

DellEMC Isilon OneFS 8.2: Cloudera Enterprise Upgrade Guide

Upgrading Cloudera CDH 5.14.4 to CDH 6.3.1 on DellEMC Isilon OneFS 8.2 and later versions.

Abstract

The official DellEMC recommended process to upgrade Cloudera CDH 5.14.4 to CDH 6.3.1 on Isilon OneFS 8.2.

October 2019

Revisions

Date	Description
October 2019	Initial release

Acknowledgements

This paper was produced by the following:

Author: Kirankumar Bhusanurmath (Kirankumar.bhusanurmath@dell.com), Analytics Solutions Architect.

Support:

Other:

The information in this publication is provided "as is." Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

Copyright © <pub date – rev date> Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners. [11/6/2019] [Deployment Guide] [H18026]

Table of contents

Revisions	2
Acknowledgements	2
Table of contents	3
Introduction	5
1 Upgrade Cluster Requirements	6
1.1 Upgrade Overview	6
2 Upgrading Cloudera Manager	7
2.1 Getting Started Upgrading Cloudera Manager	7
2.1.1 My Environment	7
2.1.2 Collect Information	7
2.1.3 Preparing to Upgrade Cloudera Manager	8
2.2 Back Up Cloudera Manager	9
2.3 Upgrading the Server	9
2.3.1 Step1: Establish Access to the Software	9
2.3.2 Step2: Install Oracle JDK 8	9
2.3.3 Step3: Upgrade the Cloudera Manager Server	10
2.4 Upgrading the Agents	11
2.4.1 Step1: Upgrade the Agent using Cloudera Manager	11
2.5 After You Upgrade	11
2.5.1 Perform Post Upgrade Steps	11
2.6 Troubleshooting	11
2.7 Reverting a Failed Upgrade	11
3 Upgrading a CDH Cluster	12
3.1 Getting Started	12
3.1.1 My Environment	12
3.1.2 Collect Information	12
3.1.3 Preparing to Upgrade CDH	12
3.2 Backing Up CDH	13
3.3 CDH 6 Pre-Upgrade Migration	13
3.4 Upgrading the Cluster	14
3.4.1 Back Up Cloudera Manager	14
3.4.2 Enter Maintenance Mode	14
3.4.3 Complete Pre-Upgrade Migration Steps	14
3.4.4 Run Hue Document Cleanup	14

- 3.4.5 Check Oracle Database Initialization 14
 - 3.4.6 Download and Distribute Parcels..... 14
 - 3.4.7 Run the Upgrade CDH Wizard..... 15
 - 3.4.8 Remove the Previous CDH Version Packages and Refresh Symlinks..... 16
 - 3.4.9 Finalize the HDFS Upgrade..... 16
 - 3.4.10 Complete Post-Upgrade Migration Steps..... 17
 - 3.4.11 Exit Maintenance Mode..... 17
- 3.5 CDH 6 Post-Upgrade Migration 17
- 3.6 Manual Upgrade Steps..... 17
- 3.7 Troubleshooting 17
- A Technical support and resources 18
 - A.1 Related resources 18

Introduction

This Deployment and Configuration guide walks you through the process of upgrading Cloudera Distribution Hadoop (CDH) 5.14.4 to CDH 6.3.1 on DellEMC Isilon OneFS 8.2. This is intended for systems administrators, IT program managers, IT architects, and IT managers who are upgrading Cloudera Enterprise Distribution of Hadoop on OneFS 8.2 or later versions.

We will demonstrate in-place upgrades. Make sure the cluster is ready and meets all the success criteria as mentioned in the official [Cloudera Enterprise Upgrade Guide](#).

1 Upgrade Cluster Requirements

We will follow the Cloudera official Cloudera Enterprise Upgrade guide; the following topics provide overview of the Cloudera Enterprise upgrade process and include complete procedures for upgrading Cloudera Manager and CDH clusters.

Upgrade CDH cluster requirements

1. Hadoop cluster running Cloudera Manager 5.14 and CDH 5.14.4
2. DellEMC Isilon OneFS updated to 8.2.0

Note: This upgrade process is specifically from CDH 5.14.4 to CDH 6.3.1 on DellEMC Isilon, please be informed the intermediary CDH version do not contain Isilon supporting code bits.

1.1 Upgrade Overview

Upgrading consists of two major steps.

1. Upgrading Cloudera Manager
2. Upgrading a CDH cluster

2 Upgrading Cloudera Manager

When you upgrade Cloudera Manager, use RPM based package command to upgrade the software on the Cloudera Manager server host and Cloudera Manager manages upgrading the Cloudera Manager Agents on the remaining managed hosts.

Follow below step by step process or refer [here](#).

2.1 [Getting Started Upgrading Cloudera Manager](#)

Before upgrade Cloudera Manager, you need to gather some information and review the limitations and release notes.

2.1.1 My Environment

Fill in the following form to create a customized set of instructions for your environment.

Figure 1 Upgrade Cluster Details

2.1.2 Collect Information

1. Login into Cloudera manager Server host
2. Check current OS and Database parameters

```
[root@pipe-cdm ~]# lsb_release
LSB Version:      :core-4.1-amd64:core-4.1-noarch
[root@pipe-cdm ~]# lsb_release -a
LSB Version:      :core-4.1-amd64:core-4.1-noarch
Distributor ID:   CentOS
Description:      CentOS Linux release 7.7.1908 (Core)
Release:          7.7.1908
Codename:         Core
[root@pipe-cdm ~]# cat /etc/cloudera-scm-server/db.properties
# Auto-generated by scm_prepare_database.sh on Wed Oct 16 03:25:01 EDT 2019
#
# For information describing how to configure the Cloudera Manager Server
# to connect to databases, see the "Cloudera Manager Installation Guide."
#
com.cloudera.cmf.db.type=postgresql
com.cloudera.cmf.db.host=localhost
com.cloudera.cmf.db.name=scm
com.cloudera.cmf.db.user=scm
com.cloudera.cmf.db.setupType=EXTERNAL
com.cloudera.cmf.db.password=scm
[root@pipe-cdm ~]#
```

Figure 2 Upgrade Cluster OS and DB parameters

3. Log in to Cloudera Manager Admin Console and find CM and JDK versions from Support→ About

About

Version: Cloudera Enterprise Trial 5.14.4 (#3 built by jenkins on 20180707-0445 git: 0971e84bdceb60db9b96533f46451f40ed8cbdf9)

Java VM Name: Java HotSpot(TM) 64-Bit Server VM

Java VM Vendor: Oracle Corporation

Java Version: 1.7.0_67

Server Time: Oct 17, 2019 2:39:57 AM, Eastern Daylight Time (EDT)

Copyright © 2011-2019 Cloudera, Inc. All rights reserved.
Hadoop and the Hadoop elephant logo are trademarks of the Apache Software Foundation.

Close

Figure 3 Upgrade Cluster Cloudera Manager and JDK

2.1.3 Preparing to Upgrade Cloudera Manager

1. SSH access to Cloudera Manager Server host and able to log in using root into all cluster hosts.
2. Review the CDH upgrade documents, OS, Database and JAVA requirements.

2.2 [Back Up Cloudera Manager](#)

Cloudera recommends that you perform these backup steps before upgrading. The backups will allow you to rollback your Cloudera Manager upgrade if needed. These steps are out of scope of this process, but it is mandatory to follow these steps from the Cloudera website, refer [here](#).

1. Collect information
2. Back up Cloudera Manager Agent
3. Back Up the Cloudera Management Service
4. Stop Cloudera Manager Server and Cloudera Management Service
5. Back Up the Databases
6. Back Up Cloudera Manager Server
7. (Optional) Start Cloudera Manager Server and Cloudera Management Service

2.3 [Upgrading the Server](#)

This topic provides procedures for upgrading the Cloudera Manager Server.

2.3.1 [Step1: Establish Access to the Software](#)

1. Cloudera Manager needs access to a package repository that contains the updated software packages
2. Login into Cloudera Manager Server host and remove existing repository directory

```
sudo rm /etc/yum.repos.d/cloudera*manager.repo*
```

3. Create a repository file so that the package manager can locate and download the binaries.

Create a file named `/etc/yum.repos.d/cloudera-manager.repo` with the following content:

```
[cloudera-manager]
# Packages for Cloudera Manager
name=Cloudera Manager
baseurl=https://archive.cloudera.com/cm6/6.3.1/redhat7/yum/
gpgkey=https://archive.cloudera.com/cm6/6.3.1/redhat7/yum/RPM-GPG-KEY-cloudera
gpgcheck=1
```

4. A Cloudera Manager upgrade can introduce new package dependencies. Your organization may have restrictions or require prior approval for installation of new packages. You can determine which packages may be installed or upgraded:

```
yum deplist cloudera-manager-agent
```

2.3.2 [Step2: Install Oracle JDK 8](#)

Oracle JDK 1.8 is required on all cluster hosts managed by Cloudera Manager 6.0.0 or higher. As part of the Cloudera Manager upgrade process, you can specify that Cloudera Manager upgrade the JDK on the remaining hosts.

1. Login in to the Cloudera Manager Server host
2. Install JDK 1.8

```
sudo yum install oracle-j2sdk1.8.x86_64
```

3. Open the following file in a text editor:

```
/etc/default/cloudera-scm-server
```

4. Edit the line that begins with export JAVA_HOME (if this line does not exist, add it) and change the path to the path of the new JDK (you can find the path under /usr/java).

```
export JAVA_HOME="/usr/java/jdk1.8.0_181-cloudera"
```

5. Save the file.

2.3.3 Step3: Upgrade the Cloudera Manager Server

1. Log in to the Cloudera Manager Server host
2. Stop the Cloudera Management Service
3. Stop the Cloudera Manager Server

```
sudo systemctl stop cloudera-scm-server
```

4. Stop the Cloudera Manager Agent

```
sudo systemctl stop cloudera-scm-agent
```

5. Upgrade the packages.

```
sudo yum clean all
sudo yum upgrade cloudera-manager-server cloudera-manager-daemons
cloudera-manager-agent
```

6. Verify that you have the correct packages installed.

```
rpm -qa 'cloudera-manager-*'
cloudera-manager-daemons-6.3.1-1466458.el7.x86_64
cloudera-manager-agent-6.3.1-1466458.el7.x86_64
cloudera-manager-server-6.3.1-1466458.el7.x86_64
```

7. Start the Cloudera Manager Agent

```
sudo systemctl start cloudera-scm-agent
```

8. Start the Cloudera Manager Server

```
sudo systemctl start cloudera-scm-server
```

9. Open the Cloudera Manager Admin Console on the Web browser.

```
http://cloudera_Manager_server_hostname:7180/cm/upgrade
```

2.4 [Upgrading the Agents](#)

We will upgrade the agents using Cloudera Manager which is the recommend method.

2.4.1 [Step1: Upgrade the Agent using Cloudera Manager](#)

1. Login into Cloudera Manager Server hosts admin console

```
https://my_cloudera_manager_server_host:port/cm/upgrade
```

2. Select the group from the drop-down list labeled Upgrade Cloudera Manager Agent Packages running on:
3. Click Upgrade Cloudera Manager Agent packages.
4. Click Continue
The **Accept JDK License** page displays.
5. Install JDK 8 on all the hosts by selecting **Install Oracle Java SE Development Kit**.
6. Click Continue.
The Enter Login Credentials page displays.
7. Specify the credentials and initiate Agent installation:
8. Click Continue.
The Cloudera Manager Agent packages and, if selected, the JDK are installed.
9. When the installations complete, click **Finish**.
The **Upgrade Cloudera Manager** page displays the status of the upgrade.
10. Click **Run Host Inspector**, start the Cloudera Management Service.
11. Go to Home Page.
12. If required restart services or redeploy stale client configurations.

2.5 [After You Upgrade](#)

2.5.1 [Perform Post Upgrade Steps](#)

1. Start the **Cloudera Management Service** and adjust any configurations when prompted.
2. Redeploy any stale client configurations.

2.6 [Troubleshooting](#)

2.7 [Reverting a Failed Upgrade](#)

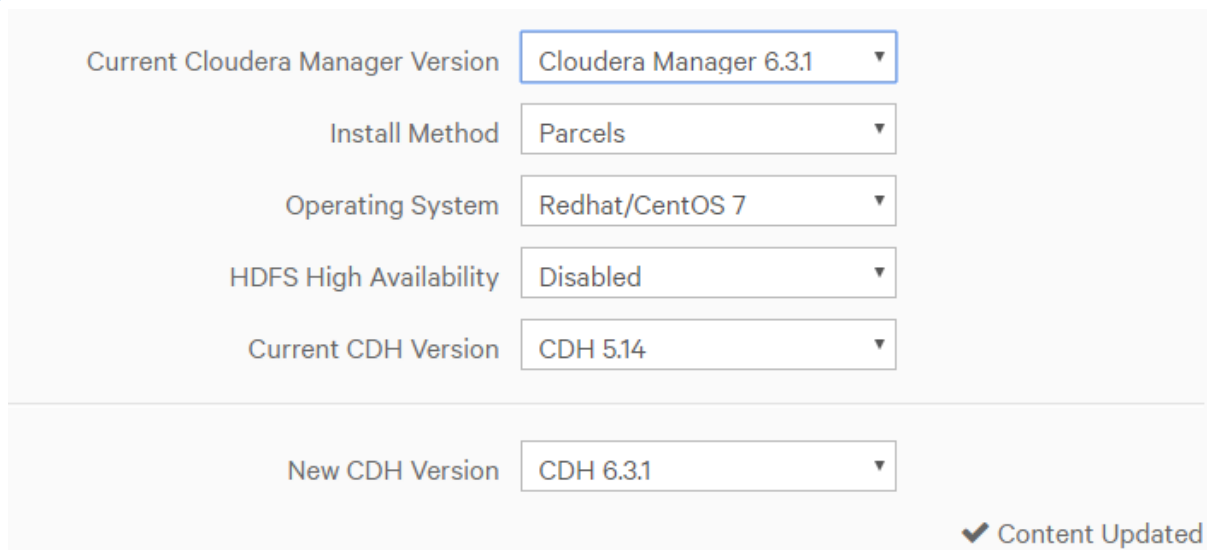
3 Upgrading a CDH Cluster

This topic describes how to upgrade CDH from 5.14.4 version to a higher version of CDH 6.3.1, using Cloudera Manager and parcels.

Follow below step by step process or refer [here](#).

3.1 [Getting Started](#)

3.1.1 My Environment



Current Cloudera Manager Version	Cloudera Manager 6.3.1
Install Method	Parcels
Operating System	Redhat/CentOS 7
HDFS High Availability	Disabled
Current CDH Version	CDH 5.14
New CDH Version	CDH 6.3.1

✓ Content Updated

Figure 4 Upgrade CDH Cluster

3.1.2 Collect Information

1. Login into the Cloudera Manager Server host
2. Check current OS versions

```
lsb_release -a
```

3. Login into Cloudera Manager Console and confirm Cloudera Manager and versions
4. Also find **Install Method** and current CDH version displayed on the Cloudera Manager home page.

3.1.3 Preparing to Upgrade CDH

1. Should be able to SSH as root to Cloudera Manager Server and all hosts.
2. Review [Cloudera Enterprise 6 Requirements and Supported Versions](#)
3. Ensure supported Java is installed on all hosts in the cluster, refer [link](#).
4. Review [CDH 6 Release Notes](#) and [Cloudera Security Bulletins](#).
5. Review the upgrade procedure and reserve a maintenance window with enough time allotted to perform all steps.

6. If the cluster uses **Impala**, check your SQL against the newest reserved words listed in [incompatible changes](#).
7. Run the [Security Inspector](#) and fix any reported errors.
8. If your cluster uses HBase, see [Migrating Apache HBase Before Upgrading to CDH 6](#).
9. If you are upgrading to CDH 6.0 or higher, and Hue is deployed in the cluster, and Hue is using PostgreSQL as its database, you must manually install psycopg2. [See Installing Dependencies for Hue](#).
10. The following services are no longer supported as of Enterprise 6.0.0:
 - a. Accumulo
 - b. Sqoop 2
 - c. MapReduce 1
 - d. Spark 1.6
 - e. Record Service

You must stop and delete these services before upgrading CDH. See [Stopping a Service on All Hosts and Deleting Services](#).

11. Back up Cloudera Manager before beginning the upgrade. See [Backing Up Cloudera Manager](#).
12. Review all CDH 6 pre-upgrade migration steps. There are steps you must perform before beginning the upgrade for the following components: Sentry, Cloudera Search, Apache Spark, HBase, Hue, Key Trustee KMS, HSM KMS.

3.2 [Backing Up CDH](#)

This topic describes how to back up a CDH cluster managed by Cloudera Manager prior to upgrading the cluster. These procedures *do not back up the data stored in the cluster*. Backing up CDH process is out of scope of this blog post, but it is mandatory to perform these steps before proceeding ahead with the upgrade.

1. [Back Up Databases](#)
2. [Back Up ZooKeeper](#)
3. [Back Up HDFS](#)
4. [Back Up Key Trustee Server and Clients](#)
5. [Back Up HSM KMS](#)
6. [Back Up Navigator Encrypt](#)
7. [Back Up HBase](#)
8. [Back Up Search](#)
9. [Back Up Sqoop 2](#)
10. [Back Up Hue](#)

3.3 [CDH 6 Pre-Upgrade Migration](#)

If you have deployed the Sentry, HBase, Cloudera Search, Spark, Key Trustee KMS, or HSM KMS services in a CDH 5.x cluster that you want to upgrade to CDH 6, CDH 6 Pre-Upgrade Migration process is out of scope of this blog post, but it is mandatory to perform these steps before proceeding ahead with the upgrade. See the following topics for additional pre-upgrade steps:

1. [Sentry Policy Files](#)
2. [Cloudera Search Configuration](#)
3. [Apache Spark](#)

4. [Apache HBase](#)
5. [Hue](#)
6. [Key Trustee KMS](#)
7. [HSM KMS](#)

3.4 [Upgrading the Cluster](#)

The version of CDH you can upgrade to depends on the version of Cloudera Manager that is managing the cluster. You may need to upgrade Cloudera Manager before upgrading CDH.

3.4.1 [Back Up Cloudera Manager](#)

Before you upgrade a CDH cluster, back up Cloudera Manager. Even if you just backed up Cloudera Manager before an upgrade, you should now back up your upgraded Cloudera Manager deployment. See [Backing Up Cloudera Manager](#).

3.4.2 [Enter Maintenance Mode](#)

To avoid unnecessary alerts during the upgrade process, enter maintenance mode on your cluster before you start the upgrade.

3.4.3 [Complete Pre-Upgrade Migration Steps](#)

Make sure [CDH 6 Pre-Upgrade Migration](#) step is completed successfully.

3.4.4 [Run Hue Document Cleanup](#)

If your cluster uses Hue, perform the following. These steps clean up the database tables used by Hue and can help improve performance after an upgrade.

3.4.5 [Check Oracle Database Initialization](#)

If your cluster uses Oracle for any databases, before upgrading from CDH 5 to CDH 6, check the value of the COMPATIBLE initialization parameter in the Oracle Database using the following SQL query: Refer Step 5 of the [Upgrading the cluster](#)

3.4.6 [Download and Distribute Parcels](#)

1. Log in to the Cloudera Manager Admin Console.
2. Click Hosts > Parcels
3. Update the parcel Repository for CDH using the following remote parcel repository URL:

```
https://archive.cloudera.com/cdh6/6.3.1/parcels/
```

4. If your cluster has **GPLEXTRAS** installed, update the version of the GPLEXTRAS parcel to match the CDH version using the following remote parcel repository URL:

```
https://archive.cloudera.com/gplextras6/6.3.1/parcels/
```

5. After all the parcels are distributed, click on the Upgrade button next to the chosen CDH. The chosen CDH should be selected automatically.

3.4.7 Run the Upgrade CDH Wizard

1. Get to the Upgrade CDH page from the Home > Status tab, click dropdown and select Upgrade Cluster.
2. In the **Choose CDH(Parcels)** section, select the **CDH 6.3.1 version**.
3. Click **Continue**.
4. A page displays the version you are upgrading to and asks you to confirm that you have completed some additional steps.
5. Click **YES, I have performed these steps**.
6. Click **Continue**.
7. Cloudera Manager verifies that the agents are responsive and that the correct software is installed. When you see the **No Errors Found** message, click **Continue**.
8. The selected parcels are downloaded, distributed, and unpacked.

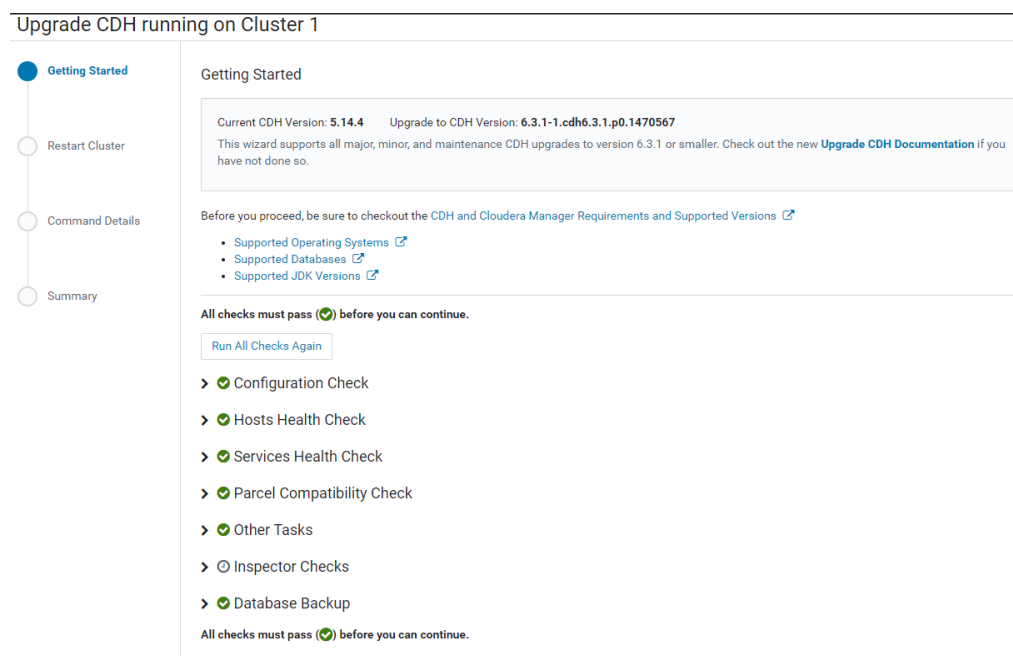


Figure 5 Upgrade CDH Wizard

9. Click **Continue**.

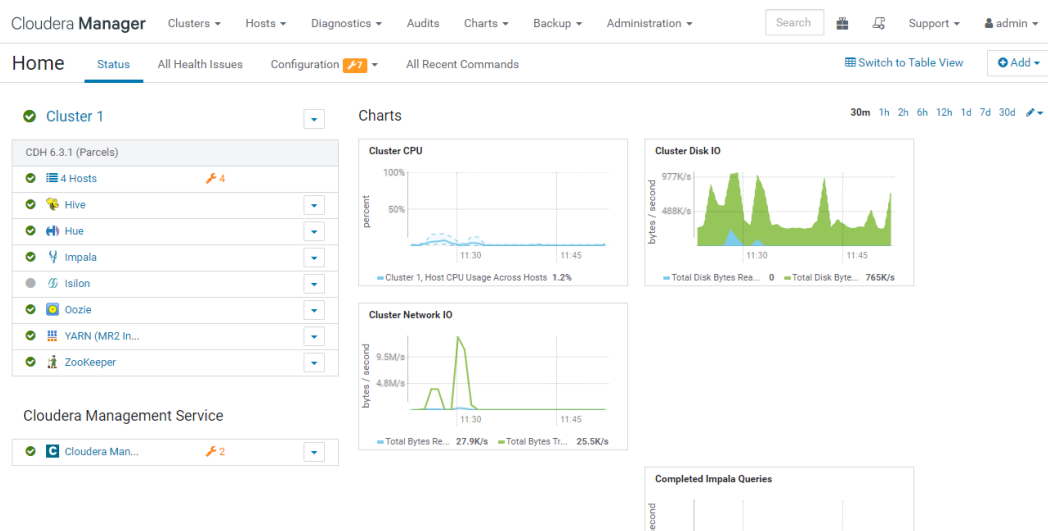


Figure 6 Upgraded CDH Cluster

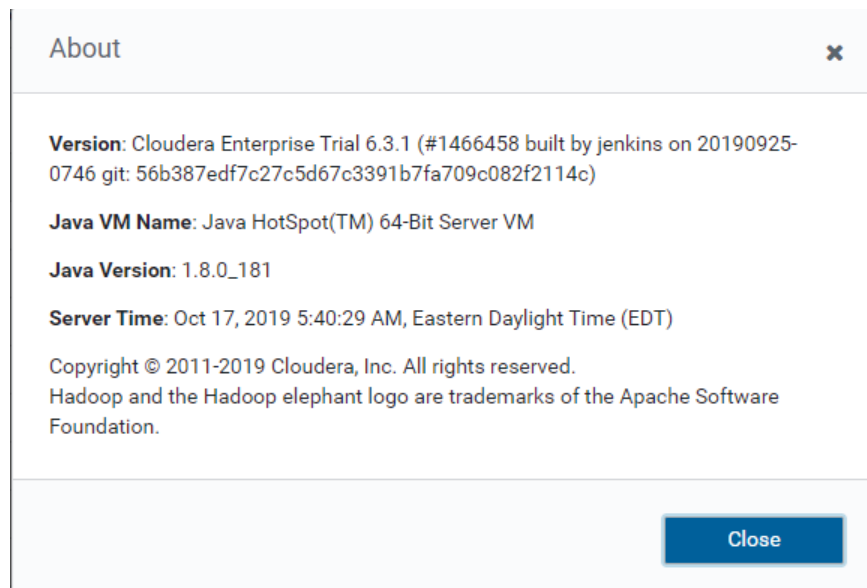


Figure 7 Upgraded CDH and JDK version

3.4.8 Remove the Previous CDH Version Packages and Refresh Symlinks

[Not required for CDH maintenance release upgrades.]

Skip this step if your previous installation or upgrade used parcels.

3.4.9 Finalize the HDFS Upgrade

To determine if you can finalize the upgrade, run important workloads and ensure that they are successful.

3.4.10 Complete Post-Upgrade Migration Steps

Several components require additional migrations steps after you complete the CDH upgrade:

1. Impala – Refer [Impala Upgrade Considerations](#)
2. Cloudera Search – After upgrading to CDH 6, you must re-index your collections, see [Re-Indexing Solr Collections After Upgrading to CDH 6](#).
3. Spark -- See [Apache Spark Post Upgrade Migration Steps](#).
4. MapReduce 1 to MapReduce 2 – See [Migrating from MapReduce 1 \(MRv1\) to MapReduce 2 \(MRv2\)](#)
5. Kafka

3.4.11 Exit Maintenance Mode

If you entered maintenance mode during this upgrade, exit maintenance mode.

On the **Home > Status** tab, click dropdown next to the cluster name and select **Exit Maintenance Mode**.

3.5 [CDH 6 Post-Upgrade Migration](#)

If you have deployed the HBase, Cloudera Search, or Spark services in a CDH 5.x cluster that you want to upgrade to CDH 6, see the following topics for additional post-upgrade steps: This process is out scope and request to refer official Cloudera Documentation.

1. [Impala](#)
2. [Re-Indexing Solr Collections](#)
3. [Apache Spark](#)
4. [MapReduce1 to MapReduce2](#)

3.6 [Manual Upgrade Steps](#)

3.7 [Troubleshooting](#)

A Technical support and resources

[Dell.com/support](https://dell.com/support) is focused on meeting customer needs with proven services and support.

[Storage technical documents and videos](#) provide expertise that helps to ensure customer success on Dell EMC storage platforms.

A.1 Related resources

List of documents and other assets that are referenced in the paper; include other resources that may be helpful.

[Cloudera Enterprise Upgrade Guide](#)

[Using CDH with Isilon Storage](#)

[Cloudera Product Compatibility for Dell EMC Isilon](#)

[OneFS Upgrades](#)

[DellEMC Isilon OneFS Upgrade Planning and Process Guide](#)

[Isilon OneFS8.2 HDFS Reference Guide](#)