White-box Modelling of Parallel Computing Dynamics

Ayesha Afzal, Georg Hager (advisor), and Gerhard Wellein (advisor)
{ayesha.afzal,georg.hager,gerhard.wellein}@fau.de
Erlangen National High Performance Computing Center (NHR@FAU)
Friedrich-Alexander University Erlangen-Nürnberg, Germany

1 PROBLEM AND MOTIVATION

"White-box" performance modeling of distributed-memory programs is often imprecise due to a wide spectrum of disturbances in the application and the system, which cause multi-faceted performance impacts. Even for applications having the regular, homogeneous compute-communicate phases and a perfect translational symmetry across processes, simply adding communication time (via Hockney, LogGP or other model) to computation time (via Roofline, ECM or other model) does not often give a good prediction of total parallel runtime. This is due to dynamics in the parallel system that destroy the "lock-step" pattern of applications. A comprehensive theory of such dynamics is a great challenge.

2 WHITE-BOX MODELLING OF DYNAMICS

To achieve better insight into this issue, we developed a validated analytic model [1] of the propagation speed (in ranks per second) of "idle waves". Idle waves are initiated when delays emanating from different sources on the individual MPI processes propagate across the other processes depending on the execution and communication properties of the code. We use a spectrum of HPC clusters and widespread applications to explore how idle waves interact nonlinearly within a parallel application on a cluster. Furthermore, idle waves decay in time due to three key ingredients: differences in communication characteristics among parts of the system, imbalanced application load, or system noise. Although one might assume that idle waves are eliminated by collective communication, we could show that some implementation variants of collectives are permeable to them [3]. In addition, the presence of bottleneck(s) among processes (in terms of intra-node memory bandwidth, per-node network injection bandwidth, and full-system bisection bandwidth) may break the inherent symmetry of the underlying software and hardware and enable a transition into a structured, stable desynchronized state [2]. In this state, different cores can execute different loop kernels and facilitate an automatic communication and computation overlap that we investigate further for the code performance.

3 PERFORMANCE ENGINEERING

We have devised a performance model [4] for overlapping execution of memory-bound loop kernels, which is also applicable to task-based programs. The model can predict the memory bandwidth share per kernel on a memory contention domain depending on three parameters: the number of active cores, the single-core bandwidth and the saturated bandwidth characteristics of all paired kernels. Although an overlaping execution can cause increased waiting time per process, it can also boost the available per-core bandwidth via bottleneck evasion, improving the overall time to solution without explicit programming. To model spontaneous communication overlap, the minimum number of processes per memory domain required to achieve full memory bandwidth is one of the decisive factors in memory-bound parallel programs.

4 PHYSICAL OSCILLATOR MODEL

We propose a novel physical oscillator model [5] along the lines of the well-known Kuramoto model that suggests a physical interpretation of processes as a set of coupled oscillators whose inherent frequencies are influenced by an interaction potential and coupling topology. Model saves both resources and time via a system of first-order coupled ordinary differential equations that are easily solved by standard numerical methods. We validate the correspondence of model and real behavior of parallel programs: At medium-scale, the delay propagation across processes reflected via the communication topology of programs is mapped by the interaction matrix in the model. Idle wave decay due to application and system noise is mimicked via the appropriate incorporation of (local and interaction) fine-grained noise in the model. At large-scale, bottleneck-free parallel programs exhibit self-synchronization across processes and are well described by a long-range hyperbolic tangent interaction potential in the model. In contrast, in resource-bottlenecked parallel programs, a short-range repulsive interaction potential is more appropriate, shifting stability points away from the translationally symmetric, unstable "lock-step" pattern. The bottleneck evasion emerging from the symmetry-breaking desynchronization is an indication of a potential Goldstone-like mode.

5 CONCLUSION

We explore various aspects and application scenarios of our innovative "white-box" modelling approach that addresses the analysis and simulation challenges of the dynamics of parallel computing and connects them to the physical world. Our analysis encompasses the performance-limiting bottlenecks and is useful in guiding appropriate code changes and can produce thorough insight into the hardware-software interaction. Our oscillator model of parallel computing is a viable candidate for a high-level cluster characterization tool for both regular and irregular massively parallel programs.

REFERENCES

[1] Ayesha Afzal, Georg Hager, and Gerhard Wellein. 2019. Propagation and Decay of Injected One-Off Delays on Clusters: A Case Study. In *Proceedings - IEEE International Conference on Cluster Computing, ICCC.*

[2] Ayesha Afzal, Georg Hager, and Gerhard Wellein. 2020. Desynchronization and Wave Pattern Formation in MPI-Parallel and Hybrid Memory-Bound Programs. In *Lecture Notes in Computer Science.* Springer International Publishing.

[3] Ayesha Afzal, Georg Hager, and Gerhard Wellein. 2021. Analytic Modeling of Idle Waves in Parallel Programs: Communication, Cluster Topology, and Noise Impact. In *Lecture Notes in Computer Science.* Springer International Publishing.

[4] Ayesha Afzal, Georg Hager, and Gerhard Wellein. 2021. Analytic Performance Model for Parallel Overlapping Memory-Bound Kernels. arXiv:2011.00243 Accepted for Concurrency and Computation: Practice and Experience.

[5] Ayesha Afzal, Georg Hager, and Gerhard Wellein. 2021. Physical Oscillator Model for Parallel Distributed Computing. Poster at ISC High Performance.