Scalable multi-physics simulations will require new discretization and numerical methods research

Veselin Dobrev, Tzanio Kolev, Robert Rieben, Bert Still
Lawrence Livermore National Laboratory
Contact: Tzanio Kolev, Box 808, L-561, Livermore, CA 94551, tzanio@llnl.gov, (925)423-9797

The ability to accurately model complex, interacting physical processes is a key component in developing insight and understanding across a broad range of scientific and engineering disciplines. The DOE has a long history as a leader in predictive science by developing world-class simulation tools. These codes were developed in an era when floating-point operations (FLOPs) were the performance-limiting factor influencing the choice of the discretization method and the overall design of the numerical algorithm. As computing architectures are changing from FLOP-centric to data-centric designs, the computing paradigm on which the codes were designed is being effectively inverted, since the cost of FLOPs is considerably smaller than the cost of data motion. This is illustrated in Figure 1, where the theoretical peak and realizable performance on two different types of data-centric architectures are compared (one tuned for enhanced floating-point, and another more balanced design) based on industry roadmaps. The realizable performance of refactored multi-physics codes based on current algorithms falls significantly short of peak indicating that the DOE codes cannot take advantage of new and emerging hardware features due to memory capacity and bandwidth constraints. Moreover, this performance gap is projected to grow with time. Research into code and data transformations on compressible shock hydrodynamics and diffusion algorithms has shown that perhaps a factor of two improvement of current codes is possible, but the performance gap cannot be closed.

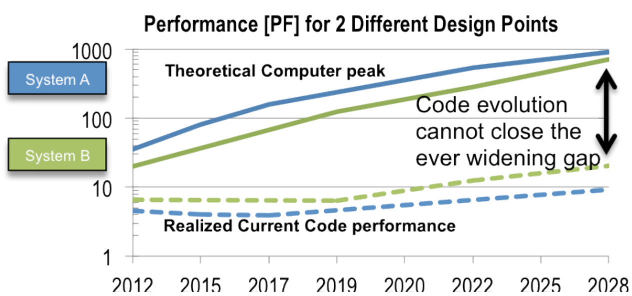

Figure 1. Peak versus realizable performance for two different computing architectures

We therefore argue that code optimization of current algorithms will not be sufficient for scalable performance on future computing architectures. A viable path to extreme parallel computing and exascale architectures will require fundamental mathematical research in discretization algorithms and numerical methods, which are designed from inception to take advantage of the data-centric hardware.

One example illustrating this position is a multi-material shock hydrodynamics discretization method based on (arbitrarily) high-order curvilinear finite elements, developed under LDRD funding at LLNL. While high-order methods have a long history and nowadays are widely recognized to be better suited for the current and future heterogeneous architectures, they were not used previously in computational shock hydrodynamics due to their high computational cost. Instead, many codes across the DOE national labs employ classical low-order schemes to discretize the Arbitrary Lagrangian-Eulerian (ALE) formulation of fluid dynamics. Recent research on the Lagrange phase of ALE shows that

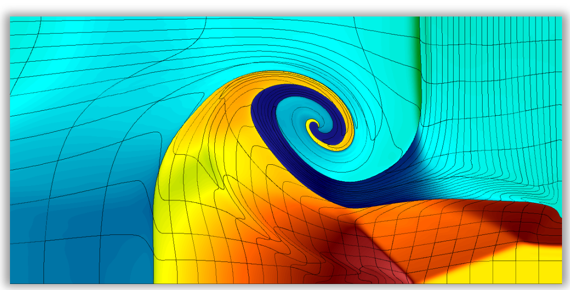

Figure 2. High-order finite element Lagrangian simulation

high-order methods can lead to significant improvements in robustness, accuracy and numerical conservation of physical quantities. They can also enable previously intractable simulations, such as the one shown in Figure 2, through highly curved zones and sub-zonal resolution of the shock waves, which are simply not possible with a low-order method. These improvements come at a cost of additional FLOPs per element, however the ratio to the required memory (FLOPs/bytes ratio or arithmetic intensity) is well aligned with the emerging data-centric architectures and can be tuned by adjusting the order of the method (introducing a new performance parameter in addition to mesh resolution). This is illustrated in Figure 3, showing that the increased computational intensity of high-order finite elements leads to excellent strong scaling in the research code BLAST on the Sequoia machine, paving the way to closing the gap from Figure 1. High-order methods also excel on heterogeneous (e.g., GPU+CPU) architectures, where one can take advantage of additional levels of parallelism (e.g., threading over quadrature points per element). Despite their appeal from both mathematical and computer science perspective, the application of high-order finite elements to coupled multi-physics problems involving ALE hydrodynamics, multi-group radiation-diffusion and magneto-hydrodynamics (MHD) requires substantial new discretization research in each of these multi-physics components, as well as in the high-order coupling between them. This is a challenge, but also an opportunity to improve the quality of the solution, while increasing the arithmetic intensity. Such high-order finite element research can lead to a new breed of DOE multi-physics codes, which are well-suited for emerging architectures and expand the state-of-the-art in high-fidelity modeling.

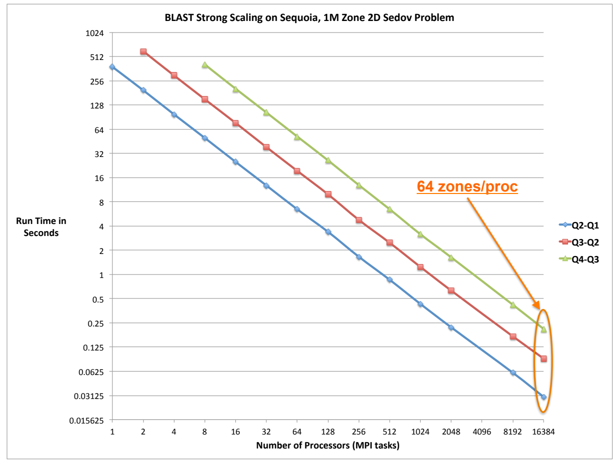

Figure 3. Strong scaling for quadratic (Q2-Q1), cubic (Q3-Q2) and quartic (Q4-Q3) finite elements

Another example of fundamental mathematical research leading to scalable performance in multi-physics problems is the development of the Auxiliary-space Maxwell Solver (AMS), which is the first provably-scalable linear solver for the definite Maxwell systems arising in MHD simulations of electromagnetic diffusion. This class of matrices has been challenging for traditional solution methods due to the large nullspace of the curl operator, limiting the problem resolutions that can be used in practice. In contrast, the AMS research demonstrated that better mathematical understanding of the nature of the problem could enable scalability on large problems with complex unstructured geometries and multiple materials with orders of magnitude jumps in the coefficients. This approach was enabled by additional discretization information and a recent Hiptmair-Xu stable decomposition of the Nedelec finite element space, leading to explicit handling of the curl-related near-nullspace and reduction of the edge-based problem to several nodal-based problems that can be treated with classical Algebraic Multigrid (AMG). In practice, AMS has been scaled up to 125K cores (see Figure 4) and has had a significant impact in large-scale MHD simulations in several major multi-physics DOE codes, exemplifying how research motivated by the properties of the application and deep mathematical insight can lead to a more scalable solver.

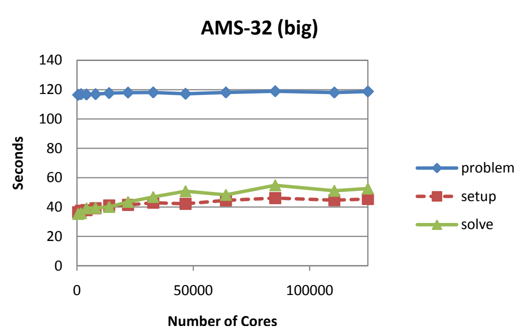

Figure 4. Weak scaling of the AMS setup and solve phases up to 125,000 cores (12 billion unknowns)