Project Milestone-- Data Storage Implementation: KV + relational

Lab 3

Shanjay Kailayanathan

1.

* Sink and Source connectors.
  + A sink connector is used as a pipeline for transferring data between Kafka and consumers whereas a source connector acts as a pipeline between producer and Kafka.
* The applications/advantages of using Kafka Connectors with data storage.
  + The advantages of using Kafka connectors with data storage is that it’s easily maintainable, it offers loose coupling and high latency & throughput.
* How do Kafka connectors maintain availability?
  + Kafka connectors maintains availability by ensuring that a connector restarts itself should the connection ever terminate.
* List the popular Kafka converters for values and the properties/advantages of each.
  + Avro – Fast, compact, best compatibility, rich & extensible schema.
  + JSON – Commonly used.

2.

* What’s a Key-Value (KV) database?
  + A key-value database stores data as a collection of key-value pairs in which a key serves as a unique identifier. Key-value databases are highly partitionable and allow horizontal scaling at scales that other types of databases cannot achieve.
* What are KV databases’ advantages and disadvantages?
  + Advantages - Simple data format makes write and read operations fast. Value can be anything, including JSON, flexible schemas.
  + Disadvantages - Optimized only for data with single key and value. A parser is required to store multiple values. Not optimized for lookup. Lookup requires scanning the whole collection or creating separate index values.
* List some popular KV databases.
  + Amazon DynamoDB, Redis, Oracle NoSQL Database, Microsoft SQL Server.

3. Video link - <https://drive.google.com/file/d/10COz4B1A_98qQmPUZQJTZfJPYivXe4BK/view?usp=sharing>

4.

* List some possible applications that can be implemented by using the uploaded dataset.
  + Quick and easy processing of the dataset can be done using various microservices.