IBM SPSS Analytic Server
バージョン 3.2.1

インストールと構成のガイド

IBM

注記

本書および本書で紹介する製品をご使用になる前に、 83 ページの『特記事項』に記載されている情報をお読みください。

本書は、IBM SPSS Analytic Server バージョン 3、リリース 2、モディフィケーション 1、および新しい版で明記されていない限り、以降のすべてのリリースおよびモディフィケーションに適用されます。

お客様の環境によっては、資料中の円記号がバックスラッシュと表示されたり、バックスラッシュが円記号と表示されたりする場合があります。

原典： IBM SPSS Analytic Server
Version 3.2.1
Installation and Configuration Guide

発行： 日本アイ・ビー・エム株式会社

担当： トランスレーション・サービス・センター

目次

第 1 章 前提条件

Analytic Server をインストールする前に、以下の情報を確認してください。

システム要件

最新のシステム要件情報については、IBM Technical Support サイトの http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html にある「Detailed system requirements」レポートを使用してください。このページでは以下を行うことができます。

1. 製品名として SPSS Analytic Server を入力して、「**Search**」をクリックします。
2. 目的のバージョンとレポートの範囲を選択して、「**Submit**」をクリックします。

WebSocket トラフィック

クライアントと Analytic Server の間の WebSocket トラフィックがファイアウォール、VPN、またはその他のポート・ブロッキング方法によってブロックされていないことを確認する必要があります。WebSocket ポートは、一般的な Analytic Server ポートと同じです。

SuSE Linux (SLES) 12

SuSE Linux 12 に Analytic Server をインストールする前に、以下のタスクを実行します。

1. 以下の URL から、ご使用のホストに公開鍵をダウンロードします。

   ```
   https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.1.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
   ```

2. ご使用のホストで以下のコマンドを実行して、公開鍵をインポートします。

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

Power Systems

クラスター内のすべてのホストに IBM XLC コンパイラーおよび XLF コンパイラーがインストールされており、PATH に含まれていることを確認してください。

これらのコンパイラー用のライセンスの取得の詳細については、以下の Web サイトで確認できます。

- 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)

Analytic Server をインストールする前に、1 つ以上の HDP クライアントがクラスター環境にデプロイされていることを確認する必要があります。Ambari Manager をホストするノードでは `/usr/hdp` ディレクトリーが想定されるため、HDP クライアントが存在しない場合は Analytic Server に障害が発生します。

Hive/HCatalog

NoSQL データ・ソースを使用する予定の場合、Hive および HCatalog をリモート・アクセス用に構成します。さらに、`hive-site.xml` に、アクティブな Thrift Hive Metastore サーバーを示す *hive.metastore.uris* プロパティーが `thrift://<host_name>:<port>` の形式で設定されていることを確認します。詳しくは、使用している Hadoop ディストリビューションの資料を参照してください。

注**:** Analytic Server Metastore は、Hive Metastore と同じマシンにインストールすることはできません。

Hive 2.1 を使用する場合は、Ambari コンソールで「**Interactive Query**」設定を有効にして Hive 2.1 を有効化してから、Analytic Server インストール時に `hive.version` プロパティーとして `2.x` と入力する必要があります。

1. Ambari コンソールを開き、「**Analytic Server Advanced analytics.cfg**」セクションで以下のプロパティーを追加します。
   - キー: `hive.version`
   - 値: 適切な Hive バージョン (2.x など) を入力します。
2. 構成を保存します。

注**:** Hive 2.1 は、HDP 2.5 以降でサポートされています (Spark 2.x を使用している場合)。

メタデータ・リポジトリー

デフォルトでは、Analytic Server は MySQL データベースをインストールして使用します。あるいは、既存の Db2 インストール済み環境を使用するように Analytic Server を構成することもできます。選択するデータベースのタイプにかかわらず、データベースには UTF-8 のエンコードが必要です。

MySQL

MySQL のデフォルト文字セットはバージョンとオペレーティング・システムによって異なります。ご使用の MySQL インストール済み環境が UTF-8 に設定されているかどうかを確認するには、以下の手順を使用してください。

1. MySQL のバージョンを確認します。

   ```
   mysql -V
   ```

2. MySQL のコマンド・ライン・インターフェースから以下の照会を実行して、MySQL のデフォルト文字セットを確認します。

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

   文字セットが既に UTF-8 に設定されている場合、追加の変更は不要です。

3. MySQL のコマンド・ライン・インターフェースから以下の照会を実行して、MySQL のデフォルト照合を確認します。

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

   照合が既に UTF-8 に設定されている場合、追加の変更は不要です。

4. デフォルトの文字セットまたは照合が UTF-8 でない場合、文字セットを UTF-8 に変更するために /etc/my.cnf を編集して MySQL デーモンを再始動する方法について、MySQL の資料を参照してください。

Db2 Db2 の構成について詳しくは、Knowledge Center (http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html) を参照してください。

高可用性クラスター

ロード・バランサー

高可用性クラスターには、セッション・アフィニティー (スティッキー・セッションと呼ばれることもあります) をサポートするロード・バランサーが必要です。Analytic Server は、Cookie「request-token」でセッションを識別します。これにより、アプリケーションによって制御されるセッション・アフィニティーで使用するために、ユーザー・ログイン

の期間にわたってセッションが識別されます。セッション・アフィニティーがどのようにサポートされるかについて詳しくは、ご使用の特定のロード・バランサーの資料を参照してください。

Analytic Server ジョブの失敗

クラスター・メンバーの障害が原因で Analytic Server ジョブが失敗した場合、通常そのジョブは他のクラスター・メンバー上で自動的に再開されます。ジョブが再開されない場合、高可用性クラスター内に少なくとも 4 つのクラスター・メンバーが存在することを確認してください。

第 2 章 Ambari のインストールおよび構成

Ambari 固有の前提条件

一般的な前提条件に加えて、以下の情報を確認してください。

サービス

Analytic Server は Ambari サービスとしてインストールされます。Analytic Server をインストールする前に、以下のクライアントが Ambari サービスとしてインストールされていることを確認する必要があります。

- HDFS/HDFS_CLIENT
- MAPREDUCE2/MAPREDUCE2_CLIENT
- HIVE/HIVE_CLIENT
- SPARK/SPARK_CLIENT (Spark 1.x が使用される場合)
- SPARK2/SPARK2_CLIENT (Spark 2.x が使用される場合)
- HBASE/HBASE_CLIENT (HBASE が使用される場合)
- YARN
- Zookeeper

パスワード無し **SSH**

root ユーザー用に Analytic Metastore ホストとクラスター内のすべてのホストの間にパスワード無し SSH をセットアップしてください。

インストールの事前チェック・ツールと事後チェック・ツール - Ambari

事前チェック・ツールの概要

Analytic Server インストールの事前チェック・ツールは、Analytic Server インストールの前に潜在的な環境問題を特定することにより、インストール問題と実行時エラーを減らすために役立ちます。

事前チェック・ツールは、以下の項目を検証します。

- ローカル・システム上の OS と Ambari のバージョン
- ローカル・システム上の OS の ulimit 設定
- ローカル・システム上の使用可能なディスク・スペース
- Hadoop バージョン
- Ambari サービスの可用性 (HDFS、HCatalog、Spark、Hive、MapReduce、Yarn、Zookeeper など)
- Analytic Server 固有の Ambari 設定

注**:** 事前チェック・ツールは、自己解凍型 Analytic Server バイナリー・ファイルを実行した後で使用できます。

事後チェック・ツールの概要

Analytic Server インストールの事後チェック・ツールは、Analytic Server インストールの後で、以下を処理するための REST API 要求をサブミットすることにより、構成問題を特定します。

- HDFS 内のデータ
- Hive/HCatalog 内のデータ
- 圧縮データ (deflate、bz2、snappy を含む)
- PySpark でのデータ
- ネイティブ SPSS コンポーネントを使用するデータ (alm、tree、neuralnet、scoring、tascoring を含む)
- MapReduce でのデータ
- メモリー内の MapReduce でのデータ

ツールの場所と前提条件

Analytic Server サービスをインストールする前に、Analytic Server サービスの一部となるすべてのノード上で事前チェック・ツールを実行し、Linux 環境に Analytic Server をインストールする準備が整っているか確認します。

事前チェックツールは、インストールの一部として自動的に起動されます。このツールは、各ホスト上でインストールを実行する前に、Analytic Metastore および各 Analytic Server ノードをチェックします。Ambari Server ノードで事前チェック・ツールを手動で起動することもできます。これにより、サービスをインストールする前にマシンが検証されます。

自己解凍型 Analytic Server バイナリー・ファイルを実行した後で、事前チェック・ツールは、以下のディレクトリーにあります。

- **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
```

Analytic Server のインストール後に、事後チェック・ツールは以下のディレクトリーにあります。

- **HDP**

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

ツールは root として実行する必要があり、Python 2.6.X 以上を必要とします。

事前チェック・ツールが失敗を報告した場合は、Analytic Server インストールを続行する前に、それらの失敗に対処する必要があります。

`chktool` ディレクトリーは、Analytic Server 自己解凍型バイナリーの実行 (8 ページの『Ambari でのインストール』セクションのステップ 2) の後で使用可能になります。 12 ページの『オフライン・インストール』の実行を選択した場合、`chktool` ディレクトリーは、メタデータ RPM のインストールの後で使用可能になります。

事前チェック・ツールの実行

自動

事前チェックツールは、Ambari コンソールを介して Analytic Server をインストールするときに、Analytic Server インストールの一部として自動的に起動できます。Ambari サーバーのユーザー名とパスワードを手動で入力する必要があります。

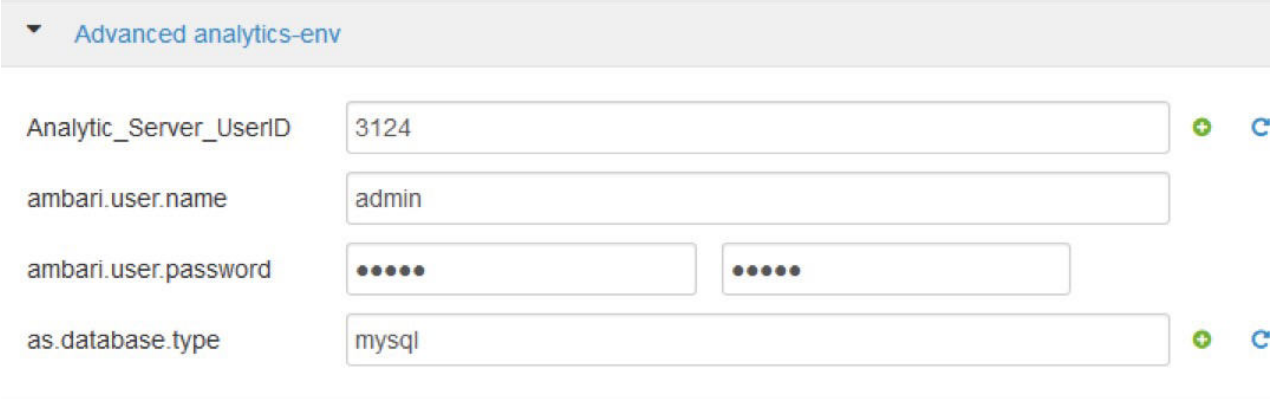

図 1. 「*Advanced analytics-env*」設定

手動

Ambari Server ノードで事前チェック・ツールを手動で起動できます。

以下の事前チェックの例は、Ambari クラスター `MyCluster` をチェックします。このクラスターは、`myambarihost.ibm.com:8080` で実行され、SSL が有効になっていて、ログイン資格情報 `admin:admin` を使用します。

```
python ./precheck.py --target B --cluster MyCluster --username admin
--password admin --host myambarihost.ibm.com --port 8080 --as_host myashost.ibm.com --ssl
```

注:

- `as_host` 値は、IP アドレスまたは完全修飾ドメイン名のいずれかによって指定する必要があります。
- パスワード引数が省略されると、ツールはパスワードの入力を求めるプロンプトを出します。
- `precheck.py` コマンドに含まれている使用法ヘルプは、`-h` 引数 (`python ./precheck.py -help`) と表示されます。
- `--cluster` 引数はオプションです (`--cluster` が使用されていない場合は、現在のクラスターが指定されます)。

事前チェック・ツールがチェックを実行しているときには、各チェックの状況がコマンド・ウィンドウに表示されます。失敗が発生した場合は、ログ・ファイル内の詳細情報を参照できます (ログ・ファイルの具体的な場所は、コマンド・ウィンドウで指示されます)。追加のサポートが必要な場合は、ログ・ファイルを IBM Technical Support に提供できます。

事後チェック・ツールの実行

事後チェック・ツールは、Analytic Server が適切に実行されていること、および単純なジョブを処理できることを検証します。以下の事後チェックの例は、特定の Analytic Server インスタンスをチェックします。このインスタンスは、`myanalyticserverhost.ibm.com:9443` で実行され、SSL が有効になっていて、ログイン資格情報 `admin:ibmspss` を使用します。

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

Knox が Analytic Server と共に使用される場合、コマンドは以下のとおりです。

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

単一のチェックを実行するには、以下のコマンドを使用します。

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

注:

- パスワード引数が省略されると、ツールはパスワードの入力を求めるプロンプトを出します。
- `postcheck.py` コマンドに含まれている使用法ヘルプは、`--h` 引数 (`python ./postcheck.py --help`) と表示されます。

事後チェック・ツールがチェックを実行しているときには、各チェックの状況がコマンド・ウィンドウに表示されます。失敗が発生した場合は、ログ・ファイル内の詳細情報を参照できます (ログ・ファイルの具体的な場所は、コマンド・ウィンドウで指示されます)。追加のサポートが必要な場合は、ログ・ファイルを IBM Technical Support に提供できます。

Ambari でのインストール

基本的なプロセスは、次のとおりです。Analytic Server ファイルを Ambari クラスター内のホストにインストールし、続いて Analytic Server を Ambari サービスとして追加します。

『オンライン・インストール』
: Ambari サーバー・ホストおよびクラスター内のすべてのノードが https://ibm-open-platform.ibm.com にアクセス可能な場合は、オンライン・インストールを選択してください。

12 ページの『オフライン・インストール』
: ご使用の Ambari サーバー・ホストがインターネットにアクセスできない場合は、オフラインを選択します。

オンライン・インストール

Ambari サーバー・ホストおよびクラスター内のすべてのノードが https://ibm-open-platform.ibm.com にアクセス可能な場合は、オンライン・インストールを選択してください。

1. IBM パスポート・アドバンテージ Web サイトに移動し、ご使用のスタック、スタック・バージョン、およびハードウェア・アーキテクチャーに固有の自己解凍型バイナリー・ファイルを Ambari Manager ノードにダウンロードします。使用可能な Ambari バイナリーは以下のとおりです。

表 *1. Analytic Server* 自己解凍型バイナリー・ファイル

説明	バイナリー・ファイル名
IBM® SPSS® Analytic Server 3.2.1 for Hortonworks Data Platform 2.5、2.6、3.0、および 3.1 Linux x86-64 英語	`spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.1 for Hortonworks Data Platform 2.6、3.0、および 3.1 Linux on System p LE 英語	`spss_as-3.2.1.0-hdp2.6-3.1-lppc64.bin`

2. 自己解凍型バイナリー・ファイルを実行し、指示に従ってライセンスを表示し、ライセンスを受け入れて、オンライン・インストールを選択し、Analytic Server が使用するデータベース・タイプのインストール・プロセスを選択します。以下のデータベース・タイプ・オプションが提供されています。
   - 新規 MySQL インスタンス
   - 既存の MySQL インスタンスまたは Db2 インスタンス

3. /var/lib/ambari-server/resources/stacks/<stack-name>/<stack-version>/services/ANALYTICSERVER/package/scripts ディレクトリーから、update_clientdeps.sh スクリプトを適切な引数を指定 (例えば、--help 引数を使用) して実行します。
4. Ambari サーバーを再起動します。

   ```
   ambari-server restart
   ```
5. Ambari サーバーにログオンし、Ambari UI を使用して Analytic Server をサービスとしてインストールします。

   メタデータ・リポジトリー

   デフォルトでは、Analytic Server は MySQL を使用して、データ・ソース、プロジェクト、およびテナントに関する情報を追跡します。インストール時に、Analytic Server と MySQL の間の JDBC 接続で使用されるユーザー名 (**metadata.repository.user.name**) およびパスワード (**metadata.repository.password**) を指定する必要があります。インストーラーは MySQL データベースにそのユーザーを作成しますが、そのユーザーは MySQL データベースに固有であり、既存の Linux ユーザーや Hadoop ユーザーである必要はありません。

   注**:** インストール中に新しい MySQL インスタンスをインストールする場合、MySQL がインストールされていないマシンに Analytic Server Metastore をインストールする必要があります。

   メタデータ・リポジトリーを Db2 に変更するには、以下のステップを実行します。

   注**:** インストールの完了後にメタデータのリポジトリーを変更することはできません。

   a. 別のマシンに Db2 がインストールされていることを確認します。詳しくは、トピック 1 ページの『第 1 章 前提条件』 の『メタデータ・リポジトリー』セクションを参照してください。
   b. Ambari の「Services」タブで、Analytic Server サービスの「Configs」タブに移動します。
   c. 「**Advanced analytics-env**」セクションを開きます。
   d. **as.database.type** の値を mysql から db2 に変更します。
   e. 「**Advanced analytics-meta**」セクションを開きます。
   f. **metadata.repository.driver** の値を com.mysql.jdbc.Driver から com.ibm.db2.jcc.DB2Driver に変更します。
   g. **metadata.repository.url** の値を jdbc:db2://{Db2_HOST}:{PORT}/{DBName}:currentSchema={SchemaName}; に変更します。ここで、
      - {Db2_HOST} は、Db2 がインストールされているサーバーのホスト名です。
      - {PORT} は、Db2 が listen しているポートです。
      - {SchemaName} は、使用可能な、未使用のスキーマです。

      入力する値がわからない場合は、Db2 管理者に協力を求めてください。
   h. **metadata.repository.user.name** および **metadata.repository.password** に、有効な Db2 資格情報を入力します。
   i. 「**Save**」をクリックします。

   LDAP 構成

   Analytic Server は、LDAP サーバーを使用して、ユーザーおよびグループを保管および認証します。必要な LDAP 構成情報を Analytic Server のインストール中に指定します。

表 2. *LDAP* 構成設定

LDAP 設定	説明
as.ldap.type	LDAP タイプ。値は、ads、ad、または openladp にすることができます。 • ads - Apache Directory Server (デフォルト設定) • ad - Microsoft Active Directory • openladp - OpenLDAP
as.ldap.host	LDAP ホスト
as.ldap.port	LDAP ポート番号
as.ldap.binddn	LDAP バインド DN
as.ldap.bindpassword	LDAP バインド DN パスワード
as.ldap.basedn	LDAP ベース DN
as.ldap.filter	LDAP ユーザーおよびグループのフィルター・ルール
as.ldap.ssl.enabled	Analytic Server と LDAP の間の通信に SSL を使用するかどうかを指定します。値は true または false にすることができます。
as.ldap.ssl.reference	LDAP SSL 参照 ID
as.ldap.ssl.content	LDAP SSL 構成

- デフォルトでは、as.ldap.type は ads に設定され、その他の関連設定にはデフォルト設定が含まれます。ただし、例外として、as.ldap.bindpassword 設定のパスワードは独自に指定する必要があります。Analytic Server は、構成設定を使用して、Apache Directory Server (ADS) をインストールし、サーバーの初期化を実行します。デフォルトの ADS プロファイルには、admin というパスワードを持つユーザー admin が含まれています。Analytic Server コンソールを使用してユーザー管理を実行するか、<Analytic Root>/bin フォルダー内にある importUser.sh スクリプトを使用して XML ファイルからユーザーおよびグループの情報をインポートすることができます。
- Microsoft Active Directory や OpenLDAP などの外部 LDAP サーバーを使用する予定がある場合は、実際の LDAP 値に従って構成設定を定義する必要があります。詳しくは、『Liberty での LDAP ユーザー・レジストリーの構成』を参照してください。
- Analytic Server がインストールされた後で LDAP 構成を変更できます (例えば、Apache Directory Server から OpenLDAP に変更します)。ただし、最初に Microsoft Active Directory または OpenLDAP で開始してから、後で Apache Directory Server に切り替えることを決定した場合は、Analytic Server がインストール中に Apache Directory Server をインストールすることはありません。Apache Directory Server は、それが最初の Analytic Server のインストール中に選択されていた場合にのみインストールされます。

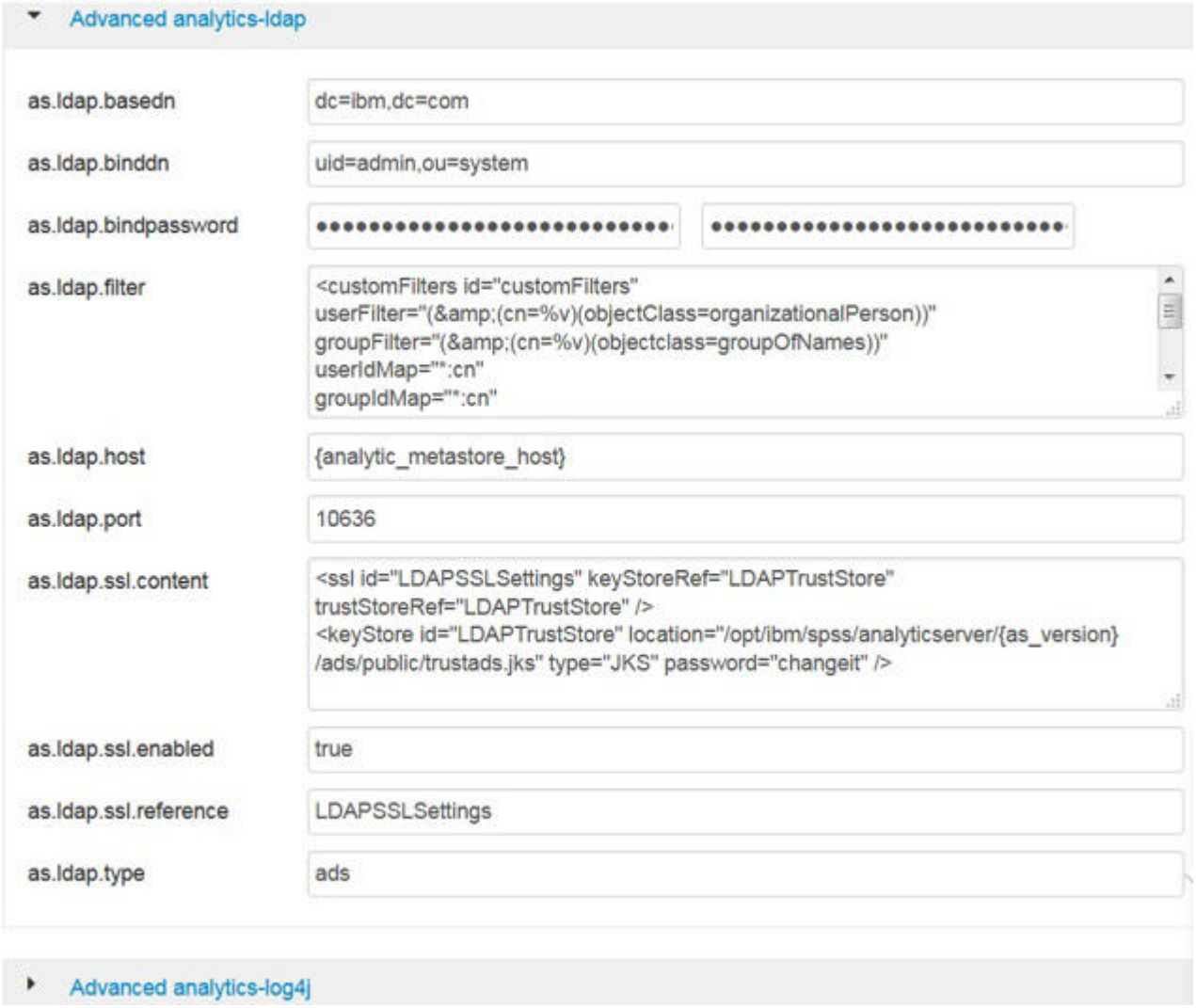

図 2. *LDAP* 構成設定の例

インストール後に変更してはならない構成設定

インストール後に以下の設定は変更しないでください。変更すると Analytic Server が動作しなくなります。

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

6. これで、Analytic Server のインスタンスが機能するようになりました。追加の構成はオプションです。Analytic Server の構成と管理について詳しくは、 18 ページの『構成』 のトピックを参照してください。既存構成の新規インストール済み環境へのマイグレーションについては、 37 ページの『Ambari での IBM SPSS Analytic Server のマイグレーション』のトピックを参照してください。
7. Web ブラウザーを開き、アドレス `http://<host>:<port>/analyticserver/admin/ibm` を入力します。ここで、<host> は Analytic Server ホストのアドレスであり、<port> は Analytic Server が listen している ポートです。デフォルトではこれは 9080 です。この URL にアクセスすると、Analytic

Server コンソールのログイン・ダイアログが開きます。 Analytic Server 管理者としてログインします。デフォルトでは、このユーザー ID は admin であり、パスワードは admin です。

オフライン・インストール

IBM SPSS Analytic Server オフライン・インストールは、自動的に実行することも、手動で実行することもできます。

『**HDP** での自動インストール』

自動インストール・プロセスでは Ambari REST API が使用されます。この方法によるインストールが推奨されています。

13 ページの『**HDP (RHEL、SLES)** での手動インストール』
Analytic Server を Hortonworks Data Platform に手動でインストールする場合

16 ページの『**HDP (Ubuntu)** での手動インストール』
Analytic Server を Ubuntu Linux に手動でインストールする場合

HDP での自動インストール

自動インストール・プロセスでは Ambari REST API が使用されます。この方法によるインストールが推奨されています。

重要**:**

- オフライン自動インストール手順では、組み込み Apache Directory Server (ADS) をインストールします。サード・パーティー LDAP サーバーを使用する場合は、IBM SPSS Analytic Server のインストールが完了した後で、LDAP 設定を構成できます。
- オフライン自動インストール手順では、単一の Analytic Server サービス・インスタンスのみをインストールできます。最初のインストールが完了した後で、さらにインスタンスを追加できます。
- オフライン自動インストール手順では、Kerberos が有効になっているクラスターへの Analytic Server のインストールはサポートされません。
- オフライン自動インストール手順では、HDP 3.0 または 3.1 への Analytic Server のインストールはサポートされません。

これらの制限は、HDP または Ubuntu での手動インストールには適用されません。

1. IBM パスポート・アドバンテージ Web サイトに移動し、自己解凍型バイナリー・ファイルを、https://ibm-open-platform.ibm.com にアクセス可能なコンピューターにダウンロードします。

表 *3. Analytic Server* 自己解凍型バイナリー・ファイル

説明	バイナリー・ファイル名
IBM SPSS Analytic Server 3.2.1 for Hortonworks Data Platform 2.5、2.6、3.0、および 3.1 Linux x86-64 英語	`spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin`
IBM SPSS Analytic Server 3.2.1 for Hortonworks Data Platform 2.6、3.0、および 3.1 Linux on System p LE 英語	`spss_as-3.2.1.0-hdp2.6-3.1-lppc64.bin`

2. ステップ 1 でダウンロードした実行可能バイナリーを実行し、オフライン・インストールを指定します。オフライン・インストールでは、インストール・プロセスの後続のステップで必要になる RPM ファイルまたは DEB ファイルをダウンロードします。そのため、オフライン・インストールは、

https://ibm-open-platform.ibm.com にアクセス可能なコンピューターで実行する必要があります。ダウンロードされたファイルは、現行の実行可能バイナリー・ディレクトリー ./IBM-SPSS-AnalyticServer にあります。

3. 実行可能バイナリー・ディレクトリー ./IBM-SPSS-AnalyticServer のすべての内容を、インターネットにアクセスできるマシンから、(ファイアウォールで保護されている) Ambari Manager ノードにコピーします。
4. Ambari Manager ノードで、以下のコマンドを使用して、Ambari サーバーが実行されているかどうかを確認します。

   ```
   ambari-server status
   ```

5. Ambari Manager ノード、および Analytic Server をデプロイする他のすべてのノードで、ローカル yum リポジトリーを作成するツールをインストールします。

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

   または

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

6. Ambari Manager ノードで、実行可能バイナリー・ファイル ./IBM-SPSS-AnalyticServer/packages/spss_as-ambari-offlineinstall.bin を実行します。インストール中に、実行可能バイナリーは、必要な Analytic Server RPM/DEB ファイルが packages ディレクトリーにあることを確認します。必要な RPM ファイルは、ご使用のディストリビューション、バージョン、およびアーキテクチャーによって異なります。

   HDP 2.5、2.6、3.0、および 3.1 (x86_64)
   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.x86_64.rpm

   HDP 2.6、3.0、および 3.1 (PPC64LE)
   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.ppc64le.rpm

   HDP 2.5、2.6、3.0、および 3.1 (Ubuntu)
   IBM-SPSS-AnalyticServer-ambari-2.x_3.2.1.0_amd64.deb

   IBM-SPSS-AnalyticServer_1_amd64.deb

インストール中に、Analytic Server バージョン、JDBC ドライバー、Spark バージョン、Hive バージョンなどの入力を求めるプロンプトが出されます。

HDP (RHEL、SLES) での手動インストール

HDP (RHEL、SLES) での手動オフライン・インストールの一般的なワークフローは、以下のとおりです。

1. IBM パスポート・アドバンテージ Web サイトに移動し、自己解凍型バイナリー・ファイルを、https://ibm-open-platform.ibm.com にアクセス可能なコンピューターにダウンロードします。

表 *4. Analytic Server* 自己解凍型バイナリー・ファイル

説明	バイナリー・ファイル名
IBM SPSS Analytic Server 3.2.1 for Hortonworks Data Platform 2.5、2.6、3.0、および 3.1 Linux x86-64 英語	spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin
IBM SPSS Analytic Server 3.2.1 for Hortonworks Data Platform 2.6、3.0、および 3.1 Linux on System p LE 英語	spss_as-3.2.1.0-hdp2.6-3.1-lppc64.bin

2. ステップ 1 でダウンロードした実行可能バイナリーを実行し、オフライン・インストールを指定します。オフライン・インストールでは、インストール・プロセスの後続のステップで必要になる RPM ファイルをダウンロードします。そのため、オフライン・インストールは、https://ibm-open-platform.ibm.com にアクセス可能なコンピューターで実行する必要があります。ダウンロードされたファイルは、現行の実行可能バイナリー・ディレクトリー ./IBM-SPSS-AnalyticServer にあります。
3. 実行可能バイナリー・ディレクトリー ./IBM-SPSS-AnalyticServer のすべての内容を、インターネットにアクセスできるマシンから、Ambari Manager ノードの <AS_INSTALLABLE_HOME> ディレクトリーにコピーします (Ambari Manager ノードはファイアウォールで保護されています)。
4. Ambari Manager ノードで、以下のコマンドを使用して、Ambari サーバーが実行されているかどうかを確認します。

   ```
   ambari-server status
   ```

5. ローカル yum リポジトリーを作成するツールをインストールします。

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

   または

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

6. Analytic Server の RPM ファイルのリポジトリーとして機能するディレクトリーを作成します。以下の例を参照してください。

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

7. 新規ディレクトリーに、必要な Analytic Server の RPM ファイルをコピーします。必要な RPM ファイルは、ご使用のディストリビューション、バージョン、およびアーキテクチャーによって異なります。

   HDP 2.5、2.6、3.0、および 3.1 (x86_64)

   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.x86_64.rpm

   HDP 2.6、3.0、および 3.1 (PPC64LE)

   IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

   IBM-SPSS-AnalyticServer-3.2.1.0-1.ppc64le.rpm

8. ローカル・リポジトリーの定義を作成します。例えば、IBM-SPSS-AnalyticServer-3.2.1.0.repo というファイルを、/etc/yum.repos.d/ (RHEL、CentOS の場合) または /etc/zypp/repos.d/ (SLES の場合) に、以下の内容を指定して作成します。

   ```
   [IBM-SPSS-AnalyticServer]
   name=IBM-SPSS-AnalyticServer
   baseurl=file:///{path to local repository}
   enabled=1
   gpgcheck=0
   protect=1
   ```

9. ローカル yum リポジトリーを作成します。

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

10. root ユーザー・コマンド・ウィンドウから、cd を実行して <AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer ディレクトリーに移動し、./offLineInstall.sh を実行します。スクリプトは、以前に実行されたバイナリー実行可能インストール・コマンドに対する永続化された応答を読み取り、(rpm をインストールするための) 該当するプラットフォーム・コマンドを発行します。

    注: ステップ 11 は、外部管理される MySQL 環境を使用する場合にのみ適用されます。

11. AS_MetaStore として使用される MySQL インスタンスがインストールされているノード/ホストで add_mysql_user.sh スクリプトを実行します。
    a. add_mysql_user.sh スクリプトを、<AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer から、AS_MetaStore として使用される MySQL インスタンスがインストールされているノード/ホストにコピーします。
        - MySQL ノード/ホスト上で add_mysql_user.sh スクリプトを実行します。例: ./add_mysql_user.sh -u as_user -p spss -d aedb

        注:
        - ユーザー名およびパスワードは、Ambari 構成画面の AS_Metastore で入力されたデータベース・ユーザー名およびパスワードと一致する必要があります。
        - コマンドを発行するように add_mysql_user.sh スクリプトを手動で更新できます (希望する場合)。
        - セキュアな (root ユーザーによってアクセスされる) MySQL データベースに対して add_mysql_user.sh スクリプトを実行する場合は、-r パラメーターおよび -t パラメーターを使用して、dbuserid および dbuserid_password を渡します。スクリプトは、dbuserid および dbuserid_password を使用して、MySQL 操作を実行します。

    注: 「**AS_Configuration**」画面の metadata.repository.url 設定 (「**Advanced analytics-meta**」) が MySQL データベース・ホストを指すように変更する必要があります。例えば、JDBC 設定 mysql://{analytic_metastore_host}/aedb?createDatabaseIfNotExist=true を mysql://{MySQL_DB}/aedb?createDatabaseIfNotExist=true に変更します。
12. ご使用の Ambari リポジトリー・ファイル repoinfo.xml (通常は /var/lib/ambari-server/resources/stacks/$stackName/$stackVersion/repos/ に配置されています) に以下の行を追加して、ローカル yum リポジトリーを使用するように更新します。

```
<os type="host_os">
   <repo>
        <baseurl>file:///{path to local repository}/</baseurl>
        <repoid>IBM-SPSS-AnalyticServer</repoid>
        <reponame>IBM-SPSS-AnalyticServer-3.2.1.0</reponame>
   </repo>
</os>
```

    例の中の {path to local repository} は次のようになります。

    /home/root/repos/IBM-SPSS-AnalyticServer/x86_64/
13. 各 Ambari 非サーバー・クラスター・ノードに対して以下の手順を繰り返します。
    a. 該当する <AS_INSTALLABLE_HOME> ディレクトリーのすべての内容を、インターネットにアクセスできるマシンから、Ambari 非サーバー・クラスター・ノードにコピーします。
    b. ローカル yum リポジトリーを作成するツールをインストールします。

        yum install createrepo (RHEL, CentOS)

        または

        zypper install createrepo (SLES)
    c. Analytic Server の RPM ファイルのリポジトリーとして機能するディレクトリーを作成します。以下の例を参照してください。

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

d. 新規ディレクトリーに、必要な Analytic Server の RPM ファイルをコピーします。必要な RPM ファイルは、ご使用のディストリビューション、バージョン、およびアーキテクチャーによって異なります。

HDP 2.5、2.6、3.0、および 3.1 (x86_64)

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.1.0-1.x86_64.rpm

HDP 2.6、3.0、および 3.1 (PPC64LE)

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.1.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.1.0-1.ppc64le.rpm

e. ローカル・リポジトリーの定義を作成します。例えば、IBM-SPSS-AnalyticServer-3.2.1.0.repo というファイルを、/etc/yum.repos.d/ (RHEL、CentOS の場合) または /etc/zypp/repos.d/ (SLES の場合) に、以下の内容を指定して作成します。

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer
baseurl=file:///{path to local repository}
enabled=1
gpgcheck=0
protect=1
```

f. ローカル yum リポジトリーを作成します。

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

14. 8 ページの『オンライン・インストール』のセクションのステップ 3 に進みます。

HDP (Ubuntu) での手動インストール

HDP (Ubuntu) での手動オフライン・インストールの一般的なワークフローは、以下のとおりです。

1. IBM パスポート・アドバンテージ Web サイトに移動し、適切な Ubuntu 自己解凍型バイナリー・ファイルを、https://ibm-open-platform.ibm.com にアクセス可能なコンピューターにダウンロードします。

表 5. *Analytic Server* 自己解凍型バイナリー・ファイル

説明	バイナリー・ファイル名
IBM SPSS Analytic Server 3.2.1 for Hortonworks Data Platform 2.5、2.6、3.0、および 3.1 Linux x86-64 英語	spss_as-3.2.1.0-hdp2.5-3.1-lx86.bin

2. ステップ 1 でダウンロードした実行可能バイナリーを実行し、オフライン・インストールを指定します。オフライン・インストールでは、インストール・プロセスの後続のステップで必要になる DEB ファイルをダウンロードします。そのため、オフライン・インストールは、https://ibm-open-platform.ibm.com にアクセス可能なコンピューターで実行する必要があります。ダウンロードされたファイルは、現行の実行可能バイナリー・ディレクトリー ./IBM-SPSS-AnalyticServer にあります。
3. 実行可能バイナリー・ディレクトリー ./IBM-SPSS-AnalyticServer のすべての内容を、インターネットにアクセスできるマシンから、Ambari Manager ノードの <AS_INSTALLABLE_HOME> ディレクトリーにコピーします (Ambari Manager ノードはファイアウォールで保護されています)。
4. Ambari Manager ノードで、以下のコマンドを使用して、Ambari サーバーが実行されているかどうかを確認します。

```
ambari-server status
```

5. Analytic Server の DEB ファイルのリポジトリーとして機能する <local_repo> ディレクトリーを作成します。以下に例を示します。

```
mkdir –p /usr/local/mydebs
```

6. <local_repo> ディレクトリーに、必要な Analytic Server の DEB ファイルをコピーします。
   - IBM-SPSS-AnalyticServer-ambari-2.x_3.2.1.0_amd64.deb
   - IBM-SPSS-AnalyticServer_1_amd64.deb
7. ローカル・リポジトリーを作成します。
   a. ローカル・リポジトリーを作成するツールをインストールします。

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

   b. 以下のようにソース・パッケージ・ファイルを生成します。

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

   c. ローカル・リポジトリーのコンポーネント (メイン) およびアーキテクチャー (例えば、binary-i386、binary-amd64) を作成します。

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

   d. 以下のようにソース・パッケージをコピーします。

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

8. ローカル・リポジトリーの定義を作成します。例えば、IBM-SPSS-AnalyticServer-3.2.1.0.list というファイルを、/etc/apt/sources.list.d に、以下の内容を指定して作成します。

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

9. 以下のコマンドを実行して、リポジトリー・リストを更新します。

   ```
   apt-get update
   ```

10. IBM SPSS Analytic Server 3.2.1 をインストールするには、以下のコマンドを実行します。

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

    注: ローカル・リポジトリーが正しくセットアップされていることを確認するには、<local_repo> ディレクトリーで上記のコマンドを実行しないでください。インストール中にパッケージが見つからない場合は、ローカル・リポジトリーが正しくセットアップされていないことを意味します (この場合は、上記のステップをすべて確認する必要があります)。
11. 各 Ambari 非サーバー・クラスター・ノードに対して以下の手順を繰り返します。
    a. Analytic Server の DEB ファイルのリポジトリーとして機能する <local_repo> ディレクトリーを作成します。以下に例を示します。

       ```
       mkdir –p /usr/local/mydebs
       ```

    b. <local_repo> ディレクトリーのすべての内容を、Ambari Manager ノード・マシンから、Ambari 非サーバー・クラスター・ノードの <local_repo> ディレクトリーにコピーします。このディレクトリーには、以下のファイルが含まれている必要があります。
       - <local_repo>/IBM-SPSS-AnalyticServer-ambari-2.x_3.2.1.0_amd64.deb
       - <local_repo>/IBM-SPSS-AnalyticServer_1_amd64.deb
       - <local_repo>/Packages.gz
       - <local_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/Packages
       - <local_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/Packages
    c. ローカル・リポジトリーの定義を作成します。例えば、IBM-SPSS-AnalyticServer-3.2.1.0.list というファイルを、/etc/apt/sources.list.d に、以下の内容を指定して作成します。

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

12. 8 ページの『オンライン・インストール』のセクションのステップ 3 に進みます。

外部管理される MySQL 環境に対する Analytic Server のインストール

外部管理される MySQL 環境に対してインストールする場合、Analytic Server のインストール・プロセスは、通常のインストールとは異なります。

以下のステップでは、外部管理される MySQL 環境に対して Analytic Server をインストールするプロセスについて説明します。

1. IBM パスポート・アドバンテージ Web サイトに移動し、ご使用のスタック、スタック・バージョン、およびハードウェア・アーキテクチャーに固有の自己解凍型バイナリー・ファイルを Ambari クラスター内のホストにダウンロードします。
2. 自己解凍型バイナリー・ファイルを実行し、指示に従って (オプションで) ライセンスを表示し、ライセンスを受け入れます。
   a. オンライン・オプションを選択します。
   b. プロンプトが出されたら、「外部 **MySQL** データベース **(External MySQL Database)**」オプションを選択します。
3. `add_mysql_user.sh` スクリプトを、`<AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer` から、`AS_MetaStore` として使用される MySQL インスタンスがインストールされているノード/ホストにコピーします。
   - MySQL ノード/ホスト上で `add_mysql_user.sh` スクリプトを実行します。例:
     `./add_mysql_user.sh -u as_user -p spss -d aedb`

   注:
   - ユーザー名およびパスワードは、Ambari 構成画面の `AS_Metastore` で入力されたデータベース・ユーザー名およびパスワードと一致する必要があります。
   - コマンドを発行するように `add_mysql_user.sh` スクリプトを手動で更新できます (希望する場合)。
   - セキュアな (root ユーザーによってアクセスされる) MySQL データベースに対して `add_mysql_user.sh` スクリプトを実行する場合は、`-r` パラメーターおよび `-t` パラメーターを使用して、`dbuserid` および `dbuserid_password` を渡します。スクリプトは、`dbuserid` および `dbuserid_password` を使用して、MySQL 操作を実行します。
4. Ambari サーバーを再起動します。
   ```
   ambari-server restart
   ```
5. Ambari コンソールから、`AnalyticServer` サービスを通常として追加します (ステップ 3 で入力したのと同じデータベース・ユーザー名およびパスワードを入力します)。

   注: 「**AS_Configuration**」画面の `metadata.repository.url` 設定 (「**Advanced analytics-meta**」) が MySQL データベース・ホストを指すように変更する必要があります。例えば、JDBC 設定 `mysql://{analytic_metastore_host}/aedb?createDatabaseIfNotExist=true` を `mysql://{MySQL_DB}/aedb?createDatabaseIfNotExist=true` に変更します。

構成

インストール後、オプションで Ambari UI を使用して Analytic Server を構成し、管理することができます。

注: Analytic Server ファイル・パスには以下の規則が使用されます。

- {AS_ROOT} は、Analytic Server がデプロイされている場所を示します (例えば、`/opt/IBM/SPSS/AnalyticServer/3.2`)。

- {AS_SERVER_ROOT} は、構成ファイル、ログ・ファイル、およびサーバー・ファイルの場所を示します (例えば、/opt/IBM/SPSS/AnalyticServer/3.2/ae_wlpserver/usr/servers/aeserver)。
- {AS_HOME} は、Analytic Server がルート・フォルダーとして使用する HDFS 上の場所を示します。

セキュリティー

LDAP レジストリーの構成

LDAP レジストリーは、Active Directory や OpenLDAP などの外部 LDAP サーバーを使用してユーザーを認証できるようにします。

重要: LDAP ユーザーを Ambari 内の Analytic Server 管理者として指定する必要があります。

以下に、OpenLDAP の ldapRegistry の例を示します。

```
<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>
```

以下の例は、Active Directory を使用した Analytic Server の認証を提供します。

```
<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>
```

注: 多くの場合、LDAP 構成を確認するには、サード・パーティーの LDAP ビューアー・ツールを使用すると便利です。

以下の例は、Active Directory を使用した WebSphere Liberty Profile の認証を提供します。

```
<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" >
    </activedFilters>
</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=" />
```

注:

- Analytic Server での LDAP のサポートは、WebSphere Liberty によって制御されます。詳しくは、『Liberty での LDAP ユーザー・レジストリーの構成』を参照してください。
- LDAP が SSL で保護されている場合は、以下の『Analytic Server から LDAP への Secure Sockets Layer (SSL) 接続の構成』セクションの指示に従ってください。

Analytic Server から LDAP への Secure Sockets Layer (SSL) 接続の構成

Analytic Server のインストール中に Apache Directory Server (ads) LDAP オプションを選択した場合 (デフォルト構成を使用する場合)、Apache Directory Server は、SSL が構成されて有効になっている状態でインストールされます (Analytic Server は、自動的に SSL を使用して Apache Directory Server と通信します)。

Analytic Server のインストール中に他のいずれかの LDAP オプションを選択した場合 (例えば、外部 LDAP サーバーを使用する場合) は、以下のステップを使用して SSL を構成します。

1. Analytic Server マシンのそれぞれに Analytic Server ユーザーとしてログインし、SSL 証明書の共通ディレクトリーを作成します。

   注: デフォルトでは、Analytic Server ユーザーは as_user です。Ambari コンソールの「Admin」タブの下の「**Service accounts**」を参照してください。

2. 鍵ストア・ファイルおよびトラストストア・ファイルを、すべての Analytic Server マシンの共通ディレクトリーにコピーします。また、LDAP クライアントの CA 証明書をトラストストアに追加します。以下に、手順例を示します。

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

   注: JAVA_HOME は、Analytic Server の起動に使用するのと同じ JRE です。

3. securityUtility ツール ({AS_ROOT}/ae_wlpserver/bin にあります) を使用してパスワードをエンコードすることで、パスワードの値を難読化できます。次に例を示します。

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

4. Ambari コンソールにログインし、Analytic Server の構成設定 **ssl.keystore.config** を、正しい SSL 構成設定に更新します。次に例を示します。

```
<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="/>
```

注**:** 鍵ファイルおよびトラストストア・ファイルについては、絶対パスを使用してください。

5. Analytic Server の構成設定 **security.config** を、正しい LDAP 構成設定に更新します。例えば、**ldapRegistry** 要素の場合、**sslEnabled** 属性を true に設定し、**sslRef** 属性を defaultSSLConfig に設定します。

Kerberos の構成

Analytic Server は、Ambari を使用した Kerberos をサポートします。

注**:** Kerberos シングル・サインオン (SSO) が Apache Knox と組み合わせて使用される場合、IBM SPSS Analytic Server は Kerberos SSO をサポートしません。

1. Analytic Server へのアクセス権を付与する予定のすべてのユーザーについて、Kerberos ユーザー・リポジトリー内にアカウントを作成します。
2. LDAP サーバー上に (前のステップと) 同じアカウントを作成します。
3. 前のステップで、Analytic Server および Hadoop の各ノードで作成したそれぞれのユーザーについて、OS ユーザー・アカウントを作成します。
   - これらのユーザーの UID は、すべてのマシンで一致させてください。kinit コマンドを使用して各アカウントにログオンして、これをテストすることができます。
   - UID が、Yarn の「ジョブをサブミットするための最小ユーザー ID (Minimum user ID for submitting job)」設定に従っていることを確認してください。これは、container-executor.cfg 内の **min.user.id** パラメーターです。例えば、**min.user.id** が 1000 の場合、作成される各ユーザー・アカウントの UID は 1000 以上でなければなりません。
4. Analytic Server のすべてのプリンシパルについて、HDFS 上にユーザーのホーム・フォルダーを作成します。例えば、Analytic Server システムに testuser1 を追加した場合、HDFS 上に /user/testuser1 のようなホーム・フォルダーを作成し、testuser1 がこのフォルダーに対する読み取り権限と書き込み権限を持つようにします。
5. [オプション] HCatalog データ・ソースを使用する予定であり、Analytic Server が Hive Metastore とは別のマシンにインストールされている場合、HDFS で Hive クライアント名を使用する必要があります。
   a. Ambari コンソールで、HDFS サービスの「Configs」タブに移動します。
   b. **hadoop.proxyuser.hive.groups** パラメーターを編集して値 * を設定するか、すべてのユーザーが Analytic Server へのログインを許可されているグループを指定します。
   c. **hadoop.proxyuser.hive.hosts** パラメーターを編集して値 * を設定するか、サービスとして Hive Metastore および Analytic Server の各インスタンスがインストールされているホストのリストを指定します。
   d. HDFS サービスを再起動します。

これらのステップの実行を完了した後、Analytic Server がインストールされていると、Analytic Server がサイレントかつ自動的に Kerberos の構成を行います。

Kerberos を使用したシングル・サインオン (SSO) 用の HAProxy の構成

1. HAProxy の資料 (http://www.haproxy.org/#docs) に従って HAProxy を構成して開始します。

2. HAProxy ホスト用の Kerberos プリンシパル (HTTP/<proxyHostname>@<realm>) およびキータブ・ファイルを作成します。ここで、<proxyHostname> は HAProxy ホストの完全な名前、<realm> は Kerberos レルムです。
3. キータブ・ファイルを各 Analytic Server ホストに /etc/security/keytabs/spnego_proxy.service.keytab としてコピーします。
4. このファイルのアクセス許可を各 Analytic Server ホストで更新します。次に例を示します。

```
chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
```

5. Amabri コンソールを開き、Analytic Server の「Custom analytics.cfg」セクションで以下のプロパティーを更新します。

```
web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
web.authentication.kerberos.principal=HTTP/<proxy machine full name>@<realm>
```

6. 構成を保存し、Amabri コンソールからすべての Analytic Server サービスを再始動します。

これで、ユーザーが IBM SPSS Analytic Server のログイン画面で「シングル・サインオンでのログイン **(Single sign on log in)**」オプションを使用して Analytic Server にログインできるようになりました。

Kerberos 偽名の使用の有効化

偽名を使用すると、スレッドを所有しているプロセスのセキュリティー・コンテキストとは異なるセキュリティー・コンテキスト内で、そのスレッドを実行できます。例えば、偽名の使用は、標準 Analytic Server ユーザー (as_user) 以外のユーザーとして Hadoop ジョブを実行する手段を提供します。Kerberos 偽名の使用を有効にするには、以下を行います。

1. Kerberos が有効になっているクラスター内で実行する場合は、HDFS (または Hive サービス構成) に偽名の使用構成属性を追加します。HDFS の場合は、以下のプロパティーを HDFS core-site.xml ファイルに追加する必要があります。

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

   ここで、<analytic_server_service_principal_name> は、Analytic Server 構成の Analytic_Server_User フィールドで指定されているデフォルトの as_user 値です。

   HDFS から Hive/HCatalog を経由してデータにアクセスする場合は、以下のプロパティーも HDFS core-site.xml ファイルに追加する必要があります。

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

2. as_user 以外のユーザー名を使用するように Analytic Server が構成されている場合は、そのユーザー名を反映するようにプロパティー名を変更する必要があります (例えば、hadoop.proxyuser.xxxxx.hosts です。ここで、xxxxx は、Analytic Server 構成で指定されている構成済みのユーザー名です)。
3. Analytic Server ノード上のコマンド・シェルから以下のコマンドを実行します。

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

複数レルムの有効化

複数のレルムを定義する場合は、**as.kdc.realms** の設定が必要です。**as.kdc.realms** の値は、Amabri コンソールの Analytic Server「Advanced analytics.cfg」セクションにあります。

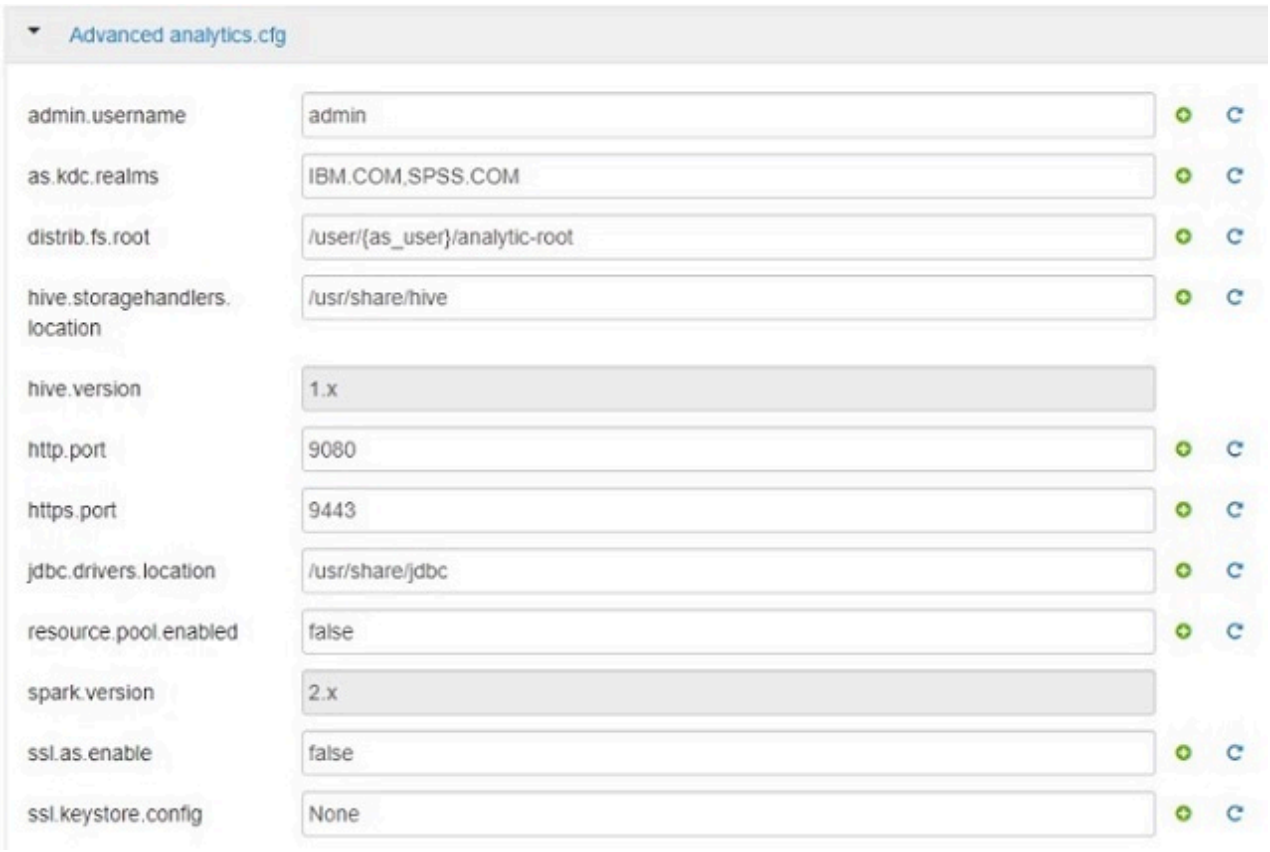

図 3. *Advanced analytics.cfg* 設定

コンマ文字で区切ると、複数のレルム名がサポートされます。指定された Kerberos レルム名はユーザー名に対応し、ユーザー名に関連付けられます。例えば、ユーザー名 `UserOne@us.ibm.com` および `UserTwo@eu.ibm.com` は、レルム `us.ibm.com`、`eu.ibm.com` に対応します。

「**Kerberos** レルム名 **(Kerberos Realm Name)**」として複数のレルムを指定する場合は、Kerberos クロスレルム・トラストを構成する必要があります。Analytic Server コンソールのログイン・プロンプトに入力されるユーザー名は、レルム名の接尾辞なしで入力されます。このため、複数のレルムが指定されている場合は、ユーザーに「レルム **(Realms)**」ドロップダウン・リストが表示され、該当するレルムを選択できます。

注: レルムが 1 つしか指定されていない場合は、Analytic Server へのサインイン時に「レルム **(Realms)**」ドロップダウン・リストは表示されません。

Kerberos の無効化

1. Ambari コンソールで Kerberosを無効化します。
2. Analytic Server サービスを停止します。
3. Custom `analytics.cfg` から、以下のパラメーターを削除します。

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

4. 「**Save**」をクリックし、Analytic Server サービスを再起動します。

Analytic Server コンソールへの Secure Sockets Layer (SSL) 接続の有効化

デフォルトでは、Analytic Server は自己署名証明書を生成して Secure Socket Layer (SSL) を有効にします。自己署名証明書を受け入れることにより、セキュア・ポートを使用して Analytic Server コンソールにアクセスできるようになります。HTTPS によるアクセスの安全性をさらに強化するには、サード・パーティー・ベンダーの証明書をインストールする必要があります。

サード・パーティー・ベンダーの証明書をインストールするには、以下のステップを実行します。

1. サード・パーティー・ベンダーの鍵ストア証明書およびトラストストア証明書を、すべての Analytic Server ノードで、同じディレクトリーにコピーします。例えば、`/home/as_user/security` です。

   注: Analytic Server ユーザーには、このディレクトリーの読み取りアクセス権限が必要です。
2. Ambari の「Services」タブで、Analytic Server サービスの「Configs」タブに移動します。
3. **`ssl.keystore.config`** パラメーターを編集します。

```
<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>"/>
```

   置き換える値:

   - <KEYSTORE-LOCATION> に、鍵ストアの絶対位置を指定します。例: `/home/as_user/security/mykey.jks`
   - <TRUSTSTORE-LOCATION> に、トラストストアの絶対位置に指定します。例: `/home/as_user/security/mytrust.jks`
   - <TYPE> に、証明書のタイプを指定します。例: JKS、PKCS12、その他。
   - <PASSWORD> に、Base64 暗号化形式の暗号化パスワードを指定します。エンコードには、securityUtility を使用できます。例: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility encode <password>`。

   自己署名証明書を生成する場合は、securityUtility を使用できます。例: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility createSSLCertificate --server=myserver --password=mypassword --validity=365 --subject=CN=mycompany,O=myOrg,C=myCountry`。

   注: `CN` 値に適切なホスト・ドメイン名を指定する必要があります。

   securityUtility およびその他の SSL 設定について詳しくは、WebSphere Liberty Profile の資料を参照してください。
4. 「**Save**」をクリックし、Analytic Server サービスを再起動します。

SSL を介した Apache Hive との通信

SSL 接続を介して Apache Hive と通信するためには、hive.properties ファイルを更新する必要があります。あるいは、ご使用の Apache Hive 環境で高可用性が有効になっている場合は、メインの Analytic Server「データ・ソース」ページ上で高可用性パラメーターを選択できます。

hive.properties ファイルの更新

1. hive.properties ファイルを開きます。このファイルは、/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/database にあります。
2. 以下の行を見つけます。

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

3. 以下の太字の情報を追加して、行を更新します。

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

4. hive.properties ファイルを保存します。

Essentials for R に対するサポートの有効化

Analytic Server は、R モデルのスコアリング、および R スクリプトの実行をサポートしています。

R に対するサポートを構成するには、Analytic Server が正常にインストールされた後で、以下を行います。

1. Essentials for R のサーバー環境をプロビジョンします。

RedHat Linux x86_64

以下のコマンドを実行します。

```
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

以下のコマンドを実行します。

```
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

SUSE での Essentials for R のインストールには、構成された ZYPPER リポジトリーでは通常は使用できない、互換性のある FORTRAN が必要です (SUSE SDK メディアからのみ使用可能です)。結果として、SUSE サーバーで Essentials for R の Ambari インストールを実行すると、FORTRAN をインストールできないため、失敗します。SUSE でプロビジョンするには、以下のステップを使用します。

a. GCC C++ をインストールします。

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

b. GCC FORTRAN をインストールします。必要な RPM ファイルを SUSE SDK メディアからコピーできますが、以下の順序でインストールする必要があります。

```
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. 以下のコマンドを実行して、Essentials for R ライブラリーをインストールします。

```
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. IBM SPSS Modeler Essentials for R の RPM または DEB 用の自己解凍型アーカイブ (BIN) をダウンロードします。Essentials for R は、https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp からダウンロードできます。ご使用のスタック、スタックのバージョン、およびハードウェア・アーキテクチャーに固有のファイルを選択します。
3. 自己解凍型バイナリー・ファイルを実行し、指示に従って (オプションで) ライセンスを表示し、ライセンスを受け入れて、オンライン・インストールまたはオフライン・インストールを選択します。

   オンライン・インストール

   Ambari サーバー・ホストおよびクラスター内のすべてのノードが https://ibm-open-platform.ibm.com にアクセス可能な場合は、オンライン・インストールを選択してください。

   オフライン・インストール

   ご使用の Ambari サーバー・ホストがインターネットにアクセスできない場合は、オフラインを選択します。オフライン・インストールでは必要な RPM ファイルをダウンロードするため、https://ibm-open-platform.ibm.com にアクセス可能なマシンで実行する必要があります。その後、RPM ファイルを Ambari サーバー・ホストにコピーできます。

   a. 必要な Essentials for R の RPM または DEB ファイルを Ambari サーバー・ホスト上の任意の場所にコピーします。必要な RPM/DEB ファイルは、ご使用のディストリビューション、バージョン、およびアーキテクチャーによって以下のように異なります。

HDP 2.5、2.6、3.0、および 3.1 (x86_64)
IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.1.0.0-1.x86_64.rpm

HDP 2.6、3.0、および 3.1 (PPC64LE)
IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.1.0.0-1.ppc64le.rpm

HDP 2.5、2.6、3.0、および 3.1 (Ubuntu)
IBM-SPSS-ModelerEssentialsR-ambari-3.2.1.0_3.2.1.0_amd64.deb

b. RPM または DEB をインストールします。以下の例では、コマンドは Essentials for R を HDP 2.6 (x86_64) にインストールします。

```
rpm -i IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.1.0.0-1.x86_64.rpm
```

以下の例では、コマンドは Essentials for R を HDP 2.5 (Ubuntu) にインストールします。

```
dpkg -i IBM-SPSS-ModelerEssentialsR-ambari-3.2.1.0_3.2.1.0_amd64.deb
```

4. Ambari サーバーを再起動します。

```
ambari-server restart
```

5. Ambari サーバーにログオンし、Ambari コンソールを使用して SPSS Essentials for R をサービスとしてインストールします。SPSS Essentials for R は、Analytic Server および Analytic Metastore がインストールされているすべてのホストにインストールする必要があります。

注**:** Ambari は R をインストールする前に gcc-c++ および gcc-gfortran (RHEL)、および gcc-fortran (SUSE) のインストールを試行します。これらのパッケージは、R の Ambari サービス定義で依存関係として宣言されています。R のインストールおよび実行場所となるサーバーが、gcc-c++ および gcc-[g]fortran の RPM をダウンロードするように構成されていること、あるいは、そのサーバーに GCC コンパイラーおよび FORTRAN コンパイラーがインストールされていることを確認してください。Essentials for R のインストールが失敗する場合は、Essentials for R をインストールする前にこれらのパッケージを手動でインストールしてください。

6. Analytic Server サービスをリフレッシュします。
7. 32 ページの『クライアント依存関係の更新』の手順に従って `update_clientdeps` スクリプトを実行します。
8. SPSS Modeler Server をホストするマシンに Essentials for R をインストールすることも必要です。詳しくは、SPSS Modeler の資料を参照してください。

リレーショナル・データベース・ソースの有効化

各 Analytic Server Metastore と各 Analytic Server ホストの共有ディレクトリー内に JDBC ドライバーを配置すると、Analytic Server でリレーショナル・データベース・ソースを使用できます。デフォルトでは、このディレクトリーは `/usr/share/jdbc` です。

共有ディレクトリーを変更するには、以下のステップを実行します。

1. Ambari の「Services」タブで、Analytic Server サービスの「Configs」タブに移動します。
2. 「**Advanced analytics.cfg**」セクションを開きます。
3. **`jdbc.drivers.location`** で、JDBC ドライバーの共有ディレクトリーを指定します。
4. 「**Save**」をクリックします。
5. Analytic Server サービスを停止します。
6. 「**Refresh**」をクリックします。
7. Analytic Server サービスを開始します。

表 6. サポート対象データベース

データベース	サポート対象バージョン	**JDBC** ドライバー **jar**	ベンダー
Amazon Redshift	8.0.2 以降	`RedshiftJDBC41-1.1.6.1006.jar` 以降	Amazon
BigSQL	4.1.0.0 以降	db2jcc.jar	IBM
dashDB	Bluemix サービス	`db2jcc.jar`	IBM
Db2 for Linux、UNIX、および 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	1.2, 2.1	`hive-jdbc-*.jar`	Apache
MySQL	5.6, 5.7	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7、6.x	`nzjdbc.jar`	IBM
Oracle	12c, 11g R2 (11.2)	`ojdbc6.jar`、`orai18n.jar`	Oracle
SQL Server	2014, 2012, 2008 R2	`sqljdbc4.jar`	Microsoft
Teradata	15, 15.1	`tdgssconfig.jar`、`terajdbc4.jar`	Teradata

注

- Analytic Server をインストールする前に Redshift データ・ソースを作成した場合、Redshift データ・ソースを使用するには以下のステップを実行する必要があります。
  1. Analytic Server コンソールで Redshift データ・ソースを開きます。
  2. `Redshift` データベース・データ・ソースを選択します。
  3. Redshift のサーバー・アドレスを入力します。
  4. データベース名とユーザー名を入力します。パスワードは自動的に入力されます。
  5. データベース表を選択します。
- BigSQL は、Apache Hadoop 環境の IBM SQL インターフェースです。BigSQL はリレーショナル・データベースではありませんが、Analytic Server は、JDBC を経由した BigSQL へのアクセスをサポートします (JDBC jar ファイルは、Db2 で使用される jar ファイルと同じです)。

  Analytic Server での BigSQL の一般的な使用法の 1 つは、HCatalog データ・ソースを経由した BigSQL Hadoop/HBase 表へのアクセスです。

HCatalog データ・ソースの有効化

Analytic Server は、Hive/HCatalog を介して複数のデータ・ソースをサポートしています。一部のソースでは、手動での構成ステップが必要です。

1. データ・ソースを有効にするために必要な JAR ファイルを収集します。Apache HBase および Apache Accumulo のサポートを有効にするために、追加のステップは必要ありません。その他の NoSQL データ・ソースについては、データベース・ベンダーに連絡して、該当するストレージ・ハンドラーおよび関連する jar を取得してください。サポートされる HCatalog データ・ソースについては、「IBM SPSS Analytic Server 3.2.1 ユーザーズ・ガイド」の『HCatalog データ・ソースの使用』セクションを参照してください。

2. これらの JAR ファイルを、各 Analytic Server Metastore と各 Analytic Server ノードの {HIVE_HOME}/auxlib ディレクトリーおよび /usr/share/hive ディレクトリーに追加します。
3. Hive Metastore サービスを再起動します。
4. Analytic Metastore サービスをリフレッシュします。
5. Analytic Server サービスの各インスタンスを再起動します。

注:

- Analytic Server Metastore は、Hive Metastore と同じマシンにインストールすることはできません。
- Analytic Server HCatalog データ・ソースを経由して HBase データにアクセスする場合、アクセスするユーザーは、HBase 表に対する読み取り権限を持っている必要があります。
  - Kerberos 以外の環境では、Analytic Server は as_user を使用して HBase にアクセスします (as_user は、HBase に対する読み取り権限を持っている必要があります).
  - Kerberos 環境では、as_user とログイン・ユーザーの両方が、HBase 表に対する読み取り権限を持っている必要があります。

NoSQL データベース

Analytic Server は、ベンダーから Hive ストレージ・ハンドラーが提供されている任意の NoSQL データベースをサポートします。

Apache HBase および Apache Accumulo のサポートを有効にするために、追加のステップは必要ありません。

その他の NoSQL データベースについては、データベース・ベンダーに連絡して、該当するストレージ・ハンドラーおよび関連する jar を取得してください。

ファイル・ベース Hive 表

Analytic Server は、組み込みまたはカスタムの Hive SerDe (serializer-deserializer) が利用可能な任意のファイル・ベース Hive 表をサポートします。

XML ファイルを処理するための Hive XML SerDe は Maven の Central Repository (http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde) にあります。

MapReduce v2 ジョブ

Analytic Server の「**Custom analytic.cfg**」セクション内の **preferred.mapreduce** 設定を使用して、MapReduce ジョブの処理方法を制御します。

表 7. *Custom analytics.cfg* プロパティー

プロパティー	説明
preferred.mapreduce	MapReduce ジョブが実行される方法を制御します。有効な値は以下のとおりです。 • spark • m3r • hadoop 例: preferred.mapreduce=spark

Apache Spark

Spark (バージョン 1.5 以降) を使用する場合は、Analytic Server インストール時に spark.version プロパティーを手動で追加する必要があります。

1. Amabri コンソールを開き、Analytic Server の「**Advanced analytics.cfg**」セクションで以下のプロパティーを追加します。
   - キー: spark.version
   - 値: 適切な Spark バージョン番号 (例えば、1.x、2.x、または None) を入力します。
2. 構成を保存します。

注**:** Custom analytics.cfg 設定を使用して、HCatalog が Spark を使用しないように強制できます。

1. Amabri コンソールを開き、Analytic Server の「**Custom analytics.cfg**」セクションで以下のプロパティーを追加します。
   - キー: spark.hive.compatible
   - 値: false

Kerberos 対応 HDP 3.0 (またはそれ以降の) 環境

Kerberos 対応 HDP 3.0 (またはそれ以降の) 環境では、追加のセキュリティー構成設定が必要な場合があります。HDFS では、ファイル・システムの FACL が /warehouse/tablespace/managed/hive ディレクトリーで使用されます。次の例外が messages.log ファイルまたは as_trace.log ファイルに出力されたときに、Hive Metastore に FACL を設定する要件を特定できます。

```
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)
```

広範に (この例では、hadoop グループ内のすべてのメンバーに対して) Hive の warehouse ディレクトリーへのアクセス権を付与する **setfacl** コマンドの例を次に示します。

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

より細分化されたアクセス制御が必要なときは、その他の制限が強いコマンドを使用する必要があります。

追加の参照情報が、次のサイトにあります。

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

Analytic Server で使用するポートの変更

デフォルトでは、Analytic Server はポート 9080 (HTTP 用) および 9443 (HTTPS 用) を使用します。ポートの設定を変更するには、以下のステップを実行します。

1. Ambari の「Services」タブで、Analytic Server サービスの「Configs」タブに移動します。
2. 「**Advanced analytics.cfg**」セクションを開きます。
3. 使用するポートを、**http.port** (HTTP ポート) および **https.port** (HTTPS ポート) に指定します。

4. 「**Save**」をクリックします。
5. Analytic Server サービスを再始動します。

高可用性 Analytic Server

クラスター内の複数のノードに Analytic Server をサービスとして追加することにより、高可用性構成にすることができます。

1. Ambari コンソールで、「Hosts」タブに移動します。
2. Analytic Server をまだサービスとして実行していないホストを選択します。
3. 「Summary」タブで、「**Add**」をクリックし、Analytic Serverを選択します。
4. 「追加の確認 **(Confirm Add)**」をクリックします。

複数クラスターのサポート

複数クラスター機能は、IBM SPSS Analytic Server の高可用性機能の拡張であり、複数テナント環境での独立性を強化します。デフォルトでは、(Ambari または ClouderaManager のいずれかで) Analytic Server サービスをインストールすると、結果として、単一の Analytic Server クラスターが定義されます。

クラスター仕様では、Analytic Server クラスター・メンバーシップが定義されます。クラスター仕様の変更は、(Ambari Analytic Server 構成の `analytics-cluster` フィールドで、または Cloudera Manager の `configuration/analytics-cluster.xml` ファイルを手動で編集して) XML コンテンツを使用して実行されます。複数の Analytic Server クラスターを構成する際は、それぞれの Analytic Server クラスターに独自のロード・バランサーを提供する必要があります。

複数クラスター機能を使用することで、あるテナントに対する作業が、別のテナントのクラスターで実行されている作業にマイナスの影響を与えることがなくなります。高可能性ジョブについては、ジョブのフェイルオーバーは、タスクが開始された Analytic Server クラスターの範囲内のみで発生します。以下の例は、複数クラスター XML 仕様を提供します。

注**:** クラスター内の複数のノードに Analytic Server をサービスとして追加することにより、それを高可用性にすることができます。

```
<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>
```

上記の例では、2 つのロード・バランサーが必要です。一方のロード・バランサーは `cluster1` のメンバー (`one.cluster` および `two.cluster`) に要求を送信し、もう一方のロード・バランサーは `cluster2` のメンバー (`three.cluster` および `four.cluster`) に要求を送信します。

以下の例は、単一クラスター XML 仕様 (デフォルト構成) を提供します。

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

上記の例では、構成されたクラスター・メンバーが複数ある場合に対応するために、1 つのロード・バランサーが必要です。

注

- シングルトン・クラスターのみが、**memberName** 要素でのワイルドカードの使用をサポートしています (例えば、クラスター・カーディナリティー = "1")。カーディナリティー要素の有効な値は、1 および 1+ です。
- **memberName** は、Analytic Server 役割の割り当て先のホスト名と同じように指定する必要があります。
- クラスター構成の変更が適用された後は、すべてのクラスター内のすべてのサーバーを再起動する必要があります。
- Cloudera Manager では、すべての Analytic Server ノードの analytics-cluster.xml ファイルを変更して維持する必要があります。すべてのノードが同じ内容を含むように維持する必要があります。

スモールデータ向けの JVM オプションの最適化

小規模な (M3R) ジョブの実行時にご使用のシステムを最適化するために、JVM プロパティーを編集できます。

Ambari コンソールで、Analytic Server サービスの「Configs」タブの Advanced analytics-jvm-options セクションを参照します。以下のパラメーターを変更して、Analytic Server (Hadoop ではなく) をホストするサーバーで実行されるジョブのヒープ・サイズを設定します。これは小規模な (M3R) ジョブを実行する場合に重要です。システムを最適化するために、これらの値を調整する必要がある場合があります。

```
-Xms512M
-Xmx2048M
```

クライアント依存関係の更新

このセクションでは、update_clientdeps スクリプトを使用して Analytic Server サービスの依存関係を更新する方法を説明します。

1. Ambari サーバー・ホストに root としてログインします。
2. ディレクトリーを /var/lib/ambari-server/resources/stacks/<stack-name>/<stack-version>/services/ANALYTICSERVER/package/scripts に変更します。例を示します。

   ```
   cd "/var/lib/ambari-server/resources/stacks/HDP/2.6/services/ANALYTICSERVER/package/scripts"
   ```
3. 以下の引数を指定して、update_clientdeps スクリプトを実行します。

   -u <ambari-user>
   Ambari アカウント・ユーザー名。

   -p <ambari-password>
   Ambari アカウント・ユーザーのパスワード。

   -h <ambari-host>
   Ambari サーバーのホスト名。

   -x <ambari-port>
   Ambari が listen しているポート。

   以下の例を参照してください。

   ```
   ./update_clientdeps.sh -u admin -p admin -h host.domain -x 8080
   ```
4. 以下のコマンドを使用して Ambari サーバーを再始動します。

   ```
   ambari-server restart
   ```

Apache Knox の構成

Apache Knox Gateway は、Apache Hadoop サービスにセキュア・アクセスの単一ポイントを提供するシステムです。このシステムにより、ユーザー (クラスター・データにアクセスし、ジョブを実行する人) およびオペレーター (アクセスを制御し、クラスターを管理する人) の両者の Hadoop セキュリティーが簡素化されます。この Gateway は、1 つ以上の Hadoop クラスターに機能やサービスを提供するサーバー (またはサーバーのクラスター) として実行されます。

注: Apache Knox が Kerberos シングル・サインオン (SSO) と組み合わせて使用される場合、IBM SPSS Analytic Server は Apache Knox をサポートしません。

Apache Knox Gateway は、Hadoop クラスター・トポロジーの詳細を効果的に非表示にし、エンタープライズ LDAP および Kerberos と統合されます。以下のセクションでは、Apache Knox および Analytic Server の必要な構成タスクについての情報を提供します。

前提条件

- 既知の Apache Knox 問題では、HTTP Cookie およびヘッダーに含まれているセキュリティー情報が伝搬されません (詳しくは、https://issues.apache.org/jira/browse/KNOX-895 を参照してください)。この問題は、Knox 0.14.0 (またはそれ以降) では解決されています。Knox を Analytic Server と共に使用する前に、Knox 0.14.0 (またはそれ以降) が含まれている更新済みの Hortonworks ディストリビューションを取得する必要があります。詳しくは、Hortonworks プロバイダーにお問い合わせください。
- Analytic Server ノードは、パスワードなしの SSH 接続を使用して Knox サーバーと接続する必要があります。パスワードなしの SSH 接続は、Analytic Server から Knox に移動します (「**Analytic Server**」 > 「**Knox**」)。
- Analytic Server は、Knox サービスがインストールされた後でインストールされる必要があります。

場合によっては、予期しない問題の結果、構成ファイルが自動的にコピーされないことがあります。このような場合には、以下の構成ファイルを手動でコピーする必要があります。

- `com.ibm.spss.knox_0.6-3.2.1.0.jar`: このファイルは、Analytic Server の以下の場所からコピーする必要があります。

  ```
  <Analytic_Server_Installation_Path>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib
  ```

  Knox サーバー・ノードの以下の場所にコピーします。

  ```
  /KnoxServicePath/ext
  ```

  例: `/usr/iop/4.1.0.0/knox/ext`
- `rewrite.xml` および `service.xml`: これらのファイルは、Analytic Server の以下の場所からコピーする必要があります。

  ```
  <Analytic_Server_Installation_Path>/ae_wlpserver/usr/servers/aeserver/configuration/knox
  ```

  Knox サーバー・ノードの以下の場所にコピーします。

  ```
  /KnoxServicePath/data/services
  ```

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

注: rewrite.xml ファイルと service.xml ファイルの 2 つのセット (http://rest トラフィック用に 1 セットと、ws://websocket トラフィック用に 1 セット) があります。analyticserver および analyticserver_ws の両方のすべての rewrite.xml ファイルおよび service.xml ファイルを Knox サーバー・ノードにコピーします。

Ambari の構成

Analytic Server サービスは、Ambari ユーザー・インターフェースで構成する必要があります。

1. Ambari ユーザー・インターフェースで、「**Knox**」 > 「**Configs**」 > 「**Advanced topology**」に移動します。現在の Knox 構成設定が「**content**」ウィンドウに表示されます。
2. 以下の 2 つのサービスを Knox 構成内の「**Advanced topology**」セクションに追加します。

```
<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} および {analyticserver-port} は、Analytic Server の対応するサーバー名およびポート番号に置き換える必要があります。

   - {analyticserver-host} URL は、Ambari ユーザー・インターフェース (「**SPSS Analytic Server**」 > 「**Summary**」 > 「**Analytic Server**」) にあります。
   - {analyticserver-port} 番号は、Ambari ユーザー・インターフェース (「**SPSS Analytic Server**」 > 「**Configs**」 > 「**Advanced analytics.cfg**」 > 「**http.port**」) にあります。

   注: Analytic Server が複数のノードにデプロイされていて、LoadBalancer が使用される場合、{analyticserver-host} および {analyticserver-port} は LoadBalancer の URL およびポート番号に対応する必要があります。
3. Knox サービスを再起動します。

LDAP が使用される場合、Knox のデフォルトは、指定された「Demo」LDAP になります。エンタープライズ LDAP サーバー (Microsoft LDAP や OpenLDAP など) に変更できます。

Analytic Server の構成

Analytic Server に LDAP を使用するには、Apache Knox で使用されたものと 同じ LDAP サーバーを使用するように、Analytic Server を構成する必要があります。以下の Ambari 設定の <value> 項目は、対応する Knox LDAP サーバー設定を反映するように更新する必要があります。

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

その値は、Ambari ユーザー・インターフェース (「**Knox**」 > 「**Configs**」 > 「**Advanced topology**」) で使用可能です。以下に例を示します。

```
<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>
```

Knox LDAP 設定を更新した後で Knox サービスを再起動します。

重要: Analytic Server の管理者パスワードは、Knox の管理者パスワードと同じである必要があります。

Apache Knox の構成

1. 以下のように Knox gateway.jks ファイルをリフレッシュします。
   a. Knox サーバーで、Knox サービスを停止します。
   b. gateway.jks を /var/lib/knox/data-2.6.2.0-205/security/keystores から削除します。
   c. Knox サービスを再起動します。
2. Knox サーバーで、サブディレクトリー <knox_server>/data/service/analyticserver/3.2.1.0 を作成し、service.xml ファイルおよび rewrite.xml ファイルを新規ディレクトリーにアップロードします。これらの 2 つのファイルは、Analytic Server 上の <analytic_server>/configuration/knox/analyticserver/ (例えば、/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/*.xml) にあります。
3. <knox_server>/bin で、スクリプト ./knoxcli.sh redeploy --cluster default を実行します。
4. com.ibm.spss.knoxservice_0.6-*.jar ファイルを <knox_server>/ext にアップロードします。このファイルは、Analytic Server 上の <analytic_server>/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.1.0.jar (例えば、/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.1.0.jar) にあります。
5. Ambari ユーザー・インターフェースで、「**Knox**」 > 「**Configs**」 > 「**Advanced topology**」から以下の要素を追加します。

```
<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>
```

   注: デフォルトでは、WebSocket 機能は無効になっています。これを有効にするには、/conf/gateway-site.xml ファイル内で gateway.websocket.feature.enabled プロパティーを true に変更します。
6. Ambari ユーザー・インターフェースで、「**Knox**」 > 「**Configs**」 > 「**Advanced users-ldif**」からユーザーを追加または更新します (例えば、admin、qauser1、qauser2)。
7. 「**Knox**」 > 「**Service Actions**」 > 「**Start Demo LDAP**」から LDAP を再起動します。
8. Knox サービスを再起動します。

Apache Knox 対応の Analytic Server の URL 構造

Knox 対応の Analytic Server のユーザー・インターフェース URL は、https://{knox-host}:{knox-port}/gateway/default/analyticserver/admin です。

- https プロトコル - ユーザーは証明書を受け入れて、Web ブラウザーに進む必要があります。
- knox-host は、Knox のホストです。
- knox-port は、Knox のポート番号です。
- URI は、gateway/default/analyticserver です。

IBM SPSS Analytic Server テナントごとに別個の YARN キューの構成 - HDP

Yarn キューの構成は、Spark 動的リソース割り振り技術を使用して行われます。

Hortonworks Data Platform 2.x

1. Ambari ユーザー・インターフェースで、「**SPSS Analytic Server サービス (SPSS Analytic Server service)**」 > 「**Configs**」 > 「**Advanced analytics.cfg**」タブに移動します。
2. 「**resource.pool.enabled**」値を `true` に変更します。
3. 「**Custom analytics.cfg**」タブで以下のプロパティーを追加します。

```
config.folder.path=/etc/spark2/conf
resource.pool.mapping=tenant1:test,tenant2:production
resource.pool.default=default
spark.scheduler.mode=FAIR
spark.yarn.queue=default
```

表 *8. Custom analytics.cfg* プロパティー

プロパティー	説明
`config.folder.path`	このディレクトリーには、Spark プール・プロパティー情報を含む `fairscheduler.xml` ファイルが含まれています。このファイルは必須であり、手動で作成する必要があります。詳しくは、『**fairscheduler.xml** の例』セクションを参照してください。
`resource.pool.mapping`	**Spark:** `fairscheduler.xml` ファイル内で定義されているプールにテナントをマップします。テナントのペアは、コンマで区切る必要があります (例えば、`tenant1:test,tenant2:production`)。プールを指定する前に、そのプールが `fairscheduler.xml` ファイル内で構成されていることを確認してください。 **MapReduce:** YARN Queue Manager 内で定義されているキューにテナントをマップします。テナントのペアは、コンマで区切る必要があります (例えば、tenant1:test,tenant2:production)。キューを指定する前に、そのキューを使用してシステムが構成されていること、およびジョブをキューにサブミットするためのアクセスが許可されていることを確認してください。 **注:** Spark ジョブと MapReduce ジョブを両方とも実行する場合、テナント・マップ値は、`fairscheduler.xml` ファイル内および YARN Queue Manager 内で同じ名前である必要があります。
`resource.pool.default`	**Spark:** デフォルト・リソース・プールを定義します。この値は、`default`、または `fairscheduler.xml` ファイル内で定義されているプール名にすることができます。テナントが構成されていない (または間違って構成されている) 場合は、`default` 設定を使用します。 **MapReduce:** ジョブのサブミット先となるデフォルト・キューを定義します。
`spark.scheduler.mode=FAIR`	**Spark:** fair scheduler を有効にします。このプロパティーを変更しないでください。
`spark.yarn.queue`	**Spark:** アプリケーションのサブミット先となる YARN キューの名前。YARN Queue Manager 内で、カスタマイズされた YARN キュー名を指定できます。

4. 構成を保存し、Analytic Server サービスを再始動します。

fairscheduler.xml の例

`fairscheduler.xml` ファイルには、Spark プール・プロパティー情報が含まれています。このファイルは必須であり、手動で作成する必要があります。

```
<?xml version="1.0"?>
<allocations>
  <pool name="production">
    <schedulingMode>FAIR</schedulingMode>
    <weight>1</weight>
    <minShare>2</minShare>
  </pool>
```

```
  <pool name="test">
    <schedulingMode>FIFO</schedulingMode>
    <weight>2</weight>
    <minShare>3</minShare>
  </pool>
</allocations>
```

参照情報

詳しくは、以下のサイトを参照してください。

- https://spark.apache.org/docs/latest/job-scheduling.html#dynamic-resource-allocation
- https://spark.apache.org/docs/latest/running-on-yarn.html

Ambari での IBM SPSS Analytic Server のマイグレーション

Analytic Server は、既存の Analytic Server インストール済み環境から新規インストール済み環境にデータおよび構成設定をマイグレーションすることができます。マイグレーションは、同じクラスター環境と新規クラスター環境のどちらに対しても実行できます。

同じサーバー・クラスター上での Analytic Server 3.1.2 から 3.2.1 へのマイグレーション

Analytic Server 3.1.2 の既存のインストール済み環境がある場合、ご使用の 3.1.2 構成設定を同じサーバー・クラスター上の 3.2.1 インストール済み環境にマイグレーションできます。

1. 以前の Analytic Server バージョン (Analytic Server 3.1.2) から構成設定を収集します。
   a. {AS_ROOT}¥tools¥unzip configcollector.zip アーカイブを解凍します (configcollector という名前の新規フォルダーが作成されます)。
   b. configcollector フォルダー内の configcollector.sh スクリプトを実行します。生成される圧縮 (ZIP) ASConfiguration_3.1.2.0.xxx.zip ファイルを別のフォルダーの場所に (バックアップとして) コピーします。
2. 以前の Analytic Server 3.1.2 バージョンのインストール済み環境から新規の場所に Analytic ルートをバックアップします。
   a. Analytic ルートの場所が不明な場合は、**hadoop fs -ls** コマンドを実行します。Analytic ルートのパスは /user/as_user/analytic-root/analytic-workspace と類似しています。ここで、as_user は、Analytic ルートを所有するユーザー ID です。
   b. **hadoop fs -copyToLocal** コマンドおよび **hadoop fs -copyFromLocal** コマンドを使用して、以前の Analytic Server バージョンの analytic-workspace フォルダーを新規の場所 (例えば、/user/as_user/analytic-root/AS31Location) にコピーします。
3. 組み込み Apache Directory Server を使用する場合は、サード・パーティー LDAP クライアント・ツールを使用して現在のユーザー/グループ構成をバックアップします。Analytic Server 3.2.1 がインストールされた後で、バックアップのユーザー/グループ構成を Apache Directory Server にインポートします。

   注**:** 外部 LDAP サーバーを使用する場合は、このステップをスキップできます。
4. Ambari コンソールを開き、**Analytic Server** サービスを停止します。
5. 以前の Analytic Server バージョン (Analytic Server 3.1.2) をアンインストールしてから、Analytic Server 3.2.1 をインストールします。インストール手順については、5 ページの『第 2 章 Ambari のインストールおよび構成』を参照してください。

6. Ambari コンソールを開き、**Analytic Server** サービスを停止します (Ambari では、**Analytic Metastore** サービスが実行されていることを確認します)。
7. ステップ 2 でバックアップされた Analytic Server 3.1.2 Analytic ルートを新規の Analytic Server バージョンの場所にコピーします。
   a. 新規にインストールされた Analytic Server バージョンから `analytic-workspace` を削除します。
   b. バックアップされた Analytic Server 3.1.2 Analytic ワークスペース・フォルダー (`/user/as_user/analytic-root/AS31Location`) を新規のバージョンの場所 (例えば、`/user/as_user/analytic-root/analytic-workspace`) にコピーします。Analytic ワークスペース所有者が `as_user` として定義されていることを確認する必要があります。
8. Zookeeper の状態をクリアします。Zookeeper の bin ディレクトリー (Hortonworks 上の `/usr/hdp/current/zookeeper-client` など) で、以下のコマンドを実行します。

   ```
   ./zkCli.sh rmr /AnalyticServer
   ```
9. ステップ 1 のバックアップ・アーカイブ `ASConfiguration_3.1.2.0.xxx.zip` を新規の Analytic Server バージョンの場所 (例えば、`/opt/ibm/spss/analyticserver/3.2/`) にコピーします。
10. **`migrationtool.sh`** スクリプトを実行し、(構成収集ツールによって作成された) `ASConfiguration_3.1.2.0.xxx.zip` アーカイブ・ファイルのパスを引数として渡すことで、マイグレーション・ツールを実行します。以下に例を示します。

    ```
    migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.1.2.0.xxx.zip
    ```
11. Analytic Server ノード上のコマンド・シェルから以下のコマンドを実行します。

    ```
    hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
    ```
12. Ambari コンソールで、**Analytic Server** サービスを開始します。

新規サーバー・クラスター上での Analytic Server 3.1.2 から 3.2.1 へのマイグレーション

Analytic Server 3.1.2 の既存のインストール済み環境がある場合、ご使用の 3.1.2 構成設定を新規サーバー・クラスター上の 3.2.1 インストール済み環境にマイグレーションできます。

1. 8 ページの『Ambari でのインストール』 の指示に従って、Analytic Server の新規バージョンをインストールします。
2. 以前のインストール済み環境から新規インストール済み環境に Analytic ワークスペースをコピーします。
   a. Analytic ワークスペースの場所が不明な場合は、`hadoop fs -ls` を実行します。Analytic ワークスペースのパスは `/user/as_user/analytic-root/analytic-workspace` と類似しています。ここで、`as_user` は、Analytic ワークスペースを所有するユーザー ID です。
   b. 新規サーバーの `analytic-workspace` を削除します。
   c. `hadoop fs -copyToLocal` および `hadoop fs -copyFromLocal` を使用して、以前のサーバーの Analytic ワークスペースを新規サーバーの `/user/as_user/analytic-root/analytic-workspace` フォルダーにコピーします (所有者が `as_user` として設定されていることを確認します)。
3. 組み込み Apache Directory Server を使用する場合は、サード・パーティー LDAP クライアント・ツールを使用して現在のユーザー/グループ構成をバックアップします。Analytic Server 3.2.1 がインストールされた後で、バックアップのユーザー/グループ構成を Apache Directory Server にインポートします。

   注: 外部 LDAP サーバーを使用する場合は、このステップをスキップできます。

4. 新規サーバーで、Ambari コンソールを開き、Analytic Server サービスを停止します (Ambari では、Analytic Metastore サービスが実行されていることを確認します)。
5. 古いインストール済み環境から構成設定を収集します。
   a. 新規インストール済み環境の configcollector.zip アーカイブを、古いインストール済み環境の {AS_ROOT}¥tools にコピーします。
   b. コピーした configcollector.zip を解凍します。これにより、以前のインストール済み環境内に新規の configcollector サブディレクトリーが作成されます。
   c. {AS_ROOT}¥tools¥configcollector 内の **configcollector** スクリプトを実行して、以前のインストール済み環境内の構成収集ツールを実行します。その結果生成された圧縮ファイル (ZIP) を、新規インストール済み環境をホストするサーバーにコピーします。

   重要**:** 指定された **configcollector** スクリプトは、最新バージョンの Analytic Server と互換性がない場合があります。**configcollector** スクリプトに関する問題が発生した場合は、IBM 技術サポート担当員にお問い合わせください。
6. Zookeeper の状態をクリアします。Zookeeper の bin ディレクトリー (Hortonworks 上の /usr/hdp/current/zookeeper-client など) で、以下のコマンドを実行します。

   ```
   ./zkCli.sh rmr /AnalyticServer
   ```
7. **migrationtool** スクリプトを実行し、構成収集ツールによって作成された圧縮ファイルのパスを引数として渡すことで、マイグレーション・ツールを実行します。次に例を示します。

   ```
   migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.1.2.0.xxx.zip
   ```
8. Analytic Server ノード上のコマンド・シェルから以下のコマンドを実行します。

   ```
   hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
   ```
9. Ambari コンソールで、Analytic Server サービスを開始します。

   注**:** 既存の Analytic Server インストール済み環境で使用するように R を構成している場合、新規 Analytic Server インストール済み環境で R を構成するステップに従います。

アンインストール

重要**:** Essentials for R がインストールされている場合、まず remove_R.sh スクリプトを実行する必要があります。Analytic Server をアンインストールする前に、Essentials for R のアンインストールに失敗すると、後から Essentials for R をアンインストールできなくなります。Analytic Server がアンインストールされると、remove_R.sh スクリプトは削除されます。Essentials for R の アンインストールについては、 40 ページの『Essentials for R のアンインストール』を参照してください。

1. Analytic Metastore ホストで、{AS_ROOT}/bin ディレクトリーにある remove_as.sh スクリプトを、以下のパラメーターを指定して実行します。

   u 必須。Ambari サーバー管理者のユーザー ID。

   p 必須。Ambari サーバー管理者のパスワード。

   h 必須。Ambari サーバー・ホスト名。

   x 必須。Ambari サーバー・ポート。

   l オプション。セキュア・モードを有効にします。

   以下に例を示します。

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

クラスター内の Ambari ホスト one.cluster から Analytic Server を削除します。

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

クラスター内の Ambari ホスト one.cluster から Analytic Server をセキュア・モードで削除します。

注: この操作により、HDFS 上の Analytic Server フォルダーが削除されます。

注: この操作では、Analytic Server に関連付けられた Db2 スキーマは一切削除されません。スキーマを手動で削除する方法については、Db2 の資料を参照してください。

Essentials for R のアンインストール

1. Essentials for R ホストで、{AS_ROOT}/bin ディレクトリーにある remove_R.sh スクリプトを、以下のパラメーターを指定して実行します。

   u 必須。Ambari サーバー管理者のユーザー ID。

   p 必須。Ambari サーバー管理者のパスワード。

   h 必須。Ambari サーバー・ホスト名。

   x 必須。Ambari サーバー・ポート。

   l オプション。セキュア・モードを有効にします。

   以下に例を示します。

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

   クラスター内の Ambari ホスト one.cluster から Essentials for R を削除します。

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

   クラスター内の Ambari ホスト one.cluster から Essentials for R をセキュア・モードで削除します。
2. Ambari サーバー・サービス・ディレクトリーから R サービス・ディレクトリーを削除します。例えば、HDP 2.6 の場合、ESSENTIALR ディレクトリーは /var/lib/ambari-server/resources/stacks/HDP/2.6/services に配置されています。
3. Ambari コンソールで、Essentials for R サービスが存在していないことを確認します。

第 3 章 Cloudera のインストールおよび構成

Cloudera の概要

Cloudera は、オープン・ソースの Apache Hadoop ディストリビューションです。Cloudera Distribution Including Apache Hadoop (CDH) は、当該テクノロジーのエンタープライズ・クラスのデプロイメントを対象としています。

Analytic Server は CDH プラットフォームで実行できます。CDH には、大規模なデータ・セット (主に MapReduce および HDFS) の確実かつスケーラブルな分散データ処理を実現する Hadoop の主要なコア要素が含まれています。また、セキュリティー、高可用性、およびハードウェアや他のソフトウェアとの統合を実現するその他のエンタープライズ向けのコンポーネントも含まれています。

Cloudera 固有の前提条件

一般的な前提条件に加えて、以下の情報を確認してください。

サービス

各 Analytic Server ホストに以下のインスタンスがインストールされていることを確認してください。

- HDFS: Gateway、DataNode または NameNode
- Hive: Gateway、Hive Metastore Server または HiveServer2
- Yarn: Gateway、ResourceManager または NodeManager

以下のインスタンスは、それらの機能が使用される場合にのみ必要です。

- Accumulo: Gateway
- HBase: Gateway、Master または RegionServer
- Spark: Gateway
- Spark 2: Gateway

メタデータ・リポジトリー

Db2 および MySQL を Analytic Server メタデータ・リポジトリーとして使用できます。MySQL を Analytic Server メタデータ・リポジトリーとして使用する場合は、 43 ページの『Analytic Server 用の MySQL の構成』の手順に従ってください。

Kerberos が有効になっている Cloudera 環境

Kerberos が有効になっている Cloudera 環境に Analytic Server をインストールする予定がある場合は、Kerberos が Analytic Server との互換性を持つように適切に構成されていることを確認する必要があります。

以下の各セクションは、Kerberos が既にインストールされている Cloudera 環境に適用されます。Cloudera に Analytic Server をインストールする前に、以下の各セクションの指示に従う必要があります。Kerberos 固有の用語 (例えば、**kinit**、**kadmin** など) が含まれているため、ここでは基本的な Kerberos 認証の知識がある読者が想定されています。

注**:** Analytic Server は、認証に使用する Kerberos 関連の値について HDFS 構成を検査します。

Kerberos 認証

Analytic Server をインストールする前に、Kerberos 認証が各 Cloudera クラスター・ノード上で構成されていることを確認してください。詳しくは、Cloudera の製品資料の『Configuring Authentication in Cloudera Manager』を参照してください。

注**:** Kerberos 認証を各 Cloudera クラスター・ノード上で構成した後で、Analytic Server をインストールする前に、**cloudera-scm-server** サービスおよび **cloudera-scm-agent** サービスを再始動する必要があります。**cloudera-scm-agent** サービスは、すべてのクラスター・ノード上で再始動する必要があります。

Kerberos での必要なアカウントの作成

1. Analytic Server へのアクセス権を付与する予定のすべてのユーザーについて、Kerberos ユーザー・リポジトリー内にアカウントを作成します。
2. LDAP サーバー上に (前のステップと) 同じアカウントを作成します。
3. 前のステップで、Analytic Server および Hadoop の各ノードで作成したそれぞれのユーザーについて、OS ユーザー・アカウントを作成します。
   - これらのユーザーの UID は、すべてのマシンで一致させてください。`kinit` コマンドを使用して各アカウントにログオンして、これをテストすることができます。
   - UID が、Yarn の「ジョブをサブミットするための最小ユーザー **ID (Minimum user ID for submitting job)**」設定に従っていることを確認してください。これは、`container-executor.cfg` 内の **min.user.id** 設定です。例えば、**min.user.id** が 1000 の場合、作成される各ユーザー・アカウントの UID は 1000 以上でなければなりません。
4. Analytic Server 管理者ユーザーについて、HDFS 上にユーザーのホーム・フォルダーを作成します。フォルダー・アクセス許可を `777` に、所有者を `admin` にして、さらにユーザー・グループを `hdfs` として設定する必要があります。以下の太字の例を参照してください。

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

Found 9 items

drwxrwxrwx   - hdfs     supergroup      0 2017-07-26 03:41 /user/AE
drwxrwxrwx   - admin    hdfs            0 2017-06-08 01:33 /user/admin
drwxr-x--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. HCatalog データ・ソースを使用する予定であり、Analytic Server が Hive Metastore とは別のマシンにインストールされている場合、HDFS で Hive クライアント名を使用する必要があります。
   a. Cloudera Manager で、HDFS サービスの「構成 (Configuration)」タブに移動します。

      注**:** 以下の設定がまだ設定されていない場合、「構成 **(Configuration)**」タブにそれらの設定が表示されない可能性があります。その場合は、検索を実行して見つけてください。

   b. **hadoop.proxyuser.hive.groups** 設定を編集して値 `*` を設定するか、すべてのユーザーが Analytic Server へのログインを許可されているグループを指定します。
   c. **hadoop.proxyuser.hive.hosts** 設定を編集して値 `*` を設定するか、サービスとして Hive Metastore および Analytic Server の各インスタンスがインストールされているホストのリストを指定します。
   d. HDFS サービスを再起動します。

これらのステップの実行を完了した後、Analytic Server がインストールされていると、Analytic Server がサイレントかつ自動的に Kerberos の構成を行います。

Kerberos 偽名の使用の有効化

偽名を使用すると、スレッドを所有しているプロセスのセキュリティー・コンテキストとは異なるセキュリティー・コンテキスト内で、そのスレッドを実行できます。例えば、偽名の使用は、標準 Analytic Server ユーザー (as_user) 以外のユーザーとして Hadoop ジョブを実行する手段を提供します。Kerberos 偽名の使用を有効にするには、以下を行います。

1. Cloudera Manager を開き、「**Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml**」領域 (「**HDFS (サービス全体) (HDFS (Service-Wide))**」 > 「構成 **(Configuration)**」タブにあります) で以下のプロパティーを追加するか、更新します。
   - 名前: `hadoop.proxyuser.as_user.hosts`
   - 値: `*`
   - 名前: `hadoop.proxyuser.as_user.groups`
   - 値: `*`

   注: **core-site.xml** 設定は、Hadoop 構成に適用されます (Analytic Server には適用されません)。
2. Analytic Server ノード上のコマンド・シェルから以下のコマンドを実行します。

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

Analytic Server 用の MySQL の構成

Cloudera Manager で IBM SPSS Analytic Server を構成するには、MySQL サーバー・データベースをインストールして構成する必要があります。

1. MySQL データベースが格納されているノードのコマンド・ウィンドウから以下のコマンドを実行します。

   ```
   yum install mysql-server
   ```

   注: SuSE Linux の場合は `zypper install mysql` を使用してください。
2. 各 Cloudera クラスター・ノードのコマンド・ウィンドウから以下のコマンドを実行します。

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

   注: SUSE Linux の場合は `sudo zypper install mysql-connector-java` を使用してください。
3. Analytic Server が MySQL データベースへのアクセス時に使用する Analytic Server のデータベース名、データベースのユーザー名、およびデータベースのパスワードを決定し、メモを取ります。
4. 46 ページの『Cloudera でのインストール』の手順に従って Analytic Server をインストールします。
5. Cloudera によって管理されているいずれかのサーバーから、MySQL データベースがインストールされているノードに `/opt/cloudera/parcels/AnalyticServer/bin/add_mysql_user.sh` スクリプトをコピーします。ユーザー固有の構成に適したパラメーターを指定してそのスクリプトを実行します。以下に例を示します。

   ```
   ./add_mysql_user.sh -u <database_user_name> -p <database_password> -d
   <database_name>
   ```

   注: データベースがセキュア・モード (root ユーザー・パスワードが設定される) で実行される場合は `a -r <dbRootPassword)` パラメーターが必須になります。

root 以外のユーザー名を使用してデータベースがセキュア・モードで実行されている場合は -r <dbUserPassword> パラメーターおよび -t <dbUserName> パラメーターが必須になります。

インストールの事前チェック・ツールと事後チェック・ツール - Cloudera

ツールの場所と前提条件

Analytic Server サービスをインストールする前に、Analytic Server サービスの一部となるすべてのノード上で事前チェック・ツールを実行し、Linux 環境に Analytic Server をインストールする準備が整っているか確認します。

事前チェックツールは、インストールの一部として自動的に起動されます。このツールは、各ホスト上でインストールを実行する前に、各 Analytic Server ノードをチェックします。各ノードで事前チェック・ツールを手動で起動することもできます。これにより、サービスをインストールする前にマシンを検証できます。

自己解凍型 Analytic Server バイナリー・ファイルを実行した後で、事前チェック・ツールは、以下のディレクトリーにあります。

- **Cloudera**

```
/opt/cloudera/parcels/AnalyticServer-3.2.1.0/tools/com.spss.ibm.checker.zip

[root@servername ~]# cd /opt/cloudera/parcels/AnalyticServer-3.2.1.0/tools/
[root@servername tools]# ls
com.spss.ibm.checker.zip configcollector.zip regex-files
```

注: 実行可能バイナリー・ファイルを実行し、Cloudera Manager の「パーセル (Parcels)」ページ内で Analytic Server を配布 (「ダウンロード **(Download)**」 > 「配布 **(Distribute)**」) してアクティブにするまで、事前チェック・ツールは tools ディレクトリーにありません。

Analytic Server のインストール後に、事後チェック・ツールは以下のディレクトリーにあります。

- **Cloudera**

```
/opt/cloudera/parcels/AnalyticServer-3.2.1.0/tools/com.spss.ibm.checker.zip
```

ツールは root として実行する必要があり、Python 2.6.X 以上を必要とします。

事前チェック・ツールが失敗を報告した場合は、Analytic Server インストールを続行する前に、それらの失敗に対処する必要があります。

事前チェック・ツールの実行

自動

事前チェックツールは、Cloudera Manager コンソールを介して Analytic Server をインストールするときに、Analytic Server インストールの一部として自動的に起動できます。Cloudera Manager の管理者ユーザー名とパスワードを手動で入力する必要があります。

図 *4. Cloudera Manager* 管理者設定

手動

各クラスター・ノードで事前チェック・ツールを手動で起動できます。

以下の事前チェックの例は、Cloudera クラスター `MyCluster` をチェックします。このクラスターは、`myclouderahost.ibm.com:7180` で実行され、ログイン資格情報 `admin:admin` を使用します。

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

注:

- `as_host` 値は、IP アドレスまたは完全修飾ドメイン名のいずれかによって指定する必要があります。
- パスワード引数が省略されると、ツールはパスワードの入力を求めるプロンプトを出します。
- `precheck.py` コマンドに含まれている使用法ヘルプは、`-h` 引数 (`python ./precheck.py -help`) と表示されます。
- `--cluster` 引数はオプションです (`--cluster` が使用されていない場合は、現在のクラスターが指定されます)。

事前チェック・ツールがチェックを実行しているときには、各チェックの状況がコマンド・ウィンドウに表示されます。失敗が発生した場合は、ログ・ファイル内の詳細情報を参照できます (ログ・ファイルの具体的な場所は、コマンド・ウィンドウで指示されます)。追加のサポートが必要な場合は、ログ・ファイルを IBM Technical Support に提供できます。

事後チェック・ツールの実行

事後チェック・ツールは、Analytic Server が適切に実行されていること、および単純なジョブを処理できることを検証します。以下の事後チェックの例は、特定の Analytic Server インスタンスをチェックします。このインスタンスは、`myanalyticserverhost.ibm.com:9443` で実行され、SSL が有効になっていて、ログイン資格情報 `admin:ibmspss` を使用します。

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

Knox が Analytic Server と共に使用される場合、コマンドは以下のとおりです。

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

単一のチェックを実行するには、以下のコマンドを使用します。

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

注:

- パスワード引数が省略されると、ツールはパスワードの入力を求めるプロンプトを出します。
- `postcheck.py` コマンドに含まれている使用法ヘルプは、`--h` 引数 (`python ./postcheck.py --help`) と表示されます。

事後チェック・ツールがチェックを実行しているときには、各チェックの状況がコマンド・ウィンドウに表示されます。失敗が発生した場合は、ログ・ファイル内の詳細情報を参照できます (ログ・ファイルの具体的な場所は、コマンド・ウィンドウで指示されます)。追加のサポートが必要な場合は、ログ・ファイルを IBM Technical Support に提供できます。

Cloudera でのインストール

以下のステップでは、Cloudera Manager で IBM SPSS Analytic Server を手動でインストールするプロセスについて説明します。

Analytic Server 3.2.1

オンライン・インストール

1. IBM パスポート・アドバンテージ Web サイトに移動し、ご使用のスタック、スタック・バージョン、およびハードウェア・アーキテクチャーに固有の自己解凍型バイナリー・ファイルを Cloudera クラスター内のホストにダウンロードします。使用可能な Cloudera バイナリーは以下のとおりです。

表 *9. Analytic Server* 自己解凍型バイナリー・ファイル

説明	バイナリー・ファイル名
IBM SPSS Analytic Server 3.2.1 for Cloudera 5.11、5.12、5.13、5.14、5.15、6.0、および 6.1 Ubuntu 英語	`spss_as-3.2.1.0-cdh5.11-6.1-ubun.bin`
IBM SPSS Analytic Server 3.2.1 for Cloudera 5.11、5.12、5.13、5.14、5.15、6.0、および 6.1 Linux x86-64 英語	`spss_as-3.2.1.0-cdh5.11-6.1-lx86.bin`

2. Cloudera の自己解凍型 `*.bin` インストーラーを Cloudera Manager マスター・クラスター・ノードで実行します。ご使用条件に同意し、デフォルトの CSD インストール・ディレクトリーを維持してインストールのプロンプトに従ってください。

   注: CSD ディレクトリーをデフォルトの場所から変更した場合は、別の CSD ディレクトリーを指定する必要があります。

3. インストールの完了後に Cloudera Manager を再始動するには、以下のコマンドを使用します。

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

4. Cloudera Manager インターフェース (例えば、`http://${CM_HOST}:7180/cmf/login`) を、デフォルトのログイン資格情報 `admin/admin` を使用して開き、「リモート・パーセル・リポジトリーの **URL**

(Remote Parcel Repository URLs)」(「ホスト」 > 「パーセル **(Parcels)**」 > 「構成」にあります) をリフレッシュし、URL が正しいことを確認します。以下に例を示します。

```
https://ibm-open-platform.ibm.com
```

注**:** 「パーセルの更新頻度 **(Parcel Update Frequency)**および「リモート・パーセル・リポジトリーの **URL (Remote Parcel Repository URLs)**」は、ユーザーの固有のニーズに合わせて更新できます。

5. Cloudera Manager がパーセル・ファイルをリフレッシュした後で (「新しいパーセルの確認 **(Check for New Parcels)**」をクリックすることでパーセル・ファイルを手動でリフレッシュできます)、「**AnalyticServer**」パーセルの状況が「リモートで使用可能 **(Available Remotely)**」に設定されていることが分かります。
6. 「ダウンロード **(Download)**」 > 「配布 **(Distribute)**」 > 「アクティブ化 **(Activate)**」を選択します。「**AnalyticServer**」パーセルの状況が「配布済み、アクティブ化済み **(Distributed, Activated)**」に更新されます。
7. Cloudera Manager で Analytic Server をサービスとして追加し、Analytic Server を配置する場所を決定します。以下の情報を「サービスの追加ウィザード (Add Service Wizard)」に指定する必要があります。

   注**:** 「サービスの追加ウィザード (Add Service Wizard)」には、サービス作成プロセスの各フェーズにおける全体の進行状況が表示されます。また、クラスターでサービスが正常にインストールおよび構成されたときに最終確認メッセージが表示されます。

   - Analytic Server metastore ホスト名
   - Analytic Server metastore データベース名
   - Analytic Server metastore ユーザー名
   - Analytic Server metastore パスワード

   MySQL を **Analytic Server** メタデータ・リポジトリーとして使用

   - Analytic Server metastore ドライバー・クラス: `com.mysql.jdbc.Driver`
   - Analytic Server metastore リポジトリー URL: `jdbc:mysql://${MySQL_DB}/{DBName}?createDatabaseIfNotExist=true`

     `{MySQL_DB}` は、MySQL がインストールされているサーバーのホスト名です。

   Db2 を **Analytic Server** メタデータ・リポジトリーとして使用

   - Analytic Server metastore ドライバー・クラス: `com.ibm.db2.jcc.DB2Driver`
   - Analytic Server metastore リポジトリー URL: `jdbc:db2://{Db2_HOST}:{PORT}/{DBName}:currentSchema={SchemaName};`

     `{Db2_HOST}` は、Db2 がインストールされているサーバーのホスト名です。

     `{PORT}` は、Db2 が listen しているポートです。

     `{SchemaName}` は、使用可能な、未使用のスキーマです。

     入力する値がわからない場合は、Db2 管理者に協力を求めてください。

   LDAP 構成

   Analytic Server は、LDAP サーバーを使用して、ユーザーおよびグループを保管および認証します。必要な LDAP 構成情報を Analytic Server のインストール中に指定します。

表 *10. LDAP* 構成設定

LDAP 設定	説明
as.ldap.type	LDAP タイプ。値は、ads、ad、または openladp にすることができます。 • ads - Apache Directory Server (デフォルト設定) • ad - Microsoft Active Directory • openladp - OpenLDAP
as.ldap.host	LDAP ホスト
as.ldap.port	LDAP ポート番号
as.ldap.binddn	LDAP バインド DN
as.ldap.bindpassword	LDAP バインド DN パスワード
as.ldap.basedn	LDAP ベース DN
as.ldap.filter	LDAP ユーザーおよびグループのフィルター・ルール **注:** この値に縦棒文字 | が含まれている場合は、円記号文字を使用して縦棒文字をエスケープする必要があります (例えば、¥|)。
as.ldap.ssl.enabled	Analytic Server と LDAP の間の通信に SSL を使用するかどうかを指定します。値は true または false にすることができます。
as.ldap.ssl.reference	LDAP SSL 参照 ID
as.ldap.ssl.content	LDAP SSL 構成

- デフォルトでは、as.ldap.type は ads に設定され、その他の関連設定にはデフォルト設定が含まれます。ただし、例外として、as.ldap.bindpassword 設定のパスワードは独自に指定する必要があります。Analytic Server は、構成設定を使用して、Apache Directory Server (ADS) をインストールし、サーバーの初期化を実行します。デフォルトの ADS プロファイルには、admin というパスワードを持つユーザー admin が含まれています。Analytic Server コンソールを使用してユーザー管理を実行するか、<Analytic Root>/bin フォルダー内にある importUser.sh スクリプトを使用して XML ファイルからユーザーおよびグループの情報をインポートすることができます。
- Microsoft Active Directory や OpenLDAP などの外部 LDAP サーバーを使用する予定がある場合は、実際の LDAP 値に従って構成設定を定義する必要があります。詳しくは、『Liberty での LDAP ユーザー・レジストリーの構成』を参照してください。
- Analytic Server がインストールされた後で LDAP 構成を変更できます (例えば、Apache Directory Server から OpenLDAP に変更します)。ただし、最初に Microsoft Active Directory または OpenLDAP で開始してから、後で Apache Directory Server に切り替えることを決定した場合は、Analytic Server がインストール中に Apache Directory Server をインストールすることはありません。Apache Directory Server は、それが最初の Analytic Server のインストール中に選択されていた場合にのみインストールされます。

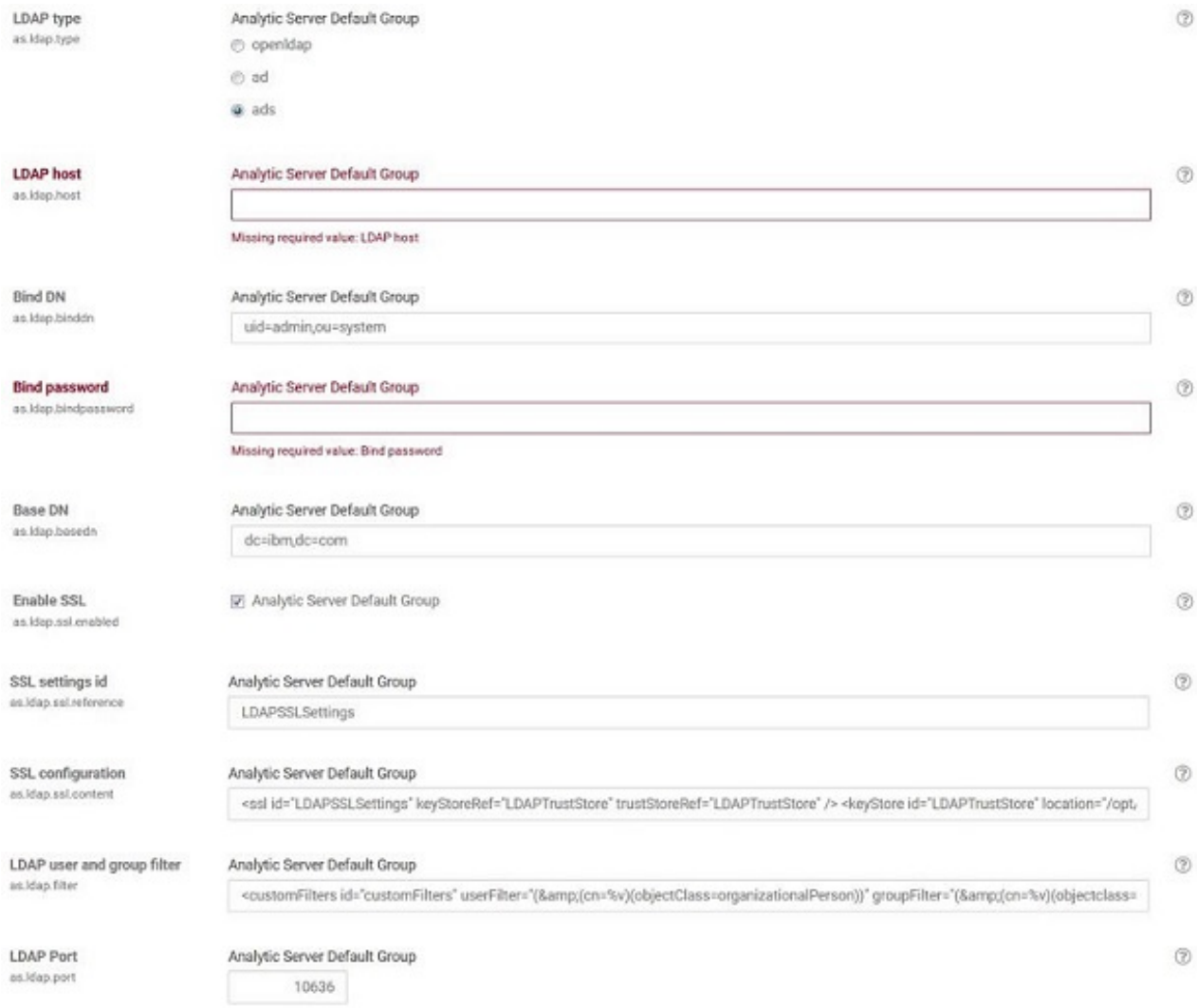

図 *5. LDAP* 構成設定の例

8. Kerberos が有効になっている Cloudera 環境に Analytic Server をインストールする場合は、以下の設定を「サービスの追加ウィザード (Add Service Wizard)」で構成する必要もあります。

   注**:** Analytic Server は、認証に使用する Kerberos 関連の値について HDFS 構成を検査します。

   - Analytic Server コンソールにログインするときに Kerberos 認証を有効にする場合は、「Kerberos」を「**Analytic Server** セキュリティー **(Analytic Server security)**」設定として選択します。「**Kerberos**」が「**Analytic Server** セキュリティー **(Analytic Server security)**」設定として選択されている場合、Analytic Server コンソールはデフォルトで Kerberos ログイン・モードになります。
   - Kerberos が有効になっているデータベースに接続する場合は、「Kerberos」を「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」設定として選択します。「**Kerberos**」が「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」設定として選択されている場合、Analytic Server コンソールはデータベースに接続するときに Kerberos モードを使用します。

- 「**Kerberos レルム名 (Kerberos Realm Name)**」および「**KDC ホスト (KDC host)**」の設定が必要です。「**Kerberos レルム名 (Kerberos Realm Name)**」(**as.kdc.realms**) および「**KDC ホスト (KDC host)**」(**kdcserver**) の値は、Kerberos 鍵配布センター (KDC) サーバー上の `krb5.conf` ファイルにあります。

  コンマ文字で区切ると、複数のレルム名がサポートされます。指定された Kerberos レルム名はユーザー名に対応し、ユーザー名に関連付けられます。例えば、ユーザー名 `UserOne@us.ibm.com` および `UserTwo@eu.ibm.com` は、レルム `us.ibm.com`、`eu.ibm.com` に対応します。

  「**Kerberos レルム名 (Kerberos Realm Name)**」として複数のレルムを指定する場合は、Kerberos クロスレルム・トラストを構成する必要があります。Analytic Server コンソールのログイン・プロンプトに入力されるユーザー名は、レルム名の接尾辞なしで入力されます。このため、複数のレルムが指定されている場合は、ユーザーに「**レルム (Realms)**」ドロップダウン・リストが表示され、該当するレルムを選択できます。

  注: レルムが 1 つしか指定されていない場合は、Analytic Server へのサインイン時に「**レルム (Realms)**」ドロップダウン・リストは表示されません。

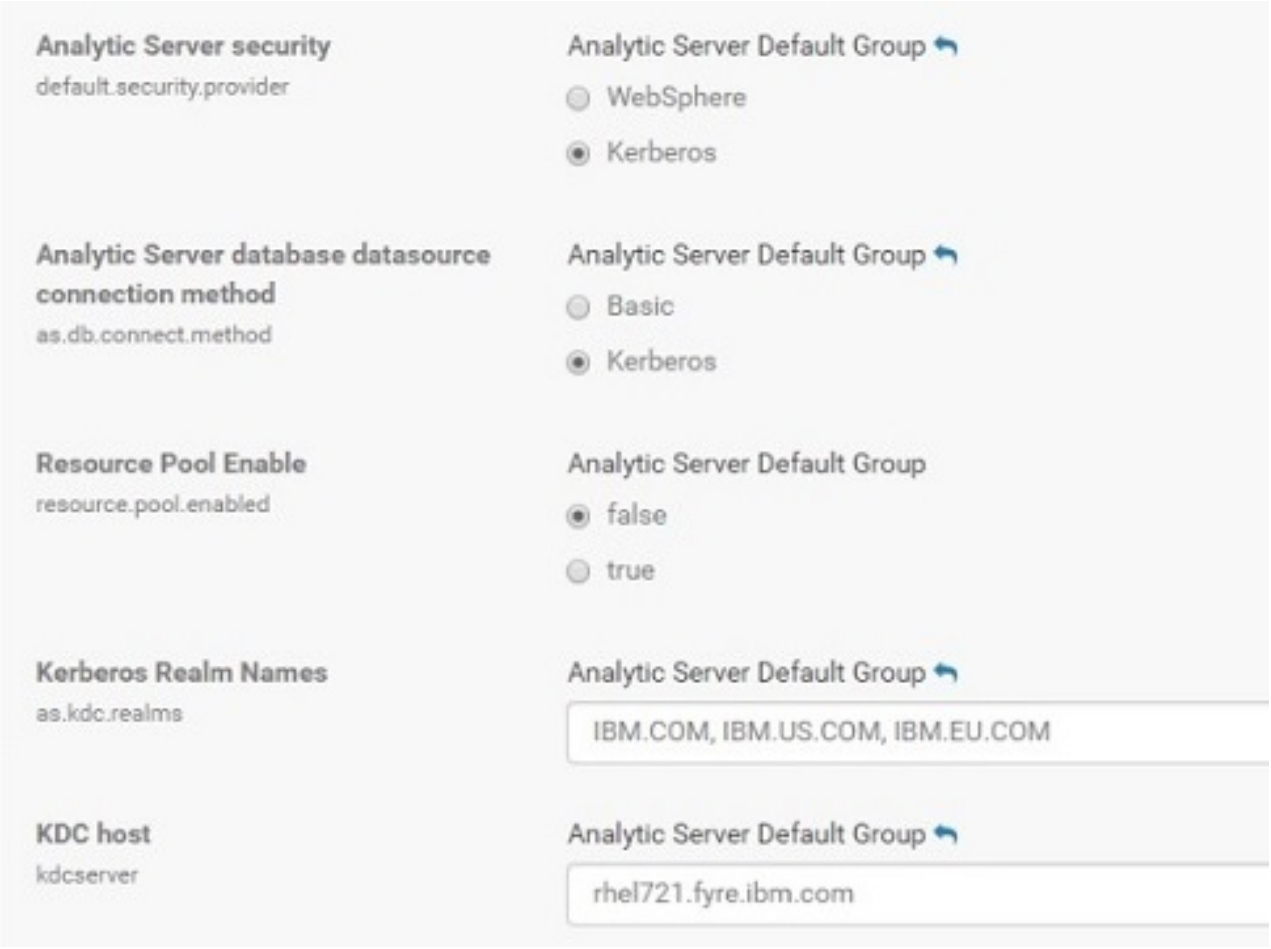

図 *6. Kerberos* 設定の例

注:

– 「**Analytic Server** セキュリティー **(Analytic Server security)**」および「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」の設定は、IBM SPSS Modeler クライアント認証および Analytic Server コンソール認証に適用されます。
– 「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」が「Kerberos」に設定されている場合は、ターゲット・データベースでも Kerberos が有効になっていることを確認する必要があります。
– 「**Analytic Server** セキュリティー **(Analytic Server security)**」および「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」の設定では、Hadoop クラスター上の Kerberos 認証は構成されません。詳しくは、『Kerberos 偽名の使用の有効化』セクションを参照してください。
– Kerberos 認証をログイン時に有効にする場合は、IBM SPSS Modeler クライアントを有効な Kerberos クライアントとしてデプロイする必要があります。これを行うには、Kerberos 鍵配布センター (KDC) サーバーで **addprinc** コマンドを使用します。詳しくは、IBM SPSS Modeler の資料を参照してください。

Kerberos が有効になっている Cloudera 環境に Analytic Server をインストールする場合は、Kerberos で必要なアカウントを作成し、Kerberos 偽名の使用を有効にする必要もあります。詳しくは、53 ページの『Kerberos の構成』を参照してください。

注**:** Analytic Server が正常にインストールされた後に、Cloudera Manager の Analytic Server サービス・ページの「アクション (Actions)」リストで「**Analytic Server Metastore** の作成 **(Create Analytic Server Metastore)**」をクリックしないでください。Metastore を作成すると、既存のメタデータ・リポジトリーが上書きされます。

オフライン・インストール

オフライン・インストールのステップは、特定のオペレーティング・システムに適したパーセル・ファイルとメタデータをユーザーが手動でダウンロードする必要がある点を除いて、オンラインのステップと同じです。

RedHat Linux では、以下のファイルが必要です。

- AnalyticServer-3.2.1.0-el7.parcel
- AnalyticServer-3.2.1.0-el7.parcel.sha
- manifest.json

SuSE Linux では、以下のファイルが必要です。

- AnalyticServer-3.2.1.0-sles11.parcel
- AnalyticServer-3.2.1.0-sles11.parcel.sha
- manifest.json

または

- AnalyticServer-3.2.1.0-sles12.parcel
- AnalyticServer-3.2.1.0-sles12.parcel.sha

Ubuntu Linux 14.04 では、以下のファイルが必要です。

- AnalyticServer-3.2.1.0-trusty.parcel
- AnalyticServer-3.2.1.0-trusty.parcel.sha

Ubuntu Linux 16.04 では、以下のファイルが必要です。

- AnalyticServer-3.2.1.0-xenial.parcel
- AnalyticServer-3.2.1.0-xenial.parcel.sha

1. Cloudera の自己解凍型 `*.bin` インストーラーをダウンロードして Cloudera Manager マスター・クラスター・ノードで実行します。ご使用条件に同意し、デフォルトの `CSD` インストール・ディレクトリーを維持してインストールのプロンプトに従ってください。

   注: `CSD` ディレクトリーがデフォルトの場所とは異なる場合は、別の CSD ディレクトリーを指定する必要があります。
2. 必要なパーセル・ファイルとメタデータ・ファイルを、Cloudera Manager マスター・クラスター・ノード上のローカル Cloudera `repo` パスにコピーします。デフォルトのパスは `/opt/cloudera/parcel-repo` です (このパスは Cloudera Manager ユーザー・インターフェースで構成可能です)。
3. Cloudera Manager を再始動するには、以下のステップを使用します。

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

   Cloudera Manager が「**AnalyticServer**」パーセルをリフレッシュした後でそのパーセルが「ダウンロード済み **(downloaded)**」と表示されます。「新しいパーセルの確認 **(Check for New Parcels)**」をクリックすると強制的にリフレッシュできます。
4. 「配布 **(Distribute)**」 > 「アクティブ化 **(Activate)**」をクリックします。

   「**AnalyticServer**」パーセルが「配布済み」および「アクティブ化済み」と表示されます。
5. Cloudera Manager で、Analytic Server をサービスとして追加します。詳しくは、『オンライン・インストール』セクションのステップ 7 と 8 を参照してください。

Cloudera の構成

インストール後、オプションで Cloudera Manager を使用して Analytic Server を構成し、管理することができます。

注: Analytic Server ファイル・パスには以下の規則が使用されます。

- {AS_ROOT} は、Analytic Server がデプロイされている場所を示します (例えば、`/opt/cloudera/parcels/AnalyticServer`)。
- {AS_SERVER_ROOT} は、構成ファイル、ログ・ファイル、およびサーバー・ファイルの場所を示します (例えば、`/opt/cloudera/parcels/AnalyticServer/ae_wlpserver/usr/servers/aeserver`)。
- {AS_HOME} は、Analytic Server がルート・フォルダーとして使用する HDFS 上の場所を示します (例えば、`/user/as_user/analytic-root`)。

セキュリティー

IBM SPSS Modeler `options.cfg` ファイル内のデフォルト **`tenant_id`** 値は **`ibm`** です。Analytic Server コンソールでテナントを表示できます。テナント管理について詳しくは、「*IBM SPSS Analytic Server* 管理者ガイド」を参照してください。

LDAP レジストリーの構成

LDAP は、Analytic Server のインストール中に構成されます。Analytic Server のインストール後に、別の LDAP サーバー方式に変更できます。

注: Analytic Server での LDAP のサポートは、WebSphere Liberty によって制御されます。詳しくは、『Liberty での LDAP ユーザー・レジストリーの構成』を参照してください。

Analytic Server から LDAP への Secure Sockets Layer (SSL) 接続の構成

1. Analytic Server マシンのそれぞれに Analytic Server ユーザーとしてログインし、SSL 証明書の共通ディレクトリーを作成します。

   注: Cloudera では、Analytic Server ユーザーは常に as_user になり、これは変更できません。

2. 鍵ストア・ファイルおよびトラストストア・ファイルを、すべての Analytic Server マシンの共通ディレクトリーにコピーします。また、LDAP クライアントの CA 証明書をトラストストアに追加します。以下に、手順例を示します。

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

   注: JAVA_HOME は、Analytic Server の起動に使用するのと同じ JRE です。

3. securityUtility ツール ({AS_ROOT}/ae_wlpserver/bin にあります) を使用してパスワードをエンコードすることで、パスワードの値を難読化できます。次に例を示します。

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

4. Cloudera Manager にログインし、Analytic Server の構成設定 **ssl_cfg** を、正しい SSL 構成設定で更新します。次に例を示します。

   ```
   <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="/>
   ```

   注: 鍵ファイルおよびトラストストア・ファイルについては、絶対パスを使用してください。

5. Analytic Server の構成設定 **security_cfg** を、正しい LDAP 構成設定で更新します。例えば、**ldapRegistry** 要素の場合、**sslEnabled** 属性を true に設定し、**sslRef** 属性を defaultSSLConfig に設定します。

Kerberos の構成

Analytic Server は、Cloudera での Kerberos をサポートします。以下の各セクションでは、Kerberos が Analytic Server に適合するように適切に構成するための構成設定について説明します。

注: Analytic Server は、認証に使用する Kerberos 関連の値について HDFS 構成を検査します。

Analytic Server および Kerberos の設定

Kerberos が有効になっている Cloudera 環境に Analytic Server をインストールする場合は、以下の設定に注意してください。

- Analytic Server コンソールにログインするときに Kerberos 認証を有効にする場合は、「Kerberos」を「**Analytic Server** セキュリティー **(Analytic Server security)**」設定として選択します。「**Kerberos**」が「**Analytic Server** セキュリティー **(Analytic Server security)**」設定として選択されている場合、Analytic Server コンソールはデフォルトで Kerberos ログイン・モードになります。

- Kerberos が有効になっているデータベースに接続する場合は、「Kerberos」を「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」設定として選択します。「**Kerberos**」が「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」設定として選択されている場合、Analytic Server コンソールはデータベースに接続するときに Kerberos モードを使用します。
- 「**Kerberos** レルム名 **(Kerberos Realm Name)**」および「**KDC** ホスト **(KDC host)**」の設定が必要です。「**Kerberos** レルム名 **(Kerberos Realm Name)**」(**as.kdc.realms**) および「**KDC** ホスト **(KDC host)**」(**kdcserver**) の値は、Kerberos 鍵配布センター (KDC) サーバー上の krb5.conf ファイルにあります。

  コンマ文字で区切ると、複数のレルム名がサポートされます。指定された Kerberos レルム名はユーザー名に対応し、ユーザー名に関連付けられます。例えば、ユーザー名 UserOne@us.ibm.com および UserTwo@eu.ibm.com は、レルム us.ibm.com、eu.ibm.com に対応します。

  「**Kerberos** レルム名 **(Kerberos Realm Name)**」として複数のレルムを指定する場合は、Kerberos クロスレルム・トラストを構成する必要があります。Analytic Server コンソールのログイン・プロンプトに入力されるユーザー名は、レルム名の接尾辞なしで入力されます。このため、複数のレルムが指定されている場合は、ユーザーに「レルム **(Realms)**」ドロップダウン・リストが表示され、該当するレルムを選択できます。

  注**:** レルムが 1 つしか指定されていない場合は、Analytic Server へのサインイン時に「レルム **(Realms)**」ドロップダウン・リストは表示されません。

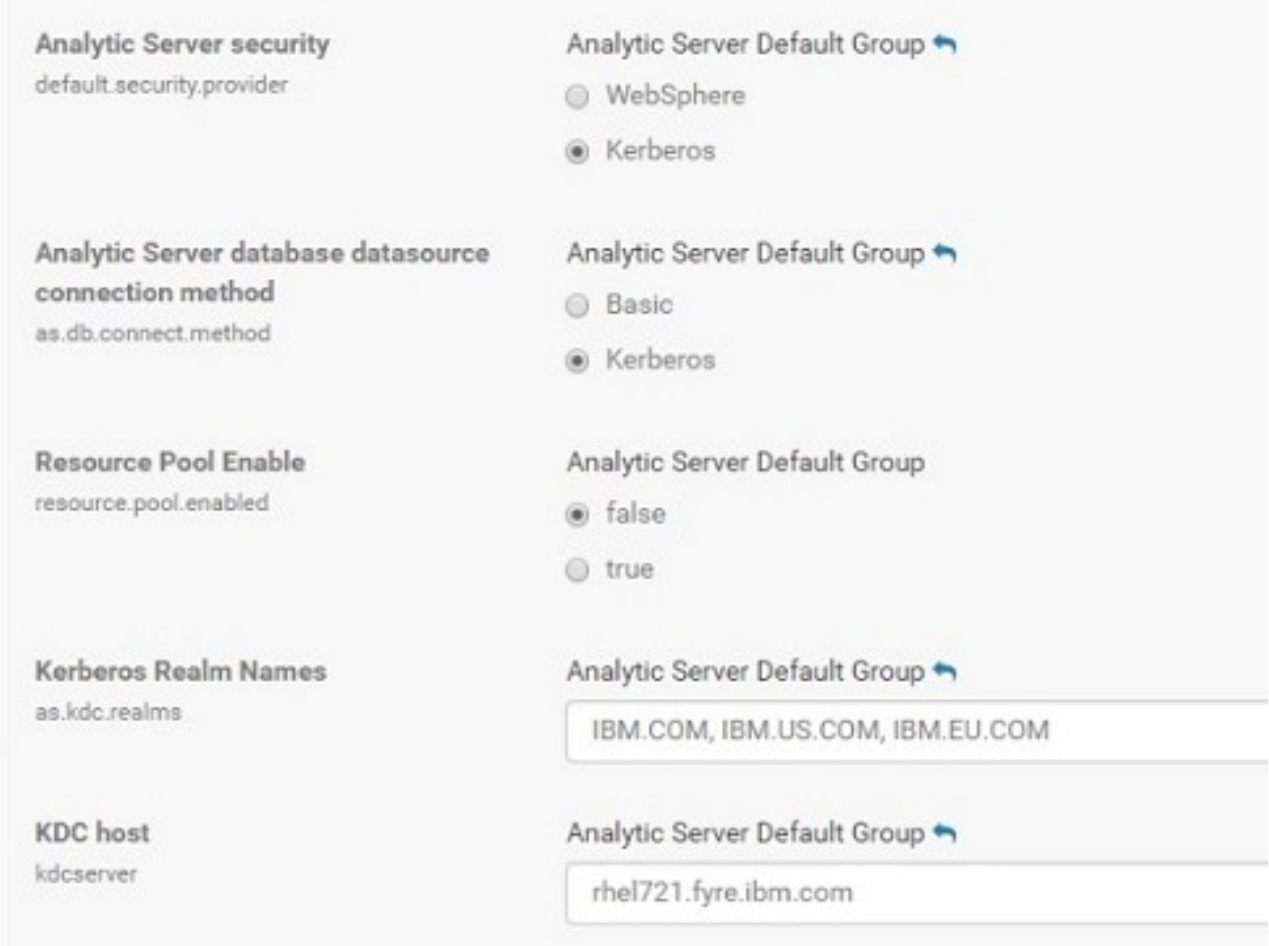

図 *7. Kerberos* 設定の例

注**:**

- 「**Analytic Server** セキュリティー **(Analytic Server security)**」および「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」の設定は、IBM SPSS Modeler クライアント認証および Analytic Server コンソール認証に適用されます。
- 「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」が「`Kerberos`」に設定されている場合は、ターゲット・データベースでも Kerberos が有効になっていることを確認する必要があります。
- 「**Analytic Server** セキュリティー **(Analytic Server security)**」および「**Analytic Server** データベース・データ・ソース接続方式 **(Analytic Server database data source connection method)**」の設定では、Hadoop クラスター上の Kerberos 認証は構成されません。詳しくは、『Kerberos 偽名の使用の有効化』セクションを参照してください。
- Kerberos 認証をログイン時に有効にする場合は、IBM SPSS Modeler クライアントを有効な Kerberos クライアントとしてデプロイする必要があります。これを行うには、Kerberos 鍵配布センター (KDC) サーバーで **`addprinc`** コマンドを使用します。詳しくは、IBM SPSS Modeler の資料を参照してください。

Kerberos での必要なアカウントの作成

1. Analytic Server へのアクセス権を付与する予定のすべてのユーザーについて、Kerberos ユーザー・リポジトリー内にアカウントを作成します。
2. LDAP サーバー上に (前のステップと) 同じアカウントを作成します。
3. 前のステップで、Analytic Server および Hadoop の各ノードで作成したそれぞれのユーザーについて、OS ユーザー・アカウントを作成します。
   - これらのユーザーの UID は、すべてのマシンで一致させてください。`kinit` コマンドを使用して各アカウントにログオンして、これをテストすることができます。
   - UID が、Yarn の「ジョブをサブミットするための最小ユーザー **ID (Minimum user ID for submitting job)**」設定に従っていることを確認してください。これは、`container-executor.cfg` 内の **`min.user.id`** 設定です。例えば、**`min.user.id`** が 1000 の場合、作成される各ユーザー・アカウントの UID は 1000 以上でなければなりません。
4. Analytic Server 管理者ユーザーについて、HDFS 上にユーザーのホーム・フォルダーを作成します。フォルダー・アクセス許可を `777` に、所有者を `admin` にして、さらにユーザー・グループを `hdfs` として設定する必要があります。以下の太字の例を参照してください。

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

Found 9 items

drwxrwxrwx   - hdfs     supergroup      0 2017-07-26 03:41 /user/AE
drwxrwxrwx   - admin    hdfs            0 2017-06-08 01:33 /user/admin
drwxr-x--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. HCatalog データ・ソースを使用する予定であり、Analytic Server が Hive Metastore とは別のマシンにインストールされている場合、HDFS で Hive クライアント名を使用する必要があります。
   a. Cloudera Manager で、HDFS サービスの「構成 (Configuration)」タブに移動します。

      注**:** 以下の設定がまだ設定されていない場合、「構成 **(Configuration)**」タブにそれらの設定が表示されない可能性があります。その場合は、検索を実行して見つけてください。

   b. **`hadoop.proxyuser.hive.groups`** 設定を編集して値 `*` を設定するか、すべてのユーザーが Analytic Server へのログインを許可されているグループを指定します。
   c. **`hadoop.proxyuser.hive.hosts`** 設定を編集して値 `*` を設定するか、サービスとして Hive Metastore および Analytic Server の各インスタンスがインストールされているホストのリストを指定します。
   d. HDFS サービスを再起動します。

これらのステップの実行を完了した後、Analytic Server がインストールされていると、Analytic Server がサイレントかつ自動的に Kerberos の構成を行います。

Kerberos 偽名の使用の有効化

偽名を使用すると、スレッドを所有しているプロセスのセキュリティー・コンテキストとは異なるセキュリティー・コンテキスト内で、そのスレッドを実行できます。例えば、偽名の使用は、標準 Analytic Server ユーザー (`as_user`) 以外のユーザーとして Hadoop ジョブを実行する手段を提供します。Kerberos 偽名の使用を有効にするには、以下を行います。

1. Cloudera Manager を開き、「**Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml**」領域 (「**HDFS (サービス全体) (HDFS (Service-Wide))**」 > 「構成 **(Configuration)**」タブにあります) で以下のプロパティーを追加するか、更新します。
   - 名前: `hadoop.proxyuser.as_user.hosts`
   - 値: `*`
   - 名前: `hadoop.proxyuser.as_user.groups`
   - 値: `*`

   注: **core-site.xml** 設定は、Hadoop 構成に適用されます (Analytic Server には適用されません)。
2. Analytic Server ノード上のコマンド・シェルから以下のコマンドを実行します。

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

Kerberos を使用したシングル・サインオン (SSO) 用の HAProxy の構成

1. HAProxy の資料 (http://www.haproxy.org/#docs) に従って HAProxy を構成して開始します。
2. HAProxy ホスト用の Kerberos プリンシパル (HTTP/<proxyHostname>@<realm>) およびキータブ・ファイルを作成します。ここで、<proxyHostname> は HAProxy ホストの完全な名前、<realm> は Kerberos レルムです。
3. キータブ・ファイルを各 Analytic Server ホストに `/etc/security/keytabs/spnego_proxy.service.keytab` としてコピーします。
4. このファイルのアクセス許可を各 Analytic Server ホストで更新します。次に例を示します。

   ```
   chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
   chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
   ```
5. Cloudera Manager を開き、Analytic Server の「**Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**」領域で以下のプロパティーを追加するか、更新します。

   ```
   web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
   web.authentication.kerberos.principal=HTTP/<proxy machine full name>@<realm>
   ```
6. 構成を保存し、Cloudera Manager からすべての Analytic Server サービスを再始動します。
7. Kerberos を使用するようにブラウザーを構成することをユーザーに指示します。

これで、ユーザーが IBM SPSS Analytic Server のログイン画面で「シングル・サインオンでのログイン **(Single sign on log in)**」オプションを使用して Analytic Server にログインできるようになりました。

Kerberos の無効化

1. Cloudera Manager コンソールで Kerberosを無効化します。
2. Analytic Server サービスを停止します。
3. 「**Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**」領域から以下の設定を削除します。

   ```
   default.security.provider
   hdfs.keytab
   hdfs.user
   java.security.krb5.conf
   as.db.connect.method
   web.authentication.kerberos.keytab
   web.authentication.kerberos.principal
   ```
4. 「変更の保存 **(Save Changes)**」をクリックし、Analytic Server サービスを再始動します。

Analytic Server コンソールへの Secure Sockets Layer (SSL) 接続の有効化

デフォルトでは、Analytic Server は自己署名証明書を生成して Secure Socket Layer (SSL) を有効にします。自己署名証明書を受け入れることにより、セキュア・ポートを使用して Analytic Server コンソールにアクセスできるようになります。HTTPS によるアクセスの安全性をさらに強化するには、サード・パーティー・ベンダーの証明書をインストールする必要があります。

サード・パーティー・ベンダーの証明書をインストールするには、以下のステップを実行します。

1. サード・パーティー・ベンダーの鍵ストア証明書およびトラストストア証明書を、すべての Analytic Server ノードで、同じディレクトリーにコピーします。例えば、`/home/as_user/security` です。

   注: Analytic Server ユーザーには、このディレクトリーの読み取りアクセス権限が必要です。
2. Cloudera Manager で、Analytic Server サービスの「構成 (Configuration)」タブに移動します。
3. **`ssl_cfg`** パラメーターを編集します。

```
<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>"/>
```

   置き換える値:

   - <KEYSTORE-LOCATION> に、鍵ストアの絶対位置を指定します。例: `/home/as_user/security/mykey.jks`
   - <TRUSTSTORE-LOCATION> に、トラストストアの絶対位置に指定します。例: `/home/as_user/security/mytrust.jks`
   - <TYPE> に、証明書のタイプを指定します。例: JKS、PKCS12、その他。
   - <PASSWORD> に、Base64 暗号化形式の暗号化パスワードを指定します。エンコードには、securityUtility を使用できます。例: `{AS_ROOT}/ae_wlpserver/bin/securityUtility encode <password>`

   自己署名証明書を生成する場合は、securityUtility を使用できます。例: `{AS_ROOT}/ae_wlpserver/bin/securityUtility createSSLCertificate --server=myserver --password=mypassword --validity=365 --subject=CN=mycompany,O=myOrg,C=myCountry`。securityUtility およびその他の SSL 設定について詳しくは、WebSphere Liberty Profile の資料を参照してください。

   注: `CN` 値に適切なホスト・ドメイン名を指定する必要があります。
4. 「変更の保存 **(Save Changes)**」をクリックし、Analytic Server サービスを再始動します。

SSL を介した Apache Hive との通信

SSL 接続を介して Apache Hive と通信するためには、`hive.properties` ファイルを更新する必要があります。あるいは、ご使用の Apache Hive 環境で高可用性が有効になっている場合は、メインの Analytic Server「データ・ソース」ページ上で高可用性パラメーターを選択できます。

hive.properties ファイルの更新

1. hive.properties ファイルを開きます。このファイルは、/opt/cloudera/parcels/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/database にあります。
2. 以下の行を見つけます。

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

3. 以下の太字の情報を追加して、行を更新します。

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

4. hive.properties ファイルを保存します。

Essentials for R に対するサポートの有効化

Analytic Server は、R モデルのスコアリング、および R スクリプトの実行をサポートしています。

Cloudera Manager で Analytic Server が正常にインストールされた後で Essentials for R をインストールするには、以下を行います。

1. Essentials for R のサーバー環境をプロビジョンします。詳しくは、 25 ページの『Essentials for R に対するサポートの有効化』のステップ 1 を参照してください。
2. IBM SPSS Modeler Essentials for R の RPM 用の自己解凍型アーカイブ (BIN) をダウンロードします。Essentials for R は、https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp からダウンロードできます。ご使用のスタック、スタックのバージョン、およびハードウェア・アーキテクチャーに固有のファイルを選択します。
3. Cloudera Manager サーバー・ホストで root ユーザーまたは sudo ユーザーとして自己解凍型アーカイブを実行します。以下のパッケージがインストール済みであるか、構成済みのリポジトリーから使用可能である必要があります。
   - Red Hat Linux: gcc-gfortran、zip、gcc-c++
   - SUSE Linux: gcc-fortran、zip、gcc-c++
   - Ubuntu Linux: gcc-fortran、zip、gcc-c++
4. この自己解凍型インストーラーは、以下のタスクを行います。
   a. 必要なライセンスを表示し、それに同意するようにインストーラーでプロンプトを出します。
   b. R のソースの場所を入力するか、またはデフォルトの場所で続行するようにインストーラーでプロンプトを出します。インストールされているデフォルトの R バージョンは 3.3.2 です。別のバージョンをインストールするには、以下を行います。
      - オンライン・インストール: 必要な R バージョン・アーカイブの URL を指定します。例えば、R 2.15.3 の場合は https://cran.r-project.org/src/base/R-2/R-2.15.3.tar.gz です。
      - オフライン・インストール: 必要な R バージョン・アーカイブをダウンロードし、Cloudera Manager サーバー・ホストにコピーします。アーカイブの名前を変更しないでください (デフォルトの名前は R-x.x.x.tar.gz です)。コピーした R アーカイブの URL を file://<R_archive_directory>/R-x.x.x.tar.gz のように指定します。R-2.15.3.tar.gz アーカイブをダウンロードして /root にコピーした場合、その URL は file:///root/R-2.15.3.tar.gz になります。

      注: その他の R バージョンは https://cran.r-project.org/src/base/ にあります。
   c. R が必要とするパッケージをインストールします。
   d. R および Essentials for R プラグインをダウンロードしてインストールします。

e. パーセルおよび parcel.sha ファイルを作成し、それらを /opt/cloudera/parcel-repo にコピーします。この場所が変更されている場合は、正しい場所を入力します。

5. インストールが完了したら、Cloudera Manager で「**Essentials for R**」パーセルを配布およびアクティブ化します (「新しいパーセルの確認 **(Check for New Parcels)**」をクリックしてパーセル・リストをリフレッシュします)。
6. Analytic Server サービスが既にインストールされている場合は、以下を行います。
   a. サービスを停止します。
   b. Analytic Server バイナリーをリフレッシュします。
   c. サービスを開始して Essentials for R のインストールを完了します。
7. Analytic Server サービスがインストールされていない場合は、そのインストールを進めます。

注**:** すべての Analytic Server ホストに適切なアーカイブ (zip および unzip) パッケージがインストールされている必要があります。

リレーショナル・データベース・ソースの有効化

各 Analytic Server Metastore と各 Analytic Server ホストの共有ディレクトリー内に JDBC ドライバーを配置すると、Analytic Server でリレーショナル・データベース・ソースを使用できます。デフォルトでは、このディレクトリーは /usr/share/jdbc です。

共有ディレクトリーを変更するには、以下のステップを実行します。

1. Cloudera Manager で、Analytic Server サービスの「構成 (Configuration)」タブに移動します。
2. **jdbc.drivers.location** で、JDBC ドライバーの共有ディレクトリーを指定します。
3. 「変更の保存 **(Save Changes)**」をクリックします。
4. 「アクション **(Action)**」ドロップダウンから「停止 **(Stop)**」を選択して Analytic Server サービスを停止します。
5. 「アクション **(Action)**」ドロップダウンから「**Analytic Server** バイナリーのリフレッシュ **(Refresh Analytic Server Binaries)**」を選択します。
6. 「アクション **(Action)**」ドロップダウンから「開始 **(Start)**」を選択して Analytic Server サービスを開始します。

表 *11.* サポート対象データベース

データベース	サポート対象バージョン	**JDBC** ドライバー **jar**	ベンダー
Amazon Redshift	8.0.2 以降	RedshiftJDBC41-1.1.6.1006.jar 以降	Amazon
Apache Impala	JDBC 4 (2.5.5 以降)	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	Bluemix サービス	db2jcc.jar	IBM

表 11. サポート対象データベース (続き)

データベース	サポート対象バージョン	**JDBC** ドライバー **jar**	ベンダー
Db2 for Linux、UNIX、および 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
MySQL	5.6, 5.7	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7、6.x	nzjdbc.jar	IBM
Oracle	12c, 11g R2 (11.2)	ojdbc6.jar、orai18n.jar	Oracle
SQL Server	2014, 2012, 2008 R2	sqljdbc4.jar	Microsoft
Teradata	15, 15.1	tdgssconfig.jar、terajdbc4.jar	Teradata

注

- Analytic Server をインストールする前に Redshift データ・ソースを作成した場合、Redshift データ・ソースを使用するには以下のステップを実行する必要があります。
  1. Analytic Server コンソールで Redshift データ・ソースを開きます。
  2. Redshift データベース・データ・ソースを選択します。
  3. Redshift のサーバー・アドレスを入力します。
  4. データベース名とユーザー名を入力します。パスワードは自動的に入力されます。
  5. データベース表を選択します。

HCatalog データ・ソースの有効化

Analytic Server は、Hive/HCatalog を介して複数のデータ・ソースをサポートしています。一部のソースでは、手動での構成ステップが必要です。

1. データ・ソースを有効にするために必要な JAR ファイルを収集します。詳しくは、下のセクションを参照してください。
2. これらの JAR ファイルを、各 Analytic Server Metastore と各 Analytic Server ノードの {HIVE_HOME}/auxlib ディレクトリーおよび /usr/share/hive ディレクトリーに追加します。
3. Hive Metastore サービスを再起動します。
4. Analytic Server サービスの各インスタンスをすべて再起動します。

注:

Analytic Server HCatalog データ・ソースを経由して HBase データにアクセスする場合、アクセスするユーザーは、HBase 表に対する読み取り権限を持っている必要があります。

- Kerberos 以外の環境では、Analytic Server は as_user を使用して HBase にアクセスします (as_user は、HBase に対する読み取り権限を持っている必要があります).
- Kerberos 環境では、as_user とログイン・ユーザーの両方が、HBase 表に対する読み取り権限を持っている必要があります。

NoSQL データベース

Analytic Server は、ベンダーから Hive ストレージ・ハンドラーが提供されている任意の NoSQL データベースをサポートします。

Apache HBase および Apache Accumulo のサポートを有効にするために、追加のステップは必要ありません。

その他の NoSQL データベースについては、データベース・ベンダーに連絡して、該当するストレージ・ハンドラーおよび関連する jar を取得してください。

ファイル・ベース Hive 表

Analytic Server は、組み込みまたはカスタムの Hive SerDe (serializer-deserializer) が利用可能な任意のファイル・ベース Hive 表をサポートします。

XML ファイルを処理するための Hive XML SerDe は Maven の Central Repository (http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde) にあります。

MapReduce v2 ジョブ

「**Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**」領域内の **preferred.mapreduce** 設定を使用して、MapReduce ジョブの処理方法を制御します。

表 12. *Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties*

プロパティー	説明
preferred.mapreduce	MapReduce ジョブが実行される方法を制御します。有効な値は以下のとおりです。 • spark • m3r • hadoop 例: preferred.mapreduce=spark

Apache Spark

Spark (バージョン 1.5 以降) を使用する場合は、Analytic Server インストール時に spark.version を選択する必要があります。

1. Cloudera Manager を開き、「**Analytic Server Spark Version**」領域で適切な spark.version (None、1.x、2.x など) を選択します。

   注**:** Spark 1.x を使用している場合は、「**Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**」領域内に以下の行を追加する必要もあります。

   ```
   spark.extraListeners=org.apache.spark.JavaSparkListener
   ```

2. 構成を保存します。

Apache Impala の構成

Apache Impala は、(Impala が SSL 対応であるかどうかにかかわらず) Analytic Server データベース・データ・ソースまたは HCatalog データ・ソースに対して Cloudera 上で実行する場合にサポートされます。

Apache Impala データのデータベース・データ・ソースの作成

1. メインの Analytic Server「データ・ソース」ページで、「新規」をクリックして新規データ・ソースを作成します。「新規データ・ソース (New data source)」ダイアログが表示されます。
2. 「新規データ・ソース **(New data source)**」フィールドに適切な名前を入力し、**「データベース」**を「内容タイプ **(Content type)**」値として選択して、「**Ok**」をクリックします。
3. 「データベース選択 **(Database Selections)**」セクションを開き、以下の情報を入力します。

   データベース：
   ドロップダウン・メニューから「**Impala**」を選択します。

   サーバー・アドレス **(Server address):**
   Impala デーモンをホストするサーバーの URL を入力します。Analytic Server に対して Kerberos が有効になっている場合、完全修飾ドメイン名が必要です。

   サーバー・ポート**:**
   Impala データベースが listen するポートの番号を入力します。

   データベース名**:**
   接続先データベースの名前を入力します。

   ユーザー名**:**
   Impala データベースにログインする権限を持つユーザー名を入力します。

   パスワード**:**
   適切なユーザー名のパスワードを入力します。

   表名**:** 使用するデータベースの表の名前を入力します。「選択」をクリックしてファイルを手動で選択します。

   最大同時読み取り数 **(Maximum concurrent reads):**
   データ・ソースで指定された表からデータを読み込むために、Analytic Server からデータベースに送信することができる同時クエリー数の制限を入力します。
4. 必要な情報を入力した後、「保存」をクリックします。

Apache Impala データの HCatalog データ・ソースの作成

1. メインの Analytic Server「データ・ソース」ページで、「新規」をクリックして新規データ・ソースを作成します。「新規データ・ソース (New data source)」ダイアログが表示されます。
2. 「新規データ・ソース **(New data source)**」フィールドに適切な名前を入力し、「HCatalog」を「内容タイプ **(Content type)**」値として選択して、「**Ok**」をクリックします。
3. 「データベース選択 **(Database Selections)**」セクションを開き、以下の情報を入力します。

   データベース：
   ドロップダウン・メニューから「デフォルト」を選択します。

   表名**:** 使用するデータベースの表の名前を入力します。

   HCatalog スキーマ
   「**HCatalog** の要素 **(HCatalog Element)**」オプションを選択し、適切な「**HCatalog** のフィールドのマッピング **(HCatalog Field Mappings)**」オプションを選択します。
4. 必要な情報を入力した後、「保存」をクリックします。

Apache Impala 有効データへの接続

1. Cloudera Manager コンソールで以下の Impala SSL 設定を定義します。

Impala の TLS/SSL を有効にする (client_services_ssl_enabled)
「**Impala (サービス全体) (Impala (Service-Wide))**」オプションを選択します。

Impala TLS/SSL サーバー証明書ファイル (PEM 形式) (Impala TLS/SSL Server Certificate File (PEM Format)) (ssl_server_certificate)
PEM 形式の自己署名証明書の場所とファイル名を入力します (例: `/tmp/<user_name>/ssl/ll4200v21.crt`)。

Impala TLS/SSL サーバー秘密鍵ファイル (PEM 形式) (Impala TLS/SSL Server Private Key File (PEM Format)) (ssl_private_key)
PEM 形式の秘密鍵の場所とファイル名を入力します (例: `/tmp/<user_name>/ssl/ll4200v21.key`)。

2. Analytic Server ホストで、`*.crf` ファイル (Impala SSL を有効にするために使用される) を `*.jks` ファイルにインポートします。このファイルは、cacerts ファイル (`/etc/pki/java/cacerts` など) であっても他の `*.jks` ファイルであっても構いません。
3. Analytic Server ホストで、以下の `jdbcurl` キー値を付加することにより、Impala 構成ファイル (`impala.properties`) を更新します。

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

   注: `*.jks` ファイル (cacerts 以外) を使用した場合、以下を指定する必要もあります。

   ```
   SSLTrustStore=<your_pks_file>;SSLTrustStorePwd=<password_for_pks_file>;
   ```
4. Cloudera Manager コンソールで Analytic Server を再始動します。

Analytic Server で使用するポートの変更

デフォルトでは、Analytic Server はポート 9080 (HTTP 用) および 9443 (HTTPS 用) を使用します。ポートの設定を変更するには、以下のステップを実行します。

1. Cloudera Manager で、Analytic Server サービスの「構成 (Configuration)」タブに移動します。
2. 使用する HTTP ポートおよび HTTPS ポートを、それぞれ **http.port** パラメーターおよび **https.port** パラメーターに指定します。

   注: これらのパラメーターを表示するには、「フィルター (Filters)」セクションで「ポートおよびアドレス **(Ports and Addresses)**」カテゴリーを選択する必要がある場合があります。
3. 「変更の保存 **(Save Changes)**」をクリックします。
4. Analytic Server サービスを再始動します。

高可用性 Analytic Server

クラスター内の複数のノードに Analytic Server をサービスとして追加することにより、高可用性構成にすることができます。

1. Cloudera Manager で、Analytic Server サービスの「インスタンス (Instances)」タブに移動します。
2. 「役割インスタンスの追加 **(Add Role Instances)**」をクリックし、 Analytic Server をサービスとして追加するホストを選択します。

複数クラスターのサポート

複数クラスター機能は、IBM SPSS Analytic Server の高可用性機能の拡張であり、複数テナント環境での独立性を強化します。デフォルトでは、(Ambari または ClouderaManager のいずれかで) Analytic Server サービスをインストールすると、結果として、単一の Analytic Server クラスターが定義されます。

クラスター仕様では、Analytic Server クラスター・メンバーシップが定義されます。クラスター仕様の変更は、(Ambari Analytic Server 構成の analytics-cluster フィールドで、または Cloudera Manager の configuration/analytics-cluster.xml ファイルを手動で編集して) XML コンテンツを使用して実行されます。複数の Analytic Server クラスターを構成する際は、それぞれの Analytic Server クラスターに独自のロード・バランサーを提供する必要があります。

複数クラスター機能を使用することで、あるテナントに対する作業が、別のテナントのクラスターで実行されている作業にマイナスの影響を与えることがなくなります。高可能性ジョブについては、ジョブのフェイルオーバーは、タスクが開始された Analytic Server クラスターの範囲内のみで発生します。以下の例は、複数クラスター XML 仕様を提供します。

注: クラスター内の複数のノードに Analytic Server をサービスとして追加することにより、それを高可用性にすることができます。

```
<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>
```

上記の例では、2 つのロード・バランサーが必要です。一方のロード・バランサーは cluster1 のメンバー (one.cluster および two.cluster) に要求を送信し、もう一方のロード・バランサーは cluster2 のメンバー (three.cluster および four.cluster) に要求を送信します。

以下の例は、単一クラスター XML 仕様 (デフォルト構成) を提供します。

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

上記の例では、構成されたクラスター・メンバーが複数ある場合に対応するために、1 つのロード・バランサーが必要です。

注

- シングルトン・クラスターのみが、**memberName** 要素でのワイルドカードの使用をサポートしています (例えば、クラスター・カーディナリティー = "1")。カーディナリティー要素の有効な値は、1 および 1+ です。
- **memberName** は、Analytic Server 役割の割り当て先のホスト名と同じように指定する必要があります。
- クラスター構成の変更が適用された後は、すべてのクラスター内のすべてのサーバーを再起動する必要があります。
- Cloudera Manager では、すべての Analytic Server ノードの analytics-cluster.xml ファイルを変更して維持する必要があります。すべてのノードが同じ内容を含むように維持する必要があります。

スモールデータ向けの JVM オプションの最適化

小規模な (M3R) ジョブの実行時にご使用のシステムを最適化するために、JVM プロパティーを編集できます。

Cloudera Manager で、Analytic Server サービスの「構成 (Configuration)」タブの「**Jvm オプション (jvm.options)**」コントロールを確認します。以下のパラメーターを変更して、Analytic Server (Hadoop ではなく) をホストするサーバーで実行されるジョブのヒープ・サイズを設定します。これは小規模な (M3R) ジョブを実行する場合に重要です。システムを最適化するために、これらの値を調整する必要がある場合があります。

```
-Xms512M
-Xmx2048M
```

IBM SPSS Analytic Server テナントごとに別個の YARN キューの構成 - Cloudera

Yarn キューの構成は、Spark 動的リソース割り振り技術を使用して行われます。

Cloudera 5.x

SPSS Analytic Server サービスを既存のクラスターに追加する場合は、以下のステップを実行します。

1. Cloudera Manager で、「**SPSS Analytic Server サービス (SPSS Analytic Server Service)**」 > 「構成」に移動します。
2. 「**Resource Pool Enable: resource.pool.enabled**」値を `true` に変更します。
3. 以下のプロパティーを「**Analytic Server Advanced Configuration Snippet (Safety Valve))**」 > 「**analyticserver-conf.config.properties**」に追加します。

```
config.folder.path=/etc/spark2/conf
resource.pool.mapping=tenant1:test,tenant2:production
resource.pool.default=default
spark.scheduler.mode=FAIR
spark.yarn.queue=default
```

表 *13. analyticserver-conf.config.properties* 設定

プロパティー	説明
`config.folder.path`	このディレクトリーには、Spark プール・プロパティー情報を含む `fairscheduler.xml` ファイルが含まれています。このファイルは必須であり、手動で作成する必要があります。詳しくは、『**fairscheduler.xml** の例』セクションを参照してください。
`resource.pool.mapping`	**Spark:** `fairscheduler.xml` ファイル内で定義されているプールにテナントをマップします。テナントのペアは、コンマで区切る必要があります (例えば、`tenant1:test,tenant2:production`)。プールを指定する前に、そのプールが `fairscheduler.xml` ファイル内で構成されていることを確認してください。 **MapReduce:** 動的リソース・プール構成内で定義されているキューにテナントをマップします。テナントのペアは、コンマで区切る必要があります (例えば、tenant1:test,tenant2:production)。キューを指定する前に、そのキューを使用してシステムが構成されていること、およびジョブをキューにサブミットするためのアクセスが許可されていることを確認してください。 **注:** Spark ジョブと MapReduce ジョブを両方とも実行する場合、テナント・マップ値は、`fairscheduler.xml` ファイル内および動的リソース・プール構成内で同じ名前である必要があります。
`resource.pool.default`	**Spark:** デフォルト・リソース・プールを定義します。この値は、`default`、または `fairscheduler.xml` ファイル内で定義されているプール名にすることができます。テナントが構成されていない (または間違って構成されている) 場合は、`default` 設定を使用します。 **MapReduce:** ジョブのサブミット先となるデフォルト・キューを定義します。

表 13. *analyticserver-conf.config.properties* 設定 (続き)

プロパティー	説明
spark.scheduler.mode=FAIR	**Spark:** fair scheduler を有効にします。このプロパティーを変更しないでください。
spark.yarn.queue	**Spark:** アプリケーションのサブミット先となる YARN キューの名前。動的リソース・プール構成内のカスタマイズされた YARN キュー名を指定できます。

4. 構成を保存し、Analytic Server サービスを再始動します。

fairscheduler.xml の例

fairscheduler.xml ファイルには、Spark プール・プロパティー情報が含まれています。このファイルは必須であり、手動で作成する必要があります。

```
<?xml version="1.0"?>
<allocations>
  <pool name="production">
    <schedulingMode>FAIR</schedulingMode>
    <weight>1</weight>
    <minShare>2</minShare>
  </pool>
  <pool name="test">
    <schedulingMode>FIFO</schedulingMode>
    <weight>2</weight>
    <minShare>3</minShare>
  </pool>
</allocations>
```

参照情報

詳しくは、以下のサイトを参照してください。

- https://spark.apache.org/docs/latest/job-scheduling.html#dynamic-resource-allocation
- https://spark.apache.org/docs/latest/running-on-yarn.html

マイグレーション

Analytic Server では、既存の Analytic Server インストール済み環境から新規インストール済み環境へのデータおよび構成設定のマイグレーションが可能です。

Analytic Server の新規バージョンへのアップグレード

Analytic Server 3.1.2 の既存のインストール済み環境がある状態で新規バージョンを購入した場合、3.1.2 の構成設定を新規インストール済み環境にマイグレーションできます。

制約事項**:** 3.1.2 インストール済み環境と新規インストール済み環境は、同じ Hadoop クラスター内には共存できません。3.1.2 インストール済み環境と同じ Hadoop クラスターを使用するように新規インストール済み環境を構成すると、3.1.2 インストール済み環境は動作しなくなります。

3.1.2 から新規バージョンへのマイグレーション・ステップ

1. 46 ページの『Cloudera でのインストール』 の手順に従って、Analytic Server の新規インストールを実行します。
2. 以前のインストール済み環境から新規インストール済み環境に Analytic ワークスペースをコピーします。
   a. Analytic ワークスペースの場所が不明な場合は、hadoop fs -ls を実行します。Analytic ワークスペースのパスの形式は /user/as_user/analytic-root/analytic-workspace です。ここで、as_user は、Analytic ワークスペースを所有するユーザー ID です。

b. Analytic Server の新規インストール済み環境のホストに as_user としてログインします。/user/as_user/analytic-root/analytic-workspace ディレクトリーが存在する場合は削除します。

c. 以下のコピー・スクリプトを実行します。

```
hadoop distcp hftp://{host of 3.1.2 namenode}:50070/{path to 3.1.2 analytic-workspace}
hdfs://{host of 3.2.1 namenode}/user/as_user/analytic-root/analytic-workspace
```

3. 組み込み Apache Directory Server を使用する場合は、サード・パーティー LDAP クライアント・ツールを使用して現在のユーザー/グループ構成をバックアップします。Analytic Server 3.2.1 がインストールされた後で、バックアップのユーザー/グループ構成を Apache Directory Server にインポートします。

   注: 外部 LDAP サーバーを使用する場合は、このステップをスキップできます。

4. Cloudera Manager で、Analytic Server サービスを停止します。
5. 古いインストール済み環境から構成設定を収集します。
   a. 新規インストール済み環境の configcollector.zip アーカイブを、古いインストール済み環境の {AS_ROOT}¥tools にコピーします。
   b. コピーした configcollector.zip を解凍します。これにより、古いインストール済み環境内に新規の configcollector サブディレクトリーが作成されます。
   c. {AS_ROOT}¥tools¥configcollector 内の **configcollector** スクリプトを実行して、古いインストール済み環境内の構成収集ツールを実行します。その結果生成された圧縮ファイル (ZIP) を、新規インストール済み環境をホストするサーバーにコピーします。

   重要: 指定された **configcollector** スクリプトは、最新バージョンの Analytic Server と互換性がない場合があります。**configcollector** スクリプトに関する問題が発生した場合は、IBM 技術サポート担当員にお問い合わせください。

6. Zookeeper の状態をクリアします。Zookeper の bin ディレクトリー (例えば、Cloudera 上の /opt/cloudera/parcels/CDH-5.4...../lib/zookeeper/bin) で、以下のコマンドを実行します。

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

7. **migrationtool** スクリプトを実行し、構成収集ツールによって作成された圧縮ファイルのパスを引数として渡すことで、マイグレーション・ツールを実行します。次に例を示します。

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.1.2.xxx.zip
```

8. Analytic Server ノード上のコマンド・シェルから以下のコマンドを実行します。

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

9. Cloudera Manager で、Analytic Server サービスを開始します。

   注: 既存の Analytic Server インストール済み環境で使用するように R を構成している場合、新規 Analytic Server インストール済み環境で R を構成するステップに従う必要があります。

Cloudera での Analytic Server のアンインストール

Cloudera は、Analytic Server のサービスおよびパーセルをアンインストールするために必要なステップの大部分を自動的に処理します。

Cloudera 環境から Analytic Server をクリーンアップするには、以下のステップを実行する必要があります。

1. Analytic Server サービスを停止してから削除します。
2. Analytic Server パーセルの非アクティブ化、ホストからの削除、および削除を行います。

3. HDFS の Analytic Server ユーザー・ディレクトリーを削除します。デフォルトの場所は `/user/as_user/analytic-root` です。
4. Analytic Server が使用するデータベース (スキーマ) を削除します。
5. Analytic Server インストール・パッケージの残りの内容をすべてクリーンアップします。これを行うには、以下を削除します。
   - `csd` フォルダー
   - `parcels`、`parcel-cache`、および `parcel-repo` フォルダー内にあるすべての既存の 3.2.1 ファイル

第 4 章 IBM SPSS Analytic Server で使用するための IBM SPSS Modeler の構成

SPSS Modeler を Analytic Server で使用できるようにするには、SPSS Modeler Server インストール済み環境に対する更新をいくつか行う必要があります。

1. SPSS Modeler Server を構成して、Analytic Server インストール済み環境と関連付けます。
   a. メインサーバーのインストール・ディレクトリーの config サブディレクトリーにある options.cfg ファイルを編集して、以下の行を追加または編集します。

```
as_ssl_enabled, {Y|N}
as_host, "{AS_SERVER}"
as_port, PORT
as_context_root, "{CONTEXT-ROOT}"
as_tenant, "{TENANT}"
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

Analytic Server でセキュア通信が構成されている場合は Y を指定して、それ以外の場合には N を指定してください。

as_host

Analytic Server をホストするサーバーの IP アドレス/ホスト名。

注: Analytic Server で SSL が有効になっている場合は、適切な IP アドレス/ホスト・ドメイン名を指定する必要があります。

as_port

Analytic Server が listen するポート (デフォルトは 8080)。

as_context_root

Analytic Server コンテキスト・ルート (デフォルトは analyticserver)。

as_tenant

SPSS Modeler Server インストール済み環境がメンバーになっているテナント (デフォルトのテナントは ibm)。

as_prompt_for_password

SPSS Modeler Server が、Analytic Server で使用されているユーザーおよびパスワードの認証システムと同じ認証システムを使用して構成されている場合 (例えば、Kerberos 認証を使用している場合) は N を指定します。そうでない場合は、Y を指定します。

SPSS Modeler をバッチ・モードで実行している場合、clemb コマンドの引数として -analytic_server_username {ASusername} -analytic_server_password {ASpassword} を追加します。

as_kerberos_auth_mode

SPSS Modeler からの Kerberos SSO を有効にする場合は Y を指定します。

as_kerberos_krb5_conf
Analytic Server で使用する Kerberos 構成ファイルへのパスを指定します (例: ¥etc¥krb5.conf)。

as_kerberos_krb5_spn
Analytic Server Kerberos SPN を指定します (例: HTTP/ashost.mydomain.com@MYDOMAIN.COM)。

b. SPSS Modeler Server サービスを再開します。

SSL/TLS が有効になっている Analytic Server インストール済み環境に接続するには、SPSS Modeler Server とクライアントのインストール済み環境を構成するための追加のステップがいくつかあります。

a. http{s}://{HOST}:{PORT}/{CONTEXT-ROOT}/admin/{TENANT} にナビゲートして、Analytic Server コンソールにログオンします。

b. ブラウザーから認証ファイルをダウンロードして、ファイル・システムに保存します。

c. 認証ファイルを SPSS Modeler Server と SPSS Modeler Client の両方のインストール済み環境の JRE に追加します。更新する場所は、SPSS Modeler インストール・パスの /jre/lib/security/cacerts サブディレクトリーで見つかります。

1) cacerts ファイルが読み取り専用でないことを確認します。

2) Modeler に付属の keytool プログラムを使用します。これは、SPSS Modeler インストール・パスの /jre/bin/keytool サブディレクトリーにあります。

次のコマンドを実行します。

```
keytool -import -alias <as-alias> -file <cert-file> -keystore "<cacerts-file>"
```

<as-alias> は cacerts ファイルの別名であることに注意してください。 cacerts ファイルに固有のものである限り、任意の名前を使用できます。

以下にコマンドの例を示します。

```
keytool -import -alias MySSLCertAlias -file C:¥Download¥as.cer
        -keystore "c:¥Program Files¥IBM¥SPSS¥Modeler¥{ModelerVersion}¥jre¥lib¥security¥cacerts"
```

d. SPSS Modeler Server および SPSS Modeler Client を再起動します。

2. [オプション] Analytic Server データ・ソースを使用してストリーム内の R モデルのスコアリングを行う予定の場合、IBM SPSS Modeler - Essentials for R をインストールします。IBM SPSS Modeler - Essentials for R は、https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp からダウンロードできます。

第 5 章 UDF Hive プッシュバックの構成

プッシュバックが可能なすべての IBM SPSS Analytic Server ノードは、可能な限り、UDF Hive にプッシュパックします。Hive UDF が HiveDB に登録されると、Analytic Server は新しい UDF 機能を使用してプッシュバックを実行できます。

UDF Hive は、デフォルトでは無効になっているため、ASModules.xml ファイル内の **udfmodule** 設定を使用して、手動で有効にする必要があります (**disabled** 値を **enabled** に変更)。設定を有効にした後、Analytic Server を再始動し、UDF を Hive に手動で登録する必要があります。

HDP 環境と Cloudera 環境で、UDF を Hive に登録/登録解除する方法を次の例に示します。

HDP 上での UDF の登録/登録解除

UDF の登録

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

UDF の登録解除

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

Cloudera 上での UDF の登録/登録解除

UDF の登録

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

UDF の登録解除

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

第 6 章 SLM タグを使用したライセンス交付の追跡

SLM タグは、Resource Utilization Measurement の ISO/IEC 19770-4 規格のドラフトに基づいています。SLM タグは、製品がライセンス・メトリック (ソフトウェア資産の使用に関連するリソース) の使用量を報告するための標準化機能を提供します。製品の SLM を有効にすると、そのライセンス使用を自己報告するためのランタイム XML ファイルが生成されます。

Analytic Server が開始されると、slmtag ファイルが <as_installation_path>/logs/slmtag フォルダー内に作成されます。

2 つのライセンス・タイプがあるため、異なる 2 つのメトリックが定期的に記録されます。

- Analytic Server の現行バージョンの場合、ライセンス交付は、(仮想サーバーに基づいた) Hadoop クラスター内のデータ・ノードの総数によって決まります。ノード数は以下の slmtag ファイル・セクションに記録されます。

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

- Analytic Server の 3.1 より前のバージョンの場合、ライセンス交付は、(RVU に基づいた) Hadoop クラスター内の HDFS ストレージのサイズによって決まります。例えば、ストレージ・サイズ (単位: テガバイト) は以下の slmtag ファイル・セクションに記録されます。

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

SLM タグ出力がスレッド内で開始されますが、それは SlmTagOutput.properties ファイル内で定義されたプロパティーの影響を受けます。このファイルは、<as_installation_path>/configuration フォルダー内にあります。

表 *14. SLM* タグのプロパティー

プロパティー	説明
license.metric.logger.output.enabled	SLM ログ・ファイル生成を制御します。デフォルト値は False です。
license.metric.logger.output.dir	SLM タグ・ファイルを格納するディレクトリーの相対パス。デフォルト・ディレクトリーは <as_installation_path>/logs です。
license.metric.logger.output.SLMLogFrequency	SLM ログ収集の時間間隔 (単位: ミリ秒)。
icense.metric.logger.file.size	SML タグ・ファイルの最大サイズ (単位: バイト)。
license.metric.logger.file.number	1 つのソフトウェア ID インスタンスに対する SLM タグ・ファイルの最大数。

第 7 章 トラブルシューティング

このセクションでは、インストールおよび構成の一般的な問題とその解決方法を説明します。

一般的な問題

インストールが警告付きで成功するが、「要求を完了できません。理由: 権限が拒否されました **(Permission denied)**」のエラーが発生してユーザーがデータ・ソースを作成できない

distrib.fs.root パラメーターを Analytic Server ユーザー (デフォルトでは as_user) がアクセス権限を持たないディレクトリーに設定すると、エラーが発生します。Analytic Server ユーザーが **distrib.fs.root** ディレクトリーに対して読み取り、書き込み、および実行を許可されるようにしてください。

Analytic Server のパフォーマンスが徐々に低下している

Analytic Server のパフォーマンスが予期されるレベルに到達しない場合は、Knox サービス・デプロイメント・パス `/<KnoxServicePath>/data/deployments` からすべての `*.war` ファイルを削除します。例: `/usr/hdp/3.1.0.0-78/knox/data/deployments`

Ambari での **Analytic Server** または **Essentials for R** のアンインストール

場合によっては、Ambari で Analytic Server または Essentials for R をアンインストールするときに、アンインストール・プロセスがハングすることがあります。この問題が発生した場合は、Ambari サーバーのプロセス ID を手動で停止する必要があります。

OpenJDK を使用する **POWER Systems** に **Analytic Server** をインストールする際の問題

OpenJDK を使用する POWER Systems で Analytic Server を実行する際は、以下の構成ステップを手動で実行して、座標系 API が予期されたとおりに機能することを確認する必要があります。

注: 座標系 API を使用しない場合は、構成要件を無視して構いません。

1. Ambari コンソールで、「**Analytic Server service**」 > 「**Configs tab**」 > 「**Advanced analytics-jvm-options**」に移動し、以下の行をコンテンツ領域に追加します。

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

2. Ambari コンソールで、「**Custom analytics.cfg**」セクションに移動し、以下の 3 つの構成を追加します。

   spark.executor.extraJavaOptions

   値を `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*` に設定します。

   spark.driver.extraJavaOptions

   値を `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*` に設定します。

   mapred.child.java.opts

   値を `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*` に設定します。

SuSE Linux 12 に **Analytic Server** をインストールするときのエラー

SuSE Linux 12 に Analytic Server をインストールするときに、以下のエラーが発生することがあります。

```
Signature verification failed [4-Signatures public key is not available]
```

この問題は、SuSE Linux 12 に Analytic Server をインストールする前に、以下のタスクを実行することによって解決できます。

1. 以下の URL から、ご使用のホストに公開鍵をダウンロードします。

   ```
   https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.1.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
   ```

2. ご使用のホストで以下のコマンドを実行して、公開鍵をインポートします。

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

特定の Hadoop ディストリビューションに関する問題

Analytic Server サービスに対するリフレッシュ・アクションが **Hortonworks 2.3 - 2.6** で無効になっている

Hortonworks 2.3 - 2.6 上の Analytic Server ライブラリーを手動でリフレッシュするには、以下のステップを使用します。

1. Analytic Metastore を実行しているホストに Analytic Server ユーザー (デフォルトでは as_user) としてログオンします。

   注: このホスト名は Ambari コンソールから確認できます。

2. {AS_ROOT}/bin ディレクトリーにある **refresh** スクリプトを実行します。例えば、次のようにします。

   ```
   cd /opt/ibm/spss/analyticserver/3.2/bin
   ./refresh
   ```

3. Ambari コンソールで Analytic Server サービスを再始動します。

外部サイトからダウンロードされたパッケージが、**Cloudera Manager** 内のハッシュ検査で不合格になる

ハッシュ検査エラーがパーセル・リストに表示されます。ダウンロード・プロセスが完了するまで待機してから、cloudera-scm-server サービスを使用して Cloudera を再始動することにより、この問題を解決できます。サービスが再始動されると、エラーは発生しなくなります。

HDFS supergroup プロパティー

as_user が HDFS グループ・プロパティー **dfs.permissions.supergroup/ dfs.permissions.superusergroup** のメンバーでない場合、Analytic Server は始動中に例外をログに記録します。以下に例を示します。

```
[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:694)
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)
```

hdfs-site 構成プロパティー **dfs.permissions.supergroup/dfs.permissions.superusergroup** で定義されている OS グループに as_user を手動で追加する必要があります。

- Cloudera の場合、デフォルトのプロパティー値は **supergroup** であり、実際に存在する OS グループに変更する必要があります。Cloudera での supergroup 設定については、Cloudera の資料を参照してください。

- Ambari の場合、デフォルトのプロパティー値は **hdfs** です。デフォルトでは、Ambari のインストール中に、Analytic Server は、as_user を HDFS グループと Hadoop グループに追加します。

Linux で、**usermod** コマンドを使用して、HDFS **superusergroup** に as_user を追加します (まだ存在していない場合)。

HDFS アクセス許可に関する一般情報については、「HDFS Permissions Guide」を参照してください。

HDP 3.0 上で **MapReduce** ジョブが失敗する

HDP 3.0 上の MapReduce ジョブで次のエラーが発生することがあります。

```
要求を完了できません。
理由: java.lang.IllegalStateException: ジョブの状態が RUNNING ではなく DEFINE です (Job in state DEFINE instead of RUNNING) (as_trace.log)
```

エラー状態は、次のように解決できることがあります。

1. Custom analytics.cfg ファイルに次の構成を追加します。

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

2. Analytic Server を再始動します。

Analytic Server の再始動後、MapReduce ジョブは正常に実行されます。

メタデータ・リポジトリーに関する問題

add_mysql_user スクリプトの実行時に **CREATE USER** 操作が失敗する

add_mysql_user スクリプトを実行する前に、まず、追加しようとしているユーザーを mysql データベースから手動で削除する必要があります。MySQL Workbench UI または MySQL コマンドによってユーザーを削除できます。以下に例を示します。

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

上記のコマンドで、削除するユーザー名で $AEDB_USERNAME_VALUE を置換し、データベースがインストールされているホスト名で $METASTORE_HOST を置換してください。

Apache Spark に関する問題

Spark プロセス内で実行されるストリームに関する問題

Spark プロセス内で強制的に実行された SPSS Modeler ストリームを完了できませんでした。失敗した SPSS Modeler ストリームは、Analytic Server ソース・ノード (HDFS ファイル) を使用して作成されていました。このソース・ノードは、Sort ノードにリンクされていて、さらにその後に別の Analytic Server データ・ソースにエクスポートするように設定されています。ストリームが実行された後で、リソース・マネージャー・ユーザー・インターフェースは、新規アプリケーションが実行されているが、ストリームはいつまでも完了せず Running 状態のままであることを示します。Analytic Server ログ、YARN ログ、および Spark ログには、ストリームを完了できなかった理由を示すメッセージはありません。

Analytic Server 構成内の Custom analytics.cfg ファイルに spark.executor.memory 設定を追加することにより、この問題を解決できます。メモリー値を 4GB に設定すると、以前に失敗した SPSS Modeler ストリームを 2 分未満で完了できます (単一ノード・クラスター環境の場合)。

Cloudera 5.x および **Spark 1.x** で **Spark** ジョブを実行できない

Cloudera 5.x および Spark 1.x を使用している場合に、以下の例外が発生することがあります。

```
org.apache.spark.SparkException: Exception when registering SparkListener
```

この例外は、java.lang.ClassCastException: com.cloudera.spark.lineage.ClouderaNavigatorListener が org.apache.spark.scheduler.SparkListener をキャストできないために発生します。

この例外を回避するには、「**Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties**」領域内に以下の行を追加する必要があります。

```
spark.extraListeners=org.apache.spark.JavaSparkListener
```

エラー「**HdfsAuthcom.spss.utilities.i18n.LocException** でエラーが発生しました **(Exception during HdfsAuthcom.spss.utilities.i18n.LocException)**。実行が失敗しました。理由: **com.spss.ae.filesystem.exception.FileSystemException:** ファイル・システムへのアクセスを初期化できません。」が **SparkML** ケースの実行中に発生する。

このエラーは、Spark の lineage ログ・ディレクトリーが見つからないときに発生します。この問題の回避策は、spark.lineage.log.dir を /ae_wlpserver/usr/servers/aeserver/logs/spark にリダイレクトすることです。

高可用性クラスター

依存関係の変更が原因で **Analytic Server** を追加ホストに追加できない

32 ページの『クライアント依存関係の更新』の手順に従って update_clientdeps スクリプトを実行します。

「分析クラスター・サービスと **Zookeeper** との接続が予期せず切断されました。クラスターの整合性を保つため、この **JVM** を終了しています。」

この状態が発生する可能性のある原因の 1 つとして、Zookeeper に書き込まれるデータの量が多すぎることがあります。 Zookeeper のログに以下のような例外がある場合:

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

あるいは、Analytic Server のログに以下のようなメッセージがある場合:

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

1. Ambari コンソールで、Zookeeper サービスの「Configs」タブにナビゲートし、以下の行を env-template に追加してから、Zookeeper サービスを再始動します。

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

2. Ambari コンソールで、Analytic Server サービスの「Configs」タブに移動し、以下の行を Advanced analytics-jvm-options に追加してから、Analytic クラスター・サービスを再始動します。

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

jute.maxbuffer 設定に対して指定する数値は、例外メッセージで示されている数値よりも大きくする必要があります。

Zookeeper のトランザクション・データが管理不能になる

zoo.cfg の **autopurge.purgeInterval** パラメーターを 1 に設定して、Zookeeper トランザクション・ログの自動消去を有効にします。

Analytic クラスター・サービスが **Zookeeper** との接続を失う

zoo.cfg の **tickTime**、**initLimit**、および **syncLimit** の各パラメーターを確認して変更します。以下に例を示します。

```
# The number of milliseconds of each tick
tickTime=2000
# The number of ticks that the initial
# synchronization phase can take
initLimit=30
# The number of ticks that can pass between
# sending a request and getting an acknowledgement
syncLimit=15
```

詳細については、Zookeeper の資料 (https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html) を参照してください。

Analytic Server ジョブが再開されない

Analytic Server ジョブが再開されない一般的な状態があります。

- クラスター・メンバーの障害が原因で Analytic Server ジョブが失敗した場合、通常そのジョブは他のクラスター・メンバー上で自動的に再開されます。ジョブが再開されない場合、高可用性クラスター内に少なくとも 4 つのクラスター・メンバーが存在することを確認してください。

特記事項

本書は米国 IBM が提供する製品およびサービスについて作成したものです。この資料の他の言語版を IBM から入手できる場合があります。ただし、これを入手するには、本製品または当該言語版製品を所有している必要がある場合があります。

本書に記載の製品、サービス、または機能が日本においては提供されていない場合があります。日本で利用可能な製品、サービス、および機能については、日本 IBM の営業担当員にお尋ねください。本書で IBM 製品、プログラム、またはサービスに言及していても、その IBM 製品、プログラム、またはサービスのみが使用可能であることを意味するものではありません。これらに代えて、IBM の知的所有権を侵害することのない、機能的に同等の製品、プログラム、またはサービスを使用することができます。ただし、IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

IBM は、本書に記載されている内容に関して特許権 (特許出願中のものを含む) を保有している場合があります。本書の提供は、お客様にこれらの特許権について実施権を許諾することを意味するものではありません。実施権についてのお問い合わせは、書面にて下記宛先にお送りください。

〒*103-8510*
東京都中央区日本橋箱崎町*19*番*21*号
日本アイ・ビー・エム株式会社
法務・知的財産
知的財産権ライセンス渉外

IBM およびその直接または間接の子会社は、本書を特定物として現存するままの状態で提供し、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任を負わないものとします。国または地域によっては、法律の強行規定により、保証責任の制限が禁じられる場合、強行規定の制限を受けるものとします。

この情報には、技術的に不適切な記述や誤植を含む場合があります。本書は定期的に見直され、必要な変更は本書の次版に組み込まれます。IBM は予告なしに、随時、この文書に記載されている製品またはプログラムに対して、改良または変更を行うことがあります。

本書において IBM 以外の Web サイトに言及している場合がありますが、便宜のため記載しただけであり、決してそれらの Web サイトを推奨するものではありません。それらの Web サイトにある資料は、この IBM 製品の資料の一部ではありません。それらの Web サイトは、お客様の責任でご使用ください。

IBM は、お客様が提供するいかなる情報も、お客様に対してなんら義務も負うことのない、自ら適切と信ずる方法で、使用もしくは配布することができるものとします。

本プログラムのライセンス保持者で、(i) 独自に作成したプログラムとその他のプログラム (本プログラムを含む) との間での情報交換、および (ii) 交換された情報の相互利用を可能にすることを目的として、本プログラムに関する情報を必要とする方は、下記に連絡してください。

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
US

本プログラムに関する上記の情報は、適切な使用条件の下で使用することができますが、有償の場合もあります。

本書で説明されているライセンス・プログラムまたはその他のライセンス資料は、IBM 所定のプログラム契約の契約条項、IBM プログラムのご使用条件、またはそれと同等の条項に基づいて、IBM より提供されます。

記載されている性能データとお客様事例は、例として示す目的でのみ提供されています。実際の結果は特定の構成や稼働条件によって異なります。

IBM 以外の製品に関する情報は、その製品の供給者、出版物、もしくはその他の公に利用可能なソースから入手したものです。IBM は、それらの製品のテストは行っておりません。したがって、他社製品に関する実行性、互換性、またはその他の要求については確証できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給者にお願いします。

IBM の将来の方向または意向に関する記述については、予告なしに変更または撤回される場合があり、単に目標を示しているものです。

表示されている IBM の価格は IBM が小売り価格として提示しているもので、現行価格であり、通知なしに変更されるものです。卸価格は、異なる場合があります。

本書はプランニング目的としてのみ記述されています。記述内容は製品が使用可能になる前に変更になる場合があります。

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名前はすべて架空のものであり、名前や住所が類似する個人や企業が実在しているとしても、それは偶然にすぎません。

著作権使用許諾:

本書には、日常の業務処理で用いられるデータや報告書の例が含まれています。より具体性を与えるために、それらの例には、個人、企業、ブランド、あるいは製品などの名前が含まれている場合があります。これらの名前はすべて架空のものであり、名前や住所が類似する個人や企業が実在しているとしても、それは偶然にすぎません。

それぞれの複製物、サンプル・プログラムのいかなる部分、またはすべての派生的創作物にも、次のように、著作権表示を入れていただく必要があります。

商標

IBM、IBM ロゴおよび ibm.com は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては、http://www.ibm.com/legal/copytrade.shtml をご覧ください。

Adobe、Adobe ロゴ、PostScript、PostScript ロゴは、Adobe Systems Incorporated の米国およびその他の国における登録商標または商標です。

IT Infrastructure Library は AXELOS Limited の登録商標です。

インテル、Intel、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Centrino、Intel Centrino ロゴ、Celeron、Xeon、Intel SpeedStep、Itanium、および Pentium は、Intel Corporation または子会社の米国およびその他の国における商標または登録商標です。

Linux は、Linus Torvalds の米国およびその他の国における登録商標です。

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

ITIL は AXELOS Limited の登録商標です。

UNIX は The Open Group の米国およびその他の国における登録商標です。

Cell Broadband Engine は、Sony Computer Entertainment, Inc. の米国およびその他の国における商標であり、同社の許諾を受けて使用しています。

Linear Tape-Open、LTO、LTO ロゴ、Ultrium および Ultrium ロゴは、HP、IBM Corp. および Quantum の米国およびその他の国における商標です。

Printed in Japan

日本アイ・ビー・エム株式会社
〒103-8510 東京都中央区日本橋箱崎町19-21