The two big advantages of MapReduce are

1. Parallel processing
2. Data Locality

**Parallel Processing:**

In the process of MapReduce, we divide the job into number of sub multiple jobs which has their own activity to carry out. Map reduce helps in processing the data. The time taken to complete multiple jobs is very less as the data is divided into number of multiple jobs of each size 128mb and processed by number of machines.

We have master and slaves in this process, when the master is fully loaded, the job is assigned to the slaves.

**Data Locality:**

The process of moving large of chunks of data to the processing unit is called Data Locality, Traditionally the data is moved to the data processing unit but now the data is moved to the data processing unit as the income of the data is large and can’t be handled all at once.

* So is the Data Locality being one of the applications of the map reduce helps in reducing the cost as the data is moved to the data processing unit.
* The master node may fail when it’s over loaded.
* Processing time takes more as the data is processed by only single unit.

In order to overcome the above problems Map reduces helps in avoiding the above challenges as the data processing units are moved to the data.

**Problem:** For Example, we have 100,000 source files and each file has to be processed which has no similar relation among them, if processing one file takes one min then it would take so many days to process the data, **so how do you process it?**

**Here comes the methodology of the Parallel Processing**

The Data is divided into number of small chunks of jobs of each 128 mb of similar data, So it will be easy to process the similar data and the data is processed by number machines, The data processing machines are brought to the data which is cost effective.