ADIOS2: Performance benefits over MPIIO and HDF5?

Shrey Bhardwaj
EPCC,
University of Edinburgh
shrey.bhardwaj@ed.ac.uk

Paul Bartholomew
EPCC,
University of Edinburgh
p.bartholomew@epcc.ed.ac.uk

Mark Parsons
EPCC,
University of Edinburgh
m.parsons@epcc.ed.ac.uk

1. INTRODUCTION

The I/O bottleneck is a key problem faced by HPC applications. A potential solution to this bottleneck is by using optimised I/O libraries which can be more efficient while writing data. ADIOS2 is a relatively new I/O library which is designed to provide scalable I/O performance [1]. It does this by using a writer-per-node functionality and by writing to a directory rather than one single file. It is able to call I/O backends such as HDF5 and its own proprietary BP4 and BP5 formats with minimal changes to the source code. Additionally it can use a runtime xml file to change the backends without having to recompile the original application. In this work ADIOS2, HDF5 and MPIIO will be benchmarked while writing the same amout of data to compare the resulting I/O bandwidths. Further, the effects of writing to a burst buffer storage will be analysed as compared to writing to disk on ARCHER2. These results will be analysed using the DARSHAN [2] I/O profiler and presented in this poster.

2. METHODS

A C-based benchmarking tool, benchmark_c [3] was created to compare the I/O bandwidths and other userful metrics obtained from the different I/O backends by writing the same size of data to disk. Different parameters are compared against such as stripe count, stripe size, number of ranks and problem size. Both weak and strong scalings were used to compare the scaling performance of the backends. The results are obtained by submitting jobs to ARCHER2 [4] by adding the DARSHAN I/O profiler [2] path. The results from DARSHAN were used to analyse different file sizes and number of files written by the different backends. The I/O performance obtained by writing to the newly installed ARCHER2's burst buffer storage was compared wrt writing to disk.

3. RESULTS

Results have shown that the ADIOS2 proprietary formats BP4 and BP5 display far better scaling properties than MPIIO and HDF5. From the DARSHAN results, it is possible that the higher bandwidth is due to the smaller file sizes and the higher frequency file writes. Interestingly, ADIOS2 calling HDF5 performed worse-off than by calling HDF5 directly.

It was found that the speedup obtained by writing to the burst buffer as compared to writing to disk, was not as much as advertised. From the preliminary results, it was found that ADIOS2 calling HDF5 had the highest speedup as compared to writing to disk. These results were obtained when the burst buffer was first introduced and could be inaccurate due to teething problems with new hardware and software. However, these results will be re-run to get more accuracy.

From the DARSHAN results, the different behaviours for conducting I/O operations were observed when the different I/O libraries were used. Using a plotting tool developed in python, bar plots were created to display the frequency and size of file writes used by different I/O libraries. A detailed plot will be presented explaining the different results obtained.

4. ACKNOWLEDGMENTS

The ARCHER2 UK National Supercomputing Service was used to obtain the results used in this submission [4]. This work was supported by an EPCC funded studentship as part of the ASiMoV project. Funding and support from EPCC is gratefully acknowledged.

REFERENCES

[1] ADIOS2. [Online]. Available: https://adios2.readthedocs.io/en/latest/.

[2] DARSHAN. [Online]. Available: https://github.com/darshan-hpc/darshan.

[3] S. Bhardwaj, "benchmark_c," [Online]. Available: https://github.com/sb15895/benchmark_c.git.

[4] EPCC, "ARCHER2," [Online]. Available: https://www.archer2.ac.uk.