Benchmarking and Continuous Performance Monitoring of HPC Resources using the XDMoD Application Kernel Module

1st Nikolay A. Simakov
Center for Computational Research
SUNY University at Buffalo
Buffalo, NY, USA
nikolays@buffalo.edu

***Abstract*—High-Performance Computing (HPC) resources are integral to scientific and engineering computations, yet they are associated with substantial initial and operational costs. Consequently, it is essential to ensure their optimal utilization throughout their lifecycle. Typically, significant attention is devoted to performance optimization during the initial stages of the HPC resource lifecycle. However, performance assessments are conducted less frequently during the production phase, which can result in prolonged periods of suboptimal hardware and software performance.**

The Continuous Performance Monitoring Module [1] of XDMoD [2] addresses this issue by automating the execution of a set of applications and benchmarks on a regular basis, typically daily, to proactively identify performance degradation. The input data are carefully designed to yield crucial performance insights within a short wall time, enabling efficient monitoring. The module incorporates an algorithm that detects performance degradation, and it can be configured to send email reports either regularly or upon detecting performance issues.

Since 2011, this Performance Monitoring Module has been employed to oversee the performance of XSEDE and, subsequently, ACCESS resources. It has also been used to compare the performance of different hardware [3], [4]. In our presentation, we detailed the functionality of the Continuous Performance Monitoring Module and provided illustrative use cases demonstrating its application.

***Index Terms*—High-Performance Computing, Performance Monitoring, Benchmarks, Regression Tests**

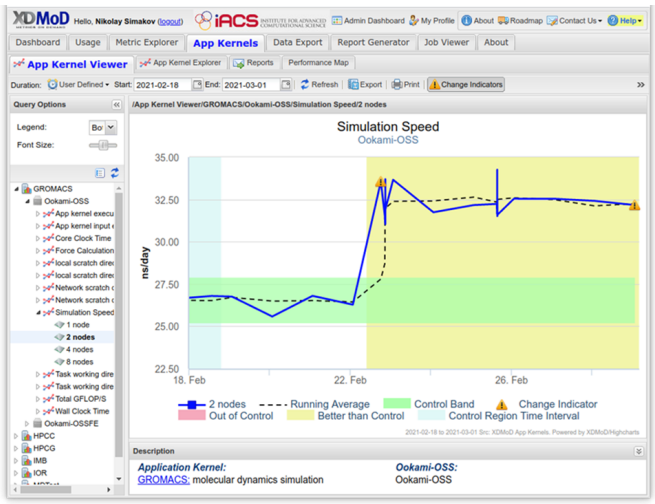

Fig. 1. XDMoD portal showing Gromacs application performance. The green horizontal slab in the control region, the vertical blue shows the region during which the control region was established, and the yellow shows the region where the application is over-performing. The exclamation mark in the triangle indicates a change in the application signature (i.e., the hash sum on binary or used shared libraries has changed).

ACKNOWLEDGMENT

This work is supported by the National Science Foundation under awards OAC 2137603.

REFERENCES

[1] N. A. Simakov, J. P. White, R. L. DeLeon, A. Ghadersohi, T. R. Furlani, M. D. Jones, S. M. Gallo, and A. K. Patra, "Application kernels: Hpc resources performance monitoring and variance analysis," *Concurrency and Computation: Practice and Experience*, vol. 27, no. 17, pp. 5238–5260, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpe.3564

[2] J. T. Palmer, S. M. Gallo, T. R. Furlani, M. D. Jones, R. L. DeLeon, J. P. White, N. Simakov, A. K. Patra, J. Sperhac, T. Yearke, R. Rathsam, M. Innus, C. D. Cornelius, J. C. Browne, W. L. Barth, and R. T. Evans, "Open xdmod: A tool for the comprehensive management of high-performance computing resources," *Computing in Science & Engineering*, vol. 17, no. 4, pp. 52–62, 7 2015.

[3] N. A. Simakov, M. D. Jones, T. R. Furlani, E. Siegmann, and R. J. Harrison, "First impressions of the NVIDIA Grace CPU Superchip and NVIDIA Grace Hopper Superchip for scientific workloads," ser. HPCAsia '24 Workshops. New York, NY, USA: Association for Computing Machinery, 2024, p. 36–44.

[4] E. Siegmann, R. J. Harrison, D. Carlson, S. Chheda, A. Curtis, F. Coskun, R. Gonzalez, D. Wood, and N. A. Simakov, "First impressions of the sapphire rapids processor with hbm for scientific workloads," *SN Computer Science*, vol. 5, no. 5, p. 623, 2024.