

The NASTJA Framework: Non-Collective Scalable Global Communications

MARCO BERGHOFF, Karlsruhe Institute of Technology, Germany

IVAN KONDOV, Karlsruhe Institute of Technology, Germany

In recent years, simulations in various areas of science and engineering have proven to be very useful. They are used to investigate various phenomena, often by solving partial differential equations. To efficiently deploy simulation codes on current and future high-performance computer systems with thousands of processor cores, high node level performance, scalable communication and the exclusion of unnecessary calculations are an absolute must when developing new solvers.

We have introduced the NASTJA framework, a block-based MPI parallel solver for algorithms, based on regular grid methods, i.e., stencil codes. NASTJA decomposes the domain into small cuboid blocks. A characteristic feature of NASTJA is a dynamic block adaptation, which modifies the calculation domain around the region in which the calculation is currently taking place, thus avoiding unnecessary calculations. Among other, this can be useful in phase-field simulations. The creation and deletion of blocks are autonomously managed within local neighborhoods. Collective all-gather communication is avoided by using a multi-hop network to distribute information across the entire domain that greatly improves the application scaling.

In this contribution, we present applications that can benefit from this adaptive method and scaling tests demonstrating the excellent scalability.

CCS Concepts: • **Computing methodologies** → **Massively parallel algorithms**; **Multiscale systems**; **Massively parallel and high-performance simulations**; • **Mathematics of computing** → *Partial differential equations*; • **Networks** → *Topology analysis and generation*; *Network algorithms*;

Additional Key Words and Phrases: stencil code, distributed memory, scalable parallel algorithms, massive-parallel performance, phase-field method, cellular Potts model

ACM Reference Format:

Marco Berghoff and Ivan Kondov. 2018. The NASTJA Framework: Non-Collective Scalable Global Communications. 1, 1 (October 2018), 2 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

Computer simulations help us to understand the theory behind various phenomena. Especially, when the studied effects are difficult to access experimentally or when experiments are too complex or expensive. With the opportunity to perform calculations on ever-larger machines, users' thirst for knowledge is also growing. Engineers, physicists, and other scientists consider now models with increasing accuracy, complexity and size employing multiscale and

Authors' addresses: Marco Berghoff, Karlsruhe Institute of Technology, Steinbuch Centre for Computing, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Germany, marco.berghoff@kit.edu; Ivan Kondov, Karlsruhe Institute of Technology, Steinbuch Centre for Computing, Hermann-von-Helmholtz-Platz 1, Eggenstein-Leopoldshafen, 76344, Germany, ivan.kondov@kit.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

XXXX-XXXX/2018/10-ART

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

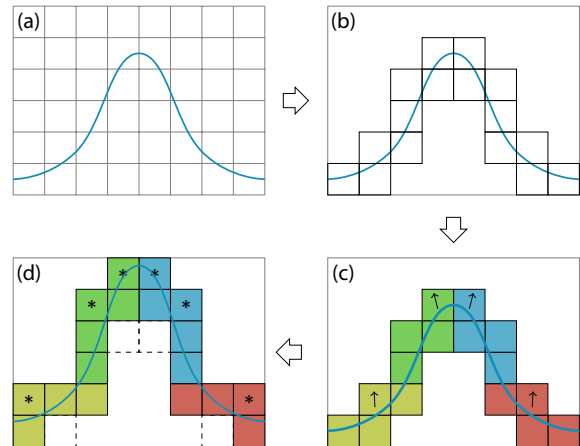


Fig. 1. Representation of a phase-field simulation with the interface (blue line) (a) A block-structured grid covers the whole domain. (b) Only blocks holding a part of the interface are allocated. (c) Allocated blocks are distributed to the MPI processes, marked by color. During the simulation, the interface changes in space. (d) The blocks are adapted correspondingly, as denoted by the stars for newly created blocks, and the dashes for deleted blocks.

multiphysics approaches and thus posing great challenges for simulation software. In particular, high node level performance, scalable parallel communication and omitting of unnecessary calculations are indispensable for efficient utilization of current and future HPC systems.

We have introduced NASTJA [Berghoff and Kondov 2018; Berghoff et al. 2018], a framework for simulations based on regular grids. From its outset, NASTJA is designed for high performance and flexibility to including new solvers and models. Blocks play a central role for the domain decomposition in NASTJA. A block-structured grid covers the whole simulation domain. Only the blocks that contain the interface, i.e., the part that has to be calculated, are allocated. These blocks are distributed to the MPI processes. As the interface is moving during the simulation, the blocks are adapted correspondingly. Figure 1 illustrates allocating and adapting blocks in a phase-field simulation with NASTJA.

In the following example, we will demonstrate the benefit of using NASTJA for a phase-field simulation of a water drop on a chemically structured surface depicted in Figure 2. Only the diffuse interface region between the water and gas phases has to be calculated. Bridging the scales between a few micrometers for the structured surface and a few millimeters for the diameter of a water drop requires large domain sizes. If the water drop is covered with a cuboid simulation domain, most of the domain contains the bulk of one phase that does not need to be calculated. The interface area to be calculated is

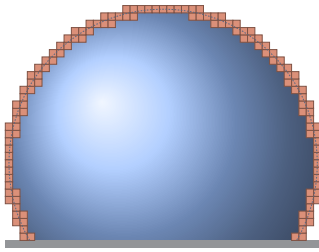


Fig. 2. A water drop on a chemically structured surface. The interface (dashed) is covered by allocated blocks (red).

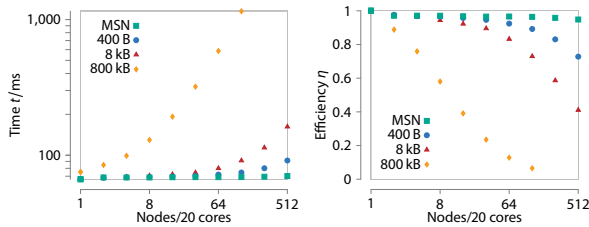


Fig. 3. The time and the efficiency of the MSN in comparison to MPI_Alltoall communications with different message sizes.

very small. For a smooth discretization of the interface area, a width of about ten grid points is required. Water drops on a chemically structured surface, with 700 lamellae, have been studied experimentally. A typical raindrop of about 3 mm in size covers all 700 lamellae, each with a size of about $5\ \mu\text{m}$ —the typical diameter of spider silk. In a phase-field simulation, the lamellae should be significantly wider than the diffuse interface. For 15 grid points per lamella, this results in a resolution of around 10^{12} grid points, in three dimensions. The regions outside the interface do not have to be calculated and thus are omitted in NASTJA. The computational effort and memory consumption can be reduced to only 3.5%, compared to the classical full-domain reference simulation as illustrated in Figure 2.

In certain situations, the local neighborhood is not sufficient to ensure the consistency of block adaptation. When two interfaces processed by two disjoint neighborhoods move towards each other during simulation, both of these neighborhoods could allocate the same new block. To prevent this inconsistency, a global exchange of information across all local neighborhoods of MPI processes is necessary. For a large number of MPI processes, both collective communication and master-worker communication are not scalable. With particular properties of the phase-field method, we have gained time until the next block is created from one created block. This time can be used to manage the global exchange of information. NASTJA uses a multi-hop network for this. We present the scaling with a multidimensional Manhattan Street Network in Figure 3. The diameter of the Manhattan Street Network can be reduced by increasing dimension so that it can be guaranteed that for many MPI processes the given time slot is sufficient to exchange the necessary information globally.

Figure 4 shows simulation of different crystals simulated with the phase-field crystal [Berghoff and Nestler 2015; Elder et al. 2007]

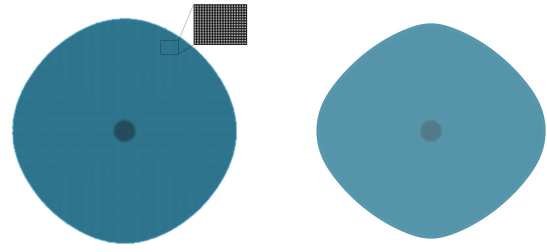


Fig. 4. Anisotropic crystal growth with a parallel semi-implicit spectral method for phase-field crystal simulations. The initial crystal with orientation 100 (left) and 110 (right) is marked dark.

module in NASTJA. A parallel semi-implicit spectral method is used to speed up the simulations. The anisotropy of the crystals is compared to molecular dynamics and phase-field simulations [Guerdane and Berghoff 2018].

Besides materials science simulations, the NASTJA framework can also cover other areas. Thus, a module is implemented that simulates biological cells with the cellular Potts model [Graner and Glazier 1992]. For this purpose, the individual grid points are re-colored to a fraction of the individual cells according to a Metropolis Monte Carlo algorithm that minimizes the energy such that the individual cells reach their target volume and surface.

While focusing on high scalability, the NASTJA framework is generic by design and supports any methods based on regular grids with a stencil sweeping through the simulation domain. Therefore, the framework provides a beneficial infrastructure not only for many applications in computational materials science (e.g., phase-field models) but also in other emerging fields of computational science, for example in computational systems biology (e.g., cellular automata) because all these applications are implemented on regular grids. Specifically, it already supports methods for droplets, as done by [Ben Said et al. 2014], handles phase-field crystal models and the cellular Potts model. In addition, the framework can be simply extended with a wide range of algorithms that work with finite difference schemes or other regular grid methods.

REFERENCES

- Marouen Ben Said, Michael Selzer, Britta Nestler, Daniel Braun, Christian Greiner, and Harald Garcke. 2014. A phase-field approach for wetting phenomena of multiphase droplets on solid surfaces. *Langmuir* 30, 14 (2014), 4033–4039.
- Marco Berghoff and Ivan Kondov. 2018. Scalable Global Network Based on Local Non-Collective Communications. *SC Workshop: Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA'18)* (2018).
- Marco Berghoff, Ivan Kondov, and Johannes Hötzer. 2018. Massively parallel Stencil Code Solver with Autonomous Adaptive Block Distribution. *IEEE Transactions on Parallel and Distributed Systems* 29, 10 (Oct 2018), 2282–2296.
- Marco Berghoff and Britta Nestler. 2015. Phase field crystal modeling of ternary solidification microstructures. *Computational Condensed Matter* 4 (2015), 46–58.
- Ken R. Elder, Nikolas Provatas, Joel Berry, Peter Stefanovic, and Martin Grant. 2007. Phase-field crystal modeling and classical density functional theory of freezing. *Physical Review B* 75, 6 (2007), 064107.
- François Graner and James A. Glazier. 1992. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Physical Review Letters* 69, 13 (1992), 2129–2132.
- Mohammed Guerdane and Marco Berghoff. 2018. Crystal-melt interface mobility in bcc Fe: Linking molecular dynamics to phase-field and phase-field crystal modeling. *Physical Review B* 97, 14 (2018), 144105.

Received July 2018