**SUMMARY:**

Jeffrey Dean and Sanjay Ghemawat.

**MapReduce: Simplified Data Processing on Large Clusters**

**DATE: 3 January 2018**

This paper presents MapReduce, a simplified and abstract programming model for processing large clusters of data in a distributed computing system. MapReduce contains two functions; a map function that processes each key/value pair and a reduce function that processes

MapReduce is a library that helps people write distributed programs that operate on large clusters of data. The library spreads the computations across a lot of machines, run all the computations, deal with failures, collect all the results and then save them away.