Github Link: <https://github.com/HamzaFarhat/CloudProject1Group8>

Video Link: <https://drive.google.com/drive/folders/1BUQvZQCUqcvLiAJe11c_s0D8AVdqwRJd?usp=sharing>

Sink and Source connectors

* A sink reads data from Kafka topics and inserts the data into a MySQL database
* A source reads data from a MySQL database and inserts into Kafka topics

The applications/advantages of using Kafka Connectors with data storage.

* Allows programs that don’t natively talk to Kafka to talk to Kafka. Particularly important when one is unable to modify the application.
* Use it to link two or more applications which normally don’t talk to each other.

How do Kafka connectors maintain availability?

* Kafka connectors are stateless?

List the popular Kafka converters for values and the properties/advantages of each

* Avro
  + Advantages
    - Default converter
    - Space efficient binary encoding
    - Uses Schemas to enforce message formats
  + Disadvantages
    - Serialized messages are not human readable
    - Difficult to implement compared to String/JSON when used for simple tasks
* Protobuf
  + Advantages
    - Space efficient and fast encoding of data in binary format
    - Supported in many languages
    - Uses Schemas to enforce message formats
  + Disadvantages
    - More difficult to implement compared to String/JSON when used for simple tasks
* String
  + Advantages
    - Simple
    - Well supported in many languages and easily understood by developers
  + Disadvantages
    - Not good when multiple fields are required
* JSON
  + Advantages
    - Human readable
    - Supported natively by many databases
  + Disadvantages
    - High overhead (storage wise)

What’s a Key-Value (KV) database?

* A Key-Value database is a database that uses key-value pairs to store data rather than the traditional row-column structure of tables.

What are KV databases’ advantages and disadvantages?

Advantages:

* KV databases are simple to implement
* Very fast

Disadvantages:

* Not structured
* Inability to do complex queries

List some popular KV databases.

* DynamoDB
* MongoDB
* Redis
* BerkeleyDB

List some possible applications that can be implemented by using the uploaded dataset.

* Navigation systems
* Obstacle detection and tracking