Synopsis: Google Cloud Pub/Sub

- Google Cloud Pub/Sub is a fully managed messaging service used for real-time data communication.
- It follows the publish–subscribe (pub/sub) messaging model.
- Publishers send messages to a topic without knowing who will receive them.
- Subscribers receive messages asynchronously from the topic.
- It enables loose coupling between producers and consumers.
- Pub/Sub automatically scales to handle high message volumes.
- It provides high availability and fault tolerance.
- Messages are delivered with at-least-once delivery guarantee.
- Supports event-driven architectures and streaming data pipelines.
- Integrates easily with other Google Cloud services like Cloud Functions, Cloud Run, Big Query, and Dataflow.
- Useful for real-time analytics, log processing, notifications, and microservices communication.
- Reduces system dependency and improves performance and reliability.

TASK: Pull messages from a Pub/Sub topic and store them in a Big Query table using Cloud Functions Gen2.

1. Introduction

- Google Cloud Pub/Sub is a fully managed messaging service used for real-time event delivery.
- It enables asynchronous communication between distributed systems.
- This task uses Pub/Sub to stream events into Big Query for analysis.

2. Problem Statement

- Traditional systems struggle to handle real-time event ingestion at scale.
- Tight coupling between data producers and consumers reduces system flexibility.
- There is a need for a scalable, serverless solution to ingest and store event data.

3. Objective of the Task

- To capture event messages in real time using Pub/Sub.
- To process incoming messages using Cloud Functions Gen2.
- To store processed event data into Big Query for analytics.

4. System Architecture

- Publisher sends event messages to a Pub/Sub topic.
- Pub/Sub Topic acts as an event buffer.
- Cloud Function (Gen2) is triggered automatically on message arrival.
- Big Query stores the structured event data.

5. Working Principle

- Events are published to a Pub/Sub topic in JSON format.
- Pub/Sub delivers the message as a CloudEvent.
- Cloud Function decodes and processes the message.
- Extracted fields are inserted into a Big Query table with a timestamp.

6. Technologies Used

- Google Cloud Pub/Sub
- Cloud Functions (Gen2)
- Big Query
- Python
- Google Cloud IAM

7. Advantages of the System

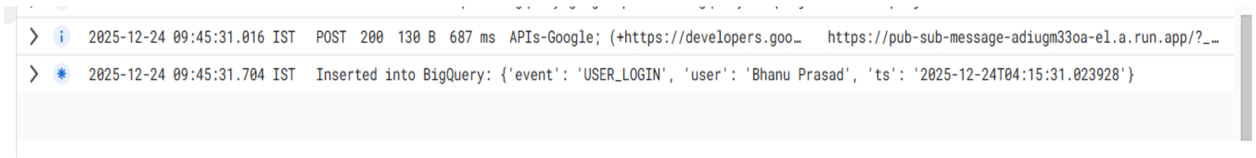
- Fully serverless and auto-scaling
- Loose coupling between components
- High reliability and fault tolerance
- Supports real-time analytics
- No infrastructure management required

8. Security and Permissions

- IAM roles control access to Big Query and Pub/Sub.
- Cloud Function uses a service account with limited permissions.
- Ensures secure and controlled data access.

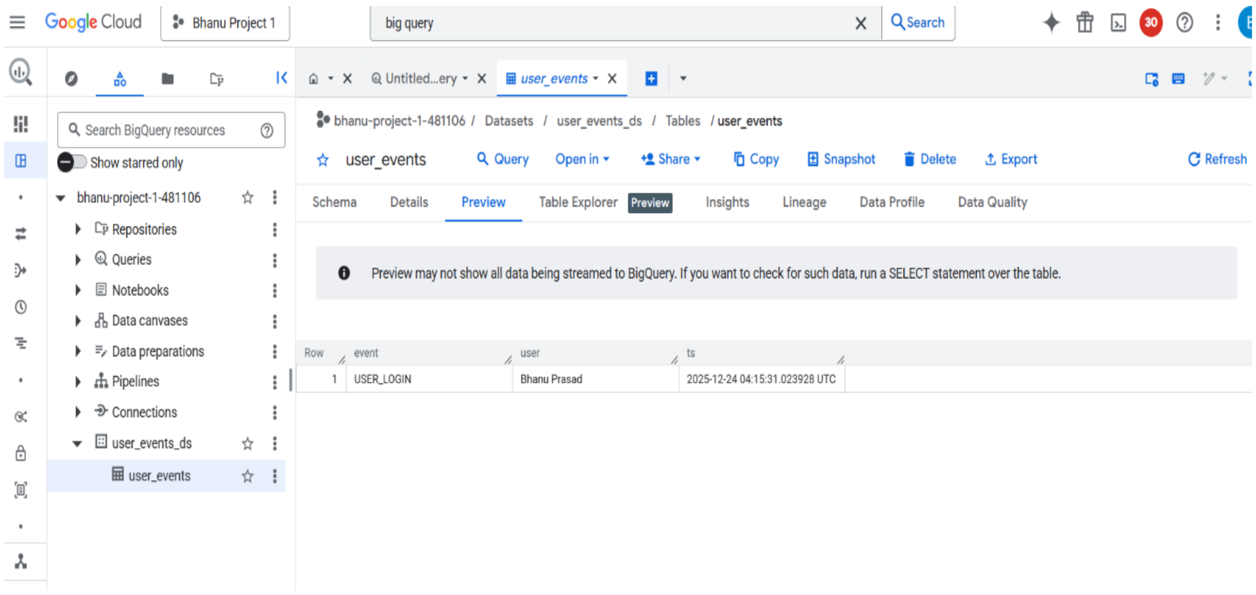
9. Implementation Results and Screenshots

- The Pub/Sub topic was successfully created and tested using sample event messages.
- Cloud Functions Gen2 was deployed and triggered automatically upon message publication.
- Event messages were decoded and processed without errors.
- Processed data was successfully inserted into the Big Query table.
- Screenshots are included to demonstrate:
    - Pub/Sub message publishing
    - Cloud Function execution logs
    - Successful data insertion into Big Query



```
2025-12-24 09:45:31.016 IST   POST  200  130 B  687 ms  APIs-Google; (+https://developers.goo…   https://pub-sub-message-adiugm33oa-el.a.run.app/?_…
2025-12-24 09:45:31.704 IST   Inserted into BigQuery: {'event': 'USER_LOGIN', 'user': 'Bhanu Prasad', 'ts': '2025-12-24T04:15:31.023928'}
```

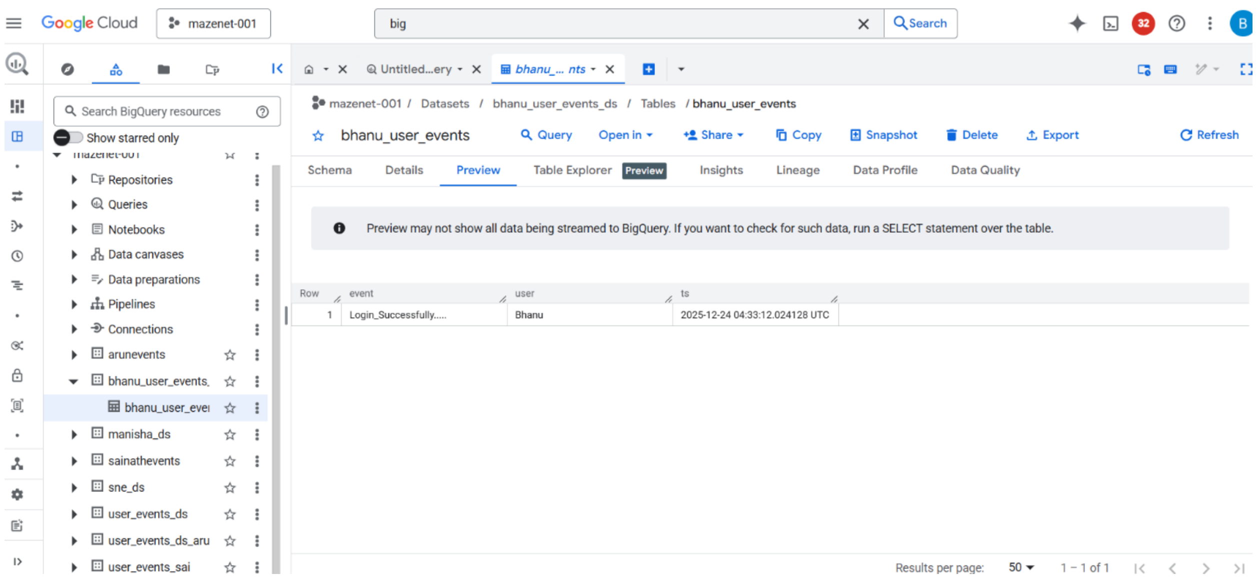
Cloud Function Logs



Preview may not show all data being streamed to BigQuery. If you want to check for such data, run a SELECT statement over the table.

Row	event	user	ts
1	USER_LOGIN	Bhanu Prasad	2025-12-24 04:15:31.023928 UTC

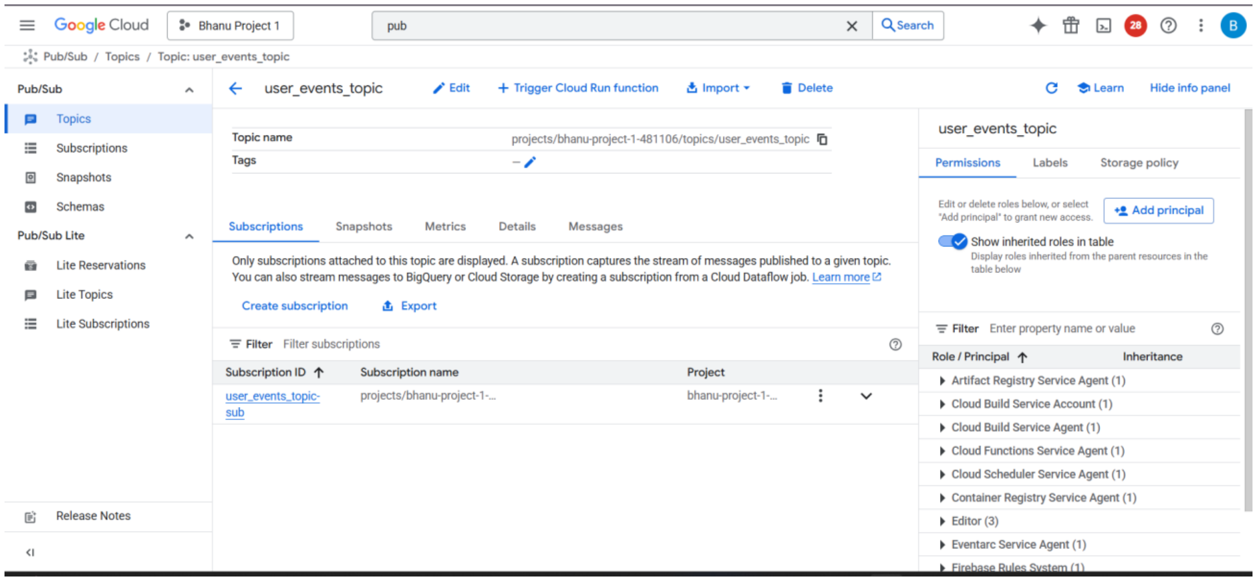
Big Query Table



Preview may not show all data being streamed to BigQuery. If you want to check for such data, run a SELECT statement over the table.

Row	event	user	ts
1	Login_Successfully.....	Bhanu	2025-12-24 04:33:12.024128 UTC

Results per page: 50 1 – 1 of 1



Topic name	projects/bhanu-project-1-481106/topics/user_events_topic
Tags	–

Only subscriptions attached to this topic are displayed. A subscription captures the stream of messages published to a given topic. You can also stream messages to BigQuery or Cloud Storage by creating a subscription from a Cloud Dataflow job. Learn more

Subscription ID	Subscription name	Project
user_events_topic-sub	projects/bhanu-project-1-...	bhanu-project-1-...

user_events_topic

Edit or delete roles below, or select "Add principal" to grant new access.

Show inherited roles in table
Display roles inherited from the parent resources in the table below

Role / Principal	Inheritance
Artifact Registry Service Agent (1)	
Cloud Build Service Account (1)	
Cloud Build Service Agent (1)	
Cloud Functions Service Agent (1)	
Cloud Scheduler Service Agent (1)	
Container Registry Service Agent (1)	
Editor (3)	
Eventarc Service Agent (1)	
Firebase Rules System (1)	

Pub/Sub Topic

10. Conclusion

- The Pub/Sub based architecture provides an efficient solution for real-time event ingestion.
- Integrating Pub/Sub with Cloud Functions and Big Query enables scalable data processing.