# IBM SPSS Analytic Server
Versão 3.2.2

*Guia de Instalação e Configuração*

**Nota**

Antes de usar estas informações e o produto suportado por elas, leia as informações em “Avisos” na página 79.

**Informações do produto**

Esta edição aplica-se à versão 3, liberação 2, modificação do 2 do IBM® SPSS Analytic Server e a todas as liberações e modificações subsequentes até que seja indicado de outra forma em novas edições.

# Índice

# Capítulo 1. Pré-requisitos

Antes de instalar o Analytic Server, revise as informações a seguir.

**Requisitos do sistema**
Para obter informações mais atualizadas sobre os requisitos do sistema, use os relatórios de requisitos do sistema Detalhados no site de Suporte Técnico da IBM: http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html. Nesta página:

1. Digite SPSS Analytic Server como o nome do produto e clique em **Procurar**.
2. Selecione a versão e o escopo do relatório desejados e, em seguida, clique em **Enviar**.

**Tráfego do WebSocket**
Deve-se assegurar que o tráfego do WebSocket entre os clientes e o Analytic Server não seja bloqueado por firewalls, VPNs ou outros métodos de bloqueio de porta. A porta do WebSocket é a mesma que a porta geral do Analytic Server.

**SuSE Linux (SLES) 12**
Execute as tarefas a seguir antes de instalar o Analytic Server no SuSE Linux 12:

1. Faça download de uma chave pública para o seu host a partir da URL a seguir: https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
2. Importe a chave pública executando o comando a seguir em seu host:

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

**Ubuntu 18.04**
Execute as tarefas a seguir em todos os nós do cluster antes de instalar o Analytic Server no Ubuntu 18.04:

1. Faça download de uma chave pública para o seu host a partir da URL a seguir: https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
2. Importe a chave pública executando o comando a seguir em seu host:

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

**Power Systems**
Assegure-se de que os compiladores IBM XLC e XLF estejam instalados e incluídos no PATH em todos os hosts no cluster.

É possível localizar mais informações sobre como obter uma licença para esses compiladores nos websites a seguir:

- 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)**
Antes de instalar o Analytic Server, deve-se assegurar que pelo menos um cliente HDP tenha sido implementado em seu ambiente em cluster. Como o nó que hospeda o Ambari Manager espera o diretório `/usr/hdp` o Analytic Server falhará na ausência de um cliente HDP.

**Hive/HCatalog**
Se você planeja usar origens de dados NoSQL, configure Hive and HCatalog para acesso remoto. Além disso, assegure-se de que `hive-site.xml` contenha uma propriedade *hive.metastore.uris* no formato `thrift://<host_name>:<port>` que aponta para o servidor Thrift Hive Metastore ativo. Consulte a documentação da distribuição do Hadoop para obter detalhes.

Se você deseja usar o Hive 2.1, deve-se ativar o Hive 2.1 ativando a configuração **Consulta interativa**no console do Ambari e, em seguida, inserir `2.x` como a propriedade `hive.version` durante a instalação do Analytic Server.

1. Abra o console do Ambari e inclua a propriedade a seguir na seção **analytics.cfg avançado do Analytic Server**.
   - Chave: `hive.version`
   - Valor: insira a versão apropriada do Hive (por exemplo, 2.x)
2. Salve a configuração.

**Nota:** O Hive 2.1 é suportado com o HDP 2.6 ou posterior com o Spark 2.x. Para HDP 2.x, o `hive.version` padrão é 1.x; para HDP 3.x, o `hive.version` padrão é 3.x.

**Repositório de metadados**

Por padrão, o Analytic Server instala e usa um banco de dados MySQL. Como alternativa, é possível configurar Analytic Server para usar uma instalação do Db2 existente. Independentemente do tipo de banco de dados escolhido, ele deverá ter uma codificação UTF-8.

**MySQL**

O conjunto de caracteres padrão para MySQL depende da versão e do sistema operacional. Use as etapas a seguir para determinar se sua instalação do MySQL está configurada para UTF-8.

1. Determine a versão do MySQL.

   ```
   mysql -V
   ```

2. Determine o conjunto de caracteres padrão para o MySQL ao executar a seguinte consulta a partir da interface da linha de comandos MySQL.

   ```
   mysql > mostrar variáveis como 'char%';
   ```

   Se os conjuntos de caracteres já estiverem configurados para UTF-8, nenhuma mudança adicional será necessária.

3. Determine a ordenação padrão para o MySQL ao executar a seguinte consulta a partir da interface da linha de comandos MySQL.

   ```
   mysql > mostrar variáveis como 'coll%';
   ```

   Se a ordenação já estiver configurada para UTF-8, nenhuma mudança adicional será necessária.

4. Se o conjunto de caracteres ou a ordenação padrão não for UTF-8, consulte a documentação do MySQL para obter detalhes sobre como editar o arquivo /etc/my.cnf e reinicie o daemon do MySQL para alterar o conjunto de caracteres para UTF-8.

**Db2**

Para obter mais informações sobre como configurar o Db2, consulte o Knowledge Center http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html.

**Clusters de alta disponibilidade**

**Balanceador de carga**

Seu cluster de alta disponibilidade deve ter um balanceador de carga que suporta afinidade de sessão, conhecida também às vezes como sessões persistentes. O Analytic Server identifica as sessões com o cookie "request-token". Isso identifica uma sessão para a duração de um login de usuário para uso em afinidade de sessão controlada pelo aplicativo. Consulte a documentação do seu balanceador de carga particular para obter os detalhes de como ela suporta afinidade de sessão.

**Falha na tarefa Analytic Server**

Quando uma tarefa do Analytic Server falhar porque um membro de cluster falhou, a tarefa será normalmente reiniciada automaticamente em outro membro de cluster. Se a tarefa não for

retomada, verifique para assegurar se há pelo menos quatro membros de cluster no cluster de alta disponibilidade.

# Capítulo 2. Instalação e configuração do Ambari

## Pré-requisitos específicos do Ambari

Além dos pré-requisitos gerais, revise as informações a seguir.

**Serviços**
O Analytic Server é instalado como um serviço Ambari. Antes de instalar o Analytic Server, deve-se assegurar que os clientes a seguir estejam instalados como serviços do Ambari:

- HDFS/HDFS_CLIENT
- MAPREDUCE2/MAPREDUCE2_CLIENT
- HIVE/HIVE_CLIENT
- SPARK2/SPARK2_CLIENT (quando o Spark 2.x for usado)
- HBASE/HBASE_CLIENT (quando o HBASE for usado)
- YARN
- Zookeeper

**SSH sem senha**
Configure o SSH sem senha para o usuário raiz entre o host do Analytic Server e todos os hosts no cluster.

## Ferramentas de pré-verificação e pós-verificação de instalação - Ambari

### Visão geral da ferramenta de pré-verificação

A ferramenta de pré-verificação de instalação do Analytic Server ajuda a reduzir problemas de instalação e erros de tempo de execução identificando potenciais problemas de ambiente antes da instalação do Analytic Server.

A ferramenta de pré-verificação verifica:

- Versões do OS e do Ambari no sistema local
- Configurações de ulimit do OS no sistema local
- Espaço em disco disponível no sistema local
- Versão do Hadoop
- Disponibilidade do serviço Ambari (HDFS, HCatalog, Spark, Hive, MapReduce, YARN, Zookeeper e assim por diante)
- Configurações específicas do Ambari do Analytic Server

**Nota:** A ferramenta de pré-verificação pode ser usada após a execução do arquivo binário autoextrator do Analytic Server.

### Visão geral da ferramenta de pós-verificação

A ferramenta de pós-verificação de instalação do Analytic Server identifica problemas de configuração, após a instalação do Analytic Server, enviando solicitações de API de REST para processamento:

- Dados no HDFS
- Dados no Hive/HCatalog
- Dados compactados (incluindo o deflate, o bz2, o snappy)
- Dados com o PySpark

- Dados que usam componentes do SPSS nativos (incluindo alm, árvore, rede neural, pontuação, tascoring)
- Dados com o MapReduce
- Dados com o MapReduce em memória

### Local e pré-requisitos da ferramenta

Antes de instalar o serviço do Analytic Server, execute a ferramenta de pré-verificação em todos os nós que farão parte do serviço do Analytic Server para verificar se seu ambiente Linux está pronto para instalar o Analytic Server.

A ferramenta de pré-verificação é chamada automaticamente como parte da instalação. A ferramenta verifica o Analytic Metastore e cada nó do Analytic Server antes de executar a instalação em cada host. Também é possível chamar manualmente a ferramenta de pré-verificação no nó do Ambari Server, que validará a máquina antes da instalação do serviço.

Após executar o arquivo binário autoextrator do Analytic Server, a ferramenta de pré-verificação estará localizada nos diretórios a seguir:

- **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
```

Depois de instalar o Analytic Server, a ferramenta de pós-verificação está localizada no seguinte diretório:

- **HDP**

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

As ferramentas devem ser executadas como raiz e requerem o Python 2.6.X (ou superior).

Se a ferramenta de pré-verificação relatar alguma falha, as falhas deverão ser resolvidas antes de você continuar com a instalação do Analytic Server.

O diretório `chktool` estará disponível após o binário autoextrator do Analytic Server ser executado (etapa 2 na seção do “Instalação no Ambari” na página 7). Se você escolher executar um “Instalação offline” na página 11, o diretório `chktool` estará disponível após o RPM de metadados ser instalado.

### Executando a ferramenta de pré-verificação

#### Automática

A ferramenta de pré-verificação pode ser chamada automaticamente como parte da instalação do Analytic Server quando o Analytic Server é instalado como um serviço por meio do console do Ambari. deve-se inserir manualmente o nome do usuário e a senha do servidor Ambari:

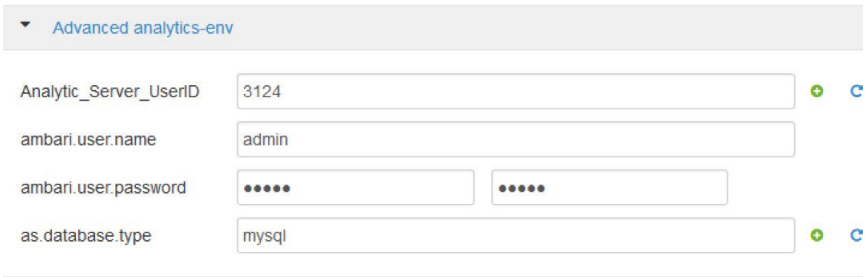



*Figura 1. Configurações avançadas de analytics-env*

#### Manual

É possível chamar manualmente a ferramenta de pré-verificação no nó do Ambari Server.

O exemplo de pré-verificação a seguir verifica o cluster `MyCluster` do Ambari que está em execução em `myambarihost.ibm.com:8080`, com SSL ativado e usa as credenciais de login `admin:admin`:

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

**Notes:**

- Os argumentos `--target`, `--host`, `--port` e `--username` são necessários.
- O valor `--host` deve ser fornecido por um endereço IP ou por um nome completo do domínio.
- A ferramenta solicitará uma senha quando o argumento de senha for omitido.
- O comando `precheck.py` inclui ajuda de uso, que é exibida com o argumento `--h` (`python ./precheck.py --help`).
- O argumento `--cluster` é opcional (o atual cluster é identificado quando `--cluster` não é usado).

Conforme a ferramenta de pré-verificação executa as suas verificações, o status de cada verificação é exibido na janela de comando. Quando uma falha ocorre, informações detalhadas ficam disponíveis no arquivo de log (o local exato do arquivo de log é fornecido na janela de comando). O arquivo de log poderá ser fornecido para o suporte técnico IBM quando for necessário mais suporte.

### Executando a ferramenta de pós-verificação

A ferramenta de pós-verificação verifica se o Analytic Server está sendo executado adequadamente e pode processar tarefas simples. O exemplo de pós-verificação a seguir verifica uma instância do Analytic Server que está em execução em `myanalyticserverhost.ibm.com:9443`, com SSL ativado e usa as credenciais de login `admin:ibmspss`:

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

Quando o Knox é usado com o Analytic Server, o comando é como segue:

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

Para executar uma verificação única, use o comando a seguir:

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

**Notes:**

- Os argumentos `--target`, `--host`, `--port` e `--username` são necessários.
- O valor `--host` deve ser fornecido por um endereço IP ou por um nome completo do domínio.
- A ferramenta solicitará uma senha quando o argumento de senha for omitido.
- O comando `postcheck.py` inclui ajuda de uso, que é exibida com o argumento `--h` (`python ./postcheck.py --help`).

Conforme a ferramenta de pós-verificação executa as suas verificações, o status de cada verificação é exibido na janela de comando. Quando uma falha ocorre, informações detalhadas ficam disponíveis no arquivo de log (o local exato do arquivo de log é fornecido na janela de comando). O arquivo de log poderá ser fornecido para o suporte técnico IBM se for necessário mais suporte.

## Instalação no Ambari

O processo básico é instalar os arquivos do Analytic Server em um host no cluster do Ambari e, em seguida, incluir o Analytic Server como um serviço do Ambari.

**“Instalação online” na página 8**

Escolha instalação on-line se o seu host do servidor Ambari e todos os nós no cluster puderem acessar o https://ibm-open-platform.ibm.com.

**“Instalação offline” na página 11**
Escolha off-line se o seu host do servidor Ambari não tiver acesso à internet.

## Instalação online

Escolha instalação on-line se o seu host do servidor Ambari e todos os nós no cluster puderem acessar o https://ibm-open-platform.ibm.com.

1. Navegue até o Website do IBM Passport Advantage® e faça download do arquivo binário autoextrator específico de sua pilha, versão de pilha e arquitetura de hardware para o nó do Ambari Manager. Os binários disponíveis do Ambari são:

| *Tabela 1. Arquivos binários autoextratores do Servidor analítico* | |
|---|---|
| **descrição** | **Nome do arquivo binário** |
| IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0 e 3.1 Linux x86-64 English | `spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin` |
| IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0 e 3.1 Linux on System p LE em inglês | `spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin` |

2. Execute o arquivo binário autoextrator e siga as instruções para visualizar a licença, aceitar a licença, escolher a instalação on-line e selecionar o processo de instalação para o tipo de banco de dados que o Analytic Server usa. As opções de tipo de banco de dados a seguir são fornecidas:
   - Nova instância do MySQL
   - Instância MySQL ou Db2 preexistente
3. No diretório `/var/lib/ambari-server/resources/stacks/<stack-name>/<stack-version>/services/ANALYTICSERVER/package/scripts`, execute o script `update_clientdeps.sh` com os argumentos adequados (use o argumento `--help` para obter exemplos).
4. Reinicie o servidor Ambari.

```
ambari-server restart
```

5. Efetue logon no servidor Ambari e instale o Analytic Server como um serviço por meio da UI do Ambari.

   **Repositório de metadados**

   Por padrão, o Analytic Server usa o MySQL para controlar informações sobre origens de dados, projetos e locatários. Durante a instalação, é necessário fornecer um nome de usuário (**`metadata.repository.user.name`**) e uma senha **`metadata.repository.password`** usados na conexão JDBC entre o Analytic Server e o MySQL. O instalador cria o usuário no banco de dados do MySQL, mas esse usuário é específico para o banco de dados do MySQL e não precisa ser um usuário Linux ou Hadoop existente.

   **Nota:** Se desejar que o instalador do Analytic Server crie uma nova instância do MySQL, deve-se instalar o Metastore do Analytic Server em uma máquina onde o MySQL ainda não esteja instalado.

   Para mudar o repositório de metadados para Db2, siga estas etapas.

   **Nota:** Não é possível mudar o repositório de metadados após a instalação ser concluída.

   a. Assegure-se de que o Db2 esteja instalado em outra máquina. Para obter informações adicionais, consulte a seção do repositório de metadados do tópico Capítulo 1, “Pré-requisitos”, na página 1.
   b. Na guia Serviços do Ambari, navegue para a guia Configurações do serviço do Analytic Server.
   c. Abra a seção **analytics-env avançado**.
   d. Mude o valor de **`as.database.type`** de `mysql` para db2.

e. Abra a seção **analytics-meta avançado**.

f. Altere o valor de **metadata.repository.driver** de `com.mysql.jdbc.Driver` para `com.ibm.db2.jcc.DB2Driver`.

g. Mude o valor de **metadata.repository.url** para `jdbc:db2://{Db2_HOST}:{PORT}/{DBName}:currentSchema={SchemaName};`, em que

- {Db2_HOST} é o nome do host do servidor no qual o Db2 está instalado
- {PORT} é a porta na qual o Db2 está atendendo
- {SchemaName} é um esquema disponível, não utilizável.

Se você não estiver certo quanto aos valores que deve inserir, trabalhe com o administrador do Db2.

h. Forneça credenciais válidas do Db2 em **metadata.repository.user.name** e **metadata.repository.password**.

i. Clique em **Salvar**.

**Configuração LDAP**

O Analytic Server usa um servidor LDAP para armazenar e autenticar usuários e grupos. Você fornece as informações de configuração LDAP necessárias durante a instalação do Analytic Server.

*Tabela 2. Definições de configuração LDAP*

| **Configuração LDAP** | **descrição** |
|---|---|
| `as.ldap.type` | Tipo de LDAP. O valor pode ser `ads`, `ad` ou `openladp`.<br>• `ads` - Apache Directory Server (configuração padrão)<br>• `ad` - Microsoft Active Directory<br>• `openladp` - OpenLDAP |
| `as.ldap.host` | Host LDAP |
| `as.ldap.port` | Número da porta LDAP |
| `as.ldap.binddn` | DN de ligação de LDAP |
| `as.ldap.bindpassword` | Senha do DN de ligação de LDAP |
| `as.ldap.basedn` | DN base de LDAP |
| `as.ldap.filter` | Regra de filtro de grupo e de usuário LDAP |
| `as.ldap.ssl.enabled` | Especifica se o SSL deve ser usado para comunicação entre o Analytic Server e o LDAP. O valor pode ser `true` ou `false`. |
| `as.ldap.ssl.reference` | ID de referência SSL LDAP |
| `as.ldap.ssl.content` | Configuração SSL LDAP |

- Por padrão, `as.ldap.type` é configurado para `ads` e as outras configurações relacionadas contêm configurações padrão. A exceção é que se deve fornecer uma senha para a configuração de `as.ldap.bindpassword`. O Analytic Server usa as definições de configuração para instalar um Apache Directory Server (ADS) e executar a inicialização do servidor. O perfil padrão do ADS inclui o usuário `admin` com uma senha `admin`. É possível executar o gerenciamentogerenciamento de usuários pelo Analytic Server Console ou importar informações sobre o usuário e o grupo de um arquivo XML por meio do script `importUser.sh`, localizado na pasta `<AnalyticRoot>/bin`.

- Se você planeja usar um servidor LDAP externo, como o Microsoft Active Directory ou o OpenLDAP, deve-se definir as definições de configuração de acordo com os valores reais de LDAP. Para obter mais informações, consulte Configurando registros de usuários LDAP no Liberty.
- É possível mudar a configuração de LDAP após a instalação do Analytic Server (por exemplo, mudando do Apache Directory Server para o OpenLDAP). No entanto, se você iniciar primeiramente com o Microsoft Active Directory ou o OpenLDAP e decidir alternar posteriormente para o Apache Directory Server, o Analytic Server não instalará um Apache Directory Server durante a instalação. O Apache Directory Server é instalado apenas quando ele é selecionado durante a instalação inicial do Analytic Server.

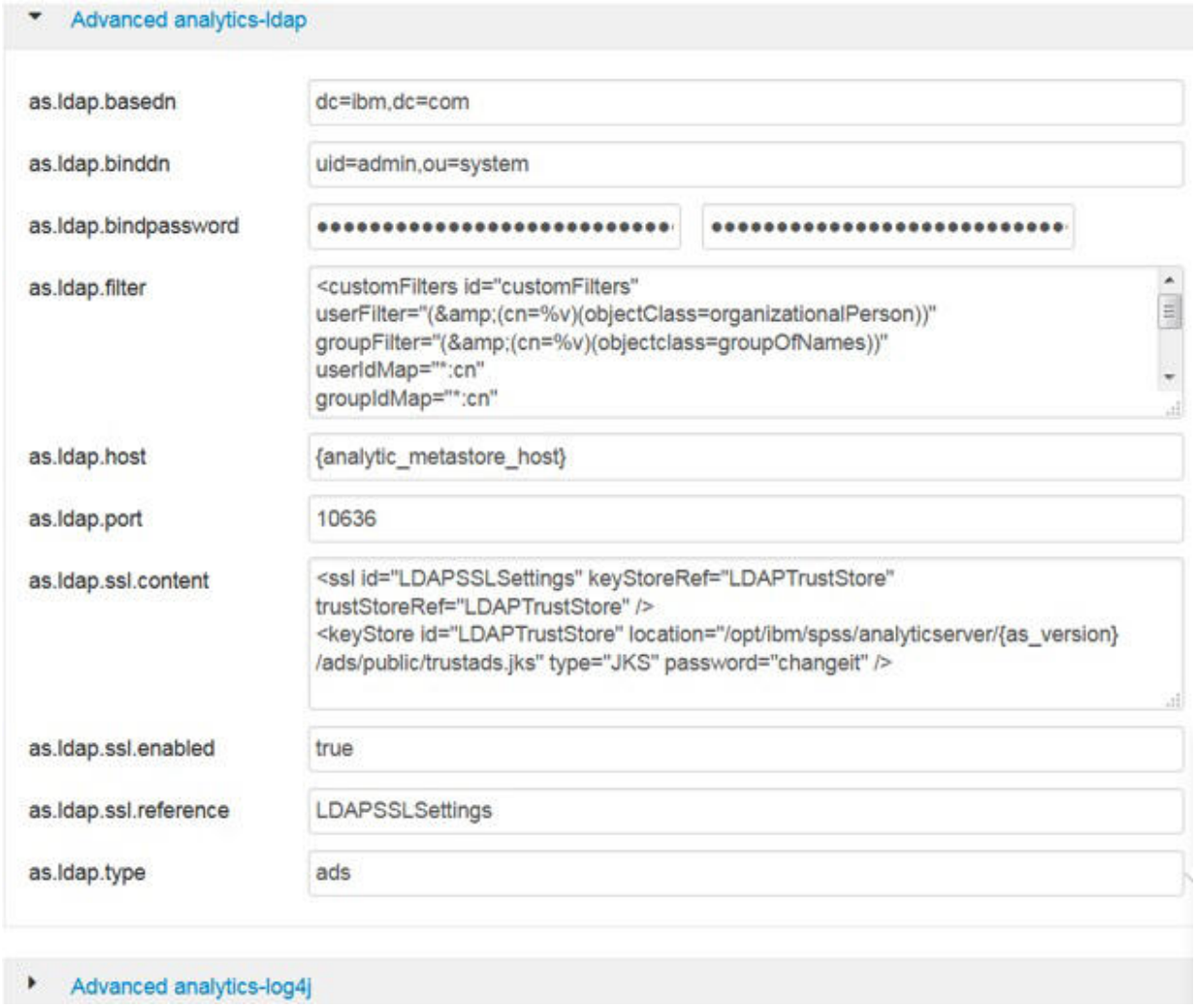



*Figura 2. Exemplo de definições de configuração de LDAP*

**Definições de configuração que não devem ser alteradas após a instalação**

Não altere as configurações a seguir após a instalação, ou o Analytic Server falhará ao funcionar.

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

6. Agora você tem uma instância em funcionamento do Analytic Server. A configuração adicional é opcional. Para obter informações adicionais sobre como configurar e administrar o Analytic Server, consulte o tópico: “Configuração” na página 18. Para obter informações sobre como migrar uma

configuração existente para uma nova instalação, consulte o tópico: “Migrando IBM SPSS Analytic Server no Ambari” na página 36.

7. Abra um navegador da web e insira o endereço `http://<host>:<port>/analyticserver/admin/ibm`, em que <host> é o endereço do host do Analytic Server e <port> é a porta na qual o Analytic Server está atendendo. Por padrão, este é 9080. Essa URL abre o diálogo de login para o console do Analytic Server. Efetue login como o administrador do Analytic Server. Por padrão, esse ID do usuário é admin e tem a senha admin.

## Instalação offline

Uma instalação off-line do IBM SPSS Analytic Server pode ser feita automática ou manualmente.



### Instalação automática no HDP

O processo de instalação automática utiliza a API de REST do Ambari e é o método preferencial para instalação.

**Importante:**

- O procedimento de instalação automática off-line instala um Apache Directory Server (ADS) integrado. Se desejar usar um servidor LDAP de terceiros, será possível configurar as definições de LDAP após a conclusão da instalação do IBM SPSS Analytic Server.
- O procedimento de instalação automática off-line pode instalar apenas uma única instância de serviço do Analytic Server. É possível incluir mais instâncias após a instalação inicial ser concluída.
- O procedimento de instalação automática off-line não suporta a instalação do Analytic Server em um cluster ativado para Kerberos.

Essas limitações não se aplicam às instalações manuais do HDP ou do Ubuntu.

1. Navegue para o website do IBM Passport Advantage® e faça download do arquivo binário autoextrator para um computador que pode acessar o https://ibm-open-platform.ibm.com.

*Tabela 3. Arquivo binário autoextrator do Analytic Server*

| **descrição** | **Nome do arquivo binário** |
|---|---|
| IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0 e 3.1 Linux x86-64 English | `spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin` |
| IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0 e 3.1 Linux on System p LE English | `spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin` |

2. Execute o binário executável transferido por download na etapa 1 e especifique uma instalação off-line. Uma instalação off-line faz download dos arquivos RPM ou DEB que são necessários posteriormente no processo de instalação e devem ser executados em um computador que possa acessar o https://ibm-open-platform.ibm.com. Os arquivos transferidos por download estão localizados no diretório binário executável atual `./IBM-SPSS-AnalyticServer`.
3. Copie todos os conteúdos do diretório binário executável `./IBM-SPSS-AnalyticServer` da máquina com acesso à Internet para o nó do Ambari Manager (localizado atrás do firewall).
4. No nó do Ambari Manager, use o comando a seguir para verificar se o servidor Ambari está em execução:

```
ambari-server status
```

5. No nó do Ambari Manager e em todos os outros nós nos quais você deseja implementar o Analytic Server, instale a ferramenta que cria um repositório local yum.

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

ou o

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

6. No nó do Ambari Manager, execute o arquivo binário executável `./IBM-SPSS-AnalyticServer/packages/spss_as-ambari-offlineinstall.bin`. Durante a instalação, o binário executável verifica se os arquivos RPM/DEB necessários do Analytic Server estão localizados no diretório de pacotes. Os arquivos do RPM dos quais você precisa dependem da sua distribuição, versão e arquitetura.

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

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

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

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

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

IBM-SPSS-AnalyticServer_1_amd64.deb

Durante a instalação, você é solicitado a inserir a versão do Analytic Server, o driver JDBC, a versão do Spark, a versão do Hive e assim por diante.

### Instalação manual no HDP (RHEL, SLES)

O fluxo de trabalho geral para uma instalação manual off-line no HDP (RHEL, SLES) é o seguinte:

1. Navegue para o website do IBM Passport Advantage® e faça download do arquivo binário autoextrator para um computador que pode acessar o https://ibm-open-platform.ibm.com.

*Tabela 4. Arquivos binários autoextratores do Servidor analítico*

| **descrição** | **Nome do arquivo binário** |
|---|---|
| IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0 e 3.1 Linux x86-64 English | `spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin` |
| IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0 e 3.1 Linux on System p LE em inglês | `spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin` |

2. Execute o binário executável transferido por download na etapa 1 e especifique uma instalação off-line. Uma instalação off-line faz download dos arquivos RPM que são necessários posteriormente no processo de instalação e deve ser executada em um computador que possa acessar https://ibm-open-platform.ibm.com. Os arquivos transferidos por download estão localizados no diretório binário executável atual `./IBM-SPSS-AnalyticServer`.
3. Copie todo o conteúdo do diretório binário executável `./IBM-SPSS-AnalyticServer` da máquina com acesso à Internet ao diretório do nó do Ambari Manager <AS_INSTALLABLE_HOME> (o nó do Ambari Manager está atrás do firewall).
4. No nó do Ambari Manager, use o comando a seguir para verificar se o servidor Ambari está em execução:

```
ambari-server status
```

5. Instale a ferramenta que cria um repositório yum local.

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

ou o

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

6. Crie um diretório que atenda como o repositório para os arquivos do RPM do Analytic Server. Veja o exemplo a seguir.

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

7. Copie os arquivos necessários do RPM do Analytic Server para o novo diretório. Os arquivos do RPM dos quais você precisa dependem da sua distribuição, versão e arquitetura.

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

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

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

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

8. Crie a definição do repositório local. Por exemplo, crie um arquivo que é denominado `IBM-SPSS-AnalyticServer-3.2.2.0.repo` em `/etc/yum.repos.d/` (para RHEL, CentOS) ou `/etc/zypp/repos.d/` (para SLES) com os conteúdos a seguir.

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

9. Crie o repositório yum local.

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

10. Em uma janela de comando do usuário raiz, cd para o diretório <AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer e `run./offLineInstall.sh`. O script lê respostas persistidas para o comando de instalação executável binário que foi executado anteriormente e emite o comando de plataforma apropriado (para instalar o rpm).

**Nota:** A Etapa 11 se aplicará somente se você usar um ambiente MySQL gerenciado externamente.

11. Execute o script `add_mysql_user.sh` no nó/host onde a instância MySQL, que será usada como a `AS_MetaStore`, está instalada.
    a. Copie o script `add_mysql_user.sh` de <AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer para o nó/host no qual a instância do MySQL, que será usada como o `AS_MetaStore`, está instalada.
        - Execute o script `add_mysql_user.sh` no nó/host do MySQL. Por exemplo, `./add_mysql_user.sh -u as_user -p spss -d aedb`

        **Notes:**

        - O nome de usuário e a senha devem corresponder ao nome de usuário e à senha do banco de dados inseridos para o `AS_Metastore` na tela de configuração do Ambari.
        - O script `add_mysql_user.sh` pode ser atualizado manualmente para emitir comandos (se assim desejado).

- Ao executar o script `add_mysql_user.sh` com relação a um banco de dados MySQL protegido (acesso de usuário raiz), use os parâmetros `-r` e `-t` para passar `dbuserid` e `dbuserid_password`. O script usa `dbuserid` end `dbuserid_password` para executar operações do MySQL.

**Nota:** A configuração `metadata.repository.url` na tela **AS_Configuration** (**Analítica avançada - meta**) deve ser modificada para apontar para o host do banco de dados MySQL. Por exemplo, mude a configuração de JDBC `mysql://{analytic_metastore_host}/aedb?createDatabaseIfNotExist=true` para `mysql://{MySQL_DB}/aedb?createDatabaseIfNotExist=true`

12. Atualize o seu arquivo de repositório do Ambari `repoinfo.xml`, normalmente localizado em `/var/lib/ambari-server/resources/stacks/$stackName/$stackVersion/repos/`, para usar o repositório `yum` local, incluindo as linhas a seguir.

```
< os type="host_os">
   <repo>
        < baseurl> file:/// { / < /baseurl>
        < repoid> IBM-SPSS-AnalyticServer < /repoid>
        <reponame>IBM-SPSS-AnalyticServer-3.2.2.0</reponame>
   < /repo>
< /os>
```

Um exemplo de `{caminho para o repositório local}` seria:

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

13. Repita os passos a seguir para cada nó de cluster do Ambari que não seja do servidor.
    a. Copie todo o conteúdo do diretório <AS_INSTALLABLE_HOME> adequado da máquina que tem acesso à Internet para o nó do cluster que não é do servidor Ambari.
    b. Instale a ferramenta que cria um repositório `yum` local.

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

ou o

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

    c. Crie um diretório que atenda como o repositório para os arquivos do RPM do Analytic Server. Veja o exemplo a seguir.

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

    d. Copie os arquivos necessários do RPM do Analytic Server para o novo diretório. Os arquivos do RPM dos quais você precisa dependem da sua distribuição, versão e arquitetura.

**HDP 2.6, 3.0 e 3.1 (x86_64)**

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

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

**HDP 2.6, 3.0 e 3.1 (PPC64LE)**

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

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

    e. Crie a definição do repositório local. Por exemplo, crie um arquivo que é denominado `IBM-SPSS-AnalyticServer-3.2.2.0.repo` em `/etc/yum.repos.d/` (para RHEL, CentOS) ou `/etc/zypp/repos.d/` (para SLES) com os conteúdos a seguir.

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

f. Crie o repositório yum local.

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

14. Continue com a etapa 3 na seção “Instalação online” na página 8.

**Instalação manual no HDP (Ubuntu)**

O fluxo de trabalho geral para uma instalação manual off-line no HDP (Ubuntu) é o seguinte:

1. Navegue até o Website do IBM Passport Advantage® e faça download do arquivo binário autoextrator apropriado do Ubuntu para um computador que possa acessar o https://ibm-open-platform.ibm.com.

*Tabela 5. Arquivos binários autoextratores do Servidor analítico*

| **descrição** | **Nome do arquivo binário** |
|---|---|
| IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6, 3.0 e 3.1 Linux x86-64 English | `spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin` |

2. Execute o binário executável transferido por download na etapa 1 e especifique uma instalação off-line. Uma instalação off-line faz download dos arquivos DEB que serão necessários posteriormente no processo de instalação e devem ser executados em um computador que possa acessar o https://ibm-open-platform.ibm.com. Os arquivos transferidos por download estão localizados no diretório binário executável atual `./IBM-SPSS-AnalyticServer`.
3. Copie todo o conteúdo do diretório binário executável `./IBM-SPSS-AnalyticServer` da máquina com acesso à Internet ao diretório do nó do Ambari Manager <AS_INSTALLABLE_HOME> (o nó do Ambari Manager está atrás do firewall).
4. No nó do Ambari Manager, use o comando a seguir para verificar se o servidor Ambari está em execução:

```
ambari-server status
```

5. Crie um diretório <local_repo> que atue como o repositório para os arquivos DEB do Analytic Server. Por exemplo:

```
mkdir –p /usr/local/mydebs
```

6. Copie os arquivos DEB necessários do Analytic Server para o diretório <local_repo>.
   - `IBM-SPSS-AnalyticServer-ambari-2.x_3.2.2.0_amd64.deb`
   - `IBM-SPSS-AnalyticServer_1_amd64.deb`
7. Crie o repositório local.
   a. Instale a ferramenta que cria um repositório local:

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

   b. Gerar o arquivo do pacote de origem:

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

   c. Crie o componente (principal) e a arquitetura (por exemplo, `binary-i386`, `binary-amd64`) de seu repositório local:

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

   d. Copie o pacote de origem:

```
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. Crie a definição do repositório local. Por exemplo, crie um arquivo denominado `IBM-SPSS-AnalyticServer-3.2.2.0.list` em `/etc/apt/sources.list.d` com o conteúdo a seguir.

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

   **Importante:** No Ubuntu 18.04, use: `deb [trusted=yes] file:/usr/local/mydebs ./`
9. Execute o comando a seguir para atualizar a lista de repositórios:

   ```
   apt-get update
   ```
10. Execute o comando a seguir para instalar o IBM SPSS Analytic Server 3.2.2:

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

    **Nota:** Para verificar se o repositório local está configurado corretamente, não execute o comando anterior no diretório `<local_repo>`. Se a instalação não puder localizar o pacote, isso significa que o repositório local não está configurado corretamente (nesse caso, deve-se verificar todas as etapas anteriores).
11. Repita os passos a seguir para cada nó de cluster do Ambari que não seja do servidor.
    a. Crie um diretório `<local_repo>` que atue como o repositório para os arquivos DEB do Analytic Server. Por exemplo:

       ```
       mkdir –p /usr/local/mydebs
       ```
    b. Copie todo o conteúdo do diretório `<local_repo>` da máquina do nó do Ambari Manager para o diretório `<local_repo>` do nó do cluster que não é do servidor Ambari. O diretório deve conter os arquivos a seguir:
       - `<local_repo>/IBM-SPSS-AnalyticServer-ambari-2.x_3.2.2.0_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. Crie a definição do repositório local. Por exemplo, crie um arquivo denominado `IBM-SPSS-AnalyticServer-3.2.2.0.list` em `/etc/apt/sources.list.d` com o conteúdo a seguir.

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

       **Importante:** No Ubuntu 18.04, use: `deb [trusted=yes] file:/usr/local/mydebs ./`
12. Continue com a etapa 3 na seção “Instalação online” na página 8.

## Instalando o Analytic Server com relação a um ambiente MySQL gerenciado externamente

O processo de instalação do Analytic Server difere de uma instalação normal quando realizado com relação a um ambiente MySQL gerenciado externamente.

As etapas a seguir explicam o processo de instalação do Analytic Server com relação a um ambiente MySQL gerenciado externamente.

1. Navegue para o Website do IBM Passport Advantage® e faça o download do arquivo binário autoextrator específico para a sua pilha, versão da pilha e arquitetura de hardware para um host no cluster do Ambari.
2. Execute o arquivo binário autoextrator e siga as instruções para (como opção) visualizar a licença e aceitá-la.
   a. Escolha a opção on-line.
   b. Selecione a opção **Banco de dados MySQL externo** quando solicitado.

3. Copie o script add_mysql_user.sh de <AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer para o nó/host no qual a instância do MySQL, que será usada como o AS_MetaStore, está instalada.
   - Execute o script add_mysql_user.sh no nó/host do MySQL. Por exemplo, ./add_mysql_user.sh -u as_user -p spss -d aedb

   **Notes:**
   - O nome de usuário e a senha devem corresponder ao nome de usuário e à senha do banco de dados inseridos para o AS_Metastore na tela de configuração do Ambari.
   - O script add_mysql_user.sh pode ser atualizado manualmente para emitir comandos (se assim desejado).
   - Ao executar o script add_mysql_user.sh com relação a um banco de dados MySQL protegido (acesso de usuário raiz), use os parâmetros -r e -t para passar dbuserid e dbuserid_password. O script usa dbuserid end dbuserid_password para executar operações do MySQL.
4. Reinicie o servidor Ambari.

```
ambari-server restart
```

5. No console do Ambari, inclua o serviço AnalyticServer como normal (insira os mesmos nome de usuário e senha do banco de dados inseridos na etapa 3).

   **Nota:** A configuração metadata.repository.url na tela **AS_Configuration** (**Analítica avançada - meta**) deve ser modificada para apontar para o host do banco de dados MySQL. Por exemplo, mude a configuração de JDBC mysql://{analytic_metastore_host}/aedb?createDatabaseIfNotExist=true para mysql://{MySQL_DB}/aedb?createDatabaseIfNotExist=true

## Permitindo agentes do Ambari não raiz

O processo de instalação do Analytic Server difere de uma instalação normal quando o Ambari Server e o Ambari Agent estão em execução como um usuário não raiz.

### Pré-requisitos

Efetue login como root inclua o usuário não raiz em cada host em seu cluster e configure o usuário não raiz com o acesso sudo. O exemplo a seguir inclui o usuário não raiz asuser ao arquivo suoders (o local do arquivo padrão é /etc/sudoers):

```
## Permitir que o root execute qualquer comando em qualquer lugar
asuser ALL=(ALL)  ALL

## Permitir que o root execute qualquer comando em qualquer lugar sem uma senha
asuser ALL=(ALL)  NOPASSWD: ALL
```

Consulte as seções “Instalação online” na página 8 ou “Instalação offline” na página 11 para obter informações de instalação detalhadas.

### requisito do sudo

Deve-se incluir sudo antes do texto de comando para todos os comandos que são executados como um usuário não raiz.

### Problemas do Ambari não raiz

Talvez você veja o erro a seguir depois de incluir o Analytic Server como um serviço através da interface com o usuário do Ambari:

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

O erro é o resultado de um problema do Ambari não raiz. Deve-se mudar o proprietário da pasta `/var/run/ambari-server` para o usuário não raiz e, em seguida, incluir o Analytic Server como um serviço através da interface com o usuário do Ambari. O exemplo a seguir demonstra o processo de mudança do proprietário da pasta `/var/run/ambari-server` para o usuário não raiz `asuser`.

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

**Ssh sem senha**

Quando o ssh sem senha não é configurado, o aviso a seguir é mostrado durante a instalação:

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

Deve-se criar manualmente `as_user` em cada nó. `as_user` é uma conta de usuário com autoridade de instalação do Analytic Server. Por exemplo:

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

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

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

# Make primary group user_group
sudo usermod -g hadoop as_user

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

**Nota:** `{as_user_id}` encontrado no nó principalAmbari através do comando `id as_user`.

## Configuração

Após a instalação, deve-se criar as contas necessárias no sistema operacional do cluster.

1. Crie contas de usuário do sistema operacional para todos os usuários aos quais você planeja dar acesso ao Analytic Server em cada nó do Analytic Server e Hadoop (esses usuários também são configurados como registros de usuário LDAP). O grupo de usuários deve ser configurado como `hadoop`.
   - Certifique-se de que o UID para esses usuários corresponda em todas as máquinas. É possível testar isso usando o comando **`kinit`** para efetuar login em cada uma das contas.
   - Certifique-se de que o ID do usuário esteja de acordo com a configuração do YARN **ID de usuário mínimo para o envio da tarefa**. Este é o parâmetro **`min.user.id`** em `container-executor.cfg`. Por exemplo, se **`min.user.id`** for 1000, então cada conta do usuário criada deverá ter um UID maior ou igual a 1000.
2. Crie uma pasta inicial do usuário no HDFS para o usuário administrador do Analytic Server. A permissão da pasta deve ser configurada como 755, o proprietário deve ser configurado como `admin` e o grupo de usuários deve ser configurado como `hdfs`. Veja o exemplo em **negrito** a seguir:

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

3. Crie pastas iniciais do usuário no HDFS para todos os usuários padrão do Analytic Server (por exemplo, `user1`). O proprietário da pasta é o usuário real e o grupo de usuários deve ser configurado como `hdfs`.

Após a instalação, opcionalmente, é possível configurar e administrar o Analytic Server por meio da UI do Ambari.

**Nota:** As convenções a seguir são utilizadas para os caminhos de arquivo do Analytic Server.

- {AS_ROOT} refere-se ao local em que o Analytic Server foi implementado; por exemplo, `/opt/ibm/spss/analyticserver/3.2`.
- {AS_SERVER_ROOT} refere-se à localização dos arquivos de configuração, de log e de servidor; por exemplo, `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver`.
- {AS_HOME} refere-se ao local no HDFS usado pelo Analytic Server como uma pasta raiz; por exemplo, `/user/as_user/analytic-root`.

## Segurança

### Configurando um Registro LDAP

O registro LDAP permite autenticar usuários com um servidor LDAP externo, como o Active Directory ou OpenLDAP.

**Importante:** Um usuário LDAP deve ser designado como um administrador do Analytic Server no Ambari.

Veja aqui um exemplo de ldapRegistry para OpenLDAP.

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

O exemplo a seguir fornece autenticação do Analytic Server com o Active Directory:

```
<ldapRegistry id="Microsoft Active Directory" realm="ibm"
  host="host"
  port="389"
  baseDN="cn=users,dc=adtest,dc=mycompany,dc=com"
  bindDN="cn=administrator,cn=users,dc=adtest,dc=mycompany,dc=com"
  bindPassword ="adminpassword"
  ldapType="Custom"
    <customFilters
    userFilter="(&amp;(sAMAccountName=%v)(objectcategory=user))"
    groupFilter="(&amp;(cn=%v)(objectcategory=group))"
    userIdMap="user:sAMAccountName"
    groupIdMap="*:cn"
    groupMemberIdMap="memberOf:member " />
< /ldapRegistry>
```

**Nota:** É frequentemente útil usar uma ferramenta de terceiro como visualizador de LDAP para verificar a configuração de LDAP.

O exemplo a seguir fornece autenticação do perfil do WebSphere Liberty com o Active Directory:

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

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

<keyStore id="LDAPKeyStore" location="${server.config.dir}/LdapSSLKeyStore.jks"
          type="JKS " password=" {9}Hgw=" />

<keyStore id="LDAPTrustStore" location="${server.config.dir}/LdapSSLTrustStore.jks"
          type="JKS " password=" {9}Hgw=" />
```

**Notes:**

- O suporte para LDAP no Analytic Server é controlado pelo WebSphere Liberty. Para obter mais informações, consulte Configurando registros de usuários de usuários LDAP no Liberty.
- Quando LDAP estiver protegido com SSL, siga as instruções na seção "Configurar uma conexão secure socket layer (SSL) do Analytic Server para LDAP" a seguir.

#### Configurando uma conexão secure socket layer (SSL) do Analytic Server para LDAP

Se você selecionar a opção LDAP do Apache Directory Server (ads) durante a instalação do Analytic Server e usar a configuração padrão, o Apache Directory Server será instalado com SSL configurado e ativado (o Analytic Server usará SSL automaticamente para se comunicar com o Apache Directory Server).

Configure SSL usando as etapas a seguir quando uma das outras opções de LDAP for selecionada durante a instalação do Analytic Server (por exemplo, ao usar um servidor LDAP externo).

1. Efetue login em cada uma das máquinas do Analytic Server como o usuário do Analytic Server e crie um diretório comum para certificados SSL.

   **Nota:** Por padrão, as_user é o usuário do Analytic Server; consulte **Contas de serviço** na guia Admin no console do Ambari.

2. Copie os arquivos keystore e truststore para algum diretório comum em todas as máquinas do Analytic Server. Inclua também o certificado CA do cliente LDAP no truststore. A seguir estão algumas instruções de amostra.

   ```
   mkdir -p /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
   ```

   **Nota:** JAVA_HOME é o mesmo JRE usado para inicialização do Analytic Server.

3. As senhas podem ser codificadas para ofuscar seus valores com a ferramenta securityUtility, que está em {AS_ROOT}/ae_wlpserver/bin. A seguir está um exemplo.

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

4. Efetue login no console do Ambari e atualize a definição de configuração **ssl.keystore.config** do Analytic Server com as definições de configuração SSL corretas. A seguir está um exemplo.

   ```
   <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=" {6}Nis= " />
   ```

   **Nota:** Use o caminho absoluto para os arquivos key e truststore.

5. Atualize a definição de configuração **security.config** do Analytic Server com as definições de configuração LDAP corretas. Por exemplo, no elemento **ldapRegistry**, configure o atributo **sslEnabled** como true e o atributo **sslRef** como defaultSSLConfig.

**Configurando o Kerberos**
O Analytic Server suporta o Kerberos usando o Ambari.

**Nota:** O IBM SPSS Analytic Server não suporta conexão única (SSO) do Kerberos quando usado juntamente com o Apache Knox.

1. Crie contas no repositório do usuário do Kerberos para todos os usuários aos quais você planeja conceder acesso ao Analytic Server.
2. Crie as mesmas contas (da etapa anterior) no servidor LDAP.
3. Crie uma conta do usuário do S.O. para cada um dos usuários criados em uma etapa anterior em cada um dos nós do Analytic Server e do nó do Hadoop. O grupo de usuários deve ser configurado como `hadoop`.
   - Certifique-se de que o UID para esses usuários corresponda em todas as máquinas. É possível testar isso usando o comando **`kinit`** para efetuar login em cada uma das contas.
   - Certifique-se de que o ID do usuário esteja de acordo com a configuração do YARN **ID de usuário mínimo para o envio da tarefa**. Este é o parâmetro **`min.user.id`** em `container-executor.cfg`. Por exemplo, se **`min.user.id`** for 1000, então cada conta do usuário criada deverá ter um UID maior ou igual a 1000.
4. Crie uma pasta inicial do usuário no HDFS para o usuário administrador do Analytic Server. A permissão da pasta deve ser configurada como 755, o proprietário deve ser configurado como `admin` e o grupo de usuários deve ser configurado como `hdfs`. Veja o exemplo em **negrito** a seguir:

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

Localizados 9 itens

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

5. Crie pastas iniciais do usuário no HDFS para todos os usuários padrão do Analytic Server (por exemplo, `user1`). O proprietário da pasta é o usuário real e o grupo de usuários deve ser configurado como `hdfs`.
6. [Opcional] Se você planeja usar origens de dados HCatalog e o Analytic Server estiver instalado em uma máquina diferente do Hive metastore, será necessário personificar o cliente Hive no HDFS.
   a. Navegue para a guia Configurações do serviço HDFS no console do Ambari.
   b. Edite o parâmetro **`hadoop.proxyuser.hive.groups`** para ter o valor `*`, ou um grupo que contém todos os usuários que podem efetuar login no Analytic Server.
   c. Edite o parâmetro **`hadoop.proxyuser.hive.hosts`** para ter o valor `*`, ou a lista de hosts na qual o Hive metastore e cada instância do Analytic Server são instalados como serviços.
   d. Reinicie o serviço HDFS.

Após a execução dessas etapas e a instalação do Analytic Server, o Analytic Server configura o Kerberos de forma silenciosa e automática.

**Configurando o HAProxy para Conexão Única (SSO) usando Kerberos**

1. Configure e inicie o HAProxy seguindo o guia da documentação do HAProxy: http://www.haproxy.org/#docs
2. Crie o princípio (HTTP/<proxyHostname>@<realm>) e o arquivo keytab do Kerberos para o host do HAProxy, no qual <proxyHostname> é o nome completo do host do HAProxy e <realm> é a região do Kerberos.
3. Copie o arquivo keytab para cada um dos hosts do Analytic Server como `/etc/security/keytabs/spnego_proxy.service.keytab`

4. Atualize as permissões para esse arquivo em cada um dos hosts do Analytic Server. A seguir está um exemplo.

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

5. Abra o console do Amabri e atualize as seguintes propriedades na seção 'Custom analytics.cfg' do Analytic Server.

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

6. Salve a configuração e reinicie todos os serviços do Analytic Server a partir do console do Amabri.

Agora, os usuários podem efetuar login no Analytic Server usando a opção **Conexão única no login** na tela de login do IBM SPSS Analytic Server.

#### Ativando a personificação do Kerberos

A personificação permite que um encadeamento seja executado em um contexto de segurança que difere do contexto de segurança do processo que possui o encadeamento. Por exemplo, a personificação fornece um meio de as tarefas do Hadoop serem executadas como usuário que não o usuário padrão do Analytic Server (`as_user`). Para ativar a personificação do Kerberos:

1. Inclua atributos de configuração de personificação no HDFS (ou as configurações de serviço do Hive) ao executar em um cluster ativado do Kerberos. No caso de HDFS, as propriedades a seguir devem ser incluídas no arquivo `core-site.xml` do HFDS:

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

   em que `<analytic_server_service_principal_name>` é o valor padrão de `as_user` especificado no campo `Analytic_Server_User` da configuração do Analytic Server.

   As propriedades a seguir também devem ser incluídas no arquivo `core-site.xml` do HDFS em casos em que os dados são acessados por meio do HDFS via Hive/HCatalog:

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

2. Se o Analytic Server é configurado para usar um nome de usuário diferente de `as_user`, deve-se modificar os nomes de propriedade para refletir o outro nome de usuário (por exemplo, `hadoop.proxyuser.xxxxx.hosts`, em que `xxxxx` é o nome de usuário configurado que é especificado na configuração do Analytic Server).
3. Execute o comando a seguir de um shell de comando no nó do Analytic Server:

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

#### Ativando Diversas Regiões

A configuração **as.kdc.realms** é necessária ao definir várias regiões. Os valores **as.kdc.realms** estão localizados na seção 'Advanced analytics.cfg' do Analytic Server do console Ambari.

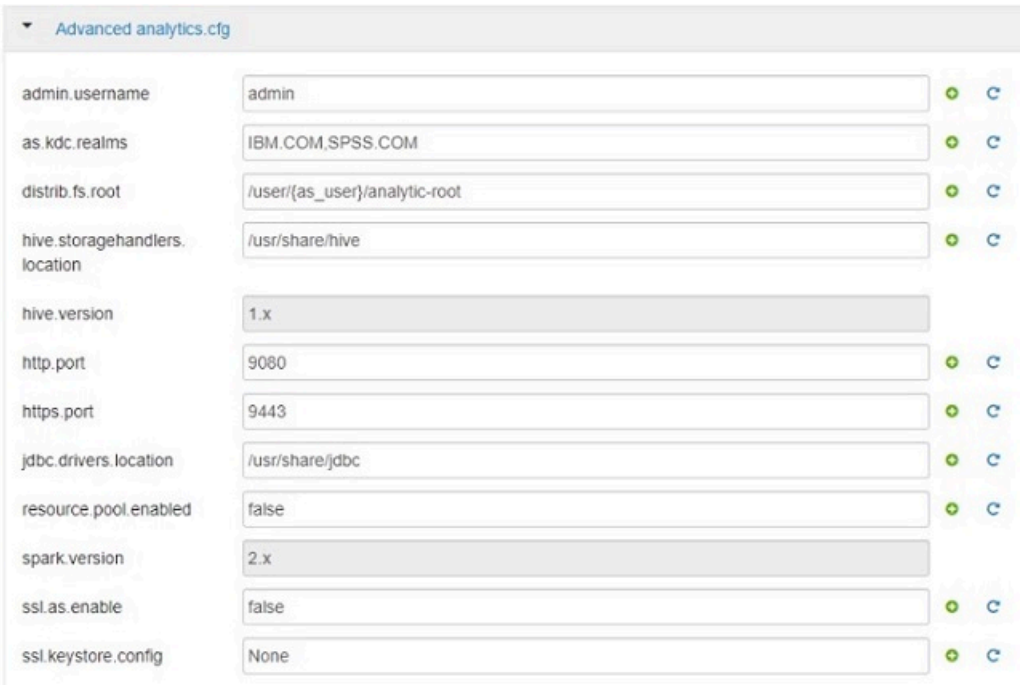



*Figura 3. Configurações avançadas de analytics.cfg*

Vários nomes de região são suportados quando separados por caracteres de vírgula. Os nomes de região do Kerberos especificados correspondem a, e estão associados a, nomes de usuário. Por exemplo, os nomes de usuário `UserOne@us.ibm.com` e `UserTwo@eu.ibm.com` corresponderão às regiões `us.ibm.com,eu.ibm.com`.

A confiança entre regiões do Kerberos deve ser configurada quando mais de uma região for especificada como um **Nome de região do Kerberos**. O nome do usuário inserido durante o prompt de login do console do Analytic Server é inserido sem o sufixo de nome de região. Como resultado, quando várias regiões são especificadas, uma lista suspensa de **Regiões** é apresentada aos usuários, permitindo-lhes selecionar a região.

**Nota:** Quando apenas uma região for especificada, os usuários não verão uma lista suspensa de **Regiões** ao se inscreverem no Analytic Server.

### Desativando o Kerberos

1. Desative o Kerberos no console do Ambari.
2. Pare o serviço do Analytic Server.
3. Clique em **Salvar** e reinicie o serviço do Analytic Server.

### Ativando conexões Secure Socket Layer (SSL) com o console do Analytic Server

Por padrão, o Analytic Server gera certificados autoassinados para ativar o Secure Socket Layer (SSL), para que seja possível acessar o console do Analytic Server por meio da porta segura, aceitando certificados autoassinados. Para tornar o acesso HTTPS mais seguro, é necessário instalar certificados de fornecedores terceiros.

### Instalando certificados de fornecedores terceiros

1. Copie os certificados de keystore e de truststore de fornecedores terceiros para o mesmo diretório em todos os nós do Analytic Server; por exemplo, `/home/as_user/security`.

   **Nota:** O usuário do Analytic Server deve ter acesso de leitura a esse diretório.
2. Na guia Serviços do Ambari, navegue para a guia Configurações do serviço do Analytic Server.

3. Edite o parâmetro **ssl.keystore.config**.

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

   Substitua

   - <KEYSTORE-LOCATION> pelo local absoluto do keystore; por exemplo: `/home/as_user/security/mykey.jks`
   - <TRUSTSTORE-LOCATION> pelo local absoluto do armazenamento confiável; por exemplo: `/home/as_user/security/mytrust.jks`
   - <TYPE> com o tipo do certificado, por exemplo: JKS, PKCS12, etc.
   - <PASSWORD> com a senha criptografada no formato de criptografia Base64. Para codificação, é possível usar o securityUtility; por exemplo: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility encode <password>`

   Se você deseja gerar um certificado autoassinado, é possível usar o securityUtility; por exemplo: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility createSSLCertificate --server=aeserver --password=mypassword --validity=365 --subject=CN=myfqdnserver,O=myorg,C=mycountry`.

   **Notes:**

   - Deve-se fornecer um nome de domínio de host apropriado para o valor `CN`.
   - Substitua **mypassword**, **myfqdnserver**, **myorg** e **mycounty** por suas credenciais específicas. Observe que **myfqdnserver** é o nome completo do domínio para o nó do Analytic Server.
   - **aeserver** é o nome do servidor Liberty (o valor deve ser **aeserver**).

   Para obter mais informações sobre o securityUtility e outras configurações SSL, consulte as documentações Perfil Liberty do WebSphere e comando securityUtility.
4. Clique em **Salvar** e reinicie o serviço do Analytic Server.

#### Gerando um certificado autoassinado

É possível usar o `securityUtility` para gerar certificados autoassinados. Por exemplo:

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

**Notes:**

- Deve-se fornecer um nome de domínio de host apropriado para o valor **CN**.
- Copie as informações que estão em `key.jks` para `trust.jks` (os dois arquivos devem ser idênticos).
- Edite o parâmetro `ssl.keystore.config`. Por exemplo:

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

**Comunicando-se com o Apache Hive por meio de SSL**
Deve-se atualizar o arquivo `hive.properties` para se comunicar com o Apache Hive por uma conexão SSL. Como alternativa, se o ambiente do Apache Hive estiver ativado para alta disponibilidade, será possível selecionar os parâmetros de alta disponibilidade na página principal Origens de dados do Analytic Server.

**Atualizando o arquivo hive.properties**

1. Abra o arquivo `hive.properties`. O arquivo está localizado em: `/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/database`
2. Localize a linha a seguir:

```
jdbcurl = jdbc:hive2: // {: { / {; user = {; password = {
```

3. Atualize a linha incluindo as informações abaixo em **negrito**:

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

4. Salve o arquivo `hive.properties` .

## Ativando o suporte para Essentials for R

Analytic Server suporta modelos R de pontuação e scripts R de execução.

Para configurar o suporte para R após uma instalação bem-sucedida do Analytic Server:

1. Provisão do ambiente do servidor para o Essentials for R.

**RedHat Linux x86_64**
Execute os comandos a seguir:

```
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**
Execute os comandos a seguir:

```
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**
A instalação do Essentials for R no SUSE requer FORTRAN compatível, que normalmente não está disponível nos repositórios ZYPPER configurados (disponível apenas a partir da mídia SUSE SDK). Como resultado, a execução de uma instalação do Ambari para o Essentials for R no servidor SUSE falhará, já que não poderá instalar o FORTRAN. Use as etapas a seguir para provisão no SUSE:

a. Instale o GCC C++.

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

b. Instale o GCC FORTRAN. Os arquivos RPM necessários podem ser copiados da mídia SUSE SDK e devem ser instalados na seguinte ordem.

```
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. Execute o comando a seguir para instalar as bibliotecas do Essentials for R.

```
R_PREFIX=/opt/ibm/spss/R
cd $R_PREFIX
rm -fr $R_PREFIX/r_libs
mkdir -p $R_PREFIX/r_libs
cd $R_PREFIX/r_libs
wget https://zlib.net/fossils/zlib-1.2.11.tar.gz --no-check-certificate
tar zxvf zlib-1.2.11.tar.gz
cd zlib-1.2.11/
./configure
make && make install
cd $R_PREFIX/r_libs
wget http://www.bzip.org/1.0.6/bzip2-1.0.6.tar.gz
tar xzvf bzip2-1.0.6.tar.gz
cd bzip2-1.0.6
sed "s|^CC=gcc|CC=gcc -fPIC|" -i ./Makefile
make -f Makefile-libbz2_so
make
clean
fazer
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
fazer
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
fazer
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. Faça download do archive autoextrator (BIN) para o IBM SPSS Modeler Essentials for R RPM or DEB. O Essentials for R está disponível para download (https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp). Escolha o arquivo específico para sua pilha, versão de pilha e arquitetura de hardware.
3. Execute o arquivo binário autoextrator e siga as instruções para (opcionalmente) visualizar a licença, aceite a licença e escolha instalação online ou offline.

   **Instalação online**
   Escolha instalação on-line se o seu host do servidor Ambari e todos os nós no cluster puderem acessar o https://ibm-open-platform.ibm.com.

   **Instalação offline**
   Escolha off-line se o seu host do servidor Ambari não tiver acesso à internet. A instalação off-line fará download dos arquivos do RPM necessários e deverá ser executada em uma máquina que possa acessar o https://ibm-open-platform.ibm.com. Os arquivos RPM podem, então, ser copiados para o host do servidor Ambari.

   a. Copie os arquivos do Essentials for R RPM or DEB necessários para qualquer local em seu host do servidor Ambari. Os arquivos do RPM/DEB dos quais você precisa dependem da sua distribuição, versão e arquitetura, mostradas abaixo.

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

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

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

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

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

b. Instale o RPM ou o DEB. No exemplo a seguir, o comando instala o Essentials for R on HDP 2.6 (x86_64).

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

No exemplo a seguir, o comando instala o Essentials for R on HDP 2.6 (Ubuntu).

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

4. Reinicie o servidor Ambari.

```
ambari-server restart
```

5. Efetue logon em seu servidor Ambari e instale o SPSS Essentials for R como um serviço via console do Ambari. O SPSS Essentials for R deve ser instalado em cada host no qual o Analytic Server e o Analytic Metastore estiverem instalados.

   **Nota:** O Ambari tentará instalar o gcc-c++ e gcc-gfortran (RHEL) e o gcc-fortran (SUSE) antes de instalar o R. Esses pacotes são declarados como dependências na definição de serviço do Ambari do R. Assegure-se de que os servidores em que o R deve ser instalado e executado estejam configurados para fazer o download de RPMs gcc-c++ e gcc-[g]fortran ou possuam os compiladores GCC e FORTRAN instalados. Se a instalação do Essentials for R falhar, instale esses pacotes manualmente antes de instalar o Essentials for R.
6. Atualize o serviço do Analytic Server.
7. Execute o script `update_clientdeps` usando as instruções "Atualizando as dependências do cliente" na página 32.
8. Deve-se também instalar o Essentials for R na máquina que hospeda o SPSS Modeler Server. Veja a a documentação do SPSS Modeler para obter detalhes.

## Ativando origens de base de dados relacional

O Analytic Server pode usar origens do banco de dados relacional se você fornecer os drivers JDBC em um diretório compartilhado no metastore do Analytic Server e em cada host do Analytic Server. Por padrão, esse diretório é `/usr/share/jdbc`.

Para alterar o diretório compartilhado, siga essas etapas.

1. Na guia Serviços do Ambari, navegue para a guia Configurações do serviço do Analytic Server.
2. Abra a seção **analytics.cfg avançado**.
3. Especifique o caminho do diretório compartilhado de drivers JDBC em **`jdbc.drivers.location`**.
4. Clique em **Salvar**.
5. Pare o serviço do Analytic Server.
6. Clique em **Atualizar**.
7. Inicie o serviço Analytic Server.

*Tabela 6. Bancos de Dados Suportados*

| Database | Versões suportadas | Jars de driver JDBC | Vendor |
|---|---|---|---|
| Amazon Redshift | 8.0.2 ou posterior | `RedshiftJDBC41-1.1.6.1006.jar` ou mais recente | Amazônio |
| BigSQL | 4.1.0.0 ou mais recente | db2jcc.jar | IBM |
| DashDB | Serviço Bluemix | `db2jcc.jar` | IBM |

| *Tabela 6. Bancos de Dados Suportados (continuação)* | | | |
|---|---|---|---|
| **Database** | **Versões suportadas** | **Jars de driver JDBC** | **Vendor** |
| Db2 for Linux, UNIX e Windows | 11.1, 10.5, 10.1, 9.7 | `db2jcc.jar` | IBM |
| Db2 z/OS | 11, 10 | `db2jcc.jar`, `db2_license_cisuz.jar` | IBM |
| Greenplum | 5 | postgresql.jar | Greenplum |
| Hive | 2.1, 1.2 | `hive-jdbc-*.jar` | Apache |
| MySQL | 5.7, 5.6 | mysql-connector-java-commercial-5.1.25-bin.jar | MySQL |
| Netezza | 7, 6.x | `nzjdbc.jar` | IBM |
| Oracle | 19c, 12c, 11g R2 (11.2) | 19c:<br>`ojdbc8.jar`, `orai18n.jar`<br>12c and 11g R2 (11.2):<br>`ojdbc6.jar`, `orai18n.jar` | Oracle |
| Servidor SQL | 2017, 2014, 2012, 2008 R2 | `sqljdbc4.jar` | Microsoft |
| Teradata | 15.1, 15 | `tdgssconfig.jar`, `terajdbc4.jar` | Teradata |

**Notes**

- Se você tiver criado uma origem de dados Redshift antes da instalação do Analytic Server, será necessário executar as etapas a seguir para usar a origem de dados Redshift.
  1. No console do Analytic Server, abra a origem de dados Redshift.
  2. Selecione a origem de dados do banco de dados `Redshift`.
  3. Insira o endereço do servidor Redshift.
  4. Insira o nome do banco de dados e o nome do usuário. A senha deve ser preenchida automaticamente.
  5. Selecione a tabela de banco de dados.
- O BigSQL é a interface do IBM SQL para o ambiente do Apache Hadoop. BigSQL não é um banco de dados relacional, mas o Analytic Server suporta o acesso a ele por meio de JDBC (o arquivo jar JDBC é o mesmo que o usado para o Db2).

  Um uso comum para o BigSQL com o Analytic Server está acessando tabelas Hadoop/HBase do BigSQL por meio de uma origem de dados do HCatalog.

## Ativação das Origem de Dados HCatalog

O Analytic Server fornece suporte para várias origens de dados por meio do Hive/HCatalog. Algumas origens requerem etapas de configuração manual.

1. Colete os arquivos JAR necessários para ativar a origem de dados. Não são necessárias etapas adicionais para ativar o suporte para o Apache HBase e o Apache Accumulo. Para outras origens de dados NoSQL, entre em contato com o fornecedor do banco de dados e obtenha o manipulador de

armazenamento e os jars relacionados. Para obter informações sobre origens de dados HCatalog compatíveis, consulte a seção "Usando origens de dados HCatalog" no Guia do Usuário do IBM SPSS Analytic Server 3.2.2.

2. Inclua esses arquivos JAR ao diretório `{HIVE_HOME}/auxlib` e ao diretório `/usr/share/hive` no metastore do Analytic Server e em cada nó do Analytic Server.
3. Reinicie o serviço Hive Metastore.
4. Atualize o Serviço de metastore analítico.
5. Reinicie cada instância do serviço Analytic Server.

**Notes:**

- O Analytic Server Metastore não pode ser instalado na mesma máquina que o Hive Metastore.
- Ao acessar dados do HBase por meio de uma origem de dados do HCatalog do Analytic Server, o usuário de acesso deve ter permissão de leitura para as tabelas do HBase.
  - Em ambientes não kerberos, o Analytic Server acessa o HBase usando `as_user` (`as_user` deve ter permissão de leitura para o HBase).
  - Em ambientes do Kerberos, tanto o `as_user` quanto o usuário de login devem ter permissão de leitura para tabelas do HBase.

#### Bancos de dados NoSQL

O Analytic Server suporta qualquer banco de dados NoSQL para o qual um manipulador de armazenamento Hive está disponível no fornecedor.

Não são necessárias etapas adicionais para ativar o suporte para o Apache HBase e o Apache Accumulo.

Para outros bancos de dados NoSQL, entre em contato com o fornecedor de base de dados e obtenha o manipulador de armazenamento e os jars relacionados.

#### Tabelas Hive baseadas em arquivo

O Analytic Server suporta tabelas Hive baseadas em arquivo para as quais um Hive SerDe integrado ou customizado (serializador-desserializador) está disponível.

O Hive XML SerDe para processar arquivos XML está localizado no Maven Central Repository em http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde.

#### Tarefas MapReduce v2

Use a configuração **`preferred.mapreduce`** na seção **analytic.cfg customizado** do Analytic Server para controlar como as tarefas MapReduce são manipuladas:

*Tabela 7. Propriedades do analytics.cfg customizado*

| **Propriedade** | **descrição** |
|---|---|
| `preferred.mapreduce` | Controla o método no qual as tarefas MapReduce são executadas. Valores válidos incluem:<br>• `spark`<br>• `m3r`<br>• `hadoop`<br>Por exemplo: `preferred.mapreduce=spark` |

#### Apache Spark

Se desejar usar o Spark (versão 2.x ou posterior), deve-se incluir manualmente a propriedade `spark.version` durante a instalação do Analytic Server.

1. Abra o console do Amabri e inclua a propriedade a seguir na seção **analytics.cfg avançado** do Analytic Server.
   - **Chave**: `spark.version`
   - **Valor**: insira o número da versão do Spark apropriado (por exemplo, `2.x` ou `Nenhum`).
2. Salve a configuração.

**Nota:** É possível forçar o HCatalog a nunca usar o Spark por meio de uma configuração do `analytics.cfg` customizado.

1. Abra o console do Amabri e inclua a propriedade a seguir na seção **analytic.cfg customizado** do Analytic Server.
   - **Chave**: `spark.hive.compatible`
   - **Valor**: `false`

#### Amibentes HDP 3.0 (ou mais recente) ativados para Kerberos

Os ambientes HDP 3.0 (ou superior) ativados para Kerberos podem requerer definições de configuração de segurança adicionais. No HDFS, o facl do sistema de arquivos é usado no diretório `/warehouse/tablespace/managed/hive`. É possível identificar o requisito para configurar os aspectos no metastore do Hive quando as exceções a seguir aparecerem nos arquivos `messages.log` ou `as_trace.log`:

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

O exemplo a seguir mostra um comando **`setfacl`** que fornece acesso amplo (nesse exemplo, a todos os membros do grupo hadoop) para o diretório `warehouse` do Hive:

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

Outras variações mais restritivas devem ser usadas quando o controle de acesso mais granular for necessário.

Os sites a seguir fornecem informações de referência adicionais.

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

## Alterando portas usadas pelo Analytic Server

O Analytic Server usa a porta 9080 para HTTP e a porta 9443 para HTTPS, por padrão. Para alterar as configurações de porta, siga essas etapas.

1. Na guia Serviços do Ambari, navegue para a guia Configurações do serviço do Analytic Server.
2. Abra a seção **analytics.cfg avançado**.
3. Especifique as portas HTTP e HTTPS requeridas em **`http.port`** e em **`https.port`**, respectivamente.
4. Clique em **Salvar**.
5. Reinicie o serviço Analytic Server.

## Analytic Server de alta disponibilidade

É possível tornar o Analytic Server altamente disponível incluindo-o como um serviço para vários nós em seu cluster.

1. No console do Ambari, navegue para a guia Hosts.
2. Selecione um host que ainda não esteja executando o Analytic Server como um serviço.

3. Na guia Sumarização, clique em **Incluir** e selecione Analytic Server.
4. Clique em **Confirmar inclusão**

**Suporte para vários clusters**
O recurso de vários clusters é um aprimoramento do recurso Alta Disponibilidade do IBM SPSS Analytic Server e fornece isolamento melhorado em ambientes de locatários múltiplos. Por padrão, a instalação do serviço Analytic Server (no Ambari ou no ClouderaManager) resulta na definição de um único cluster de servidores analíticos.

A especificação de cluster define a associação de cluster do Analytic Server. A modificação da especificação de cluster é feita com conteúdo XML (no campo `cluster de analítica` da configuração do Ambari Analytic Server ou editando manualmente o arquivo `configuration/analytics-cluster.xml` do Cloudera Manager). Durante a configuração de vários clusters do Analytic Server, é necessário alimentar solicitações para cada cluster do Analytic Server com seu próprio balanceador de carga.

O uso do recurso de vários clusters garante que o trabalho para um locatário não possa afetar negativamente o trabalho sendo realizado no cluster de outro locatário. Com relação a tarefas altamente disponíveis, o failover de tarefa ocorre apenas dentro do escopo do cluster do Analytic Server no qual o trabalho foi iniciado. O exemplo a seguir fornece uma especificação XML para vários clusters.

**Nota:** Analytic Server pode ser transformado em altamente disponível incluindo-o como um serviço em vários nós em seu cluster.

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

No exemplo anterior, dois balanceadores de carga são necessários. Um balanceador de carga envia solicitações para os membros do `cluster1` (`one.cluster` e `two.cluster`) e o outro envia solicitações para os membros do `cluster2` (`three.cluster` e `four.cluster`).

O exemplo a seguir fornece uma única especificação XML de cluster (a configuração padrão).

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

No exemplo anterior, um único balanceador de carga é necessário para tratar casos nos quais há mais de um membro de cluster configurado.

**Notes**

- Apenas clusters singleton suportam o uso de curingas no elemento **`memberName`** (por exemplo, cardinalidade de cluster = "1"). Os valores válidos para o elemento de cardinalidade são 1 e 1+.
- O **`memberName`** deve ser especificado da mesma maneira que o nome do host ao qual a função Analytic Server é designada.
- Todos os servidores em todos os clusters devem ser reiniciados após as mudanças na configuração do cluster serem aplicadas.
- No Cloudera Manager, deve-se modificar e manter o arquivo `analytics-cluster.xml` em todos os nós do Analytic Server. Todos os nós devem ser mantidos para garantir que eles tenham o mesmo conteúdo.

## Otimizando opções de JVM para dados pequenos

É possível editar propriedades JVM para otimizar seu sistema ao executar tarefas pequenas (M3R).

No console do Ambari, veja a seção analytics-jvm-options avançadas da guia Configurações no serviço Analytic Server. Modificar os parâmetros a seguir configura o tamanho do heap para execução de tarefas no servidor que hospeda o Analytic Server; ou seja, não Hadoop. Isso será importante se você estiver executando pequenas tarefas (M3R), e talvez seja necessário experimentar esses valores para otimizar seu sistema.

```
-Xms512M
-Xmx2048M
```

## Upgrade do Python - HDP

Esta seção descreve o processo de upgrade manual do Python 2.x para o Python 3.7

1. Instalar o Python 3.7 em cada nó do cluster. Consulte o site do Python para obter mais informações.
2. Instalar o NumPy em cada nó do cluster. Consulte as instruções de instalação do NumPy https://numpy.org/install/ para obter mais informações.
3. Instalar o pandas em cada nó do cluster. Consulte as instruções de instalação do pandas https://pandas.pydata.org/getting_started.html para obter mais informações.
4. Incluir `spark.driver.python=<python3.7 executable path>` na seção **Custom analytics.cfg** da configuração do Ambari. Por exemplo:

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

## Atualizando as dependências do cliente

Esta seção descreve como atualizar as dependências do serviço do Servidor analítico usando o script `update_clientdeps`.

1. Efetue login no host do servidor Ambari como raiz.
2. Mude o diretório para `/var/lib/ambari-server/resources/stacks/<stack-name>/<stack-version>/services/ANALYTICSERVER/package/scripts`; consulte o exemplo a seguir.

   ```
   cd "/var/lib/ambari-server/resources/stacks/HDP/2.6/services/ANALYTICSERVER/package/scripts"
   ```

3. Execute o script `update_clientdeps` com os argumentos a seguir.

   **-u < ambari-user>**
   O nome de usuário da conta do Ambari

   **-p < ambari-password>**
   A senha para o usuário da conta do Ambari.

   **-h < ambari-host>**
   O nome do host do servidor Ambari.

   **-x < ambari-port>**
   A porta na qual o Ambari está atendendo.

   Veja o exemplo a seguir.

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

4. Reinicie o servidor Ambari usando o comando a seguir.

   ```
   ambari-server restart
   ```

## Configurando o Apache Knox

O Apache Knox Gateway é um sistema que fornece um único ponto de acesso seguro para os serviços do Apache Hadoop. O sistema simplifica a segurança do Hadoop para ambos os usuários (que acessam os dados do cluster e executam tarefas) e os operadores (que controlam acesso e gerenciam o cluster). O Gateway executa como um servidor (ou cluster de servidores) que serve um ou mais clusters do Hadoop.

**Nota:** O IBM SPSS Analytic Server não suporta o Apache Knox quando usado juntamente com a conexão única (SSO) do Kerberos.

O Apache Knox Gateway oculta, com eficiência, os detalhes de topologia de cluster do Hadoop e integra-se ao Enterprise LDAP e Kerberos. As seções a seguir fornecem informações sobre o Apache Knox necessário e as tarefas de configuração do Analytic Server.

### Pré-requisitos

- Um problema conhecido do Apache Knox não propaga as informações de segurança contidas em cookies e cabeçalhos HTTP (para obter mais informações, consulte https://issues.apache.org/jira/browse/KNOX-895). O problema está resolvido no Knox 0.14.0 (ou posterior). Deve-se obter uma distribuição atualizada do Hortonworks que inclua o Knox 0.14.0 (ou posterior) antes que o Knox funcione com o Analytic Server. Entre em contato com o fornecedor do Hortonworks para obter mais informações.
- Os nós do Analytic Server devem se conectar com o servidor Knox por meio de uma conexão SSH (shell seguro) sem senha. A conexão SSH (shell seguro) sem senha vai do Analytic Server para o Knox (**Analytic Server** > **Knox**).
- O Analytic Server deve ser instalado após o serviço Knox ser instalado.

Em alguns casos, problemas inesperados resultam em os arquivos de configuração não serem copiados automaticamente. Nesses casos, deve-se copiar manualmente os arquivos de configuração a seguir:

- `com.ibm.spss.knox_0.6-3.2.2.0.jar`: o arquivo deve ser copiado a partir do local do Analytic Server:

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

  para o nó do servidor Knox:

  `/KnoxServicePath/ext`

  Por exemplo: `/usr/iop/4.1.0.0/knox/ext`
- `rewrite.xml` e `service.xml`: os arquivos devem ser copiados do local do Analytic Server:

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

  para o nó do servidor Knox:

  `/KnoxServicePath/data/services`

  Por exemplo: `/usr/iop/4.1.0.0/knox/data/services`

  **Nota:** Há dois conjuntos de arquivos `rewrite.xml` e `service.xml` (um conjunto para o tráfego de `http://rest` e um conjunto para o tráfego de `ws://websocket`). Copie todos os arquivos `rewrite.xml` e `service.xml` do `analyticserver` e do `analyticserver_ws` para o nó do servidor Knox.

### Configurando o Ambari

O serviço do Analytic Server deve ser configurado na interface com o usuário do Ambari:

1. Na interface com o usuário do Ambari, navegue para **Knox** > **Configs** > **Topologia avançada**. As definições atuais de configuração do Knox são exibidas na janela **conteúdo**.
2. Inclua os dois serviços a seguir na seção **Topologia avançada** na configuração do Knox:

```
< service>
    < role> ANALYTICSERVER < /role>
    < url> http:// {: { /analyticserver < /url>
 < /service>
< service>
    < role> ANALYTICSERVER_WS < /role>
    < url> ws: // {: { /analyticserver < /url>
 < /service>
```

{analyticserver-host} e {analyticserver-port} devem ser substituídos pelo nome do servidor e número da porta do Analytic Server apropriados:

- A URL {analyticserver-host} pode ser encontrada na interface com o usuário do Ambari (**SPSS Analytic Server** > **Sumarização** > **Analytic Server**).
- O número de {analyticserver-port} pode ser localizado na interface com o usuário do Ambari (**SPSS Analytic Server** > **Configurações** > **analytics.cfg avançado** > **http.port**).

**Nota:** Quando o Analytic Server for implementado em vários nós e o LoadBalancer for usado, o {analyticserver-host} e o {analyticserver-port} deverão corresponder à URL do LoadBalancer e ao número da porta.

3. Reinicie o serviço Knox.

Quando o LDAP for usado, o Knox será padronizado para o LDAP "Demo" fornecido. É possível mudar para um servidor LDAP corporativo (como o Microsoft LDAP ou o OpenLDAP).

### Configurando o Analytic Server

Para usar LDAP para Analytic Server, o Analytic Server deve estar configurado para usar o mesmo servidor LDAP usado pelo Apache Knox. As entradas <value> para as configurações do Ambari a seguir devem ser atualizadas para refletir as configurações adequadas do servidor LDAP do Knox:

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

Os valores estão disponíveis na interface com o usuário do Ambari em: **Knox** > **Configs** > **Topologia avançada**. Por exemplo:

```
< param>
   < name>main.ldapRealm.userDnTemplate < /name>
   < value>  uid = {0}, ou=pessoas, dc=hadoop, dc=apache, dc=org  < /value>
</param>
< param>
   < name>main.ldapRealm.contextFactory.url < /name>
   < value>  ldap:// {3}{3}{3}{3}{8}{9} < /value>
</param>
```

Reinicie o serviço Knox após atualizar as configurações de LDAP do Knox.

**Importante:** A senha do administrador do Analytic Server deve ser a mesma que a senha do administrador do Knox.

### Configurando o Apache Knox

1. Atualize o arquivo gateway.jks do Knox:
   a. No servidor Knox, pare o serviço do Knox.
   b. Exclua o gateway.jks de /var/lib/knox/data-2.6.2.0-205/security/keystores.
   c. Reinicie o serviço Knox.
2. No servidor Knox, crie o subdiretório <knox_server>/data/service/analyticserver/3.2.2.0 e, em seguida faça upload dos arquivos service.xml e rewrite.xml para o novo diretório. Os dois arquivos estão no Analytic Server em <analytic_server>/configuration/knox/analyticserver/ (por exemplo, /opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/*.xml)
3. Em <knox_server>/bin, execute o script ./knoxcli.sh redeploy--cluster default
4. Faça upload do arquivo com.ibm.spss.knoxservice_0.6-*.jar para <knox_server>/ext. O arquivo está no Analytic Server em <analytic_server>/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.2.0.jar (por exemplo, /opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.2.2.0.jar).

5. Na interface com o usuário do Ambari, inclua o elemento a seguir em **Knox** > **Configurações** > **Topologia avançada**:

```
< service>
   < role> ANALYTICSERVER < /role>
   < url> http:// {: { /analyticserver < /url>
   < role> ANALYTICSERVER_WS < /role>
   < url> ws: // {: { /analyticserver < /url>
< /service>
```

   **Nota:** Por padrão, a funcionalidade do WebSocket está desativada. Ela pode ser ativada mudando a propriedade `gateway.websocket.feature.enabled` para `true` no arquivo `/conf/gateway-site.xml`.
6. Na interface com o usuário do Ambari, inclua ou atualize os usuários em **Knox** > **Configurações** > **Usuários avançados - ldif** (por exemplo, `admin`, `qauser1`, `qauser2`).
7. Reinicie LDAP a partir de **Knox** > **Ações de serviço** > **Iniciar LDAP demo**.
8. Reinicie o serviço Knox.

#### Estrutura da URL para o Apache Knox ativado do Analytic Server

A URL da interface com o usuário do Analytic Server ativada pelo Knox é `https://{knox-host}:{knox-port}/gateway/default/analyticserver/admin`

- Protocolo de https - os usuários devem aceitar um certificado para continuar no navegador da web.
- `knox-host` é o host do Knox.
- `knox-port` é o número da porta do Knox.
- O URI é `gateway/default/analyticserver`.

## Configurando uma Alocação de Recursos Dinâmicos separada para cada fila YARN - HDP

É possível configurar uma Alocação de Recursos Dinâmicos separada para cada fila YARN.

#### Mapeamento de modo de usuário e locatário - Hortonworks Data Platform

As tarefas de usuário e locatário podem ser enviadas a diferentes filas YARN e cada usuário ou locatário mapeia para uma fila YARN diferente (para aproveitar as vantagens da Alocação de Recursos Dinâmicos). O modo **`user`** ou o modo **`tenant`** podem ser definidos para mapeamento para filas YARN. Antes do Analytic Server 3.2.1 Fix Pack 1, todas as tarefas do Spark eram limitadas a uma única fila YARN.

A partir do IBM SPSS Analytic Server 3.2.1 Fix Pack 1, quando o fluxo de um usuário/locatário resulta em tarefas Spark sendo executadas no sistema, uma fila YARN separada será executada como o usuário/locatário que enviou o fluxo ao Analytic Server. Múltiplas filas YARN podem ser executadas simultaneamente para as diferentes tarefas de usuário/locatário.

Cada fila YARN continua a ser executada enquanto o usuário estiver conectado ao Analytic Server (e por algum tempo após o usuário efetuar logout e não houver mais tarefas de usuário ativas). A quantidade de tempo após o logout pode ser controlada pela variável de configuração: **`as.spark.driver.cleanup.delay`**.

Um processo **`SparkDriver`** é criado para cada usuário que envia a tarefa Spark. O processo **`SparkDriver`** de cada usuário é finalizado depois que o usuário não possuir tarefas ativas por cerca de 2 minutos (o valor padrão) e nenhuma atividade **`HTTPSession`**.

**Nota:** Todos os processos **`SparkDriver`** são finalizados quando o Analytic Server é encerrado.

Use as etapas a seguir para incluir o Analytic Server a um cluster existente:

1. Na interface com o usuário do Ambari, navegue até a guia **Serviço do SPSS Analytic Server** > **Configurações** > **analytics.cfg avançado**.
2. Altere o valor de **resource.pool.enabled** para `true`.

3. Inclua as propriedades a seguir na guia **analytics.cfg customizado**:

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

*Tabela 8. Propriedades do analytics.cfg customizado*

| Propriedade | descrição |
|---|---|
| `yarn.queue.mode` | Configure o modo de mapeamento para filas YARN. Quando `yarn.queue.mode=user`, uma fila YARN separada é executada para cada usuário que enviou uma tarefa/fluxo para o Analytic Server. Múltiplas filas YARN podem ser executadas simultaneamente para as tarefas/fluxos de diferentes usuários.<br>Quando `yarn.queue.mode=tenant`, uma fila YARN separada é executada para cada locatário que enviou uma tarefa/fluxo para o Analytic Server. Múltiplas filas YARN podem ser executadas simultaneamente para os diferentes fluxos/tarefas do locatário. |
| `yarn.queue.mapping` | Mapeia os pares de usuário ou locatário para as filas YARN que são definidas no Gerenciador de Filas YARN. Os pares devem ser separados por vírgulas (por exemplo, `tenant1:test,tenant2:production` para locatários ou `user1:test,user2:production`) para usuários. |
| `yarn.queue.default` | O nome da fila YARN padrão para a qual o aplicativo é enviado. É possível especificar um nome de fila customizado do YARN no Gerenciador de Filas do YARN. |
| `as.spark.driver.cleanup.delay` | Um número inteiro que representa o número de minutos após o logout antes de finalizar a fila YARN de um usuário. O valor padrão é **2**. Esta propriedade é opcional. |
| `as.sparkdriver.max.memory` | Configura a quantidade de memória que é usada por cada processo **`SparkDriver`**. O valor padrão é **`1G`**. Essa propriedade é opcional. |

4. Salve a configuração e reinicie o serviço do Analytic Server.

**Referência**

Consulte os sites a seguir para obter mais informações:

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

## Migrando IBM SPSS Analytic Server no Ambari

O Analytic Server pode migrar dados e definições de configuração de uma instalação do Analytic Server existente para uma nova instalação. A migração pode ocorrer no mesmo ambiente em cluster ou em um novo ambiente em cluster.

### Migrando a partir do Analytic Server 3.2.1.1 para o 3.2.2 no mesmo cluster de servidores

Se você tiver uma instalação existente do Analytic Server 3.2.1.1, é possível migrar as suas definições de configuração do 3.2.1.1 para a sua instalação do 3.2.2 no mesmo cluster de servidores.

1. Colete as definições de configuração a partir da versão antiga do Analytic Server (Analytic Server 3.2.1.1).
   a. Expanda o archive `{AS_ROOT}\tools\unzip configcollector.zip` (ele criará uma nova pasta denominada `configcollector`).

b. Execute o script `configcollector.sh` na pasta `configcollector`. Copie o arquivo resultante compactado (ZIP) `ASConfiguration_3.2.1.1.xxx.zip` para um local da pasta diferente (como um backup).

2. Faça backup da raiz analítica da sua instalação da versão antiga do Analytic Server 3.2.1.1 para um novo local.

   a. Se você não tiver certeza sobre a localização da raiz analítica, execute o comando **hadoop fs -ls**. O caminho para a raiz analítica é semelhante a `/user/as_user/analytic-root/analytic-workspace`, em que `as_user` é o ID do usuário proprietário da raiz analítica.

   b. Use os comandos **hadoop fs -copyToLocal** e **hadoop fs -copyFromLocal** para copiar a pasta do Analytic Server `analytic-workspace` versão antiga para um novo local (por exemplo, `/user/as_user/analytic-root/AS3211Location`).

3. Se você usar o Apache Directory Server integrado, faça backup da configuração atual de usuário/grupo com uma ferramenta de cliente LDAP de terceiros. Após o Analytic Server 3.2.2 ser instalado, importe a configuração do usuário/grupo de backup no Apache Directory Server.

   **Nota:** Esta etapa pode ser ignorada se você usa um servidor LDAP externo.

4. Abra o console do Ambari e pare o **Serviço do Analytic Server**.
5. Desinstale a versão antiga do Analytic Server (Analytic Server 3.2.1.1) e, em seguida, instale o Analytic Server 3.2.2. Para obter instruções de instalação, consulte Capítulo 2, “Instalação e configuração do Ambari”, na página 5.
6. Abra o console do Ambari e pare o **Serviço do Analytic Server** (no Ambari, assegure que o **Serviço de metastore analítico** esteja em execução).
7. Copie o backup da raiz analítica Analytic Server 3.2.1.1, da etapa 2, para o local da nova versão do Analytic Server.

   a. Remova `analytic-workspace` da nova versão instalada do Analytic Server.

   b. Copie o backup da pasta da área de trabalho analítica do Analytic Server 3.2.1.1 (`/user/as_user/analytic-root/AS3211Location`) para o local da nova versão (por exemplo, `/user/as_user/analytic-root/analytic-workspace`). Deve-se assegurar que o proprietário da área de trabalho analítica esteja definido como `as_user`.

8. Limpe o estado do Zookeeper. No diretório bin do Zookeeper (por exemplo, `/usr/hdp/current/zookeeper-client` no Hortonworks), execute o comando a seguir:

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

9. Copie o archive de backup `ASConfiguration_3.2.1.1.xxx.zip` da etapa 1 para o novo local da versão Analytic Server (por exemplo, `/opt/ibm/spss/analyticserver/3.2/`).
10. Execute a ferramenta de migração executando o script **migrationtool.sh** e passando o caminho do archive `ASConfiguration_3.2.1.1.xxx.zip` (que foi criado pelo coletor de configuração) como um argumento. Por exemplo:

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

11. Execute o comando a seguir de um shell de comando no nó do Analytic Server:

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

12. No console do Ambari, inicie o **Serviço do Analytic Server**.

### Migrando a partir do Analytic Server 3.2.1.1 para o 3.2.2 em um novo cluster de servidores

Se você tiver uma instalação existente do Analytic Server 3.2.1.1, é possível migrar as definições de configuração do 3.2.1.1 para a instalação do 3.2.2 em um novo cluster de servidores.

1. Instale a nova versão do Analytic Server de acordo com as instruções em “Instalação no Ambari” na página 7.
2. Copie a área de trabalho analítica da sua instalação antiga para a nova.

a. Se você não tiver certeza sobre a localização da área de trabalho analítica, execute `hadoop fs -ls`. O caminho para a área de trabalho analítica é semelhante a `/user/as_user/analytic-root/analytic-workspace`, em que `as_user` é o ID do usuário proprietário da área de trabalho analítica.

b. Remova `analytic-workspace` do novo servidor.

c. Use `hadoop fs -copyToLocal` e `hadoop fs -copyFromLocal` para copiar a área de trabalho analítica do servidor antigo para a pasta `/user/as_user/analytic-root/analytic-workspace` do novo servidor (assegure que o proprietário esteja configurado como `as_user`).

3. Se você usar o Apache Directory Server integrado, faça backup da configuração atual de usuário/grupo com uma ferramenta de cliente LDAP de terceiros. Após o Analytic Server 3.2.2 ser instalado, importe a configuração do usuário/grupo de backup no Apache Directory Server.

   **Nota:** Esta etapa pode ser ignorada se você usa um servidor LDAP externo.

4. No novo servidor, abra o console do Ambari e pare o serviço do Analytic Server (no Ambari, assegure que o Serviço do servidor analítico esteja em execução).

5. Colete as definições de configuração a partir da instalação antiga.

   a. Copie o archive `configcollector.zip` em sua nova instalação para `{AS_ROOT}\tools` em sua antiga instalação.

   b. Extraia a cópia do `configcollector.zip`, que cria um novo subdiretório `configcollector` na antiga instalação.

   c. Execute a ferramenta coletora de configuração na antiga instalação executando o script **`configcollector`** em `{AS_ROOT}\tools\configcollector`. Copie o arquivo compactado resultante (ZIP) no servidor que hospeda sua nova instalação.

   **Importante:** O script **`configcollector`** fornecido pode não ser compatível com a versão do Analytic Server mais recente. Entre em contato com o representante de suporte técnico IBM se você encontrar problemas com o script **`configcollector`**.

6. Limpe o estado do Zookeeper. No diretório bin do Zookeeper (por exemplo, `/usr/hdp/current/zookeeper-client` no Hortonworks), execute o comando a seguir.

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

7. Execute a ferramenta de migração executando o script **`migrationtool`** e passando o caminho do arquivo compactado que foi criado pelo coletor de configuração como um argumento. A seguir está um exemplo.

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

8. Execute o comando a seguir de um shell de comando no nó do Analytic Server:

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

9. No console do Ambari, inicie o serviço Analytic Server.

   **Nota:** Se você configurou R para usar com a instalação do Analytic Server existente, siga as etapas para configurá-lo com a nova instalação do Analytic Server.

## Desinstalando

**Importante:** Quando o Essentials for R estiver instalado, deve-será executar primeiro o script `remove_R.sh`. Falha ao desinstalar o Essentials for R, antes de desinstalar o Analytic Server, o que resulta na incapacidade de desinstalar o Essentials for R posteriormente. O script `remove_R.sh` será removido quando o Analytic Server for desinstalado. Para obter informações sobre a desinstalação do Essentials for R, consulte “Desinstalando o Essentials for R” na página 39.

1. No host Analytic Metastore, execute o script `remove_as.sh` no diretório `{AS_ROOT}/bin` com os parâmetros a seguir.

**u**
Obrigatório. O ID do usuário do administrador do Ambari Server.

**p**
Obrigatório. A senha do administrador do Ambari Server.

**h**
Obrigatório. O nome do host do Ambari Server.

**x**
Obrigatório. A porta do Ambari Server.

**l**
Opcional. Ativa o modo seguro.

Os exemplos são os seguintes.

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

Remove o Analytic Server de um cluster com o host Ambari one.cluster.

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

Remove o Analytic Server de um cluster com o host Ambari one.cluster, no modo seguro.

**Nota:** Essa operação remove a pasta do Analytic Server no HDFS.

**Nota:** Essa operação não remove nenhum esquema do Db2 associado ao Analytic Server. Consulte a documentação do Db2 para obter informações sobre como remover esquemas manualmente

## Desinstalando o Essentials for R

1. No host do Essentials for R, execute o script `remove_R.sh` no diretório `{AS_ROOT}/bin` com os parâmetros a seguir.

   **u**
   Obrigatório. O ID do usuário do administrador do Ambari Server.

   **p**
   Obrigatório. A senha do administrador do Ambari Server.

   **h**
   Obrigatório. O nome do host do Ambari Server.

   **x**
   Obrigatório. A porta do Ambari Server.

   **l**
   Opcional. Ativa o modo seguro.

   Os exemplos são os seguintes.

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

   Remove o Essentials for R de um cluster com o host do Ambari one.cluster.

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

   Remove o Essentials for R de um cluster com o host do Ambari one.cluster, no modo seguro.
2. Remova o diretório de serviços R do diretório de serviços do servidor do Ambari. Por exemplo, no HDP 2.6, o diretório ESSENTIALR está localizado em `/var/lib/ambari-server/resources/stacks/HDP/2.6/services`.
3. No console do Ambari, verifique se o serviço do Essentials for R não existe mais.

# Capítulo 3. Instalação e configuração do Cloudera

## Visão geral do Cloudera

O Cloudera é uma distribuição de software livre do Apache Hadoop. A Distribuição do Cloudera, incluindo o Apache Hadoop (CDH), destina-se às implementações de classe corporativa dessa tecnologia.

Analytic Server pode ser executado na plataforma CDH. O CDH contém os elementos principais do Hadoop que fornecem processamento de dados confiável, escalável e distribuído de grandes conjuntos de dados (principalmente MapReduce e HDFS), assim como outros componentes orientados pela empresa que fornecem segurança, alta disponibilidade e integração com hardware e outro software.

## Pré-requisitos específicos do Cloudera

Além dos pré-requisitos gerais, revise as informações a seguir.

**Serviços**

Assegure-se de que as instâncias a seguir foram instaladas em cada host do Analytic Server.

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

As instâncias a seguir são necessárias somente quando seus recursos são usados.

- Accumulo: Gateway
- HBase: Gateway, Principal ou RegionServer
- Spark 2: Gateway

**Repositório de metadados**

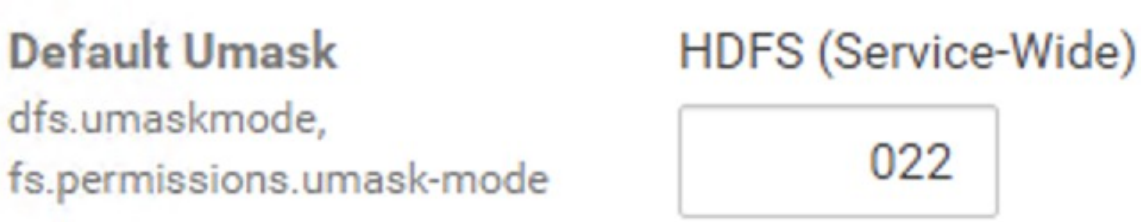
É possível usar o Db2 e o MySQL como o repositório de metadados do Analytic Server. Se você planeja usar o MySQL como repositório de metadados do Analytic Server, siga as instruções para "Configurando MySQL para Analytic Server" na página 43.

**SSH sem senha**

Configure o SSH sem senha para o usuário raiz entre o host do Analytic Server e todos os hosts no cluster.

**Umask padrão**

A configuração de Umask padrão deve ser configurada como 022. Por exemplo:




A configuração 022 é a Umask mais restritiva que permite que o Analytic Server funcione.

### Ambientes Cloudera ativados para Kerberos

Se você planeja instalar o Analytic Server em um ambiente Cloudera ativado para Kerberos, deve-se verificar se o Kerberos está configurado corretamente de uma maneira compatível com o Analytic Server.

As seções a seguir se aplicam aos ambientes Cloudera em que o Kerberos já está instalado. As seções a seguir devem ser seguidas antes da instalação do Analytic Server no Cloudera. Supõe-se que você tenha

conhecimento básico de autenticação do Kerberos, já que as seções incluem terminologia específica do Kerberos (por exemplo, **kinit**, **kadmin**, etc.).

**Nota:** O Analytic Server inspeciona a configuração de HDFS para valores relacionados do Kerberos a serem usados para autenticação.

### Autenticação do Kerberos

Verifique se a autenticação do Kerberos está configurada em cada nó do cluster do Cloudera antes de instalar o Analytic Server. Para obter mais informações, consulte Configurando autenticação no Cloudera Manager na documentação do produto Cloudera.

**Nota:** Após configurar a autenticação do Kerberos em cada nó do cluster do Cloudera, os serviços **cloudera-scm-server** e **cloudera-scm-agent** deverão ser reiniciados antes da instalação do Analytic Server. O serviço **cloudera-scm-agent** deve ser reiniciado em todos os nós do cluster.

### Criando as Contas Necessárias no Kerberos

1. Crie contas no repositório do usuário do Kerberos para todos os usuários aos quais você planeja conceder acesso ao Analytic Server.
2. Crie as mesmas contas (da etapa anterior) no servidor LDAP.
3. Crie uma conta do usuário do S.O. para cada um dos usuários criados em uma etapa anterior em cada um dos nós do Analytic Server e do nó do Hadoop. O grupo de usuários deve ser configurado como `hadoop`.
   - Certifique-se de que o UID para esses usuários corresponda em todas as máquinas. É possível testar isso usando o comando `kinit` para efetuar login em cada uma das contas.
   - Certifique-se de que o ID do usuário esteja de acordo com a configuração do YARN de **ID de usuário mínimo para o envio da tarefa**. Essa é a configuração de **min.user.id** em `container-executor.cfg`. Por exemplo, se **min.user.id** for 1000, então cada conta do usuário criada deverá ter um UID maior ou igual a 1000.
4. Crie uma pasta inicial do usuário no HDFS para o usuário administrador do Analytic Server. A permissão da pasta deve ser configurada como 755, o proprietário deve ser configurado como `admin` e o grupo de usuários deve ser configurado como `hdfs`. Veja o exemplo em **negrito** a seguir:

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

Localizados 9 itens

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

5. Crie pastas iniciais do usuário no HDFS para todos os usuários padrão do Analytic Server (por exemplo, `user1`). O proprietário da pasta é o usuário real e o grupo de usuários deve ser configurado como `hdfs`.
6. Se planejar usar origens de dados do HCatalog e o Analytic Server estiver instalado em uma máquina diferente do Hive metastore, será necessário personificar o cliente Hive no HDFS.
   a. Navegue para a guia Configuração do serviço do HDFS no Cloudera Manager.

      **Nota:** As configurações a seguir podem não aparecer na guia **Configuração** se elas ainda não tiverem sido configuradas. Neste caso, execute uma procura para encontrá-los.

   b. Edite a configuração de **hadoop.proxyuser.hive.groups** para ter o valor * ou um grupo que contenha todos os usuários com permissão para efetuar login no Analytic Server.
   c. Edite a configuração de **hadoop.proxyuser.hive.hosts** para ter o valor * ou a lista de hosts nos quais o metastore Hive e cada instância do Analytic Server estão instalados como serviços.
   d. Reinicie o serviço HDFS.

Após a execução dessas etapas e a instalação do Analytic Server, o Analytic Server configura o Kerberos de forma silenciosa e automática.

#### Ativando a personificação do Kerberos

A personificação permite que um encadeamento seja executado em um contexto de segurança que difere do contexto de segurança do processo que possui o encadeamento. Por exemplo, a personificação fornece um meio de as tarefas do Hadoop serem executadas como usuário que não o usuário padrão do Analytic Server (`as_user`). Para ativar a personificação do Kerberos:

1. Abra o Cloudera Manager e inclua ou atualize as propriedades a seguir na área **Fragmento de configuração avançada para todo o cluster (válvula de segurança) para core-site.xml** (localizada na guia **HDFS (para todo o serviço)** > **Configuração**).
   - **Nome:** `hadoop.proxyuser.as_user.hosts`
   - **Valor:** `*`
   - **Nome:** `hadoop.proxyuser.as_user.groups`
   - **Valor:** `*`

   **Nota:** As configurações de **core-site.xml** se aplicam à configuração do Hadoop (não do Analytic Server).
2. Execute o comando a seguir de um shell de comando no nó do Analytic Server:

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

## Configurando MySQL para Analytic Server

Configurar o IBM SPSS Analytic Server no Cloudera Manager requer a instalação e configuração de um banco de dados do servidor do MySQL.

1. Execute o comando a seguir a partir de uma janela de comando no nó em que o banco de dados do MySQL está armazenado:

```
yum install mysql-server
```

   **Nota:** Use `zypper install mysql` para o SuSE Linux.
2. Execute o comando a seguir a partir de uma janela de comando em cada nó do cluster do Cloudera:

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

   **Nota:** Use `sudo zypper install mysql-connector-java` para o SuSE Linux.
3. Decida sobre e anote o nome do banco de dados do Analytic Server, nome de usuário do banco de dados e a senha do banco de dados que o Analytic Server usa ao acessar o banco de dados do MySQL.
4. Instale o Analytic Server de acordo com as instruções no “Instalação no Cloudera” na página 45.
5. Copie o script `/opt/cloudera/parcels/AnalyticServer/bin/add_mysql_user.sh` de um dos servidores gerenciados pelo Cloudera no nó em que o banco de dados do MySQL está instalado. Execute o script com os parâmetros apropriados para a sua configuração específica. Por exemplo:

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

   **Notes:** O parâmetro a `-r <dbRootPassword>` é requerido quando o banco de dados for executado no modo seguro (a senha do usuário raiz é configurada).

   Os parâmetros `-r <dbUserPassword>` e `-t <dbUserName>` são necessários quando o banco de dados está em execução no modo seguro com um nome de usuário diferente de `root`.

# Ferramentas de pré-verificação e pós-verificação de instalação - Cloudera

### Local e pré-requisitos da ferramenta

Antes de instalar o serviço do Analytic Server, execute a ferramenta de pré-verificação em todos os nós que farão parte do serviço do Analytic Server para verificar se seu ambiente Linux está pronto para instalar o Analytic Server.

A ferramenta de pré-verificação é chamada automaticamente como parte da instalação. A ferramenta verifica cada nó do Analytic Server antes de executar a instalação em cada host. Também é possível chamar manualmente a ferramenta de pré-verificação em cada nó, que pode validar a máquina antes da instalação do serviço.

Após executar o arquivo binário autoextrator do Analytic Server, a ferramenta de pré-verificação estará localizada nos diretórios a seguir:

- **Cloudera**

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

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

**Nota:** A ferramenta de pré-verificação não está disponível no diretório `tools` até que você execute o arquivo binário executável e, em seguida, distribua (**Fazer download** > **Distribuir**) e ative o Analytic Server na página Pacotes do Cloudera Manager.

Depois de instalar o Analytic Server, a ferramenta de pós-verificação está localizada no seguinte diretório:

- **Cloudera**

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

As ferramentas devem ser executadas como raiz e requerem o Python 2.6.X (ou superior).

Se a ferramenta de pré-verificação relatar alguma falha, as falhas deverão ser resolvidas antes de você continuar com a instalação do Analytic Server.

### Executando a ferramenta de pré-verificação

#### Automática

A ferramenta de pré-verificação pode ser chamada automaticamente como parte da instalação do Analytic Server quando o Analytic Server é instalado como um serviço por meio do console do Cloudera Manager. deve-se inserir manualmente o nome do usuário e a senha do administrador do Cloudera Manager:


*Figura 4. Configurações do Administrador do Cloudera Manager*

#### Manual

É possível chamar manualmente a ferramenta de pré-verificação em cada nó do cluster.

O exemplo de pré-verificação a seguir verifica o cluster do Cloudera, `MyCluster`, que está em execução em `myclouderahost.ibm.com:7180` e usa as credenciais de login `admin:admin`:

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

**Notes:**

- Os argumentos `--target`, `--host`, `--port` e `--username` são necessários.
- O valor `--host` deve ser fornecido por um endereço IP ou por um nome completo do domínio.
- A ferramenta solicitará uma senha quando o argumento de senha for omitido.
- O comando `precheck.py` inclui ajuda de uso, que é exibida com o argumento `--h` (`python ./precheck.py --help`).
- O argumento `--cluster` é opcional (o atual cluster é identificado quando `--cluster` não é usado).

Conforme a ferramenta de pré-verificação executa as suas verificações, o status de cada verificação é exibido na janela de comando. Quando uma falha ocorre, informações detalhadas ficam disponíveis no arquivo de log (o local exato do arquivo de log é fornecido na janela de comando). O arquivo de log poderá ser fornecido para o suporte técnico IBM quando for necessário mais suporte.

#### Executando a ferramenta de pós-verificação

A ferramenta de pós-verificação verifica se o Analytic Server está sendo executado adequadamente e pode processar tarefas simples. O exemplo de pós-verificação a seguir verifica uma instância do Analytic Server que está em execução em `myanalyticserverhost.ibm.com:9443`, com SSL ativado e usa as credenciais de login `admin:ibmspss`:

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

Quando o Knox é usado com o Analytic Server, o comando é como segue:

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

Para executar uma verificação única, use o comando a seguir:

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

**Notes:**

- Os argumentos `--target`, `--host`, `--port` e `--username` são necessários.
- O valor `--host` deve ser fornecido por um endereço IP ou por um nome completo do domínio.
- A ferramenta solicitará uma senha quando o argumento de senha for omitido.
- O comando `postcheck.py` inclui ajuda de uso, que é exibida com o argumento `--h` (`python ./postcheck.py --help`).

Conforme a ferramenta de pós-verificação executa as suas verificações, o status de cada verificação é exibido na janela de comando. Quando uma falha ocorre, informações detalhadas ficam disponíveis no arquivo de log (o local exato do arquivo de log é fornecido na janela de comando). O arquivo de log poderá ser fornecido para o suporte técnico IBM se for necessário mais suporte.

## Instalação no Cloudera

As etapas a seguir explicam o processo de instalação manual do IBM SPSS Analytic Server no Cloudera Manager.

#### Analytic Server 3.2.2

#### Instalação online

1. Navegue para o Website do IBM Passport Advantage® e faça o download do arquivo binário autoextrator específico para a sua pilha, versão de pilha e arquitetura de hardware para um host dentro do cluster do Cloudera. Os binários disponíveis do Cloudera são:

| *Tabela 9. Arquivos binários autoextratores do Servidor analítico* | |
|---|---|
| **descrição** | **Nome do arquivo binário** |
| IBM SPSS Analytic Server 3.2.2 for Cloudera 5.11, 5.12, 5.13, 5.14, 5.15, 6.0, 6.1, 6.2 e 6.3 Ubuntu English | `spss_as-3.2.2.0-cdh5.11-6.3-ubun.bin` |
| IBM SPSS Analytic Server 3.2.2 for Cloudera 5.11, 5.12, 5.13, 5.14, 5.15, 6.0, 6.1, 6.2 e 6.3 Linux x86-64 English | `spss_as-3.2.2.0-cdh5.11-6.3-lx86.bin` |

2. Execute o instalador de autoextração `*.bin` do Cloudera no nó do cluster principal do Cloudera Manager. Siga os prompts da instalação aceitando o contrato de licença e mantendo o diretório de instalação do CSD padrão.

   **Nota:** Deve-se especificar um diretório do CSD diferente se ele for alterado a partir do local padrão.

3. Use o comando a segui para reiniciar o Cloudera Manager após a instalação ser concluída:

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

4. Abra a interface do Cloudera Manager (por exemplo, `http://${CM_HOST}:7180/cmf/login` com as credenciais de login padrão de `admin`/`admin`), atualize as **URLs dos repositórios de pacotes remotos** (localizadas em **Host** > **Pacotes** > **clique em Configuração**) e verifique se a URL está correta. Por exemplo:

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

   **Nota:** A **Frequência de atualização do pacote** e as **URLs do repositório do pacote remoto** podem ser atualizadas para atender às suas necessidades específicas.

5. Após a atualização dos arquivos do pacote pelo Cloudera Manager (é possível atualizar os arquivos do pacote manualmente clicando em **Verificar se há novos pacotes**), você verá que o status do pacote **AnalyticServer** está configurado como **Disponível remotamente**.
6. Selecione **Download** > **Distribuir** > **Ativar**. O status do pacote **AnalyticServer** é atualizado para **Distribuído, ativado**.
7. No Cloudera Manager, inclua Analytic Server como um serviço e decida onde colocar o Analytic Server. É necessário fornecer as informações a seguir em **Incluir assistente de serviço**:

   **Nota:** O **Incluir assistente de serviço** mostra o progresso geral durante cada fase do processo de criação de serviço e fornece uma mensagem de confirmação final quando o serviço for instalado e configurado com êxito no cluster.

   - Nome do host de metastore do Analytic Server
   - Nome do banco de dados de metastore do Analytic Server
   - Nome do usuário de metastore do Analytic Server
   - Senha de metastore do Analytic Server

   **MySQL como o repositório de metadados Analytic Server**

   - Classe do driver de metastore Analytic Server : `com.mysql.jdbc.Driver`
   - URL do repositório de metastore Analytic Server : `jdbc:mysql://${/{?createDatabaseIfNotExist=true`

     `{MySQL_DB}` é o nome do host do servidor no qual o MySQL está instalado

   **Db2 como o repositório de metadados Analytic Server**

   - Classe do driver de metastore Analytic Server : `com.ibm.db2.jcc.DB2Driver`

- URL do repositório de metastore Analytic Server : `jdbc:db2://{2}_HOST}:{:{/{:currentSchema={;`

  `{Db2_HOST}` é o nome do host do servidor no qual o Db2 está instalado.

  `{PORT}` é a porta na qual o Db2 está atendendo.

  `{SchemaName}` é um esquema disponível não utilizado.

  Trabalhe com seu administrador do Db2 se você não tiver certeza dos valores a serem inseridos.

**Configuração LDAP**

O Analytic Server usa um servidor LDAP para armazenar e autenticar usuários e grupos. Você fornece as informações de configuração LDAP necessárias durante a instalação do Analytic Server.

*Tabela 10. Definições de configuração LDAP*

| **Configuração LDAP** | **descrição** |
|---|---|
| `as.ldap.type` | Tipo de LDAP. O valor pode ser `ads`, `ad` ou `openladp`.<br>• `ads` - Apache Directory Server (configuração padrão)<br>• `ad` - Microsoft Active Directory<br>• `openladp` - OpenLDAP |
| `as.ldap.host` | Host LDAP |
| `as.ldap.port` | Número da porta LDAP |
| `as.ldap.binddn` | DN de ligação de LDAP |
| `as.ldap.bindpassword` | Senha do DN de ligação de LDAP |
| `as.ldap.basedn` | DN base de LDAP |
| `as.ldap.filter` | Regra de filtro de grupo e de usuário LDAP<br>**Nota:** Quando esse valor contiver caracteres de barra vertical \|, os caracteres deverão ser escapados com caracteres de barra invertida (por exemplo, \\\|). |
| `as.ldap.ssl.enabled` | Especifica se o SSL deve ser usado para comunicação entre o Analytic Server e o LDAP. O valor pode ser `true` ou `false`. |
| `as.ldap.ssl.reference` | ID de referência SSL LDAP |
| `as.ldap.ssl.content` | Configuração SSL LDAP |

- Por padrão, `as.ldap.type` é configurado para `ads` e as outras configurações relacionadas contêm configurações padrão. A exceção é que se deve fornecer uma senha para a configuração de `as.ldap.bindpassword`. O Analytic Server usa as definições de configuração para instalar um Apache Directory Server (ADS) e executar a inicialização do servidor. O perfil padrão do ADS inclui o usuário `admin` com uma senha `admin`. É possível executar o gerenciamentogerenciamento de usuários pelo Analytic Server Console ou importar informações sobre o usuário e o grupo de um arquivo XML por meio do script `importUser.sh`, localizado na pasta `<AnalyticRoot>/bin`.
- Se você planeja usar um servidor LDAP externo, como o Microsoft Active Directory ou o OpenLDAP, deve-se definir as definições de configuração de acordo com os valores reais de LDAP. Para obter mais informações, consulte Configurando registros de usuários LDAP no Liberty.

- É possível mudar a configuração de LDAP após a instalação do Analytic Server (por exemplo, mudando do Apache Directory Server para o OpenLDAP). No entanto, se você iniciar primeiramente com o Microsoft Active Directory ou o OpenLDAP e decidir alternar posteriormente para o Apache Directory Server, o Analytic Server não instalará um Apache Directory Server durante a instalação. O Apache Directory Server é instalado apenas quando ele é selecionado durante a instalação inicial do Analytic Server.

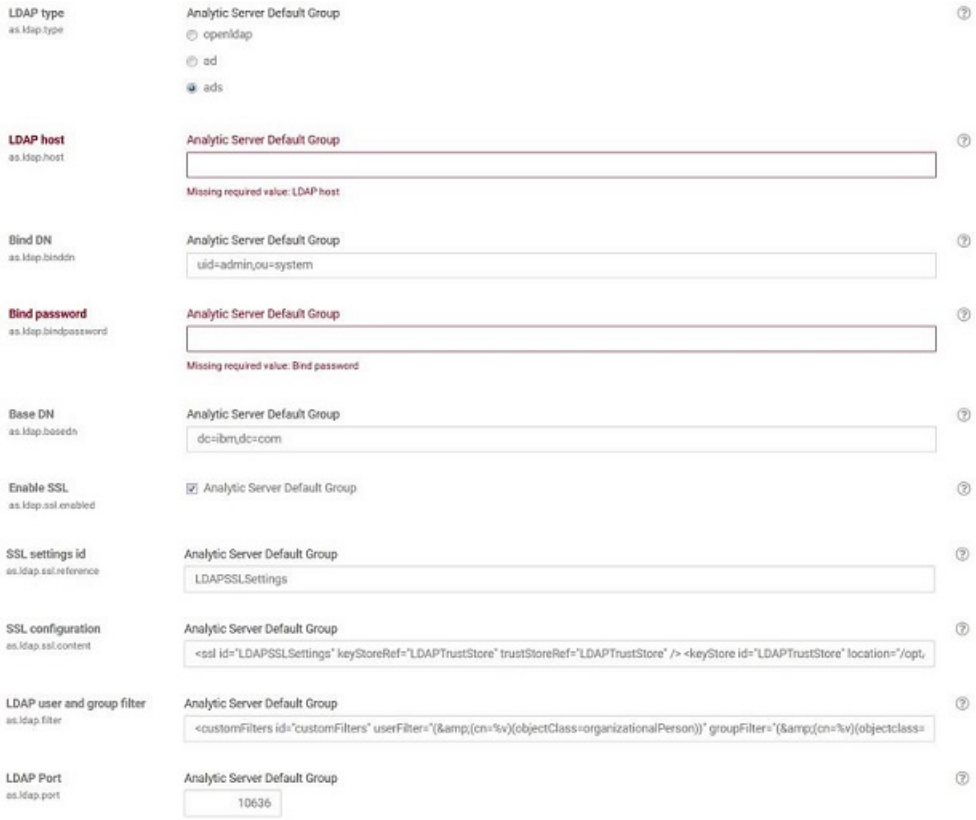



*Figura 5. Exemplo de definições de configuração de LDAP*

8. Ao instalar o Analytic Server em um ambiente Cloudera ativado para Kerberos, as configurações a seguir também devem ser configuradas no **Assistente para incluir serviço**:

   **Nota:** O Analytic Server inspeciona a configuração de HDFS para valores relacionados do Kerberos a serem usados para autenticação.

   - Selecione `Kerberos` como a configuração de **Segurança do Analytic Server** se desejar ativar a autenticação do Kerberos ao efetuar login no console do Analytic Server. Quando **Kerberos** for selecionado como a configuração de **Segurança do Analytic Server**, o console do Analytic Server será padronizado para o modo de login do Kerberos.
   - Selecione `Kerberos` como a configuração de **Método de conexão de origem de dados do banco de dados do Analytic Server** quando desejar se conectar aos bancos de dados ativados para Kerberos. Quando **Kerberos** for selecionado como a configuração de **Método de conexão de origem de dados do banco de dados do Analytic Server**, o console do Analytic Server usará o modo do Kerberos ao conectar-se a um banco de dados.
   - As configurações de **Nome da região do Kerberos** e de **Host do KDC** são necessárias. Os valores **Nome da região do Kerberos** (**`as.kdc.realms`**) e **Host do KDC** (**`kdcserver`**) estão localizados no arquivo `krb5.conf` no servidor Key Distribution Center (KDC) do Kerberos.

     Vários nomes de região são suportados quando separados por caracteres de vírgula. Os nomes de região do Kerberos especificados correspondem a, e estão associados a, nomes de usuário. Por exemplo, os nomes de usuário `UserOne@us.ibm.com` e `UserTwo@eu.ibm.com` corresponderão às regiões `us.ibm.com,eu.ibm.com`.

A confiança entre regiões do Kerberos deve ser configurada quando mais de uma região for especificada como um **Nome de região do Kerberos**. O nome do usuário inserido durante o prompt de login do console do Analytic Server é inserido sem o sufixo de nome de região. Como resultado, quando várias regiões são especificadas, uma lista suspensa de **Regiões** é apresentada aos usuários, permitindo-lhes selecionar a região.

**Nota:** Quando apenas uma região for especificada, os usuários não verão uma lista suspensa de **Regiões** ao se inscreverem no Analytic Server.

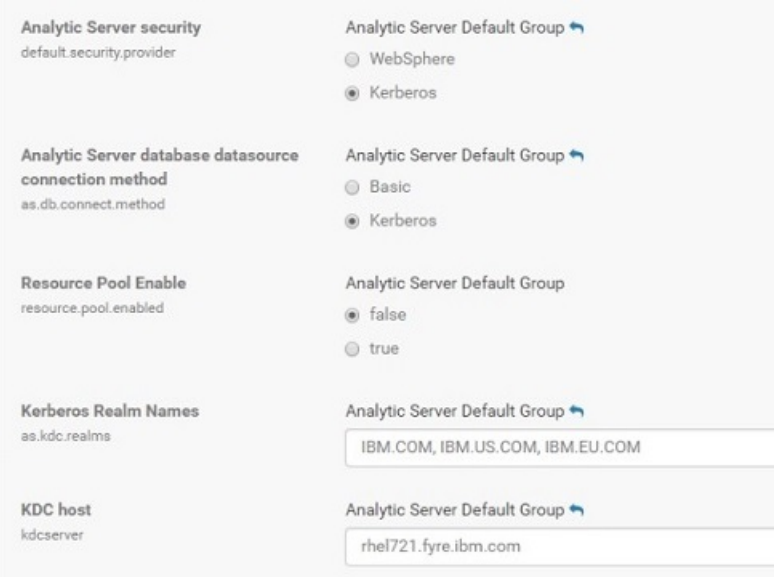



*Figura 6. Configurações de Kerberos de Exemplo*

**Notes:**

- As configurações de **Segurança do Analytic Server** e de **Método de conexão de origem de dados do banco de dados do Analytic Server** são aplicáveis à autenticação do cliente do IBM SPSS Modeler e do console do Analytic Server.
- Quando o **Método de conexão de origem de dados do banco de dados do Analytic Server** é configurado para `Kerberos`, deve-se assegurar que os bancos de dados de destino também sejam ativados para o Kerberos.
- As configurações de **Segurança do Analytic Server** e de **Método de conexão de origem de dados do banco de dados do Analytic Server** não configuram a autenticação do Kerberos no cluster do Hadoop. Para obter mais informações, consulte a seção "Ativando a personificação do Kerberos".
- Se você deseja que a autenticação do Kerberos seja ativada no login, deve-se implementar o cliente do IBM SPSS Modeler como um cliente Kerberos válido. Isso é realizado usando o comando **`addprinc`** no servidor do centro de distribuição de chaves (KDC) do Kerberos. Para obter mais informações, consulte a documentação do IBM SPSS Modeler.

Ao instalar o Analytic Server em um ambiente Cloudera ativado para Kerberos, deve-se também criar as contas necessárias no Kerberos e ativar a personificação do Kerberos. Para obter mais informações, consulte “Configurando o Kerberos” na página 52.

**Aviso:** Após a instalação bem-sucedida do Analytic Server, não clique em **Criar metastore do Analytic Server** na lista Ações da página de serviços do Analytic Server no Cloudera Manager. A criação de um metastore sobrescreve o repositório de metadados existente.

**Instalação offline**

As etapas da instalação off-line são as mesmas que as etapas on-line, exceto que deve-se fazer download manualmente dos arquivos do pacotes e dos metadados apropriados para seu sistema operacional específico.

O RedHat Linux requer os arquivos a seguir:

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

O SuSE Linux requer os arquivos a seguir:

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

O Ubuntu Linux 16.04 requer os arquivos a seguir:

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

Ubuntu Linux 18 requer os arquivos a seguir:

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

1. Faça download e execute o instalador de autoextração `*.bin` do Cloudera no nó do cluster principal do Cloudera Manager. Siga os prompts de instalação aceitando o contrato de licença e mantendo o diretório de instalação CSD padrão.

   **Nota:** Deve-se especificar um diretório CSD diferente se diferir do local padrão.
2. Copie o pacote e os arquivos de metadados necessários para o caminho do repositório `repo` do Cloudera local no nó do cluster principal do Cloudera Manager. O caminho padrão é `/opt/cloudera/parcel-repo` (o caminho é configurável na interface com o usuário do Cloudera Manager).
3. Use o comando a seguir para reiniciar o Cloudera Manager:

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

   O pacote **AnalyticServer** é mostrado como **transferido por download** após o Cloudera Manager atualizar o pacote. É possível clicar em **Verificar novos pacotes** para forçar uma atualização.
4. Clique em **Distribuir** > **Ativar**.

   O pacote **AnalyticServer** é mostrado como distribuído e ativado.
5. No Cloudera Manager, inclua o Analytic Server como um serviço. Consulte as etapas 7 e 8 da seção "Instalação on-line" para obter mais informações.

## Configurando o Cloudera

Após a instalação, deve-se criar as contas necessárias no sistema operacional do cluster.

1. Crie contas de usuário do sistema operacional para todos os usuários aos quais você planeja dar acesso ao Analytic Server em cada nó do Analytic Server e Hadoop (esses usuários também são configurados como registros de usuário LDAP). O grupo de usuários deve ser configurado como hadoop.
   - Certifique-se de que o UID para esses usuários corresponda em todas as máquinas. É possível testar isso usando o comando **`kinit`** para efetuar login em cada uma das contas.
   - Certifique-se de que o ID do usuário esteja de acordo com a configuração do YARN **ID de usuário mínimo para o envio da tarefa**. Este é o parâmetro **`min.user.id`** em `container-executor.cfg`. Por exemplo, se **`min.user.id`** for 1000, então cada conta do usuário criada deverá ter um UID maior ou igual a 1000.

2. Crie uma pasta inicial do usuário no HDFS para o usuário administrador do Analytic Server. A permissão da pasta deve ser configurada como 755, o proprietário deve ser configurado como admin e o grupo de usuários deve ser configurado como hdfs. Veja o exemplo em **negrito** a seguir:

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

3. Crie pastas iniciais do usuário no HDFS para todos os usuários padrão do Analytic Server (por exemplo, user1). O proprietário da pasta é o usuário real e o grupo de usuários deve ser configurado como hdfs.

Opcionalmente, após a instalação, será possível configurar e administrar o Analytic Server por meio do Cloudera Manager.

**Nota:** As convenções a seguir são utilizadas para os caminhos de arquivo do Analytic Server.

- {AS_ROOT} refere-se ao local em que Analytic Server é implementado; por exemplo, /opt/IBM/SPSS/AnalyticServer/{version}.
- {AS_SERVER_ROOT} refere-se ao local dos arquivos de configuração, de log e de servidor ; por exemplo, /opt/cloudera/parcels/AnalyticServer/ae_wlpserver/usr/servers/aeserver.
- {AS_HOME} refere-se ao local no HDFS usado pelo Analytic Server como uma pasta raiz; por exemplo, /user/as_user/analytic-root.

## Segurança

Por padrão, o valor **tenant_id** no arquivo options.cfg do IBM SPSS Modeler é **ibm**. É possível visualizar Locatários no console do Analytic Server. Consulte o *IBM SPSS Analytic Server Guia do Administrador* para obter detalhes sobre o gerenciamento de locatário.

### Configurar um registro LDAP

O LDAP é configurado durante a instalação do Analytic Server . É possível alterar para outro método de servidor LDAP após a instalação do Analytic Server.

**Nota:** O suporte para LDAP no Analytic Server é controlado pelo WebSphere Liberty. Para obter mais informações, consulte Configurando registros de usuários de usuários LDAP no Liberty.

#### Configure uma conexão secure socket layer (SSL) do Analytic Server com o LDAP

1. Efetue login em cada uma das máquinas do Analytic Server como o usuário do Analytic Server e crie um diretório comum para certificados SSL.

   **Nota:** No Cloudera, o usuário do Analytic Server é sempre as_user e ele não pode ser alterado.

2. Copie os arquivos keystore e truststore para algum diretório comum em todas as máquinas do Analytic Server. Inclua também o certificado CA do cliente LDAP no truststore. A seguir estão algumas instruções de amostra.

```
mkdir -p /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
```

   **Nota:** JAVA_HOME é o mesmo JRE usado para inicialização do Analytic Server.

3. As senhas podem ser codificadas para ofuscar seus valores com a ferramenta securityUtility, que está em {AS_ROOT}/ae_wlpserver/bin. A seguir está um exemplo.

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

4. Efetue login no Cloudera Manager e atualize a definição de configuração **ssl_cfg** do Analytic Server com as definições de configuração SSL corretas. A seguir está um exemplo.

```
<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=" {6}Nis= " />
```

   **Nota:** Use o caminho absoluto para os arquivos key e truststore.

5. Atualize a definição de configuração **security_cfg** do Analytic Server com as definições de configuração LDAP corretas. Por exemplo, no elemento **ldapRegistry**, configure o atributo **sslEnabled** como `true` e o atributo **sslRef** como `defaultSSLConfig`.

**Configurando o Kerberos**

Analytic Server suporta o Kerberos no Cloudera. As seções a seguir fornecem as definições de configuração para assegurar que o Kerberos esteja configurado adequadamente de maneira compatível com o Analytic Server.

**Nota:** O Analytic Server inspeciona a configuração de HDFS para valores relacionados do Kerberos a serem usados para autenticação.

### Configurações do Analytic Server e do Kerberos

Use as configurações a seguir ao instalar o Analytic Server em um ambiente Cloudera ativado para Kerberos.

- Selecione `Kerberos` como a configuração de **Segurança do Analytic Server** se desejar ativar a autenticação do Kerberos ao efetuar login no console do Analytic Server. Quando **Kerberos** for selecionado como a configuração de **Segurança do Analytic Server**, o console do Analytic Server será padronizado para o modo de login do Kerberos.
- Selecione `Kerberos` como a configuração de **Método de conexão de origem de dados do banco de dados do Analytic Server** quando desejar se conectar aos bancos de dados ativados para Kerberos. Quando **Kerberos** for selecionado como a configuração de **Método de conexão de origem de dados do banco de dados do Analytic Server**, o console do Analytic Server usará o modo do Kerberos ao conectar-se a um banco de dados.
- As configurações de **Nome da região do Kerberos** e de **Host do KDC** são necessárias. Os valores **Nome da região do Kerberos** (**as.kdc.realms**) e **Host do KDC** (**kdcserver**) estão localizados no arquivo `krb5.conf` no servidor Key Distribution Center (KDC) do Kerberos.

  Vários nomes de região são suportados quando separados por caracteres de vírgula. Os nomes de região do Kerberos especificados correspondem a, e estão associados a, nomes de usuário. Por exemplo, os nomes de usuário `UserOne@us.ibm.com` e `UserTwo@eu.ibm.com` corresponderão às regiões `us.ibm.com,eu.ibm.com`.

  A confiança entre regiões do Kerberos deve ser configurada quando mais de uma região for especificada como um **Nome de região do Kerberos**. O nome do usuário inserido durante o prompt de login do console do Analytic Server é inserido sem o sufixo de nome de região. Como resultado, quando várias regiões são especificadas, uma lista suspensa de **Regiões** é apresentada aos usuários, permitindo-lhes selecionar a região.

  **Nota:** Quando apenas uma região for especificada, os usuários não verão uma lista suspensa de **Regiões** ao se inscreverem no Analytic Server.

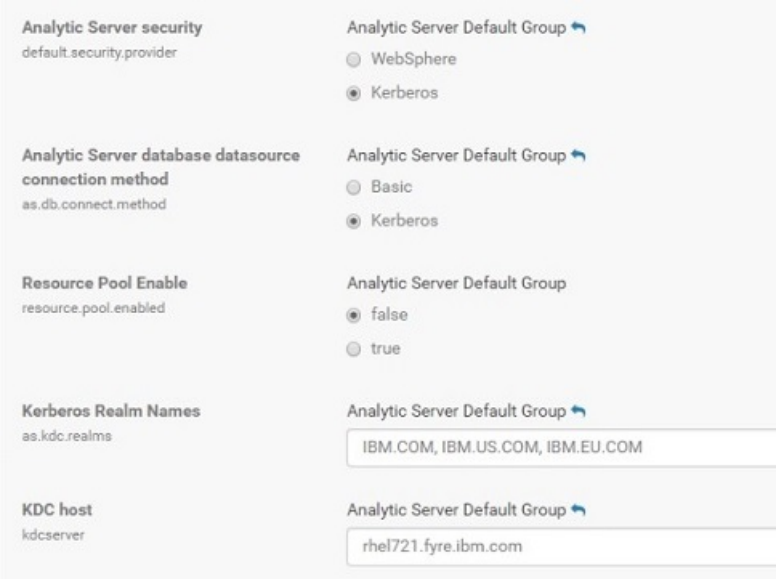



*Figura 7. Configurações de Kerberos de Exemplo*

**Notes:**

- As configurações de **Segurança do Analytic Server** e de **Método de conexão de origem de dados do banco de dados do Analytic Server** são aplicáveis à autenticação do cliente do IBM SPSS Modeler e do console do Analytic Server.
- Quando o **Método de conexão de origem de dados do banco de dados do Analytic Server** é configurado para `Kerberos`, deve-se assegurar que os bancos de dados de destino também sejam ativados para o Kerberos.
- As configurações de **Segurança do Analytic Server** e de **Método de conexão de origem de dados do banco de dados do Analytic Server** não configuram a autenticação do Kerberos no cluster do Hadoop. Para obter mais informações, consulte a seção "Ativando a personificação do Kerberos".
- Se você deseja que a autenticação do Kerberos seja ativada no login, deve-se implementar o cliente do IBM SPSS Modeler como um cliente Kerberos válido. Isso é realizado usando o comando **`addprinc`** no servidor do centro de distribuição de chaves (KDC) do Kerberos. Para obter mais informações, consulte a documentação do IBM SPSS Modeler.

### Criando as Contas Necessárias no Kerberos

1. Crie contas no repositório do usuário do Kerberos para todos os usuários aos quais você planeja conceder acesso ao Analytic Server.
2. Crie as mesmas contas (da etapa anterior) no servidor LDAP.
3. Crie uma conta do usuário do S.O. para cada um dos usuários criados em uma etapa anterior em cada um dos nós do Analytic Server e do nó do Hadoop. O grupo de usuários deve ser configurado como `hadoop`.
   - Certifique-se de que o UID para esses usuários corresponda em todas as máquinas. É possível testar isso usando o comando `kinit` para efetuar login em cada uma das contas.
   - Certifique-se de que o ID do usuário esteja de acordo com a configuração do YARN de **ID de usuário mínimo para o envio da tarefa**. Essa é a configuração de **`min.user.id`** em `container-executor.cfg`. Por exemplo, se **`min.user.id`** for 1000, então cada conta do usuário criada deverá ter um UID maior ou igual a 1000.
4. Crie uma pasta inicial do usuário no HDFS para o usuário administrador do Analytic Server. A permissão da pasta deve ser configurada como 755, o proprietário deve ser configurado como `admin` e o grupo de usuários deve ser configurado como `hdfs`. Veja o exemplo em **negrito** a seguir:

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

```
Localizados 9 itens

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

5. Crie pastas iniciais do usuário no HDFS para todos os usuários padrão do Analytic Server (por exemplo, `user1`). O proprietário da pasta é o usuário real e o grupo de usuários deve ser configurado como `hdfs`.
6. Se planejar usar origens de dados do HCatalog e o Analytic Server estiver instalado em uma máquina diferente do Hive metastore, será necessário personificar o cliente Hive no HDFS.
   a. Navegue para a guia Configuração do serviço do HDFS no Cloudera Manager.

      **Nota:** As configurações a seguir podem não aparecer na guia **Configuração** se elas ainda não tiverem sido configuradas. Neste caso, execute uma procura para encontrá-los.
   b. Edite a configuração de **hadoop.proxyuser.hive.groups** para ter o valor * ou um grupo que contenha todos os usuários com permissão para efetuar login no Analytic Server.
   c. Edite a configuração de **hadoop.proxyuser.hive.hosts** para ter o valor * ou a lista de hosts nos quais o metastore Hive e cada instância do Analytic Server estão instalados como serviços.
   d. Reinicie o serviço HDFS.

Após a execução dessas etapas e a instalação do Analytic Server, o Analytic Server configura o Kerberos de forma silenciosa e automática.

#### Ativando a personificação do Kerberos

A personificação permite que um encadeamento seja executado em um contexto de segurança que difere do contexto de segurança do processo que possui o encadeamento. Por exemplo, a personificação fornece um meio de as tarefas do Hadoop serem executadas como usuário que não o usuário padrão do Analytic Server (`as_user`). Para ativar a personificação do Kerberos:

1. Abra o Cloudera Manager e inclua ou atualize as propriedades a seguir na área **Fragmento de configuração avançada para todo o cluster (válvula de segurança) para core-site.xml** (localizada na guia **HDFS (para todo o serviço)** > **Configuração**).
   - **Nome:** `hadoop.proxyuser.as_user.hosts`
   - **Valor:** *
   - **Nome:** `hadoop.proxyuser.as_user.groups`
   - **Valor:** *

   **Nota:** As configurações de **core-site.xml** se aplicam à configuração do Hadoop (não do Analytic Server).
2. Execute o comando a seguir de um shell de comando no nó do Analytic Server:

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

#### Configurando o HAProxy para Conexão Única (SSO) usando Kerberos

1. Configure e inicie o HAProxy seguindo o guia da documentação do HAProxy: http://www.haproxy.org/#docs
2. Crie o princípio (HTTP/<proxyHostname>@<realm>) e o arquivo keytab do Kerberos para o host do HAProxy, no qual <proxyHostname> é o nome completo do host do HAProxy e <realm> é a região do Kerberos.
3. Copie o arquivo keytab para cada um dos hosts do Analytic Server como `/etc/security/keytabs/spnego_proxy.service.keytab`

4. Atualize as permissões para esse arquivo em cada um dos hosts do Analytic Server. A seguir está um exemplo.

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

5. Abra o Cloudera Manager e inclua ou atualize as propriedades a seguir na área do Analytic Server **Snippet de configuração avançada do Servidor analítico (válvula de segurança) para analyticserver-conf/config.properties**.

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

6. Salve a configuração e reinicie todos os serviços do Analytic Server do Cloudera Manager.
7. Instrua os usuários a configurar o seu navegador para usar o Kerberos.

Agora, os usuários podem efetuar login no Analytic Server usando a opção **Conexão única no login** na tela de login do IBM SPSS Analytic Server.

#### Desativando o Kerberos

1. Desative o Kerberos no console do Cloudera Manager.
2. Pare o serviço do Analytic Server.
3. Modifique as configurações a seguir no **fragmento de configuração avançada do servidor analítico (válvula de segurança) para a área analyticserver-conf/config.properties**:

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

   **Método de conexão da origem de dados do banco de dados do Analytic Server (as.db.connect.method)** > **Básico**
4. Clique em **Salvar mudanças** e reinicie o serviço do Analytic Server.

### Ativando conexões Secure Socket Layer (SSL) com o console do Analytic Server

Por padrão, o Analytic Server gera certificados autoassinados para ativar o Secure Socket Layer (SSL), para que seja possível acessar o console do Analytic Server por meio da porta segura, aceitando certificados autoassinados. Para tornar o acesso HTTPS mais seguro, é necessário instalar certificados de fornecedores terceiros.

#### Instalando certificados de fornecedores terceiros

1. Copie os certificados de keystore e de truststore de fornecedores terceiros para o mesmo diretório em todos os nós do Analytic Server; por exemplo, `/home/as_user/security`.

   **Nota:** O Usuário do Servidor analítico deve ter acesso de leitura para este diretório.
2. No Cloudera Manager, navegue para a guia Configuração do serviço Analytic Server.
3. Edite o parâmetro **`ssl_cfg`**.

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

   Substitua

- <KEYSTORE-LOCATION> pelo local absoluto do keystore; por exemplo: `/home/as_user/security/mykey.jks`
- <TRUSTSTORE-LOCATION> pelo local absoluto do armazenamento confiável; por exemplo: `/home/as_user/security/mytrust.jks`
- <TYPE> com o tipo do certificado, por exemplo: JKS, PKCS12, etc.
- <PASSWORD> com a senha criptografada no formato de criptografia Base64. Para codificação, é possível usar o securityUtility; por exemplo: `{AS_ROOT}/ae_wlpserver/bin/securityUtility encode <password>`

Se você deseja gerar um certificado autoassinado, é possível usar o securityUtility; por exemplo: `{AS_ROOT}/ae_wlpserver/bin/securityUtility createSSLCertificate --server=aeserver --password=mypassword --validity=365 --subject=CN=myfqdnserver,O=myorg,C=mycountry`. Para obter informações adicionais sobre o securityUtility e outras configurações de SSL, consulte a documentação do WebSphere Liberty Profile.

**Notes:**

- Deve-se fornecer um nome de domínio de host apropriado para o valor `CN`.
- Substitua **mypassword**, **myfqdnserver**, **myorg** e **mycounty** por suas credenciais específicas. Observe que **myfqdnserver** é o nome completo do domínio para o nó do Analytic Server.
- **aeserver** é o nome do servidor Liberty (o valor deve ser **aeserver**).

Para obter mais informações sobre o **securityUtility** e outras configurações SSL, consulte as documentações Perfil Liberty do WebSphere e comando securityUtility.

4. Clique em **Salvar mudanças** e reinicie o serviço do Analytic Server.

#### Gerando certificados autoassinados

É possível usar o `securityUtility` para gerar certificados autoassinados. Por exemplo:

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

**Notes:**

- Deve-se fornecer um nome de domínio de host apropriado para o valor **CN**.
- Copie as informações que estão em `key.jks` para `trust.jks` (os dois arquivos devem ser idênticos).
- Edite o parâmetro `ssl.keystore.config`. Por exemplo:

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

### Comunicando-se com o Apache Hive por meio de SSL

Deve-se atualizar o arquivo `hive.properties` para se comunicar com o Apache Hive por uma conexão SSL. Como alternativa, se o ambiente do Apache Hive estiver ativado para alta disponibilidade, será possível selecionar os parâmetros de alta disponibilidade na página principal Origens de dados do Analytic Server.

#### Atualizando o arquivo hive.properties

1. Abra o arquivo `hive.properties`. O arquivo está localizado em: `/opt/cloudera/parcels/AnalyticServer/ae_wlpserver/usr/servers/aeserver/configuration/database`
2. Localize a linha a seguir:

```
jdbcurl = jdbc:hive2: // {: { / {; user = {; password = {
```

3. Atualize a linha incluindo as informações abaixo em **negrito**:

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

4. Salve o arquivo `hive.properties` .

## Ativando o suporte para Essentials for R

Analytic Server suporta modelos R de pontuação e scripts R de execução.

Para instalar o Essentials for R após uma instalação do Analytic Server com êxito no Cloudera Manager:

1. Forneça o ambiente de servidor para o Essentials for R. Para obter mais informações, consulte a etapa 1 em “Ativando o suporte para Essentials for R” na página 25.
2. Faça download do archive autoextrator (BIN) para o IBM SPSS Modeler Essentials for R RPM. O Essentials for R está disponível para download (https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp). Escolha o arquivo específico para sua pilha, versão de pilha e arquitetura de hardware.
3. Execute o archive autoextrator como um usuário root ou sudo no host do servidor do Cloudera Manager. Os pacotes a seguir devem ser instalados ou estar disponíveis a partir dos repositórios configurados:
   - Red Hat Linux: gcc-gfortran, zip, gcc-c++
   - SUSE Linux: gcc-fortran, zip, gcc-c++
   - Ubuntu Linux: gcc-fortran, zip, gcc-c++
4. O instalador autoextrator executa as tarefas a seguir:
   a. Exibe as licenças necessárias e solicita que o instalador aceite-as.
   b. Solicita ao instalador para inserir o local de origem R ou continuar com o local padrão. A versão R padrão que está instalada é a 3.5.1. Para instalar uma versão diferente:
      - Instalação on-line: forneça a URL para o archive da versão R requerida. Por exemplo, https://cran.r-project.org/src/base/R-3/R-3.4.4.tar.gz para R 3.4.4.
      - Instalação off-line: faça o download e, em seguida, copie o archive da versão R requerida no host do servidor do Cloudera Manager. Não renomeie o archive (por padrão, ele será chamado `R-x.x.x.tar.gz`). Forneça a URL para o archive R copiado, conforme a seguir: `file://<R_archive_directory>/R-x.x.x.tar.gz`. Se o archive `R-3.4.4.tar.gz` foi transferido por download e depois copiado para `/root`, a URL é `file:///root/R-3.4.4.tar.gz`.

      **Nota:** Outras versões R podem ser localizadas em https://cran.r-project.org/src/base/.
   c. Instala os pacotes que requerem o R.
   d. Faz o download e instala o R, mais o plug-in Essentials for R.
   e. Cria o pacote e o arquivo `parcel.sha` e os copia em `/opt/cloudera/parcel-repo`. Insira o local correto se o local foi alterado.
5. Após a conclusão da instalação, distribua e ative o pacote **Essentials for R** no Cloudera Manager (clique em **Verificar novos pacotes** para atualizar a lista de pacotes).
6. Se o serviço do Analytic Server já estiver instalado:
   a. Pare o serviço.
   b. Atualize os binários do Analytic Server.
   c. Inicie o serviço para concluir a instalação do Essentials for R.
7. Se o serviço do Analytic Server não estiver instalado, então continue com a sua instalação.

**Nota:** Todos os hosts do Analytic Server devem ter os pacotes de archive apropriados (`zip` e `unzip`) instalados.

## Ativando origens de base de dados relacional

O Analytic Server pode usar origens do banco de dados relacional se você fornecer os drivers JDBC em um diretório compartilhado no metastore do Analytic Server e em cada host do Analytic Server. Por padrão, esse diretório é `/usr/share/jdbc`.

Para alterar o diretório compartilhado, siga essas etapas.

1. No Cloudera Manager, navegue para a guia Configuração do serviço Analytic Server.
2. Especifique o caminho do diretório compartilhado de drivers JDBC em **`jdbc.drivers.location`**.
3. Clique em **Salvar Mudanças**.
4. Selecione **Parar** na lista suspensa **Ações** para parar o serviço do Analytic Server.
5. Selecione **Atualizar os binários do Servidor analítico** na lista suspensa **Ações**.
6. Selecione **Iniciar** na lista suspensa **Ações** para iniciar o serviço do Analytic Server.

*Tabela 11. Bancos de Dados Suportados*

| **Database** | **Versões suportadas** | **Jars de driver JDBC** | **Vendor** |
|---|---|---|---|
| Amazon Redshift | 8.0.2 ou posterior | `RedshiftJDBC41-1.1.6.1006.jar` ou mais recente | Amazônio |
| Apache Impala | JDBC 4 com 2.5.5 ou posterior | `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 | Serviço Bluemix | `db2jcc.jar` | IBM |
| Db2 for Linux, UNIX e 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 | 5.7, 5.6 | mysql-connector-java-commercial-5.1.25-bin.jar | MySQL |
| Netezza | 7, 6.x | `nzjdbc.jar` | IBM |

| *Tabela 11. Bancos de Dados Suportados (continuação)* | | | |
|---|---|---|---|
| **Database** | **Versões suportadas** | **Jars de driver JDBC** | **Vendor** |
| Oracle | 19c, 12c, 11g R2 (11.2) | 19c:<br>`ojdbc8.jar`, `orai18n.jar`<br>12c e 11g R2 (11.2):<br>`ojdbc6.jar`, `orai18n.jar` | Oracle |
| Servidor SQL | 2017, 2014, 2012, 2008 R2 | `sqljdbc4.jar` | Microsoft |
| Teradata | 15.1, 15 | `tdgssconfig.jar`, `terajdbc4.jar` | Teradata |

**Notes**

- Se você tiver criado uma origem de dados Redshift antes da instalação do Analytic Server, será necessário executar as etapas a seguir para usar a origem de dados Redshift.
  1. No console do Analytic Server, abra a origem de dados Redshift.
  2. Selecione a origem de dados do banco de dados `Redshift`.
  3. Insira o endereço do servidor Redshift.
  4. Insira o nome do banco de dados e o nome do usuário. A senha deve ser preenchida automaticamente.
  5. Selecione a tabela de banco de dados.

## Ativação das Origem de Dados HCatalog

O Analytic Server fornece suporte para várias origens de dados por meio do Hive/HCatalog. Algumas origens requerem etapas de configuração manual.

1. Colete os arquivos JAR necessários para ativar a origem de dados. Consulte as seções abaixo para obter detalhes.
2. Inclua esses arquivos JAR ao diretório `{HIVE_HOME}/auxlib` e ao diretório `/usr/share/hive` no metastore do Analytic Server e em cada nó do Analytic Server.
3. Reinicie o serviço Hive Metastore.
4. Reinicie toda instância do serviço do Analytic Server.

**Nota:**

Ao acessar dados do HBase por meio de uma origem de dados do HCatalog do Analytic Server, o usuário de acesso deve ter permissão de leitura para as tabelas do HBase.

- Em ambientes não kerberos, o Analytic Server acessa o HBase usando `as_user` (`as_user` deve ter permissão de leitura para o HBase).
- Em ambientes do Kerberos, tanto o `as_user` quanto o usuário de login devem ter permissão de leitura para tabelas do HBase.

### Bancos de dados NoSQL

O Analytic Server suporta qualquer banco de dados NoSQL para o qual um manipulador de armazenamento Hive está disponível no fornecedor.

Não são necessárias etapas adicionais para ativar o suporte para o Apache HBase e o Apache Accumulo.

Para outros bancos de dados NoSQL, entre em contato com o fornecedor de base de dados e obtenha o manipulador de armazenamento e os jars relacionados.

#### Tabelas Hive baseadas em arquivo

O Analytic Server suporta tabelas Hive baseadas em arquivo para as quais um Hive SerDe integrado ou customizado (serializador-desserializador) está disponível.

O Hive XML SerDe para processar arquivos XML está localizado no Maven Central Repository em http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde.

#### Tarefas MapReduce v2

Use a configuração de **preferred.mapreduce** na área **Fragmento de configuração avançada do Analytic Server (válvula de segurança) para analyticserver-conf/config.properties** para controlar como as tarefas MapReduce são manipuladas:

*Tabela 12. Fragmento de configuração avançada do Analytic Server (válvula de segurança) para analyticserver-conf/config.properties*

| Propriedade | descrição |
|---|---|
| `preferred.mapreduce` | Controla o método no qual as tarefas MapReduce são executadas. Valores válidos incluem:<br>• `spark`<br>• `m3r`<br>• `hadoop`<br>Por exemplo: `preferred.mapreduce=spark` |

#### Apache Spark

Se você deseja usar o Spark (versão 2.x ou posterior), deve-se selecionar a `spark.version` durante a instalação do Analytic Server.

1. Abra o Cloudera Manager e selecione a `spark.version` apropriada (por exemplo, `Nenhum` ou `2.x`) na área **Versão do Analytic Server Spark**.
2. Salve a configuração.

## Configurando o Apache Impala

O Apache Impala é compatível com o Cloudera em uma origem de dados do banco de dados do Analytic Server ou uma origem de dados HCatalog (independentemente se o Impala for ativado para SSL).

#### Criando uma origem de dados do banco de dados para dados do Apache Impala

1. Na página principal de **Origens de dados** do Analytic Server, clique em **Nova** para criar uma nova origem de dados. O diálogo **Nova origem de dados** é exibido.
2. Insira um nome apropriado no campo **Nova origem de dados**, selecione `Banco de dados` como o valor **Tipo de conteúdo** e, em seguida, clique em **Ok**.
3. Abra a seção **Seleções de banco de dados** e insira as informações a seguir.

   **Banco de dados:**
   Selecione **Impala** no menu suspenso.

   **Endereço do Servidor:**
   Insira a URL do servidor que hospeda o daemon Impala. Um nome completo do domínio é necessário quando o Kerberos é ativado para o Analytic Server.

   **Porta do Servidor:**
   Insira o número da porta em que o banco de dados Impala atende.

**Nome do Banco de Dados:**
Insira o nome do banco de dados ao qual deseja se conectar.

**Nome do Usuário:**
Insira um nome do usuário com autoridade para efetuar login no banco de dados Impala.

**Senha:**
Insira a senha do nome do usuário apropriado.

**Nome da Tabela:**
Insira o nome de uma tabela do banco de dados que deseja usar. Clique em **Selecionar** para selecionar manualmente um arquivo.

**Máximo de leituras simultâneas:**
Insira o limite no número de consultas paralelas que podem ser enviadas do Analytic Server para o banco de dados para ler a tabela especificada na origem de dados.

4. Clique em **Salvar** após inserir as informações necessárias.

#### Criando uma origem de dados de HCatalog para dados do Apache Impala

1. Na página principal de **Origens de dados** do Analytic Server, clique em **Nova** para criar uma nova origem de dados. O diálogo **Nova origem de dados** é exibido.
2. Insira um nome apropriado no campo **Nova origem de dados**, selecione `HCatalog` como o valor **Tipo de conteúdo** e, em seguida, clique em **Ok**.
3. Abra a seção **Seleções de banco de dados** e insira as informações a seguir.

   **Banco de dados:**
   Selecione **default** no menu suspenso.

   **Nome da Tabela:**
   Insira o nome de uma tabela do banco de dados que deseja usar.

   **Esquema do HCatalog**
   Selecione a opção **Elemento de HCatalog** e, em seguida, selecione as opções apropriadas **Mapeamentos de campo do HCatalog**.
4. Clique em **Salvar** após inserir as informações necessárias.

#### Conectando-se a dados Apache Impala habilitados para SSL

1. Defina as configurações de SSL a seguir do Impala no console do Cloudera Manager.

   **Ative TLS/SSL para Impala (client_services_ssl_enabled)**
   Selecione a opção **Impala (Service-Wide)**.

   **Arquivo de certificado do servidor Impala TLS/SSL (Formato PEM) (ssl_server_certificate)**
   Insira o local do certificado de formato PEM autoassinado e o nome do arquivo (por exemplo: `/tmp/<user_name>/ssl/ll4200v21.crt`).

   **Arquivo de chave privada do servidor Impala TLS/SSL (Formato PEM) (ssl_private_key)**
   Insira a chave privada, no formato PEM, o local e o nome do arquivo (por exemplo: `/tmp/<user_name>/ssl/ll4200v21.key`).
2. No host do Analytic Server, importe o arquivo `*.crf` (que é usado para ativar o Impala SSL) em um arquivo `*.jks`. O arquivo pode ser um arquivo cacerts (por exemplo, `/etc/pki/java/cacerts`) ou qualquer outro arquivo `*.jks`.
3. No host do Analytic Server, atualize o arquivo de configuração do Impala (`impala.properties`) anexando o valor da chave `jdbcurl` a seguir:

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

   **Nota:** Quando um arquivo `*.jks` (que não cacerts) é usado, é necessário especificar também o seguinte:

   ```
   SSLTrustStore = < your_pks_file>; SSLTrustStorePwd = < password_for_pks_file>;
   ```

4. Reinicie Analytic Server no console do Cloudera Manager.

## Alterando portas usadas pelo Analytic Server

O Analytic Server usa a porta 9080 para HTTP e a porta 9443 para HTTPS, por padrão. Para alterar as configurações de porta, siga essas etapas.

1. No Cloudera Manager, navegue para a guia Configuração do serviço Analytic Server.
2. Especifique as portas HTTP e HTTPS requeridas nos parâmetros **`http.port`** e **`https.port`**, respectivamente.

   **Nota:** Pode ser necessário selecionar a categoria **Portas e endereços** na seção Filtros para poder ver esses parâmetros.
3. Clique em **Salvar Mudanças**.
4. Reinicie o serviço Analytic Server.

## Analytic Server de alta disponibilidade

É possível tornar o Analytic Server altamente disponível incluindo-o como um serviço para vários nós em seu cluster.

1. No Cloudera Manager, navegue para a guia Instâncias do serviço do Analytic Server.
2. Clique em **Incluir instâncias de função** e selecione os hosts nos quais incluir o Analytic Server como um serviço.

**Suporte para vários clusters**
O recurso de vários clusters é um aprimoramento do recurso Alta Disponibilidade do IBM SPSS Analytic Server e fornece isolamento melhorado em ambientes de locatários múltiplos. Por padrão, a instalação do serviço Analytic Server (no Ambari ou no ClouderaManager) resulta na definição de um único cluster de servidores analíticos.

A especificação de cluster define a associação de cluster do Analytic Server. A modificação da especificação de cluster é feita com conteúdo XML (no campo `cluster de analítica` da configuração do Ambari Analytic Server ou editando manualmente o arquivo `configuration/analytics-cluster.xml` do Cloudera Manager). Durante a configuração de vários clusters do Analytic Server, é necessário alimentar solicitações para cada cluster do Analytic Server com seu próprio balanceador de carga.

O uso do recurso de vários clusters garante que o trabalho para um locatário não possa afetar negativamente o trabalho sendo realizado no cluster de outro locatário. Com relação a tarefas altamente disponíveis, o failover de tarefa ocorre apenas dentro do escopo do cluster do Analytic Server no qual o trabalho foi iniciado. O exemplo a seguir fornece uma especificação XML para vários clusters.

**Nota:** Analytic Server pode ser transformado em altamente disponível incluindo-o como um serviço em vários nós em seu cluster.

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

No exemplo anterior, dois balanceadores de carga são necessários. Um balanceador de carga envia solicitações para os membros do `cluster1` (`one.cluster` e `two.cluster`) e o outro envia solicitações para os membros do `cluster2` (`three.cluster` e `four.cluster`).

O exemplo a seguir fornece uma única especificação XML de cluster (a configuração padrão).

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

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

No exemplo anterior, um único balanceador de carga é necessário para tratar casos nos quais há mais de um membro de cluster configurado.

#### Notes

- Apenas clusters singleton suportam o uso de curingas no elemento **memberName** (por exemplo, cardinalidade de cluster = "1"). Os valores válidos para o elemento de cardinalidade são 1 e 1+.
- O **memberName** deve ser especificado da mesma maneira que o nome do host ao qual a função Analytic Server é designada.
- Todos os servidores em todos os clusters devem ser reiniciados após as mudanças na configuração do cluster serem aplicadas.
- No Cloudera Manager, deve-se modificar e manter o arquivo `analytics-cluster.xml` em todos os nós do Analytic Server. Todos os nós devem ser mantidos para garantir que eles tenham o mesmo conteúdo.

## Upgrade do Python - CDH

Esta seção descreve o processo de upgrade manual do Python 2.x para o Python 3.7

1. Instalar o Python 3.7 em cada nó do cluster. Consulte o site do Python para obter mais informações.
2. Instalar o NumPy em cada nó do cluster. Consulte as instruções de instalação do NumPy https://numpy.org/install/ para obter mais informações.
3. Instalar o pandas em cada nó do cluster. Consulte as instruções de instalação do pandas https://pandas.pydata.org/getting_started.html para obter mais informações.
4. No Cloudera Manager, atualize a seção **Analytic Server Advanced Configuration Snippet (Safety Valve) for analyticserver-conf/config.properties** para incluir o caminho executável do Python 3.7. Por exemplo:

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

## Otimizando opções de JVM para dados pequenos

É possível editar propriedades JVM para otimizar seu sistema ao executar tarefas pequenas (M3R).

No Cloudera Manager, consulte o controle **Opções de Jvm (jvm.options)** na guia Configuração no serviço do Analytic Server. Modificar os parâmetros a seguir configura o tamanho do heap para execução de tarefas no servidor que hospeda o Analytic Server; ou seja, não Hadoop. Isso será importante se você estiver executando pequenas tarefas (M3R), e talvez seja necessário experimentar esses valores para otimizar seu sistema.

```
-Xms512M
-Xmx2048M
```

## Configurando uma alocação de recursos dinâmicos separada para cada conjunto de recursos YARN - Cloudera

É possível configurar uma Alocação de recursos dinâmicos separada para cada conjunto de recursos YARN.

#### Mapeamento de modo de usuário e locatário - Cloudera

As tarefas de usuário e locatário podem ser enviadas a diferentes conjuntos de recursos YARN e cada usuário ou locatário mapeia para um conjunto de recursos YARN diferente (para aproveitar as vantagens da Alocação de Recursos Dinâmicos). O modo **user** ou o modo **tenant** pode ser definido para mapeamento para conjuntos de recursos YARN. Antes do Analytic Server 3.2.1 Fix Pack 1, todas as tarefas do Spark eram limitadas a um único conjunto de recursos YARN.

A partir do IBM SPSS Analytic Server 3.2.1 Fix Pack 1, quando o fluxo de um usuário/locatário resulta em tarefas Spark sendo executadas no sistema, um conjunto de recursos YARN separado será executado como o usuário/locatário que enviou o fluxo para Analytic Server. Múltiplos conjuntos de recursos YARN podem ser executados simultaneamente para as diferentes tarefas de usuário/locatário.

Cada conjunto de recursos YARN continua em execução, desde que o usuário esteja conectado ao Analytic Server (e por algum tempo depois que o usuário efetuar logout e não houver mais tarefas de usuário ativas). A quantidade de tempo após o logout pode ser controlada pela variável de configuração: **as.spark.driver.cleanup.delay**.

Um processo **SparkDriver** é criado para cada usuário que envia a tarefa Spark. O processo **SparkDriver** de cada usuário é finalizado depois que o usuário não possuir tarefas ativas por cerca de 2 minutos (o valor padrão) e nenhuma atividade **HTTPSession**.

**Nota:** Todos os processos **SparkDriver** são finalizados quando o Analytic Server é encerrado.

Use as etapas a seguir para incluir o Analytic Server a um cluster existente:

1. No Cloudera Manager, navegue até **Serviço do SPSS Analytic Server** > **Configuração**.
2. Mude o valor de **Ativação do conjunto de recursos: resource.pool.enabled** para `true`.
3. Inclua as propriedades a seguir em **Fragmento de configuração avançada do Analytic Server (válvula de segurança)** > **analyticserver-conf.config.properties**:

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

*Tabela 13. Customizar as configurações do analyticserver-conf.config.properties.*

| **Propriedade** | **descrição** |
|---|---|
| `yarn.queue.mode` | Configura o modo de mapeamento para conjuntos de recursos YARN. Quando `yarn.queue.mode=user`, um um aplicativo YARN separado é executado para cada usuário que enviou uma tarefa/fluxo para o Analytic Server. Múltiplos aplicativos YARN podem ser executados simultaneamente para as tarefas/fluxos de diferentes usuários.<br>Quando `yarn.queue.mode=tenant`, um aplicativo YARN separado é executado para cada locatário que enviou uma tarefa/fluxo para o Analytic Server. Múltiplos aplicativos YARN podem ser executados simultaneamente para as diferentes tarefas/fluxos de locatário. |
| `yarn.queue.mapping` | Mapeia os pares de usuário ou locatário para os conjuntos de recursos YARN que são definidos na configuração do conjunto de recursos dinâmicos do Cloudera Manager. Os pares devem ser separados por vírgulas (por exemplo, `tenant1:test,tenant2:production` para locatários ou `user1:test,user2:production`) para usuários. |
| `yarn.queue.default` | O nome do conjunto de recursos YARN padrão para o qual o aplicativo é enviado. Você pode especificar um nome de conjunto de recursos YARN customizado na Configuração do Conjunto de Recursos Dinâmicos. |
| `as.spark.driver.cleanup.delay` | Um número inteiro que representa o número de minutos após o logout antes de finalizar um aplicativo YARN de um usuário. O valor padrão é **2**. Esta propriedade é Opcional. |
| `as.sparkdriver.max.memory` | Configura a quantidade de memória que é usada por cada processo **SparkDriver**. O valor padrão é **1G**. Essa propriedade é opcional. |

**Referência**

Consulte os sites a seguir para obter mais informações:

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

## Migração

O Analytic Server permite migrar dados e definições de configuração de uma instalação do Analytic Server existente para uma nova instalação.

**Fazer upgrade para uma nova versão do Analytic Server**

Se você tiver uma instalação existente do Analytic Server 3.2.1.1 e tiver comprado uma versão mais recente, será possível migrar suas definições de configuração do 3.2.1.1 para a sua nova instalação.

**Restrição:** O seu 3.2.1.1 e as novas instalações não podem coexistir no mesmo cluster Hadoop. Se você configurar a sua nova instalação para usar o mesmo cluster Hadoop que sua instalação do 3.2.1.1, a instalação do 3.2.1.1 não funcionará mais.

**Etapas de migração, 3.2.1.1 para versão mais recente**

1. Instale a nova instalação do Analytic Server de acordo com as instruções em "Instalação no Cloudera" na página 45.
2. Copie a área de trabalho analítica da sua instalação antiga para a nova.
   a. Se você não tiver certeza sobre a localização da área de trabalho analítica, execute `hadoop -fs ls`. O caminho para a área de trabalho analítica terá o formato `/user/as_user/analytic-root/analytic-workspace`, em que `as_user` é o ID do usuário que proprietário da área de trabalho analítica.
   b. Efetue login no host da nova instalação do Analytic Server como `as_user`. Exclua o diretório `/user/as_user/analytic-root/analytic-workspace`, se ele existir.
   c. Use `hadoop fs -copyToLocal` e `hadoop fs -copyFromLocal` para copiar a área de trabalho analítica do servidor antigo para a pasta `/user/as_user/analytic-root/analytic-workspace` do novo servidor (assegure que o proprietário esteja configurado como `as_user`).
3. Se você usar o Apache Directory Server integrado, faça backup da configuração atual de usuário/grupo com uma ferramenta de cliente LDAP de terceiros. Após o Analytic Server 3.2.2 ser instalado, importe a configuração do usuário/grupo de backup no Apache Directory Server.

   **Nota:** Esta etapa pode ser ignorada se você usa um servidor LDAP externo.
4. No Cloudera Manager, pare o serviço do Analytic Server.
5. Colete as definições de configuração a partir da instalação antiga.
   a. Copie o archive `configcollector.zip` em sua nova instalação para `{AS_ROOT}\tools` em sua antiga instalação.
   b. Extraia a cópia de `configcollector.zip`. Isso cria um novo subdiretório `configcollector` em sua antiga instalação.
   c. Execute a ferramenta coletora de configuração em sua antiga instalação, executando o script **`configcollector`** em `{AS_ROOT}\tools\configcollector`. Copie o arquivo compactado resultante (ZIP) no servidor que hospeda sua nova instalação.

   **Importante:** O script **`configcollector`** fornecido pode não ser compatível com a versão do Analytic Server mais recente. Entre em contato com o representante de suporte técnico IBM se você encontrar problemas com o script **`configcollector`**.
6. Limpe o estado do Zookeeper. No diretório bin do Zookeper (por exemplo, `/opt/cloudera/parcels/CDH-5.4...../lib/zookeeper/bin` no Cloudera), execute o comando a seguir.

   ```
   ./zkCli.sh rmr /AnalyticServer
   ```
7. Execute a ferramenta de migração ao executar o script **`migrationtool`** e transmitir o caminho do arquivo compactado criado pelo coletor de configuração como um argumento. A seguir está um exemplo.

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

8. Execute o comando a seguir de um shell de comando no nó do Analytic Server:

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

9. No Cloudera Manager, inicie o serviço do Analytic Server.

   **Nota:** Se você configurou R para usar com a instalação do Analytic Server existente, será necessário seguir os passos para configurá-lo com a nova instalação do Analytic Server.

## Desinstalando o Analytic Server no Cloudera

O Cloudera manipula automaticamente a maioria das etapas requeridas para desinstalar o serviço e o pacote do Analytic Server.

As etapas a seguir são requeridas para a limpeza do Analytic Server do ambiente do Cloudera:

1. Pare e exclua o Serviço Analytic Server.
2. **Desativar**, **Remover dos Hosts** e **Excluir** os pacotes Analytic Server.
3. Exclua o diretório do usuário do Analytic Server no HDFS. O local padrão é `/user/as_user/analytic-root`.
4. Exclua o banco de dados ou o esquema usado pelo Analytic Server.
5. Limpe quaisquer remanescências do pacote de instalação do Analytic Server. Isso é feito excluindo o seguinte:
   - Pasta `csd`
   - Quaisquer arquivos do 3.2.2 existentes localizados nas pastas `parcels`, `parcel-cache` e `parcel-repo`.

# Capítulo 4. Configurando o IBM SPSS Modeler para Utilização com o IBM SPSS Analytic Server

Para ativar o SPSS Modeler para uso com o Analytic Server, é necessário fazer algumas atualizações para a instalação do SPSS Modeler Server.

1. Configure o SPSS Modeler Server para associá-lo a uma instalação do Analytic Server.
   a. Edite o arquivo `options.cfg` no subdiretório `config` do diretório de instalação de servidor principal e inclua ou edite as linhas a seguir:

```
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**
   Especifique Y se a comunicação segura estiver configurada no Analytic Server; caso contrário, N.

   **as_host**
   O endereço IP/nome do host do servidor que hospeda o Analytic Server.

   **Nota:** Deve-se fornecer um endereço IP/nome de domínio do host quando SSL estiver ativado para o Analytic Server.

   **as_port**
   A porta na qual o Analytic Server está atendendo (por padrão isso é 9080).

   **as_context_root**
   O contexto raiz Analytic Server (por padrão, esse é analyticserver).

   **as_tenant**
   O locatário da instalação do SPSS Modeler Server é um membro de (o locatário padrão é ibm).

   **as_prompt_for_password**
   Especifique N se o SPSS Modeler Server estiver configurado com o mesmo sistema de autenticação para usuários e senhas que o usado no Analytic Server; por exemplo, ao usar a autenticação do Kerberos. Caso contrário, especifique Y.

   Ao executar o SPSS Modeler em modo em lote, inclua `-analytic_server_username {ASusername} -analytic_server_password {ASpassword}` como argumentos para o comando `clemb`.

   **as_kerberos_auth_mode**
   Especifique Y para ativar o Kerberos SSO a partir do SPSS Modeler.

   **as_kerberos_krb5_conf**
   Especifique o caminho para o arquivo de configuração do Kerberos que o Analytic Server deve usar; por exemplo, `\etc\krb5.conf`.

   **as_kerberos_krb5_spn**
   Especifique o Kerberos SPN do Analytic Server; por exemplo, HTTP/ashost.mydomain.com@MYDOMAIN.COM.

   b. Reinicie o serviço do SPSS Modeler Server.

Para se conectar a uma instalação do Analytic Server que tem SSL/TLS ativado, há alguns passos adicionais para a configuração de suas instalações de cliente e SPSS Modeler Server.

a. Navegue para `http{s}://{HOST}:{PORT}/{CONTEXT-ROOT}/admin/{TENANT}` e efetue logon no console do Analytic Server.

b. Faça download do arquivo de certificação do navegador e salve-o em seu sistema de arquivos.

c. Inclua o arquivo de certificação para o JRE das instalações do SPSS Modeler Server e SPSS Modeler Client. O local para atualizar pode ser localizado no subdiretório `/jre/lib/security/cacerts` do caminho da instalação do SPSS Modeler.

   1) Certifique-se de que o arquivo `cacerts` não seja de somente leitura.

   2) Use o programa keytool Modeler fornecido com – isso pode ser localizado no subdiretório `/jre/bin/keytool` do caminho da instalação do SPSS Modeler.

      Execute o comando a seguir

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

      Observe que <as-alias> é um alias para o arquivo `cacerts`. É possível usar qualquer nome que você gostaria, contanto que seja exclusivo para o arquivo `cacerts`.

      Portanto, um exemplo de comando seria semelhante ao seguinte.

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

d. Reinicie seu SPSS Modeler Server e SPSS Modeler Client .

2. [opcional] Instale o IBM SPSS Modeler - Essentials for R se planeja pontuar modelos R em fluxos com origens de dados do Analytic Server. O IBM SPSS Modeler - Essentials for R está disponível para download (https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp).

# Capítulo 5. Configurando o pushback do Hive UDF

Depois que o Hive UDF for registrado no HiveDB, o Analytic Server poderá usar as novas funções do UDF para realizar pushback.

O pushback do Hive UDF está desativado por padrão e deve ser ativado manualmente por meio da configuração **udfmodule** no arquivo ASModules.xml (mude o valor **disabled** para **enabled**). Após ativar a configuração, deve-se reiniciar o Analytic Server e registrar manualmente o UDF para o Hive.

**Notes:**

- Ao usar a origem de dados Hive no HDP 3.x, você poderá encontrar o seguinte erro:

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

  1. Abra o console Ambari e mude a propriedade a seguir na seção **Hive** > **Configurações** > **Avançado** > **Site Hive avançado**.

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

  2. Salve a configuração.
- Ao usar uma origem de dados do Hive em um ambiente não Kerberos, assegure-se de que o nome do usuário inserido na seção **Seleções de banco de dados** seja o mesmo que o login do usuário do Analytic Server.

Os exemplos a seguir demonstram como registrar/cancelar o registro do UDF para o Hive em ambientes HDP e Cloudera.

### Registrar/cancelar o registro do UDF para o HDP

#### Registrar UDF

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

#### Cancelar o registro de UDF

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

### Registrar/cancelar o registro de UDF para Cloudera

#### Registrar UDF

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

#### Cancelar o registro de UDF

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

# Capítulo 6. Usando tags SLM para controlar o licenciamento

As tags SLM são baseadas no rascunho padrão ISO/IEC 19770-4 para Medição de Utilização de Recurso. As tags SLM fornecem um recurso padronizado para um produto com o objetivo de relatar seu uso das métricas de licença (recursos relacionados ao uso de um ativo de software). Após ativar o SLM em um produto, um arquivo XML de tempo de execução é gerado para relatar automaticamente seu uso de licença.

Quando o Analytic Server é iniciado, os arquivos `slmtag` são criados na pasta `<as_installation_path>/logs/slmtag`.

Como existem dois tipos de licença, duas métricas diferentes são registradas periodicamente:

- Para a versão atual do Analytic Server, o licenciamento é baseado no número total de nós de dados no cluster Hadoop (com base no Virtual Server). O número de nós é registrado na seguinte seção do arquivo `slmtag`.

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

- Para versões do Analytic Server anteriores à 3.1, o licenciamento era baseado no tamanho do armazenamento HDFS no cluster Hadoop (com base em RVU). Por exemplo, o tamanho do armazenamento (em tegabytes) é registrado na seguinte seção do arquivo `slmtag`.

```
< Type> RESOURCE_VALUE_UNIT < /Type>
   < SubType> Armazenamento HDFS (Unidade: Tega byte) < /SubType>
   < Value> 0,21 < /Value>
```

A saída da tag SLM é iniciada em um encadeamento e afetada pelas propriedades definidas no arquivo `SlmTagOutput.properties`. O arquivo está localizado na pasta `<as_installation_path>/configuration`.

*Tabela 14. Propriedades da Tag SLM*

| **Propriedades** | **descrição** |
|---|---|
| license.metric.logger.output.enabled | Controla a geração do arquivo de log SLM. O valor padrão é `False`. |
| license.metric.logger.output.dir | O caminho relativo para o diretório que armazena arquivos de tags SLM. O diretório padrão é `<as_installation_path>/logs`. |
| license.metric.logger.output.SLMLogFrequency | O intervalo de tempo (unit:milliseconds) para coletar logs SLM. |
| icense.metric.logger.file.size | O tamanho máximo do arquivo de tags SML, em bytes. |
| license.metric.logger.file.number | O número máximo de arquivos de tags SLM para uma instância de identidade de software. |

# Capítulo 7. Resolução de Problemas

Esta seção descreve alguns problemas comuns de instalação e configuração e como corrigi-los.

### Problemas gerais

**A instalação é concluída com avisos, mas os usuários não conseguem criar origens de dados com o erro "Não é possível concluir a solicitação. Motivo: Permissão negada"**

Configurar o parâmetro **`distrib.fs.root`** para um diretório ao qual o usuário Analytic Server (por padrão, as_user) não possui acesso resultará em erros. Assegure-se de que o usuário do Analytic Server esteja autorizado a ler, gravar e executar o diretório **`distrib.fs.root`**.

**O desempenho do Analytic Server está piorando progressivamente.**

Quando o desempenho do Analytic Server não atender às expectativas, remova todos os arquivos `*.war` do caminho de implementação do serviço Knox: `/<KnoxServicePath>/data/deployments`. Por exemplo: `/usr/hdp/3.1.0.0-78/knox/data/deployments`.

**Desinstalando o Analytic Server ou o Essentials for R no Ambari**

Em alguns casos, o processo de desinstalação é interrompido ao desinstalar o Analytic Server ou o Essentials for R no Ambari. Quando o problema ocorre, deve-se parar manualmente o ID do processo do servidor Ambari.

**Problemas quando o Analytic Server é instalado no POWER System usando OpenJDK**

Quando o Analytic Server estiver em execução em um POWER System que usa OpenJDK, deve-será executar manualmente as etapas de configuração a seguir para garantir que a API do sistema de coordenadas funcione conforme esperado

**Nota:** É possível desconsiderar o requisito de configuração se você não usar a API do sistema de coordenadas.

1. No console do Ambari, navegue para **Serviço do Analytic Server** > **Guia de Configurações** > **analytics-jvm-options avançado** e inclua a seguinte linha na área de conteúdo:

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

2. No console do Ambari, navegue até a seção **analytics.cfg customizado** e inclua as três configurações a seguir:

   **spark.executor.extraJavaOptions**
   Configure o valor para: `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

   **spark.driver.extraJavaOptions**
   Configure o valor para: `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

   **mapred.child.java.opts**
   Configure o valor para: `-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

**Erro ao instalar o Analytic Server no SuSE Linux 12**

Você poderá encontrar o erro a seguir ao instalar o Analytic Server no SuSE Linux 12:

```
A verificação de assinatura falhou [4-A chave pública de assinaturas não está disponível]
```

O problema poderá ser resolvido ao executar as tarefas a seguir antes de instalar o Analytic Server no SuSE Linux 12:

1. Faça download de uma chave pública para seu host da URL a seguir:

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

2. Importe a chave pública executando o comando a seguir em seu host:

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

### Problemas com distribuições específicas do Hadoop

**A ação de atualização para o serviço do Analytic Server fica desativada no Hortonworks 2.3-2.6**

Para atualizar manualmente as bibliotecas do Analytic Server no Hortonworks 2.3-2.6, use as etapas a seguir.

1. Efetue logon no host que executa o Analytic Metastore como o usuário do Analytic Server (por padrão, as_user).

   **Nota:** É possível localizar esse nome do host no console do Ambari.

2. Execute o script **refresh** no diretório {AS_ROOT}/bin; por exemplo:

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

3. Reinicie o serviço do Analytic Server no console do Ambari.

**Os pacotes transferidos por download de um site externo falham a verificação de hash no Cloudera Manager**

O erro de verificação de hash é exibido na lista de pacotes. O problema pode ser resolvido permitindo a conclusão do processo de download e, em seguida, reiniciando o Cloudera por meio do serviço `cloudera-scm-server`. O erro não ocorre após a reinicialização do serviço.

**Propriedades do supergrupo HDFS**

O Analytic Server irá registrar uma exceção durante a inicialização se o `as_user` não for um membro das propriedades do grupo do HDFS a seguir: **dfs.permissions.supergroup/ dfs.permissions.superusergroup**. Por exemplo:

```
[ 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.
Privilégio de superusuário é necessário
em org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkSuperuserPrivilege (FSPermissionChecker.java: 93)
em org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkSuperuserPrivilege (FSNamesystem.java: 6606)
em org.apache.hadoop.hdfs.server.namenode.FSNamesystem.datanodeReport (FSNamesystem.java: 5595)
em org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getDatanodeReport (NameNodeRpcServer.java: 928)
em org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.getDatanodeReport
(AuthorizationProviderProxyClientProtocol.java: 390)
em org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getDatanodeReport
(ClientNamenodeProtocolServerSideTranslatorPB.java: 694)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
em org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call (ProtobufRpcEngine.java: 617)
em org.apache.hadoop.ipc.RPC$Server.call (RPC.java: 1073)
em org.apache.hadoop.ipc.Server$Handler$1.run (Server.java: 2141)
em org.apache.hadoop.ipc.Server$Handler$1.run (Server.java: 2137)
em java.security.AccessController.doPrivileged (Native Method)
em javax.security.auth.Subject.doAs (Subject.java: 415)
em org.apache.hadoop.security.UserGroupInformation.doAs (UserGroupInformation.java: 1912)
em org.apache.hadoop.ipc.Server$Handler.run (Server.java: 2135)
```

Deve-se incluir manualmente `as_user` no grupo de S.O. que está definido nas propriedades de configuração de `hdfs-site`: **dfs.permissions.supergroup/ dfs.permissions.superusergroup**.

- Para o Cloudera, o valor da propriedade padrão é **supergroup** e deve ser mudado para um grupo de S.O. que realmente existe. Para obter informações sobre a configuração de supergroup no Cloudera, consulte a documentação do Cloudera.
- Para o Ambari, o valor da propriedade padrão é **hdfs**. Por padrão, durante uma instalação do Ambari, o Analytic Server inclui `as_user` nos grupos HDFS e Hadoop.

No Linux, use o comando **usermod** para incluir `as_user` no **superusergroup** do HDFS (se ele ainda não existir).

Para obter informações gerais sobre as permissões do HDFS, consulte o Guia de permissões do HDFS.

**Tarefas MapReduce falham no HDP 3.0**

É possível encontrar o erro a seguir com as tarefas MapReduce no HDP 3.0:

```
Não é possível concluir a solicitação. Reason: java.lang.IllegalStateException: Job in state DEFINE instead of RUNNING
(as_trace.log)
```

O estado do erro pode ser resolvido:

1. Incluindo a seguinte configuração no arquivo `Custom analytics.cfg`:

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

2. Reinicie o Analytic Server.

Após a reinicialização do Analytic Server, as tarefas MapReduce serão executadas normalmente.

**A gravação de valores de data e de registro de data e hora em tabelas Hive falha devido a um problema do Cloudera**

Quando o Analytic Server tenta gravar valores de data ou de registro de data e hora em tabelas Hive em um ambiente do Cloudera, o processo falha devido a um problema do Cloudera conhecido (https://issues.apache.org/jira/browse/HIVE-11024.

**Nota:** O problema do valor de data não afeta o Hive 1.3.0 ou 2.0.0 (https://issues.apache.org/jira/browse/HIVE-11024); o problema do valor de registro de data e hora não afeta o Hive 2.0.0 (https://issues.apache.org/jira/browse/HIVE-11748?jql=project%20%3D%20HIVE%20AND%20text%20~%20%22jdbc%20timestamp%22). deve-se assegurar que uma versão compatível do Hive (1.3.0 ou 2.0.0) esteja presente em seu ambiente do Cloudera.

## Problemas com o repositório de metadados

**A operação CREATE USER falha ao executar o script add_mysql_user**

Antes de executar o script **`add_mysql_user`**, será necessário primeiro remover manualmente o usuário que você está tentando incluir do banco de dados mysql. É possível remover os usuários por meio da UI do ambiente de trabalho do MySQL ou dos comandos do MySQL. Por exemplo:

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

Nos comandos acima, substitua o $AEDB_USERNAME_VALUE pelo nome de usuário que você deseja remover e substitua $METASTORE_HOST com o nome do host em que o banco de dados está instalado.

## Problemas com o Apache Spark

**Problemas com fluxos que são executados em um processo do Spark**

Os fluxos do SPSS Modeler não são concluídos quando forçados a executar em um processo do Spark. Os fluxos do SPSS Modeler que falham são construídos com um nó de origem do Analytic Server (arquivo HDFS) que é vinculado a um nó `Sort` e, em seguida, configurado para exportação para outra origem de dados do Analytic Server. Após o fluxo ser executado, a interface com o usuário do Gerenciador de Recursos indica que o novo aplicativo está em execução, mas o fluxo nunca é concluído e permanece em um estado `Running`. Nos logs do Analytic Server, nos logs do YARN e nos logs do Spark, não há nenhuma mensagem que indique por que o fluxo não é concluído.

O problema pode ser resolvido incluindo a configuração de `spark.executor.memory` no arquivo `analytics.cfg` customizado na configuração do Analytic Server. Configurar o valor de memória para 4 GB permite que os fluxos do SPSS Modeler anteriormente com falha sejam concluídos em menos de dois minutos (em um ambiente em cluster de nó único).

**O erro "Exception during HdfsAuthcom.spss.utilities.i18n.LocException:Execution falhou. Motivo: com.spss.ae.filesystem.exception.FileSystemException: Não foi possível inicializar o acesso do sistema de arquivos." ocorre ao executar casos SparkML.**

O erro ocorre quando o Spark não consegue localizar o diretório de log de linhagem. Uma solução alternativa para o problema é redirecionar o `spark.lineage.log.dir` para / `ae_wlpserver/usr/servers/aeserver/logs/spark`.

### Clusters de alta disponibilidade

**O Analytic Server não pode ser incluído em mais hosts devido a mudanças nas dependências**

Execute o script `update_clientdeps` usando as instruções "Atualizando as dependências do cliente" na página 32.

**"O Analytic Cluster Service perdeu inesperadamente o contato com o Zookeeper, essa JVM está sendo finalizada para manter a integridade do cluster".**

Uma coisa que pode causar isso é se a quantidade dos dados que está sendo gravada no Zookeeper é muito grande. Se, nos logs do Zookeeper, houver exceções como:

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

ou nos logs do Analytic Server são mensagens como:

```
Causado por: java.io.UTFDataFormatException: cadeia
codificada muito longa: 2054758 bytes
  at java.io.DataOutputStream.writeUTF(DataOutputStream.java:375)
```

1. No console do Ambari, navegue para a guia Configurações de serviço do Zookeeper, inclua a seguinte linha no env-template e, em seguida, reinicie o serviço do Zookeeper.

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

2. No console do Ambari, navegue para a guia Configs de serviço do Analytic Server e inclua o seguinte no analytics-jvm-options Avançado e, em seguida, reinicie o serviço do Analytic Cluster.

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

O número para especificar a configuração jute.maxbuffer deve ser maior que o número indicado nas mensagens de exceção.

**Dados de transação do Zookeeper não podem ser gerenciados**

Configure o parâmetro **`autopurge.purgeInterval`** em `zoo.cfg` para 1 para ativar limpezas automáticas do log de transação do Zookeeper.

**Serviço de cluster analítico perde contato com Zookeeper**

Revise e modifique os parâmetros **`tickTime`**, **`initLimit`** e **`syncLimit`** no `zoo.cfg`. Por exemplo:

```
# O número de milissegundos de cada marcação
tickTime=2000
# O número de marcações que a fase de sincronização # inicial pode obter
initLimit=30
# O número de marcações que podem ser passadas entre
# o envio de uma solicitação e a obtenção de reconhecimento
syncLimit=15
```

Consulte a documentação do Zookeeper para obter detalhes: https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html

**Tarefas do Analytic Server não continuam**

Há uma situação comum na qual as tarefas do Analytic Server não são continuadas.

- Quando uma tarefa do Analytic Server falhar porque um membro de cluster falhou, a tarefa será normalmente reiniciada automaticamente em outro membro de cluster. Se a tarefa não for retomada, verifique para assegurar se há pelo menos quatro membros de cluster no cluster de alta disponibilidade.

### Hive pushback

Você pode receber a mensagem de erro a seguir quando o pushback do Hive estiver ativado:

```
(AEQAE2103E) Falha na execução do SQL - Erro ao compilar a instrução: FALHA: SemanticException [Erro 10014]: Linha 3:47 Argumentos
errados '9223372036854775808': Comparações inseguras entre diferentes tipos são desativadas por razões de segurança. Se você sabe
o que
está fazendo, configure hive.strict.checks.type.safety como false e certifique-se de que hive.mapred.mode não esteja configurado
como 'strict' para continuar.  Observe que você pode obter erros ou resultados incorretos se cometer um erro ao usar alguns dos
recursos não seguros. (as_trace.log)
```

O erro pode ser resolvido usando um dos métodos a seguir:

- Inclua **`hive.sql.check=true`** ao arquivo `config.properties` do Analytic Server.
- Mude a configuração de **`hive.strict.checks.type.safety`** na configuração do Hive para **`false`**.

# Avisos

Estas informações foram desenvolvidas para produtos e serviços oferecidos nos EUA. Este material pode estar disponível através da IBM em outros idiomas. Entretanto, pode ser necessário possuir uma cópia do produto ou da versão do produto nesse idioma a fim de acessá-lo.

É possível que a IBM não ofereça os produtos, serviços ou recursos discutidos neste documento em outros países. Consulte um representante IBM local para obter informações sobre produtos e serviços disponíveis atualmente em sua área. Qualquer referência a produtos, programas ou serviços IBM não significa que apenas produtos, programas ou serviços IBM possam ser utilizados. Qualquer produto, programa ou serviço funcionalmente equivalente, que não infrinja nenhum direito de propriedade intelectual da IBM poderá ser usado em substituição a este produto, programa ou serviço. Entretanto, a avaliação e verificação da operação de qualquer produto, programa ou serviço não IBM são de responsabilidade do usuário.

A IBM pode ter patentes ou solicitações de patentes pendentes relativas a assuntos tratados neste documento. O fornecimento desse documento não garante a você nenhum direito sobre tais patentes. Pedidos de licença devem ser enviados, por escrito, para:

*Gerência de Relações Comerciais e Industriais da IBM Brasil*
*Av. Pasteur, 138-146*
*Botafogo,*
*Rio de Janeiro, RJ*
*CEP  22290-240*

Para consultas sobre licença relacionadas a informações de DBCS (Conjunto de Caracteres de Byte Duplo), entre em contato com o Departamento de Propriedade Intelectual da IBM em seu país ou envie pedidos de licença, por escrito, para:

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

A INTERNATIONAL BUSINESS MACHINES CORPORATION FORNECE ESTA PUBLICAÇÃO "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA DE NENHUM TIPO, SEJA EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS NÃO SE LIMITANDO ÀS GARANTIAS IMPLÍCITAS DE NÃO-VIOLAÇÃO, DE MERCADO OU DE ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Alguns países não permitem a exclusão de garantias explícitas ou implícitas em certas transações; portanto, esta disposição pode não se aplicar ao Cliente.

Essas informações podem incluir imprecisões técnicas ou erros tipográficos. Periodicamente, são feitas mudanças nas informações aqui contidas; tais mudanças serão incorporadas em novas edições da publicação. A IBM pode, a qualquer momento, aperfeiçoar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

Referências nestas informações a Web sites não IBM são fornecidas apenas por conveniência e não representam de forma alguma um endosso a esses Web sites. Os materiais contidos nesses Web sites não fazem parte dos materiais desse produto IBM e a utilização desses Web sites é de inteira responsabilidade do Cliente.

A IBM por usar ou distribuir as informações fornecidas da forma que julgar apropriada sem incorrer em qualquer obrigação para com o Cliente.

Licenciados deste programa que desejam obter informações sobre este assunto com objetivo de permitir: (i) a troca de informações entre programas criados independentemente e outros programas (incluindo este) e (ii) a utilização mútua das informações trocadas, devem entrar em contato com:

*Gerência de Relações Comerciais e Industriais da IBM Brasil*
*Av. Pasteur, 138-146*
*Botafogo,*
*Rio de Janeiro, RJ*
*CEP 22290-240*

Tais informações podem estar disponíveis, sujeitas a termos e condições apropriadas, incluindo em alguns casos o pagamento de uma taxa.

O programa licenciado descrito neste documento e todo o material licenciado disponível são fornecidos pela IBM sob os termos do IBM Customer Agreement, do Contrato de Licença do Programa Internacional da IBM ou de qualquer outro contrato equivalente.

Os dados de desempenho e os exemplos dos clientes citados são apresentados para fins ilustrativos apenas. Os resultados reais do desempenho podem variar dependendo das configurações específicas e condições operacionais.

As informações relativas a produtos não IBM foram obtidas junto aos fornecedores dos respectivos produtos, de seus anúncios publicados ou de outras fontes disponíveis publicamente. A IBM não testou estes produtos e não pode confirmar a precisão de seu desempenho, compatibilidade nem qualquer outra reivindicação relacionada a produtos não IBM. As dúvidas sobre os recursos de produtos não IBM devem ser encaminhadas diretamente aos seus fornecedores.

Declarações relacionadas aos objetivos e intenções futuras da IBM estão sujeitas a alterações ou cancelamento sem aviso prévio e representam apenas metas e objetivos.

Todos os preços da IBM mostrados são preços de varejo sugeridos pela IBM, são atuais e estão sujeitos a mudanças sem aviso prévio. Os preços para o revendedor podem variar.

Estas informações têm apenas o propósito de planejamento. As informações aqui contidas estão sujeitas a mudanças antes que os produtos descritos estejam disponíveis.

Estas informações contêm exemplos de dados e relatórios utilizados em operações comerciais diárias. Para ilustrá-los da forma mais completa possível, os exemplos podem incluir nomes de indivíduos, empresas, marcas e produtos. Todos esses nomes são fictícios e qualquer semelhança com pessoas ou empresas reais é mera coincidência.



## Marcas comerciais

IBM, o logotipo IBM e ibm.com são marcas comerciais ou marcas registradas da International Business Machines Corp., registradas em vários países no mundo todo. Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresas. ma lista atual de marcas comerciais da IBM está disponível na web em "Copyright and trademark information" em www.ibm.com/legal/copytrade.shtml.

Adobe, o logotipo Adobe, PostScript e o logotipo PostScript são marcas comerciais ou marcas registradas da Adobe Systems Incorporated nos Estados Unidos, e/ou outros países.

IT Infrastructure Library é uma marca registrada da Central Computer and Telecommunications Agency, a qual agora é parte do departamento de comércio do governo.

Intel, logotipo Intel, Intel Inside, logotipo Intel Inside, Intel Centrino, logotipo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium são marcas comerciais ou marcas registradas da Intel Corporation ou de suas subsidiárias nos Estados Unidos e em outros países.

Linux é uma marca registrada de Linus Torvalds nos Estados Unidos e/ou em outros países.

Microsoft, Windows, Windows NT e o logotipo Windows são marcas comerciais da Microsoft Corporation nos Estados Unidos, e/ou em outros países.

ITIL é uma marca registrada e uma marca comercial de comunidade registrada do The Minister for the Cabinet Office e está registrada no U.S. Patent and Trademark Office.

UNIX é uma marca registrada da The Open Group nos Estados Unidos e em outros países.

Cell Broadband Engine é uma marca comercial da Sony Computer Entertainment, Inc. nos Estados Unidos e/ou em outros países e é usada sob licença desta empresa.

Linear Tape-Open, LTO, o logotipo LTO, Ultrium e o logotipo Ultrium são marcas comerciais da HP, IBM Corp. e Quantum nos Estados Unidos e em outros países.