Yang Wang   yangwan2

P1: Red Mean TwoTwo   h("Red") = 0
P2: Mine Mine Mean   h("Mean") = 1
P3: Red TwoTwo Mean   h("TwoTwo") = 2
                             h("Mine") = 3

h: 0   1   2   3

Use HashMap to store the status of the partition, avoid duplication work

Master	Worker 1	Worker 2	Worker 3	Worker 4	Description
	P1, P2	P1, P2	P3	P2, P3	**Initial Contents**
Send Map Tasks → mapTask → mapTask → mapTask → mapTask →					**Start Map** Master tells worker servers to start the map phase, providing the map function.
←← map Result Path	mapResult intermediate result P1 --- Red 1 Mean 1 TwoTwo 1	mapResult intermediate result P2 --- Mine 1 Mine 1 Mean 1	mapResult intermediate result P3 --- Red 1 TwoTwo 1 Mean 1	mapResult Empty intermediate result ---	**Workers Perform Map** Workers execute the map task and store their results in the mapResult file. Then, they inform the master server that the map task has been completed. If a worker server is not used, leave the box blank.
Send Shuffle Tasks → shuttleTask → shuttleTask → shuttleTask →					**Start Shuffle** Master tells worker servers to start the shuffle phase.
Transfer keyPairs among workers Pair(Mean,1) → Pair(TwoTwo,1) → 2×Pair(Mine,1) ← Pair(Red, 0)	shuffleResult Red 1 Red 1	shuffleResult Mean 1 Mean 1	shuffleResult TwoTwo 1 TwoTwo 1	shuffleResult Mine 1 Mine 1	**Workers Perform Shuffle** Workers send the map results to each other, based on the hashcode of the key. Workers store these results in another file. Then, they inform the master server when they are done. If a worker server is not used, leave the box blank.
←← shuttle done shuttle done shuttle done					
Send Reduce Tasks → ReduceTask → reduceTask → reduceTask → reduceTask →					**Start Reduce** Master tells worker servers to start the reduce phase, providing the reduce function.
← final Path1 ← final Path2 ← final Path 3 ← final Path 4	finalResult Red 2	finalResult Mean 2	finalResult TwoTwo 2	finalResult Mine 2	**Workers Perform Reduce** Workers execute the reduce task, creating a new file with the final results of the map-reduce task. Then, they inform the master server they are done, providing the file location of the result. In this case, the file location is always the same for simplicity.