January 21, 2015 | By Jan Zverina

UC San Diego Physicist Frank Würthwein Joins SDSC

Open Science Grid Pioneer to Focus on High-Throughput Computing UC-wide

University of California, San Diego Professor Frank Würthwein, an expert in high-energy particle physics and advanced computation, has joined the university's San Diego Supercomputer Center (SDSC) to help implement a high-capacity data cyberinfrastructure across all UC campuses.

Würthwein, who joined UC San Diego as a physics professor in 2003, was recently named executive director of the Open Science Grid (OSG) project, a multi-disciplinary research partnership funded by the U.S. Department of Energy and the National Science Foundation. He was OSG's founding executive during 2005. His appointment to SDSC is effective this month.

Frank Würthwein. Photo: Ben Tolo SDSC.

Würthwein is no stranger to processing extremely large data sets. In 2013, he and his team used SDSC's data-intensive *Gordon* supercomputer to provide auxiliary computing capacity to OSG by processing massive data sets generated by the Compact Muon Solenoid (CMS), one of two large general-purpose particle detectors at the Large Hadron Collider (LHC) near CERN, Switzerland.

The LHC is being used by researchers to first find, and now study in detail, the elusive Higgs particle. Details of that project, which was one of *Gordon's* most data-intensive assignments to-date, can be viewed here. *Gordon* is part of the NSF's XSEDE (eXtreme Science and

Engineering Discovery Environment) program, under which academic researchers can request computer time on a full range of systems.

"I am delighted that Frank has joined us here at SDSC," said SDSC Director Michael Norman. "Frank's appointment will help create a seamless interface between the nation's two leading open scientific computing infrastructures – OSG and XSEDE – which will directly benefit a broad spectrum of researchers. In addition, Frank's expertise in developing, deploying, and operating a worldwide distributed high-throughput computing cyberinfrastructure paves the way for him to pioneer a shared data and compute platform across the entire UC system anchored at SDSC. Frank's appointment is just one of many ways we are committed to strengthening our UC engagement efforts."

One of the key benefits of such a compute and storage network is that individual PIs across all UC campuses will have direct access to SDSC's expertise and resources from their home institutions. Added Norman: "We view this network as a key solution and enabler for data-enabled research, and we can see the day when other university systems and research enterprises follow suit with similar systems."

While Würthwein's primary physics interest is searching for new phenomena at the high energy frontier with the CMS detector at the LHC, as an experimentalist he also is interested in instrumentation and data analysis. His research group has focused on software development, integration, and operations of services that implement the central core of globally federated systems. Through his appointment, Würthwein's group will be more integrated with SDSC's other research groups and Data-Enabled Scientific Computing (DESC) division.

"Core to the global federation is the notion that policy and control regarding resource access is ultimately defined and controlled at the local level, by local resource owners," said Würthwein. "However, mechanisms for delegating trust on a global scale allow for global sharing of resources via the creation of an overarching cyberinfrastructure that also allows for distributed, domain science-specific, shared virtual infrastructures."

XSEDE Meta-Scheduler

Würthwein's group currently operates a meta-scheduler that aggregates CPU resources across more than 100 computer and storage clusters worldwide. It serves about a dozen different communities, one of which functions as a service provider for the NSF's XSEDE program. Würthwein has served on XSEDE's User Advisory Committee since 2012.

"Moreover, research communities may attach the meta-scheduler to their local clusters, and leave it up to the individual scientists to specify whether their application is self-contained to the point that it can on-ramp onto the nationally distributed Open Science Grid facility," said Würthwein.

To solve the data input problem, Würthwein's group and collaborators also created a global data federation that presently comprises about 100 petabytes of storage capacity across dozens of clusters worldwide. The federation implements a global namespace, and files can be opened via an X509 authenticated WAN access protocol. A mix of partial file reads plus intelligent data prefetching makes WAN reads not just possible but practical.

"We now measure the largest data volumes in petabytes, with one PB equaling one quadrillion bytes of information," said Würthwein. "In as little as five years from now, we expect them to be measured in exabytes, with one EB equal to one quintillion bytes."

Prior to joining UC San Diego, Würthwein was a professor at the Massachusetts Institute of Technology (MIT), and a senior research associate with the California Institute of Technology. He received his Ph.D. in Physics from Cornell University in 1995.

About SDSC

As an Organized Research Unit of UC San Diego, SDSC is considered a leader in data-intensive computing and cyberinfrastructure, providing resources, services, and expertise to the national research community, including industry and academia. Cyberinfrastructure refers to an accessible, integrated network of computer-based resources and expertise, focused on accelerating scientific inquiry and discovery. SDSC supports hundreds of multidisciplinary programs spanning a wide variety of domains, from earth sciences and biology to astrophysics, bioinformatics, and health IT. SDSC is a partner in XSEDE (eXtreme Science and Engineering Discovery Environment), the most advanced collection of integrated digital resources and services in the world.

MEDIA CONTACT

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.