

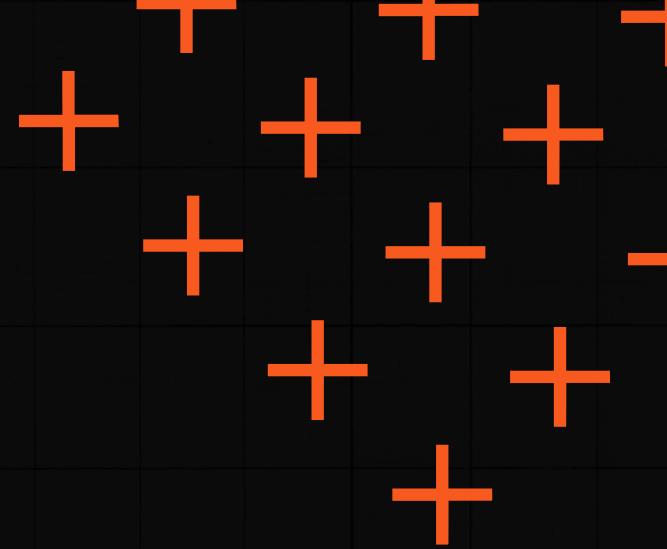
BigData Series

PART 17/xx

Consumer
Replicas

@Abhijith





Recap :

topics:

- named event streams
producers write,
consumers read
independently.

partitions:

- parallel lanes;
ordered logs
enabling horizontal
throughput scaling.

Offsets:

- per-partition
positions; commits
enable replay,
recovery, control.



Consumers:

1.

- read messages from specific offset.

2.

- read from any chosen offset point.

3.

- join any cluster at any point in time.





Consumer Groups:

1.

- consumers can join any group

2.

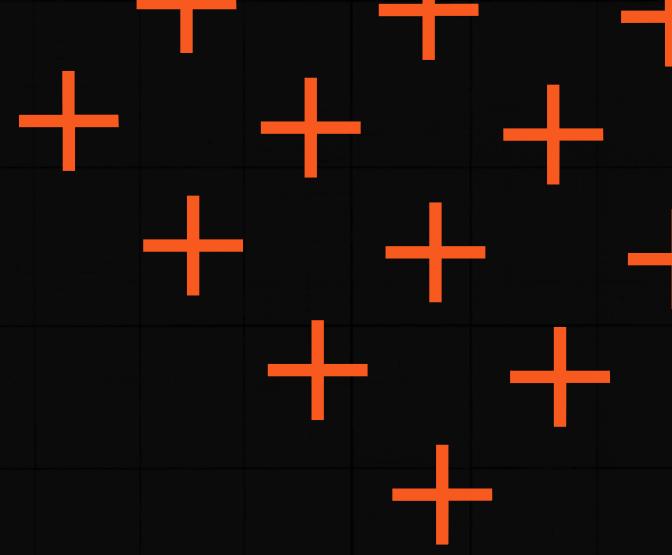
- can consume from same topic or same cluster.

3.

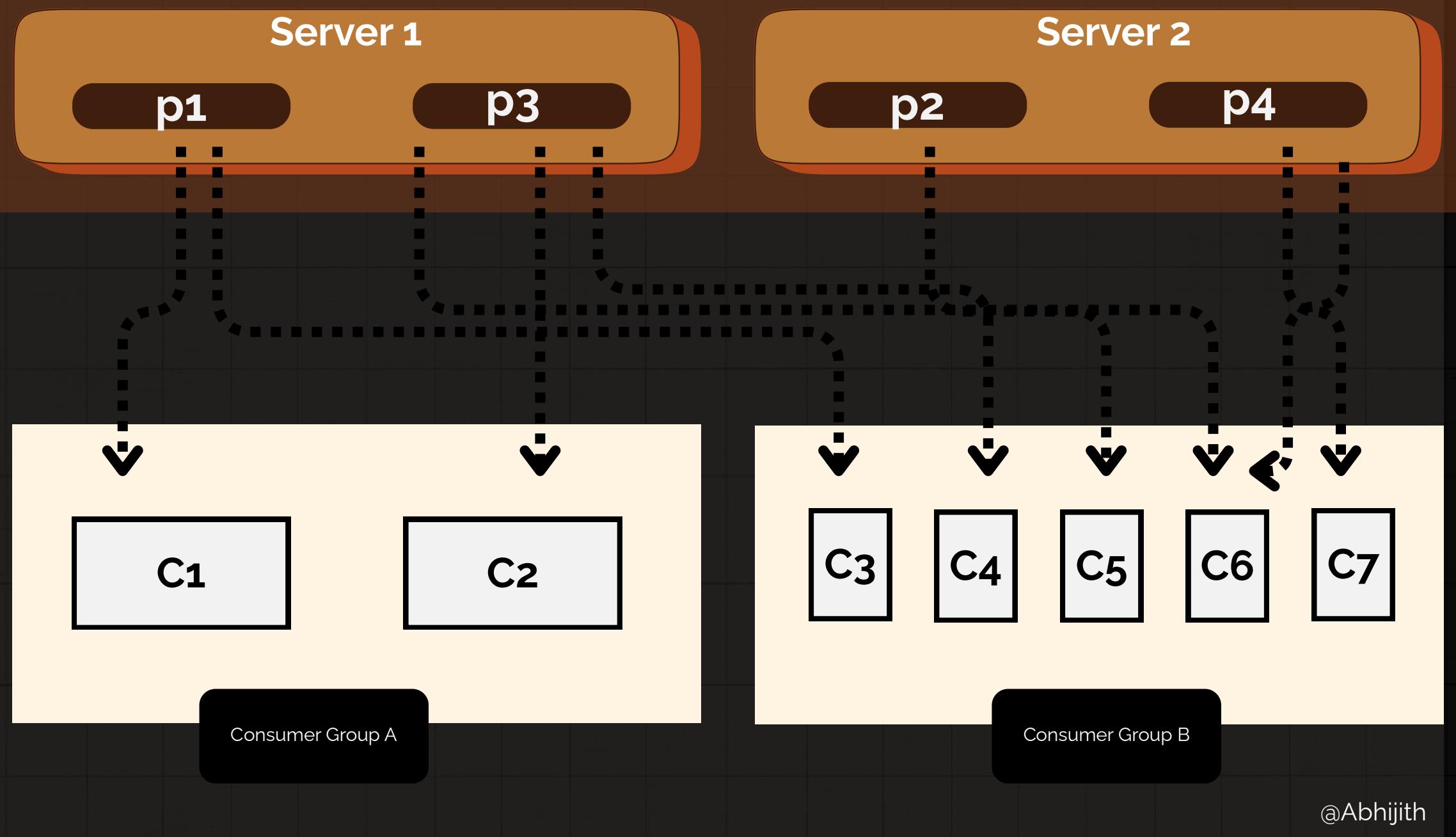
- set of consumers subscribed to same topic.



66



Kafka Cluster: Topic T1





Replication:

1.

- **VERY IMPORTANT** for fault tolerant systems.

2.

- replication implemented at partition level.

3.

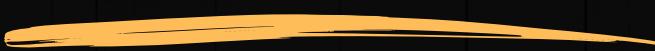
- prevents data loss

4.

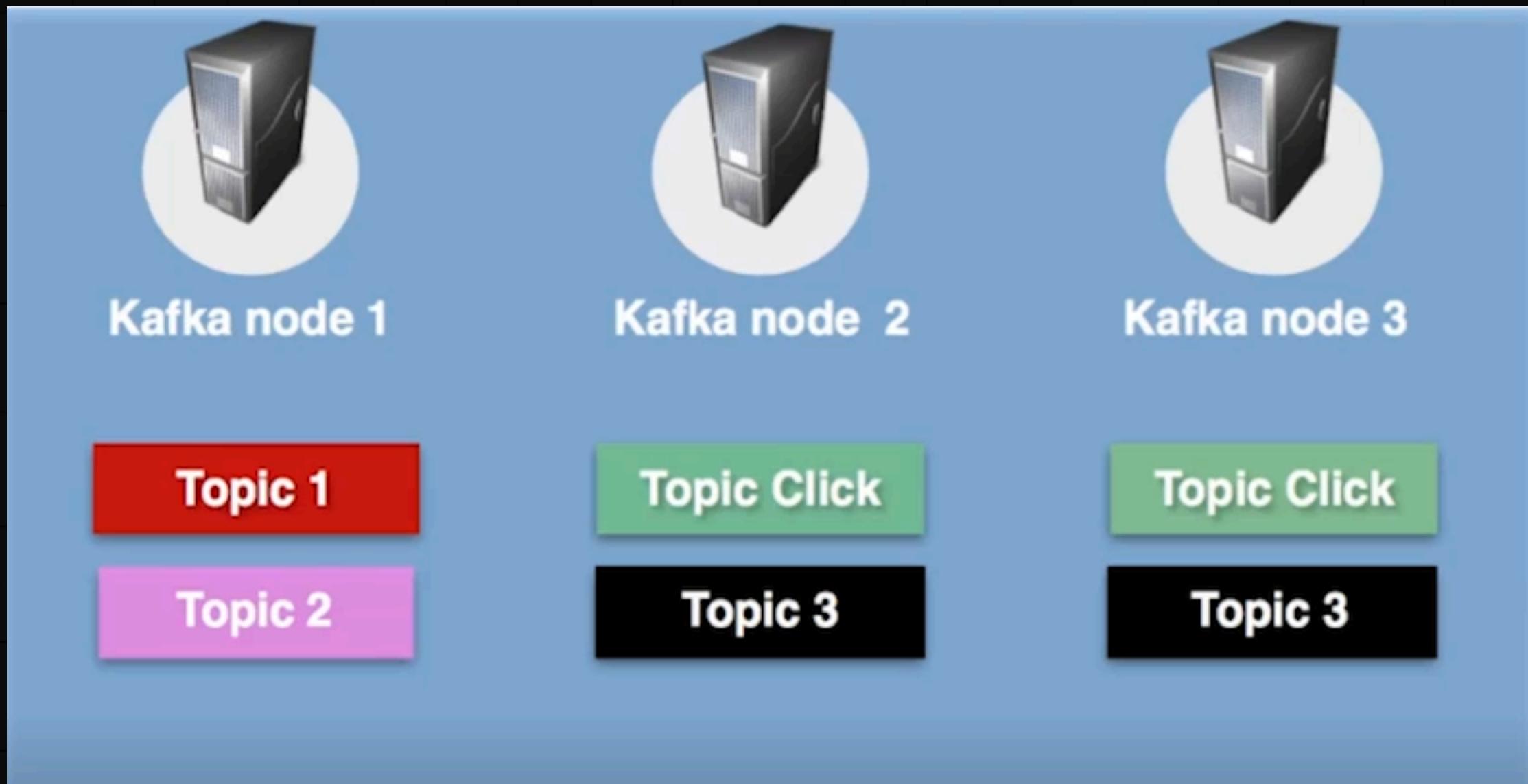
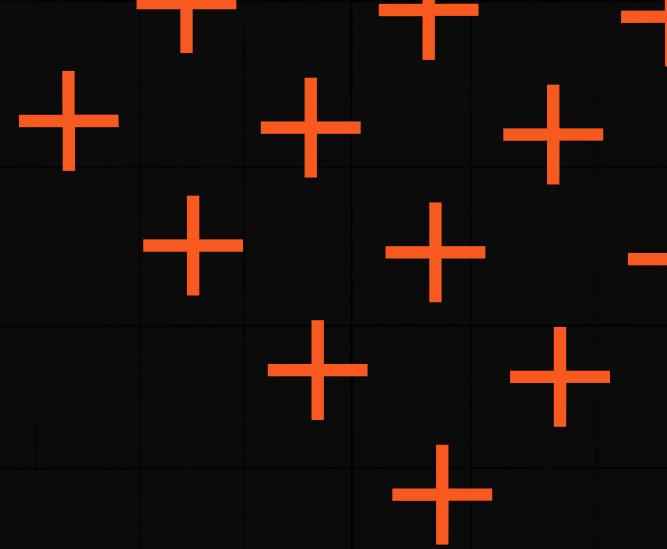
- redundant unit of topic is called a replica.

5.

- each partition has one or more replicas.



66

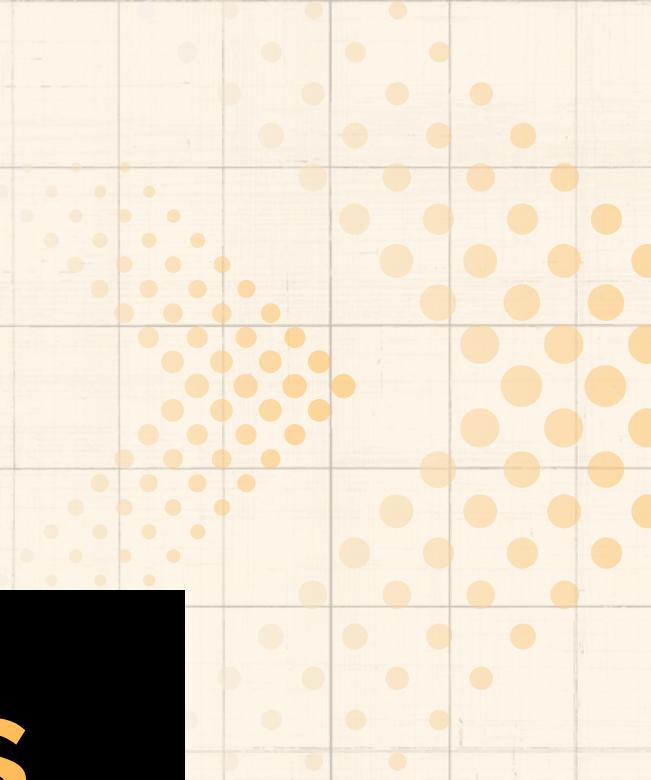


source: <https://kafka.apache.org>

@Abhijith



66



Up Next:

Kafka API & Clusters

**Follow for more and
share this post with
others who might find
it helpful!**

#HappyLearning #BigDataSeries #ML