“The PRP and Its Applications”

Panel Presentation

With Tom DeFanti and Frank Wuerthwein

Nautilus and the National Research Platform

CENIC 2022 Annual Conference

September 27, 2022

Dr. Larry Smarr

Founding Director Emeritus, California Institute for Telecommunications and Information Technology;

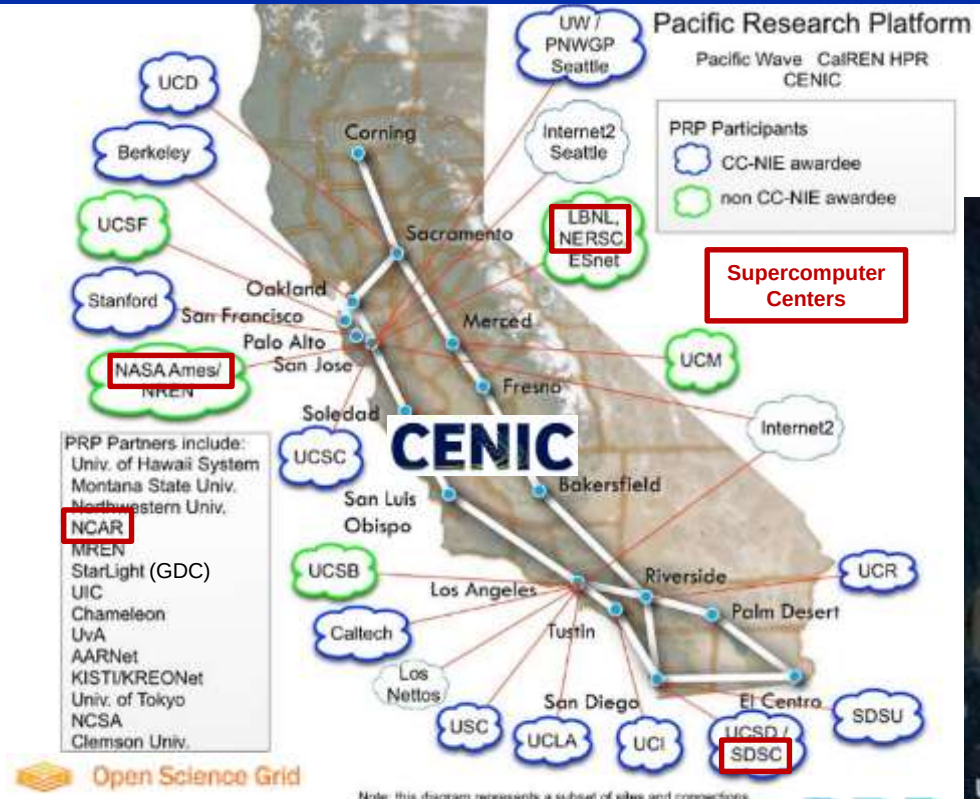
Distinguished Professor Emeritus, Dept. of Computer Science and Engineering

Jacobs School of Engineering, UCSD

http://lsmarr.calit2.net

PACIFIC RESEARCH PLATFORM

2015 Vision: The Pacific Research Platform Will Connect Science DMZs Creating a Regional End-to-End Science-Driven Community Cyberinfrastructure



NSF CC*DNI Grant
$7.3M 10/2015-10/2020
Extended - In Year 7 Now



Source: John Hess, CENIC

2015-2022: UCSD Designs PRP Data Transfer Nodes (DTNs) -- Flash I/O Network Appliances (FIONAs)

FIONAs Solved the Disk-to-Disk Data Transfer Problem at Near Full Speed on Best-Effort 10G, 40G and 100G

FIONA

Software Specifications

Installation Instructions

Hardware Specifications

2021 – 2020 – 2019 – 2018 – 2017 – 2016 – 2015

https://pacificresearchplatform.org/fiona/



Add Up to 8 Nvidia GPUs Per 2U FIONA
To Add Machine Learning Capability
Up to 240TB Storage

PRP PACIFIC RESEARCH PLATFORM

NSF

FIONAs Designed by UCSD's Phil Papadopoulos, John Graham, Joe Keefe, and Tom DeFanti

SDSC

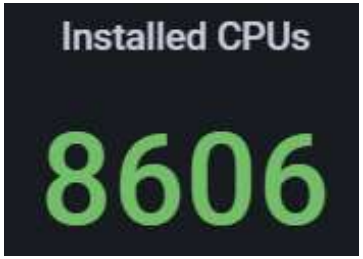
QUALCOMM INSTITUTE

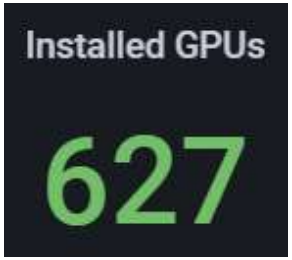
PRP's Nautilus is a Multi-Institution Hypercluster Connected by Optical Networks

Nautilus

160 GPU & Storage FIONAs on 27 Partner Campuses Networked Together at 10-100Gbps

Grafana







PRP
PACIFIC RESEARCH PLATFORM

CENIC

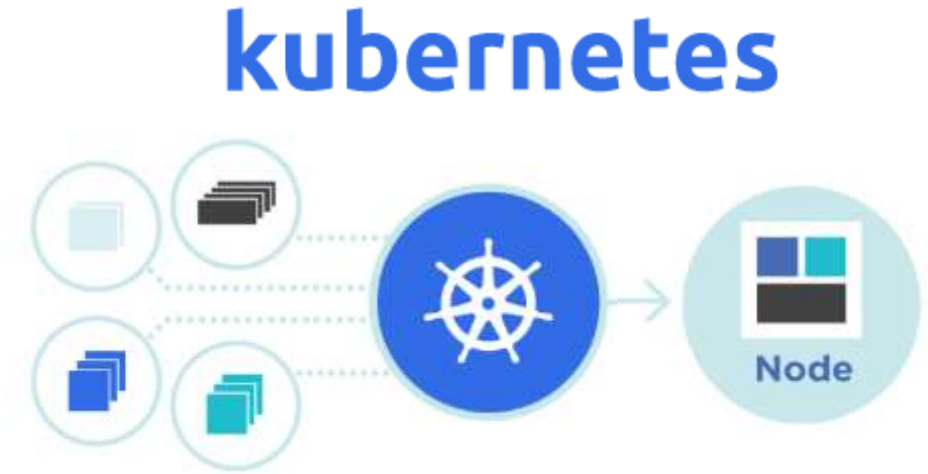
THE QUILT

INTERNET 2

As of Sept 21, 2022

2018/2019: PRP Game Changer!
Using Google's Kubernetes to Orchestrate Containers Across the PRP

User Applications



Containers

PRP PACIFIC RESEARCH PLATFORM

Clouds

PRP PACIFIC RESEARCH PLATFORM

CLOUD NATIVE COMPUTING FOUNDATION

PRP's Nautilus Hypercluster Adopted Open-Source Kubernetes and Rook to Orchestrate Software Containers and Manage Distributed Storage

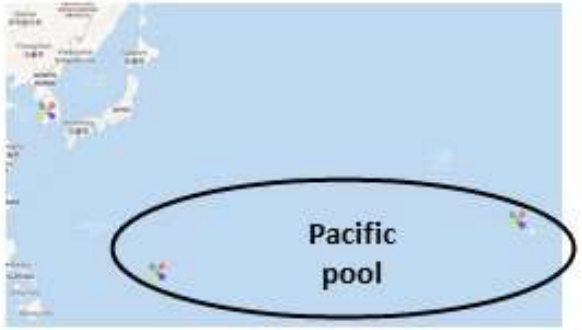


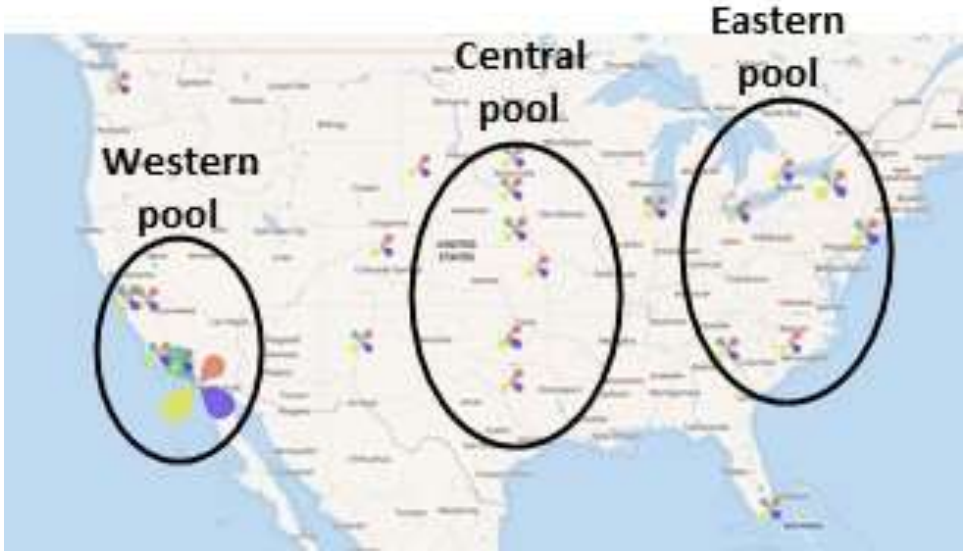


"Kubernetes with Rook/Ceph Allows Us to Manage Petabytes of Distributed Storage and GPUs for Data Science, While We Measure and Monitor Network Use."
--John Graham, UC San Diego

PRP
PACIFIC RESEARCH PLATFORM

Nautilus Has Established a Distributed Set of Ceph Storage Pools





PRP forms optimal-scale Ceph pools with best performance and lowest latency

Allows users to select the placement for compute jobs relative to the storage pools.

The PRP Web Site Provides Widely-Used Open-Source Services For How to Join, Application Research, Development, and Collaboration



Nautilus documentation

Get Access

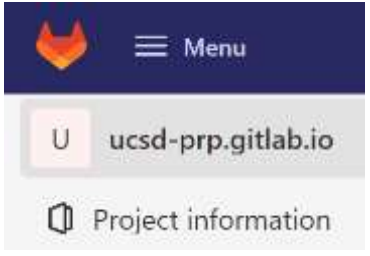
Docker and containers

Basic kubernetes

PRP PACIFIC RESEARCH PLATFORM

[matrix] Chat for Nautilus Users

JupyterLab for Nautilus





Five Major Components of Nautilus Security

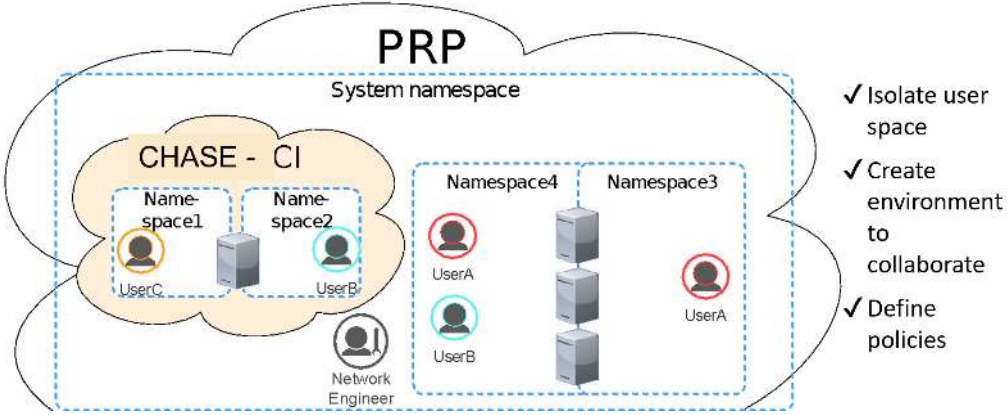
ESnet
ENERGY SCIENCES NETWORK

Science DMZ Security - Firewalls vs. Router ACLs

https://fasterdata.es.net/science-dmz/science-dmz-security/

Nautilus Uses CILogon for Federated Identity
Already Adopted by Most Campuses

CILogon



- ✓ Isolate user space
- ✓ Create environment to collaborate
- ✓ Define policies

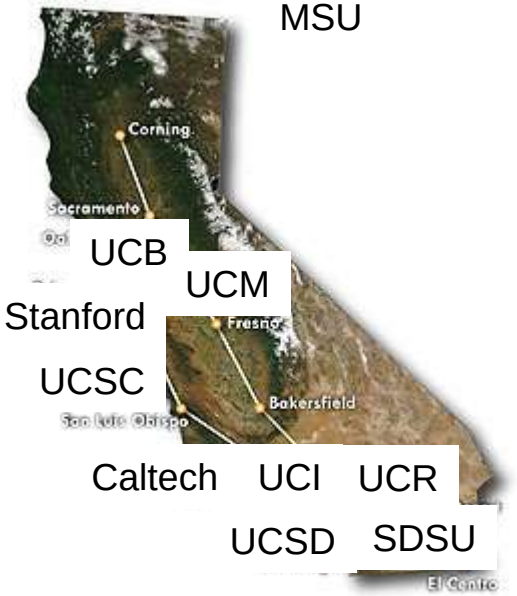
Policy-driven network security



PRP
PACIFIC RESEARCH PLATFORM

ElastiFlow
Network Performance and Security Analytics for Open Data Platforms

2017-2020: NSF CHASE-CI Grant Adds a Machine Learning Layer Built on Top of the Pacific Research Platform



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

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, Calit2, UCSD
Co-PIs:

- Tajana Rosing, CSE, UCSD
- Ken Kreutz-Delgado, ECE, UCSD
- Ilkay Altintas, SDSC, UCSD
- Tom DeFanti, QI, UCSD

NSF Has Funded Two Extensions:
CHASE-CI ABR-Smarr PI &
CHASE-CI ENS-DeFanti PI
$2.8M

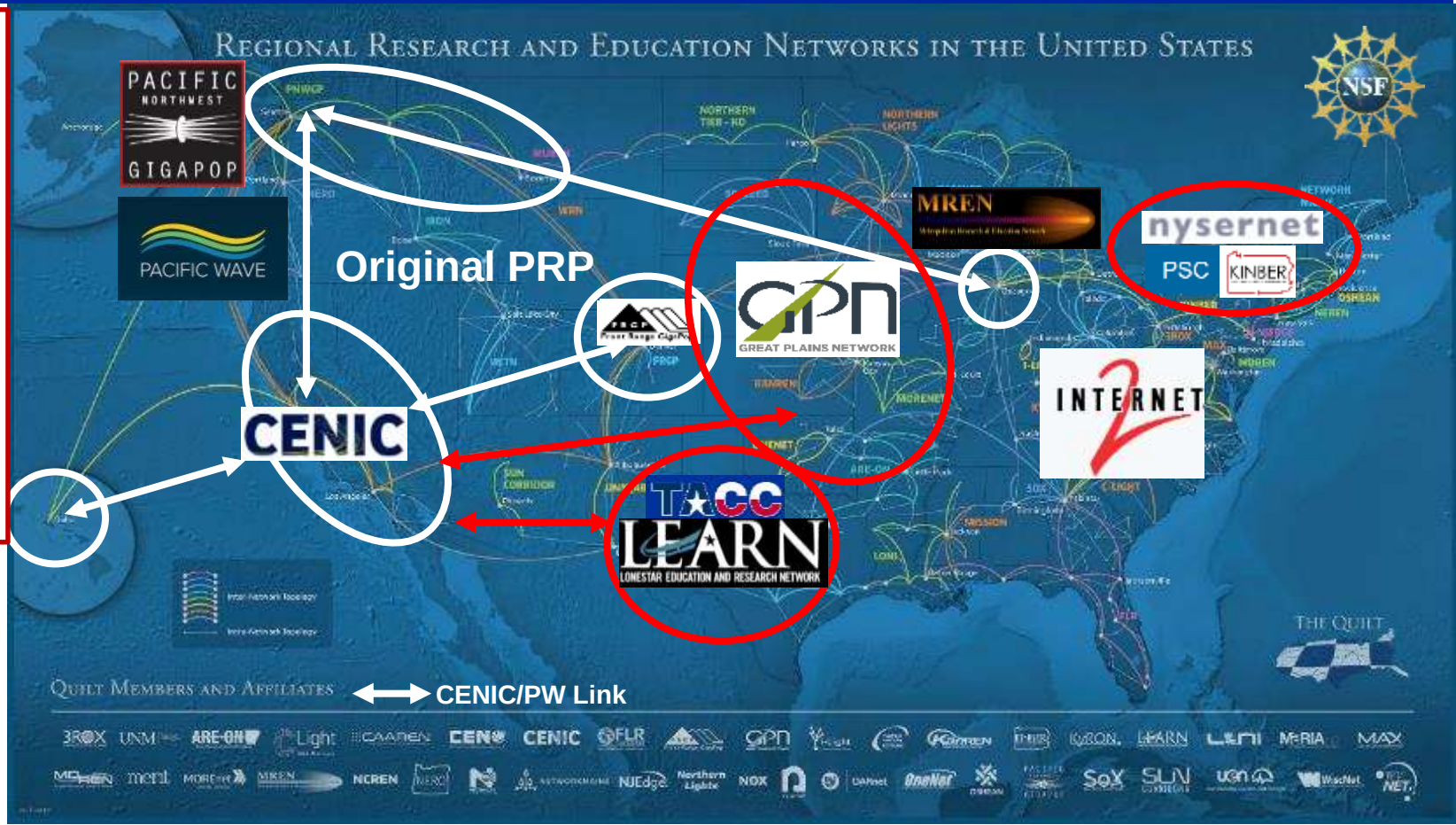
PACIFIC RESEARCH PLATFORM

JACOBS SCHOOL OF ENGINEERING
SDSC UC San Diego QUALCOMM INSTITUTE

2018-2021: Toward the National Research Platform (TNRP) - Using CENIC & Internet2 to Connect Quilt Regional R&E Networks

"Towards The NRP"
3-Year Grant Funded by NSF
$2.5M
October 2018
Award #1826967

PI Smarr
Co-PIs Altintas
Papadopoulos
Wuerthwein
Rosing
DeFanti



PACIFIC RESEARCH PLATFORM

The Open Science Grid (OSG) Has Been Integrated With the PRP

Source: Frank Würthwein,
OSG Exec Director; PRP co-PI; UCSD/SDSC

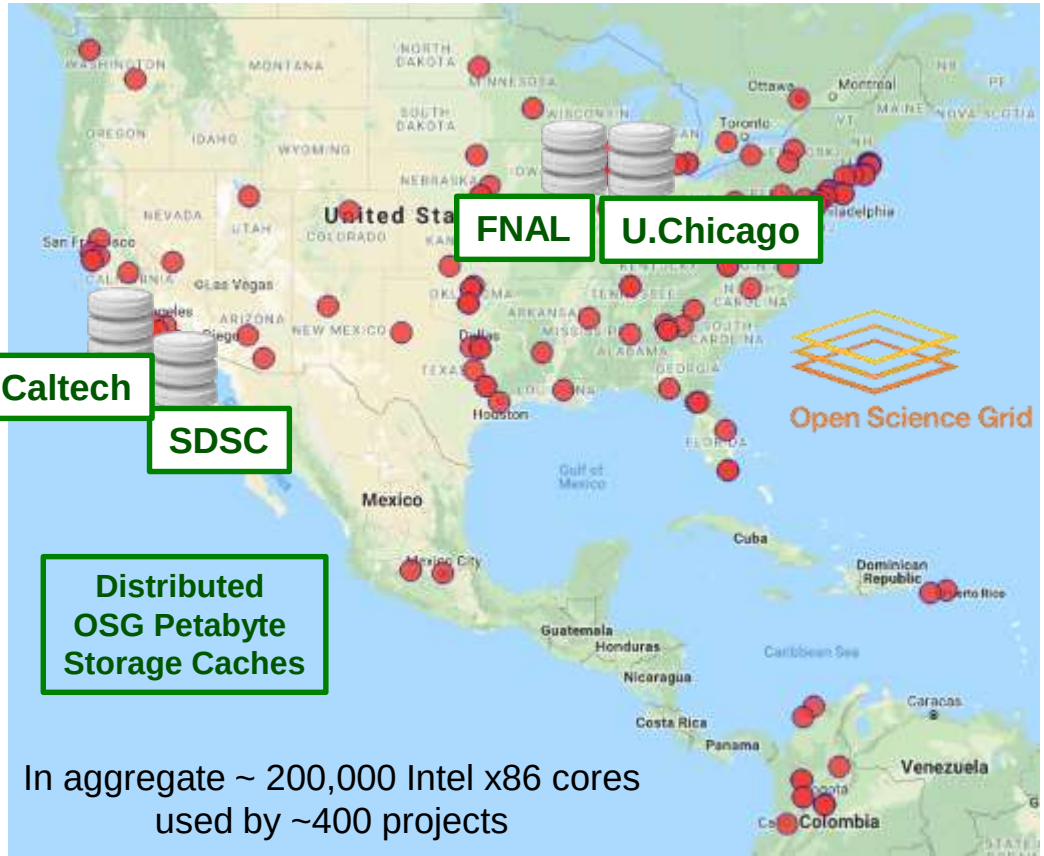
All OSG User Communities
Use HTCondor for Resource Orchestration

HTCondor
High Throughput Computing

NSF

PRP
PACIFIC RESEARCH PLATFORM

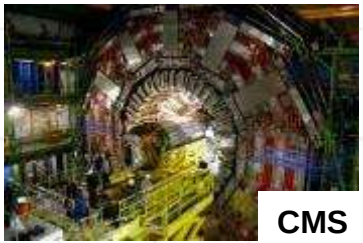
OSG Federates ~100 Clusters Worldwide



The Open Science Grid Delivers to Over 50 Fields of Science
2.4 Billion Core-Hours Per Year of Distributed High Throughput Computing

NCSA Delivered ~35,000 Core-Hours Per Year in 1990





PACIFIC RESEARCH PLATFORM

HTCondor Software Suite

OSG

More Than 1 Million GPU-Hours on PRP Used via OSG Integration Within the Last 2 Years

https://gracc.opensciencegrid.org/dashboard/db/gracc-home

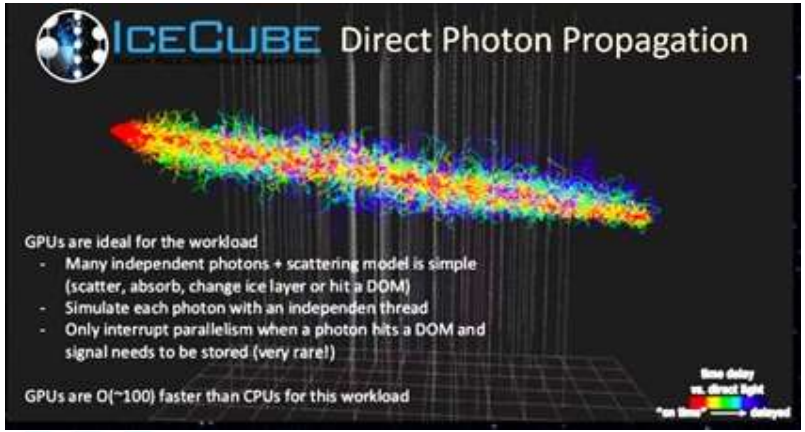
The New Pacific Research Platform Video Highlights 3 Different Applications Out of 700 Nautilus Namespace Projects



Pacific Research Platform Video:
https://pacificresearchplatform.org/media/pacific-research-platform-video/

NSF Large-Scale Observatories Are Using PRP and OSG as a Cohesive, Federated, National-Scale Research Data Infrastructure





GPU Simulations Needed *to Improve Ice Model*.
=> Results in ***Significant Improvement***
in Pointing Resolution for Multi-Messenger Astrophysics

WISCONSIN WIPAC Wisconsin IceCube Particle Astrophysics Center NSF

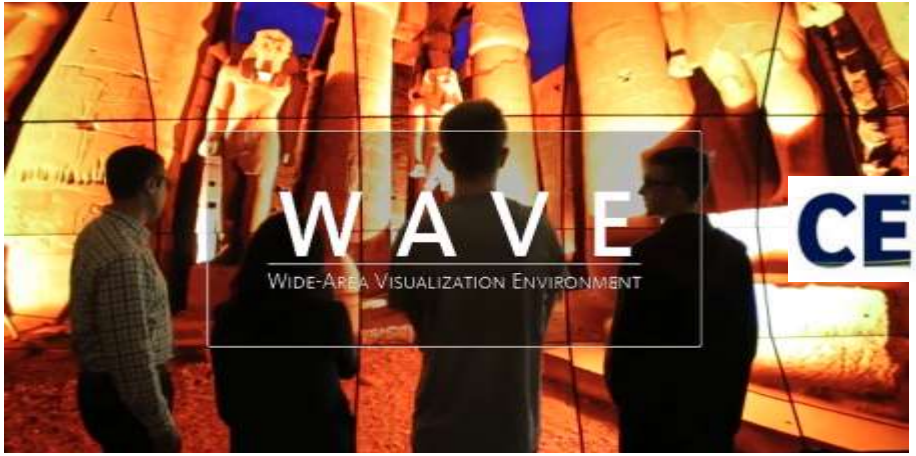
IceCube Used Up to 300 of PRP's 500 GPUs in 2021!

PRP Pacific Research Platform

NSF's IceCube & LIGO Both See Nautilus as Just Another OSG Resource

Open Science Grid

2017: PRP 20Gbps Connection of UCSD SunCAVE and UCM WAVE Over CENIC
2018-2019: Added Their 90 GPUs to PRP for Machine Learning Computations



UC Merced WAVE (20 Screens, 20 GPUs)



UCSD SunCAVE (70 Screens, 70 GPUs)

CENIC

See These VR Facilities in Action in the PRP Video

PRP
PACIFIC RESEARCH PLATFORM

Leveraging UCM Campus Funds and NSF CNS-1456638 & CNS-1730158 at UCSD

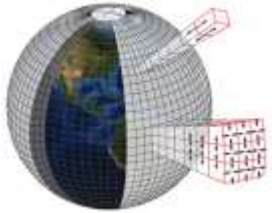
NSF-Funded WIFIRE Uses PRP/CENIC to Couple Wireless Edge Sensors With Supercomputers, Enabling Fire Modeling Workflows



HPWREN

Real-Time Meteorological Sensors

Weather Forecasts



PRP

CENIC

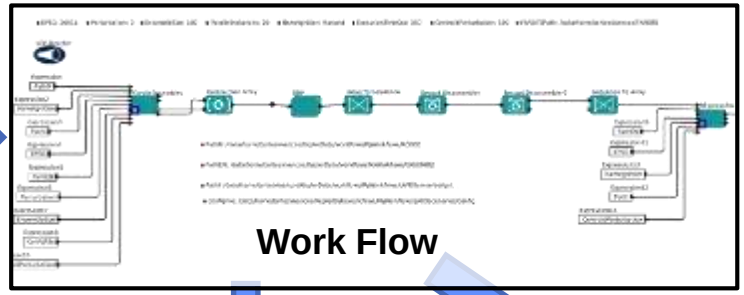


Fire Perimeter

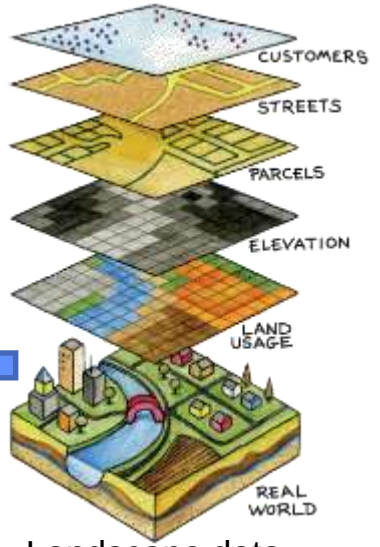
Source: Ilkay Altintas, SDSC

WIFIRE





Work Flow



Landscape data



SDSC





PRP

PACIFIC RESEARCH PLATFORM

NSF

The PRP Has Emphasized Expanding Diversity and Inclusion

- **When the PRP Grant Was Funded in 2015, It Started With:**
  - **6 States ➡ Now 40 States**
  - **19 Campuses ➡ Now 95 Campuses**
  - **9 Minority Serving Institutions ➡ Now 20 MSIs**
  - **2 NSF EPCoR States ➡ Now 19 EPSCoR States, 2 Territories, and Wash DC**

PRP
PACIFIC RESEARCH PLATFORM

The Pacific Research Platform Is Only Possible Because of the Prior Existence of CENIC!



PRP
PACIFIC RESEARCH
PLATFORM

https://cenic.org/blog/prp-fulfilling-the-promise-of-collapsing-space-and-time