DSC 650 – Big Data

Week 7 Assignment

Eyram Kueviakoe

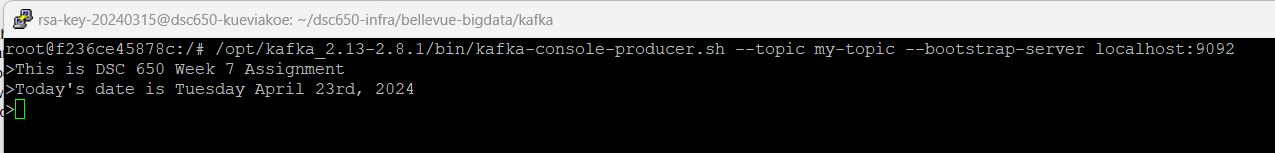
04/23/2024

Screenshot showing the ‘my-topic’ listed amongst the topics.

A screenshot of a computer

Description automatically generated

Screenshot of the producer



Screenshot of the consumer terminal

A computer screen with a black screen

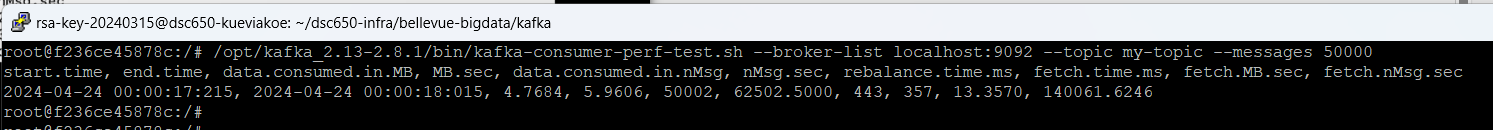
Description automatically generated

Screenshot of the producer performance test results

A screen shot of a computer

Description automatically generated

Screenshot of the consumer performance test results



Discussions of the results

Producer Test: Sent 50,000 records with an average throughput of around 999.82 records/sec and an average latency of 0.89 ms, with occasional spikes in max latency up to 401 ms.

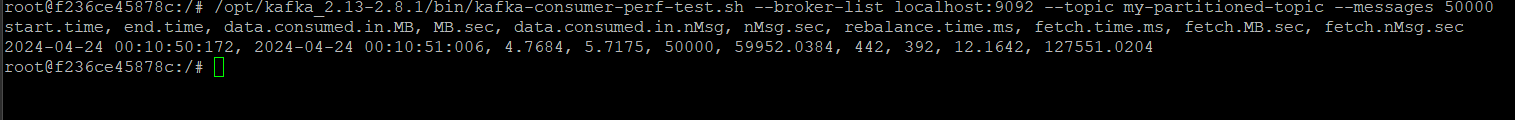
Consumer Test: Consumed 50,000 messages with a consumption rate of 62,502.5 messages/sec and a fetch throughput of 13.3570 MB/sec, showing efficient message retrieval and processing.

The test results indicate that the Kafka producer was able to achieve the target throughput with relatively low latency. The consumer was also able to keep up with the producer, consuming all messages efficiently.

Screenshots of the performance tests on the partitioned topic

A screen shot of a computer screen

Description automatically generated



Observations on performance variations between a single Kafka instance and the scaled setup

Producer Test: Sent 50,000 records with an average throughput of approximately 999.72 records/sec and an average latency of 4.12 ms, with occasional spikes in max latency up to 423 ms.

Consumer Test: Consumed 50,000 messages with a consumption rate of 59,952.0384 messages/sec and a fetch throughput of 12.1642 MB/sec, indicating efficient message retrieval and processing from the partitioned topic.

There is a significant difference in average latency between the single partition test (0.89ms) and the portioned topic test (4.12ms). This variation in performance is due to how the workload is distributed across the 3 partitions in the partitioned topic, which reduces the pressure on any single partition.