

First tryouts with MPI-IO

Source code for this homework can be found in the archive directory, along with a *README* file.

Introduction

The goal of this project is to familiarize with MPI-IO in the first part and to stress the used file system then.

Implementation

No specific difficulty has been met during the implementation of the two parts. We just had to choose the right type to store the target data (`int` for the rank and `char` for single bytes).

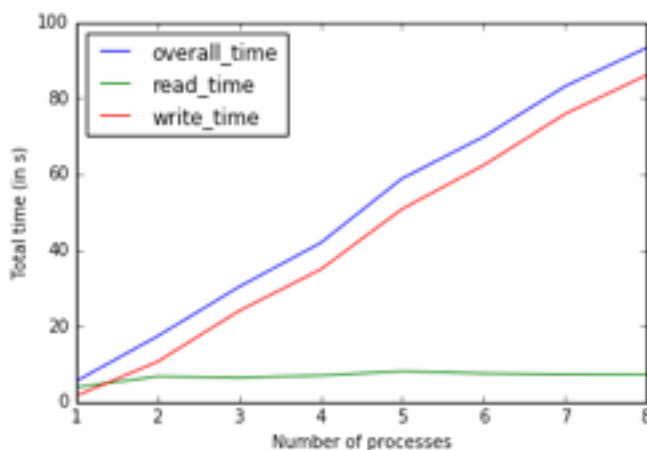
Then the real difficulty appeared to be in testing this benchmark.

OrangeFS

As jarvis cluster has been very busy all along the last week, I preferred to start testing on my own computer (under OS-X). However, when came the time of testing, I had to use OrangeFS that I was not able to build on my computer. And at this time jarvis was totally down (no space on the drives).

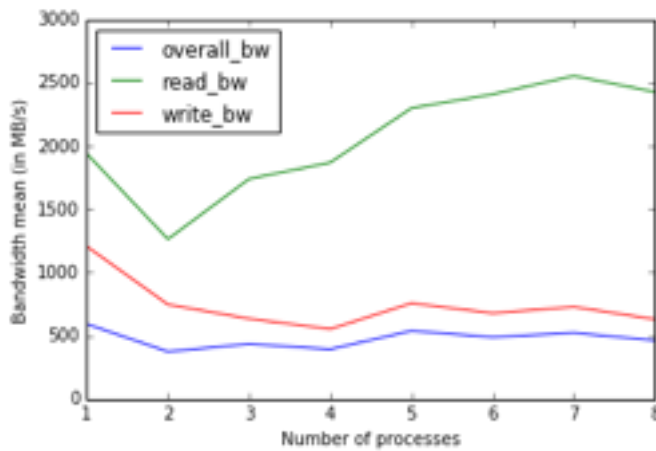
Thus, I was only able to test on my computer (1 node of 8 cores). In order to get a better rendering of these test and because my PC cannot offer the stability of the cluster, I launched my tests several times to gather more data.

The first result of the test is about the difference of speed between reading and writing. In order to analyze it, I summed, for each number of processes, the different times, as shown in the next figure.



We can see not only it takes more time to write data than to read it, but the required time to write increases linearly depending on the number of processes.

The other important measure is the maximum bandwidth of the file system. In order to compute it, I took the mean, for each number of processes, of the computed bandwidth, as shown in the next figure.



Here, we see again that the fastest phase is to read the data. Besides, reading bandwidth seems to increase with the number of processes, when writing bandwidth doesn't seem to be affected by that.

Conclusion

In conclusion, MPI-IO appeared to be easier than I thought. Unfortunately, I really could not use OrangeFS for the tests because of the overload on jarvis during the last days of the homework.

Besides, I cannot really explain why the parallel IO here does not impact the writing bandwidth. Maybe it would have appeared on bigger amount of data (which was not really possible to test on my computer).