

How to resolve cluster configuration errors while adding Hive and Iceberg data sources and choosing catalog type as "Hive" in EzPresto

 Table of contents

Issue

Environment

Cause

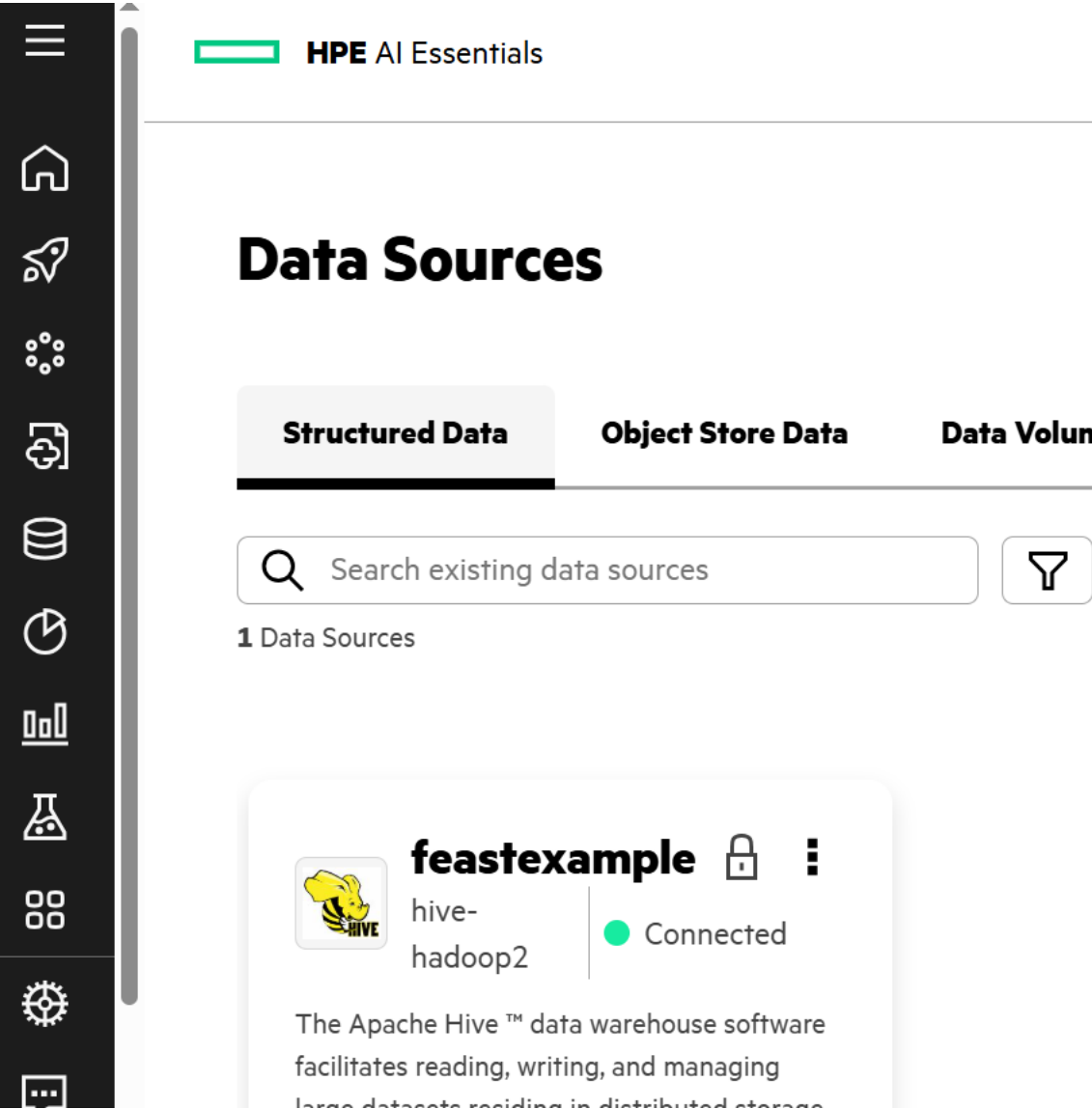
Resolution

Issue

When adding a Hive Data Source connection in EzPresto, user can face below issue which will not allow the connection to be added.

Example:

User is going to EZUA portal and going to Data Sources tab and then clicking on Add New Data Source as shown in below image.



Then click on Hive or Iceberg as shown in the below image.

Data Sources



9 data sources

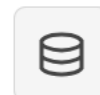


Hive



The Apache Hive™ data warehouse software facilitates reading, writing, and managing large datasets residing in distributed storage using SQL. The Hive execution engine communicates with...

Create Connection



Delta

Delta Lake is an open-source storage layer that enables building a data lake on top of any cloud compute engines, such as Amazon EMR, Databricks, or Snowflake.

Create Connection

After filling in the required details and clicking the Connect button you may get the below error.

Connect Iceberg



Name*

test1

Iceberg Catalog Type*

The catalog type for Iceberg tables.

hive



Hive Metastore*

The type of Hive metastore to use

thrift



Hive Metastore Uri*

Hive metastore URIs (comma separated).

thrift://10.163.173.108:9083

Hive Metastore Advanced Settings



Enables local copy of database table for accelerated query performance



Enable Local Snapshot Table

HR

Error:

Unable to add data source "rahul test".
test connection unsuccessful: MapR Cluster config does not match existing config

Environment

EZUA 1.5

AI Essentials 1.6

Cause

This issue arises when the customer passes in EDF configurations like the cluster name, CLDB IPs and mapr security ticket and these conflict with existing EDF configurations already present in the EzPresto pod.

Selected 3 fields

Df Cluster Details

The DF cluster entry from mapr-clusters.conf that Presto will be connecting to.

my.cluster.com secure=true m2-maprts-vm107

Hive Hdfs Authentication Type

HDFS authentication type.

MAPRSASL

Hive Hdfs Df Ticket

The DF cluster Mapr SASL ticket

.....

To resolve this configuration conflict, exec into the EzPresto mst pod and verify the contents of the EDF configuration files.

```
# kubectl get pods -n ezpresto
NAME                                READY   STATUS    RESTARTS   AGE
ezpresto-dep-mysql-6c87969c6f-xxksg 1/1     Running   0           15d
ezpresto-dep-web-7fbfb95cd6-827kr    1/1     Running   0           15d
ezpresto-sts-mst-0                    3/3     Running   0           15d
ezpresto-sts-wrk-0                    1/1     Running   0           15d

# kubectl exec -it ezpresto-sts-mst-0 -n ezpresto -- bash

[presto@ezpresto-sts-mst-0 ~]$ cd /etc/presto/catalog/
[presto@ezpresto-sts-mst-0 catalog]$ ls -lthr
total 4.5K
drwxrwxrwx. 2 5000 root    2 Nov 15 13:14 maprconf
-r--r--r--. 1 5000 root   206 Nov 15 13:14 cache.properties
-r--r--r--. 1 5000 root    19 Nov 15 13:14 jmx.properties
-r--r--r--. 1 5000 root   522 Nov 15 13:14 network.properties
drwxr--r--. 2 presto presto  0 Nov 26 11:05 tmp
-rw-r--r--. 1 presto presto 215 Jan 23 12:11 shmysql.properties
-rw-r--r--. 1 presto presto 340 Jan 23 12:47 shsnowflake.properties
-rw-r--r--. 1 presto presto 219 Jan 24 10:26 shhive.properties
-r--r--r--. 1 5000 root     6 Mar 26 10:25 masterstatus.txt

[presto@ezpresto-sts-mst-0 catalog]$ cd maprconf/
[presto@ezpresto-sts-mst-0 maprconf]$ ls -lthr
total 0
-rw-rw-rw-. 1 5000 root  0 Mar 26 10:25 mapr-clusters.conf
-rw-rw-rw-. 1 5000 root  0 Mar 26 10:25 maprtickets
```

Resolution

We can solve this problem by correcting the mapr-conf details present in the mst-pod.

Step 1. List the ezpresto pod

```
[root@m2-maprts-vm60-172 ~]# kubectl get pods -n ezpresto
NAME                                READY   STATUS    RESTARTS   AGE
ezpresto-dep-mysql-6c87969c6f-xxksg 1/1     Running   0           15d
ezpresto-dep-web-7fbfb95cd6-827kr    1/1     Running   0           15d
ezpresto-sts-mst-0                    3/3     Running   0           15d
ezpresto-sts-wrk-0                    1/1     Running   0           15d
```

```
[root@m2-maprts-vm60-172 ~]#
```

Step 2. Exec inside the mst pod

```
[root@m2-maprts-vm60-172 ~]# kubectl exec -it ezpresto-sts-mst-0 -n ezpresto -- bash
```

Step 3. cd to /etc/presto/catalog folder

```
[presto@ezpresto-sts-mst-0 /]$ cd /etc/presto/catalog/
[presto@ezpresto-sts-mst-0 catalog]$
[presto@ezpresto-sts-mst-0 catalog]$ ls -lthr
total 4.5K
drwxrwxrwx. 2 5000 root  2 Nov 15 13:14 maprconf
-r-xr-xr-x. 1 5000 root 206 Nov 15 13:14 cache.properties
-r-xr-xr-x. 1 5000 root  19 Nov 15 13:14 jmx.properties
-r-xr-xr-x. 1 5000 root 522 Nov 15 13:14 network.properties
drwxr-xr-x. 2 presto presto  0 Nov 26 11:05 tmp
-rw-r--r--. 1 presto presto 215 Jan 23 12:11 shmysql.properties
-rw-r--r--. 1 presto presto 340 Jan 23 12:47 shsnowflake.properties
-rw-r--r--. 1 presto presto 219 Jan 24 10:26 shhive.properties
-r-xr-xr-x. 1 5000 root  6 Mar 26 10:25 masterstatus.txt
```

Step 4. Check the mapr-conf folder and see the details of mapr cluster

```
[presto@ezpresto-sts-mst-0 catalog]$ cd maprconf/
[presto@ezpresto-sts-mst-0 maprconf]$ ls -lthr
total 0
-rw-rw-rw-. 1 5000 root 0 Mar 26 10:25 mapr-clusters.conf
-rw-rw-rw-. 1 5000 root 0 Mar 26 10:25 maprtickets
```

Step 5. Change the correct details

```
[presto@ezpresto-sts-mst-0 maprconf]$ vi mapr-clusters.conf
```

```
ster.com secure=true m2-maprts-vm107-173.mip.storage.hpecorp.net:7222 m2-maprts-vm108-173.mip.storage.hpecorp.net:7222 m2-
maprts-vm37-172.mip.storage.hpecorp.net:7222
~
```

Step 6. Save the file and try it again

