

# Project Milestone-- Data Storage Implementation: KV + relational

**1. Describe the following:**

**Sink and Source connectors.**

Connectors are responsible for the interaction between Kafka Connect and the external technology being integrated with. Source is input whilst Sink is output. A source connector collects data from a system, the sources could be a message queue or a Db. A sink connector delivers data from Kafka topics into other systems, such as Dbs or Elasticsearch

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

The advantage of using Kafka Connectors is that the developer does not have to write multiple pieces of code each time they want to do some data integration. Instead, with Kafka Connectors, you just get to use that piece of code after declaring what the connector is used for using a JSON config file. Kafka Connectors also support REST interface – we can manage connectors using a REST API.

**How do Kafka connectors maintain availability?**

Kafka connectors can be used as load balancers also, meaning that there can be multiple connectors responsible for doing the same task. So in-case one of the connectors goes down, another one can be deployed with the same specifications as the config/specifications will be stored in Kafka.

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

- **String** - converts to string. Good for immediate processing. Not legible as it is not structured
- **JSON** - default format. Very legible and supported by many systems. Lower character count compared to String
- **Protobuf** - Protocol Buffers, or Protobuf, are a smaller/faster/simpler replacement for XML files. Supported by many systems.

**2. Search the internet to answer the following question:**

**What's a Key-Value (KV) database?**

Key-Value Dbs are a type of NoSQL Dbs that consists of 1 key and a value. Note that the value can be any sort of data and the key will be its unique identifier. This means that the data can vary drastically between different entries.

**What are KV databases' advantages and disadvantages?**

- Fast, Simple, Scalable
- No defined rules (Lack of Standardization), Consistency

**List some popular KV databases.**

MongoDB, redis, DynamoDB, Voldemort