

HOMEWORK 04

UMANG MUKESH MEHTA

FRIDAY 1:35 – 3:15 PM

Performance Comparison and Analysis

A) SECONDARY

Configuration 1: 155 seconds

Configuration 2: 117 seconds

Answer to Question 2: (Based on just the SECONDARY program)

To achieve good Load Balance in the SECONDARY Program I used Key Comparator and Group Comparator to compare values and perform the task of Secondary Sorting and later on used a partition based on the Unique key of Airline ID so my program was Balanced.

All the keys and values are sorted initially so the reducer just needs to write the sorted values the Map and Reducer have to perform the basic operations all the other Sorting and Partitioning is done by the other classes.

There are no files generated in the stderr for the SECONDARY program hence enclosing both the configuration syslog files.