

UC San Diego

UC San Diego News Center

February 02, 2017 | By Jan Zverina

SDSC's 'Comet' Supercomputer Surpasses '10,000 Users' Milestone



Image: Ben Tolo, SDSC

Comet, the petascale supercomputer at the San Diego Supercomputer Center (SDSC), an Organized Research Unit of UC San Diego, has easily surpassed its target of serving at least 10,000 researchers across a diverse range of science disciplines, from astrophysics to redrawing the “tree of life”.

In fact, about 15,000 users have used *Comet* to run science gateways jobs alone since the system went into production less than two years ago. A science gateway is a community-developed set of tools, applications, and data services and collections that are integrated through a web-based portal or suite of applications. Another 2,600 users have accessed the high-performance computing (HPC) resource via traditional runs. The target was established by SDSC as part of its cooperative agreement with the National Science Foundation (NSF), which awarded funding for *Comet* in late 2013.

“*Comet* was designed to meet the needs of what is often referred to as the ‘long tail’ of science – the idea that the large number of modest-sized computationally-based research projects represent, in aggregate, a tremendous amount of research that can yield scientific advances and discovery,” said SDSC Director Michael Norman, principal investigator for the *Comet* project.

Comet, which went into operation in mid-2015, has been one of the most widely used supercomputers in the NSF’s XSEDE (Extreme Science and Engineering Discovery Environment) program, which provides researchers with an advanced collection of integrated digital resources and services.

SDSC will hold a webinar on February 15 to provide a detailed overview of *Comet*’s capabilities and system upgrades. First-time and current users are invited to attend. More information about the webinar [can be found here](#).

“Now in its second year of serving the national research community, *Comet* is exceeding our expectations and we encourage new users to learn more about how *Comet* can support their research,” said Norman. “Feedback from our current user base – both anecdotally and through their expressed use on the system, as well as examining the data we’ve been collecting – underscores a strong need for systems such as *Comet* that serve what we call the ‘99 percent’ of the research community.”

In addition to *Comet*’s design, its allocation and operational policies are geared toward rapid access, quick turnaround, and an overall focus on scientific productivity. *Comet* also features large memory nodes, GPUs and local flash, which taken together, provide a highly usable and flexible computing environment for a wide range of domains.

Providing 'Science Gateways' for Researchers

Surpassing the 10,000-user milestone in less than two years of operations is due in large part to researchers accessing *Comet* via science gateways, which provide scientists with access to many of the tools used in cutting-edge research – telescopes, seismic shake tables, supercomputers, sky surveys, undersea sensors, and more – and connect often diverse resources in easily accessible ways that save researchers and institutions time and money.

Science gateways make it possible to run the available applications on supercomputers such as *Comet* so results come quickly, even with large data sets. Moreover, browser access offered by gateways allows researchers to focus on their scientific problem without having to learn the details of how supercomputers work and how to access and organize the data needed.

In mid-2016, a collaborative team led by SDSC was awarded a five-year \$15 million NSF grant to establish a Science Gateways Community Institute to accelerate the development and application of highly functional, sustainable science gateways that address the needs of researchers across the full spectrum of NSF directorates. The award was part of a larger NSF announcement in which the agency committed \$35 million to create two Scientific Software Innovation Institutes (S2I2) that will serve as long-term hubs for scientific software development, maintenance and education.

“It’s possible to support gateways across many disciplines because of the variety of hardware and support for complex, customized software environments on *Comet*,” said Nancy Wilkins-Diehr, an associate director of SDSC and co-director of XSEDE’s Extended Collaborative Support Services. “This is a great benefit to researchers who value the ease of use of high-end resources via such gateways.”

One of the most popular science gateways across the entire XSEDE resource portfolio is the CIPRES science gateway, created as a portal under the NSF-funded Cyberinfrastructure for Phylogenetic Research (CIPRES) project in late 2009. The gateway is used by scientists to explore evolutionary relationships by comparing DNA sequence information between species.

In 2013, SDSC received a \$1.5 million NSF award to extend the project to make supercomputer access simpler and more flexible for phylogenetics researchers. Typically, about 200 CIPRES jobs are running simultaneously on *Comet*.

“The scheduling policy on *Comet* allows us to make big gains in efficiency because we can use anywhere between one and 24 cores on each node,” said Mark Miller, a bioinformatics researcher with SDSC and principal investigator of the CIPRES gateway. “When you are

running 200 small jobs 24/7, those savings really add up in a hurry.”

To date, the CIPRES science gateway has supported more than 20,000 users conducting phylogenetic studies involving species in every branch of the “[tree of life](#)”. The gateway is used by researchers on six continents, and their results have appeared in more than 3,000 scientific publications since 2010, including *Cell*, *Nature*, and *PNAS*.

I-TASSER: A New Protein Structure Gateway Available via Comet

In late 2016, a new science gateway called [I-TASSER](#) (Iterative Threading ASSEmbly Refinement), developed by researchers at the [Zhang Lab](#) at the University of Michigan’s Medical School, began accepting users. I-TASSER is a hierarchical approach to protein structure and function prediction. Structural templates are first identified from the [Protein Data Bank](#) using LOMETS (Local Meta-Threading-Server), an on-line web service for protein structure prediction. Full-length atomic models are then constructed by iterative template fragment assembly simulations. Finally, function insights of the target are derived by threading the 3D models through the protein function database called [BioLiP](#).

Since October 2016, I-TASSER has been accessed via *Comet* – the only resource within the XSEDE portfolio to do so – by more than 8,000 unique users, according to Yang Zhang, a U-M professor of computational medicine and bioinformatics as well as biological chemistry, and the I-TASSER’s principal investigator. In total, I-TASSER currently has more than 76,000 registered users from 130 countries.

“With the increasing requests from the community for protein structure and function modeling, one of the major bottlenecks of the I-TASSER Server has been the limit in supporting computer resources of our laboratory that was originally funded by the Department of Computational Medicine and Bioinformatics at the University of Michigan,” said Zhang. “The generous grant of computing resources from XSEDE is very helpful in improving the capacity of the I-TASSER system to serve the broader biomedical community by providing faster and higher quality simulations of protein models.”

Including I-TASSER, [29 science gateways are available](#) via XSEDE’s resources, each one designed to address the computational needs of a particular community such as computational chemistry, phylogenetics, and the neurosciences. SDSC alone has delivered 77 percent of all gateway cycles since the start of the XSEDE project in 2011.

Jan Zverina, 858-534-5111, jzverina@sdsc.edu

UC San Diego's [Studio Ten 300](#) offers radio and television connections for media interviews with our faculty, which can be coordinated via studio@ucsd.edu. To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit <https://ucsdnews.ucsd.edu/media-resources/faculty-experts>.