Configure for KMS for Ranger

Contents:

- *Configure Hadoop Cluster*
  - *Add Trifacta user properties to KMS site file*
  - *Configuration for Ranger*
- *Validate*

This section describes how to configure the Trifacta® platform for integration with KMS system for Hortonworks Data Platform. It assumes that access to the cluster is gated by Ranger.

Before you begin, please verify the pre-requisites. See *Configure for KMS*.

Configure Hadoop Cluster

> **NOTE:** These changes should be applied through the management console for the Hadoop cluster before pushing the client configuration files to the nodes of the cluster.

In the following sections:

- `[hadoop.user` (default=`trifacta)]` - the userID accessing the cluster component
- `[hadoop.group` (default=`trifactausers)]` -the appropriate group of user accessing the cluster component

Add Trifacta user properties to KMS site file

In Ambari on the Hortonworks cluster, navigate to **KMS > Configs > Advanced > kms-site**. Add the following properties:

```
<property>
  <name>hadoop.kms.proxyuser.[hadoop.user].users</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.kms.proxyuser.[hadoop.user].groups</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.kms.proxyuser.[hadoop.user].hosts</name>
  <value>*</value>
</property>
```

Configuration for Ranger

If you are using Ranger's Key Management System, additional configuration is required.

- For more information on installing KMS, see *http://docs.hortonworks.com/HDPDocuments/HDP2/HDP-2.3.2/bk_Ranger_KMS_Admin_Guide/content/ch02s01.html* .

NOTE: These changes apply to the Hortonworks cluster only. Make changes through Ambari; avoid editing configuration files directly. Configuration files do not need to be shared with the Trifacta platform.

KMS Configuration for Hive

If you are using Hive, please add the Hive users and groups information to `kms-site.xml`:

```
<property>
  <name>hadoop.kms.proxyuser.[hadoop.user].users</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.kms.proxyuser.[hadoop.user].groups</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.kms.proxyuser.[hadoop.user].hosts</name>
  <value>*</value>
</property>
```

Verify Kerberos authentication for KMS

If Kerberos is deployed, edit `kms-site.xml` and verify the following properties in `kms-site.xml`:

```
<property>
  <name>hadoop.kms.authentication.type</name>
  <value>kerberos</value>
  <description> Authentication type for the KMS. Can be either &quot;simple&quot; or &quot;kerberos&quot;.<
/description>
</property>
<property>
  <name>hadoop.kms.authentication.kerberos.keytab</name>
  <value>/etc/security/keytabs/spnego.service.keytab</value>
  <description> Path to the keytab with credentials for the configured Kerberos principal.</description>
</property>
<property>
  <name>hadoop.kms.authentication.kerberos.principal</name>
  <value>HTTP/FQDN for KMS host@YOUR HADOOP REALM</value>
  <description> The Kerberos principal to use for the HTTP endpoint. The principal must start with 'HTTP/' as
per the Kerberos HTTP SPNEGO specification.</description>
</property>
```

Verify users for KMS

If you are using Kerberos KMS authentication, verify the following properties in `kms-site.xml`:

```
<property>
  <name>hadoop.kms.proxyuser.hdfs.users</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.kms.proxyuser.hdfs.groups</name>
  <value>*</value>
</property>
<property>
  <name>hadoop.kms.proxyuser.hdfs.hosts</name>
  <value>*</value>
</property>
```

Configure connection to the KMS node

> **NOTE:** The following changes need to be applied to the Hortonworks cluster configuration files and then shared with the Trifacta node. For more information on the files required by the platform, see *Configure for Hadoop*.

Changes to `core-site.xml` for KMS

Edit `core-site.xml` and make the following change:

```
hadoop.security.key.provider.path=kms://http@<KMS_HOST>:9292/kms
```

Changes to `hdfs-site.xml` for KMS

Edit `hdfs-site.xml` and make the following change:

```
dfs.encryption.key.provider.uri=kms://http@<KMS_HOST>:9292/kms
```

Changes `dbks-site.xml` for KMS

> **NOTE:** The following changes is required only if Ranger's KMS system is enabled. If so, this change enables access to files that are stored in secured folders.

Edit `dbks-site.xml` and make the following change:

> **NOTE:** If the existing value is `hdfs`, you may leave it as-is.

```
update property hadoop.kms.blacklist.DECRYPT_EEK='-'
```

Save the files.

Validate

After the configuration is complete, you can try to import a dataset from a source stored in a cluster location managed by KMS, assuming that any required authentication configuration has been completed. See *Import Data Page*.

For more information, see *Configure Hadoop Authentication*.