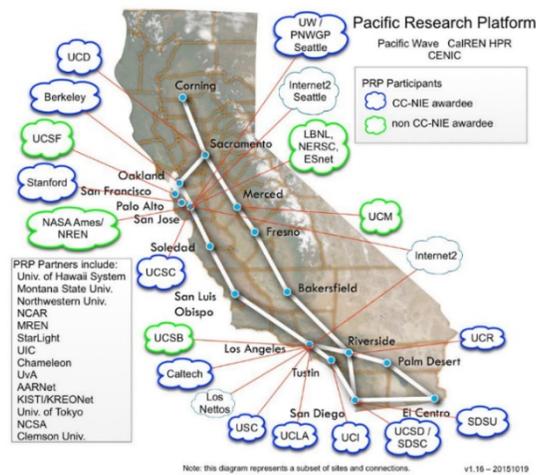


# Pacific Research Platform Version 2 (PRPv2) Workshop

## Summary of Workshop Presentations *Richard Moore and Tom DeFanti*



**Workshop held February 21, 2017**

**Hosted by Calit2's Qualcomm Institute (Calit2/QI) at  
University of California San Diego (UCSD)**

This workshop was supported by NSF award ACI-1541349.

Additional support is from the

University of California Office of the President (UCOP), the Corporation for  
Education Network Initiatives in California (CENIC), and UC San Diego Calit2

# Table of Contents

|   |    |
|---|----|
| Preface .....   | 3  |
| Agenda .....  | 3  |
| Workshop Context and Objectives.....  | 4  |
| Session Notes .....   | 5  |
| Welcome and PRP Two Years In – Larry Smarr (UCSD) .....   | 5  |
| Session 1: Fast and Secure Data Transfer Panel – Moderator: Tom DeFanti (UCSD).....                           | 6  |
| 1.1 Data Placement with HTCondor – Phil Papadopoulos (UCSD/SDSC).....   | 6  |
| 1.2 IPv6 and CENIC/ESnet/I2 Goals – John Hess (CENIC) and Tom Hutton (UCSD/SDSC) .....                        | 7  |
| 1.3 Fast Data Transfer: SC16 results, etc. – Azher Mughal (Caltech) .....                                     | 7  |
| 1.4 Trusted Platform Modules – John Graham (UCSD/Calit2/QI) .....   | 8  |
| 1.5 Session 1 Panel Q&A .....   | 8  |
| Session 2: Science Engagement Future Panel – Moderator: Frank Würthwein (UCSD).....                           | 9  |
| 2.1 PRP Weather Data Transfer – Scott Sellars (UCSD/SIO) .....  | 9  |
| 2.2 PRP Engagement Foci – Camille Crittenden (UCB/CITRIS) & Tom DeFanti (UCSD/Calit2/QI) ..                   | 9  |
| 2.3 PRP from the UCSC Point of View – Shawfeng Dong (UCSC).....   | 10 |
| 2.4 PRP from the UC Merced Point of View – Jeff Weekley (UCM).....  | 10 |
| 2.5 Session 2 Panel Q&A .....   | 11 |
| Session 3: Network Architecture for Globalization of PRPv2 Panel – Moderator: Cees de Laat (UvA)..            | 12 |
| 3.1 Cooperative Work Groups in PRPv2 – Leon Gommans (Air France/KLM).....                                     | 12 |
| 3.2 Science DMZs at KISTI – Jeonghoon Moon (KISTI).....   | 12 |
| 3.3 Fast Data Transfer: Brazil – Michael Stanton (RNP, Brazil).....   | 13 |
| 3.4 Observations on Expanding the PRP – Harvey Newman (Caltech).....  | 13 |
| 3.5 Session 3 Panel Q&A .....   | 13 |
| Session 4: SDN/SDX in PRPv2 Panel – Moderator: Tom Hutton (UCSD/SDSC).....                                    | 14 |
| 4.1 SDN in PRPv2 for secure flows – Phil Papadopoulos (UCSD/SDSC) .....                                       | 14 |
| 4.2 100Gb Services for Data-Intensive Science Enabled by SDN/SDX – Joe Mambretti<br>(Northwestern Univ) ..... | 14 |
| 4.3 SDN/SDX in CENIC/Pacific Wave – John Hess (CENIC).....  | 14 |
| 4.4 Session 4 Panel Q&A .....   | 15 |

## Preface

This Summary of Workshop Presentations provides a brief overview of the workshop, including notes from the talks and panel question-and-answer sessions. It is not intended to reproduce or summarize the information in the speakers' presentations, which are available at [pacificresearchplatform.org](http://pacificresearchplatform.org). The workshop video-recording is also available there as well.

Recommendations and findings from the workshop will be incorporated into the planning for the remainder of the PRP project, and may be added to this report at a later time.

## Agenda

### **10:00 Welcome and PRP Two Years In – Larry Smarr (UCSD)**

#### **Session 1: Fast and Secure Data Transfer Panel – Moderator: Tom DeFanti (UCSD)**

- 10:30 1.1 Data Placement with HTCondor – Phil Papadopoulos (UCSD/SDSC)
- 10:42 1.2 IPv6 and CENIC/ESnet/I2 Goals – John Hess (CENIC) and Tom Hutton (UCSD/SDSC)
- 10:54 1.3 Fast Data Transfer: SC16 results, etc. – Azher Mughal (Caltech)
- 11:06 1.4 Trusted Platform Modules – John Graham (UCSD/Calit2/QI)
- 11:18 1.5 Q&A

### **11:45 Lunch, with optional walk to see Library CAVEkiosk**

#### **Session 2: Science Engagement Future Panel – Moderator: Frank Würthwein (UCSD)**

- 1:45 2.1 PRP Weather Data Transfer – Scott Sellars (UCSD/SIO)
- 1:55 2.2 PRP Engagement Focii – Camille Crittenden (UCB) and Tom DeFanti (UCSD/Calit2/QI)
- 2:05 2.3 PRP from the UCSC Point of View – Shawfeng Dong (UCSC)
- 2:15 2.4 PRP from the UC Merced Point of View – Jeff Weekley (UCM)
- 2:25 2.5 Q&A

#### **Session 3: Network Architecture for Globalization of PRPv2 Panel – Moderator: Cees de Laat (Uva)**

- 3:00 3.1 Cooperative Work Groups in PRPv2 – Leon Gommans (Air France/KLM)
- 3:12 3.2 DMZs at KISTI – Jeonghoon Moon (KISTI)
- 3:24 3.3 Fast Data Transfer: Brazil – Michael Stanton (RNP)
- 3:36 3.4 Observations on Expanding the PRP – Harvey Newman (Caltech)
- 3:48 3.5 Q&A

### **4:00 Break**

#### **Session 4: SDN/SDX in PRPv2 Panel – Moderator: Tom Hutton (UCSD)**

- 4:30 4.1 SDN in PRPv2 for secure flows – Phil Papadopoulos (UCSD/SDSC)
- 4:45 4.2 100Gb Services for Data-Intensive Science Enabled by SDN/SDX – Joe Mambretti (Northwestern U)
- 5:00 4.3 SDN/SDX in CENIC/Pacific Wave – John Hess (UCSD/Calit2/QI)
- 5:15 4.4 Q&A

### **5:30 Reception**

## Workshop Context and Objectives

The Pacific Research Platform (PRP) is an ambitious project driven by the high-speed networking needs of collaborative, big-data science. Many research disciplines are increasingly multi-investigator and multi-institutional and need ever more rapid access to their ultra-large heterogeneous and widely distributed datasets. In response to this challenge, the Department of Energy's ESnet developed the Science DMZ model, a network system optimized for high-performance scientific applications rather than for general-purpose or enterprise computing. The PRP extends the campus Science DMZ model, which is widely funded on individual campuses by NSF's CC-NIE and CC-NII programs, to a regional model for data-intensive networking. It enables researchers to quickly and easily move data between collaborator labs, supercomputer centers, and data repositories, creating a big-data freeway that allows the data to traverse multiple, heterogeneous networks without performance degradation. The PRP's data sharing architecture, with end-to-end 10–100Gb/s connections, also enables region-wide virtual co-location of data with computing.

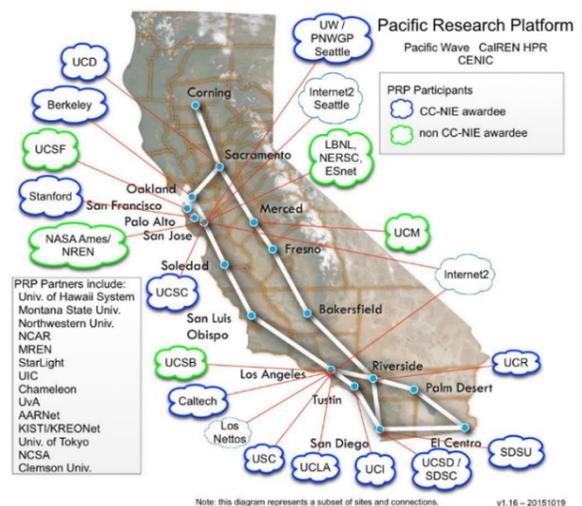
The main focus of the PRP project is to build a researcher-defined and data-focused network whose requirements are driven by direct engagements with sophisticated, cyberinfrastructure-knowledgeable Science Teams chosen from the fields of particle physics, astronomy, biomedical sciences, earth sciences, and scalable data visualization. It is a partnership of more than 20 institutions, including four National Science Foundation, Department of Energy, and NASA supercomputer centers.

A PRP workshop was held in October 2015, shortly after the NSF award was initiated, to bring together representatives from all PRP partner institutions, participating Science Team members, technical staff, and network architects and implementers to discuss deployment of the PRP and address interoperability of Science DMZs at a regional level (<http://prp.ucsd.edu/presentations/2015-prp-workshop-report>). The primary objectives of that workshop were:

- To engage the PRP team and collaborators and facilitate in-person interactions.
- Gain a common understanding of the science engagement process.
- Explore and capture data storage, computing, and networking requirements across five scientific domains identified as initial applications for the PRP.
- Identify common science-driven technical requirements for the PRP.
- Exchange technical ideas for the PRP's technical implementation.

The technical design and build-out of the PRP project is being conducted in two phases – PRP Version 1 in years 1-2 and PRP Version 2 in years 3-5 – with the aim of achieving the following goals:

- PRPv1: Create a scalable network design for optimizing data transfer.
- PRPv2: Evolve to IPv6 with Cooperating Research Groups.
- PRPv2: Create OpenFlow as a firewall and explore other trust and security features.



Considerable progress has been made during the first 17 months of the project in deploying PRPv1 and successfully applying that system to a number of scientific applications. The current workshop is focused on the transition for the PRP project from Version 1 to Version 2. The PRPv1 capabilities will not be replaced by PRPv2, but rather PRPv2 is an exploration of means to improve PRPv1 and possibly address the increasing challenges of security from the campuses' point of view. The PRPv2 goals are not set in stone. This workshop's goal was meant to offer an open discussion with alpha users, long-term national and international partners in networking and data transfer, and the implementers of the PRPv1 and PRPv2. This workshop's purpose is to debate and discuss the goals for PRPv2.

The PRPv2 workshop was held at UCSD/Calit2 on February 22, 2017, in conjunction with the ON\*VECTOR International Photonics workshop February 23-24. The PRPv2 workshop was attended by 48 people, including 12 international attendees representing five countries (Brazil, Czech Republic, Korea, Netherlands and Japan).

## Session Notes

The following section is based on contemporaneous notes from the talks and question-and-answer sessions. These notes are not intended to be comprehensive, nor to reproduce or summarize the speakers' presentation materials, which are available as source material at [pacificresearchplatform.org](http://pacificresearchplatform.org). The workshop video-recording is also available there as well. These notes have not been reviewed by the participants/speakers, and as such they may inadvertently misrepresent the speakers' statements or intents; please contact the participants directly if you have any questions.

### Welcome and PRP Two Years In – Larry Smarr (UCSD)

The foundation for the PRP project started with ESnet developing the Science DMZ and Data Transfer Node (DTN) concepts.

PRP's DTN boxes, the 10-100Gb/s Flash I/O Network Appliances (FIONAs), are a critical enabler for the project. Recently the team has developed a less expensive version, FIONette, for ~1 Gbps sites.

The project monitors data transfer performance across the entire the system 4 times per day. John Graham, PRP chief engineer, has played a critical role in the development/deployment of FIONAs and the monitoring capabilities.

The PRP project would not have been possible without the long-term substantial investments in CENIC. John Hess has played a critical role representing CENIC within the PRP project. CENIC hooks up campus gateways in a reliable fashion. PRP goes from campus network gateway to the researcher on campus. There has historically been a gap in this connectivity that the PRP is trying to fill – the campus CIO is responsible for the campus network, but often the department has responsibility for the researcher's connectivity to that network.

There have been great strides already in PRPv1 to establish, use, and monitor the high-speed end-to-end connectivity between partner sites. (See the color matrices of end-point connectivity performance from January 2016 to December 2016 in the presentation, as well as real-time MaDDash monitoring results at

<http://prp-maddash.calit2.optiputer.net/maddash-webui/> ). More campuses/end-points have been added since the early demonstrations of the project, and there is improved bandwidth performance between those sites. These advances are a tribute to the collective, sustained activities – mostly on a volunteer basis in regards to this NSF grant - by the networking people at these sites.

There are already a number of science application successes during the first 1.5 years of the project. Existing teams include the high-energy physics group at Caltech (Harvey Newman), Shawfeng Dong from UCSC connecting their campus Hyades cluster to NERSC/LBNL and getting data from the Dark Energy Spectroscopic Instrument (located at Kitt Peak ~800GB/night), the virtual reality visualization capability at UC Merced (Jeff Weekley), and the PEER earthquake project led out of UC Berkeley. Some of these will be discussed in this workshop.

And there are new science use cases that were not part of the original proposal, including connecting the microscope off Scripps Pier (Jules Jaffe UCSD/SIO), regional downscaling of climate models (Dan Cayan/Mike Dettinger at UCSD/SIO), expansion of HPWREN (Frank Vernon UCSD), and Scott Sellars and Marty Ralph from the Center for Western Weather and Water Extremes (UCSD/SIO) using the system (see later talk). There is a possible expansion to include the Ocean Observatory Initiative in 2017, connecting sensors on the sea floor off the state of Washington.

The work to date has also spurred new opportunities that couldn't have previously been done, with PRP providing high-speed access to emerging technologies. For example, there are new capabilities based on non-Von Neumann architectures – TrueNorth by IBM, neural processors by KnuEdge, field-programmable gate array technologies, or ARM processors like Qualcomm's Snapdragon – that PRP can facilitate access to. A proposal has been developed to support machine learning capabilities, using a cloud of GPU-heavy FIONA boxes in a Condor-managed cloud. And the concept is expanding to international capabilities, led by some of the workshop's international participants.

This talk has reviewed the past work – i.e. PRP V1. This workshop focuses on PRP V2 – looking ahead to SDN/SDX, IPv6, and other technologies.

We should thank the PRP partners – in Missouri terms, the team has been 'raising a barn,' based largely on partners' volunteer labor contributing to these goals.

In response to a question about sustained funding, Larry indicated that he has talked with program officers at NSF and the funding prospects are uncertain both at an agency level and at the Office of Advanced CyberInfrastructure (ACI). There is some reorganization within the Office of ACI – NSF is looking for a new Office director, and there are priority debates within the office across supercomputers, data efforts, software and networking. He hopes there will be ongoing funding opportunities that would leverage the large investments NSF has already made, and notes that Internet2 has issued white papers citing networking's role in major infrastructure projects for the country.

## Session 1: Fast and Secure Data Transfer Panel – Moderator: Tom DeFanti (UCSD)

### 1.1 Data Placement with HTCondor – Phil Papadopoulos (UCSD/SDSC)

The talk discusses a project that is related to PRP with Dr. Miron Livny at U Wisconsin-Madison and other partners at Beihang University and the Computer Network Information Center (CNIC) in China.

PRP is often focused on raw performance (bandwidth). This project extends the metrics to reliability and access to data. More security issues are involved when touching disks, and is network measurement performance a good proxy for disk-to-disk performance? This project measures end-to-end and disk-to-disk performance among a set of international endpoints using different data-movement protocols and IPv4 and IPv6.

There are a lot of known network traversals, but it's tricky to know what the actual paths are.

A 'test manifest' is submitted as a routine cron job to a Condor pool, with HTCondor handling errors/recovery/reporting/iteration. A limited Condor pool is built across sites, with nodes needing to trust each other 'enough.' Servers are set up and torn down in user space for experiments.

See performance slides in the presentation. Many of the end-point/protocol results show substantial variability in performance, which is not necessarily understood at this point. The testing did indicate that iRODS performance is ~10% of the raw network performance.

Commonly-used components include: HTCondor, Graphite, Carbon and Whisper database (open-sourced from Orbitz), Python >2.6.x, Git; see slides for websites. All components have V4 and V6 implementations.

## 1.2 IPv6 and CENIC/ESnet/I2 Goals – John Hess (CENIC) and Tom Hutton (UCSD/SDSC)

PRPv1 is not unique in terms of networking architecture.

There is a general concern on campuses about Science DMZs bypassing campus firewalls. We don't want to inadvertently create insecure paths. Also the project and campuses want to ensure that the PRP network is appropriately used for research purposes only.

What's envisioned for PRPv2? PRPv2 will reflect a different strategy, using Border Gateway Protocol (BGP) communities for tagging classes of DMZ networks. The plan is to have a pilot implementation across six sites (UCSD, SDSC, Caltech, NCSA, U Chicago ... and one Northern California site). Pilot BGP peering will be native IPv6 only. A stretch goal is to incorporate SDN/SDX type signaling for paths or super-channels.

Take the ~20 sites in PRPv1 and take ~6 PRPv2 pilot sites. Use route servers located at Pacific Wave sites. Tag routes for this purpose. Dynamic signaling. Data plane will traverse existing paths.

Proposing IPv6 as a mechanism to deal with establishing specific routes connecting groups – generally there is IPv6 address space available. And one of the objectives here is to accelerate adoption of IPv6.

ESNet has four public DTNs that are available for read-only transfers. PRP may want to consider a similar approach.

Tstat is a way to instrument data transfers (think of it as Netflow++). They would like to get some major centers to incorporate tstat, but it won't be everyone. And would like to be able to generate reports on the various connections via tstat.

Tom DeFanti clarifies that PRPv1 is not going away, but rather will persist through the next phase of the project, with PRPv2 being a parallel testbed.

## 1.3 Fast Data Transfer: SC16 results, etc. – Azher Mughal (Caltech)

Talk focuses on demonstrations conducted at SC'16 and plans for SC'17. See presentation materials for goals of demonstrations.

The team had a few 1 Tpbs links across the show floor (Caltech, StarLight, SCinet), more 100 Gbs links .

SC16 demos across CENIC: PRP based on top of the CENIC network backbone. One of the objectives is to 'energize' the science teams so that researchers can see what can be achieved.

1 Tbps booth-to-booth transfers:

- Looked at various NVMe drives for high-speed transfers. Built a low-cost DTN server that could drive ~100 Gbps. Maximum throughput at 14 drives (7 drives/processor).

Need processor with more PCIe lanes in order to get higher speeds. Intel SkyLake?

Have achieved transmission across 4 Mellanox VPI NICs – close to 400 Gbps.

SC17 goals

- East West integration with other controllers along with state, recovery, provisioning, monitoring
- Demonstrating SENSE project for DTN auto tuning (SENSE = SDN for end-to-end networked science at exascale)
- NVMe over Fabrics across the WAN
- DTN design using 200G NICs (Mellanox/Chelsio)

#### 1.4 Trusted Platform Modules – John Graham (UCSD/Calit2/QI)

Trusted Platform Modules (TPMs) offer facilities for the secure generation of cryptographic keys, and limitation of their use, in addition to a random number generator. TPMs are low-cost devices, allowing systems to interact securely.

Have successfully implemented UC-Jupyter on SDSC's Comet HPC system using TPMs. JupyterHub authenticates a user with CILogon and spawns kernels on Comet. Trusted Platform Modules (TPM) on the JupyterHub FIONAs secure the keys from the CILogon member organization. These keys are used to connect jupyter.calit2.optiputer.net to comet.sdsc.edu.

Globus and XSEDE have CILogon authentication, so approach can extend to many systems/users.

How do you scale access to Jupyter notebooks campus-wide? Berkeley has been doing it. Doing a pilot now with Prof Coleman at UCSD.

Developing next-generation GPU Jupyter box. Dual 8-core CPUs, 8 GPUs/box, 6x480GB SSDs, 2 NVMe bays, ~\$13K.

Distributed Trusted Computing. Using Kubernetes, a spinoff of Google infrastructure tools; see slides for list of features. Take container, sign container with TPM, then no one can tamper with anything in container.

KubeSpawner (jupyterhub-kubernetes-spawner). Can spawn single-user notebook servers on a Kubernetes cluster – multiple entities resident on same hardware.

#### 1.5 Session 1 Panel Q&A

There was a question for John Graham about the SSDs and cooling issues in the next-gen GPU box. John indicated there are 6 SSD drives and 2 NVMe bays. There are eight GPUs and four fans. John notes that the unit has been ordered, but not run yet – it will be interesting to see how cool and how loud the box is. It will be interesting to evaluate the cooling, but they have had similar boxes and expect it to work. Azher Mughal notes that they have had to look at high-speed fans for their boxes.

In response to a question on liquid-cooled options, John Graham replies that he has seen some liquid-cooled GPUs, but low-profile units are required to fit in the box.

There was a question for Azher Mughal about whether special configurations were required for high-speed transfers. Yes, the systems had to be configured and tuned to achieve high-speeds.

There was a question for Phil Papadopoulos why the setup for IPv6 was more painful than IPv4 in their testing. The primary issues were limitations on the UCSD campus. Because there is not that much experience with IPv6 on campus, there are not very many IPv6 services set up yet. For example, it took a while to get required network info, routes are static. This should improve with more experience.

Cees de Laat notes that the IPv6 services available to him locally in Netherlands are pretty robust. Phil was jealous.

Tom Hutton notes that because IPv6 is optional rather than required for many people, it keeps getting deferred. In contrast, Phil's Chinese collaborators only had IPv6 address space, so they had to get up to speed. The tentative plan at UCSD is probably to set up a sub-domain within the DMZ Optiputer network that's IPv6.

George Peek asked a general question – from a PRP perspective, what's the biggest headache for IPv6? Tom Hutton and Phil Papadopoulos replied 'Applications.' John Hess noted that sys admins may be reluctant to set up rule sets for IPv6.

Joe Mambretti asked John Hess about IPv4 and IPv6 issues for various institutions in PRPv2. John replies that among exercises with NCSA or U Chicago, the collection of routes have great resources but we're only interested in the subset of routes that map to addresses for Science DMZ and high-performance resources. Most large institutions may have routes that are both v4 and v6. Want to make network IPv6 friendly enough that researchers don't need to worry about IPv4 or IPv6 configurations – i.e. users initiate request to move data and the network works. The technical end is often easier, while the science engagement piece can be more challenging.

## Session 2: Science Engagement Future Panel – Moderator: Frank Würthwein (UCSD)

Frank Würthwein posed three questions for the speakers in this session – noting that the speakers have not been informed of these questions in advance of their talks!

- How can PRPv2 add value to your work?
- Is PRPv2 going to serve IT or science needs, or both?
  - Are you concerned that IT and science needs could be at odds with each other in PRPv2?
- How can we make change least disruptive to existing PRPv1 science use?

### 2.1 PRP Weather Data Transfer – Scott Sellars (UCSD/SIO)

His group is handling large climatology datasets (see presentation for list).

They had a recent workflow to identify features in water transfer models – took about 20 days for a 4-step process with much of that being network data transfer time.

They have worked with PRP and others to expedite this workflow and improve their scientific productivity. The download speed increased ~4X using PRP network and FIONAs. They have also brought in Kepler workflow software, got connected to SDSC's Comet system, and are using Jupyter.

### 2.2 PRP Engagement Foci – Camille Crittenden (UCB/CITRIS) & Tom DeFanti (UCSD/Calit2/QI)

PRP will be participating in a number of workshops:

- Add-on to annual CENIC conference March 20, 2017.
- Big Data and Earth Sciences May 30-June 2 with Scott Sellars.
- In planning: Cryo-EM, machine learning, HPWREN.
- There is some additional participant support funding available in the grant.

There are a number of advantages of participating in these workshops, including exposure and potential experience to student participants.

The PRP also convenes regular PRP networking telcons – while some partners participate regularly, all partners have access to notices and minutes of these meetings.

The project would like to do better at an institutional level (compared to networking level) and in cracking the nut of science engagement across campuses. PRP wants more faculty, postdocs and staff on board.

Science engagement efforts do better when you can highlight science opportunities, e.g. in context of workshops. Engagements need person-to-person contact, and need ‘evangelists’ (beyond Larry Smarr) to talk with people. The project tried to recruit someone to fill this role (e.g. a postdoc) and ran into obstacles. Evangelism remains a challenge for the project.

Workshop participants are encouraged to reach out to scientists on their campuses. Networking problems are tough – and are getting attention of talented people; still need help on the social engineering side.

How does the project team know it’s being successful – when emails about problems stop? That’s an important measure, but there are broader issues beyond that metric. Generally IT people on a campus are in a different realm than professors and research staff. It’s important to introduce faculty to the IT staff that can help them.

### 2.3 PRP from the UCSC Point of View – Shawfeng Dong (UCSC)

He works in ‘Research and Faculty Partnerships (RFP),’ a division in the campus IT organization that is responsible for working with faculty/researchers at UCSC. He himself is an astrophysicist, with expertise in HPC simulations for his own research. It helps him in his role as interface to faculty that he has a strong research background.

See presentation for a slide on related Recent Awards at UCSC – they include a CC-NIE \$500K award for Science DMZ and a \$400K CC-DNI award for CI Engineer (funding him).

Their partners across campus: Computational Astrophysics, Astronomy, Particle Physics, Genomics, Chemistry, Biology, Data Science. Dominated by astrophysics, but are getting other people on campus engaged.

### 2.4 PRP from the UC Merced Point of View – Jeff Weekley (UCM)

He is the Director of Cyberinfrastructure and Research Computing at UC Merced.

One of his objectives is to elevate UCM to a first-class participant in PRP.

They are building cyberinfrastructure at UC Merced:

- Built a WAVE. Interestingly, in the School of Humanities and the Arts.
- Recipient of CC-\* award for Science DMZ. Intend to support IPv6.
- Campus is planning to double its size in Project 2020 - \$1.3B construction project. Should be able to build networking in the right way.

- Planning wireless/microwave links outside the campus.
- Maybe HPWREN 2.0 – Yosemite-Monterey Bay. National park service, UC, State of California, other agencies.

UC Merced has faculty that have significant networking needs in a number of pioneering applications.

Another objective is to build Human Capacity. They are removing roadblocks, repeating workflows, fostering touchstone projects across UC campuses, partnering with CITRIS and Calit2, and uncovering ‘force-multipliers.’

## 2.5 Session 2 Panel Q&A

Scott Sellars was asked about the computing environment for analysis he’s doing on his data. At this point it is primarily PC-level hardware with tools such as Matlab and Python. They may move to central HPC resources to speed up the analysis steps. He was also asked about the total reduction in time from the original 20-day workflow. The networking improvements reduced it to ~10 days. They may alter algorithms and move to central computing to reduce it further.

Tom DeFanti was asked about PRP demonstrations planned, in addition to workshops, as a means to engage with science groups. Tom notes that there have been some key demonstrations, including the efforts at the annual Supercomputing conferences.

The panel was asked about more specific ideas for workshop participants to better engage faculty, including efforts that would scale to large numbers of faculty. Camille Crittenden highlighted the concept of force multipliers – e.g. researchers that are involved in research IT efforts already. Speaking at conferences is good for achieving scale. Another idea is to ask researchers that have used PRP to acknowledge it in publications, just as they would grant funding. Tom DeFanti notes that one approach is to pick someone in a sector and focus resources on them to ‘make them famous.’ Then that person’s colleagues will hear what can be achieved by leveraging something like PRP. Jeff Weekley comments that he pitches PRP and big data movers and Science DMZ – no matter what the topic when he’s talking with faculty. Also, if he can’t get to faculty, he gets to grad students. Camille Crittenden noted that a good time to engage is when new faculty members are just joining the university. Shawfeng Dong commented that their engagement has been helped by the fact that his boss is an adjunct faculty member, and Shawfeng personally has been part of research groups.

Later in the discussion Harvey Newman noted that part of science engagement is to show people what others are able to achieve compared to what they’ve been doing ... and hence how much time they’re wasting.

Tim Lance asked about engagement/evangelization via ‘listening versus talking.’ Jeff Weekley commented that he listens to what people need and doesn’t build what isn’t needed. Tom DeFanti noted that he talked with researchers in machine learning in order to come up with idea for the hardware to support it.

The session moderator Frank Würthwein was asked about his progress in establishing high-speed end-to-end connectivity for LHC high-energy physics researchers across UC and other campuses. His team has put boxes at ~5 campuses (Davis, Irvine, Riverside San Diego, soon Santa Barbara) and they are all connected at 10 Gbps. They have benchmarked the infrastructure and just recently submitted a research paper on the project. UC Irvine is probably making the most science use of the capability right now. In terms of lessons learned, he had thought it would be ‘cookie cutter’ deployments, but found out that all the cutters and cookies were a little different. It was also harder to get adoption than expected -

there is another step between getting working equipment at a site, and training the local researchers how to operate and use it.

Shawfeng Dong and Jeff Weekley were asked to what extent multi-campus or multi-nodal collaborations (within campus) are important to research on their campuses. Shawfeng Dong replied that the LHC collaboration was important to campus effort, and the PRP effort really made a difference in reducing significant bottlenecks. Jeff Weekley characterized multi-campus collaborations as an example of a force multiplier that he had noted in his talk. He has made investments that enable campus->edge capabilities and that will make a difference. He's seen examples of other campuses that just have awful networking connectivity and it's a real barrier to collaboration/research.

The panel was asked about using public DTNs that are close to POPs (similar to what ESNNet has done). The general reply was that DTNs at POPS are a good idea, and that it would be an excellent idea for there to be centralized data repositories with high performance. One participant noted that users still need to be informed how to discover and access those assets.

Scott Sellars was asked whether his primary challenge is getting the data in the first place or where it goes after he gets it. Scott replied that it's a challenge that the central datasets are so large, since one has to make decisions in advance which portion of the data to download. If the infrastructure existed to easily pull data down entire datasets and then prune data, he would prefer that approach.

### Session 3: Network Architecture for Globalization of PRPv2 Panel – Moderator: Cees de Laat (UvA)

#### 3.1 Cooperative Work Groups in PRPv2 – Leon Gommans (Air France/KLM)

'Trusted Sharing of Big Data Assets in Cooperative Working Groups.'

There are barriers for business organizations to share data – e.g. would others that you're sharing with get more value relative to your organization, intellectual property and legal concerns, and privacy protections.

'Secure Digital Market Place' concept allows community members to share data according to well-defined market rules.

Using high-speed bandwidth of the PRP to enable in-memory analyses of shared data, without storing the shared data, would help alleviate some of the data-sharing concerns. Could also utilize 'containers' on that data within the shared space.

Wants to develop a 'future internet field lab' (see slides).

#### 3.2 Science DMZs at KISTI – Jeonghoon Moon (KISTI)

The speaker notes that DMZ has a very real – and non-scientific - meaning in Korea.

Have an SDN-based Data Transfer Node and a Layer 3-based DTN (FIONA box). The latter was used to join the SC'16 demonstration described earlier by Azher Mughal, and this node was just added to the PRP 'MaDDash' matrix of end-points.

They have achieved 9.2 Gbps (FTP) and 9.1 (GridFTP) on 10Gbps line from Seoul to Daejeon. Daejeon to PRP universities in California achieves 7-8Gbps both directions with GridFTP.

One of their science use cases is AgroMeteorology Federation (AMF) in Korea, storing climate data and analyzing/sharing results.

Doing research to develop an “OpenScienceDMZ Cloud” with DTN/DMZ as a Service (DDaaS). They are also working on a project to develop a “SaaS OverCloud on DTN.”

### 3.3 Fast Data Transfer: Brazil – Michael Stanton (RNP, Brazil)

RNP is Brazil’s Research and Education Network. The RNP’s national backbone connects capitals of each state and federal capital (26 capitals and 13 other cities); most links are at 10G, except in the Amazon region. May move some of these to 100G in the next year+. There are also shared links to the US via ANSP and AmLight EXP, Europe and Latin America via RedClara, and terrestrial links with Argentina. There is a concentration of universities/research centers/population in Sao Paulo/Rio de Janeiro regions.

RNP has participated in SC demonstrations since 2004. In SC’16 (with Caltech group), RNP demonstrated >95% usage of the 100G access to US, via ANSP.

It’s important to talk about using data, not just moving it (and not just demos). A Brazilian e-astronomy lab collaborates with NCSA and JHU/Sloan data via high-speed links. When they started, they got 20-30 Mbps on a 1 Gbps line, and it took two weeks to transfer a 5.7 TB collection. Now they get ~1 Gbps using GridFTP (and a DTN box essentially at the POP). And there is a new project to enable efficient remote access to LNCC’s supercomputer (in Petropolis). Potential clients are a light synchrotron sources and a Weather/Climate Lab. This capability may save cost of purchasing a separate computer for the synchrotron.

He notes that he has seen interesting work here on high-speed DTNs that he will be taking back to Brazil.

### 3.4 Observations on Expanding the PRP – Harvey Newman (Caltech)

This is the time to expand horizons, to think in the large. There are programs on the horizon that will be even bigger than LHC – e.g. LSST and SKA.

Historically, ESnet traffic has grown ~10X over 4 years. Much of ESNet traffic is LHC data. In past year alone with LHC Run 2, LHC data growth ~2.7X.

On the other hand, we have not been in a period of high technological advancement. We’re in 4<sup>th</sup> - 5<sup>th</sup> year of ~100 Gbps technology. Next technology step coming is ~400 Gb/s. That rate of technology improvement is not keeping pace with traffic growth rates.

SC16 demo – up to 1 Tbps local, many 100Gbps connections.

There is a major opportunity for a new CPU/storage/network ecosystem + Large Computing Facilities (LCFs) as focal points.

SENSE effort: “SDN for end-to-end networked science at exascale” with partners ESNet, Caltech, Fermilab, Argonne, Maryland and LBNL.

PRP has become a locus for activity in this area.

### 3.5 Session 3 Panel Q&A

The panel was asked about future major drivers. Harvey Newman provided examples such as LSST, genomics, and the Internet of Things.

There was a question about the latency of the path to Korea in the SC16 demonstration - it approached 400 msec round-trip.

For many workflows, data in a repository is shipped somewhere for analysis, and then results are returned to the researcher. The question was posed - what are the computing environments at the

endpoints? Harvey Newman replied regarding the LHC computing environment – e.g. ~20 Tier 1 sites, ~150 Tier 2 sites (each with several petabytes storage and several thousand job slots). Data transfers are typically ~2TB; if you could move 10s of TBs efficiently, that would significantly increase productivity.

There was a question about the Brazilian Amazon rivers program, that was primarily deferred to the following ON\*VECTOR workshop. Michael Stanton said that they have completed ~250km the past year, and expect to set up next 400km within next 4-5 months.

Leon Gommans was asked for additional information on airline data sharing. Leon cited the example that when an aircraft lands, there is ~500GB of data available per flight. With a fleet of aircraft over a long period of time, there could be exabytes of data theoretically available. Can this data be shared effectively?

## Session 4: SDN/SDX in PRPv2 Panel – Moderator: Tom Hutton (UCSD/SDSC)

### 4.1 SDN in PRPv2 for secure flows – Phil Papadopoulos (UCSD/SDSC)

The talk addresses using an OpenFlow Switch as a low-impact firewall.

In Linux, every packet is looked at, meaning there is overhead on every packet. But Linux can't make decisions based on other 'meta' data.

Purpose here is to use an OpenFlow switch to make decisions based on flows, reducing overhead per packet. But lose observability of flow from controller. Limited flow-table space. And what happens if controller goes offline – network goes away.

They tested how much delay could be introduced in initiating rules for the flow without impacting network performance: 0.5 sec (a lot!) was almost no impact, while 5 sec crippled network. So it looks reasonable to use an OpenFlow switch as a firewall.

There are additional possibilities with IPv6.

### 4.2 100Gb Services for Data-Intensive Science Enabled by SDN/SDX – Joe Mambretti (Northwestern Univ)

'OpenFlow is dead.' Orchestration is critical. Highly programmable networks. Lots of important work coming out of GENI.

SDX's are in US, but where else? Some at SURFNet, one soon in Germany, one other in Europe. Will be in Taiwan, Singapore, Korea, Tokyo.

There were many demos at SC16, most dependent on SDN techniques.

He likes mdtmFTP compared to GridFTP.

Wants to move towards the Global Research Platform – a worldwide PRP.

### 4.3 SDN/SDX in CENIC/Pacific Wave – John Hess (CENIC)

Pacific Wave is a joint collaboration CENIC and Pacific Northwest Gigapop (PNWGP), supporting both commercial and R&E peers, serving countries across the Pacific connecting to the western US.

Pacific Wave received an NSF/ACI IRNC award September 2015 to support more 100G connections and to deploy SDN/SDX on parallel infrastructure to enable experimentation while maintaining production use of the Pacific Wave exchange. An SDX node is being established in Los Angeles, and an existing SDX node in Seattle is being leveraged; control nodes are located in Seattle and Los Angeles.

The group is participating in the worldwide 'AutoGOLE/NSI Pilot' project. (GOLE = GLIF Open Lightpath Exchanges, GLIF = Global Lambda Integrated Facility.)

#### 4.4 Session 4 Panel Q&A

John Graham noted that the Globus features for sharing data, cited in one of the charts, cost money compared to the basic GridFTP data-movement capabilities. John Hess replied that yes, Globus has a licensing model that enables some additional features like multiple streams, data sharing, and third-party authentication. ESnet has a network-wide license for Globus, and Globus has allowed them to use this license for the pilot project. They hope to get a similar arrangement ... but not clear yet how far it can be extended.

Phil Papadopoulos was asked whether the OpenFlow switch-as-firewall could be used for all data. He replied that it's the 'nice flows' (not the short housekeeping flows) that would go thru this system.

John Hess was asked how new institutions can join the SDN/SDX effort. He indicated that if an institution is interested, best to start the process by talking with him.