Analysis Report
Database Systems Assignment 2
Two Phase Merge Sort

Jayitha .C
20171401

September 23, 2018

Abstract

Objective of the assignment is to implement Two Phase Merge Sort and observe it's performance by varying the input file size and the main memory allowed

1 System Configuration

- cpu MHz : 1271.424

- cache size : 3072 KB

- Number of Processors : 4

2 Varying Input Size

Observations made by varying the Input File Size and fixing the main memory size to be 100MB

Table 1: Varying Input Size

Input File Size	Time Taken (in secs)
1MB	17.8684
2MB	77.418
3MB	162.881
4MB	294.665
5MB	448.481
50MB	44053.1
500MB	4220368.1
1GB	17657771.7
2GB	-
3GB	-

From these observations we can see that the program gets a lot slower as the input file size increases as the number if I/O reads increases. I/O reads are

dominant as the file size is increased as the block size decreases i.e. block size given to each sublist now decreases.

The time values are reported to be high because of the redundant parsing and usage of insertion sort in the code.

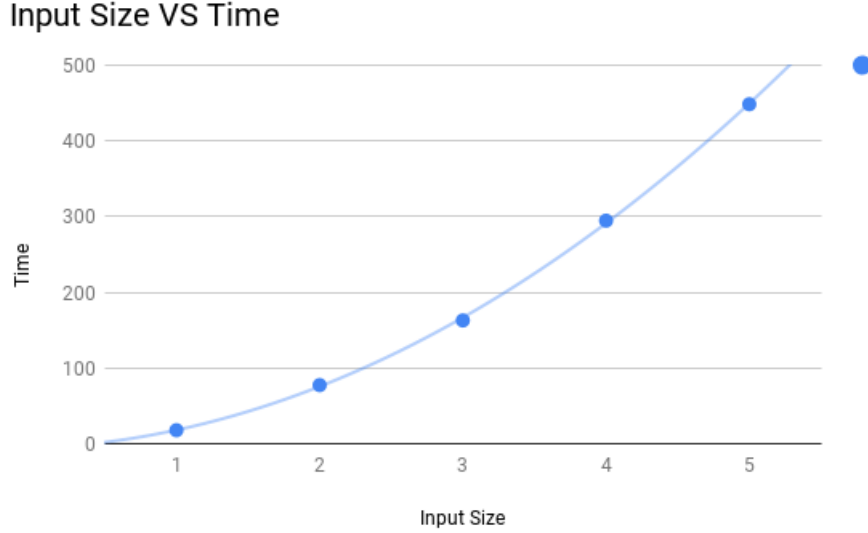Thus as the input size increases the time taken to sort increases polynomially.

Input Size VS Time

Figure 1: The time taken increases polynomially w.r.t. the input size

3 Varying Memory Size

Observations made by varying the Memory Size and fixing the Input size to be 5MB

Table 2: Varying Memory Size

Memory Size	Time Taken (in secs)
1MB	76.42
2MB	168.16
3MB	300.5
4GB	472.8
5MB	685.3

These values don't characterize two phase merge sort. The sorting technique implemented is Insertion sort which is a lot slower than Merge sort and as the main memory size increases more records get sorted using insertion sort than merge sort and hence the time taken increases

If the sorting technique implemented was maybe merge sort or quick sort then the graph shown in Fig 3 would be obtained. The time taken should decrease as the number of I/O operations would decrease
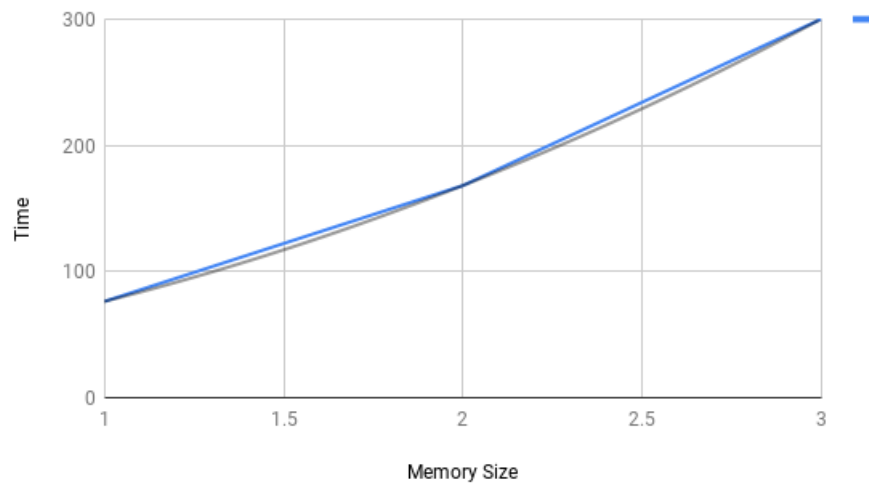
Time VS Memory Size

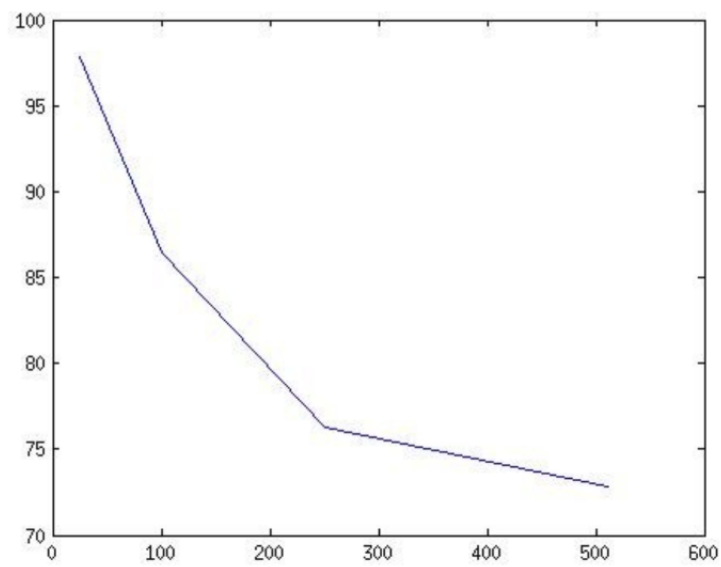Figure 2: The time taken increases polynomially w.r.t. the memory size

Figure 3: The time taken decreases as memory size increases and number of I/Os decrease