An Abstraction Layer for I/O Characterization of Large-Scale Applications

Advisors

- François Tessier - francois.tessier@inria.fr
- Julien Monniot - julien.monniot@inria.fr
- Gabriel Antoniu - gabriel.antoniu@inria.fr

Keywords

HPC, I/O, Simulation, Characterization

Subject

Nowadays, there are many scientific fields where the need for computing power and data processing capacity goes beyond what current machines can provide. In radio astronomy, for example, the international SKA project aims to create the largest telescope in the world in order to observe a part of the Universe. A very large volume of data is generated at the telescope level, pre-processed on local clusters (filtering, reduction) in real time and sent to a supercomputer at a rate of 10TB/s. This data feeds numerical simulation, generating 1PB of daily output data that needs to be saved. At this stage, the computing power and storage resources required are such that machines capable of reaching the **exascale** become necessary. To date, only a few supercomputers such as Frontier at Oak Ridge National Laboratory (USA) have this capability, but in the coming months, new systems will be deployed. However, the efficient use of these systems raises new challenges, especially regarding **data management**.

Indeed, even though HPC systems are increasingly powerful, there has been a relative decline in I/O bandwidth. Over the past ten years, the ratio of I/O bandwidth to computing power of the top three supercomputers has been divided by 10 while in some scientific computing centers the volume of data stored has been multiplied by 40 [1]. To mitigate this congestion, **new tiers of memory and storage** have been added to recently deployed supercomputers, increasing their complexity. Current work is looking at how to efficiently allocate these storage resources to applications [2, 3]. Yet, information on what these applications are going to do with them is often limited to read or written data volume. However, making the most of these new resources requires a **better understanding of how applications and workflows perform their I/O** [4]. While very fine-grained monitoring is not realistic (overhead, analysis costs), characterizing application I/O patterns at a higher level can enable appropriate optimizations significantly improving storage allocation decisions, leading to improved I/O performance.

The aim of this internship will be to work on an **abstraction layer for the I/O behavior** of applications. This abstraction will take the form of a configuration interface enabling a high-level description of the I/O pattern (compute-write, read-compute, interval-writes, ...). This information will be used to feed a job scheduler and help it allocate the right storage resources to applications in order to reduce I/O concurrency at the machine level. The evaluation of this implementation will be done in an existing WRENCH-based [5] **simulator**, called StorAlloc [2], developed in the team.

The selected student will have the opportunity to join a very dynamic international research team in a stimulating work environment with a lot of active collaborations. This internship comes with an important opportunity to pursue a thesis co-supervised by the CEA and Inria as part of the national NumPEx project [6], the aim of which is to prepare for the arrival of the first French Exascale system in 2025.

Skills and abilities

- Programming skills (Bash, C/C++, Python)
- Knowledge of computer networks and distributed systems
- Familiarity with high-performance computing or cloud computing is an advantage

Bibliography

[1] GK. Lockwood, D. Hazen, Q. Koziol, RS. Canon, K. Antypas, and J. Balewski. "Storage 2020: A Vision for the Future of HPC Storage". In: Report: LBNL-2001072. Lawrence Berkeley National Laboratory, 2017.

[2] Monniot, J, Tessier, F, Robert, M, Antoniu, G. Supporting dynamic allocation of heterogeneous storage resources on HPC systems. Concurrency Computat Pract Exper. 2023;e7890. doi: 10.1002/cpe.7890

[3] F. Tessier, M. Martinasso, M. Chesi, M. Klein, M. Gila. "Dynamic Provisioning of Storage Resources: A Case Study with Burst Buffers". In: IPDPSW 2020 - IEEE International Parallel and Distributed Processing Symposium Workshops, May 2020, New Orleans, United States.

[4] J. L. Bez, F. Z. Boito, R. Nou, A. Miranda, T. Cortes and P. O. A. Navaux, "Detecting I/O Access Patterns of HPC Workloads at Runtime," *2019 31st International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)*, Campo Grande, Brazil, 2019, pp. 80-87, doi: 10.1109/SBAC-PAD.2019.00025.

[5] H. Casanova, R. Ferreira da Silva, R. Tanaka, S. Pandey, G. Jethwani, W. Koch, S. Albrecht, J. Oeth, and F. Suter. "Developing Accurate and Scalable Simulators of Production Workflow Management Systems with WRENCH". In: Future Generation Computer Systems, vol. 112, p. 162-175, 2020.

[6] https://numpex.irisa.fr/