Understanding Parallel I/O Performance and Tuning

Suren Byna
sbyna@lbl.gov
Lawrence Berkeley National Laboratory
Berkeley, California, USA

ABSTRACT

Performance of parallel I/O is critical for large-scale scientific applications to store and access data from parallel file systems on high-performance computing (HPC) systems. These applications use HPC systems often to generate and analyze large amounts of data. They use the parallel I/O software stack for accessing and retrieving data. This stack includes several layers of software libraries – high-level I/O libraries such as HDF5, middleware (MPI-IO), and low-level I/O libraries (POSIX, STD-IO). Each of these layers have complex inter-dependencies among them that impact the I/O performance significantly. As a result, scientific applications frequently spend a large fraction of their execution time in reading and writing data on parallel file systems. These inter-dependencies also complicate tuning parallel I/O performance.

A typical parallel I/O performance tuning approach includes collecting performance logs or traces, identifying performance bottlenecks, attributing root causes, and devising optimization strategies. Toward this systematic process, we have done research in collecting Darshan traces for I/O [7], studying logs on production supercomputing systems [4, 6], attributing root cause analysis by zooming into application I/O performance [5], visualizing parallel I/O performance [3, 5], and applying performance tuning [1–3].

We will introduce parallel I/O basics, I/O monitoring using various profiling tools, analysis of logs collected on production class supercomputers to identify performance bottlenecks, and application of performance tuning options. We will also describe numerous application use cases and performance improvements.

CCS CONCEPTS

• **Information systems** → *Distributed storage.*

KEYWORDS

parallel I/O, I/O performance, performance tuning, parallel file systems

ACM Reference Format:
Suren Byna. 2022. Understanding Parallel I/O Performance and Tuning. In *Proceedings of the Fifth International Workshop on Systems and Network Telemetry and Analytics (SNTA '22), June 30, 2022, Minneapolis, MN, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3526064.3534114

1 BIOGRAPHY

Suren Byna is a Staff Scientist in the Scientific Data Management (SDM) Group at Lawrence Berkeley National Lab. His research interests are in scalable scientific data management. More specifically, he works on optimizing parallel I/O performance, developing systems for managing scientific data, and supporting scientists to find data of their interest efficiently. He is an author or co-author of more than 150 publications in the high-performance computing area. He is the PI of the ECP funded ExaIO project, and ASCR funded object-centric data management systems (Proactive Data Containers - PDC) and experimental and observational data management (EOD-HDF5) projects.

ACKNOWLEDGMENTS

Contributions from numerous researchers and engineers that worked on the ExaHDF5, EOD-HDF5, ExaIO, and TOKIO projects were used in this presentation. The speaker would like to thank all members of these projects and in specific Glenn Lockwood, Babak Behzad, Quincey Koziol, Teng Wang, Nick Wright, and Devesh Tiwari. The speaker would also like to thank collaborators from Argonne National Laborator (Rob Ross, Phil Carns, Shane Snyder, and Venkat Vishwanath) and those from Oak Ridge National Laboratory (Sarp Oral and Bing Xie). This work is supported by the Director, Office of Science, Office of Advanced Scientific Computing Research, of the U.S. Department of Energy under Contract No.DE-AC02-05CH11231.

SNTA '22, June 30, 2022, Minneapolis, MN, USA
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9315-7/22/06…$15.00
https://doi.org/10.1145/3526064.3534114

REFERENCES

[1] Megha Agarwal, Divyansh Singhvi, Preeti Malakar, and Suren Byna. 2019. Active Learning-based Automatic Tuning and Prediction of Parallel I/O Performance. In *2019 IEEE/ACM Fourth International Parallel Data Systems Workshop (PDSW)*. 20–29. https://doi.org/10.1109/PDSW49588.2019.00007

[2] Babak Behzad, Surendra Byna, Prabhat, and Marc Snir. 2019. Optimizing I/O Performance of HPC Applications with Autotuning. *ACM Trans. Parallel Comput.* 5, 4, Article 15 (March 2019), 27 pages. https://doi.org/10.1145/3309205

[3] Jean Luca Bez, Houjun Tang, Bing Xie, David Williams-Young, Rob Latham, Rob Ross, Sarp Oral, and Suren Byna. 2021. I/O Bottleneck Detection and Tuning: Connecting the Dots using Interactive Log Analysis. In *2021 IEEE/ACM Sixth International Parallel Data Systems Workshop (PDSW)*. 15–22. https://doi.org/10.1109/PDSW54622.2021.00008

[4] Glenn K. Lockwood, Shane Snyder, Teng Wang, Suren Byna, Philip Carns, and Nicholas J. Wright. 2018. A Year in the Life of a Parallel File System. In *SC18: International Conference for High Performance Computing, Networking, Storage and Analysis*. 931–943. https://doi.org/10.1109/SC.2018.00077

[5] Teng Wang, Suren Byna, Glenn K. Lockwood, Shane Snyder, Philip Carns, Sunggon Kim, and Nicholas J. Wright. 2019. A Zoom-in Analysis of I/O Logs to Detect Root Causes of I/O Performance Bottlenecks. In *2019 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID)*. 102–111. https://doi.org/10.1109/CCGRID.2019.00021

[6] Teng Wang, Shane Snyder, Glenn Lockwood, Philip Carns, Nicholas Wright, and Suren Byna. 2018. IOMiner: Large-Scale Analytics Framework for Gaining Knowledge from I/O Logs. In *2018 IEEE International Conference on Cluster Computing (CLUSTER)*. 466–476. https://doi.org/10.1109/CLUSTER.2018.00062

[7] Cong Xu, Shane Snyder, Omkar Kulkarni, Vishwanath Venkatesan, Phillip Carns, Surendra Byna, Robert Sisneros, and Kalyana Chadalavada. 2017. DXT: Darshan eXtended Tracing. In *Cray User Group (CUG) meeting*.