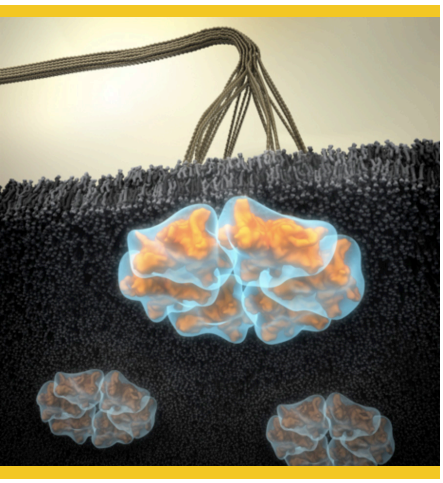
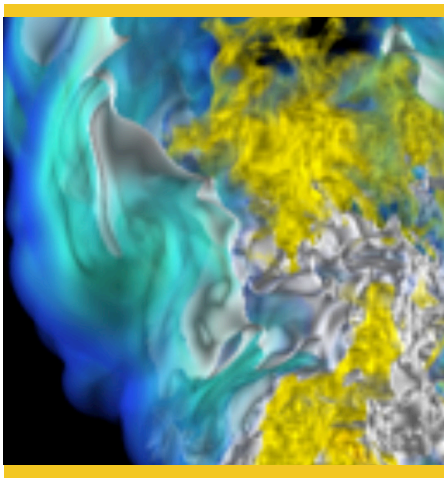
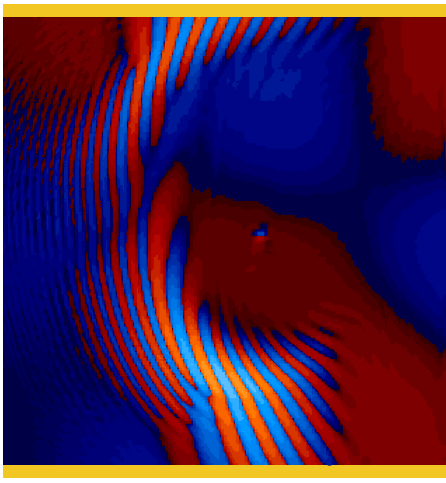
U.S. DEPARTMENT OF ENERGY

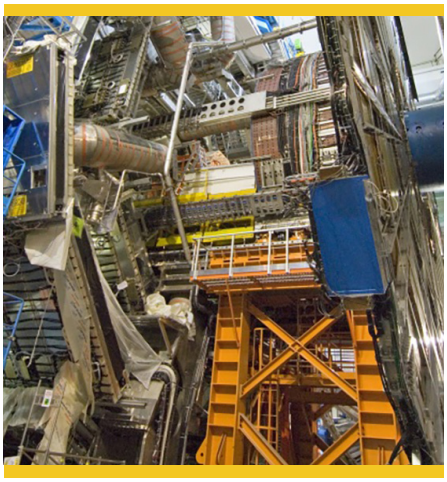
CROSSCUT

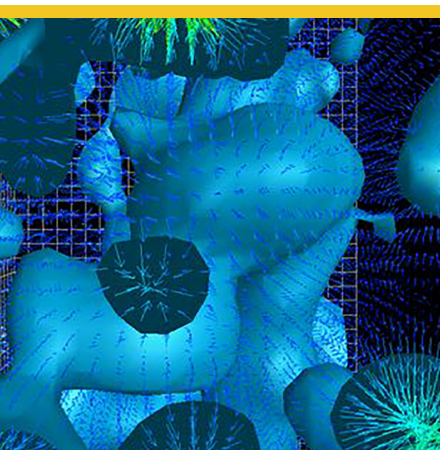
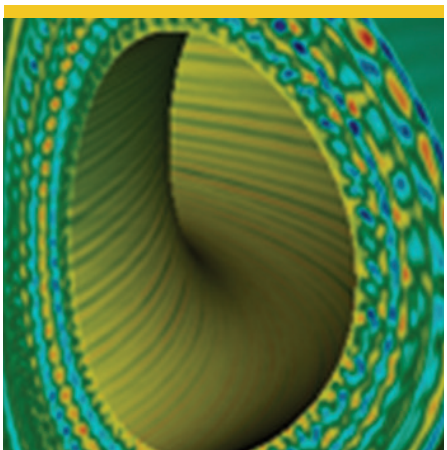
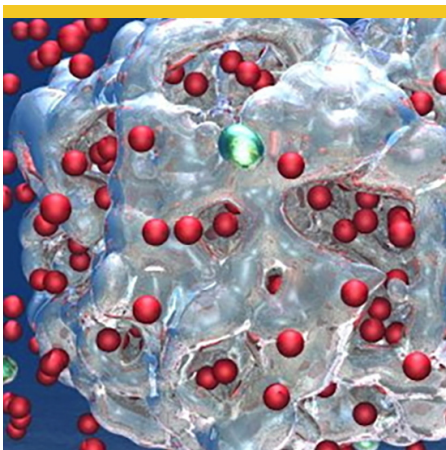
EXASCALE REQUIREMENTS REVIEW

An Office of Science review sponsored jointly by Advanced Scientific Computing Research, Basic Energy Sciences, Biological and Environmental Research, Fusion Energy Sciences, High Energy Physics, and Nuclear Physics

March 9–10, 2017
Tysons Corner, Virginia



600

SCIENTISTS AND **DOE PLANNERS** MET THROUGHOUT 2015 AND 2016

1,496

PAGES OF **FINDINGS**

100

CROSS-CUT MEETING **PARTICIPANTS** DEBATED AND REVIEWED FOUR KEY AREAS

Exascale Requirements Reviews and the Cross-Cut Meeting

During 2015 and 2016, the U.S. Department of Energy (DOE) conducted Exascale Requirements Reviews for each of its six Office of Science (SC) program offices.

On March 9–10, 2017, 100 attendees including DOE program managers, members of the DOE SC scientific community, and ASCR facilities staff met in Tyson's Corner, Virginia, to begin identifying key findings from the six exascale reports.

Key Findings

Computing

To meet the mission, DOE SC researchers have substantial needs that vary widely in complexity. The demand for both extreme-scale, high-performance computing (HPC) and HPC computing at a somewhat smaller scale exceeds the supply at DOE's User Facilities. For example, computational needs are expected to increase by a factor of 1,000 by 2025 at the frontiers of SC's initiatives. At the same time, large numbers of medium-scale runs are needed to support uncertainty quantification (UQ) and the verification and validation (VV) of models.

A successful DOE exascale ecosystem will support a range of workflows, including providing wide and early access to testbed and prototype systems for code and workflow development prior to deployment on leading-edge production systems at scale. Scheduling and allocation policies must support these workflow needs, especially concerning data analysis requirements at experimental facilities. Multiyear commitments of resources are needed to assure the success of long-range scientific campaigns.

Data

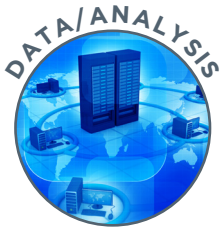
The volumes, the rapidity, and the complexity of data generated from experiments and simulations require data management, archiving, and curation capabilities that will not be satisfied by the best practices employed today. Sharing, transferring, and accessing data remotely and securely are becoming increasingly important in today's community-focused efforts.

As workflows related to data movement in both simulation and analysis become more complex, scientists use HPC to help guide experiments in real time — Advanced Scientific Computing Research (ASCR) and other DOE SC user facilities will be challenged to work out co-scheduling and data management approaches that will optimize the productivity of all facility resources.

The input/output (I/O) capabilities of large HPC systems need to scale with their computational capabilities, and sometimes grow faster. Simulations cannot spend excessive amounts of time blocking on I/O, and data read/write rates can limit performance of data analysis pipelines.

Software and Application Development

The development of new mathematics, software, and algorithms to take advantage of exascale computing architectures and to meet data analysis, management, and workflow needs was clearly highlighted by the SC communities. Standard mathematical libraries and frameworks must transition to new architectures in a predictable and timely fashion. The reports also express the urgent need to identify which common tools and libraries will be supported and sustained, especially those sponsored by DOE.



Scalable data processing, data analysis, machine learning, discrete algorithms, and multiscale/multiphysics simulations will be crucial to reducing and understanding the large-scale data generated. Developer productivity needs be improved, and researchers must be able to package DOE research artifacts for broad reuse. All SC communities expressed a desire for sufficiently performance-portable and expressive programming models.

Broad consensus exists that, to advance the sustainability of DOE software infrastructure and the productivity of DOE researchers, a common computing and data analysis environment must be established across platforms and facilities — highlighting the need for community planning; policies; and processes to improve software quality, interoperability, testing, and deployment.

Training and Workforce Development

All six requirements reports emphasized the importance of investing significantly in workforce development and training, including via ASCR partnerships. Communities across DOE SC need individuals who can develop methods and algorithms optimized for exascale systems; use and configure those systems effectively and efficiently; and write, provide, and maintain applications software.

Addressing this need will involve creating HPC career paths, funding training on better workflows, and transitioning codes to next-generation platforms. In addition, the community would be served well by ASCR facilities developing and providing access to prototype and large pre-production systems.

Additional Directions

Successful Cross-Cut Facility Efforts

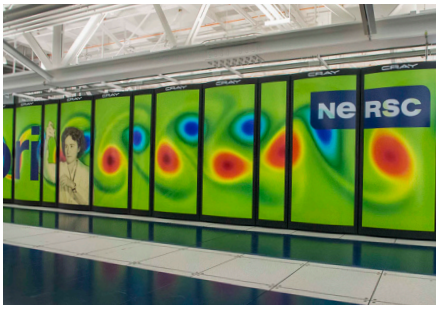
The SC community highlighted several current activities conducted by the ASCR Facilities as being of high value for engaging the community to achieve the exascale mission for science.

- Application readiness programs: ALCF's Early Science Program, OLCF's Center for Accelerated Readiness, and NERSC's Exascale Science Application Program
- Exascale Computing Project elements: applications, system software, hardware and architectures, and training and outreach
- ESnet's research and development efforts in advanced networking, network automation, and next-generation network services
- Center of Excellence collaborations
- Procurement activities (e.g., benchmarking, nonrecurring engineering development)

Opportunities for Collaboration

The SC communities discussed opportunities for collaboration in the exascale era. Many current facilities' efforts already advance these areas (highlighted above) and could be leveraged to further support:

- Creating an easy-to-use HPC environment that is integrated across ASCR centers and other work environments
- Co-evolving ASCR Facilities and DOE User Facilities including new methods of analysis and data management
- Creating new or expanding current programs that incentivize integration of SC application scientists, ASCR Facilities, and ASCR Research, which will help develop the workforce and build tools for the exascale ecosystem
- Conducting training and workforce development efforts with closer collaboration between ASCR Facilities and the science programs, including efforts aimed at DOE communities at an earlier stage of HPC capabilities



Cori – NERSC


Titan – OLCF

Theta – ALCF


ESnet Network

The Exascale Requirements Review reports and supporting materials can be found at http://exascaleage.org

DOE's HPC centers are based at Argonne National Laboratory (ALCF), Lawrence Berkeley National Laboratory (NERSC), and Oak Ridge National Laboratory (OLCF), with networking services provided by ESnet; all of these facilities operate under the direction of ASCR.

Argonne NATIONAL LABORATORY · ESnet ENERGY SCIENCES NETWORK · NERSC · OAK RIDGE National Laboratory