PART – 1 REPORT

Real-time Underwater Exploration and Monitoring

In this use case, an innovative project called "MarineGuard" is dedicated to the real-time monitoring and conservation of marine ecosystems. MarineGuard aims to deploy a network of advanced sensors and devices in various marine locations to collect crucial environmental data, ensuring the protection and sustainable management of ocean resources.

MarineGuard plants a range of underwater sensors, including underwater cameras, hydrophones, and water quality probes, distributed strategically across different marine environments across the globe. These sensors continuously monitor and gather real-time data on water quality, marine species behavior, and environmental conditions. The underwater sensor data will be securely transmitted to AWS IoT Core. Collected data is seamlessly transmitted as records to AWS Kinesis Data Streams, which acts as the central data gathering service.

To optimize data processing and analysis, AWS Kinesis Data Streams partition the incoming data into multiple shards based on unique partition keys derived from identifiers associated with each marine location or specific aspects of marine life. This ensures even data distribution and streamlines subsequent analysis and decision-making processes for each targeted area or marine species. The partitioned data undergoes parallel processing and efficient distribution, maximizing the value of the collected information.

The streaming data is further processed and automatically transformed by AWS Kinesis Data Firehose, a service that prepares the data before delivery. It handles essential data transformations such as compression, encoding, and conversion to various formats. AWS Kinesis Data Analytics plays a pivotal role in MarineGuard's real-time data analysis. It ingests data from AWS Kinesis Data Firehose enabling comprehensive and in-depth analytics. The prepared data is stored in Amazon S3 for long-term archival, facilitating historical analysis and data retention. Additionally, the data can be seamlessly transferred to Amazon Redshift for data warehousing and advanced analytics.

Following the data analytics process, AWS QuickSight steps in, providing an interactive and user-friendly platform for data visualization and business intelligence. By creating interactive dashboards, charts, graphs, and maps from the processed data, AWS QuickSight empowers our team to communicate findings, identify trends, and make informed decisions. The seamless integration of AWS QuickSight into the data pipeline enhances our ability to optimize marine conservation efforts, safeguard marine ecosystems, and promote sustainable management practices globally.

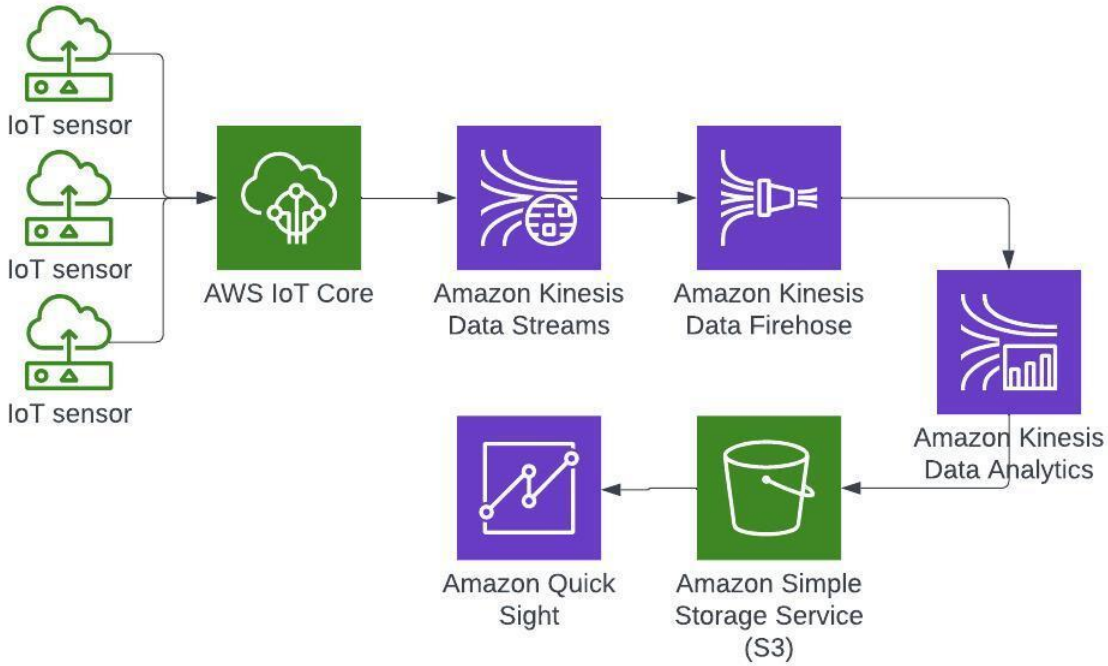

Figure 1:Architecture for developing ocean data analysis using amazon cloud services [12]

PART – 2 REPORT

Github Link : https://git.cs.dal.ca/pusuluru/csci-5410-summer-23-b00913674.git

1. My application uploading files to S3 bucket (sampledata-b00913674) for every 100 milli-seconds

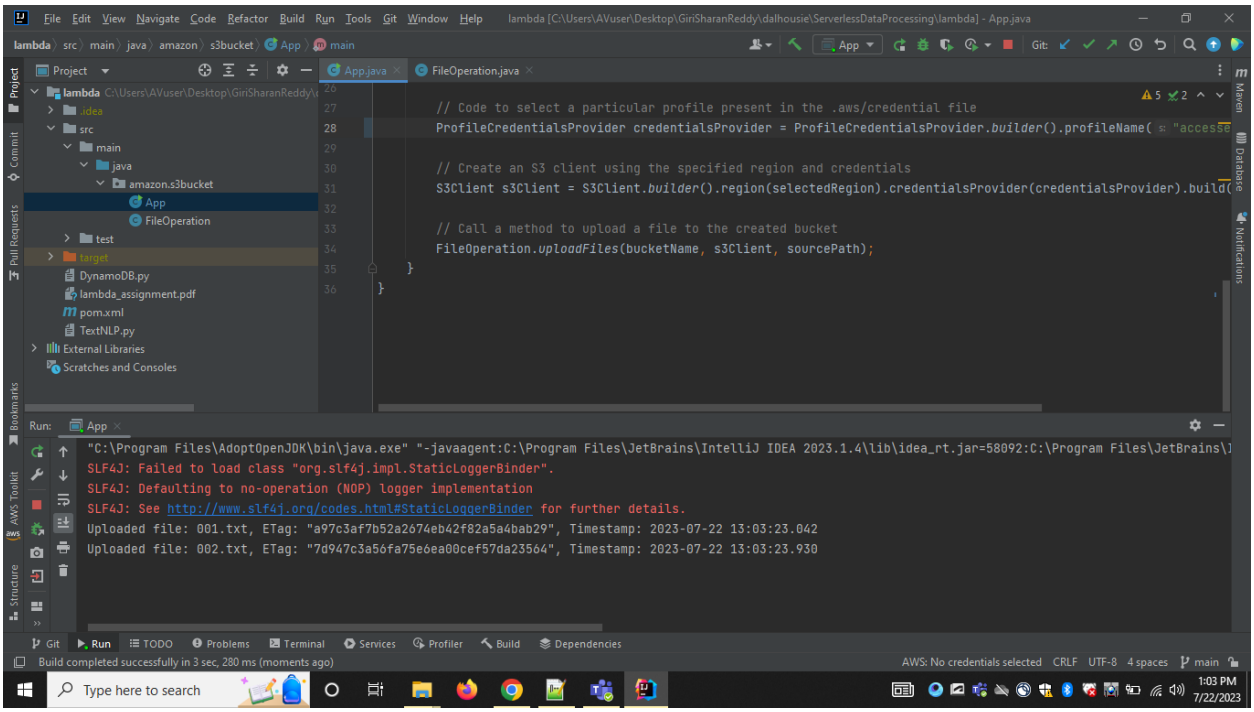

Figure 2:Application uploading files to s3 bucket

2. Initially there were no files in my bucket (sampledata-b00913674)

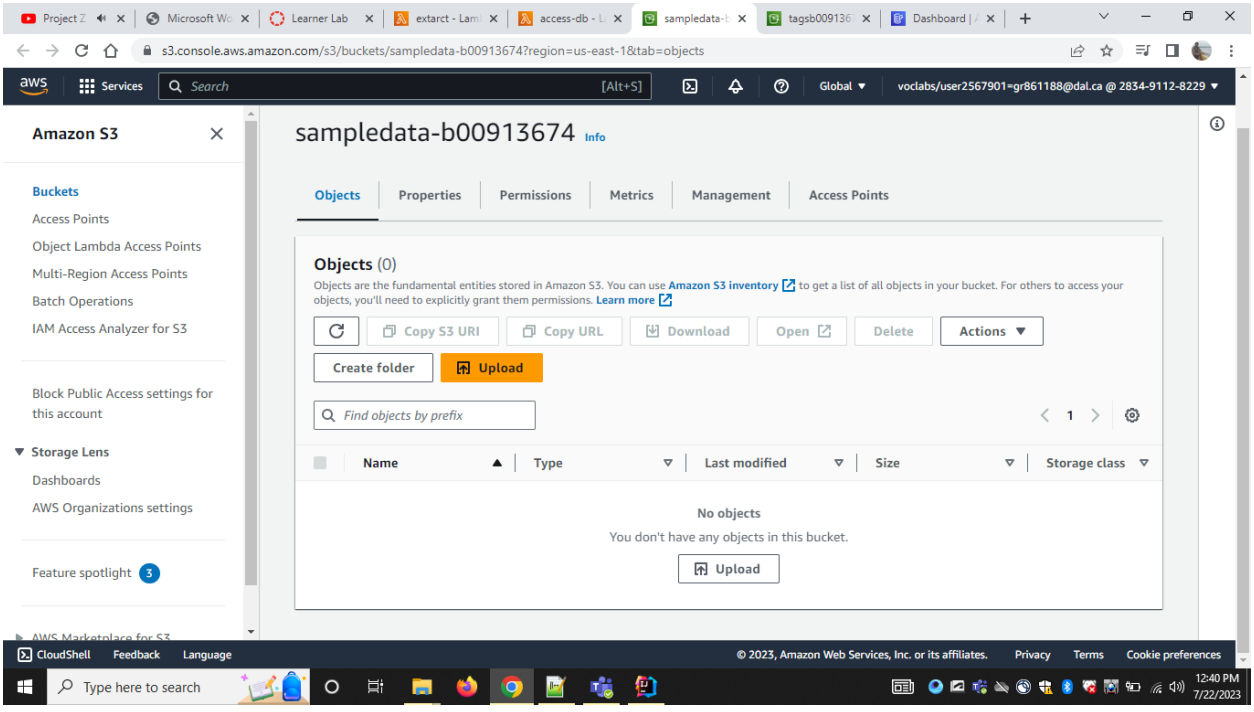

Figure 3:No objects present in s3 bucket initially

3. Initially there were no files in S3 Bucket (tagsb00913674)

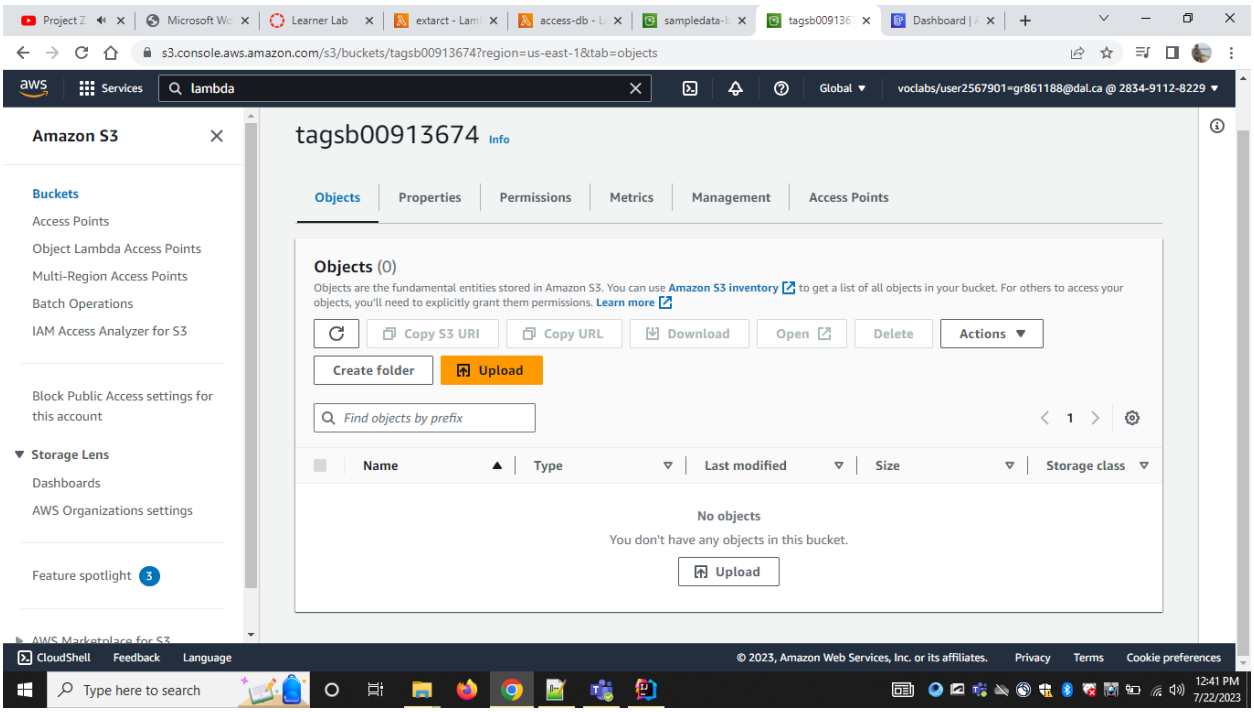

Figure 4:No objects present in s3 bucket initially

4. My application code successfully uploaded 401 files to S3 bucket (sampledata-b00913674)

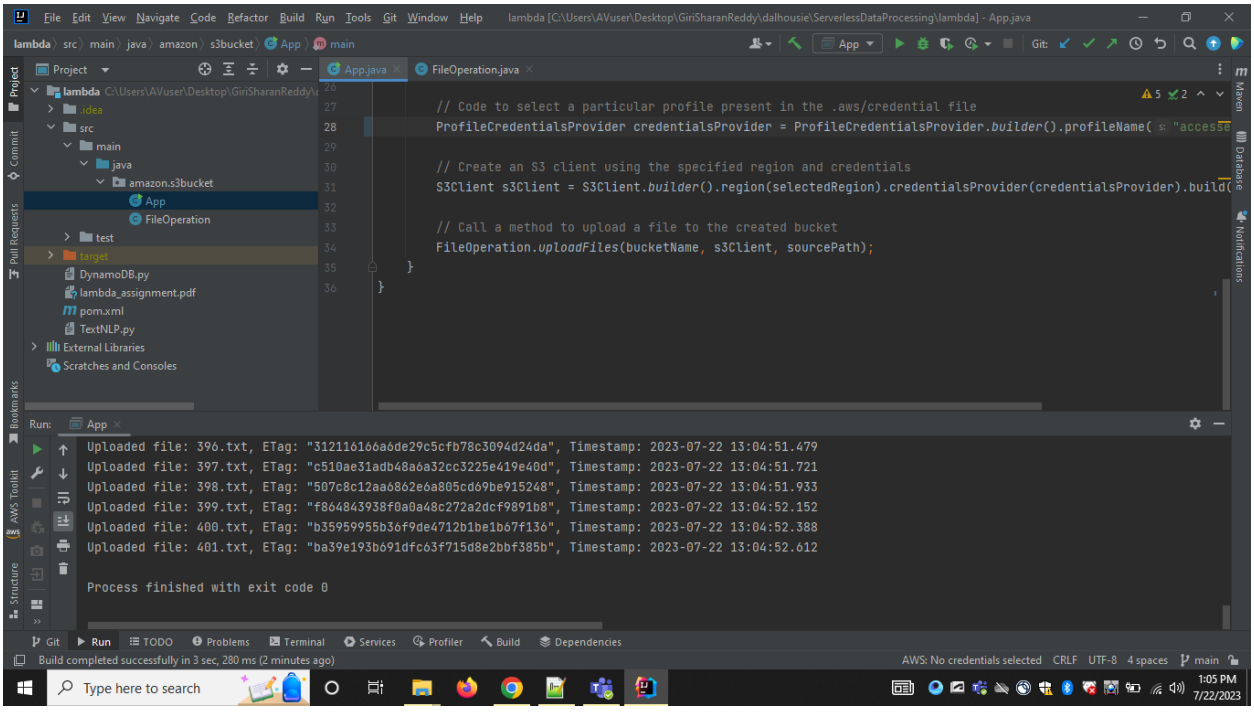

Figure 5:Application successfully uploaded files to s3 bucket

5. Total 401 files uploaded to sampledata-b00913674 S3 bucket

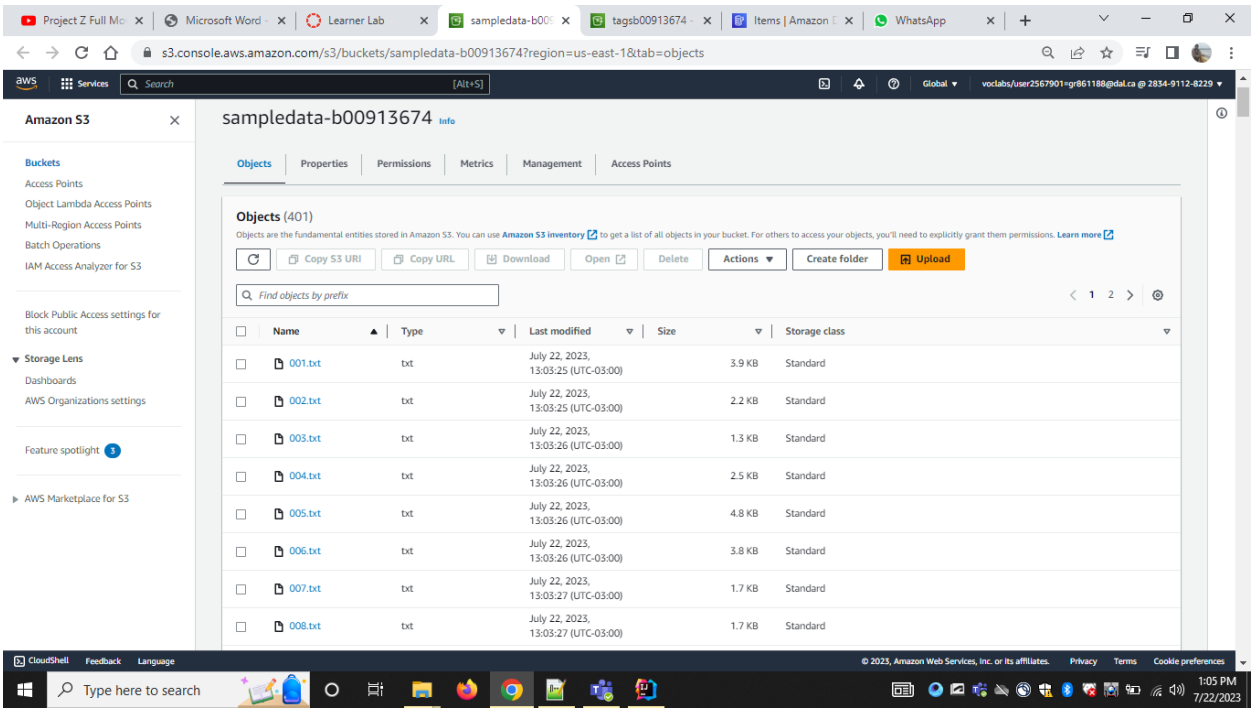

Figure 6:401 objects present in s3 bucket

6. 401 new files created by "extarct" lambda function and uploaded to tagsb00913674 S3 bucket

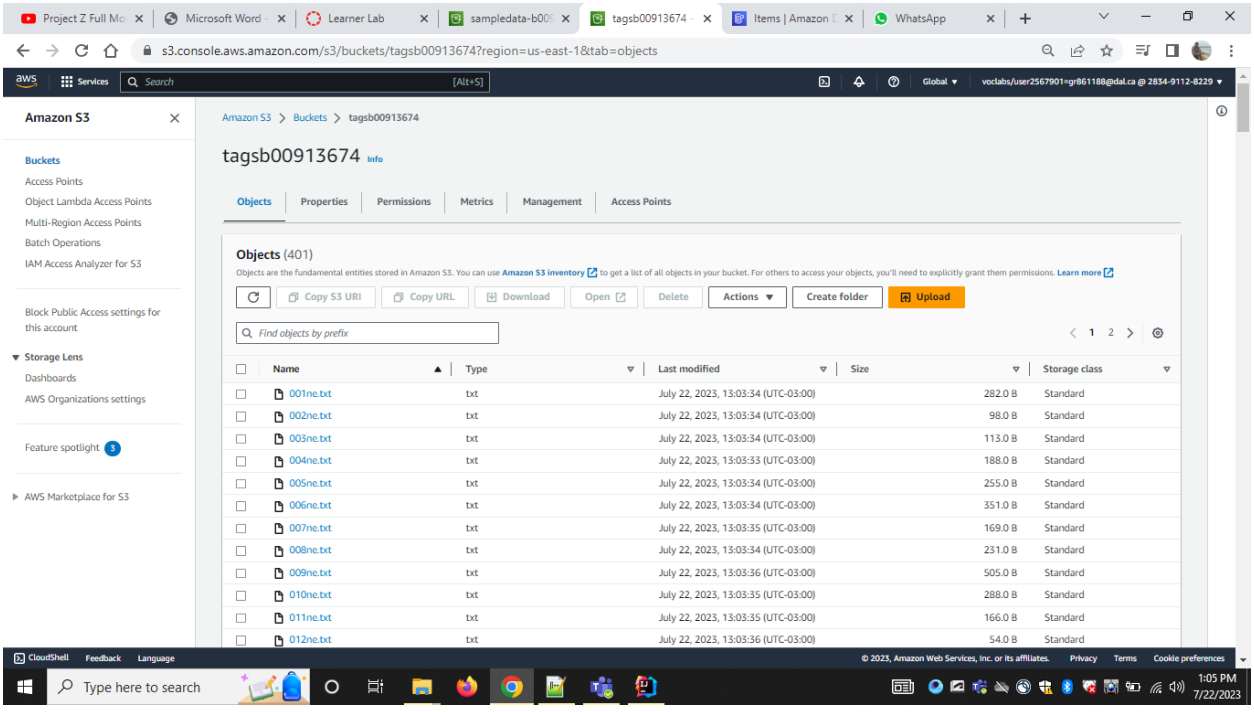

Figure 7:401 new objects present in s3 bucket

7. Data present in newly created files (e.g., data present in 001ne.txt)

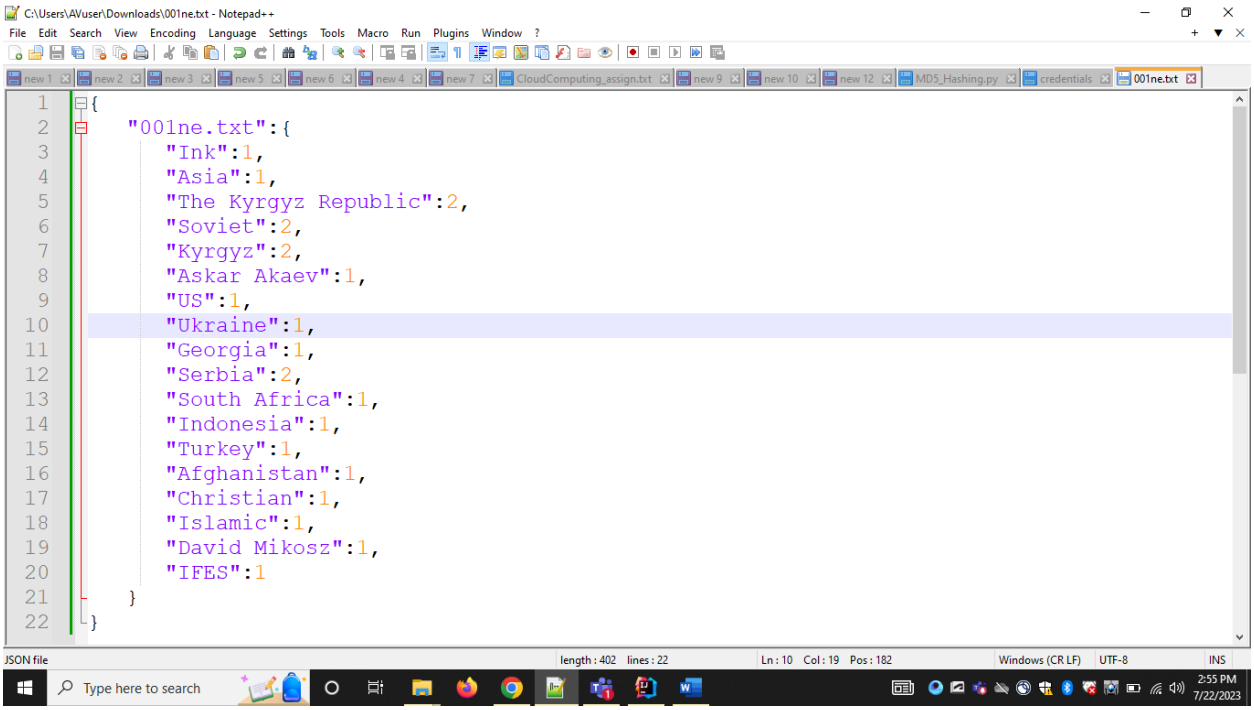

Figure 8:Sample data present in new objects created

8. Initally there are no items in Dynamodb (serverless-assignment)

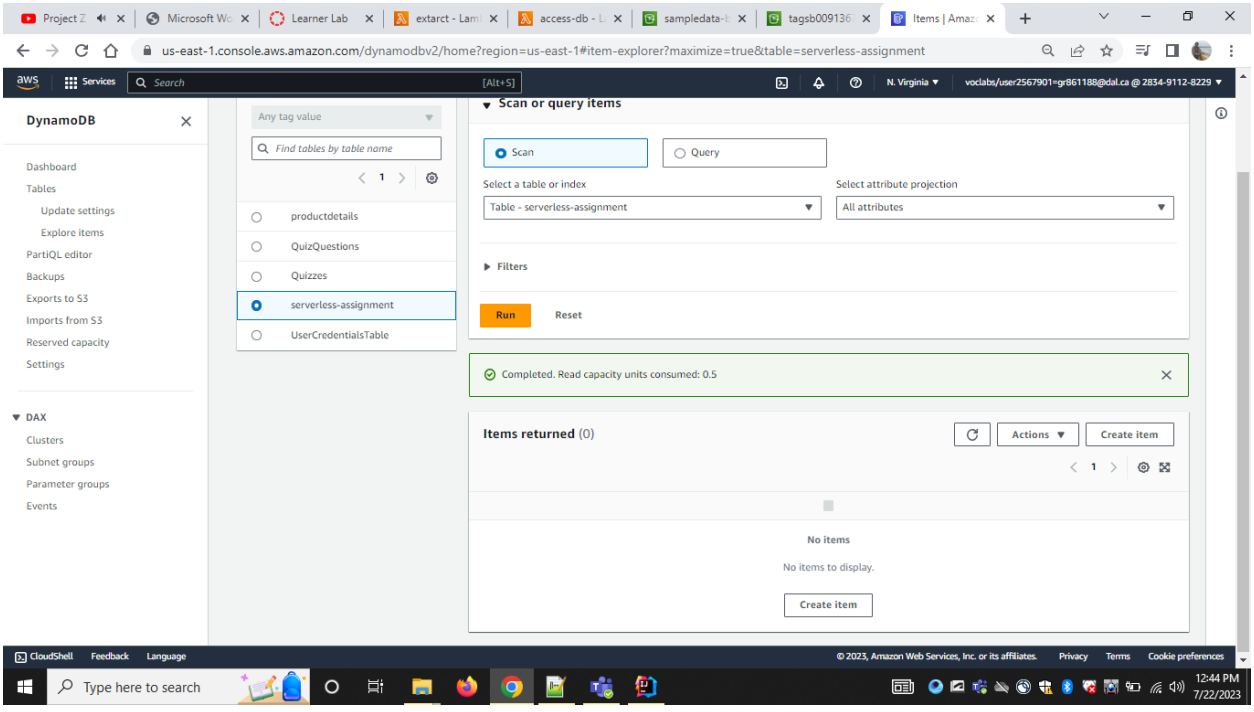

Figure 9:No items present in dynamodb

Extarct Lambda code to extract Important entities like location, Nouns and in this code, I have removed some unwanted entities like percentages, time, date formats and moreover I removed the entities which are not starting with capital letters even though they belong to entities like location, nouns. Once the important entities are fetched from the file, then a new file is created with name like 001ne.txt and kept the entities in that file as json format

```
/*
[1] “SpaCy · industrial-strength Natural Language Processing in Python,” Spacy.io. [Online]. Available: https://github.com/minway/NlpTest. [Accessed: 22-Jul-2023].
[2] NlpTest: Test CoreNLP on AWS Lambda. Available: https://spacy.io/. [Accessed: 22-Jul-2023].
 */

import json
import spacy
import boto3
from collections import defaultdict

def lambda_handler(event, context):

nlp = spacy.load('/opt/en_core_web_sm-2.1.0')

# Get the S3 bucket name and object key from the event
bucket_name = event['Records'][0]['s3']['bucket']['name']
object_key = event['Records'][0]['s3']['object']['key']

print('Bucket Name: {} - Object Name: {}'.format(bucket_name, object_key))

# Create an S3 client
s3_client = boto3.client('s3')

# Read the content of the object
response = s3_client.get_object(Bucket=bucket_name, Key=object_key)
content = response['Body'].read().decode('utf-8')

word_frequencies = {}
my_set = {'apple'}  # Existing set

nlp_test = nlp(content)

doc = nlp(content)

# for entity in doc.ents:
# print(entity.text, entity.label_)


unwanted_entity_types = ["CARDINAL", "PERCENT", "QUANTITY", "TIME", "DATE", "ORDINAL", "MONEY"]
for entity in nlp_test.ents:
if entity.label_ in unwanted_entity_types:
my_set.add(str(entity))
continue
# checking if first letter is capital or not
if entity.text[0].isupper() or entity.text.isupper():
if str(entity) in word_frequencies:
```

```
word_frequencies[str(entity)] += 1
else:
word_frequencies[str(entity)] = 1

print(word_frequencies)
print(my_set)

newFileName = object_key.split(".")[0] + 'ne.txt'


# Prepare data in the required format
data = {newFileName: word_frequencies}


file_path = f"/tmp/{newFileName}"  # Use /tmp directory for writable storage

with open(file_path, "w") as f:
json.dump(data, f)

new_bucket_name = 'tagsb00913674'

try:
response = s3_client.upload_file(file_path, new_bucket_name, newFileName)
return {
'statusCode': 200
}
except ClientError as e:
logging.error(e)
```

Accessdb lambda code, this code will extract the keywords from the files stored in tagsb00913674 bucket and keep in dynamodb along with their repeatable frequency count

```
/*
Amazon.com. [Online]. Available: https://docs.aws.amazon.com/code-
library/latest/ug/python_3_dynamodb_code_examples.html. [Accessed: 22-Jul-
2023].
 */

        import json
        import boto3

        def lambda_handler(event, context):
        # Get the S3 bucket name and object key from the event
        bucket_name = event['Records'][0]['s3']['bucket']['name']
        object_key = event['Records'][0]['s3']['object']['key']

        # Create an S3 client
        s3_client = boto3.client('s3')

        # Read the content of the object
        response = s3_client.get_object(Bucket=bucket_name, Key=object_key)
        content = response['Body'].read().decode('utf-8')
```

```
        # Parse the content and extract the required data
        content_data = json.loads(content)[object_key]

        # Create or get the DynamoDB resource
        dynamodb_resource = boto3.resource('dynamodb')
        table_name = 'serverless-assignment'
        table = dynamodb_resource.Table(table_name)

        # Update the DynamoDB table with the content data
        for key, value in content_data.items():
        response = table.update_item(
        Key={'entity-name': key},
        UpdateExpression='ADD #attrName :value',
        ExpressionAttributeNames={'#attrName': 'value'},
        ExpressionAttributeValues={':value': value},
        ReturnValues='UPDATED_NEW'
        )

        print("Data written to DynamoDB table:", table_name)
```

9. Processed data of files from tagsb00913674 bucket, stored in dynamodb along with their count

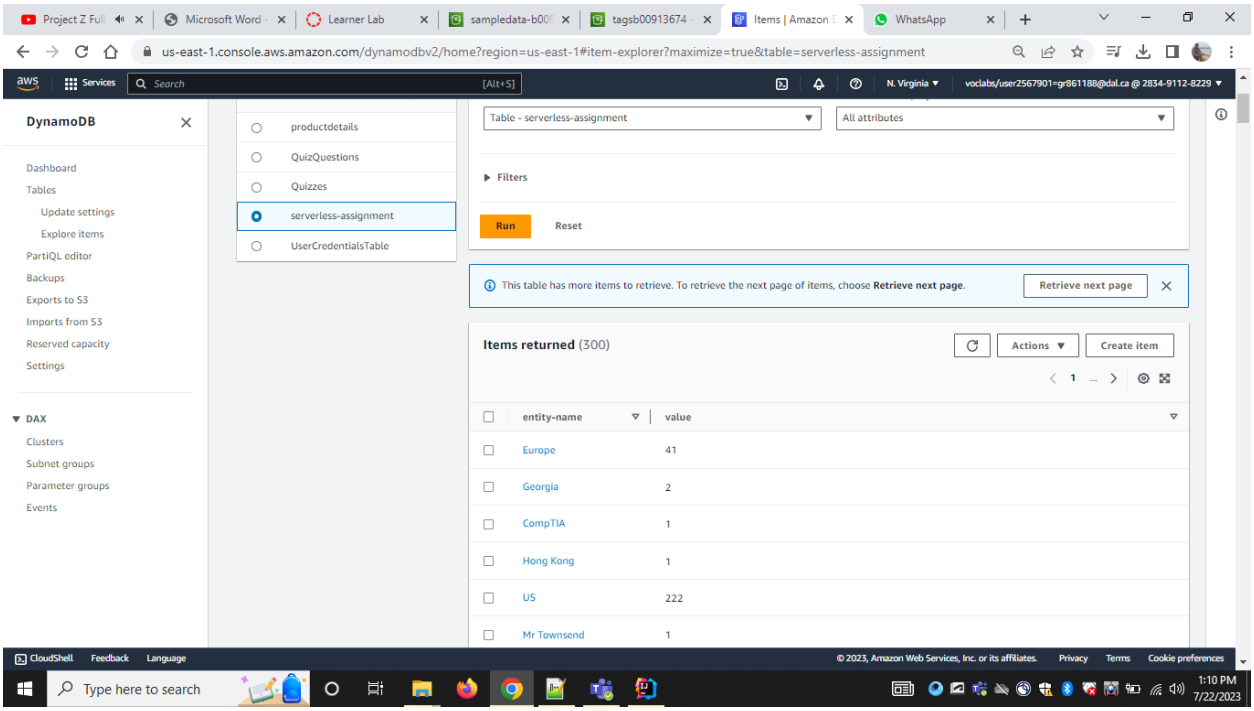

Figure 10:Items and their count in dynamodb

10. Extarct Lambda execution Logs

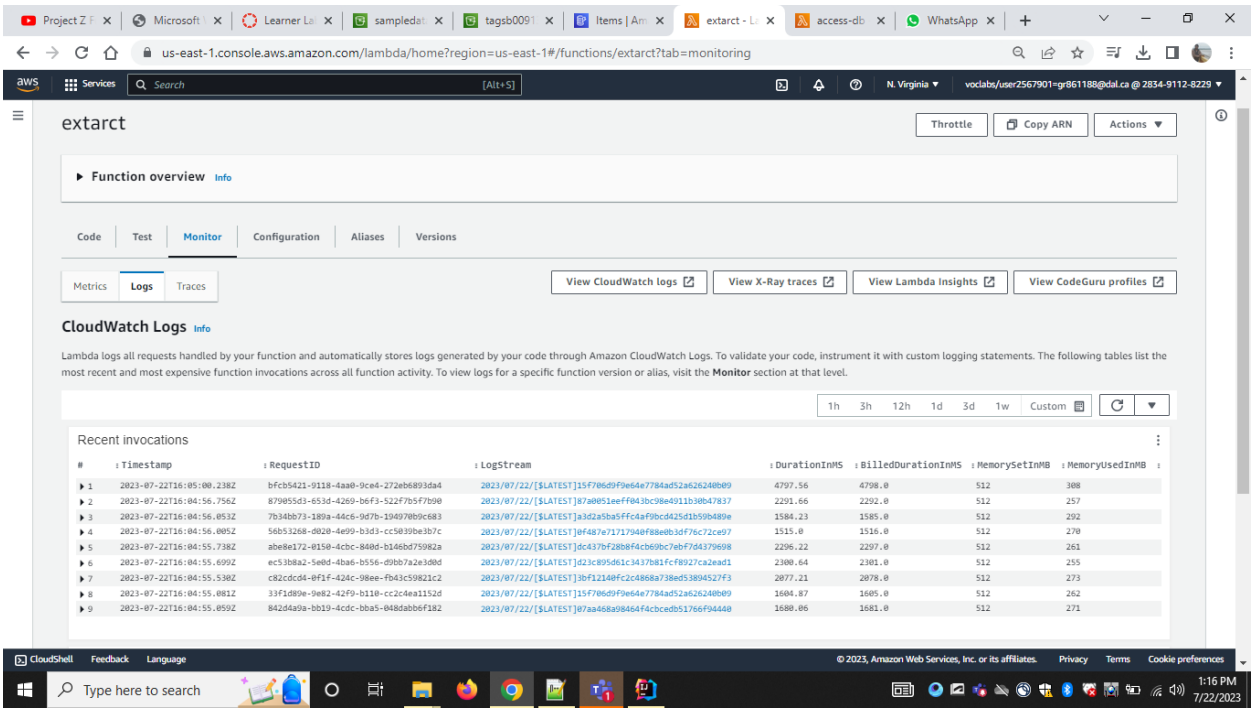

Figure 11:Extarct lambda execution logs

11. Accessdb lambda execution Logs

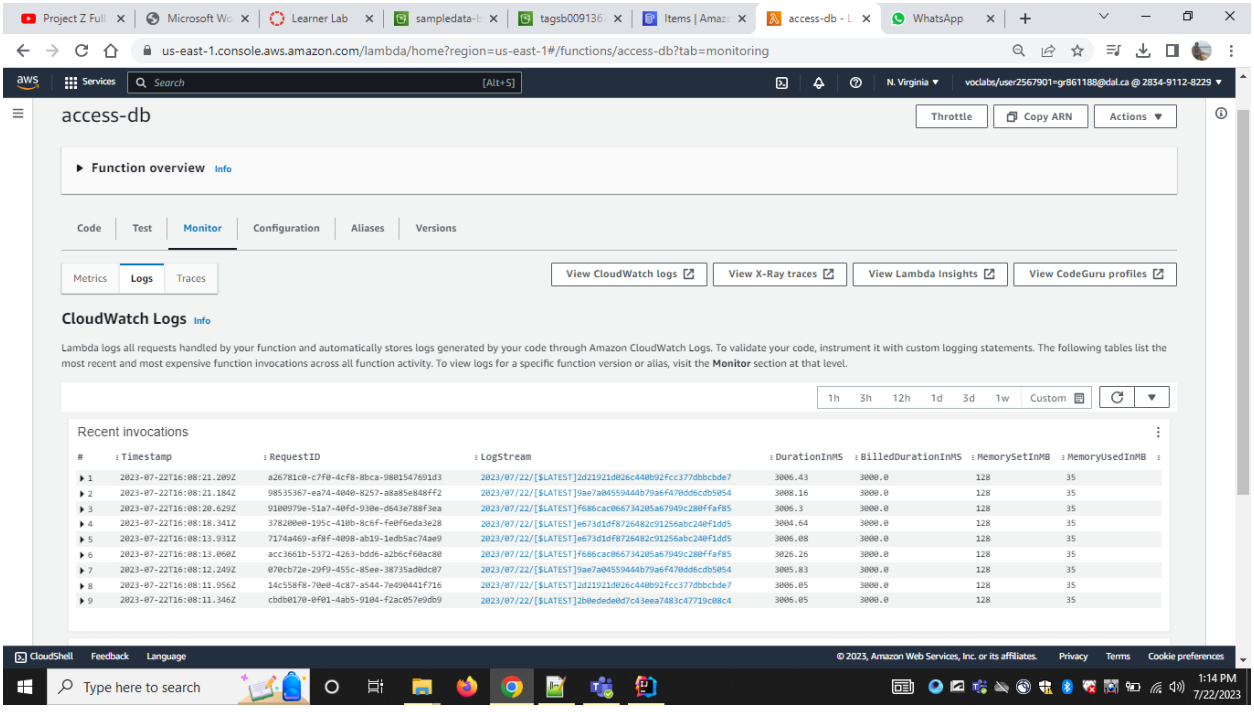

Figure 12:Accessdb lambda execution logs

12. When an object is uploaded to S3 bucket (sampledata-b00913674) lambda function (extarct) will get triggered automatically

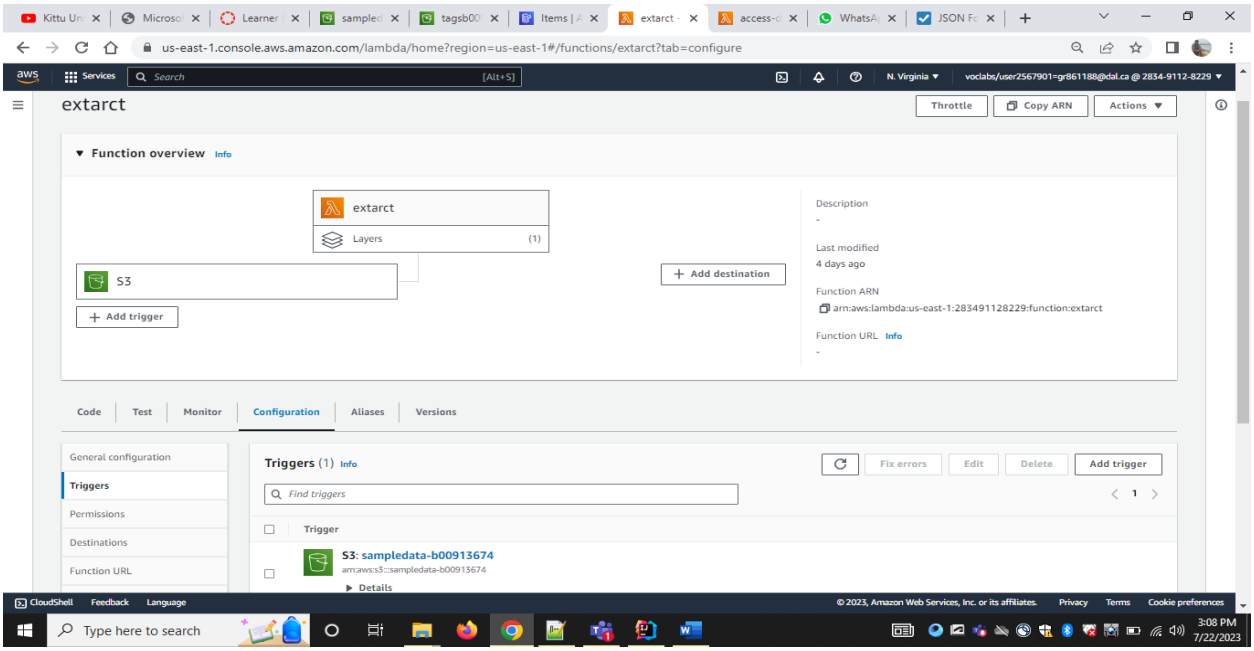

Figure 13:Trigger between s3 bucket and lambda

13. When an object is uploaded to S3 bucket (tagsb00913674) lambda function (accessdb) will get triggered automatically

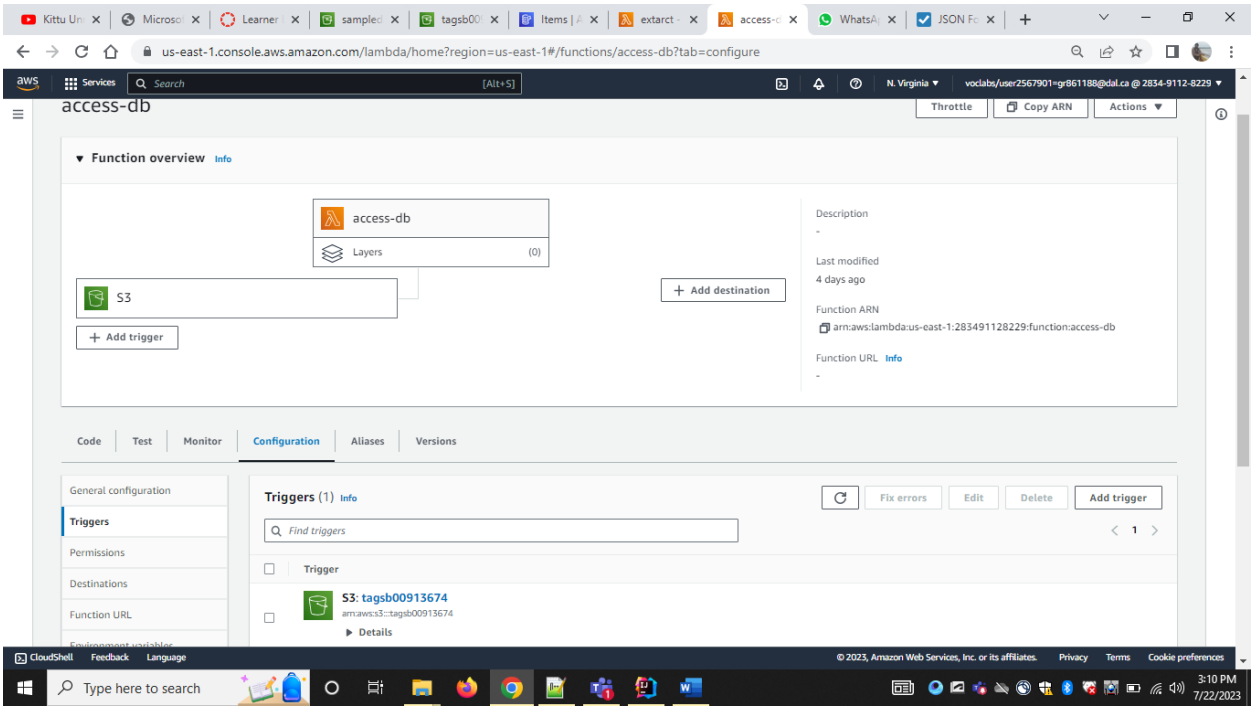

Figure 14:Trigger between s3 bucket and lambda

My application code to upload files to s3 bucket (sampledata-b00913674)

```
package amazon.s3bucket;

import software.amazon.awssdk.auth.credentials.ProfileCredentialsProvider;
import software.amazon.awssdk.regions.Region;
import software.amazon.awssdk.services.s3.S3Client;
import software.amazon.awssdk.services.s3.model.BucketAlreadyExistsException;
import
software.amazon.awssdk.services.s3.model.BucketAlreadyOwnedByYouException;
import software.amazon.awssdk.services.s3.model.CreateBucketRequest;
import software.amazon.awssdk.services.s3.model.CreateBucketResponse;

public class App {

    /*[1]   Amazon.com. [Online].
        Available: https://docs.aws.amazon.com/sdk-for-java/latest/developer-
guide/get-started.html. [Accessed: 22-Jul-2023].*/

    public static void main(String[] args) {

        final String sourcePath = "C:\\Users\\AVuser\\Downloads\\tech\\";

        // Specify the name of the bucket
        final String bucketName = "sampledata-b00913674";

        // Specify the region for the bucket
        Region selectedRegion = Region.US_EAST_1;

        // Code to select a particular profile present in the .aws/credential
file
        ProfileCredentialsProvider credentialsProvider =
ProfileCredentialsProvider.builder().profileName("accessed-June-22-
2023").build();

        // Create an S3 client using the specified region and credentials
        S3Client s3Client =
S3Client.builder().region(selectedRegion).credentialsProvider(credentialsProv
ider).build();

        // Call a method to upload a file to the created bucket
        FileOperation.uploadFiles(bucketName, s3Client, sourcePath);
    }
}
```

```
package amazon.s3bucket;

import java.io.File;
import java.time.Instant;
import java.time.LocalDateTime;
import java.time.ZoneId;
import java.time.format.DateTimeFormatter;
```

```java
import software.amazon.awssdk.core.sync.RequestBody;
import software.amazon.awssdk.services.s3.model.PutObjectResponse;
import software.amazon.awssdk.services.s3.S3Client;
import software.amazon.awssdk.services.s3.model.PutObjectRequest;

public class FileOperation {


    /*[1]   Amazon.com. [Online].
        Available: https://docs.aws.amazon.com/sdk-for-java/latest/developer-
guide/get-started.html. [Accessed: 22-Jul-2023].*/

    /**
     * Uploads multiple files to the specified S3 bucket with a 100ms delay
between uploads.
     * Adds a timestamp in milliseconds when each file is uploaded.
     *
     * @param bucketName the name of the S3 bucket
     * @param s3Client   the S3 client used for the upload
     * @param sourcePath the local directory path containing the files to
upload
     */
    public static void uploadFiles(String bucketName, S3Client s3Client,
String sourcePath) {
        // Get the list of files in the source directory
        File directory = new File(sourcePath);
        File[] files = directory.listFiles();

        if (files == null || files.length == 0) {
            System.out.println("No files found in the specified directory: "
+ sourcePath);
            return;
        }

        // Iterate over the files and upload each file with a 100ms delay
between uploads
        for (File file : files) {
            // Check if the item is a file (not a directory)
            if (file.isFile()) {
                // Specify the name of the file to upload
                String fileName = file.getName();

                // Create a request to upload the file to the specified
bucket with the specified key (file name)
                PutObjectRequest putObjectRequest =
PutObjectRequest.builder()
                        .bucket(bucketName)
                        .key(fileName)
                        .build();

                // Get the current timestamp
                Instant timestamp = Instant.now();

                // Upload the file to the S3 bucket using the S3 client and
request
                PutObjectResponse putObjectResponse =
s3Client.putObject(putObjectRequest, RequestBody.fromFile(file));
```

```
                // Format the timestamp using the desired format
                DateTimeFormatter formatter =
DateTimeFormatter.ofPattern("yyyy-MM-dd HH:mm:ss.SSS");
                String formattedTimestamp =
LocalDateTime.ofInstant(timestamp, ZoneId.systemDefault())
                        .format(formatter);

                // Print the response from the upload operation along with
the formatted timestamp
                System.out.println("Uploaded file: " + fileName + ", ETag: "
+ putObjectResponse.eTag()
                        + ", Timestamp: " + formattedTimestamp);

                try {
                    // Sleep for 100ms before uploading the next file
                    Thread.sleep(100);
                } catch (InterruptedException e) {
                    e.printStackTrace();
                }
            }
        }
    }
}
```

Functional Testing

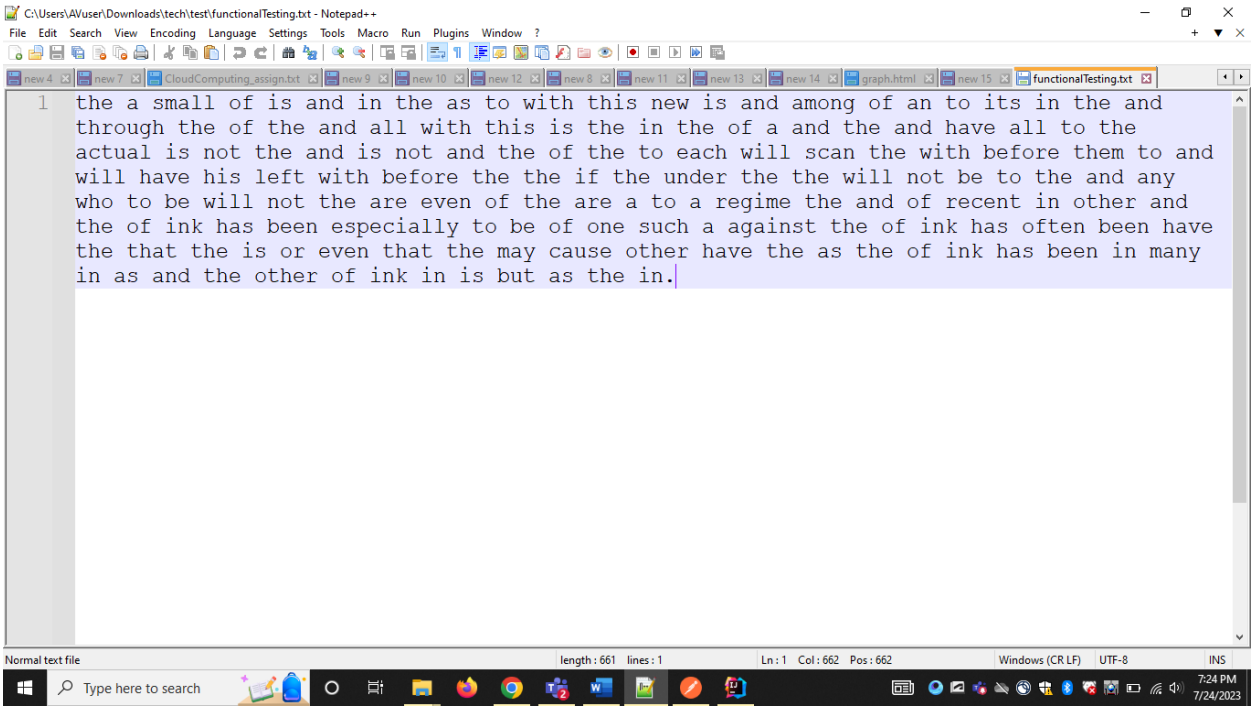

Figure 15: Text file to test functionality

In the above photo functionalTesting.txt file (which I am going to upload to the bucket), there are no words that are starting with capital letters. So, my total process should not keep any of the words in my DynamoDB table. I have deleted all the elements present in the table prior to testing.

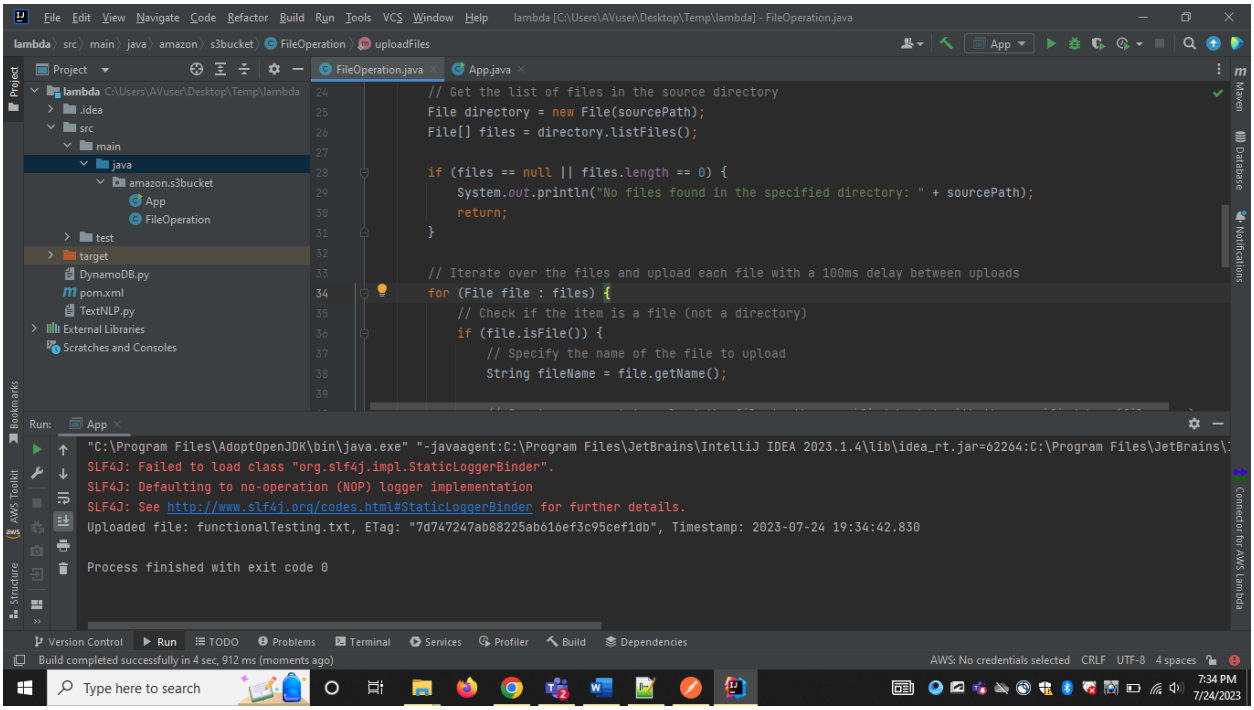

Figure 16:Testing file uploaded to s3 bucket successfully from code

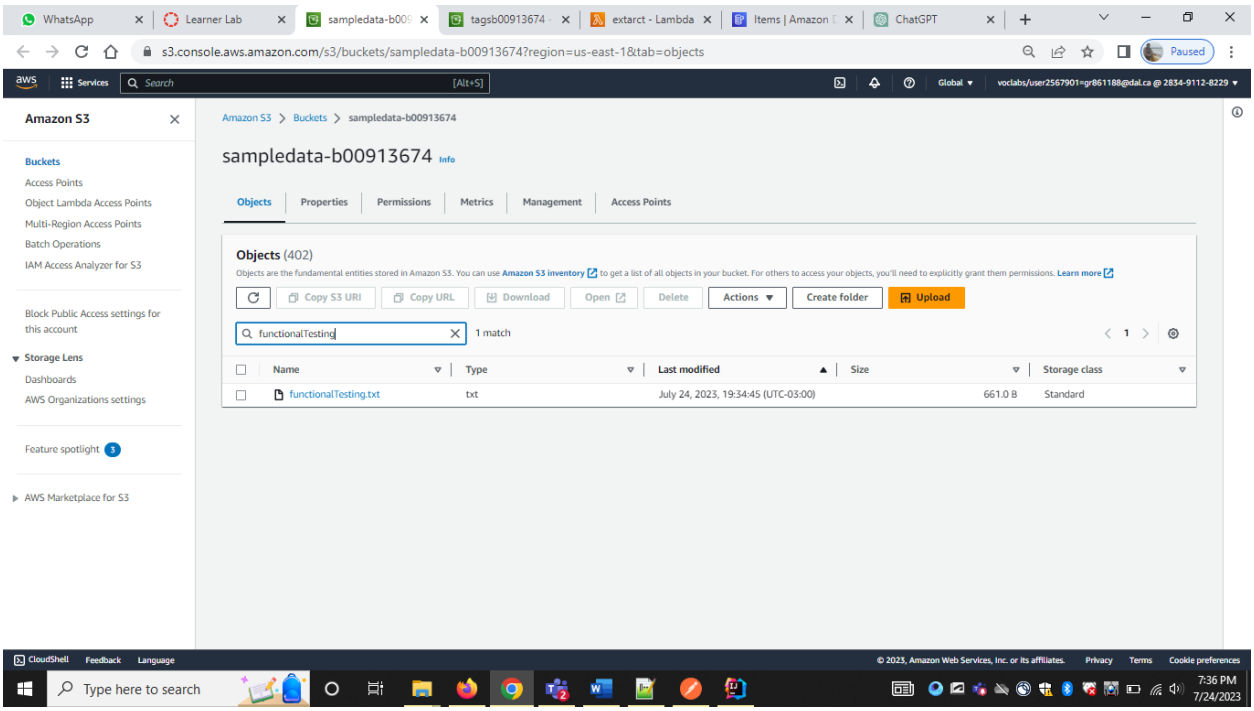

Figure 17: Testing file in s3 bucket

Extract lambda log while testing

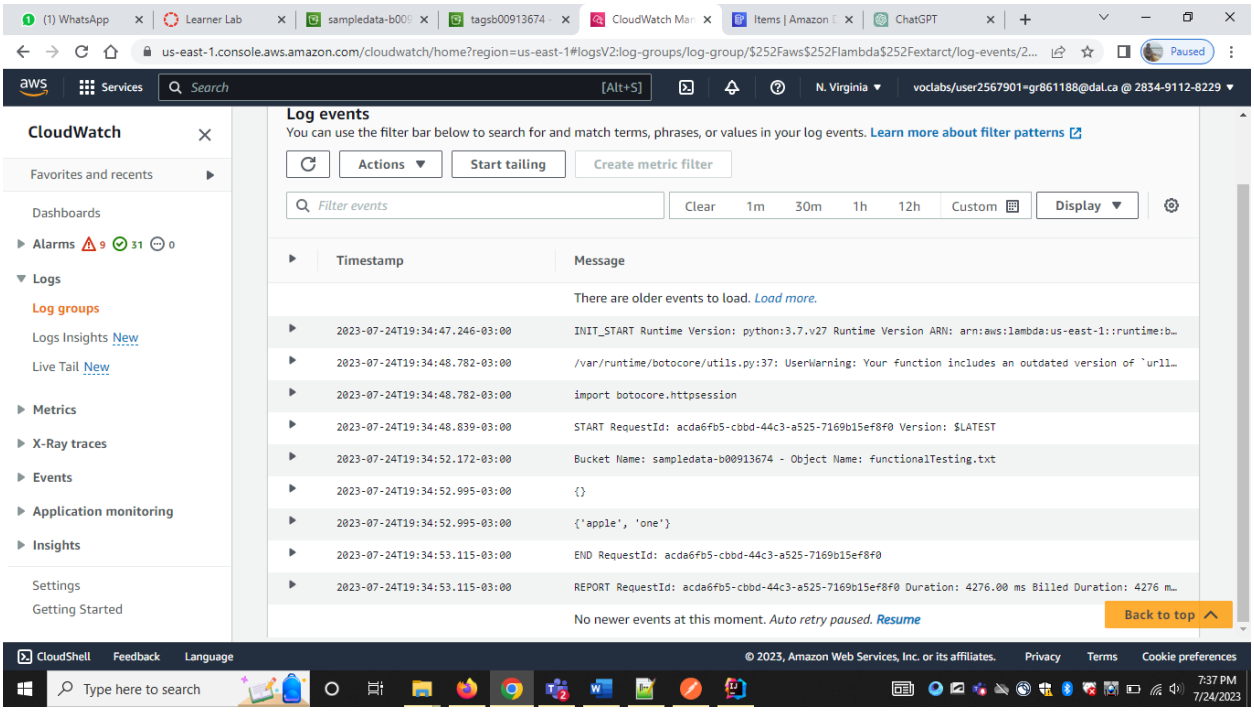

Figure 18:extarct lambda log

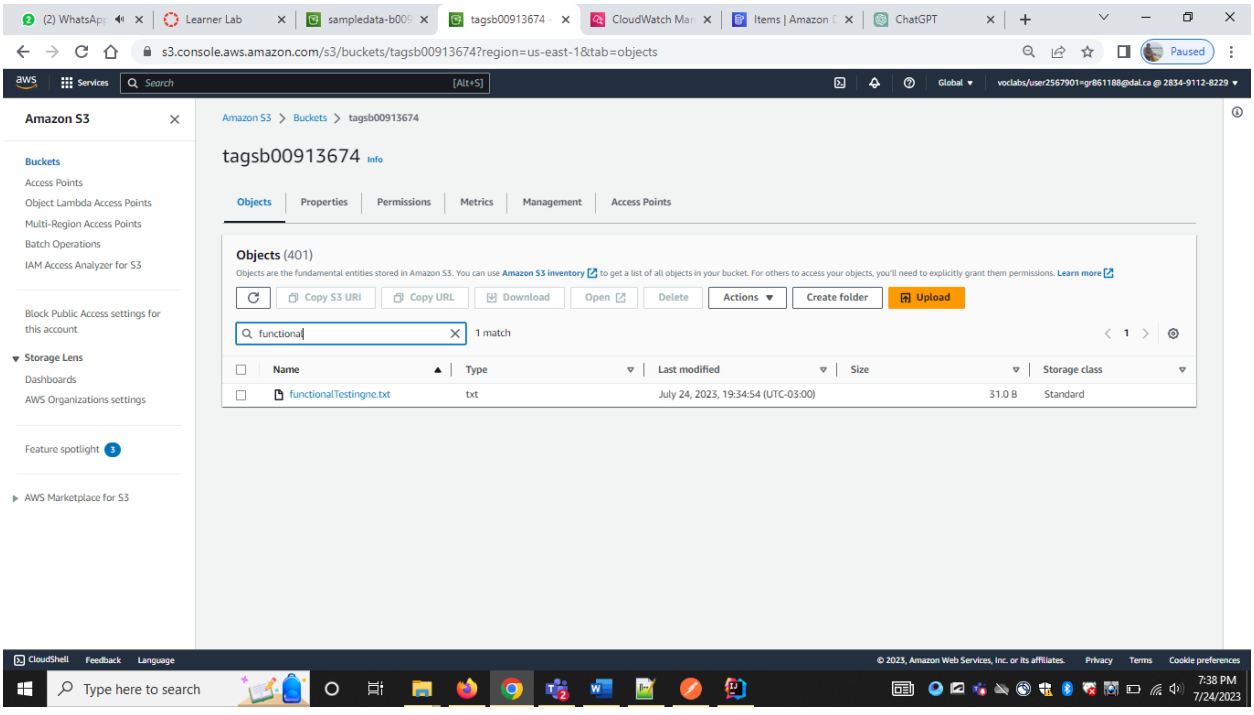

Figure 19: new file created in tags s3 bucket with extension 'ne'

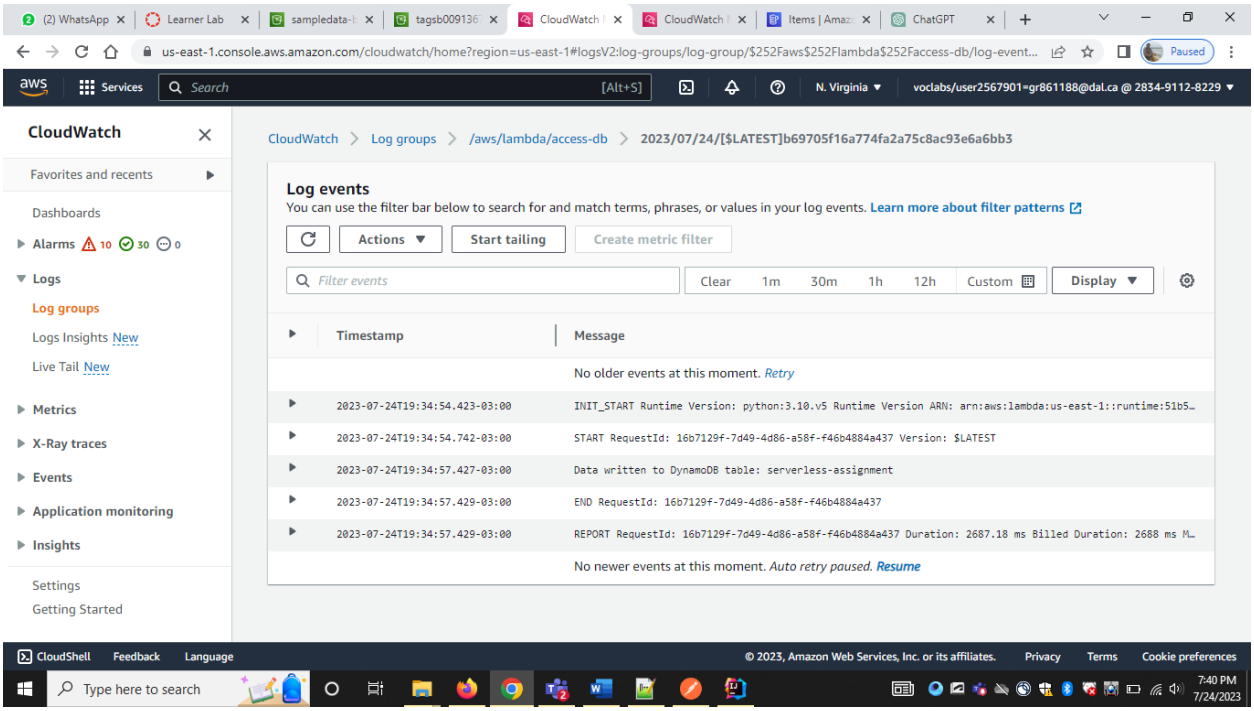

Figure 20:AccessDB lambda function log

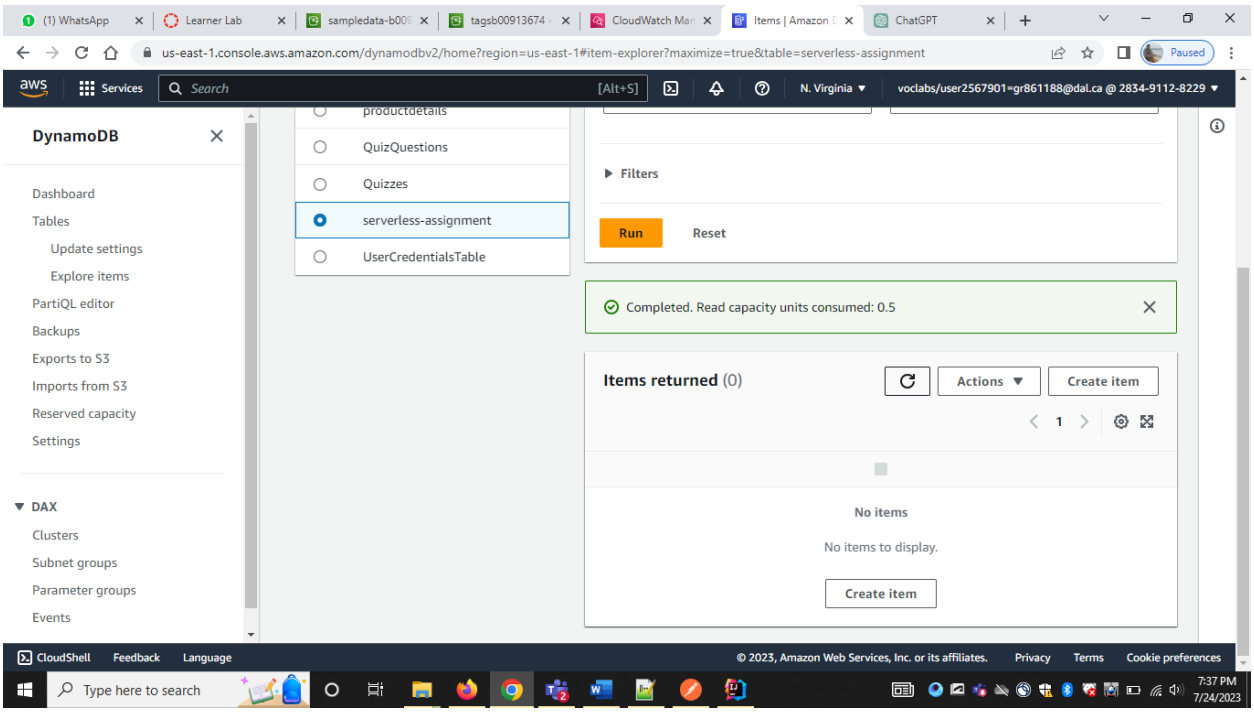

Figure 21: DynamoDB with no items created

In the above testing, we can see that I tried to upload a file with no word starting with capital letter and the extract lambda invocked once the file got uploaded to sampledata-b00913674 bucket. The lambda function processed the file and created a new file with extension ‘ne’ to filename and uploaded the file to tagsb00913674 bucket. Once file is successfully uploaded to tagsb00913674, accessdb lambda function will access that file and get the contents that are in json format and keep the entities in DynamoDB along with their updated count.

PART – 3 REPORT

1. List of SNS Topics available are (messageTopic1, messageTopic2)

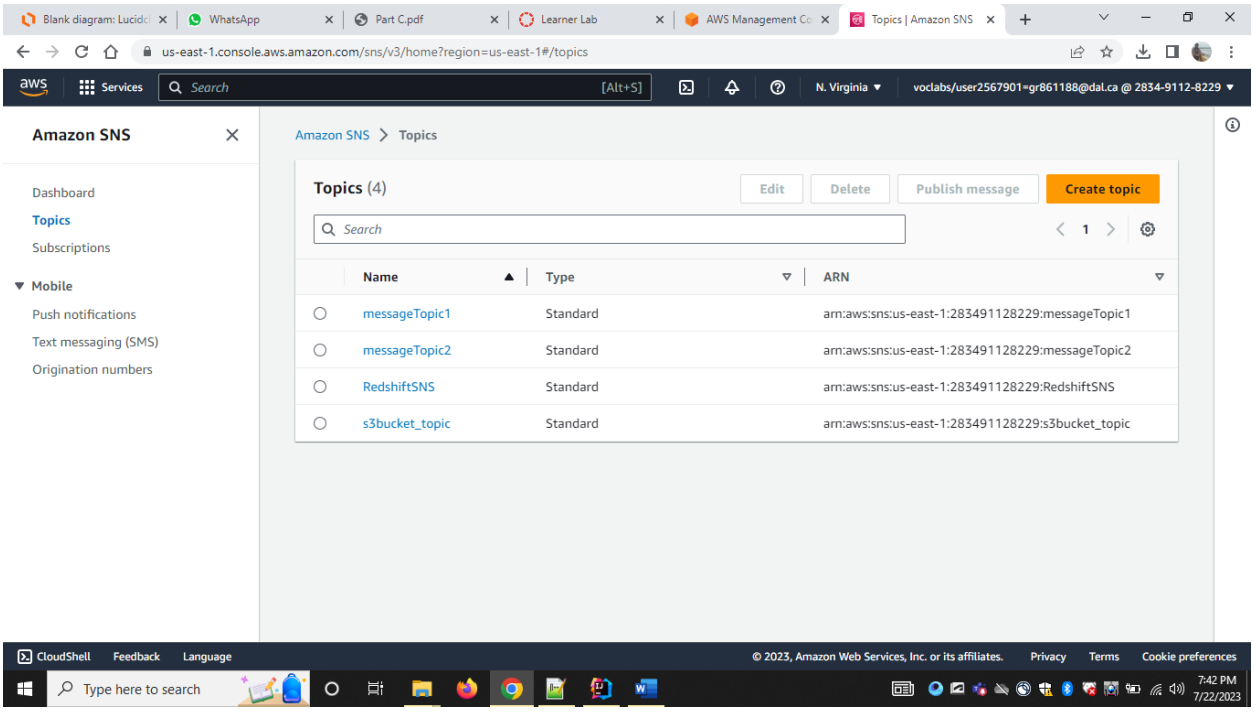

Figure 22:List of SNS Topics available

SNS topic messageTopic1 will be triggered by lambda function “mesageGenerator” where this lambda function will create a random string containing vehicle car type, addon, address. This lambda function will send publish message to SNS topic messageTopic1.

```
""" References:
[1] Amazon.com. [Online]. Available:
    https://docs.aws.amazon.com/code-
library/latest/ug/python_3_sns_code_examples.html.
    [Accessed: 22-Jul-2023]. """


import json
import random
import boto3

def lambda_handler(event, context):
    # TODO implement
    CAR = ["Compact", "mid-size Sedan", "SUV", "Luxury"]
    ADDON = ["GPS", "Camera"]
    ADDRESSES = ["6050 University Avenue", "1333 South Park St.", "5543 Clyde
St."]

    selected_car = random.choice(CAR)
    selected_addon = random.choice(ADDON)
    selected_address = random.choice(ADDRESSES)
```

```
    response_message = f"You selected {selected_car} with {selected_addon}
and location {selected_address}"

    # Replace 'YOUR_SNS_TOPIC_ARN' with the actual ARN of your SNS topic
    sns_topic_arn = 'arn:aws:sns:us-east-1:283491128229:messageTopic1'

    # Create an SNS client
    sns = boto3.client('sns')

    # Publish the message to the SNS topic
    snsResponse = sns.publish(
        TopicArn=sns_topic_arn,
        Message=response_message
    )

    print(snsResponse)

    return {
        'statusCode': 200,
        'body': response_message
    }
```

SNS topic “messageTopic1” is created and SQS “firstQueue” is subscribed to it

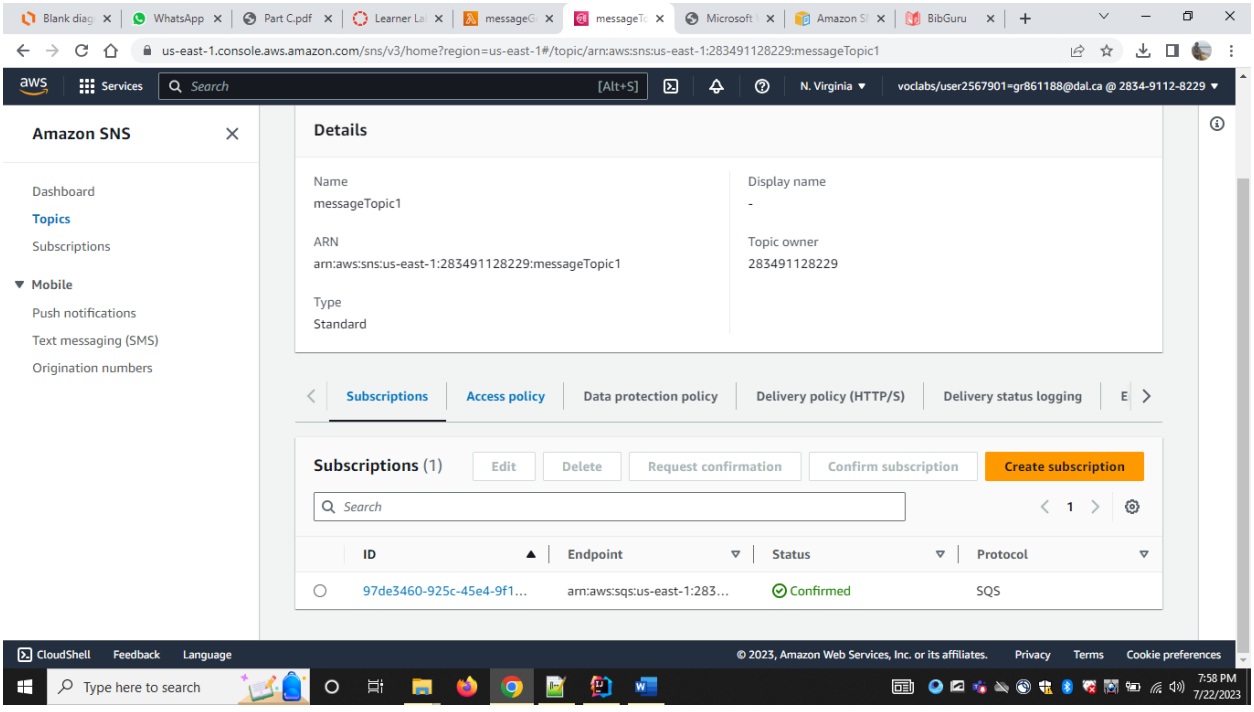

Figure 23:SNS Topic messageTopic1

SQS “firstQueue” is created and we can see that it is subscribed to SNS Topic “messageTopic1”

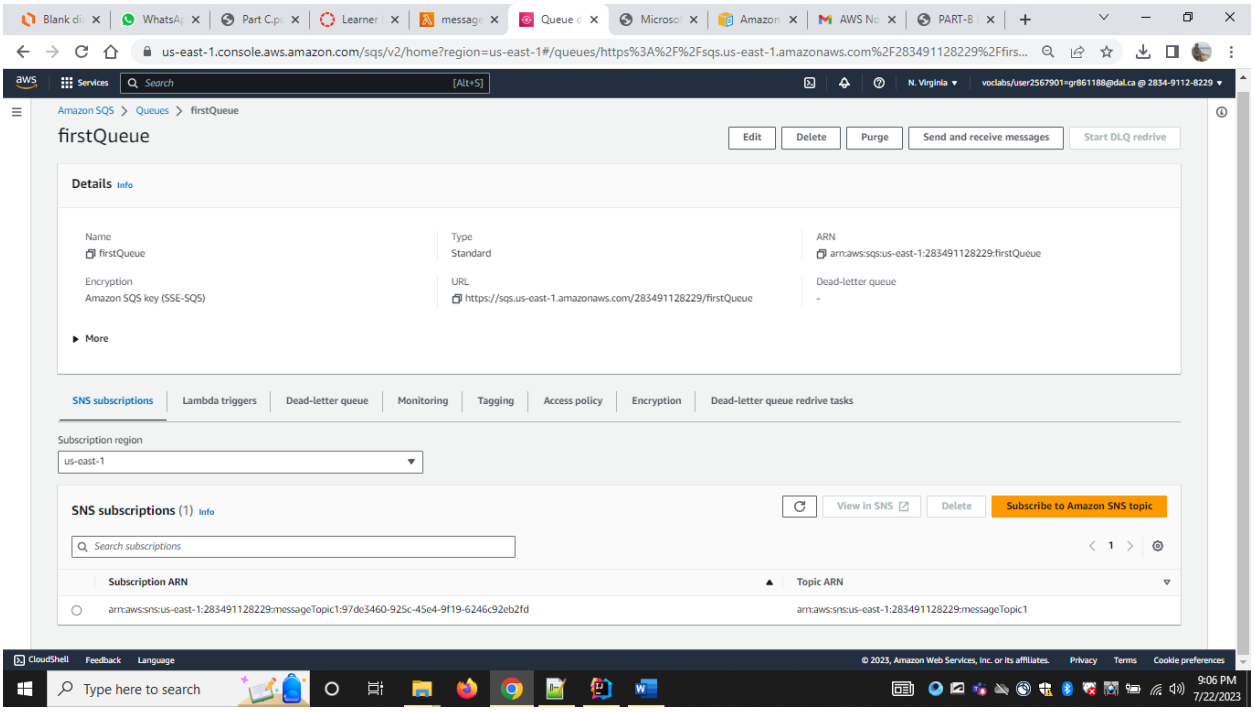

Figure 24:SQS firstQueue and its subscription

Creating a one more SNS Topic “messageTopic2” and subscribing it to my email id to receive the emails

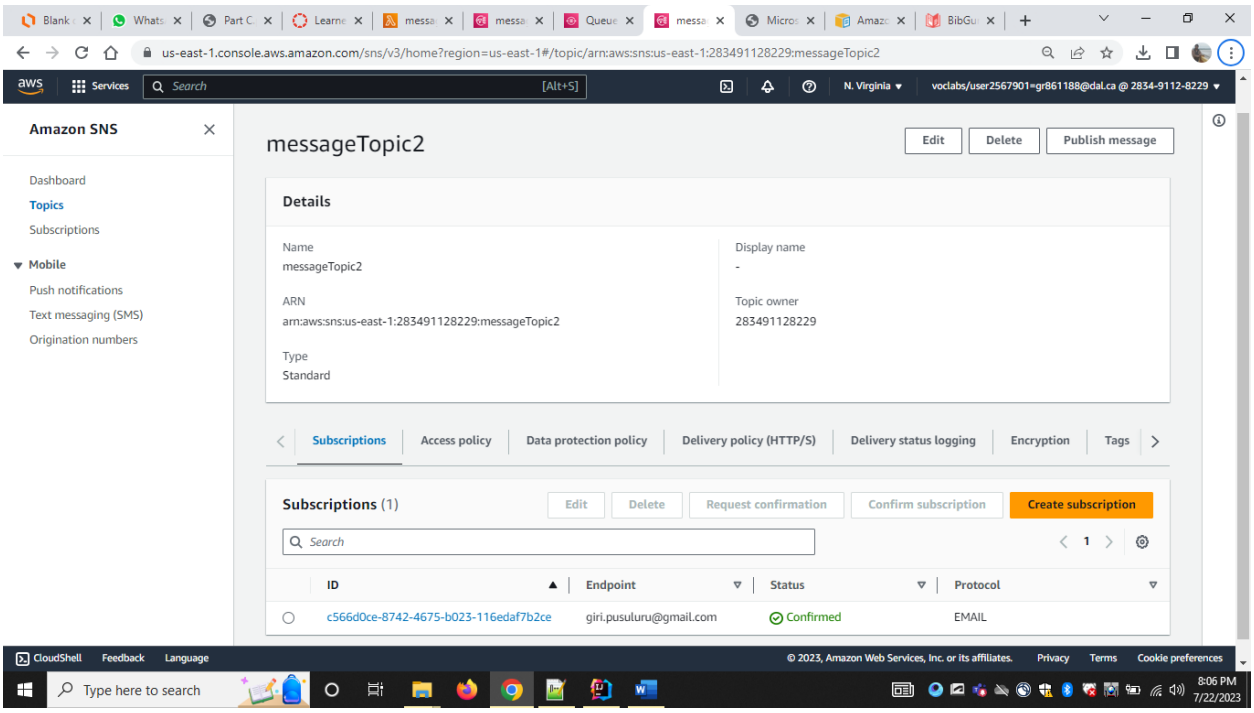

Figure 25:SNS Topic messageTopic2 and its subscription

Email I got to subscribe for SNS Topic messageTopic2 and then confirming by clicking the link provided

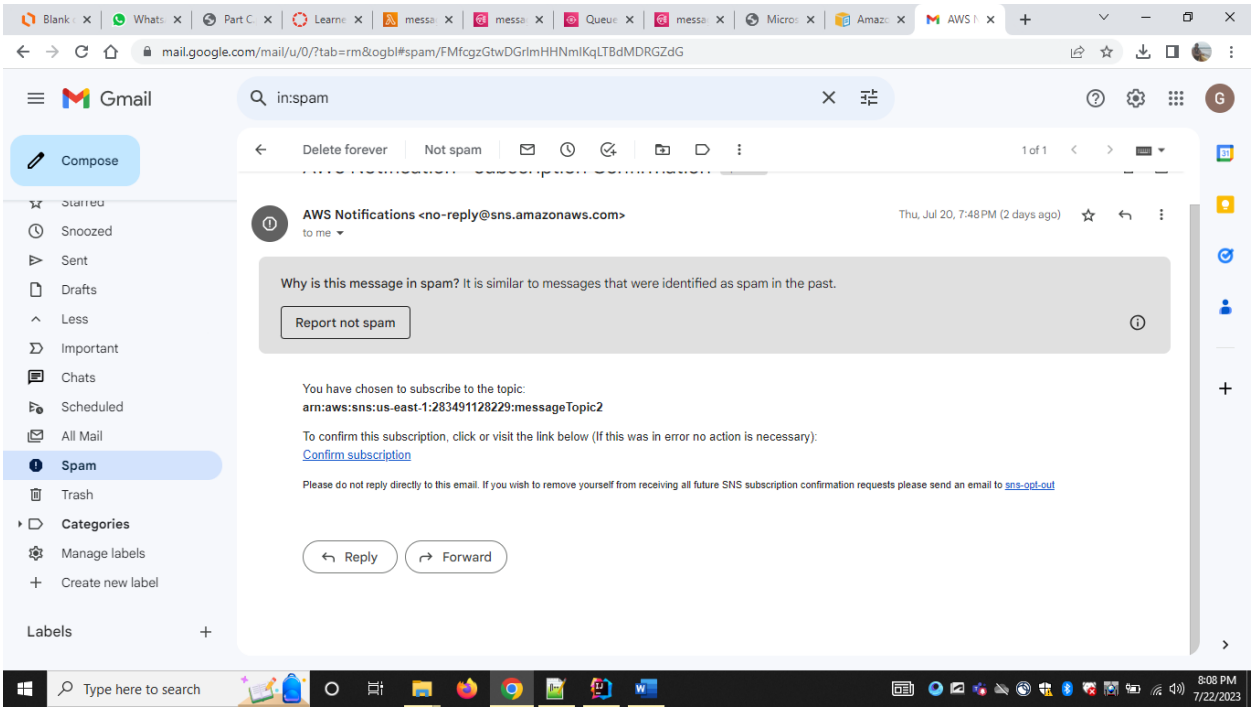

Figure 26:Email notification to subscribe

Lambda “messageGetter” code that will poll on SQS “firstQueue” to get the message from the queue based on queue url and then publish it to SNS Topic “messageTopic2” and after that deleting the message from the queue.

```
""" References:
[1] Amazon.com. [Online]. Available:
    https://docs.aws.amazon.com/code-
library/latest/ug/python_3_sns_code_examples.html.
    [Accessed: 22-Jul-2023].
[2] “Sending and receiving messages in Amazon SQS - Boto3 1.28.9
documentation,” Amazonaws.com. [Online].
    Available:
https://boto3.amazonaws.com/v1/documentation/api/latest/guide/sqs-example-
sending-receiving-msgs.html.
    [Accessed: 22-Jul-2023].
"""

import json
import boto3

def lambda_handler(event, context):

    sns = boto3.client('sns')

    sns_topic_arn = 'arn:aws:sns:us-east-1:283491128229:messageTopic2'
```

```
    # Create an SQS client
    sqs = boto3.client('sqs')

    # Receive messages from the queue
    response = sqs.receive_message(
        QueueUrl=
'https://sqs.us-east-1.amazonaws.com/283491128229/firstQueue',
        MaxNumberOfMessages=1,  # Number of messages to retrieve at once
        WaitTimeSeconds=10  # Wait time for receiving messages
    )

    print(response)

    # Check if there are messages in the response
    if 'Messages' in response:
        message = response['Messages'][0]  # Get the first message

        # Extract the body of the message
        body = json.loads(message['Body'])

        # Getting the message sent by messageGenerator Code
        messageReceived = body['Message']

        print(messageReceived)

        snsResponse = sns.publish(
            TopicArn='arn:aws:sns:us-east-1:283491128229:messageTopic2',
            Message=messageReceived
            )

        print(snsResponse)

        # Delete the message from the queue after processing (optional,
depending on your use case)
        receipt_handle = message['ReceiptHandle']
        sqs.delete_message(
            QueueUrl="https://sqs.us-east-
1.amazonaws.com/283491128229/firstQueue",
            ReceiptHandle=receipt_handle
            )
        print('Message Deleted Successfully')

    return {
        'statusCode': 200,
        'body': json.dumps('SQS message processed successfully!')
    }
```

Amazon Event Bridge rule “eventBridge” is created where it will invoke the lambda function “messageGetter” for every 1 minutes

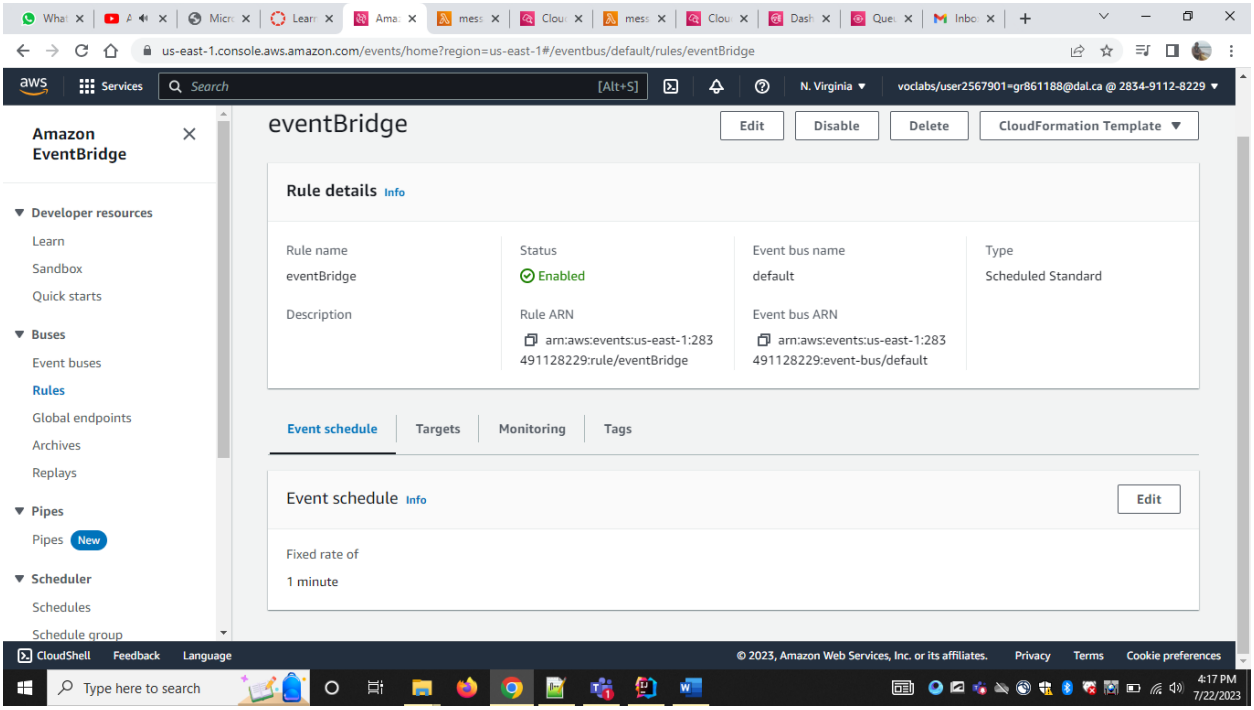

Figure 27:EventBridge to schedule

Procedure I am following to do this task

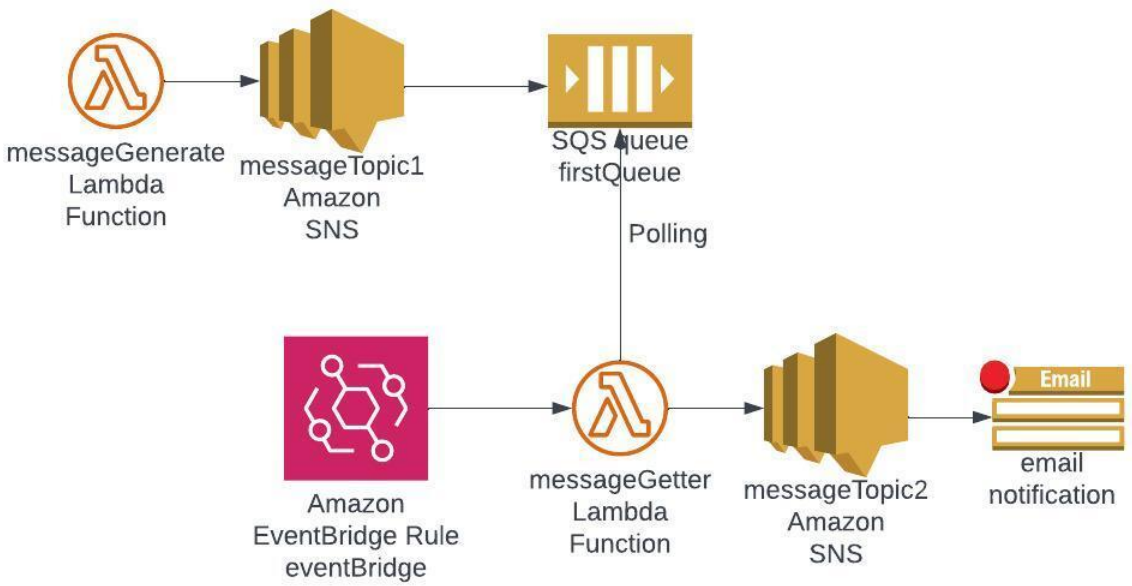

Figure 28: Architecture followed to send an email [12]

I have invoked messageGenerator lambda function

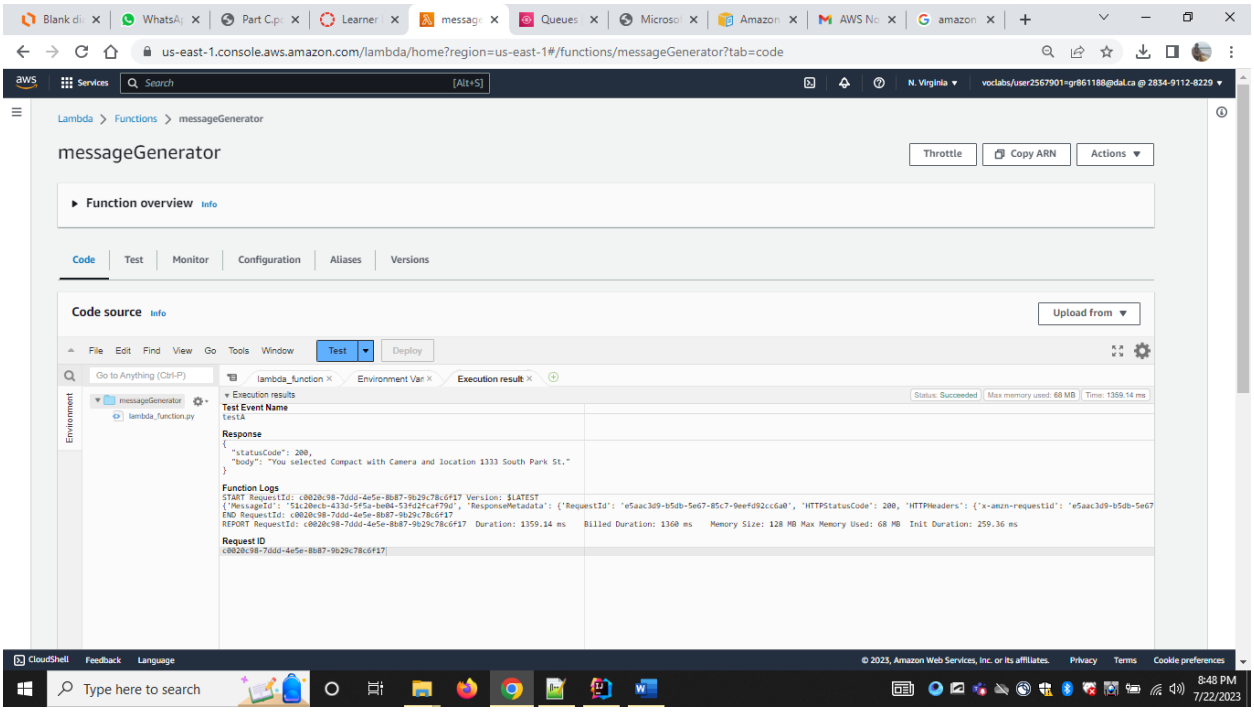

Figure 29: Invoked lambda to generate a random message

The message is in queue now

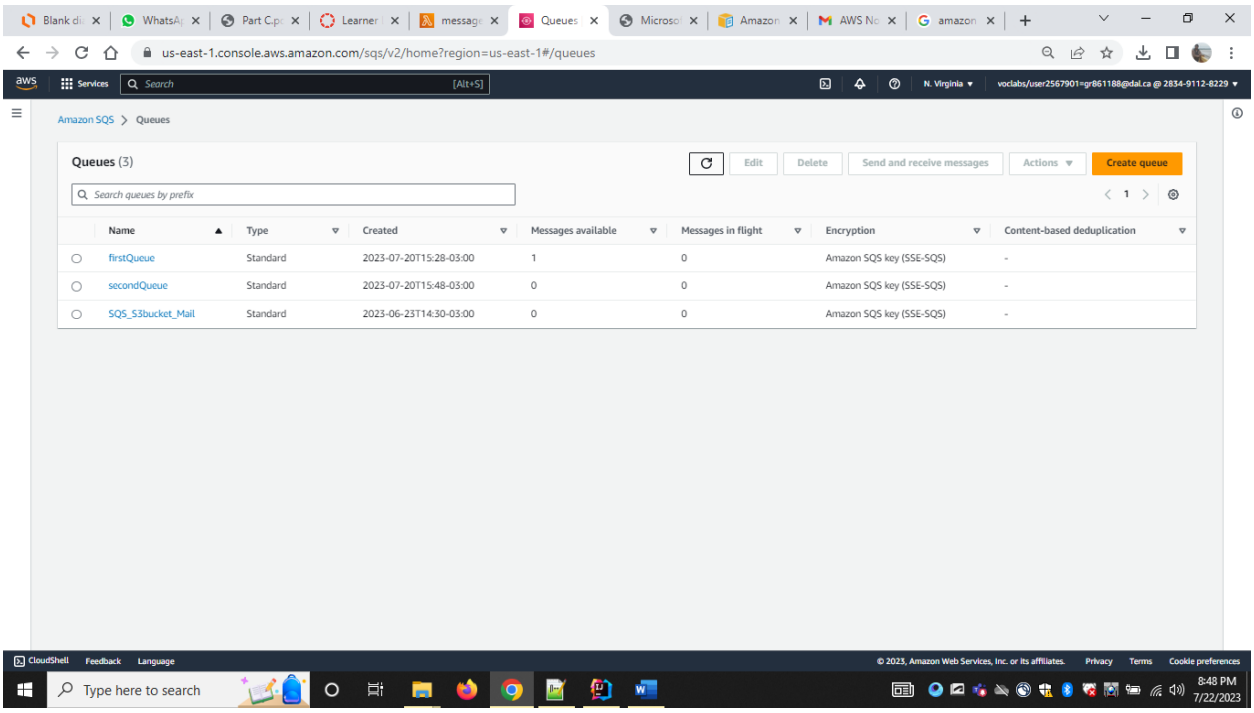

Figure 30: Message arrived into queue

Email Notification received

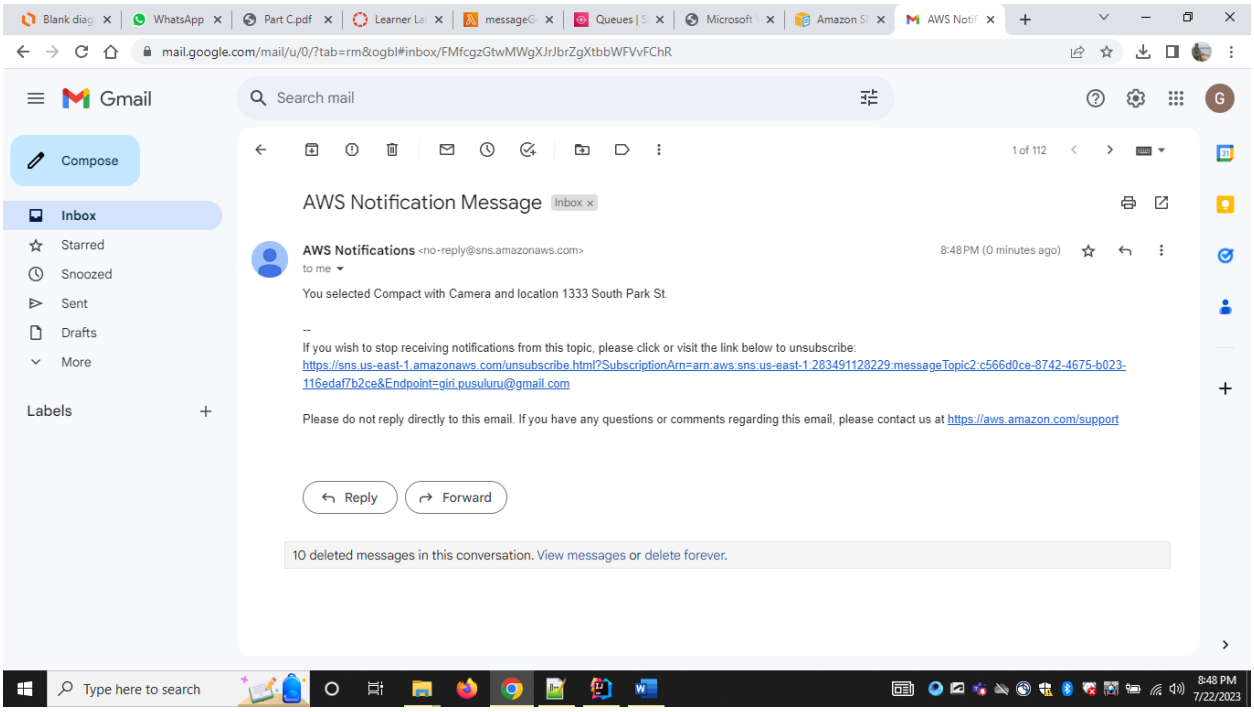

Figure 31: Email notification received

messageGenerator Lambda function Log

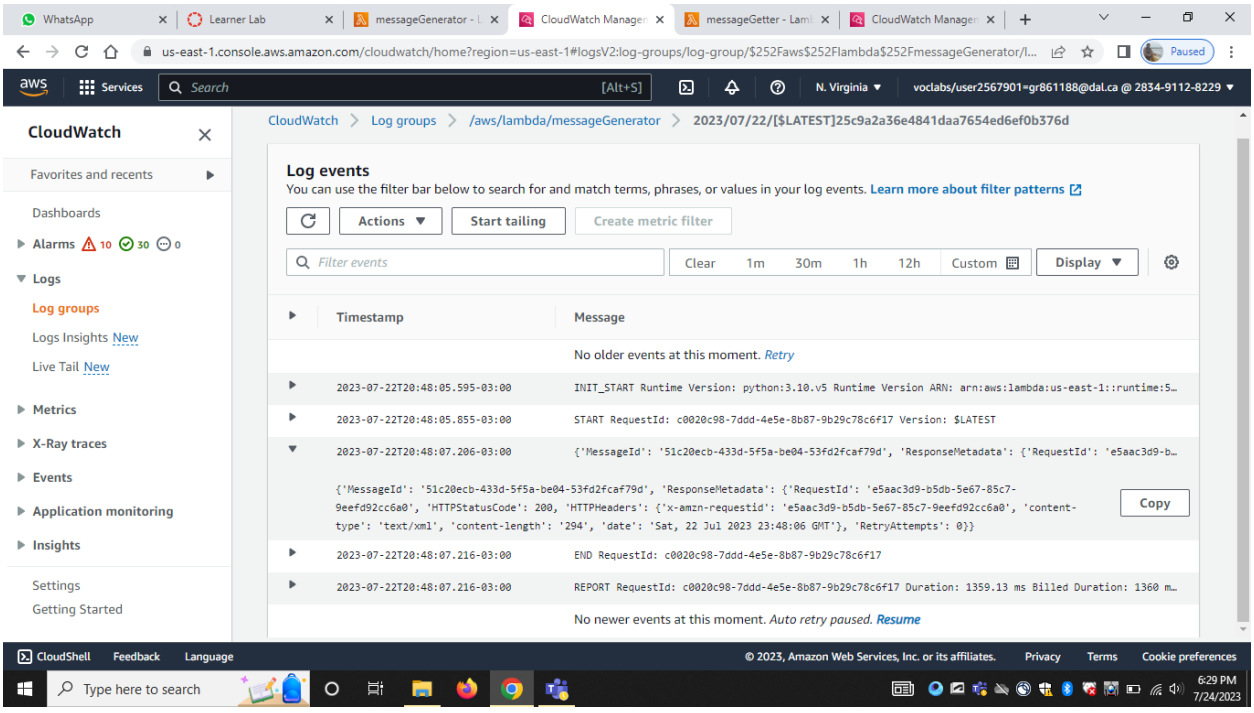

Figure 32:Log of messageGenerator lambda

messageGetter Lambda function Log

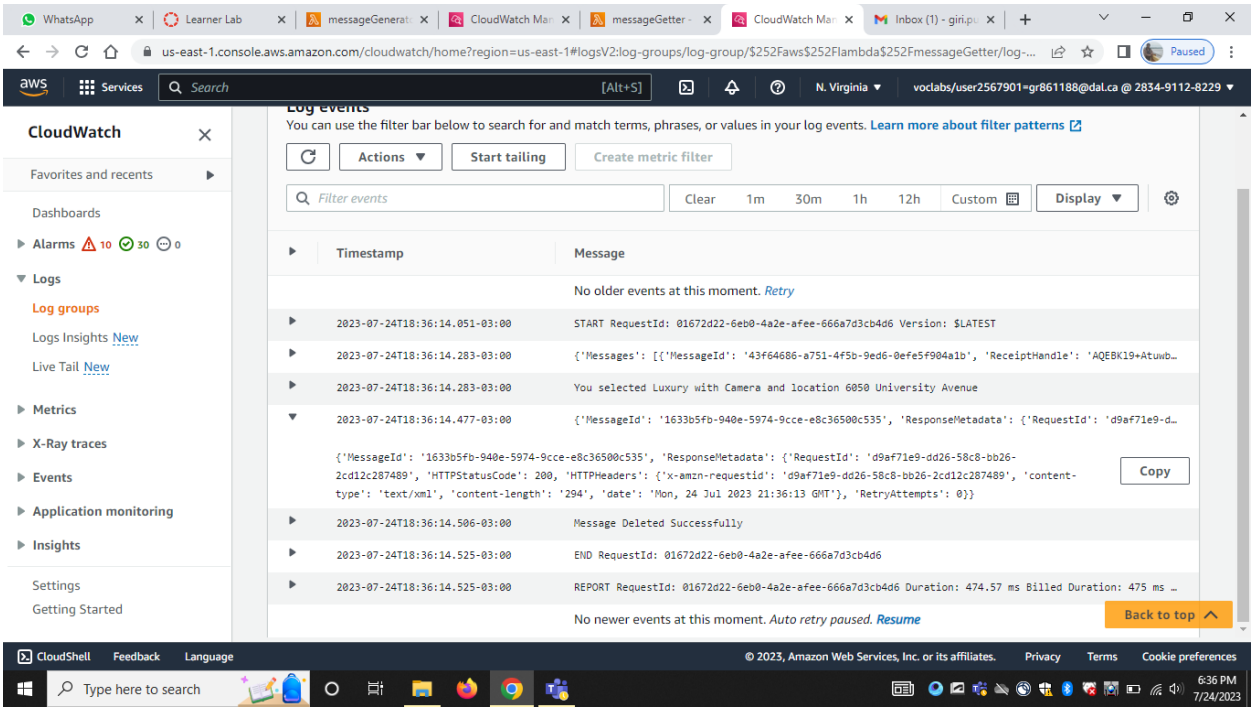

Figure 33:Log of messageGetter lambda function

Message got deleted in the queue once it is processed

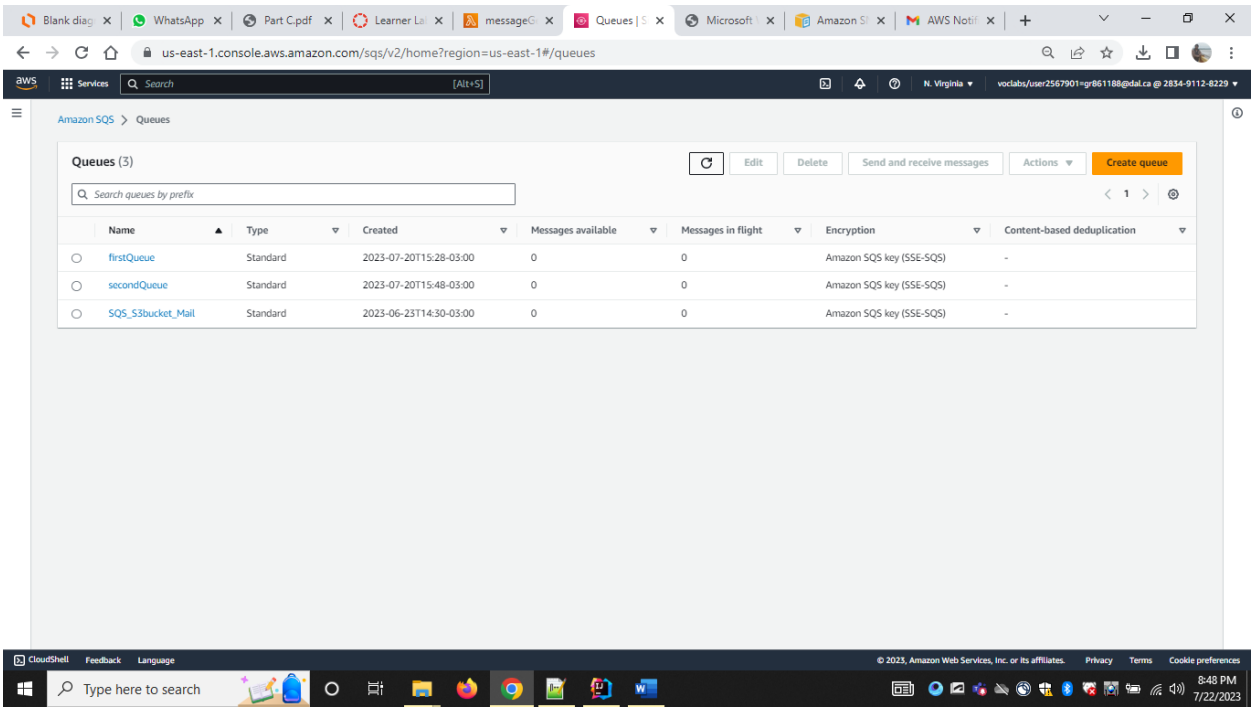

Figure 34: Message got deleted in the queue

References

[1] "Sending and receiving messages in Amazon SQS - Boto3 1.28.9 documentation," *Amazonaws.com*. [Online]. Available: https://boto3.amazonaws.com/v1/documentation/api/latest/guide/sqs-example-sending-receiving-msgs.html. [Accessed: 22-Jul-2023].

[2] "SNS - Boto3 1.28.9 documentation," *Amazonaws.com*. [Online]. Available: https://boto3.amazonaws.com/v1/documentation/api/latest/reference/services/sns.html. [Accessed: 22-Jul-2023].

[3] M. Ozkaya, "Amazon SQS Queue Polling From AWS Lambda with Event Source Mapping Invocations," *AWS Lambda & Serverless — Developer Guide with Hands-on Labs*, 06-Oct-2022. [Online]. Available: https://medium.com/aws-lambda-serverless-developer-guide-with-hands/amazon-sqs-queue-polling-from-aws-lambda-with-event-source-mapping-invocations-2bf34108cef0. [Accessed: 22-Jul-2023].

[4] A. Maksimov, "Boto3 SQS - complete tutorial 2023 2023," *Hands-On.Cloud*, 02-Sep-2021. [Online]. Available: https://hands-on.cloud/boto3-sqs-tutorial/. [Accessed: 22-Jul-2023].

[5] *Amazon.com*. [Online]. Available: https://docs.aws.amazon.com/eventbridge/latest/userguide/eb-create-rule.html. [Accessed: 22-Jul-2023].

[6] *Amazon.com*. [Online]. Available: https://docs.aws.amazon.com/AWSSimpleQueueService/latest/SQSDeveloperGuide/step-create-queue.html. [Accessed: 22-Jul-2023].

[7] *Amazon.com*. [Online]. Available: https://docs.aws.amazon.com/sns/latest/dg/sns-getting-started.html. [Accessed: 22-Jul-2023].

[8] R. Python, "Natural language processing with spaCy in Python," *Realpython.com*, 02-Jan-2023. [Online]. Available: https://realpython.com/natural-language-processing-spacy-python/. [Accessed: 22-Jul-2023].

[9] *Amazon.com*. [Online]. Available: https://docs.aws.amazon.com/code-library/latest/ug/python_3_dynamodb_code_examples.html. [Accessed: 22-Jul-2023].

[10] "SpaCy · industrial-strength Natural Language Processing in Python," *Spacy.io*. [Online]. Available: https://github.com/minway/NlpTest. [Accessed: 22-Jul-2023].

[11] *NlpTest: Test CoreNLP on AWS Lambda*. Available: https://spacy.io/. [Accessed: 22-Jul-2023].

[12] "*Lucid visual collaboration suite: Log in*," *Lucid.app*. [Online]. Available: https://lucid.app/documents#/documents?folder_id=recent. [Accessed: 22-Jul-2023].