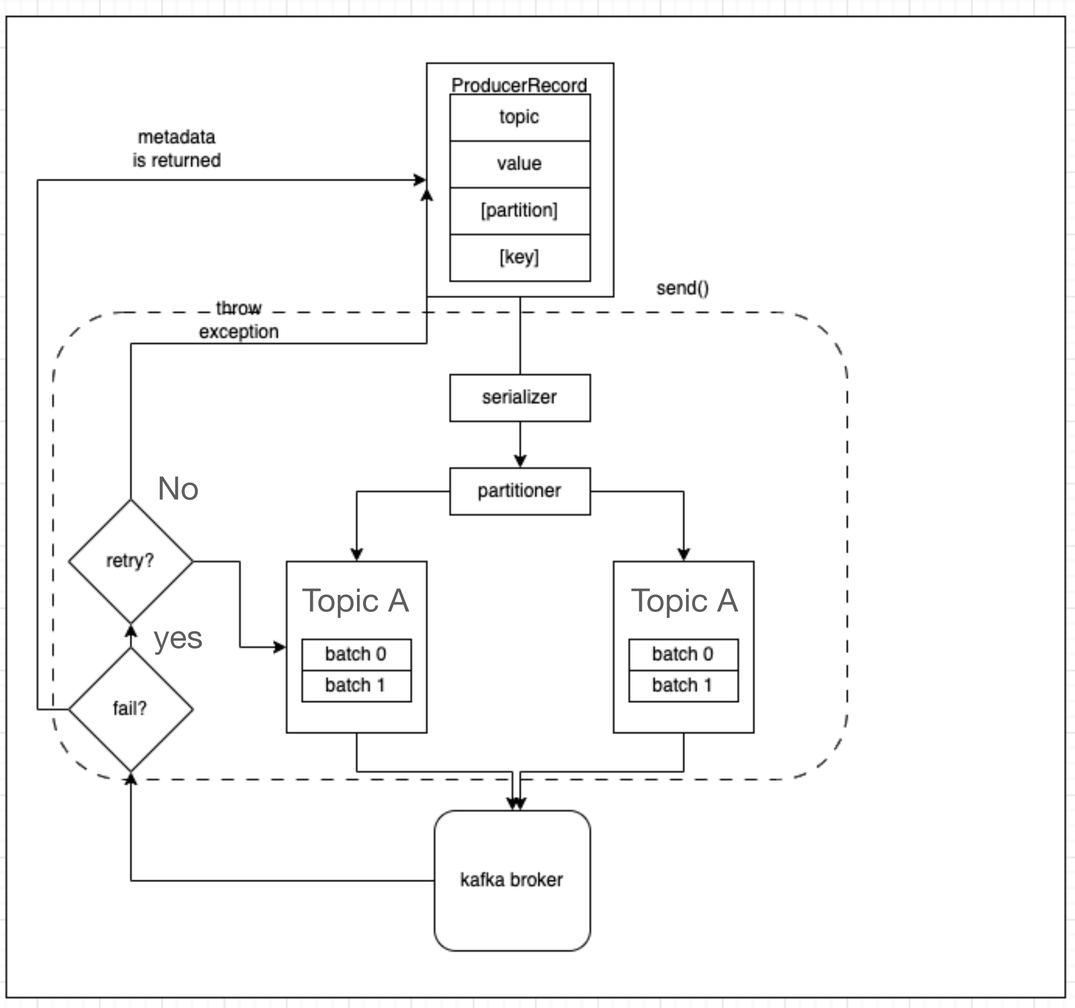
Kafka producers

Overview, types of producers,callback functions, producer configurations, serializers, partitioning, quotas and throttling

Sairam Elangovan

Producer Overview



Constructing a Kafka Producer

- bootstrap.servers
- key.serializer
- value.serializer

Methods of sending producer messages

- Fire and forget
- Synchronous send (use of get() method to wait for future object)
- Asynchronous send

Producer configurations

- client.id
- acks
- linger.ms
- batch.ms
- Compression.type
- max.in.flight.requests.per.connection

Idempotence in kafka producers

- max.in.flight.requests.per.connection <= 5
- Retries greater than zero
- acks=all

Partitioning in producers

- Default partitioner
- In addition to the default partitioner, Apache Kafka clients also provide RoundRobin Partitioner and UniformStickyPartitioner.(used in ETL data pipelines where a certain key value data can be skewed)
- Custom partitioning strategy can also be implemented for special use cases.

Quotas and Throttling

- Kafka brokers have the ability to limit the rate at which messages are produced and consumed. This is done via the quota mechanism
- Quotas can be applied to all clients by setting default quotas, specific client-ids, specific users, or both. User-specific quotas are only meaningful in clusters where security is configured and clients authenticate.
- It is important to plan and monitor to make sure that the broker capacity over time will match the rate at which producers are sending data. (Client memory management)