**Map Reduce Paradigm**

You've probably heard that MapReduce, the programming model for processing large data sets with a parallel and distributed algorithm on a cluster, the cornerstone of the Big Data eclosion, was invented by Google.

One of Google’s first challenges was to figure out how to index the exploding volume of content on the web.

MapReduce is a programming model and an associated implementation for processing and generating large data sets.

Users specify a map function that processes a key/value pair to generate a set of intermediate key/value pairs, and a reduce function that merges all intermediate values associated with the same intermediate key. Programs written in this functional style are automatically parallelized and executed on a large cluster of commodity machines.

MapReduce programming framework is used to perform distributed and parallel processing with large data sets in a distributed environment.