Haddop – open source software platform for distributed storage and distributed processing of very large data sets on computer clusters built from commodity hardware.

Distributed Storage – Big Data, Terabytes of info. Adding more and more storage spaces as one single file system. It is redundant.

Distributed Processing – can distribute the data. Do it all in the Parallel Manner. Get through all the data very quickly. Parallel Computing.

Large Dataset

Commodity hardware – AWS, throw into Hadoop cluster and go forward with it.

GFS and Map Reduce from Google