Enabling Fine-grained Access Control from S3 File Browser in Hue (Preview)

Date published: 2021-10-21
Date modified: 2021-10-21

Legal Notice

This document has been released as part of a technical preview for features described herein. Technical preview components are provided as a convenience to our customers for their evaluation and trial usage. These components are provided ‘as is’ without warranty or support. Further, Cloudera assumes no liability for the usage of technical preview components, which should be used by customers at their own risk.

Contents

Introduction to fine-grained access control in Hue

CDP Public Cloud defaults to using cloud storage which introduces new challenges around managing data access across teams and individual users. The Ranger Authorization Service (RAZ) addresses these challenges, enabling users to have the equivalent fine-grained and audit capabilities in Apache Ranger they enjoyed with HDFS files in an on-prem or IaaS deployment.

Hue offers you the capability to browse S3 buckets, upload files to S3, and create tables by importing files from S3. With RAZ, you can grant fine-grained access to per-user home directories.

With RAZ enabled, only admin users can access the S3 bucket till the root. Non-admin users can only access their individual user directories to which they are granted access.

In CDP, you can enable the S3 File Browser in Hue in the following three ways:

- With IDBroker
- Without IDBroker
- With Ranger Authorization Service

Note: If you have enabled RAZ while registering your AWS environment with CDP, then Hue uses RAZ as the default mechanism for enabling the S3 File Browser.

Prerequisites for enabling S3 File Browser using RAZ

RAZ for S3 is a preview feature and is under development. You must enable RAZ while registering your environment with CDP. Follow the instructions listed in the "Fine-grained Access Control for Amazon S3 (Preview)" document to:

1. Register an AWS environment with the "**Enable Ranger authorization for AWS S3**" option enabled. You can use the CDP web interface or the CDP CLI to complete this task.
2. Create a Data Hub cluster with Data Engineering or Data Mart cluster template.
3. Create the following Ranger policies:
    a. Hadoop SQL policy (`all - database, table, column, all - url`).
    Note: You must grant permissions to individual users or groups in these Ranger policies. To grant permissions to all users, you can specify `{USER}` in the **Permission** section.
    b. S3 (cm_S3) policy (`Default: User Home`).

You must also grant appropriate permissions to the users in CDP User Management Service (UMS). For example, EnvironmentUser.

Next, you must create a “user” directory within your S3 bucket and also create sub-directories for individual users. For example, `MY-S3-BKT/user/csso_john.smith`, `MY-S3-BKT/user/csso_nick.fry`, and so on. Specify the bucket name in the **S3 Bucket** field and the directory path in the **Path** field of the cm_S3 Ranger policy.

Enabling S3 File Browser in Hue using RAZ

The S3 File Browser in Hue is not enabled by default. You can enable it by setting the properties in the Advanced Configuration Snippet (hue_safety_valve.ini file).

Steps

1. Log in to the CDP Management Console and go to your environment as an Administrator.
2. From the Data Lake tab, open Cloudera Manager.
3. Go to **Clusters** > **Ranger RAZ service** > **Instances** > **RAZ server** > **Processes** and note the value of the `fs.s3a.ext.raz.rest.host.url` property from the `core-site.xml` file. You need this to specify the value of the `api_url` property in the Hue configuration.
4. Go to **Clusters** > **Hue service** > **Configuration** > **Hue Service Advanced Configuration Snippet (Safety Valve) for hue_safety_valve.ini** and add the following lines:
```
[desktop]
app_blacklist=spark,zookeeper,hbase,impala,search,sqoop,securit
y,pig
[[raz]]
is_enabled=true
api_url=[***fs.s3a.ext.raz.rest.host.url***]

[aws]
has_iam_detection=false
[[aws_accounts]]
[[[default]]]
region=[***AWS-REGION***]
host=s3.[***AWS-REGION***].amazonaws.com
allow_environment_credentials=false

[filebrowser]
remote_storage_home=s3a://[***S3-BUCKET-NAME***]/user
```
5. Click **Save Changes**.
6. Restart the Hue service.

You should be able to view the icon for the S3 File Browser on the left assist pane on the Hue web interface.

Important: When a non-admin user clicks on the S3 File Browser from the left assist pane, they may be redirected to a 403 page. This is a known issue. To resolve this issue, they must

manually specify the path to their user home directory in the URL which should be provided by the administrator. For example: `https://[***HUE-INSTANCE-URL***]/cdp-proxy/hue/hue/filebrowser/view=**s3a://[***S3-BUCKET-NAME***]/user/csso_john.smith**`. The same path (**`s3a://[***S3-BUCKET-NAME***]/user/csso_john.smith`**) needs to be specified when using the Hue Importer.