

Cloud Support

Date published: 2024-01-01

Date modified: 2024-08-15

CLOUDERA

Legal Notice

© Cloudera Inc. 2024. 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

- Public cloud provider support..... 4**
- Amazon Web Services support in Cloudera Data Warehouse Public Cloud..... 4**
- Microsoft Azure support in Cloudera Data Warehouse Public Cloud.....6**
- Custom image repositories.....7**
 - Limitations..... 8
 - Downloading images..... 9
 - Setting up a custom repository.....9
 - Handling CDW upgrades..... 9
 - Disabling image verification..... 10

Public cloud provider support

Cloudera Data Warehouse (CDW) Public Cloud service supports Amazon Web Services (AWS) and Microsoft Azure. Basic configuration for environments that use these public cloud providers is performed during registration in Management Console. To configure specific CDW features, you use the CDW service UI.

Amazon Web Services support in Cloudera Data Warehouse Public Cloud

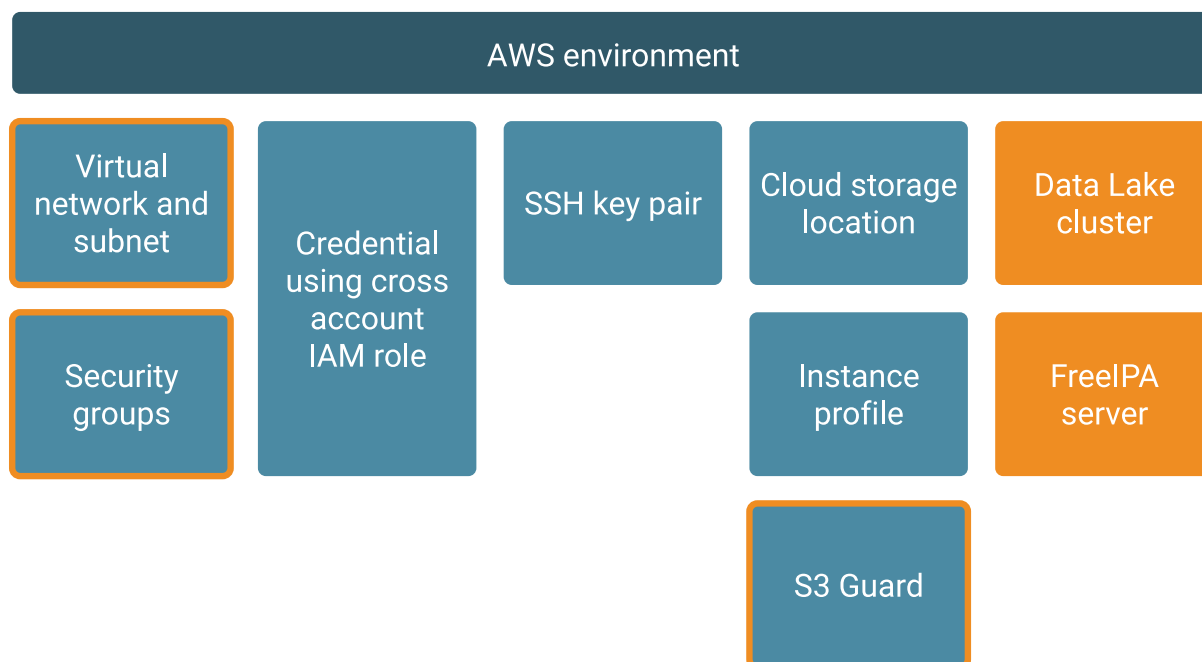
This topic provides an overview of Amazon Web Services (AWS) support in Cloudera Data Warehouse (CDW) Public Cloud.






Important: Ensure your AWS environment satisfies the items listed in the [AWS environments requirements checklist](#).

The "environment" concept in Cloudera Data Platform (CDP) is closely related to the virtual private network in your AWS account. Registering an environment provides CDP with access to your AWS account and identifies the resources there that CDP services can access or provision. A single environment is contained within a single AWS region, so all resources deployed by CDP are deployed within that region within one specific virtual network. Once you've registered an environment with Management Console, you can start provisioning resources such as clusters, which run on the physical infrastructure in an AWS data center.

The following diagram shows the components of a CDP environment on AWS:



-
-  Can optionally be created by CDP admin or CDP can create automatically
 -  Always created by CDP admin
 -  Always created by CDP

The diagram includes all major user-created and CDP-created components of an AWS environment:

- Items in dark blue boxes with orange outlines can be automatically provisioned by CDP in your AWS account. Or you can optionally pre-create them in your AWS account and then provide them when registering an environment in CDP.
- Items in dark blue boxes must be pre-created by your CDP administrator prior to environment registration and then provided during environment registration in CDP.
- Items in orange boxes are automatically provisioned on AWS by CDP as part of environment registration.



Note: Items that are user-created are not terminated when a CDP environment is deleted.

Information that describes how to register an AWS environment with CDP is linked to at the end of this page.

Data Warehouse service features for AWS environments

CDP Data Warehouse service offers the following additional features for AWS environments that are used for Database Catalogs and Virtual Warehouses:

- Restrict access to Kubernetes endpoints and to service endpoints of the Kubernetes cluster at the load balancer lever by specifying a list of IP CIDRs that are allowed access. For more information, see "Restricting endpoint access in AWS," which is linked to at the end of this page.
- Set up private networking in AWS environments, which uses AWS private subnets. Private subnets receive no direct inbound connections from the internet, providing private network connectivity for workload endpoints in

Data Warehouse service. For more information, see "Set up private networking in AWS," which is linked to at the end of this page.

These features require additional configuration in the Data Warehouse service UI to use them.

Related Information

[AWS account requirements](#)

[AWS credentials](#)

[How to register AWS environments](#)

[Restricting endpoint access in AWS](#)

[Set up private networking in AWS](#)

Microsoft Azure support in Cloudera Data Warehouse Public Cloud

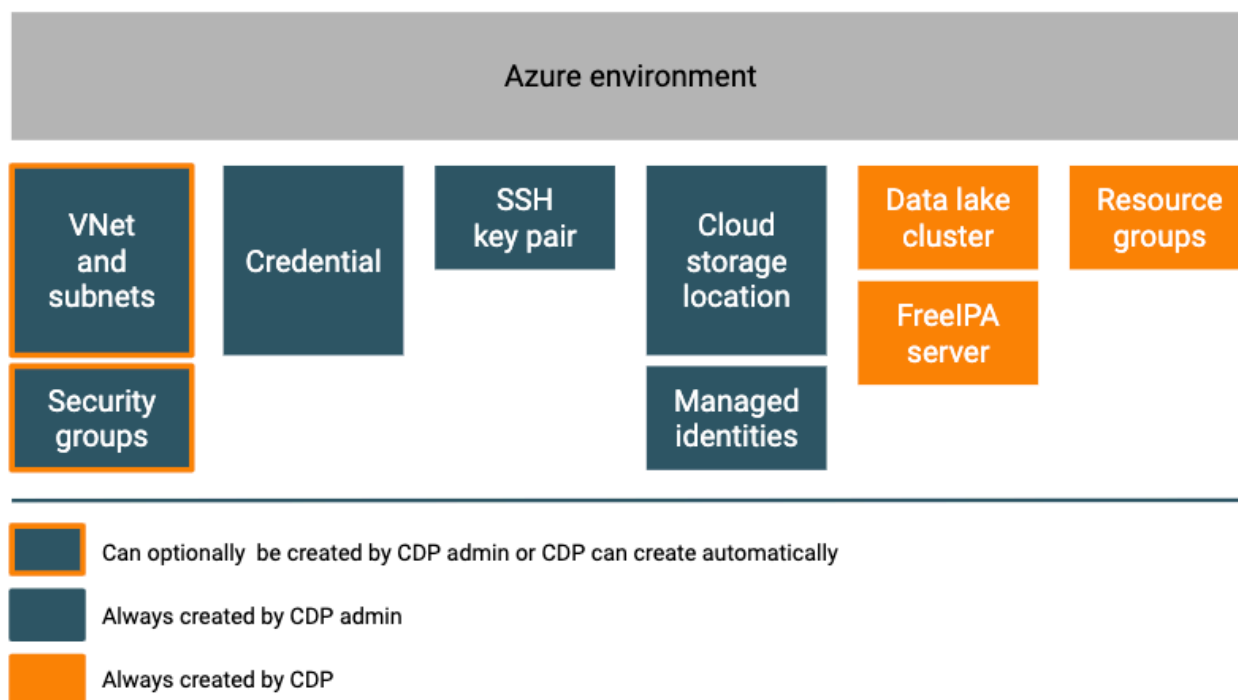
This topic provides an overview of Azure support in Cloudera Data Warehouse (CDW) Public Cloud.



Important: Ensure your Azure environment satisfies the items listed in the [Azure environments requirements checklist](#).

Like Amazon Web Services support, Cloudera Data Platform (CDP) environments are closely related to the virtual private network concept in your Azure account as well. Also registering an environment provides CDP access to resources in your Azure account and a single CDP environment is contained within a single region in Azure. All resources that are deployed by CDP are deployed within that region and within one specific virtual network. After you have registered an environment in CDP, you can start provisioning CDP resources such as clusters, which run on the physical infrastructure in an Azure data center.

The following diagram shows the components of a CDP environment on Azure:



The diagram illustrates all major user-created and CDP-created components of an environment:

- The items in dark blue boxes with orange outlines can either be automatically provisioned by CDP on your Azure account, or you can optionally pre-create them and specify them when registering an environment.
- The items in dark blue boxes must be pre-created by your CDP administrator prior to environment registration and then specified when registering an environment.
- The items in orange boxes are automatically provisioned on Azure by CDP as part of environment provisioning.



Note: The items that are user-created are not terminated during environment deletion.

Information that describes how to register an Azure environment with CDP is linked to at the end of this page.

Related Information

[Azure account requirements](#)

[Azure credentials](#)

[How to register Azure environments](#)

Custom image repositories

If your organization must control the acquisition and provisioning of images in your cloud account custom image repositories are available. If your organization does not allow internet access, or restricts image repositories to only those within your virtual private network (VPC in AWS or VNet in Azure), you can bring your own repository to CDP.

Using your custom repository, you can host and scan Cloudera Data Warehouse (CDW) images. You gain complete control over which images are provisioned in your cloud account and how you acquire the images. You can choose to verify, or not, that the image targeted by a query is present.

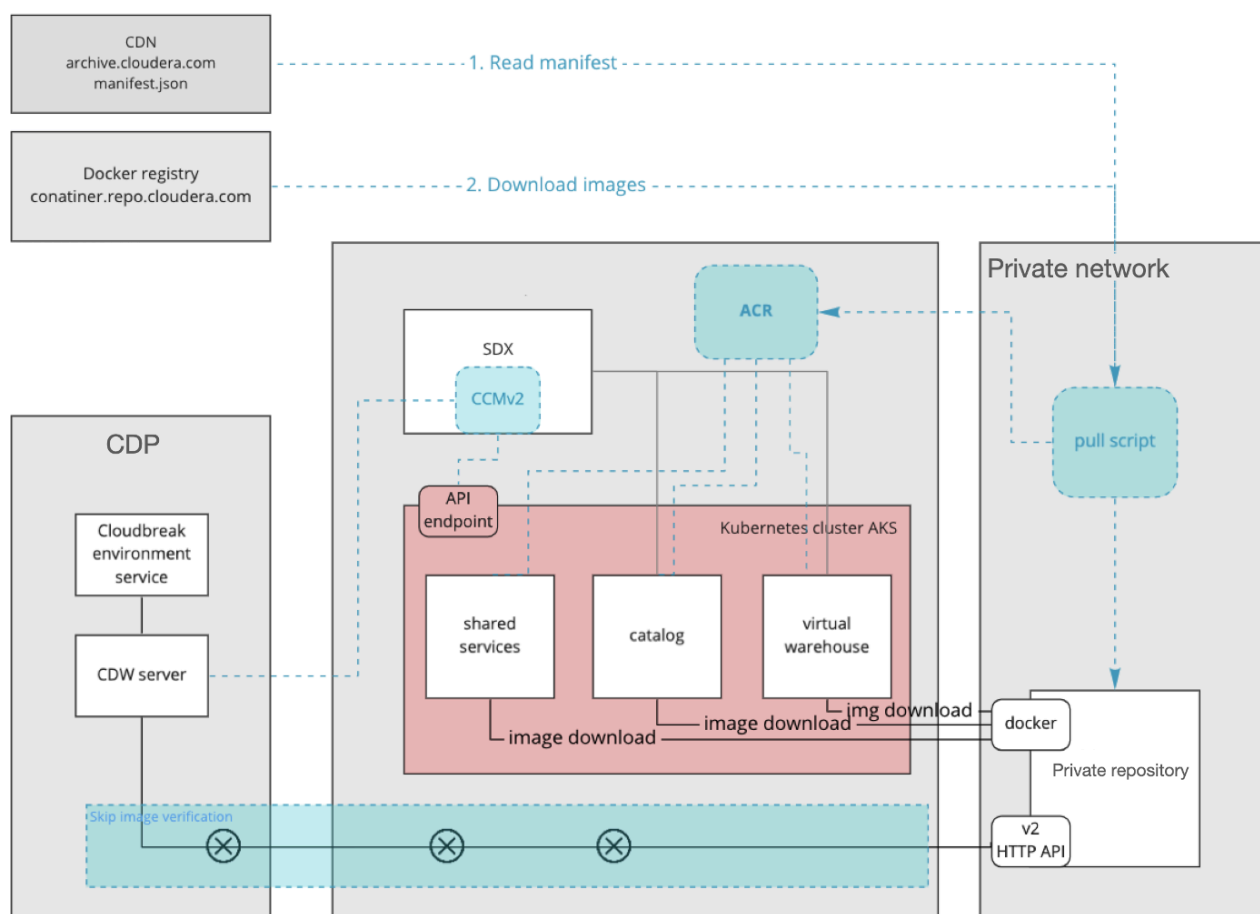
The following registries are supported for use as a custom image repository in CDW:

- Azure Container Registry (ACR)
- Amazon Elastic Container Registry (ECR)
- Docker registry V2

To use this feature, you must obtain the CDP_CUSTOM_REPO entitlement.

CDW architecture

The following diagram shows the CDW architecture using a custom image repository based on ACR. The architecture is similar for ECR-based repositories.



When you set up a custom repository, you must read the manifest.json file using your payroll credentials from Cloudera to obtain a list of images. Cloudera provides an image list with each release. You must pull images each time a new release of CDW occurs. Cloudera recommends that you create an automated way to pull these images.

You can set the Disable image version verification option. You might not set this option, thereby enabling verification, for networking or security reasons. You might set this option to disable verification if you need time to onboard new images. The option works as follows:

- If you do not set this option, Cloudera attempts to verify the images in your repository, and if the latest version of the images is not available in the registry, some filtering occurs, and old images are started. CDW does not fully support starting older images, so pulling new versions of images promptly is recommended.
- If you do set this option, when the new releases occur, the server starts the new images without verification. If the images are not present in the registry, provisioning fails.

Limitations

- Image namespaces must not include arbitrary paths, such as `.../mycompany/myhierarchy/hive:...`. Only the following namespace format is allowed:

```
container.repo.cloudera.com/cloudera/hive:...
```

- CLI integration is not available.
- Azure private registry support is available, but image verification is disabled. In this release, CDW does not support a repository that is set up as a private registry from the networking perspective.

Downloading images

Learn how to download Cloudera Data Warehouse container images from the registry. You can run a security scan and approve (or block) each image before copying the image into your repository

Before you begin

Obtain the following entitlements:

- Salesforce entitlement
- CDP_CUSTOM_REPO entitlement

Procedure

1. Login to cloudera container registry (<https://container.repo.cloudera.com>) using your Salesforce entitlement token and password.

For example:

```
$ docker login container.repo.cloudera.com -u 634ae2db-f926-4530-842c-95d3401f37e5
```

2. Download container images using the docker pull command. For example:

```
docker pull container.repo.cloudera.com/cloudera/hive:7.2.3.0-22
```

For example:

```
Pulling from cloudera/hive Digest:sha256:29c635b8612c770c8089d4ab134c10e599dcf61553dfa4147975d6af33c835a9
Status: Downloaded newer image for container.repo.cloudera.com/cloudera/hive:7.2.3.0-22container.repo.cloudera.com/cloudera/hive:7.2.3.0-22
```

Setting up a custom repository

When you activate the environment from Cloudera Data Warehouse (CDW), you can choose to use a custom repository.

Procedure

1. Log in to the Data Warehouse service as DWAdmin.
2. Go to the Environments tab, locate the environment you want to activate, and click Activate.
The **Activate Environment** modal is displayed.
3. Select the Use Custom repository option.
Depending in your custome repository type, input ecr, acr, or docker in the Registry Type field.
4. Specify the URL of your custom repository.
5. Fields for your user name and password appear.
You do not need to enter these credentials.
6. Click Activate.

Handling CDW upgrades

About this task

You must update your custom repository to contain New Cloudera Data Warehouse images when available before you upgrade your Virtual Warehouse or Database Catalog. The Virtual Warehouse UI indicates when an upgrade is available. If the new images are not available in your custom repository the upgrade fails. Currently there is no rollback option available.

If all required images for the new version of a Virtual Warehouse or Database catalog are not in the repository, CDW falls back to a previous version. You can then create a Virtual Warehouse based on the earlier version of CDW.

Disabling image verification

About this task

By default, when you authorize CDW to run a query on your custom repository, CDW performs the following verifications:

- Searches tags in repository to verify image availability.
- Populates the image registry secrets on the Kubernetes namespaces for registry type docker if credentials are required.
- Fills in the registry address in the Helm chart template.