Mapreduce: Summary

- Mapreduce uses parallelization + aggregation to schedule applications across clusters

- Need to deal with failure

- Plenty of ongoing research work in scheduling and fault-tolerance for Mapreduce and Hadoop