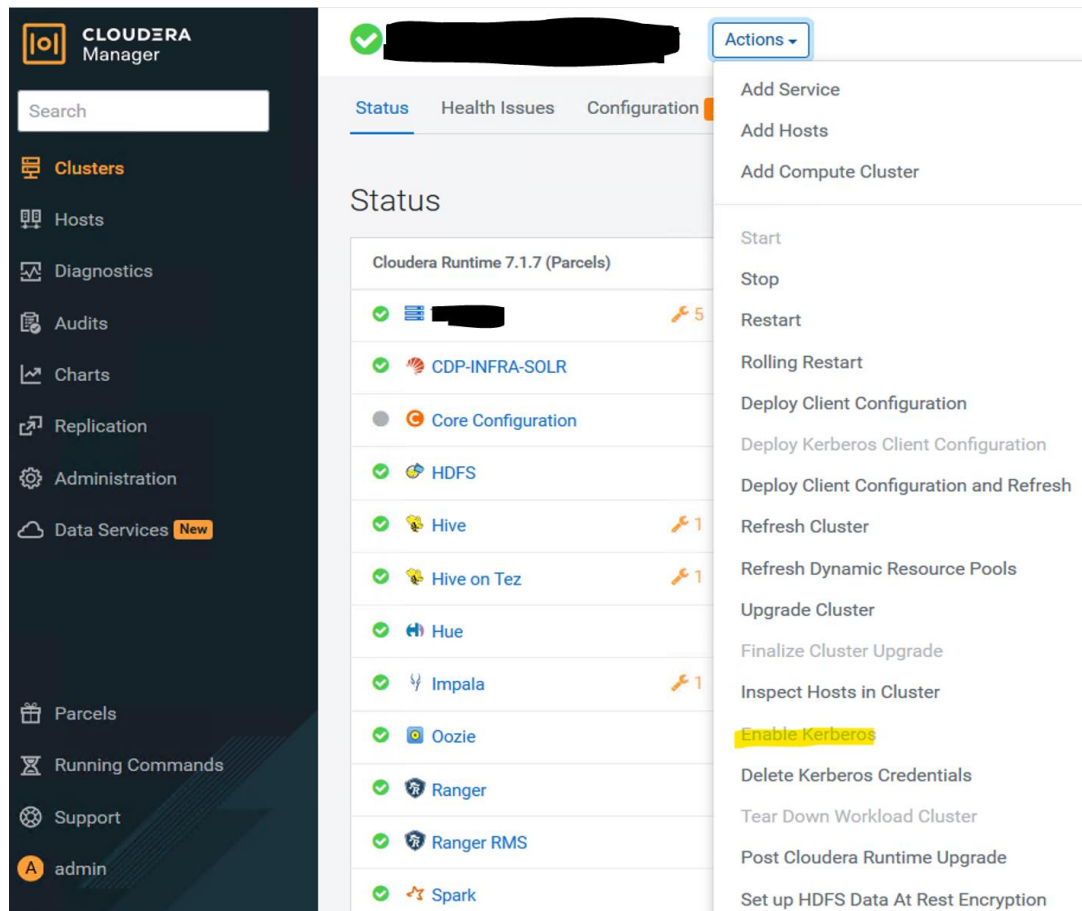


## Task1: How to enable Kerberos in cloudera

To start the Kerberos wizard, open the Cloudera Manager Admin Console, click the options menu for the applicable cluster, then click Enable Kerberos.



After opening the Kerberos wizard, you'll encounter a "Getting Started" page where you'll select your KDC type (like MIT KDC or Active Directory) to see tailored configuration steps. Follow these steps for your KDC type to set up Kerberos. Once all steps are completed, check the box confirming you've finished, then click "Continue" to proceed.

1 Getting Started

2 Enter KDC Information

3 Manage krb5.conf

4 Enter Account Credentials

5 Command Details

6 Configure Kerberos

7 Command Details

8 Summary

## Getting Started

1 This wizard walks you through the steps to configure Cloudera Manager and the cluster to use Kerberos for authentication. All services in the cluster, as well as the Cloudera Management Service, are restarted as part of the wizard.

Before using the wizard, ensure that you have performed the following steps:

1. Read the [documentation](#) about enabling Kerberos.
2. Set up a working KDC (Key Distribution Center) and specify the **KDC Type**:

KDC Type

☒ MIT KDC
 ☐ Active Directory
 ☐ Red Hat IPA

3. Configure the KDC to have **non-zero ticket lifetime and renewal lifetime**. Clusters will not work properly if tickets are not renewable.
4. Configure the KDC to have an account that has **permissions to create other accounts**.
5. Install OpenLdap client libraries on the **Cloudera Manager Server host** if you want to use Active Directory.

6.

```
# RHEL / CentOS
$ yum install openldap-clients krb5-workstation krb5-libs

# if Red Hat IPA is used as the KDC
$ yum install freeipa-client

# SUSE
$ zypper install openldap2-client krb5-client

# if Red Hat IPA is used as the KDC
$ zypper install freeipa-client

# Ubuntu
```

Once all the necessary libraries are installed on every server within the cluster, proceed by selecting the Active Directory option among the choices provided.

`yum install openldap-clients krb5-workstation krb5-libs`

```
yum install openldap-clients krb5-workstation krb5-libs
```

In the Active Directory KDC example below, we entered values for the Kerberos Security Realm, the KDC Server Host, and the Active Directory Suffix, and also selected the Active Directory Delete Accounts on Credential Regeneration check box.

Request the System team to create an organizational unit (OU) and provide the following details:

Kerberos Security Realm : mbk.com.uk

Active Directory Suffix: ou=hdp,DC=mbk,DC=uk

Getting Started

2 Enter KDC Information

3 Manage krb5.conf

4 Enter Account Credentials

5 Command Details

6 Configure Kerberos

7 Command Details

8 Summary

### Enter KDC Information

Specify information about the KDC. The properties below are used by Cloudera Manager to generate principals for daemons running on the cluster.

Kerberos Encryption Types

rc4-hmac

ⓘ

ⓘ

krb\_enc\_types

Kerberos Security Realm

HADOOP.COM

ⓘ

ⓘ

default\_realm

ⓘ

security\_realm

KDC Server Host

ⓘ

ⓘ

kdc

ⓘ

kdc\_host

KDC Admin Server Host

ⓘ

ⓘ

admin\_server

ⓘ

kdc\_admin\_host

Domain Name(s)

ⓘ

ⓘ

ⓘ

krb\_domain

Active Directory Suffix

ou=hadoop,DC=hadoop,DC=com

ⓘ

ⓘ

ad\_kdc\_domain

Active Directory Delete Accounts on Credential

☐

ⓘ

Cancel

← Back

Continue →

## Manage krb5.conf

To configure multiple Key Distribution Centers (KDCs) as mentioned earlier, you can specify the additional Domain Controllers by utilizing the Advanced Configuration Snippet (Safety Valve) for the Default Realm in the krb5.conf property box. For example:

Enable Kerberos for bdaktprod-cluster

Getting Started

Enter KDC Information

3 Manage krb5.conf

4 Enter Account Credentials

5 Command Details

6 Configure Kerberos

7 Command Details

8 Summary

### Manage krb5.conf

Specify the properties needed for generating the krb5.conf file for the cluster. You can use the Advanced Configuration Snippet to specify configuration of an advanced KDC setup; for example, with cross-realm authentication.

krb5.conf file path

/etc/krb5.conf

ⓘ

⚠ Requires Server Restart

ⓘ

krb\_krb5\_conf\_path

Manage krb5.conf through Cloudera Manager

☒ Undo

ⓘ

⚠ Requires Server Restart

ⓘ

krb\_manage\_krb5\_conf

Kerberos Ticket Lifetime

1

day(s)

ⓘ

ⓘ

ticket\_lifetime

ⓘ

krb\_ticket\_lifetime

Kerberos Renewable Lifetime

7

day(s)

ⓘ

ⓘ

renew\_lifetime

ⓘ

krb\_renew\_lifetime

DNS Lookup KDC

☐

ⓘ

ⓘ

dns\_lookup\_kdc

ⓘ

krb\_dns\_lookup\_kdc

Forwardable Tickets

☒

ⓘ

Cancel

← Back

Continue →

Forwardable Tickets

forwardable

krb\_forwardable

KDC Timeout

kdc\_timeout

krb\_kdc\_timeout

Advanced Configuration Snippet (Safety Valve) for [libdefaults] section of krb5.conf

krb\_libdefaults\_safety\_valve

Advanced Configuration Snippet (Safety Valve) for the Default Realm in krb5.conf

krb\_realms\_safety\_valve

Advanced Configuration Snippet (Safety Valve) for remaining krb5.conf

krb\_other\_safety\_valve

☒

kdc = [REDACTED]

kdc = [REDACTED]

Undo

Cancel

← Back

Continue →

After entering the credentials, the following screen will appear.

## Enable Kerberos for bdaktprod-cluster

✓ Getting Started

✓ Enter KDC Information

✓ Manage krb5.conf

✓ Enter Account Credentials

**5 Command Details**

6 Configure Kerberos

7 Command Details

8 Summary

Command Details

Import KDC Account Manager Credentials Command

Status **Running** Jul 7, 3:59:41 PM 

Abort

In the next step, all servers and services in the cluster will be Kerberized. This means that the wizard will create service principals for all hosts and services within the Cloudera Manager.

## Enable Kerberos for Cluster 1

Getting Started

Enter KDC Information

Manage krb5.conf

Enter Account Credentials

Command Details

**6 Configure Kerberos**

7 Command Details

8 Summary

### Configure Kerberos

Install Kerberos client libraries on all hosts before proceeding.

```
# RHEL / CentOS
$ yum install krb5-workstation krb5-libs

# if Red Hat IPA is used as the KDC
$ yum install freeipa-client

# SUSE
$ zypper install krb5-client

# if Red Hat IPA is used as the KDC
$ zypper install freeipa-client

# Ubuntu
$ apt-get install krb5-user

# if Red Hat IPA is used as the KDC
$ apt-get install freeipa-client
```

Configure the privileged ports required by DataNodes in a secure HDFS service.

DataNode Transceiver Port

Port for DataNode's Xceiver Protocol. Combined with the DataNode's hostname to build its address.

DataNode HTTP Web UI Port

Port for the DataNode HTTP web UI. Combined with the DataNode's hostname to build its HTTP address.

After successfully completing all the steps, all principals were added.

Getting Started

Enter KDC Information

Manage krb5.conf

Enter Account Credentials

Command Details

Configure Kerberos

**7 Command Details**

8 Summary

### Command Details

#### Enable Kerberos Command

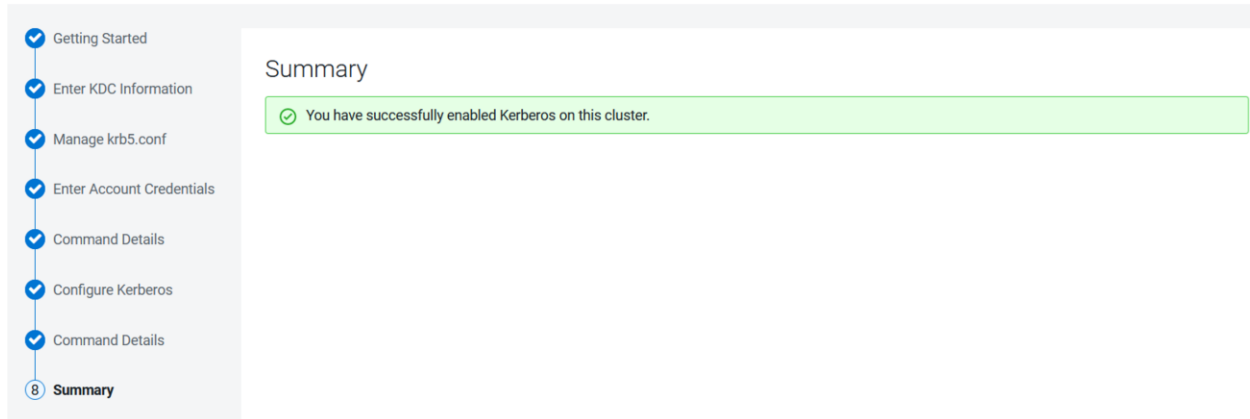
Status **Running** Context [bdaktprod-cluster](#) Jul 7, 7:17:35 PM [Abort](#)

Completed 6 of 7 step(s).

☒ Show All Steps ☐ Show Only Failed Steps ☐ Show Only Running Steps

>  Stop cluster	<a href="#">bdaktprod-cluster</a>	Jul 7, 7:17:35 PM	3.91s
>  Stop Cloudera Management Services	<a href="#">Cloudera Management Service</a>	Jul 7, 7:17:39 PM	3.13s
>  Configure all services to use Kerberos	<a href="#">bdaktprod-cluster</a>	Jul 7, 7:17:42 PM	112ms
>  Wait for credentials to be generated		Jul 7, 7:17:43 PM	12.27s
>  Deploy client configuration	<a href="#">bdaktprod-cluster</a>	Jul 7, 7:17:55 PM	21.31s
>  Start Cloudera Management Services	<a href="#">Cloudera Management Service</a>	Jul 7, 7:18:16 PM	38.6s
>  Start cluster	<a href="#">bdaktprod-cluster</a>	Jul 7, 7:18:55 PM	<a href="#">Abort</a>

## Enable Kerberos for bdaktprod-cluster



<https://docs.cloudera.com/cdp-private-cloud-base/7.1.6/security-kerberos-authentication/topics/cm-security-kerberos-enabling-step4-kerberos-wizard.html>

<https://www.youtube.com/watch?v=n1gjlwm438>

## Task2: How to create the key tab file for specific user

Following are the Steps to Create Keytab file for a new user;

su - user

ktutil

ktutil: addent -password -p user1@mbk.com.uk -k 1 -e RC4-HMAC

Password for user1@mbk.com.uk:

ktutil: wkt user1.keytab



ktutil: q

Validate the principal.

klist -kt /etc/security/keytabs/user1.keytab

### Task3: How to fetch the data through MySQL source using the spark script

First need to download the specific mysql jar file based on the spark version

	mysql-connector-java-8.0.27	12/22/2023 7:08 AM	Executable Jar File	2,418 KB
	protobuf-java-3.11.4	12/22/2023 7:08 AM	Executable Jar File	1,623 KB

PFB the code for fetching the data from mysql source

```
pyspark --jars /u01/software/mysql_files/mysql-connector-java-5.1.48/mysql-connector-java-8.0.27.jar,/u01/software/mysql_files/mysql-connector-java-5.1.48/protobuf-java-3.11.4.jar,,/u01/software/jar_files/kudu-spark2_2.11-1.15.0.7.1.7.2000-305.jar,/u01/software/jar_files/hive-kudu-handler-3.1.3000.7.1.7.2000-305.jar,/u01/software/jar_files/ojdbc7.jar --num-executors 4 --executor-memory 20G
```

```
from pyspark.sql import SparkSession
```

```
spark = SparkSession.builder \
    .appName("Mysql Read")\
    .getOrCreate()
```

```
database = 'abl_infobip'
driver = 'com.mysql.cj.jdbc.Driver'
username = 'add the username'
hostname = 'add the host'
port = '3306'
table_name = 'table_name'
password = '#####'
```

```
mysql_url = "jdbc:mysql://" + hostname + ":" + port + "/" + database
```

```
df = spark.read.format("jdbc")\
    .option("url", mysql_url)\
    .option("driver", driver)\
    .option("dbtable", table_name)\
    .option("user", username)\
    .option("password", password)\
    .load()
```