IBM SPSS Analytic Server
Version 3.2.2

Guide d'installation et de configuration

Important

Avant d'utiliser le présent document et le produit associé, prenez connaissance des informations générales figurant à la section «Remarques», à la page 79.

Certaines illustrations de ce manuel ne sont pas disponibles en français à la date d'édition.

Cette édition s'applique à la version 3.2.2 de IBM® SPSS Analytic Server, et à toutes les éditions et modifications ultérieures sauf mention contraire dans les nouvelles éditions.

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

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

Illustrations

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

Terminologie

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

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

Claviers

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

OS/2 et Windows - Paramètres canadiens

Au Canada, on utilise :

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

Nomenclature

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

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

Brevets

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

Assistance téléphonique

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

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**.

Trafic WebSocket
Vous devez vous assurer que le trafic WebSocket entre les clients et Analytic Server n'est pas bloqué par des pare-feux, des VPN ou d'autres méthodes de blocage de port. Le port WebSocket est le même que le port Analytic Server général.

SuSE Linux (SLES) 12
Procédez comme suit avant d'installer Analytic Server sur SuSE Linux 12 :

1. Téléchargez une clé publique sur votre hôte depuis l'URL suivante :https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
2. Importez la clé publique en exécutant la commande suivante sur votre hôte :

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

Ubuntu 18.04
Avant d'installer Analytic Server sur Ubuntu 18.04, procédez comme suit sur tous les noeuds de cluster :

1. Téléchargez une clé publique sur votre hôte depuis l'URL suivante : https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
2. Importez la clé publique en exécutant la commande suivante sur votre hôte :

```
apt-key add IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
```

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

Hortonworks Data Platform (HDP)
Avant d'installer Analytic Server, vous devez vérifier qu'au moins un client HDP a été déployé dans votre environnement en cluster. Le noeud qui héberge Ambari Manager recherche le répertoire `/usr/hdp`. Analytic Server ne fonctionnera donc pas en l'absence d'un client HDP.

Hive/HCatalog
Si vous comptez utiliser des sources de données NoSQL, configurez Hive et HCatalog pour accès distant. Vérifiez également que `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.

Si vous souhaitez utiliser Hive 2.1, vous devez activer le paramètre **Interactive Query** dans la console Ambari, puis entrer `2.x` dans la propriété `hive.version` lors de l'installation d'Analytic Server.

1. Ouvrez la console Ambari et ajoutez la propriété suivante dans la section **Analytic Server Advanced analytics.cfg**.
   - Clé : `hive.version`
   - Valeur : Entrez la version de Hive (par exemple 2.x)
2. Sauvegardez la configuration.

Remarque : Hive 2.1 est pris en charge sur HDP 2.6 ou version ultérieure avec Spark 2.x. Pour HDP 2.x, `hive.version` par défaut est 1.x ; pour HDP 3.x, `hive.version` par défaut est 3.x.

Référentiel de métadonnées

Par défaut, Analytic Server installe et utilise une base de données MySQL. Vous pouvez aussi 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, reportez-vous 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

Equilibreur 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.

Echec d'un travail Analytic Server

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.

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 vérifier que les clients suivants ont été installés en tant que services Ambari :

- HDFS/HDFS_CLIENT
- MAPREDUCE2/MAPREDUCE2_CLIENT
- HIVE/HIVE_CLIENT
- SPARK2/SPARK2_CLIENT (lorsque Spark 2.x est utilisé)
- HBASE/HBASE_CLIENT (lorsque HBASE est utilisé)
- YARN
- Zookeeper

Connexion SSH sans mot de passe

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

Outils precheck et postcheck d'installation - Ambari

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 utilisé après l'exécution du fichier binaire autoextractible 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 d'API REST pour traiter :

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

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

Emplacement des outils et prérequis

Avant d'installer le service Analytic Server, exécutez l'outil precheck sur tous les noeuds qui feront partie du service afin de vérifier que votre environnement Linux est prêt pour l'installation d'Analytic Server.

L'outil precheck est appelé automatiquement dans le cadre de l'installation. Il analyse le service Analytic Metastore et tous les noeuds Analytic Server avant de procéder à l'installation sur chaque hôte. Vous pouvez également l'appeler manuellement sur le noeud Ambari Server afin de vérifier la machine avant d'installer le service.

Après l'exécution du fichier binaire autoextractible d'Analytic Server, l'outil precheck est situé sous les répertoires suivants :

- **HDP**

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

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

Après l'installation d'Analytic Server, l'outil postcheck est situé sous le répertoire suivant :

- **HDP**

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

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

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 9, le répertoire `chktool` est disponible après l'installation du fichier RPM de métadonnées.

Exécution de l'outil precheck

Automatique

L'outil precheck peut être appelé automatiquement dans le cadre de l'installation d'Analytic Server lorsque Analytic Server est installé en tant que service via la console Ambari. Vous devez pour cela saisir manuellement le nom d'utilisateur et le mot de passe du serveur Ambari :

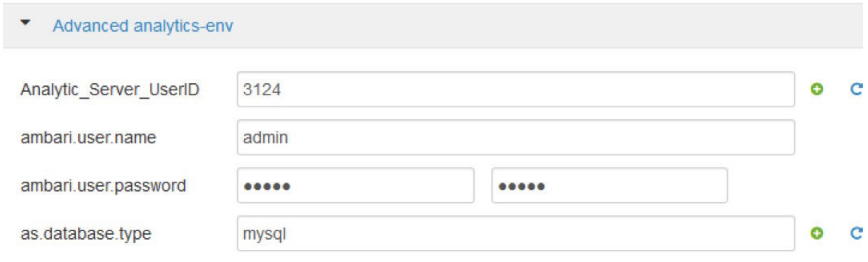


Figure 1. Paramètres Advanced analytics-env

Manuelle

Vous pouvez appeler manuellement l'outil precheck sur le noeud Ambari Server.

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 H --cluster MyCluster --username admin
--password admin --host myambarihost.ibm.com --port 8080 --ssl
```

Remarques :

- Les arguments `--target`, `--host`, `--port` et `--username` sont requis.
- La valeur `--host` doit correspondre à l'adresse IP ou à 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` inclut une aide pour son utilisation, laquelle est affichée en spécifiant l'argument `--h` (`python ./precheck.py --help`).
- L'argument `--cluster` est facultatif (le cluster actuel est utilisé si `--cluster` n'est pas spécifié).

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 --target H --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 --target H --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 --target H --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default --check Modèle_génération_PYSPARK_AS
```

Remarques :

- Les arguments `--target`, `--host`, `--port` et `--username` sont requis.
- La valeur `--host` doit correspondre à l'adresse IP ou à un nom de domaine complet.
- L'outil vous invite à saisir un mot de passe lorsque l'argument -password est omis.
- La commande `postcheck.py` inclut une aide pour son utilisation, laquelle est affichée en spécifiant 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.

«Installation en ligne», à la page 6
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.

«Installation hors ligne», à la page 9
Sélectionnez l'installation hors ligne si votre serveur Ambari n'a pas accès à Internet.

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.

1. Accédez au site Web IBM Passport Advantage® et téléchargez le fichier binaire autoextractible correspondant à votre pile, version de pile et architecture matérielle sur le noeud Ambari Manager. Les fichiers binaires Ambari disponibles sont :

   Tableau 1. Fichiers binaires autoextractibles Analytic Server

Description	**Nom du fichier binaire**
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0, and 3.1 Linux x86-64 (en anglais)	`spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0, and 3.1 Linux on System p LE (en anglais)	`spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin`

2. Exécutez le fichier binaire autoextractible et suivez les instructions pour afficher la licence, l'accepter, choisissez l'installation en ligne et sélectionnez la procédure d'installation correspondant au type de base de données utilisé par Analytic Server. Les options de type de base de données sont les suivantes :
   - Nouvelle instance MySQL
   - Instance MySQL ou Db2 préexistante
3. Depuis le répertoire `/var/lib/ambari-server/resources/stacks/<stack-name>/<stack-version>/services/ANALYTICSERVER/package/scripts`, exécutez le script `update_clientdeps.sh` avec les arguments appropriés (utilisez l'argument `--help` pour consulter des exemples).
4. Redémarrez votre serveur Ambari.

   ```
   ambari-server restart
   ```

5. 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 suivre les 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.

   Remarque : Si vous souhaitez que le programme d'installation Analytic Server crée une nouvelle instance MySQL, vous devez installer Analytic Server Metastore sur une machine sur laquelle MySQL n'est pas installé.

   Pour modifier le référentiel de métadonnées en désignant 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. Modifiez la valeur de **metadata.repository.url** en `jdbc:db2://{hôte_Db2}:{port}/{nom_BD}:currentSchema={nom_schéma};`, où :

- {hôte_Db2} est le nom d'hôte du serveur sur lequel Db2 est installé
- {port} est le port sur lequel Db2 est à l'écoute
- {nom_schéma} est un schéma disponible, non utilisé.

Si vous n'êtes pas sûr des valeurs à entrer, interrogez votre administrateur Db2.

h. Indiquez des données d'identification Db2 valides dans **metadata.repository.user.name** et **metadata.repository.password**.

i. Cliquez sur **Save**.

Configuration LDAP

Analytic Server utilise un serveur LDAP pour stocker et authentifier les utilisateurs et les groupes. Vous devez fournir les informations de configuration LDAP lors de l'installation d'Analytic Server.

Tableau 2. Paramètres de configuration LDAP

Paramètre LDAP	**Description**
`as.ldap.type`	Type LDAP. Valeurs admises : `ads`, `ad` ou `openladp`. • `ads` - Apache Directory Server (valeur par défaut) • `ad` - Microsoft Active Directory • `openladp` - OpenLDAP
`as.ldap.host`	Hôte LDAP
`as.ldap.port`	Numéro de port LDAP
`as.ldap.binddn`	Nom distinctif de liaison LDAP
`as.ldap.bindpassword`	Mot de passe du nom distinctif de liaison LDAP
`as.ldap.basedn`	Nom distinctif de base LDAP
`as.ldap.filter`	Règle de filtrage d'utilisateurs et de groupes LDAP
`as.ldap.ssl.enabled`	Indique si SSL doit être utilisé pour la communication entre Analytic Server et LDAP. Valeurs admises ; `true` ou `false`.
`as.ldap.ssl.reference`	ID de référence SSL LDAP
`as.ldap.ssl.content`	Configuration SSL LDAP

- Par défaut, `as.ldap.type` est défini sur `ads` et les autres paramètres associés reçoivent les valeurs par défauts. Une exception cependant : vous devez fournir un mot de passe pour le paramètre `as.ldap.bindpassword`. Analytic Server utilise les paramètres de configuration pour installer un serveur ADS (Apache Directory Server) et lancer son initialisation. Le profil ADS par défaut inclut l'utilisateur `admin` avec pour mot de passe `admin`. Vous pouvez gérer les utilisateurs via la Console Analytic Server ou importer les informations sur les utilisateurs et les groupes depuis un fichier XML en utilisant le script `importUser.sh` situé sous le dossier `<Analytic Root>/bin`.
- Si vous comptez utiliser un serveur LDAP externe, tel que Microsoft Active Directory ou OpenLDAP, vous devez définir les informations de configuration d'après les valeurs LDAP réelles. Pour plus d'informations, voir Configuration de registres utilisateur LDAP dans Liberty.
- Vous pouvez modifier la configuration LDAP après qu'Analytic Server a été installé (par exemple, passer d'Apache Directory Server à OpenLDAP). Toutefois, si votre installation initiale utilisait Microsoft Active Directory ou OpenLDAP et que vous décidez plus tard de passer à Apache Directory Server, Analytic Server n'installe pas un serveur Apache Directory Server lors de l'installation. Le serveur Apache Directory Server n'est installé que si vous le sélectionnez lors de l'installation initiale d'Analytic Server.

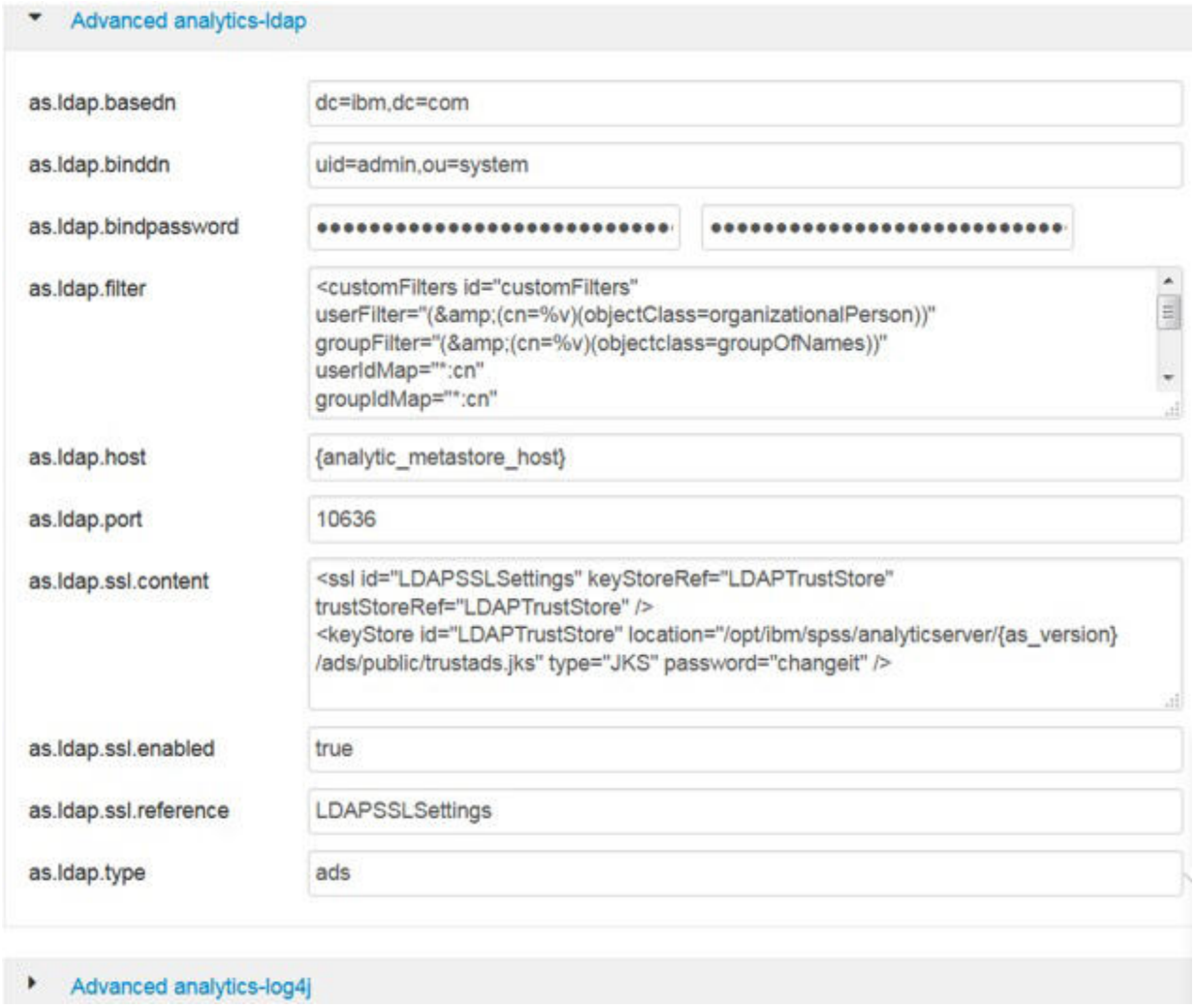


Figure 2. Exemple de paramètres de configuration LDAP

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

6. 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 16. Pour plus d'informations sur la migration d'une configuration existante vers une nouvelle installation, reportez-vous à la rubrique : «Migration d'IBM SPSS Analytic Server sur Ambari», à la page 35.
7. Ouvrez un navigateur Web et entrez l'adresse `http://<hôte>:<port>/analyticserver/admin/ibm`, où <hôte> désigne l'adresse de l'hôte Analytic Server et <port> 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

Une installation hors ligne d'IBM SPSS Analytic Server peut être effectuée automatiquement ou manuellement.

«Installation automatique sur HDP», à la page 9

La procédure d'installation automatique utilise l'API REST Ambari et constitue le mode d'installation recommandé

«Installation manuelle sur HDP (RHEL, SLES)», à la page 10
Pour installation manuelle d'Analytic Server sur Hortonworks Data Platform

«Installation manuelle sur HDP (Ubuntu)», à la page 13
Pour installation manuelle d'Analytic Server sur Ubuntu Linux.

Installation automatique sur HDP

La procédure d'installation automatique utilise l'API REST Ambari et constitue le mode d'installation recommandé.

Important :

- La procédure d'installation automatique hors ligne installe un serveur ADS (Apache Directory Server) intégré. Si vous désirez utiliser un serveur LDAP tiers, vous pouvez configurer vos paramètres LDAP une fois l'installation d'IBM SPSS Analytic Server terminée.
- La procédure d'installation automatique hors ligne ne peut installer qu'une seule instance de service Analytic Server. Vous pouvez en ajouter d'autres une fois l'installation initiale terminée.
- La procédure d'installation automatique hors ligne ne prend pas en charge l'installation d'Analytic Server sur un cluster activé pour Kerberos.

Ces limitations ne s'appliquent pas à des installations HDP ou Ubuntu manuelles.

1. Accédez au site Web IBM Passport Advantage® et téléchargez le fichier binaire autoextractible vers un ordinateur ayant accès à https://ibm-open-platform.ibm.com.

Tableau 3. Fichier binaire autoextractible d'Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0, and 3.1 Linux x86-64 (en anglais)	`spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0, and 3.1 Linux on System p LE (en anglais)	`spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin`

2. Lancez le fichier binaire exécutable téléchargé à l'étape 1 et spécifiez une installation hors ligne. L'installation hors ligne télécharge les fichiers REM ou DEB requis plus tard dans le processus d'installation et doit être exécutée sur un ordinateur pouvant accéder à https://ibm-open-platform.ibm.com. Les fichiers téléchargés sont situés sous le répertoire actuel des binaires exécutables, `./IBM-SPSS-AnalyticServer`.
3. Copiez le contenu complet du répertoire des binaires exécutables, `./IBM-SPSS-AnalyticServer`, depuis l'ordinateur avec accès Internet vers le noeud Ambari Manager (ce noeud est situé derrière le pare-feu).
4. Sur le noeud Ambari Manager, utilisez la commande suivante pour vérifier si le serveur Ambari est en opération :

   ```
   ambari-server status
   ```

5. Sur le noeud Ambari Manager, et tous les autres noeuds sur lesquels vous désirez déployer Analytic Server, installez l'outil qui crée un répertoire yum local.

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

   ou

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

6. Sur le noeud Ambari Manager, lancez le fichier binaire exécutable `./IBM-SPSS-AnalyticServer/packages/spss_as-ambari-offlineinstall.bin`. Lors de l'installation, le binaire exécutable vérifie que les fichiers RPM/DEB requis pour Analytic Server sont présents dans le répertoire des packages. Les fichiers RPM dont vous avez besoin varient en fonction de votre distribution, de votre version et de votre architecture.

   HDP 2.6, 3.0 et 3.1 (x86_64)
   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.2.0-1.x86_64.rpm

   HDP 2.6, 3.0 et 3.1 (PPC64LE)
   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.2.0-1.ppc64le.rpm

   HDP 2.6, 3.0 et 3.1 (Ubuntu)
   IBM-SPSS-AnalyticServer-ambari-2.x_3.2.2.0_amd64.deb

   IBM-SPSS-AnalyticServer_1_amd64.deb

Au cours de l'installation, vous êtes invité à entrer la version d'Analytic Server, le pilote JDBC, la version de Spark, celle de Hive, etc.

Installation manuelle sur HDP (RHEL, SLES)

Le flux général pour une installation hors ligne manuelle sur HDP (RHEL, SLES) est le suivant :

1. Accédez au site Web IBM Passport Advantage® et téléchargez le fichier binaire autoextractible vers un ordinateur ayant accès à https://ibm-open-platform.ibm.com.

Tableau 4. Fichiers binaires autoextractibles Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0, and 3.1 Linux x86-64 (en anglais)	`spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0, and 3.1 Linux on System p LE (en anglais)	`spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin`

2. Lancez le fichier binaire exécutable téléchargé à l'étape 1 et spécifiez une installation hors ligne. L'installation hors ligne télécharge les fichiers RPM requis plus tard dans le processus d'installation et doit être exécutée sur un ordinateur pouvant accéder à https://ibm-open-platform.ibm.com. Les fichiers téléchargés sont situés sous le répertoire actuel des binaires exécutables, `./IBM-SPSS-AnalyticServer`.
3. Copiez l'intégralité du contenu du répertoire des binaires exécutables `./IBM-SPSS-AnalyticServer`, depuis l'ordinateur avec accès Internet vers le répertoire <AS_INSTALLABLE_HOME> du noeud Ambari Manager (ce noeud est situé derrière le pare-feu).
4. Sur le noeud Ambari Manager, utilisez la commande suivante pour vérifier si le serveur Ambari est en opération :

```
ambari-server status
```

5. Installez l'outil qui crée un référentiel yum local.

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

ou

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

6. Créez un répertoire utilisé comme référentiel pour les fichiers RPM d'Analytic Server. Voir l'exemple ci-après.

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

7. 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.

HDP 2.6, 3.0 et 3.1 (x86_64)

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.x86_64.rpm

HDP 2.6, 3.0 et 3.1 (PPC64LE)

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.ppc64le.rpm

8. Créez la définition du référentiel local. Par exemple, créez un fichier nommé `IBM-SPSS-AnalyticServer-3.2.2.0.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
```

9. Créez le référentiel yum local.

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

10. Depuis une fenêtre de commande d'utilisateur root, exécutez la commande `cd` sur le répertoire <AS_INSTALLABLE_HOME>`/IBM-SPSS-AnalyticServer` et tapez `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 : L'étape 11 ne s'applique que si vous utilisez un environnement MySQL géré en externe.

11. 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 de <AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer dans le noeud/l'hôte dans lequel l'instance MySQL qui sera utilisée comme AS_MetaStore est installée.

   - 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 sur 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

12. 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:///{chemin vers le référentiel local}/</baseurl>
        <repoid>IBM-SPSS-AnalyticServer</repoid>
        <reponame>IBM-SPSS-AnalyticServer-3.2.2.0</reponame>
   </repo>
</os>
```

Voici un exemple de {chemin du référentiel local} :

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

13. Répétez la procédure suivante sur tous les noeuds de cluster Ambari qui ne sont pas un serveur.

a. Copiez l'intégralité du contenu du répertoire <AS_INSTALLABLE_HOME> approprié, à partir de l'ordinateur avec accès Internet vers 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 -p /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.

HDP 2.6, 3.0 et 3.1 (x86_64)
IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.x86_64.rpm

HDP 2.6, 3.0 et 3.1 (PPC64LE)

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.ppc64le.rpm

e. Créez la définition du référentiel local. Par exemple, créez un fichier nommé `IBM-SPSS-AnalyticServer-3.2.2.0.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)
```

14. Passez à l'étape 3 de la rubrique «Installation en ligne», à la page 6.

Installation manuelle sur HDP (Ubuntu)

Le flux général pour une installation hors ligne manuelle sur HDP (Ubuntu) est le suivant :

1. Accédez au site Web IBM Passport Advantage® et téléchargez le fichier binaire autoextractible Ubuntu approprié sur un ordinateur pouvant se connecter à https://ibm-open-platform.ibm.com.

Tableau 5. Fichiers binaires autoextractibles Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0, and 3.1 Linux x86-64 (en anglais)	`spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin`

2. Lancez le fichier binaire exécutable téléchargé à l'étape 1 et spécifiez une installation hors ligne. L'installation hors ligne télécharge les fichiers DEB requis plus tard dans le processus d'installation et doit être exécutée sur un ordinateur pouvant accéder à https://ibm-open-platform.ibm.com. Les fichiers téléchargés sont situés sous le répertoire actuel des binaires exécutables, `./IBM-SPSS-AnalyticServer`.
3. Copiez l'intégralité du contenu du répertoire des binaires exécutables `./IBM-SPSS-AnalyticServer`, depuis l'ordinateur avec accès Internet vers le répertoire <AS_INSTALLABLE_HOME> du noeud Ambari Manager (ce noeud est situé derrière le pare-feu).
4. Sur le noeud Ambari Manager, utilisez la commande suivante pour vérifier si le serveur Ambari est en opération :

```
ambari-server status
```

5. Créez un répertoire <`référentiel_local`> qui fera office de référentiel pour les fichiers DEB d'Analytic Server. Exemple :

```
mkdir –p /usr/local/mydebs
```

6. Copiez les fichiers DEB requis pour Analytic Server vers le répertoire <`référentiel_local`>.
   - `IBM-SPSS-AnalyticServer-ambari-2.x_3.2.2.0_amd64.deb`
   - `IBM-SPSS-AnalyticServer_1_amd64.deb`
7. Créez le référentiel local.
   a. Installez l'outil qui crée un référentiel local :

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

b. Générez le fichier du package source :

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

c. Créez le composant (principal) et l'architecture (par exemple, `binary-i386`, `binary-amd64`) de votre référentiel local :

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

d. Copiez le package source :

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

8. Créez la définition du référentiel local. Par exemple, créez un fichier nommé `IBM-SPSS-AnalyticServer-3.2.2.0.list` dans `/etc/apt/sources.list.d` avec le contenu suivant.

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

Important : Sous Ubuntu 18.04, utilisez : `deb [trusted=yes] file:/usr/local/mydebs ./`

9. Exécutez la commande suivante pour mettre à jour la liste des référentiels :

```
apt-get update
```

10. Exécutez la commande suivante pour installer IBM SPSS Analytic Server 3.2.2 :

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

Remarque : Pour vérifier que votre référentiel local est configuré correctement, n'exécutez pas la commande précédente dans votre répertoire `<référentiel_local>`. Si l'installation ne parvient pas à trouver le package, ceci signifie que votre répertoire local n'est pas configuré correctement (auquel cas, il vous faut vérifier toutes les étapes précédentes).

11. Répétez la procédure suivante sur tous les noeuds de cluster Ambari qui ne sont pas un serveur.

a. Créez un répertoire `<référentiel_local>` qui fera office de référentiel pour les fichiers DEB d'Analytic Server. Exemple :

```
mkdir –p /usr/local/mydebs
```

b. Copiez l'intégralité du contenu du répertoire `<référentiel_local>`, à partir de l'ordinateur du noeud Ambari Manager vers le répertoire `<référentiel_local>` du noeud du cluster non serveur Ambari. Le répertoire doit contenir les fichiers suivants :

- `<local_repo>/IBM-SPSS-AnalyticServer-ambari-2.x_3.2.2.0_amd64.deb`
- `<référentiel_local>/IBM-SPSS-AnalyticServer_1_amd64.deb`
- `<référentiel_local>/Packages.gz`
- `<référentiel_local>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/Packages`
- `<référentiel_local>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/Packages`

c. Créez la définition du référentiel local. Par exemple, créez un fichier nommé `IBM-SPSS-AnalyticServer-3.2.2.0.list` dans `/etc/apt/sources.list.d` avec le contenu suivant.

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

Important : Sous Ubuntu 18.04, utilisez : `deb [trusted=yes] file:/usr/local/mydebs ./`

12. Passez à l'étape 3 de la rubrique «Installation en ligne», à la page 6.

Installation d'Analytic Server dans 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 dans 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` de `<AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer` dans le noeud/l'hôte dans lequel l'instance MySQL qui sera utilisée comme `AS_MetaStore` est installée.
   - 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` sur 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.

   ```
   ambari-server restart
   ```
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`

Autorisation des agents Ambari non root

La procédure d'installation d'Analytic Server diffère d'une installation normale lorsque le serveur et l'agent Ambari s'exécutent en tant qu'utilisateur non root.

Prérequis

Connectez-vous en tant que `root`, ajoutez l'utilisateur non root sur chaque hôte de votre cluster et configurez l'utilisateur non root avec l'accès `sudo`. Dans l'exemple suivant, on ajoute l'utilisateur non root `asuser` au fichier `suoders` (l'emplacement par défaut du fichier est `/etc/sudoers`) :

```
## Allow root to run any commands anywhere
asuser ALL=(ALL)  ALL
```

```
## Allow root to run any commands anywhere without a password
asuser ALL=(ALL)  NOPASSWD: ALL
```

Pour obtenir des informations détaillées sur l'installation, reportez-vous aux sections «Installation en ligne», à la page 6 ou «Installation hors ligne», à la page 9.

Exigence pour sudo

Pour toutes les commandes exécutées en tant qu'utilisateur non root, vous devez ajouter `sudo` devant le texte de la commande.

Problèmes liés à la propriété non root d'Ambari

L'erreur suivante peut survenir après l'ajout d'Analytic Server en tant que service dans l'interface utilisateur Ambari :

```
Error: 500 status code received on POST method for API: /api/v1/stacks/HDP/versions/2.6/
recommendations
```

L'erreur provient de la propriété non root d'Ambari. Vous devez remplacer le propriétaire du dossier `/var/run/ambari-server` par l'utilisateur non root, puis ajouter Analytic Server en tant que service dans l'interface utilisateur Ambari. L'exemple suivant montre comment remplacer le propriétaire du dossier `/var/run/ambari-server` par l'utilisateur non root `asuser`.

```
sudo chown asuser:asuser /var/run/ambari-server/
```

SSH sans mot de passe

Lorsque le mode SSH sans mot de passe n'est pas configuré, l'avertissement suivant s'affiche au cours de l'installation :

```
UserWarning: Failure to add as_user. This must be done manually on each node.
warnings.warn("Failure to add as_user. This must be done manually on each node.")
```

Vous devez créer manuellement `as_user` sur chaque noeud. `as_user` est un compte utilisateur disposant des droits d'installation d'Analytic Server. Exemple :

```
# Create the as_user user (whatever id possible first) and note the id for use on susequent
nodes
sudo useradd as_user

# set the user for nologin
sudo usermod -s /sbin/nologin as_user

# Mod to as_user user id
sudo usermod -u {as_user_id} as_user

# Make primary group user_group
sudo usermod -g hadoop as_user

# Make extends group hdfs
sudo usermod -G hdfs as_user
```

Remarque : Vous pouvez trouver `{as_user_id}` sur le noeud maître Ambari par la commande `id as_user`.

Configuration

Après l'installation, vous devez créer les comptes requis sur le système d'exploitation de cluster.

1. Créez des comptes utilisateur du système d'exploitation pour tous les utilisateurs auxquels vous prévoyez d'accorder un accès à Analytic Server sur chaque Analytic Server et noeud Hadoop (ces utilisateurs sont également configurés en tant que registres utilisateur LDAP). Le groupe d'utilisateurs doit être défini comme 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 est conforme au paramètre YARN **Minimum user ID for submitting job**. Il s'agit du paramètre **min.user.id** défini dans le fichier `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.

2. Créez un dossier de base sur HDFS pour l'administrateur Analytic Server. Le droit d'accès au dossier doit être défini sur 755, le propriétaire sur `admin` et le groupe d'utilisateurs sur `hdfs`. Voir l'exemple **en gras** suivant :

```
[root@xxxxx configuration]# hadoop fs -ls /user
drwxr-xr-x   - admin    hdfs           0 2017-06-08 01:33 /user/admin
drwxr-xr-x   - user1    hdfs          0 2017-06-06 01:00 /user/user1
```

3. Créez des dossiers de base utilisateur sur HDFS pour tous les utilisateurs standard Analytic Server (par exemple, `user1`). Le propriétaire du dossier est l'utilisateur réel et le groupe d'utilisateurs doit être défini sur `hdfs`.

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.

- {AS_ROOT} désigne l'emplacement dans lequel Analytic Server est déployé ; par exemple, `/opt/ibm/spss/analyticserver/3.2`.
- {AS_SERVER_ROOT} désigne l'emplacement de la configuration, du journal et des fichiers serveur ; par exemple, `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver`.
- {AS_HOME} désigne 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

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

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="(&amp;(sAMAccountName=%v)(objectcategory=user))"
```

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

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

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

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

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) d'Analytic Server vers LDAP

Si vous sélectionnez l'option LDAP Apache Directory Server (ads) lors de l'installation d'Analytic Server et utilisez la configuration par défaut, Apache Directory Server est installé avec SSL configuré et activé (Analytic Server utilise alors automatiquement SSL pour communiquer avec Apache Directory Server).

Configurez SSL en suivant les instructions ci-dessous lorsque l'une des autres options LDAP est sélectionnée lors de l'installation d'Analytic Server (par exemple, utilisation d'un serveur LDAP externe).

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 -p /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}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. 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.
2. Créez les mêmes comptes que ceux de l'étape précédente sur le serveur LDAP.
3. 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. Le groupe d'utilisateurs doit être défini comme 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 est conforme au paramètre YARN **Minimum user ID for submitting job**. Il s'agit du paramètre **min.user.id** défini dans le fichier 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.
4. Créez un dossier de base sur HDFS pour l'administrateur Analytic Server. Le droit d'accès au dossier doit être défini sur 755, le propriétaire sur admin et le groupe d'utilisateurs sur hdfs. Voir l'exemple **en gras** suivant :

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

Found 9 items

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

5. Créez des dossiers de base utilisateur sur HDFS pour tous les utilisateurs standard Analytic Server (par exemple, user1). Le propriétaire du dossier est l'utilisateur réel et le groupe d'utilisateurs doit être défini sur hdfs.
6. [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 nom de principal du service Kerberos (HTTP/<nom_hôte_proxy>@<domaine>) et le fichier de clés de l'hôte HAProxy, où <nom_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>@<realm>
```

6. Enregistrez la configuration et redémarrez tous les services Analytic Server depuis la console Ambari.

Les utilisateurs peuvent maintenant se connecter à Analytic Server à l'aide de l'option **Single sign on log in** dans l'écran de connexion à IBM SPSS Analytic Server.

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` spécifiée dans la zone `Analytic_Server_User` de la configuration d'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).

3. Depuis un shell de commande, exécutez la commande suivante sur le noeud Analytic Server :

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

Activation de plusieurs domaines

Le paramètre **as.kdc.realms** est requis lorsque vous définissez plusieurs domaines. Les valeurs **as.kdc.realms** sont situées dans la section Analytic Server 'Advanced analytics.cfg' de la console Ambari.

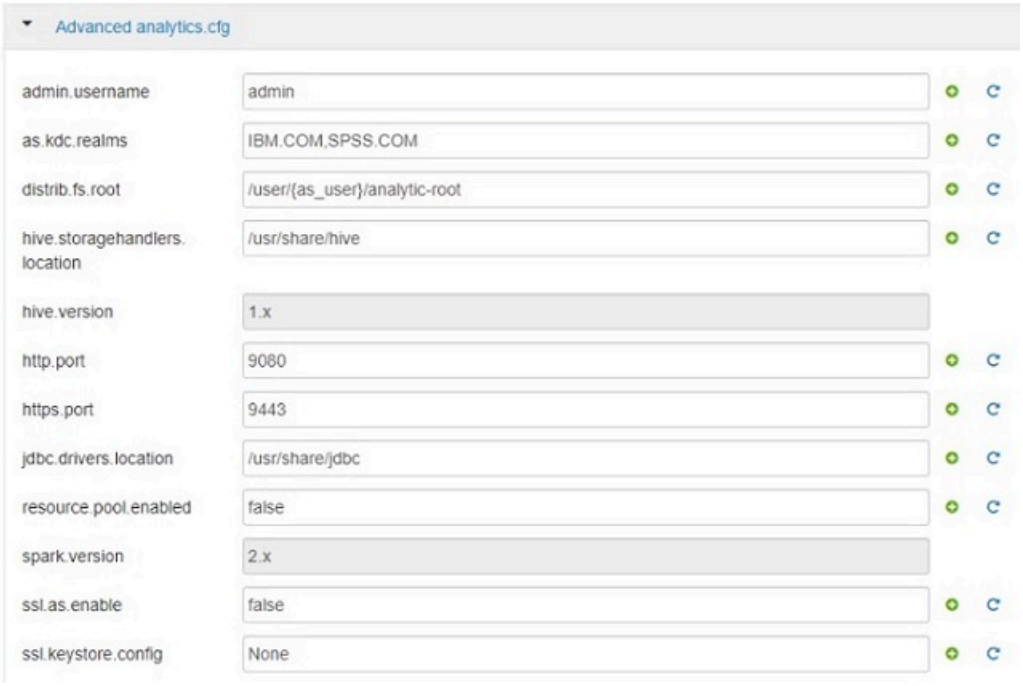


Figure 3. Paramètres Advanced analytics.cfg

Les noms de domaines multiples sont pris en charge lorsqu'ils sont séparés par des virgules. Les noms de domaines Kerberos correspondent et sont associés aux noms d'utilisateur. Par exemple, les noms d'utilisateur `UserOne@us.ibm.com` et `UserTwo@eu.ibm.com` correspondent respectivement aux domaines `us.ibm.com,eu.ibm.com`.

Des relations d'approbation inter-domaines Kerberos doivent être configurées lorsque plusieurs domaines sont spécifiés comme **Domaine Kerberos**. Le nom d'utilisateur saisi à l'invite de connexion à la console Analytic Server ne doit pas inclure le suffixe du nom de domaine. Par conséquent, lorsque plusieurs domaines sont spécifiés, les utilisateurs doivent sélectionner le domaine de leur choix dans la liste déroulante **Domaines**.

Remarque : Lorsqu'un seul domaine est spécifié, la liste déroulante **Domaines** n'apparaît pas lors de la connexion à Analytic Server.

Désactivation de Kerberos

1. Désactivez Kerberos dans la console Ambari.
2. Arrêtez le service Analytic Server.
3. 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.

Installation de certificats de fournisseur tiers

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 au format Base64. Pour le codage, vous pouvez utiliser l'outil securityUtility, par exemple `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility encode <mot_de_passe>`

   Si vous souhaitez générer un certificat autosigné, vous pouvez utiliser l'outil securityUtility ; par exemple, `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility createSSLCertificate --server=aeserver --password=mypassword --validity=365 --subject=CN=myfqdnserver,O=myorg,C=mycountry`.

   Remarques :

   - Vous devez fournir un nom de domaine hôte approprié pour la valeur `CN`.
   - Remplacez **`mypassword`**, **`myfqdnserver`**, **`myorg`** et **`mycounty`** par vos données d'identification. Notez que **`myfqdnserver`** est le nom de domaine complet du noeud Analytic Server.
   - **`aeserver`** est le nom du serveur Liberty (la valeur doit être **`aeserver`**).

   Pour plus d'informations sur l'utilitaire securityUtility et d'autres paramètres SSL, reportez-vous à la documentation WebSphere Liberty Profile et securityUtility command.
4. Cliquez sur **Save** et redémarrez le service Analytic Server.

Génération d'un certificat autosigné

Vous pouvez utiliser `securityUtility` pour générer des certificats autosignés. Exemple :

```
/opt/ibm/spss/analyticserver/3.2.2/ae_wlpserver/bin/securityUtility createSSLCertificate
--server=<myserver> --password=<mypassword> --validity=365 --subject=CN=<mycompany>,O=<myOrg>,C=<myCountry>
```

Remarques :

- Vous devez fournir un nom de domaine hôte approprié pour la valeur **`CN`**.
- Copiez les informations de `key.jks` dans `trust.jks` (les deux fichiers doivent être identiques).
- Editez le paramètre `ssl.keystore.config`. Exemple :

```
<ssl id="defaultSSLConfig"
keyStoreRef="defaultKeyStore"
```

```
trustStoreRef="defaultTrustStore"
clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
location="/opt/ibm/spss/analyticserver/3.2.2
/ae_wlpserver/usr/servers/aeserver/resources/security/key.jks"
type="JKS"
password="{xor}Dz4sLG5tbGs="/>
<keyStore id="defaultTrustStore"
location="/opt/ibm/spss/analyticserver/3.2.2
/ae_wlpserver/usr/servers/aeserver/resources/security/trust.jks"
type="JKS" password="{xor}Dz4sLG5tbGs="/>
```

Communication avec Apache Hive via SSL

Vous devez mettre à jour le fichier `hive.properties` afin de communiquer avec Apache Hive via une connexion SSL. Sinon, si votre environnement Apache Hive est activé pour haute disponibilité, vous pouvez aussi sélectionner les paramètres de haute disponibilité sur la page Analytic Server Sources de données principale.

Mise à jour du fichier hive.properties

1. Ouvrez le fichier `hive.properties`. Ce fichier est situé sous le répertoire : `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/database`
2. Localisez la ligne suivante :

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

3. Mettez à jour la ligne en lui ajoutant les informations **en gras** ci-dessous :

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password};
ssl=true;sslTrustStore=pathtotheirtruststorefile;trustStorePassword=xxxtheirTrustStorePassword
```

4. Enregistrez le fichier `hive.properties`.

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. Provisionnez l'environnement de serveur pour Essentials for R.

RedHat Linux x86_64

Exécutez les commandes suivantes :

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

Ubuntu Linux

Exécutez les commandes suivantes :

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

SUSE Linux

L'installation d'Essentials for R sur SUSE requiert un logiciel FORTRAN compatible qui n'est normalement pas disponible dans les répertoires ZYPPER configurés (mais est disponible sur le média du SDK SUSE). Par conséquent, l'exécution d'une installation d'Ambari pour Essentials for R sur un serveur SUSE échouera car elle ne pourra pas installer FORTRAN. Procédez comme suit pour le rendre disponible sur SUSE :

a. Installez GCC C++.

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

b. Installez GCC FORTRAN. Les fichiers RPM requis peuvent être copiés depuis le média SDK SUSE et doivent être installés dans l'ordre suivant.

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

c. Exécutez la commande suivante pour les bibliothèques Essentials for R.

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

2. 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.
3. 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.

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.

HDP 2.6 (x86_64)
IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.2.0.3-1.x86_64.rpm

HDP 3.0 et 3.1 (x86_64)
IBM-SPSS-ModelerEssentialsR-ambari-2.7-HDP-3.0-9.2.0.3-1.x86_64.rpm

HDP 2.6 (PPC64LE)
IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.2.0.3-1.ppc64le.rpm

HDP 3.0 et 3.1 (PPC64LE)
IBM-SPSS-ModelerEssentialsR-ambari-2.7-HDP-3.0-9.2.0.3-1.ppc64le.rpm

HDP 2.6, 3.0 et 3.1 (Ubuntu)
IBM-SPSS-ModelerEssentialsR-ambari-3.2.2.0_3.2.2.0_amd64.deb

b. Installez le fichier RPM ou DEB. Dans l'exemple ci-dessous, la commande installe Essentials for R sur HDP 2.6 (x86_64).

```
rpm -i IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.2.0.3-1.x86_64.rpm
```

Dans l'exemple ci-dessous, la commande installe Essentials for R sur HDP 2.6 (Ubuntu).

```
dpkg -i IBM-SPSS-ModelerEssentialsR-ambari-3.2.2.0_3.2.2.0_amd64.deb
```

4. Redémarrez votre serveur Ambari.

```
ambari-server restart
```

5. 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.

6. Actualisez le service Analytic Server.
7. Exécutez le script `update_clientdeps` en suivant les instructions figurant dans «Mise à jour des dépendances de client», à la page 30.
8. 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 dans le métamagasin Analytic Server et dans chaque noeud 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 6. 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 pour Linux, UNIX et Windows	11.1, 10.5, 10.1, 9.7	`db2jcc.jar`	IBM
Db2 z/OS	11, 10	`db2jcc.jar`, `db2_license_cisuz.jar`	IBM
Greenplum	5	postgresql.jar	Greenplum
Hive	2.1, 1.2	`hive-jdbc-*.jar`	Apache
MySQL	5.7, 5.6	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7, 6.x	`nzjdbc.jar`	IBM

Oracle	19c, 12c, 11g R2 (11.2)	19c :
		`ojdbc8.jar`, `orai18n.jar`
		12c et 11g R2 (11.2) :
		`ojdbc6.jar`, `orai18n.jar`
			Oracle

SQL Server	2017, 2014, 2012, 2008 R2	`sqljdbc4.jar`	Microsoft
Teradata	15.1, 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 prend en charge l'accès à cette base de données via JDBC (le fichier jar JDBC est le même que celui utilisé pour Db2).

  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. La prise en charge d'Apache HBase et d'Apache Accumulo ne demande aucune opération particulière. Pour les autres sources de données NoSQL, procurez-vous le gestionnaire d'espace de stockage et les fichiers JAR associés auprès du fournisseur de la base de données. Pour plus d'informations sur les sources de données HCatalog prises en charge, reportez-vous à la section Utilisation de sources de données HCatalog dans le manuel IBM SPSS Analytic Server 3.2.2 - Guide d'utilisation.
2. Ajoutez ces fichiers JAR au répertoire `{HIVE_HOME}/auxlib` et au répertoire `/usr/share/hive` sur le métamagasin Analytic Server et 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.

Remarques :

- Le méta-magasin Analytic Server Metastore ne peut pas être installé sur la même machine que le méta-magasin Hive Metastore.
- Lors de l'accès aux données HBase data via une source de données Analytic Server HCatalog, l'utilisateur doit disposer d'un droit de lecture sur les tables HBase.
  - Dans les environnements non Kerberos, Analytic Server accède à HBase en tant que `as_user` (`as_user` doit disposer d'un droit de lecture dans HBase).
  - Dans les environnements kerberos, `as_user` et l'utilisateur qui se connecte doivent tous deux disposer d'un droit de lecture sur les tables HBase.

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.

Travaux MapReduce v2

Utilisez le paramètre **`preferred.mapreduce`** dans la section Analytic Server **Custom analytic.cfg** pour indiquer comment traiter les travaux MapReduce :

Tableau 7. Propriétés analytics.cfg personnalisées	
Propriété	**Description**
`preferred.mapreduce`	Contrôle la méthode dans laquelle exécuter les travaux MapReduce. Les valeurs valides sont les suivantes : • `spark` • `m3r` • `hadoop` Par exemple : `preferred.mapreduce=spark`

Apache Spark

Si vous désirez utiliser Spark (version 2.x ou ultérieure), vous devez ajouter manuellement la propriété `spark.version` pendant l'installation d'Analytic Server.

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, `2.x` ou `None`).
2. Sauvegardez la configuration.

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`

Environnements Kerberos HDP 3.0 (ou version ultérieure)

Les environnements Kerberos HDP 3.0 (ou version ultérieure) peuvent avoir besoin de paramètres de sécurité supplémentaires. Dans HDFS, les FACL du système de fichiers sont utilisées dans le répertoire `/warehouse/tablespace/managed/hive`. Vous pouvez déterminer que des FACL ont besoin d'être configurées dans le service Hive Metastore lorsque les exceptions ci-dessous se produisent dans les fichiers `messages.log` ou `as_trace.log` :

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

L'exemple suivant présente une commande **`setfacl`** qui offre un large accès (à l'ensemble des membres du groupe `hadoop` dans notre cas) au répertoire Hive `warehouse` :

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

Sinon, des variations plus restrictives doivent être utilisées lorsqu'un contrôle d'accès plus élevé est requis.

Les sites suivants contiennent des informations de référence complémentaires.

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

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

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**.

Prise en charge de plusieurs clusters

La fonctionnalité de prise en charge de clusters multiples est un approfondissement des capacités de haute disponibilité d'IBM SPSS Analytic Server et fournit un isolement amélioré dans les environnements à plusieurs titulaires. Par défaut, l'installation du service Analytic Server (dans Ambari ou ClouderaManager) entraîne la définition d'un serveur analytique unique.

La spécification du cluster définit l'appartenance au cluster Analytic Server. La modification de la spécification du cluster est réalisée via un contenu XML (dans la zone `analytics-cluster` de la configuration d'Ambari Analytic Server ou en modifiant manuellement le fichier `configuration/analytics-cluster.xml` de Cloudera Manager). Si vous configurez plusieurs clusters Analytic Server, vous devez acheminer les demandes à chaque cluster Analytic Server avec son propre équilibreur de charge.

L'utilisation du dispositif de clusters multiples assure que le travail concernant un titulaire n'affecte pas négativement celui en cours d'exécution dans le cluster d'un autre titulaire. Quant à la haute disponibilité des travaux, le basculement des travaux se produit uniquement au sein du cluster Analytic Server où le travail a été déclenché. L'exemple suivant décrit une spécification XML de clusters multiples.

Remarque : La haute disponibilité d'Analytic Server peut être réalisée en l'ajoutant en tant que service dans plusieurs noeuds dans votre cluster.

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

Dans l'exemple précédent, deux équilibreurs de charge sont requis. L'un envoie des demandes aux membres de `cluster1` (`one.cluster` et `two.cluster`), l'autre aux membres de `cluster2` (`three.cluster` et `four.cluster`).

L'exemple suivant décrit la spécification XML d'un cluster unique (configuration par défaut).

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

Dans l'exemple suivant, un seul équilibreur de charge est requis pour gérer les cas où plusieurs membres du cluster ont été configurés.

Remarques

- Seules les instances de cluster unique prennent en charge l'utilisation de caractères génériques dans l'élément **memberName** (par exemple, cluster cardinality = "1"). Les valeurs valides pour l'élément cardinality sont 1 et 1+.
- **memberName** doit être spécifié de la même manière que le nom d'hôte auquel le rôle Analytic Server est affecté.
- Tous les serveurs de chaque cluster doivent être redémarrés après une modification de la configuration de cluster.
- Dans Cloudera Manager, vous devez modifier et gérer le fichier `analytics-cluster.xml` sur tous les noeuds d'Analytic Server. Tous les noeuds doivent être contrôlés pour vérifier que leur contenu reste identique.

Optimisation des options JVM pour le 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 niveau de Python - HDP

Cette section décrit la mise à niveau manuelle de Python 2.x à Python 3.7

1. Installez Python 3.7 sur chaque noeud du cluster. Pour plus d'informations, reportez-vous au site Python.
2. Installez NumPy sur chaque noeud du cluster. Pour plus d'informations, reportez-vous aux instructions d'installation NumPy.
3. Installez pandas sur chaque noeud du cluster. Pour plus d'informations, reportez-vous aux instructions d'installation pandas.
4. Ajoutez `spark.driver.python=<python3.7 executable path>` à la section **Custom analytics.cfg** de la configuration Ambari. Exemple :

   ```
   spark.driver.python=/opt/python3/bin/python3.7
   ```

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/2.6/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

- Un problème Apache Knox connu est qu'il ne propage par les informations de sécurité dans les cookies et les en-têtes HTTP (pour plus d'informations, voir https://issues.apache.org/jira/browse/KNOX-895). Ce problème est résolu dans Knox 0.14.0 (ou version ultérieure). Vous devez vous procurer une distribution Hortonworks mise à jour et qui inclut Knox 0.14.0 (ou version ultérieure), pour que Knox puisse opérer avec Analytic Server. Contactez votre fournisseur Hortonworks pour plus d'informations.
- 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_0.6-3.2.2.0.jar` : ce fichier doit être copié depuis l'emplacement Analytic Server suivant :

  `<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`

Remarque : Il existe deux jeux de fichiers `rewrite.xml` et `service.xml` (l'un pour le trafic `http://rest` et l'autre pour le trafic `ws://websocket`). Copiez tous les fichiers `rewrite.xml` et `service.xml` pour `analyticserver` tout comme pour `analyticserver_ws` vers le noeud du serveur Knox.

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 les deux services suivants dans la section **Advanced topology** de la configuration Knox :

```
<service>
    <role>ANALYTICSERVER</role>
    <url>http://{analyticserver-host}:{analyticserver-port}/analyticserver</url>
 </service>
<service>
    <role>ANALYTICSERVER_WS</role>
    <url>ws://{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 mis 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. Actualisez le fichier Knox `gateway.jks` :
   a. Sur le serveur Knox, arrêtez le service Knox.
   b. Supprimez le fichier `gateway.jks` du dossier `/var/lib/knox/data-2.6.2.0-205/security/keystores`.
   c. Redémarrez le service Knox.
2. Sur le serveur Knox, créez le sous-répertoire `<knox_server>/data/service/analyticserver/3.2.2.0`, puis téléchargez les fichiers `service.xml` et `rewrite.xml` dans le nouveau répertoire. Les deux fichiers sont situés sur Analytic Server sous `<analytic_server>/configuration/knox/analyticserver/` (par exemple, `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/*.xml`)
3. Depuis le répertoire `<serveur_knox>/bin`, exécutez le script `./knoxcli.sh redeploy --cluster default`
4. Téléchargez le fichier `com.ibm.spss.knoxservice_0.6-*.jar` dans `<knox_server>/ext`. Le fichier se trouve dans Analytic Server sous `<analytic_server>/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.2.0.jar` (par exemple,`/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.2.0.jar`).
5. Depuis l'interface utilisateur Ambari, ajoutez l'élément suivant dans **Knox** > **Configs** > **Advanced topology**:

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

   Remarque : La fonctionnalité WebSocket est désactivée par défaut. Vous pouvez l'activer en changeant en `true` la valeur de la propriété `gateway.websocket.feature.enabled` dans le fichier `/conf/gateway-site.xml`.
6. Depuis l'interface utilisateur Ambari, ajoutez ou mettez à jour les utilisateurs dans **Knox** > **Configs** > **Advanced users-ldif** (par exemple, `admin`, `qauser1`, `qauser2`).
7. Redémarrez LDAP depuis **Knox** > **Service Actions** > **Start Demo LDAP**.
8. Redémarrez le service Knox.

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`.

Configuration d'une allocation de ressource dynamique distincte pour chaque file d'attente YARN - HDP

Vous pouvez configurer une allocation de ressource dynamique distincte pour chaque file d'attente YARN.

Mappage des modes utilisateur et titulaire - Hortonworks Data Platform

Les tâches utilisateur et titulaire peuvent être soumises à différentes files d'attente YARN, et chaque utilisateur ou titulaire est mappé à une file d'attente YARN différente (pour tirer parti de l'allocation de ressource dynamique). Le mode **`utilisateur`** ou le mode **`titulaire`** peut être défini pour le mappage

aux files d'attente YARN. Avant Analytic Server 3.2.1 groupe de correctifs 1, tous les travaux Spark étaient limités à une seule file d'attente YARN.

A partir de IBM SPSS Analytic Server 3.2.1, groupe de correctifs 1, lorsque le flux d'un utilisateur/titulaire entraîne l'exécution de travaux Spark sur le système, une file d'attente YARN distincte s'exécute en tant qu'utilisateur/titulaire ayant soumis le flux à Analytic Server. Plusieurs files d'attente YARN peuvent s'exécuter simultanément pour différentes tâches utilisateur/titulaire.

Chaque file d'attente YARN continue à s'exécuter tant que l'utilisateur est connecté à Analytic Server (et pendant un certain temps après la déconnexion de l'utilisateur et l'absence de travaux utilisateur actifs). La durée après déconnexion peut être contrôlée par la variable de configuration : **as.spark.driver.cleanup.delay**.

Un processus **SparkDriver** est créé pour chaque utilisateur qui soumet le travail Spark. Le processus **SparkDriver** de chaque utilisateur se termine lorsque l'utilisateur n'a plus de travaux actifs pendant environ 2 minutes (la valeur par défaut) et aucune activité **HTTPSession**.

Remarque : Tous les processus **SparkDriver** se terminent lorsque Analytic Server s'arrête.

Pour ajouter Analytic Server à un cluster existant, procédez comme suit :

1. Dans l'interface utilisateur Ambari, accédez à l'onglet **SPSS Analytic Server service** > **Configs** > **Advanced analytics.cfg**.
2. Remplacez la valeur de **resource.pool.enabled** par `true`.
3. Ajoutez les propriétés suivantes dans l'onglet **Custom analytics.cfg** :

```
//Using user/tenant mapping to YARN pool
yarn.queue.mode=<user/tenant>
yarn.queue.mapping=<user1:test,user2:production>/<tenant1:test,tenant2:production>
yarn.queue.default=default
as.spark.driver.cleanup.delay=2
as.sparkdriver.max.memory=1G
```

Tableau 8. Propriétés analytics.cfg personnalisées

Propriété	Description
`yarn.queue.mode`	Définit le mode de mappage pour les files d'attente YARN. Lorsque `yarn.queue.mode=user`, une file d'attente YARN distincte est exécutée pour chaque utilisateur qui a soumis un travail/flux à Analytic Server. Plusieurs files d'attente YARN peuvent s'exécuter simultanément pour les différents travaux/flux des utilisateurs. Lorsque `yarn.queue.mode=tenant`, une file d'attente YARN distincte est exécutée pour chaque titulaire qui a soumis un travail/flux à Analytic Server. Plusieurs files d'attente YARN peuvent s'exécuter simultanément pour les différents travaux/flux des titulaires.
`yarn.queue.mapping`	Mappe les paires utilisateur ou titulaire vers les files d'attente YARN définies dans le gestionnaire de files d'attente YARN. Les paires doivent être séparées par des virgules (par exemple, `tenant1:test,tenant2:production` pour les titulaires ou `user1:test,user2:production` pour les utilisateurs).
`yarn.queue.default`	Nom de la file d'attente YARN par défaut à laquelle l'application est soumise. Vous pouvez spécifier une file d'attente YARN personnalisée depuis le gestionnaire de file d'attente YARN.
`as.spark.driver.cleanup.delay`	Entier représentant le nombre de minutes à s'écouler entre la déconnexion et l'arrêt de la file d'attente YARN d'un utilisateur. La valeur par défaut est **2**. Cette propriété est facultative.
`as.sparkdriver.max.memory`	Définit le volume de mémoire utilisé par chaque processus **SparkDriver**. La valeur par défaut est **1G**. Cette propriété est facultative.

4. Enregistrez la configuration et redémarrez le service Analytic Server.

Références

Reportez-vous aux sites suivants pour plus d'informations :

- https://spark.apache.org/docs/latest/running-on-yarn.html

Migration d'IBM SPSS Analytic Server sur Ambari

Analytic Server peut migrer des données et des paramètres de configuration depuis une installation Analytic Server existante vers une nouvelle installation. La migration peut être effectuée sur le même environnement de cluster ou sur un nouveau.

Migration de Analytic Server 3.2.1.1 vers la version 3.2.2 sur le même cluster de serveurs

Si vous disposez d'une installation Analytic Server 3.2.1.1 existante, vous pouvez migrer ses paramètres de configuration 3.2.1.1 vers votre installation 3.2.2 sur le même cluster de serveurs.

1. Collectez les paramètres de configuration de votre ancienne version d'Analytic Server (Analytic Server 3.2.1.1).
   a. Décompressez l'archive `{AS_ROOT}\tools\unzip configcollector.zip` (ceci crée un nouveau dossier nommé `configcollector`).
   b. Exécutez le script `configcollector.sh` situé sous le dossier `configcollector`. Copiez le fichier compressé (au format ZIP) ASConfiguration_3.2.1.1.xxx.zip dans un dossier différent (en tant que sauvegarde).
2. Sauvegardez la racine d'analyse de l'ancienne version d'Analytic Server 3.2.1.1 dans un nouvel emplacement.
   a. Si vous ne connaissez pas l'emplacement exact de la racine d'analyse, exécutez la commande **`hadoop fs -ls`**. Le chemin de la racine d'analyse est similaire à `/user/utilisateur_AS/analytic-root/analytic-workspace`, où `utilisateur_AS` correspond à l'ID de l'utilisateur propriétaire de la racine d'analyse.
   b. Utilisez les commandes **`hadoop fs -copyToLocal`** et **`hadoop fs -copyFromLocal`** pour copier le dossier `analytic-workspace` de l'ancienne version d'Analytic Server dans le nouvel emplacement (par exemple, `/user/as_user/analytic-root/AS3211Location`).
3. Si vous utilisez le serveur Apache Directory Server intégré, sauvegardez la configuration utilisateurs/groupes actuelle à l'aide d'un outil client LDAP tiers. Après l'installation d'Analytic Server 3.2.2, importez dans le serveur Apache Directory Server la configuration utilisateurs/groupes que vous aviez sauvegardée.

   Remarque : Vous pouvez ignorer cette étape si vous utilisez un serveur LDAP externe.
4. Ouvrez la console Ambari et arrêtez le **service Analytic Server**.
5. Désinstallez l'ancienne version d'Analytic Server (Analytic Server 3.2.1.1), puis installez Analytic Server 3.2.2. Pour les instructions d'installation, voir Chapitre 2, «Installation et configuration d'Ambari», à la page 3.
6. Ouvrez la console Ambari et arrêtez le **service Analytic Server** (dans Ambari, vérifiez que le **service Analytic Metastore** est en opération).
7. Copiez la racine d'analyse Analytic Server 3.2.1.1 sauvegardée à l'étape 2 dans l'emplacement de la nouvelle version d'Analytic Server.
   a. Supprimez le dossier `analytic-workspace` de la nouvelle version Analytic Server installée.
   b. Copiez le dossier d'espace de travail d'Analytic Server 3.2.1.1 sauvegardé (`/user/as_user/analytic-root/AS3211Location`) dans l'emplacement de la nouvelle version (par exemple, `/user/as_user/analytic-root/analytic-workspace`). Vous devez vérifier que `as_user` est bien défini comme propriétaire de l'espace de travail d'analyse.

8. Effacez l'état de Zookeeper. Depuis le répertoire bin de Zookeeper (par exemple, `/usr/hdp/current/zookeeper-client` sur Hortonworks), exécutez la commande suivante :

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

9. Copiez l'archive de sauvegarde ASConfiguration_3.2.1.1.xxx.zip de l'étape 1 dans l'emplacement de la nouvelle version d'Analytic Server (par exemple, `/opt/ibm/spss/analyticserver/3.2/`).
10. Lancez l'outil de migration en exécutant le script **migrationtool.sh** et en transmettant sous forme d'argument le chemin du fichier archive ASConfiguration_3.2.1.1.xxx.zip (créé par l'outil de collecte de configuration). Exemple :

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.2.1.1.xxx.zip
```

11. Depuis un shell de commande, exécutez la commande suivante sur le noeud Analytic Server :

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

12. Dans la console Ambari, démarrez le **service Analytic Server**.

Migration de Analytic Server 3.2.1.1 vers la version 3.2.2 sur un nouveau cluster de serveurs

Si vous disposez d'une installation Analytic Server 3.2.1.1 existante, vous pouvez migrer ses paramètres de configuration 3.2.1.1 vers votre installation 3.2.2 sur un nouveau cluster de serveurs.

1. Installez la nouvelle version d'Analytic Server d'après les instructions figurant dans «Installation dans Ambari», à la page 5.
2. Copiez l'espace de travail d'analyse de votre ancienne installation vers la nouvelle.
   a. Si vous ne connaissez pas l'emplacement exact de l'espace d'analyse, exécutez la commande `hadoop fs -ls`. Le chemin de l'espace d'analyse est similaire à `/user/as_user/analytic-root/analytic-workspace`, où `utilisateur_AS` est l'ID utilisateur qui en est le propriétaire.
   b. Supprimez le dossier `analytic-workspace` sur le nouveau serveur.
   c. Utilisez les commandes `hadoop fs -copyToLocal` et `hadoop fs -copyFromLocal` pour copier l'espace de travail d'analyse de l'ancien serveur vers le dossier `/user/as_user/analytic-root/analytic-workspace` du nouveau serveur (vérifiez que le propriétaire est défini comme `as_user`).
3. Si vous utilisez le serveur Apache Directory Server intégré, sauvegardez la configuration utilisateurs/groupes actuelle à l'aide d'un outil client LDAP tiers. Après l'installation d'Analytic Server 3.2.2, importez dans le serveur Apache Directory Server la configuration utilisateurs/groupes que vous aviez sauvegardée.

   Remarque : Vous pouvez ignorer cette étape si vous utilisez un serveur LDAP externe.
4. Sur le nouveau serveur, ouvrez la console Ambari et arrêtez le service Analytic Server (sur Ambari, vérifiez que le service Analytic Metastore est en opération).
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`, ce qui créera un nouveau sous-répertoire `configcollector` dans votre ancienne installation.
   c. Exécutez l'outil de collecte de configuration dans votre ancienne installation en exécutant le script **configcollector** dans `{AS_ROOT}\tools\configcollector`. Copiez le fichier compressé (ZIP) résultant sur le serveur qui héberge la nouvelle installation.

   Important : Il se peut que le script **configcollector** fourni ne soit pas compatible avec la version Analytic Server la plus récente. Contactez votre interlocuteur du support technique IBM si vous rencontrez des problèmes avec le script **configcollector**.

6. Effacez l'état de Zookeeper. Depuis le répertoire bin de Zookeeper (par exemple, `/usr/hdp/current/zookeeper-client` sur Hortonworks), exécutez la commande suivante.

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

7. Lancez le script 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/3.2/ASConfiguration_3.2.1.1.xxx.zip
```

8. Depuis un shell de commande, exécutez la commande suivante sur le noeud Analytic Server :

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

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

   Remarque : Si vous aviez configuré R pour son utilisation avec l'installation Analytic Server existante, suivez les étapes destinées à le configurer pour son utilisation avec la nouvelle installation d'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 37.

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. Consultez la documentation Db2 pour les instructions de 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 HDP 2.6, le répertoire ESSENTIALR est situé sous `/var/lib/ambari-server/resources/stacks/HDP/2.6/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 disponible 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
- Spark 2 : Gateway

Référentiel de métadonnées

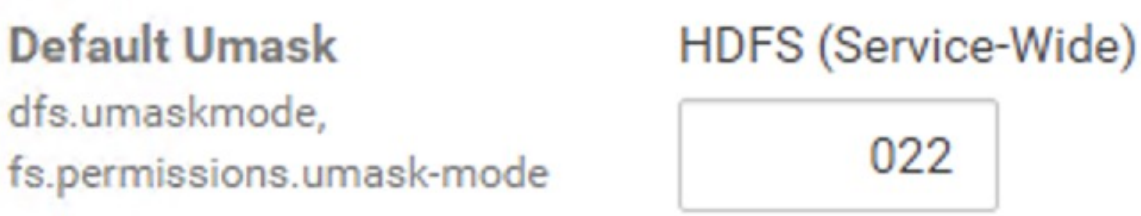
Vous pouvez utiliser Db2 et MySQL comme référentiel de métadonnées Analytic Server. Si vous comptez utiliser MySQL comme référentiel de métadonnées Analytic Server, suivez les instructions de la rubrique «Configuration de MySQL pour Analytic Server», à la page 41.

Connexion SSH sans mot de passe

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

Default Umask

Le paramètre Default Umask doit être défini sur 022. Exemple :



Le paramètre 022 est le umask le plus restrictif qui permette à Analytic Server de fonctionner.

Environnements Cloudera activés pour Kerberos

Si vous comptez installer Analytic Server dans un environnement Cloudera activé pour Kerberos, vous devez vérifier que Kerberos est correctement configuré de manière compatible avec Analytic Server.

Les sections ci-après s'appliquent aux environnements Cloudera où Kerberos est déjà installé. Les instructions suivantes doivent être suivies avant d'installer Analytic Server dans Cloudera. Elles supposent une connaissance de base de l'authentification Kerberos vu que les sections contiennent une terminologie spécifique à Kerberos (par exemple, **kinit**, **kadmin**, etc.).

Remarque : Analytic Server inspecte la configuration HDFS pour les valeurs associées à Kerberos à utiliser pour l'authentification.

Authentification Kerberos

Vérifiez que l'authentification kerberos est configurée sur chaque noeud du cluster Cloudera avant d'installer Analytic Server. Pour plus d'informations, voir Configuration de l'authentification dans Cloudera Manager dans la documentation du produit Cloudera.

Remarque : Après avoir configuré l'authentification Kerberos sur chaque noeud du cluster Cloudera, les services **cloudera-scm-server** et **cloudera-scm-agent** doivent être redémarrés avant d'installer Analytic Server. Le service **cloudera-scm-agent** doit être redémarré sur tous les noeuds du cluster.

Création des comptes requis dans Kerberos

1. Créez des comptes dans le référentiel utilisateur de Kerberos pour tous les utilisateurs auxquels vous souhaitez donner accès à Analytic Server.
2. Créez les mêmes comptes que ceux de l'étape précédente sur le serveur LDAP.
3. 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. Le groupe d'utilisateurs doit être défini comme `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 est conforme au paramètre YARN **Minimum user ID for submitting job**. Il s'agit du paramètre **min.user.id** dans le fichier `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.
4. Créez un dossier de base sur HDFS pour l'administrateur Analytic Server. Le droit d'accès au dossier doit être défini sur 755, le propriétaire sur `admin` et le groupe d'utilisateurs sur `hdfs`. Voir l'exemple **en gras** suivant :

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

Found 9 items

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

5. Créez des dossiers de base utilisateur sur HDFS pour tous les utilisateurs standard Analytic Server (par exemple, `user1`). Le propriétaire du dossier est l'utilisateur réel et le groupe d'utilisateurs doit être défini sur `hdfs`.
6. 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. Attribuez au paramètre **hadoop.proxyuser.hive.groups** la valeur * ou spécifiez un groupe englobant tous les utilisateurs habilités à se connecter à Analytic Server.

c. Attribuez au paramètre **hadoop.proxyuser.hive.hosts** la valeur * ou spécifiez une liste d'hôtes sur laquelle le métamagasin Hive et chaque instance de 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.

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 zone **Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml** située sous l'onglet **HDFS (Service-Wide)** > **Configuration**.
   - **Name:** `hadoop.proxyuser.as_user.hosts`
   - **Value:** *
   - **Name:** `hadoop.proxyuser.as_user.groups`
   - **Value:** *

   Remarque : les paramètres **core-site.xml** s'appliquent à la configuration Hadoop (et non pas à Analytic Server).
2. Depuis un shell de commande, exécutez la commande suivante sur le noeud Analytic Server :

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

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 44.

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`.

Outils precheck et postcheck d'installation - Cloudera

Emplacement des outils et prérequis

Avant d'installer le service Analytic Server, exécutez l'outil precheck sur tous les noeuds qui feront partie du service afin de vérifier que votre environnement Linux est prêt pour l'installation d'Analytic Server.

L'outil precheck est appelé automatiquement dans le cadre de l'installation. Il analyse chaque noeud Analytic Server avant d'exécuter l'installation sur chaque noeud. Vous pouvez appeler manuellement l'outil precheck sur chaque noeud afin de vérifier la machine avant d'installer le service.

Après l'exécution du fichier binaire autoextractible d'Analytic Server, l'outil precheck est situé sous les répertoires suivants :

- **Cloudera**

```
/opt/cloudera/parcels/AnalyticServer-3.2.2.0/tools/com.spss.ibm.checker.zip

[root@servername ~]# cd /opt/cloudera/parcels/AnalyticServer-3.2.2.0/tools/
[root@servername tools]# ls
com.spss.ibm.checker.zip configcollector.zip regex-files
```

Remarque : L'outil precheck n'est pas disponible dans le répertoire `tools` tant que vous n'avez pas exécuté le fichier binaire exécutable, puis distribué (**Download** > **Distribute**) et activé Analytic Server sur la page Parcels de Cloudera Manager.

Après l'installation d'Analytic Server, l'outil postcheck est situé sous le répertoire suivant :

- **Cloudera**

```
/opt/cloudera/parcels/AnalyticServer-3.2.2.0/tools/com.spss.ibm.checker.zip
```

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

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

Exécution de l'outil precheck

Automatique

L'outil precheck peut être appelé automatiquement dans le cadre de l'installation d'Analytic Server lorsque Analytic Server est installé en tant que service via la console Cloudera Manager. Vous devez pour cela saisir manuellement le nom d'utilisateur et le mot de passe de l'administrateur Cloudera Manager :

Figure 4. Paramètres administrateur de Cloudera Manager

Manuelle

Vous pouvez appeler manuellement l'outil precheck sur chaque noeud de cluster.

L'exemple suivant vérifie au préalable le cluster Cloudera `MyCluster` qui s'exécute sur `myclouderahost.ibm.com:7180`, et utilise les données d'identification et de connexion `admin:admin` :

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

Remarques :

- Les arguments `--target`, `--host`, `--port` et `--username` sont requis.
- La valeur `--host` doit correspondre à l'adresse IP ou à 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` inclut une aide pour son utilisation, laquelle est affichée en spécifiant l'argument `--h` (`python ./precheck.py --help`).
- L'argument `--cluster` est facultatif (le cluster actuel est utilisé si `--cluster` n'est pas spécifié).

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 --target C --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 --target C --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 --target C --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default --check Modèle_génération_PYSPARK_AS
```

Remarques :

- Les arguments `--target`, `--host`, `--port` et `--username` sont requis.
- La valeur `--host` doit correspondre à l'adresse IP ou à un nom de domaine complet.
- L'outil vous invite à saisir un mot de passe lorsque l'argument -password est omis.
- La commande `postcheck.py` inclut une aide pour son utilisation, laquelle est affichée en spécifiant 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 Cloudera

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

Analytic Server 3.2.2

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 9. Fichiers binaires autoextractibles Analytic Server

Description	Nom du fichier binaire
IBM SPSS Analytic Server 3.2.2 for Cloudera 5.11, 5.12, 5.13, 5.14, 5.15, 6.0, 6.1, 6.2, and 6.3 Ubuntu (en anglais)	`spss_as-3.2.2.0-cdh5.11-6.3-ubun.bin`
IBM SPSS Analytic Server 3.2.2 for Cloudera 5.11, 5.12, 5.13, 5.14, 5.15, 6.0, 6.1, 6.2, and 6.3 Linux x86-64 (en anglais)	`spss_as-3.2.2.0-cdh5.11-6.3-lx86.bin`

2. Exécutez le programme d'installation autoextractible `*.bin` Cloudera sur le noeud de cluster principal 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 si l'emplacement par défaut a été modifié.

3. Utilisez la commande suivante pour redémarrer Cloudera Manager une fois l'installation terminée :

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

4. Ouvrez l'interface de Cloudera Manager (par exemple, `http://${CM_HOST}:7180/cmf/login` avec les données d'identification par défaut `admin`/`admin`), actualisez la liste **Remote Parcel Repository URLs** (située sous **Host** > **Parcels** > **cliquez sur Configuration**), et vérifiez que l'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. 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** :

   Remarque : 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.

   - 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

   Utilisation de MySQL comme référentiel des métadonnées d'Analytic Server

   - Classe de pilote de métamagasin Analytic Server : `com.mysql.jdbc.Driver`
   - URL du référentiel de métamagasin Analytic Server : `jdbc:mysql://${MySQL_DB}/{DBName}?createDatabaseIfNotExist=true`

     `{MySQL_DB}` est le nom d'hôte du serveur sur lequel MySQL est installé

   Utilisation de Db2 comme référentiel des métadonnées d'Analytic Server

   - Classe de pilote de métamagasin Analytic Server : `com.ibm.db2.jcc.DB2Driver`
   - URL du référentiel de métamagasin Analytic Server : `jdbc:db2://{Db2_HOST}:{PORT}/{DBName}:currentSchema={SchemaName};`

     `{Db2_HOST}` est le nom d'hôte du serveur sur lequel Db2 est installé.

     `{PORT}` est le port sur lequel Db2 est à l'écoute.

     `{SchemaName}` est un schéma disponible non utilisé.

     Consultez votre administrateur Db2 si vous n'êtes pas sûr des valeurs à entrer.

   Configuration LDAP

   Analytic Server utilise un serveur LDAP pour stocker et authentifier les utilisateurs et les groupes. Vous devez fournir les informations de configuration LDAP lors de l'installation d'Analytic Server.

   Tableau 10. Paramètres de configuration LDAP

Paramètre LDAP	Description
`as.ldap.type`	Type LDAP. Valeurs admises : `ads`, `ad` ou `openladp`. • `ads` - Apache Directory Server (valeur par défaut) • `ad` - Microsoft Active Directory • `openladp` - OpenLDAP
`as.ldap.host`	Hôte LDAP
`as.ldap.port`	Numéro de port LDAP
`as.ldap.binddn`	Nom distinctif de liaison LDAP
`as.ldap.bindpassword`	Mot de passe du nom distinctif de liaison LDAP
`as.ldap.basedn`	Nom distinctif de base LDAP

Tableau 10. Paramètres de configuration LDAP (suite)	
Paramètre LDAP	**Description**
`as.ldap.filter`	Règle de filtrage d'utilisateurs et de groupes LDAP **Remarque :** Si cette valeur contient une barre verticale |, celle-ci doit être précédée du caractère d'échappement barre oblique inversée (par exemple, \\\|).
`as.ldap.ssl.enabled`	Indique si SSL doit être utilisé pour la communication entre Analytic Server et LDAP. Valeurs admises ; `true` ou `false`.
`as.ldap.ssl.reference`	ID de référence SSL LDAP
`as.ldap.ssl.content`	Configuration SSL LDAP

- Par défaut, `as.ldap.type` est défini sur `ads` et les autres paramètres associés reçoivent les valeurs par défauts. Une exception cependant : vous devez fournir un mot de passe pour le paramètre `as.ldap.bindpassword`. Analytic Server utilise les paramètres de configuration pour installer un serveur ADS (Apache Directory Server) et lancer son initialisation. Le profil ADS par défaut inclut l'utilisateur `admin` avec pour mot de passe `admin`. Vous pouvez gérer les utilisateurs via la Console Analytic Server ou importer les informations sur les utilisateurs et les groupes depuis un fichier XML en utilisant le script `importUser.sh` situé sous le dossier `<Analytic Root>/bin`.
- Si vous comptez utiliser un serveur LDAP externe, tel que Microsoft Active Directory ou OpenLDAP, vous devez définir les informations de configuration d'après les valeurs LDAP réelles. Pour plus d'informations, voir Configuration de registres utilisateur LDAP dans Liberty.
- Vous pouvez modifier la configuration LDAP après qu'Analytic Server a été installé (par exemple, passer d'Apache Directory Server à OpenLDAP). Toutefois, si votre installation initiale utilisait Microsoft Active Directory ou OpenLDAP et que vous décidez plus tard de passer à Apache Directory Server, Analytic Server n'installe pas un serveur Apache Directory Server lors de l'installation. Le serveur Apache Directory Server n'est installé que si vous le sélectionnez lors de l'installation initiale d'Analytic Server.

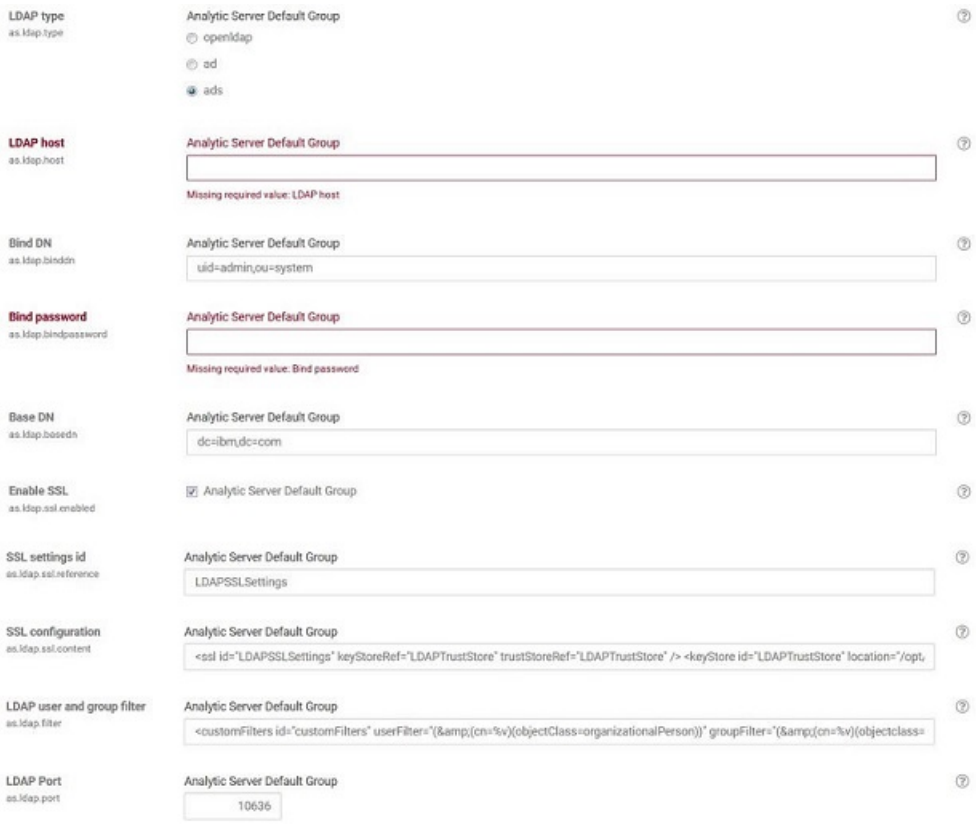


Figure 5. Exemple de paramètres de configuration LDAP

8. Lors de l'installation d'Analytic Server dans un environnement Cloudera activé pour Kerberos, les paramètres suivants doivent également être configurés dans L'assistant **Add Service Wizard**:

   Remarque : Analytic Server inspecte la configuration HDFS pour les valeurs associées à Kerberos à utiliser pour l'authentification.

   - Sélectionnez `Kerberos` pour le paramètre **Sécurité Analytic Server** si vous désirez activer l'authentification Kerberos lors de la connexion à la console Analytic Server. Lorsque **Kerberos** est sélectionné pour le paramètre **Sécurité Analytic Server**, la console Analytic Server adopte par défaut le mode de connexion Kerberos.
   - Sélectionnez `Kerberos` pour le paramètre **Méthode de connexion à la source de données de la base de données Analytic Server** si vous désirez vous connecter à des bases de données activées pour Kerberos. Lorsque **Kerberos** est sélectionné pour le paramètre **Méthode de connexion à la source de données de la base de données Analytic Server**, la console Analytic Server utilise le mode Kerberos lors de la connexion à une base de données.
   - Les paramètres **Domaine Kerberos** et **Hôtes KDC** sont obligatoires. Les valeurs de **Domaine Kerberos** (**`as.kdc.realms`**) et **Hôte KDC** (**`kdcserver`**) sont consignées dans le fichier `krb5.conf` sur le serveur KDC (Kerberos Key Distribution).

     Les noms de domaines multiples sont pris en charge lorsqu'ils sont séparés par des virgules. Les noms de domaines Kerberos correspondent et sont associés aux noms d'utilisateur. Par exemple, les noms d'utilisateur `UserOne@us.ibm.com` et `UserTwo@eu.ibm.com` correspondent respectivement aux domaines `us.ibm.com,eu.ibm.com`.

     Des relations d'approbation inter-domaines Kerberos doivent être configurées lorsque plusieurs domaines sont spécifiés comme **Domaine Kerberos**. Le nom d'utilisateur saisi à l'invite de connexion à la console Analytic Server ne doit pas inclure le suffixe du nom de domaine. Par conséquent, lorsque plusieurs domaines sont spécifiés, les utilisateurs doivent sélectionner le domaine de leur choix dans la liste déroulante **Domaines**.

     Remarque : Lorsqu'un seul domaine est spécifié, la liste déroulante **Domaines** n'apparaît pas lors de la connexion à Analytic Server.

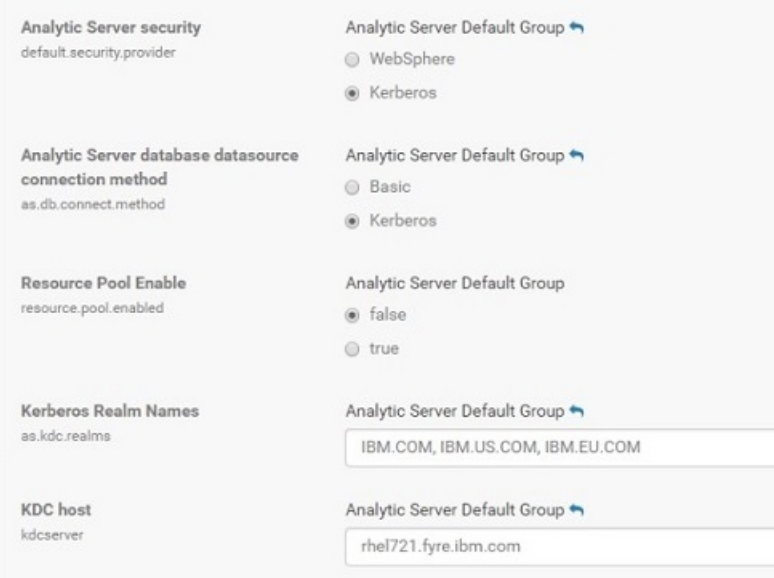


Figure 6. Exemple de paramètres Kerberos

Remarques :

- Les paramètres **Sécurité Analytic Server** et **Méthode de connexion à la source de données de la base de données Analytic Server** sont applicables au client IBM SPSS Modeler et à l'authentification de la console Analytic Server.
- Lorsque **Méthode de connexion à la source de données de la base de données Analytic Server** est définie à `Kerberos`, vous devez vérifier que les bases de données cibles sont également activées pour Kerberos.
- Les paramètres **Sécurité Analytic Server** et **Méthode de connexion à la source de données de la base de données Analytic Server** ne configurent pas l'authentification Kerberos sur le cluster Hadoop. Pour plus d'informations, reportez-vous à la section "Activation de l'emprunt d'identité Kerberos".
- Si vous désirez que l'authentification Kerberos soit activée lors d'une connexion, vous devez déployer le client IBM SPSS Modeler en tant que client Kerberos valide. Ceci est réalisé en utilisant la commande **`addprinc`** sur le serveur KDC (Kerberos Key Distribution Center). Pour plus d'informations, reportez-vous à votre documentation IBM SPSS Modeler.

Lors de l'installation d'Analytic Server dans un environnement Cloudera activé pour Kerberos, vous devez également créer les comptes requis dans Kerberos et autoriser l'emprunt d'identité Kerberos. Pour plus d'informations, voir «Configuration de Kerberos», à la page 51.

Avertissement : 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-3.2.2.0-el7.parcel
- AnalyticServer-3.2.2.0-el7.parcel.sha
- manifest.json

SuSE Linux requiert les fichiers suivants :

- AnalyticServer-3.2.2.0-sles12.parcel

- AnalyticServer-3.2.2.0-sles12.parcel.sha
- manifest.json

Ubuntu Linux 16.04 requiert les fichiers suivants :

- AnalyticServer-3.2.2.0-xenial.parcel
- AnalyticServer-3.2.2.0-xenial.parcel.sha

Ubuntu Linux 18 requiert les fichiers suivants :

- AnalyticServer-3.2.2.0-bionic.parcel
- AnalyticServer-3.2.2.0-bionic.parcel.sha

1. Téléchargez et exécutez le programme d'installation autoextractible `*.bin` Cloudera sur le noeud de cluster principal 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 principal Cloudera Manager. Le chemin par défaut est `/opt/cloudera/parcel-repo` (il peut être configuré dans l'interface utilisateur de Cloudera Manager).
3. Utilisez la commande suivante pour redémarrer Cloudera Manager :

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

   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.
4. Cliquez sur **Distribute** > **Activate**.

   Le fichier parcel **AnalyticServer** est signalé comme ayant été distribué et activé.
5. Dans Cloudera Manager, ajoutez Analytic Server en tant que service. Pour plus d'informations, reportez-vous aux étapes 7 et 8 de la rubrique "Installation en ligne".

Configuration de Cloudera

Après l'installation, vous devez créer les comptes requis sur le système d'exploitation de cluster.

1. Créez des comptes utilisateur du système d'exploitation pour tous les utilisateurs auxquels vous prévoyez d'accorder un accès à Analytic Server sur chaque Analytic Server et noeud Hadoop (ces utilisateurs sont également configurés en tant que registres utilisateur LDAP). Le groupe d'utilisateurs doit être défini comme `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 est conforme au paramètre YARN **Minimum user ID for submitting job**. Il s'agit du paramètre **`min.user.id`** défini dans le fichier `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.
2. Créez un dossier de base sur HDFS pour l'administrateur Analytic Server. Le droit d'accès au dossier doit être défini sur 755, le propriétaire sur `admin` et le groupe d'utilisateurs sur `hdfs`. Voir l'exemple **en gras** suivant :

   ```
   [root@xxxxx configuration]# hadoop fs -ls /user
   drwxr-xr-x   - admin    hdfs            0 2017-06-08 01:33 /user/admin
   drwxr-xr-x   - user1    hdfs          0 2017-06-06 01:00 /user/user1
   ```
3. Créez des dossiers de base utilisateur sur HDFS pour tous les utilisateurs standard Analytic Server (par exemple, `user1`). Le propriétaire du dossier est l'utilisateur réel et le groupe d'utilisateurs doit être défini sur `hdfs`.

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

La valeur par défaut de **`tenant_id`** dans le fichier IBM SPSS Modeler `options.cfg` est **`ibm`**. Vous pouvez identifier les titulaires dans la console Analytic Server. Voir le manuel *IBM SPSS Analytic Server - Guide d'administration* pour plus d'informations sur la gestion des titulaires.

Configuration d'un registre LDAP

LDAP est configuré lors de l'installation d'Analytic Server. Vous pouvez opter pour un autre serveur LDAP après l'installation d'Analytic Server.

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 -p /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 Cloudera Manager et mettez à jour le paramètre de configuration d'Analytic Server **`ssl_cfg`** avec les paramètres de configuration SSL appropriés. 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}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 appropriés. 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. Les sections suivantes indiquent les paramètres de sécurité requis afin que Kerberos soit configuré correctement et de manière compatible avec Analytic Server.

Remarque : Analytic Server inspecte la configuration HDFS pour les valeurs associées à Kerberos à utiliser pour l'authentification.

Analytic Server et paramètres Kerberos

Lorsque vous installez Analytic Server dans un environnement Cloudera activé pour Kerberos, les paramètres suivants sont de rigueur.

- Sélectionnez `Kerberos` pour le paramètre **Sécurité Analytic Server** si vous désirez activer l'authentification Kerberos lors de la connexion à la console Analytic Server. Lorsque **Kerberos** est sélectionné pour le paramètre **Sécurité Analytic Server**, la console Analytic Server adopte par défaut le mode de connexion Kerberos.
- Sélectionnez `Kerberos` pour le paramètre **Méthode de connexion à la source de données de la base de données Analytic Server** si vous désirez vous connecter à des bases de données activées pour Kerberos. Lorsque **Kerberos** est sélectionné pour le paramètre **Méthode de connexion à la source de données de la base de données Analytic Server**, la console Analytic Server utilise le mode Kerberos lors de la connexion à une base de données.
- Les paramètres **Domaine Kerberos** et **Hôtes KDC** sont obligatoires. Les valeurs de **Domaine Kerberos** (**as.kdc.realms**) et **Hôte KDC** (**kdcserver**) sont consignées dans le fichier `krb5.conf` sur le serveur KDC (Kerberos Key Distribution).

  Les noms de domaines multiples sont pris en charge lorsqu'ils sont séparés par des virgules. Les noms de domaines Kerberos correspondent et sont associés aux noms d'utilisateur. Par exemple, les noms d'utilisateur `UserOne@us.ibm.com` et `UserTwo@eu.ibm.com` correspondent respectivement aux domaines `us.ibm.com,eu.ibm.com`.

  Des relations d'approbation inter-domaines Kerberos doivent être configurées lorsque plusieurs domaines sont spécifiés comme **Domaine Kerberos**. Le nom d'utilisateur saisi à l'invite de connexion à la console Analytic Server ne doit pas inclure le suffixe du nom de domaine. Par conséquent, lorsque plusieurs domaines sont spécifiés, les utilisateurs doivent sélectionner le domaine de leur choix dans la liste déroulante **Domaines**.

  Remarque : Lorsqu'un seul domaine est spécifié, la liste déroulante **Domaines** n'apparaît pas lors de la connexion à Analytic Server.

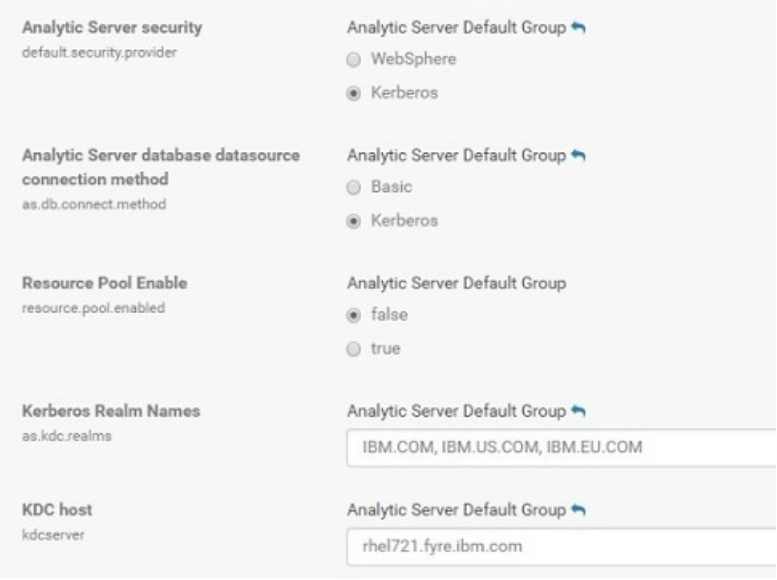


Figure 7. Exemple de paramètres Kerberos

Remarques :

- Les paramètres **Sécurité Analytic Server** et **Méthode de connexion à la source de données de la base de données Analytic Server** sont applicables au client IBM SPSS Modeler et à l'authentification de la console Analytic Server.
- Lorsque **Méthode de connexion à la source de données de la base de données Analytic Server** est définie à `Kerberos`, vous devez vérifier que les bases de données cibles sont également activées pour Kerberos.
- Les paramètres **Sécurité Analytic Server** et **Méthode de connexion à la source de données de la base de données Analytic Server** ne configurent pas l'authentification Kerberos sur le cluster Hadoop. Pour plus d'informations, reportez-vous à la section "Activation de l'emprunt d'identité Kerberos".
- Si vous désirez que l'authentification Kerberos soit activée lors d'une connexion, vous devez déployer le client IBM SPSS Modeler en tant que client Kerberos valide. Ceci est réalisé en utilisant la commande **`addprinc`** sur le serveur KDC (Kerberos Key Distribution Center). Pour plus d'informations, reportez-vous à votre documentation IBM SPSS Modeler.

Création des comptes requis dans Kerberos

1. Créez des comptes dans le référentiel utilisateur de Kerberos pour tous les utilisateurs auxquels vous souhaitez donner accès à Analytic Server.
2. Créez les mêmes comptes que ceux de l'étape précédente sur le serveur LDAP.
3. 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. Le groupe d'utilisateurs doit être défini comme `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 est conforme au paramètre YARN **Minimum user ID for submitting job**. Il s'agit du paramètre **`min.user.id`** dans le fichier `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.

4. Créez un dossier de base sur HDFS pour l'administrateur Analytic Server. Le droit d'accès au dossier doit être défini sur 755, le propriétaire sur `admin` et le groupe d'utilisateurs sur `hdfs`. Voir l'exemple **en gras** suivant :

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

Found 9 items

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

5. Créez des dossiers de base utilisateur sur HDFS pour tous les utilisateurs standard Analytic Server (par exemple, `user1`). Le propriétaire du dossier est l'utilisateur réel et le groupe d'utilisateurs doit être défini sur `hdfs`.
6. 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. Attribuez au paramètre **hadoop.proxyuser.hive.groups** la valeur `*` ou spécifiez un groupe englobant tous les utilisateurs habilités à se connecter à Analytic Server.
   c. Attribuez au paramètre **hadoop.proxyuser.hive.hosts** la valeur `*` ou spécifiez une liste d'hôtes sur laquelle le métamagasin Hive et chaque instance de 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.

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 zone **Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml** située sous l'onglet **HDFS (Service-Wide)** > **Configuration**.
   - **Name:** `hadoop.proxyuser.as_user.hosts`
   - **Value:** `*`
   - **Name:** `hadoop.proxyuser.as_user.groups`
   - **Value:** `*`

   Remarque : les paramètres **core-site.xml** s'appliquent à la configuration Hadoop (et non pas à Analytic Server).
2. Depuis un shell de commande, exécutez la commande suivante sur le noeud Analytic Server :

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

Configuration de HAProxy pour la 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 nom de principal du service Kerberos (HTTP/<nom_hôte_proxy>@<domaine>) et le fichier de clés de l'hôte HAProxy, où <nom_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>@<realm>
```

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 maintenant se connecter à Analytic Server à l'aide de l'option **Single sign on log in** dans l'écran de connexion à IBM SPSS Analytic Server.

Désactivation de Kerberos

1. Désactivez Kerberos dans la console Cloudera Manager.
2. Arrêtez le service Analytic Server.
3. Modifiez les paramètres suivants dans la zone **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties** :

   Analytic Server security(default.security.provider) > **WebSphere**

   Analytic Server database datasource connection method(as.db.connect.method) > **Basic**
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.

Installation de certificats de fournisseur tiers

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 au format Base64. Pour le codage, vous pouvez utiliser l'outil securityUtility, par exemple `{AS_ROOT}/ae_wlpserver/bin/securityUtility encode <mot_de_passe>`

Si vous souhaitez générer un certificat autosigné, vous pouvez utiliser l'utilitaire securityUtility ; par exemple `{AS_ROOT}/ae_wlpserver/bin/securityUtility createSSLCertificate --server=aeserver --password=mypassword --validity=365 --subject=CN=myfqdnserver,O=myorg,C=mycountry`. Pour plus d'informations sur securityUtility et sur les autres paramètres SSL, reportez-vous à la documentation WebSphere Liberty Profile.

Remarques :

- Vous devez fournir un nom de domaine hôte approprié pour la valeur `CN`.
- Remplacez **`mypassword`**, **`myfqdnserver`**, **`myorg`** et **`mycounty`** par vos données d'identification. Notez que **`myfqdnserver`** est le nom de domaine complet du noeud Analytic Server.
- **`aeserver`** est le nom du serveur Liberty (la valeur doit être **`aeserver`**).

Pour plus d'informations sur l'utilitaire **`securityUtility`** et d'autres paramètres SSL, reportez-vous à la documentation WebSphere Liberty Profile et securityUtility command.

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

Génération de certificats autosignés

Vous pouvez utiliser `securityUtility` pour générer des certificats autosignés. Exemple :

```
/opt/cloudera/parcels/AnalyticServer-3.2.2.0/ae_wlpserver/bin/securityUtility createSSLCertificate
--server=<myserver> --password=<mypassword> --validity=365 --subject=CN=<mycompany>,O=<myOrg>,C=<myCountry>
```

Remarques :

- Vous devez fournir un nom de domaine hôte approprié pour la valeur **`CN`**.
- Copiez les informations de `key.jks` dans `trust.jks` (les deux fichiers doivent être identiques).
- Editez le paramètre `ssl.keystore.config`. Exemple :

```
<ssl id="defaultSSLConfig"
keyStoreRef="defaultKeyStore"
trustStoreRef="defaultTrustStore"
clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
location="/opt/cloudera/parcels/AnalyticServer-3.2.2.0
/ae_wlpserver/usr/servers/aeserver/resources/security/key.jks"
type="JKS"
password="{xor}Dz4sLG5tbGs="/>
<keyStore id="defaultTrustStore"
location="/opt/cloudera/parcels/AnalyticServer-3.2.2.0
/ae_wlpserver/usr/servers/aeserver/resources/security/trust.jks"
type="JKS"
password="{xor}Dz4sLG5tbGs="/>
```

Communication avec Apache Hive via SSL

Vous devez mettre à jour le fichier `hive.properties` afin de communiquer avec Apache Hive via une connexion SSL. Sinon, si votre environnement Apache Hive est activé pour haute disponibilité, vous pouvez aussi sélectionner les paramètres de haute disponibilité sur la page Analytic Server Sources de données principale.

Mise à jour du fichier hive.properties

1. Ouvrez le fichier `hive.properties`. Ce fichier se trouve dans le répertoire : `/opt/cloudera/parcels/AnalyticServer/ae_wlpserver/usr/servers/aeserver/configuration/database`
2. Localisez la ligne suivante :

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

3. Mettez à jour la ligne en lui ajoutant les informations **en gras** ci-dessous :

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password};
ssl=true;sslTrustStore=pathtotheirtruststorefile;trustStorePassword=xxxtheirTrustStorePassword
```

4. Enregistrez le fichier `hive.properties`.

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. Provisionnez l'environnement de serveur pour Essentials for R. Pour plus d'informations, reportez-vous à l'étape 1 dans «Activation de la prise en charge d'Essentials for R», à la page 23.
2. 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.
3. 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++
4. 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 la version 3.5.1. 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-3/R-3.4.4.tar.gz pour R 3.4.4.
      - 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-3.4.4.tar.gz` a été téléchargée, puis copiée dans `/root`, l'adresse URL est `file:///root/R-3.4.4.tar.gz`.

      Remarque : D'autres versions de R sont disponibles à l'adresse https://cran.r-project.org/src/base/.
   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.
5. 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).
6. 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.

7. 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 dans le métamagasin Analytic Server et dans chaque noeud 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 11. 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
Apache Impala	JDBC 4 avec 2.5.5 ou version ultérieure	`ImpalaJDBC4.jar`, `commons-codec-*.jar`, `commons-logging-*.jar`, `httpclient-*.jar`, `httpcore-*.jar`, `log4j-*.jar`, `libthrift-*.jar`, `libfb303-*.jar`, `slf4j-api-*.jar`, `ql.jar`, `zookeeper-*.jar`, `TCLIServiceClient.jar`	Apache
DashDB	Service Bluemix	`db2jcc.jar`	IBM
Db2 pour Linux, UNIX et Windows	11.1, 10.5, 10.1, 9.7	`db2jcc.jar`	IBM
Db2 z/OS	11, 10	`db2jcc.jar`, `db2_license_cisuz.jar`	IBM
Greenplum	5.x	postgresql.jar	Greenplum
Hive	1.1	`hive-jdbc-*.jar`	Apache

Tableau 11. Bases de données prises en charge (suite)

Base de données	Versions prises en charge	Fichiers jar du pilote JDBC	Fournisseur
MySQL	5.7, 5.6	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7, 6.x	`nzjdbc.jar`	IBM

Oracle	19c, 12c, 11g R2 (11.2)	19c :
		`ojdbc8.jar`, `orai18n.jar`
		12c and 11g R2 (11.2):
		`ojdbc6.jar`, `orai18n.jar`
			Oracle

SQL Server	2017, 2014, 2012, 2008 R2	`sqljdbc4.jar`	Microsoft
Teradata	15.1, 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 `{HIVE_HOME}/auxlib` et au répertoire `/usr/share/hive` sur le métamagasin Analytic Server et sur chaque noeud Analytic Server.
3. Redémarrez le service Hive Metastore.
4. Redémarrez chaque instance du service Analytic Server.

Remarque :

Lors de l'accès aux données HBase data via une source de données Analytic Server HCatalog, l'utilisateur doit disposer d'un droit de lecture sur les tables HBase.

- Dans les environnements non Kerberos, Analytic Server accède à HBase en tant que `as_user` (`as_user` doit disposer d'un droit de lecture dans HBase).
- Dans les environnements kerberos, `as_user` et l'utilisateur qui se connecte doivent tous deux disposer d'un droit de lecture sur les tables HBase.

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.

Travaux MapReduce v2

Utilisez le paramètre **`preferred.mapreduce`** dans la zone **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties** pour indiquer comment traiter les travaux MapReduce :

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

Propriété	Description
`preferred.mapreduce`	Contrôle la méthode dans laquelle exécuter les travaux MapReduce. Les valeurs valides sont les suivantes : • `spark` • `m3r` • `hadoop` Par exemple : `preferred.mapreduce=spark`

Apache Spark

Si vous souhaitez utiliser Spark (version 2.x ou ultérieure), vous devez définir la propriété `spark.version` pendant l'installation d'Analytic Server.

1. Ouvrez Cloudera Manager et sélectionnez la propriété `spark.version` appropriée (par exemple, None ou `2.x`) dans la zone **Analytic Server Spark Version**.
2. Sauvegardez la configuration.

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. Sur 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. Sur 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 aux données Apache Impala compatibles SSH

1. Définissez les paramètres SSL Impala ci-après dans la console Cloudera Manager.

   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/ll4200v21.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/ll4200v21.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.

Prise en charge de plusieurs clusters

La fonctionnalité de prise en charge de clusters multiples est un approfondissement des capacités de haute disponibilité d'IBM SPSS Analytic Server et fournit un isolement amélioré dans les environnements à plusieurs titulaires. Par défaut, l'installation du service Analytic Server (dans Ambari ou ClouderaManager) entraîne la définition d'un serveur analytique unique.

La spécification du cluster définit l'appartenance au cluster Analytic Server. La modification de la spécification du cluster est réalisée via un contenu XML (dans la zone `analytics-cluster` de la configuration d'Ambari Analytic Server ou en modifiant manuellement le fichier `configuration/analytics-cluster.xml` de Cloudera Manager). Si vous configurez plusieurs clusters Analytic Server, vous devez acheminer les demandes à chaque cluster Analytic Server avec son propre équilibreur de charge.

L'utilisation du dispositif de clusters multiples assure que le travail concernant un titulaire n'affecte pas négativement celui en cours d'exécution dans le cluster d'un autre titulaire. Quant à la haute disponibilité des travaux, le basculement des travaux se produit uniquement au sein du cluster Analytic Server où le travail a été déclenché. L'exemple suivant décrit une spécification XML de clusters multiples.

Remarque : La haute disponibilité d'Analytic Server peut être réalisée en l'ajoutant en tant que service dans plusieurs noeuds dans votre cluster.

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

```
    </cluster>
</analayticServerClusterSpec>
```

Dans l'exemple précédent, deux équilibreurs de charge sont requis. L'un envoie des demandes aux membres de `cluster1` (`one.cluster` et `two.cluster`), l'autre aux membres de `cluster2` (`three.cluster` et `four.cluster`).

L'exemple suivant décrit la spécification XML d'un cluster unique (configuration par défaut).

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

Dans l'exemple suivant, un seul équilibreur de charge est requis pour gérer les cas où plusieurs membres du cluster ont été configurés.

Remarques

- Seules les instances de cluster unique prennent en charge l'utilisation de caractères génériques dans l'élément **memberName** (par exemple, cluster cardinality = "1"). Les valeurs valides pour l'élément cardinality sont 1 et 1+.
- **memberName** doit être spécifié de la même manière que le nom d'hôte auquel le rôle Analytic Server est affecté.
- Tous les serveurs de chaque cluster doivent être redémarrés après une modification de la configuration de cluster.
- Dans Cloudera Manager, vous devez modifier et gérer le fichier `analytics-cluster.xml` sur tous les noeuds d'Analytic Server. Tous les noeuds doivent être contrôlés pour vérifier que leur contenu reste identique.

Mise à niveau de Python - CDH

Cette section décrit la mise à niveau manuelle de Python 2.x à Python 3.7

1. Installez Python 3.7 sur chaque noeud du cluster. Pour plus d'informations, reportez-vous au site Python.
2. Installez NumPy sur chaque noeud du cluster. Pour plus d'informations, reportez-vous aux instructions d'installation NumPy.
3. Installez pandas sur chaque noeud du cluster. Pour plus d'informations, reportez-vous aux instructions d'installation pandas.
4. Dans Cloudera Manager, mettez à jour la section **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties** de manière à y inclure le chemin de l'exécutable Python 3.7. Exemple :

   ```
   spark.driver.python=/opt/python3/bin/python3.7
   ```

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
```

Configuration d'une allocation de ressource dynamique distincte pour chaque pool de ressources YARN - Cloudera

Vous pouvez configurer une allocation de ressource dynamique distincte pour chaque pool de ressources YARN.

Mappage des modes utilisateur et titulaire - Cloudera

Les tâches utilisateur et titulaire peuvent être soumises à différents pools de ressources YARN, et chaque utilisateur ou titulaire est mappé à un pool de ressources YARN différent (pour tirer parti de l'allocation de ressource dynamique). Le mode **utilisateur** ou le mode **titulaire** peut être défini pour le mappage aux pools de ressources YARN. Avant Analytic Server 3.2.1 groupe de correctifs 1, tous les travaux Spark étaient limités à un seul pool de ressources YARN.

A partir de IBM SPSS Analytic Server 3.2.1, groupe de correctifs 1, lorsque le flux d'un utilisateur/titulaire entraîne l'exécution de travaux Spark sur le système, un pool de ressources YARN distinct s'exécute en tant qu'utilisateur/titulaire ayant soumis le flux à Analytic Server. Plusieurs pools de ressources YARN peuvent s'exécuter simultanément pour différentes tâches utilisateur/titulaire.

Chaque pool de ressources YARN continue à s'exécuter tant que l'utilisateur est connecté à Analytic Server (et pendant un certain temps après la déconnexion de l'utilisateur et l'absence de travaux utilisateur actifs). La durée après déconnexion peut être contrôlée par la variable de configuration : **as.spark.driver.cleanup.delay**.

Un processus **SparkDriver** est créé pour chaque utilisateur qui soumet le travail Spark. Le processus **SparkDriver** de chaque utilisateur se termine lorsque l'utilisateur n'a plus de travaux actifs pendant environ 2 minutes (la valeur par défaut) et aucune activité **HTTPSession**.

Remarque : Tous les processus **SparkDriver** se terminent lorsque Analytic Server s'arrête.

Pour ajouter Analytic Server à un cluster existant, procédez comme suit :

1. Dans Cloudera Manager, accédez à **SPSS Analytic Server Service** > **Configuration**.
2. Remplacez la valeur de **Resource Pool Enable: resource.pool.enabled** par `true`.
3. Ajoutez les propriétés suivantes à la zone **Analytic Server Advanced Configuration Snippet (Safety Valve)** > **analyticserver-conf.config.properties**:

```
//Using user/tenant mapping to YARN pool
yarn.queue.mode=<user/tenant>
yarn.queue.mapping=<user1:test,user2:production>/<tenant1:test,tenant2:production>
yarn.queue.default=default
as.spark.driver.cleanup.delay=2
as.sparkdriver.max.memory=1G
```

Tableau 13. Propriétés analyticserver-conf.config.properties personnalisées

Propriété	Description
`yarn.queue.mode`	Définit le mode de mappage pour les pools de ressources YARN. Lorsque `yarn.queue.mode=user`, une application YARN distincte est exécutée pour chaque utilisateur qui a soumis un travail/flux à Analytic Server. Plusieurs pools de ressources YARN peuvent s'exécuter simultanément pour les différents travaux/flux des utilisateurs. Lorsque `yarn.queue.mode=tenant`, une application YARN distincte est exécutée pour chaque titulaire qui a soumis un travail/flux à Analytic Server. Plusieurs pools de ressources YARN peuvent s'exécuter simultanément pour les différents travaux/flux des titulaires.
`yarn.queue.mapping`	Mappe les paires utilisateur ou titulaire vers les pools de ressources YARN définis dans la configuration du pool de ressources dynamiques du gestionnaire Cloudera. Les paires doivent être séparées par des virgules (par exemple, `tenant1:test,tenant2:production` pour les titulaires ou `user1:test,user2:production` pour les utilisateurs).
`yarn.queue.default`	Nom du pool de ressources YARN par défaut auquel l'application est soumise. Vous pouvez spécifier un nom de pool de ressources YARN personnalisé dans la configuration du pool de ressources dynamiques.

Tableau 13. Propriétés analyticserver-conf.config.properties personnalisées (suite)	
Propriété	**Description**
`as.spark.driver.cleanup.delay`	Entier représentant le nombre de minutes à s'écouler entre la déconnexion et l'arrêt de l'application YARN d'un utilisateur. La valeur par défaut est **2**. Cette propriété est facultative.
`as.sparkdriver.max.memory`	Définit le volume de mémoire utilisé par chaque processus **`SparkDriver`**. La valeur par défaut est **`1G`**. Cette propriété est facultative.

Références

Reportez-vous aux sites suivants pour plus d'informations :

- https://spark.apache.org/docs/latest/running-on-yarn.html

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 3.2.1.1 existante et que vous avez fait l'acquisition d'une version plus récente, vous pouvez migrer vos paramètres de la version 3.2.1.1 vers votre nouvelle installation.

Restriction : Votre version 3.2.1.1 et les nouvelles installations ne peuvent pas coexister sur le même cluster Hadoop. Si vous configurez votre nouvelle installation de manière qu'elle utilise le même cluster Hadoop que votre installation 3.2.1.1, cette dernière ne fonctionnera plus.

Etapes de migration d'une version 3.2.1.1 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 44.
2. Copiez l'espace de travail d'analyse de votre ancienne installation vers la nouvelle.
   a. Si vous n'êtes pas certain de l'emplacement de l'espace de travail d'analyse, exécutez la commande `hadoop -fs ls`. Le chemin de l'espace de travail d'analyse suit la forme `/user/as_user/analytic-root/analytic-workspace`, où `as_user` est l'ID utilisateur propriétaire de cet espace.
   b. Connectez-vous à l'hôte de la nouvelle installation Analytic Server en tant qu'`as_user`. Supprimez le répertoire `/user/as_user/analytic-root/analytic-workspace`, s'il existe.
   c. Utilisez les commandes `hadoop fs -copyToLocal` et `hadoop fs -copyFromLocal` pour copier l'espace de travail d'analyse de l'ancien serveur vers le dossier `/user/as_user/analytic-root/analytic-workspace` du nouveau serveur (vérifiez que le propriétaire est défini comme `as_user`).
3. Si vous utilisez le serveur Apache Directory Server intégré, sauvegardez la configuration utilisateurs/groupes actuelle à l'aide d'un outil client LDAP tiers. Après l'installation d'Analytic Server 3.2.2, importez dans le serveur Apache Directory Server la configuration utilisateurs/groupes que vous aviez sauvegardée.

   Remarque : Vous pouvez ignorer cette étape si vous utilisez un serveur LDAP externe.
4. Dans Cloudera Manager, arrêtez le service Analytic Server.
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.

Important : Il se peut que le script **configcollector** fourni ne soit pas compatible avec la version Analytic Server la plus récente. Contactez votre interlocuteur du support technique IBM si vous rencontrez des problèmes avec le script **configcollector**.

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

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.2.1.1.xxx.zip
```

8. Depuis un shell de commande, exécutez la commande suivante sur le noeud Analytic Server :

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

9. 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**, **retirez des hôtes** et **supprimez** les fichiers parcel d'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.
5. Supprimez les séquelles éventuelles du package d'installation d'Analytic Server. Pour ce faire, supprimez les éléments suivants :
   - Dossier `csd`
   - Tous les fichiers 3.2.2 situés dans les dossiers `parcels`, `parcel-cache` et `parcel-repo`.

Chapitre 4. 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/nom d'hôte du serveur qui héberge Analytic Server.

   Remarque : Vous devez fournir une adresse IP/un nom de domaine hôte approprié lorsque SSL est activé pour Analytic Server.

   as_port
   Port sur lequel Analytic Server est à l'écoute (par défaut, il s'agit du port 9080).

   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 parAnalytic 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 <as-alias> -file <cert-file> -keystore "<cacerts-file>"
      ```

      Notez que <as-alias> 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 5. Configuration du pushback UDF Hive

Après l'enregistrement de Hive UDF dans HiveDB, Analytic Server peut utiliser les fonctions UDF pour exécuter le pushback.

Le pushback UDF Hive est désactivé par défaut et vous devez l'activer manuellement à l'aide du paramètre **udfmodule** dans le fichier ASModules.xml (pour ce faire, remplacez la valeur **disabled** par **enabled**). Une fois le paramètre activé, redémarrez Analytic Server et enregistrez manuellement la fonction UDF dans Hive.

Remarques :

- Lors de l'utilisation d'une source de données Hive sur HDP 3.x, les erreurs suivantes peuvent se produire :

  ```
  Error: The file that you are trying to load does not match the file format of the destination
  table.
  ```

  1. Ouvrez la console Ambari et modifiez la propriété suivante dans la section **Hive** > **Configs** > **Advanced** > **Advanced hive-site**.

     ```
     Key: hive.default.fileformat.managed
     Value: TextFile (change the default value from ORC to TextFile)
     ```

  2. Sauvegardez la configuration.
- Si vous utilisez une source de données Hive dans un environnement non-Kerberos, vérifiez que le nom d'utilisateur entré à la section **Database Selections** est identique à l'utilisateur qui se connecte à Analytic Server.

Les exemples ci-dessous montrent comment enregistrer ou supprimer l'enregistrement d'une fonction UDF dans Hive au sein des environnements HDP et Cloudera.

Enregistrement et suppression de l'enregistrement d'une fonction UDF sur HDP

Enregistrer une fonction UDF

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

Supprimer l'enregistrement d'une fonction UDF

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

Enregistrement et suppression de l'enregistrement d'une fonction UDF sur Cloudera

Enregistrer une fonction UDF

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

Supprimer l'enregistrement d'une fonction UDF

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

Chapitre 6. Utilisation de balises SLM pour le suivi des licences

Les balises SLM sont basées sur l'ébauche de norme ISO/IEC 19770-4 pour Mesure de l'utilisation des ressources. Les balises SLM fournissent à un produit une capacité normalisée pour rendre compte de ses métriques de consommation de licence (ressources associées à l'utilisation d'un actif de type logiciel). Après l'activation de SLM dans un produit, un fichier XML d'exécution est généré pour rendre compte automatiquement de son utilisation de licence.

Au démarrage d'Analytic Server, des fichiers `slmtag` sont créés dans le dossier `<chemin_installation_as>/logs/slmtag`.

Comme il existe deux types de licence, deux métriques différentes sont enregistrées périodiquement :

- Pour la version Analytic Server actuelle, l'octroi de licence est basé sur le nombre total de noeuds dans le cluster Hadoop (basé sur Virtual Server). Le nombre de noeuds est enregistré dans la section suivante du fichier `slmtag`.

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

- Pour les versions Analytic Server antérieures à la version 3.1, l'octroi de licence était basé sur la taille du stockage HDFS dans le cluster Hadoop (basée sur RVU). Par exemple, la taille du stockage (en tegabytes) est enregistrée dans la section suivante du fichier `slmtag`.

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

La sortie de la balise SLM est lancée dans une unité d'exécution et est affectée par les propriétés définies dans le fichier `SlmTagOutput.properties`. Ce fichier est situé dans le dossier `<chemin_installation_as>/configuration`.

Tableau 14. Propriétés de la balise SLM

Propriétés	**Description**
license.metric.logger.output.enabled	Contrôle la génération de fichier journal SLM. Valeur par défaut : `False`.
license.metric.logger.output.dir	Chemin relatif du répertoire qui contient les fichiers de balise SLM. Le répertoire par défaut est `<chemin_installation_as>/ logs`.
license.metric.logger.output. SLMLogFrequency	Intervalle de temps (unité :millisecondes) pour collecte des journaux SLM.
icense.metric.logger.file.size	Taille maximale de fichier de balise SML, en octets.
license.metric.logger.file.number	Nombre maximal de fichiers de balise SLM pour une instance d'identité de logiciel.

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 ne sont pas satisfaisantes, supprimez tous les fichiers `*.war` du chemin de déploiement du service Knox : `/<KnoxServicePath>/data/deployments`. Par exemple : `/usr/hdp/3.1.0.0-78/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 lorsqu'Analytic Server est installé sur un serveur POWER System qui utilise OpenJDK

Lorsqu'Analytic Server opère sur un serveur POWER System qui utilise OpenJDK, vous devez réaliser manuellement les étapes de configuration suivantes pour garantir que l'API de système de coordonnées fonctionne comme prévu.

Remarque : Vous pouvez ignorer cette exigence si vous n'utilisez pas l'API de système de coordonnées.

1. Dans la console Ambari, accédez à **Analytic Server service** > **onglet Configs** > **Advanced analytics-jvm-options** et ajoutez la ligne suivante à la zone de contenu :

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

2. Dans la console Ambari, accédez à la section **Custom analytics.cfg** et ajoutez les 3 configurations suivantes :

   spark.executor.extraJavaOptions
   Définissez sa valeur à : `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

   spark.driver.extraJavaOptions
   Définissez sa valeur à : `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

   mapred.child.java.opts
   Définissez sa valeur à : `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

Erreur lors de l'installation d'Analytic Server sur SuSE Linux 12

Vous pouvez rencontrer l'erreur suivante lors de l'installation d'Analytic Server sur SuSE Linux 12 :

```
Echec de la vérification de signature [clé publique à 4 signatures non disponible]
```

Vous pouvez résoudre le problème en effectuant les tâches suivantes avant d'installer Analytic Server sur SuSE Linux 12 :

1. Téléchargez sur votre hôte une clé publique depuis l'URL suivante :

   ```
   https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
   ```

2. Importez la clé publique en exécutant la commande suivante sur votre hôte :

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

Problèmes concernant des distributions Hadoop spécifiques

L'action d'actualisation pour le service Analytic Server est désactivée sur Hortonworks 2.3-2.6

Pour actualiser manuellement les bibliothèques Analytic Server sur Hortonworks 2.3-2.6, 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/3.2/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.

Propriétés du supergroupe HDFS

Analytic Server consigne une exception au démarrage si `as_user` n'est pas membre des propriétés de groupe HDFS suivantes : **dfs.permissions.supergroup/ dfs.permissions.superusergroup**. Exemple :

```
[11/15/17 7:32:35:510 PST] 000000bf SystemOut
O 2017-11-15 07:32:35,510 | : |  |  |  |  | ERROR | slmtagoutput.SlmOuputAgent | SLM Logger => Error in performing callback function when
calculating number
of nodes in kerberos environment: org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.AccessControlException): Access denied for user
as_user.
Superuser privilege is required
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkSuperuserPrivilege(FSPermissionChecker.java:93)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkSuperuserPrivilege(FSNamesystem.java:6606)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.datanodeReport(FSNamesystem.java:5595)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getDatanodeReport(NameNodeRpcServer.java:928)
at
org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.getDatanodeReport(AuthorizationProviderProxyClientProtocol.java:390)
at
org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getDatanodeReport(ClientNamenodeProtocolServerSideTranslatorPB.java:6
94)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:617)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1073)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2141)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2137)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:415)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1912)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2135)
```

Vous devez ajouter manuellement l'utilisateur `as_user` au groupe OS défini dans les propriétés de configuration `hdfs-site` : **dfs.permissions.supergroup/ dfs.permissions.superusergroup**.

- Pour Cloudera, la valeur par défaut de la propriété est **supergroup** et doit être modifiée en indiquant un groupe OS qui existe réellement. Pour les informations sur le paramètre supergroup dans Cloudera, reportez-vous à la documentation Cloudera.
- Pour Ambari, la valeur par défaut de la propriété est **hdfs**. Par défaut, lors d'une installation d'Ambari, Analytic Server ajoute `as_user` aux groupes HDFS et Hadoop.

Sur Linux, utilisez la commande **usermod** pour ajouter `as_user` au **superusergroup** HDFS (s'il n'existe pas déjà).

Pour des informations générales concernant les autorisations HDFS, reportez-vous au manuel HDFS Permissions Guide.

Echec des travaux MapReduce sur HDP 3.0

Vous pouvez rencontrer l'erreur suivante lors de l'exécution des travaux MapReduce sur HDP 3.0 :

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

L'erreur peut être résolue comme suit :

1. Ajoutez la configuration ci-dessous au fichier `Custom analytics.cfg` :

```
exclude.mapreduce.jars=icu4j-
```

2. Redémarrez Analytic Server.

Une fois Analytic Server redémarré, les travaux MapReduce s'exécutent normalement.

Echec de l'écriture des valeurs de date ou d'horodatage dans les tables Hive en raison d'un problème Cloudera

Lorsque Analytic Server tente d'écrire des valeurs de date ou d'horodatage dans des tables Hive dans un environnement Cloudera, le processus échoue en raison d'un problème Cloudera connu (https://issues.apache.org/jira/browse/HIVE-11024.

Remarque : Le problème de valeur de date n'affecte pas Hive 1.3.0 ou 2.0.0 (https://issues.apache.org/jira/browse/HIVE-11024) ; le problème de valeur d'horodatage n'affecte pas Hive 2.0.0 (https://issues.apache.org/jira/browse/HIVE-11748?jql=project%20%3D%20HIVE%20AND%20text%20~%20%22jdbc%20timestamp%22). Vous devez vous assurer que votre environnement Cloudera comporte une version Hive prise en charge (1.3.0 ou 2.0.0).

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 liés à Apache Spark

Problèmes affectant les flux exécutés au sein d'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).

L'erreur "Erreur pendant HdfsAuthcom.spss.utilities.i18n.LocException: L'exécution a échoué. Motif : com.spss.ae.filesystem.exception.FileSystemException: Impossible d'initialiser l'accès au système de fichiers." se produit lors de l'exécution des cas SparkML.

L'erreur est générée lorsque Spark ne trouve pas le répertoire des journaux de lignée. Une solution de contournement consiste à rediriger `spark.lineage.log.dir` vers `/ae_wlpserver/usr/servers/aeserver/logs/spark`.

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 30.

"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`. 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.

Hive pushback

Il est possible que vous receviez le message d'erreur suivant lorsque le pushback Hive est activé :

```
(AEQAE2103E) SQL execution failed - Error while compiling statement:
FAILED: SemanticException [Error 10014]: Line 3:47 Wrong arguments '9223372036854775808':
```

```
Unsafe compares between different types are disabled for safety reasons. If you know what
you are doing, please set hive.strict.checks.type.safety to false and make sure that
hive.mapred.mode is not set to 'strict' to proceed. Note that you may get errors or
incorrect results if you make a mistake while using some of the unsafe features. (as_trace.log)
```

L'erreur peut être résolue en employant l'une des méthodes suivantes :

- Ajoutez **hive.sql.check=true** au fichier Analytic Server `config.properties`.
- Définissez le paramètre **hive.strict.checks.type.safety** dans la configuration Hive sur **false**.

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 IBM 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

Pour toute demande au sujet des licences concernant les produits utilisant un jeu de caractères codé sur deux octets, contactez le service de propriété intellectuelle IBM de votre pays ou envoyez vos questions par courrier 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 du document intitulé IBM Customer Agreement, 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. Aucune réclamation relative à des produits non IBM ne pourra être reçue par 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 :

Marques

IBM, le logo IBM et 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.

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 licence.

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.