“Toward the Global Research Platform”

Keynote Presentation SC Asia
Singapore
March 27, 2018
Dr. Tom DeFanti
Research Scientist, Co-PI The Pacific Research Platform and CHASE-CI
California Institute for Telecommunications and Information Technology’s Qualcomm Institute
University of California San Diego
Distinguished Professor Emeritus, University of Illinois at Chicago
Abstract

Abstract: The US National Science Foundation-funded (award # 1541349) “The Pacific Research Platform (PRP)” to the University of California San Diego 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 third year of building a broad base of support from application scientists, campus CIOs, regional network leaders, and network engineers, and continues to successfully bring in new, unanticipated science applications, as well as test new means to dramatically improve throughput. The PRP is, in fact, a grand volunteer community in an ever-expanding region where 35 CIOs and 50 application scientists initially signed letters of support for the original NSF proposal, all as unfunded partners. 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 to verify that the technology used is not limited to the size and homogeneity of CENIC, the regional network serving California. There is pent-up demand from the high-performance networking and scientific communities to extend the PRP nationally, and indeed worldwide. This motivated the PRP to host The First National Research Platform Workshop in Bozeman, MT, in August 2017. At that meeting, a strong US and international community emerged, well documented in the report published on the PRP website (pacificresearchplatform.org). This presentation will discuss will cover lessons learned from PRP applications, technology, and science engagement activities, as well as how best to align future PRP networking strategies with the GRP’s emerging groundswell of enthusiasm. The goal is to prototype a future in which a fully-funded multi-national Global Research Platform emerges.

This presentation includes ideas, words and visuals from many sources, Most prominently: the PI of the PRP and CHASE-CI, Larry Smarr, UCSD
Thirty Years After US NSF Adopts US DOE Supercomputer Center Model
NSF Adopts DOE ESnet’s Science DMZ for High Performance Applications

- 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 that are regularly used to characterize and troubleshoot the network
  - Security policies and enforcement mechanisms that are tailored for high performance science environments

Science DMZ
Coined 2010

http://fasterdata.es.net/science-dmz/
Based on Community Input and on ESnet’s Science DMZ Concept, NSF Has Funded Over 100 US Campuses to Build DMZs

Source: NSF
Logical Next Step: The Pacific Research Platform Networks Campus DMZs to Create a Regional End-to-End Science-Driven “Big Data Superhighway” System

NSF CC*DNI DIBBs Cooperative Agreement
$6M 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, UCSD SDSC,
• Frank Wuerthwein, UCSD Physics and SDSC

Letters of Commitment from:
• 50 Researchers from 15 Campuses
• 32 IT/Network Organization Leaders

Source: John Hess, CENIC
Key Innovation: UCSD Designed FIONAs To Solve the Disk-to-Disk Data Transfer Problem at Full Speed on 10/40/100G Networks

- **FIONAs PCs [ESnet DTNs]:**
  - ~$8,000 Big Data PC with:
    - 10/40 Gbps Network Interface Cards
    - 3 TB SSDs
  - **Higher Performance at higher cost:**
    - +NVMe SSDs & 100Gbps NICs Disk-to-Disk
    - +Up to 8 GPUs [4M GPU Core Hours/Week]
    - +Up to 196 TB of Disks used as Data Capacitors
    - +Up to 38 Intel CPU cores or AMD Epyc cores
  - **US$1,100 10Gbps FIONA (if 10G is fast enough)**

- **FIONettes are US$300 EL-30-based FIONAs**
  - 1Gbps NIC With USB-3 for Flash Storage or SSD
  - Perfect for Training and smaller campuses

Phil Papadopoulos, SDSC &
Tom DeFanti, Joe Keefe & John Graham, Calit2
FIONAs on the PRP and Partners

- ~40 FIONAs are on the PRP as GridFTP (MaDDash) + perfSONAR Systems
  - PRP Partners: all 10 UCs, Caltech, Stanford, USC, SDSC, UW, UIC
  - Plus U Utah, Montana State, U Chicago, Clemson U, U Hawaii, NCAR, Guam
  - Plus Internationals: Uv Amsterdam, KISTI (Korea), Singapore

- Many States and Regionals Building FIONAs and Creating MaDDashes
  - FIONA Build Specs on pacificresearchplatform.org Website
  - Weekly Engineering Calls with Notes Going to 60+ Technical Participants
  - Fasterdata.es.net has lots of DTN and DMZ wisdom and data
We Measure Disk-to-Disk Throughput with 10GB File Transfer 4 Times Per Day in Both Directions for All PRP Sites

From Start of Monitoring 12 DTNs to 24 DTNs Connected at 10-40G in 1 ½ Years

Source: John Graham, Calit2/QI
We Use Kubernetes to Manage FIONAs Across the PRP

“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
Rook is Ceph Cloud-Native Object Storage ‘Inside’ Kubernetes

https://rook.io/

Open source file, block and object storage for your cloud-native environment.

Battle-tested, production storage

Rook is based on an embedded version of Ceph, which has 10+ years of production deployments and runs some of the world's largest clusters.

Cloud-native environment integration

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

Source: John Graham, Calit2/QI
We Built Nautilus - A Multi-Tenant Containerized PRP HyperCluster for Big Data Applications Running Kubernetes with Rook/Ceph Cloud Native Storage and GPUs for Machine Learning.

Rook/Ceph - Block/Object/FS Swift API compatible with SDSC, AWS, and Rackspace.

March 2018 John Graham, Calit2/QI
New NSF CHASE-CI Grant Creates a Community Cyberinfrastructure: Adding a Machine Learning Layer Built on Top of the Pacific Research Platform

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 ML Faculty & Their Students at 10 Campuses
for Training AI Algorithms on Big Data

JACOBS SCHOOL OF ENGINEERING

SDSC UC San Diego
“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
FIONA8: a FIONA with 8 GPUs
Supports PRP Data Science Machine Learning--4M GPU Core Hours/Week

8 Nvidia GTX-1080 Ti GPUs (11 GB)
Testing AMD Radeon Vega (16 GB)

24 CPU Cores, 32,000 GPU cores, 96 GB RAM, 2TB SSD, Dual 10Gbps ports
3” High; ~$16,000
## Single vs. Double Precision GPUs: Gaming vs. Supercomputing

<table>
<thead>
<tr>
<th>Nvidia Card</th>
<th>~Cost</th>
<th>32-bit GF</th>
<th>GB</th>
<th>per GF</th>
<th>per GB</th>
<th>cores</th>
<th>8-GPU PC</th>
<th>160 GPU rack</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTX 1080 Ti 11GB</td>
<td>$726</td>
<td>10609</td>
<td>11</td>
<td>$0.07</td>
<td>$66</td>
<td>3584</td>
<td>$13,804</td>
<td>$276,090</td>
</tr>
<tr>
<td>P100 16GB</td>
<td>$8,304</td>
<td>8071</td>
<td>16</td>
<td>$1.03</td>
<td>$519</td>
<td>3584</td>
<td>$74,432</td>
<td>$1,488,640</td>
</tr>
<tr>
<td>AWS p2.xlarge EC2 (8) K-80 GPUs+disk for 3 years +55% ICR</td>
<td>$370,512</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AWS p2.xlarge EC2 (8) K-80 GPUs+disk for 3 years</td>
<td>$239,040</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8 x 1080 Ti: 1 million GPU core hours every two days.

700 million GPU core hours for $16K in 4 yrs

$22/million GPU core hours.

Plus power, admin costs
UCSD Game GPUs for Data Sciences Cyberinfrastructure - Devoted to Data Analytics and Machine Learning Research and Teaching

UC San Diego

IT SERVICES

88 GPUs for Students

UCSD Cognitive Science

JACOBS SCHOOL OF ENGINEERING

FIONA with 8-Game GPUs

Open Science Grid

GPUs for OSG Applications

SunCAVE 70 GPUs

WAVE + VROOM 48 GPUs

CHASE-CI Grant Provides
256 GPUs to 32 Researchers on 10 Campuses:
>22B GPU Core Hours over 4 years
Running Kubernetes/Rook/Ceph On PRP
Allows Us to Deploy a Distributed PB+ of Storage for Posting Science Data
Expanding to the Global Research Platform
Via CENIC/Pacific Wave, Internet2, and International Links

Asia to US Shows Distance is Not the Barrier
to Above 5Gb/s Disk-to-Disk Performance

PRP's Current International Partners

Korea, Guam, Singapore, Australia, Netherlands
PRP Held
The First National Research Platform Workshop on August 7-8, 2017

Co-Chairs:
Larry Smarr, Calit2
& Jim Bottum, Internet2
Program Chair:
Tom DeFanti

See agenda, reports, video on pacificresearchplatform.org

135 Attendees
Coming: The Second National Research Platform Workshop (2NRP)
Bozeman, MT August 6-7, 2018—Register Soon at CENIC.ORG!

Local Hosts: Jerry Sheehan, MSU and CENIC

Steering Committee:
Larry Smarr, Calit2
Inder Monga, ESnet
Ana Hunsinger, Internet2

Program Committee:
Jim Bottum
Maxine Brown
Sherilyn Evans
Marla Meehl
Wendy Huntoon
Kate Mace
Thank You for Your Kind Attention!
Our Support Comes From:

- 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’s Qualcomm Institute
- CENIC, PacificWave and StarLight
- DOE ESnet
PRP’s First 2 Years:
Connecting Multi-Campus Application Teams and Devices

Particle Physics

Biomedical ‘omics

Telescope Surveys

Earth Sciences

Visualization, Virtual Reality, Collaboration
Data Transfer Rates From 40 Gbps DTN in UCSD Physics Building, Across Campus on PRISM DMZ, Then to Chicago’s Fermilab Over CENIC/ESnet

Based on This Success,
Upgrading 40G DTN to 100G
For Bandwidth Tests & Kubernetes
to OSG, Caltech, and UCSC

Source: Frank Wuerthwein, UCSD, SDSC
LHC Data Analysis Running on PRP

Two Projects:
- OSG Cluster-in-a-Box for “T3”
- Distributed Xrootd Cache for “T2”

Source: Frank Würthwein, OSG, UCSD/SDSC, PRP
PRP Over CENIC
Couples UC Santa Cruz Astrophysics Cluster to LBNL NERSC Supercomputer

A 2D superluminous supernova simulation generated with the CASTRO code. (Image: Ken Chen, National Astronomical Observatory of Japan)


NERSC Project PI: S. Woolsey, UC Santa Cruz
100 Gbps FIONA at UCSC Allows for Downloads to the UCSC Hyades Cluster from the LBNL NERSC Supercomputer for DESI Science Analysis

Precursors to LSST and NCSA

300 images per night.
100MB per raw image
120GB per night

Source: Peter Nugent, LBNL
Professor of Astronomy, UC Berkeley

250 images per night.
530MB per raw image
800GB per night

NSF-Funded Cyberengineer
Shaw Dong @UCSC
Receiving FIONA
Feb 7, 2017
Simulating the Injection of CO$_2$ in Brine-Saturated Reservoirs: Poroelastic & Pressure-Velocity Fields Solved In Parallel With MPI Using Domain Decomposition Across Containers
PRP Enables Distributed Walk-in Virtual Reality CAVEs

Pr transferring 5 CAVEcam images from UCSD to UC Merced:
2 Gigabytes now takes 2 Seconds (8 Gb/sec)
The Prototype PRP Has Attracted New Application Drivers

Frank Vernon, Graham Kent, & Ilkay Altintas, Wildfires

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

Tom Levy At-Risk Cultural Heritage

Jules Jaffe – Undersea Microscope
PRP Links At-Risk Cultural Heritage and Archaeology Datasets at UCB, UCLA, UCM and UCSD with CAVEkiosks

UC President Napolitano’s Research Catalyst Award to UC San Diego (Tom Levy), UC Berkeley (Benjamin Porter), UC Merced (Nicola Lercari) and UCLA (Willeke Wendrich)
Technology Projects to Combat California Wildfires Are Recognized with CENIC Innovation Award

New PRP Application:
Coupling Wireless Wildfire Sensors to Computing

Church Fire, San Diego CA
Alert SD&ECameras/HPWREN
October 21, 2017

Thomas Fire, Ventura, CA
Firemap Tool, WIFIRE
December 10, 2017

CENIC 2018
Innovations in Networking Award for Experimental Applications
Mount Laguna Meteorological Sensor Instrumentation Provides Real-Time Data Flows Over HPWREN to PRP-Connected Servers

Source: Hans-Werner Braun, SDSC
HPWREN-Connected SoCal Weather Stations:
Giving High-Resolution Weather Data in San Diego County
PRP/CENIC Backbone Sets Stage for 2018 Expansion of HPWREN Wireless Connectivity Into Orange and Riverside Counties

- **PRP CENIC 100G Links UCSD, SDSU & UCI HPWREN Servers**
  - FIONAs Endpoints
  - Data Redundancy
  - Disaster Recovery
  - High Availability
  - Kubernetes Handles Software Containers and Data

- **Potential Future UCR CENIC Anchor**

Source: Frank Vernon, Hans Werner Braun HPWREN

UCI Antenna Dedicated June 27, 2017
Once a Wildfire is Spotted, PRP Brings High-Resolution Weather Data to Fire Modeling Workflows in WIFIRE

Source: Ilkay Altintas, SDSC
Some Machine Learning Case Studies
To Improve on WIFIRE

- Smoke and fire perimeter detection based on imagery
- Prediction of Santa Ana and fire conditions specific to location
- Prediction of fuel build up based on fire and weather history
- NLP for understanding local conditions based on radio communications
- Deep learning on multi-spectra imagery for high resolution fuel maps
- Classification project to generate more accurate fuel maps (using Planet Labs satellite data)

All Require Periodic, Dynamic, and Programmatic Access to Data!

Source: Ilkay Altintas, SDSC; Co-PI CHASE-CI
Collaboration on Atmospheric Water in the West Between UC San Diego and UC Irvine

Director: F. Martin Ralph  Website: cw3e.ucsd.edu

Big Data Collaboration with:

Director, Soroosh Sorooshian, UCSD  Website: http://chrs.web.uci.edu

Source: Scott Sellers, CW3E
Major Speedup in Scientific Work Flow Using the PRP

Pacific Research Platform (10-100 Gb/s)

Complete workflow time: 20 days → 20 hrs → 20 Minutes!

UC, Irvine

GPUs

SDSC’s COMET

UC, San Diego

GPUs

Calit2’s FIONA

Source: Scott Sellers, CW3E
Using Machine Learning to Determine the Precipitation Object Starting Locations

*Sellars et al., 2017 (in prep)
UC San Diego Jaffe Lab (SIO) Scripps Plankton Camera
Off the SIO Pier with Fiber Optic Network
Over 300 Million Images So Far! Requires Machine Learning for Automated Image Analysis and Classification

Phytoplankton: Diatoms

Zooplankton: Copepods

Zooplankton: Larvaceans

"We are using the FIONAs for image processing... this includes doing Particle Tracking Velocimetry that is very computationally intense." - Jules Jaffe

Source: Jules Jaffe, SIO