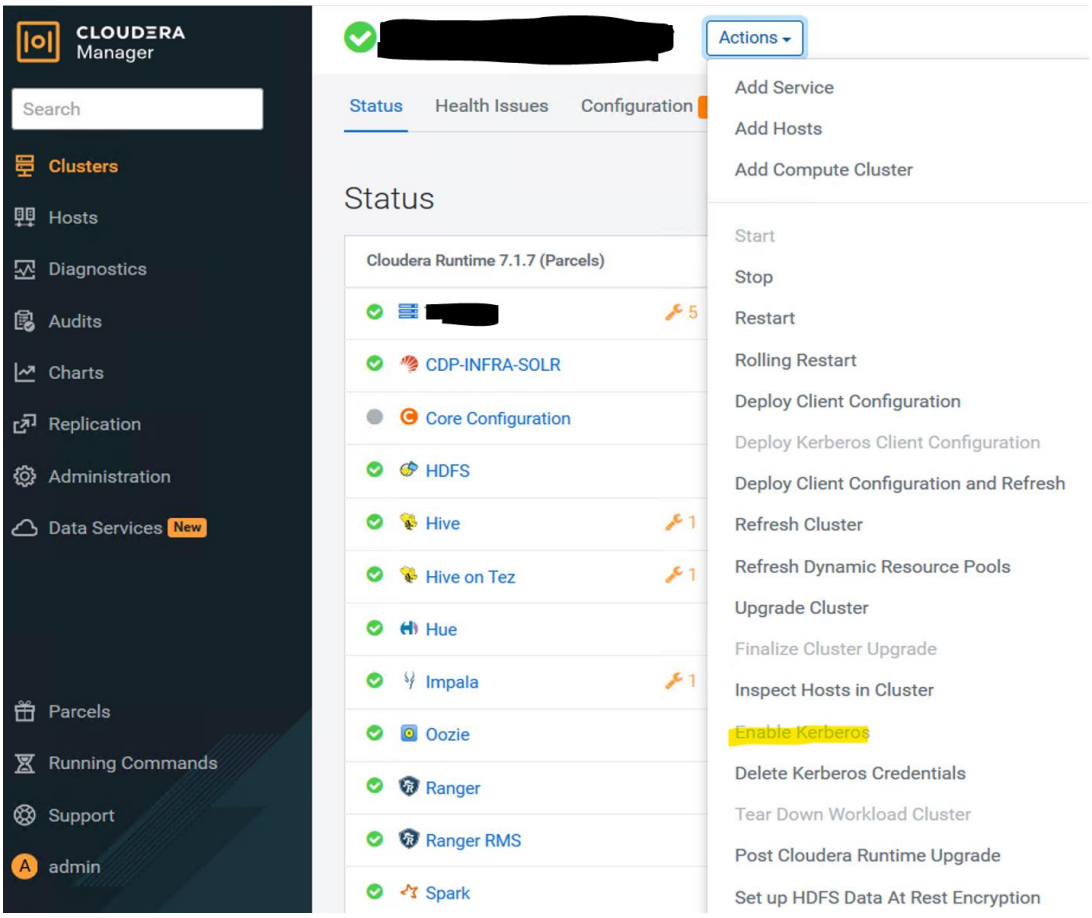
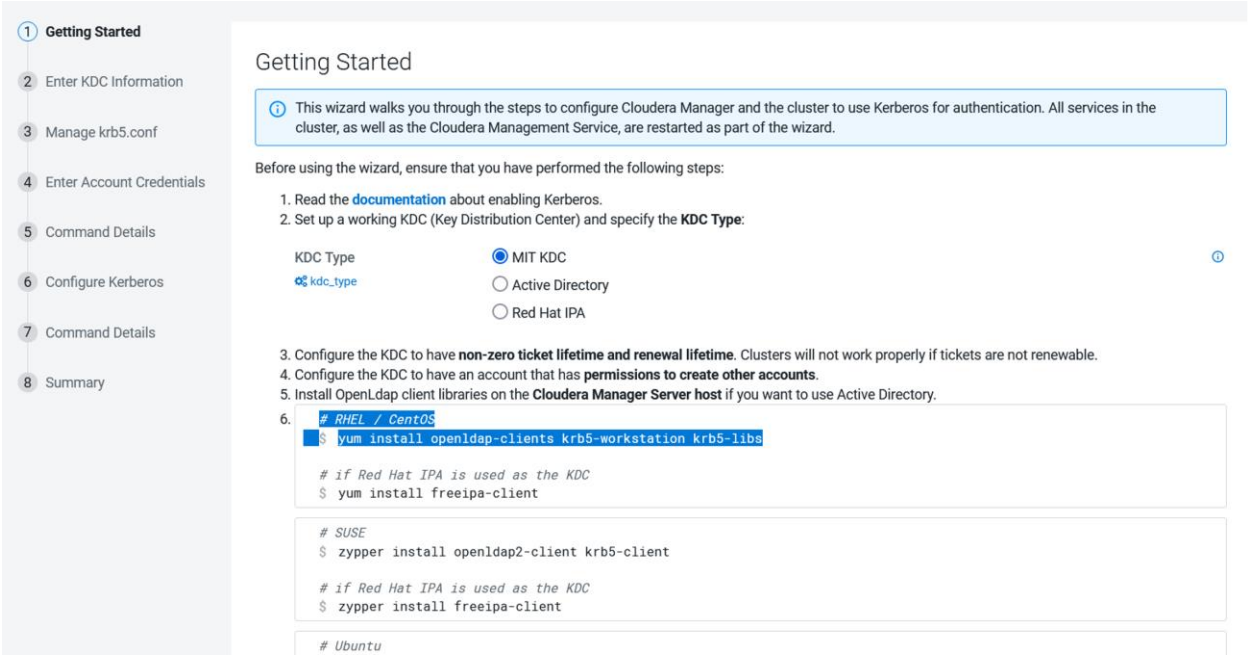
Task1: How to enable Kerberos in cloudera

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.



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

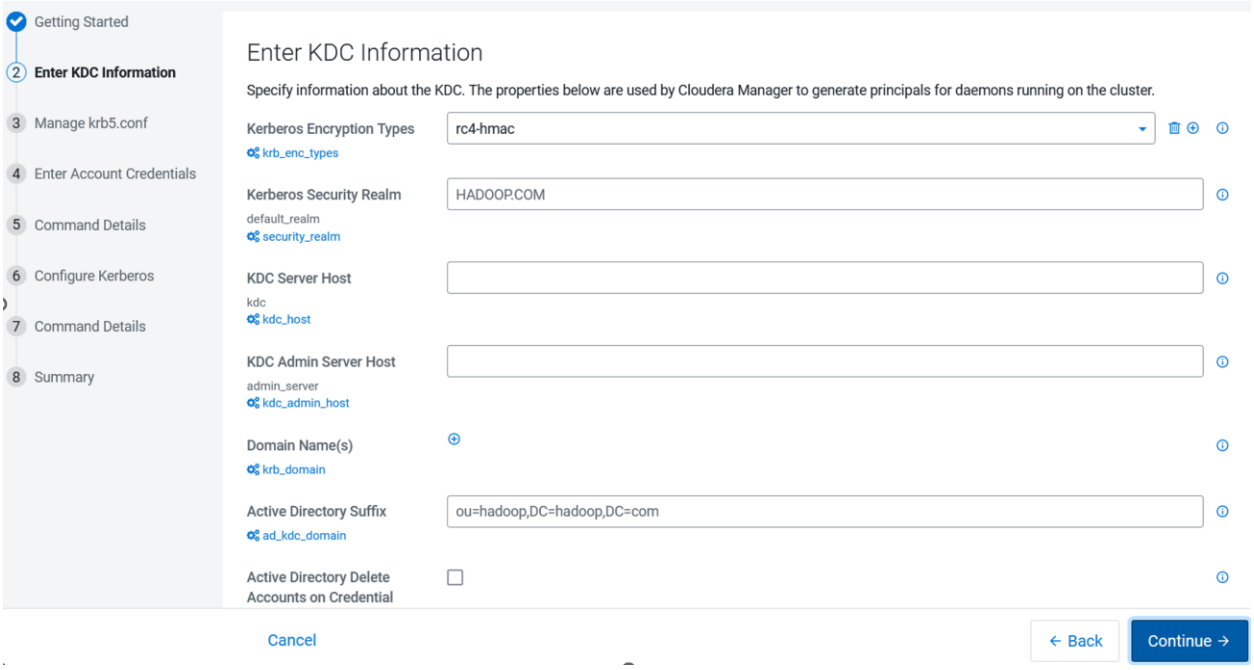
yum install openldap-clients krb5-workstation krb5-libs

```
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.

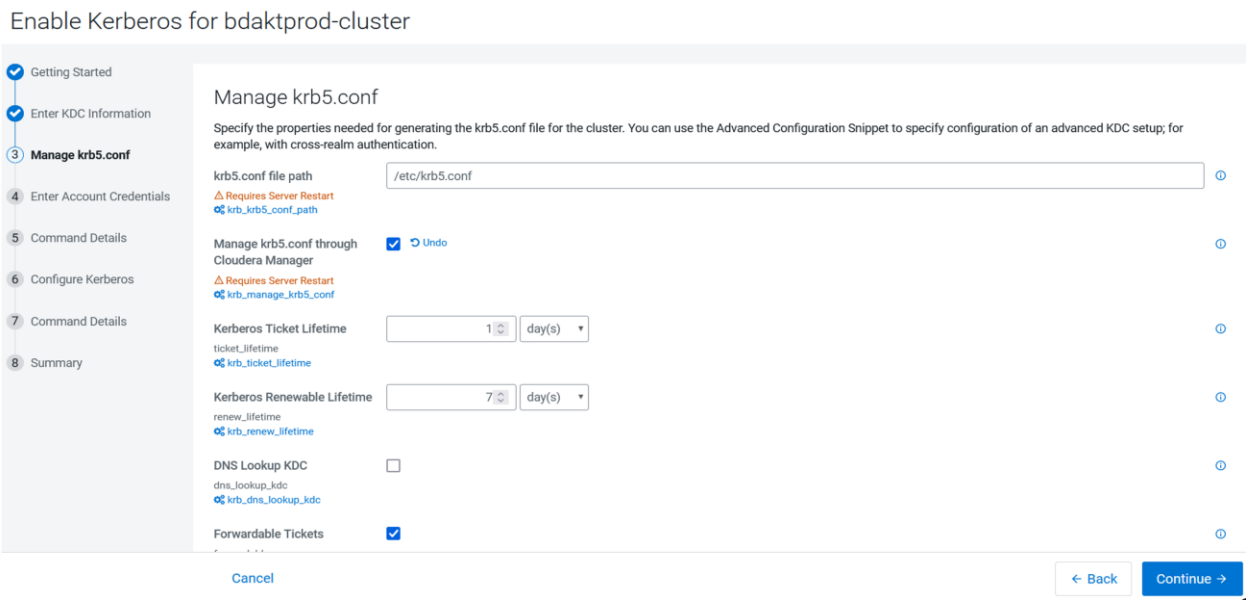
Request the System team to create an organizational unit (OU) and provide the following details:

Kerberos Security Realm : mbk.com.uk
Active Directory Suffix: ou=hdp,DC=mbk,DC=uk



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:



Forwardable Tickets
forwardable
krb_forwardable

KDC Timeout
kdc_timeout
krb_kdc_timeout
3 second(s)

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
kdc = [redacted]
kdc = [redacted]
Undo

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

Cancel ← Back Continue →

After entering the credentials, the following screen will appear.

Enable Kerberos for bdaktprod-cluster

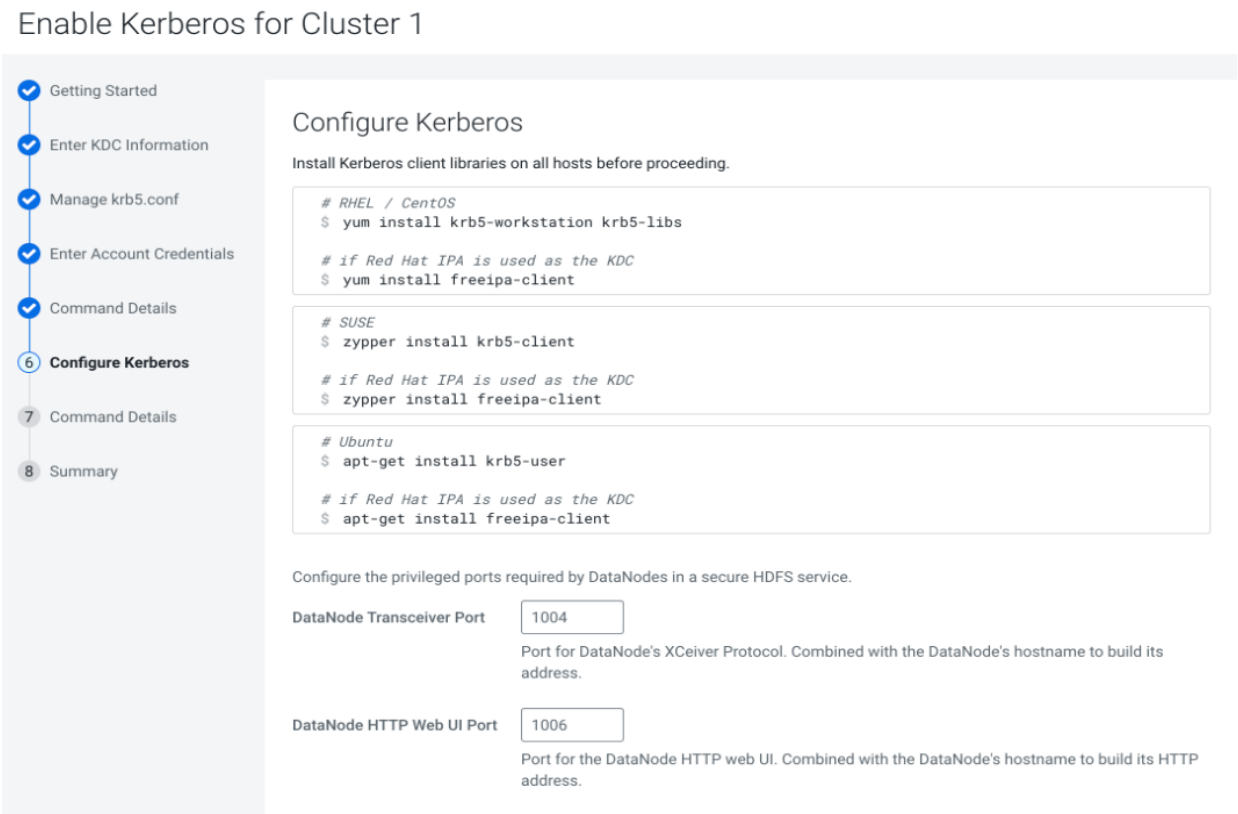
- Getting Started
- Enter KDC Information
- Manage krb5.conf
- Enter Account Credentials
- 5 **Command Details**
- 6 Configure Kerberos
- 7 Command Details
- 8 Summary

Command Details

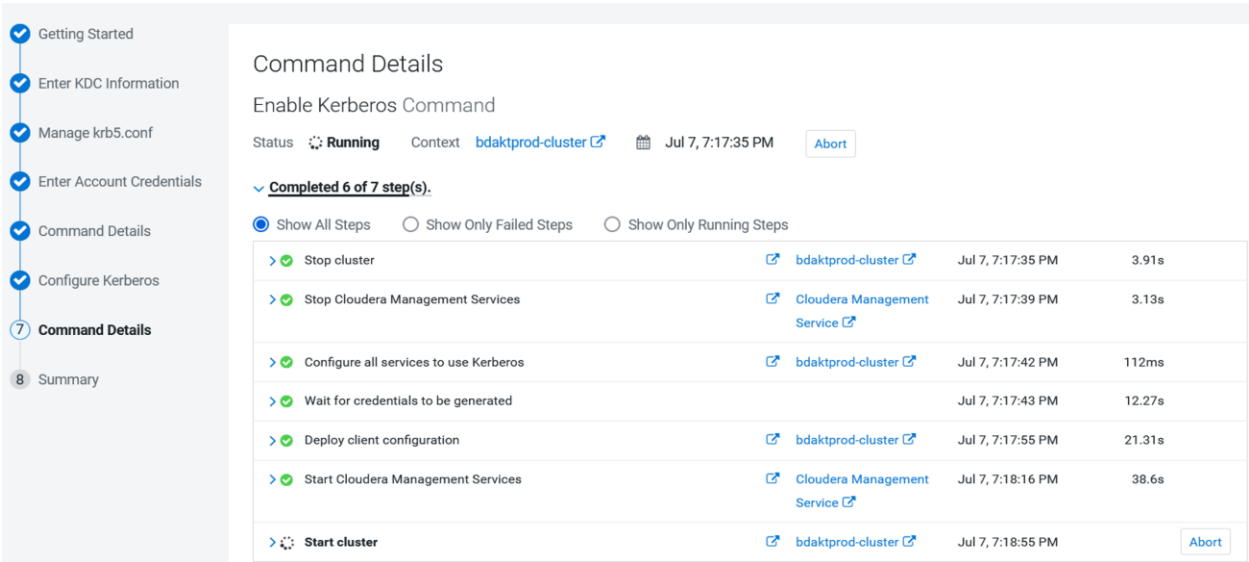
Import KDC Account Manager Credentials Command

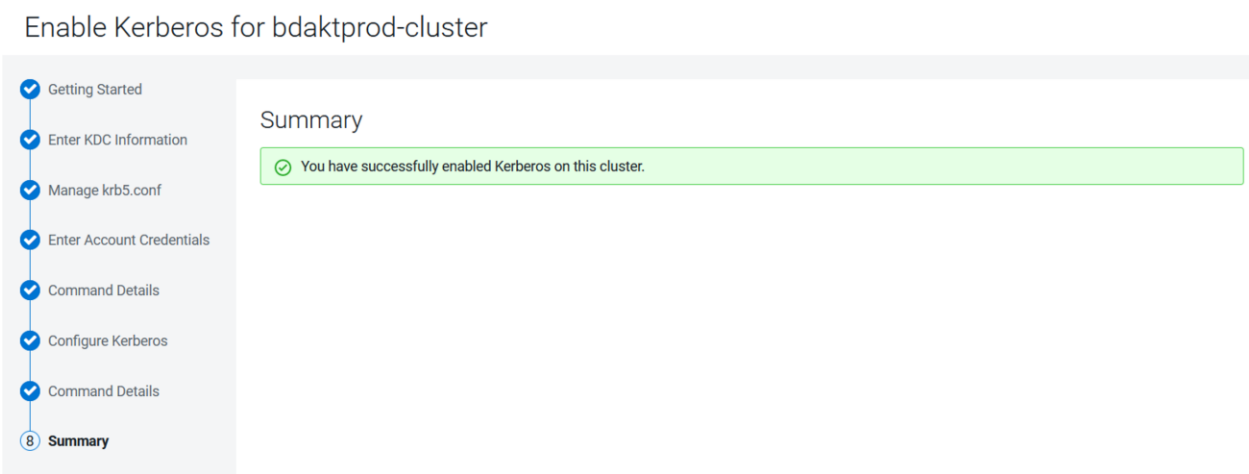
Status **Running** Jul 7, 3:59:41 PM Abort

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.



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





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=n1gjvlwm438

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/softwares/mysql_files/mysql-connector-java-5.1.48/mysql-connector-java-
8.0.27.jar,/u01/softwares/mysql_files/mysql-connector-java-5.1.48/protobuf-java-
3.11.4.jar,,/u01/softwares/jar_files/kudu-spark2_2.11-1.15.0.7.1.7.2000-
305.jar,/u01/softwares/jar_files/hive-kudu-handler-3.1.3000.7.1.7.2000-
305.jar,/u01/softwares/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()
```