

CS4830 - BIG DATA LABORATORY

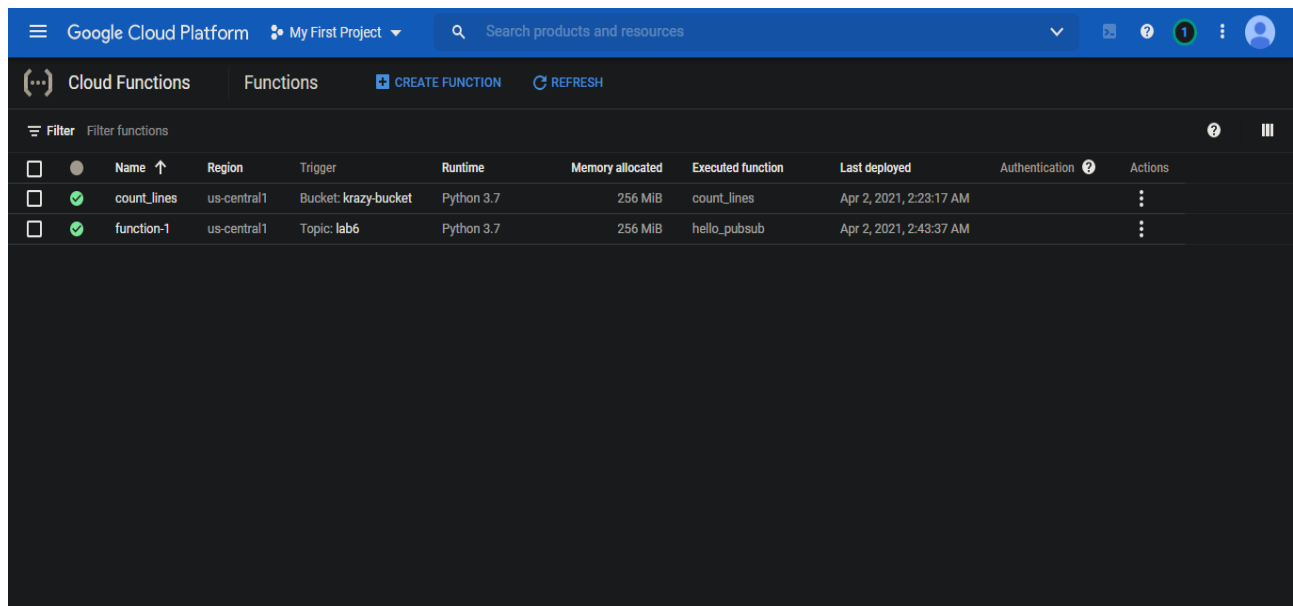
LAB 6 ASSIGNMENT

REPORT

Sahil Ahmad Ansari
ME17B065

1. Question 1

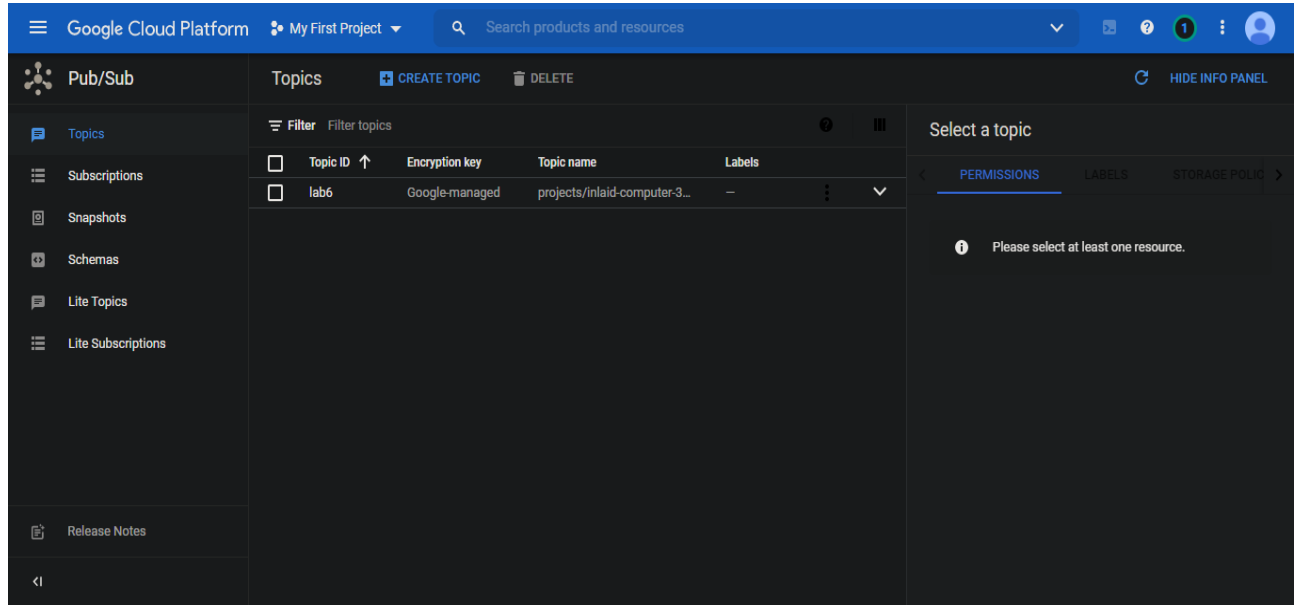
Two cloud functions were created for this task. The first one gets triggered whenever a file is added to the respective bucket and the second one gets triggered whenever a message is posted to the respective topic. The first cloud function simply pushes the name of the file added to the bucket to the pub/sub topic named 'lab6'. The second cloud function is a subscriber to the topic 'lab6' and it downloads the file pushed to the bucket using the filename and counts the number of lines in it. The following are the required screenshots:



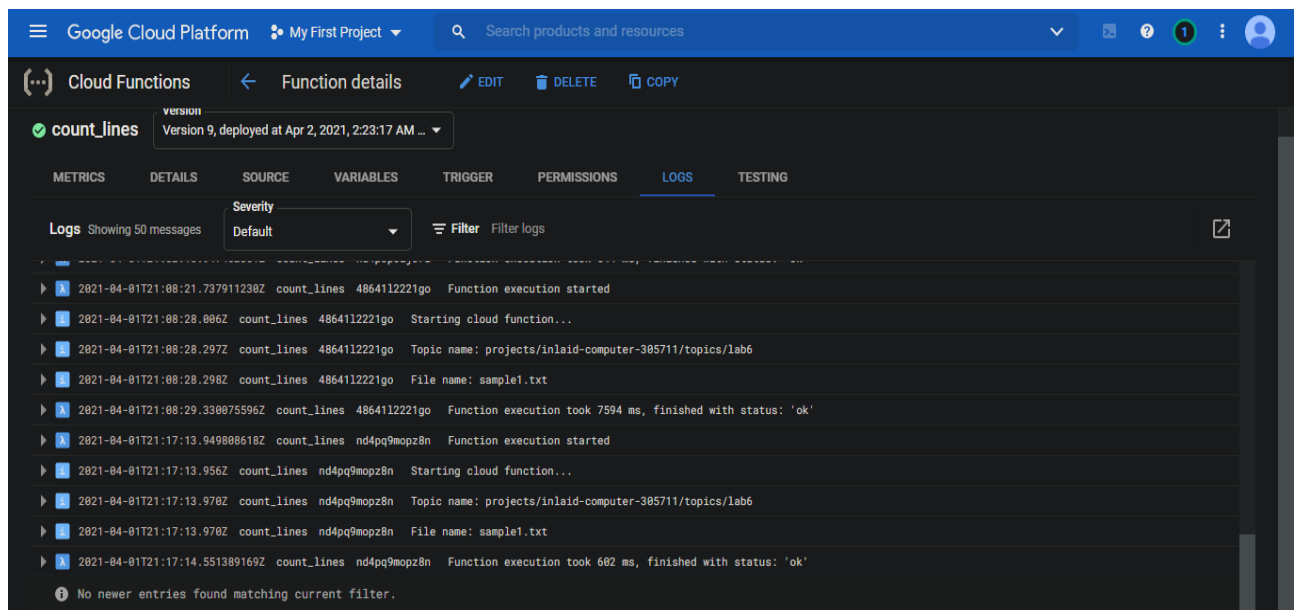
The screenshot shows the Google Cloud Platform interface for Cloud Functions. The top navigation bar includes 'Google Cloud Platform', 'My First Project', and a search bar. Below the navigation bar, the 'Cloud Functions' section is active, showing a list of functions. The table below lists the functions:

	Name	Region	Trigger	Runtime	Memory allocated	Executed function	Last deployed	Authentication	Actions
<input type="checkbox"/>	count_lines	us-central1	Bucket: crazy-bucket	Python 3.7	256 MiB	count_lines	Apr 2, 2021, 2:23:17 AM		⋮
<input type="checkbox"/>	function-1	us-central1	Topic: lab6	Python 3.7	256 MiB	hello_pubsub	Apr 2, 2021, 2:43:37 AM		⋮

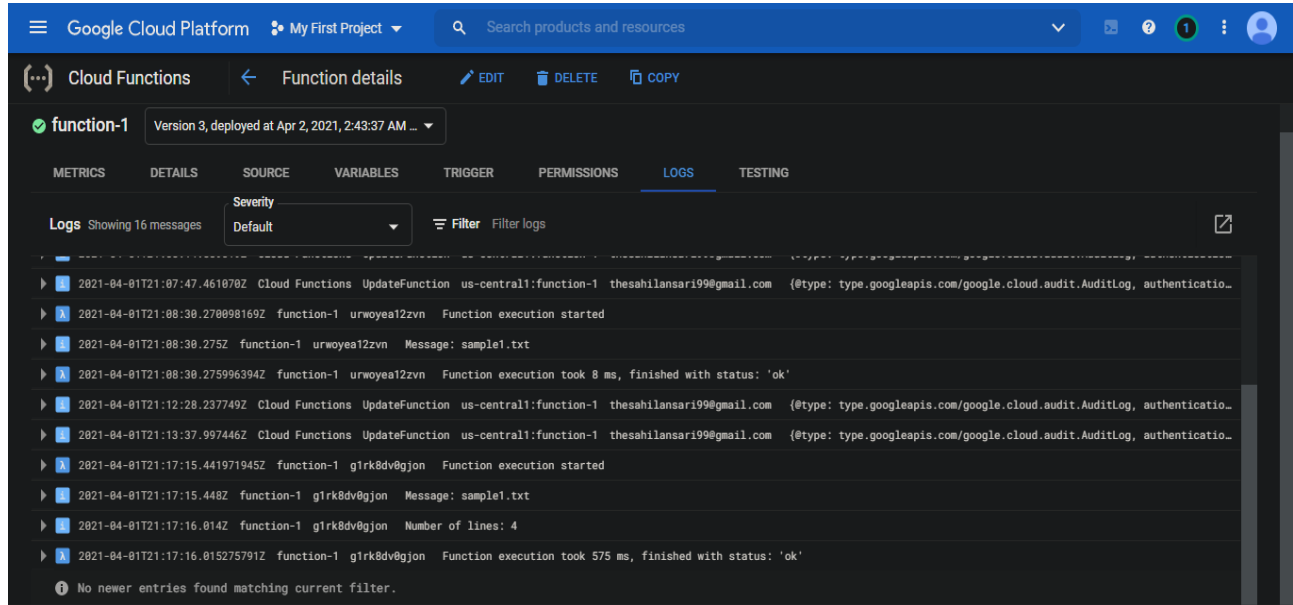
(a) Cloud Functions



(a) Screenshot of topic in pub/sub



(a) Screenshot of logs for first cloud function



(a) Screenshot of logs for second cloud function

The file named sample1.txt was uploaded to the bucket named 'krazy-bucket'. The second cloud function detected 4 lines in the file uploaded.

2. Question 2

Pull and Push subscriptions are two mechanisms for message delivery in Pubsub.

a. Pull Subscription

In pull subscription, the subscriber application initiates requests to the Pub/Sub server to retrieve messages. The subscribing application explicitly calls the pull method, which requests messages for delivery. The Pull subscription should be used when there is a large volume of messages and the efficiency and throughput of message processing is critical. A pull subscription will be useful in applications such as a news website, which will pull all the news posted to a topic.

b. Push Subscription

In push subscription, Pub/Sub initiates requests to the subscriber application to deliver messages. The Pub/Sub server sends each message as an HTTPS request to the subscriber application at a pre-configured endpoint. The endpoint acknowledges the message by returning an HTTP success status code. A push subscription must be used when multiple topics must be processed by the same webhook. The push endpoint can also be a load balancer. A push subscription will be useful in applications such as Notification engine, where message needs to be pushed to the notification engine to send out notifications to the respective clients.