

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

🗧 🔶 C 🛔 https://console.cloud.google.com/ome/dashboard?project=dark-lexicon-228111						
III Ар	ps 🖾 Rackspace Webr 📙	BizViz : Enterpri 📑 BDB: Decision P 🏾 🌒 192.1	168.1.6:9200 🧧 http://192.168.1 🍵 192.168.1.6:9200 🔇 Online JSON	Vic 📓 JSON Editor O 🗃 Apache Tomcat 🕴 🏄 java - "PKIX path 👘	💭 Hibernate - Cann 🛛 🔹 »	
	Google Cloud Platform	🔹 My First Project 👻	٩	. (#) #	994:0	
A	Home				CUSTOMISE	
۲	Getting started	act info	I NPI APIS	Google Cloud Platform status		
Ŕ	Marketplace	t name st Project	Requests (requests/sec)	All services normal		
RPI	APIs & Services >	Dashboard	· · · · · ·	-> Go to Cloud status dashboard		
0	IAM & admin	Library Credentials	m MMMMMMMM			
	Billing	project settings		Estimated charges		
÷	Support >			For the billing period 1–16 Jan 2019		
COMP	UTE	e ce data from the last 7 days	 Requests: 4.733 	View detailed charges		
·@·	App Engine >	artad with Stanledding Trans	-> Go to APIs overview	Error Reporting		
۲	Compute Engine >			No sign of any errors. Have you set up Error Reporting?		

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.

≡	Google Cloud Platform	🔹 My First Project 👻 🔍		
API	APIs & Services	Credentials		
<≎	Dashboard	Credentials OAuth consent screen Domain verification		
ш	Library	Create credentials Delete		
0+	Credentials	API key Identifies your project using a simple API key to check quota and access.		
		OAuth client ID Requests user consent so your app can access the user's data.		
		Service account key Enables server-to-server, app-level authentication using robot accounts.		
		Help me choose Asks a few questions to help you decide which type of credential to use		



- 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.

Google Cloud Platform Solution My First	t Project 👻			٩
- Create service account key				
ervice account				
New service account	*			
Service account name 💿	Role 💿			
bizviz	Storage Object Ad 👻			
Service account ID	Selected			
bizviz @dark-lexicon-228111.iam.gs	Storage Object Admin			
y type bwnloads a file that contains the private key. Store the file secure unnot be recovered if lost. JSON Recommended P12 For backward compatibility with code using the P12 format Create Cancel Cancel	IAM Kubernetes Engine Logging Monitoring Organisation Policy Pub/Sub Resource Manager Roles Service Accounts Service Accounts Service Management Service Usage Stackdriver			
	Storage	Storage Adm	hin	
	Manage roles	Storage Obje	ct Admin Full contro	I of GCS objects.

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.

Pipeline Editor > GCS Monitor Doc 🖉				
Components	Pallet			
System	Custom			
Reader	Θ			
S3 Reader	HDES Reader	GCS Monitor GCS Reader		
	Ś			
Cassandra Reader	JDBC Reader	rootbr.vent 20		
ES Reader	SFTP Reader	5 80		
0	9			
Mongodb Reader	GCS Reader			

- 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

GCS Reader			🙆 🔋 : 🕂 🕹
Basic Information Meta Information			
Bucket Name* Spartaguru	Directory Path* monitor_copy/	Choose file	File Name* My First Project-49751c672059.json

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.