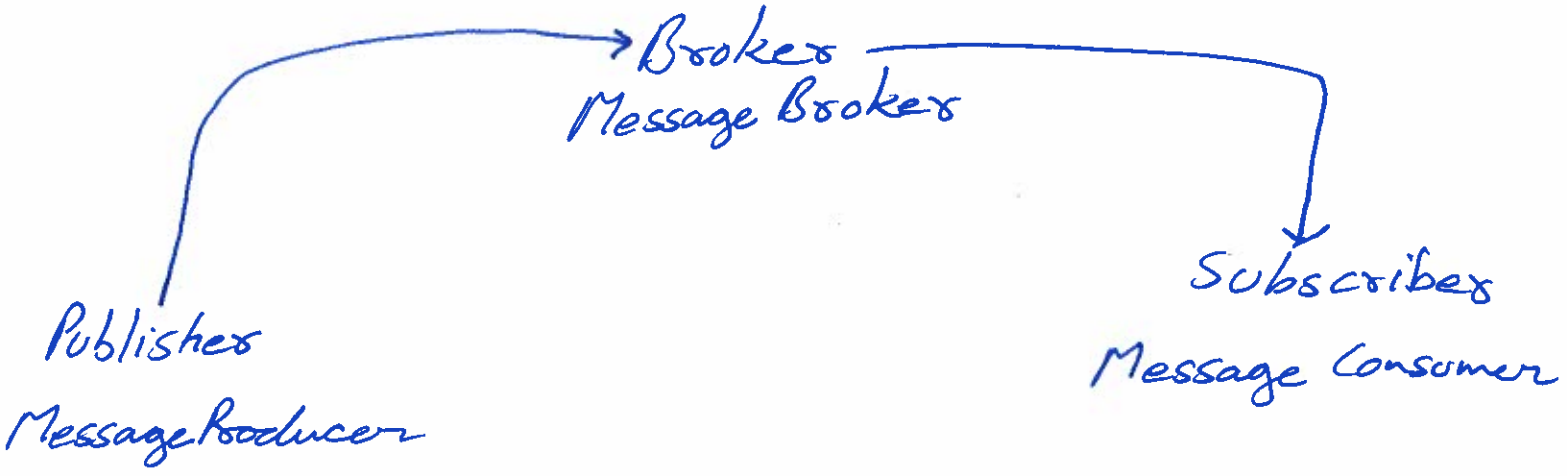
Kafka Work:-



Kafka Components:-

1. Server Software - Brokers
2. Client API - Java Library
   a. Producer API
   b. Consumer API
3. Kafka Connect
4. Kafka Streams
5. KSQL

Kafka Core Components:-

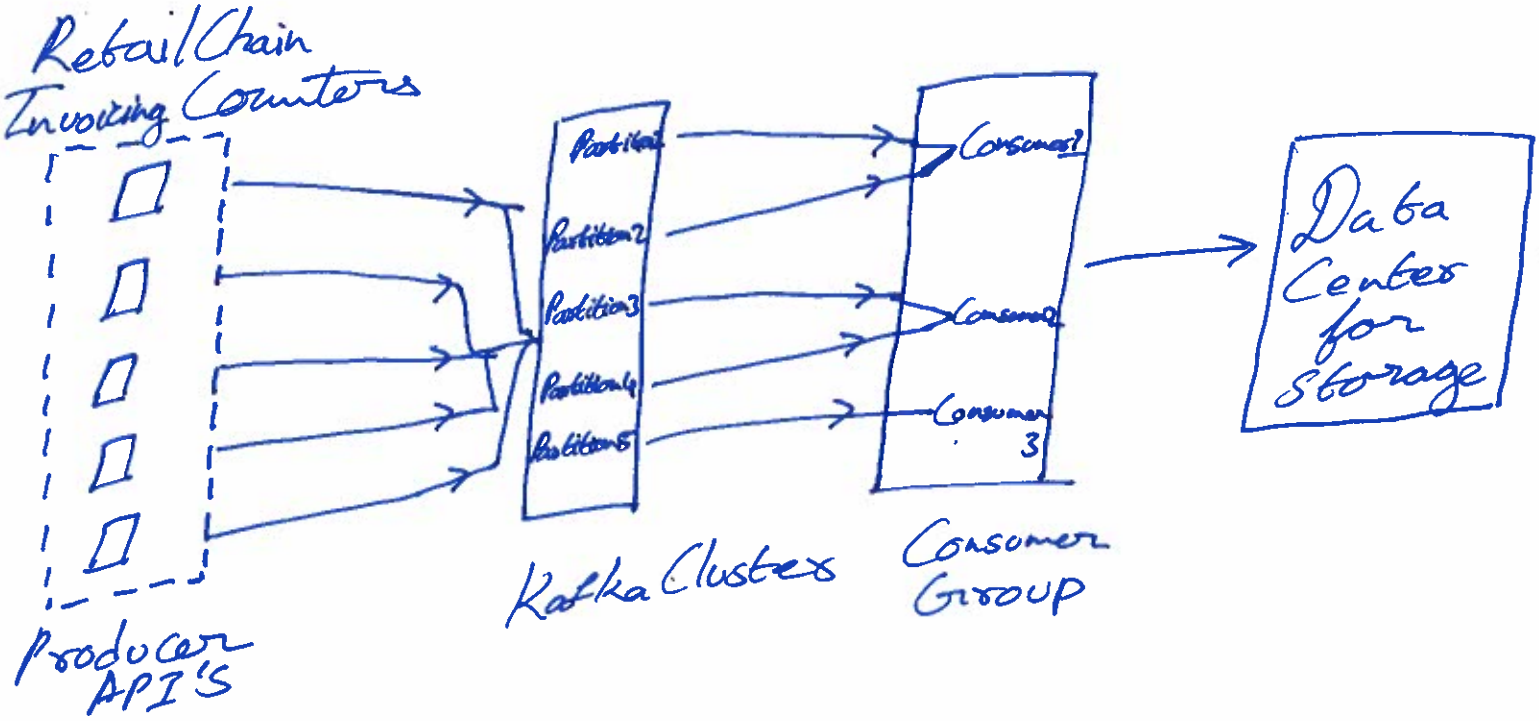
Topic based Streaming is the core component for Apache Kafka.

All interactions are via Kafka Brokers

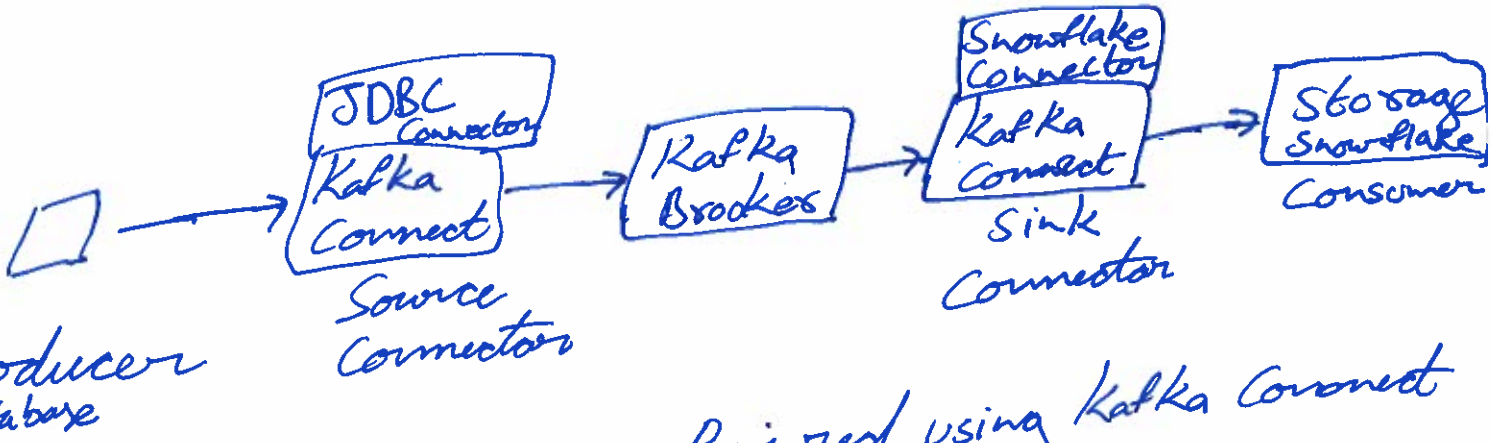
To locate a message the "Topic Name", "Partition Number" & "Offset Number" were needed to be known. Kafka Topic partitions are the core idea of making Kafka a distributed & Scalable System.

Sample Implementation :-

The maximum number of possible parallel consumers are limited by the number of partition of the kafka-topic. Kafka doesn't allow more than one consumer to read & process data from the same partition simultaneously.



Kafka Connect :-



Kafka Connect can be configured using Kafka Connect framework and can be linked to various sources like Relational Databases, Teradata, IoT Hubs, Salesforce, Twitter, Reddit, FileSystem, Cloud Storage, Hadoop Storage, Elastic search, Cassandra, MongoDB & Google Firebase.

Kafka Connect Framework:

1. Source Connector
   - i. SourceConnector
   - ii. Source Task
2. Sink Connector
   - i. Sink Connector
   - ii. Sink Task

Kafka Connect is configured has a cluster, each individual unit in the cluster is called a "Connect worker". Connect worker run "Source Task" & "Sink Task"

Kafka Streams :-

Data Streams are unbounded (No definite starting or ending) often Infinite & ever growing sequence of data in small packets (KB).

Common Examples are :-

1. Sensors
2. Log Entries
3. Click streams
4. Transactions
5. Data Feeds.

Stream processing should produce most recent version of results from realtime streams all the time. "Kafka Streams" is an effective solution for stream processing.

Kafka Streams is a Java/Scala library where the input Data must be in Kafka Topic.

KSQL :—

SQL interface to Kafka Streams

has two modes :-

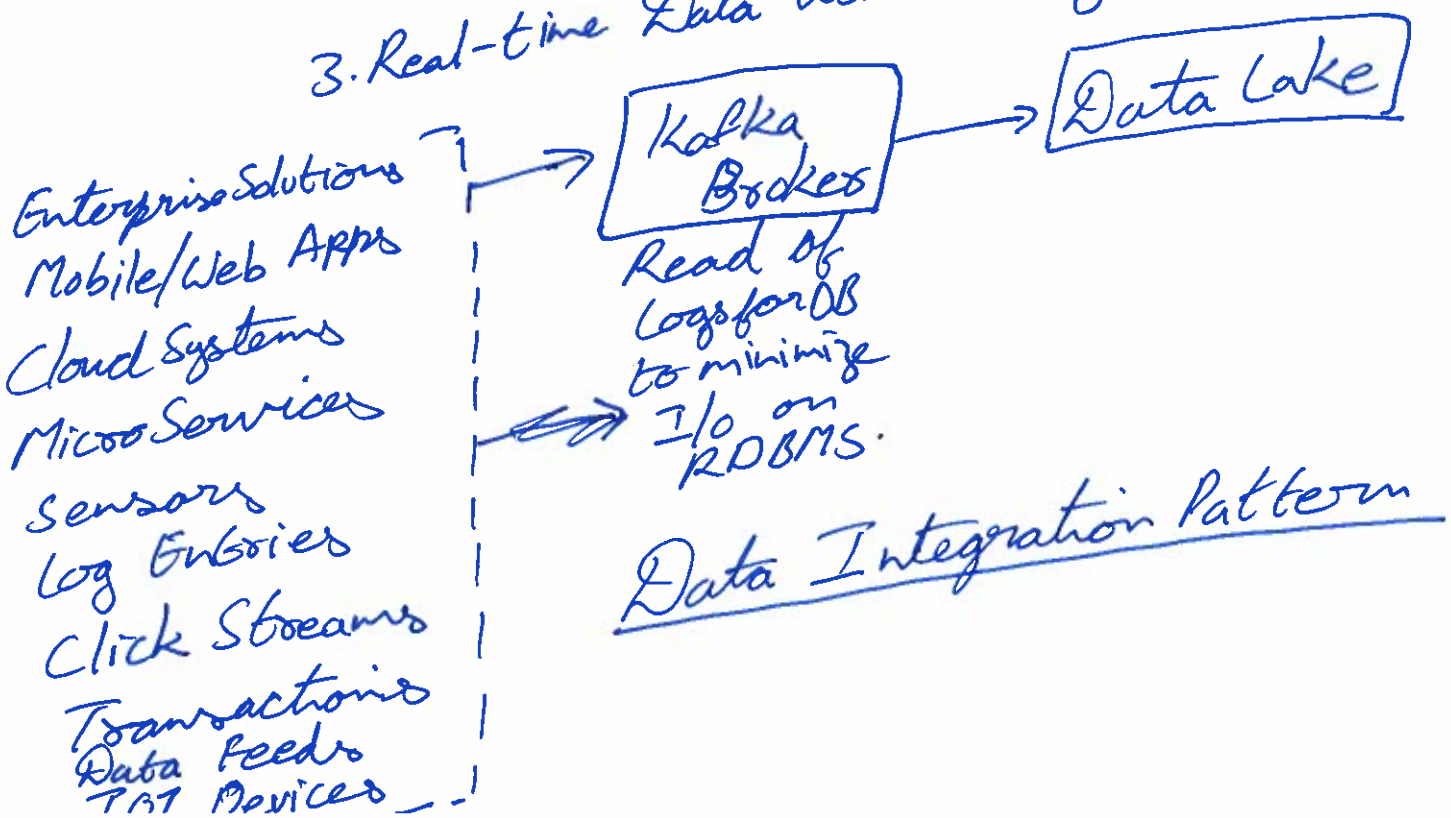
1. Interactive Mode
2. Headless Mode

KSQL Components :-

1. KSQL engine
2. REST interface
3. KSQL Client (CLI/UI)

Kafka Solutions Patterns :-

1. Data Integration Pattern
2. Microservice Architecture for Stream Processing
3. Real-time Data Warehousing



Data Integration Pattern