Coordinating Government Investments in High-end Computing File Systems and I/O

Gary A. Grider, HPC-DO; James A. Nunez, John M. Bent, HPC-5

Million-way parallelism is around the corner and, with it, bandwidth needs to storage will go from tens of gigabytes/s to terabytes/s. Disk access times are not growing at a commensurate rate. These major challenges in file systems and I/O for high-end computing are being addressed through coordinated government agency planning designed to maximize the benefit of R&D investments.

Shortly after the deployment of Japan's Earth System Simulator Supercomputer and due to concerns with programming and optimizing applications to achieve substantial fractions of peak hardware computational power, the United States government realized that the U.S. no longer had clear leadership in the supercomputing arena. Due to this realization, federal budgets included language that required planning activities to guide future investment in high-end computing (HEC). In response, the President's Information Technology Advisory Council formed the High-End Computing Revitalization Task Force (RTF) and, in 2003, convened a workshop: The Roadmap for the Revitalization of High-End Computing. Gary Grider, from the LANL High Performance Computing Division Office, participated in this workshop and helped to craft a set of recommendations [1] to provide guidance in future government spending, and in all aspects of HEC including specification, procurement, and deployment of future HEC systems. One outcome of these recommendations was the formation of the HEC Inter-Agency Working Group (HECIWG) to develop a 5-year plan to find ways to coordinate government R&D in order to make a much more coordinated investment strategy for HEC. The HECIWG is composed of federal agencies including the Department of Commerce, the Department of Defense, the Department of Energy, the Department of Health and Human Services, the Department of Homeland Security, and the National Science Foundation. The HECIWG decided to pilot Technical Advisory Groups (TAG) in a few important R&D areas to see if coordinating government funding of HEC R&D could help the nation regain leadership in HEC. Due to a persuasive multigovernmental agency white paper [2], file systems and I/O was designated as the first technical advisory group for national focus.

Under the direction of Gary Grider, the HECIWG Technical Advisory Group for File Systems and I/O (HECIWG FSIO) was formed from a variety of government agencies, including SNL, ORNL, ANL, PNNL, NASA, and DARPA. The purpose of the HECIWG FSIO is to coordinate government spending on file systems and I/O R&D by all of the government agencies involved in the HECIWG. The mandate of the HECIWG FSIO is to list, categorize, and prioritize HEC I/O and File Systems R&D needs in particular, and to understand the gap areas that currently exist between the R&D that is being done in industry and the research that is being done in academia.

To collect a broader set of research needs in this area, the first HEC File Systems and I/O workshop was organized by Gary Grider, James Nunez of High Performance Computing Systems Integration (HPC-5) group, and others and held in August 2005 in Grapevine, TX. Government agencies, top universities in the I/O area, and commercial entities that fund file systems and I/O research were invited to help the HECIWG FSIO determine and rank the topics within the FSIO area that in most need of research. The recommended research topics were organized around these themes: metadata, measurement and understanding, quality of service, security, next-generation I/O architectures, communication and protocols, archive, and management and RAS. Additionally, university I/O center support in the forms of computing and simulation equipment availability, and availability of operational data to enable research, and HECIWG institutions involvement in the educational process were called out as areas needing assistance.

As a result of the information gathered at the 2005 workshop, the National Science Foundation (NSF) issued a call to fund research in I/O, and in file and storage systems for the high-end computing environment under the High End Computing University Research Activity (HECURA) program. In the first year, the NSF HECURA solicitation resulted in 62 proposals from over 80 universities with 19 HECURA awards made in areas ranging from autonomics and self-managing, self-healing storage systems to design theoretic approach to data placement.

Grider, Nunez, and John Bent of HPC-5 continue to help organize and set the agenda for the HECIWG FSIO workshop, held every year in August in Washington, D.C. The HECIWG FSIO workshop is very well attended by leaders in the FSIO area and a variety of government agencies. The recommendations and work done through the HECIWG FSIO have been so successful that a second NSF HECURA call focused on FSIO was released in 2009 and will continue to fund R&D in this vital HEC technical area—because without a pipeline of R&D, breakthrough concepts will not emerge. The solicitation will concentrate on areas that are still considered to be gaps and will seek to round out the overall portfolio of R&D to cover the gaps as well as seek proposals that are both evolutionary and revolutionary. A paper in the January 2009 issue of the *Operating Systems Review* [3] reviews the 2008 HECIWG FSIO identified research gaps and previews the forthcoming HECURA solicitation.

For further information contact James A. Nunez at jnunez@lanl.gov.

[1] NITRD High End Computing Revitalization Task Force (HECRTF). *Report of the Workshop on the Roadmap for the Revitalization of High-End Computing.* Daniel A. Reed, Ed. Washington, D.C., June 16-20, 2003.
[2] R. Ross et al., 2005 *HPC File Systems and Scalable I/O: Suggested Research and Development Topics for the Fiscal 2005-2009 Time Frame,* (2005) http://institutes.lanl.gov/hec-fsio/docs/FileSystems-DTS-SIO-FY05-FY09-R&D-topics-final.pdf
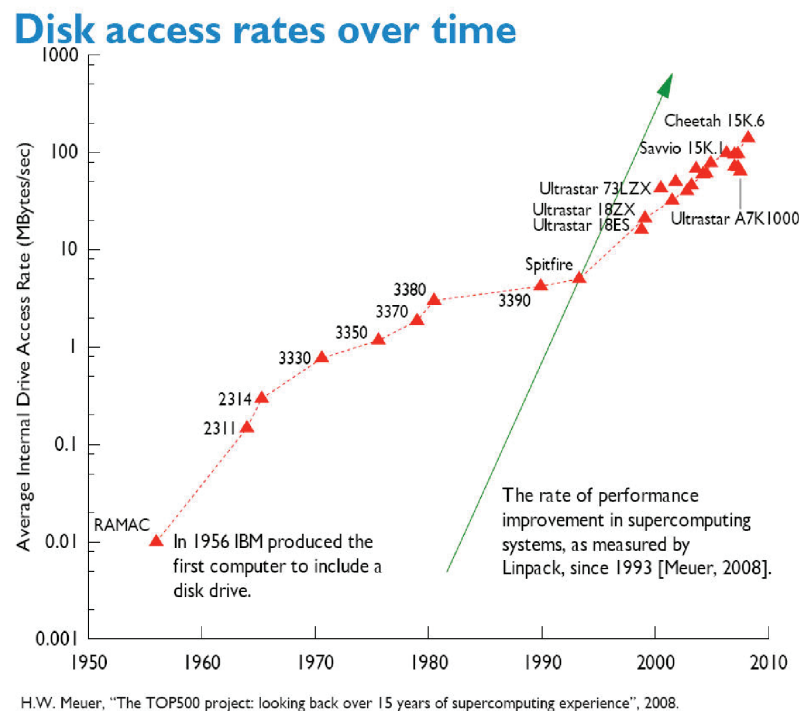[3] G.A. Grider et al., *SIGOPS Oper. Syst. Rev.* **43**:1, 2-7 (2009).

Disk access rates over time

H.W. Meuer, "The TOP500 project: looking back over 15 years of supercomputing experience", 2008.

Fig. 1. Disk access rates lag behind increase in supercomputer performance. Chart provided by Rob Ross, ANL.