligne : téléchargez, puis copiez l'archive de version R requise sur l'hôte de serveur Cloudera Manager. Ne renommez pas l'archive (par défaut, elle s'appelle R-x.x.x.tar.gz). Indiquez l'adresse URL de l'archive R copiée comme suit : `file://<répertoire_archive_R>/R-x.x.x.tar.gz`. Si l'archive R-2.15.3.tar.gz a été téléchargée, puis copiée dans /root, l'adresse URL est `file:///root/R-2.15.3.tar.gz`.
  - c. Il installe les packages que R requiert.
  - d. Il télécharge et installe R, ainsi que le plug-in Essentials for R.
  - e. Il crée le fichier parcel et le fichier parcel.sha et les copie dans /opt/cloudera/parcel-repo. Entrez l'emplacement correct si l'emplacement a été modifié.
4. Une fois l'installation terminée, distribuez et activez le fichier parcel **Essentials for R** dans Cloudera Manager (cliquez sur **Check for New Parcels** pour actualiser la liste des parcels).
5. Si le service Analytic Server est déjà installé :
  - a. Arrêtez le service.
  - b. Actualisez les fichiers binaires d'Analytic Server.
  - c. Démarrez le service pour terminer l'installation d'Essentials for R.
6. Si le service Analytic Server n'est pas installé, installez-le.

**Remarque :** Les packages d'archive appropriés (zip et unzip) doivent être installés sur tous les hôtes Analytic Server.

## Activation des sources de base de données relationnelle

Analytic Server peut utiliser des sources de base de données relationnelle si vous rendez disponibles les pilotes JDBC dans un répertoire partagé sur chaque hôte Analytic Server. Par défaut, ce répertoire est /usr/share/jdbc.

Pour utiliser un autre répertoire partagé, procédez comme suit.

1. Dans Cloudera Manager, accédez à l'onglet Configuration du service Analytic Server.
2. Entrez le chemin du répertoire partagé des pilotes JDBC dans **jdbc.drivers.location**.
3. Cliquez sur **Save Changes**.
4. Sélectionnez **Stop** dans la liste déroulante **Actions** pour arrêter le service Analytic Server.
5. Sélectionnez **Refresh Analytic Server Binaries** dans la liste déroulante **Actions**.
6. Sélectionnez **Start** dans la liste déroulante **Actions** pour démarrer le service Analytic Server.

Tableau 7. Bases de données prises en charge

Base de données	Versions prises en charge	Fichiers jar du pilote JDBC	Fournisseur
Amazon Redshift	8.0.2 ou version ultérieure	RedshiftJDBC41-1.1.6.1006.jar ou version ultérieure	Amazon
DashDB	Service Bluemix	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

## Remarques

- Si vous avez créé une source de données Redshift avant d'installer Analytic Server, vous devez effectuer les opérations ci-dessous pour pouvoir utiliser cette source de données.
  1. Dans la console Analytic Server, ouvrez la source de données Redshift.
  2. Sélectionnez la source de données de base de données Redshift.
  3. Entrez l'adresse du serveur Redshift.
  4. Entrez le nom de la base de données et le nom d'utilisateur. Le mot de passe devrait être renseigné automatiquement.
  5. Sélectionnez la table de base de données.

## Activation des sources de données HCatalog

Analytic Server prend en charge différentes sources de données par l'intermédiaire de Hive et de HCatalog. Certaines nécessitent des opérations de configuration manuelles.

1. Collectez les fichiers JAR nécessaires pour activer la source de données. Voir les sections ci-dessous pour plus de détails.
2. Ajoutez ces fichiers JAR au répertoire {DOSSIER\_PRINCIPAL\_HIVE}/auxlib et au répertoire /usr/share/hive sur chaque noeud Analytic Server.
3. Redémarrez le service Hive Metastore.
4. Redémarrez chaque instance du service Analytic Server.

## Bases de données NoSQL

Analytic Server prend en charge toutes les bases de données NoSQL pour lesquelles un gestionnaire d'espace de stockage Hive est disponible chez le fournisseur.

La prise en charge d'Apache HBase et d'Apache Accumulo ne demande aucune opération particulière.

Pour les autres bases de données NoSQL, procurez-vous le gestionnaire d'espace de stockage et les fichiers JAR associés auprès du fournisseur de la base.

## Tables Hive sous forme de fichiers

Analytic Server prend en charge toutes les tables Hive sous forme de fichiers pour lesquelles un sérialiseur-désérialiseur (SerDe) Hive intégré ou personnalisé est disponible.

Le sérialiseur-désérialiseur XML Hive pour le traitement des fichiers XML est stocké dans le référentiel Maven Central à l'adresse <http://search.maven.org/#search%7Cga%7C1%7Cchivexmlserde>.

## Apache Spark

Si vous souhaitez utiliser Spark (version 1.5 ou suivante) avec une source de données d'entrée HCatalog, vous devez ajouter manuellement la propriété `spark.version=X.X.0` (par exemple `spark.version=2.0.0`)

1. Ouvrez Cloudera Manager et ajoutez ou mettez à jour les propriétés suivantes dans la section Analytic Server **Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**.

```
spark.version=2.0.0
```

2. Sauvegardez la configuration et redémarrez tous les services Analytic Server depuis Cloudera Manager.

## Configuration d'Apache Impala

Apache Impala est pris en charge lors de l'exécution sur Cloudera avec une source de données de base de données Analytic Server ou une source de données HCatalog (qu'Impala soit configuré pour SSL ou non).

## Création d'une source de données de base de données pour les données Apache Impala

1. Dans la page principale d'Analytic Server intitulée **Data sources**, cliquez sur **New** pour créer une source de données. La boîte de dialogue **New data source** s'ouvre.
2. Entrez un nom approprié dans la zone **New data source**, sélectionnez **Database** dans la zone **Content type**, puis cliquez sur **OK**.
3. Ouvrez la section **Database Selections** et entrez les informations ci-dessous.

### Database :

Sélectionnez **Impala** dans le menu déroulant.

### Server address :

Entrez l'adresse URL du serveur qui héberge le démon Impala. Un nom de domaine qualifié complet est requis lorsque Kerberos est activé pour Analytic Server.

### Server port :

Entrez le numéro de port sur lequel la base de données Impala est à l'écoute.

### Database name :

Entrez le nom de la base de données à laquelle vous voulez vous connecter.

### Username :

Entrez un nom d'utilisateur pouvant se connecter à la base de données Impala.

**Password :**

Entrez le mot de passe de l'utilisateur approprié.

**Table name :**

Entrez le nom d'une table de la base de données à utiliser. Cliquez sur **Select** pour sélectionner un fichier manuellement.

**Maximum concurrent reads :**

Entrez le nombre maximal de requêtes parallèles pouvant être envoyées depuis Analytic Server à la base de données pour lecture de la table spécifiée dans la source de données.

4. Cliquez sur **Save** une fois que vous avez fini d'entrer les informations requises.

## Création d'une source de données HCatalog pour les données Apache Impala

1. Dans la page principale d'Analytic Server intitulée **Data sources**, cliquez sur **New** pour créer une source de données. La boîte de dialogue **New data source** s'ouvre.
2. Entrez un nom approprié dans la zone **New data source**, sélectionnez **HCatalog** dans la zone **Content type**, puis cliquez sur **OK**.
3. Ouvrez la section **Database Selections** et entrez les informations ci-dessous.

**Database :**

Sélectionnez **default** dans le menu déroulant.

**Table name :**

Entrez le nom d'une table de la base de données à utiliser.

**HCatalog Schema**

Sélectionnez l'option **HCatalog Element**, puis sélectionnez les options **HCatalog Field Mappings** appropriées.

4. Cliquez sur **Save** une fois que vous avez fini d'entrer les informations requises.

## Connexion à des données compatibles avec Apache Impala

1. Définissez les paramètres SSL Impala ci-après dans la console Analytic Server.

**Enable TLS/SSL for Impala (client\_services\_ssl\_enabled)**

Sélectionnez l'option **Impala (Service-Wide)**.

**Impala TLS/SSL Server Certificate File (PEM Format) (ssl\_server\_certificate)**

Entrez le nom de fichier et l'emplacement du certificat autosigné au format PEM (par exemple /tmp/<nom\_utilisateur>/ssl/114200v21.crt).

**Impala TLS/SSL Server Private Key File (PEM Format) (ssl\_private\_key)**

Entrez le nom de fichier et l'emplacement de la clé privée au format PEM (par exemple /tmp/<nom\_utilisateur>/ssl/114200v21.key).

2. Sur l'hôte Analytic Server, importez le fichier \*.crf (qui est utilisé afin d'activer le protocole SSL pour Impala) dans un fichier \*.jks. Le fichier peut être un fichier cacerts (par exemple /etc/pki/java/cacerts) ou un autre fichier \*.jks.
3. Sur l'hôte Analytic Server, mettez à jour le fichier de configuration d'Impala (impala.properties) en ajoutant la valeur de clé jdbcurl suivante :

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

**Remarque :** lorsqu'un fichier \*.jks (autre que cacerts) est utilisé, vous devez aussi spécifier les éléments suivants :

```
SSLTrustStore=<votre_fichier_pks>;SSLTrustStorePwd=<mot_de_passe_fichier_pks>;
```

4. Redémarrez Analytic Server dans la console Cloudera Manager.

## Modification des ports utilisés par Analytic Server

Analytic Server utilise par défaut le port 9080 pour HTTP et 9443 pour HTTPS. Pour modifier les paramètres de port, procédez comme suit.

1. Dans Cloudera Manager, accédez à l'onglet Configuration du service Analytic Server.
2. Spécifiez les ports HTTP et HTTPS de votre choix dans les paramètres **http.port** et **https.port** respectivement.

**Remarque :** Il peut être nécessaire de sélectionner la catégorie **Ports and Addresses** dans la section Filters pour que ces paramètres s'affichent.

3. Cliquez sur **Save Changes**.
4. Redémarrez le service Analytic Server.

## Haute disponibilité d'Analytic Server

Vous pouvez garantir la haute disponibilité d'Analytic Server en le définissant en tant que service à plusieurs noeuds de votre cluster.

1. Dans Cloudera Manager, accédez à l'onglet Instances du service Analytic Server.
2. Cliquez sur **Add Role Instances** et sélectionnez les hôtes sur lesquels ajouter Analytic Server en tant que service.

## Optimisation des options JVM pour les petits volumes de données (Small Data)

Vous pouvez éditer les propriétés JVM pour optimiser votre système en cas d'exécution de petits travaux (M3R).

Dans Cloudera Manager, reportez-vous au contrôle **Jvm Options (jvm.options)** dans l'onglet Configuration du service Analytic Server. La modification des paramètres ci-après définit la taille de segment de mémoire des travaux s'exécutant sur le serveur hébergeant Analytic Server (pas le serveur Hadoop). Cette option est importante pour l'exécution de petits travaux (M3R) et vous devrez éventuellement tester différentes valeurs afin d'optimiser votre système.

```
-Xms512M
-Xmx2048M
```

---

## Migration

Analytic Server vous permet de migrer des données et des paramètres de configuration d'une installation Analytic Server existante vers une nouvelle installation.

### Mise à niveau vers une nouvelle version d'Analytic Server

Si vous disposez d'une installation Analytic Server 2.0/2.1 existante et avez fait l'acquisition d'une version plus récente, vous pouvez migrer vos paramètres de configuration 2.0/2.1 vers votre nouvelle installation.

**Restriction :** Si votre installation est antérieure à la version 2.0, vous devez d'abord migrer depuis cette version vers la version 2.0/2.1, puis de la version 2.0/2.1 vers la nouvelle version.

**Restriction :** Votre installation 2.0/2.1 et votre nouvelle installation ne peuvent pas coexister dans le même cluster Hadoop. Si vous configurez votre nouvelle installation pour utiliser le même cluster Hadoop que votre installation 2.0/2.1, l'installation 2.0/2.1 ne fonctionnera plus.

### Étapes de migration depuis la version 2.1 vers une version plus récente

1. Installez la nouvelle installation d'Analytic Server en suivant les instructions figurant dans «Installation dans Cloudera», à la page 30.
2. Copiez la racine d'analyse de l'ancienne installation dans la nouvelle.

a. Si vous n'êtes pas certain de l'emplacement de la racine d'analyse, exécutez la commande `hadoop -fs ls`. Le chemin de la racine d'analyse aura le format `/user/aeuser/analytic-root`, où `aeuser` est l'ID utilisateur auquel appartient la racine d'analyse.

b. Remplacez le propriétaire de la racine d'analyse (`aeuser`) par `as_user`.

```
hadoop dfs -chown -R {as_user:{groupe}} {chemin analytic-root 2.1}
```

**Remarque :** Si vous prévoyez d'utiliser l'installation Analytic Server existante après la migration, déposez une copie du répertoire `analytic-root` dans HDFS, puis changez le propriétaire de la copie.

c. Connectez-vous à l'hôte de la nouvelle installation Analytic Server en tant que `as_user`. Supprimez le répertoire `/user/as_user/analytic-root` s'il existe.

d. Exécutez le script de copie suivant :

```
hadoop distcp hftp://{hôte namenode 2.1}:50070/{chemin analytic-root 2.1}
hdfs://{hôte namenode 3.1}/user/as_user/analytic-root
```

3. Dans Cloudera Manager, arrêtez le service Analytic Server.

4. Collectez les paramètres de configuration de l'ancienne installation.

a. Copiez l'archive `configcollector.zip` de la nouvelle installation dans le répertoire `{AS_ROOT}\tools` de l'ancienne installation.

b. Décompressez la copie de `configcollector.zip`. Cette opération crée un sous-répertoire `configcollector` dans l'ancienne installation.

c. Exécutez l'outil de collecte de la configuration dans votre ancienne installation en lançant le script **configcollector** situé sous `{RACINE_AS}\tools\configcollector`. Copiez le fichier compressé (ZIP) généré sur le serveur qui héberge la nouvelle installation.

5. Lancez l'outil de migration en exécutant le script **migrationtool** et en transmettant sous forme d'argument le chemin du fichier compressé créé par le collecteur de configuration. Par exemple :

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

6. Effacez l'état de Zookeeper. Dans le répertoire `bin` de Zookeeper (par exemple `/opt/cloudera/parcels/CDH-5.4...../lib/zookeeper/bin` on Cloudera), exécutez la commande suivante :

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

7. Dans Cloudera Manager, démarrez le service Analytic Server.

**Remarque :** Si vous avez configuré R pour son utilisation avec l'installation Analytic Server existante, vous devez suivre les étapes permettant de le configurer avec la nouvelle installation Analytic Server.

---

## Désinstallation d'Analytic Server dans Cloudera

Cloudera gère automatiquement la plupart des étapes requises pour désinstaller le service et le fichier `parcel` Analytic Server.

Les étapes suivantes sont requises pour supprimer Analytic Server de l'environnement Cloudera :

1. Arrêtez et supprimez le service Analytic Server.

2. **Désactivez et supprimez des hôtes** les fichiers `parcel` Analytic Server.

3. Supprimez le répertoire de l'utilisateur Analytic Server dans le système de fichiers HDFS. L'emplacement par défaut est `/user/as_user/analytic-root`.

4. Supprimez la base de données ou le schéma qu'Analytic Server utilise.



---

## Chapitre 4. Installation et configuration de MapR

---

### Présentation de MapR

MapR est une distribution complète pour Apache Hadoop qui regroupe plus d'une douzaine de projets de l'écosystème Hadoop afin de fournir un large éventail de capacités pour les mégadonnées.

Le système de fichiers MapR n'est pas accessible hors du cluster de serveurs. Par conséquent, IBM SPSS Analytic Server doit être déployé sur les noeuds de cluster MapR. Dans ce scénario de déploiement, Analytic Server doit être exécuté par un utilisateur qui dispose du droit d'accès au système de fichier MapR ainsi que du droit permettant de soumettre des travaux à yarn afin de procéder au déploiement dans Analytic Server (en tant que <utilisateur\_as>).

---

### Installation d'Analytic Server dans MapR

Les étapes ci-après détaillent le processus d'installation manuelle d'IBM SPSS Analytic Server dans un cluster MapR.

#### Installation d'Analytic Server 3.1.0 sur MapR 5.0 ou 5.1

1. Accédez au [site Web d'IBM Passport Advantage](#) et téléchargez le fichier binaire autoextractible MapR.

Tableau 8. Fichiers binaires autoextractibles MapR

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.1 pour MapR 5.0 et 5.1 Linux x86-64 (Anglais)	spss_as-3.1.0-mapr5.0-5.1-lx86_en.bin

2. Exécutez le programme d'installation d'Analytic Server à l'aide d'un ID superutilisateur (root) ou sudo. Suivez les invites d'installation afin d'accepter la licence et choisissez d'installer Analytic Server en ligne ou hors ligne.
  - a. Sélectionnez l'option d'installation en ligne lorsque le serveur qui héberge Analytic Server dispose d'une connexion à <https://ibm-open-platform.ibm.com> via Internet. Le programme d'installation installe Analytic Server automatiquement.
  - b. Sélectionnez l'option d'installation hors ligne si le serveur qui héberge Analytic Server ne possède pas de connexion à <https://ibm-open-platform.ibm.com> via Internet. Exécutez le programme d'installation sur un autre serveur ayant accès à l'adresse URL et choisissez d'installer Analytic Server hors ligne. Le programme d'installation télécharge automatiquement le package RPM ou DEB.
3. Recherchez et exécutez le package RPM ou DEB pour Analytic Server:

- RedHat ou SuSe Linux :

```
rpm -ivh IBM-SPSS-AnalyticServer-3.1.0-1.x86_64.rpm
```
- Ubuntu Linux :

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

Pour les modes d'installation en ligne et hors ligne, Analytic Server est installé dans /opt/ibm/spss/analyticserver/3.1 (dans <chemin\_installation\_as>).

4. Attribuez tous les fichiers qui se trouvent dans le chemin d'installation à l'utilisateur qui exécute Analytic Server:

```
chown
-R <as_user> <chemin_installation_as>
```

Basculez vers l'utilisateur <as\_user> ; toutes les étapes suivantes utilisent <as\_user>.

- Configurez la propriété HTTP. Créez un fichier nommé `http_endpoint.xml` dans le chemin `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver` et ajoutez-y les lignes suivantes :

```
<server>
 <httpEndpoint host="*" id="defaultHttpEndpoint" httpPort="<port_http>" httpsPort="<port_https>" onError="FAIL"/>
</server>
```

`<port_http>` et `<port_https>` sont les ports utilisés par Analytic Server via les protocoles HTTP et HTTPS. Remplacez-les par des ports disponibles .

- Ajoutez des utilisateurs et des groupes. Créez un fichier nommé `security_cfg.xml` dans le chemin `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver` et ajoutez-y les lignes suivantes :

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

A l'état par défaut, le fichier XML ne contient que l'utilisateur admin. Vous devez ajouter manuellement d'autres utilisateurs et groupes dans le paramètre `<basicRegistry>` ou remplacer le paramètre par `ldapRegistry`.

- Configurez la base de données des métadonnées. Analytic Server prend en charge les bases de données DB2 et MySQL.
  - Configurez les utilisateurs de base de données. Lorsque la base de données MySQL est utilisée, exécutez le script SQL dans l'interpréteur de commandes MySQL :

```
DROP DATABASE IF EXISTS <nom_bd>;
CREATE DATABASE <nom_bd> DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8_bin;
CREATE USER '<nom_utilisateur_bd>'@'%' IDENTIFIED BY '<mot_de_passe_bd>';
CREATE USER '<nom_utilisateur_bd>'@'localhost' IDENTIFIED BY '<mot_de_passe_bd>';
GRANT ALL PRIVILEGES ON *.* TO '<nom_utilisateur_bd>'@'%;
GRANT ALL PRIVILEGES ON *.* TO '<nom_utilisateur_bd>'@'localhost';
```

- Chiffrez le mot de passe. Les mots de passe des utilisateurs de base de données doivent être chiffrés pour pouvoir être transmis à Analytic Server. Exécutez la commande suivante :

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

**Remarque :** Si la commande est exécutée directement dans un interpréteur de commandes Linux, il se peut que vous deviez ajouter un caractère d'échappement à \* et indiquer \\*.

Le résultat de la commande est le suivant : Le mot de passe chiffré est '`<mot_de_passe_bd_chiffré>`'. Enregistrez le mot de passe de base de données chiffré.

- Supprimez le fichier `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`, s'il existe, et créez un autre fichier du même nom. Changez les propriétés suivantes si la base de données DB2 est utilisée :

```
jndi.aedb=jdbc/aeds
jndi.aedb.url=jdbc:db2://<hôte_bd>:<port_bd>/<nom_bd>;currentSchema=<nom_schéma_bd>;
jndi.aedb.driver=com.ibm.db2.jcc.DB2Driver
jndi.aedb.username=<nom_utilisateur_bd>
jndi.aedb.password=<mot_de_passe_bd_chiffré>
```

Si le schéma `<nom_schéma_bd>` n'existe pas, l'utilisateur `<nom_utilisateur_bd>` doit disposer du droit implicite de création du schéma. Changez les propriétés suivantes si la base de données MySQL est utilisée :

```
jndi.aedb=jdbc/aeds
jndi.aedb.url=jdbc:mysql://<hôte_bd>:<port_bd>/<nom_bd>;createDatabaseIfNotExist=true
jndi.aedb.driver=com.mysql.jdbc.Driver
jndi.aedb.username=<nom_utilisateur_bd>
jndi.aedb.password=<mot_de_passe_bd_chiffré>
```

- Le pilote JDBC MySQL doit être installé lorsque la base de données MySQL est utilisée. Exécutez la commande suivante :

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

- Exécutez la commande suivante pour créer les tables requises :

```
cd <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/sql/<type_bd>
java -Xmx128m -Xms128m -cp <chemin_installation_as>/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
```

`<type_bd>` est `db2` ou `mysql`, selon la base de données utilisée.

**Remarque :** Lorsque la base de données MySQL est utilisée avec le moteur MYISAM, la deuxième commande signale les messages d'erreur suivants, que vous pouvez ignorer :

```
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. Exécutez la commande suivante pour décompresser la bibliothèque cf :

```
cd <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration
unzip cf.zip
```

9. Configurez le chemin d'accès aux classes des modules de connexion JAAS en créant un fichier nommé `private_library.xml` dans le chemin `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver` et entrez les informations suivantes dans le fichier :

```
<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>
```

**Remarque :** L'exemple précédent permet de configurer le module de connexion `hadoop_simple`. Vous devez modifier la configuration si MapR utilise d'autres modules de connexion.

10. Vérifiez si le fichier `ASModules.xml` existe dans le chemin `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/`. S'il n'existe pas, renommez le fichier `ASModules.xml.template` (dans le même chemin) en `ASModules.xml`.
11. Configurez les informations de cluster en ajoutant les propriétés suivantes dans le fichier `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`.

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

La propriété `ae.cluster.zookeeper.connect.string` indique la liste des noeuds zookeeper séparés par une virgule. Elle peut partager le cluster zookeeper utilisé par MapR. `ae.cluster.member.name` indique le nom du noeud qui héberge Analytic Server.

L'exemple suivant présente le format de la propriété `ae.cluster.zookeeper.connect.string` :

```
ae.cluster.zookeeper.connect.string=<hôte zookeeper 1>:<port zookeeper 1>,<hôte zookeeper 2>:<port zookeeper 2>,<hôte zookeeper 3>:<port zookeeper 3>...
```

Lorsqu'Analytic Server partage le cluster zookeeper avec MapR, la valeur de la propriété `ae.cluster.zookeeper.connect.string` doit être identique à celle de la propriété `zookeeper.servers` dans le fichier `warden.conf` de MapR. (L'emplacement par défaut du fichier est `/opt/mapr/conf/`.)

12. Ouvrez le fichier `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/server.env` et ajoutez-y les lignes suivantes :

```
JAVA_HOME=<rép_base_java>
PATH=<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<java_home>/jre/lib/amd64:/usr/sbin:/usr/bin:/sbin:/bin
IBM_SPSS_AS_NATIVE_PATH=<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64
LD_LIBRARY_PATH=<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<java_home>/jre/lib/amd64:/opt/mapr/hadoop/hadoop-2.7.0/lib/native
```

Remplacez <chemin\_installation\_as> et <r ep\_base\_java> par le chemin d'installation et le chemin du r epertoire de base Java r eels.

- Editez la racine d'analyse en ouvrant le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/configuration/config.properties et en ajoutant la ligne suivante :

```
distrib.fs.root=<racine_analyse>
```

<racine\_analyse> est un chemin dans le syst eme de fichiers MapR qui h eberge les fichiers distants d'Analytic Server essentiels. Le chemin recommand e est /user/<as\_user>/analytic-root.

- D efinissez l'administrateur en ouvrant le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/configuration/config.properties et en ajoutant la ligne suivante :

```
admin.username=admin
```

La valeur doit  tre un nom d'administrateur Analytic Server et l'un des utilisateurs qui est configur e dans le fichier security\_cfg.xml.

- T el echargez les d ependances d'Analytic Server dans le syst eme de fichiers MapR en ajoutant la ligne suivante   la ligne 69 dans le fichier <chemin\_installation\_as>/bin/hdfsUpdate.sh :

```
JAVA_CLASS_PATH="hadoop classpath":$JAVA_CLASS_PATH
```

Ex ecutez les commandes suivantes pour cr eer <racine\_analyse> :

```
cd
<chemin_installation_as>/bin
./hdfsUpdate.sh
```

<as\_user> doit disposer du droit d'acc es en  criture au r epertoire parent <racine\_analyse>.

- D emarrez et arr etez Analytic Server.
  - Ex ecutez la commande suivante pour d emarrer Analytic Server :

```
cd
<chemin_installation_as>/ae_wlpserver/bin
./server start aeserver
```

- Ex ecutez la commande suivante pour arr eter Analytic Server :

```
cd
<chemin_installation_as>/ae_wlpserver/bin
./server stop aeserver
```

## Installation d'Analytic Server 3.1.0 sur MapR 5.2

- Acc edez au [site Web d'IBM Passport Advantage](#) et t el echargez le fichier binaire autoextractible MapR.

Tableau 9. Fichiers binaires autoextractibles MapR

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.1.0.0 pour MapR 5.0, 5.1 et 5.2 Linux x86-64 (Anglais)	spss_as3.1.0.0-mapr5-5.2-1x86_en.bin

- Les  tapes restantes pour installer Analytic Server sont principalement les m emes que celles permettant d'installer Analytic Server 3.1.0 sur MapR 5.0 ou 5.1. En revanche, les informations de «Activation d'Apache HBase»,   la page 48 et de «Activation d'Apache Spark»,   la page 49 diff erent entre MapR 5.1 et 5.2. Pour plus d'informations sur l'installation sur MapR 5.2, reportez-vous aux rubriques suivantes :

---

## Configuration de MapR

Apr es l'installation, si vous le souhaitez, vous pouvez configurer et administrer des fonctions MapR pour Analytic Server.

## Activation de la r epercussion de base de donn ees

La r epercussion de base de donn ees est la pratique qui consiste   lire des donn ees depuis une base de donn ees et   les traiter directement.

IBM SPSS Analytic Server prend en charge la r epercussion pour les bases de donn ees suivantes :

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

Suivez les étapes ci-dessous pour activer la répercussion de base de données.

1. Copiez les fichiers JAR de pilote JDBC appropriés dans <chemin\_installation\_as>/jdbc.
2. Ouvrez le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/private\_library.xml, localisez la bibliothèque de balises dont l'ID est maprLib, puis ajoutez la ligne suivante dans la balise :

```
<fileset dir="<chemin_installation_as>/jdbc" includes="*.jar"/>
```

3. Exécutez les commandes suivantes :

```
cd <chemin_installation_as>/jdbc
hadoop fs -put *.jar <racine_analyse_as>/cluster1/classpath
```

4. Redémarrez Analytic Server.

## Activation d'Apache Hive

Apache Hive est une infrastructure d'entrepôt de données qui est construite sur Hadoop afin de fournir des fonctions de récapitulatif de données, d'interrogation et d'analyse.

**Remarque :** Hive doit être configuré en vue de l'utilisation de MySQL comme métamagasin. Le fichier hive-site.xml qui existe sur le noeud qui héberge IBM SPSS Analytic Server doit être identique au fichier qui se trouve sur le noeud qui exécute le métamagasin Hive.

Pour activer la prise en charge d'Apache Hive après une installation de MapR réussie :

1. Téléchargez les dépendances Hive et hcatalog dans le système de fichiers MapR en exécutant les commandes suivantes :

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

<racine\_analyse\_as> est le chemin d'accès à la racine d'analyse défini dans «Installation d'Analytic Server dans MapR», à la page 43.

2. Ouvrez le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/private\_library.xml, localisez la bibliothèque de balises dont l'ID est maprLib, puis ajoutez les lignes suivantes dans la balise :

```
<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. Exécutez les commandes suivantes pour créer les liens vers les fichiers de configuration Hive et hcatalog :

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

4. Ajoutez la ligne suivante dans le fichier private\_library.xml lorsqu'il existe des fichiers jar supplémentaires dans le dossier auxlib pour Hive :

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

Exécutez les commandes suivantes après avoir ajouté la ligne précédente :

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

5. Redémarrez Analytic Server.

## Exécution de Hive en mode HTTP

Par défaut, Hive opère en mode binaire (mode TCP). Pour exécuter Hive en mode HTTP, vous devez mettre à jour les propriétés de configuration Hive suivantes (en particulier la propriété `hive.server2.transport.mode`).

**Remarque :** Pour plus d'informations sur chaque propriété, voir Hive Configuration Properties.

Tableau 10. Propriétés Hive pour mode HTTP

Nom de propriété	Valeur par défaut	Description
<code>hive.server2.transport.mode</code>	binaire	Mode de transport du serveur. La valeur peut être <code>binary</code> ou <code>http</code> . Affectez-lui la valeur <code>http</code> pour activer le mode de transport HTTP.
<code>hive.server2.thrift.http.port</code>	10001	Numéro de port en mode HTTP.
<code>hive.server2.thrift.http.path</code>	<code>cliservice</code>	Composant de chemin du noeud final d'URL en mode HTTP.
<code>hive.server2.thrift.http.min.worker.threads</code>	5	Nombre minimum d'unités d'exécution de tâche dans le pool de serveurs en mode HTTP.
<code>hive.server2.thrift.http.max.worker.threads</code>	500	Nombre maximum d'unités d'exécution de tâche dans le pool de serveurs en mode HTTP.

**Remarque :** Hive doit être redémarré après la mise à jour des propriétés.

## Activation d'Apache HBase

Apache HBase est une base de données répartie open source non relationnelle écrite en Java. Elle est développée dans le cadre du projet Apache Hadoop d'Apache Software Foundation et s'exécute sur le système de fichiers HDFS (Hadoop Distributed Filesystem).

Pour activer la prise en charge d'Apache HBase après une installation de MapR réussie :

### IBM SPSS Analytic Server 3.1.0 on MapR 5.0/5.1

1. Téléchargez les dépendances HBase dans le système de fichiers MapR et exécutez les commandes suivantes :

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

<racine\_analyse\_as> est le chemin d'accès à la racine d'analyse défini dans «Installation d'Analytic Server dans MapR», à la page 43.

2. Ouvrez le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/private\_library.xml, localisez la bibliothèque de balises dont l'ID est `maprLib`, puis ajoutez la ligne suivante dans la balise :

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

3. Exécutez les commandes suivantes pour créer les liens vers les fichiers de configuration HBase et hcatalog :

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

4. Redémarrez IBM SPSS Analytic Server.

## IBM SPSS Analytic Server 3.1.0 on MapR 5.2

1. Téléchargez les dépendances HBase sur le système de fichiers MapR en exécutant les commandes suivantes :

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

<racine\_analyse\_as> est le chemin défini à la 12e étape de la rubrique «Installation d'Analytic Server dans MapR», à la page 43.

2. Ouvrez <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/private\_library.xml et recherchez la bibliothèque de balises associée à l'ID maprLib. Ajoutez la balise suivante :

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

3. Exécutez les commandes suivantes pour créer des liens pour les fichiers de configuration Hive et Hcatalog :

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

4. Ajoutez la ligne suivante à <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/configuration/config.properties

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

5. Redémarrez Analytic Server.

## Activation d'Apache Spark

Apache Spark est une norme ouverte de traitement souple des données en mémoire pour l'analytique évoluée, par lots et en temps réel.

Pour activer la prise en charge d'Apache Spark après une installation de MapR réussie :

## IBM SPSS Analytic Server 3.1.0 on MapR 5.0/5.1

1. Copiez le fichier spark-assembly-1.4.1-hadoop2.5.1-mapr-1501.jar de /opt/mapr/spark/spark-1.4.1/lib dans <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/modules/spark/.
2. Téléchargez les dépendances Spark dans le système de fichiers MapR et exécutez les commandes suivantes :

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

<racine\_analyse\_as> est le chemin d'accès à la racine d'analyse défini dans «Installation d'Analytic Server dans MapR», à la page 43.

3. Ouvrez le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/private\_library.xml, localisez la bibliothèque de balises dont l'ID est maprLib, puis ajoutez la ligne suivante dans la balise :

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

4. Exécutez les commandes suivantes pour créer des liens vers les fichiers de configuration Spark :

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

5. Ajoutez la ligne suivante dans le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/server.env :

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

6. Ajoutez la ligne suivante dans le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/configuration/config.properties :

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

7. Redémarrez IBM SPSS Analytic Server.

8. Pour activer la fonction PySpark, ajoutez la ligne suivante dans le fichier yarn-env.sh, puis redémarrez les gestionnaires de ressources et les gestionnaires de noeuds :

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

## IBM SPSS Analytic Server 3.1.0 on MapR 5.2

Les étapes varient en fonction de la version de Spark.

### Spark 1.x

1. Copiez le fichier spark-assembly-1.4.1-hadoop2.5.1-mapr-1501.jar stocké dans :  
`/opt/mapr/spark/spark-1.4.1/lib`  
  
dans  
`<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/modules/spark/`
2. Ouvrez le fichier suivant :  
`<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/private_library.xml`  
  
et recherchez la bibliothèque de balises associée à l'ID maprLib. Ajoutez la balise suivante :  
`<fileset dir="/opt/mapr/spark/spark-1.4.1/lib" includes="spark-assembly-*.jar"/>`
3. Supprimez le fichier suivant :  
`<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/modules/spark/com.ibm.spss.sparkmapreduce_2-3.1.0.0.jar`

### Spark 2.x

1. Supprimez le fichier suivant :  
`<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/modules/spark/com.ibm.spss.sparkmapreduce-3.1.0.0.jar`
2. Ouvrez le fichier suivant :  
`<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/private_library.xml`  
  
et recherchez la bibliothèque de balises associée à l'ID maprLib. Ajoutez les lignes suivantes à la balise  
:  
`<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_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark" includes="*.jar"/>`
3. Ajoutez la ligne spark.version=2.0 au fichier suivant :  
`<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`

Les étapes ci-dessous sont communes à Spark 1.x et 2.x et doivent être effectuées après les étapes Spark 1.x ou 2.x précédentes.

**Remarque :** Toutes les références à `<spark_version>` doivent être remplacées par la version Spark en cours (par exemple 1.4.1 ou 2.0.1).

1. Importez les dépendances Spark sur le système de fichiers MapR en exécutant les commandes suivantes :  
`cd <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/modules/spark/`  
`hadoop fs -put *.jar <racine_analyse_as>/cluster1/classpath`  
  
`<racine_analyse_as>` est le chemin défini à la 12e étape de la rubrique «Installation d'Analytic Server dans MapR», à la page 43.
2. Exécutez les commandes suivantes pour créer des liens pour les fichiers de configuration Spark :  
`mkdir <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf`  
`ln -s /opt/mapr/spark/spark-<version_spark>/conf/*`  
`<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf`
3. Ajoutez la ligne suivante au fichier `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/server.env`  
`SPARK_HOME=/opt/mapr/spark/spark-<spark_version>`



- Ajoutez la ligne suivante dans le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/configuration/config.properties  

```
spark.executor.extraLibraryPath=/opt/mapr/hadoop/hadoop-2.7.0/lib/native
```
- Redémarrez Analytic Server.
- Si vous devez activer la fonction PySpark, ajoutez les lignes suivantes au fichier yarn-env.sh :  

```
export SPARK_HOME=/opt/mapr/spark/spark-<version_spark>
```

Redémarrez ResourceManagers et NodeManagers.

## Activation d'indicateurs de fonction

Les indicateurs de fonction permettent d'activer et de désactiver des fonctions d'application spécifiques.

Pour activer la prise en charge des indicateurs de fonction après une installation de MapR réussie :

- Ajoutez la ligne suivante dans le fichier <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/configuration/config.properties :  

```
load.feature.flags.on.msg=true
```
- Redémarrez IBM SPSS Analytic Server.

## Activation de R

R est un langage et un environnement pour l'analyse statistique et les graphiques.

Pour activer la prise en charge de R après une installation de MapR réussie :

**Remarque :** Vous devez installer le package suivant pour pouvoir exécuter le programme d'installation sur tous les noeuds de cluster qui hébergent le gestionnaire de noeuds et IBM SPSS Analytic Server :

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

- Exécutez le programme d'installation spss\_er-8.4.0.0-mapr5-1x86\_64\_en.bin sur tous les noeuds de cluster hébergeant le gestionnaire de noeuds et Analytic Server. L'utilisateur qui exécute le programme d'installation doit disposer du droit d'accès en écriture dans les chemins d'installation de R et Analytic Server.
- Suivez les instructions d'installation en acceptant le contrat de licence et entrez les informations requises. Si Analytic Server est installé sur le serveur d'installation, choisissez Oui lorsque vous y êtes invité et entrez <chemin\_installation\_as>. Si Analytic Server n'est pas installé sur le serveur d'installation, choisissez Non lorsque vous y êtes invité.
- Lorsqu'Analytic Server est installé, Essentials for R est installé automatiquement dans le chemin d'installation d'Analytic Server.
  - Si Analytic Server n'est pas installé, Essentials for R est installé dans le chemin <chemin\_programme\_installation>/IBM\_SPSS\_ModelerEssentialsR/linux.
  - Si Analytic Server est installé ultérieurement, utilisez la commande suivante pour copier Essentials for R dans le chemin de configuration d'Analytic Server dans lequel Analytic Server est installé :  

```
cp -r <chemin_programme_installation>/IBM_SPSS_ModelerEssentialsR/linux
<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration
```
- Supprimez le fichier cf.zip dans le chemin <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver/configuration et générez un nouveau fichier avec les commandes suivantes :  

```
cd <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration
zip -r cf.zip linux
```
- Exécutez les commandes suivantes :  

```
cd <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration
hadoop fs -rm <racine_analyse_as>/cluster1/configuration/cf.zip
hadoop fs -put cf.zip <racine_analyse_as>/cluster1/configuration/
```
- Redémarrez Analytic Server.

## Activation de LZO

LZO est une bibliothèque de compression de données sans perte qui favorise la vitesse plutôt que le taux de compression. MapR doit être configuré manuellement pour fournir la prise en charge de LZO.

Le site suivant fournit des instructions d'installation et de configuration de LZO : <https://github.com/twitter/hadoop-lzo/>

Les étapes ci-après détaillent le processus d'importation d'une bibliothèque LZO dans MapR.

1. Copiez le fichier `hadoop-lzo-<version>.jar` dans le chemin d'accès aux classes Hadoop. Le chemin recommandé est `/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/mapreduce/lib`.
2. Copiez les fichiers natifs `libgplcompression.so` et `liblzo2.so.2` dans `/opt/mapr/hadoop/hadoop-2.7.0/lib/native`, puis ajoutez les propriétés suivantes dans le fichier `core-site.xml` :

```
<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. Ouvrez le fichier `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/server.env` et ajoutez `<chemin_natif_lzo>` au paramètre `LD_LIBRARY_PATH`. `<chemin_natif_lzo>` est le dossier contenant la bibliothèque native Hadoop-LZO.

```
LD_LIBRARY_PATH=<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<rép_base_java>/jre/lib/amd64:/opt/mapr/hadoop/hadoop-2.7.0/lib/native:<chemin_natif_lzo>
```

4. Redémarrez IBM SPSS Analytic Server.

## Configuration d'un cluster IBM SPSS Analytic Server pour MapR

Suivez les étapes ci-dessous pour configurer un environnement de cluster IBM SPSS Analytic Server pour la prise en charge de MapR.

1. Ajoutez la ligne suivante dans le fichier `<chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`.  
`enable.resume=true`
2. Copiez le chemin d'installation sur les autres noeuds de cluster et remplacez la valeur de la propriété `ae.cluster.member.name` dans le fichier `config.properties` par le nom d'hôte correct.
3. Démarrez tous les noeuds de cluster.

---

## Désinstallation de MapR

Les étapes suivantes expliquent le processus de désinstallation de MapR :

1. Arrêtez IBM SPSS Analytic Server.
2. Supprimez la base de données des métadonnées.
  - a. Exécutez les commandes suivantes :

```
cd <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/sql/<type_bd>
java -Xmx128m -Xms128m -cp <chemin_installation_as>/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. Exécutez l'instruction SQL suivante pour supprimer la base de données :

```
drop database <nom_bd>
```

3. Désinstallez le package RPM :

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

4. Supprimez le chemin d'installation :

```
rm -r <chemin_installation_as>
```

5. Supprimez la racine d'analyse :

```
hadoop fs -rm -r <racine_analyse>
```

## 6. Supprimez les données zookeeper :

```
/opt/mapr/zookeeper/zookeeper-3.4.5/bin/zkCli.sh -server
<hôte_zookeeper>:<port_zookeeper>
rmr /AnalyticServer
```

---

## Migration d'IBM SPSS Analytic Server dans MapR

IBM SPSS Analytic Server peut être migré dans MapR.

Utilisez les étapes suivantes pour migrer IBM SPSS Analytic Server 2.0 ou 2.1 vers la version 3.1.0 sur MapR.

1. Installez Analytic Server 3.1.0 sur un cluster MapR en suivant les instructions d'installation de la rubrique «Installation d'Analytic Server dans MapR», à la page 43.
2. Copiez la racine d'analyse.

**Remarque :** Vous pouvez ignorer cette étape si la racine d'analyse n'a pas été modifiée.

- Exécutez la commande suivante sur l'un des noeuds de données si la racine d'analyse d'Analytic Server version 2.0/2.1 et 3.1.0 se trouve sur le même cluster MapR :

```
hadoop fs -cp <ancienne_racine_analyse>/analytic-workspace/* <nouvelle_racine_analyse>/analytic-workspace
```

- Les services WEBHDFS et NFS installés déterminent quand la racine d'analyse d'Analytic Server version 2.0/2.1 et 3.1.0 se trouve sur des clusters MapR différents. Ils sont requis pour copier les données de racine d'analyse car le système de fichiers MapR n'est pas accessible directement hors du cluster.

- a. Exécutez la commande suivante sur l'un des nouveaux noeuds de cluster Analytic Server 3.1.0 lorsque l'ancien cluster Analytic Server 2.0/2.1 inclut le service WEBHDFS :

```
hadoop distcp
webhdfs://<serveur_webhdfs>:<port_webhdfs>/<ancienne_racine_analyse>/analytic-workspace/*
maprfs://<nouvelle_racine_analyse>/analytic-workspace
```

- b. Exécutez la commande suivante sur l'un des anciens noeuds de cluster Analytic Server 2.0/2.1 lorsque le nouveau cluster Analytic Server 3.1.0 inclut le service WEBHDFS :

```
hadoop distcp maprfs://<ancienne_racine_analyse>/analytic-workspace/*
webhdfs://<serveur_webhdfs>:<port_webhdfs>/<nouvelle_racine_analyse>/analytic-workspace
```

- c. Exécutez la commande suivante sur l'un des anciens noeuds de cluster Analytic Server 2.0/2.1 lorsque l'ancien cluster inclut le système NFS et que le système NFS est également monté sur l'un des nouveaux noeuds de cluster Analytic Server 3.1.0 :

```
hadoop distcp file:///<chemin_montage>/<ancienne_racine_analyse>/analytic-workspace/*
maprfs://<nouvelle_racine_analyse>/analytic-workspace
```

- d. Exécutez la commande suivante sur l'un des nouveaux noeuds de cluster Analytic Server 3.1.0 lorsque le nouveau cluster inclut le système NFS et que le système NFS est également monté sur l'un des anciens noeuds de cluster Analytic Server 2.0/2.1 :

```
hadoop discp maprfs://<ancienne_racine_analyse>/analytic-workspace/*
file:///<chemin_montage>/<nouvelle_racine_analyse>/analytic-workspace
```

Consultez le site de MapR Data Migration pour des informations sur la migration de données entre différents clusters MapR.

3. Exécutez les commandes suivantes pour changer le propriétaire de la nouvelle racine d'analyse et les droits :

```
hadoop fs -chown -R <utilisateur_as> <racine_analyse>
hadoop fs -chmod -R 755 <>
```

4. Arrêtez Analytic Server 3.1.0, mais assurez-vous que la base de données des métadonnées continue de s'exécuter.
5. Collectez les paramètres de configuration de l'ancienne installation de cluster Analytic Server 2.0/2.1.
  - a. Copiez l'archive configcollector.zip depuis la nouvelle installation de cluster Analytic Server 3.1.0 dans <ancien\_chemin\_installation\_as>/tools dans l'ancienne installation de cluster Analytic Server 2.0/2.1.

- b. Décompressez le contenu du fichier configcollector.zip sur l'ancienne installation de cluster Analytic Server 2.0/2.1. Un nouveau sous-répertoire configcollector est créé dans l'ancienne installation de cluster Analytic Server 2.0/2.1.
  - c. Lancez l'outil de collecte de configuration dans l'ancienne installation de cluster Analytic Server 2.0/2.1 en exécutant le script configcollector situé sous <ancien\_chemin\_d'installation\_as>/tools/configcollector. Copiez le fichier compressé (ZIP) généré dans la nouvelle installation de cluster Analytic Server 3.1.0.
6. Exécutez l'outil de migration dans le nouveau cluster Analytic Server 3.1.0 en exécutant le script migrationtool et en transmettant sous forme d'argument le chemin d'accès au fichier compressé qui a été créé par le collecteur de configuration. Exemple :
 

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.1/ASConfiguration_2.1.0.0.xxx.zip
```
  7. Démarrez Analytic Server 3.1.0.

---

## Traitement des incidents de MapR

Cette section décrit certains problèmes d'installation et de configuration de MapR courants et explique comment les résoudre.

### Problèmes liés au script hdfsUpdate.sh

Le script hdfsUpdate.sh ne doit être exécuté qu'une fois car il supprime tous les fichiers qui se trouvent à la racine d'analyse avant d'en télécharger de nouveaux. Si le script est exécuté plusieurs fois, vous devez recharger les dépendances pour la répercussion de base de données, Hive, HBase et Spark. Exécutez les commandes suivantes pour recharger les dépendances requises :

```
cd <chemin_installation_as>/jdbc

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

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

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

cd <chemin_installation_as>/ae_wlpserver/usr/servers/aeserver/modules/spark/

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

### Un conflit entre les versions de MapR et de Spark entraîne l'échec de l'exécution des travaux Spark

Un conflit entre des classes de MapR et de Spark (1.6.1) se produit lorsque la version de MapR est la version 5.1 ou suivante. Le conflit entraîne l'échec de l'exécution des travaux Spark. Vous pouvez résoudre le problème en modifiant le fichier private\_library.xml dans <chemin\_installation\_as>/ae\_wlpserver/usr/servers/aeserver. L'exemple suivant identifie la modification à effectuer :

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

---

# Chapitre 5. Installation et configuration de Huawei FusionInsight HD

---

## Présentation de FusionInsight HD

Huawei FusionInsight HD fournit une plateforme logicielle de mégadonnées complète pour l'analyse par lots et en temps réel à l'aide de technologies Hadoop et Spark open source. Le système optimise HDFS, HBase, MapReduce et YARN/Zookeeper pour la mise en cluster Hadoop ainsi qu'Apache Spark pour une analyse en temps réel plus rapide et des requêtes interactives.

Analytic Server peut s'exécuter sur la plateforme FusionInsight HD. FusionInsight HD contient les éléments de base principaux pour Hadoop, qui permettent une informatique répartie fiable et évolutive pour les ensembles de données volumineux (principalement MapReduce et HDFS), ainsi que d'autres composants orientés entreprise qui assurent la sécurité, la haute disponibilité et l'intégration au matériel et à d'autres logiciels.

---

## Installation dans Huawei FusionInsight HD

Les étapes ci-après décrivent le processus d'installation manuelle d'IBM SPSS Analytic Server dans Huawei FusionInsight HD.

### Analytic Server 3.1.0

1. Accédez au [site Web d'IBM Passport Advantage](#) et téléchargez le fichier binaire autoextractible ci-après sur un hôte dans le cluster FusionInsight HD.

Tableau 11. Fichier binaire autoextractible d'Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.1 pour FusionInsight HD 2.6 Linux x86-64 (Anglais)	spss_as-3.1-fhd2.6-1x86_en.bin

2. Exécutez le programme d'installation \*.bin sur le noeud de cluster maître FusionInsight Manager. Suivez les invites d'installation en acceptant le contrat de licence et en conservant le répertoire d'installation par défaut. Le programme d'installation télécharge les fichiers RPM nécessaires et doit être exécuté sur un ordinateur ayant accès à <https://ibm-open-platform.ibm.com>. Le fichier binaire exécutable se trouve dans le répertoire de distribution <AS\_INSTALLABLE\_HOME> de FusionInsight HD disponible.

3. Utilisez la commande suivante pour installer Analytic Server 3.1.0 :

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

4. Connectez-vous avec omm et créez le fichier 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
```

Le mot de passe par défaut de kadmin est Admin@123. Vous devez le changer au moment de sa première utilisation. Dans les commandes suivantes, remplacez \_HOTE par votre nom d'hôte.

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

5. Installez MYSQL et créez manuellement l'élément aedb. Exemple :

```
cd /etc/yum.repos.d
wget http://dev.mysql.com/get/mysql157-community-release-e17-9.noarch.rpm
yum -y install mysql157-community-release-e17-9.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
```

Procédez à l'extraction du mot de passe de l'utilisateur racine MYSQL :

```
grep 'temporary password' /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. Créez un script install\_as.sh dans /opt et exécutez-le avec l'utilisateur omm :

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

```
su omm
/opt/install_as.sh
```

```
install_as.sh
```

install\_as.sh contient le script suivant :

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

```
cd /opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/
echo "JAVA_HOME=/opt/ Huawei/Bigdata/jdk/jre" > server.env
echo "PATH=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:/opt/ Huawei/Bigdata/jdk/jre/lib/amd64:/usr/sbin:/usr/bin:/usr/local/bin:/usr/local/sbin" >> server.env
echo "IBM_SPSS_AS_NATIVE_PATH=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64" >> server.env
echo "LD_LIBRARY_PATH=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:/opt/ Huawei/Bigdata/jdk/jre/lib/amd64:/usr/local/lib" >> server.env
echo "SPARK_HOME=/opt/ibm/spss/analyticsserver/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\" rejectedWorkPolicy=\"CALLER_RUNS\" />" >> server.xml
echo " <webContainer deferServletLoad=\"false\" disallowAllFileServing=\"false\" fileServingEnabled=\"true\" trusted=\"false\" directoryBrowsingEnabled=\"false\" />" >> server.xml
echo " <classloading useJarUrls=\"true\" />" >> server.xml
echo " <applicationMonitor updateTrigger=\"mbean\" />" >> server.xml
echo " <mimeType>" >> server.xml
echo " <type>svg=image/svg+xml</type>" >> server.xml
echo " </mimeType>" >> 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\" />" >> http_endpoint.xml
echo "</server>" >> http_endpoint.xml
```

```
touch private_library.xml
```

```
echo "<server>" > private_library.xml
echo " <application context-root=\"/analyticsserver\" id=\"AS_BOOT\" location=\"/AE_BOOT.war\" name=\"/AS_BOOT\" type=\"war\">" >> 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\" />" >> 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"/>" >> private_library.xml
echo " <fileset dir="\${wlp.install.dir}/../jdbc" includes="*.jar"/>" >> 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="/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="om/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" >> 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=om/huawei-1@HADOOP.COM" >> config.properties
echo "hdfs.user=om/huawei-1@HADOOP.COM" >> config.properties
echo "web.authentication.kerberos.principal=HTTP/huawei-1@HADOOP.COM" >> config.properties
echo "java.security.krb5.conf=/home/om/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=FEFFUy9FQ0IvUEtDUzVQYWRkaW5nAMDJul7PVsvdIyLlzeS8ws=" >> config.properties
echo "encryption.keystore.base64=zs7OzgAAAAIAAABAAAAAwA6Y29tLnNwc3MuYWUuZW5jcmlwdG1vbi5wcm92aWR1ci51bmNyeXB0aW9ucHJvdmlkZXJpbXBsLmFlcwAAAUg2Ahyr00ABX" >> 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. Téléchargez FusionInsight SPARK et le client Zookeeper, procédez à l'extraction du contenu, et copiez les fichiers de configuration Spark dans le dossier /opt/ibm/spss/analyticserver/3.1/ae\_wlpserver/usr/servers/aeserver/configuration/hadoop-conf.
8. Copiez les fichiers Hadoop natifs dans le dossier suivant : /opt/ibm/spss/analyticserver/3.1/ae\_wlpserver/usr/servers/aeserver/configuration/linux/lib\_64
9. Ajoutez la ligne suivante au fichier /opt/ibm/spss/analyticserver/3.1/ae\_wlpserver/usr/servers/aeserver/jvm.option :
 

```
-Dconfig.folder.path=/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration
```
10. Exécutez le script hdfsUpdate.sh (/opt/ibm/spss/analyticserver/3.1/bin/hdfsUpdate.sh).

11. Exécutez le script `start.sh` (`/opt/ibm/spss/analyticserver/3.1/bin/start.sh`) pour démarrer le service Analytic Server ; exécutez le script `stop.sh` (`/opt/ibm/spss/analyticserver/3.1/bin/stop.sh`) pour arrêter le service Analytic Server.



---

## Chapitre 6. Configuration d'IBM SPSS Modeler pour son utilisation avec IBM SPSS Analytic Server

Pour activer SPSS Modeler pour son utilisation avec Analytic Server, vous devez effectuer certaines mises à jour dans l'installation de SPSS Modeler Server.

1. Configurez SPSS Modeler Server en l'associant à une installation Analytic Server.
  - a. Ouvrez le fichier `options.cfg` situé sous le sous-répertoire `config` du répertoire d'installation racine du serveur et ajoutez ou éditez les lignes suivantes :

```
as_ssl_enabled, {Y|N}
as_host, "{SERVEUR_AS}"
as_port, PORT
as_context_root, "{RACINE_CONTEXTE}"
as_tenant, "{TITULAIRE}"
as_prompt_for_password, {Y|N}
as_kerberos_auth_mode, {Y|N}
as_kerberos_krb5_conf, {CONF-PATH}
as_kerberos_krb5_spn, {AS-SPN}
```

### **as\_ssl\_enabled**

Spécifiez Y si la communication sécurisée est configurée sur Analytic Server ; sinon, spécifiez N.

### **as\_host**

Adresse IP du serveur sur lequel réside Analytic Server.

### **as\_port**

Port sur lequel Analytic Server est à l'écoute (par défaut, il s'agit du port 8080).

### **as\_context\_root**

Racine de contexte Analytic Server (par défaut, il s'agit de `analyticserver`).

### **as\_tenant**

Titulaire dont l'installation SPSS Modeler Server est membre (par défaut, il s'agit de `ibm`).

### **as\_prompt\_for\_password**

Spécifiez N si SPSS Modeler Server est configuré avec le même système d'authentification pour les utilisateurs et les mots de passe que celui utilisé sur Analytic Server ; par exemple, lors de l'utilisation de l'authentification Kerberos. Sinon, spécifiez Y.

Si vous exécutez SPSS Modeler en mode de traitement par lots, ajoutez les arguments `-analytic_server_username {nom_utilisateur_AS}` `-analytic_server_password {mot_de_passe_AS}` à la commande `clemb`.

### **as\_kerberos\_auth\_mode**

Entrez Y pour activer l'authentification unique Kerberos depuis SPSS Modeler.

### **as\_kerberos\_krb5\_conf**

Entrez le chemin du fichier de configuration Kerberos qui doit être utilisé par Analytic Server ; par exemple, `\etc\krb5.conf`.

### **as\_kerberos\_krb5\_spn**

Indiquez le SPN Kerberos pour Analytic Server ; par exemple, `HTTP/ashost.mydomain.com@MYDOMAIN.COM`.

- b. Redémarrez le service SPSS Modeler Server.

Pour pouvoir vous connecter à une installation Analytic Server sur laquelle SSL/TLS est activée, d'autres étapes de configuration de vos installations du serveur et du client SPSS Modeler sont requises.

- a. Accédez à `http{s}://{HOTE}:{PORT}/{RACINE_CONTEXTE}/admin/{TITULAIRE}` et connectez-vous à la console Analytic Server.
  - b. Téléchargez le fichier de certification depuis le navigateur et enregistrez-le sur votre système de fichiers.
  - c. Ajoutez le fichier de certification à l'environnement d'exécution Java (JRE) de votre installation SPSS Modeler Server et de votre installation SPSS Modeler Client. L'emplacement à mettre à jour est celui sous le sous-répertoire `/jre/lib/security/cacerts` du chemin d'installation SPSS Modeler.
    - 1) Vérifiez que le fichier `cacerts` n'est pas en lecture seule.
    - 2) Utilisez l'outil de clés Modeler livré avec le produit. Cet outil est situé dans le sous-répertoire `/jre/bin/keytool` du chemin d'installation SPSS Modeler.

Exécutez la commande suivante :

```
keytool -import -alias <alias_as> -file <fichier_cert> -keystore "<fichier_cacerts>"
```

Notez que `<alias_as>` est un alias pour le fichier `cacerts`. Vous pouvez utiliser n'importe quel nom dans la mesure où il est unique au fichier `cacerts`.

Un exemple de commande serait similaire à ceci :

```
keytool -import -alias MySSLCertAlias -file C:\Download\as.cer
-keystore "c:\Program Files\IBM\SPSS\Modeler\{VersionModeler}\jre\lib\security\cacerts"
```
  - d. Redémarrez SPSS Modeler Server et SPSS Modeler Client .
2. [facultatif] Installez IBM SPSS Modeler - Essentials for R si vous prévoyez d'évaluer les modèles R dans les flux avec sources de données Analytic Server. Vous pouvez télécharger IBM SPSS Modeler - Essentials for R depuis le site <https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>.

---

## Chapitre 7. Traitement des incidents

Cette section décrit certains problèmes d'installation et de configuration fréquents et explique comment les résoudre.

### Problèmes généraux

**L'installation aboutit tout en étant accompagnée d'avertissements, mais les utilisateurs ne peuvent pas créer des sources de données (renvoi d'une erreur "Impossible de faire aboutir la demande. Motif : permission refusée")**

La définition du paramètre **distrib.fs.root** sur un répertoire auquel l'utilisateur Analytic Server (par défaut, `as_user`) n'a pas accès entraîne des erreurs. Assurez-vous que l'utilisateur Analytic Server dispose d'un accès en lecture, écriture et exécution sur le répertoire **distrib.fs.root**.

**Les performances d'Analytic Server se dégradent progressivement.**

Si les performances d'Analytic Server sont insatisfaisantes, supprimez tous les fichiers `*.war` du chemin de déploiement du service Knox : `<chemin_service_knox>/data/ deployments`. Par exemple : `/usr/iop/4.1.0.0/knox/data/deployments`.

**Désinstallation d'Analytic Server ou d'Essentials for R on Ambari**

Dans certains cas, la procédure de désinstallation d'Analytic Server ou d'Essentials for R on Ambari se bloque. Dans ce cas, vous devez arrêter manuellement l'ID processus du serveur Ambari.

### Problèmes concernant des distributions Hadoop spécifiques

**L'action d'actualisation pour le service Analytic Server est désactivée sur Hortonworks 2.3**

Pour actualiser manuellement les bibliothèques Analytic Server sur Hortonworks 2.3, procédez comme suit :

1. Connectez-vous en tant qu'utilisateur Analytic Server (par défaut, `as_user`) à l'hôte exécutant le service Analytic Metastore.

**Remarque :** Vous pouvez identifier ce nom d'hôte depuis la console Ambari.

2. Exécutez le script **refresh** dans le répertoire `{RACINE_AS}/bin`. Par exemple :

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

3. Redémarrez le service Analytic Server dans la console Ambari.

**Les modules téléchargés depuis un site externe ne passent pas la vérification de hachage dans Cloudera Manager**

Une erreur de vérification de hachage est signalée dans la liste des fichiers parcel. Vous pouvez résoudre ce problème en attendant que la procédure de téléchargement s'achève, puis en redémarrant Cloudera via le service `cloudera-scm-server`. L'erreur ne survient plus après le redémarrage du service.

**Le compresseur LZO ne fonctionne pas avec BigInsights 4.2.X**

Vous devez exécuter `yum install lzo* hadoop-lzo*` pour installer les composants LZO dans BigInsights 4.2.X. En raison du changement du comportement de l'installation dans l'environnement BigInsights 4.2, les composants LZO sont désormais installés dans le répertoire `/usr/lib/hadoop-lzo/lib`. (Ce répertoire n'inclut pas l'environnement d'exécution Hadoop.)

Utilisez les étapes suivantes pour configurer manuellement les composants LZO dans BigInsights 4.2.X :

1. Connectez-vous à l'hôte qui exécute Analytic Server avec l'utilisateur Analytic Server (`as_user` par défaut).

**Remarque :** Vous pouvez utiliser la console Ambari pour déterminer le nom d'hôte.

2. Copiez le fichier JAR LZO Hadoop dans le répertoire des bibliothèques Analytic Server.

Exemple :

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

3. Copiez les bibliothèques natives LZO dans le répertoire des bibliothèques natives Analytic Server. Par exemple :

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

4. Exécutez le script **refresh** stocké dans le répertoire {AS\_ROOT}/bin. Exemple :

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

5. Redémarrez le service Analytic Server dans la console Ambari.

## Problèmes liés au référentiel de métadonnées

### L'opération CREATE USER échoue lors de l'exécution du script add\_mysql\_user

Avant d'exécuter le script **add\_mysql\_user**, vous devez supprimer manuellement l'utilisateur que vous tentez d'ajouter depuis la base de données mysql. Vous pouvez supprimer les utilisateurs via l'interface utilisateur de MySQL Workbench ou via des commandes MySQL. Exemple :

```
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'@'%';"
```

Dans les commandes ci-dessus, remplacez \$AEDB\_USERNAME\_VALUE par le nom d'utilisateur à supprimer et \$METASTORE\_HOST par le nom de l'hôte sur lequel est installée la base de données.

## Problèmes affectant les flux IBM SPSS Modeler exécutés dans un processus Spark

Les flux SPSS Modeler ne s'achèvent pas lorsqu'ils sont forcés de s'exécuter dans un processus Spark. Les flux SPSS Modeler en échec sont construits avec un noeud source Analytic Server (fichier HDFS) qui est lié à un noeud Sort, puis configuré pour une exportation vers une autre source de données Analytic Server. Après l'exécution du flux, l'interface utilisateur de Resource Manager indique que la nouvelle application est en opération, mais le flux ne s'achève jamais et demeure à l'état Running (en cours d'exécution). Aucun message n'indique pourquoi le flux ne s'est pas achevé dans les journaux Analytic Server, YARN ou Spark.

Le problème peut être résolu en ajoutant le paramètre spark.executor.memory au fichier analytics.cfg dans la configuration d'Analytic Server. L'attribution d'une valeur de mémoire de 4 Go permet aux flux SPSS Modeler auparavant en échec de s'achever en moins de 2 minutes (dans un environnement de cluster avec un seul noeud).

## Clusters à haute disponibilité

### Analytic Server ne peut pas être ajouté à d'autres hôtes en raison de modifications des dépendances

Exécutez le script update\_clientdeps en suivant les instructions figurant dans «Mise à jour des dépendances de client», à la page 21.

### java.net.SocketTimeoutException : Expiration du délai d'attente de lecture

Modifiez comme suit la variable d'environnement du délai d'attente de lecture de Liberty ND :

```
export LIBERTYND_READ_TIMEOUT=<millisecondes>
```

où <millisecondes> correspond au nombre de secondes à utiliser pour expiration du délai d'attente de lecture JMX.

**java.io.IOException: CWWKX7202E: La valeur du délai d'attente (60 secondes), de la commande ./server start a expiré**

Ajoutez les informations suivantes dans le fichier server.xml de Controller Server :

```
<!-- Augmentation du délai d'attente de démarrage/d'arrêt
du serveur pour tenir compte des matériels lents -->
<serverCommands startServerTimeout="120" stopServerTimeout="120"/>
```

**java.lang.OutOfMemoryError: espace de segment mémoire Java**

Ajoutez les lignes suivantes au fichier jvm.options sur chaque membre du cluster à haute disponibilité.

```
-Xms512M
-Xmx2048M
```

**"Le service de cluster d'analyse a perdu de manière inattendue le contact avec Zookeeper, cette machine virtuelle Java (JVM) est en cours d'arrêt afin de conserver l'intégrité du cluster."**

Il se peut que la quantité de données en cours d'écriture sur Zookeeper soit trop grande. Dans ce cas, les journaux Zookeeper contiennent des exceptions telles que :

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

ou les journaux d'Analytic Server contiennent des messages tels que :

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

1. Dans la console Ambari, accédez à l'onglet Configs du service Zookeeper et ajoutez la ligne suivante à env-template, puis redémarrez le service Zookeeper.

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

2. Dans la console Ambari, accédez à l'onglet Configs du service Analytic Server et ajoutez les informations suivantes dans Advanced analytics-jvm-options, puis redémarrez le service Analytic Cluster :

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

Le nombre à spécifier pour le paramètre jute.maxbuffer doit être supérieur au nombre indiqué dans les messages d'exception.

**Les données de transactions Zookeeper deviennent ingérables**

Attribuez au paramètre **autopurge.purgeInterval** dans zoo.cfg la valeur 1 pour permettre des purges automatiques du journal de transactions Zookeeper.

**Le service cluster d'analyse perd le contact avec Zookeeper**

Examinez et modifiez les paramètres **tickTime**, **initLimit** et **syncLimit** dans zoo.cfg. Par exemple :

```
Nombre de millisecondes de chaque graduation
tickTime=2000
Nombre de graduations que la phase de
synchronisation initiale peut accepter
initLimit=30
Nombre de graduations pouvant s'écouler entre l'envoi
d'une demande et son accusé de réception
syncLimit=15
```

Pour plus d'informations, reportez-vous à la documentation Zookeeper : <https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html>

**Les travaux Analytic Server ne reprennent pas**

Il existe une situation relativement courante dans laquelle les travaux Analytic Server ne reprennent pas.

- Lorsqu'un travail Analytic Server échoue en raison de l'échec d'un membre du cluster, le travail est normalement redémarré automatiquement sur un autre membre du cluster. Si le travail ne reprend pas, vérifiez que le cluster de haute disponibilité comprend au moins 4 membres.

**Les serveurs Analytic Server se bloquent parfois lors de leur arrêt**

Arrêtez manuellement le serveur.

---

## Remarques

Le présent document a été développé pour des produits et des services proposés aux Etats-Unis. Il peut être disponible dans d'autres langues auprès d'IBM. Toutefois, il peut être nécessaire de posséder une copie du produit ou de la version du produit dans cette langue pour pouvoir y accéder.

Le présent document peut contenir des informations ou des références concernant certains produits, logiciels ou services IBM non annoncés dans ce pays. Pour plus de détails, référez-vous aux documents d'annonce disponibles dans votre pays, ou adressez-vous à votre partenaire commercial IBM. Toute référence à un produit, logiciel ou service IBM n'implique pas que seul ce produit, logiciel ou service puisse être utilisé. Tout autre élément fonctionnellement équivalent peut être utilisé, s'il n'enfreint aucun droit d'IBM. Il est de la responsabilité de l'utilisateur d'évaluer et de vérifier lui-même les installations et applications réalisées avec des produits, logiciels ou services non expressément référencés par IBM.

IBM peut détenir des brevets ou des demandes de brevet couvrant les produits mentionnés dans le présent document. La remise de ce document ne vous donne aucun droit de licence sur ces brevets ou demandes de brevet. Si vous désirez recevoir des informations concernant l'acquisition de licences, veuillez en faire la demande par écrit à l'adresse suivante :

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
U.S.A.*

Pour le Canada, veuillez adresser votre courrier à :

*IBM Director of Commercial Relations  
IBM Canada Ltd.  
3600 Steeles Avenue East  
Markham, Ontario  
L3R 9Z7 Canada*

Les informations sur les licences concernant les produits utilisant un jeu de caractères double octet peuvent être obtenues par écrit à l'adresse suivante :

*Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan*

LE PRESENT DOCUMENT EST LIVRE EN L'ETAT SANS AUCUNE GARANTIE EXPLICITE OU IMPLICITE. IBM DECLINE NOTAMMENT TOUTE RESPONSABILITE RELATIVE A CES INFORMATIONS EN CAS DE CONTREFACON AINSI QU'EN CAS DE DEFAUT D'APTITUDE A L'EXECUTION D'UN TRAVAIL DONNE. Certaines juridictions n'autorisent pas l'exclusion des garanties tacites, auquel cas l'exclusion ci-dessus ne vous sera pas applicable.

Le présent document peut contenir des inexactitudes ou des coquilles. Ce document est mis à jour périodiquement. Chaque nouvelle édition inclut les mises à jour. IBM peut, à tout moment et sans préavis, modifier les produits et logiciels décrits dans ce document.

Les références à des sites Web non IBM sont fournies à titre d'information uniquement et n'impliquent en aucun cas une adhésion aux données qu'ils contiennent. Les documents sur ces sites web ne font pas partie des documents de ce produit IBM et l'utilisation de ces sites web se fait à vos propres risques.

IBM pourra utiliser ou diffuser, de toute manière qu'elle jugera appropriée et sans aucune obligation de sa part, tout ou partie des informations qui lui seront fournies.

Les licenciés souhaitant obtenir des informations permettant : (i) l'échange des données entre des logiciels créés de façon indépendante et d'autres logiciels (dont celui-ci), et (ii) l'utilisation mutuelle des données ainsi échangées, doivent adresser leur demande à :

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785  
U.S.A.*

Ces informations peuvent être soumises à des conditions particulières, prévoyant notamment le paiement d'une redevance.

Le logiciel sous licence décrit dans ce document et tous les éléments sous licence disponibles s'y rapportant sont fournis par IBM conformément aux dispositions de l'ICA, des Conditions internationales d'utilisation des logiciels IBM ou de tout autre accord équivalent.

Les données de performance et les exemples client ne sont présentés qu'à des fins d'illustration. Les performances réelles peuvent varier en fonction des configurations et des conditions d'exploitation.

Les informations concernant des produits non IBM ont été obtenues auprès des fournisseurs de ces produits, par l'intermédiaire d'annonces publiques ou via d'autres sources disponibles. IBM n'a pas testé ces produits et ne peut confirmer l'exactitude de leurs performances ni leur compatibilité. Elle ne peut recevoir aucune réclamation concernant des produits non IBM. Toute question concernant les performances de produits non IBM doit être adressée aux fournisseurs de ces produits.

Les instructions relatives aux intentions d'IBM pour ses opérations à venir sont susceptibles d'être modifiées ou annulées sans préavis, et doivent être considérées uniquement comme un objectif.

Tous les tarifs indiqués sont les prix de vente actuels suggérés par IBM et sont susceptibles d'être modifiés sans préavis. Les tarifs appliqués peuvent varier selon les revendeurs.

Ces informations sont fournies uniquement à titre de planification. Elles sont susceptibles d'être modifiées avant la mise à disposition des produits décrits.

Le présent document peut contenir des exemples de données et de rapports utilisés couramment dans l'environnement professionnel. Ces exemples mentionnent des noms fictifs de personnes, de sociétés, de marques ou de produits à des fins illustratives ou explicatives uniquement. Tous ces noms sont fictifs, et toute ressemblance avec des noms de personnes ou de sociétés réelles serait purement fortuite.

#### LICENCE DE COPYRIGHT :

Le présent document peut contenir des exemples de données et de rapports utilisés couramment dans l'environnement professionnel. Ces exemples mentionnent des noms fictifs de personnes, de sociétés, de marques ou de produits à des fins illustratives ou explicatives uniquement. Tous ces noms sont fictifs, et toute ressemblance avec des noms de personnes ou de sociétés réelles serait purement fortuite.

Toute copie totale ou partielle de ces programmes exemples et des oeuvres qui en sont dérivées doit comprendre une notice de copyright, libellée comme suit :

© (nom de votre société) (année). Des segments de code sont dérivés des exemples de programmes d'IBM Corp.



## Marques

IBM, le logo IBM et [ibm.com](http://ibm.com) sont des marques d'International Business Machines Corp. dans de nombreux pays. Les autres noms de produits et de services peuvent être des marques d'IBM ou appartenir à des tiers. La liste actualisée de toutes les marques d'IBM est disponible sur la page Web "Copyright and trademark information" à l'adresse [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Adobe, le logo Adobe, PostScript et le logo PostScript sont des marques d'Adobe Systems Incorporated aux Etats-Unis et/ou dans certains autres pays.

IT Infrastructure Library est une marque de The Central Computer and Telecommunications Agency qui fait désormais partie de The Office of Government Commerce.

Intel, le logo Intel, Intel Inside, le logo Intel Inside, Intel Centrino, le logo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium, et Pentium sont des marques d'Intel Corporation ou de ses filiales aux Etats-Unis et dans certains autres pays.

Linux est une marque de Linus Torvalds aux Etats-Unis et/ou dans certains autres pays.

Microsoft, Windows, Windows NT et le logo Windows sont des marques de Microsoft Corporation aux Etats-Unis et/ou dans certains autres pays.

ITIL est une marque de The Minister for the Cabinet Office et est enregistrée au bureau américain Patent and Trademark Office.

UNIX est une marque enregistrée de The Open Group aux Etats-Unis et/ou dans certains autres pays.

Cell Broadband Engine est une marque de Sony Computer Entertainment, Inc., aux Etats-Unis et/ou dans certains autres pays, et est utilisée sous license.

Linear Tape-Open, LTO, le logo LTO, Ultrium et le logo Ultrium sont des marques de HP, IBM Corp. et Quantum aux Etats-Unis et/ou dans certains autres pays.





