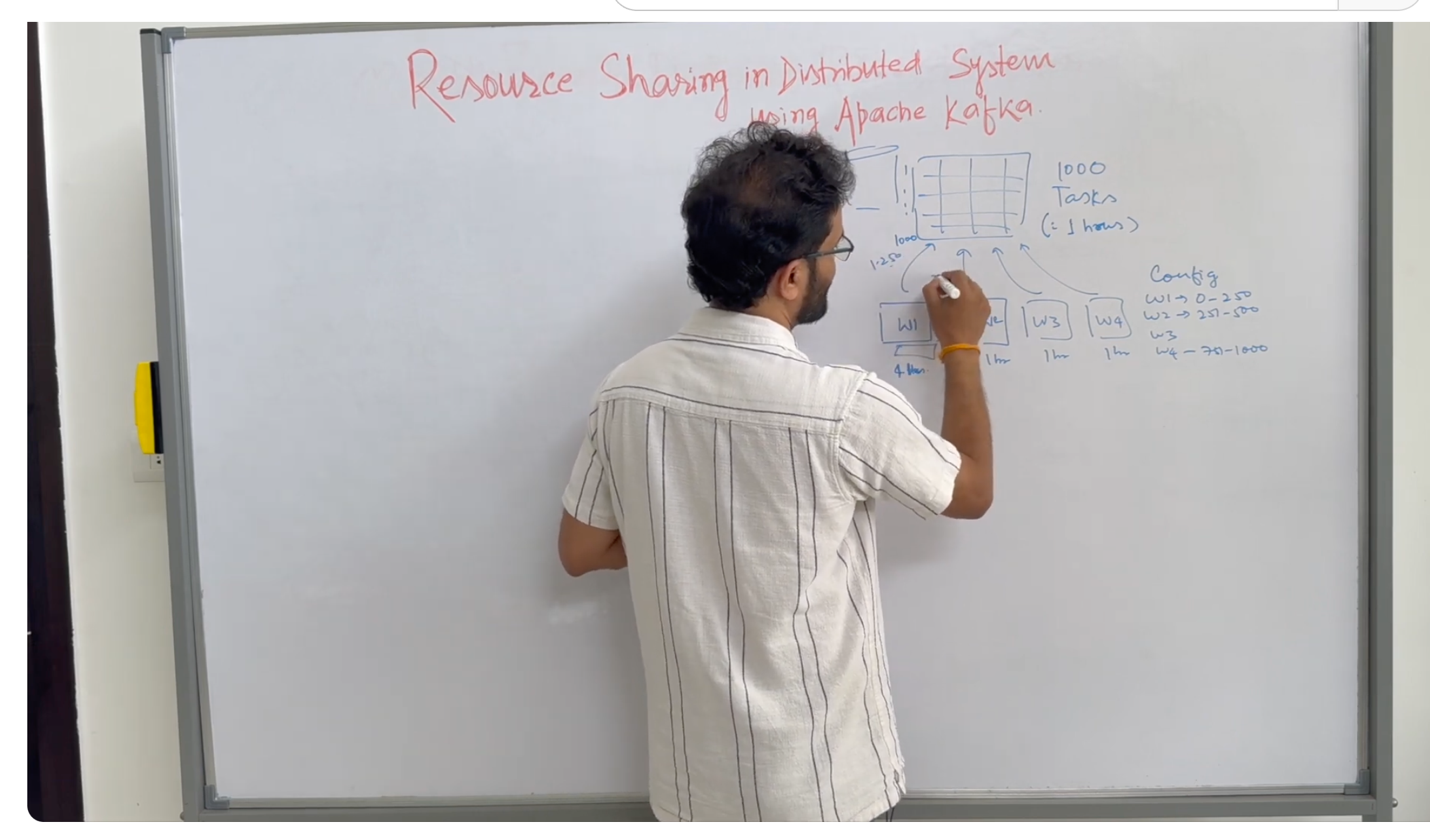
Build the distributed database here



The problem statement here

1. You have a database that has 1000 tasks so (basically 1000 rows of tasks) here,

How to distribute the tasks among the 4 nodes?

You have a config file that goes w1: 0 – 250

But what about if you add more tasks rebalancing? Montoring, these can be taken care of by Kafka.

Using kafka

N producers and then the consumers are not coupled and they are just connected to the kafka in the middle A person standing in front of a whiteboard

Description automatically generated

A new problem

What if multi producers go to 1 topic?

Low thruput will result as a result,

Solution: use partition here

: p1 goes to p0

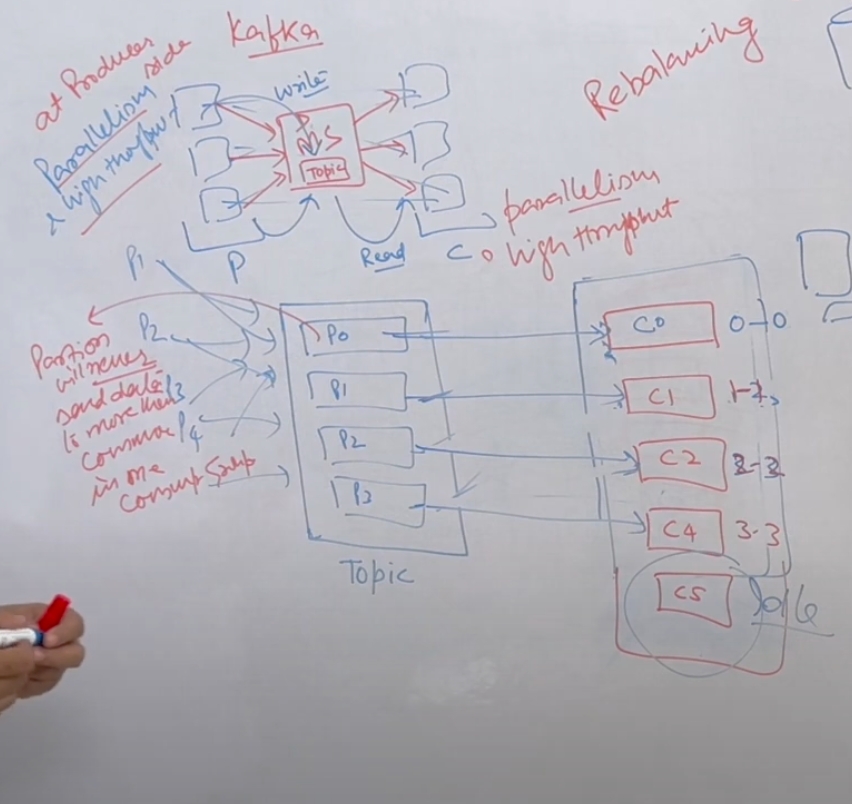
A hand writing on a white board

Description automatically generated

**How to ahceive high thruput on the consumer side?**

Consumer group here

You can have miniworkers group in the consumer node



IF you have more c then p, than it will just sit idly