

Upgrading Cloudera Data Services on premises on the Cloudera Embedded Container Service

Date published: 2023-12-16

Date modified: 2025-11-08



Legal Notice

© Cloudera Inc. 2025. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Upgrading Cloudera Data Services on premises using Cloudera Embedded Container Service.....4

- Pre-upgrade checklist.....4
- Pre-upgrade - Upgrading Cloudera Data Engineering service..... 8
 - Prerequisites for upgrading Cloudera Data Engineering service..... 8
 - Backing up Cloudera Data Engineering service using the docker image..... 11
- Cloudera Data Services on premises supported in-place upgrade paths.....12
- Upgrading Cloudera Manager..... 18
- Upgrading Cloudera Data Services on premises using Cloudera Embedded Container Service..... 18
- Post upgrade - Ozone Gateway validation.....26
- Post upgrade - Restoring Cloudera Data Engineering service.....27
 - Restoring a Cloudera Data Engineering service..... 27
 - Limitations of restoring Cloudera Data Engineering service.....30

Upgrading Cloudera Data Services on premises using Cloudera Embedded Container Service

Learn about upgrading Cloudera Data Services on premises with the Cloudera Embedded Container Service. You can upgrade Cloudera Data Services on premises from 1.5.4 or 1.5.5 versions to 1.5.5 SP1 version without uninstalling your existing installation.

Pre-upgrade checklist

There is a list of pre-upgrade checks that will run after the upgrade version has been chosen. This checklist verifies if your cluster is ready for upgrade.

About this task



Note: Pre-upgrade checklist is added in the Upgrade wizard of the Cloudera Manager UI in the Cloudera Data Services on premises 1.5.4 SP2 release and onwards.

In the Cloudera Manager UI, under Getting Started for Upgrade Step in the Upgrade wizard, select the repository URL for upgrade. A pre-upgrade checklist will appear that verifies that the hosts and services in your cluster are ready for the upgrade.

1. Select Repository

[Custom Repository](#)

Please ensure all the Data Lake clusters are running Cloudera Runtime 7.1.7.3 or greater

You are about to update CDP Private Cloud Data Services to version 1.5.5-b173. This is an update within the same version but to a newer build number.

All checks must pass (✔) before you can continue.

[Run All Checks Again](#)

> **Hosts Check**

> **Services Health Check**

✓ **Download Upgrade Validator**

The upgrade validator needs to be downloaded onto every host in the cluster.

[Download Upgrade Validator](#)

✓ **Control Plane Health Check**

The control plane health check verifies if longhorn, vault, and rke2 is ready for upgrade.

[Run Control Plane Health Check](#)

✓ **Docker Registry Health Check**

The docker registry health check verifies docker registry connection and checks if all required images are in the docker registry.

[Run Docker Registry Check](#)

All checks must pass (✔) before you can continue.

[Cancel](#) [Back](#) [Next](#)

- Click on Download Upgrade Validator, this will download the upgrade validator onto all your Cloudera Embedded Container Service hosts and is needed to run the Control Plane Health Check and Docker Registry Health Check.

- Once the Download Upgrade Validator is completed, the Control Plane Health Check and Docker Registry Health Check will automatically run.



Note: For custom docker registry setups, the docker registry health check will not run automatically. You must follow the instructions to download the new upgrade images into your custom registry first, then run the Docker Registry Health Check.

Here is the pre-upgrade checklist:

Checklist	Description
Hosts check	<p>This verifies the host health status, runs the host prerequisite inspections, and host warning inspections.</p> <ul style="list-style-type: none">• Host Health Status- This check verifies that there are no hosts in bad health or concerning health. It also checks for any stopped roles on the hosts.• Host Prerequisites Inspections - These are host inspections that must pass in order for you to proceed

Checklist	Description
	<p>to upgrade. Currently the prerequisite inspection includes:</p> <ul style="list-style-type: none"> EcsSystemConfigInspection - Checks to make sure that the <code>fs.inotify.max_user_instances</code> are set to the minimum required value 256 <p>This check MUST be fixed, and cannot be bypassed to continue install or upgrade.</p> <p>To fix this issue temporarily, you can perform the following steps:</p> <ol style="list-style-type: none"> 1. Login to host affected. 2. Enter the following command: <code>sysctl fs.inotify.max_user_instances=256</code> <p>For fixing this issue permanently, perform the following steps:</p> <ol style="list-style-type: none"> a. Login to each host affected. b. Verify that the current configuration is 128 by executing the following command: <pre>cat /proc/sys/fs/inotify/max_user_instances</pre> c. Edit the following file: <code>vi /etc/sysctl.conf</code> d. Add the following contents to the end and save: <pre>fs.inotify.max_user_instances=256</pre> e. Reload the config: <pre>sudo sysctl -p</pre> f. Verify if the config is updated, the following should return 256: <pre>cat /proc/sys/fs/inotify/max_user_instances</pre> EcsHostDnsInspection - Checks to make sure that there are less than three nameserver entries in the <code>/etc/resolv.conf</code> file, and checks the connections to the Cloudera Manager cluster and the Cloudera console. It also checks to see if <code>vault.localhost.localdomain</code>'s ping can be resolved. If not, it is likely that the host <code>/etc/nsswitch.conf</code> file is misconfigured. <p>If this inspection fails:</p> <ul style="list-style-type: none"> Check the <code>/etc/resolv.conf</code> and <code>/etc/nsswitch.conf</code> files and ensure that <code>/etc/resolv.conf</code> does not contain three or more nameservers, and that <code>/etc/nsswitch.conf</code> must contain <code>myhostname</code> under the <code>hosts</code> field.

Checklist	Description
	<ul style="list-style-type: none"> • Check to see the connections are resolved correctly. If the connection to the Cloudera console fails, check to see if your DNS wildcard is configured properly. • Host Warning Inspections - These are host inspections that are used to detect potential factors that can cause issues during an upgrade. Currently the warning inspections include: <ul style="list-style-type: none"> • SecuritySoftwareInspection - Checks to make sure that there are no security software processes running on the hosts in the cluster. • Upgrade Storage Inspection - Checks to make sure there is at least 100 GB of free space under /var/lib/ and 200 GB of free space under the docker data directory.
Services Health Check	This verifies that there are no services in bad or concerning health.
Download Upgrade Validator	This downloads the upgrade validator used to verify the control plane and docker registry health checks onto all the hosts in the cluster.
Control Plane Health Check	<p>This verifies the control plane is in a healthy state before upgrade. Here is the list of things it checks:</p> <ul style="list-style-type: none"> • Longhorn Health Check: This verifies that all the longhorn volumes are in a healthy, robust state. It also verifies that PVCs are bound. • Longhorn Engine Check: This verifies that the longhorn engine version matches the current longhorn manager version. • RKE2 Health Check: This verifies that the Kubernetes API server is reachable and the nodes are in a Ready and Schedulable state. • Pod Readiness Health Check: This verifies that all the pods in the kube-system, longhorn-system, vault-system, yunikorn, Kubernetes, ecs-webhooks, and CDP namespaces are in a Ready state. • Vault Health Check: This verifies that the vault-0 pod is running and the vault is unsealed.
Docker Registry Health Check	<p>This verifies that the selected docker registry is ready for upgrade:</p> <ul style="list-style-type: none"> • It verifies the connection to the docker registry by pulling an image. • For custom registry setups, it will also verify that the new required images stated in the manifest.json are present in your registry before upgrade.



Note: Fix any pre-check issues and run the check again to continue the upgrade.

Pre-upgrade - Upgrading Cloudera Data Engineering service

You can seamlessly upgrade a previous Cloudera Data Engineering service version to a new version.



Important:

- Upgrading Cloudera Data Engineering service from version 1.5.4 SP2 or earlier to Cloudera Data Engineering 1.5.5 does not support endpoint stability.
- After upgrading Cloudera Data Engineering from version 1.5.4 SP2 or earlier to 1.5.5, the Cloudera Data Engineering Services and Virtual Clusters that were created in the earlier version does not work in Cloudera Data Engineering 1.5.5. Cloudera recommends you not to use the old Cloudera Data Engineering Services and Virtual Clusters that were created before upgrade. Instead, create new Services and Virtual Clusters in Cloudera Data Engineering 1.5.5 and use them.
- After upgrading Cloudera Data Engineering, the upgraded Virtual Cluster retains the same base OS images that were used in the source Virtual Cluster before upgrade to ensure maximum compatibility particularly for jobs that depend on particular python and scala versions such as Spark jobs. For example, if an old Virtual Cluster is Redhat insecure based, then the new restored Virtual Cluster will also be Redhat insecure based only and if the old Virtual Cluster is security hardened based, then the new restored Virtual Cluster will also be security hardened based only. No automation path is supported from Redhat insecure to security hardened or vice versa.

Once you upgrade to Cloudera Data Engineering 1.5.5, the endpoints that you were using in the previous version are not supported. The Cloudera Data Engineering service endpoint migration process lets you migrate your resources, jobs, job run history, Spark jobs' logs, and event logs from your old cluster to the new cluster.

Prerequisites for upgrading Cloudera Data Engineering service

You must first download the docker image and create the cde-upgrade-util.properties file to back up the Cloudera Data Engineering service.

About this task



Important:

- In Cloudera Data Engineering 1.5.5 SP1 and earlier releases, if Cloudera Data Engineering is installed and you plan to enable SAML authentication, ensure that the Cloudera Data Engineering service is upgraded using LDAP first. Once the upgrade is successful, you can proceed to enable SAML authentication.
- Ensure that your cluster has enough resources to accommodate the increase in hardware requirements after the upgrade. For more information about hardware requirements, see [Additional resource requirements for Cloudera Data Engineering](#).
- If you created a Cloudera Data Engineering Virtual Cluster with Spark version 2.4.x, 3.3.x, or 3.4.x using Data Lake 7.1.9, and the Data Lake is later upgraded to version 7.3.1 (which supports only Spark 3.5), then the old Virtual Clusters created in Spark 2.4.x, 3.3.x, or 3.4.x will not be automatically restored as a Spark 3.5 Virtual Clusters after you upgrade Cloudera Data Services on premises to version 1.5.5 CHF1. You must manually create a new Virtual Cluster with Spark 3.5 and migrate the workloads. For instructions about migrating Cloudera Data Engineering Virtual Clusters to Spark 3.5, see [Migrating Virtual Clusters to Spark 3.5 after upgrading Cloudera Data Lake to 7.3.1](#).

Procedure

1. Login into the ECS Server host using SSH and create an external kubeconfig file. The following command assumes that your home directory, that is, ~/ is the working directory.

```
sed -e 's/certificate-authority-data/#&/' -e "s/server: ./server: https\n:\/\`hostname`:6443/" -e '/server/a \ \ \ insecure-skip-tls-verify: true' /etc/rancher/rke2/rke2.yaml > ~/kubeconfig && cat ~/kubeconfig
```

This command creates a file named kubeconfig in the working directory which is the external kubeconfig file.

2. Copy the CDP Credentials file named credentials of the DEAdmin user into the ECS Server host's working directory as follows:
 - a) In the Cloudera Data Platform (CDP) console, click the Management Console tile.
 - b) Click User Management and select the user.
 - c) Click Generate Access Key Download credentials file .
 - d) Copy the CDP Credentials file into the ECS Server host with the name credentials.
 - e) Verify if the credentials are present in the ECS Server host:

```
ls -l credentials
```

3. Set the environment variables in the ECS Server host by running the following command:

```
export PATH=$PATH:/opt/cloudera/parcels/ECS/installer/install/bin/linux/:/opt/cloudera/parcels/ECS/docker export KUBECONFIG=~/.kubeconfig
```

4. Download the dex-upgrade-utils docker image tarball. The file naming convention is dex-upgrade-utils-[***VERSION-NUMBER***]-[***BUILD-NUMBER***].tar.gz.



Important: To download the dex-upgrade-utils docker image tarball, see the following locations:

- [Cloudera Data Services on premises 1.5.5](#)
- [Cloudera Data Services on premises 1.5.5 SP1](#)

5. Load the downloaded docker image into the host machine docker runtime:

```
docker load < dex-upgrade-utils-[***VERSION-NUMBER***]-[***BUILD-NUMBER***].tar.gz
```

Example:

```
docker load < dex-upgrade-utils-1.20.1-b48.tar.gz
```

Sample output:

```
368243204766.dkr.ecr.us-west-2.amazonaws.com/cdp-private/cloudera/dex/dex-upgrade-utils:1.20.1-b48
```



Important: The version of the utility must be same as the version of the CDE control plane that you are upgrading to.

6. Create the required folders on the ECS Server host and copy the credentials and kubeconfig secret files.

```
mkdir /opt/backup-restore
export BASE_WORK_DIR=/opt/backup-restore

cd $BASE_WORK_DIR
mkdir backup secrets
chmod 775 backup/
```

7. Place the CDP credentials file of the *DEADMIN* user and *ADMINISTRATOR* kubeconfig file in the \$BASE_WORK_DIR/secrets directory.

```
cp ~/credentials secrets/
cp ~/.kubeconfig secrets/
```

8. Create the cde-upgrade-util.properties file as follows:

- a) Create the cde-upgrade-util.properties file and save it in the \$BASE_WORK_DIR directory.
- b) Update the following information in the cde-upgrade-util.properties file:

```
cdp_k8s_namespace:<CDP control plane k8s namespace>
cdp_endpoint:<CDP control plane endpoint>
credential_file_path:<Path to the DEAdmin user CDP credentials file>
de_admin_user:<DEAdmin user-id>
```

```

de_admin_password:<DEAdmin user's password must be in base64 encoded
format. Use the "echo -n [***PASSWORD***] | base64" command to encode
the password. >
tls_insecure:<Keep it true if you are using a self-signed certificate>
auto_unpause_jobs: <Specify it as "true" if you want to automatically re
sume the jobs that were paused during the backup phase. The jobs will be
resumed after you restore the CDE service.>
platform_type:ECS
use_stored_user:<(optional) Boolean property which can be TRUE or FALSE.
Use this property in conjunction with DO-AS described below.>
do_as:<(optional) if the value of USE_STORED_USER is set to TRUE, this v
alue is used as a fallback when the stored user is not valid. Otherwise,
this is directly used as job owner. If the USE_STORED_USER parameter i
s set to FALSE and no value is supplied in the DO_AS parameter, then no
validation will be performed for the job's username and it will be resto
red as it is.>

```

For example: The following options are the minimum recommended options that you must include in the cde-upgrade-util.properties file:

```

cdp_k8s_namespace=cdp
cdp_endpoint=https://console-cdp.apps.host-1.ecs-pvc1.kcloud.cloudera.
com
credential_file_path=/home/dex/.cdp/credentials
de_admin_user=cdpuser1
de_admin_password=VGZzdDEyMw==
tls_insecure=true
auto_unpause_jobs=true
platform_type=ECS
use_stored_user=false

```



Important:

- The cdp_k8s_namespace, cdp_endpoint, de_admin_user, and de_admin_password values must be updated based on your cluster.
- The de_admin_password password is the base64 encoded password of the de_admin_user. You can use `echo -n <pwd> | base64` to encode it.
- You must always set the value of the credential_file_path property as `/home/dex/.cdp/credentials` and must not be changed.



Warning: You can specify the `cdp_env_override:[***ENVIRONMENT-NAME***]` optional property in the `cde-upgrade-util.properties` file, if you want to change the environment of the CDE service that is being restored. But, if you change the environment during restore, it leads to loss of old spark jobs' logs and event logs that were there in old virtual clusters.

9. Make a note of the details of the Cloudera Data Engineering service that is being migrated. This information is required if you are using a CDP database that is external and is not accessible from the container which is running the cde-upgrade endpoint stability commands. Identify the cluster endpoint:
 - a. In the Cloudera Data Platform (CDP) console, click the Data Engineering tile. The Cloudera Data Engineering Home page displays.
 - b. Click Administration in the left navigation menu. The Administration page displays.
 - c. In the Services column on the left, click the Cluster Details icon corresponding to the Cloudera Data Engineering service whose endpoint you want to migrate.
 - d. Make a note of the CDE cluster ID.

Related Information

[Upgrading CDP on the Cloudera Embedded Container Service](#)

[Managing cluster resources using Quota Management](#)

Backing up Cloudera Data Engineering service using the docker image

You must run the docker image to take a backup of a Cloudera Data Engineering service. It takes backup of all the active virtual clusters in that Cloudera Data Engineering service. You can take backup of only one active Cloudera Data Engineering service at a time.

Before you begin

You must download the dex-upgrade-utils docker image and create the cde-upgrade-util.properties file before backing up jobs as described in the *Prerequisites for upgrading Cloudera Data Engineering Service with endpoint stability* section.



Warning: You must make sure to allocate sufficient downtime before you proceed further. If you start the backup procedure, you cannot create, edit, or run jobs in the existing Cloudera Data Engineering service and its associated virtual clusters until the backup is complete. The virtual clusters will be in the read-only mode after you backup the service and until you restore it.



Important: It is recommended that you copy the logs of the commands that are run from the terminal and save them on your machine. This helps you during debugging or raising a support ticket. You can also increase the terminal buffer size so that it does not throw away the logs and save the terminal logs of each command for reference.

Procedure

1. Set the following environment variables in the ECS Server host terminal:

```
export PATH=$PATH:/opt/cloudera/parcels/ECS/installer/install/bin/linux/:/
opt/cloudera/parcels/ECS/docker
export KUBECONFIG=~/.kubeconfig
export BASE_WORK_DIR=/opt/backup-restore
export BACKUP_OUTPUT_DIR=/home/dex/backup
```

2. Run the dex-upgrade-utils docker image on the host machine:

```
docker run \
-v [***KUBECONFIG_FILE_PATH***]:/home/dex/.kube/config:ro \
-v [***CDP_CREDENTIAL_FILE_PATH***]:/home/dex/.cdp/credentials:ro \
-v [***CDE-UPGRADE-UTIL.PROPERTIES_FILE_PATH***]:/opt/cde-backup-restore/
scripts/backup-restore/cde-upgrade-util.properties:ro \
-v [***LOCAL_BACKUP_DIRECTORY***]:$BACKUP_OUTPUT_DIR \
-e KUBECONFIG=/home/dex/.kube/config \
[***DOCKER_IMAGE_NAME***]:[***DOCKER_IMAGE_VERSION***] prepare-for-upgrade
-s [***CDE-CLUSTER-ID***] -o $BACKUP_OUTPUT_DIR
```





Important: All the paths to the right side of colon (:) in volume mounts, that is, paths inside the container are fixed paths and must not be changed. Here -s is the Cloudera Data Engineering service ID which is being backed up and -o is the backup output directory path in the container. The backup output directory value must always be \$BACKUP_OUTPUT_DIR and should not be changed.

Example:

```
docker run \
-v $BASE_WORK_DIR/secrets/kubeconfig:/home/dex/.kube/config:ro \
-v $BASE_WORK_DIR/secrets/credentials:/home/dex/.cdp/credentials:ro \
-v $BASE_WORK_DIR/cde-upgrade-util.properties:/opt/cde-backup-restore/s
cripts/backup-restore/cde-upgrade-util.properties:ro \
-v $BASE_WORK_DIR/backup:$BACKUP_OUTPUT_DIR \
-e KUBECONFIG=/home/dex/.kube/config \
```

```
docker-private.infra.cloudera.com/cloudera/dex/dex-upgrade-utils:1.20.1
prepare-for-upgrade -s cluster-c2dhkp22 -o $BACKUP_OUTPUT_DIR
```

To identify the *CDE-CLUSTER-ID*, do the following steps:

- a. In the Cloudera console, click the Data Engineering tile. The Cloudera Data Engineering home page displays.
- b. On the left navigation menu, click Administration.
- c.  In the Services column, click  for the Service for which you want to identify the Cluster ID.
- d. Note the CLUSTER ID displayed on the page.

Results

You have now taken the Cloudera Data Engineering service backup as a ZIP file. You can make a note of the Zip file name from the logs to use it while restoring the Cloudera Data Engineering service.

What to do next

You must now expand the resource pool, and then upgrade your Cloudera Data Platform (CDP) before you restore the Cloudera Data Engineering service. For information about configuring resource pool and capacity, see *Managing cluster resources using Quota Management*.



Important:

- During the restore operation, both old and the new Cloudera Data Engineering services use the same resources allocated to the existing Cloudera Data Engineering service. Hence, you must double the resource pool size using the Quota Management option. For example, if root.<env_pool>.cde.sales is the pool that is used for the old or existing Cloudera Data Engineering service, you must double the CPU and memory resources of this pool. Also, make sure that you have sufficient hardware when doubling the resource pool size. Consider the following conditions and plan whether to modify the resource pool size or not:
 - If the Cloudera Data Engineering service uses the default resource pool, that is root.<env_pool>.cde, then do not change the resource pool size.
 - If the Cloudera Data Engineering service uses a custom resource pool (for example, root.<env_pool>.cde.primary.secondary), the resource pool size of the last level (that is, secondary level in the example) must be doubled using the Quota Management option. The additional capacity required after doubling the last level's pool size is allocated from the levels above it, starting from the higher levels and progressing downward. In this example, when you double the secondary level (last level), the extra resource pool capacity required is initially added to the primary level pool. Then the newly added resource pool capacity is added to the secondary level pool, resulting in an overall doubling of the resource pool size of the last level.
 - The resource capacity at the Cloudera Data Engineering service and the Virtual Cluster level must not be changed. Modifying the pool size at the resource pool level is sufficient.
- In Cloudera Data Engineering 1.5.5 SP1 and higher versions, along with doubling the resource pool size, you must change the memory units of the pools to the corresponding binary memory unit, such as Ki, Mi, and Gi. For example, if the memory unit is G, change it to Gi.

Related Information

[Managing cluster resources using Quota Management](#)

[Prerequisites for upgrading Cloudera Data Engineering Service with endpoint stability](#)

Cloudera Data Services on premises supported in-place upgrade paths

Supported upgrade paths for Cloudera Data Services on premises.

The tables below detail the supported upgrade paths for upgrades to Cloudera Data Services on premises. Before upgrading, ensure that you select a compatible version of Cloudera Manager. See [Cloudera Manager Support Matrix](#).

For CDP Private Cloud Data Services

For supported versions of Operating system, Database, JDK versions, see <https://supportmatrix.cloudera.com/>.

Table 1: Upgrade paths for Cloudera Data Services on premises 1.5.5 SP1

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Fresh install of Cloudera Data Services on premises 1.5.5 SP1					
NA	NA	NA	7.13.1.501	7.3.1 Service Pack 3	
NA	NA	NA	7.13.1.501	7.1.7 SP3 CHF 10	
NA	NA	NA	7.13.1.501	7.1.9 Service Pack 1 Cumulative hotfix 11	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.5 SP1					
7.13.1.402	7.1.9 CHF5 7.1.7 SP3 CHF10	1.5.5-h103	7.13.1.501	7.1.7 SP3 CHF 10, Pack 1 Cumulative	7.1.9 Service Pack 1 Cumulative hotfix 11
7.13.1.402	7.3.1.100	1.5.5-h103	7.13.1.501	7.3.1 Service Pack 3	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.13.1 Cumulative hotfix 1	7.1.9 CHF7 7.1.9 SP1 CHF4 7.1.7 SP3	1.5.4 SP2	7.13.1.501	7.1.7 SP3 CHF 10, Pack 1 Cumulative Service Pack 3	7.1.9 Service Pack 1 Cumulative hotfix 11, 7.3.1 Service Pack 3
7.11.3 Cumulative hotfix 7 and 7.11.3 Cumulative hotfix 9.1	7.1.7 SP3 CHF 10 or 7.1.9 SP1	1.5.4 CHF3	7.13.1.501	7.1.7 SP3 CHF 10, Pack 1 Cumulative Service Pack 3	7.1.9 Service Pack 1 Cumulative hotfix 11, 7.3.1 Service Pack 3

Table 2: Upgrade paths for Cloudera Data Services on premises 1.5.5

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Fresh install of Cloudera Data Services on premises 1.5.5					
NA	NA	NA	7.13.1 Cumulative hotfix 3	7.1.9 Service Pack 1 Cumulative hotfix 7	
NA	NA	NA	7.13.1 Cumulative hotfix 3	7.1.7 SP3 CHF 10	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.5					
Cloudera Manager 7.11.3 CHF6	7.1.9 CHF6	1.5.4	7.13.1 Cumulative hotfix 3	7.1.9 Service Pack 1 Cumulative hotfix 7	
Cloudera Manager 7.11.3 CHF6	7.1.8 CHF22	1.5.4	7.13.1 Cumulative hotfix 3	7.1.9 Service Pack 1 Cumulative hotfix 7	
Cloudera Manager 7.11.3 CHF6	7.1.7 SP3	1.5.4	7.13.1 Cumulative hotfix 3	7.1.7 SP3 CHF 10	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.11.3 Cumulative hotfix 11	7.1.9 Service Pack 1 Cumulative hotfix 7	1.5.4 SP1	7.13.1 Cumulative hotfix 3	7.1.9 Service Pack 1 Cumulative hotfix 7	
7.13.1 Cumulative hotfix 1	7.1.7 SP3 CHF 10	1.5.4 SP2	7.13.1 Cumulative hotfix 3	7.1.7 SP3 CHF 10	

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
7.11.3 Cumulative hotfix 7 and 7.11.3 Cumulative hotfix 9.1	7.1.7 SP3 CHF 10	1.5.4 CHF1 and CHF3	7.13.1 Cumulative hotfix 3	7.1.7 SP3 CHF 10	
7.11.3 Cumulative hotfix 1	7.1.9 Service Pack 1 Cumulative hotfix 7	1.5.3 SP1 and 1.5.3	7.13.1 Cumulative hotfix 3	7.1.9 Service Pack 1 Cumulative hotfix 7	

Table 3: Upgrade paths for Cloudera Data Services on premises 1.5.4 SP2

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Fresh install of Cloudera Data Services on premises 1.5.4 SP2					
NA	NA	NA	7.13.1 Cumulative hotfix 1	7.1.9 Cumulative hotfix 7	
NA	NA	NA	7.13.1 Cumulative hotfix 1	7.1.9 Service Pack 1	
NA	NA	NA	7.13.1 Cumulative hotfix 1	7.1.7 SP3	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.4 SP2					
7.11.3 Cumulative hotfix 6	7.1.9 Cumulative hotfix 6	1.5.4	7.13.1 Cumulative hotfix 1	7.1.9 Cumulative hotfix 7	
7.11.3 Cumulative hotfix 6	7.1.8 Cumulative hotfix 22 or higher	1.5.4	7.13.1 Cumulative hotfix 1	7.1.9 Service Pack 1 Cumulative hotfix 4	
7.11.3 Cumulative hotfix 6	7.1.7 SP3	1.5.4	7.13.1 Cumulative hotfix 1	7.1.7 SP3	
7.11.3 Cumulative hotfix 4	7.1.7 SP2, 7.1.8 CHF 17, 7.1.9.4	1.5.3	7.13.1 Cumulative hotfix 1	7.1.7 SP 3, 7.1.9 Cumulative hotfix 7, or 7.1.9 SP1 CHF 4	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.11.3 Cumulative hotfix 4	7.1.9 Cumulative hotfix 7	1.5.3	7.13.1 Cumulative hotfix 1	7.1.9 Cumulative hotfix 7	
7.11.3 Cumulative hotfix 4	7.1.9 Service Pack 1	1.5.3	7.13.1 Cumulative hotfix 1	7.1.9 Service Pack 1	
7.11.3 Cumulative hotfix 4	7.1.7 SP3	1.5.3	7.13.1 Cumulative hotfix 1	7.1.7 SP3	
7.11.3 Cumulative hotfix 1	7.1.7 SP3	1.5.2	7.13.1 Cumulative hotfix 1	7.1.7 SP3	

Table 4: Upgrade paths for Cloudera Data Services on premises 1.5.4 SP1

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Fresh install of Cloudera Data Services on premises 1.5.4 SP1					
NA	NA	NA	7.11.3 Cumulative hotfix 11	7.1.9 Cumulative hotfix 7	
NA	NA	NA	7.11.3 Cumulative hotfix 11	7.1.9 Service Pack 1 CHF 4	
NA	NA	NA	7.11.3 Cumulative hotfix 11	7.1.7 SP3	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.4 SP1					
7.11.3 Cumulative hotfix 6	7.1.9 Cumulative hotfix 6	1.5.4	7.11.3 Cumulative hotfix 11	7.1.9 Cumulative hotfix 7	
7.11.3 Cumulative hotfix 6	7.1.8 Cumulative hotfix 22 or higher	1.5.4	7.11.3 Cumulative hotfix 11	7.1.9 Service Pack 1 CHF4	
7.11.3 Cumulative hotfix 6	7.1.8 Cumulative hotfix 22 or higher and 7.1.7 SP3	1.5.4	7.11.3 Cumulative hotfix 11	7.1.7 SP3	

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
7.11.3 Cumulative hotfix 4	7.1.9.4, 7.1.7 SP2, 7.1.8 CHF 17	1.5.3	7.11.3 Cumulative hotfix 11	7.1.7 SP3, 7.1.9 Cumulative hotfix 7, 7.1.9 Service Pack 1 CHF4	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.11.3 Cumulative hotfix 4	7.1.9 Cumulative hotfix 7	1.5.3	7.11.3 Cumulative hotfix 6	7.1.9 Cumulative hotfix 7	
7.11.3 Cumulative hotfix 4	7.1.9 Service Pack 1	1.5.3	7.11.3 Cumulative hotfix 6	7.1.9 Service Pack 1	
7.11.3 Cumulative hotfix 4	7.1.7 SP3	1.5.3	7.11.3 Cumulative hotfix 6	7.1.7 SP3	
7.11.3 Cumulative hotfix 1	7.1.7 SP3	1.5.2	7.11.3 Cumulative hotfix 6	7.1.7 SP3	

Table 5: Upgrade paths for Cloudera Data Services on premises 1.5.4

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Fresh install of Cloudera Data Services on premises 1.5.4					
NA	NA	NA	7.11.3 Cumulative hotfix 6	7.1.9 Cumulative hotfix 6	
NA	NA	NA	7.11.3 Cumulative hotfix 6	7.1.8 Cumulative hotfix 22 or higher	
NA	NA	NA	7.11.3 Cumulative hotfix 6	7.1.7 SP3	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.4					
7.11.3 Cumulative hotfix 4	7.1.9.4	1.5.3	7.11.3 Cumulative hotfix 6	7.1.9 Cumulative hotfix 6	
7.7.3 Latest cumulative hotfix	7.1.8 Cumulative hotfix 17	1.5.3	7.11.3 Cumulative hotfix 6	7.1.8 Cumulative hotfix 22 or higher	
7.11.3 Cumulative hotfix 4	7.1.7 SP2	1.5.3	7.11.3 Cumulative hotfix 6	7.1.7 SP3	
7.11.3 Cumulative hotfix 1	7.1.7 SP2, 7.1.8 Cumulative hotfix 11, 7.1.9	1.5.2	7.11.3 Cumulative hotfix 6	7.1.9 Cumulative hotfix 6, 7.1.8 Cumulative hotfix 22 or higher, and 7.1.7 SP3	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.11.3 Cumulative hotfix 4	7.1.9 Cumulative hotfix 6	1.5.3	7.11.3 Cumulative hotfix 6	7.1.9 Cumulative hotfix 6	
7.11.3 Cumulative hotfix 4	7.1.8 Cumulative hotfix 22 or higher	1.5.3	7.11.3 Cumulative hotfix 6	7.1.8 Cumulative hotfix 22 or higher	
7.11.3 Cumulative hotfix 4	7.1.7 SP3	1.5.3	7.11.3 Cumulative hotfix 6	7.1.7 SP3	
7.11.3 Cumulative hotfix 1	7.1.8 Cumulative hotfix 22 or higher	1.5.2	7.11.3 Cumulative hotfix 6	7.1.8 Cumulative hotfix 22 or higher	

Table 6: Upgrade paths for Cloudera Data Services on premises 1.5.3

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Greenfield deployment of Cloudera Data Services on premises 1.5.3					
NA	NA	NA	7.11.3 Cumulative hotfix 4	7.1.9 Cumulative hotfix 3	
NA	NA	NA	7.11.3 Cumulative hotfix 4	7.1.8 Cumulative hotfix 19 or higher	
NA	NA	NA	7.11.3 Cumulative hotfix 4	7.1.7 SP2	

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.2					
7.11.3 Cumulative hotfix 1	7.1.9 Cumulative hotfix 3	1.5.2	7.11.3 Cumulative hotfix 4	7.1.9 Cumulative hotfix 3	
7.7.3 Latest cumulative hotfix	7.1.8 Cumulative hotfix 19 or higher	1.5.2	7.11.3 Cumulative hotfix 4	7.1.8 Cumulative hotfix 19 or higher	
7.7.1 Latest cumulative hotfix	7.1.8 Cumulative hotfix 19 or higher	1.5.2	7.11.3 Cumulative hotfix 4	7.1.8 Cumulative hotfix 19 or higher	
7.6.7 Latest cumulative hotfix	7.1.7 SP2	1.5.2	7.11.3 Cumulative hotfix 4	7.1.7 SP2	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.11.3 Cumulative hotfix 1	7.1.9 Cumulative hotfix 3	1.5.2	7.11.3 Cumulative hotfix 4	7.1.9 Cumulative hotfix 3	
7.11.3 Cumulative hotfix 1	7.1.8 Cumulative hotfix 19 or higher	1.5.2	7.11.3 Cumulative hotfix 4	7.1.8 Cumulative hotfix 19 or higher	
7.11.3 Cumulative hotfix 1	7.1.7 SP2	1.5.2	7.11.3 Cumulative hotfix 4	7.1.7 SP2	
7.10.1	7.1.8 Cumulative hotfix 19 or higher	1.5.1	7.11.3 Cumulative hotfix 4	7.1.8 Cumulative hotfix 19 or higher	
7.10.1	7.1.7 SP2	1.5.1	7.11.3 Cumulative hotfix 4	7.1.7 SP2	

Table 7: Upgrade paths for Cloudera Data Services on premises 1.5.2

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Greenfield deployment of Cloudera Data Services on premises 1.5.2					
NA	NA	NA	7.11.3 Cumulative hotfix 1	7.1.9	
NA	NA	NA	7.11.3 Cumulative hotfix 1	7.1.9 Cumulative hotfix 1	
NA	NA	NA	7.11.3 Cumulative hotfix 1	7.1.8 CHF11 or higher	
NA	NA	NA	7.11.3 Cumulative hotfix 1	7.1.7 SP2	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.2					
7.11.3	7.1.9	NA	7.11.3 Cumulative hotfix 1	7.1.9	
7.11.3	7.1.9 Cumulative hotfix 1	NA	7.11.3 Cumulative hotfix 1	7.1.9 Cumulative hotfix 1	
7.7.3 Latest cumulative hotfix	7.1.8 Cumulative hotfix 11 or higher	NA	7.11.3 Cumulative hotfix 1	7.1.8 Cumulative hotfix 11 or higher	
7.7.1 Latest cumulative hotfix	7.1.8 Cumulative hotfix 11 or higher	NA	7.11.3 Cumulative hotfix 1	7.1.8 Cumulative hotfix 11 or higher	
7.6.7 Latest cumulative hotfix	7.1.7 SP2	NA	7.11.3 Cumulative hotfix 1	7.1.7 SP2	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.10.1	7.1.7 SP2	1.5.1	7.11.3 Cumulative hotfix 1	7.1.7 SP2	
7.10.1	7.1.8 Cumulative hotfix 11 or higher	1.5.1	7.11.3 Cumulative hotfix 1	7.1.8 Cumulative hotfix 11 or higher	
7.9.5	7.1.8 Cumulative hotfix 11 or higher	1.5.0	7.11.3 Cumulative hotfix 1	7.1.8 Cumulative hotfix 11 or higher	
7.9.5	7.1.8	1.5.0	7.11.3 Cumulative hotfix 1	7.1.8	
7.9.5	7.1.7 SP2	1.5.0	7.11.3 Cumulative hotfix 1	7.1.7 SP2	



Important: Upgrading from Cloudera Manager 7.7.3 version to Cloudera Manager 7.10.1 is currently not supported. Note that, Cloudera Manager 7.7.3 version is supported on Python 3 and Cloudera Manager 7.10.1 supports only Python 2. Cloudera Data Services on premises 1.5.1 support for users using Cloudera Manager 7.7.3 is intended to be made available in the future with a new version of Cloudera Manager that has support for both Python 3 and 2 versions respectively.

Table 8: Upgrade paths for Cloudera Data Services on premises 1.5.1

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Greenfield deployment of Cloudera Data Services on premises 1.5.1					
NA	NA	NA	7.10.1	7.1.7 SP2	
NA	NA	NA	7.10.1	7.1.8 CHF4	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.1					
7.6.7	7.1.7 SP2	NA	7.10.1	7.1.8 CHF4 ¹	
7.7.1	7.1.8	NA	7.10.1	7.1.8 CHF4 ²	
7.6.7	7.1.7 SP2	NA	7.10.1	7.1.8	
7.7.1	7.1.8	NA	7.10.1	7.1.8	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.8.1	7.1.7 SP1	1.4.1	7.10.1	7.1.7 SP1	
7.8.1	7.1.8	1.4.1	7.10.1	7.1.8	
7.9.5	7.1.7 SP2	1.5.0	7.10.1	7.1.7 SP2	
7.9.5	7.1.7 SP1	1.5.0	7.10.1	7.1.7 SP1	
7.9.5	7.1.8	1.5.0	7.10.1	7.1.8	



Important: Upgrading from Cloudera Manager 7.7.3 version to Cloudera Manager 7.9.5 is currently not supported. Note that, Cloudera Manager 7.7.3 version is supported on Python 3 and Cloudera Manager 7.9.5 supports only Python 2. Cloudera Data Services on premises 1.5.0 support for users using Cloudera Manager 7.7.3 is intended to be made available in the future with a new version of Cloudera Manager that has support for both Python 3 and 2 versions respectively.

Table 9: Upgrade paths for Cloudera Data Services on premises 1.5.0

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
Greenfield deployment of Cloudera Data Services on premises 1.5.0					
NA	NA	NA	7.9.5	7.1.7 SP2	
NA	NA	NA	7.9.5	7.1.7 SP1	
NA	NA	NA	7.9.5	7.1.8	
Using Cloudera Base on premises, but new to Cloudera Data Services on premises 1.5.0					
7.6.7	7.1.7 SP2	NA	7.9.5	7.1.7 SP2	
7.6.1	7.1.7 SP1	NA	7.9.5	7.1.7 SP1	

¹ Upgrade from 7.1.7 SP2 to 7.1.8 CHF4 is not a mandatory upgrade. In 7.1.8 CHF4, you can install Ozone as a parcel.

² Upgrade from 7.1.8 to 7.1.8 CHF4 is not a mandatory upgrade. In 7.1.8 CHF4, you can install Ozone as a parcel.

Source version			Target version		
Cloudera Manager	Cloudera Base on premises	Cloudera Data Services on premises	Cloudera Manager	Cloudera Base on premises	
7.7.1	7.1.8	NA	7.9.5	7.1.8	
Using Cloudera Data Services on premises and wanting to upgrade to the latest version without upgrading Cloudera Base on premises					
7.8.1	7.1.7 SP1	1.4.1	7.9.5	7.1.7 SP1	
7.8.1	7.1.8	1.4.1	7.9.5	7.1.8	
7.6.5	7.1.7 SP1	1.4.0-H1	7.9.5	7.1.7 SP1	
7.6.5	7.1.7	1.4.0-H1	7.9.5	7.1.7	

For information on the pre-requisites on the upgrade, see [Prerequisites for upgrading Cloudera Data Engineering service](#) on page 8 and [Backing up Cloudera Data Engineering service using the docker image](#) on page 11

For information on the upgrade, see [Upgrading Cloudera Data Services on premises using Cloudera Embedded Container Service](#) on page 18.

For information on the post-upgrade, see [Restoring a Cloudera Data Engineering service](#) on page 27.

Upgrading Cloudera Manager

You must use Cloudera Manager version 7.13.1.501 to install or upgrade to Cloudera Data Services on premises 1.5.5 SP1.

If you already have a Cloudera Base on premises cluster set up using an earlier version of Cloudera Manager, you must first upgrade the Cloudera Manager version to Cloudera Manager 7.13.1.501 before proceeding with the Cloudera Data Services on premises update.

Related Information

[Upgrading Cloudera Manager](#)

Upgrading Cloudera Data Services on premises using Cloudera Embedded Container Service

Learn about upgrading Cloudera Data Services on premises with the Cloudera Embedded Container Service. You can upgrade Cloudera Data Services on premises from 1.5.4 or 1.5.5 versions to 1.5.5 SP1 version without uninstalling your existing installation.

Before you begin

- Before upgrading to Cloudera Data Services on premises 1.5.5 SP1 or higher versions with Data Lake 7.3.1, perform the actions described in [Cloudera Data Engineering Prerequisites](#) to ensure compliance with the [Spark support matrix](#).
- Review the [Software Support Matrix for Cloudera Embedded Container Service](#).
- The Docker registry that is configured with the cluster must remain the same during the upgrade process. If Cloudera Data Services on premises 1.5.4 was installed using the public Docker registry, Cloudera Data Services on premises 1.5.5 SP1 should also use the public Docker registry, and not be configured to use the embedded Docker registry. To use a different configuration for the Docker registry, you must perform a new installation of Cloudera Data Services on premises.
- To upgrade Cloudera Data Engineering, make sure that you have performed [Pre-upgrade - Upgrading Cloudera Data Engineering service](#) and taken a [backup of the Cloudera Data Engineering service](#).

- If you have expanded the cdp-embedded-db volume from Longhorn UI after the initial installation of your Data Services on premises cluster, then you must perform the workaround steps before planning your upgrade to Cloudera Data Services on premises 1.5.5 Service Pack 1 to avoid a potential upgrade failure. For more information about the workaround details, see [Overcoming a possible upgrade failure](#).

About this task



Note: Cloudera Embedded Container Service services will be unavailable to users for a period of time during this upgrade procedure. However, you should not stop the Cloudera Embedded Container Service cluster prior to upgrade. Upgrade requires the Cloudera Embedded Container Service cluster to be running and in a healthy state.



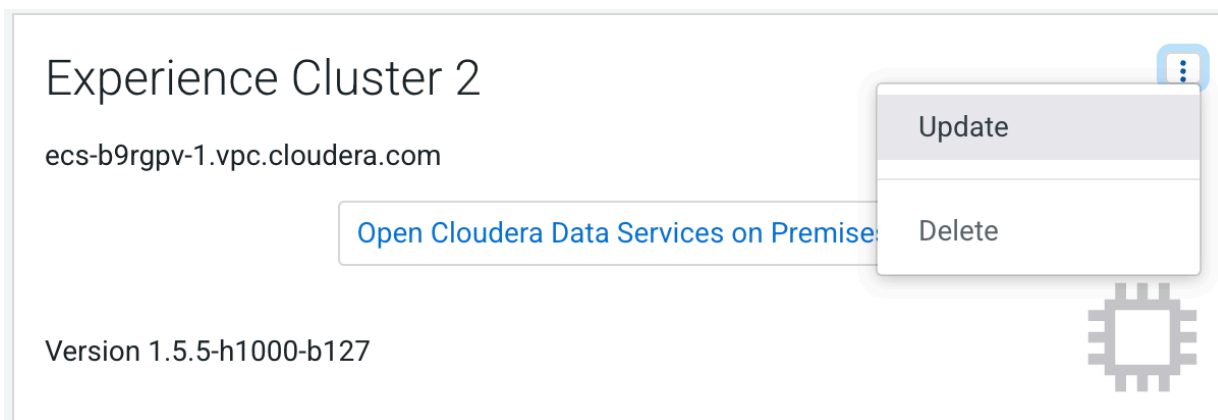
Important:

RHEL 7.x support on Cloudera Embedded Container Service has been removed in Cloudera Data Services on premises 1.5.5 and higher versions. If you are running RHEL 7.x, you must upgrade to a higher version before upgrading.

Procedure

1.

In Cloudera Manager, navigate to Cloudera on premises and click the  icon, then click Update.



2. On the Getting Started page, you can select the Install method - Air Gapped or Internet and proceed.

Internet install method

- 1 Getting Started
- 2 Collect Information
- 3 Install Parcels
- 4 Update Data Services
- 5 Summary

Getting Started

This wizard provides step-by-step guidance for updating Cloudera Data Services on Premises.

Visit the [Cloudera on Premises](#) documentation for more information.

Current Version
1.5.5-h1000-b127

Install Method
☒ Internet ☐ Air Gapped

1. Select Repository

Please ensure all the Data Lake clusters are running Cloudera Runtime 7.1.7.3 or greater

You are about to update Cloudera Data Services on Premises to version **1.5.5-h1000-b130**. This is an update within the **same version** but to a newer build number.

All checks must pass before you can continue.

> Hosts Check

> Services Health Check

> Download Upgrade Validator

The upgrade validator needs to be downloaded onto every host in the cluster.

> Control Plane Health Check

The control plane health check verifies if longhorn, vault, and rke2 is ready for upgrade.

Air Gapped install method

- 1 Getting Started
- 2 Collect Information
- 3 Install Parcels
- 4 Update Data Services
- 5 Summary

Getting Started

This wizard provides step-by-step guidance for updating Cloudera Data Services on Premises.

Visit the [Cloudera on Premises](#) documentation for more information.

Current Version
1.5.5-h1000-b127

Install Method
☐ Internet ☒ Air Gapped

Installing via a local mirror with an http server. You will need to setup a full mirror of Cloudera's repositories via a temporary http server within the perimeter network of all hosts.

1. Download everything under <https://archive.cloudera.com/p/cdp-pvc-ds/latest>

```
$ wget -l 0 --recursive --no-parent -e robots=off -nH --cut-dirs=2 --reject="index.html*" -t 10 https://<username>:<password>@archive.cloudera.com/p/cdp-pvc-ds/latest
```

2. Modify the file manifest.json inside the downloaded directory, change "http_url": "." to "http_url": "http://your_local_repo/cdp-pvc-ds/latest"

3. Mirror the downloaded directory to your local http server, e.g. http://your_local_repo/cdp-pvc-ds/latest

4. Add http://your_local_repo/cdp-pvc-ds/latest to your [Custom Repository](#) settings and select it from the dropdown below.

5. Select Repository

Please ensure all the Data Lake clusters are running Cloudera Runtime 7.1.7.3 or greater

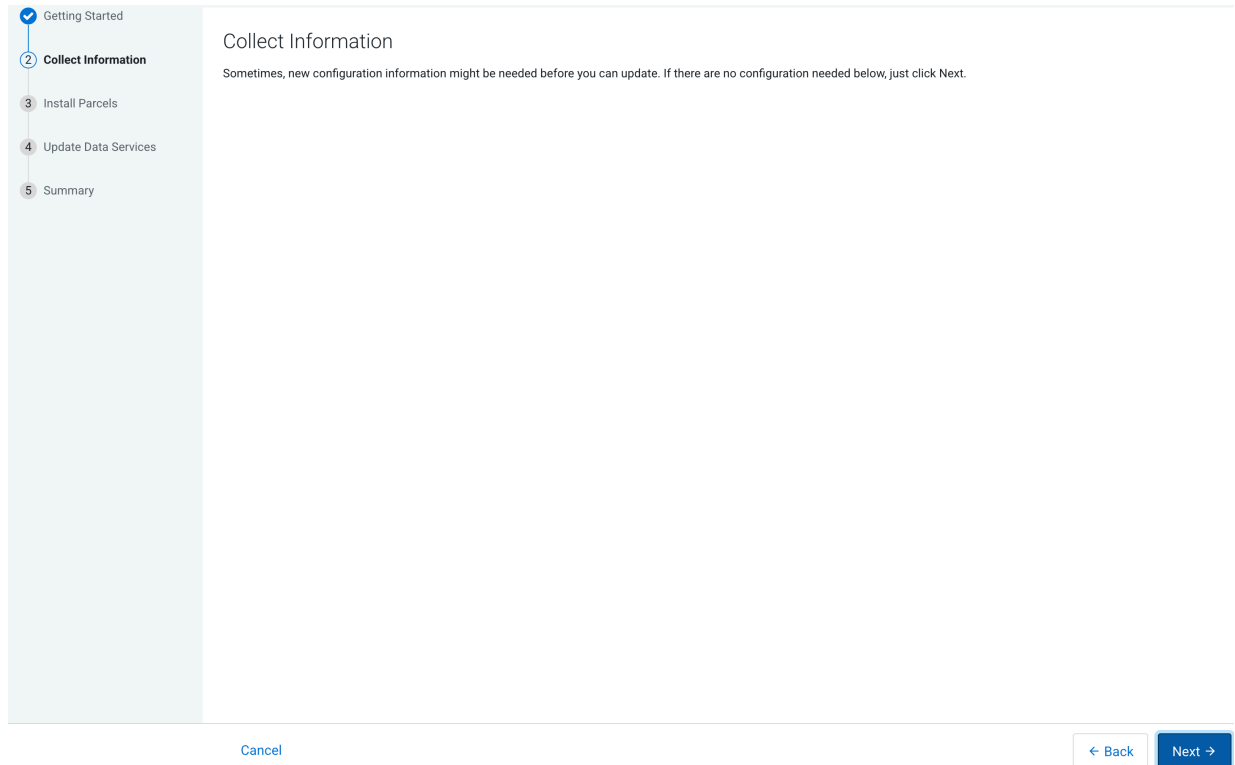
You are about to update Cloudera Data Services on Premises to version **1.5.5-h1000-b130**. This is an update within the **same version** but to a newer build number.

All checks must pass before you can continue.

> Hosts Check

Click Next.

3. On the Collect Information page, click Next.



The screenshot shows a multi-step wizard interface. On the left, a vertical sidebar contains five steps: 'Getting Started' (completed with a checkmark), '2 Collect Information' (current step, highlighted with a blue circle), '3 Install Parcels', '4 Update Data Services', and '5 Summary'. The main content area is titled 'Collect Information' and contains the text: 'Sometimes, new configuration information might be needed before you can update. If there are no configuration needed below, just click Next.' At the bottom of the sidebar is a 'Cancel' link. At the bottom right of the main area are two buttons: '← Back' and 'Next →'.

4. On the Install Parcels page, click Next.

5. On the Update Data Services page, you can see the progress of your upgrade. Post Upgrade Validation is run and these are a set of health checks added as the last step in the upgrade to verify that the upgrade was successful and the cluster is healthy. This includes verifying that the services are healthy, the hosts are upgraded to the new version, and that the control plane is in a healthy state after upgrade. For the control plane to be determined as

good health, it will verify that RKE2, Longhorn, and Vault are healthy. It will also check that the critical pods in the control plane namespace are up and running.

Update Cloudera Data Services on Premises (cdp)

Getting Started

Collect Information

Install Parcels

Update Data Services

Summary

Update Data Services

Upgrade Cluster Command

Status Running Context [Experience Cluster 2](#) Sep 30, 3:10:50 PM [Abort](#)

Completed 6 of 9 step(s).

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

> Execute pre upgrade pre stop upgrade handlers.	Experience Cluster 2	Sep 30, 3:10:50 PM	0ms
> Execute pre-upgrade upgrade handlers.	Experience Cluster 2	Sep 30, 3:10:50 PM	1ms
> Activating parcel	Experience Cluster 2	Sep 30, 3:10:50 PM	236ms
Waiting for Cloudera Manager Agents to detect release: ECS 1.5.5-h1000-b130.		Sep 30, 3:10:50 PM	15.31 s
> Converting configuration parameters	Experience Cluster 2	Sep 30, 3:11:05 PM	22ms
> Execute post upgrade pre validation upgrade handlers.	Experience Cluster 2	Sep 30, 3:11:05 PM	0ms
> Starting all services in the upgraded cluster.	Experience Cluster 2	Sep 30, 3:11:05 PM	Abort
<input type="radio"/> Execute command Post Platform Upgrade on service ECS-2			
<input type="radio"/> Execute post upgrade post validation upgrade handlers.			

[Cancel](#)

[Back](#)[Next](#)

6. Click Next after the upgrade is complete .

Note: The upgrade might occasionally fail with error messages or conditions such as the following:

- Error message: During the following step: Execute command Install Tolerations Webhook on service ECS-3 the Upgrade progress page mentions a failure waiting for kube-proxy to come up.

Workaround:

- a. Log in using ssh to one of the Cloudera Embedded Container Service Server nodes and run the following command:

```
/var/lib/rancher/rke2/bin/kubectl get nodes
```

The output looks similar to the following:

NAME	STATUS	ROLES
ecs-abc-1.vpc.myco.com 4h50m v1.21.8+rke2r2	Ready	control-plane,etcd,master
ecs-abc-2.vpc.myco.com 4h48m v1.20.8+rke2r1	NotReady	<none>
ecs-abc-3.vpc.myco.com 4h48m v1.21.8+rke2r2	Ready	<none>
ecs-abc-4.vpc.myco.com 4h48m v1.20.8+rke2r1	NotReady	<none>
ecs-abc-5.vpc.myco.com 4h48m v1.20.8+rke2r1	NotReady	<none>

If any of the version numbers in the last column are lower than the expected version, reboot those nodes. (For example, v1.20.8 in the output above.)

- b. In the Command Output window, in the step that failed, click Resume.
- Upgrade hangs on the Execute command Post upgrade configuration on service ECS step for more than an hour.

Workaround:

- a. Log in to one of the Cloudera Embedded Container Service server nodes and run the following command:

```
kubectl get nodes
```

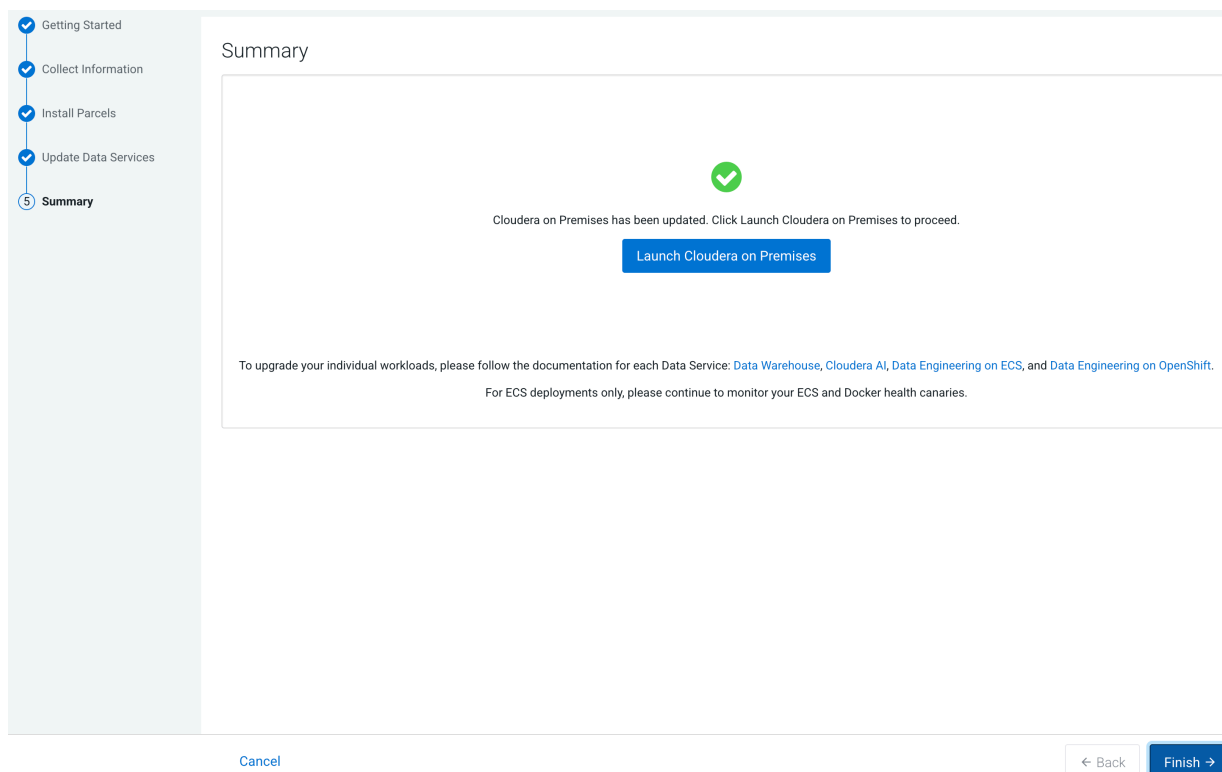
The output looks similar to the following:

NAME	STATUS	ROLES
ecs-abc-1.vpc.myco.com 3h47m v1.21.11+rke2r1	Ready	control-plane,etcd,master
ecs-abc-2.vpc.myco.com 3h45m v1.21.8+rke2r2	NotReady	<none>
ecs-abc-3.vpc.myco.com 3h45m v1.21.8+rke2r2	NotReady	<none>
ecs-abc-4.vpc.myco.com 3h45m v1.21.8+rke2r2	NotReady	<none>

If you any nodes display a status of NotReady, click the Abort option in the command output window.

- b. Reboot all nodes showing NotReady.
- c. Check the node status again as shown above. After all the nodes show Ready, click the Resume option in the command output window to continue with the upgrade.

7. After the upgrade is complete, the Summary page appears. You can now Launch Cloudera on premises from here.



If you see a Longhorn Health Test message about a degraded Longhorn volume, wait for the cluster repair to complete.

Or you can navigate to the Cloudera Data Services on premises page and click Open Cloudera on premises.

Cloudera Data Services on premises opens in a new window.

- If the upgrade stalls, do the following:

1. Check the status of all pods by running the following command on the server node:

```
export PATH=$PATH:/opt/cloudera/parcels/ECS/installer/install/bin/linux/
:/opt/cloudera/parcels/ECS/docker
export KUBECONFIG=~/.kubeconfig

kubectl get pods --all-namespaces
```

2. If there are any pods stuck in "Terminating" state, then force terminate the pod using the following command:

```
kubectl delete pods <NAME OF THE POD> -n <NAMESPACE> --grace-period=0 -f
orce
```

If the upgrade still does not resume, continue with the remaining steps.

3. If there are any pods in the "Pending" state, then you can try to reschedule the pods in the "Pending state" by restarting the yunikorn-scheduler. Run the following commands to restart yunikorn-scheduler:

```
kubectl get pods -n yunikorn

kubectl get deploy -n yunikorn

kubectl scale --replicas=0 -n yunikorn deployment/yunikorn-scheduler

kubectl get deploy -n yunikorn
```



```
kubectl scale --replicas=1 -n yunikorn deployment/yunikorn-scheduler
kubectl get deploy -n yunikorn
```

4. In the Admin Console, go to the service and click Web UI Storage UI.

The Longhorn dashboard opens.

5. Check the "In Progress" section of the dashboard to see whether there are any volumes stuck in the attaching/detaching state in. If a volume is that state, reboot its host.
6. In the LongHorn UI, go to the Volume tab and check if any of the volumes are in the "Detached" state. If any are in the "Detached" state, then restart the associated pods or reattach them to the host manually.

- You may see the following error message during the Upgrade Cluster > Reapplying all settings > kubectl-patch :

```
kubectl rollout status deployment/rke2-ingress-nginx-controller -n kube-
system --timeout=5m
error: timed out waiting for the condition
```

If you see this error, do the following:

1. Check whether all the Kubernetes nodes are ready for scheduling. Run the following command from the ECS Server node:

```
kubectl get nodes
```

You will see output similar to the following:

```
NAME STATUS ROLES AGE VERSION
<node1> Ready,SchedulingDisabled control-plane,etcd,master 103m v1.21.
11+rke2r1
<node2> Ready <none> 101m v1.21.11+rke2r1
<node3> Ready <none> 101m v1.21.11+rke2r1
<node4> Ready <none> 101m v1.21.11+rke2r1
```

2. Run the following command from the ECS Server node for the node showing a status of SchedulingDisabled:

```
kubectl uncordon <node1>
```

You must add the NODENAME to the end of the command.

You will see output similar to the following:

```
<node1>node/<node1> uncordoned
```

3. Scale down and scale up the rke2-ingress-nginx-controller pod by running the following command on the ECS Server node:

```
kubectl delete pod rke2-ingress-nginx-controller-<pod number> -n kube-s
ystem
```

4. Resume the upgrade.

- If a new release-dwx-server pod is unable to start because of an existing release-dwx-server pod failing to start:
 - Delete the pod manually by executing the following command:

```
kubectl delete -n cdp pod cdp-release-dwx-server-<pod_id>
```

- Resume the upgrade wizard if it had timed out.

What to do next

- After upgrading, the Cloudera Manager admin role may be missing the Host Administrators privilege in an upgraded cluster. The cluster administrator should run the following command to manually add this privilege to the role.

```
ipa role-add-privilege <cmadminrole> --privileges="Host Administrators"
```

- If you specified a custom certificate, select the Cloudera Embedded Container Service cluster in Cloudera Manager, then select **Actions Update Ingress Controller**. This command copies the cert.pem and key.pem files from the Cloudera Manager server host to the Cloudera Embedded Container Service Management Console host.
- After upgrading, you can enable the unified time zone feature to synchronize the Cloudera Embedded Container Service cluster time zone with the Cloudera Manager Base time zone. When upgrading from earlier versions of Cloudera Data Services on premises to 1.5.5 SP1 and higher, unified time zone is disabled by default to avoid affecting timestamp-sensitive logic. For more information, see [Cloudera Embedded Container Service unified time zone](#).

Post upgrade - Ozone Gateway validation

If you are using Cloudera Data Engineering, after upgrading Cloudera Data Services on premises you must validate that the Ozone Gateway is working as expected. This procedure applies to both 1.5.3 and 1.5.4 to 1.5.5 upgrades.

About this task

You can run the following commands to get the types of logs that are available with the job run.

Command 1

```
cde run logs --id <run_id> --show-types --vcluster-endpoint <job_api_url> --  
cdp-endpoint <cdp_control_plane_endpoint> --tls-insecure
```

For example,

```
cde run logs --id 8 --show-types --vcluster-endpoint https://76fsk4rz.cde-fm  
ttv45d.apps.apps.shared-rke-dev-01.kcloud.cloudera.com/dex/api/v1 --cdp-endp  
oint https://console-cdp-keshaw.apps.shared-rke-dev-01.kcloud.cloudera.com -  
-tls-insecure
```

Log:

TYPE	ENTITY	STREAM	ENTITY DEFAULT
driver/stderr	Driver	stderr	True
driver/stdout	Driver	stdout	False
executor_1/stderr	Executor 1	stderr	True
executor_2/stdout	Executor 2	stdout	False

Command 2

```
cde run logs --id <run_id> --type <log_type> --vcluster-endpoint <job_api_url>  
--cdp-endpoint <cdp_control_plane_endpoint> --tls-insecure
```

For example,

```
cde run logs --id 8 --type driver/stderr --vcluster-endpoint https://76fsk4r  
z.cde-fmttv45d.apps.apps.shared-rke-dev-01.kcloud.cloudera.com/dex/api/v1 --
```

```
cdp-endpoint https://console-cdp-keshaw.apps.shared-rke-dev-01.kcloud.cloudera.com --tls-insecure
```

Log:

```
Setting spark.hadoop.yarn.resourcemanager.principal to hive
23/05/22 09:27:28 INFO SparkContext: Running Spark version 3.2.3.1.20.71720
00.0-38
23/05/22 09:27:28 INFO ResourceUtils: =====
=====
23/05/22 09:27:28 INFO ResourceUtils: No custom resources configured for spark.driver.
23/05/22 09:27:28 INFO ResourceUtils: =====
=====
23/05/22 09:27:28 INFO SparkContext: Submitted application: PythonPi
23/05/22 09:27:28 INFO ResourceProfile: Default ResourceProfile created, executor resources: Map(cores -> name: cores, amount: 1, script: , vendor: , memory -> name: memory, amount: 1024, script: , vendor: , offHeap -> name: offHeap, amount: 0, script: , vendor: ), task resources: Map(cpus -> name: cpus, amount: 1.0)
23/05/22 09:27:29 INFO ResourceProfile: Limiting resource is cpus at 1 tasks per executor
23/05/22 09:27:29 INFO ResourceProfileManager: Added ResourceProfile id: 0
23/05/22 09:27:29 INFO SecurityManager: Changing view acls to: sparkuser,cdpuser1
23/05/22 09:27:29 INFO SecurityManager: Changing modify acls to: sparkuser,cdpuser1
23/05/22 09:27:29 INFO SecurityManager: Changing view acls groups to:
23/05/22 09:27:29 INFO SecurityManager: Changing modify acls groups to:
23/05/22 09:27:29 INFO SecurityManager: SecurityManager: authentication enabled; ui acls disabled; users with view permissions: Set(sparkuser, cdpuser1); groups with view permissions: Set(); users with modify permissions: Set(sparkuser, cdpuser1); groups with modify permissions: Set()
.....
.....
```

Results

- If you can see the driver pod logs, then Ozone Gateway is working as expected and you can go ahead with the upgrade.
- If the logs do not appear, then you can try restarting the Ozone Gateway and get Spark job's driver log to validate if Ozone gateway is healthy or not.
- If you do not get the Spark job driver log, then you must contact Cloudera Support.
- For more information about configuring Cloudera Data Engineering CLI, see [Using the Cloudera Data Engineering command line interface](#)

Post upgrade - Restoring Cloudera Data Engineering service

After you take backup of the Cloudera Data Engineering service and upgrade your Cloudera platform, you can restore the Cloudera Data Engineering service. Once you upgrade to Cloudera Data Engineering 1.5.5, the endpoints that you were using in the previous version are not supported.

Restoring a Cloudera Data Engineering service

You can restore the Cloudera Data Engineering service with its jobs, resources, job run history, and job logs from a backed-up ZIP file.

Before you begin

You must back up the Cloudera Data Engineering service, expand the resource pool, and then upgrade your Cloudera on premises to restore the Cloudera Data Engineering service. Also, you must validate that the Ozone Gateway is working as expected by performing the steps listed in the *Post upgrade - Ozone Gateway validation* topic.



Important: It is recommended to copy the logs of the commands run from the terminal and save them on your machine. This will be helpful in debugging or when raising a support ticket. You can also increase the terminal buffer size so that it does not throw away the logs and save the terminal logs of each command for reference.

Procedure

1. If you have exited from the previous terminal where the pre-upgrade commands were run for the Cloudera Data Engineering service being upgraded, then you have to export these variables before running any docker command.

```
export BASE_WORK_DIR=[ ***HOST_MACHINE_PATH*** ]
export BACKUP_OUTPUT_DIR=/home/dex/backup
```

2. Set the following environment variables in case you have exited from the ECS Server host:

```
export PATH=$PATH:/opt/cloudera/parcels/ECS/installer/install/bin/linux:/opt/cloudera/parcels/ECS/docker
export KUBECONFIG=~/.kubeconfig
export BASE_WORK_DIR=/opt/backup-restore
export BACKUP_OUTPUT_DIR=/home/dex/backup
```

3. Run the dex-upgrade-utils docker image on the ECS Server host to restore the service.

```
docker run \
-v [***KUBECONFIG_FILE_PATH***]:/home/dex/.kube/config:ro \
-v [***CDP_CREDENTIAL_FILE_PATH***]:/home/dex/.cdp/credentials:ro \
-v [***CDE-UPGRADE-UTIL.PROPERTIES_FILE_PATH***]:/opt/cde-backup-restore/scripts/backup-restore/cde-upgrade-util.properties:ro \
-v [***LOCAL_BACKUP_DIRECTORY***]:$BACKUP_OUTPUT_DIR \
-e KUBECONFIG=/home/dex/.kube/config \
[***DOCKER_IMAGE_NAME***]:[***DOCKER_IMAGE_VERSION***] clone-service
-s [***CDE-CLUSTER-ID***] -f $BACKUP_OUTPUT_DIR/[***BACKUP-ZIP-FILE-NAME***]
```

Where -s is the Cloudera Data Engineering service ID and -f is the backup output directory path in the container.

**Important:**

- Starting from Cloudera Data Services on premises 1.5.5 release, the restore-service command is changed to clone-service.
- Upgrading Cloudera Data Engineering service from version 1.5.4 or earlier to Cloudera Data Engineering 1.5.5 does not support endpoint stability.

Example:

```
docker run \
-v $BASE_WORK_DIR/secrets/kubeconfig:/home/dex/.kube/config:ro \
-v $BASE_WORK_DIR/secrets/credentials:/home/dex/.cdp/credentials:ro \
-v $BASE_WORK_DIR/cde-upgrade-util.properties:/opt/cde-backup-restore/scripts/backup-restore/cde-upgrade-util.properties:ro \
-v $BASE_WORK_DIR/backup:$BACKUP_OUTPUT_DIR \
-e KUBECONFIG=/home/dex/.kube/config \
docker-private.infra.cloudera.com/cloudera/dex/dex-upgrade-utils:1.20.1-b48 clone-service -s cluster-c2dhkp22 -f $BACKUP_OUTPUT_DIR/cluster-c2dhkp22-2023-03-10T06_00_05.zip
```

4. If you are using a Cloudera database that is external and is not accessible from the container which is running the Cloudera Data Engineering upgrade command, then the following SQL statements are displayed in the logs.

Example:

```
2023-05-17 13:02:29,551 [INFO] CDP control plane database is external and
not accessible
2023-05-17 13:02:29,551 [INFO] Please rename the old & new cde service
name manually by executing below SQL statement
2023-05-17 13:02:29,551 [INFO]      update cluster set name = 'cde-base-
service-1-19-1' where id = 'cluster-c2dhkp22';
2023-05-17 13:02:29,551 [INFO]      update cluster set name = 'cde-base-
service' where id = 'cluster-92c2fkgb';
2023-05-17 13:02:29,551 [INFO] Please update the lastupdated time of ol
d cde service in db to extend the expiry interval of db entry for suppor
ting CDE CLI after old CDE service cleanup
2023-05-17 13:02:29,551 [INFO]      update cluster set lastupdated =
'2025-05-05 06:16:37.786199' where id = 'cluster-c2dhkp22';
-----
```

You must execute the above SQL statements to complete the restore process.

If you have closed the terminal or do not have this information, run the following SQL statements and specify the cluster details. Use the cluster ID that you have noted when performing the steps listed in the *Prerequisites for upgrading Cloudera Data Engineering Service* section.

- a. Rename old Cloudera Data Engineering service.

```
update cluster set name = '[***MODIFIED_SERVICE_NAME***]' where id =
'[***OLD_CDE_CLUSTER_ID***]';
```

Example:

```
update cluster set name = 'cde-base-service-1-19-1' where id = 'cluster-
c2dhkp22'
```

- b. Rename the new Cloudera Data Engineering service to the old Cloudera Data Engineering service name.

```
update cluster set name = '[***OLD_CDE_SERVICE_NAME***]' where id =
'[***NEW_CDE_CLUSTER_ID***]';
```

Example:

```
update cluster set name = 'cde-base-service' where id = 'cluster-92c2fkgb'
```

- c. Run the following command so that when the old Cloudera Data Engineering service is deleted or disabled then it is not cleared from the database for the next two years. The timestamp format must be the same and should be two years from the current time.

```
update cluster set lastupdated = '[***YYYY-MM-DD HH:MM:SS[.NNN]***]' wh
ere id = '[***OLD_CDE_CLUSTER_ID***]';
```

Example:

```
update cluster set lastupdated = '2025-05-05 06:16:37.786199' where id =
'cluster-c2dhkp22'
```

5. After the restore operation completes, validate that the jobs and resources are restored by running the `cde job list` and `cde resource list` CLI commands or check the virtual cluster job UI.

In the **Administration** page of the Cloudera Data Engineering UI, you can see the old Cloudera Data Engineering service is appended with a version number. For example, if the old Cloudera Data Engineering service name was `cde-sales`, after the restore, the old Cloudera Data Engineering service is something similar to `cde-sales-1-19.1`.

6. You can now delete the old Cloudera Data Engineering service after validating that everything is working as expected. If you delete the old Cloudera Data Engineering service, then you can shrink the resource pool size back to its initial value which you expanded in the *Prerequisite* steps. Do not delete the service if you want to rollback to the old service.



Important: If you are upgrading from Cloudera Data Services on premises 1.5.4 or earlier, make sure that you update the following after upgrading to Cloudera Data Services on premises 1.5.5.

- Update the TLS certificate. For more information about updating your TLS certificate for Services, see [Updating the Control Plane certificates in Cloudera Data Engineering services](#) and for Virtual Clusters, see [Updating the Control Plane certificates in Cloudera Data Engineering Virtual Clusters](#).
- Update the Kerberos keytab file in Hadoop Authentication. For more information about how to update the keytab file, see [Hadoop Authentication](#).

Related Information

[Upgrading CDP on the Cloudera Embedded Container Service](#)

[Ozone Gateway validation](#)

[Prerequisites for upgrading Cloudera Data Engineering Service](#)

Limitations of restoring Cloudera Data Engineering service

Learn about the limitations that you might run into while restoring your Cloudera Data Engineering service.

- Airflow job logs of the old cluster will be lost after the Restore operation.
- Cloudera Data Engineering service does not support backup and restore of an Airflow Python environment resource after upgrade. You must create a new Airflow Python environment resource after restoring Cloudera Data Engineering service and then activate it manually.