Contents

[Storage Architecture 1](#_Toc152459101)

[Topic 1](#_Toc152459102)

[Partition 2](#_Toc152459103)

[Brokers 2](#_Toc152459104)

Unsorted Starts

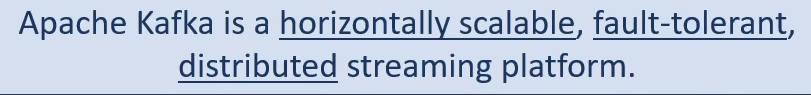
# Unsorted Starts

1. **Topic creation Command**:  
   %KAFKA\_HOME%\bin\windows\**kafka-topics.bat** --create --zookeeper localhost:2181 **--topic invoice** --partitions 5 --replication-factor 3 **--config segment.bytes**=1000000

Unsorted Ends

Definition

# Definition

1. 

##### Storage Architecture

###### Topic

1. Logical name. All the msgs can not be put in a single topic so topic is divided into partitions where each partition is physical Dir.

###### Partition

1. Topic is divided into partitions as all the msgs can not be put into a single topic. Partition is a physical dir.
2. Diagram

   Description automatically generated
3. Partitions are equally distributed over all the brokers.   
   If Partition# > Broker# then some brokers will get more partitions than the other brokers.
4. For each partition and replica, a partition and a replica directory are created under log Dir (**log.dirs**=../tmp/kafka-logs-0)

###### Replica

1. Number of copies for each partition.
2. **Command**:  
   %KAFKA\_HOME%\bin\windows\**kafka-topics.bat** --create --zookeeper localhost:2181 **--topic invoice** --partitions 5 **--replication-factor 3** --config segment.bytes=1000000  
   Graphical user interface, text

   Description automatically generated
3. Replicas can be classified into two categories.  
   Graphical user interface, text, application

   Description automatically generated

###### Logs/Segments

1. The msgs are stored in the log file in the partition directory (Data Directory).
2. However, instead of storing all the msgs into a single log file, Kafka generates several smaller log files known as **segments**.  
   Graphical user interface, text

   Description automatically generated
3. Each segment has a size and when reached, a new segment is created and next msgs are stored there.   
   **The default limit** is 1 GB or a week of data.   
   We can change it when creating a topic like:  
   A screen shot of a computer

   Description automatically generated

##### Brokers

1. **Commands**:
   1. %KAFKA\_HOME%\bin\windows\**kafka-server-start.bat** %KAFKA\_HOME%\etc\kafka\**server-0.properties**
   2. **server-0.properties**
      1. # A comma separated list of directories under which to store log files

**log.dirs**=../tmp/kafka-logs-0 🡸 Path is relative to the working Dir where it will be run.

1. This directory will maintain the state of this broker/server.
2. It will contain:
   1. Log/Data Directory which in turn will contain log files having msgs.
      1. # The id of the broker. This must be set to a unique integer for each broker.

**broker.id**=0

* + 1. # The address the socket server listens on

**listeners**=PLAINTEXT://:9092