

# Data Pipeline How to Guide

## GCS Monitor

Version: Release 2.2

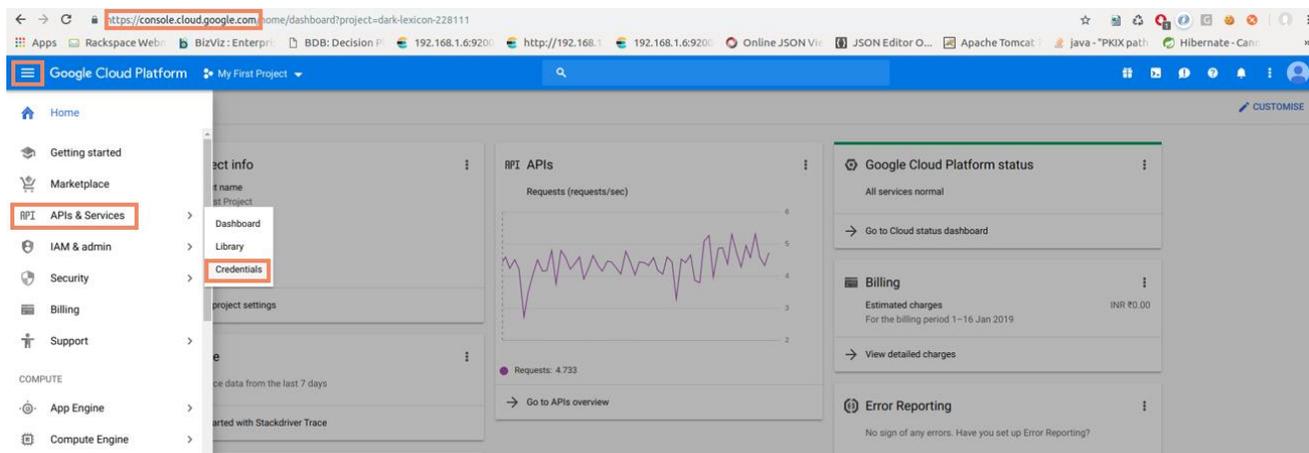
## Contents

1. Creating GCS(Google Cloud Storage) Credential .....	3
2. GCS Monitor in Data Pipeline .....	4

# 1. Creating GCS(Google Cloud Storage) Credential

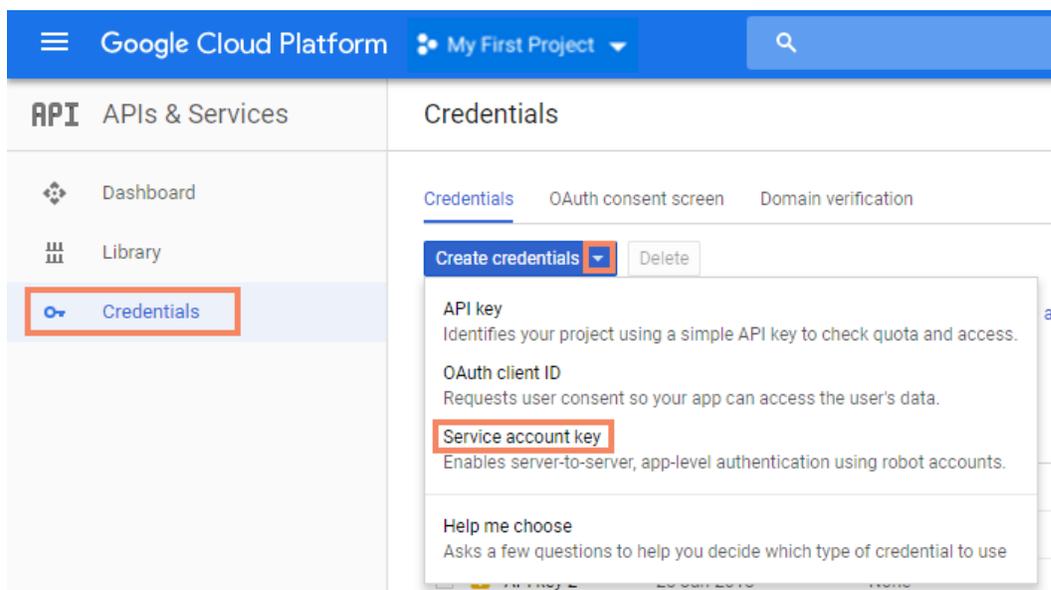
The users require Google credentials to access GCS bucket via the GCS monitor. Follow the below given steps to create a service account key:

- i) Open **Google Cloud Console** or open the link- <https://console.cloud.google.com/home>
- ii) Click on the menu button (from the top left side) to open the list of options
- iii) Select and click the **APIs & Services** option from the menu list
- iv) A context menu opens as displayed in the image below
- v) Select the **Credentials** option from the context menu

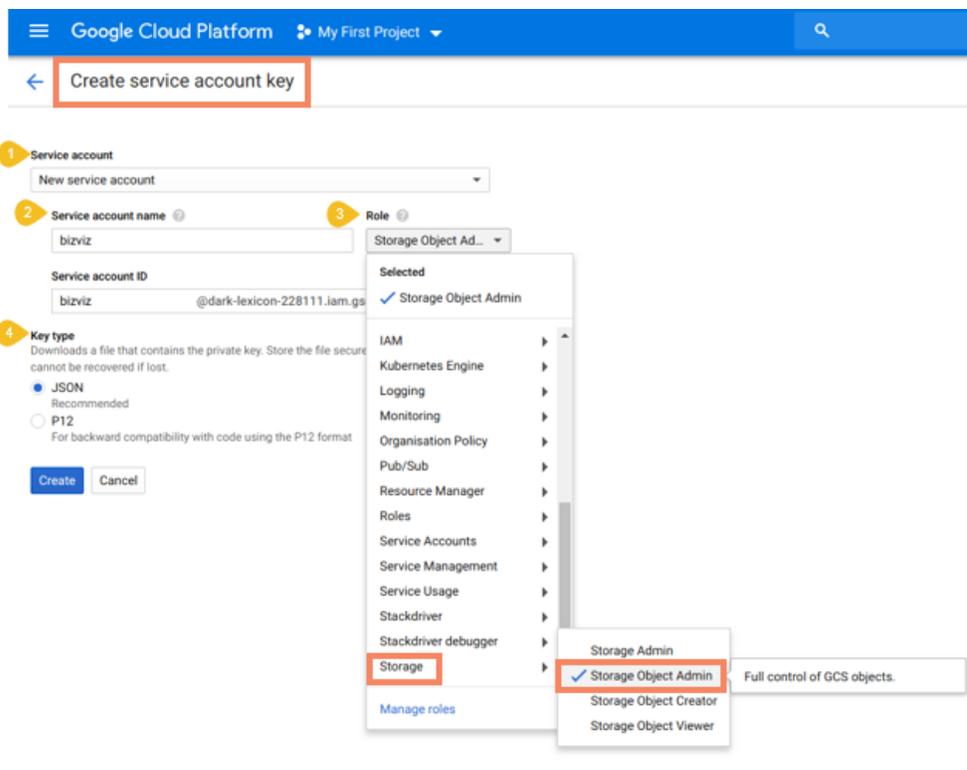


Note: The user can ignore the above-given steps to create credentials if a Service Account key is already created.

- vi) The 'Credentials' opens as displayed in the below image
- vii) Click the '**Create Credentials**' drop-down option.
- viii) A menu opens with 3 types of authorization options:
  1. API Key
  2. OAuth Client ID
  3. Service account key
- ix) Click the '**Service Account Key**' authorization option.



- x) The 'Create Service Account Key' page opens.
- xi) The users need to provide the following parameters:
  1. Services account: Select 'New Service Account' from the drop-down menu
  2. Service account name: Provide a user-defined name for the Service Account (E.g., 'bizviz' in the below given image )
  3. Role: Select the '**Storage**' option from the drop-down menu
    - a. A new context menu opens
    - b. Select '**Storage Object Admin**' from the context menu
  4. Key type: Select JSON by choosing the check mark provided next to the option
- xii) The Service Account Key gets created.



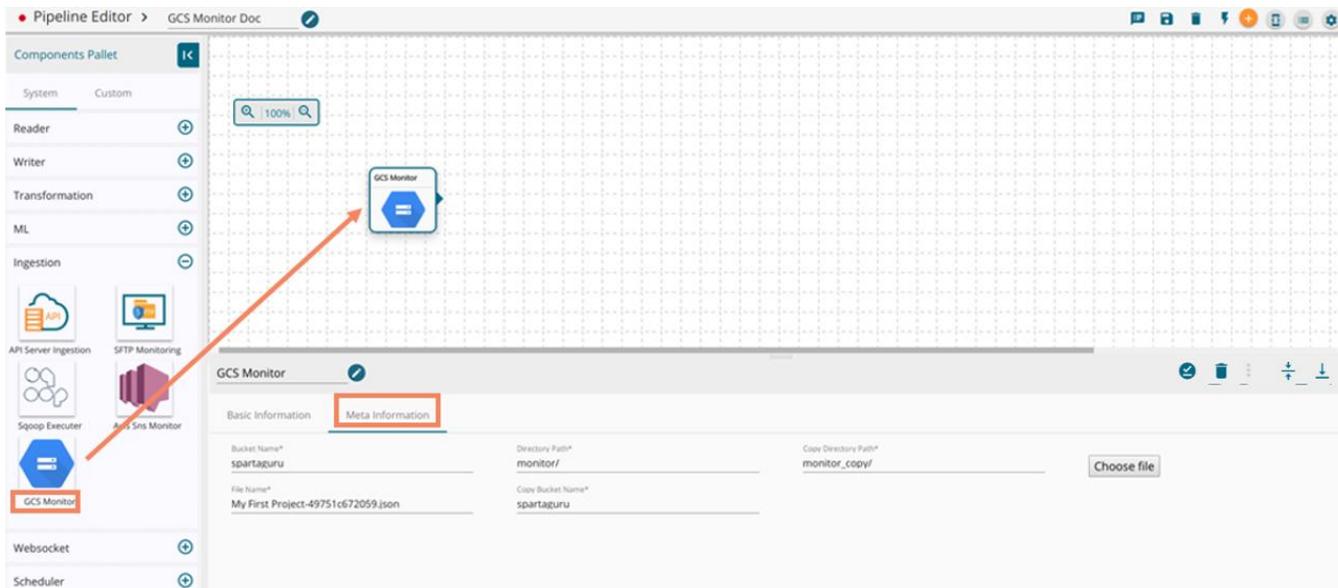
Note: After successful creation of the Service Account, the users should get Credential JSON file name according to their project name it indicates that the Service Account Key file got successfully created.

## 2. GCS Monitor in Data Pipeline

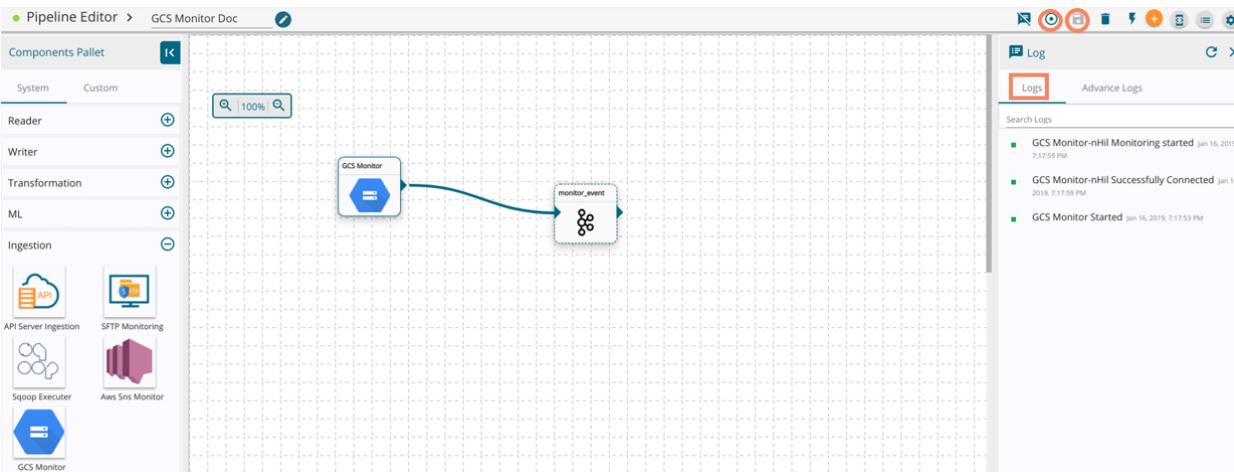
GCS Monitor monitoring one folder all the time. If any file has been uploaded, the GCS Monitor reads the file name and sends to the event. The GCS Monitor also copies that file to the copied (defined) location and deletes from the monitoring folder. It repeats the same process for all the files.

- i) Create a pipeline using the GCS Monitor component.
- ii) Navigate to the Ingestion component.
- iii) Select and Drag the GCS Monitor component to workflow editor.
- iv) Click the dragged GCS Monitor component to access the configuration fields.
- v) Fill in the required information to configure the GCS Monitor component:
  1. Bucket Name: Enter the source bucket name
  2. Directory Path: Fill monitor folder path using forward slash (/). E.g., **monitor/**

3. Copy Bucket Name: Fill the destination bucket name where you need to copy.
  4. Copy Directory Path: Fill the copy folder name where you want to copy uploaded file. E.g., **monitor\_copy/**
  5. Choose file: Upload a **Service Account Key(s)** file.
  6. File Name: After the Service Account Key file is uploaded, the file name gets auto-generated based on the uploaded file.
- vi) Click the **'Save'** option to save all the Meta Information for the GCS Monitor component.



- vii) The users need to create an event and connect it with the GCS Monitor component.
- viii) The monitor sends the filename to the event when a file gets uploaded on the monitor folder.
- ix) Save pipeline and activate the pipeline.
- x) Open the Log toggle panel to see the process of the GCS Monitor.



Note: The GCS monitor is working with GCS Reader. BDB Pipeline supports CSV and JSON file at present.