“The Emergence of the National Research Platform: With University of Missouri Leadership”

Invited Seminar

University of Missouri-Columbia

September 23, 2024

Dr. Larry Smarr

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

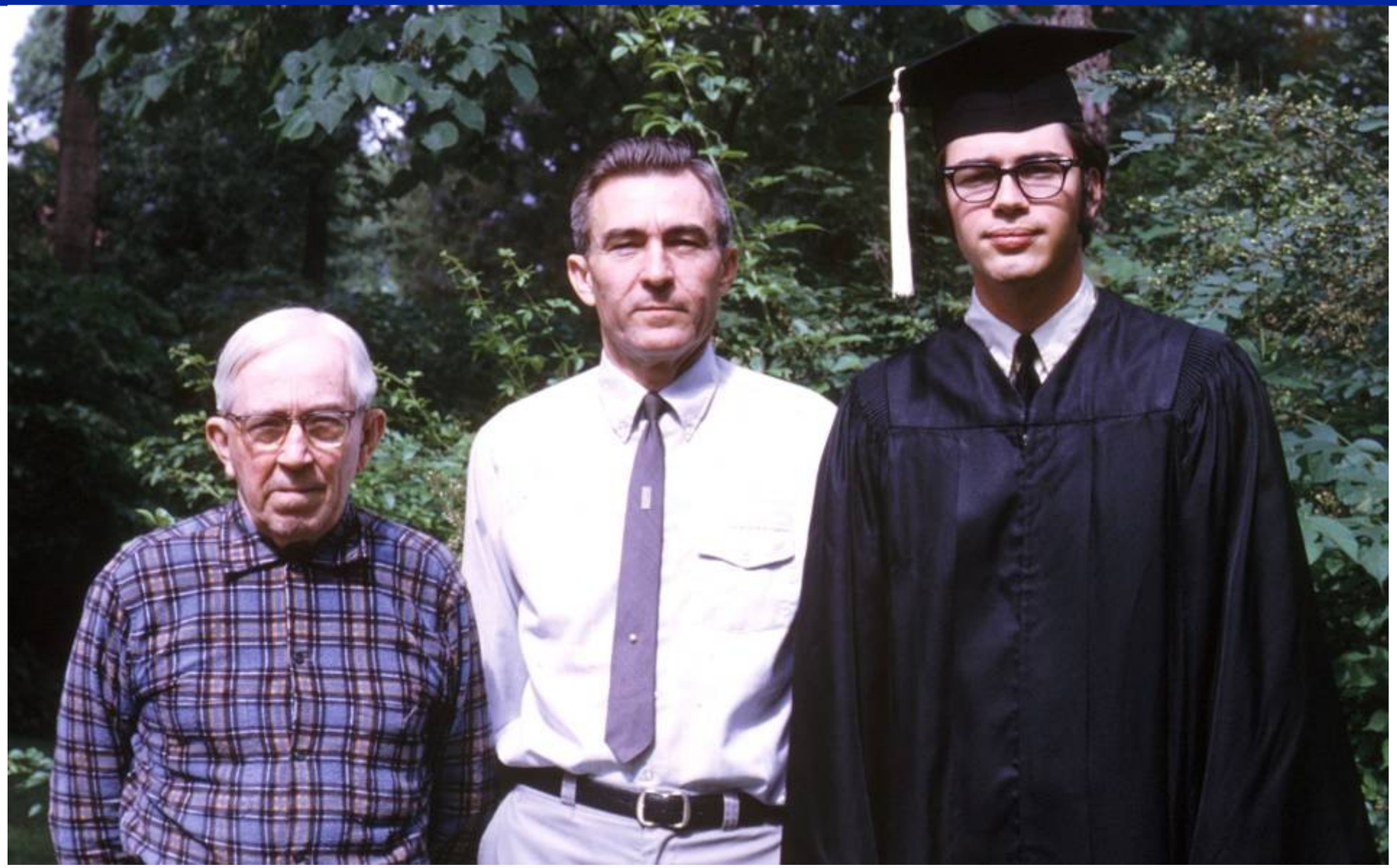
Distinguished Professor Emeritus, Dept. of Computer Science and Engineering

Jacobs School of Engineering, UCSD

http://lsmarr.calit2.net

NRP NATIONAL RESEARCH PLATFORM

I Graduated From MU in 1970



Grandpa Father Me

NRP NATIONAL RESEARCH PLATFORM

I Received an Honorary PhD From MU in May 2008



https://lsmarr.net/honorary-phd-from-the-university-of-missouri/

NRP NATIONAL RESEARCH PLATFORM

The National Research Platform is the Evolution of Metacomputer Architecture I Described in My Last MU Talk 16 Years Ago

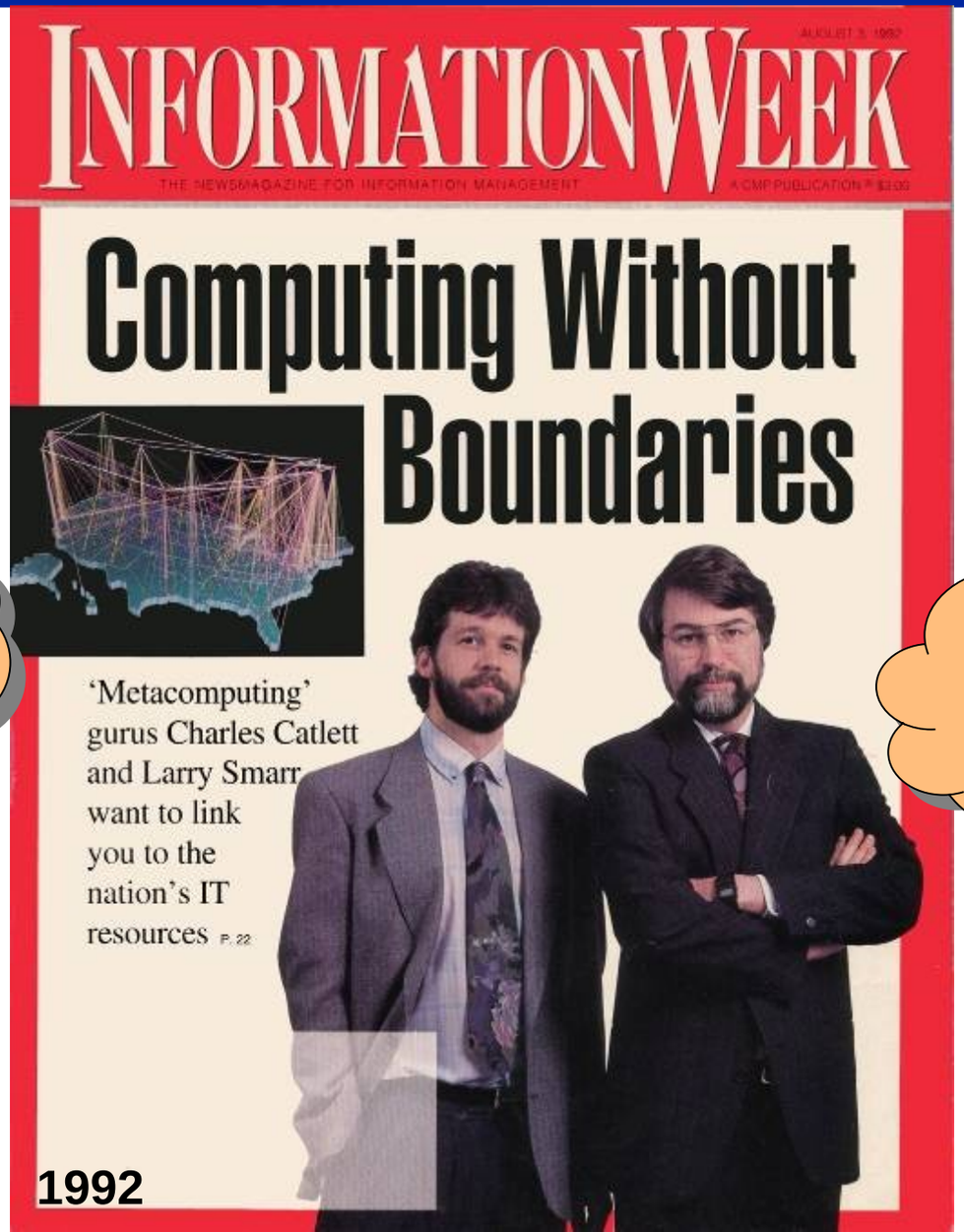
"Metacomputer Architecture of the Global LambdaGrid: How Personal Light Paths are Transforming e-Science"

Invited Talk
Departments of Computer Science / Physics and Astronomy
University of Missouri@Columbia
Columbia, MO
May 15, 2008

Dr. Larry Smarr
Director, California Institute for Telecommunications and Information Technology
Harry E. Gruber Professor,
Dept. of Computer Science and Engineering
Jacobs School of Engineering, UCSD

https://lsmarr.net/metacomputer-architecture-of-the-global-lambdagrid-how-personal-light-paths-are-transforming-e-science/

“Metacomputer” Led to the NTG, TeraGrid, OptIPuter, and NRP: 36 Years of Development



“Metacomputer” Coined by Smarr in 1988

TeraGrid PI

OptIPuter PI

Metacomputing:
L. Smarr, C. E. Catlett, Comm. of the ACM,

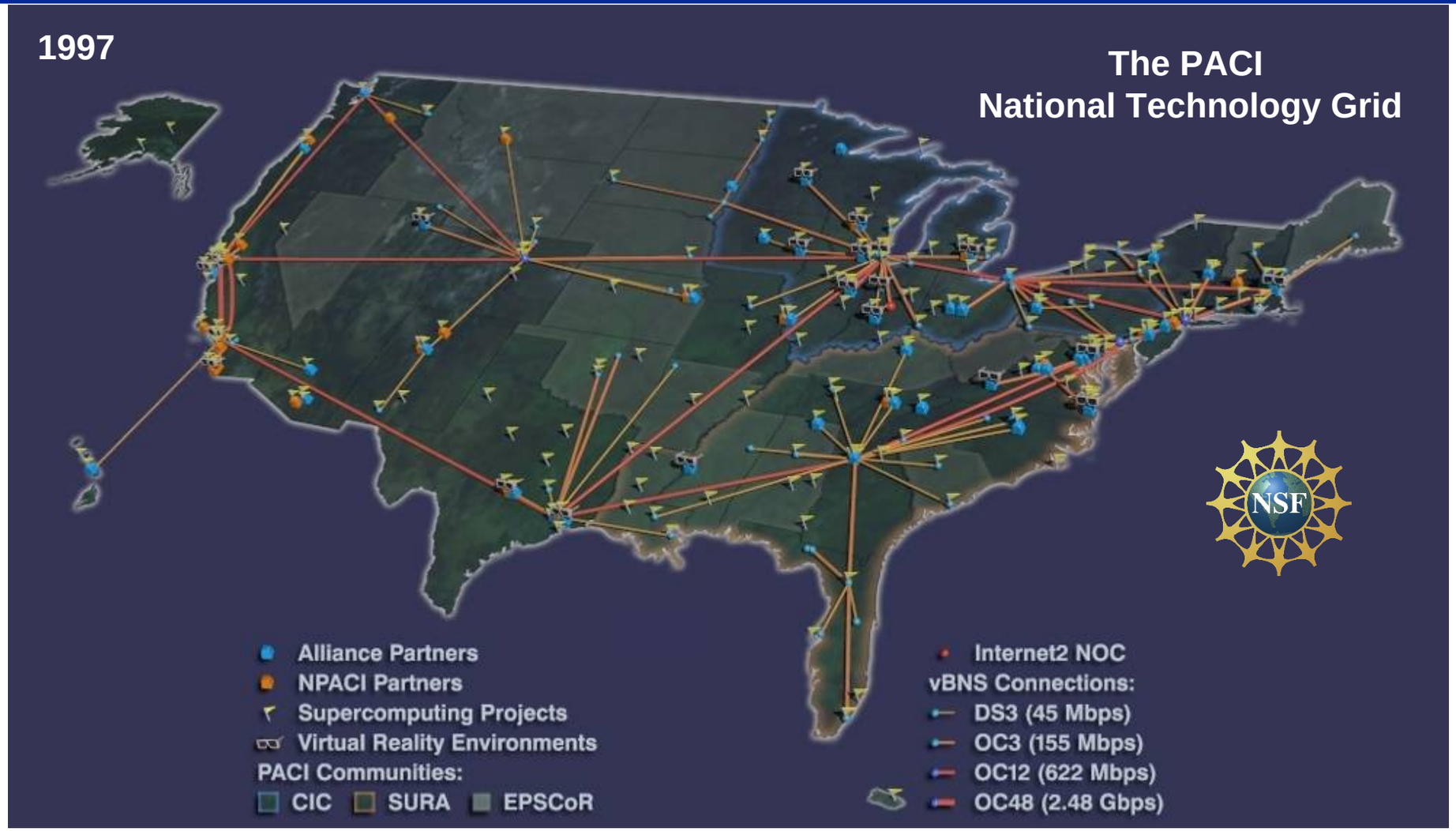
35(6): 44-52 (1992).
https://lsmarr.net/metacomputing/

NRP NATIONAL RESEARCH PLATFORM

1997: NSF's PACI Program was Built on the vBNS Network to Prototype America's 21st Century Information Infrastructure

vBNS led to

INTERNET2



Key Role
of Miron Livny
& Condor

NPACI

National Computational Science ALLIANCE

NRP NATIONAL RESEARCH PLATFORM

1999: Dave Bader Created the First Linux COTS PC Supercluster *Roadrunner* on the National Technology Grid, with the Support of NCSA and NSF

ALBUQUERQUE JOURNAL

YOUTH CRIME
Bill would fund state programs that crack down on juvenile offenders
D3

Metro & New Mexico

THURSDAY, APRIL 8, 1999
BUSINESS 4 SECTION
WEATHER 8 D
DEATHS 8

UNM To Crank Up $400,000 Supercomputer Today

Machine One of 100 Speediest in World

BY JOHN FLECK
Journal Staff Writer

A cluster of computers looking like a stack of oversized dorm room refrigerators is about to plug the University of New Mexico into an ambitious new supercomputer "mega-brain" spanning the nation.

Dignitaries from around the country will gather at UNM's Albuquerque High Performance Computing Center today to throw the switch on the "Roadrunner Supercluster."

Fresh out of the box, the Roadrunner is one of the 100 fastest computers in the world.

But its rocket-sled speed is just part of a more ambitious idea — an attempt to give scientists around the country access to a distributed supercomputing "brain" stretching from Boston to Maui, ready to solve their toughest problems.

The Roadrunner Supercluster will be part of what researchers are calling the "National Technology Grid," a collection of supercomputers across the country wired together to help handle scientists' growing demand for computer time.

Sitting at their own desks, the researchers could hook into the grid and begin solving large-scale problems without having to worry about where the computer doing the work is located, explained Larry Smarr, director of the National Computational Science Alliance, the federally funded group that paid for the new machine.

"The goal is to be completely seamless," the University of Illinois-based Smarr said in telephone interview this week.

UNM professor David Bader envisions a day when a researcher studying an ecosystem — Chesapeake Bay, say, or the Rio Grande Valley — can log in and rapidly spin through vast amounts of satellite data on climate and ground cover for their plot of land.

They wouldn't need to be located at a specialized supercomputer research center to do their work, explained Smarr — they could just plug into the grid and let the computers take care of the rest, much as you plug a toaster into the wall without having to worry about which power plant is generating the electricity.



JAIME DISPENZA/JOURNAL

BIG BRAIN: University of New Mexico professor David Bader tests the Roadrunner Supercluster. It will be dedicated today.

"That's the vision," said Bader, who has spent the last week testing the brand new Roadrunner machine in preparation for today's unveiling.

Smarr, one of the leaders of the nation's supercomputer research community, is scheduled to join Sen. Pete Domenici, R-N.M., and others to formally throw the switch turning on the computer at a ceremony this afternoon.

The $400,000 computer, built by Alta Technology Corp., bears the mark of a new breed of moderately priced machines that are making inroads in the high-performance scientific market.

Instead of using specialized high-performance computer chips made especially for supercomputing, it's built around 128 top-of-the-line Intel Pentiums, the same breed of computer chips used in desktop computers.

Programmers split up a large problem into smaller pieces, farming them out so that each of the 128 chips can work on a portion of the problem and then share results.

"You get to profit from the collapse of PC prices," explained Victor Yodaiken, a computer scientist at the New Mexico Institute of Mining and Technology who does research on a similar "cluster" machine at Tech.

Similar machines also have been built at Sandia and Los Alamos national labs for nuclear weapons and physics research, and the experience of the New Mexico research community is one reason UNM was chosen to host the new machine, Smarr said.

UNM is one of six "supernodes" on the National Technology Grid. In addition to the Roadrunner Cluster, the grid will be connected to UNM's Maui High Performance Computing Center, an Air Force-funded research complex in Hawaii run by the university for the military.

NCSA Director Larry Smarr (left),
UNM President William Gordon,
and U.S. Sen. Pete Domenici
Turn on the Roadrunner Supercomputer
in April 1999



**Roadrunner was the FIRST Linux COTS Supercomputer.
Today ALL Supercomputers and Commercial Clouds
Use This Architecture**

National Computational Science ALLIANCE

NRP NATIONAL RESEARCH PLATFORM

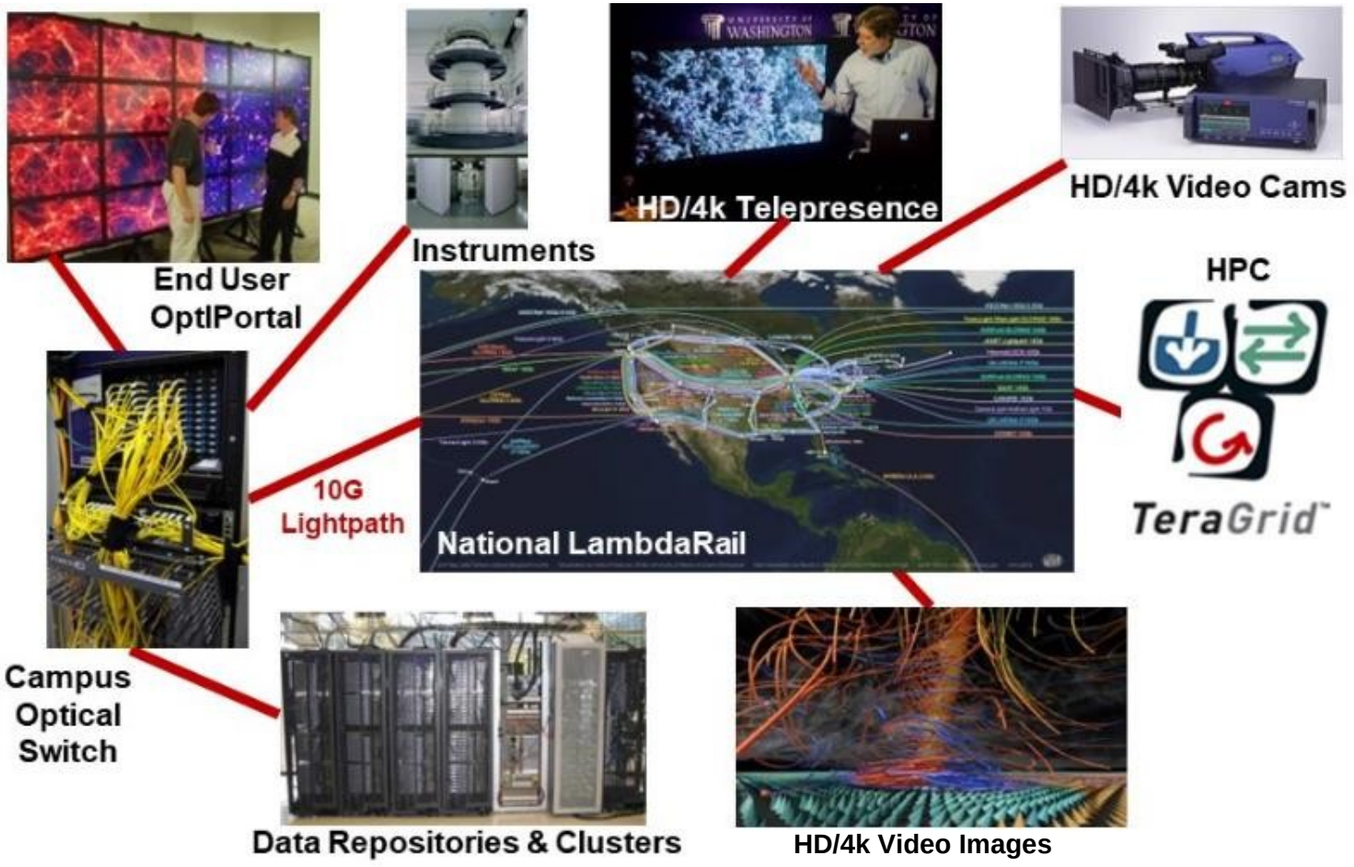
From the National Technology Grid
To the National Research Platform

> The goal is a powerful, persistent, distributed U.S. national resource as usable in the future as the NSF Supercomputer Center standalone machines were in the 1980s and 1990s.

From *I-WAY to the National Technology Grid*, CACM, 40, 51 (1997)
Rick Stevens, Paul Woodward, Tom DeFanti, and Charlie Catlett

NRP NATIONAL RESEARCH PLATFORM

2002-2009: The NSF-Funded OptIPuter Grant
Developed a Uniform Bandwidth Optical Fiber Connected Distributed System



OptIPuter
$13.5M
PI Smarr,
Co-PIs DeFanti, Papadopoulos, Ellisman, UCSD
Project Manager Maxine Brown, EVL
2002-2009

The OptIPuter
Exploits a New World
in Which
the Central Architectural Element
is Optical Networking,
Not Computers.

Demonstrating That
Wide-Area Bandwidth
Can Equal
Local Cluster Backplane Speeds

OptIPuter

it²

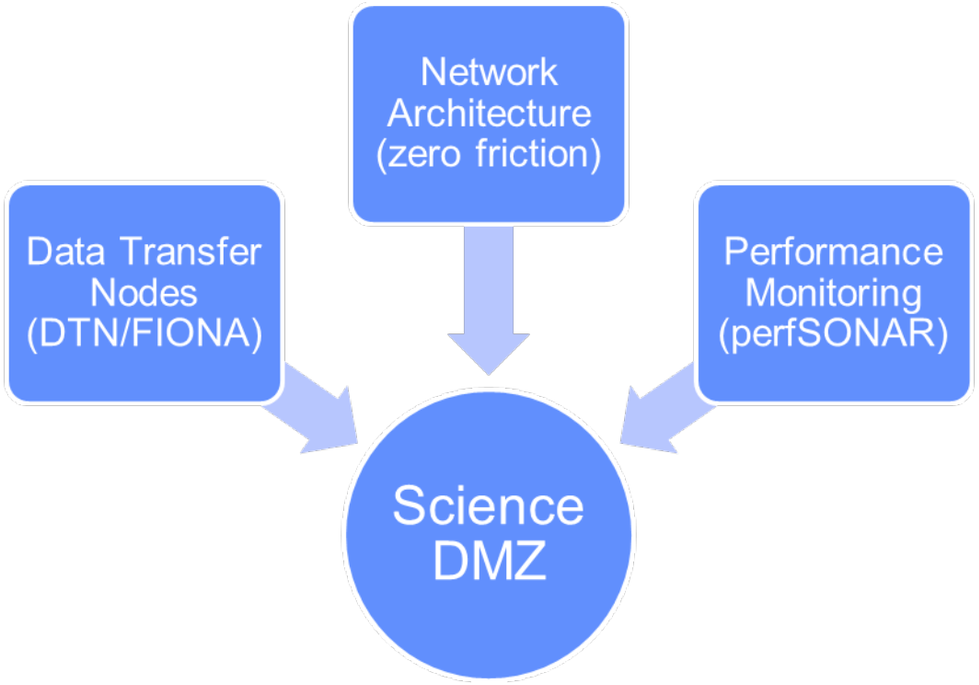
evl electronic visualization laboratory

SDSC

NRP NATIONAL RESEARCH PLATFORM

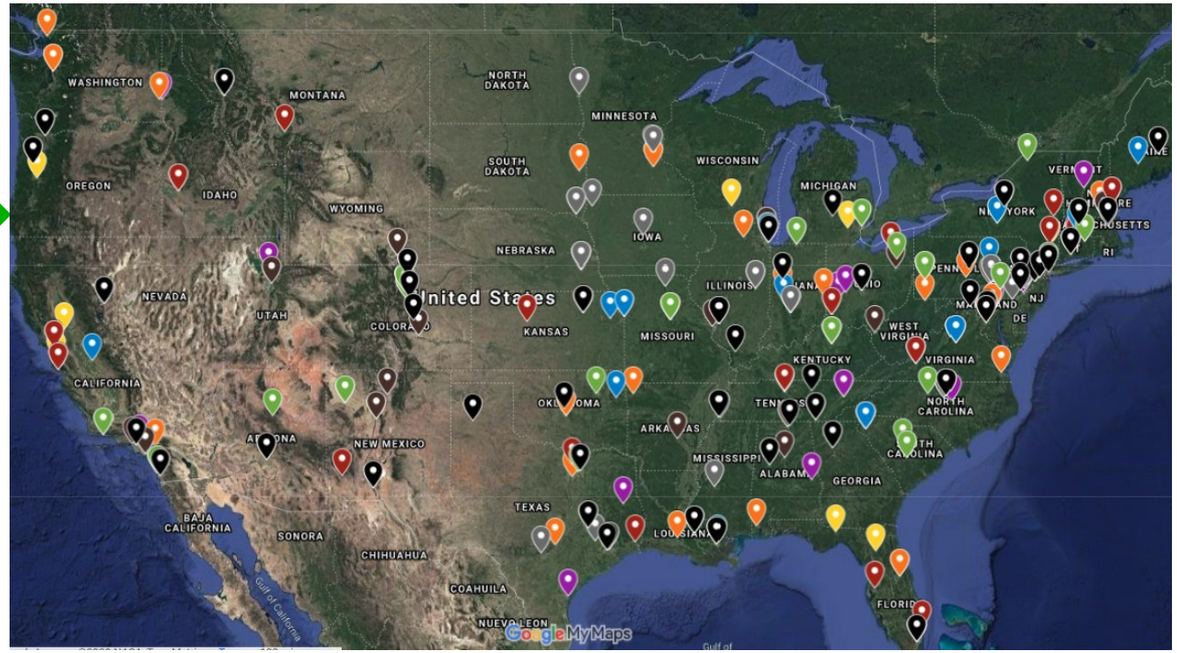
2012: NSF Adopted DOE ESnet's Science DMZ Model to Allow Campuses to Terminate SuperNetworks

ScienceDMZ Coined in 2010 by ESnet



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

NSF Campus Cyberinfrastructure Program Has Made Over 385 Awards Totaling Over $100M Since 2012



Source: Kevin Thompson, NSF

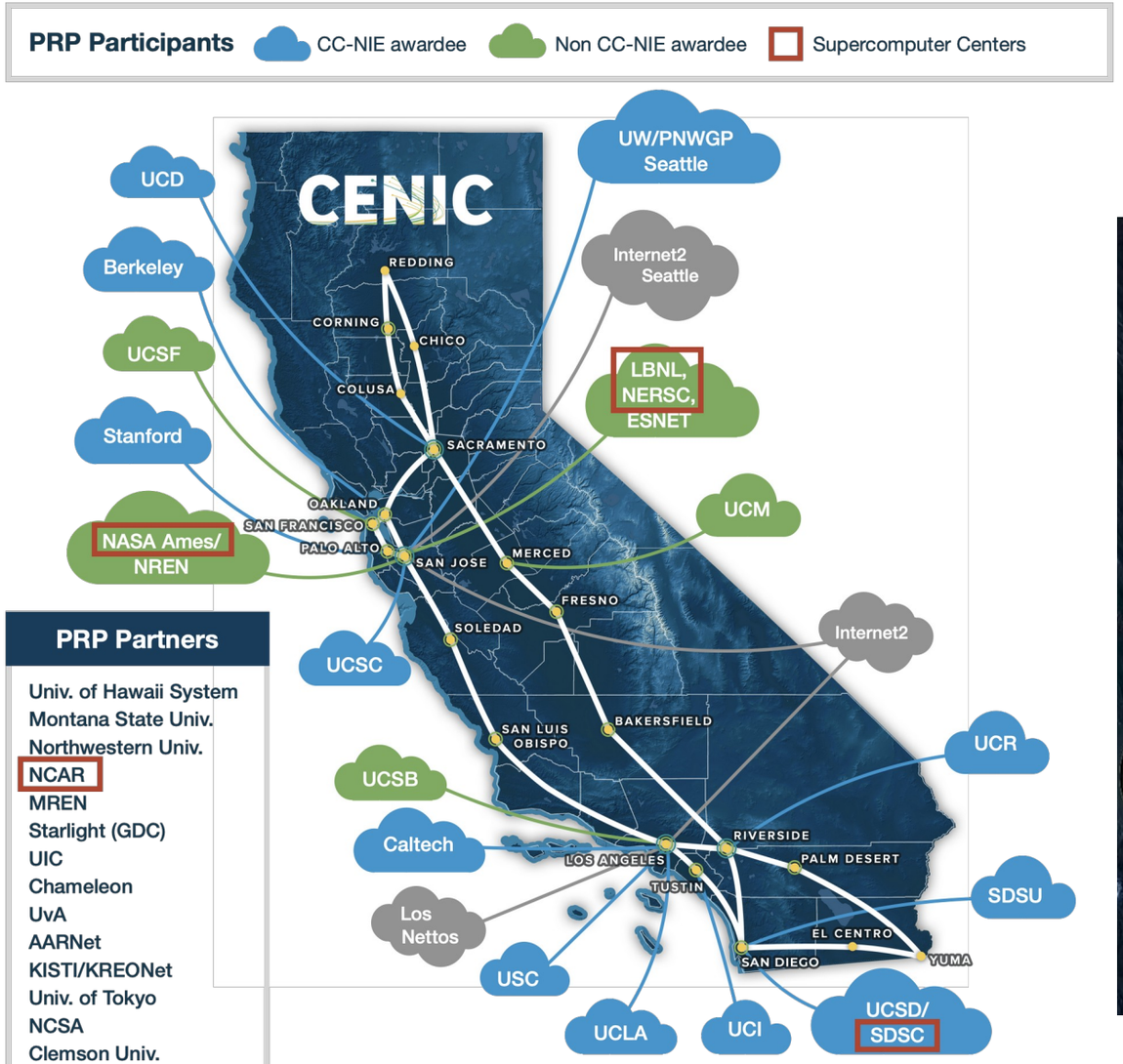
ESnet
FASTERDATA KNOWLEDGEBASE

Slide Adapted From Inder Monga, ESnet

NSF

NRP NATIONAL RESEARCH PLATFORM

2015 Vision: The Pacific Research Platform Will Build on CENIC to Connect Science DMZs Creating a Regional Community Cyberinfrastructure



Source: John Hess, & Hunter Hadaway, CENIC

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



NRP NATIONAL RESEARCH PLATFORM

Flash I/O Network Appliance (FIONA) Data Transfer Nodes (DTNs) Terminate Fiber Optic Networks on Campuses:

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



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



FIONAs Are Rack Mounted

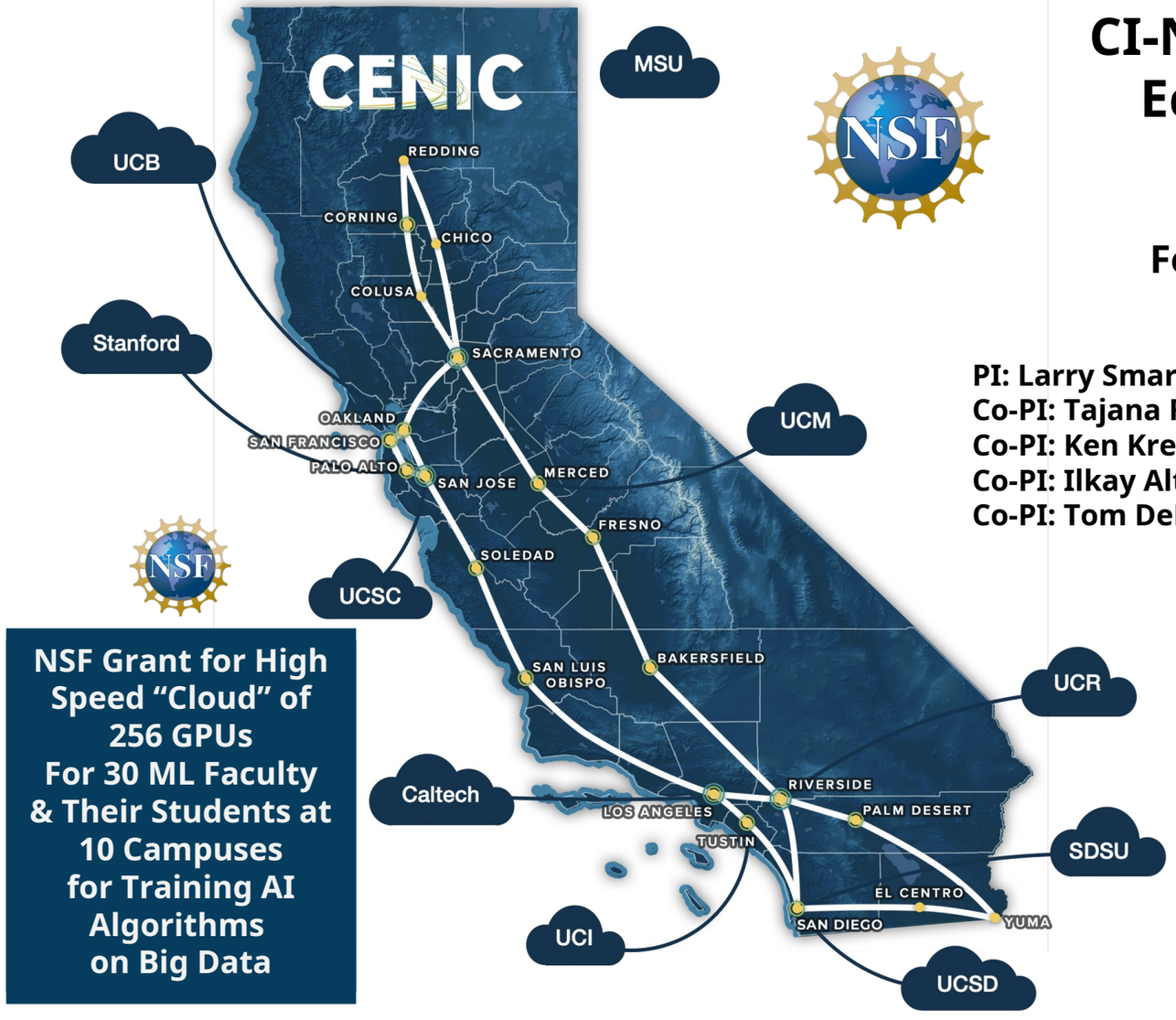
QUALCOMM INSTITUTE SDSC

NSF

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

NRP NATIONAL RESEARCH PLATFORM

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 21, 2020

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

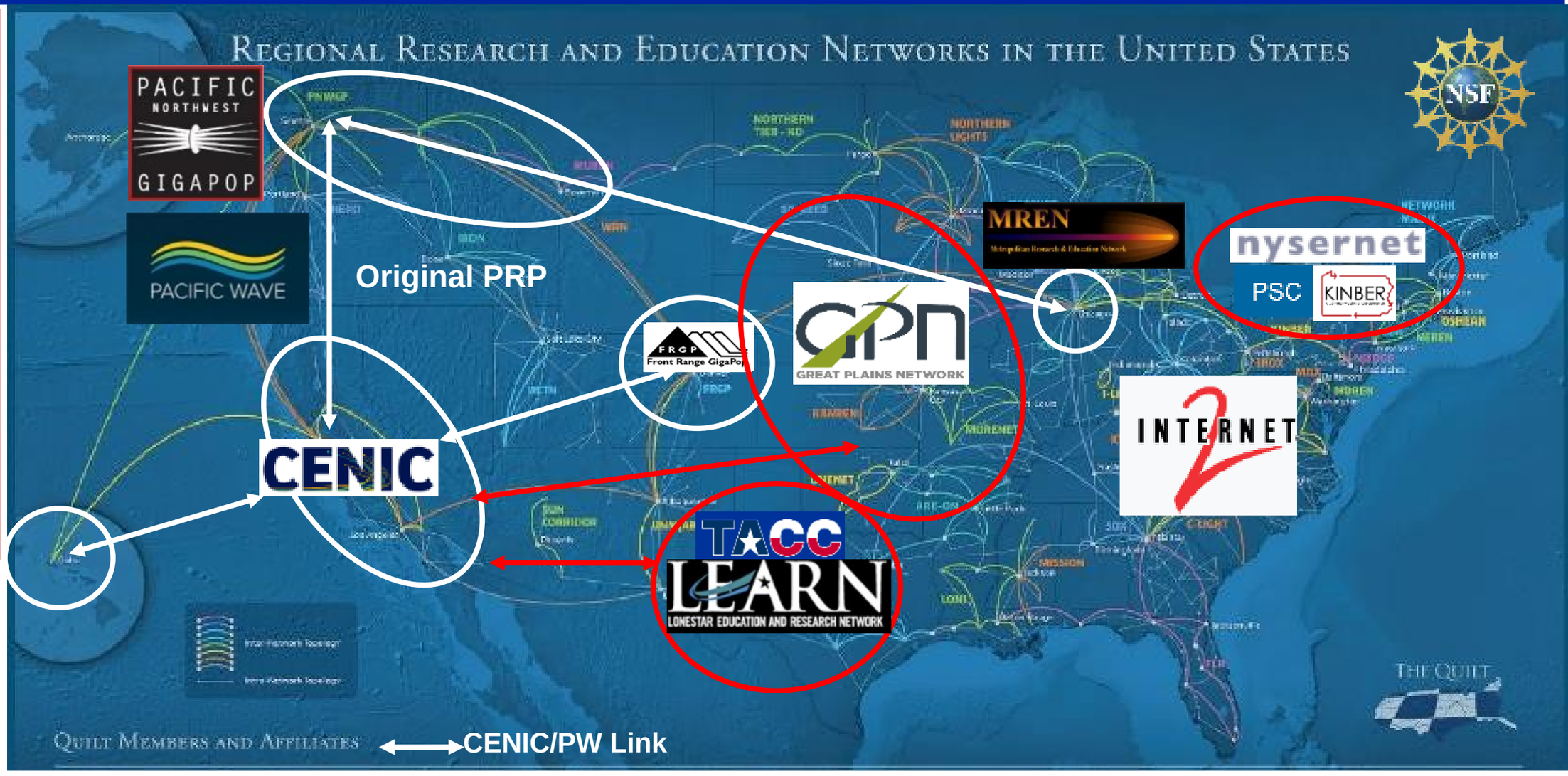
JACOBS SCHOOL OF ENGINEERING
SDSC UC San Diego QUALCOMM INSTITUTE

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

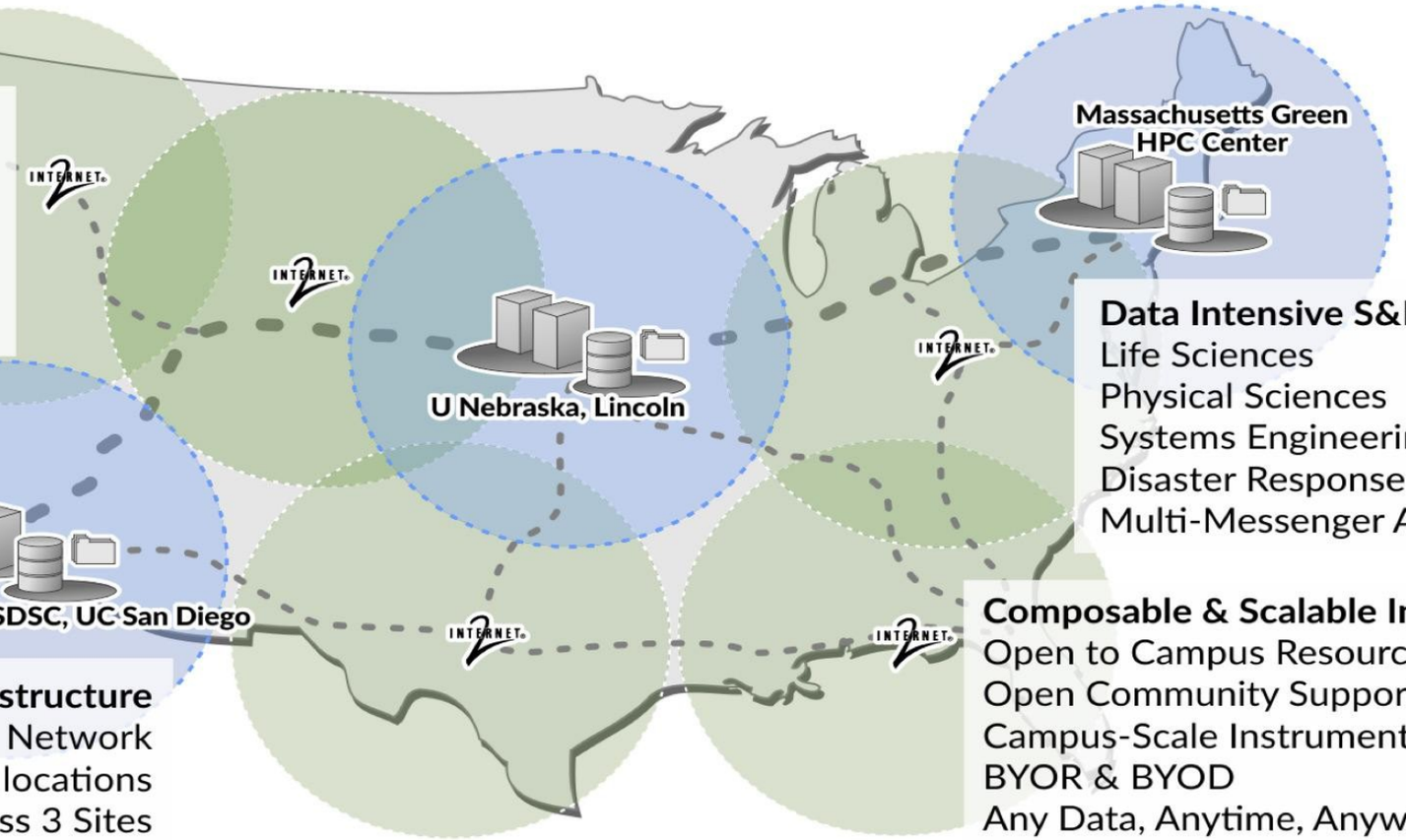


SDSC

NRP NATIONAL RESEARCH PLATFORM

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

SDSC



HPC/HTC Resource
32 ALVEO FPGAs
288 NVIDIA FP32 GPUs
48 NVIDIA FP64 GPUs
Tbps WAN IO Capabilities
Configurable Low Latency HPC Fabric

Data Intensive S&E
Life Sciences
Physical Sciences
Systems Engineering
Disaster Response
Multi-Messenger Astrophysics

Distributed Data Infrastructure
National Scale Content Delivery Network
50TB 100Gbps NVMe Caches in 8 locations
4.5PB Distributed Data Origin across 3 Sites

Composable & Scalable Innovation
Open to Campus Resource Integration
Open Community Support Model
Campus-Scale Instrument integration
BYOR & BYOD
Any Data, Anytime, Anywhere

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)

NRP NATIONAL RESEARCH PLATFORM

2023 - The National Research Platform Emerges As a Unification of 26 Years of NSF Cyberinfrastructure Grants

NRP NATIONAL RESEARCH PLATFORM

NEWS GRANTS TECHNOLOGY COMMUNITY JOIN / CONTACT



Professor Frank Würthwein

- **Director** - SDSC
- **Distributed High-Throughput Computing Lead** - SDSC

Designed for Growth and Inclusion

The National Research Platform (NRP) is a partnership of more than 50 institutions, led by researchers and cyberinfrastructure professionals at UC San Diego, supported in part by awards from the National Science Foundation.

https://

Nautilus is NRP's Multi-Institution Hypercluster Which Creates a Community Owned and Operated "AI Resource"

September 22, 2024

~200 FIONAs on 27 Partner Campuses Networked Together at 10-100Gbps

Installed GPUs

1363

Installed CPU Cores

25108

Rotating Storage

7000 TB

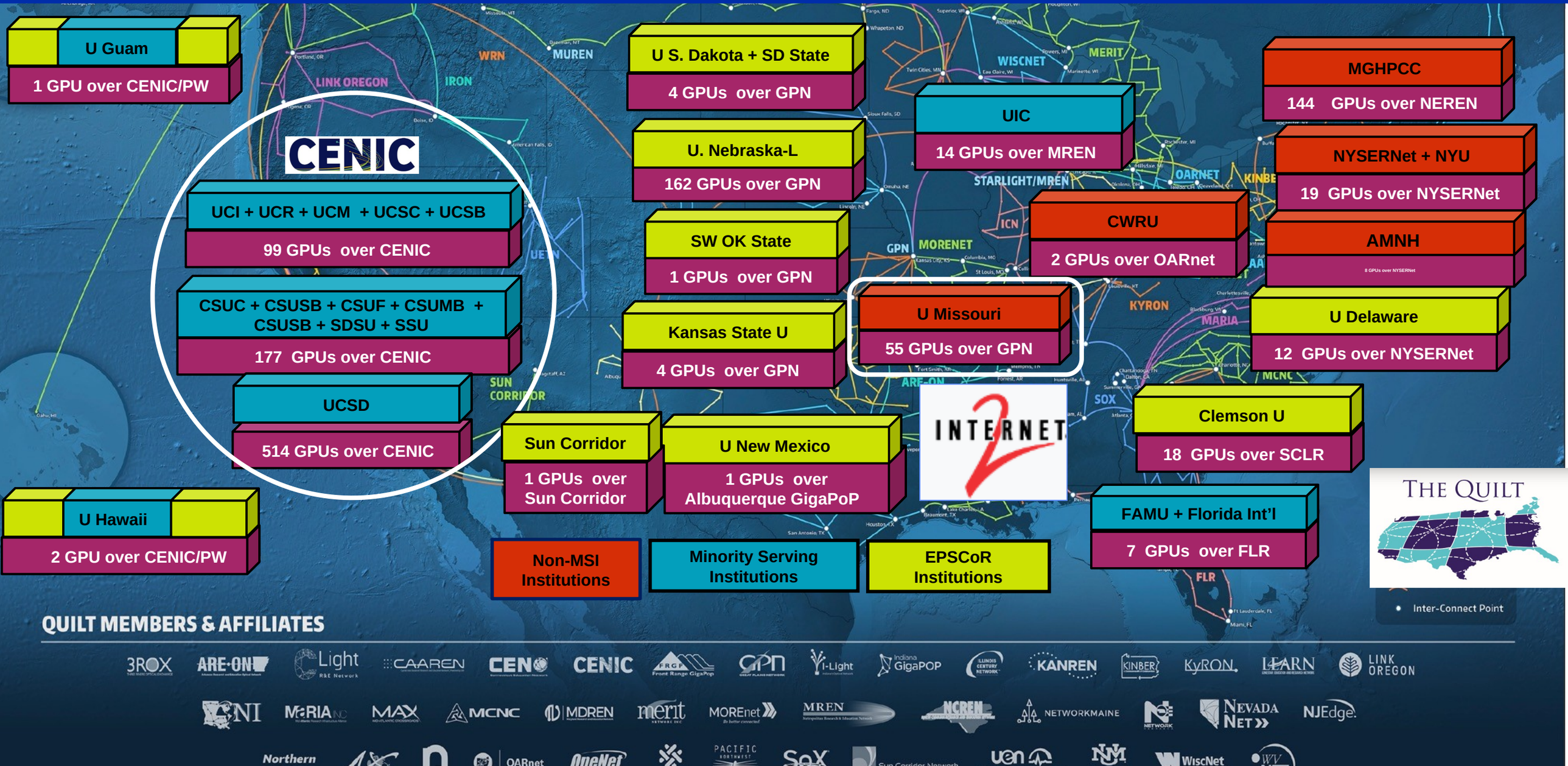
CENIC

THE QUILT

INTERNET2®

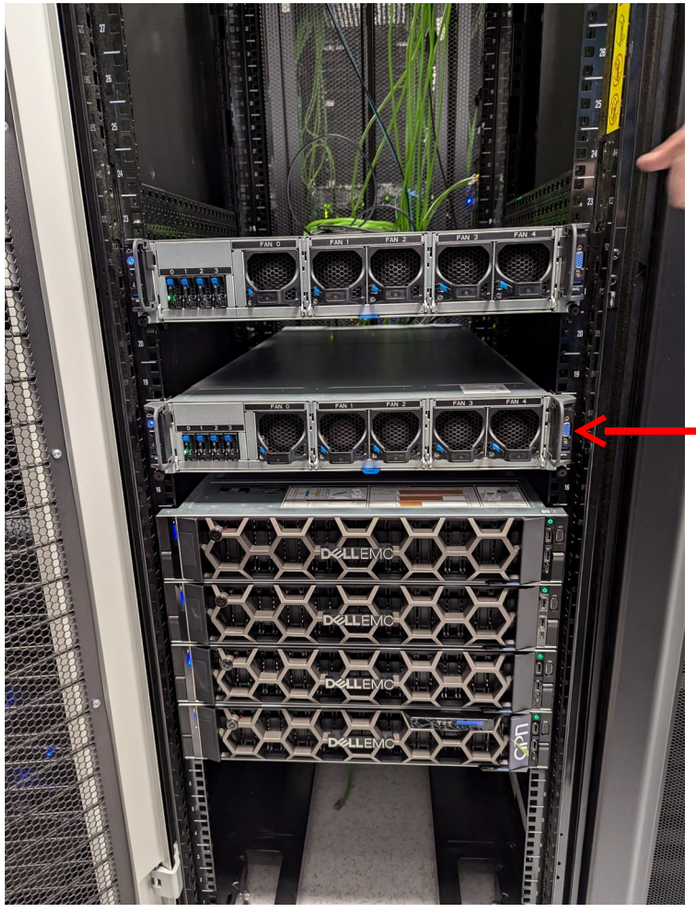
NRP NATIONAL RESEARCH PLATFORM

The NRP's Nautilus Hypercluster is Connected by Regional Optical Networks & Internet2



Alex Hurt With the MU Nautilus FIONAs





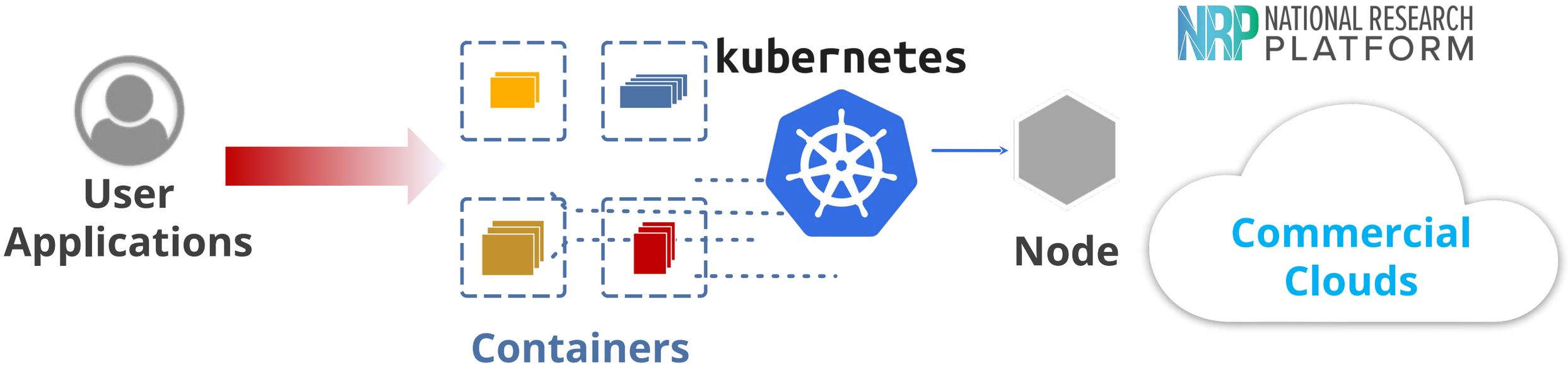
NSF GP-ENGINE Grant (Alex Hurt PI) Funded
1st NRP NVIDIA GH200 Grace Hopper AI Superchip

NRP NATIONAL RESEARCH PLATFORM

Nautilus Users Can Execute Their Containerized Applications in the NRP or in Commercial Clouds

Nautilus Containerized Applications Are "Cloud Ready"

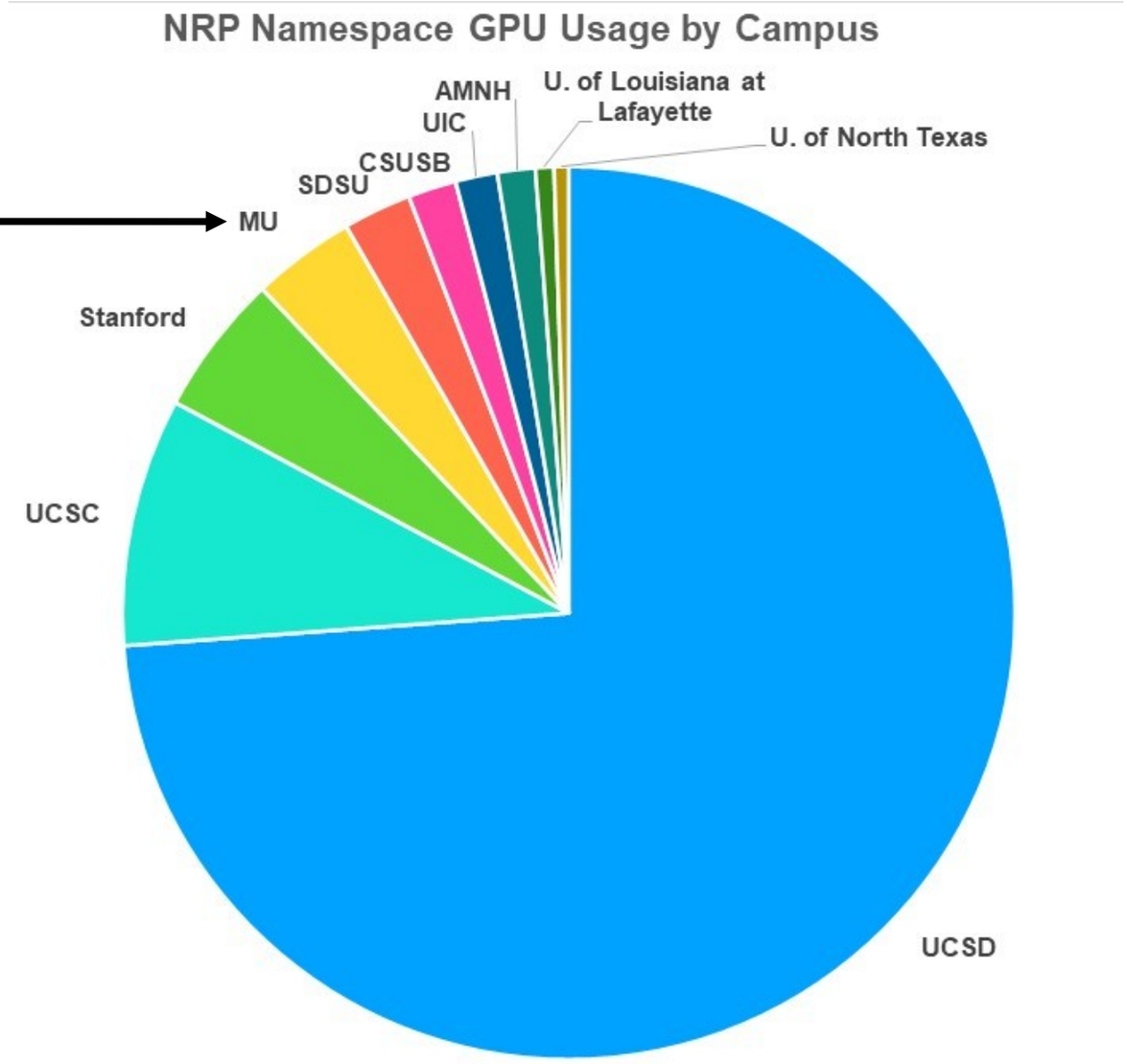
Nautilus



NRP NATIONAL RESEARCH PLATFORM

MU Had the Largest Campus Usage of NRP GPUs East of California in the Last 9 Months!

MU Used Over 100,000 GPU-Hours With Over 50 Namespaces



Only UCSD, UCSC, and Stanford Used More GPU-hours Than MU

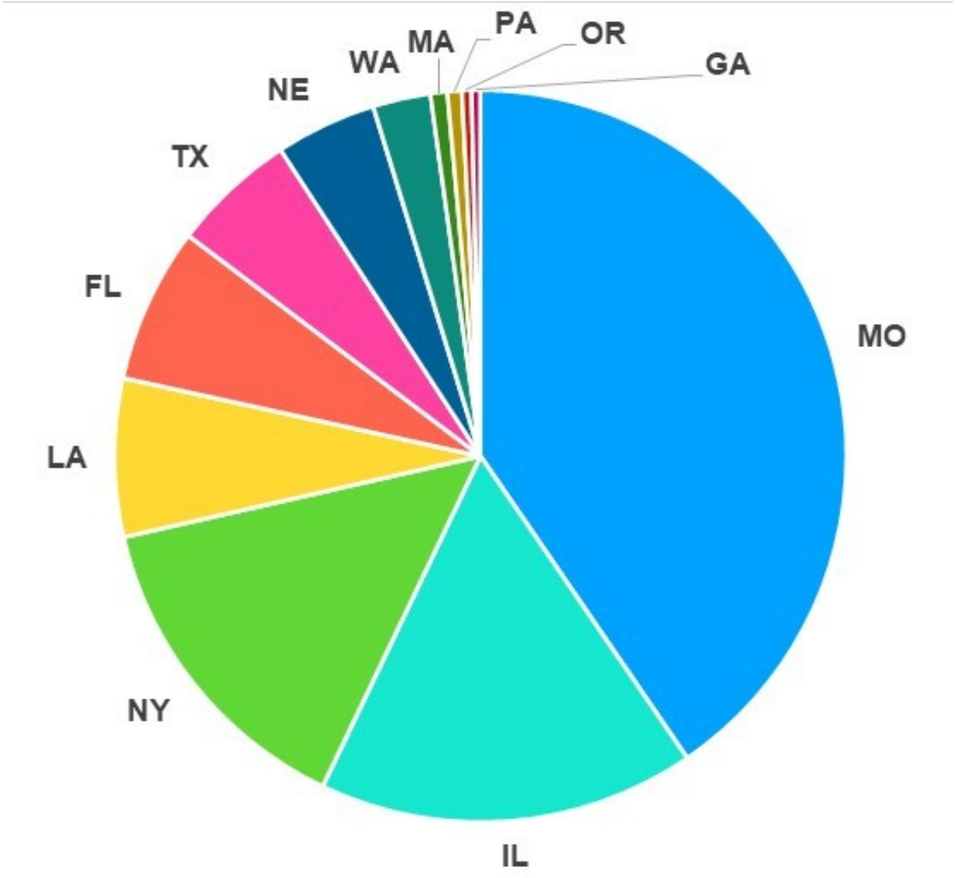
Last 9 Months: 12/1/2023 to 8/31/2024

Top 10 Campus GPU Usage

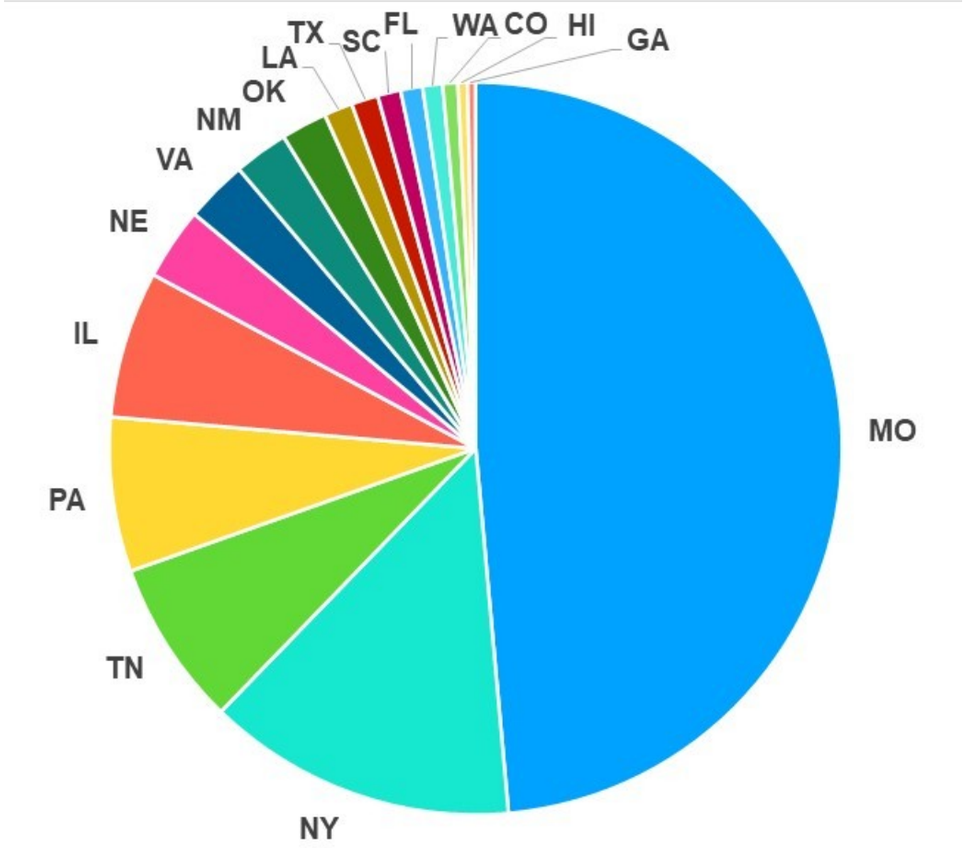
NRP NATIONAL RESEARCH PLATFORM

Missouri Dominates Non-California Nautilus Campus Namespace Usage by State: “Big MO!”

Grant Scott & Alex Hurt, UMC
Helped Organize the UMC NRP Usage



284,684 GPU-Hours

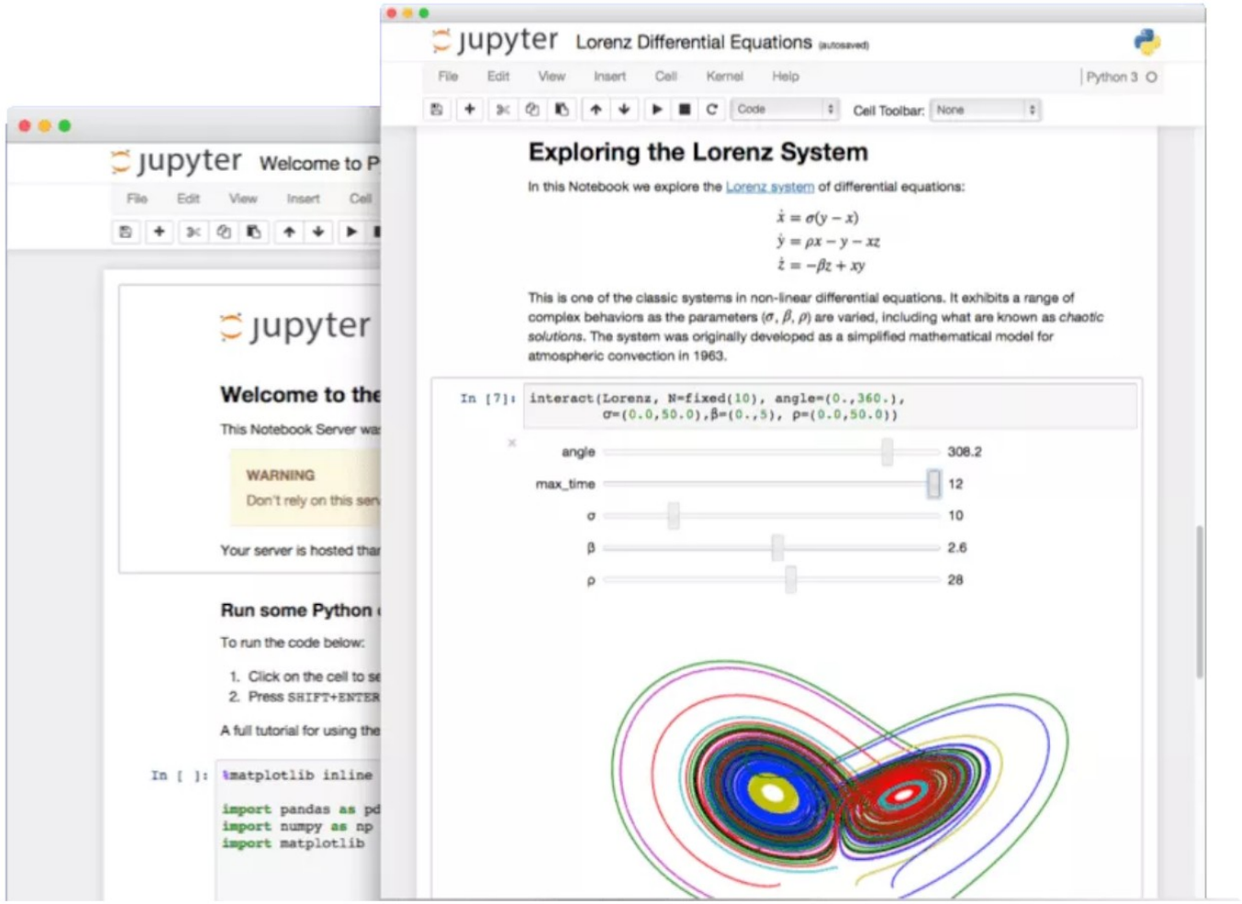


4,945,177 CPU Core-Hours

Last 9 Months: 12/1/2023 to 8/31/2024

NRP NATIONAL RESEARCH PLATFORM

jupyter Notebooks
Have Become The Popular Method of Sharing Computational Documents



Jupyter Notebook: The Classic Notebook Interface

The Jupyter Notebook is the original web application for creating and sharing computational documents. It offers a simple, streamlined, document-centric experience.

Try it in your browser | Install the Notebook

NRP NATIONAL RESEARCH PLATFORM

jupyterhub is the
Multi-User Version of the jupyter Notebook



A multi-user version of the notebook designed for companies, classrooms and research labs

NRP NATIONAL RESEARCH PLATFORM

Grant Scott and I
Gave Joint Talk to the MOREnet 2022 Conference

“Utilizing Nautilus and the National Research Platform for Big Data Research and Teaching”

Panel Presentation

Larry Smarr and Grant Scott

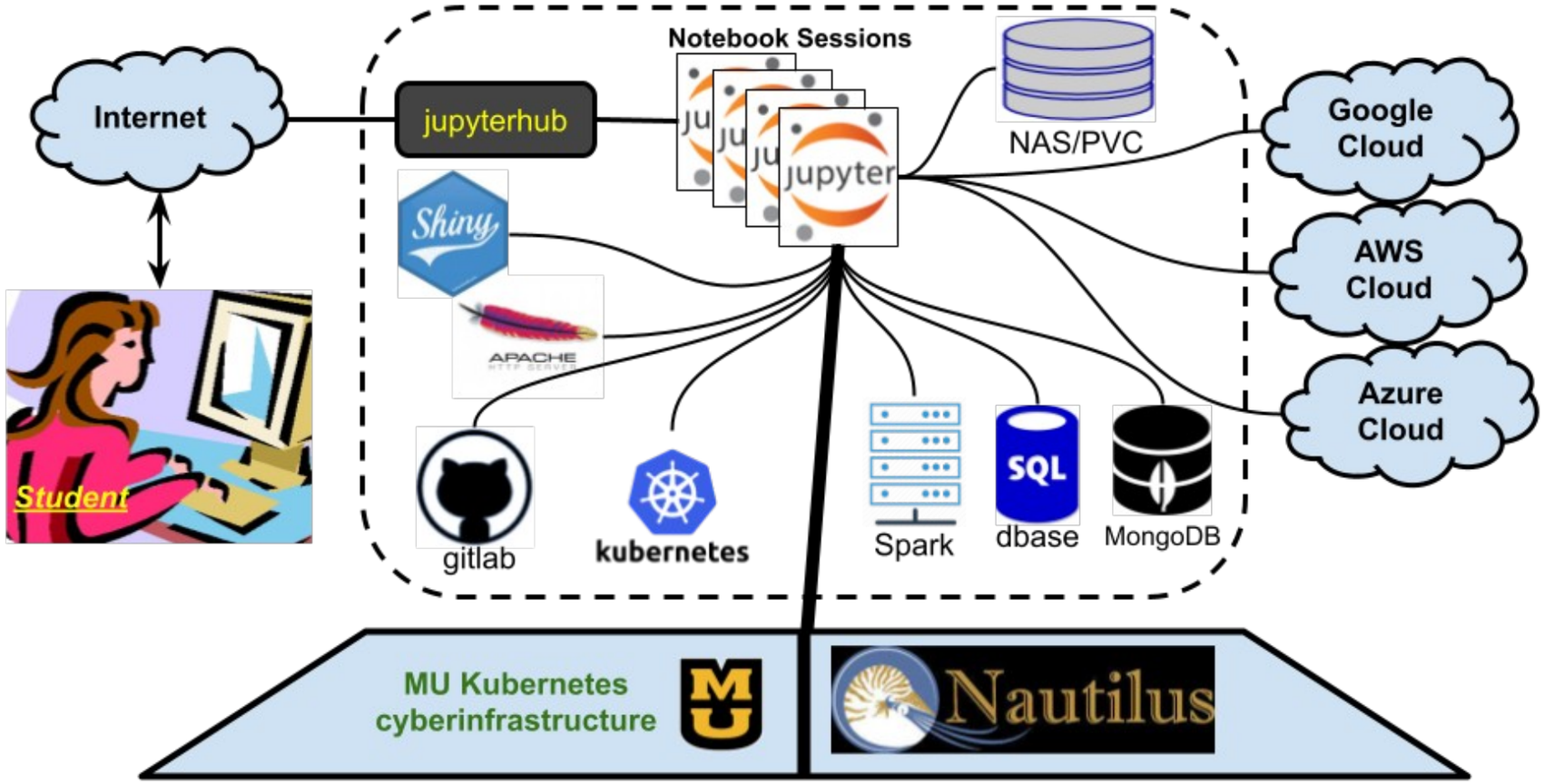
MOREnet 2022 Annual Conference

October 19, 2022

https://lsmarr.net/utilizing-nautilus-and-the-national-research-platform-for-big-data-research-and-teaching/

NRP NATIONAL RESEARCH PLATFORM

University of Missouri is Leveraging Jupyter Hubs on Nautilus for STEM & AI/ML Experiential Learning

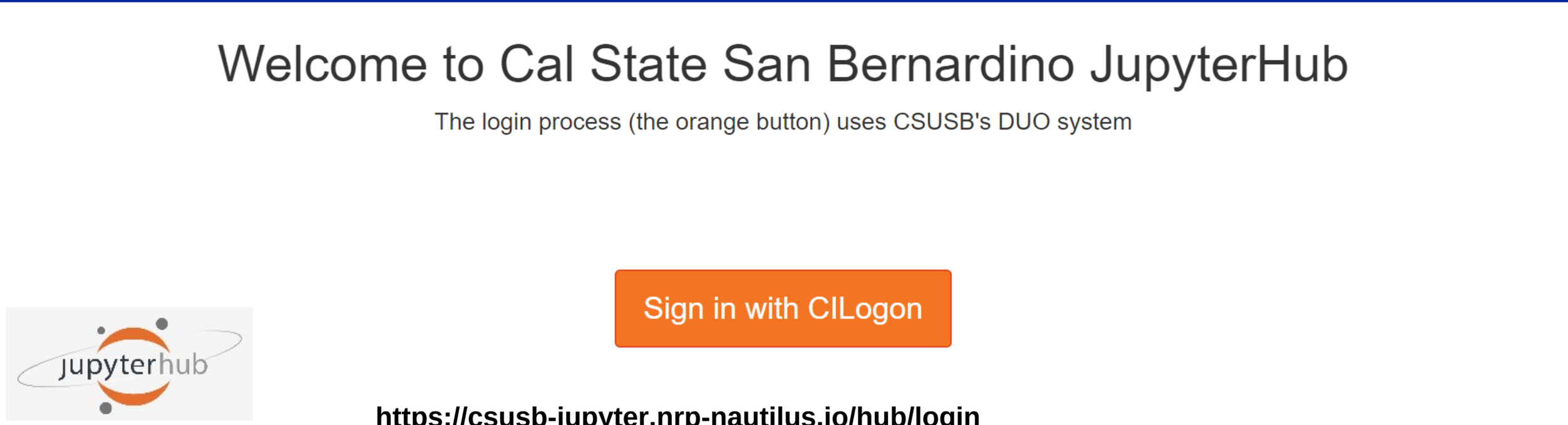


Has 16 GPUs On-Site Now -
Adding 44 with a New CC* Award

Source: Grant Scott, MU

University of Missouri

A Key Reason CSUSB Has The Largest CSU Nautilus Usage: They Installed and Publicized the JupyterHub “Easy Button”



https://csusb-jupyter.nrp-nautilus.io/hub/login

Slide Adapted from Prof. Youngsu Kim

Over 450 Total Users
Tripled in the Last Six Months!

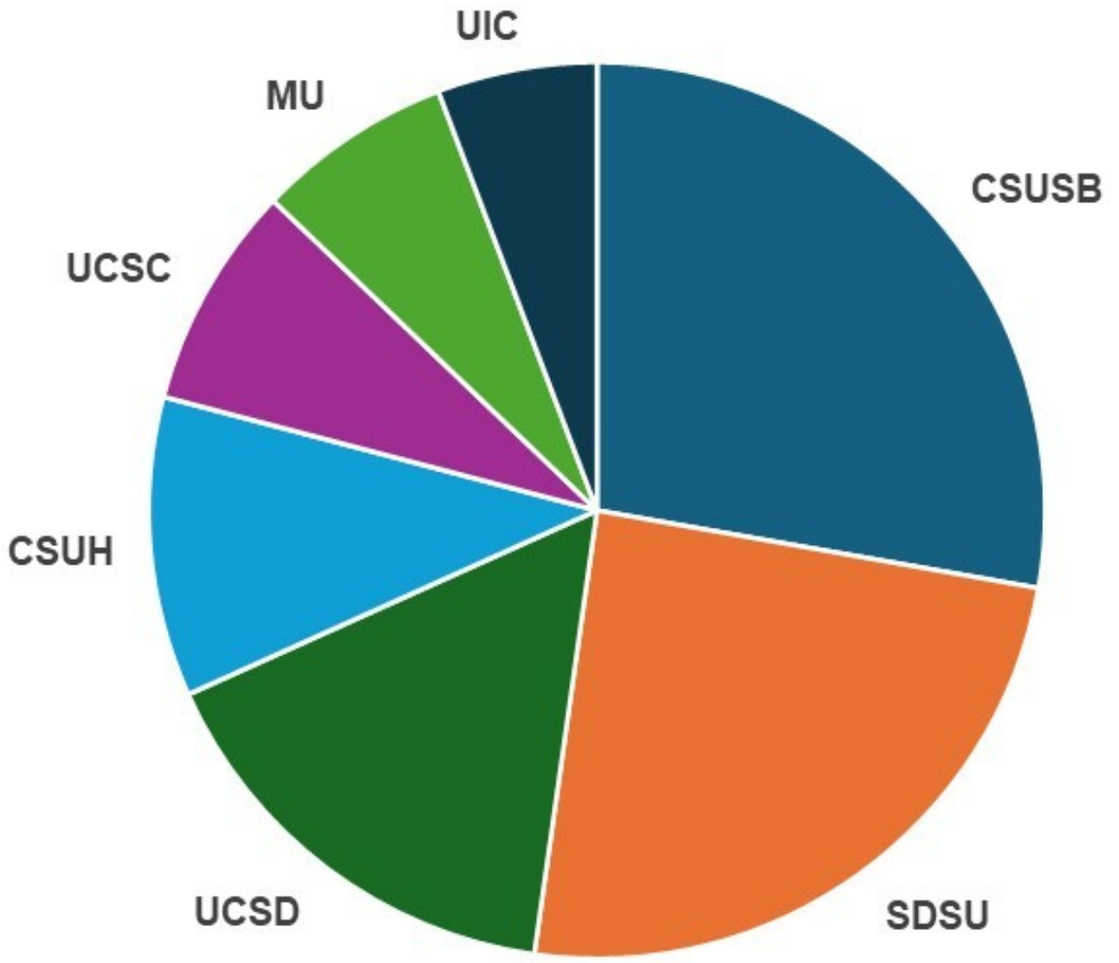
Department of Mathematics
Faculty & Staff

CSUSB | WE DEFINE THE *Future*

NRP NATIONAL RESEARCH PLATFORM

MU Also Had the Largest Campus Usage of NRP jupyterhub GPUs East of California in the Last 9 Months!

Last 9 Months: 12/1/2023 to 8/31/2024



107 NRP Namespaces Had Active Embedded jupyterhubs in the Last 9 Months

Over 100,000 GPU-Hours

NRP NATIONAL RESEARCH PLATFORM

NRP's Nautilus Cyberinfrastructure Supports a Wide Array of AI/ML Algorithms

Nautilus was Designed to Support Research in 6 Broadly Defined Families of Information Extraction and Pattern Recognition Algorithms that are Commonly Used in AI/ML Research:

1) **Deep Neural Network (DNN) and Recurrent Neural Network (RNN) Algorithms Including Layered Networks:**
   - **Convolutional layers (CNNs),**
   - **Generative adversarial networks (GANs), &**
   - **Transformer Neural Networks (e.g., LLMs)**
2) **Reinforcement Learning (RL) and Inverse-RL Algorithms & Related Markov Decision Process (MDP) Algorithms**
3) **Variational Autoencoder (VAE) and Markov Chain Monte Carlo (MCMC) Stochastic Sampling**
4) **Support Vector Machine (SVM) Algorithms and Various Ensemble ML Algorithms**
5) **Sparse Signal Processing (SSP) Algorithms, Including Sparse Bayesian Learning (SBL)**
6) **Latent Variable (LVA) Algorithms for Source Separation**

Source: CHASE-CI Proposal

NRP NATIONAL RESEARCH PLATFORM

Today's Over 1000 Nautilus Namespaces Have Utilized Many of These Algorithms

The Great Majority of Nautilus AI/ML Namespaces are Using Some Form of NNs or RL

- **For NNs PyTorch, TensorFlow, and Keras are the Preferred (in that order) Open-Source Deep Learning (DL) Frameworks Used on Nautilus.**
- **Our AI/ML Researchers Use Different Subtypes of DNNs, Including:**
  - **Deep Belief Networks (DBN),**
  - **Quantum NNs (QNN),**
  - **Graph NNs (GNNs) and**
  - **Long Short-Term Memory (LSTM) RNNs-Specifically Designed to Handle Sequential Data, such as Time Series, Speech, and Text**
- **Nautilus Namespaces Use RL and Inverse-RL Algorithms in Many Areas of Dynamic Decision-Making, Robotics, and Human/Robotic Transfer Learning**

Nautilus Namespaces with Descriptions:
https://portal.nrp-nautilus.io/namespaces-g

NRP NATIONAL RESEARCH PLATFORM

NRP's Largest GPU-Consuming AI/ML Researchers Point to the Rapid Growth of Transformer NNs

Transformer NNs Have Become the Default Architecture for Applications Involving Images, Sound, or Text

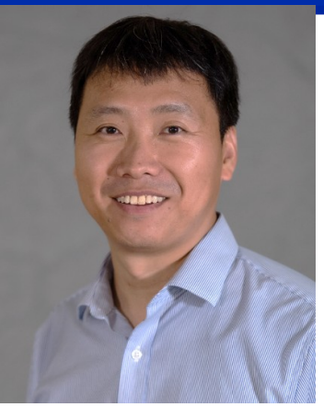
- **A Growing Number of NRP Namespaces are Using Transformer-Based Large Language Models (LLMs), Such as GPT, LLaMa, and BERT in Natural Language Processing (NLP), or Vision Language Models, Such as CLIP and ViT, for Image Understanding Research**

- **Also Popular are Generative models, Such as GANs and Diffusion Models, Which are Prevalent in Data Synthesis, Such as For Text to Image Generation, Like Stable Diffusion**

- **Finally, We See Many Namespaces Working in Fields Such as Learning for Dynamics and Control (L4DC), Computer Vision (CV), and Trustworthy ML**

NRP NATIONAL RESEARCH PLATFORM

Example: CSU Northridge Prof. Bingbing Li
Machine Learning Research Projects Utilizing Nautilus



CSUN Prof. Bingbing Li

CSUN | ARCS Autonomy Research Center for STEAHM

- **Energy Disaggregation for Manufacturing Plant**
  - Faculty Lead: Dr. Bingbing Li (CSUN), Dr. Richard Donovan (UCI)
  - **DNNs: Long Short-Term Memory (LSTM) RNN, PyTorch**
- **Graph Representation Learning for Material Prediction and Recommendation in CAD Automation**
  - Faculty Lead: Dr. Bingbing Li (CSUN)
  - Collaborator: Dr. Daniele Grandi @ Autodesk and Dr. Thomas Lu @JPL
  - **DNNs: UV-Net Graph Neural Networks (GNN), PyTorch**
- **Knowledge Graph Construction Through the Potential of Large Language Models within Manufacturing**
  - Faculty Lead: Dr. Bingbing Li (CSUN)
  - Collaborator: Dr. Jerry Fuh & Senthil Kumar @ National University of Singapore
  - **DNNs: LLMs (ChatGPT & LLaMa), Tensorflow**
- **Multi-Domain AI for Future Manufacturing**
  - Faculty Lead: Dr. Bingbing Li (CSUN)
  - Collaborator: Dr. Edward Chow & Dr. Thomas Lu @JPL
  - **DNNs: LLMs (ChatGPT & LLaMa), Tensorflow**
- **Medical Image Restoration through Optical & CT Scanning**
  - Faculty Lead: Dr. Xiyi Hang & Dr. Bingbing Li (CSUN)
  - Collaborator: Dr. Ye Pu and Prof. Demetri Psaltis @ Swiss Federal Institute of Technology Lausanne
  - **DNNs: Large-Kernel CNN, Tensorflow**

TensorFlow

PyTorch

LLM LARGE LANGUAGE MODEL

ChatGPT

Meta

Source: Bingbing Li

NRP NATIONAL RESEARCH PLATFORM

See Alex Hurt's Talk
For MU Examples of NN Applications Using Nautilus

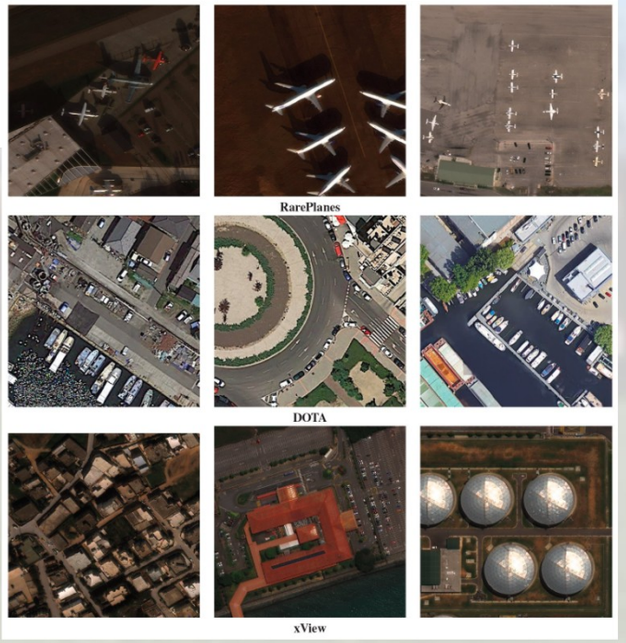
University of Missouri

Nautilus Application #1: Transformers

Evaluation and Analysis of Deep Transformers and Convolutional Neural Networks on Modern Remote Sensing Datasets

J. Alex Hurt, *Member, IEEE*, Trevor M. Bajkowski, *Student Member, IEEE*, Grant J. Scott, *Senior Member, IEEE* and Curt H. Davis, *Fellow, IEEE*

- Preparing for submission to *IEEE Journal on Selected Topics in Applied Remote Sensing*



NRP NATIONAL RESEARCH PLATFORM

The Congress Has Asked NSF to Develop a National AI Research Resource

A widely accessible AI research cyberinfrastructure **that brings together computational resources, data, testbeds, algorithms, software, services, networks, and expertise**, as described in this report **would help to democratize** the AI R&D landscape in the United States for the benefit of all.



In order to achieve its vision and goals, the Task Force estimates **the budget for the NAIRR as $2.6 billion** over an initial six-year period.

NRP NATIONAL RESEARCH PLATFORM

NSF Proposal Submitted Last Week

National Science Foundation
Directorate for Computer & Information Science & Engineering (CISE)
Division of Computer and Networked Systems



CIRC: Grand: The National Research Platform's Community Infrastructure for Research in Computer and Information Science and Engineering (NRP-CIRC)

For the Period 04/01/2025 - 03/31/2030

$5,000,000

Submitted September 12, 2024

NRP NATIONAL RESEARCH PLATFORM

NRP-CIRC is a Collaboration Between UC San Diego and MU

PI: Tom DeFanti, Recall Research Scientist at Calit2/QI, University of California, San Diego, Distinguished Professor Emeritus, University of Illinois at Chicago
Co-PI: Tajana Rosing, Professor of Computer Science and Engineering
Co-PI: Larry Smarr, Recall Distinguished Professor Emeritus of Computer Science and Engineering, University of California, San Diego
Co-PI: Frank Würthwein, Professor of Physics, SDSC Director, University of California, San Diego
Co-PI: Rose Yu, Assistant Professor of Computer Science and Engineering, University of California, San Diego
Sr. Personnel: Mary Thomas, Computational/Data Science Research Specialist, SDSC, University of California, San Diego

Subcontract PI: J. Alex Hurt, Research Assistant Professor of Electrical Engineering and Computer Science, University of Missouri
Subcontract PI: Edward L. Jones, Professor & Interim Chair of Computer & Information Sciences, Florida A&M University
Subcontract PI: Derek Weitzel, Holland Computing Center, University of Nebraska-Lincoln

NRP NATIONAL RESEARCH PLATFORM

NRP-CIRC
Overview

The National Research Platform (NRP) supports
the data-intensive AI/ML compute needs
of >3,000 researchers on over 500 projects from 129 US campuses,
including a growing number of MSIs (35 now),
as well as users at 34 campuses in 20 of the 28 EPSCoR states and territories,
among them all, 24 campuses that are neither R1 nor R2.

NRP-CIRC's personnel effort will go to community outreach activities
specifically designed to enlarge Nautilus' user base,
especially at under-resourced institutions,
by providing expert AI/ML/Data Science user training support.

U. of Missouri's J. Alex Hurt is the NRP-CIRC's Community Outreach Director.

NRP NATIONAL RESEARCH PLATFORM