SDSC readying 'Gordon' supercomputer for pre-production trials this month

August 10 2011

The San Diego Supercomputer Center (SDSC) at the University of California, San Diego, will this month launch a pre-production phase of Gordon, the first high-performance supercomputer to use large amounts of flash-based SSD (solid state drive) memory.

The installation of 64 I/O nodes, joined via an InfiniBand switched fabric communications link, is being completed early this month and will be made available to U.S. academic researchers who want to run large-scale database applications, said SDSC Director Michael Norman. University of California academic researchers interested in early access to Gordon I/O nodes are directed to request a Dash startup allocation from the XSEDE website at www.xsede.org/

Norman made the announcement at the 'Get Ready for Gordon – Summer Institute' being held this week (August 8-11) at SDSC. The four-day workshop is designed to familiarize potential users with the unique capabilities of Gordon, the result of a five-year, $20 million award from the National Science Foundation (NSF). Production startup is set for January 1, 2012.

"This year is the start of academic data-intensive supercomputing," Norman said in opening the conference, and encouraged researchers engaged in data-intensive science and data mining across a diverse range of disciplines to apply for allocations.

With about 300 trillion bytes of flash memory and those 64 I/O nodes,

Gordon will be capable of handling massive data bases while providing up to 100 times faster speeds when compared to hard drive disk systems for some queries. Flash memory is more common in smaller devices such as mobile phones and laptop computers, but unique for supercomputers, which generally use slower spinning disk technology.

"One might ask why we haven't seen flash memory in HPC (high performance computing) systems before," said Norman. "Smaller flash systems for consumers have become less expensive, but they have not been durable enough for HPC applications. Now we have enterprise MLC (multi-level cell), and it's available at both attractive prices and with very good durability (or write endurance), which is achieved by over-provisioning and wear leveling."

Gordon is being configured to aid researchers in conducting data-intensive computational science, such as visual analytics or interaction network analyses for new drug discovery, or the solution of inverse problems – converting observed measurements into information about a physical object or system – in oceanography, atmospheric science, and oil exploration.

Gordon also is suited for large-scale data mining applications, such as de novo gene assembly, or for cosmological applications, or doing modestly scalable codes in quantum chemistry or structural engineering.

Along with the well-publicized exponential growth in digitally based data, there is an explosion in the amount of research based or scientific data, noted Norman. "Data of this size is simply becoming unmanageable for analysis, so there is an urgent need for supercomputers like Gordon," said Norman, adding that "it is part of our genetics here at SDSC to do data-intensive computing."