“The National Research Platform Enables a Growing Diversity of Users and Applications”

Remote Presentation

Emerging Centers Track

Campus Research Computing Consortium (CaRCC)

June 21, 2023

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

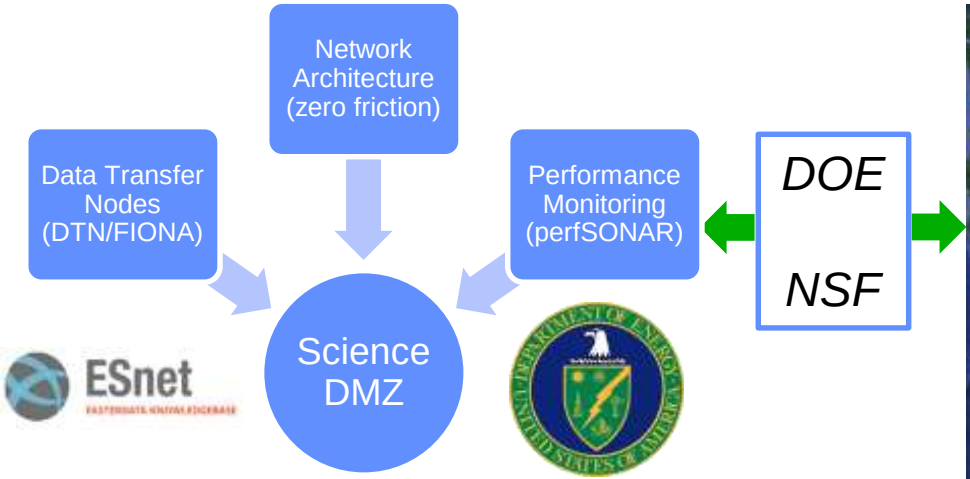
Abstract

The demand from academic institutions of all scales for equitable and inclusive access to affordable computing (CPUs, GPUs, FPGAs) and data storage to enable machine learning (ML) and artificial intelligence (AI) applications has grown rapidly in recent years. I will discuss how NSF funding of a number of cyberinfrastructure grants has led to the emergence of the National Research Platform (NRP), which is increasingly providing this access. The NRP, led by UCSD's Professor Frank Wuerthwein (SDSC Director) enables faculty and students to launch their containerized software applications from their home campus onto the widely distributed NRP cyberinfrastructure, where Google's Kubernetes automatically executes the software. I will review the broad spectrum of both data-intensive disciplinary and emerging AI/ML NRP applications. The NRP has rapidly grown as more and more campus host NRP nodes, all unified into a shared CI Commons. NRP currently enables user access to over 1,250 GPUs, nearly 20,000 CPU cores, and 32 FPGAs plus >10 PB of 40-100Gbps-connected high-speed storage. The campuses having NRP users has grown over the last ten years from 19 campuses in 6 states to 135 campuses in 45 states. Diversity has increased as well: from the original 9 Minority Serving Institutions with NRP users to now over 25 and from 2 NSF EPSCoR states to 20, plus two territories and the District of Columbia. The NRP's goal is to continue this inclusive growth by training and engaging an even broader set of researchers.

NRP NATIONAL RESEARCH PLATFORM

2010-2022:
NSF Adopted a DOE High-Performance Networking Model

ScienceDMZ Coined in 2010 by ESnet



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

Slide Adapted From Inder Monga, ESnet

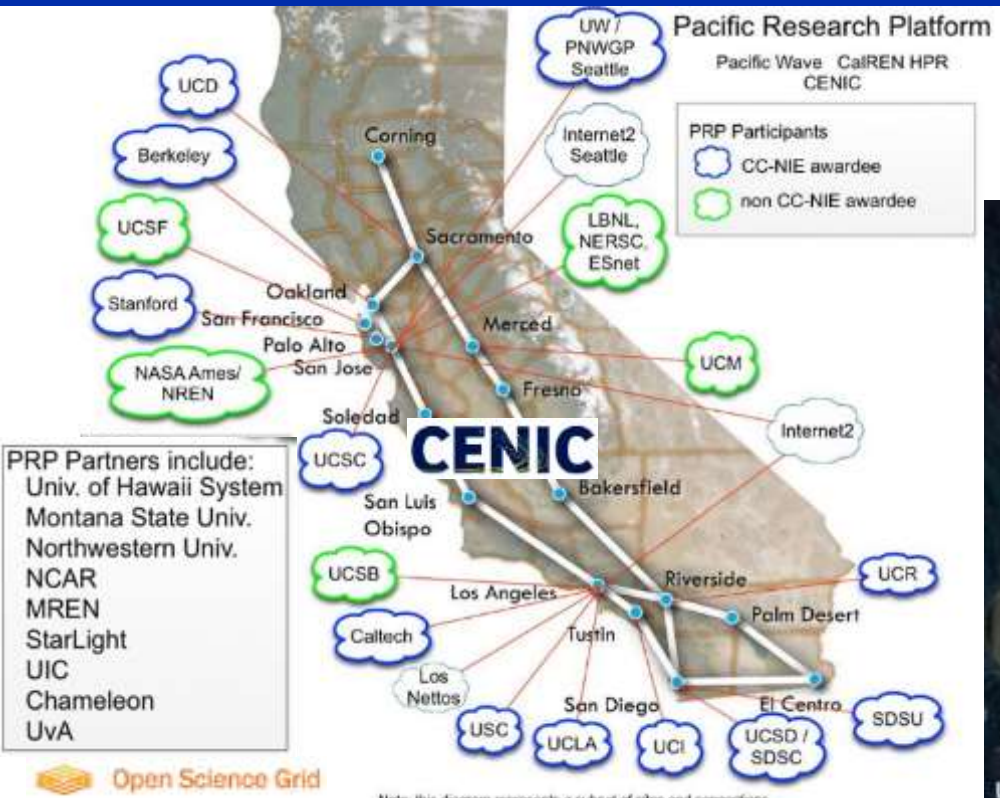
NSF Campus Cyberinfrastructure Program
2012-2022



Has Made Over 340 Awards:
Across 50 States and Territories

Slide Adapted From Kevin Thompson, NSF

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



NSF CC*DNI Grant
$6.3M 10/2015-10/2020
Extended – Ended Year 7 in Oct 2022



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://nationalresearchplatform.org/fiona/



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

NRP NATIONAL RESEARCH PLATFORM

NSF

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

SDSC

QUALCOMM INSTITUTE

2018/2019: Installing Community Shared FIONA CPU/GPU/Storage Systems on Campuses and Working With Campus CIOs on DMZs

UC Santa Cruz



UC Irvine



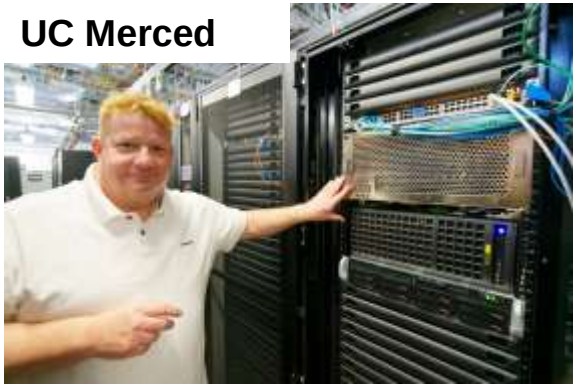
UC Santa Barbara



Stanford



UC Merced



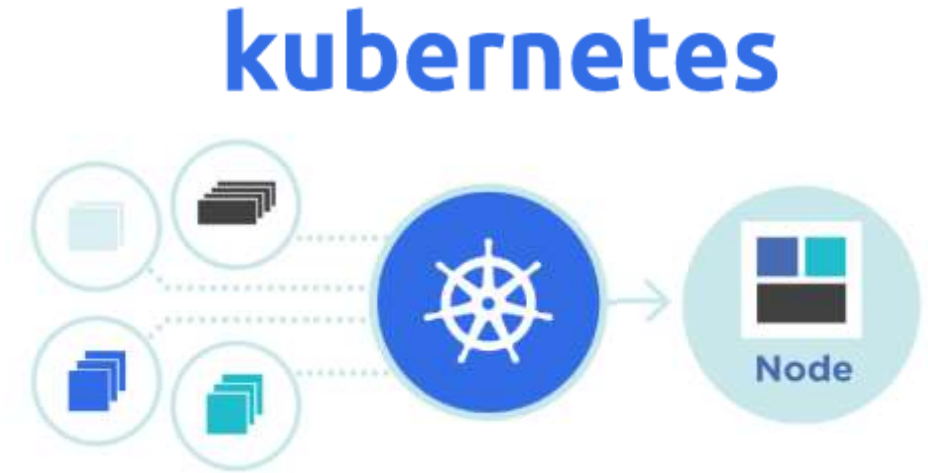
UC Riverside



NRP NATIONAL RESEARCH PLATFORM

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

User Applications



PACIFIC RESEARCH PLATFORM

Clouds

CLOUD NATIVE COMPUTING FOUNDATION

NATIONAL RESEARCH PLATFORM

PRP's Nautilus Hypercluster Adopted Kubernetes to Orchestrate Software Containers and Manage Distributed Storage





Kubernetes (K8s) is an open-source system for automating deployment, scaling, and management of containerized applications.

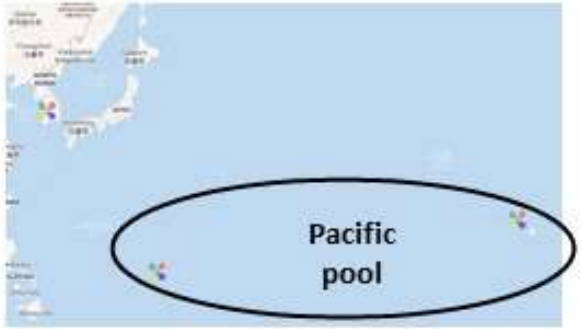




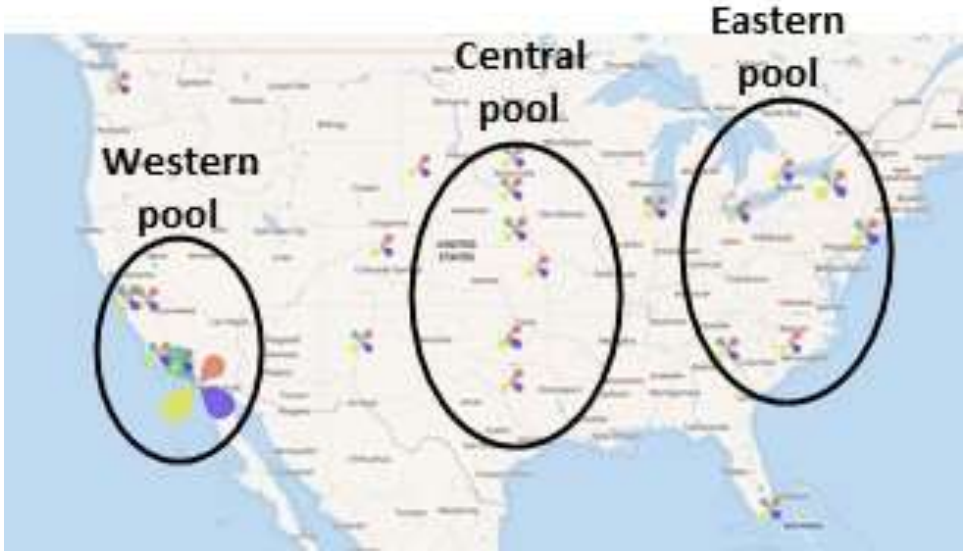
"Kubernetes with Rook/Ceph Allows Us to Manage Petabytes of Distributed Storage and GPUs for Data Science, While We Measure and Monitor Network Use."
--John Graham, Calit2/QI UC San Diego

NRP NATIONAL RESEARCH PLATFORM

Nautilus Has Established a Distributed Set of Ceph Storage Pools

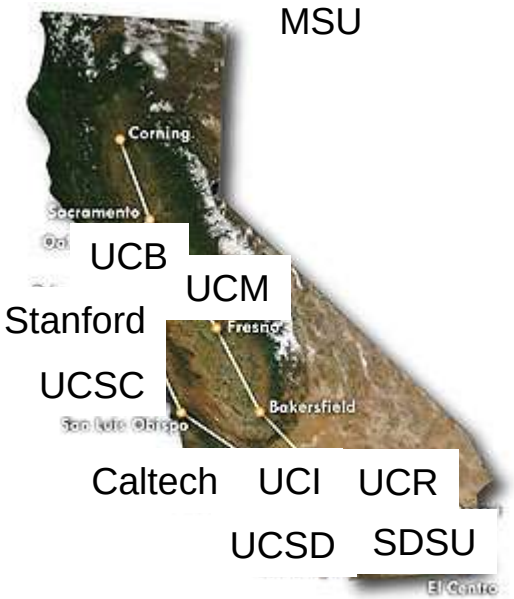


PRP forms optimal-scale Ceph pools with best performance and lowest latency



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

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
Co-PIs:

- Tajana Rosing
- Ken Kreutz-Delgado
- Ilkay Altintas
- Tom DeFanti

JACOBS SCHOOL OF ENGINEERING

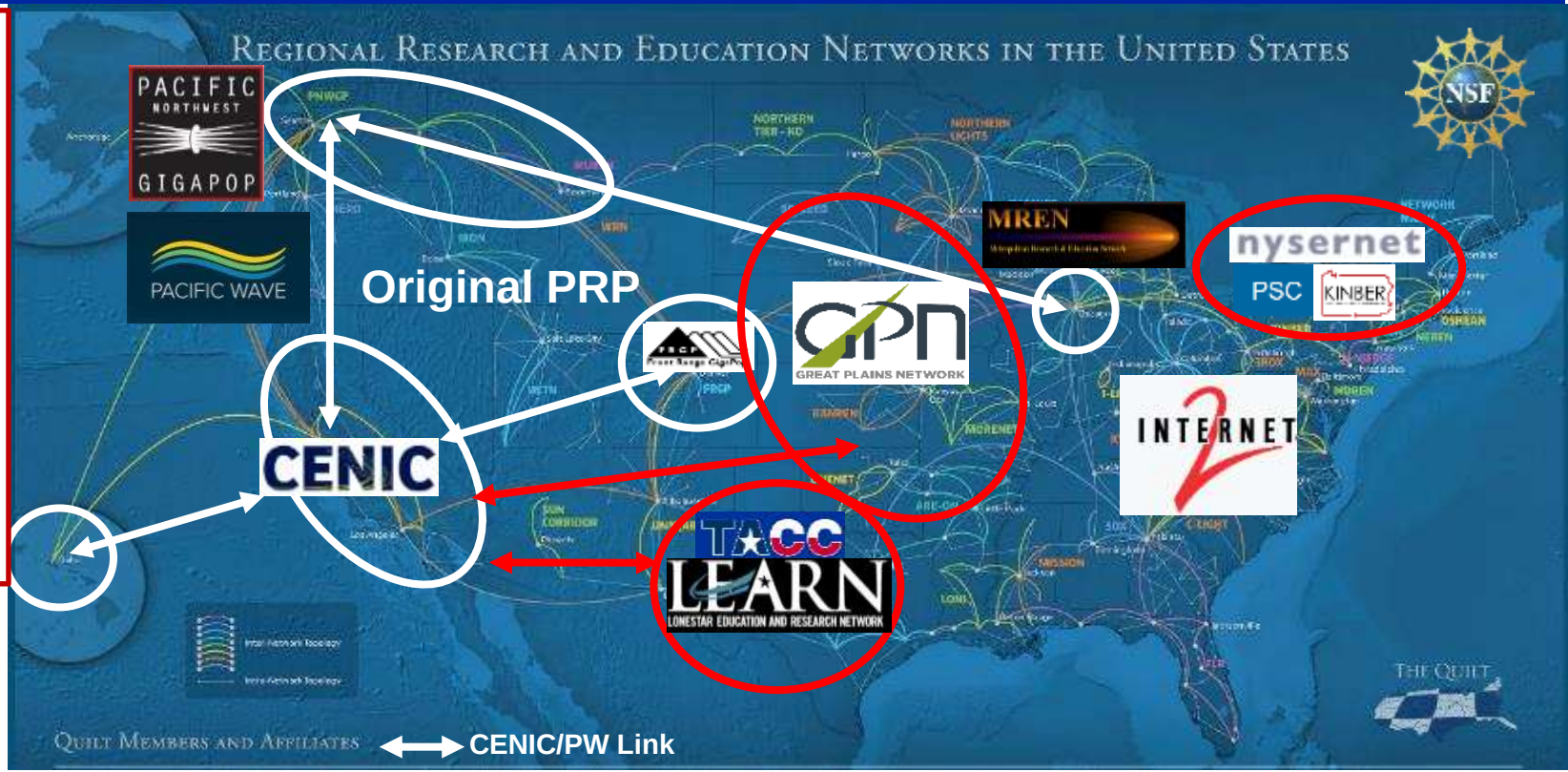
SDSC UC San Diego

NRP NATIONAL RESEARCH PLATFORM

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

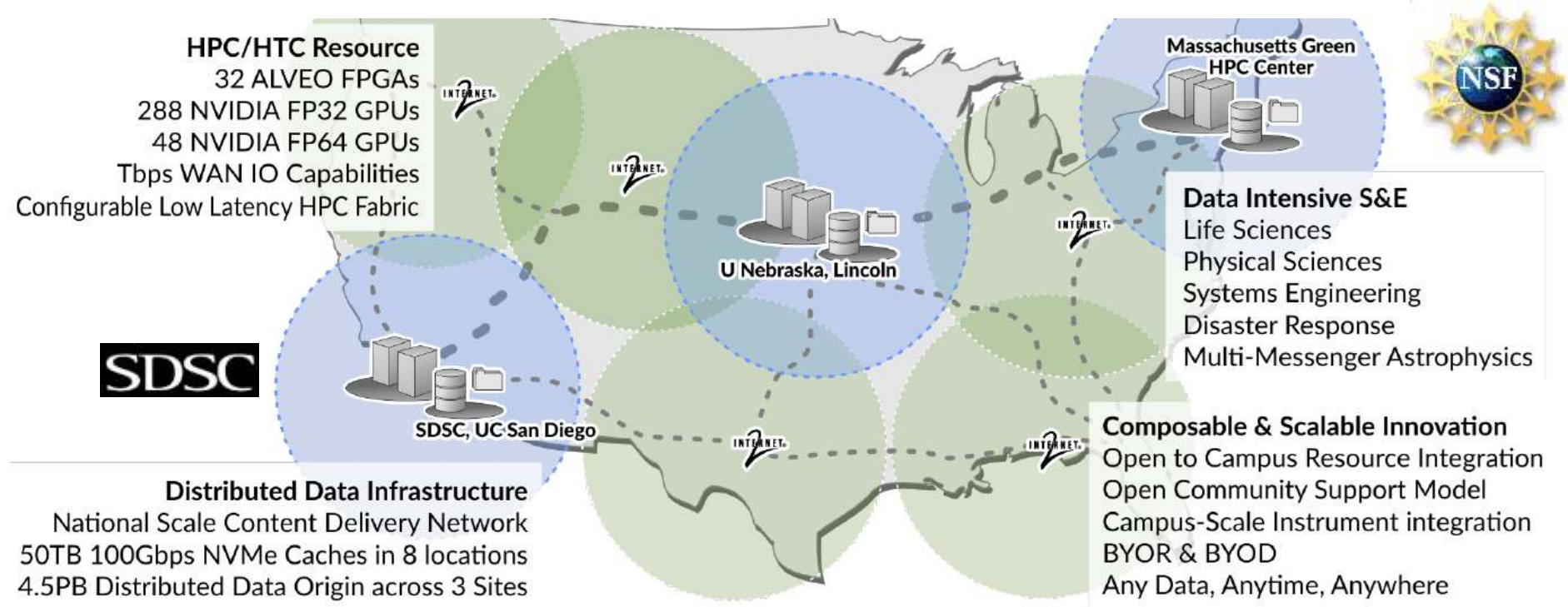
"Towards The NRP"
3-Year Grant Funded by NSF
$2.5M
October 2018
Award #1826967

PI Smarr
Co-PIs Altintas
Papadopoulos
Wuerthwein
Rosing
DeFanti



NATIONAL RESEARCH PLATFORM

2021-2026: PRP Federates with NSF-Funded *Prototype National Research Platform*



NSF Award OAC #2112167 (June 2021) [$5M Over 5 Years]
PI Frank Wuerthwein (UCSD, SDSC)
Co-PIs Tajana Rosing (UCSD), Thomas DeFanti (UCSD), Mahidhar Tatineni (SDSC), Derek Weitzel (UNL)

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

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

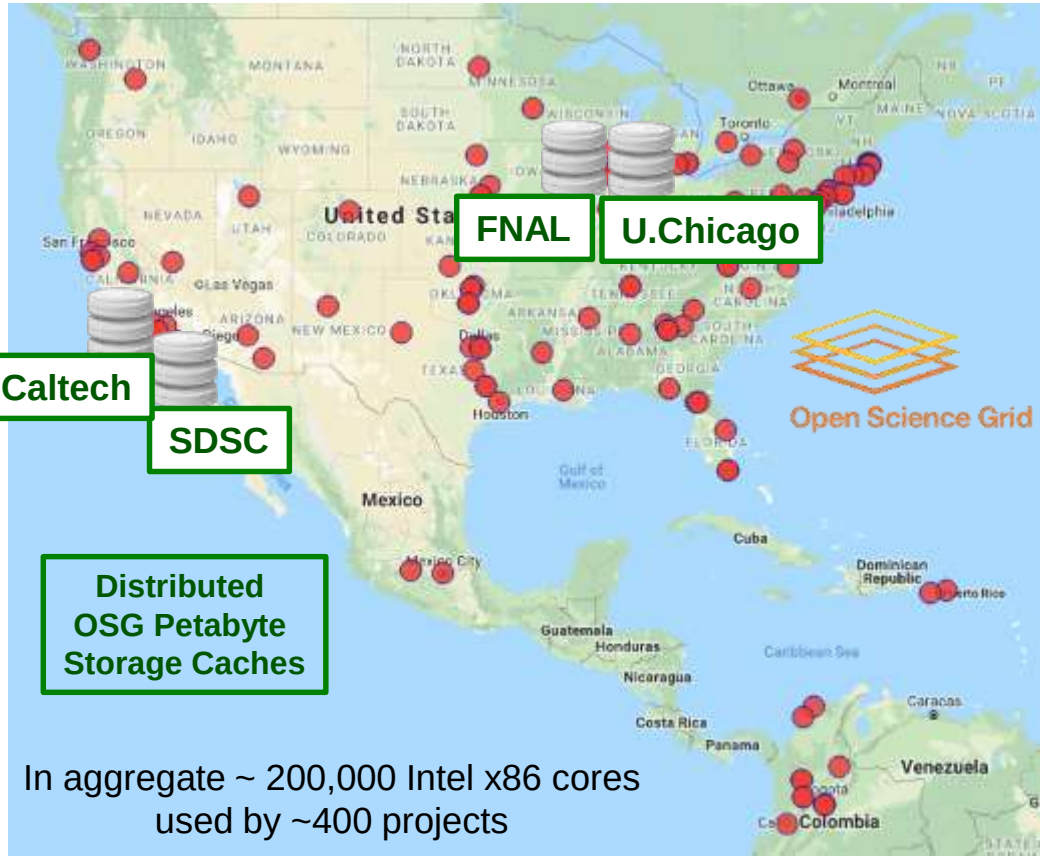
All OSG User Communities
Use HTCondor for Resource Orchestration

HTCondor
High Throughput Computing

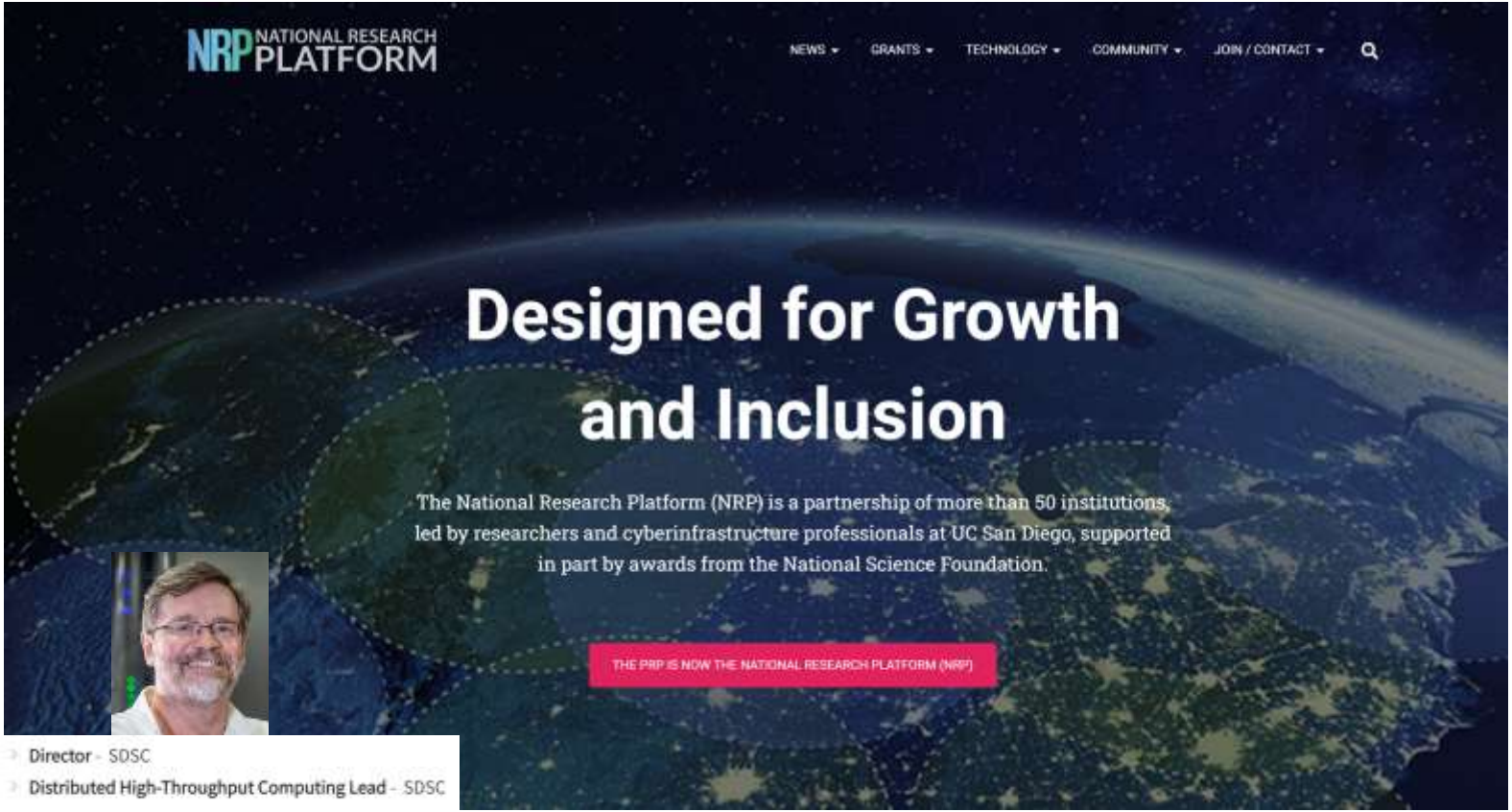
NSF

NRP NATIONAL RESEARCH PLATFORM

OSG Federates ~100 Clusters Worldwide



2023 - The National Research Platform Unifies The Previous 8 Years of NSF Grants



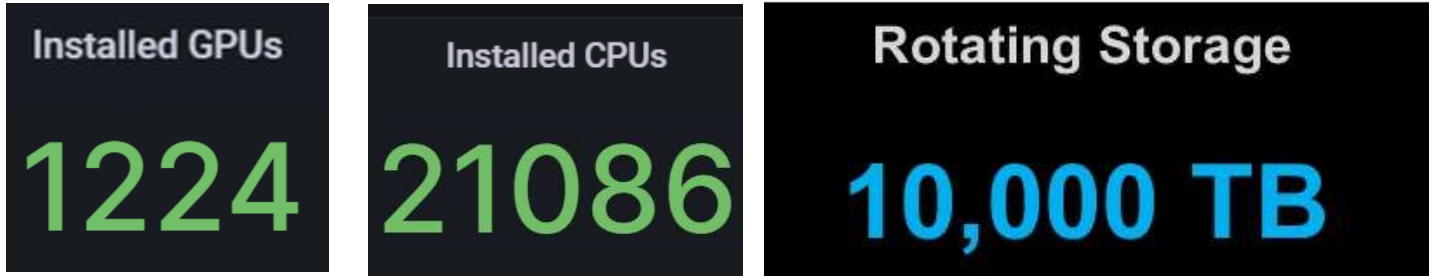
NRP NATIONAL RESEARCH PLATFORM

https://nationalresearchplatform.org/

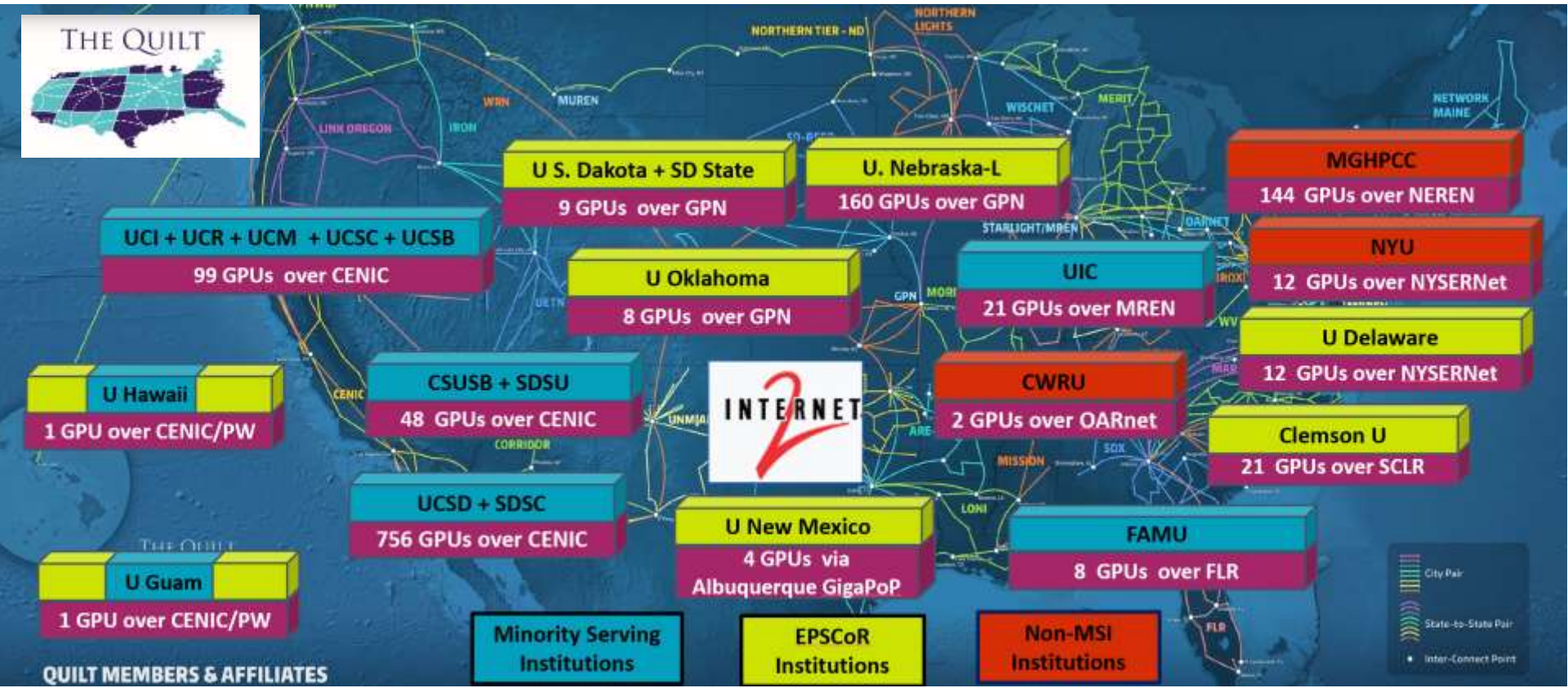
2023: NRP's Nautilus is a Multi-Institution National to Global Scale Hypercluster Connected by Optical Networks

December 14, 2023

~200 FIONAs on 30 Partner Campuses Networked Together at 10-100Gbps

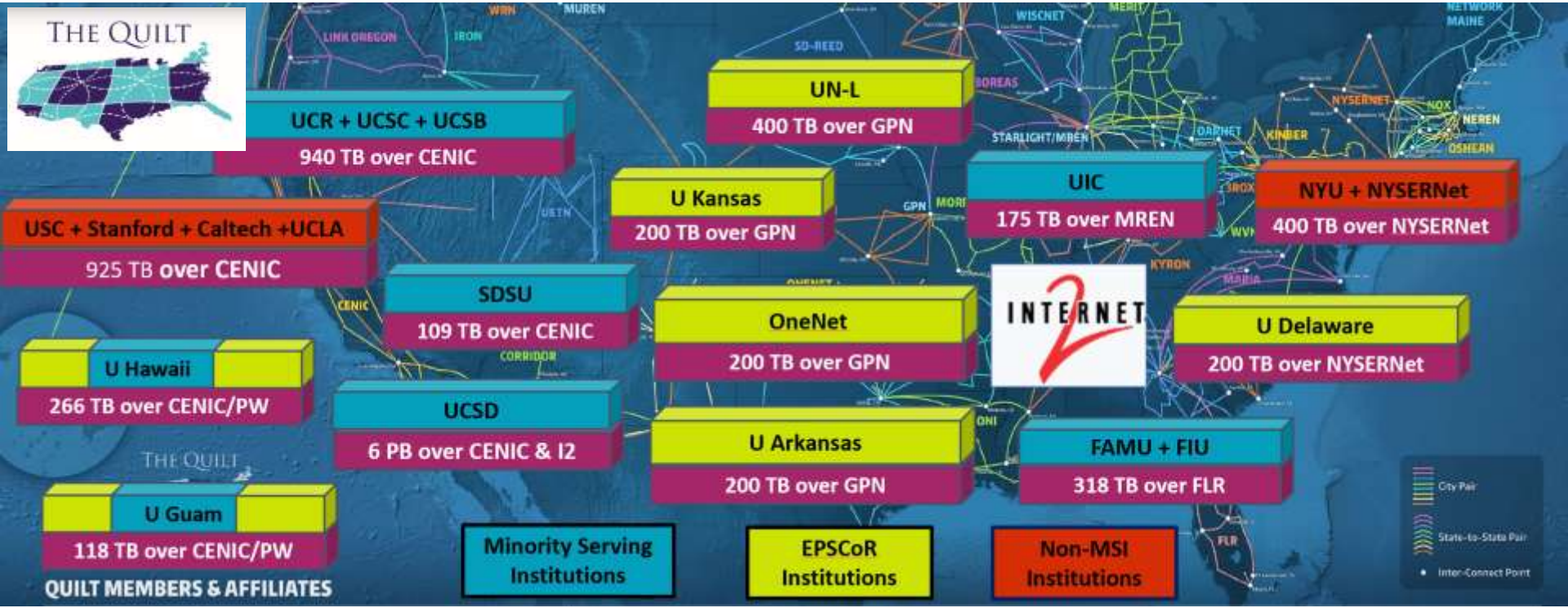


Nautilus' >21,000 CPU Cores and ~1,200 GPUs Distributed over US Networks Are Hosted on 30 Campuses, Connected by Regional Optical Networks & Internet2



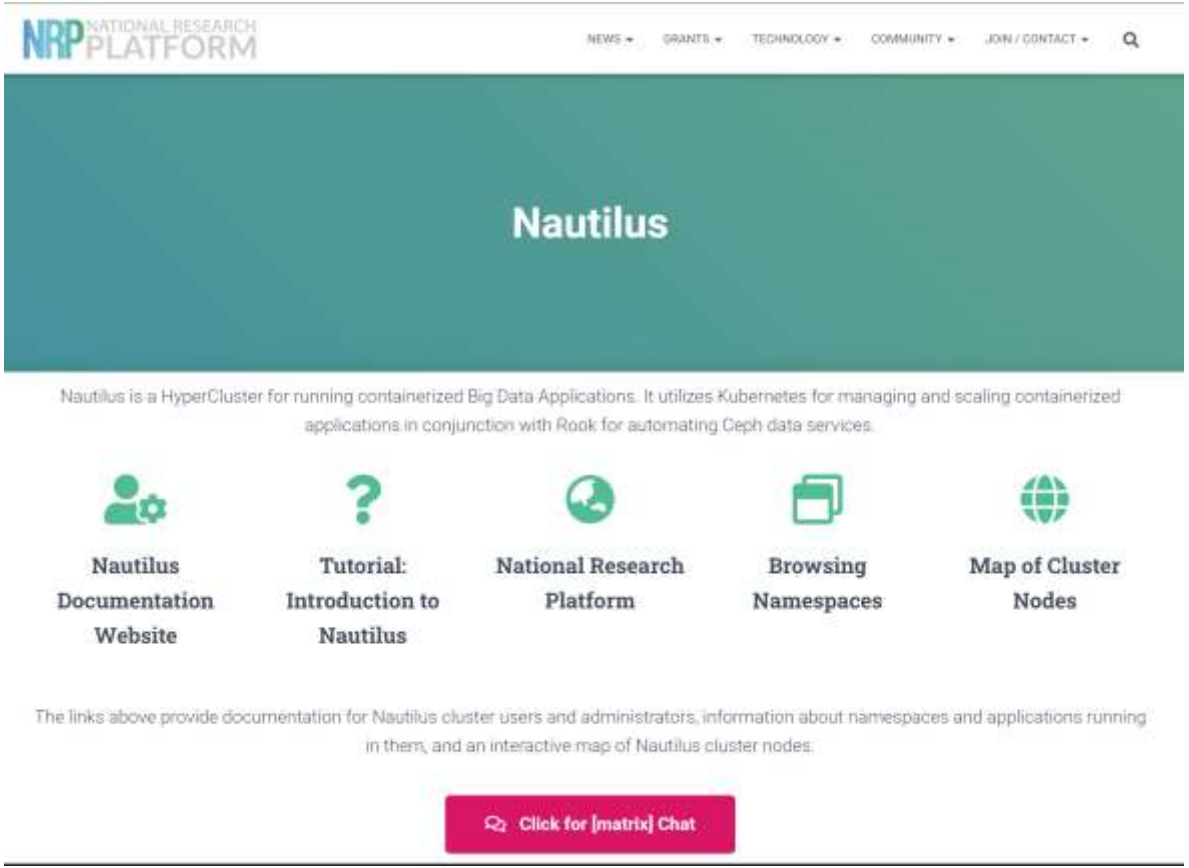
Slide Courtesy of Tom DeFanti, UCSD; PRP co-PI

Nautilus 10 PB Ceph Storage
Connected Over Regional Optical Networks & Internet2



Slide Courtesy of Tom DeFanti, UCSD; PRP co-PI

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



https://nationalresearchplatform.org/nautilus/

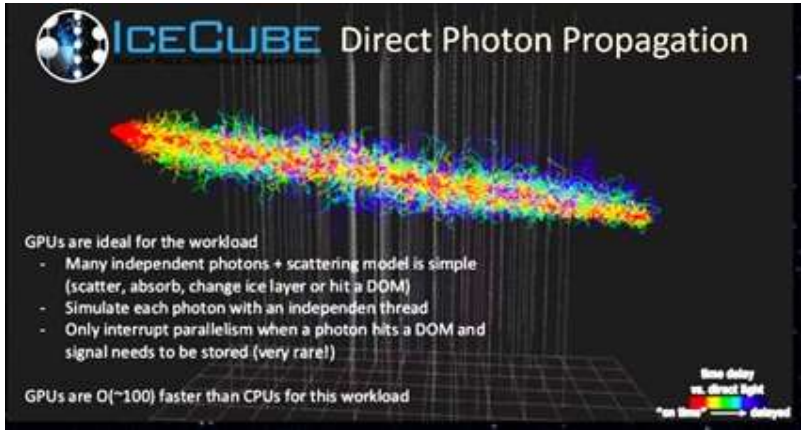
The December 2022 Pacific Research Platform Video Highlights 3 Different Applications Out of 1000 Nautilus Namespace Projects



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

NSF Large-Scale Observatories Are Using NRP 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 600 of RP's 1200
GPUs in 2022/3!

NRP National 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)

CENIC



UCSD SunCAVE (70 Screens, 70 GPUs)

See These VR Facilities in Action in the PRP Video

NRP National Research Platform

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

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



HPWREN

Real-Time Meteorological Sensors

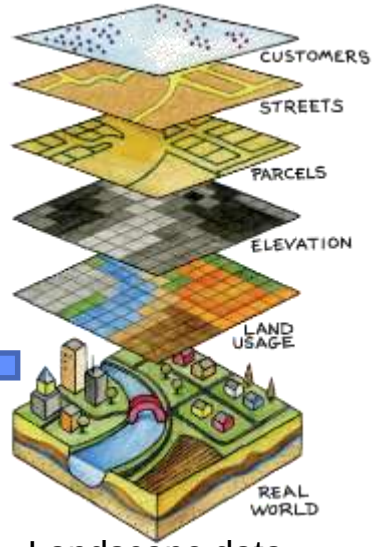


Fire Perimeter

Source: Ilkay Altintas, SDSC

WIFIRE

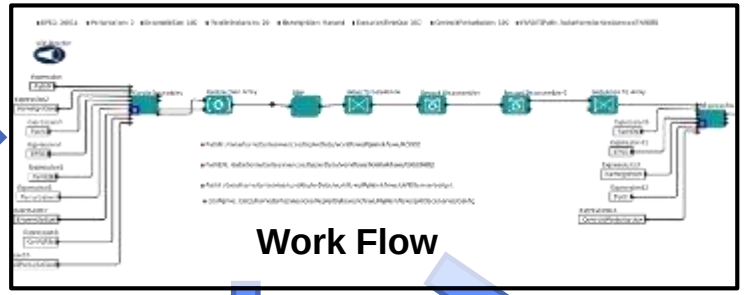




Landscape data

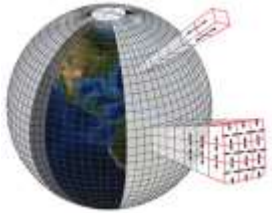
NRP

CENIC



Work Flow

Weather Forecasts





SDSC





NSF

NRP NATIONAL RESEARCH PLATFORM

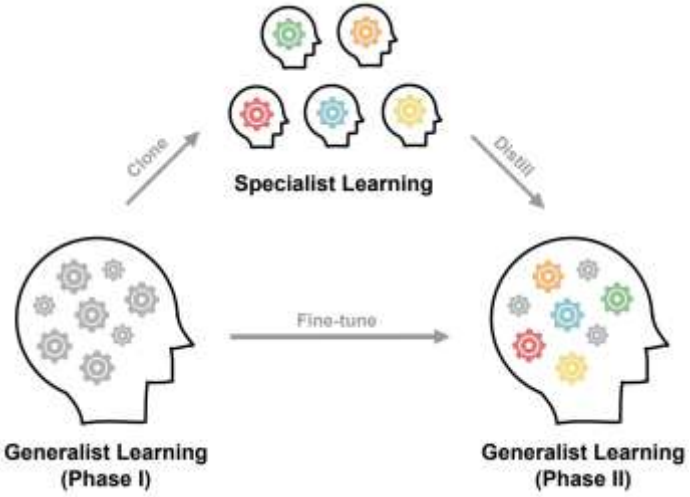
The Rise of Machine Learning ML Applications on the NRP

NRP NATIONAL RESEARCH PLATFORM

A Major Project in UCSD's Hao Su Lab is Large-Scale Robot Learning

- **We Build A Digital Twin of The Real World in Virtual Reality (VR) For Object Manipulation**
- **Agents Evolve In VR**
  - **Specialists (Neural Nets) Learn Specific Skills by Trial and Error**
  - **Generalists (Neural Nets) Distill Knowledge to Solve Arbitrary Tasks**

- On **Nautilus**:
  - *Hundreds* of specialists have been trained
  - Each specialist is trained in *millions* of environment variants
  - ~10,000 GPU hours per run



NRP NATIONAL RESEARCH PLATFORM

UCSD's Ravi Group: How to Create Visually Realistic 3D Objects or Dynamic Scenes in VR or the Metaverse

ML Computing Transforms a Series of 2D Images
Into a 3D View Synthesis





NRP NATIONAL RESEARCH PLATFORM

Source: Prof. Ravi Ramamoorthi, UCSD

Machine Learning-Based
Neural Radiance Fields for View Synthesis (NeRFs) Are Transformational!

Meta Is Investing in NeRF for 3D Object Scanning

https://youtu.be/hvfV-iGwYX8

The New York Times R&D

Spatial Journalism

Creating Workflows for NeRF Portraiture

September 27, 2022

← THE BEST INVENTIONS OF 2022

TIME

Automated 3D Rendering

NVIDIA Instant NeRF

BY JARED LINDZON
NOVEMBER 10, 2022

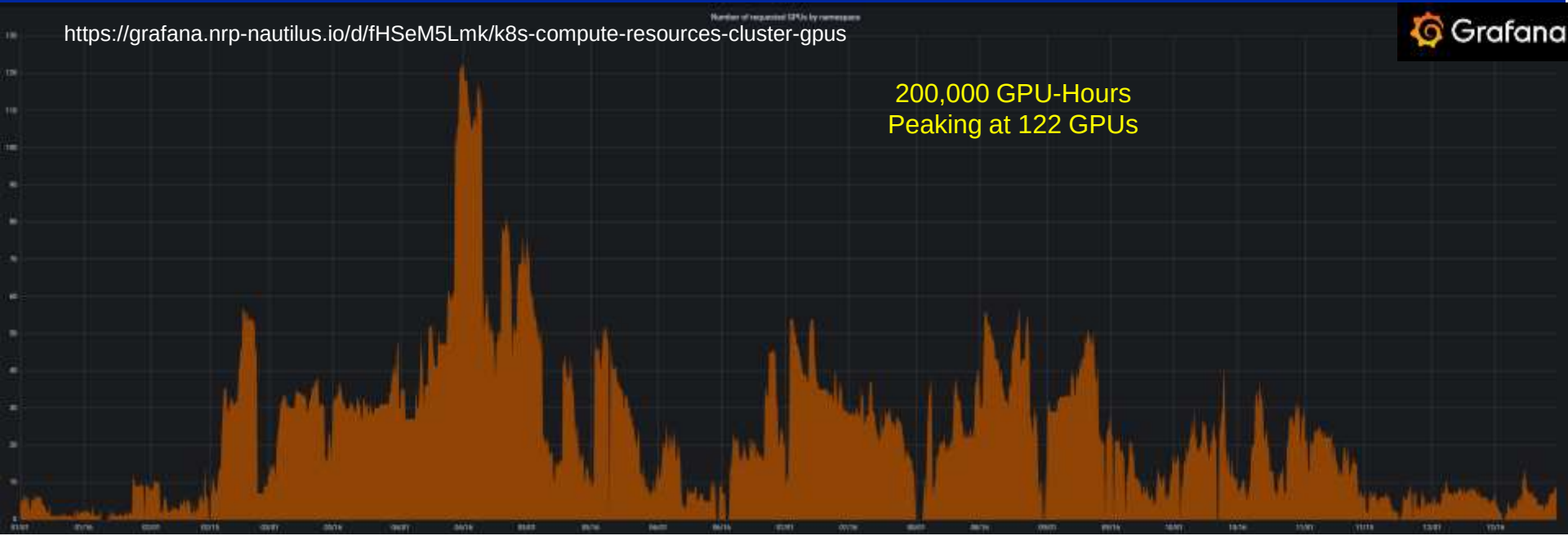
A neural radiance field (NeRF) is a fully-connected neural network that can generate novel views of complex 3D scenes, based on a partial set of 2D images.

https://datagen.tech/guides/synthetic-data/neural-radiance-field-nerf/

Source: Prof. Ravi Ramamoorthi, UCSD

Namespace *ucsd-ravigroup*
Consumed the 3nd Most Nautilus GPU-Hours in 2022

https://grafana.nrp-nautilus.io/d/fHSeM5Lmk/k8s-compute-resources-cluster-gpus



- Much of the compute involves training computationally expensive NeRFs.
- Training time to learn a representation of a single scene on a GPU can vary from seconds to a day.
- NeRFs that can see behind occlusions may require a week of training on 8 GPUs simultaneously.

NRP NATIONAL RESEARCH PLATFORM

Source: Alexander Trevithick, UCSD Ravi Group

California State University San Bernardino is an Excellent Example Of How to Help Your Faculty and Students Use NRP



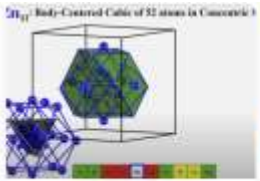
Their Campus HPC Program
Enabled CSUSB Faculty & Students
to Use More NRP GPU-Hours
In the Last 12 Months
Than 7 of the 10 UC Campuses!

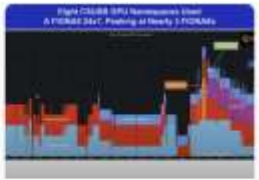


JupyterHub, NRP



ACCESS


Projects


2023 Faculty Showcase

www.csusb.edu/academic-technologies-innovation/xreal-lab-and-high-performance-computing/high-performance-computing

NRP NATIONAL RESEARCH PLATFORM

The Key Role of Regional Optical Network Meetings to Engage More Campuses in Using NRP



Jen Leasure
President & CEO,
The Quilt



See https://nationalresearchplatform.org/events/fourth-national-research-platform-4nrp/
for slides and videos of all 4NRP presentations



www.thequilt.net/quilt-circle/snapshot-scaling-a-national-research-platform/

NRP NATIONAL RESEARCH PLATFORM

University of Missouri's Grant Scott
Has Built a Model Program for Data Intensive Computing and ML/AI Using NRP

GREAT PLAINS NETWORK

ENGINEERS SHARE EXPERTISE AROUND NAUTILUS WITH GREAT PLAINS NETWORK

June 12, 2023

University of Missouri



https://engineering.missouri.edu/2023/engineers-share-expertise-around-nautilus-with-great-plains-network/

NRP NATIONAL RESEARCH PLATFORM

UCSD's Information Technology Services Has Adapted PRP FIONA8s To Support Data Science Courses

- **Student-Focused GPU/CPU Cluster For:**
  - Undergraduate & Graduate Coursework
  - For-Credit Independent Study
  - Thesis/Dissertation Research
  - Capstones & Projects
- **Research-Driven Architecture**
- **Managed by Central IT Services**
- **Racked FIONA Clusters**
  - 124 32-bit GPUs
  - 660 CPU-cores



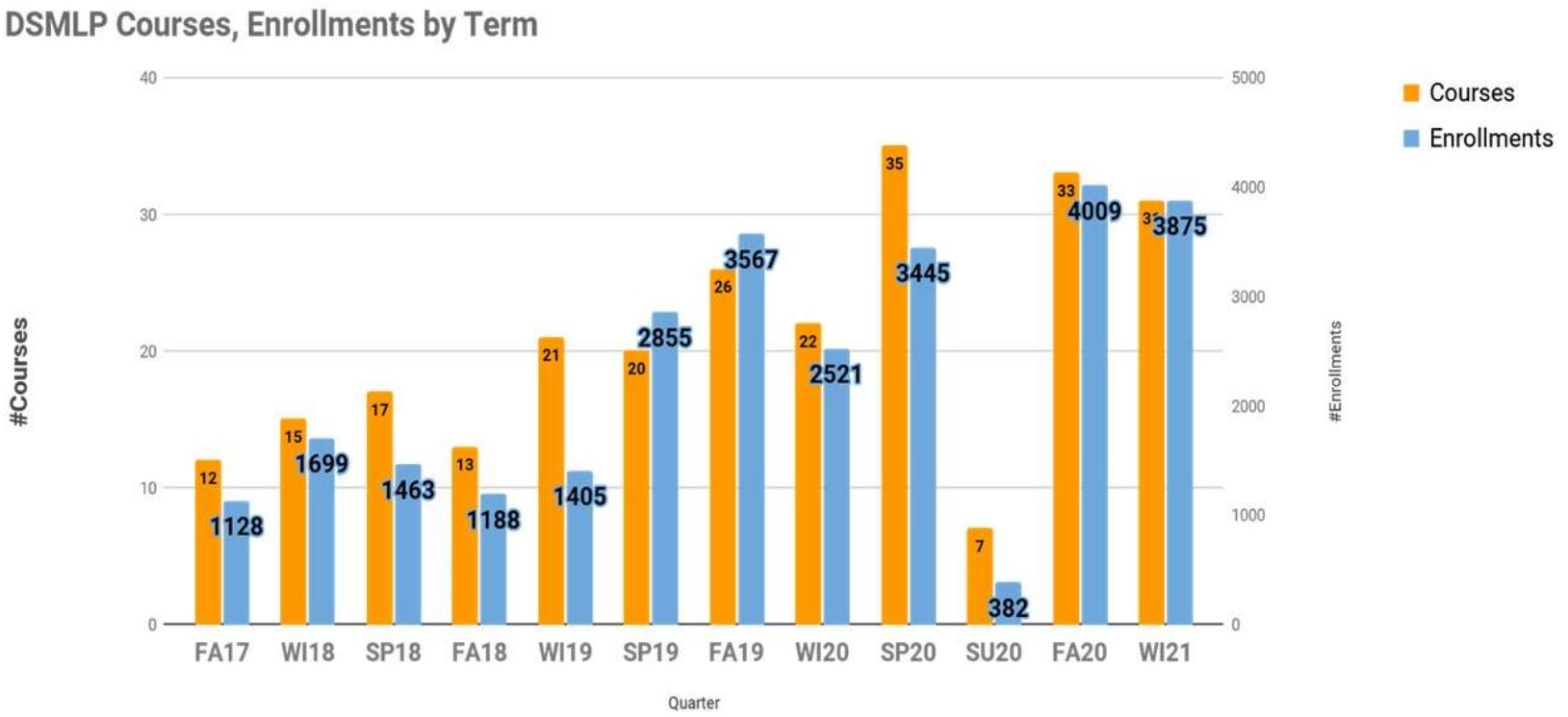
NRP NATIONAL RESEARCH PLATFORM

UC San Diego IT SERVICES

UC San Diego Jacobs School of Engineering

UCSD Cognitive Science

The UCSD Data Science and Machine Learning Platform Supports 2-3 Dozen Courses with 2500-4000 Students Per Quarter

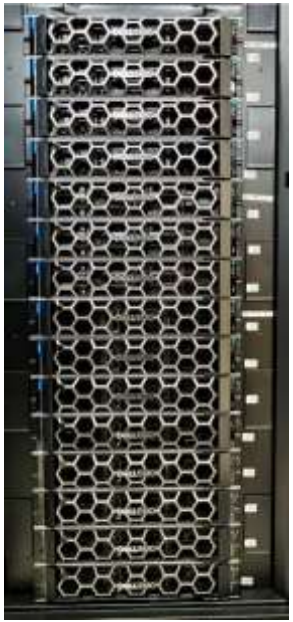


Source: Adam Tilghman, ITS
https://datahub.ucsd.edu/hub
www.hpcwire.com/off-the-wire/uc-san-diegos-jupyterhub-platform-aids-students-with-data-intensive-computing-needs/

NRP NATIONAL RESEARCH PLATFORM

SDSU's VERNE Expands the UCSD Instructional Model by Federating with NRP: Visionary, Education, Research, Network, Ecosystem

SDSU Information Technology Division



CENIC

jupyter TensorFlow PyTorch kubernetes NVIDIA docker

- Dell PowerEdge Cluster for Instructional Use, with 15 Nodes Containing 32 GPUs, 768 CPU cores, and 960 TB Storage.
- A Learning Resource that *Surges* for Research and Instruction
- System Administration as a Service from the NRP
- Friendly User mode this Spring with
  - ~90 Users in Computational Engineering 361 & Astronomical Techniques ASTR 350
  - Research in Developing Synthetic Medical Data for Machine Learning Training

NRP NATIONAL RESEARCH PLATFORM

Slide courtesy of Jerry Sheehan, SDSC CIO

The PRP Has Emphasized Expanding Diversity and Inclusion

- **When the PRP Grant Was Funded in 2015, It Started With:**
  - **6 States ➡ Now 45 States**
  - **19 Campuses ➡ Now 135 Campuses**
  - **9 Minority Serving Institutions ➡ Now 24 MSIs**
  - **2 NSF EPCoR States ➡ Now 20 EPSCoR States, 2 Territories, and Wash DC**

NRP NATIONAL RESEARCH PLATFORM

2015 - United States PRP Campus Users In the Beginning

California

- California Institute of Technology
- **San Diego State U. (MSI)**
- Stanford U.
- U. of California, Berkeley
- **U. of California, Merced (MSI)**
- **U. of California, Davis (MSI)**
- **U. of California, Riverside (MSI)**
- U. of California, San Francisco
- **U. of California, Santa Cruz (MSI)**
- **U. of California-Irvine (MSI)**
- U. of California-Los Angeles
- U. of California-San Diego
- **U. of California-Santa Barbara (MSI)**
- U. of Southern California

Hawaii (EPSCoR)

- **U. of Hawaii (MSI)**

Washington

- U. of Washington

Montana (EPSCoR)

- Montana State U- Bozeman

Colorado

- NCAR

Illinois

- Northwestern University
- **U. of Illinois at Chicago (MSI)**

Minority Serving Institutions
NSF EPSCoR States

NRP NATIONAL RESEARCH PLATFORM

Oct 2015

2023 -Expanded Diversity & Inclusion - Nautilus Campus Users Nationwide: West

California
- California Institute of Technology
- **California State Fullerton (MSI)**
- **California State Polytechnic U., Humboldt (MSI)**
- **California State Monterey Bay (MSI)**
- **California State Polytechnic U., Pomona (MSI)**
- **California State U. San Bernardino (MSI)**
- **California State U., Northridge (MSI)**
- Harvey Mudd College
- Naval Postgraduate School
- **San Diego State U. (MSI)**
- **San Jose State U. (MSI)**
- Stanford U.
- U. of California, Berkeley
- **U. of California, Davis (MSI)**
- **U. of California-Irvine (MSI)**
- U. of California-Los Angeles
- **U. of California, Merced (MSI)**
- **U. of California, Riverside (MSI)**
- U. of California-San Diego
- U. of California, San Francisco
- **U. of California-Santa Barbara (MSI)**
- **U. of California, Santa Cruz (MSI)**
- U. of Southern California

Hawaii (EPSCoR)
- **U. of Hawaii (MSI)**

Guam (EPSCoR)
- **U. of Guam (MSI)**

Oregon
- Oregon State U.
- Willamette U.

Washington
- U. of Washington
- Washington State U.

Nevada (EPSCoR)
- U. of Nevada, Reno

Montana (EPSCoR)
- Montana State U- Bozeman
- **Salish Kootenai Tribal College (MSI)**

Wyoming (EPSCoR)
- U. of Wyoming

Colorado
- Colorado School of Mines
- NCAR
- U. of Colorado at Boulder

Utah
- Brigham Young U.
- U of Utah

New Mexico (EPSCoR)
- U. of New Mexico

Arizona
- Arizona State U
- U. of Arizona

Minority Serving Institutions
NSF EPSCoR States

NRP NATIONAL RESEARCH PLATFORM

June 2023

Expanded Diversity & Inclusion - Nautilus Campus Users Across the Country: Midwest

North Dakota (EPSCoR)
- North Dakota State U.

South Dakota (EPSCoR)
- Black Hills State U.
- South Dakota School of Mines & Tech.
- South Dakota State U.
- U. of South Dakota

Nebraska (EPSCoR)
- U. of Nebraska at Kearney
- U. of Nebraska-Lincoln

Kansas (EPSCoR)
- Kansas State U.
- U. of Kansas
- Wichita State U

Oklahoma (EPSCoR)
- Oklahoma State U. System
- Southwestern Oklahoma State U.
- U. of Central Oklahoma
- U. of Oklahoma

Texas
- Baylor U.
- Southern Methodist U.
- U. of Texas at Austin
- U. of Texas at Dallas

Minnesota
- U of Minnesota

Iowa (EPSCoR)
- Iowa State U.

Missouri
- Culver-Stockton College
- Missouri Western State U
- Truman State U
- U of Central Missouri
- U. of Missouri - Columbia
- Washington U. in St. Louis

Illinois
- Illinois Institute of Technology
- Northwestern University
- U. of Chicago
- **U. of Illinois at Chicago (MSI)**
- U of Illinois, Urbana-Champaign

Arkansas (EPSCoR)
- U. of Arkansas
- U. of Arkansas for Medical Sciences
- U. of Arkansas, Little Rock

Louisiana (EPSCoR)
- Louisiana Tech U.
- Southeastern Louisiana U.

Wisconsin
- U. of Wisconsin-Madison
- U. of Wisconsin-Milwaukee

Michigan
- Michigan State U.
- U. of Michigan

Ohio
- Case Western Reserve U.
- Kent State University
- Ohio State U.
- U. of Akron
- U. of Cincinnati

Indiana
- Indiana U.
- Purdue U.
- U. of Notre Dame

Kentucky (EPSCoR)
- U. of Kentucky

Tennessee
- Tennessee Tech. U.
- Vanderbilt U.

Minority Serving Institutions
NSF EPSCoR States

NRP NATIONAL RESEARCH PLATFORM

June 2023

Expanded Diversity & Inclusion - Nautilus Campus Users Across the Country: East

Maine (EPSCoR)
- Colby College

Massachusetts
- Boston U.
- MIT
- Northeastern U.

Connecticut
- U. of Connecticut
- Yale U.

Rhode Island (EPSCoR)
- Brown U.
- U. of Rhode Island

New York
- American Museum of Natural History
- Columbia U.
- Manhattan College
- New York U.
- Rensselaer Polytechnic Institute
- Rochester Institute of Technology
- Syracuse U.
- The State U. of New York at Buffalo

Pennsylvania
- Carnegie Mellon U.
- Pennsylvania State U.
- U of Pennsylvania

New Jersey
- Princeton U.
- Rutgers U.

Delaware (EPSCoR)
- U. of Delaware

Maryland
- **Bowie State U. (MSI)**
- Johns Hopkins U.
- **Morgan State U. (MSI)**

District of Columbia
- American U.

West Virginia (EPSCoR)
- West Virginia U.

Virginia
- George Mason U.
- Virginia Commonwealth U.

North Carolina
- Duke U.
- **N. Carolina Agricultural & Tech. State U. (MSI)**
- N. Carolina State U
- U. of North Carolina at Chapel Hill

South Carolina (EPSCoR)
- Clemson U.

Georgia
- Georgia Institute of Technology
- **Georgia State U. (MSI)**
- U. of Georgia

Alabama(EPSCoR)
- Auburn U.
- U of Alabama in Huntsville

Florida
- **Florida A&M U. (MSI)**
- **Florida International U. (MSI)**
- Florida State U.
- **U. of Central Florida (MSI)**
- U. of Miami

Puerto Rico (EPSCoR)
- **U. Puerto Rico (MSI)**

Minority Serving Institutions
NSF EPSCoR States

NRP NATIONAL RESEARCH PLATFORM

June 2023

Community Building Though Inclusion and Diversity: Workshops With Minority Serving Universities

2NRP Montana State University







NRP NATIONAL RESEARCH PLATFORM

CITRIS THE BANATAO INSTITUTE

PRP/TNRP/CHASE-CI/PNRP/NRP
Support and Community

- US National Science Foundation (NSF) awards and subawards to UCSD

  --CNS-1456638, CNS-1730158, CNS-2100237, CNS-2120019

  --ACI-1540112, ACI-1541349, OAC-1826967, OAC-2029306, OAC-2112167



- DOD DURIP awards to UCSD
- UC Office of the President, Calit2 and Calit2's UCSD Qualcomm Institute
- San Diego Supercomputer Center and UCSD's Research IT and Instructional IT
- Partner Campuses: UCB, UCSC, UCI, UCR, UCLA, USC, UCD, UCSB, SDSU, Caltech, NU, UWash, UChicago, UIC, UHM, CSUSB, UMo, FAMU, MSU, NYU, UNeb, UNM, UNC, FIU, UDel, UDak, SDakSU, Stanford, UArk, UOk, UoGuam, UKansas, CWRU, Clemson, MGHPCC, KISTI, UVA, AIST
- CENIC, Pacific Wave/PNWGP, StarLight/MREN, The Quilt, Great Plains Network, NYSERNet, Open Science Grid, Internet2, DOE ESnet, NCAR/UCAR & Wyoming Supercomputing Center,
- AWS, Google, Microsoft, Cisco, Juniper, Arista

Source: Tom DeFanti, UCSD