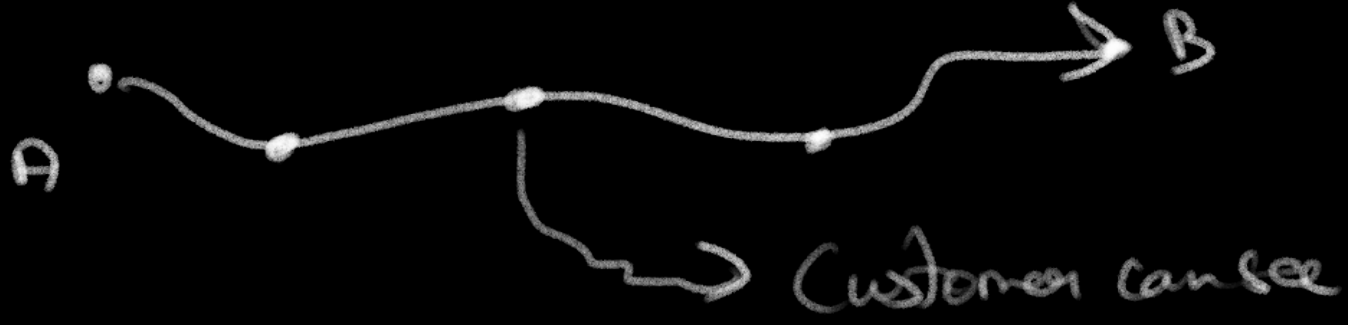
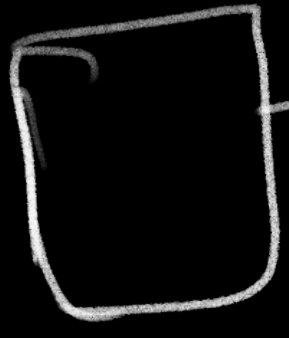
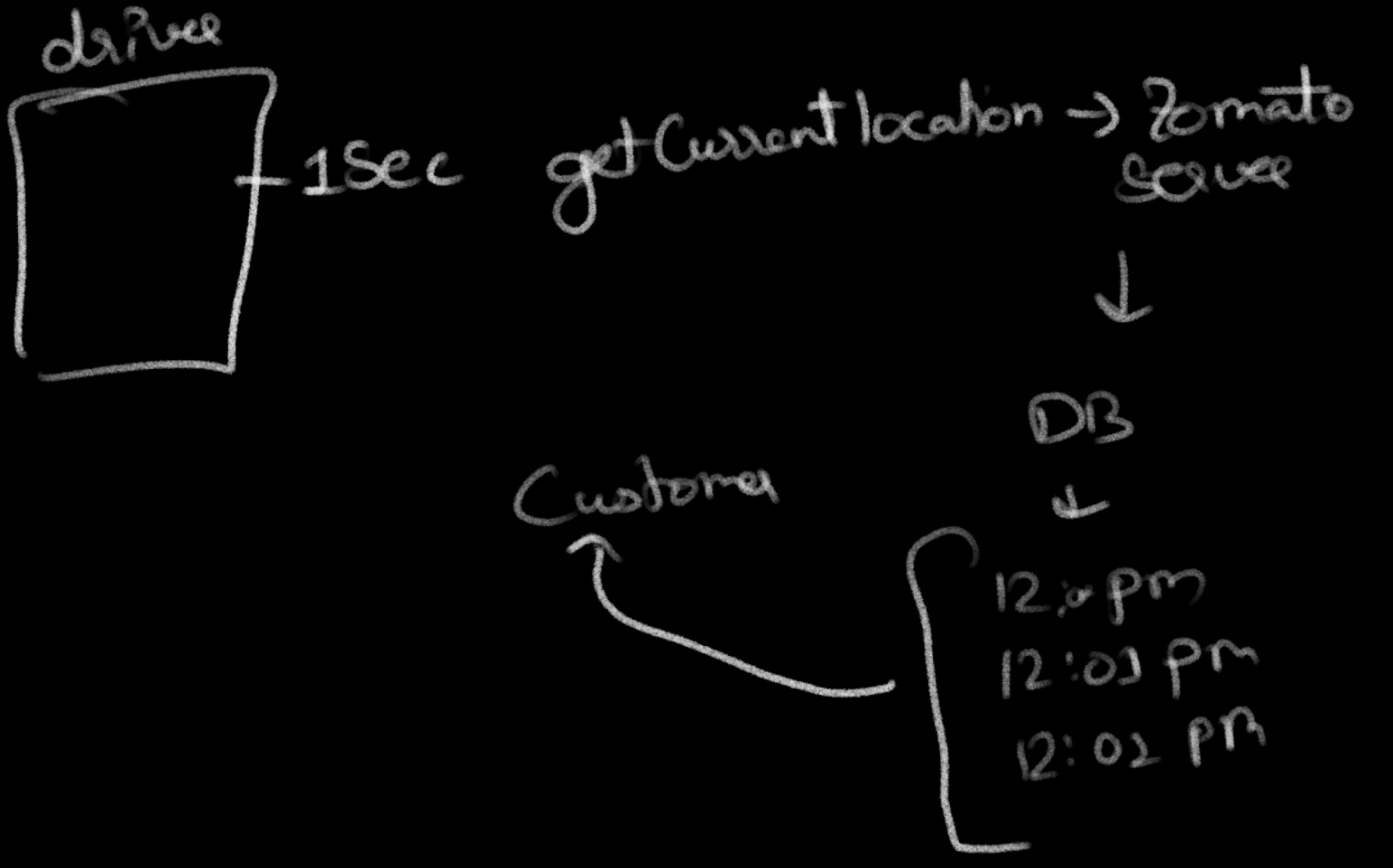
Apache KAFKA

→ Zomato → taking example into the live location of drivers from engineering perspective.



→ If we want design similar one way is to [get] the current location every second & I'll send it to Zomato Seever.



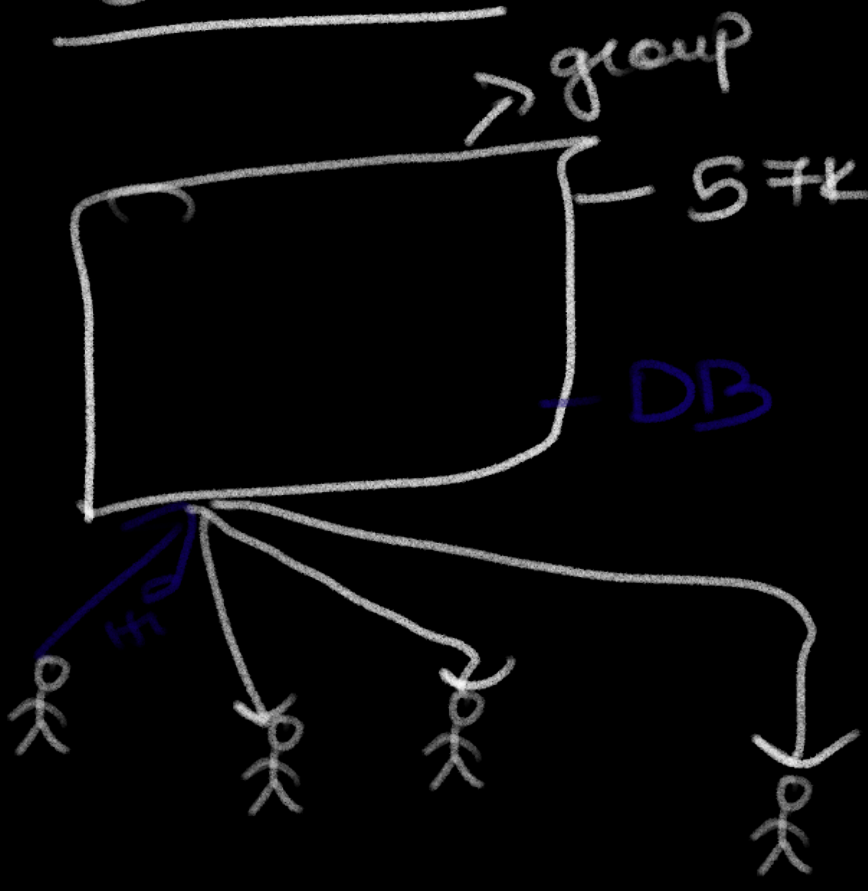


PROBLEM: On a large scale like Zomato if 1000 of drivers are there & from every driver we get location every second then DB will go down.

→ Usually DB's have less Throughput

[Throughput - no of operations per second we can do]

Discord



⇒ If 57K people are chatting operation per second increases a lot which will DB to go down

⇒ And whenever you insert anything in DB it takes sometime, let's 200ms. So, you have to wait for 200ms everytime you want to insert

OLAP Uses

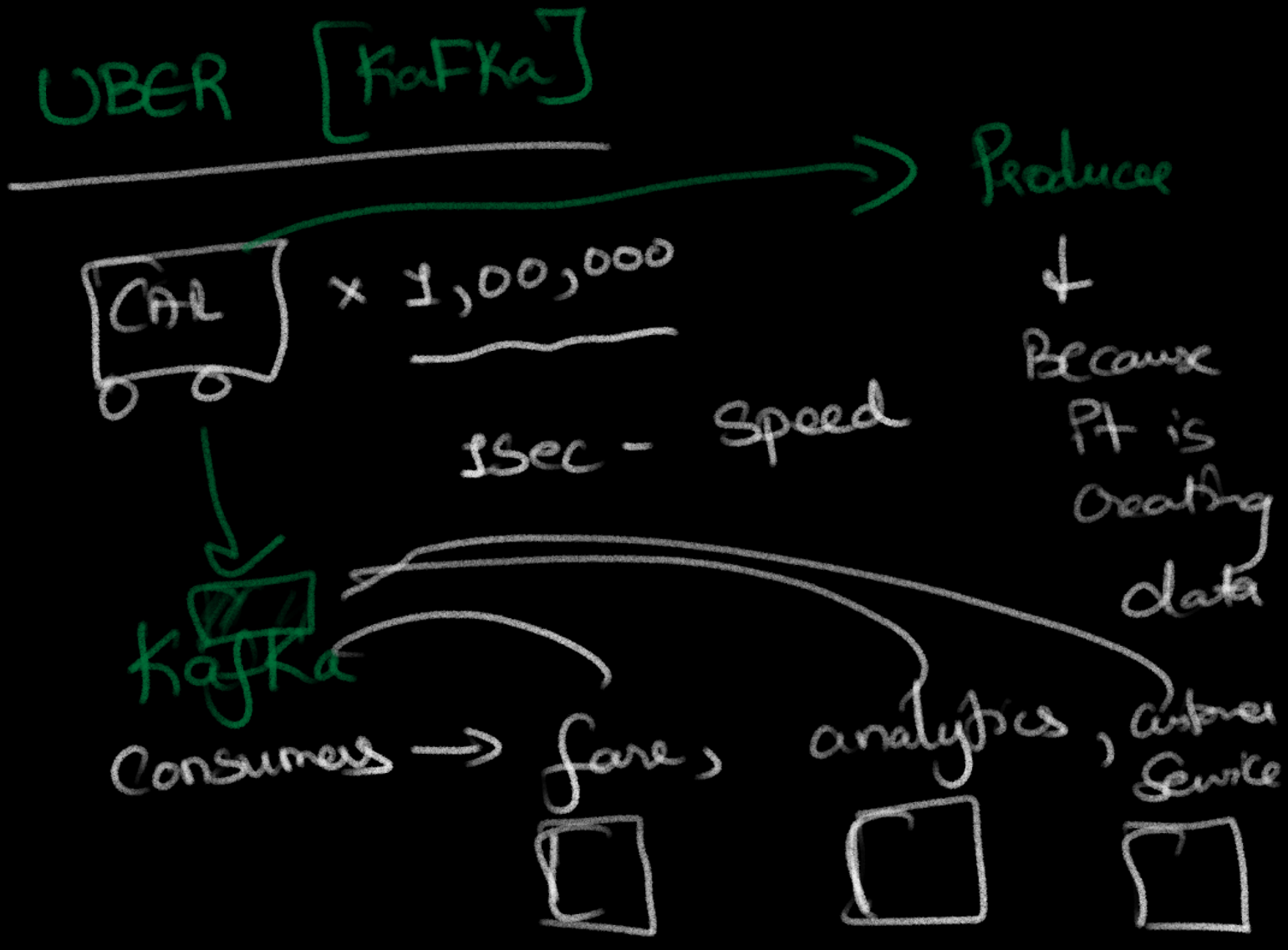
⇒ Fair Calculation
→ driver Analytics - speed, Braking
⇒ Customer service.

→ Kafka has high throughput
→ It's not an alternative to DB.

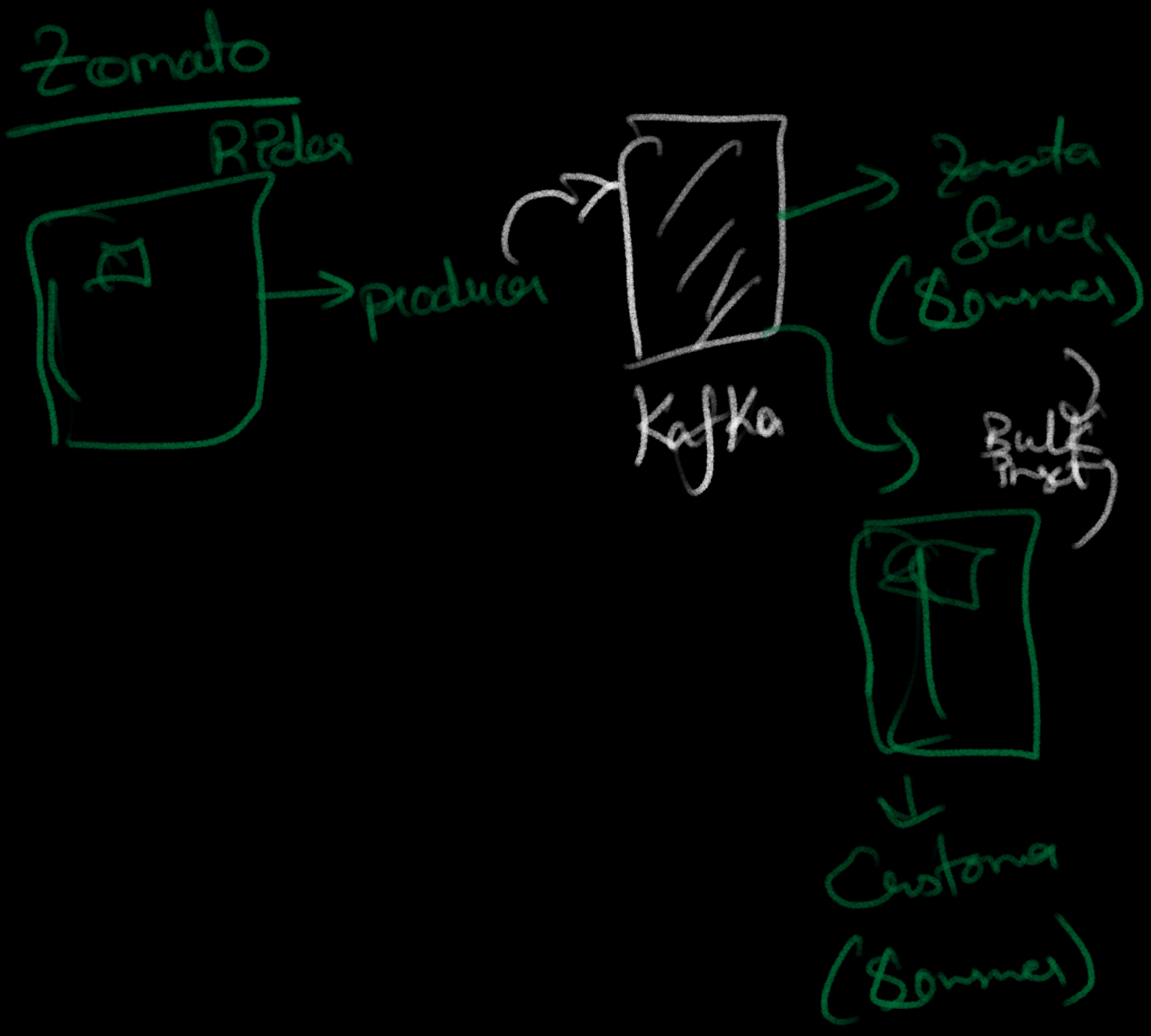
→ Kafka throughput ↑
Storage ↓
↳ very low, Temporary.

→ DB throughput ↓
Storage ↑
↳ Billions can be stored

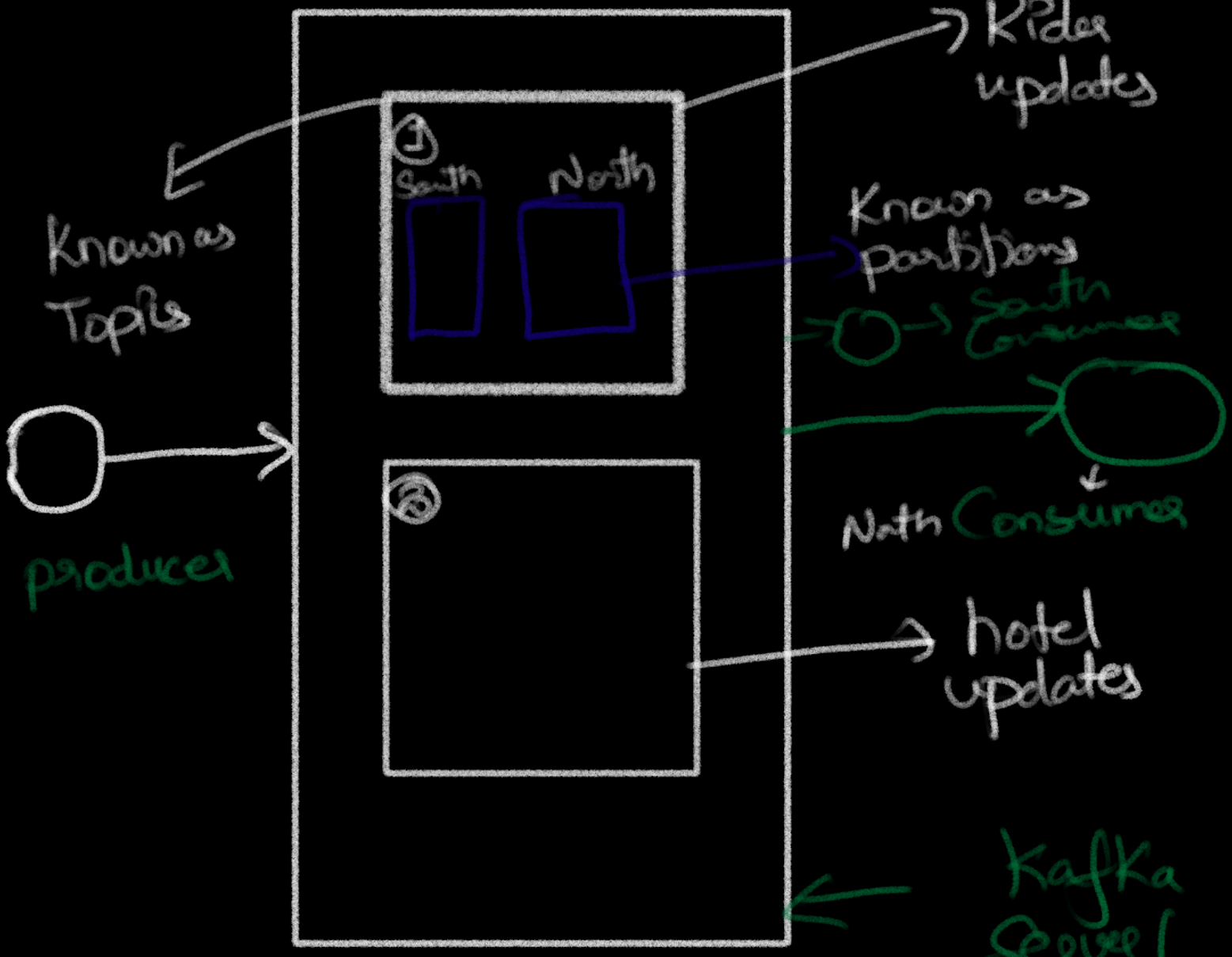
→ So, we will be using them together.



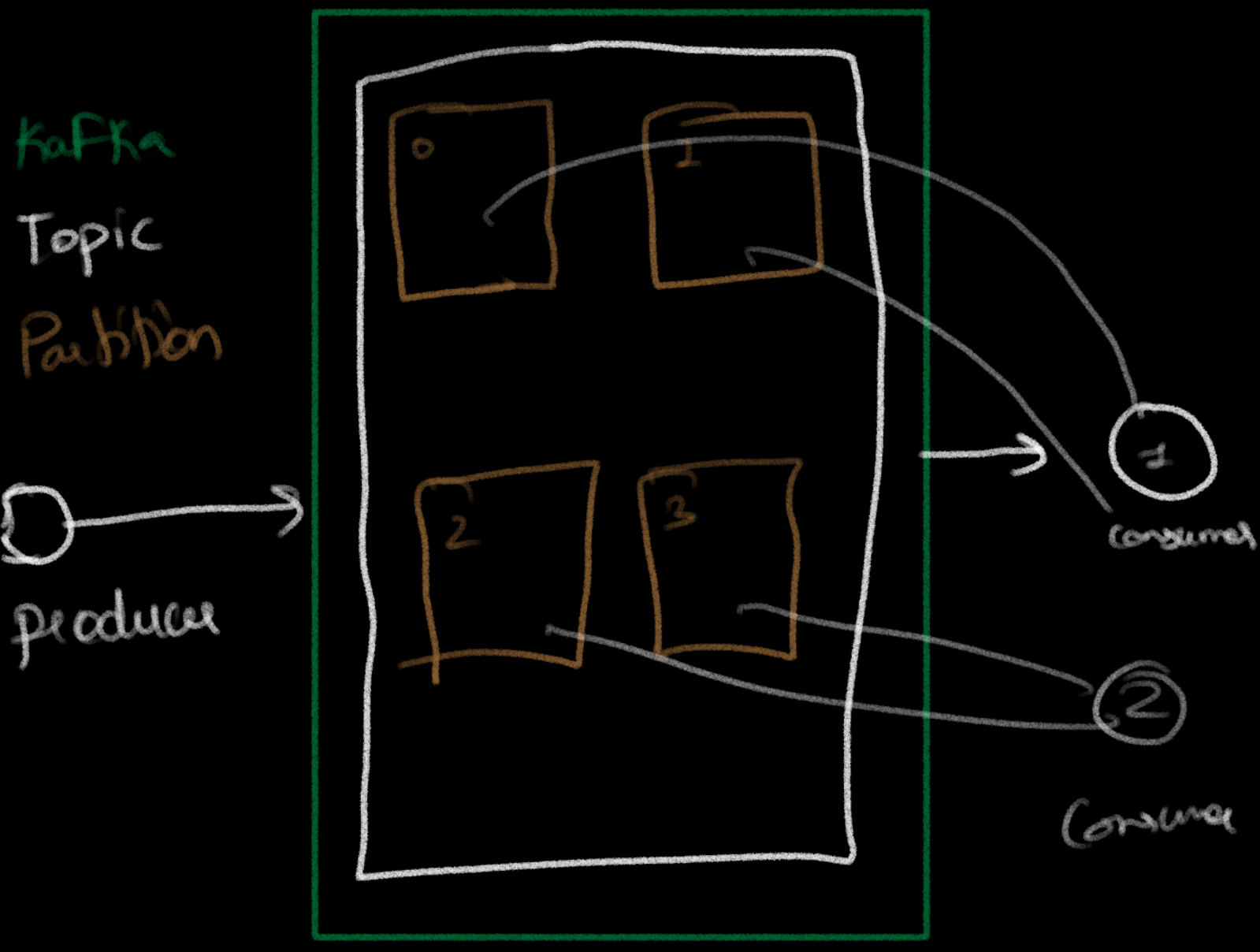
→ It might take 2 seconds


Zomato
Riders
producer
Kafka
Zomato Service (Consumer)
Bulk Insert
Customer (Consumer)

Zomato

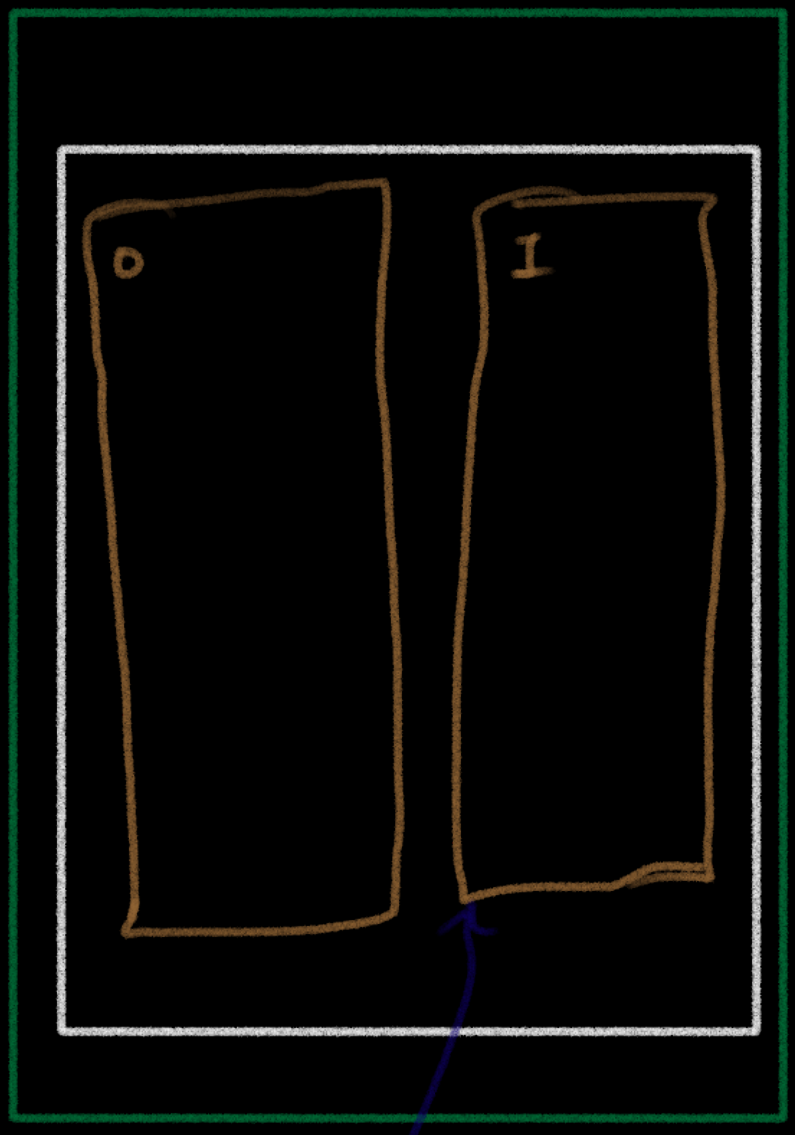


Partitions → Based on locality



→ Kafka has auto Balancing.

Let's from above example, If there are 5 Consumer for 4 partitions, 1 Consumer will not get (or) Consume any message.



only 1 Consumer
(consume Both)

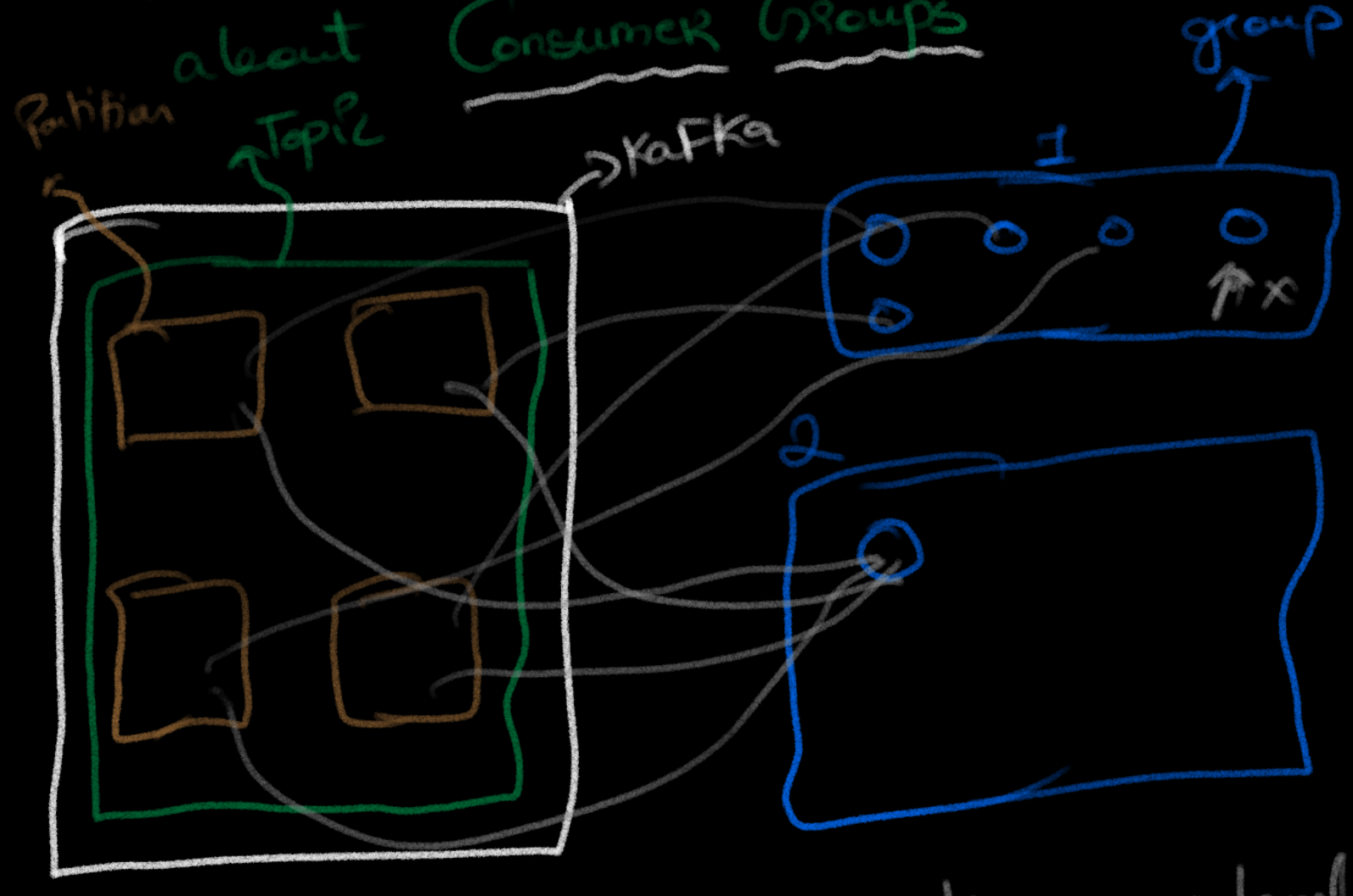
2 Consumers
1 Consumer 1 partition

3 Consumer

1 Consumer → multi° partitions Consume

1 Partition → only 1 Consumer

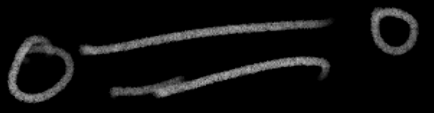
To Solve this issue, let's learn about Consumer Groups



→ Balancing happens at group level

→ It means In 1 group, 1 partition can't have multiple listeners.

→ Queue



1 producer
1 Consumer

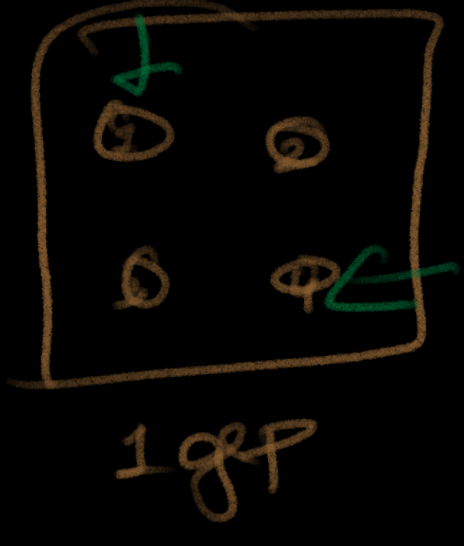
Pub/Sub

1 producer

Multiple Consumers

→ Kafka using Consumer grps can act as both queue & pub/sub.

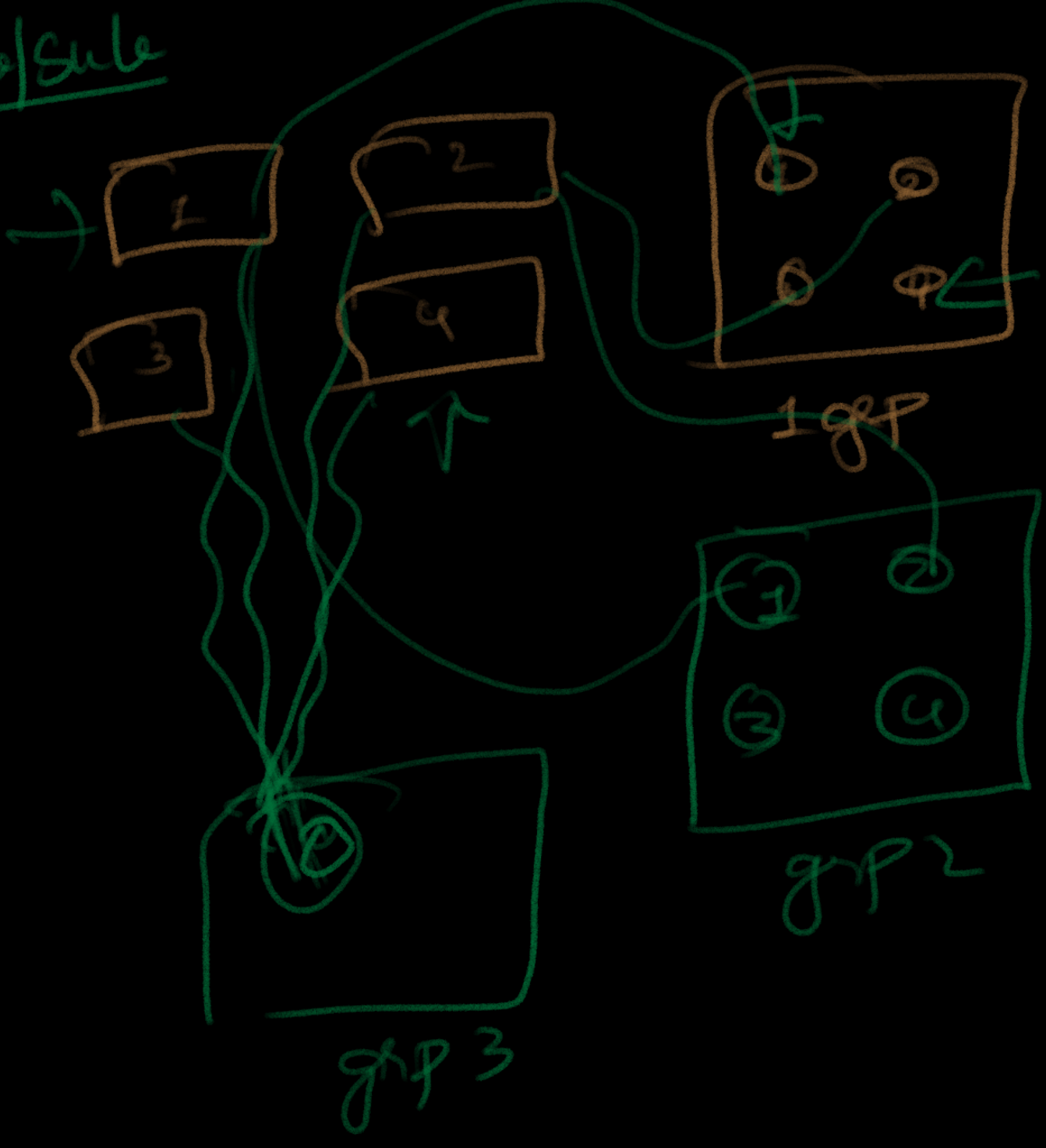
Queue



1 grp

No of Consumers = No of partitions

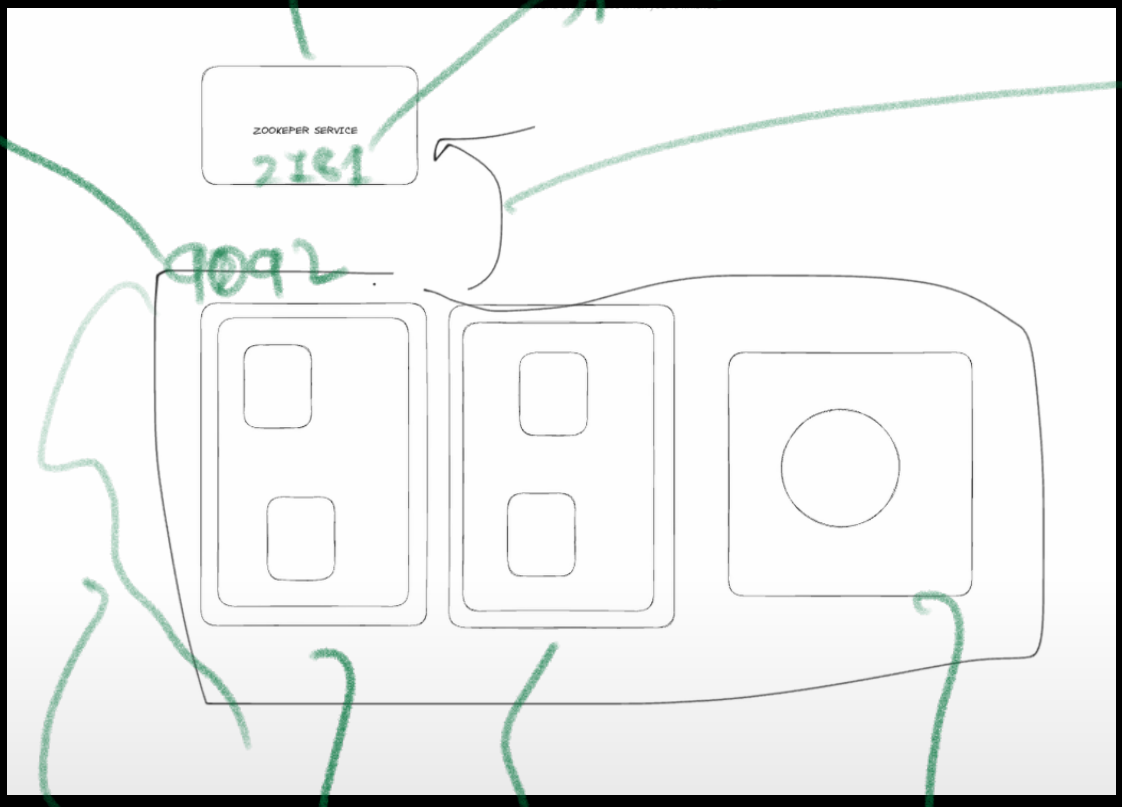
Pub/Sub



→ Kafka uses ZooKeeper Internally
So, you need to run it

Kafka port

ZooKeeper



port

Interacting

Broker
active
service is

Kafka

Consumer

Admin - Infra setup Topic/partitions

Produce - Msg produce

Consumer