

High-End Computing File Systems and I/O

I. Motivation and Approach

High-end computing (HEC) has become a cornerstone of scientific discovery and an enabling technology for a variety of industries. In the last decade alone, consistent investment in software research and development in computer science has contributed to orders-of-magnitude performance improvement in parallel applications. Among the many challenges facing current and future HEC systems is the increasing role that data storage and analysis play in numerous applications of HEC, such as data-intensive computing in Internet services (e.g. Google search) and data analysis in scientific computing. It has become apparent that the current evolutionary trajectory of research and development in this area is not adequate to deliver solutions to future storage challenges. The purpose of the HEC file systems and I/O (FSIO) effort is to ensure that research and development efforts are undertaken to fill capability gaps and foster the adoption of successful research into storage products.

The HECFSIO effort has pursued this goal via two complementary activities. First, with participation from the research community, national laboratories, and industry, the HECFSIO organizing committee assembled and maintains a living national research agenda document capturing activity and gaps in critical research areas. Second, the organizing committee organizes annual workshops to bring researchers from various data technology areas together with laboratory and industry representatives in order to discuss current activities and update the research agenda. These two activities are discussed further in Sections II and III, *Research Agenda* and *Community Building*, and have provided the spark for a new HEC storage community to develop. With the assistance of the National Science Foundation providing highly visible calls for proposals and managing proposal selection and project management, research in HEC data challenges has become part of the national research agenda.

II. Research Agenda

One of the goals of the HECFSIO effort has been to shape the national research agenda for HEC data management. A decade ago, HEC data management research was being performed by an extremely small number of universities, and only in very select topics. There were obvious and damaging gaps and overlaps in the national research agenda. In 2005 the HECFSIO committee created an inter-agency high performance computing (HPC) FSIO research and development challenges document. The goal of this document was to initiate an effort to shape the research agenda. That document, "HPC File Systems and Scalable I/O: Suggested Research and Development Topics," reported the HEC data research challenges at that time.

At the inaugural HECFSIO workshop in 2005, the research needs and existing research efforts were discussed by top experts in the field, including university faculty and industry research representatives. A set of HEC FSIO research areas were formed; Metadata, Measurement and Understanding, Quality of Service, Security, Next Generation I/O Techniques, Communication Protocols, Archive, and Management and RAS.

As a service to the research community and for the participating government agencies, the HECFSIO effort created a living national research roadmap that tracked, and continues to track, progress on the problems in the eight research areas and ranked their importance. This roadmap helps convey the most important research challenges, where the field is too crowded, and where more emphasis is needed. An example portion of this roadmap follows. Roadmaps are updated each year following discussions at the annual HECFSIO workshop, with input from government, industry, and academia. Additionally, all workshops produce publicly available reports that capture the discussion guiding the roadmap updates.

Researchers use these reports as guidelines when selecting research topics in the space of HED data challenges. The roadmaps have been used by government agencies to direct their research investments in FSIO areas that

2008 Communication and Protocols Gap Area								
Area	Researchers	FY 07	FY 08	FY 09	FY 10	FY 11	FY 12	Rankings
Active Networks	Chandy	■	■	■				⊗ ⊙ ⊕
	Maccabe/Schwan							Novel work being done, but not general enough.
Alternative I/O transport schemes	Sun	■	■	■				⊗ ⊙ ⊕
	Wyckoff	■	■	■				⊗ ⊙ ⊕
	Lustre pNFS	■	■	■	■			Most aspects are being addressed.
Coherent Schemes	ANL/CMU	■	■	■				⊗ ⊙ ⊕
	UCSC's Ceph	■	■	■				⊗ ⊙ ⊕
	Lustre	■	■	■				⊗ ⊙ ⊕
	Panasas	■	■	■				⊗ ⊙ ⊕
	PVFS	■	■	■				No consensus on how to do this correctly, but some solutions are in products.

● Very Important	● Greatly Needs Research	● Greatly Needs Commercialization
⊗ Medium Importance	⊗ Needs Research	⊗ Ready and Needs Commercialization
○ Low Importance	○ Does Not Need Research	○ Not Ready for Commercialization
■ Full Calendar Year Funding	■ Partial Calendar Year Funding	■ On-Going Work

are underfunded or overlooked and shape funding solicitations. This approach to managing the national research agenda in the HEC data management area has been remarkably successful. Universities know what is relevant, industry is aware of what universities are doing currently and plan for the future, and government HEC sites gain the benefit of guiding the agenda for the national interest.

III. Community Building

The HECFSIO effort has been instrumental and extremely successful in creating a diverse community that is working cooperatively towards a shared national research agenda. The current mechanism of annual themed workshops (e.g. the 2010 theme was exascale) has created an ongoing dialogue tying everyone’s research into a shared vision. The annual refresh of the roadmaps has ensured that funded research is relevant, timely, and balanced across gap areas. These workshops have proven successful in finding interested researchers: the initial High-End Computing University Research Activity (HECURA) solicitation in 2005 received 62 proposals from 80 universities, of which 23 were funded. The second solicitation in 2009 received 109 proposals from 54 universities, of which 20 were funded.

The HECURA projects, selected on the basis of their perceived ability to enable the solution of important problems of national interest, have also been extremely successful. Researchers in the currently funded projects alone have published approximately 250 conference and journal proceedings and have contributed 13 open-source software packages. These projects fund some portion of time for over 50 professors and 100 students and have produced 16 Ph.D. and 22 Masters graduates. Twenty-seven of these graduates have brought their expertise in these problems of national interest to industry, four to government, and eleven to academia – ensuring that future generations of students will be familiar with the area and prepared to attack key challenges. Additionally, at least three commercial companies have emerged from these projects, providing valuable and supported products as a return to the community.

IV. Summary

Although HECFSIO has created a vibrant and successful research agenda, HEC data challenges are by no means solved. The exascale era looms before us, creating new sets of problems – it is clear that current HEC data system designs will not continue to scale. Data sets, growing exponentially in size, also pose new challenges in data preservation, search, and analysis. Recent advances by the architecture community and hardware vendors have introduced many new and exciting hardware components (e.g., phase change memory). It is clear, from looking at the gap between hard drives and memory, that these new hardware components will be an important part of any new HEC data solutions for the exascale. It is, however, unclear exactly how they will be used.

Continuing the annual workshops with periodic solicitations ensures that a strong community will address these important emerging challenges and that the nation will not slide back into the situation prior to 2005, in which many important challenges were unaddressed and others were overly addressed. The HECFSIO effort has been, and should continue to be, a very effective tool to guide research towards relevant and difficult challenges.

More, and more detailed, documents can be found at <http://institute.lanl.gov/hec-fsio/docs/>