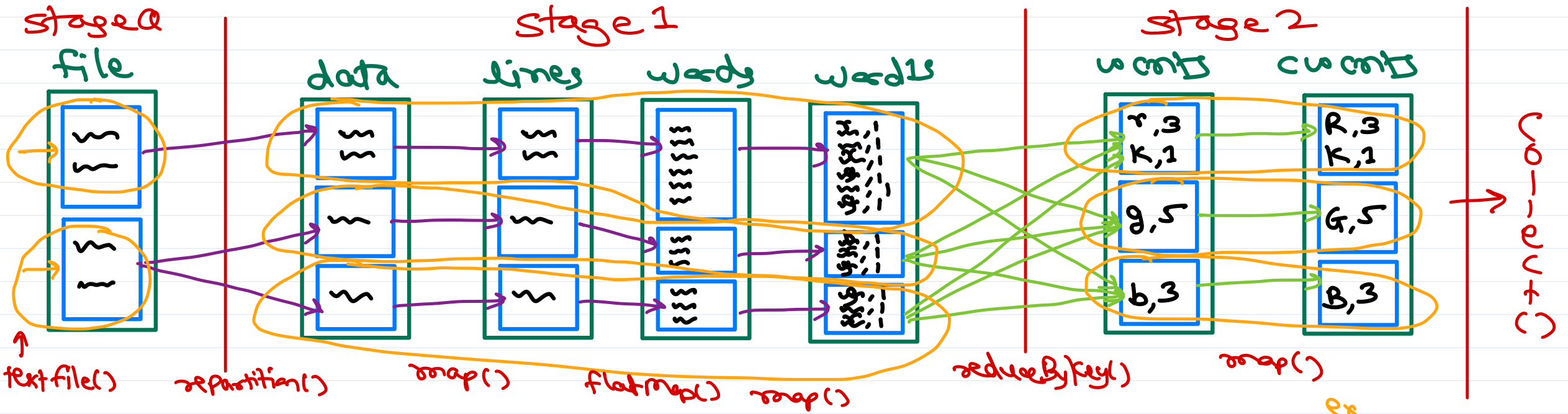




# Big Data Technologies

Trainer: Mr. Nilesh Ghule.





ex  
task 0 → 0  
task 1 → 1

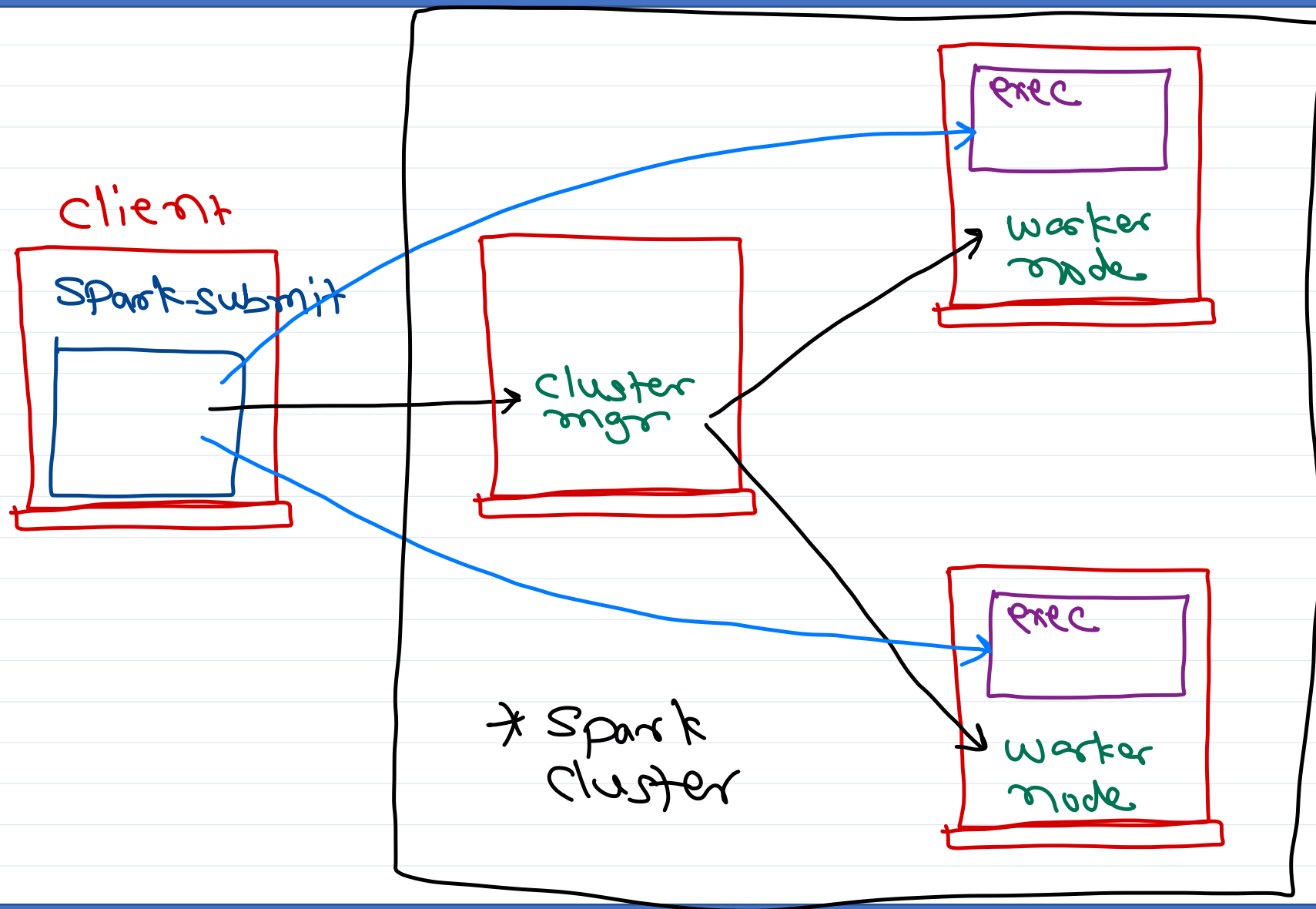
task 2 → 0  
task 3 → 0  
task 4 → 1

task 5 → 0  
task 6 → 1  
task 7 → 1

add action → job → DAG (exec plan)  
job1, job2, job3, ... → application



# Spark Architecture

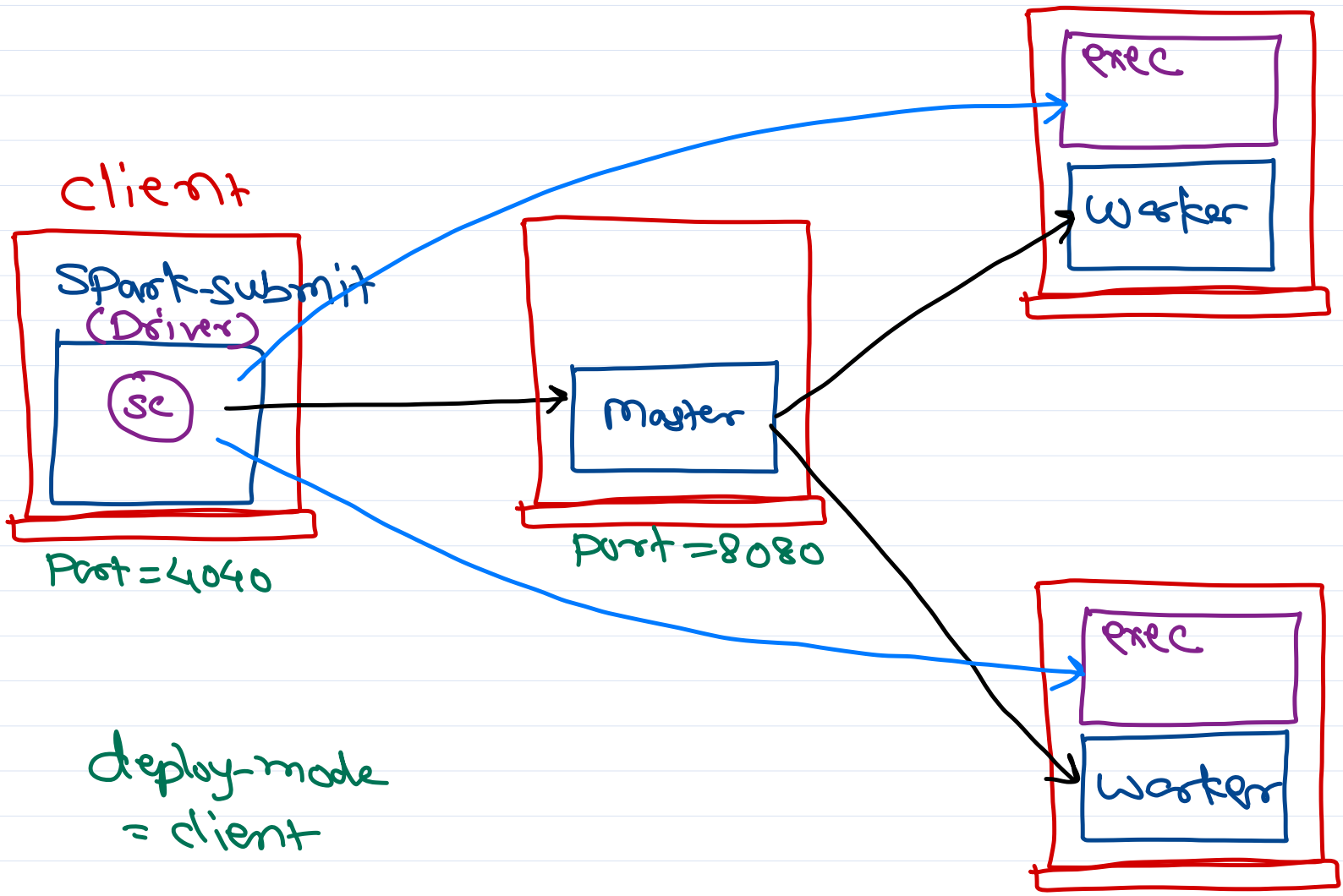


## Spark cluster mgr

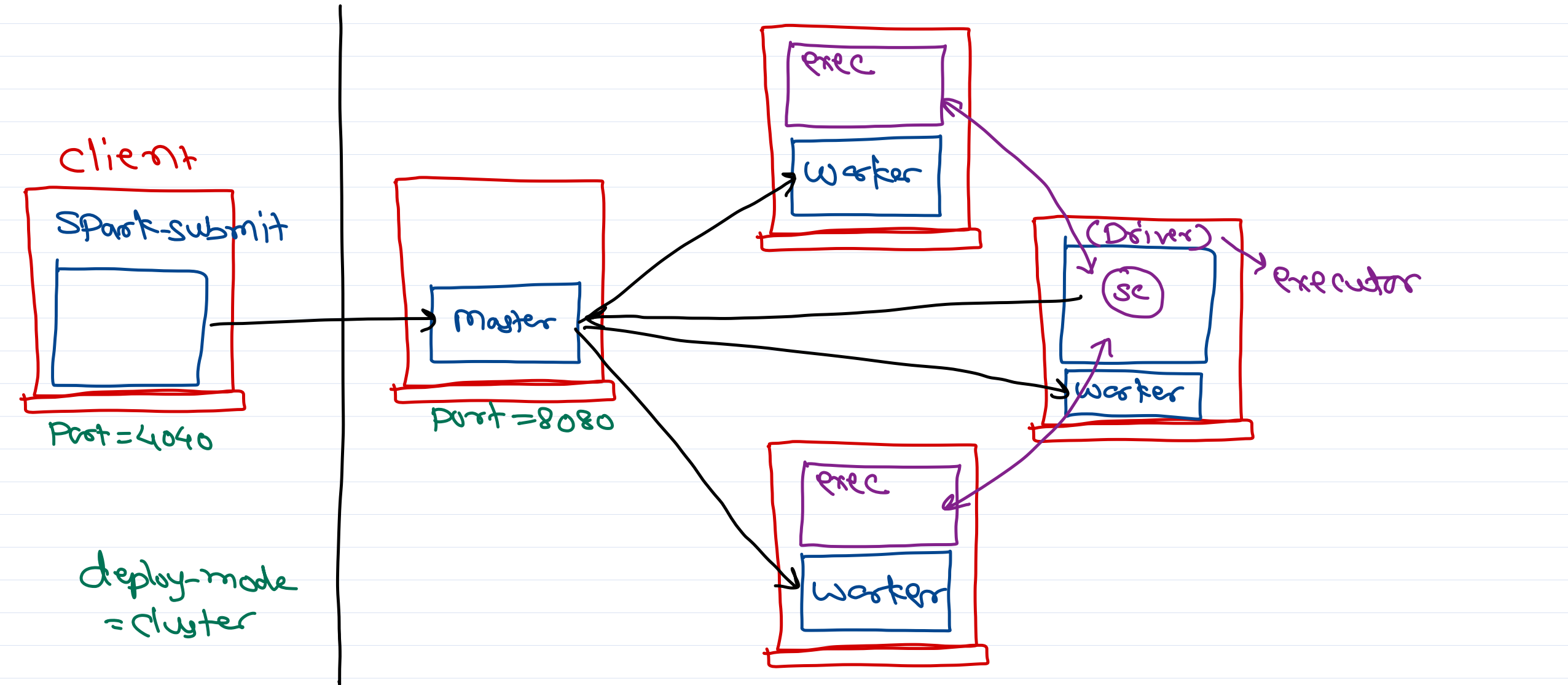
- ① Spark standalone  
↳ master, worker(s) (7077)
- ② YARN  
↳ resource manager, node manager
- ③ Mesos
- ④ Kubernetes (k8s)
- ⑤ Spark Connect



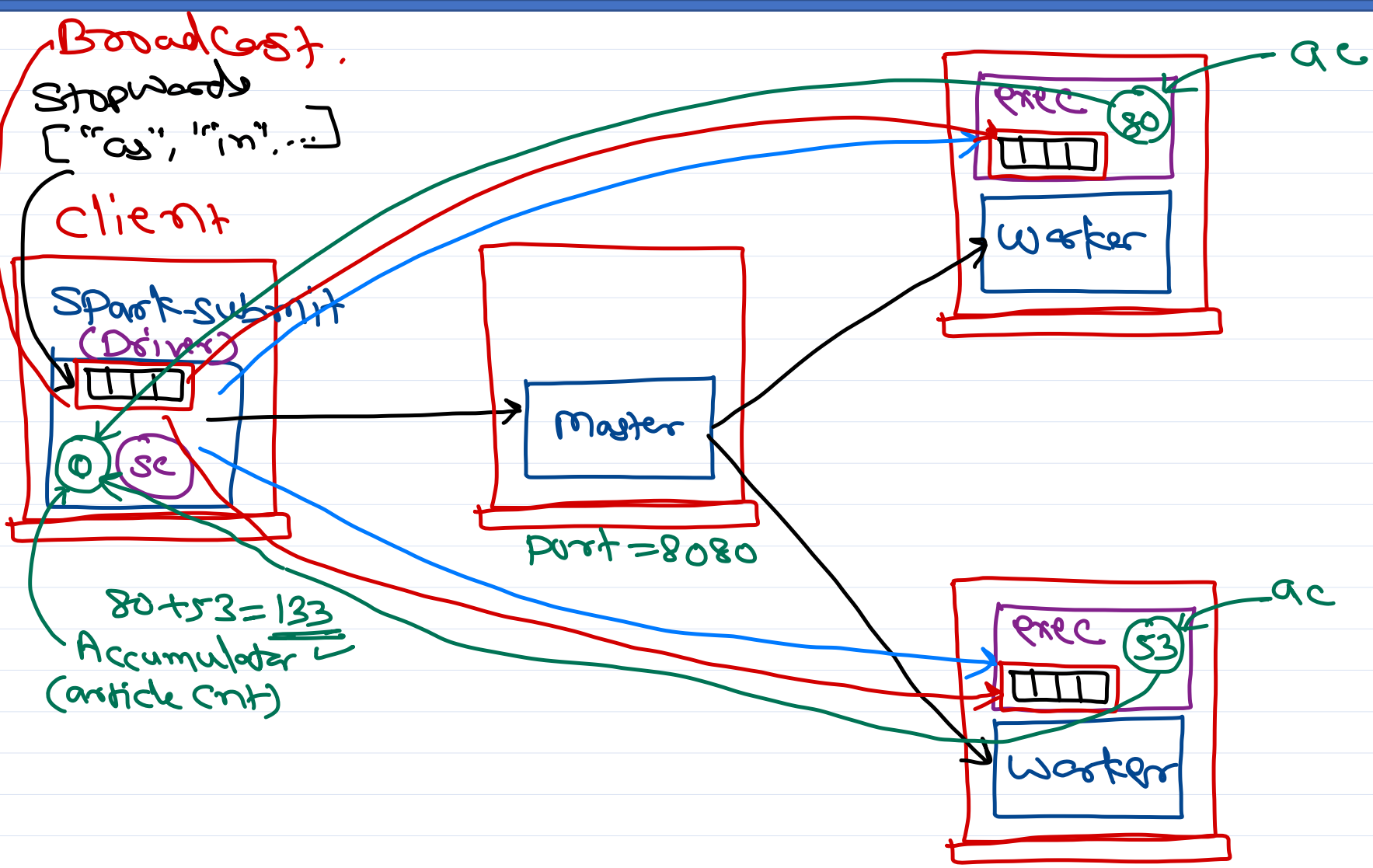
# Spark Standalone Cluster



# Spark Standalone Cluster



# Spark Standalone Cluster





Thank you!

*Nilesh Ghule <nilesh@sunbeaminfo.com>*

