As soon as a company has real-time data streaming needs, a streaming platform must be put in place. Apache Kafka is one of the most popular data streaming processing platforms in the industry today.

What are the use cases of Apache Kafka?

The use cases of Apache Kafka are many. These include stream processing for different business applications. Apache Kafka makes up the storage mechanism for some of the prominent stream processing frameworks, e.g., Apache Flink, Samza.

- Messaging systems
- Activity Tracking
- Gather metrics from many different locations, i.e., IoT devices
- Application logs analysis
- De-coupling of system dependencies
- Integration with Big Data technologies like Spark, Flink, Storm, Hadoop.
- Event-sourcing store

How is Kafka concretely being used within the industry?

UBER : The speed and flexibility of Kafka allows Uber to adjust their pricing models to the constantly evolving events in the real world (number of available drivers and their position, users and their position, weather event, other events), and bill users the right amount to manage offer and demand.
Netflix : It has integrated Kafka as the core component of its data platform. They refer to it internally as their Keystone data pipeline. As part of Netflix's Keystone, Kafka handles billions of events a day. Just to give an idea about the huge amount of data that Kafka can handle, Netflix sends about 5 hundred billion events and 1.3 petabytes of data per day into Kafka.
Kafka is at the core of lots of the services we enjoy on a daily basis from some of the world's largest tech companies such as Uber, Netflix, Airbnb, LinkedIn, Apple & Walmart.

Kafka Setup

- **Download**
  https://kafka.apache.org/downloads
  Download scala latest version in binary downloads.
- **Setup Kafka in Path**
  We are doing this to access Kafka Binaries with ease.
  - nano .zshrc
  - PATH="$PATH:/Users/shivamjaiswal/kafka_2.13-3.0.0/bin"
  - (We will not be required to provide full path for accessing Kafka binaries)

See complete setup guidance here :
https://www.conduktor.io/kafka/how-to-install-apache-kafka-on-mac

Start Zookeeper

```
zookeeper-server-start.sh   ~/kafka_2.13-3.4.0/config/zookeeper.properties
```

Start Kafka

```
kafka-server-start.sh  ~/kafka_2.13-3.4.0/config/server.properties
```

Changing the Kafka and Zookeeper data storage directory

- edit the zookeeper.properties file at ~/kafka_2.13-3.0.0/config/zookeeper.properties and set the following to your heart's desire dataDir=/your/path/to/data/zookeeper.
- edit the server.properties file at ~/kafka_2.13-3.0.0/config server.properties and set the following to your heart's desire log.dirs=/your/path/to/data/kafka.

Note:- Zookeeper is going to be completely removed from Kafka 4.0.

Install Kafka with KRaft (without Zookeeper) on Mac OS X

1. Install Java JDK version 11
2. Download Apache Kafka from https://kafka.apache.org/downloads under 'Binary Downloads'
3. Extract the contents on your Mac
4. Generate a cluster ID and format the storage using kafka-storage.sh
5. Start Kafka using the binaries
6. Setup the $PATH environment variables for easy access to the Kafka binaries

1. generate a new ID for your cluster.

```
~/kafka_2.13-3.4.0/bin/kafka-storage.sh random-uuid
```

this returns a UUID, for example bs8tpjusSZuLirvAj700Wg

2. Format your storage directory (replace <uuid> by your UUID obtained above)

```
~/kafka_2.13-3.4.0/bin/kafka-storage.sh format -t <uuid> -c ~/
kafka_2.13-3.0.0/config/kraft/server.properties
```

This will format the directory that is in the `log.dirs` in the `config/kraft/server.properties` file (by default `/tmp/kraft-combined-logs`)

3. Now you can launch the broker itself in daemon mode by running this command

```
~/kafka_2.13-3.0.0/bin/kafka-server-start.sh ~/kafka_2.13-3.4.0/config/
Kraft/server.properties
```

Kafka Message Anatomy :

Kafka messages are created by the producer. A Kafka message consists of the following elements:

Key : Key is optional in the Kafka message and it can be null. A key may be a string, number, or any object and then the key is serialized into binary format.

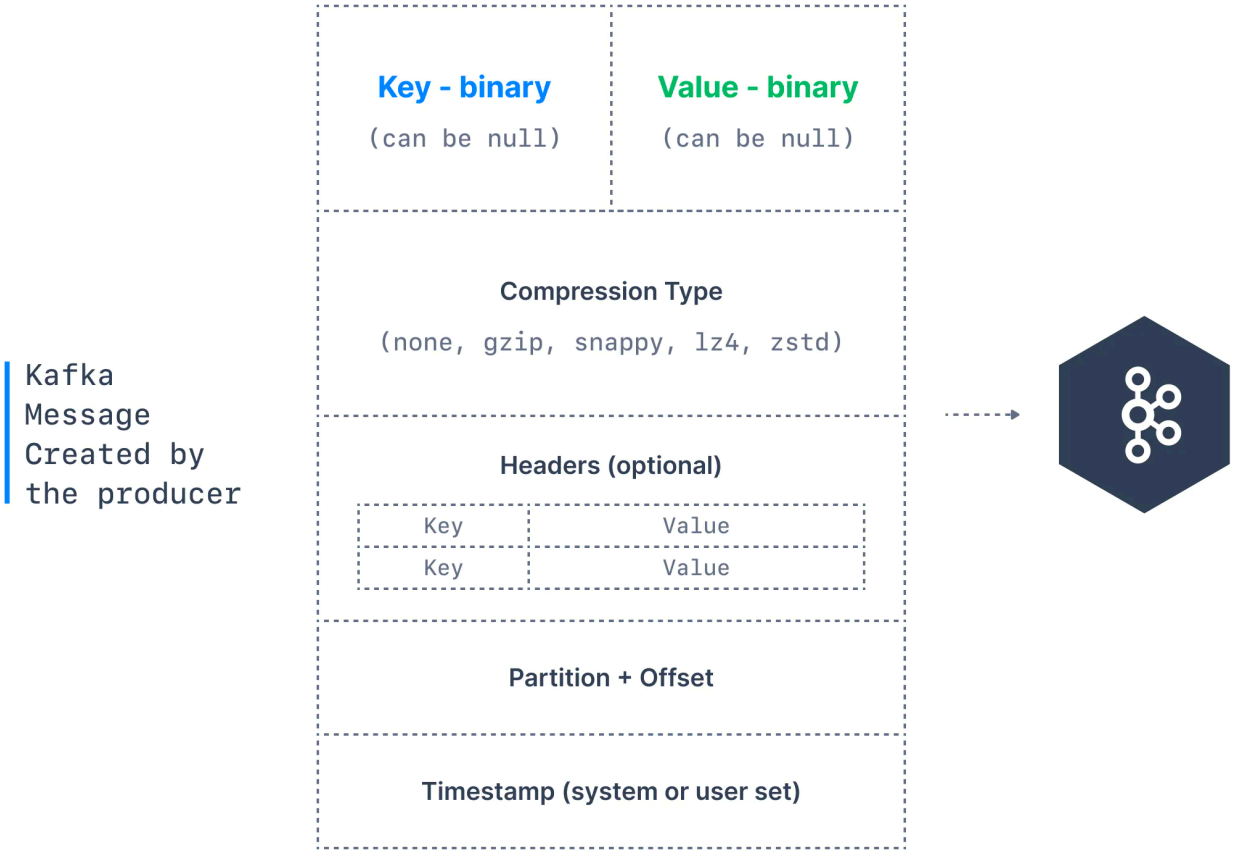
Value : The value represents the content of the message and can also be null. The value format is arbitrary and is then also serialized into binary format.

Compression Type : Kafka messages may be compressed. The compression type can be specified as part of the message. Options are none, gzip, lz4, snappy, and zstd

Headers : There can be a list of optional Kafka message headers in the form of key-value pairs. It is common to add headers to specify metadata about the message, especially for tracing.

Partition + Offset : Once a message is sent into a Kafka topic, it receives a partition number and an offset id. The combination of topic+partition+offset uniquely identifies the message.

Timestamp : A timestamp is added either by the user or the system in the message.



Kafka Terms :

Kafka Broker

- A single Kafka server is called a Kafka Broker.
- Kafka brokers store data in a directory on the server disk they run on.

Kafka Cluster

An ensemble of Kafka brokers working together is called a Kafka cluster.

Bootstrap Server

- A client that wants to send or receive messages from the Kafka cluster **may connect to any broker in the cluster.** Every broker in the cluster has metadata about all the other brokers and will help the client connect to them as well, and **therefore any broker in the cluster is also called a bootstrap server.**
- The bootstrap server will return metadata to the client that consists of a list of all the brokers in the cluster. Then, when required, the client will know which exact broker to connect to to send or receive data, and accurately find which brokers contain the relevant topic-partition.
- In practice, it is common for the Kafka client to reference at least two bootstrap servers in its connection URL

Topic

Similar to how databases have tables to organize and segment datasets, Kafka uses the concept of topics to organize related messages.

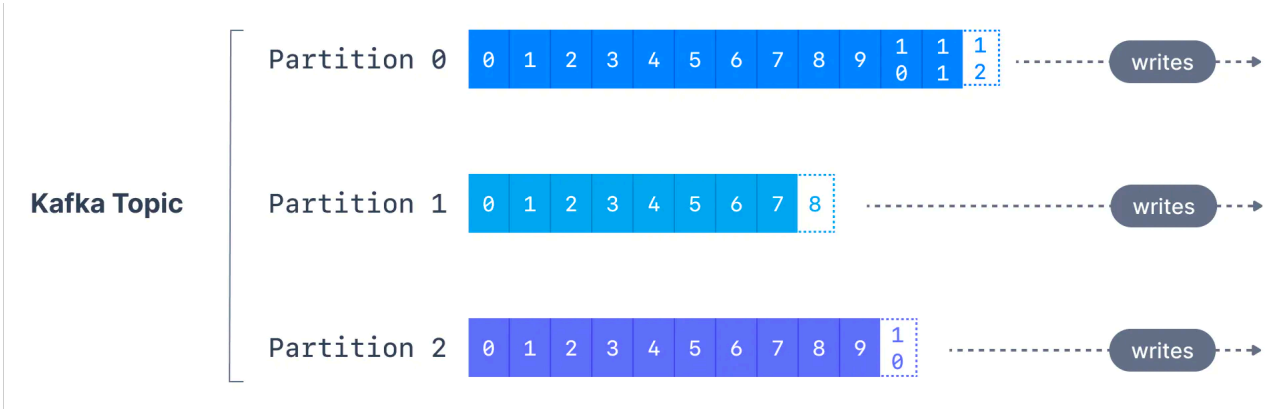
- A topic is identified by its name.
- Kafka topics can contain any kind of message in any format, and the sequence of all these messages is called a **data stream.**
- Data in Kafka topics is deleted after one week by default (also called the default message retention period), and this value is configurable. This mechanism of deleting old data ensures a Kafka cluster does not run out of disk space by recycling topics over time.
- Kafka topics are **immutable**: once data is written to a partition, it cannot be changed

Partitions

- Topics are broken down into a number of partitions.
- partitions for a given topic will be distributed among the brokers evenly, to achieve load balancing and scalability.

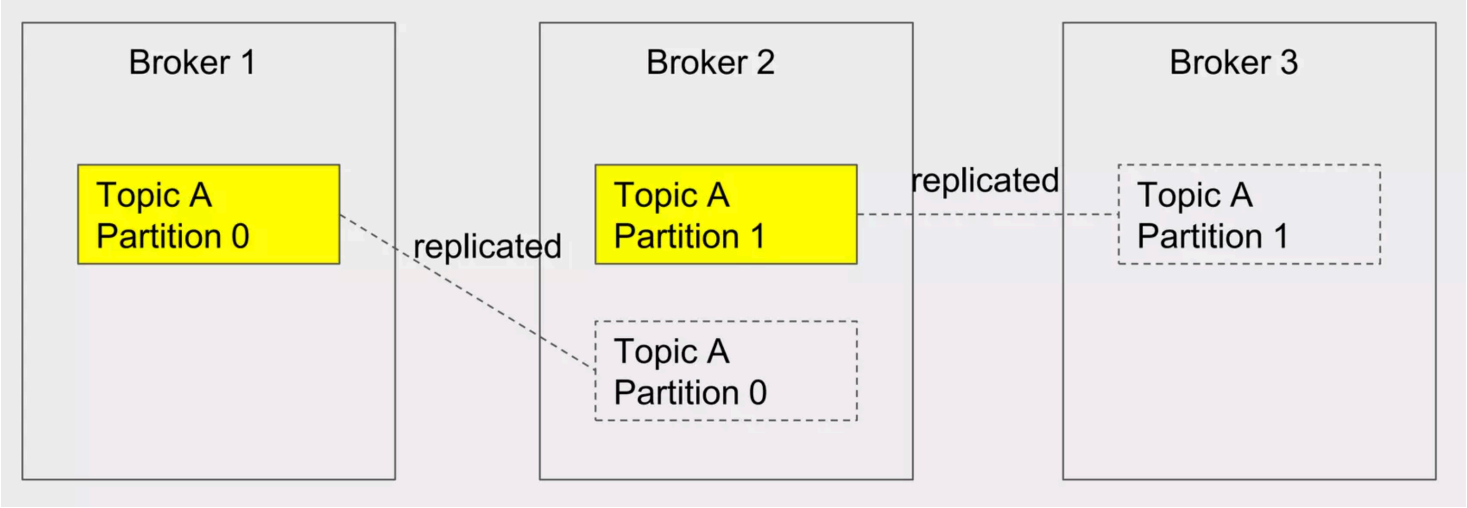
Offset

- The offset is an integer value that Kafka adds to each message as it is written into a partition.
- Offsets are 0 indexed.
- Each message in a given partition has a unique offset.
- Even though we know that messages in Kafka topics are deleted over time (as seen above), the offsets are not re-used. They are continually incremented in a never-ending sequence



Replication Factor

- Replication means that data is written down not just to one broker, but many.
- The replication factor is a topic setting and is specified at topic creation time.
- replication factor of 1 means no replication, replication factor of 3 is a commonly used .



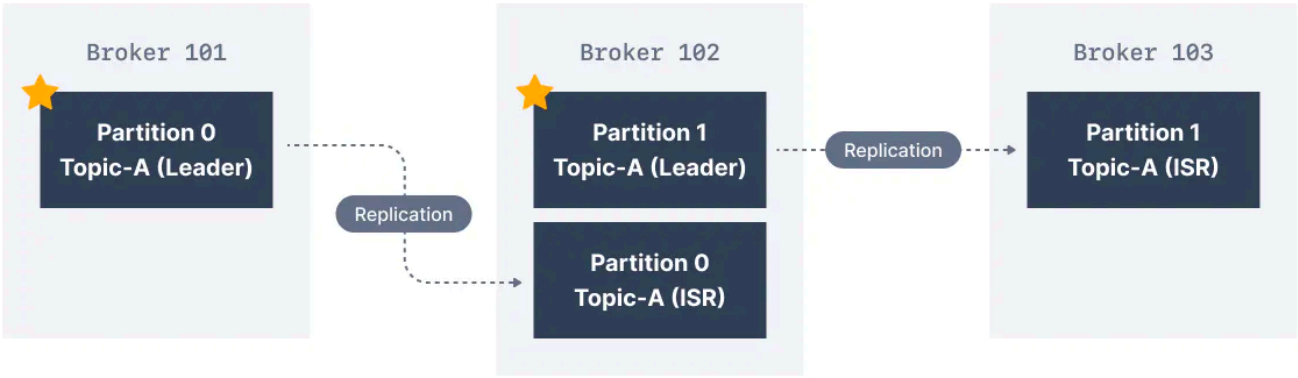
Kafka Partioner

A Kafka partitioner is a code logic that takes a record and determines to which partition to send it into.In the default Kafka partitioner, the keys are hashed using the **murmur2 algorithm.**

Messages sent to a Kafka topic that do not respect the agreed-upon serialization format are called poison pills.

Kafka Partitions Leader and Replicas

- For a given topic-partition, one Kafka broker is designated by the cluster to be responsible for sending and receiving data to clients. That broker is known as the leader broker of that topic partition. Any other broker that is storing replicated data for that partition is referred to as a replica.
- Therefore, each partition has one leader and multiple replicas.
- An ISR is a replica that is up to date with the leader broker for a partition. Any replica that is not up to date with the leader is out of sync.



Ordering

If a topic has more than one partition, Kafka guarantees the order of messages within a partition, but there is no ordering of messages across partitions while reading.

Producers

- Applications that send data into topics are known as Kafka producers. Applications typically integrate a Kafka client library to write to Apache Kafka.
- A Kafka producer sends messages to a topic, and messages are distributed to partitions according to a mechanism such as key hashing.In case the key (`key=null`) is not specified by the producer, messages are distributed evenly across partitions in a topic(Round Robin Fashion).
- For a message to be successfully written into a Kafka topic, a producer must specify a level of acknowledgment (acks).
- We need to use appropriate serialiser to send message to Kafka.

Consumers

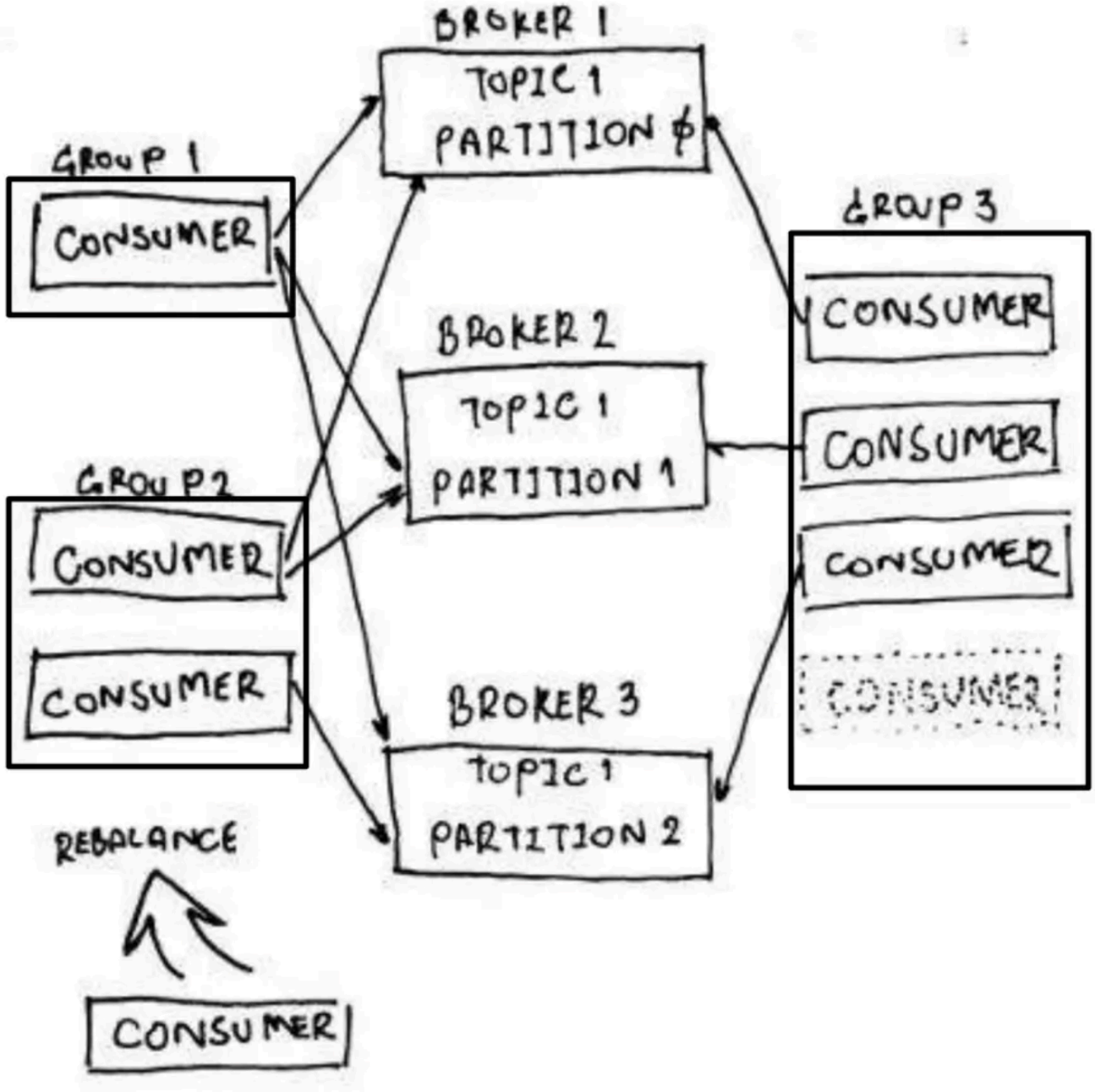
- Applications that read data from Kafka topics are known as consumers.
- Reads data from lower to higher offset.
- Consumers can read from one or more partitions at a time in Apache Kafka, and data is read in order **within each partition** .
- By default, Kafka consumers will only consume data that was produced after it first connected to Kafka.
- Data being consumed must be deserialized in the same format it was serialized in.

Consumer Offsets

- Kafka brokers use an internal topic named `__consumer_offsets` that keeps track of what messages a given **consumer group** last successfully processed.
- the consumer will regularly **commit** the latest processed message, also known as **consumer offset**.
- The process of committing offsets is not done for every message consumed (because this would be inefficient), and instead is a periodic process.
- If rebalancing happens ,assigned consumer starts reading from consumer offset.
- A consumer will not be able to consume messages from beginning if consumer offset is already set for respective consumer group.
- Reset offset cannot happen when the consumer is running.

Consumer group

- Kafka Consumers that are part of the same application and therefore performing the same "logical job" can be grouped together as a Kafka consumer group.
- Consumers automatically use a `GroupCoordinator` and a `ConsumerCoordinator` to assign consumers to a partition and ensure the load balancing is achieved across all consumers in the same group.
- Each topic partition is only assigned to one consumer within a consumer group, but a consumer from a consumer group can be assigned multiple partitions.



Delivery semantics for consumers

- By default, Java consumers automatically commit offsets (controlled by the `enable.auto.commit=true` property) every `auto.commit.interval.ms` (5 seconds by default) when `.poll()` is called.
- A consumer may opt to commit offsets by itself (`enable.auto.commit=false`).
- Depending on when it chooses to commit offsets, there are delivery semantics available to the consumer. The three delivery semantics are explained below.
  - ➡ **At most once**
    Offsets are committed as soon as the message is received.
    If the processing goes wrong, the message will be lost (it won't be read again).
  - ➡ **At Least once**
    Offsets are committed after the message is processed.
    If the processing goes wrong, the message will be read again.
    This can result in duplicate processing of messages. Therefore, it is best practice to make sure data processing is idempotent (i.e. processing the same message twice won't produce any undesirable effects
  - ➡ **Exactly once**
    This can only be achieved for Kafka topic to Kafka topic workflows using the transactions API. The Kafka Streams API simplifies the usage of that API and enables exactly once using the setting `processing.guarantee=exactly_once_v2` (`exactly_once` on Kafka < 2.5).
    For Kafka topic to External System workflows, to *effectively* achieve exactly once, you must use an idempotent consumer.In practice, at least once with idempotent processing is the most desirable and widely implemented mechanism for Kafka consumers.

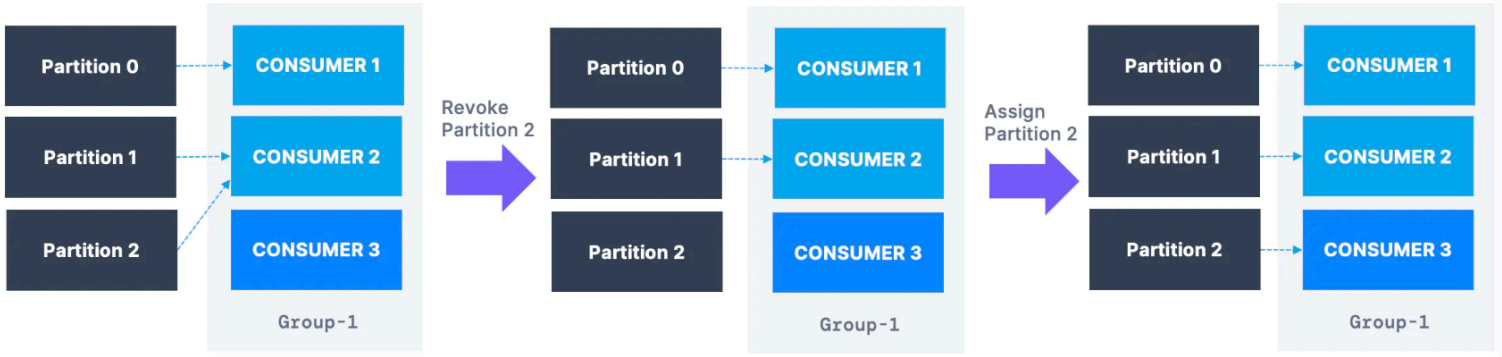
Consumer Rebalance

Consumer rebalances happen for the following events:

- Number of partitions change for any of the subscribed topics
- A subscribed topic is created or deleted
- An existing member of the consumer group is shutdown or fails
- A new member is added to the consumer group
- Rebalancing techniques
  1. **Eager rebalance** : all partitions get revoked from consumer and get

reassigned.

2. **Co-operative rebalancing** : only subset of partition(which are required to rebalance) get revoked hence other consumers keeps on reading the data from unrequited partitions.



Consumer Configurations

1. partition.assignment.strategy
   RangeAssignor: assign partitions on a per-topic basis (can lead to imbalance)
   RoundRobin: assign partitions across all topics in a round-robin fashion, optimal balance.
   StickyAssignor: balanced like RoundRobin, and then minimises partition movements when consumers join/leave the group in order to minimise movements
   CooperativeStickyAssignor: rebalance strategy is identical to StickyAssignor but supports cooperative rebalances and therefore consumers can keep on consuming from the topic.
2. group.instance.id
   it makes the consumer a static member. I.e., upon leaving consumer has session.timeout.ms to join back and get back its partition, without triggering a rebalance.

1. bootstrap.server
   A list of host and port pairs that the consumer uses to connect to the Kafka cluster.
2. group.id
   A unique identifier for the consumer group that the consumer belongs to.
3. enable.auto.commit
   Whether the consumer's offset is automatically committed to Kafka after processing a message.

4. auto.commit.interval.ms
   The interval at which the consumer's offset is automatically committed.
5. auto.offset.commit
   The policy for what to do when there is no initial offset or the current offset is not valid. The possible values are "earliest" (start from the beginning of the partition), "latest" (start from the end of the partition), and "none" (throw an exception if there is no offset).
6. max.poll.records
   The maximum number of records returned by a single call to the `poll()` method.
7. fetch.min.bytes
   The minimum amount of data in bytes that the broker must have before sending it to the consumer.
8. fetch.max.wait.ms
   The maximum amount of time that the broker will wait to accumulate data before sending it to the consumer.
9. fsession.timeout.ms
   The timeout for consumer heartbeats to the broker. If the broker does not receive a heartbeat within this timeout, it assumes that the consumer has failed and initiates a rebalance.
10. heartbeat.interval.ms
    The interval at which the consumer sends heartbeat requests to the broker.
11. max.partition.fetch.bytes
    The maximum amount of data in bytes that the consumer can fetch from a single partition at a time.
12. client.id
    A user-specified identifier for the consumer.

Producer Configurations

1. min.insync.replicas
   it is used together with acks=all , and it assures acks from minimum x brokers.i.e., min.insync.replicas=x{1,2,3,4….}.
2. acks
   Kafka producers only write data to the current leader broker for a partition.It must also specify a level of acknowledgment `acks` to specify if the message must be written to a minimum number of replicas before

being considered a successful write.producers consider messages as "written successfully" for

➡acks = 0 : the moment the message was sent without waiting for the broker to accept it at all.

➡acks =1 : when the message was acknowledged by only the leader.

➡acks=all(-1) : when the message is accepted by all in-sync replicas (ISR). Leader partition ask for acks to other ISR partitions and responds to producer together with its acks.
default values is acks = all (-1)

3. retries
   The retries setting determines how many times the producer will attempt to send a message before marking it as failed. The default values is MAX_INT, i.e., 2147483647 for Kafka >= 2.1.
4. retry.backoff.ms
   time to wait before next retry
   default values is 100ms
5. delivery.timeout.ms
   retries are bounded by a timeout
   default values is 120000ms == 2 min
6. max.in.flight.requests.per.connection
   how many produce request(message batches) can be made in parallel
   default values is 5, set it to 1 to ensure ordering of messages.
7. enable.idempotence =true
   Note : For **idempotent producer** ordering will be same even for higher values and duplicates are not introduced due to network retries.This producers are default from Kafka 3.0.
   if we enable idempotence enable.idempotence=true, then it is required for max.in.flight.requests.per.connection to be less than or equal to 5 with message ordering preserved for any allowable value!!
8. compression.type
   can be at broker or producer level , producer level is recommended.
   Note : at consumer end we don't need to decompress the batches , consumers do it by itself.
   default values is none.
9. linger.ms
   how long to wait until we send a batch. Adding a small number for example 5ms helps add more messages in a batch at the expense of latency.

default values is 0.

10. batch.size
    if a batch is filled before linger.ms, increase the batch size.
    smaller batches leads to more request and higher latency so prefer to have larger batch size.
    default values is 16kb
11. buffer.memory
    if the producer produces faster than the broker can take, the records will be buffered in memory
    default size is 33554432 (32 MB):the size of send buffer
    this buffer will fill up over time and empty back down when the throughput to the broker increases
12. max.block.ms
    the time the .send() will block until throwing the exception.
    Exception occurs when:
    if the buffer is full , then the .send() method will start to block(won't return right away).
    if broker is down or overloaded
    if 60 sec has elapsed

`flush()` and `close()` are required to ensure the producer is shut down after the message is sent to Kafka.

Seek and Assign

to read specific messages from specific partitions, the .seek() and .assign() API may help you. These APIs are also helpful to replay data from a specific offset.

Sticky Partitioner(used for performance improvement)

which means the producer that receives messages sent in time close to each other will try to fill a batch into ONE partition before switching to creating a batch for another partition.
Keys become useful when a user wants to introduce ordering and ensure the messages that share the same key end up in the same partition.

Segment File and Index file

1. As new messages are written to the partition, their offsets and positions are added to the index files.
2. When the segment files reach a certain size or time threshold, they are closed and a new segment file is created.
3. A new index file is also created for each new segment file.

Advanced Configuration

num.io.threads

Kafka uses a pool of threads to handle network I/O operations, such as accepting incoming connections, reading and writing data to sockets, and handling network timeouts. By default, Kafka sets `num.io.threads` to 8, which is typically sufficient for handling moderate levels of network traffic.
If you have a Kafka cluster that is handling a high volume of network traffic, you may need to increase `num.io.threads` to ensure that Kafka can handle the traffic efficiently. However, increasing `num.io.threads` can also increase the CPU and memory overhead of Kafka, so it's important to strike a balance between performance and resource usage.

num.network.threads

It specifies the number of threads that Kafka should use for processing network requests. This parameter determines the maximum number of concurrent client connections that Kafka can handle efficiently.

background.threads

it specifies the number of background threads that Kafka should use for performing background task\s such as log segment deletion, log retention, and quota management.
Kafka uses background threads to perform these tasks in the background without affecting the performance of client requests. By default, Kafka sets `background.threads` to 10, which is typically sufficient for handling moderate levels of background tasks.

num.recovery.threads.per.data.dir

It specifies the number of threads that Kafka should use for log recovery operations. This parameter determines the number of threads that are used to recover data from a failed broker's local disk when it is restarted.

When a Kafka broker fails, the broker's local disk may contain partially written log data that needs to be recovered when the broker is restarted. The recovery process involves scanning the log files on disk and rebuilding the index for each partition.

num.replica.fetchers
It specifies the number of threads that are used by replica fetchers to replicate data between brokers in a Kafka cluster.

unclean.leader.election.enable
It controls whether unclean leader elections are allowed for a topic partition. An unclean leader election occurs when a replica that is not fully caught up with the leader is elected as the new leader for a partition, potentially resulting in data loss.

zookeeper.session.timeout.ms
it specifies the maximum amount of time in milliseconds that a ZooKeeper session can be inactive before it is considered expired and disconnected.

broker.rack
It specify the rack location of a broker in a Kafka cluster. This information can be used by the Kafka cluster to improve fault tolerance and availability by ensuring that replicas are distributed across multiple racks.
In a typical Kafka deployment, you might have multiple brokers running on different physical machines or virtual instances. These brokers can be organized into logical groups or racks based on their physical location or network topology.
By specifying the rack location of each broker using `broker.rack`, you can ensure that replicas for a partition are distributed across different racks, which reduces the risk of data loss or unavailability in the event of a rack-level failure.

quota.producer.default
It allows you to limit the maximum amount of data that a producer can send to a broker per second. This configuration parameter sets a default value for the producer quota, which can be overridden on a per-client basis using the `client.id` configuration parameter.

The producer quota is a mechanism for controlling the rate at which data is produced by a client, which can be useful in preventing overloaded brokers or network bottlenecks. By limiting the amount of data that can be sent by a producer, you can ensure that the cluster resources are used efficiently and that other clients have a fair share of the resources.
Here's an example of how to set `quota.producer.default` using the Kafka broker configuration in Java:

```
Properties props = new Properties();
 props.put("broker.id", “1");
 props.put("listeners", “PLAINTEXT://localhost:9092");
 props.put("log.dirs", "/tmp/kafka-logs");
 props.put("quota.producer.default", "1048576"); // 1 MB/s
KafkaConfig config = new KafkaConfig(props);
KafkaServer server = new KafkaServer(config, new Time(), new
SystemTime());
server.startup();
```

In this example, `quota.producer.default` is set to `1048576`, which limits the maximum amount of data that a producer can send to 1 MB per second. Note that this configuration parameter is specified in bytes per second.
Once `quota.producer.default` is set, it applies to all producers that do not override the `client.id` configuration parameter. To set a custom quota for a specific producer, you can set the `client.id` parameter for that producer and then set the `quota.producer.<client_id>` parameter to the desired value.

For example:

```
Properties props = new Properties();
 props.put("bootstrap.servers", “localhost:9092");
 props.put("client.id", “my-producer");
 props.put("quota.producer.my-producer", "524288"); // 0.5 MB/s
KafkaProducer<String, String> producer = new KafkaProducer<>(props);
```

In this example, a producer with `client.id` set to `my-producer` is created, and the `quota.producer.my-producer` parameter is set to limit the producer to 0.5 MB per second.

Kafka Streams

Stream : a sequence of immutable data records, that fully ordered , can be replayed and is fault tolerant(think of Kafka topic as parallel).

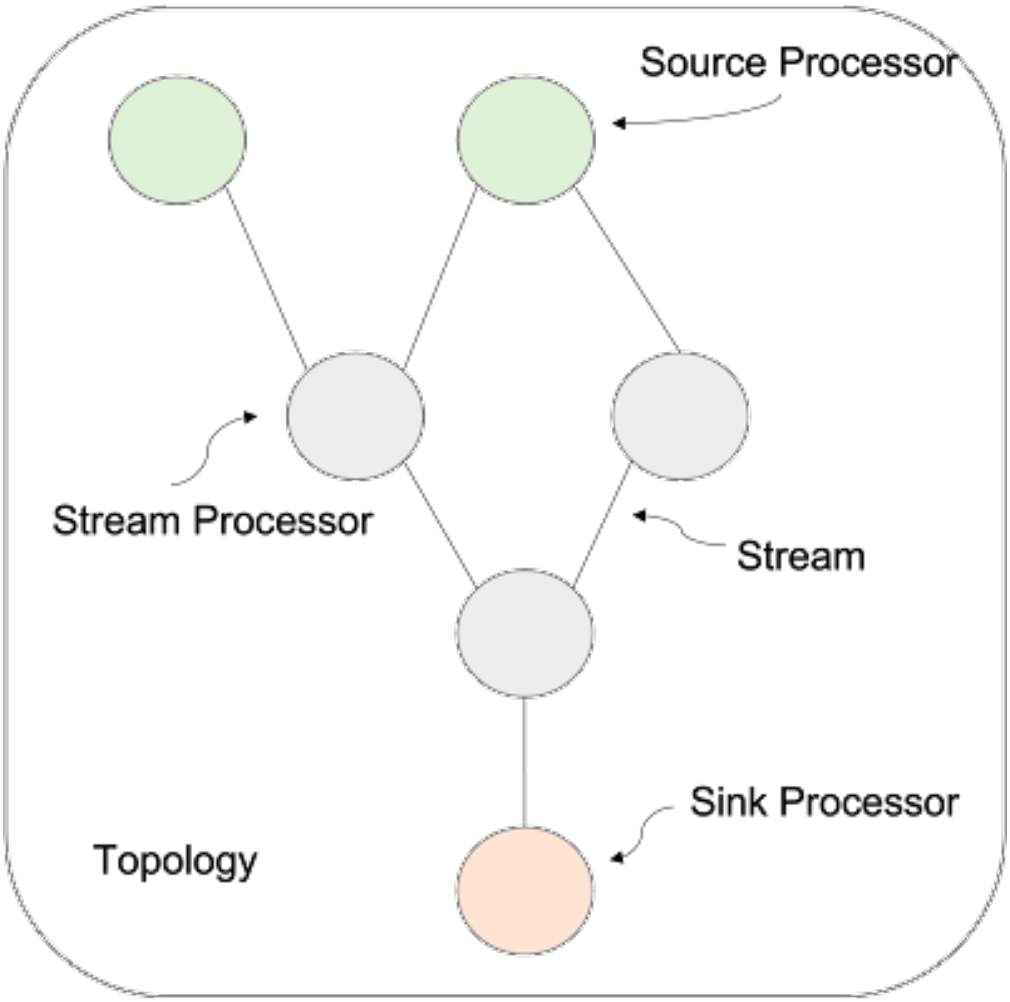
Stream Processor : It is node in the processor topology(graph).It transform incoming stream, record by record, and may create a new stream from it.

Topology : It is a graph of processors chain together by streams.

Source Processor : It is a special processor that takes data directly from Kafka topics. It has no predecessors in a topology and does not transform the data.

Sink Processor : It does not have children , it sends stream data directly to Kafka topic.

In between every other Processor, streams not necessary persisted into Kafka.It lives only within Kafka stream application.



Internal Topics : running a Kafka Streams may eventually create internal intermediary topics.Two types are :
1. Repartitioning topics : Incase we start transforming the key of our stream , a repartitioning will happen at some processor.
2. Changelog topics : Incase we perform aggregations, Kafka stream will save compacted data in these topic.
Note : we should not mess with internal topics, they are for internal use of Kafka streams, I.e., save, restore and repartition data.

We can write a topology by two ways
High Level DSL : it is the simplest. It has all the operations we need to perform.
Low Level Processor Api : It is an imperative API,Can be used to implement the most complex logic.(Rarely Used)

Some extra Configuration for Kafka Streams
1. application.id
   => Consumer group.id = application.id
   => default client.id prefix
   => prefix to internal changelog topics
2. default.[key/value].serde
   for serialisation and deserialisation of data

Scale our Application : run multiple instances for Kafka streams🫣.

Data Streams
Unbounded : no definite starting or ending often infinite and ever growing sequence of data in small packets(KB)
Examples :
- Sensors : sending data from transportation vehicles, industrial equipments , health care devices and wearables.
- Log entries : generated by mobile app, web applications, infrastructure components.
- Transactions : coming from stock market , credit cards, ATM machines, e-commerce, logistics and food delivery orders.
- Data feeds : social media activity.

Pretty much everything can be seen as sequence of small data packets.And this are unbounded and ever growing.

Now we have challenge to process them.

One common approach is to collect and store them in a storage system. Then you can do two things. Query the data to get answers to a specific question.

This is what we know as a request-response approach and usually handled through a SQLs.

This approach is all about asking one question at a time and getting the answer to the question as quickly
as possible.

The second approach is to create one big job to find answers to a bunch of queries and schedule the job to run at regular intervals.
This approach is all about asking a bunch of questions at once and repeating the question every hour or maybe every day. And this approach is known as batch processing.

The stream processing sits in-between these two approaches.
In this approach, we ask a question once, and the system should give you the most recent version of the answer all the time.
So, stream processing is a continuous process and the business reports are updated continuously based on the data available till time.

Technically stream processing is also a data processing.
But we do it in continuous and ongoing process.

Can we us database or batch processing system to perform stream processing??
yes , but it is going to make our solution too complex so we needed a tool that is specifically design for stream processing.so that's why we use Kafka streams.

What is Kafka Streams?
A java/scala library. We can embedded it in micro services
Input data are streamed in Kafka topics.

By adding this our application is just a typical application, with inherent parallel processing capability, fault tolerance and scalability.