Assignment 2 , 159.735, 2020 S2

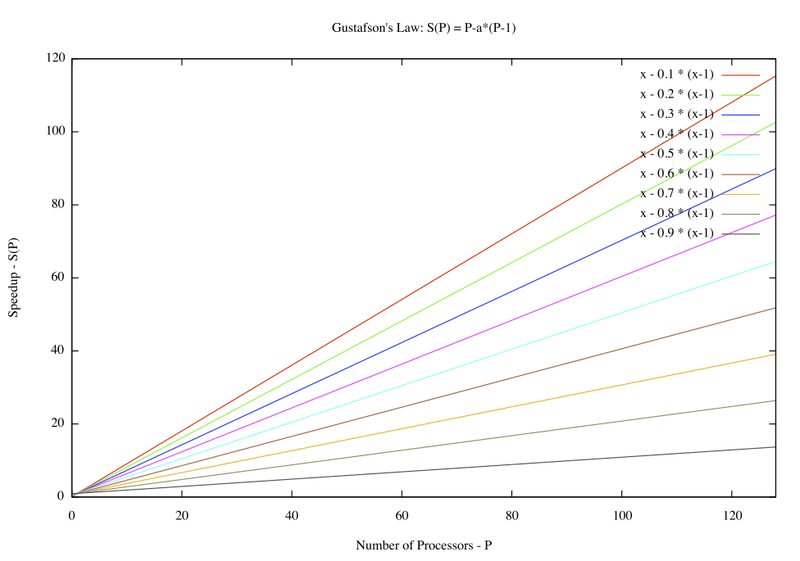
Family Name: JIANG , Given Name: TAO ,

Student ID:20004769

**Assignment2 :Parallel Bucket Sort**

The experiment uses multiple processors to sort random numbers in parallel to verify Gustafson's law.

Gustafson's law means that as the proportion of program parallelization increases, the speedup is proportional to the number of processors. It is different from Amdahl's Law. Amdahl's law emphasizes that when the serialization ratio is constant, the speedup has an upper limit. No matter how many processors are used in the calculation, the upper limit cannot be exceeded. However, Gustafson's law in an ideal environment, if the serialization ratio is small enough, and the number of processors is continuously accumulated, faster speeds can be obtained.



* **User’s Guide**

I use macOS as a development platform. And it only supports 4 processors maximum.

macOS env：

```shell  
mmpic++ ass\_2.cpp - out  
mpirun -n 4 out   
```

At the same time, I deployed the program to Mighty for testing. Get the final experimental results.

Mighty env：

Login mighty as user 20004769

/home/s20004769/ass2/

```shell  
mpic++ ass\_2.cpp -o sort  
qsub sort.pbs  
  
```

* **Experiment results**

The experiment uses a multi-group, controlled experiment. The random number range is fixed from 1 to 99999. Each processor will receive the same amount of random numbers (1000000). The experiment implements a small bucket/big bucket method. It records the efficiency of processing correspondingly more numbers when comparing more processors with the same calculation amount for each processor. Each group of experiments recorded the total sequence time, parallel time, and serial time. The results are as follows:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Total proc** | **numbers** | **Sort total** | **Serial** | **Parallel partition** |
| 1 | 1000000 | 0.380965 | 1.53E-02 | 0.365714 |
| 2 | 2000000 | 0.40783 | 3.88E-02 | 0.369079 |
| 4 | 4000000 | 0.470713 | 0.0881071 | 0.382606 |
| 8 | 8000000 | 0.688769 | 0.238497 | 0.450272 |
| 12 | 12000000 | 1.03565 | 0.623986 | 0.411668 |
| 16 | 16000000 | 1.40279 | 0.9814 | 0.421386 |
| 20 | 20000000 | 1.80809 | 1.37371 | 0.434381 |
| 24 | 24000000 | 2.20099 | 1.75284 | 0.448153 |
| 28 | 28000000 | 2.60113 | 2.13666 | 0.464474 |
| 32 | 32000000 | 3.30611 | 2.82997 | 0.476138 |

Table 1: multi-processors results

Chart 1: Multiple processors time cost

* **Conclusion**

It can be seen from the above experimental results that when more processors are used to sort more numbers, the parallel part's calculation time is not much different. It can prove Gustafson's law from another angle: the more processors, the higher the operating efficiency. It can be seen from the above experimental chart that this experiment basically conforms to Gustafson's law.

In the experiment, I tried to allocate a processor to each bucket. If most of the numbers are too concentrated in a particular storage bucket, more calculation pressure will be put on the processor where the storage bucket is located. It takes a long time for one or several processors to calculate, which reduces the overall calculation efficiency. It is the reason that using allocate a processor to each bucket is not an appropriate solution.