CDP Private Cloud Data Services 1.5.1

CDP Private Cloud Data Services Release Notes

Date published: 2020-12-16
Date modified: 2023-06-08

CLOUDERA

Legal Notice

Contents

Data Services Release Notes

CDP Private Cloud Data Services includes Management Console, Cloudera Data Warehouse (CDW), Cloudera Machine Learning (CML) and Cloudera Data Engineering (CDE). Learn about the new features and improvements in each of these services.

Note:

As of CDP Private Cloud Data Services 1.5.1, external Control Plane metadata databases are no longer supported. New installs require the use of an embedded Control Plane database. Upgrades from CDP Private Cloud Data Services 1.4.1 or 1.5.0 to 1.5.1 are supported, but there is currently no migration path from a previous external Control Plane database to the embedded database. Upgrades from 1.4.0 or 1.5.0 with external Control Plane metadata databases also require additional steps, which are described in the CDP Private Cloud Data Services 1.5.1 upgrade topics.

New features and improvements in this version of CDP Private Cloud Data Services:

- Platform:

  - Security

    - CVEs
    - DirName Kerberos support for Data Services (CDW and CML)

  - Data Services Enablements

    - Cadence support to enable CML DRS capability
    - Client configuration for enabling CDE with Kafka
    - ServiceDiscovery API to enable Data Services access to Base Storage services
    - Provide Spark3 client configurations for CML
    - Elastic Quota Management capabilities for CDE

  - Certifications

    - FreeIPA
    - OCP 4.11
    - RKE2 (v1.24) and Longhorn (1.4.2) version upgrades

- CDW:

  - Support for Iceberg V2 (Technical Preview)
  - Hive ACID compaction observability
  - Migration updates
  - Kerberos authentication for Beeline, Impala-shell, Impala, and other JDBC clients
  - Group-level access control for Impala Virtual Warehouse
  - Support for Ozone erasure coded data in Impala
  - Workload-aware autoscaling for Impala
  - Ozone File Browser support for Hue

- CDE:

  - Support for Iceberg V2
  - Ozone integration using Data Connectors
  - CDE Sessions
  - Migration updates
  - Elastic Quotas
  - Pulse metrics
  - Gang scheduling support for Spark applications
  - Support for Airflow 2.3.4

- CML:

    - Disable Engines
    - CDSW to CML migration (Technical Preview)
    - Migration preflight checklist
    - Iceberg v2 support

    - Spark 3 pushdown
    - Backup/Restore on PVC (Technical Preview)
    - Ozone add-on integration
    - Free IPA certification
    - Model Registry

Release notes for component services:

- Management Console
- Cloudera Data Warehouse (CDW)
- Cloudera Machine Learning (CML)
- Cloudera Data Engineering (CDE)
- Cloudera Manager

Cumulative hotfixes

The following cumulative hotfixes have been shipped for CDP Private Cloud Data Services 1.5.1.

CDP Private Cloud Data Services 1.5.1-CHF1

Learn about the features, fixes, and known issues in the CDP Private Cloud Data Services 1.5.1 Cumulative hotfix 1 release.

This cumulative hotfix was released on August 31, 2023.

> **Note:** Contact Cloudera Support for questions related to any specific hotfixes.

> **Note:** To upgrade to CDP Private Cloud Data Services 1.5.1-CHF1 from 1.5.1, follow the procedures described in Upgrading on the Embedded Container Service and Upgrading on OpenShift, with 1.5.1-CHF1 as the target version.

What's new

The following features have been introduced in the CDP Private Cloud Data Services 1.5.1-CHF1 hotfix release:
CDP Private Cloud Data Services is certified to work with Portworx

> The CDP Private Cloud Data Services platform and data services including Cloudera Data Warehouse (CDW), Cloudera Machine Learning (CML), and Cloudera Data Engineering (CDE) are supported with Portworx as a storage platform on OpenShift Container Platform (OCP). For more information, see Storage classes.

Added a utility to import and export CML projects

> Individual projects can now be migrated from CDSW 1.10.x to CML in both public and private clouds. For more information see Migrating Projects.

Fixed issues

The following issues were fixed in the CDP Private Cloud Data Services 1.5.1-CHF1 hotfix release:

- OPSX-4397: The domain name parameter from the environment service has an additional "apps" sub-domain
- DSE-27535: CML will have an option to always create a job in a disable or paused state
- DSE-16989: apiv2 - project creation in team context
- DSE-29305: CDSW to CML migration failing in CDSW 1.10.4
- DSE-28649: Unable to migrate jobs with Legacy Engines to MLRuntimes
- DSE-29684: Fix race condition issue when updating user namespace in CML
- DWX-15933: Make Atlas service as an optional dependency in CDW
- DWX-15869: Database Catalog edit page does not load

Known issues

Following are the platform-related known issues in the CDP Private Cloud Data Services 1.5.1-CHF1 hotfix release:
CDPVC-1137 and COMPX-15083: OpenShift Container Platform version upgrade from 4.10 to 4.11 fails due to a Pod Disruption Budget (PDB) issue

PDB can prevent a node from draining which makes the nodes to report the "Ready,SchedulingDisabled" state. As a result, the node is not updated to correct the Kubernetes version when you upgrade OCP from 4.10 to 4.11.

To resolve this issue, confirm that the upgrade has failed due to the PDB issue, and then manually delete the PDBs from the Private Cloud namespace.

1. Run the following command to check whether the nodes are stuck in the "Ready,SchedulingDisabled" state:

```
oc get nodes
```

2. Get the machine config daemon details of the particular pod as follows:

```
oc get po -n openshift-machine-config-operator -l 'k8s-app=m
achine-config-daemon' -o wide
```

3. Check the logs of the machine config operator of that particular node as follows:

```
oc logs -f -n openshift-machine-config-operator [***MACHINE-
CONFIG-DAEMON-NAME***] -c machine-config-daemon
```

Replace *[***MACHINE-CONFIG-DAEMON-NAME***]* with the actual machine config daemon name.

You may see one of the following errors in the node logs:

- error when evicting pods/cdp-release-cpx-liftie-****" -n "*[***PRIVATE-CLOUD-NAMESPACE***]* Cannot evict pod as it would violate the pod's disruption budget
- error when evicting pods/"cdp-release-cluster-proxy-[******]" -n "*[***PRIVATE-CLOUD-NAMESPACE***]* Cannot evict pod as it would violate the pod's disruption budget

Delete the PDB from the Private Cloud namespace as follows:

a. Obtain the PDB for the cdp-release-cluster-proxy namespace:

```
oc get pdb -n [***PRIVATE-CLOUD-NAMESPACE***] | grep cdp-
release-cluster-proxy
```

b. Back up the PDB:

```
oc get pdb [***PDB-NAME-OF-CLUSTER-PROXY***] -n [***PRIVATE-
CLOUD-NAMESPACE***] -o yaml >> [***BACKUP-FILE-NAME***].yaml
```

c. Delete the PDB:

```
oc delete pdb [***PDB-NAME-OF-CLUSTER-PROXY***] -
n [***PRIVATE-CLOUD-NAMESPACE***]
```

Repeat the steps to delete the cdp-release-cpx-liftie PDB as well.

For data services-related known issues, see:

- Known issues in Cloudera Data Eangineering
- Known issues in Cloudera Data Warehouse
- Known issues in Cloudera Machine Learning

Repository locations

The URLs for CDP Private Cloud Data Services 1.5.1-CHF1 are listed in the following table:

URL Type	Repository Location
Index	```https://username:password@archive.cloudera.com/p/cdp-pvc-ds/1.5.1-h2/```
Manifest	Repository: ```https://username:password@archive.cloudera.com/p/cdp-pvc-ds/1.5.1-h2/manifest.json```
Parcels	Repository: ```https://username:password@archive.cloudera.com/p/cdp-pvc-ds/1.5.1-h2/parcels/```