Cloudera Runtime 7.3.1

Application Access

Date published: 2023-09-07
Date modified: 2024-12-02

CLOUDƎRA

Legal Notice

Contents

Application Access

Learn more about the client RPMs /Debian packages, role requirements and how to access the services installed on the CDP cluster from your containerized application or edge nodes.

Note: This feature is available with limited availability.

Role Requirements

For users creating applications that connect to a CDP service(s), the Cluster Administrator must have provided the configurations required for you to connect to the services.

Client RPMs enable you to access services installed on the CDP cluster from your containerized application or edge nodes that are not managed by Cloudera Manager. You can use Client RPMs for the use cases that require thin client applications binding, without having to download and install the entire CDP parcel which requires over 10 GB of storage space on each platform node. Starting with the CDP Private Cloud Base 7.1.9 release, '-client' RPM/Debian packages are available.

Support for Application Access

Client libraries can be deployed on nodes that are not managed by Cloudera Manager. However, the burden of managing the configuration of those nodes such that the clients can connect and interact with the cluster falls outside the scope of support. Therefore, before contacting Cloudera support, the client issue must be replicated on a managed node by the customer. If the issue cannot be replicated on a managed node, then the issue is likely due to the configuration of the unmanaged node.

Services support client RPMs for CDP Private Cloud Base 7.3.1

Learn about the list of services that supports client RPMs for the CDP Private Cloud Base 7.3.1.

Services	Package
Hadoop	hadoop-client
	hadoop-libhdfs-client
	hadoop-mapreduce-client
	hadoop-yarn-client
	hadoop-hdfs-client
ZooKeeper	zookeeper-client
	zookeeper-native-client
Solr	solr-client
	search-client
	solr-crunch-client
	solr-mapreduce-client

Services	Package
Avro	avro-tools-client
HBase	hbase-client
	hbase-solr-client
	hbase_mcc-client
Impala	impala-shell-client
	impala-udf-devel-client
Oozie	oozie-client
Ozone	ozone-client
Phoenix	phoenix-client
	phoenix_queryserver-client
Spark	Cloudera recommends submitting spark jobs from edge nodes using Livy, see Submitting Spark applications using Livy.
Sqoop	sqoop-client

Prerequisites

Learn more about the configuration files and changes required to the edge node.

Copy all the required configuration files to the edge node and ensure that the configuration changes are available on the edge node.

Here are few examples:

- Copy the keytabs from the cm_host to the edge node host
- Copy the cloudera-scm-agent certificate from cm_host to edge node host
- Copy the krb5.conf files from cm_host to edge node host
- Copy the /etc/ configs files from cm_host to edge node host

Support for packages on RHEL, SLES, and Ubuntu

Understand about the support for the package on RHEL, SLES, and Ubuntu.

The client RPM package works on RHEL 7, RHEL 8, RHEL 9, SLES 12 and SLES 15. The Debian based packages should work on Ubuntu 20 and on all Debian-based Linux distributions.

Package management tools

Learn about the package management tools for RHEL, SLES, and Ubuntu.

Package management tools

- For RHEL 7 and RHEL 8, you can use the YUM package management tool and for RHEL 9, you can either use YUM or DNF package management tool.

  Note: Provide the .repo configuration file for YUM and DNF before the installation.

  The configuration file for YUM and related utilities is located in /etc/yum.conf
- For SLES 12 and SLES 15, you can use the zypper package management tool.
- For Ubuntu 20, you can use apt get install tool. The configuration file for the Debian package is called .list. The .list files contain configuration details for the Debian repository.

Repository configuration files

Learn about the repository configuration files.

Repository configuration files

Information about package repositories is stored in configuration files, the location of which varies according to the package management tool.

These are the location of the different configuration files.

- RHEL compatible (yum): /etc/yum.repos.d
- SLES (zypper): /etc/zypp/repos.d/
- Ubuntu (apt-get): /etc/apt/sources.list.d/(Additional repositories are specified using .list files in the /etc/apt/sources .list.d/ directory.)

> ⚠️ **Important:** The client RPMs are signed with the Cloudera GPG key. You must import the repository signing GPG key, and use it before starting the installation process.

Select the repository strategy

Learn about the different repository strategy and choose the preferred one.

You can choose one of the following repository strategies to access the client packages.

Method 1

If you have internet access, you can access the Cloudera public repositories that contain the client RPM packages and available through the Cloudera provided repo location.

For example: https://archive.cloudera.com/p/cdh_client/7.3.1.0/

Ensure that you add the required repository location information to your selected .repo file present in your system and access the client packages. Alteratively, you can also make your own .repo file.

Method 2

For the internally hosted repositories, you can download the RHEL, SLES, or Ubuntu tarball to your local machine and create a local repository to access the client packages. You can use one of the following options to access the client packages.

Related Information

Option 1
Option 2
Option 3

Option 1

For the internally hosted repositories, you can download the RHEL, SLES, or Ubuntu tarball in your local machine and create a local repository to access the client packages. Learn about the option 1.

You can download the tarball and host all the client packages in a locallly shared system such as NFS. You can use the shared location that contains the yum repository and RPM content and access it across multiple locations using the mount point method.

Related Information

Select the repository strategy

Option 2

For the internally hosted repositories, you can download the RHEL, SLES, or Ubuntu tarball in your local machine and create a local repository to access the client packages. Learn about the option 2.

You can download the tarball and host all the client packages in a local file system and then use it.

Related Information

Select the repository strategy

Option 3

For the internally hosted repositories, you can download the RHEL, SLES, or Ubuntu tarball in your local machine and create a local repository to access the client packages. Learn about the option 3 and how to create an internal yum repository, temporary internal repository, and configure hosts to use the internal repository.

You can have a centralized Apache server running and host all the content on that server.

Related Information

Select the repository strategy

Creating an internal yum repository for RHEL packages

Learn more about how to create an internal yum internal repository using the Apache HTTP server.

Setting up a Web server

To host an internal repository, you must install or use an existing Web server on an internal host and then download the repository files to the Web server host. The examples in this section use Apache HTTP Server as the Web server.

1. Install Apache HTTP Server:

```
sudo yum install httpd
```

2. Start Apache HTTP Server:

```
sudo systemctl start httpd
```

Downloading and publishing the package repository for the CDH client

1. Download the package repository of CDH client that you want to install.

2. Download the files for a CDP Private Cloud Base release:

- Run the following command to create a local repository directory for the Cloudera package repository:

```
sudo mkdir -p /var/www/html/cloudera-repos/cdh-client
```

- Run the following command to modify the file permission that allows you to download the files under the local repository directory:

```
sudo chmod -R ugo+rX /var/www/html/cloudera-repos/cdh-client
```

- Copy the files to the yum directory.

3. Visit the repository URL `http://<web_server>/cloudera-repos/` in your browser and verify that the files you downloaded are present.

Creating a temporary internal repository

You can create a temporary remote repository to deploy packages on a one-time basis. Cloudera recommends using the same host that runs CDP Private Cloud Base or a gateway host. This example uses Python SimpleHTTPServer as the Web server to host the /var/www/html directory, but you can use a different directory.

1. Download the repository. You need to follow the instructions in "Downloading and publishing the package repository for the CDH Client".
2. Determine a port that your system is not listening on. This example uses port 8900.
3. Start a Python SimpleHTTPServer in the /var/www/html directory:

```
cd /var/www/html
python -m SimpleHTTPServer 8900
```

4. Serving HTTP on 0.0.0.0 port 8900 ...
5. Visit the repository URL http://<web_server>:8900/cloudera-repos/ in your browser and verify the files you downloaded are present.

Configuring hosts to use the internal repository

After creating the repository, modify the client configuration to use it:

Operating System	Procedure
RHEL compatible	Create /etc/yum.repos.d/cloudera-repo.repo files on cluster hosts with the following content, where <web_server> is the hostname of the Web server: [cloudera-repo] name=cloudera-repo baseurl=http://<web_server>/<cdh-client > enabled=1 gpgcheck=0
SLES	Use the zypper utility to update client system repository information by issuing the following command: zypper addrepo http://<web_server>/<cdh-client>
Ubuntu	Create /etc/apt/sources.list.d/cloudera-repo.list files on all cluster hosts with the following content, where <web_server> is the hostname of the Web server:

Operating System	Procedure
	deb http://<web_server>/<cdh-client><codename> <components>
	You can find the <codename> and <components> variables in the ./conf/distributions file in the repository.
	After creating the .list file, run the following command:
	sudo apt-get update

Downloading and configuring the client packages

Know about the process on how to download and configure the client packages for RHEL, SLES and Ubuntu.

RHEL

Know about the process on how to download and configure the client packages for RHEL

Note: Contact the Cloudera Account Team to get access to the CDP 7.3.1 RHEL client packages.

To setup the repository for RHEL client package, you can follow the below example.

For example:

```
Create /etc/yum.repos.d/cloudera-repo.repo files on cluster hosts with
 the following content, where <web_server> is the hostname of the Web
server:
[cloudera-repo]
name=cloudera-repo
baseurl=http://<web_server>/<cdh-client >
enabled=1
gpgcheck=1
```

SLES

Know about the process on how to download and configure the client packages for SLES.

Note: Contact the Cloudera Account Team to get access to the CDP 7.3.1 SLES client packages.

To set up the repository for SLES client package, you can follow the below example.

For example:

```
sles15-builds-5gkjt:~ # cat /etc/zypp/repos.d/cdh-client.repo
[CDH-7.3.1.0-xxx]
name=CDH Version - CDH-7.3.1.x.xxxx
baseurl=file:///grid/0/jenkins/zypp/
gpgcheck=1
gpgkey=file:///grid/0/jenkins/zypp/RPM-GPG-KEY/RPM-GPG-KEY-Jenkins
enabled=1
priority=1
type=rpm-md
sles15-builds-5gkjt:~ #
```

Ubuntu

Know about the process on how to download and configure the client packages for Ubuntu.

Note: Contact the Cloudera Account Team to get access to the CDP 7.3.1 Ubuntu client packages.

Creating an APT repository for Ubuntu Packages

1. Create the following file:

```
/etc/apt/sources.list.d/cloudera-client.list
```

2. Add the client path to the arch packages.

Note: Modify the GBN value as per the builds that are being validated

3. Check if the following signature error occur:

```
The following signatures could not be verified because the public key is
not available
```

4. Run the following command to fix the signature error:

```
apt-key adv --keyserver keyserver.ubuntu.com --recv-keys B9733A7A07513CAD
```

5. Update the repository.

```
apt update
```

6. Install the package.

```
apt install <package>
```

Installing the client packages

Know more about installing the client packages for RHEL, SLES, and Ubuntu.

RHEL

Important: You must import the repository signing GPG key and use it before starting the installation process.

You can install the client RPM for the supported service and yum update to update the newly added repository.

```
yum install <service-client>
```

Here are some examples:

```
yum install spark-python-client
yum install ozone-client
```

SLES

⚠️ **Important:** You must import the repository signing GPG key and use it before starting the installation process.

You can install the client RPM for the supported service.

```
Zypper install <service-client>
```

Here are some examples:

```
Zypper install spark-python-client
Zypper install ozone-client
```

Ubuntu

You can install the client package for the supported service.

```
apt install <service-client>
```

Here are some examples:

```
apt install spark-python-client
apt install ozone-client
```