**Topic and Partitions**

1.The Kafka producers send the messages/data to a particular topic and consumers read the

data from that topic.

2. A topic can further be subdivided into partitions.

3. A record is stored in a partition.

**How DefaultPartitioner works?**

The below explains the default partitioning strategy:

1.If a partition is specified in the record, use it - this can be used when we already know the partition ,Ex-

producer.send(new ProducerRecord<String, student>("rr132",0,stdobject.getDept(),stdobject));

Targeted partition 0.

2.If no partition is specified but a key is present, choose a partition based on a hash of the key -

It can be used when we want to distribute the data based on a key. The following formula is

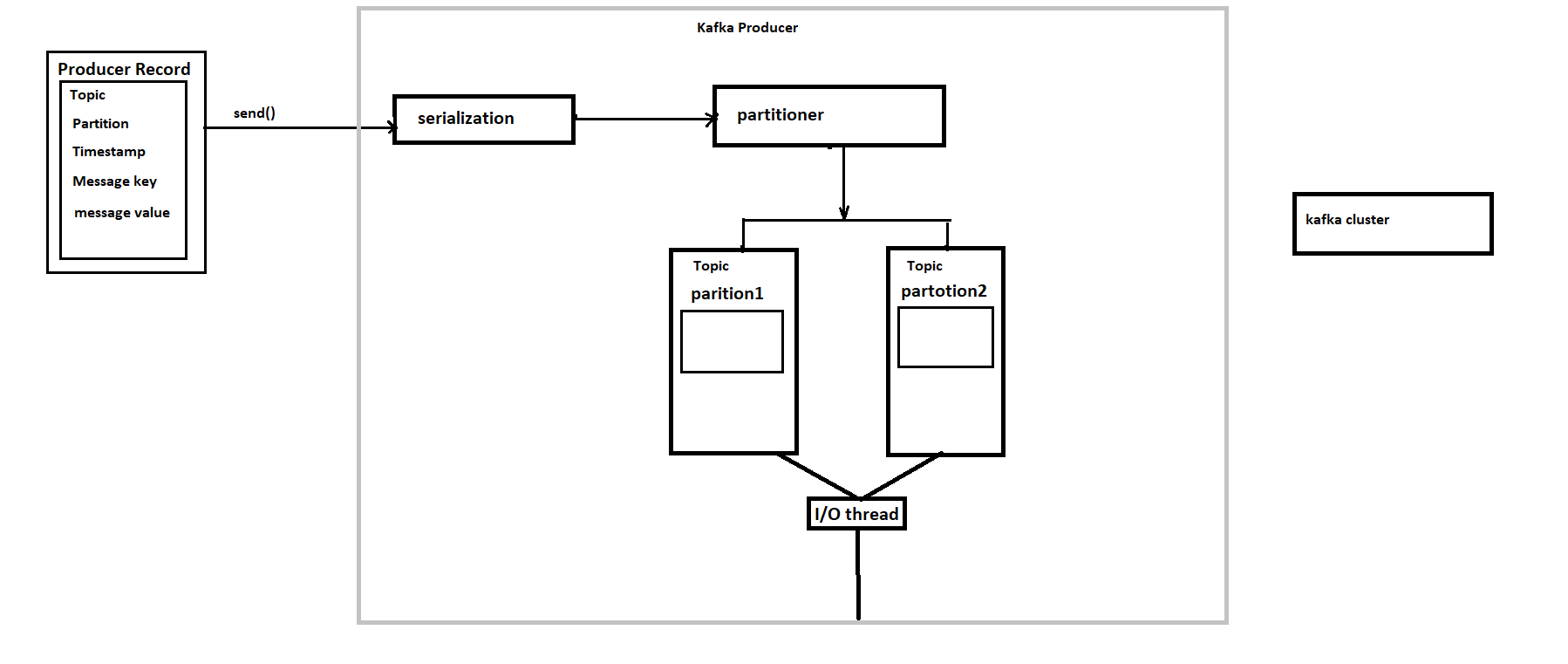
used to determine the partition:

**hashCode(key) % noOfPartitions**

3.If no partition or key is present, choose a partition in a round-robin fashion -

If we are not bothered which partition our data is going.

**If a producer doesn't specify a partition key & partition number when producing a record, Kafka will use a round-robin partition assignment**.



Q.1)**How can we store the data partition wise ?**

Ans: we can send partition number to store data into specific partition.

1.Targeted partition:

producer.send(new ProducerRecord<String, student>("rr132",0,stdobject.getDept(),stdobject));

2.key is used for partitioning:

Ex- stdobject.getDept()

3.Default - Round robin

Q.2)**How can we read or write the data partition wise ?**

Ans:

How to read from a specific offset and partition with the Kafka Console Consumer ?

If your Kafka topic is in Confluent Cloud,

use the kafka-console-consumer command with the --partition and --offset flags to read

from a specific partition and offset.

docker run -v $PWD/configuration/ccloud.properties:/tmp/ccloud.properties

confluentinc/cp-kafka:6.1.1 \

bash -c 'kafka-console-consumer \

--topic example-topic \

--bootstrap-server `grep "^\s\*bootstrap.server" /tmp/ccloud.properties | tail -1` \

--consumer.config /tmp/ccloud.properties \

--property print.key=true \

--property key.separator="-" \

--partition 1 \

--offset 3

Q.3)**What happens if the disk or kafka directory gets full ?**

Ans: the Apache Kafka broker process can't start and fails because of the disk full error.

Increase the size of the disks & Lower your retention policy .The right retention period

varies between different use cases.

**Q.4) We need to check how data can be stored on the basis of key-value pair ?**

By using dump utility command we can check.

./kafka-run-class.sh kafka.tools.DumpLogSegments --deep-iteration --print-data-log --files /u01/kafka-logs/rr132-1/00000000000000000000.log

**Q.5)Is it possible to read the data on a partition on the basis of key-value pair ?**

Ans: No, we can not read using key-value pair from partition but we can apply business logic to

sort them.

Offset are auto incremental , immutable

**Q.6)Is it possible to allocate a specific consumer for consuming the data on partitions?**

Ans: we can use assign method in consumer API to assign consumer to partition

Java one supports the assign method that you can use instead

of the subscribe for asking to be assigned specific partitions for the topic.

Of course, you lose the auto-rebalancing feature and you have to handle these use

case manually.