NSF awards $12 million to SDSC to deploy 'Comet' supercomputer

October 4 2013

The San Diego Supercomputer Center (SDSC) at the University of California, San Diego, has been awarded a $12-million grant from the National Science Foundation (NSF) to deploy *Comet*, a new petascale supercomputer designed to transform advanced scientific computing by expanding access and capacity among traditional as well as non-traditional research domains. *Comet* will be capable of an overall peak performance of nearly two petaflops, or two quadrillion operations per second.

"Supercomputers such as *Comet* and our data-intensive *Gordon* system are helping to fulfill the NSF's goal to extend the impact of advanced computational resources to a larger and more diversified user base," said UC San Diego Chancellor Pradeep K. Khosla. "Our San Diego Supercomputer Center is a key resource for our university system and has had a long track-record of leadership in high-performance computers and data-intensive computing."

While science domains such as physics, astronomy, and the earth sciences have long relied on at-scale high-performance computing (HPC) to help them create detailed simulations to accelerate discovery, there is a growing need for computing capacity for a broader set of researchers, including those in non-traditional domains such as genomics, the social sciences, and economics.

Computing for the 99 Percent

"*Comet* is designed to be part of an emerging cyberinfrastructure for what is called the 'long tail' of science, which encompasses the idea that a large number of modest-sized computationally based research projects still represents, in aggregate, a tremendous amount of research and scientific impact," said Sandra A. Brown, Vice Chancellor for Research at UC San Diego.

"*Comet* is all about computing for the 99 percent," said SDSC Director Michael Norman, the project's principal investigator. "As the world's first virtualized HPC cluster, it is designed to deliver a significantly increased level of computing capacity and customizability to support data-enabled science and engineering at the campus, regional, and national levels, and in turn support the entire science and engineering enterprise, including education as well as research."

Comet will join SDSC's *Gordon* supercomputer as a key resource within NSF's Extreme Science and Engineering Discovery Environment (XSEDE), which comprises the most advanced collection of integrated digital resources and services in the world. It is expected that *Comet* will help meet the pent-up demand for computing on up to 1,024 cores, which accounts for 98% of current jobs among XSEDE users. While *Comet* will be able to support much larger jobs, its scheduling policies will be designed to provide fast turnaround for large numbers of smaller jobs.

Comet will also be the first XSEDE production system to support high-performance virtualization. SDSC team members plan to work closely with communities and enable them to develop the customized software stacks that meet their needs by defining virtual clusters. With significant advances in Single Root IO Virtualization (SRIOV), virtual clusters will be able to attain near native hardware performance in both InfiniBand latency and bandwidth, making them suitable for MPI-style parallel computing.

"We are supporting *Comet* to provide a resource not just for the highest end-users, but for scientists and engineers across a broad spectrum of disciplines," said Barry Schneider, program director for *Comet* in NSF's Division of Advanced Cyberinfrastructure. "This so-called long tail of science is discovering the power of advanced digital resources. In this way, *Comet* complements other NSF resources such as *Blue Waters* and *Stampede*, which were designed primarily to provide power users with the ability to perform large-scale computations."

Petascale Power

Scheduled to start operations in early 2015, *Comet* will be a Dell-based cluster based on next-generation Intel Xeon processors. Each node will be equipped with two of those processors, 128 GB (gigabytes) of traditional DRAM, and 320 GB of flash memory. Since *Comet* is designed to optimize capacity for modest-scale jobs, each rack of 72 nodes will have a full bisection InfiniBand FDR interconnect, with a 4:1 bisection interconnect across the racks.

"Dell was ecstatic to partner with SDSC on this particular National Science Foundation proposal," said Tim Carroll, Executive Director of Dell Research Computing Solutions. "UC San Diego has a rich history of providing breakthrough technology to the broadest base of researchers possible. At Dell we are committed to the very same mission, so the collaboration was intuitive. We are proud to deliver this solution to the research community and to be part of the next generation of breakthroughs."

"When you look inside *Comet*, the key enabling computing capability will be the Intel Xeon processors," said Mark Seager, Chief Technology Officer for the High Performance Computing Ecosystem at Intel. "These new processors will deliver significant performance improvements to the NSF's general-purpose scientific workload, or the 99 percent, with a

robust set of virtualization features that enable a broad spectrum of high-performance applications in a large memory, virtualized environment."

In addition, *Comet* will include some large-memory nodes, each with 1.5 TB of memory, as well as nodes with NVIDIA GPUs (graphic processing units). The GPU and large-memory nodes will target specific applications, such as visualization, molecular dynamics simulations or *de novo* genome assembly.

Comet users will also have access to 7 PB (petabytes) of Lustre-based high-performance storage, as well as 6 PB of durable storage for data reliability, both based on an evolution of SDSC's *Data Oasis* storage system. UC San Diego and SDSC are also deploying new 100 Gbps (Gigabit per second) connectivity, allowing users to rapidly move data to SDSC for analysis and data sharing, and return data to their institutions for local use.

Comet will be the successor to SDSC's *Trestles* computer cluster, to be retired in 2014 after four years of service.

"*Comet* will have all of the features that made *Trestles* so popular with users, but with much more capacity and ease-of-access," said SDSC Deputy Director Richard Moore, a co-PI of the *Comet* project. "*Comet* will be particularly well-suited to science gateways that serve large communities of users, especially those new to XSEDE."

Norman and Moore are joined by three co-principal investigators from SDSC on the *Comet* project: SDSC Associate Director and XSEDE co-PI Nancy Wilkins-Diehr; SDSC Distinguished Scientist Chaitan Baru; and SDSC Chief Technical Officer Philip Papadopoulos. Geoffrey Fox, Distinguished Professor of Computer Science and Informatics at Indiana University and PI of the NSF's FutureGrid project, is a strategic partner for the project.

The *Comet* project is funded under NSF grant number ACI 1341698.