Matthew English
100704553
Cloud Computing : Project Milestone 2

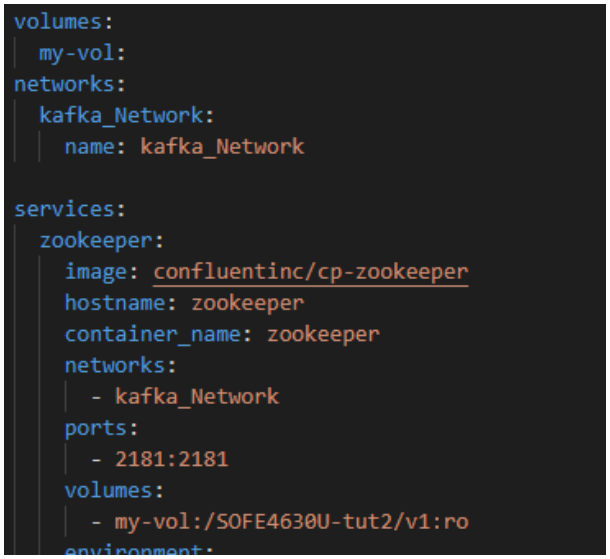
1. What is EDA? What are its advantages and disadvantages?
   EDA stands for event driven architecture which is an architectural model that utilizes events to transfer information. There is an event producer, event consumer and event router. The event producer creates and publishes the event and notifies the consumer about the event.
   Advantages:
   - Processing real-time streaming data
     - Data does not need to wait to be processed at a later time
   - Scalability
     - EDA is very scalable as multiple consumers can consumer an event
   - Reduced Costs
     - Less of a cost than other solutions

   Disadvantages:
   - Duplicated events
     - Singular events can trigger duplicate messages across multiple services if not careful
   - Error handling
     - Extensive monitoring tools are needed to track hundreds or thousands of message brokers that are constantly passing and receiving events.

2. In Kafka, what's meant by cluster, broker, topic, replica, partition, zookeeper, controller, leader , consumer, producer, and consumer group?
   Cluster: One or more servers/brokers running kafka.
   Broker: Handles requests from clients and replicates data within the cluster.
   Topic: A category name where records are stored and published.
   Replica: A copy of the data .
   Partition: A log file containing a subset of topic records.
   Zookeeper: Keeps the state of the cluster.
   Controller: Manages state of partitions and replicas within a cluster.
   Leader: Partitions have a server that acts as a leader where the other servers follow and replicate.
   Consumer: Consumers records/messages from the broker.
   Producer: Sends records/messages to the broker.
   Consumer Group: a group of consumers that divide the partitions of the topics among themselves.

   7. Video 1
   https://drive.google.com/file/d/1BJOMMtZe5bZYAWCkGuA2SPILeqz99rNg/view?usp=sharing

8. Update the YAML file for persistent data (hint: it's related to the volume options in Kafka brokers and zookeeper). Describe how this update solves the problem.
To update this file I created a volume and added it to the docker-compose.yml file as seen below.



```yaml
volumes:
  my-vol:
networks:
  kafka_Network:
    name: kafka_Network

services:
  zookeeper:
    image: confluentinc/cp-zookeeper
    hostname: zookeeper
    container_name: zookeeper
    networks:
      - kafka_Network
    ports:
      - 2181:2181
    volumes:
      - my-vol:/SOFE4630U-tut2/v1:ro
    environment:
```

This update solves the problem by allowing the volume to store persistent data by using the same volume every time it runs.

9. Video 2
https://drive.google.com/file/d/1Ecp_qyN_CF5410lGjBJVgL_Ocd0Ci7o_/view?usp=sharing