Week 6

Felipe Rodriguez

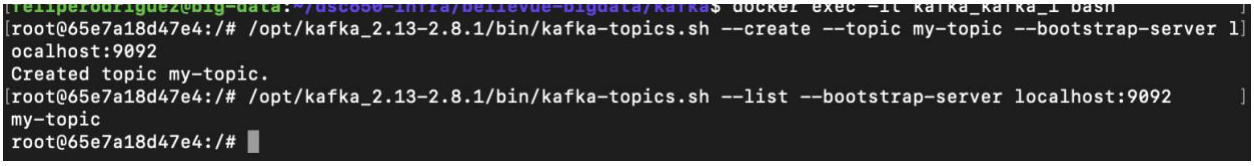
Bellevue University

DSC 650 Big Data

Professor Nasheb Ismaily

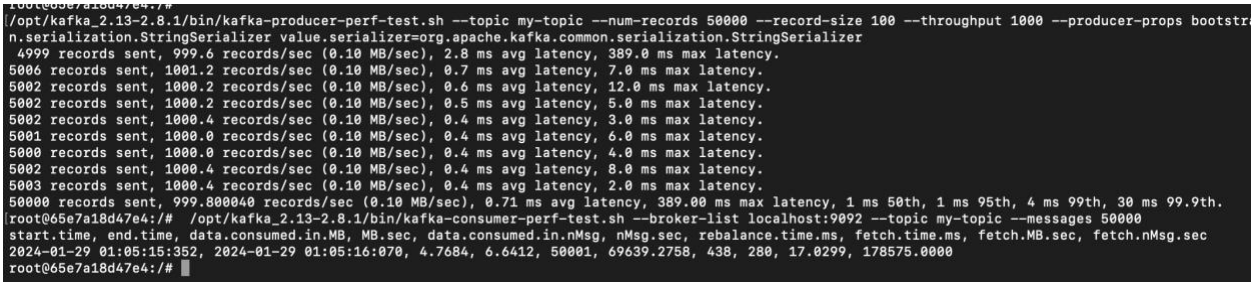
January 28, 2024

Exercise 1 and 2 Kafka topic named ‘my-topic’.



Exercise 3 and 4 Producing and Consuming Messages in Kafka

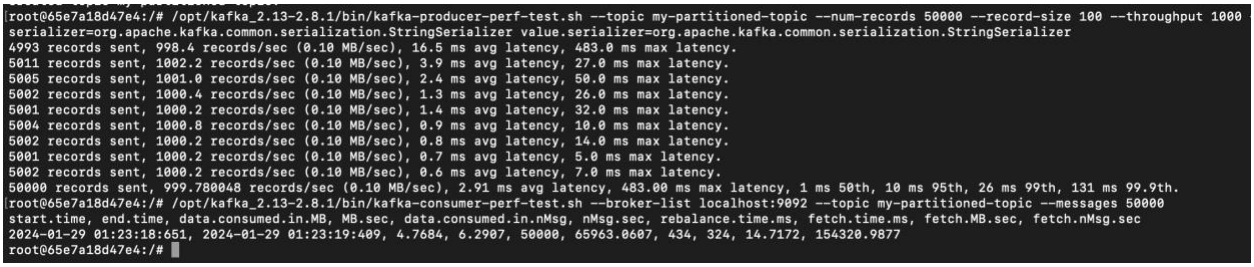
Exercise 5 and 6 Kafka Performance Tests



Interpretation

This performance test allows us to understand how the producer and consumer will act with a certain number of messages. In this example, we are using 50,000 messages. A couple of items to look at are the amount of time it took to generate the messages and the amount of data consumed. This test was small, so it performed well. This test requires us to pay attention to two important metrics; Latency measures - how long it takes to process one event and throughput measures - how many events arrive within a specific amount of time.

Exercise 5 through 8 Expanding Kafka and Running Additional Performance Tests



Interpretation

Between the two performance tests there are not many differences however when performing these tests most systems are optimized for either latency or throughput. The first test consumed 178,575 in a second while the second test was 154,320. The other factor is MB.sec or how much data is transferred in megabytes per second. The first test consumed more data than the second. This makes sense with the number of messages consumed.