Engagement of the Leadership Computing Facilities in the Exascale Math Conversation: A Critical Gap

Scientific Computing Group
Oak Ridge Leadership Computing Facility
Point of Contact: Judith Hill, hilljc@ornl.gov

Since 2004, the Oak Ridge and Argonne Leadership Computing Facilities (OLCF and ALCF, respectively) have been the home of ASCR's leadership-class supercomputers. Together with ASCR's flagship production computing facility, NERSC, these facilities serve as the external face of ASCR's "*Big Iron*" to the broader scientific community, enabling transformational science and engineering in a wide variety of fields. These facilities will be integrally involved in the design and deployment of an ASCR exascale computer, and the community of users of an exascale computer will likely have strong overlap with existing users of ASCR computing facilities. The success of the OLCF is measured, in part, by the scientific impact of the computational work of our users. **Ultimately, the impact of the exascale math research will be measured by its ability to improve and enable the productivity of scientific applications that use computing resources such as those at the OLCF.**

The symbiotic relationship between ASCR Research, encompassing the Applied Math, Computer Science, and SciDAC efforts, and ASCR Facilities is clear. As we move towards exascale computing, an opportunity exists to strengthen the partnership between the Exascale Math effort and the OLCF. Over time, the OLCF has developed strong collaborations with its user base; in part, this is due to the advanced user support model of the Scientific Computing Group at the OLCF. Comprised of computational scientists drawing from a diverse set of application domains, including astrophysics, climate science, biology, and nuclear physics, as well as applied mathematicians and computer scientists, the Scientific Computing Group is uniquely positioned to serve as a liaison between the OLCF and domain scientists. While our focus has generally been on serving the needs of our (external-to-ASCR) users, *we assert that the experience and knowledge of this unique group of researchers would also be useful in predicting the needed and useful mathematical research for exascale computing.* Engaging in a conversation with the domain scientists about their mathematical needs and computing challenges could prove to be particularly useful.

In the past, when the OLCF has prepared to field a new supercomputer, we have engaged directly with the broad scientific community regarding their application requirements [1, 2]. These requirements have encompassed a broad range of topics, including their projected hardware and software-library needs, as well as their anticipated mathematical and algorithmic bottlenecks. The articulation of these requirements has been used to directly arbitrate architectural decisions for the OLCF's procurement of computing platforms. Importantly, also embedded in the information gathered is a glimpse at the changing mathematical and algorithmic landscape of science applications.

A useful characterization of algorithms prevalent today is the so-called seven algorithmic motifs of Colella [3]; Table 1 demonstrates this approach for the algorithms expected to play a key role within scientific applications at the exascale [3]. Several trends are noteworthy in this categorization:

- The seven algorithm types are scattered broadly among science domains, with no one particular algorithm being ubiquitous and no one algorithm going unused. *This indicates an opportunity for mathematical and numerical approaches that can impact multiple science domains.*

- Compared to the seven motifs for applications projected to the petascale [reported in 4], we can expect to see a significant increase in Monte Carlo methods, and increases in unstructured grids, sparse linear algebra, and particle methods, and a relative decrease in FFTs. These projections reflect the general expectation of much-greater parallelism in architectures and the resulting need for very high on-node scalability. Load balancing, scalable sparse solvers, and random number generator algorithms will also be more important.

Application area	Structured grids	Unstructured grids	FFT	Dense linear algebra	Sparse linear algebra	Particles	Monte Carlo
Molecular Dynamics			X	X		X	X
Nanoscience	X			X		X	X
Climate	X		X		X	X	X
Combustion	X	X		X		X	
Fusion	X		X	X	X	X	X
Nuclear energy		X		X	X		
Astrophysics	X	X	X	X	X	X	X
Nuclear physics				X			
Accelerator physics		X			X		
Lattice QCD	X				X		X
Aerodynamics	X	X		X	X		

Table 1. Algorithms expected to play a key role within select scientific applications at the exascale, characterized according to a seven motifs classification

Conspicuously, several emerging algorithmic areas are absent from the seven motifs. Categories that could be expected to be of growing importance in the 2020 time frame include adaptive mesh refinement, implicit nonlinear systems, data assimilation, agent-based methods, parameter continuation, optimization, ensemble methods (e.g. stochastic methods and uncertainty quantification), and data analytics. Though application scientists are actively thinking about many of these motifs, production-ready, hardened mathematical software is a **prerequisite** to the serious exploration of new algorithmic techniques for domain experts. The realities of workforce funding and prioritization of effort mean that clear benefit must be evident before serious development work can begin. To be sure, this same sort of prioritization will also be required for ASCR research teams, as they attempt to ascertain what pieces of mathematical software will be most widely-used by application teams.

Adding to this calculus is the fact that the size and complexity of modern, large-scale scientific codes will increase at least as fast as the platforms on which these new codes will run. **To deal with the new challenges and opportunities provided by future exascale architectures, it is imperative that collaborations be forged among applied mathematicians, computer scientists, computational scientists, and domain scientists to investigate new algorithms that might be more appropriate for the highly parallel hybrid HPC architectures anticipated for these systems**. Without this sort of close interaction, currently-used algorithms that are largely based on experience with serial machines will continue their hegemony. As near-future architectures—and software implementations for them—become more and more based on ubiquitous parallelism and more complex memory hierarchies, this sort of stagnation would strongly curtail computational efficiency. As the computational home for a wide variety of relevant scientific applications already in production, the OLCF is a storehouse of knowledge and experience with the difficulty of efficiently computing at the petascale. *Building on that foundation, we hope to provide input on the mathematical challenges that application scientists will face at the exascale.*

References

1. W. Joubert, D. Kothe and H.A. Nam. Preparing for Exascale: ORNL Leadership Computing Facility Application Requirements and Strategy. ORNL Technical Report ORNL/TM-2009/308. December 2009. https://www.olcf.ornl.gov/wp-content/uploads/2010/03/olcf-requirements.pdf

2. S. Ahern, S.R. Alam, M.R. Fahey, R.J. Hartman-Baker, R.F. Barrett, R.A. Kendall, D.B. Kothe, R.T. Mills, R. Sankaran, A.N. Tharrington, and J.B. White III. Scientific Application Requirements for Leadership Computing at the Exascale. ORNL Technical Report ORNL/TM-2007/238. December 2007. https://www.olcf.ornl.gov/wp-content/uploads/2010/03/Exascale_Reqms.pdf

3. P. Colella, "Defining Software Requirements for Scientific Computing," DARPA HPCS presentation, 2004.

4. D. Kothe and R. Kendall. Computational Science Requirements for Leadership Computing (Oak Ridge National Laboratory National Center for Computational Sciences, July 2007), available at http://nccs.gov/mediacenter/nccs-reports/. https://www.olcf.ornl.gov/wp-content/uploads/2010/03/ORNL_TM-2007_44.pdf.