

Data Pipeline How to Guide

GCS Reader

Version: Release 2.2

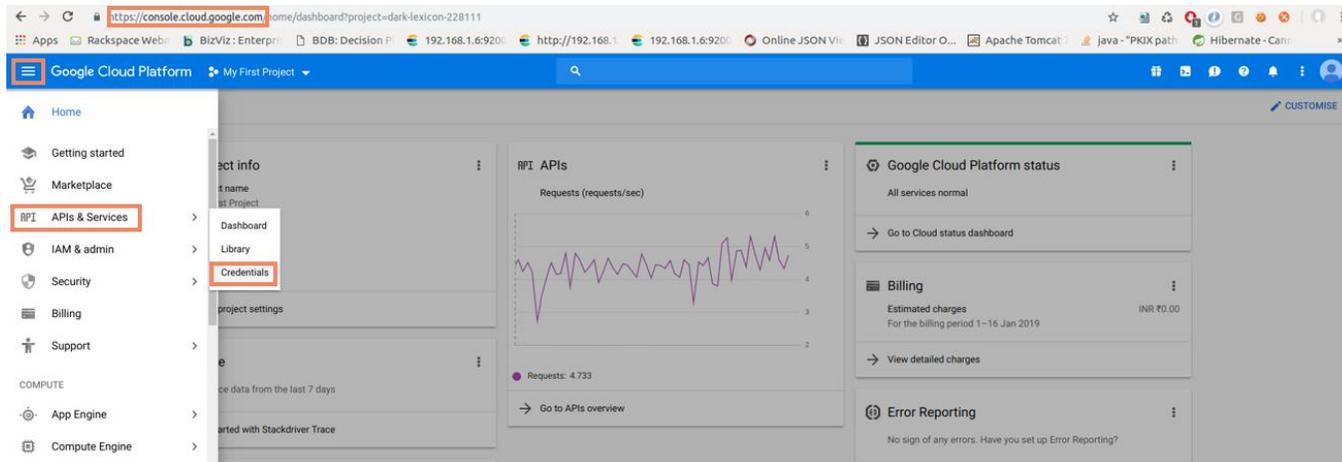
Contents

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

1. Creating GCS(Google Cloud Storage) Credential

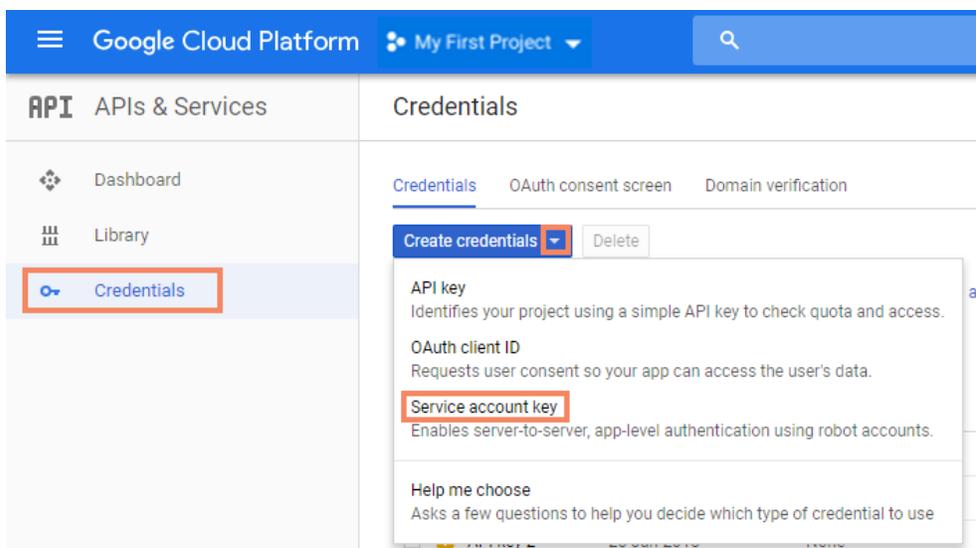
The users require Google credentials to access GCS bucket via the GCS monitor. The following are the steps to create a service account key:

- 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

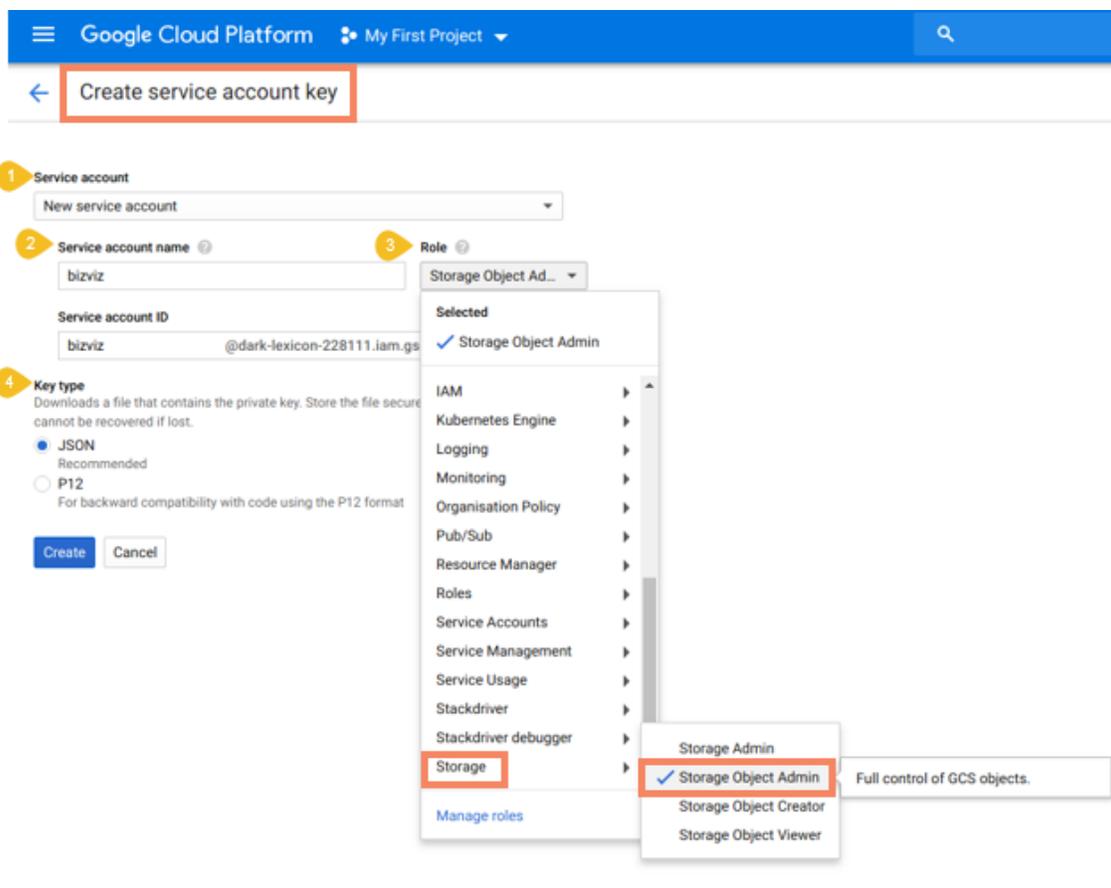


Note: The user can ignore the above-given steps to create credentials if a Service Account key is already created.

- 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 '**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 checkmark 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.

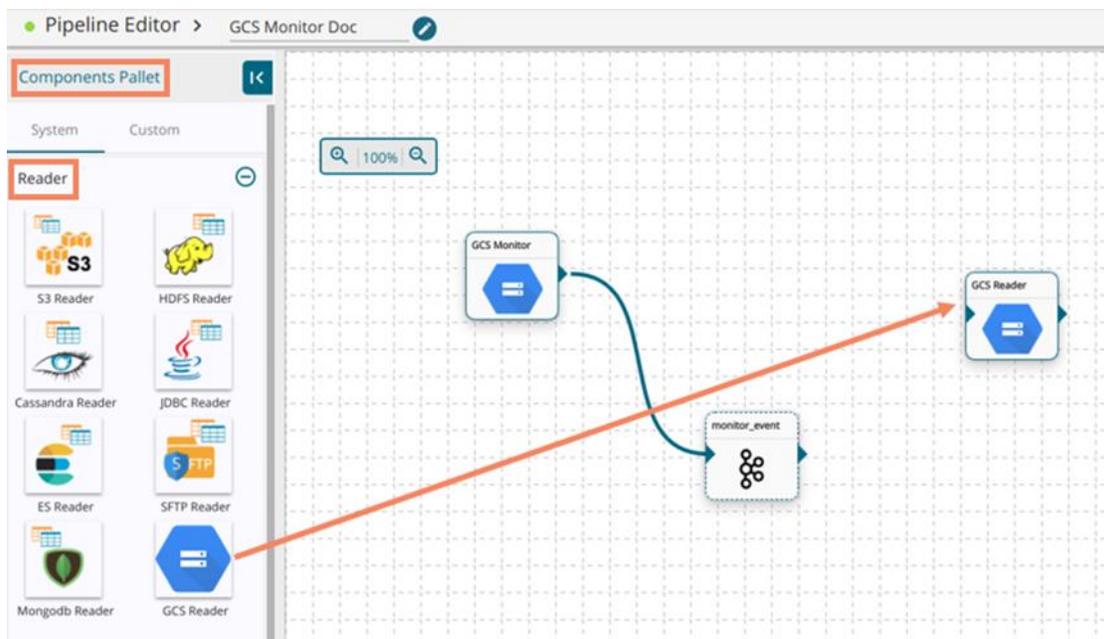
2. GCS Reader in BDB Data Pipeline

GCS Reader pulls data from the **GCS Monitor**, so the first step is to implement **GCS Monitor**.

Note: The users can refer to the GCS Monitor document for the details.

- i) Navigate to the Pipeline Workflow Editor page for an existing pipeline workflow with GCS Monitor component
- ii) Open the Reader section of the Component Pallet

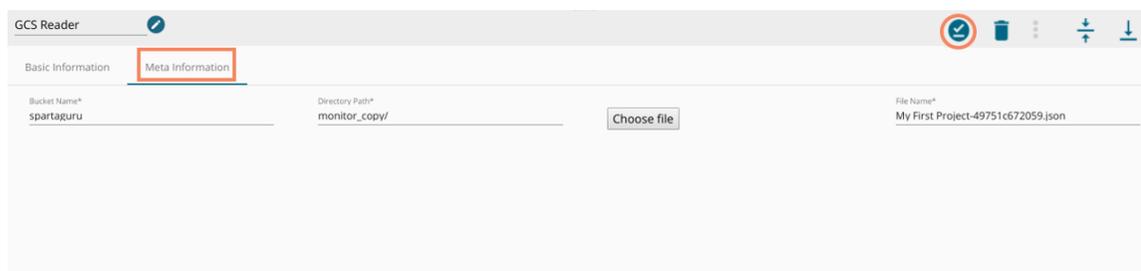
iii) Drag the GCS Reader to the workspace.



iv) Click on the dragged GCS Reader component to get configuration fields.

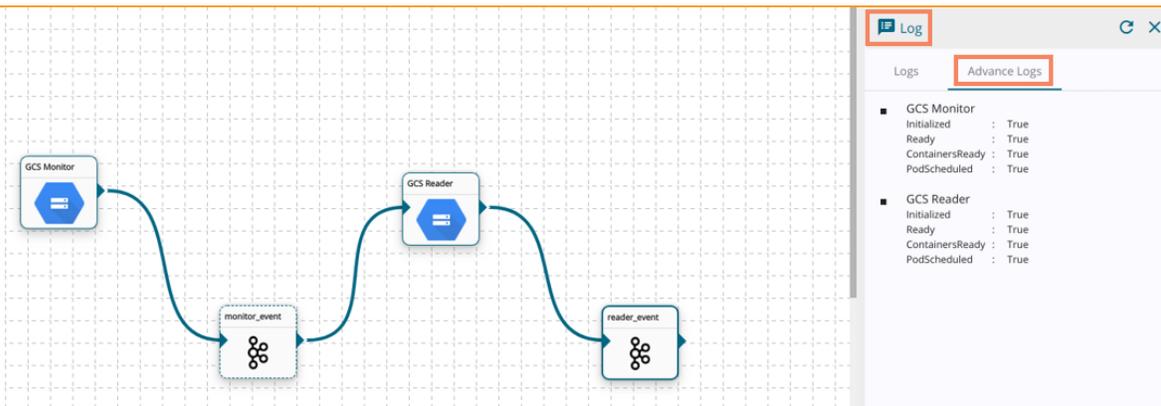
v) Provide the following information for the GCS Reader component:

- a. Bucket Name: Destination bucket name (Copy Bucket Name of the GCS Monitor)
- b. Directory Path: Destination copy folder path (Copy Directory Path of the GCS Monitor)
- c. Choose File: Upload the same **Service account keys** using the JSON file
- d. File Name: It displays the file name of the uploaded file

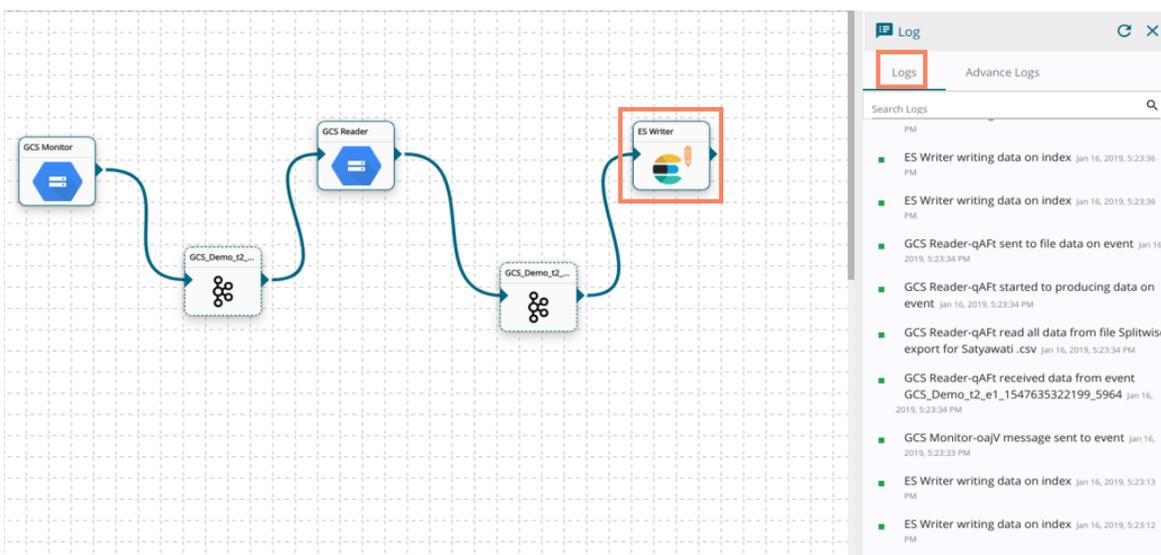


Note: Provide the same destination configuration information that was used in the **GCS Monitor**.

- vi) Create a new Event to store all the read data and connect it with the dragged GCS Reader component.
- vii) Connect the GCS Reader to the existing GCS Monitor Pipeline to form the below given workflow
- viii) Save the newly created Pipeline Workflow
- ix) Activate the Pipeline
- x) Open the Advanced Log Tab to check the component status



- xi) Upload one CSV or JSON file on monitoring location and check the Logs section displaying the ongoing operations.
- xii) Connect a Writer component to store the processed data (E.g., the ES Writer component is used in this pipeline workflow)



Note: GCS Reader component gets a message from an Event component which was initially sent by the GCS Monitor to the Event, so the users must configure the GCS Monitor component before using the GCS Reader component.