105-metastore-creation.md

©Creating a Cloud Metastore

This module includes all prerequisites for setting up the Cloud Metastore for the Serverless Spark Lab-

- 1. Prerequisites
- 2. Declaring Variables
- 3. Create a Metastore
- 4. Logging

№1. Prerequisites

№1.1. GCP Project Details

Note the project number and project ID. We will need this for the rest of the lab

№1.2. IAM Roles needed to create metastore

∞Grant the following permissions

- Dataproc Editor
- Service Account User

2. Declaring Variables

©2.1 Set the PROJECT_ID in Cloud Shell

Open Cloud shell or navigate to <u>shell.cloud.google.com</u> Run the below

gcloud config set project \$PROJECT_ID

©2.2 Verify the PROJECT_ID in Cloud Shell

Next, run the following command in cloud shell to ensure that the current project is set correctly:

gcloud config get-value project

2.3 Declare the variables

Based on the prereqs and checklist, declare the following variables in cloud shell by replacing with your values:

```
METASTORE_NAME= # Name of the metastore Service to be created.

REGION= # Region in which metastore service is to be created.

VPC= # The VPC network to be used.

port=9083 # Change the port number as per the Requirement.

tier=Developer # Change the tier as per the Requirement.

Metastore_Version=3.1.2 # Change the metastore version as per the Requirement,
```

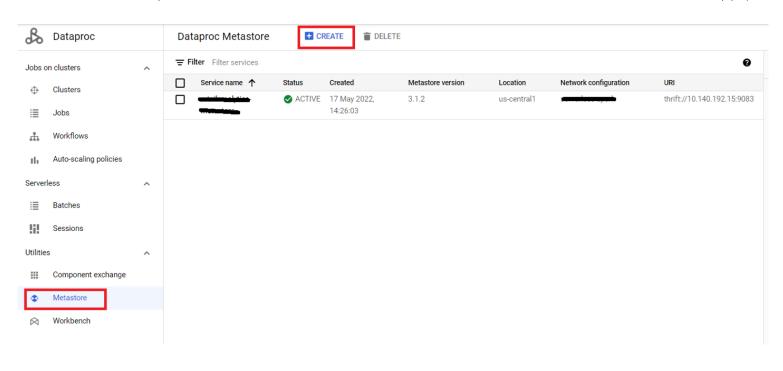
3. Create a metastore

3.1. Create a metastore through the google cloud shell

```
gcloud metastore services create $METASTORE_NAME \
    --location=$REGION \
    --network=$VPC \
    --port=$port \
    --tier=$tier \
    --hive-metastore-version=$Metastore_Version
```

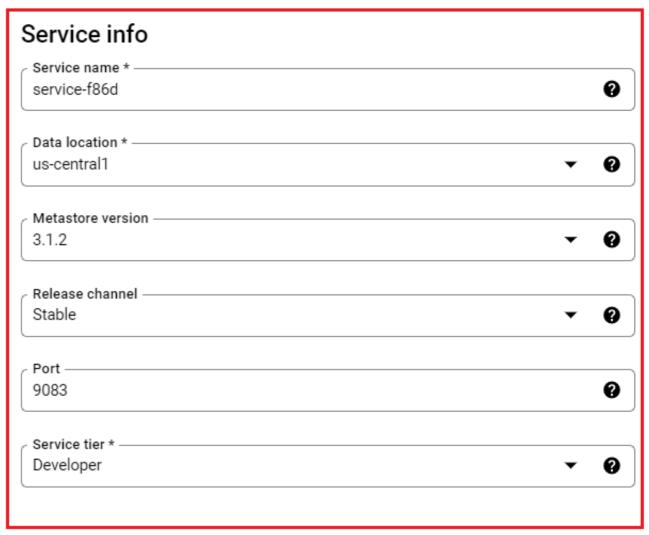
3.2. Create a metastore through the GCP console

Navigate to the Dataproc Service in your GCP project and click on Metastore Click +CREATE>Composer 2



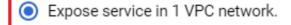


Create service



Network configuration

Use a network on which the service can be accessed. The network can belong to the same project or a different project.

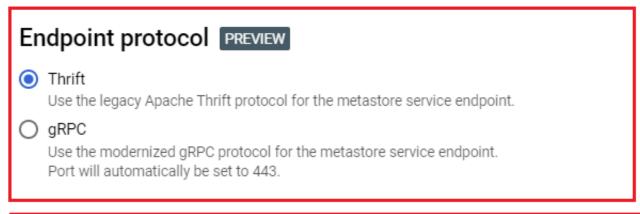


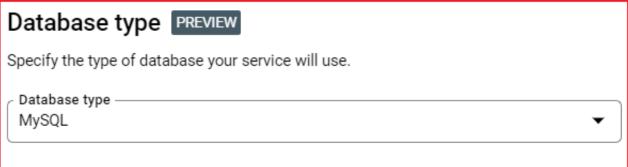


Network * —

default

②





Labels @





Next, fill in the following values in the metastore creation window:

- Service name A unique identifier for your environment
- **Data location** The region where you want to create the metastore
- Metastore Version #default
- Release channel #default
- Port #default
- Service tier #default

• Network Configuration - select the network and subnetwork with Private Google Access Enabled

- Next under **Endpoint protocol** select one of the below options: **Thrift gRPC**
- Click on **ADD LABELS** to attach the labels.
- Next, click on **Create** to create the Metastore.

4. Metastore logs

To view the metastore logs, click the 'View Logs' button on the metastore page and the logs will be shown as below:

