Toward the Global Research Platform

PACIFIC RESEARCH PLATFORM

Nadya Williams
nwilliams@ucsd.edu
University of California, San Diego

Keynote Presentations given in Singapore March 27, 2018 and 2nd NRP workshop in Bozeman, MT, August 6-7, 2018

Dr. Tom DeFanti
Research Scientist
California Institute for Telecommunications and Information Technology's
Qualcomm Institute
University of California San Diego
Distinguished Professor Emeritus, University of Illinois at Chicago

What is PRP

- The NSF funded The Pacific Research Platform (PRP) to the UCSD for 5 years starting October 1, 2015.
- It emerged out of the unmet **demand for high-performing bandwidth** to connect data generators and data consumers.
- The PRP is in its 3rd year of successfully bringing new, unanticipated science applications, as well as test new means to dramatically improve throughput.
- The PRP was scaled to be a regional program by design, mainly focusing on West Coast US institutions, although it now includes several long-distance US and transoceanic Global Lambda Integrated Facility (GLIF) partners.
- There is demand from the high-performance networking and scientific communities to extend the PRP nationally, and indeed worldwide.

The goal: to prototype a future in which a fully-funded multi-national Global Research Platform emerges

PRP PACIFIC RESEARCH PLATFORM

it²

Science DMZ

- **Based on Community Input and on ESnet's Science DMZ Concept, NSF Has Funded Over 100 Campuses to Build DMZs**
- **A Science DMZ integrates 4 key concepts into a unified whole:**
  - **A network architecture designed for high-performance applications**, with the science network distinct from the general-purpose network
  - The use of dedicated systems as **Data Transfer Nodes** (DTNs)
  - **Performance measurement and network testing systems** are regularly used to characterize and troubleshoot the network
  - **Security policies and enforcement mechanisms** are tailored for high performance science environments

Science DMZ
Coined 2010

PRP
PACIFIC RESEARCH PLATFORM

ESnet
FASTERDATA KNOWLEDGEBASE

http://fasterdata.es.net/science-dmz/

it²

Providing Needed Performance

R&E networks must meet performance needs of a diverse mix of users & applications

That includes everything from:

- **Low latency for high volume interactive K-12 student testing**
- **Through interactive video performance**
- **To "big data" flows among labs**

Reliability, flexibility and resiliency are essential

That's been hard for networks to do

- **Internet2 has tried multiple times via end-2-end performance projects (e.g. the 'user expectations' efforts), as have others.**
- **And, there are many network performance tools, and they do lots of useful things**
- **But tools used by most of R&E don't do direct, active, deterministic, measurement of actual traffic between sites**

PRP PACIFIC RESEARCH PLATFORM

it²

With NSF, CENIC & PNWGP Support

- **Active measurement is a good way to get, and to help sustaining true end-to-end user performance**
- **Based on ESnet's knowledge base fasterdata.es.net**
- **Supported by the large NSF funded 'Science DMZ' project called *Pacific Research Platform* (PRP)**
- **And by a new one called *Towards the NRP***

ESnet:
Energy Sciences Network

CENIC
The Corporation for Education Network Initiatives

PACIFIC NORTHWEST GIGAPOP

PRP
PACIFIC RESEARCH PLATFORM

it²

The Pacific Research Platform Networks Connects Campuses to Create a Regional End-to-End *Big Data Superhighway*

Source: John Hess, CENIC

NSF Grant
10/2015-10/2020

PI:

- Larry Smarr, UC San Diego Calit2

Co-PIs:

- Camille Crittenden, UC Berkeley CITRIS,
- Tom DeFanti, UC San Diego Calit2/QI,
- Philip Papadopoulos, UCI (former UCSD)
- Frank Wuerthwein, UCSD Physics and SDSC

Letters of Commitment from:

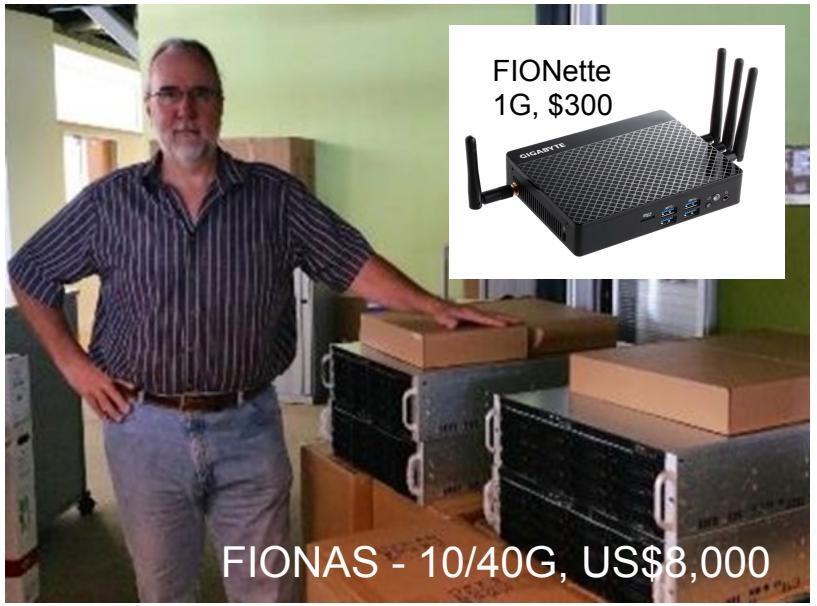
- 50 Researchers from 15 Campuses
- 32 IT/Network Organization Leaders

SDSC CITRIS AND THE BANATAO INSTITUTE it²

Big Data Science Data Transfer Nodes (DTNs) Flash I/O Network Appliances (FIONAs)

Key Innovation: UCSD Designed FIONAs to solve the Disk-to-Disk data transfer problem at full speed on 10/40/100G Networks

- FIONAs PCs [a.k.a ESnet DTNs]:
  - **~$8,000 Big Data PC with:**
    - 10/40 Gbps Network Interface Cards
    - 3 TB SSDs or 100+ TB Disk Drive
  - **Extensible for Higher Performance to:**
    - +NVMe SSDs for 100Gbps Disk-to-Disk
    - +Up to 8 GPUs [4M GPU Core Hours/Week]
    - +Up to 196 TB Disks as Data *Capacitors*
    - +Up to 38 Intel CPU cores or AMD Epyc
  - **~$1,100 10Gbps FIONAs now available**
- FIONettes are $300 EL-30-based FIONAs
  - **1Gbps NIC With USB-3 for Flash Storage or SSD**



Phil Papadopoulos, SDSC
Tom DeFanti, Calit2
Joe Keefe, Calit2
John Graham, Calit2

PACIFIC RESEARCH PLATFORM

it²

FIONAs on the PRP and Partners sites

~40 FIONAs as GridFTP (MaDDash) + perfSONAR Systems

- **PRP Partners: 10 UCs, Caltech, U of Southern California, U Washington, U of Illinois Chicago, SDSC**
- **Plus U Utah, Montana State, U Chicago, Clemson U, U Hawaii, NCAR, Guam**
- **Plus Internationals: U Amsterdam, KISTI (Korea), Singapore (soon)**

Many States and Regionals Building FIONAs and Creating MaDDashes

- **FIONA Build Specs on PRP Website http://prp.ucsd.edu**
- **Weekly Engineering Calls with Notes Going to 60+ Technical Participants**

Monitoring and Debugging Dashboard
PERFormance Service Oriented Network monitoring ARchitecture

PRP PACIFIC RESEARCH PLATFORM
CENIC ESnet SDSC CITRIS AND THE BANATAO INSTITUTE NSF it²

Performance Instrumentation

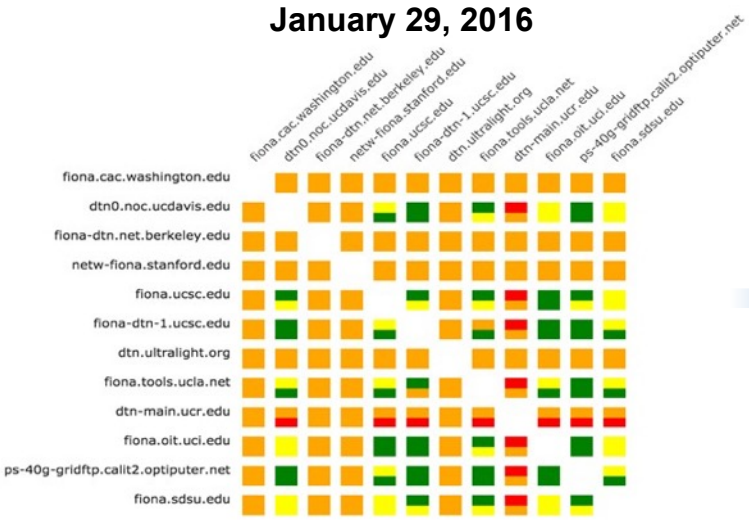
- We use purpose-built ‘FIONA’ servers that are tuned to test end-to-end 1G, 10G, 40G and 100G connections, our version of ESnet’s DTNs
- perfSONAR and GridFTP logs are then turned into visualizations
- Disk-to-disk transfers of 10GB were performed 4x a day until the PRP networks and the campus endpoints were tuned and capable of full bandwidth utilization with essentially no TCP backoff
- New efforts are expanding the types of testing and visualizations, using Kubernetes as an orchestration engine to automate this distributed cluster of FIONAs

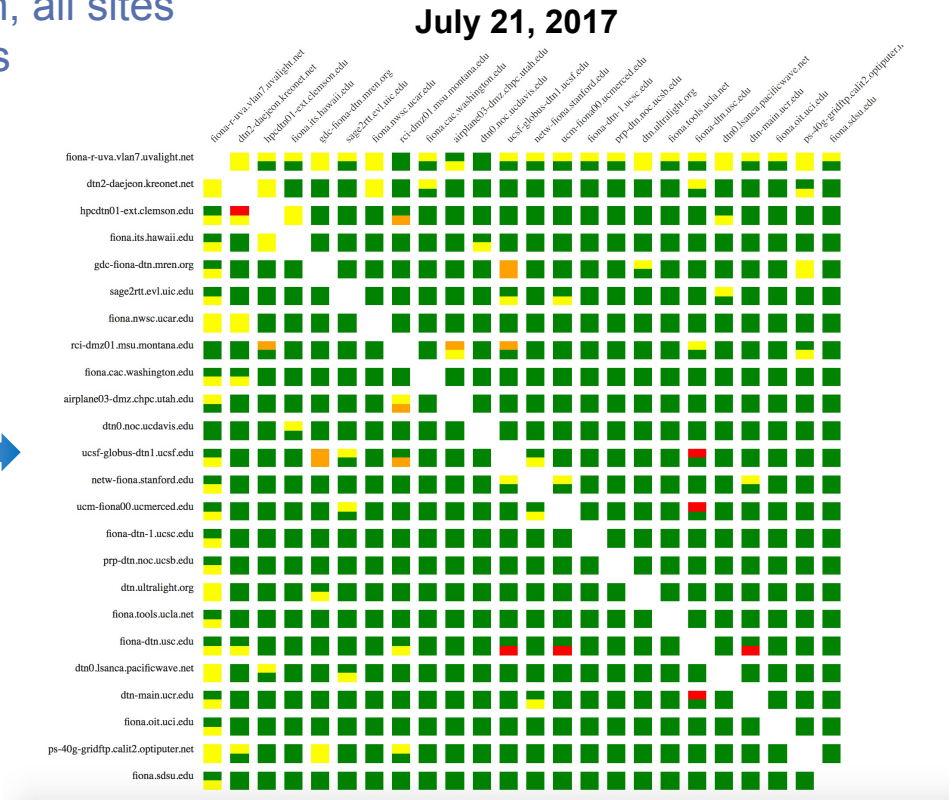
PRP PACIFIC RESEARCH PLATFORM

it²

Disk-to-Disk Throughput measurement testing

- 10Gb file transfer 4 times/day dual direction, all sites
- From monitoring start: 12 DTNs to 24 DTNs connected at 10-40G in 1.5 yrs





Source: John Graham, Calit2/QI

PRPGridFTP

Throughput >= 5000Mbps | Throughput < 5000Mbps | Throughput <= 1000Mbps | Unable to retrieve data

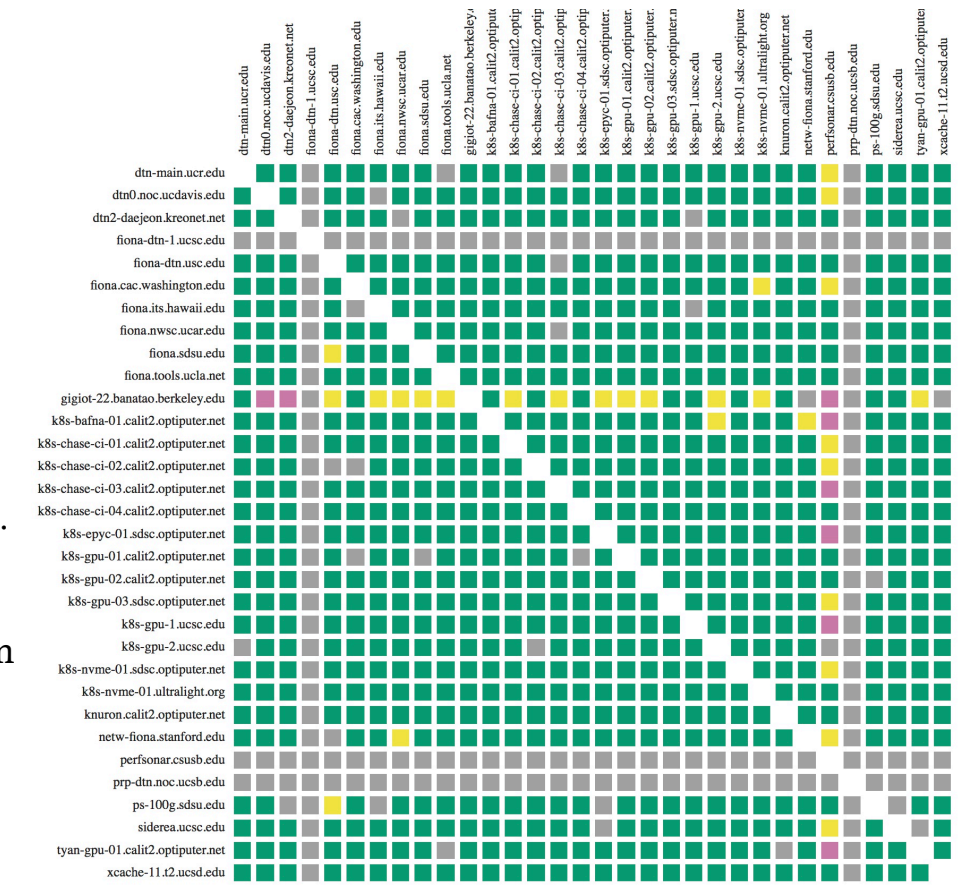
Nautilus mesh: traceroute measurement testing

Shows flows that took different paths to the same destination

- **Number of paths is <=1**
- **Number of paths is >=1**
- **Number of paths is >=2**
- **Unable to retrieve data**

Traceroute - network diagnostic tool for
- displaying the route (path)
- and measuring transit delays of packets across an Internet Protocol (IP) network.

The history of the route is recorded as the round-trip times of the packets received from each successive host (remote node) in the route (path)

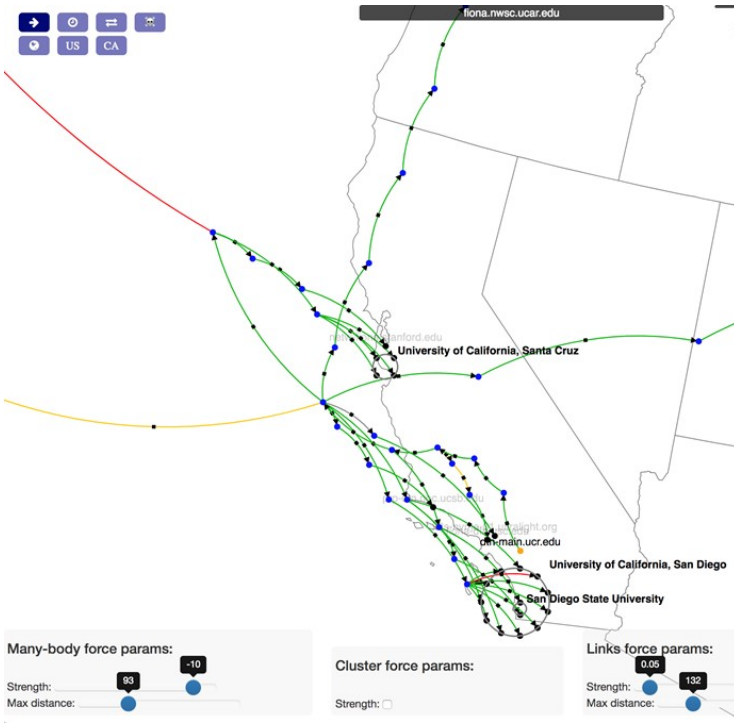


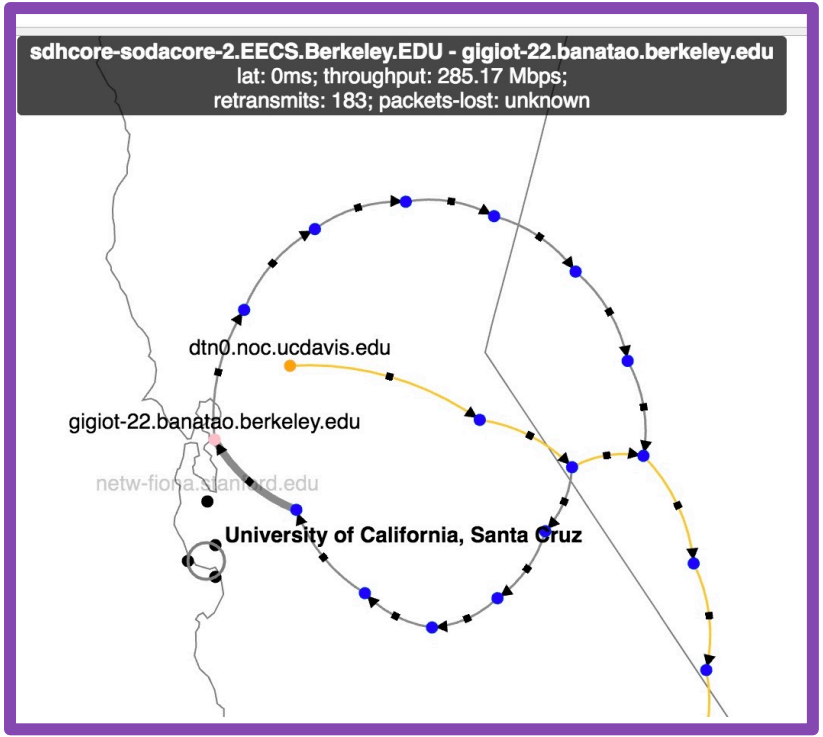
PRP PACIFIC RESEARCH PLATFORM

it²

Traceroute: Real-time Visualization of Status of Network Links

https://traceroute.nautilus.optiputer.net/



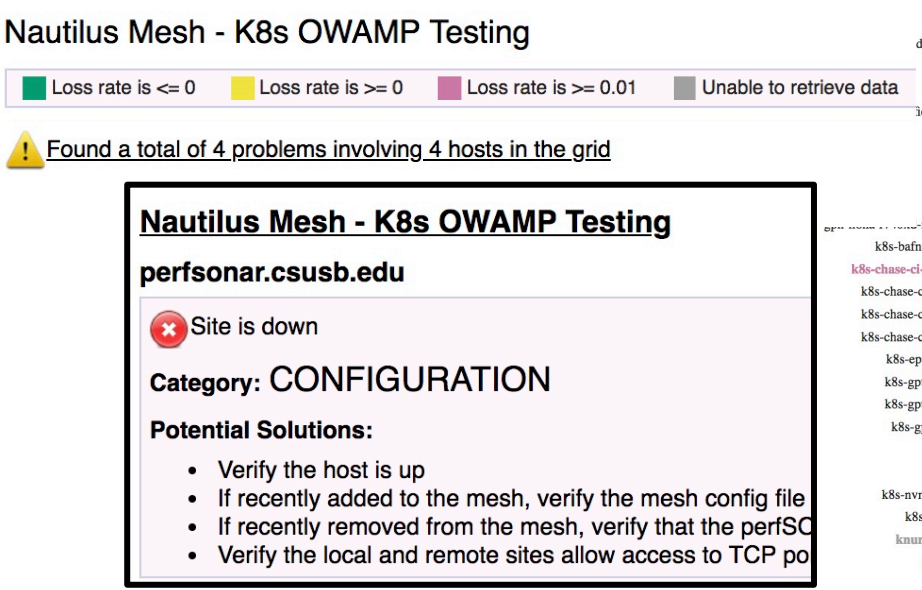


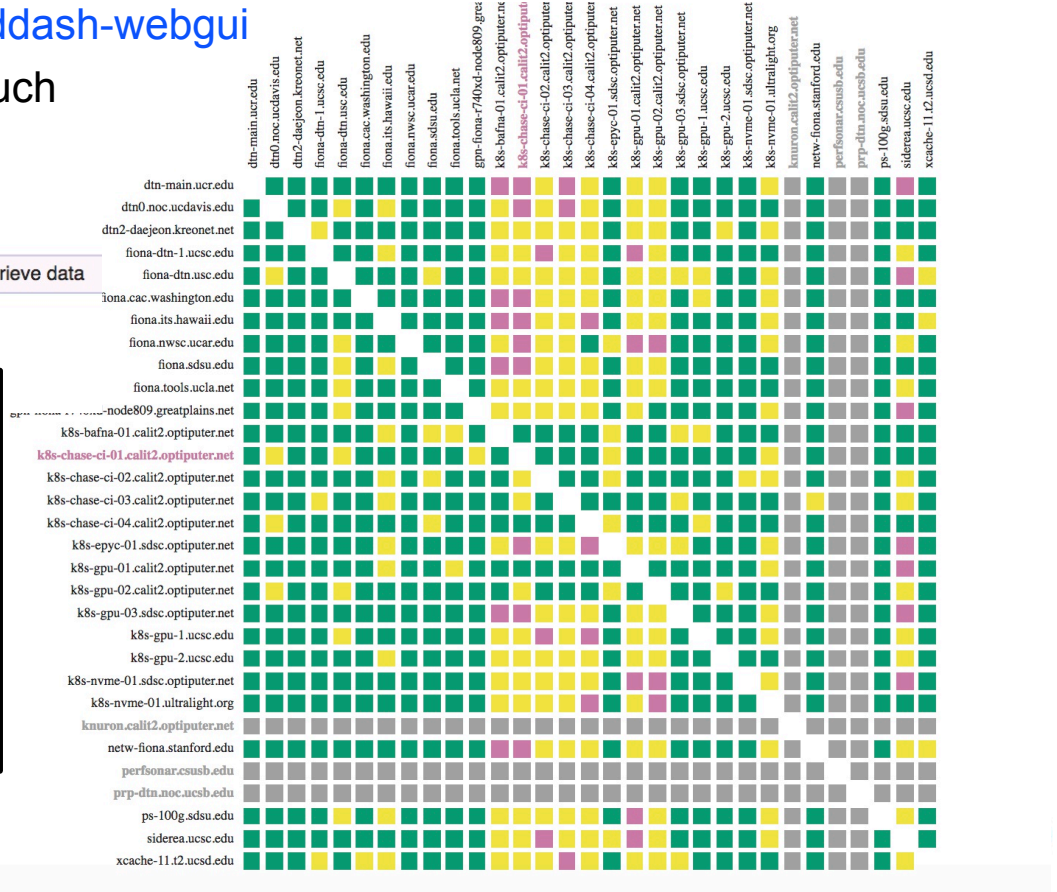
Source: Dmitry Mishin (SDSC), John Graham (Calit2)

One-Way Active Measurement Protocol (OWAMP) testing

https://perfsonar.nautilus.optiputer.net/maddash-webgui

Measures Unidirectional Characteristics such as One-Way Delay and One-Way Loss





PACIFIC RESEARCH PLATFORM

Result: Active, Fact-Based, NET Management

- Our approach gives:
  - Proactive measurements of actual performance
  - Early warning of issues to NOC and engineers
- FIONA devices can run other software to monitor for security incursions and other issues
- Further, the FIONA platform allows really inexpensive network node & end-site based R&E Cloud capabilities

PRP PACIFIC RESEARCH PLATFORM

it²

Installing 16 10&12 TB Drives in June 2018 at UC Merced, UC Riverside and Stanford

PRP
PACIFIC RESEARCH PLATFORM

it²

Use Kubernetes to Manage Across the PRPv2


CADE METZ BUSINESS 06.10.14 01:15 PM

GOOGLE OPEN SOURCES ITS SECRET WEAPON IN CLOUD COMPUTING

WIRED

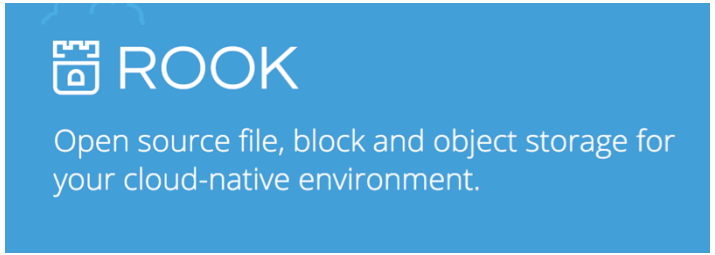
https://kubernetes.io

"Kubernetes is a way of stitching together a collection of machines into, basically, a big computer,"

--Craig Mcluckie, Google
and now CEO and Founder of Heptio

"Everything at Google runs in a container."

--Joe Beda,Google



Battle-tested, production storage:
based on an embedded version of Ceph, with 10+ years of production deployments and runs some of the worlds largest clusters

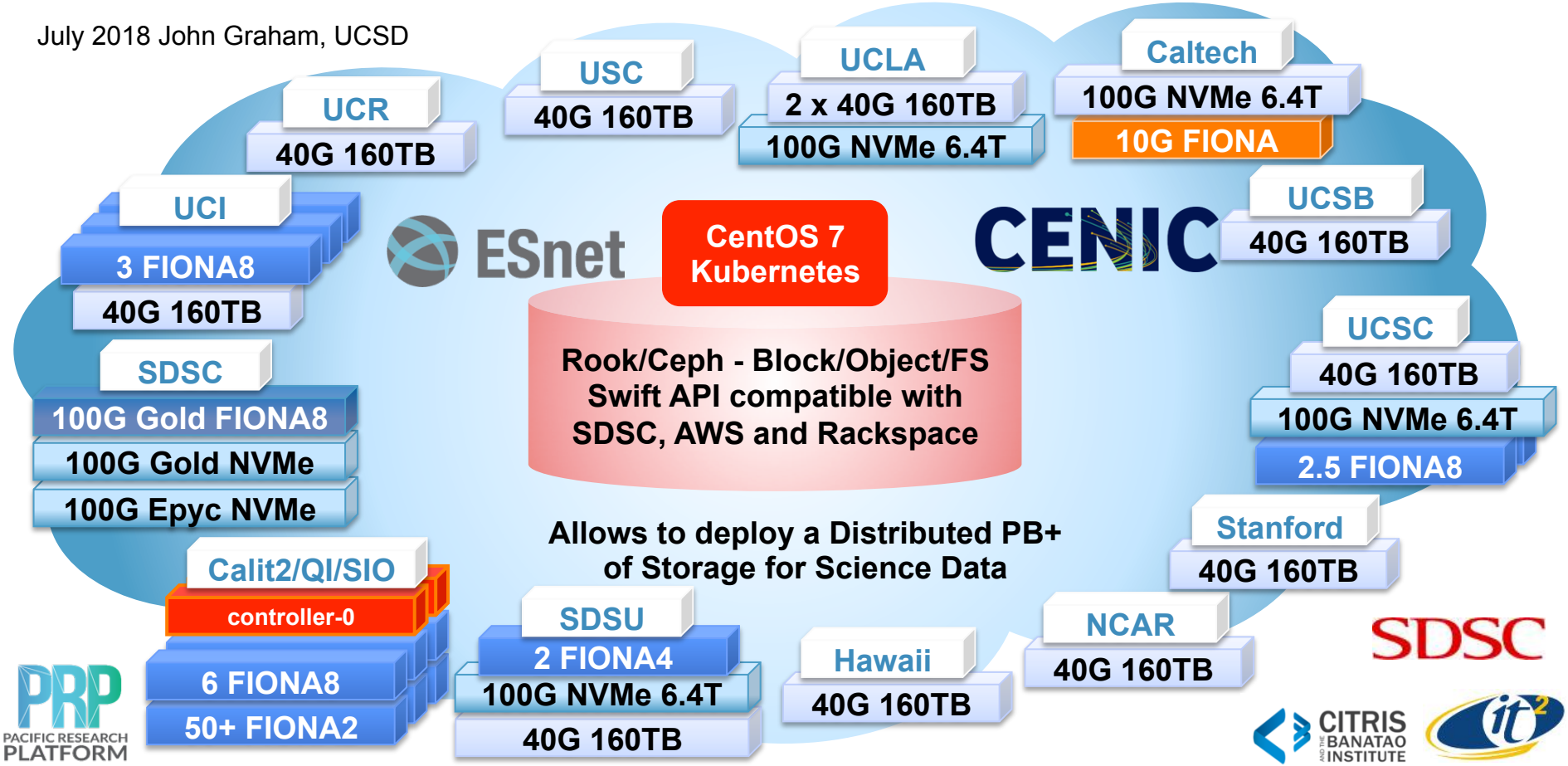
Cloud-native environment integration:
runs as a cloud native service for optimal integration with applications in need of block, object or file storage

Allows the PRP to Deploy Petabytes at $10/TB/year of Distributed Storage and GPUs for Data Science as well as Measure and Monitor Usage

Source: John Graham, Calit2/QI

Nautilus multi-tenant HyperCluster

July 2018 John Graham, UCSD



Grafana Plot of First 730 TB

This is working DMZ Scratch Space, not Archival Research Storage

https://grafana.nautilus.optiputer.net/

PRP's First 2 Years: Connecting Multi-Campus Application Teams and Devices

CMS

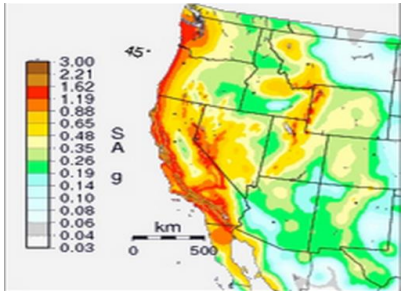
Particle Physics

Biomedical 'omics

INTERMEDIATE PALOMAR TRANSIENT FACTORY

Telescope Surveys

Earth Sciences Engineering

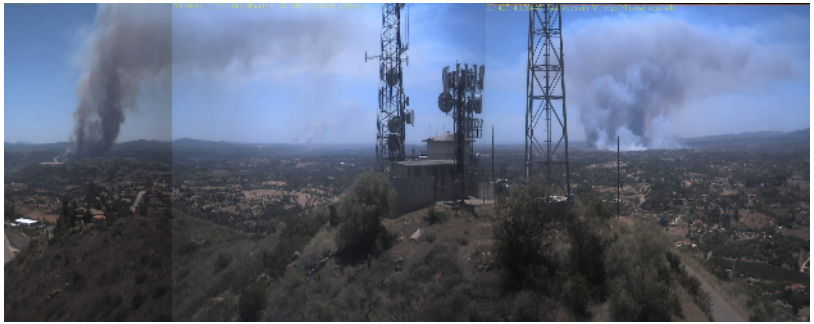


Visualization, Virtual Reality, Collaboration

PRP PACIFIC RESEARCH PLATFORM

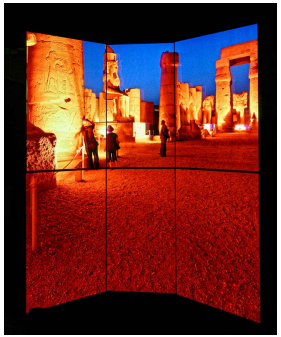
it²

The Prototype PRP Has Attracted New Application Drivers

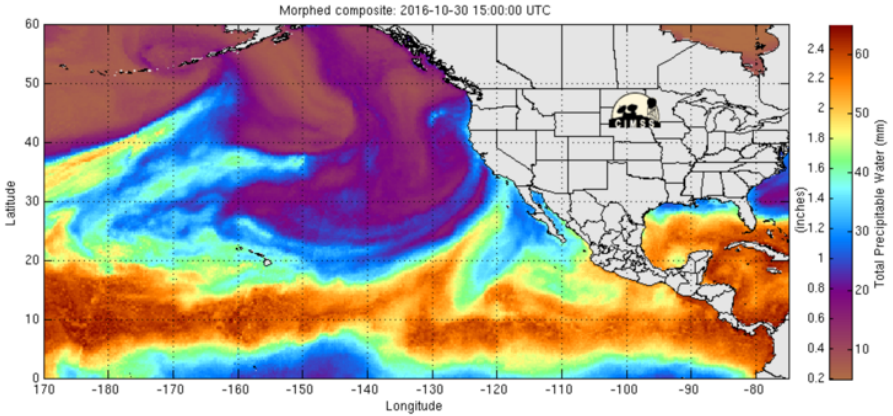


Frank Vernon, Graham Kent, & Ilkay Altintas, **Coupling Wireless Wildfire Sensors to Computing**

Jules Jaffe – **Undersea Microscope**



Tom Levy **At-Risk Cultural Heritage**



Scott Sellars, Marty Ralph **Center for Western Weather and Water Extremes**

PACIFIC RESEARCH PLATFORM

New NSF CHASE-CI Grant Creates a Community Cyberinfrastructure

MSU
UCB
Stanford
UCM
UCSC
Caltech
UCI
UCR
UCSD
SDSU

CI-New: Cognitive Hardware and Software Ecosystem Community Infrastructure (CHASE-CI)

For the Period September 1, 2017 – August 31, 2020

SUBMITTED – January 18, 2017

PI: Larry Smarr, Professor of Computer Science and Engineering, Director Calit2, UCSD
Co-PI: Tajana Rosing, Professor of Computer Science and Engineering, UCSD
Co-PI: Ken Kreutz-Delgado, Professor of Electrical and Computer Engineering, UCSD
Co-PI: Ilkay Altintas, Chief Data Science Officer, San Diego Supercomputer Center, UCSD
Co-PI: Tom DeFanti, Research Scientist, Calit2, UCSD

- NSF Grant for High Speed "Cloud" of 256 GPUs for 30 Faculty & their students at 10 Campuses for Training AI Algorithms on Big Data
- Adding a Machine Learning Layer Built on Top of the Pacific Research Platform

PRP PACIFIC RESEARCH PLATFORM
SDSC
JACOBS SCHOOL OF ENGINEERING
UC San Diego
it²

FIONA8: a FIONA with 8 GPUs
Supports PRP Data Science Machine Learning

Goal: *Machine Learning Researchers Need a New Cyberinfrastructure*

8 Nvidia GTX-1080 Ti GPUs (11 GB)

Dual 8 core 2.1Ghz CPU	**2x NVMe bays**
32GB RAM	**40G ConnectX-4**
6 480GB SSD	**Dual 10G ports**

"Until cloud providers are willing to find a solution to place commodity (32-bit) game GPUs into their servers and price services accordingly, I think we will not be able to leverage the cloud effectively."

"There is an actual scientific infrastructure need here, surprisingly unmet by the commercial market, and perhaps CHASE-CI is the perfect catalyst to break this logjam."

--UC Berkeley Professor Trevor Darrell

PRP PACIFIC RESEARCH PLATFORM

it²

Community Computing Needed to Process the Data at a Price Advantage vs. AWS

Single vs. Double Precision GPUs: Gaming vs. Supercomputing

Nvidia Card	Cost	32-bit GF	GB	Cost per GF	Cost per GB	cores	8 GPU server	160 GPU rack
GTX 1080 TI 11Gb	$726	10609	11	0.07	$66	3584	$13,804	$276,090
P100 16Gb	$8,304	8071	16	1.03	$519	3584	$74,432	$1,488,640
AWS p2.xlarge EC2 8 k80 GPUs + disk for 3 yrs							$239,040	$4,780,800

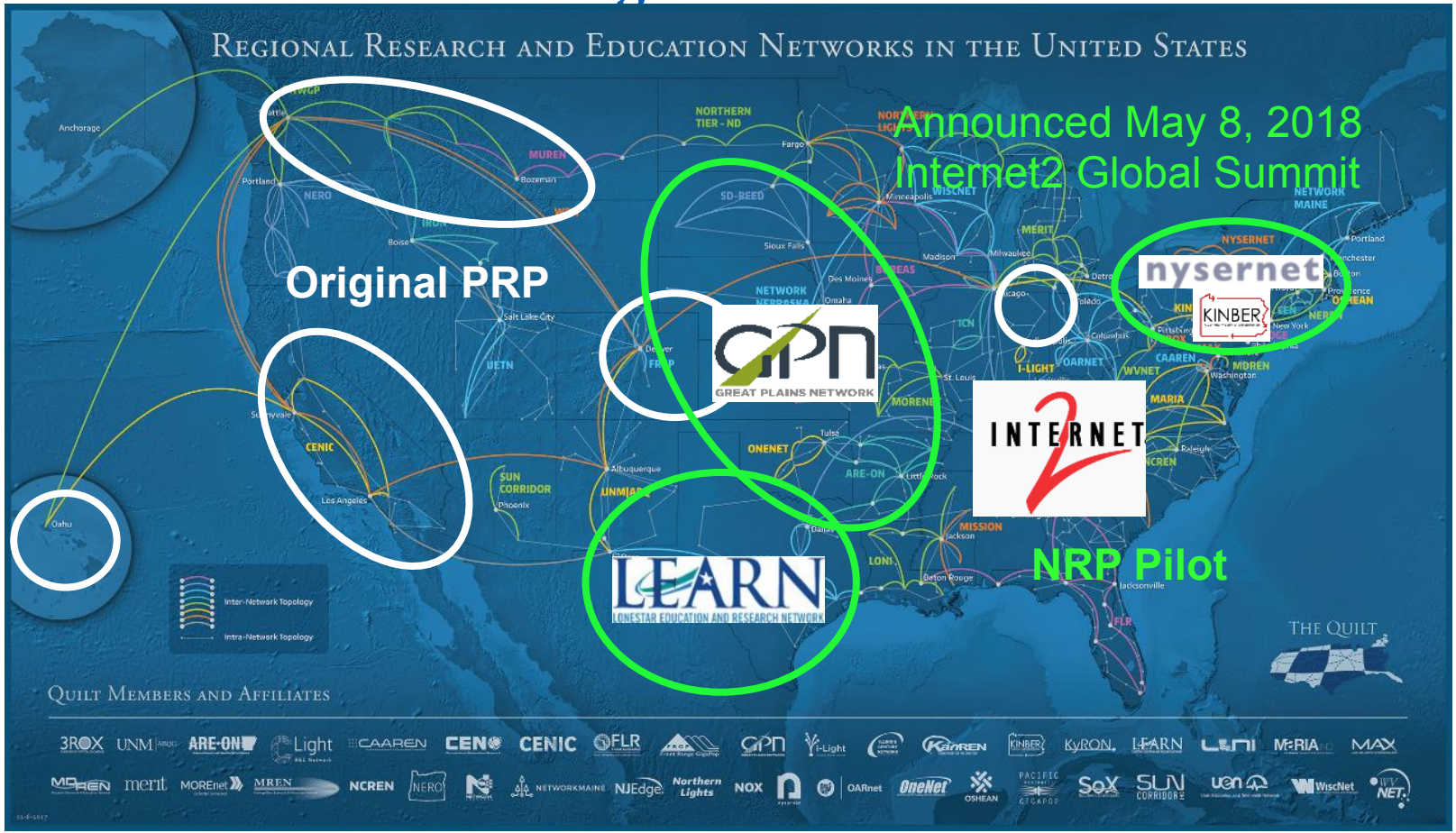
5 FIONA8 nodes:

- 40 GPUs
- 15” of rack space
- 3.44 * 10^6 GPU-hours/day

PRP PACIFIC RESEARCH PLATFORM

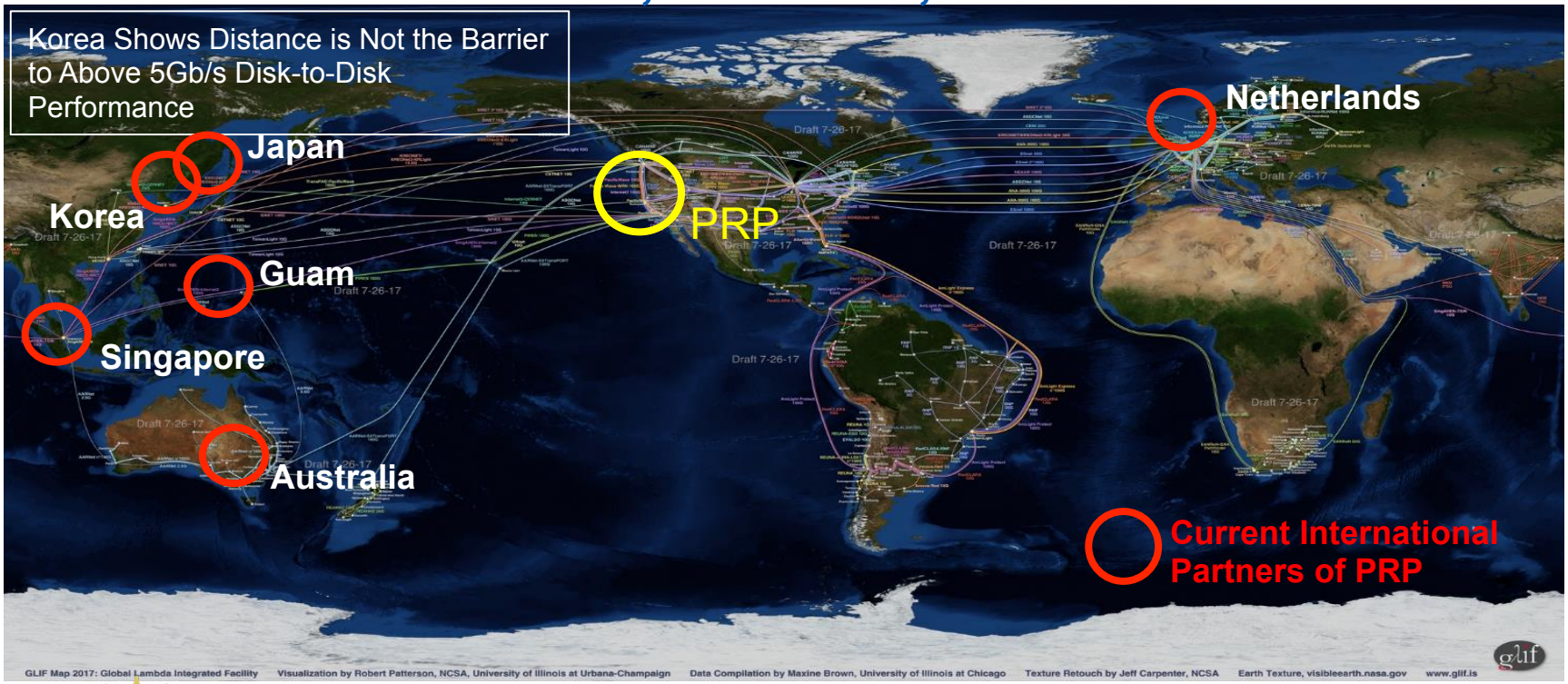
it²

Reaching beyond CENIC/Pacific Wave: Using Internet2 to Connect Regional Networks—NRP Pilot



PACIFIC RESEARCH PLATFORM

Expanding to the Global Research Platform Via CENIC/Pacific Wave, Internet2, and International Links



PRP PACIFIC RESEARCH PLATFORM

NSF CENIC ESnet CITRIS AND THE BANATAO INSTITUTE

http://pacificresearchplatform.org

it²

Trust - a key element for success

- Human trust to establish cooperation across institutions
- It took time for PRP participants to work together
  - to learn individual roles and strengths (and weaknesses),
  - to learn to rely on/trust their collaborators
- Trust is a human-intensive endeavor, one relationship at a time, not readily scalable. But can foster:
  - Identify and document successful collaborations (like PRP)
  - Emphasize peer to peer communications (at all levels)

PRP PACIFIC RESEARCH PLATFORM

it²

Key Findings:
from First NRP Workshop, August 2017

- NRP platform must be easy for scientists to implement and use
- Scientists want to do science, not networking or IT
- NRP is a social engineering project as well as a technical networking/IT project
- ScienceDMZ/DTN architecture is an effective science enabler
- Science Engagement process is crucial to scaling up

PRP PACIFIC RESEARCH PLATFORM

it²

Acknowledgements of PRP support:

- US National Science Foundation (NSF) awards
  CNS 0821155, CNS-1338192, CNS-1456638,
  CNS-1730158, ACI-1540112, ACI-1541349
- University of California Office of the President CIO
- UCSD Chancellor's Integrated Digital Infrastructure Program
- UCSD Next Generation Networking initiative
- Calit2 and Calit2 Qualcomm Institute
- CENIC, PacificWave and StarLight
- DOE ESnet

PRP PACIFIC RESEARCH PLATFORM

it²

Terima kasih!

Thank you!

PRP
PACIFIC RESEARCH PLATFORM

it²