Performance Optimization of HPC Applications in Large-Scale Cluster Systems

Longxiang Li

lilongxiang01@inspur.com

Inspur (Beijing) Electronic Information Industry Company Ltd.

Beijing, China

ABSTRACT

In modern HPC clusters, the performance of an application is a combination of several aspects. To successfully improve the application performance, all performance aspects should be analyzed and optimized. In particular, as modern CPUs contain more and more cores, the speed of floating-point computations has increased rapidly, making data access one of the main bottlenecks in most HPC applications. Furthermore, performance diagnosing for HPC applications can be extremely complex, and the performance bottlenecks of HPC applications may vary with the scale of parallelism. In this presentation, a multi-layered data access (MLDA) optimization methodology is introduced. Developers could follow this methodology to optimize the HPC applications. We provide several examples of applying the MLDA method on real-world HPC applications, including the weather, ocean, material science, CFD, and MHD areas.

CCS CONCEPTS

• **General and reference** → **Performance**.

KEYWORDS

performance analysis, multi-layer data access optimization, cluster architecture, HPC optimization methodology

ACM Reference Format:
Longxiang Li. 2022. Performance Optimization of HPC Applications in Large-Scale Cluster Systems. In *Proceedings of the 2022 ACM/SPEC International Conference on Performance Engineering (ICPE '22), April 9–13, 2022, Bejing, China.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3489525.3511696

BIOGRAPHY

Dr. Li received his degree in Engineering from the Civil Engineering Department of the Tianjin University, in 2019, with Professor Qinghe Zhang as his advisor. During his Ph.D. studies, he was working on the development of the numerical model NDFEM based on the discontinuous finite element method and applying the model on simulation of physical problems such as nearshore hydrodynamics and tsunami waves. After graduation, he entered the AI&HPC software department of Inspur Information and is working on the development of the performance analysis tool - TEYE. He has been involved in the analysis and optimization of several HPC applications. His research area includes performance model research and HPC application optimization.

Publication list:

(1) Li L., Zhang Q. Development of an efficient wetting and drying treatment for shallow water modeling using the quadrature-free RKDG method[J]. International Journal for Numerical Methods in Fluids, 2020.

(2) Li L., Zhang Q. A new vertex-based limiting approach for nodal discontinuous Galerkin methods on arbitrary unstructured meshes[J]. Computers & Fluids, 2017, 159:316-326.

(3) Li L, Zhang Q. A Quadrature-Free Scheme for Nodal Discontinuous Galerkin Method on Arbitrary Quadrilateral Unstructured Meshes[J]. Journal of Tianjin University Science and Technology, 2018.