**BDM 3203 - Hadoop Ecosystems for Big Data 01**

**Assignment 1**

**Submitted by: Group E**

**Student IDs:**

**Aadarsha Chapagain(C0825975)**

**Roshan Acharya (C0831342)**

**Anjana Kuriakose**

**(C0829580)**

**Onyinye Mbanefo (C0831578)**

**Submitted to: Prof. Teresa Zhu**

* Set up Kinesis Streams with producer and consumer.
* Come with a scenario in which the producer sends message to the consumer receive it.
* Keep all your screen shots and explanation in the pdf report.

Here We have attached the Screenshot and description of AWS Kinesis Implementation using the boto python library and AWS CLI

Install Aws cli

**A screenshot of a computer

Description automatically generated**

Create a profile named blogspot kinesis.

**Text

Description automatically generated**

The command will prompt for AWS key ID and Aws Secret Access Key.

**Creating a data stream**

Graphical user interface, text

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Once the notebook shell is successfully executed we can see that the data Stream is created in Amazon Kinesis dashboard.

Graphical user interface, text, website

Description automatically generated

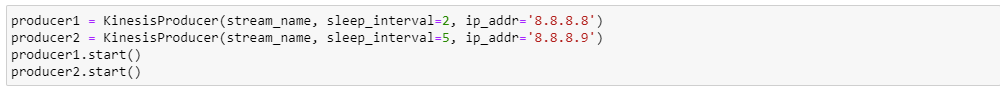
**Creating a Producer**

Here, we will create a class which will emit a record with the IP addresses as a partition key and

Timestamp as the data.

**Graphical user interface, text, application, email

Description automatically generated**

****

We have created two instances of the KinesisProdiucer class which will run at intervals of 2 and 5 seconds and started them.

**Consuming From Data stream**

To consume from the Data stream, we need consumers. Consumers receive messages from the data stream and process it. The output from these consumers can be saved in some files or forwarded to another stream.

**Creating Consumers**

**Graphical user interface, text, application

Description automatically generated**

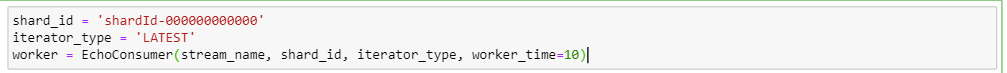
A ‘run’ method is defined here which will check the status of the stream and pass the records to the ‘process\_records’ method, the ‘process\_records’ method is implemented in another class.

Graphical user interface, text

Description automatically generated

This class will simply print the partition key which is an IP address in our case and data(timestamp).

Now, we attach the consumer to our stream. Shard ID and position of the stream is provided to the consumer so that it can start processing messages from given position.



Text

Description automatically generated

Once the consumer is started we can see that the Partition key and timestamp are displayed.

Once the streams are consumed we can delete the stream so that additional cost is not incurred.

Graphical user interface, text, application

Description automatically generated

Summary

We have demonstrated the application of Amazon Kinesis with two producers, a consumer and a data stream.

Reference

https://blog.sqreen.com/streaming-data-amazon-kinesis/