Kafka Summary :-

1. DataSource -> Kafka -> Target System
2. Kafka is composed of : Producer -> Kafka Brokers -> Consumer
3. So whole pipeline is : DataSource -> Producer -> Kafka Brokers -> Consumer -> Target System
4. Now in the above scenario:
5. Concepts come of Topics and Partitions : Topics are like RDBMS tables i.e. the set of records or rows and the partitions are like the chunks of records ( with no overlapping)
6. Partitions increase parallelism
7. Partitions have the messages in order
8. Each partition has messages in order and each message has order Id
9. Message is identified by both : partition and the topic ( of course)
10. Topic can or can not be divided into multiple partitions( Further as u will see that as one consumer in the consumer group can subscribe to a partition Very important each partition can be subscribed to only one consumer . One consumer can access the multiple partitions but one partition can be subscribed to single consumer)
11. Brokers or Bootstrap servers – They are the constituents of the kafka distributed cluster
12. They contain the paritions of the topic
13. There is a concept of replication so each partition can be replicated certain number of times in various brokers