Pacific Research Platform (PRP)
PI: Larry Smarr, Calit2, UC San Diego
Co-PIs: Camille Crittenden, CITRIS & the Banatao Institute, UC Berkeley; Tom DeFanti, Calit2, UC San Diego; Phil Papadopoulos, UC San Diego; Frank Wuerthwein, UC San Diego
NSF Project #ACI-1541349

The PRP integrates campus Science DMZs, an architecture developed by the U.S. Department of Energy's Energy Sciences Network (ESnet), into a high-capacity regional "freeway system." This system makes it possible for large amounts of scientific data to be moved between scientists' labs and their collaborators' sites, supercomputer centers or data repositories, without performance degradation.

1. What is the most significant challenge encountered in your DIBBs project and how did you overcome it?

The major goals of the PRP can be divided into two main categories: 1) Science Engagement and 2) Technology Development and Implementation. Science engagement activities are aimed at recruiting science participants, working with science teams to determine data needs, and sharing acquired knowledge with the scientific community. Technology development and implementation efforts focus on building and managing the actual network, enabling unfettered data transfer and interoperability among data clusters and instruments within the PRP, while also monitoring connections from outside the PRP.

To date, challenges in the network architecture have largely been overcome through the dedication of a growing group of network engineers who meet regularly, by phone and in person, to identify and work through data transfer bottlenecks at the local and regional scale. Challenges in science engagement have included finding qualified staff to spearhead the effort on a day-to-day basis, a leader or small team who can contact likely users and understanding their applications, then collaborate with the network engineers to improve scientific workflow. The science engagement group meets bi-weekly and has made modest progress toward this end. But although we have had some success by convening domain-specific user groups to discuss needs and opportunities, we are still exploring sustainable models to overcome this challenge (some of this work falls into a percentage of several people's jobs, including some hired under CC*DIBBS as cyberinfrastructure engineers, but it's unclear how to recruit and train new professionals into this role full time).

2. What future challenges do you envision and are there sustainability issues or other barriers to success?

Technical challenges: These include staying abreast of advances in data science, computer science (including machine vision, neural networks, and other machine learning advances in hardware and software), and networking technology, software-defined networking, Open Science Grid, etc.

Science engagement challenges: creating and sustaining an emerging professional field, incorporating these positions in a sustainable way into campus CIO/Research IT staffing infrastructures, and making faculty and researchers aware of the resource.

Crossing these two areas, a challenge remains in identifying data transfer bottlenecks--where in the process, end-to-end, is the data getting stuck or slowing dramatically? Ideally, interdisciplinary cyberinfrastructure teams could be deployed to diagnose the problem and improve data workflows. Sustaining the PRP (and scaling it beyond the original participating institutions) will require advocacy for the improved capacity from faculty and affiliated researchers, buy-in from senior campus/organizational leadership, and cross-disciplinary staffing to support the vision and its execution.

We also want to provide equitable access to the benefits of this technology. Partnering with CENIC, which is developing relationships with organizations throughout the state of California to deliver high-speed broadband to metropolitan areas as well as underserved communities in the Central Valley, will contribute to this goal.