

BigQuery Health Check Metadata Download Guide



Product Version	4.8.1
Document Type	Health Check Preparation Guide
Authors	BigQuery Data source Team
Reviewer	Red Team & Architects
Approver	CTO
Total Pages	5
Document Status	Draft

Table Of Contents

1.1	Objectives	2
1.2	Architecture	2
1.3	Pre-requisite	2
1.4	Download BigQuery Metadata for health check.	3
1.5	Output Files	5

Document Version Record

Date	Version #	Author	Remarks / Reason
08-Sep-23	1.0	Dev Team	New Document

1.1 Objectives

Health check download for bigquery unravel product.

Unravel Big query Customer Solution

The diagram illustrates the Unravel Big query Customer Solution architecture. It shows the flow of data from Google BigQuery to Unravel's internal components and finally to external alerting channels.

Components and Flow:

- Google BigQuery:** The source of data, connected to the Unravel system via **Jdbc** and **Rest-API**.
- Big query Worker:** Receives data from Google BigQuery and sends it to the **Unravel Message Bus**.
- Unravel Message Bus:** A central hub that routes data to the **Loader** and **AA / Alerting** components.
- Loader:** Processes data from the message bus and loads it into the **Datastore**.
- Datastore:** Stores the data, which is then used by the **AA / Alerting** component.
- AA / Alerting:** Generates alerts based on the data in the datastore, sending them to external channels like **E-mail**, **Slack**, and **Webhooks**.

Deployment and Configuration:

- The entire Unravel system is deployed on **AWS / Azure / VM**.
- a. Need access to GCP with Permissions:** Required for the **Big query Worker** to connect to Google BigQuery.
- b. Need instance with Specification:** Required for the **Big query Worker** to run on the cloud infrastructure.
- c. https On 3000 Port for External communication:** Required for the **AA / Alerting** component to communicate with external services.
- d. Email / Slack channel with token:** Required for the **AA / Alerting** component to send notifications.

A. BigQuery Permissions/roles required :

- a. Health check data download from bigquery
 - i. For querying INFORMATION_SCHEMA.JOBS & INFORMATION_SCHEMA.JOBS_TIMELINE view data ask your administrator to grant you the [BigQuery Resource Viewer](#) (roles/bigquery.resourceViewer) IAM role on your project. This predefined role contains the **bigquery.jobs.listAll** permission, which is required to query the INFORMATION_SCHEMA.JOBS & INFORMATION_SCHEMA.JOBS_TIMELINE view.
 - ii. For querying INFORMATION_SCHEMA.RESERVATIONS view, you need the **bigquery.reservations.list** Identity and Access Management (IAM) permission for the project. Each of the following predefined IAM roles includes the required permission:
 - A. roles/bigquery.resourceAdmin
 - B. roles/bigquery.resourceEditor
 - C. roles/bigquery.resourceViewer
 - D. roles/bigquery.user
 - E. roles/bigquery.admin

1.4 Download BigQuery Metadata for health check.

Do the following to download BigQuery metadata:

1. Download the following files from this location :
<https://github.com/unraveldata-org/BigQuery-data-loader/tree/main/BigQuery%20Queries>
 - download_bigquery_jobs_data.txt
 - download_bigquery_jobs_timeline_data.txt
 - download_bigquery_assignment_data.txt
 - download_bigquery_reservation_data.txt
 - download_billing_data.txt
2. Create a gcs bucket and replace example-bucket/folder_name/ in above txt files with your valid gcs URI (Please find the steps below on How to create a gcs bucket).
3. For each region, replace the \$region with your region in above files. Replace \$admin_project_id with your admin project id.
4. In download_billing_data.txt replace, \$project , \$dataset, and \$table_name with the project, dataset and table_name where you have exported the billing data.

For Example :

\$project = **unravel-flat-rate-test**

\$dataset = **all_billing_data**

\$table_name = **gcp_billing_export_resource_v1_016A85_2365E3_979331**

5. Execute each of the above BigQuery for each region(s) on the BigQuery console on your project.
6. Check in the gcs bucket, there must be some csv files created with prefix YOUR_REGION_JOBS_*, YOUR_REGION_JOBS_TIMELINE_*, YOUR_REGION_RESERVATIONS_*, YOUR_REGION_ASSIGNMENTS_* in the gcs bucket folder that you have defined in the export bigquery options.
7. Select each of these files and click on Download.

Steps for creating a gcs bucket :

1. Go to the Cloud Storage Buckets page:
<https://console.cloud.google.com/storage/browser> in the Google Cloud console.
2. Click + Create.

3. On the Create a bucket page, enter the following information:
 - a. Name your bucket: Enter a unique name for your bucket. The name must start with a lowercase letter or number, and it can contain up to 63 characters. It can also contain dashes and periods.
 - b. Choose where to store your data: Select a location for your bucket. You can choose a location in the same region as your project, or you can choose a different region.
 - c. Choose a storage class for your data: Select a storage class for your bucket. The storage class determines how long your data is kept and how much it costs to store.
4. Click Create.
5. Create new folder in bucket [OPTIONAL]

1.5 Output Files (Sample)

upload_csv_test_GCP_BILLING_000000000000.csv	45.8 MB	CSV Document	Today at 11:23
upload_csv_test_GCP_BILLING_000000000001.csv	1.6 MB	CSV Document	Today at 12:23
upload_csv_test_GCP_BILLING_000000000002.csv	1.6 MB	CSV Document	Today at 12:23
upload_csv_test_GCP_BILLING_000000000003.csv	1.6 MB	CSV Document	Today at 12:23
upload_csv_test_GCP_BILLING_000000000004.csv	1.6 MB	CSV Document	Today at 12:23
upload_csv_test_GCP_BILLING_000000000005.csv	1.6 MB	CSV Document	Today at 12:23
upload_csv_test_US_ASSIGNMENTS_000000000000.csv	1 KB	CSV Document	Yesterday at 02:23
upload_csv_test_US_JOBS_000000000000.csv	77.2 MB	CSV Document	Yesterday at 02:23
upload_csv_test_US_JOBS_TIMELINE_000000000000.csv	36.6 MB	CSV Document	Yesterday at 02:23
upload_csv_test_US_RESERVATIONS_000000000000.csv	800 bytes	CSV Document	Yesterday at 01:31

1.6 Video Guide

Part 1 (for downloading JOBS & JOBS_TIMELINE Data):

https://drive.google.com/file/d/1KNPY6LctE4ys9uv_SHdkGLrXeJWYnv5q/view?usp=drive_link

Part 2 (for downloading ASSIGNMENTS & RESERVATION Data):

https://drive.google.com/file/d/1YoWsnvzT2bPYAkWUnKLK1EcDJcS-gtuf/view?usp=drive_link

Part 3 (for downloading BILLING Data):

https://drive.google.com/file/d/13nSEWrUMV7oLHukemh3tod58ySfYGNiu/view?usp=drive_link