CDP Private Cloud Base 7.0.3

Cloudera Authorization

Date published: 2020-11-30
Date modified: 2020-11-30

https://docs.cloudera.com/

Legal Notice

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

Overview

Authorization is concerned with who or what has access or control over a given resource or service. Since Hadoop merges together the capabilities of multiple varied, and previously separate IT systems as an enterprise data hub that stores and works on all data within an organization, it requires multiple authorization controls with varying granularities. In such cases, Hadoop management tools simplify setup and maintenance by:

- Tying all users to groups, which can be specified in existing LDAP or AD directories.
- Providing role-based access control for similar interaction methods, like batch and interactive SQL queries. For example, Apache Ranger permissions apply to Hive (HiveServer2) and Impala.

CDP currently provides the following forms of access control:

- Traditional POSIX-style permissions for directories and files, where each directory and file is assigned a single owner and group. Each assignment has a basic set of permissions available; file permissions are simply read, write, and execute, and directories have an additional permission to determine access to child directories.
- Apache HDFS ACLs provide fine-grained control of permissions for HDFS files by allowing you to set different permissions for specific named users or named groups.
- Apache HBase uses ACLs to authorize various operations (READ, WRITE, CREATE, ADMIN) by column, column family, and column family qualifier. HBase ACLs are granted and revoked to both users and groups.
- Access control with Apache Ranger.

Cloudera Manager User Roles

Access to Cloudera Manager features is controlled by user accounts that specify an authentication mechanism and one or more user roles. User roles determine the tasks that an authenticated user can perform and the features visible to the user in the Cloudera Manager Admin Console. In addition to the default user roles, you can create user roles that apply only to specific clusters.

Documentation for Cloudera Manager administration and management tasks indicate user roles required to perform the task.

Note: All possible user roles are available with Cloudera Enterprise. Cloudera Express provides Read-Only and Full Administrator user roles only. When a Cloudera Enterprise Data Hub Edition trial license expires, only users with Read-Only and Full Administrator roles can log in to Cloudera Manager. A Full Administrator must change user accounts with other roles to Read-Only or Full Administrator before such users can log in.

Displaying Your Roles

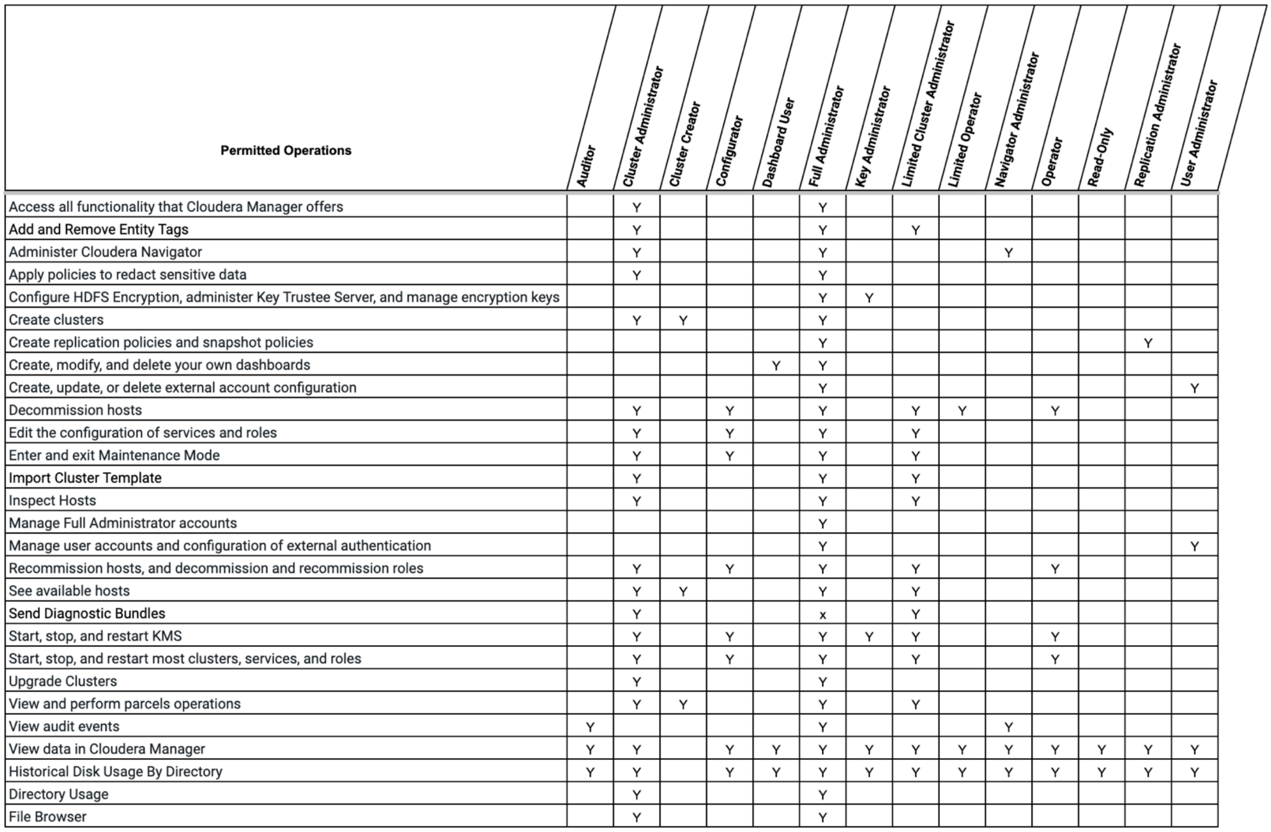
To view your roles, perform the following step:

1. In the Cloudera Manager Admin Console, select <username>My Profile.

Default User Roles

By default, Cloudera Manager ships with user roles that have privileges for all clusters managed by Cloudera Manager. You can create roles that are a combination of a default user role and privileges on a specific cluster. For more information about this type of role, see User Roles with Privileges for a Cluster on page 5.

The following table describes the actions each user role can perform:



Permitted Operations	Auditor	Cluster Administrator	Cluster Creator	Configurator	Dashboard User	Full Administrator	Key Administrator	Limited Cluster Administrator	Limited Operator	Navigator Administrator	Operator	Read-Only	Replication Administrator	User Administrator
Access all functionality that Cloudera Manager offers		Y				Y								
Add and Remove Entity Tags		Y				Y		Y						
Administer Cloudera Navigator		Y				Y				Y				
Apply policies to redact sensitive data		Y				Y								
Configure HDFS Encryption, administer Key Trustee Server, and manage encryption keys						Y	Y							
Create clusters		Y	Y			Y								
Create replication policies and snapshot policies						Y							Y	
Create, modify, and delete your own dashboards					Y	Y								
Create, update, or delete external account configuration						Y								Y
Decommission hosts		Y		Y		Y		Y	Y		Y			
Edit the configuration of services and roles		Y		Y		Y		Y						
Enter and exit Maintenance Mode		Y		Y		Y		Y						
Import Cluster Template		Y				Y		Y						
Inspect Hosts		Y				Y		Y						
Manage Full Administrator accounts						Y								
Manage user accounts and configuration of external authentication						Y								Y
Recommission hosts, and decommission and recommission roles		Y		Y		Y		Y			Y			
See available hosts		Y	Y			Y		Y						
Send Diagnostic Bundles		Y				x		Y						
Start, stop, and restart KMS		Y		Y		Y	Y	Y			Y			
Start, stop, and restart most clusters, services, and roles		Y		Y		Y		Y			Y			
Upgrade Clusters		Y				Y								
View and perform parcels operations		Y	Y			Y		Y						
View audit events	Y					Y				Y				
View data in Cloudera Manager	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Historical Disk Usage By Directory	Y	Y		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Directory Usage		Y				Y								
File Browser		Y				Y								

User Roles with Privileges for a Cluster

In addition to the default user roles, you can create user roles that apply only to specific clusters. Creating this new role is done by assigning a privilege for a specific cluster to a default role. When a user account has multiple roles, the privileges are the union of all the roles.

For example, the user account milton has the Limited Operator role and Read-Only role with a scope of Cluster 1. Additionally, milton has the Configurator role on Cluster 2.

On Cluster 1, milton can perform all the actions that a Limited Operator and Read-Only can.

On Cluster 2, milton can perform all the actions that a Configurator can.

The user account milton cannot perform these or any other actions on the other clusters that are managed by Cloudera Manager because the account does not have any other roles.

Another user account, edith, has the Configurator role with privileges for all clusters. This means that edith can perform the actions of the Configurator role on all clusters that Cloudera Manager manages since the scope is all clusters.

You can assign privileges for a specific cluster to the following user roles:

- Cluster Administrator
- Configurator
- Limited Operator
- Operator
- Read-Only

User roles that cannot be assigned privileges for a specific cluster apply to all clusters. For example, if edith has the Key Administrator user role, she can perform the actions of a Key Administrator on all clusters.

Figure 1: Cluster-Specific Permissions

Adding a User Role for a Specific Cluster

To create a role that has privileges for a specific cluster, perform the following steps:

1. In the Cloudera Manager Admin Console, navigate to AdministrationUsers & RolesRoles.
2. Click Add Role.
3. Specify the following:

   - Privilege: The user role and cluster you want to assign privileges for.
   - Users: The users you want to assign to this new role. You can assign users now or at a later time.
   - LDAP Group/External Program Exit Codes/SAML Attributes/SAML Script Exit Codes: The external mapping you want to assign this new role to. You can assign external mappings now or at a later time with the process described in Mapping External Authentication to a Role on page 7.

     This field is based on your authentication mode and does not appear for local users.

     Valid values for the External Program Exit Code and SAML Script Exit Code are between 0 and 127. You defined what users you want to associate with theses values when you configure your external authentication. For more information,

     If you are upgrading to Cloudera Manager 6 from Cloudera Manager 5, existing mappings are imported from Cloudera Manager 5. These imported mappings can be changed.

     The following list describes the LDAP groups imported from Cloudera Manager 5:

     - LDAP Full Administrator Groups
     - LDAP User Administrator Groups
     - LDAP Cluster Administrator Groups
     - LDAP BDR Administrator Groups
     - LDAP Configurator Groups
     - LDAP Key Administrator Groups
     - LDAP Navigator Administrator Groups
     - LDAP Operator Groups
     - LDAP Limited Operator Groups
     - LDAP Auditor Groups

     The following list describes the SAML and External Program codes imported from Cloudera Manager 5:

     - 0 - Full Administrator
     - 1 - Read-Only
     - 2 - Limited Operator
     - 3 - Operator
     - 4 - Configurator
     - 5 - Cluster Administrator
     - 6 - BDR Administrator
     - 7 - Navigator Administrator
     - 8 - User Administrator
     - 9 - Auditor
     - 10 - Key Administrator
     - 11 - Dashboard User
4. Click Add.

Mapping External Authentication to a Role

If you are using an external authentication, such as a SAML Script, you must map its information to Cloudera Manager user roles. Before you can map a role though, make sure that it exists. If it does not exist, create it by completing the steps described in Adding a User Role for a Specific Cluster on page 6.

Attention: If you do not map an External Authentication entity (such as an LDAP group) to a role, users that belong to that group will default to no access.

For example, you are using a SAML Script and want to assign user accounts that correspond with exit code 15 to a Cluster Administrator role with privileges for a cluster named cluster1.

To accomplish this, perform the following steps in the Cloudera Manager Admin Console:

1. Navigate to AdministrationUsers & RolesRoles.
2. Based on your authentication method, select LDAP Groups, SAML Attributes, SAML Scripts, or External Programs.
3. Click Add <authentication method> Mapping.
4. Fill in the value for your authentication method, such as SAML Script Exit Code, and select the role you want to map to that value from the dropdown menu.

   For SAML Scripts and External Programs, valid values are between 0 and 127.
5. Click Save.
6. Repeat this process for all the roles you want to map.

If you are upgrading to Cloudera Manager 6 from Cloudera Manager 5, existing mappings are imported from Cloudera Manager 5. These imported mappings can be changed.

The following list describes the LDAP groups imported from Cloudera Manager 5:

- LDAP Full Administrator Groups
- LDAP User Administrator Groups
- LDAP Cluster Administrator Groups
- LDAP BDR Administrator Groups
- LDAP Configurator Groups
- LDAP Key Administrator Groups
- LDAP Navigator Administrator Groups
- LDAP Operator Groups
- LDAP Limited Operator Groups
- LDAP Auditor Groups

The following list describes the SAML and External Program codes imported from Cloudera Manager 5:

- 0 - Full Administrator
- 1 - Read-Only
- 2 - Limited Operator
- 3 - Operator
- 4 - Configurator
- 5 - Cluster Administrator
- 6 - BDR Administrator
- 7 - Navigator Administrator
- 8 - User Administrator
- 9 - Auditor
- 10 - Key Administrator
- 11 - Dashboard User

Assigning Users to a Role

In addition to mapping groups, such as LDAP groups, to a user role, you can also assign individual users to a user role. If you do not assign a role, the local user defaults to no access. This means that the user cannot perform any actions on the cluster.

To add a user account to a role, perform the following steps:

1. In the Cloudera Manager Admin Console, navigate to AdministrationUsers & RolesRoles.
2. Click Assign for the role you want to modify.
3. Specify the Users or <Authentication Method Value> groups you want to assign to the role.
4. Save the changes.

Removing a User or External Mapping from a User Role

Perform the following steps to remove a user account or external mapping from a user role:

1. In the Cloudera Manager Admin Console, navigate to AdministrationUsers & RolesRoles.
2. Click Assign for the role you want to modify.
3. Click the X for each user or external mapping you want to remove from the user role and click Save.

Removing a Role

To remove a role with a specific privilege, you must first remove all the user accounts that have that role. Note that you cannot remove the default roles that Cloudera Manager ships with.

The following steps describe how to remove users and then delete the role:

1. In the Cloudera Manager Admin Console, navigate to AdministrationUsers & RolesRoles.
2. Click Assign for the role you want to modify.
3. Click the X for each user or external mapping you want to remove from the user role and click Save.
4. Click Remove.

Removing the Full Administrator User Role

Minimum Required Role: User Administrator (also provided by Full Administrator) This feature is not available when using Cloudera Manager to manage Data Hub clusters.

In some organizations, security policies may prohibit the use of the Full Administrator role. The Full Administrator role is created during Cloudera Manager installation, but you can remove it as long as you have at least one remaining user account with User Administrator privileges.

To remove the Full Administrator user role, perform the following steps.

1. Add at least one user account with User Administrator privileges, or ensure that at least one such user account already exists.
2. Ensure that there is only a single user account with Full Administrator privileges.
3. While logged in as the single remaining Full Administrator user, select your own user account and either delete it or assign it a new user role.

⚠ **Warning:** After you delete the last Full Administrator account, you will be logged out immediately and will not be able to log in unless you have access to another user account. Also, it will no longer be possible to create or assign Full Administrators.

A consequence of removing the Full Administrator role is that some tasks may require collaboration between two or more users with different user roles. For example:

- If the machine that the Cloudera Navigator roles are running on needs to be replaced, the Cluster Administrator will want to move all the roles running on that machine to a different machine. The Cluster Administrator can move any non-Navigator roles by deleting and re-adding them, but would need a Navigator Administrator to perform the stop, delete, add, and start actions for the Cloudera Navigator roles.
- In order to take HDFS snapshots, snapshots must be enabled on the cluster by a Cluster Administrator, but the snapshots themselves must be taken by a BDR Administrator.

Configuring LDAP Group Mappings

Each host that comprises a node in a Cloudera cluster runs an operating system, such as CentOS or Oracle Linux. At the OS-level, there are user:group accounts created during installation that map to the services running on that specific node of the cluster. The default shell-based group mapping provider, org.apache.hadoop.security.ShellBasedUnixGroupsMapping, handles the mapping from the local host system (the OS) to the specific cluster service, such as HDFS. The hosts authenticate using these local OS accounts before processes are allowed to run on the node.

For clusters integrated with Kerberos for authentication, the hosts must also provide Kerberos tickets before processes can run on the node. The cluster can use the organization's LDAP directory service to provide the login credentials, including Kerberos tickets, to authenticate transparently while the system runs. That means that the local user:group accounts on each host must be mapped to LDAP accounts. To map local user:group accounts to an LDAP service:

- Use tools such as SSSD (Systems Security Services Daemon) or Centrify Server Suite (see Identity and Access management for Cloudera).
- The Hadoop LdapGroupsMapping group mapping mechanism. The LdapGroupsMapping library may not be as robust a solution needed for large organizations in terms of scalability and manageability, especially for organizations managing identity across multiple systems and not exclusively for Hadoop clusters. Support for the LdapGroupsMapping library is not consistent across all operating systems.
- Do not use Winbind to map Linux user:group accounts to Active Directory. It cannot scale, impedes cluster performance, and is not supported.
- Use the same user:group mappings across all cluster nodes, for ease of management.
- Use either Cloudera Manager or the command-line process detailed below.

The local user:group accounts must be mapped to LDAP for group mappings in Hadoop. You must create the users and groups for your Hadoop services in LDAP.

To integrate the cluster with an LDAP service, the user:group relationships must be contained in the LDAP directory. The admin must create the user accounts and define groups for user:group relationships on each host.

The user and group names listed in the table are the default user:group values for CDP services.

Note: If the defaults have been changed for any service, use the custom values to create the users and configure the group for that service in the LDAP server, rather than the defaults listed in the table below. For example, you changed the defaults in the Cloudera Manager Admin Console to customize the System User or System Group setting for the service.

Cloudera Product or Component	User	Group
Cloudera Manager	cloudera-scm	cloudera-scm
Apache Accumulo	accumulo	accumulo
Apache Avro	(No default)	(No default)
Apache HBase (Master, RegionServer processes)	hbase	hbase
Apache HCatalog, WebHCat service	hive	hive
Apache Hive (HiveServer2, Hive Metastore)	hive	hive
Apache Kafka	kafka	kafka
Apache Oozie	oozie	oozie

Cloudera Product or Component	User	Group
Apache Spark	spark	spark
Apache Sqoop1	sqoop	sqoop
Apache Sqoop2	sqoop2	sqoop, sqoop2
Apache Whirr	(No default)	(No default)
Apache ZooKeeper	zookeeper	zookeeper
Impala	impala	impala, hive
Cloudera Search	solr	solr
HDFS (NameNode, DataNodes)	hdfs	hdfs, hadoop
HttpFS	httpfs	httpfs
Hue	hue	hue
Hue Load Balancer (needs apache2 package)	apache	apache
Java KeyStore KMS	kms	kms
Key Trustee KMS	kms	kms
Key Trustee Server	keytrustee	keytrustee
Kudu	kudu	kudu
Llama	llama	llama
MapReduce (JobTracker, TaskTracker)	mapred	mapred, hadoop
Parquet	(No default)	(No default)
YARN	yarn	yarn, hadoop

Using Cloudera Manager

Minimum Required Role: Configurator (also provided by Cluster Administrator, Limited Cluster Administrator , and Full Administrator)

Make the following changes to the HDFS service's security configuration:

1. Open the Cloudera Manager Admin Console and go to the HDFS service.
2. Click the Configuration tab.
3. Select ScopeHDFS (Service Wide)
4. Select CategorySecurity.
5. Modify the following configuration properties using values from the table below:

Configuration Property	Value
Hadoop User Group Mapping Implementation	org.apache.hadoop.security.LdapGroupsMapping
Hadoop User Group Mapping LDAP URL	ldap://<server>
Hadoop User Group Mapping LDAP Bind User	Administrator@example.com
Hadoop User Group Mapping LDAP Bind User Password	***
Hadoop User Group Mapping Search Base	dc=example,dc=com

Although the above changes are sufficient to configure group mappings for Active Directory, some changes to the remaining default configurations might be required for OpenLDAP.

Using Ranger to Provide Authorization in CDP

Apache Ranger manages access control through a user interface that ensures consistent policy administration across Cloudera Data Platform (CDP) components. Security administrators can define security policies at the database, table, column, and file levels, and can administer permissions for specific LDAP-based groups or individual users. Rules based on dynamic conditions such as time or geolocation can also be added to an existing policy rule. The Ranger authorization model is pluggable and can be easily extended to any data source using a service-based definition.

Once a user has been authenticated, their access rights must be determined. Authorization defines user access rights to resources. For example, a user may be allowed to create a policy and view reports, but not allowed to edit users and groups. You can use Ranger to set up and manage access to Hadoop services.

Ranger enables you to create services for specific resources (HDFS, HBase, Hive, etc.) and add access policies to those services. Ranger security zones enable you to organize service resources into multiple security zones. You can also create tag-based services and add access policies to those services. Using tag-based policies enables you to control access to resources across multiple components without creating separate services and policies in each component. You can also use Ranger TagSync to synchronize the Ranger tag store with an external metadata service such as Apache Atlas.

Related Information

Apache Ranger Authorization