Edward Shen

E-Portfolio Assignment

MapReduce is a programing model that is capable in processing large data sets with a parallel distributed algorithm on a cluster. MapReduce is program with a map that is capable in performing filters that sorts and reduces procedure that produces summary operations. MapReduce is a term that has two separate and distinct tasks. It’s first task is the map job, in which takes many sets of data and converts it into a different set of data, where the individual elements are then broken down into key/values pair (tuples). The reduce job is then able to take the output from the map as input and is able to combines data tuples into a smaller set of tuples. This is the sequence that is imply when using MapReduce, where the reduce job is always performed after the map job.

This program model is important because, by being able to “MapReduce” data sets we are able to precisely get the data and analyze the data without needing to go through the entire raw data and compressing it into analytical data. This topic is related to MIS2502 because; it essentially takes raw data and processes it into an analytical data set. MapReduce is able to take large data sets and creates a smaller data set and performs filters to sort and reduce data that produces a data summary for the user.

There are many cases where MapReduce is used in practice; this is because MapReduce can be used in a variety of ways. For Example MapReduce is used: in Fraud Detection for online gaming company to catch cases of fraud, Graph Analysis for social media company where it uses the SQL-MapReduce Functions, Sharing Behavior where in ShareThis uses MapReduce to reduce query times as it analyzes the items that people share online, and many more.

Also MapReduce brings efficiency in processing data within your organization and driving down costs significantly over conventional data processing technologies.

Work Cited

"How Does MapReduce Fit into My Organization and Advanced Analytics?"*MapReduce and Hadoop*

*Ecosystem*. MapReduce, n.d. Web. 28 Apr. 2014. <http://www.mapreduce.org/mapreduce-fit-

into-org.php>.

"What Is MapReduce?" *IBM*. IBM, n.d. Web. 28 Apr. 2014. <http://www-

01.ibm.com/software/data/infosphere/hadoop/mapreduce/>.