Accessing Clusters

Date of Publish: 2019-05-28

Hortonworks

Contents

Cloudbreak user accounts

Cloudbreak and its managed clusters can be accessed with the credentials provided during the installation process.

The following table describes what credentials to use to access Cloudbreak and its managed clusters:

Component	Method	Description
Cloudbreak	Web UI, CLI	Access with the username and password provided when launching Cloudbreak on the cloud provider.
Cloudbreak	SSH to VM	Access as the "cloudbreak" user with the SSH key provided when launching Cloudbreak on the cloud provider.
Cluster	SSH to VMs	Access as the "cloudbreak" user with the SSH key provided during cluster creation.
Cluster	Ambari web UI	Access with the credentials provided in the "Cluster User" parameter during cluster creation.
Cluster	Web UIs for specific cluster services	Access with the credentials provided in the "Cluster User" parameter during cluster creation.

Finding cluster information in the web UI

Once your cluster is up and running, click on the tile representing your cluster in the Cloudbreak web UI to access information related the cluster and access cluster actions.

The information presented includes:

- Cluster summary
- Cluster information
- Event history

To learn more, refer to the following documentation:

Cluster summary

The summary bar includes basic information about your cluster.
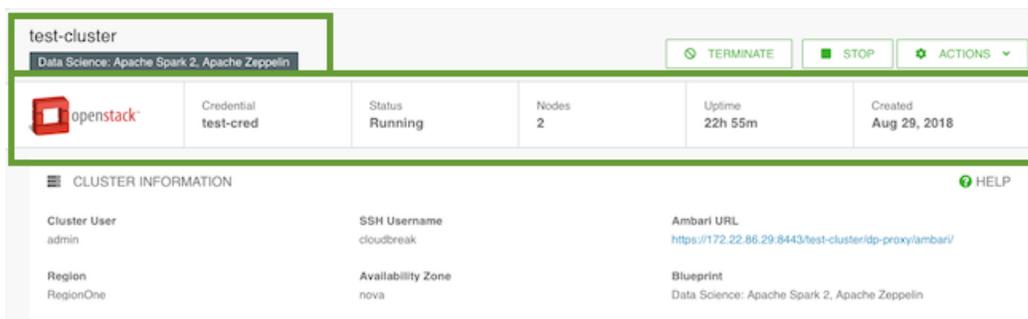
It includes the following information about your cluster:

Item	Description
Cluster name	The name that you selected for your cluster is displayed at the top of the page.
Blueprint	Below the cluster name it is the name of the blueprint used for the cluster.
Time remaining	If you enabled lifetime management for your cluster, the clock next to the cluster name indicates the amount of time that your cluster will run before it gets terminated. Note that the time remaining counter does not stop when you stop the cluster.
Cloud provider	The logo of the cloud provider on which the cluster is running.
Credential	The name of the credential used to create the cluster.
Status	Current status. When a cluster is healthy, the status is Running.
Nodes	The current number of cluster nodes, including the master node.
Uptime	The amount of time (HH:MM) that the cluster has been in the running state since it was started. Each time you stop and restart the cluster, the running time is reset to 0.
Created	The date when the cluster was created. The date format is Mon DD, YYYY. For example: Oct 27, 2017.

Cluster information

The "Cluster Information" section available from the cluster details includes detailed information about your cluster.
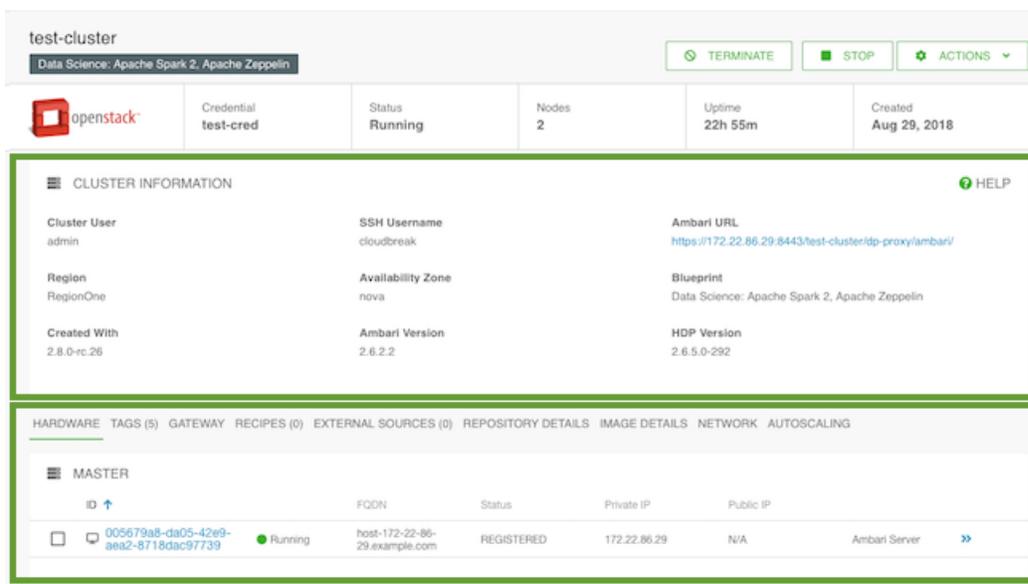
The following information is available in the cluster information section:

Item	Description
Cluster User	The name of the cluster user that you created when creating the cluster.
SSH Username	The SSH user which you must use when accessing cluster VMs via SSH. The SSH user is always "cloudbreak".
Ambari URL	Link to the Ambari web UI.
Region	The region in which the cluster is running in the cloud provider infrastructure.
Availability Zone	The availability zone within the region in which the cluster is running.
Blueprint	The name of the blueprint selected under "Cluster Type" to create this cluster.
Created With	The version of Cloudbreak used to create this cluster.
Ambari Version	The Ambari version which this cluster is currently running.
HDP/HDF Version	The HDP or HDF version which this cluster is currently running.
Authentication Source	If you are using an external authentication source (LDAP/AD) for your cluster, you can see it here.

Below this, you will see additional tabs that you can click on in order to see their content:

Item	Description
Hardware	This section includes information about your cluster instances: instance names, instance IDs, instance types, their status, fully qualified domain names (FQDNs), and private and public IPs. If you click on the ≫ , you can access more information about the instance, storage, image, and packages installed on the image.
Cloud storage	If you configured any cloud storage options, you will see them listed here.
Tags	This section lists keys and values of the user-defined tags, in the same order as you added them.
Gateway	This section is available when gateway is configured for a cluster. It includes the gateway URL to Ambari and the URLs for the service UIs.

Item	Description
Recipes	This section includes recipe-related information. For each recipe, you can see the host group on which a recipe was executed, recipe name, and recipe type.
External sources	If you are using external sources (external databases, authentication sources, or proxies), you can see them listed here.
Repository details	This section includes Ambari and HDP/HDF repository information, as you provided when creating a cluster.
Image details	This section includes information about the prewarmed or base image that was used for the cluster nodes.
Network	This section includes information about the names of the network and subnet in which the cluster is running and the links to related cloud provider console.
Security	This section is only available if you have enabled Kerberos security. It provides you with the details of your Kerberos configuration.
Autoscaling	This section includes configuration options related to autoscaling.

Event history

The "Event History" section available from the cluster details shows events logged for the cluster, with the most recent event at the top.

The Download option allows you to download the Ambari server log:
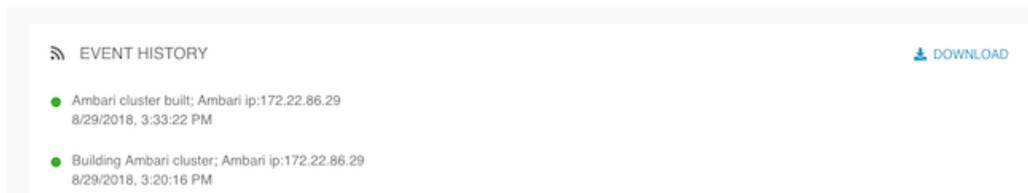
For example, after your cluster has been created, the following messages will be written to the log:

```
Ambari cluster built; Ambari ip:34.215.103.66
10/26/2017, 9:41:58 AM
Building Ambari cluster; Ambari ip:34.215.103.66
10/26/2017, 9:30:20 AM
Starting Ambari cluster services
10/26/2017, 9:27:12 AM
Setting up infrastructure metadata
10/26/2017, 9:27:11 AM
Bootstrapping infrastructure cluster
10/26/2017, 9:26:38 AM
Infrastructure successfully provisioned
10/26/2017, 9:26:37 AM
Billing started, Infrastructure successfully provisioned
10/26/2017, 9:26:37 AM
Infrastructure metadata collection finished
10/26/2017, 9:25:39 AM
Infrastructure creation took 194 seconds
10/26/2017, 9:25:37 AM
Creating infrastructure
10/26/2017, 9:22:22 AM
Setting up HDP image
10/26/2017, 9:22:21 AM
```

Access cluster via SSH

If you plan to access the cluster via the command line clients, SSH into the master node instance in the cluster.

• In order to use SSH, you must generate an SSH key pair or use an existing SSH key pair.
• You can find the cluster instance public IP addresses on the cluster details page.
• When accessing instances via SSH use the cloudbreak user.

On Mac OS, you can use the following syntax to SSH to the VM:

```
ssh -i "privatekey.pem" cloudbreak@publicIP
```

For example:

```
ssh -i "dominika-kp.pem" cloudbreak@p52.25.169.132
```

On Windows, you can access your cluster via SSH by using an SSH client such as PuTTY.

Access Ambari

You can access Ambari web UI by clicking on the links provided in the Cluster Information > Ambari URL.

Steps

1. From the cluster dashboard, click on the tile representing your cluster to navigate to cluster details.
2. Find the Ambari URL in the Cluster Information section. This URL is available once the Ambari cluster creation process has completed:
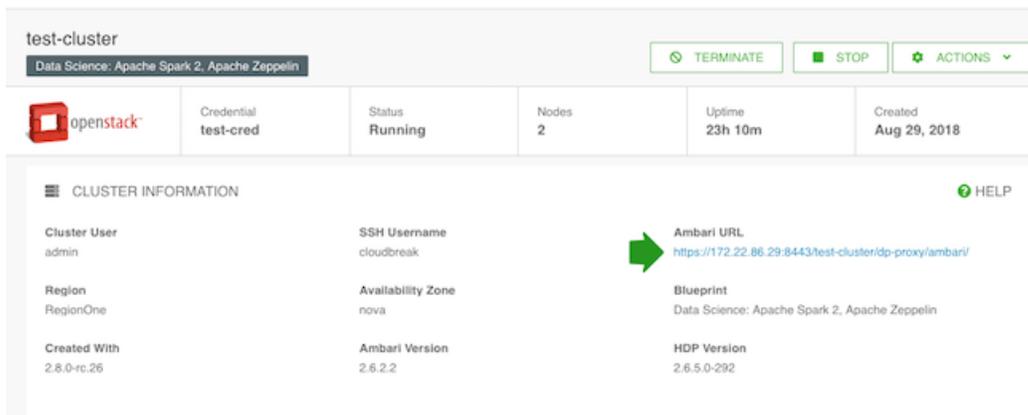
3. Click on the Ambari URL link.
4. The first time you access the server, your browser will attempt to confirm that the SSL Certificate is valid. Since Cloudbreak automatically generates a self-signed certificate, your browser will warn you about an untrusted connection and ask you to confirm a security exception. Depending on your browser, perform the steps below to proceed.

Browser	Steps
Firefox	Click Advanced > Click Add Exception… > Click Confirm Security Exception
Safari	Click Continue
Chrome	Click Advanced > Click Proceed…

5. Log in with the credentials specified during cluster creation.

Access Hive via JDBC

Hive can be accessed via JDBC through the gateway that is automatically installed and configured in your cluster. If your cluster configuration includes Hive LLAP, then Hive LLAP is configured with the gateway; otherwise, HiveServer2 is configured. In either case, the transport mode is "http" and the gateway path to Hive is "${CLUSTER_NAME}/${TOPOLOGY_NAME}/hive" (for example "test-cluster/db-proxy/hive").

Before you can start using Hive JDBC, you must download the SSL certificate to your truststore. After downloading the SSL certificate, the Hive JDBC endpoint is:

```
jdbc:hive2://
{GATEWAY_HOST}:8443/;ssl=true;sslTrustStore=gateway.jks;trustStorePassword={GATEWAY_JKS_
{TOPOLOGY_NAME}/hive
```

Related Information
Gateway Configuration

Download SSL certificate

By default, the gateway has been configured with a self-signed certificate to protect the Hive endpoint via SSL. In order to use Hive via JDBC or Beeline client, you must download the SSL certificate from the gateway and add it to your truststore.

On Linux or OSX, you can download the self-signed SSL certificate by using the following commands:

```
export GATEWAY_HOST=IP_OF_GATEWAY_NODE
export GATEWAY_JKS_PASSWORD=GATEWAY_PASSWORD
openssl s_client -servername ${GATEWAY_HOST} -connect ${GATEWAY_HOST}:8443 -
showcerts </dev/null | openssl x509 -outform PEM > gateway.pem
keytool -import -alias gateway-identity -file gateway.pem -keystore
 gateway.jks -storepass ${GATEWAY_JKS_PASSWORD}
```

Where: GATEWAY_HOST - Set this to the IP address of the instance on which gateway is running (Ambari server node). GATEWAY_JKS_PASSWORD - Create a password for the truststore that will hold the self-signed certificate. The password must be at least 6 characters long.

For example:

```
export GATEWAY_HOST=2-52-86-252-73
export GATEWAY_JKS_PASSWORD=Hadoop123!
openssl s_client -servername ${GATEWAY_HOST} -connect ${GATEWAY_HOST}:8443 -
showcerts </dev/null | openssl x509 -outform PEM > gateway.pem
keytool -import -alias gateway-identity -file gateway.pem -keystore
 gateway.jks -storepass ${GATEWAY_JKS_PASSWORD}
```

After executing these commands, gateway.pem and gateway.jks files will be downloaded onto your computer to the location where you ran the commands.

Related Information
Gateway Configuration

Example: SQL Workbench/J

SQL Workbench/J is a cross-platform SQL tool that can be used to access database systems. In this example, we provide a high-level overview of the steps required to setup SQL Workbench to access Hive via JDBC.

Prerequisite:

Download SSL certificate

Step 1: Install SQL Workbench and Hive JDBC Driver

1. Download and install SQL Workbench. Refer to http://www.sql-workbench.net/getting-started.html.
2. Download the Hortonworks JDBC Driver for Apache Hive from https://hortonworks.com/downloads/#addons. Next, extract the driver package.

Step 2: Configure SQL Workbench with Hive JDBC Driver

1. Launch SQL Workbench.
2. The Select Connection Profile window should be open by default. If it is not, you can open it from File > Connect window.
3. Click Manage Drivers. The Manage drivers window will appear.
4. Click

   to create a new driver, and enter the Name: "Hortonworks Hive JDBC".
5. Click

   and then browse to the Hortonworks JDBC Driver for Apache Hive package that you downloaded earlier. Next, select the JDBC Driver JAR files in the package.
6. When prompted, select the "com.simba.hive.jdbc41.HS2Driver" driver.
7. For the Sample URL, enter: jdbc:hive2://${GATEWAY_HOST}:8443/
8. After performing these steps, your configuration should look similar to:
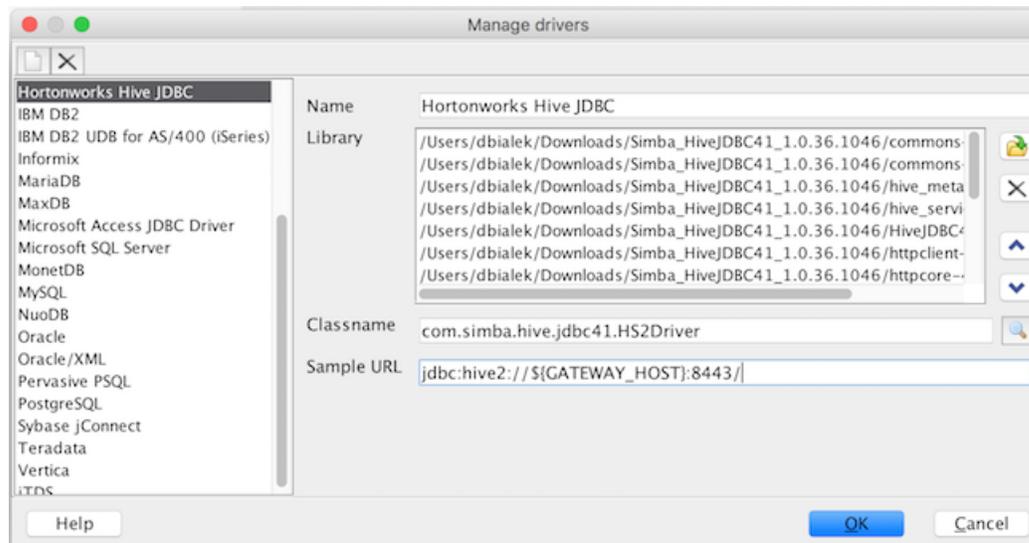
9. Click OK to save the driver.

Step 2: Create a Connection to Hive

1. From the Select Connection Profile window, select the "Hortonworks Hive JDBC" from the Driver dropdown.
2. For URL , enter the URL to the cluster instance where gateway is installed, such as jdbc:hive2://52.52.98.57:8443/ (where 52.52.98.57 is the public hostname of your gateway node).
3. For Username and Password, enter the credentials that you created when creating your cluster.
4. Click Extended Properties and add the following properties:

Property	Value
ssl	1
transportMode	http
httpPath	Provide "${CLUSTER_NAME}/${TOPOLOGY_NAME}/hive". For example hive-test/db-proxy/hive
sslTrustStore	Enter the path to the gateway.jks file. This file was generated when you downloaded the SSL certificate.
trustStorePassword	Enter the GATEWAY_JKS_PASSWORD that you specified when you downloaded the SSL certificate.

After performing these steps, your configuration should look similar to:
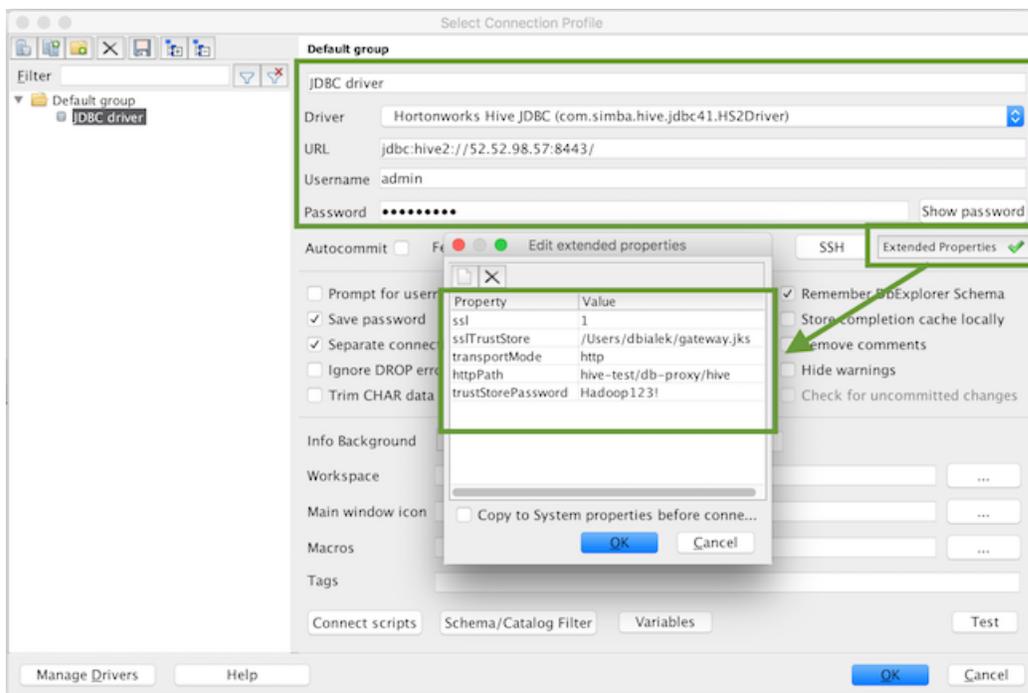
5. Click OK to save the properties.
6. Click Test to confirm a connection can be established.
7. Click OK to make the connection and start using SQL Workbench to query Hive.

Related Information
SQL Workbench
Download SSL certificate

Example: Tableau

Tableau is a business intelligence tool for interacting with and visualizing data via SQL. Connecting Tableau to Hive requires the use of an ODBC driver. In this example, we provide high-level steps required to set up Tableau to access Hive.

Prerequisite:

Download SSL certificate

Step 1: Install ODBC Driver

1. Download the Hortonworks ODBC Driver for Apache Hive from https://hortonworks.com/downloads/#addons.

2. Next, extract and install the driver.

Step 2: Launch Tableau and Connect to Hive

1. Launch Tableau. If you do not already have Tableau, you can download a trial version from https://www.tableau.com/trial/download-tableau.

2. In Tableau, create a connection to a "Hortonworks Hadoop Hive" server. Enter the following:

Property	Value
Server	Enter the public hostname of your controller node instance.
Port	8443
Type	HiveServer2
Authentication	Username and Password
Transport	HTTP
Username	Enter the cluster username created when creating your cluster
Password	Enter the cluster password created when creating your cluster
HTTP Path	Provide "${CLUSTER_NAME}/${TOPOLOGY_NAME}/hive". For example hive-test/db-proxy/hive

3. Check the Require SSL checkbox.

4. Click on the text underneath the checkbox to add a configuration file link.

5. Specify to use a custom SSL certificate, and then browse to the SSL certificate gateway.pem file that was generated when you downloaded the SSL certificate as a prerequisite.

6. Click OK.

After performing these steps, your configuration should look similar to:

7. Click Sign In and you will be connected to Hive.

Related Information
SQL Workbench
Download SSL certificate