

IT 609 - Big Data Processing Lab 6

Teaching Assistant - Saran Pandian P

March 23, 2023

1 Aim

- learn how to use apache Kafka for processing streaming Data.

2 Problem description

Kafka provides APIs for processing streaming Data. In this lab assignment, you build a producer and consumer function in python using Kafka which produces and consumes streaming data.

3 Implementation

3.1 Exercise

- Install Apache python-Kafka (zookeeper).
- start the zookeeper service.
- write the code for producer and consumer services.
- stream the current day, the timestamp from producer to consumer shell.
- Take a screenshot of the results from the producer and consumer terminal.

4 Submission guidelines

You have to submit three files

- Instructions for installing Kafka in Linux/Windows and final results in a pdf file.
- Consumer.py
- Producer.py

5 References

- https://github.com/nancyyanyu/kafka_stock/tree/master/pipeline
- <https://www.geeksforgeeks.org/how-to-install-and-run-apache-kafka-on-windows/>
- <https://betterdatascience.com/apache-kafka-in-python-how-to-stream-data-with-producer/>
- <https://www.youtube.com/watch?v=HIzOpUXhM3U>