National Research Platform (NRP)

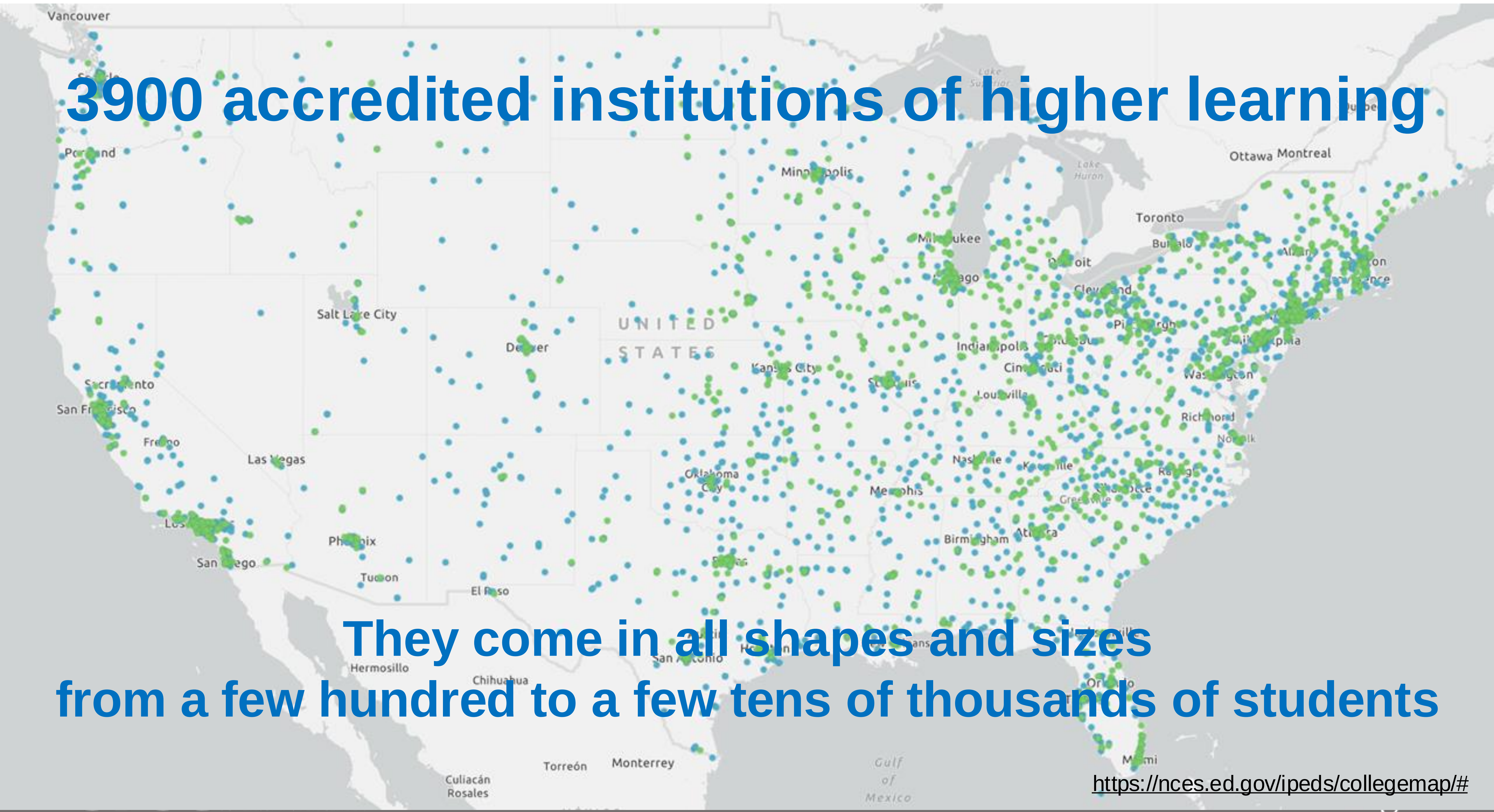
Integration and Use of PNRP NSF Cat II System

Osaka, September 16, 2024
5th GLOBAL RESEARCH PLATFORM WORKSHOP
Co-Located with 20th IEEE International Conference on eScience 2024

Mahidhar Tatineni
San Diego Supercomputer Center

Vision of NRP

3900 accredited institutions of higher learning



They come in all shapes and sizes
from a few hundred to a few tens of thousands of students

https://nces.ed.gov/ipeds/collegemap/#

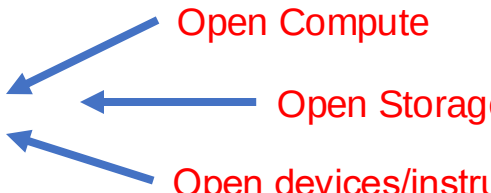
Motivation for our Vision

- **Connect every community college, every minority serving institution, and every college and university, including all urban, rural, and tribal institutions** to a world-class and secure R&E infrastructure, with particular attention to institutions that have been chronically underserved;
- **Engage and empower every student and researcher** everywhere with the opportunity to join collaborative environments of the future, because we cannot know where the next Edison, Carver, Curie, McClintock, Einstein, or Katherine Johnson will come from; and

https://mindsweneed.org

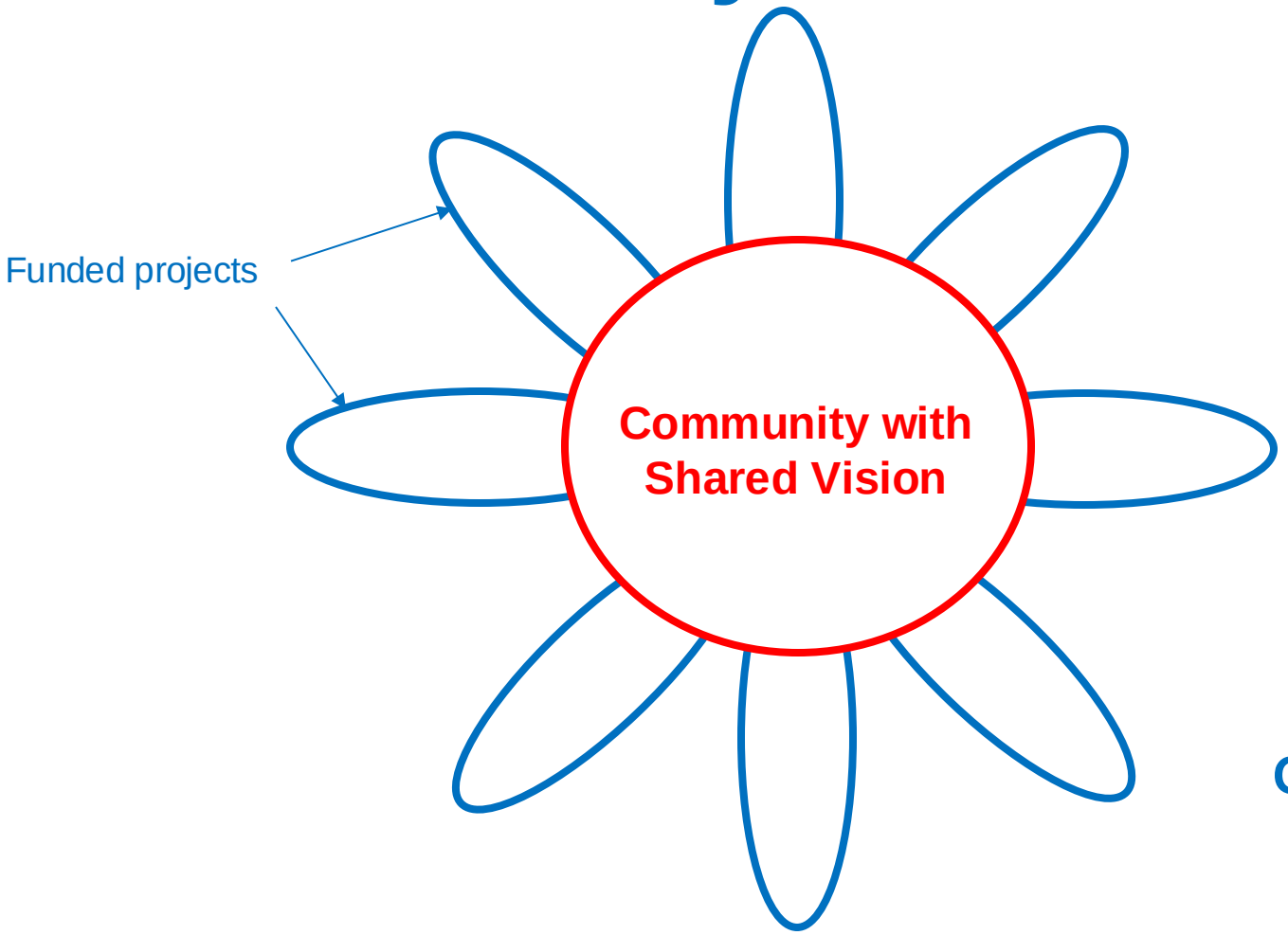
Long Term Vision

- Create an Open National Cyberinfrastructure that allows the federation of CI at all ~4,000 accredited, degree granting higher education institutions, non-profit research institutions, and national laboratories.
  - Open Science
  - Open Data
  - Open Source
  - Open Infrastructure



Openness for an Open Society

Community vs Funded Projects



Lot's of funded projects that contribute to this shared vision in different ways.

PNRP provides core services for the National Research Platform Community

Open Infrastructure is "owned" and "built" by the community for the community

PNRP Project

NSF Category II System

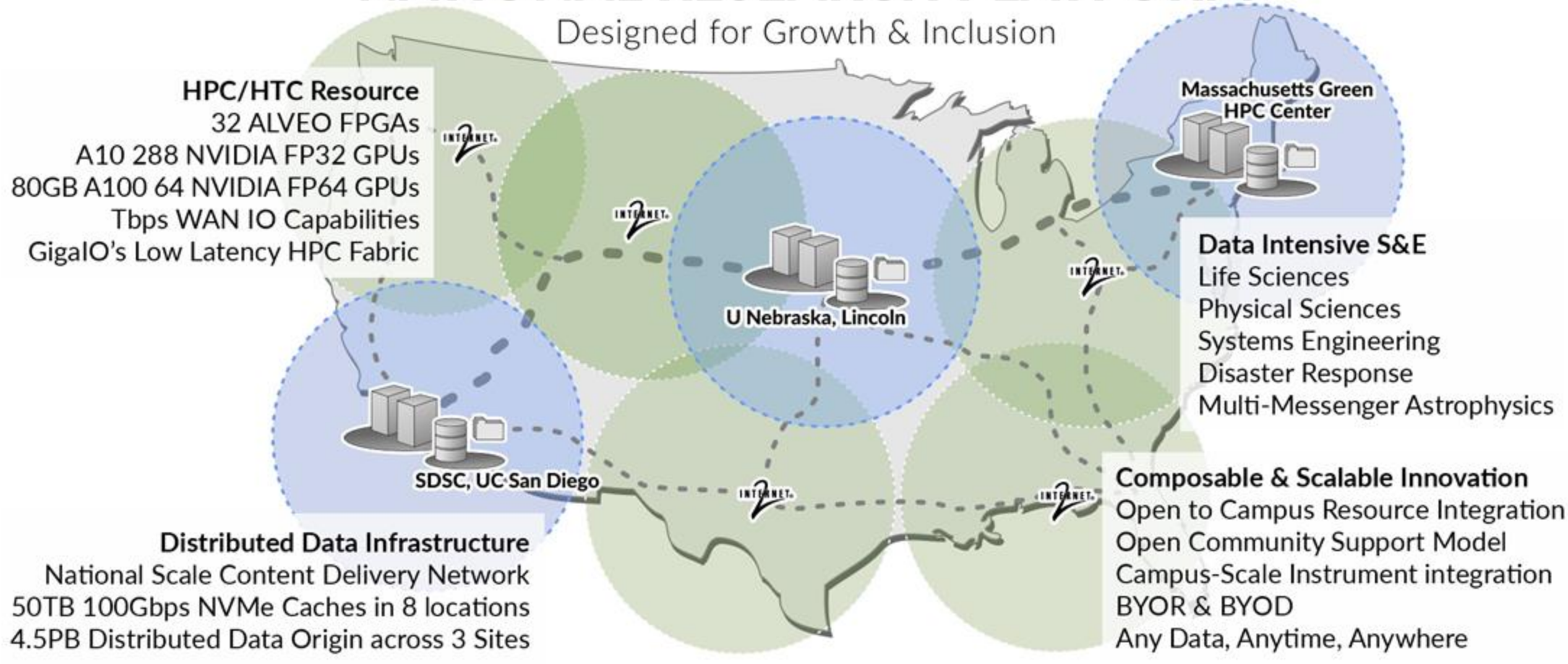
PNRP is designed & deployed as Category II innovative testbed

- NSF, Category II definition: “Resources proposed in this category will be initially deployed as a prototype/testbed system supporting S&E research through delivery of novel forward-looking capabilities and services”
  - 3-year Testbed Phase: Close collaboration with community to explore the PNRP system for transformational S&E
  - 2-year Allocation Phase: PNRP will be made available to the broader community via NSF supported allocation mechanism (e.g., ACCESS, PATh Facility, FABRIC, NAIRR, etc.)

- **Our proposal:**
  - A **nationally distributed system** with hardware deployed in 8 locations within the continental USA
  - An “**Open Cyberinfrastructure for all of Open Science**” …
    - Open horizontally via a Bring-Your-Own-Resources/Devices program
    - Open vertically for higher level service deployments on our Kubernetes layer
  - A composable hardware platform to support **R&D towards domain specific architectures**
    - PCIe hardware composability => allow for nodes that combine GPUs, FPGAs, and large volumes of NVMe

NATIONAL RESEARCH PLATFORM

Designed for Growth & Inclusion



5-year project: $5M for Acquisition and Deployment; $7.25M for Operations and Maintenance

PI = Wuerthwein; Co-PIs: DeFanti, Rosing, Tatineni, Weitzel

Funded as NSF 2112167

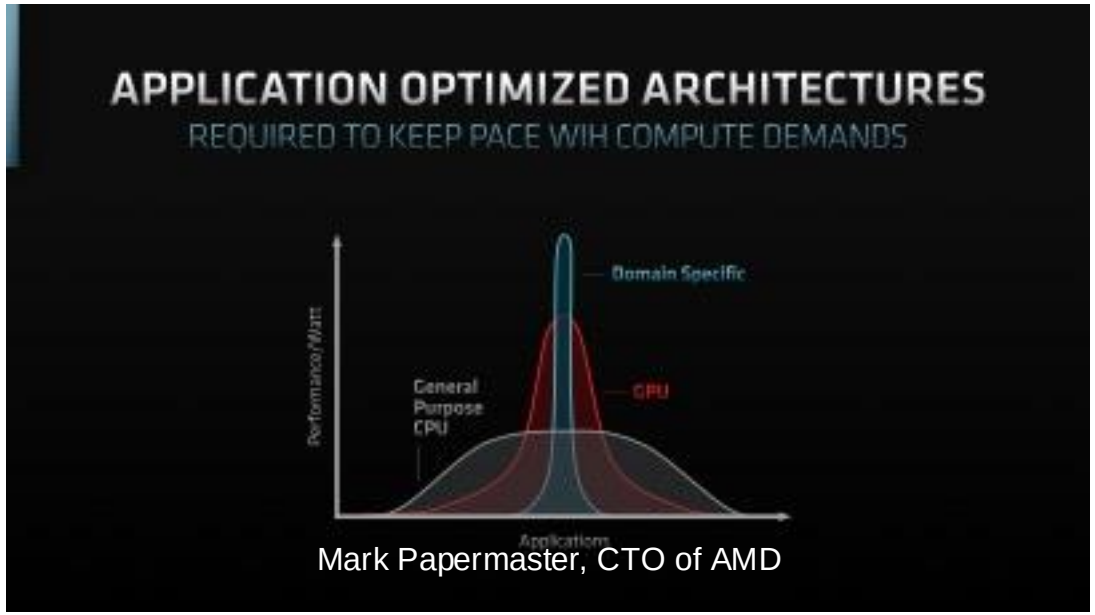
PNRP features innovations in processors, networking, system and data architecture, and operations

- I1: Innovative network fabric that allows disaggregated devices to behave like a single "node" connected via PCIe.
- I2: Innovative application libraries to expose FPGAs hardware to science apps at language constructs scientists understand (C, C++ rather than firmware)
- I3: A "**Bring Your Own Resource**" model that allows campuses nationwide to join their resources to the system.
- I4: Innovative scheduling to support urgent computing, including interactive via Jupyter.
- I5: Innovative Data Infrastructure, including national scale Content Delivery System like YouTube for science.

I3: Bring Your Own Resource (BYOR) innovation thrust of PNRP focuses on expanding the National Research Platform (NRP) Community

“end of Moore’s law” motivates new architectures

Performance improvements vs time
slowed down by O(100)



I1 & I2 motivated by “end of Moore’s law”

PI, Tajana Rosing → **PRISM, a Jump 2.0 project funded by SRC is early user of FPGAs@NRP**

NRP supports FPGAs, P4 switches, NVIDIA DPUs & DGXs

Committed to be a “Playground” of technologies, easily deployed & operated via BYOR and BYOD.

68 Institutions integrated in NRP
43 integrated at IPMI layer

17 Minority Serving Institutions
+10 in EPSCoR states

I5 motivated by the distributed nature of the hardware deployment

“Any Data Anytime Anywhere”

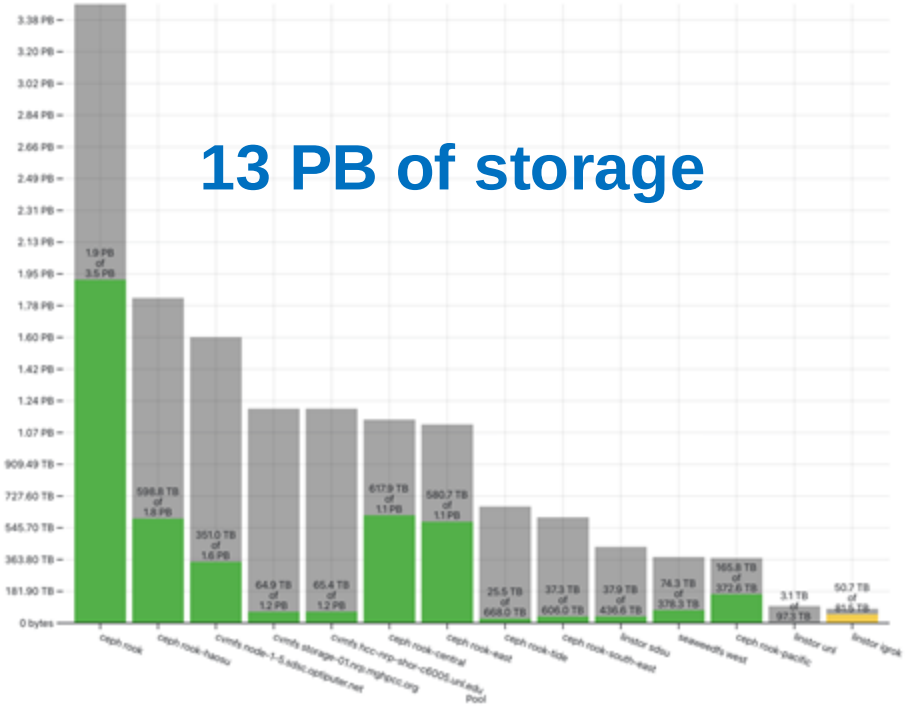
Hardware locations in our Kubernetes “cluster”

Bring Your Own Resources (BYOR) Program

811 GPUs at MSI + 212 in EPSCoR states

Type of Accelerator	# of that type in NRP
A100 80G	144
A100 40G	2
A10	288
A40	10
A6000	48
A5000	6
A4000	32

Type of Accelerator	# of that type in NRP
4090	13
3090	265
2080Ti	178
1080Ti	128
Xilinx U55C	32
Other GPUs	148
Total	**1294**



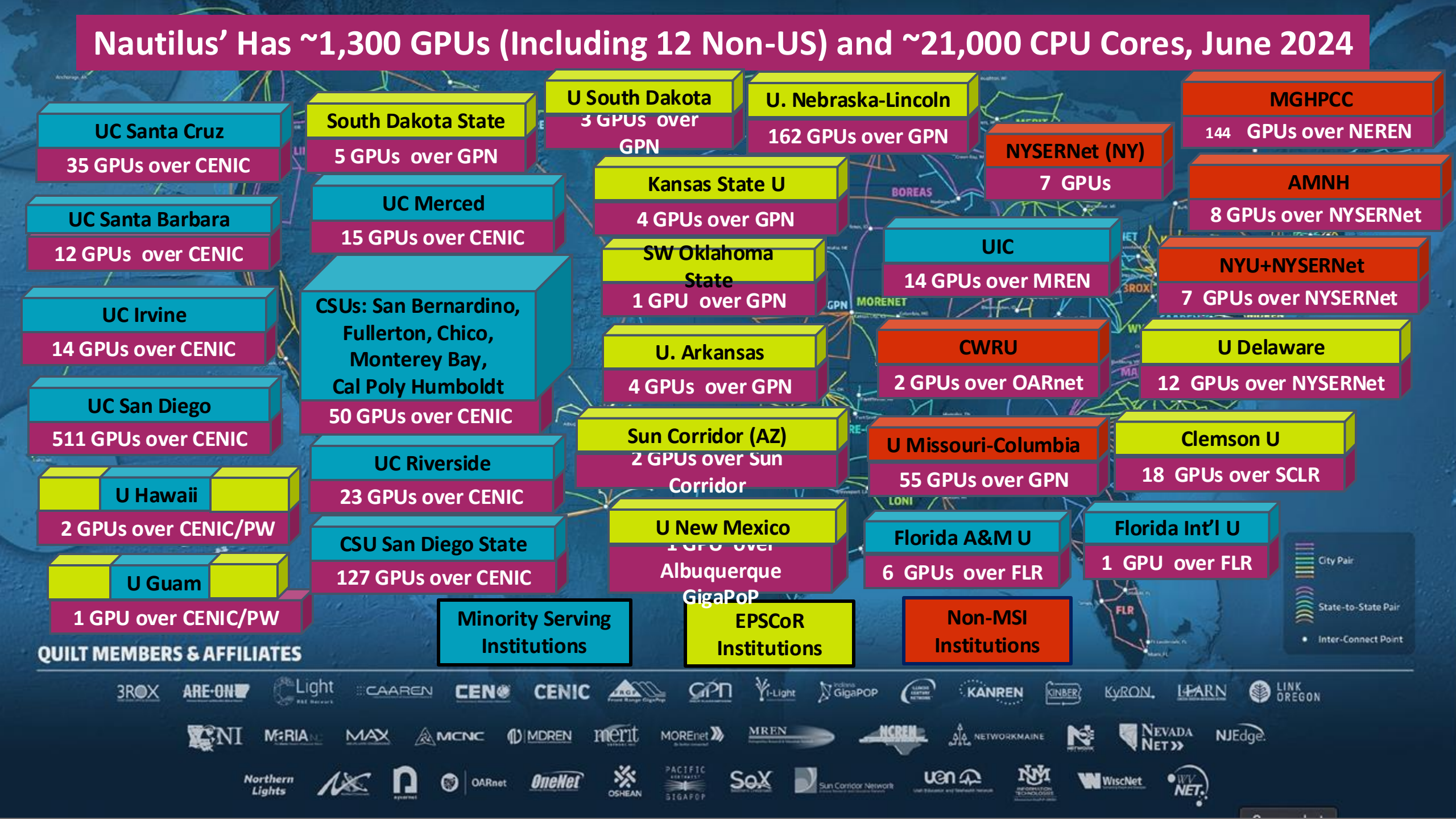
In addition, there are 23,847 x86 CPU cores, 7,406 are in nodes that have no accelerators.

Out of this total the NSF PNRP award (NSF OAC 2112167) contributes:

64 A100 80G
288 A10
32 Xilinx U55C
4 PB of storage

BYOR program has tripled the resources available to the community via the PNRP operated K8S infrastructure

Nautilus' Has ~1,300 GPUs (Including 12 Non-US) and ~21,000 CPU Cores, June 2024



PNRP Networking

- 100G aggregate available at both UNL and MGHPCC
  - 20G to each node
- SDSC 400G aggregate
  - All the GPU and CPU nodes are connected at 100G to edge.
- FPGA can have 2x 100G connections, 24 ports are active.

Supporting Users with Multiple Development Environments

- We provide several development options for our users
- Gitlab – we have instructions for:
  - Building in Gitlab

  https://docs.nrp.ai/userdocs/development/gitlab/

  How to use your private repository

  https://docs.nrp.ai/userdocs/development/private-repos/
  - Integrate Gitlab and Kubernetes vi CI/CD jobs

  https://docs.nrp.ai/userdocs/development/k8s-integration/
- Coder open-source platform for creating and managing developer workspaces - used for FPGA code development; advanced code development on k8s cluster instead of local resources; develop with profiling level access; templates available for common dev setups
  - https://coder.com
- Jupyter-based development environment for FPGA codes with Xilinx tools installed

Use of Coder Environment for ESnet SmartNIC project

- The ESnet SmartNIC Framework harnesses the power of Xilinx Alveo FPGAs, specifically U280 and U55C models, to create Smart Network Interface Cards (SmartNICs) that enhance network processing capabilities.
- Advantages:
  - Custom Logic: Enables flexible P4-programmable packet processing tailored to network requirements.
  - Seamless Integration: Interfaces with SmartNIC runtime firmware for optimized performance.
  - Enhanced Control: Offers command line tools for control plane operations and resource configuration.
  - Testing Capabilities: Provides probe counters for performance metrics evaluation.
  - Efficient Communication: Interacts with FPGAs via Data Plane Development Kit (DPDK) and pktgen for effective packet routing.
- Use of Coder:
  - Original implementation using FPGAs composed via GigaIO fabric to nodes did not work – DPDK issues
  - 12 FPGAs moved directly into compute nodes
  - Coder template developed to support DPDK

Classroom Support

I4: Innovative Scheduling enables Classroom support by NRP

- Interactive access via Jupyter is a foundational feature of PNRP
  - 343 users from 47 institutions used our JupyterHub instance last year (663 users since the start)
  - 77 namespaces have deployed their own JupyterHub instances
  - **Broader Impact & Open CI Vision & Jupyter => Education as a focus on PNRP**

  ⇒ **NAIRR Classroom offering**

    ⇒ **NSF webinar with close to 500 attendees on 6/13/24**

- Owners rule => we need to have scheduling to support owners
  - Kubernetes natively has poor scheduling support
    - We are working with YuniKorn to make it “production ready”

CSU, San Bernardino

- 1,100 faculty members
- 19,000 students
- Serves 2 of CA's largest counties
- Hispanic Serving Institution
- 57% Pell Grant recipients
- Many student oriented projects

Courses Offered utilizing NRP

CSE 5160 Machine Learning

IST 2510 Big Data Analytics

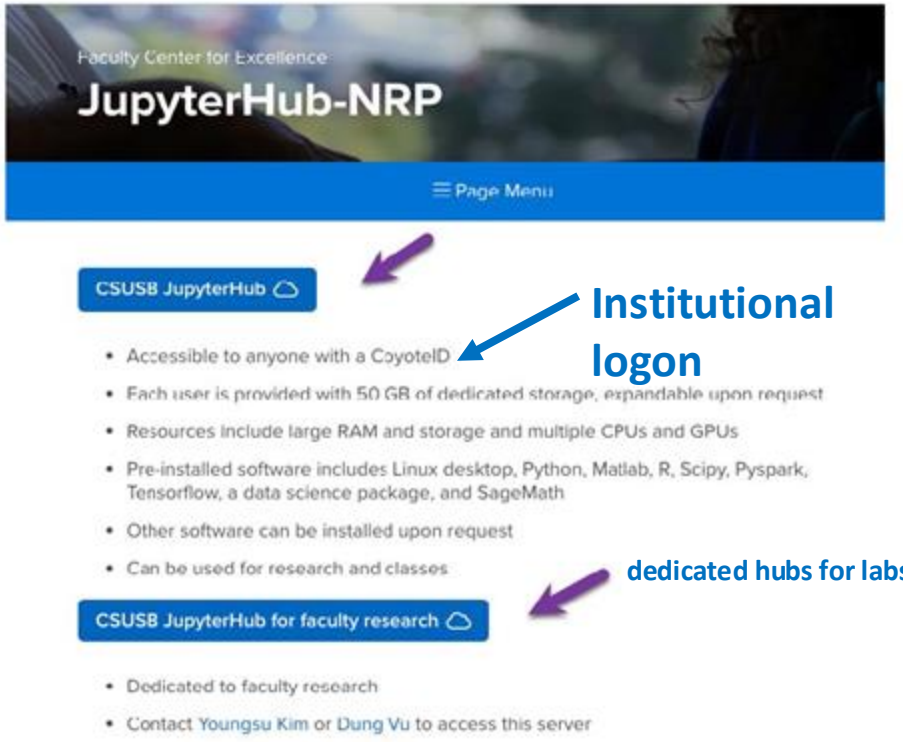
BIOL 5050 - Biostatistics and Experimental Design

IST 6110 - Foundations of Analytics and Big Data

IST 6620 - Business Analytics and Decision-Making

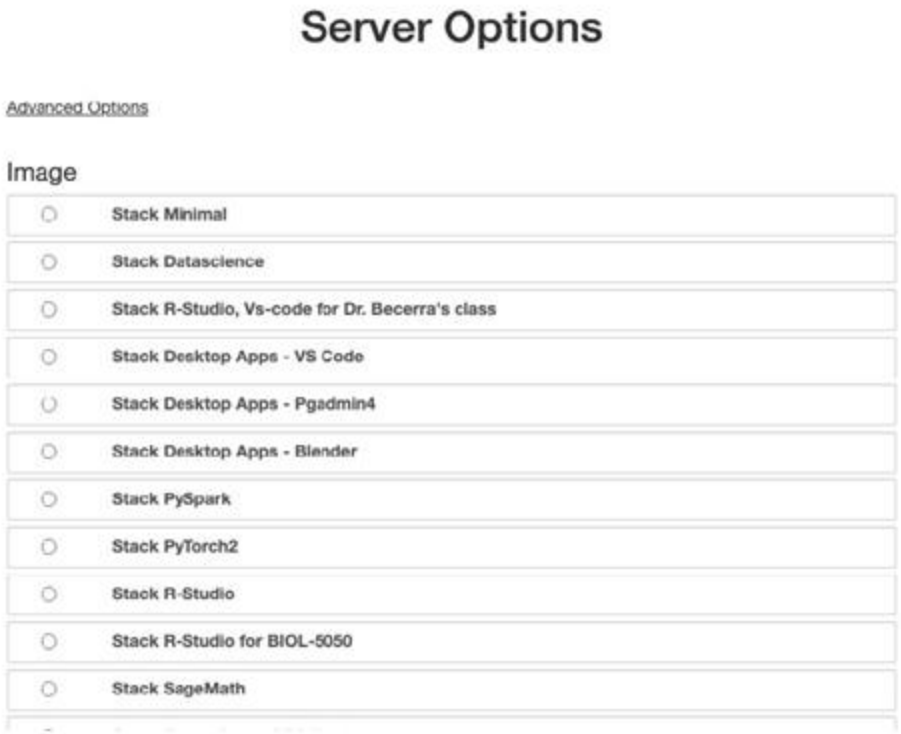
MATH 3465 - Computational Statistics

JupyterHub User Interface at CSU, San Bernardino



Institutional logon

dedicated hubs for labs



Wide range of dedicated hubs

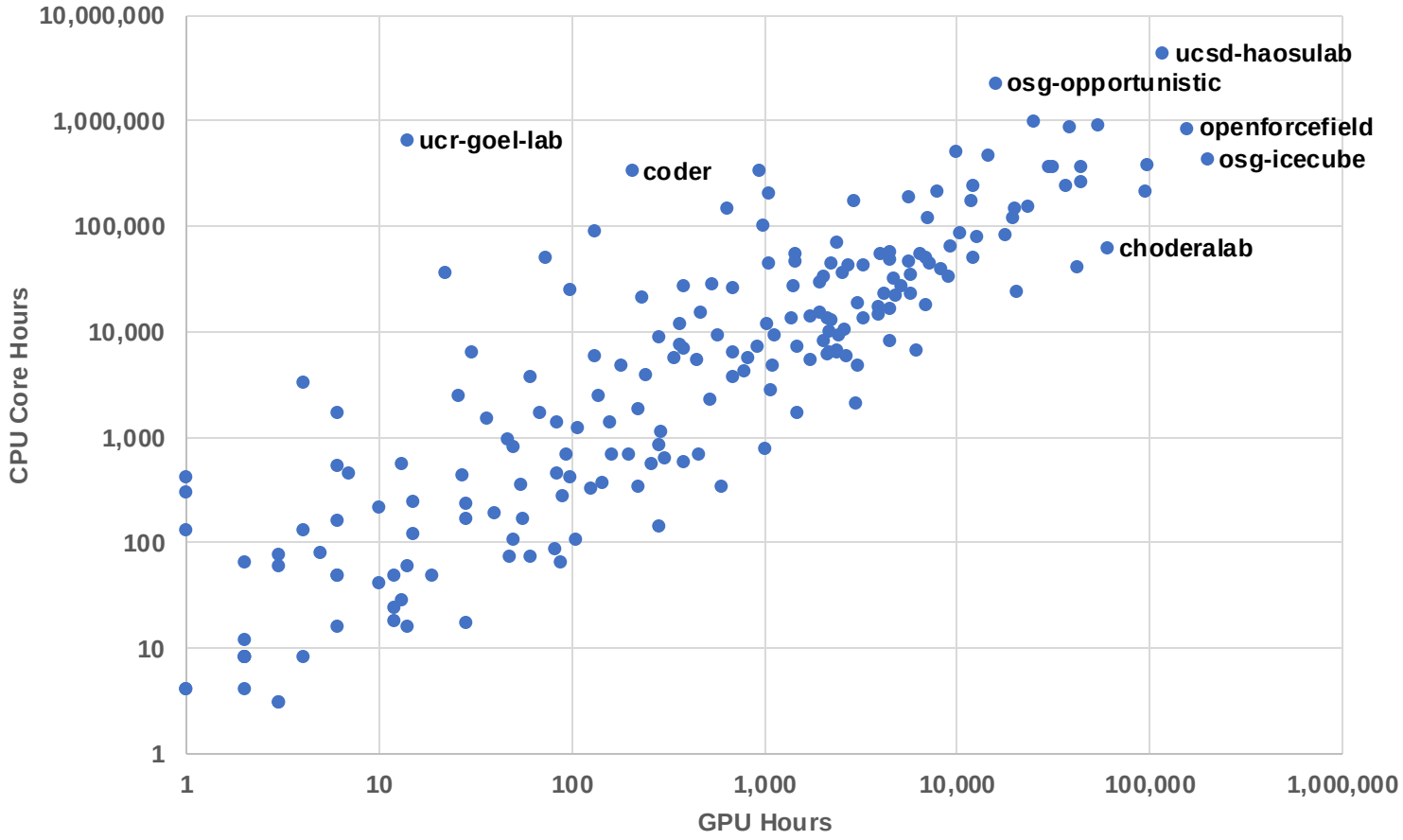
https://csusb-metashape.nrp-nautilus.io: 3D modeling
https://csusb-vasp1.nrp-nautilus.io Viena Ab initio Simulation package (VASP)
https://csusb-cousins-lab.nrp-nautilus.io: VASP simulation
https://csusb-becerra.nrp-nautilus.io AI/ML project
https://csusb-biol-5050.nrp-nautilus.io: Biology course
https://csusb-cse-salloum.nrp-nautilus.io Summer Research
https://csusb-drhamoudahub.nrp-nautilus.io Data Analytics
https://csusb-ratnasingam.nrp-nautilus.io Data Analytics
https://csusb-zhang.nrp-nautilus.io AI/ML project

PNRP NAIRR Classroom: Resources

- Jupyter resources for classes that need AI/ML – can be teaching about AI/ML or using AI/ML in domain science
- Can use institutional authentication methods leveraging InCommon
- Professor and TA will have admin status to create namespaces for classes
- Support/Community interaction through Matrix channels – one for the class that includes students so Professor/TA can help them and second for Jupyter admins that includes admins from other classes
- Direct ticketing system available for any issue needing longer follow ups
- GPU nodes each with 8 A10 GPUs, 512GB of RAM, 2 AMD EPYC 7502 CPUs, and 8TB of NVMe.
- Depending needs, other GPU resources can be provisioned via the NAIRR pilot and made available through this platform

System Usage Statistics

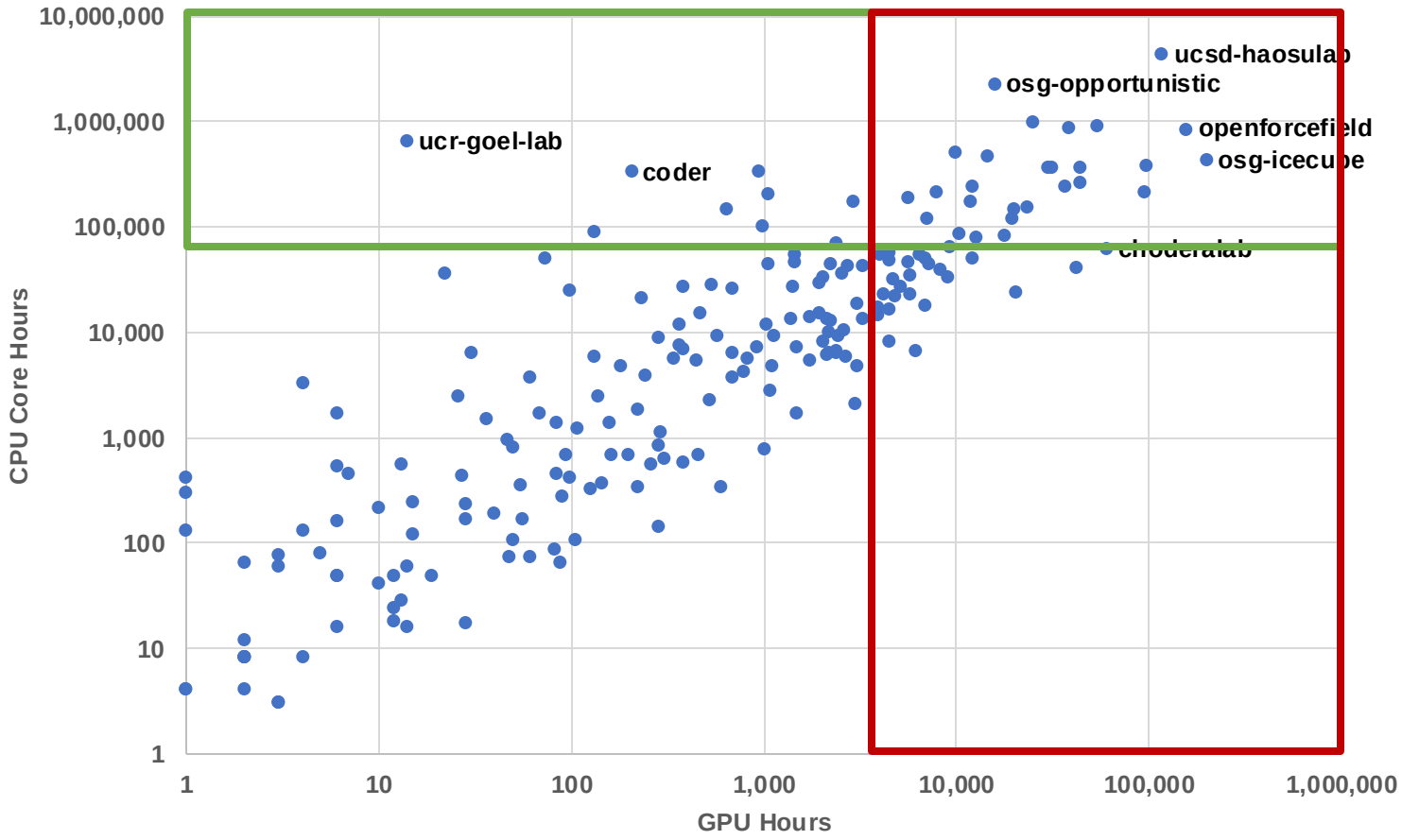
Top 200 PNRP Namespace GPU Users Also Use CPU Cores



PNRP Namespace GPU/CPU Usage 10.1.2023 to 6.1.2024

One Million-Fold Dynamic Range In Usage!

Top 200 PNRP Namespace GPU Users Also Use CPU Cores

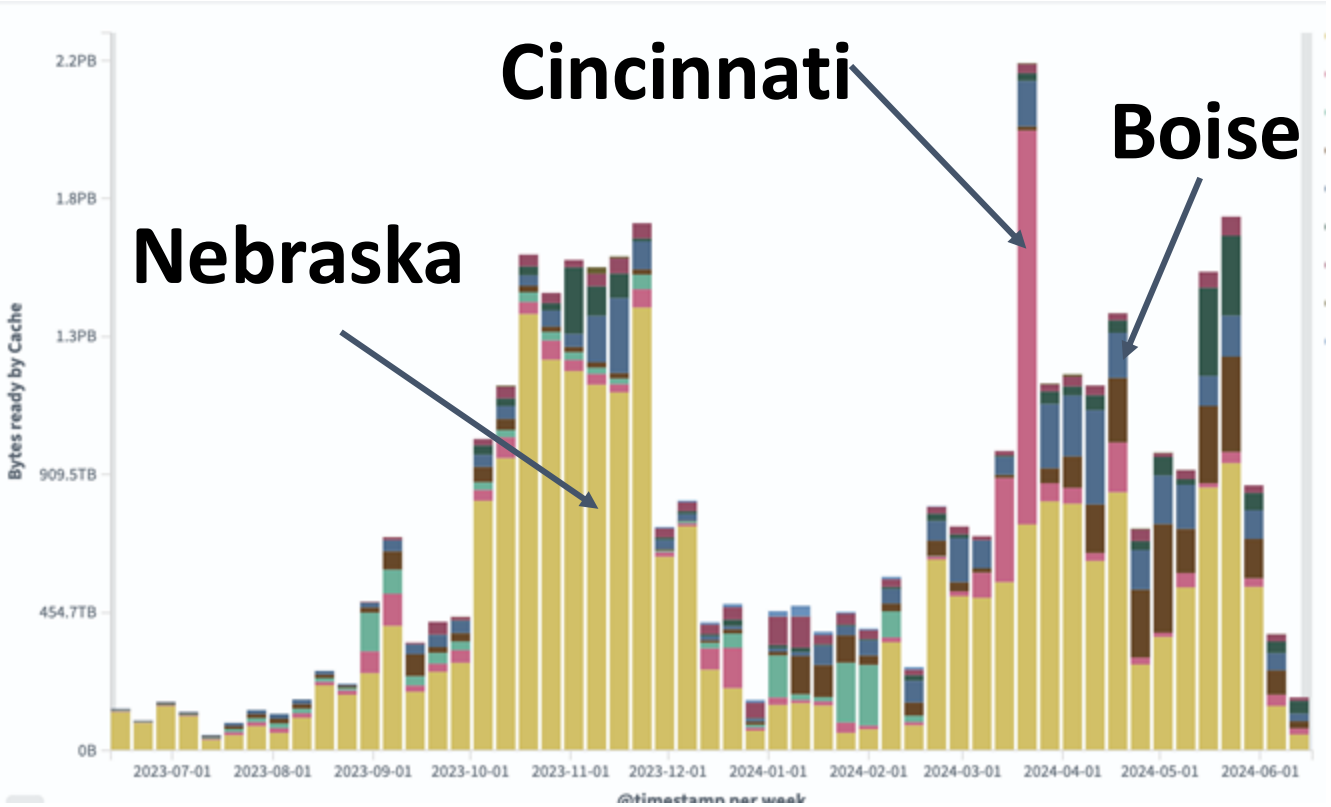


Largest GPU Namespace Users

Largest CPU Namespace Users

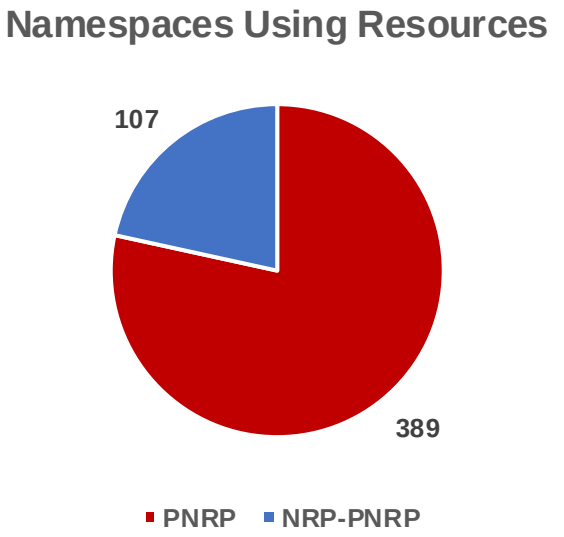
PNRP Namespace GPU/CPU Usage 10.1.2023 to 6.1.2024

Cache Usage - 1 Year

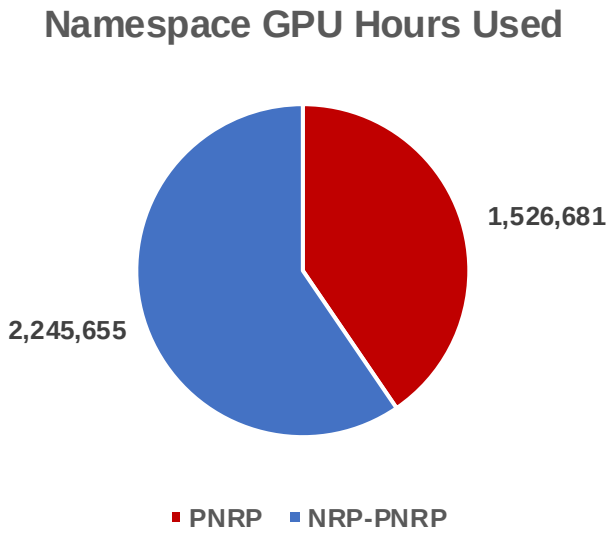


Site	Files	Transferred
Nebraska	249,608,810	24.4PB
Boise	36,637,795	3.5PB
Cincinnati	15,340,832	3.3PB
San Diego	22,103,395	3.1PB
Jacksonville	49,219,694	1.6PB
MGHPCC	29,508,257	1.3PB
Denver	8,065,565	1.3PB
Houston	4,316,063	186TB
Total	414,800,475	38.7PB

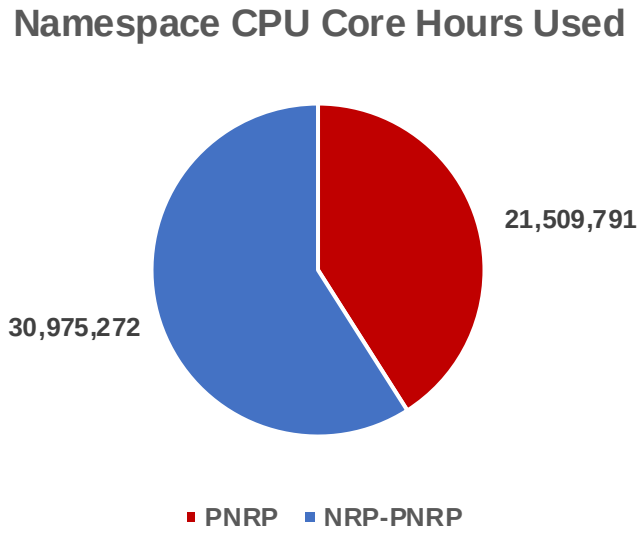
The PNRP is a Subset of the NRP



496 Total NRP Namespaces



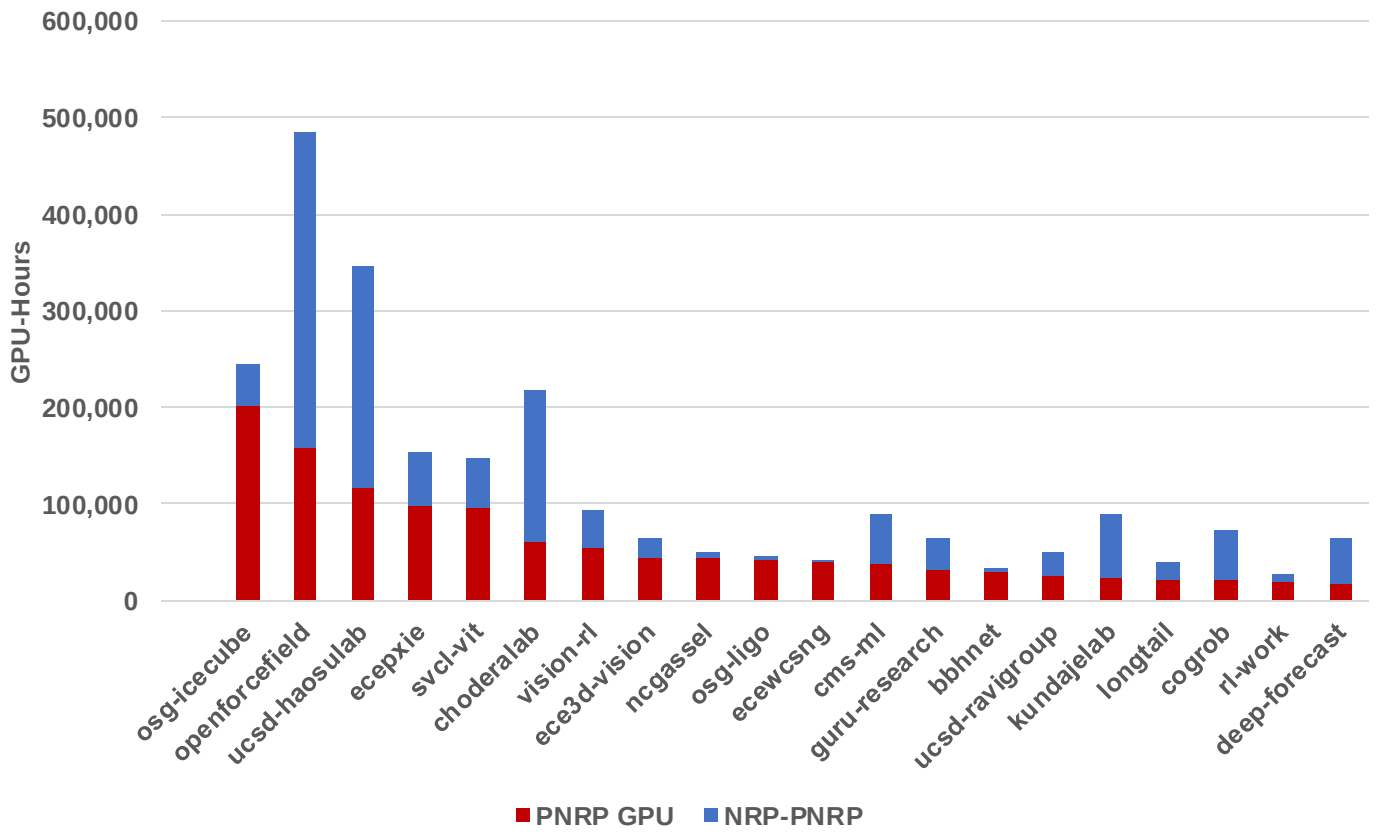
3,772,336 NRP GPU Hours Used



52,485,064 NRP CPU Core Hours Used

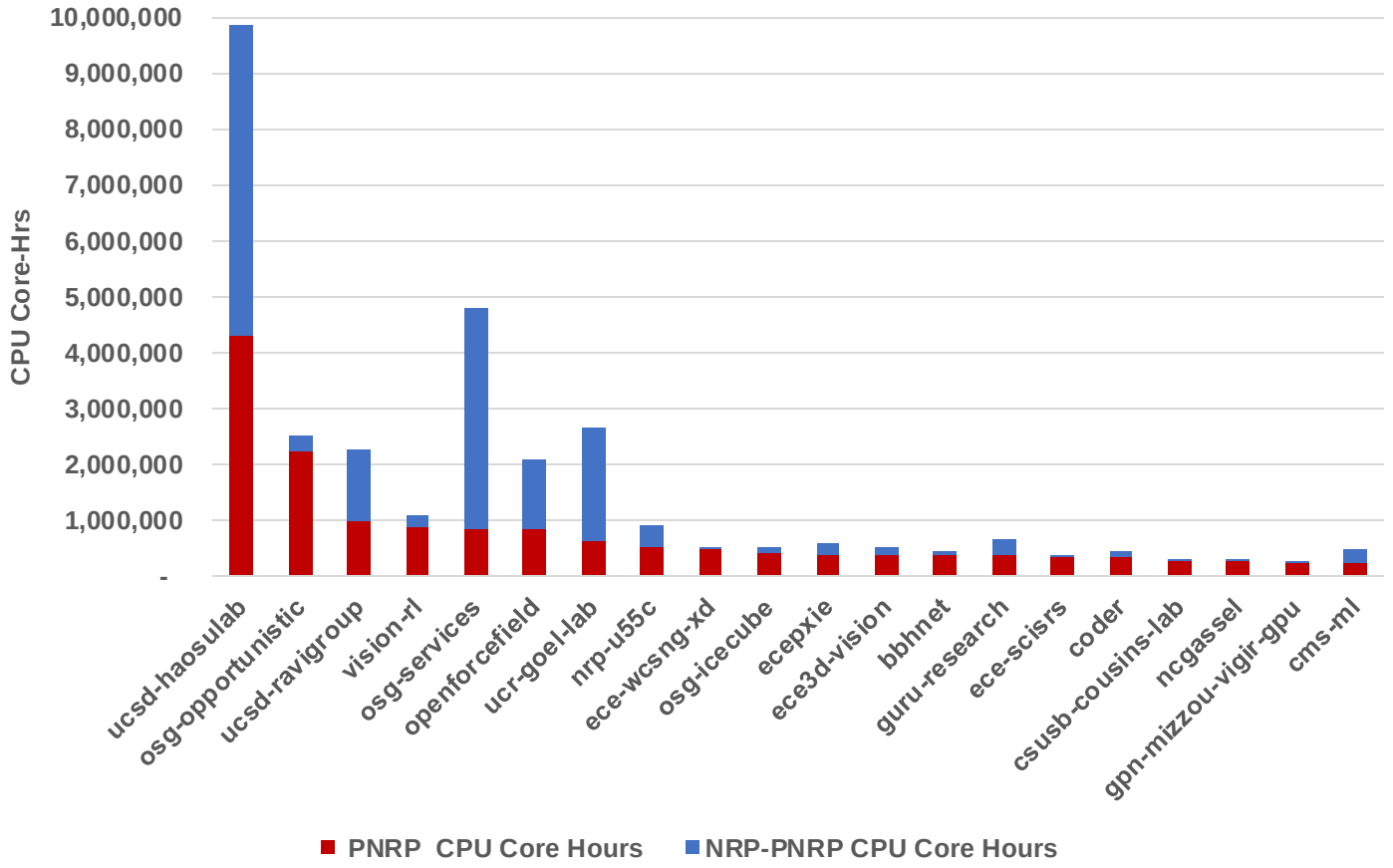
A namespace is a project with a PI. Typically, each namespace has multiple, or even many users in them.
Some PIs have multiple namespaces to track their different activities independently.
=> Total of 200 distinct PIs across 89 distinct institutions

PRNP Namespaces Can Burst Into NRP for Additional Resources – GPUs



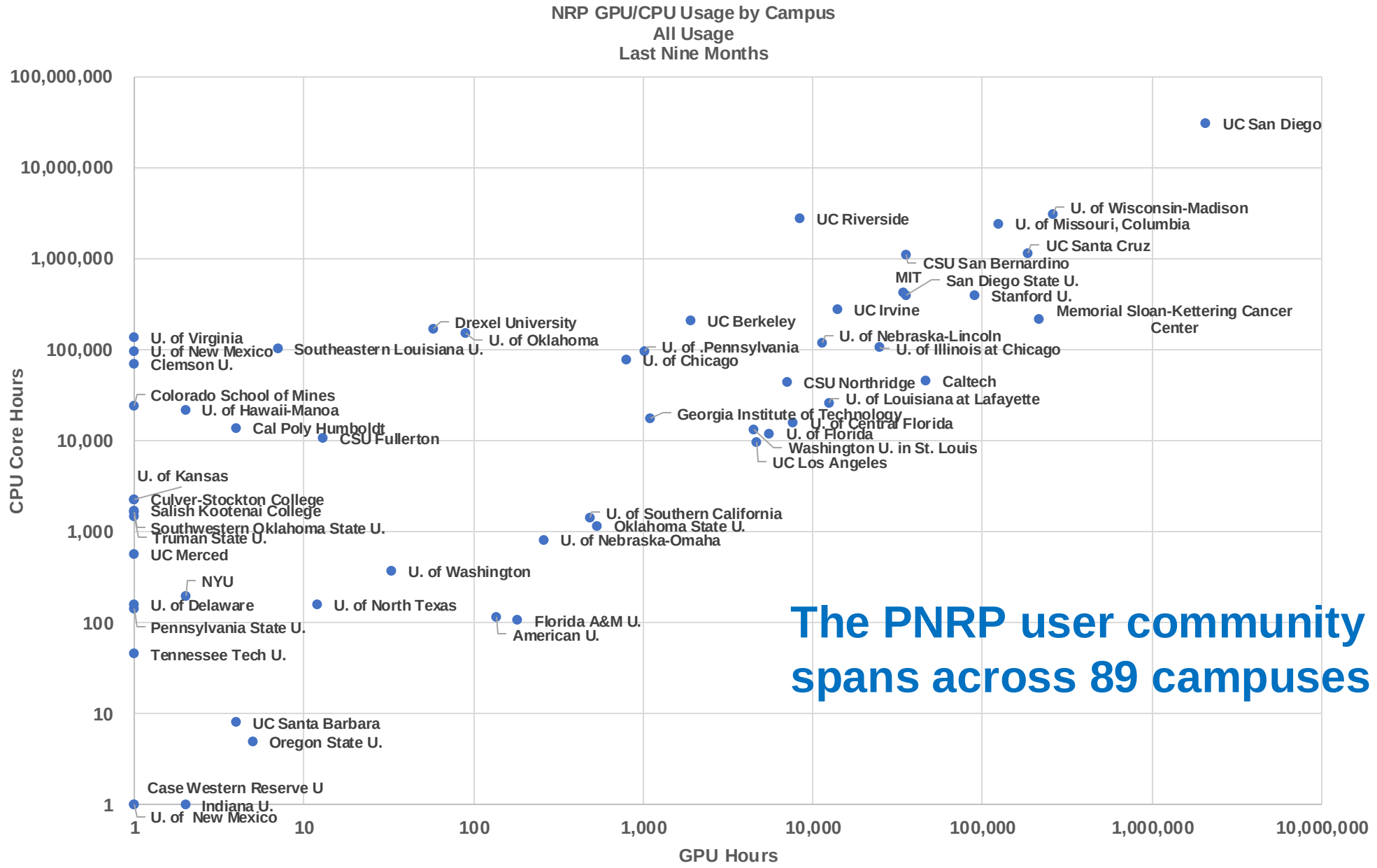
PNRP Namespace GPU-Hr Usage
& Additional GPU-Hrs from NRP
Top 20 PNRP GPU Namespace Users
10.1.2023 to 6.1.2024

PRNP Namespaces Can Burst Into NRP for Additional Resources – CPU cores



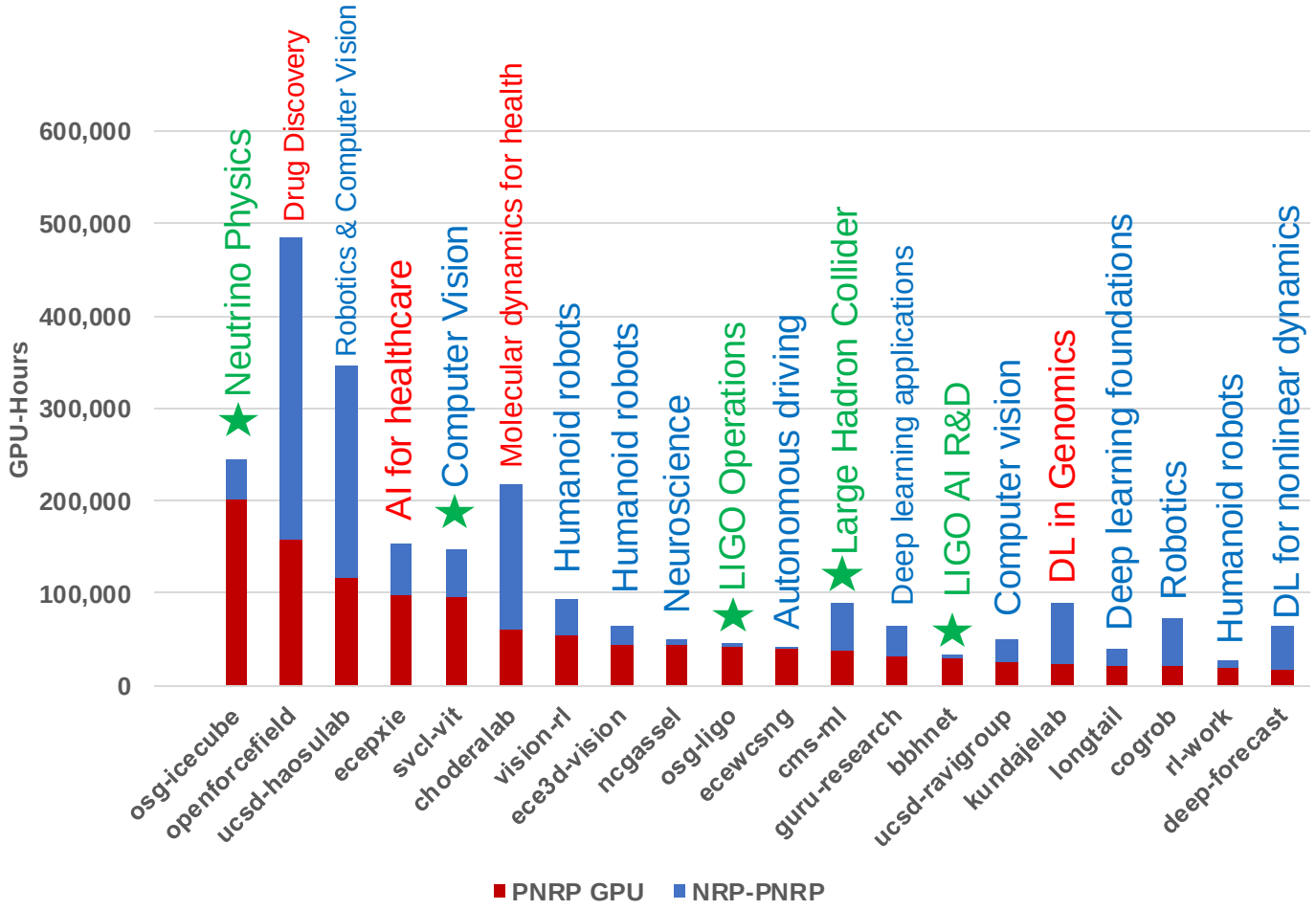
PNRP Namespace CPU Core-Hr Usage
& Additional CPU Core-Hrs from NRP
Top 20 PNRP CPU Namespace Users
10.1.2023 to 6.1.2024

Science Impact



The PNRP user community spans across 89 campuses

PNRP Science community substantially expanded since the proposal



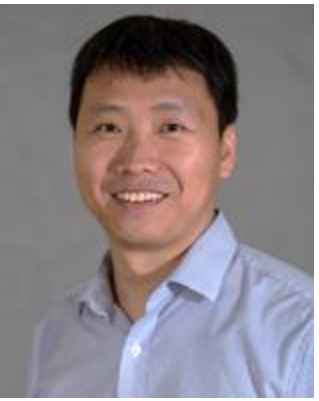
Most of our top 20 GPU users were not listed in the proposal

PNRP Namespace GPU-Hr Usage & Additional GPU-Hrs from NRP Top 20 PNRP GPU Namespace Users 10.1.2023 to 6.1.2024

Three big themes dominate:

- Big Science in Physics & Astronomy
- Engineering (Computer Vision, Robotics, AI)
- Human Health

CSU Northridge Prof. Bingbing Li
Machine Learning Research Projects Utilizing PNRP



CSUN Prof.
Bingbing Li

CSUN | ARCS Arctocomy Research Center for STEAIM

- 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

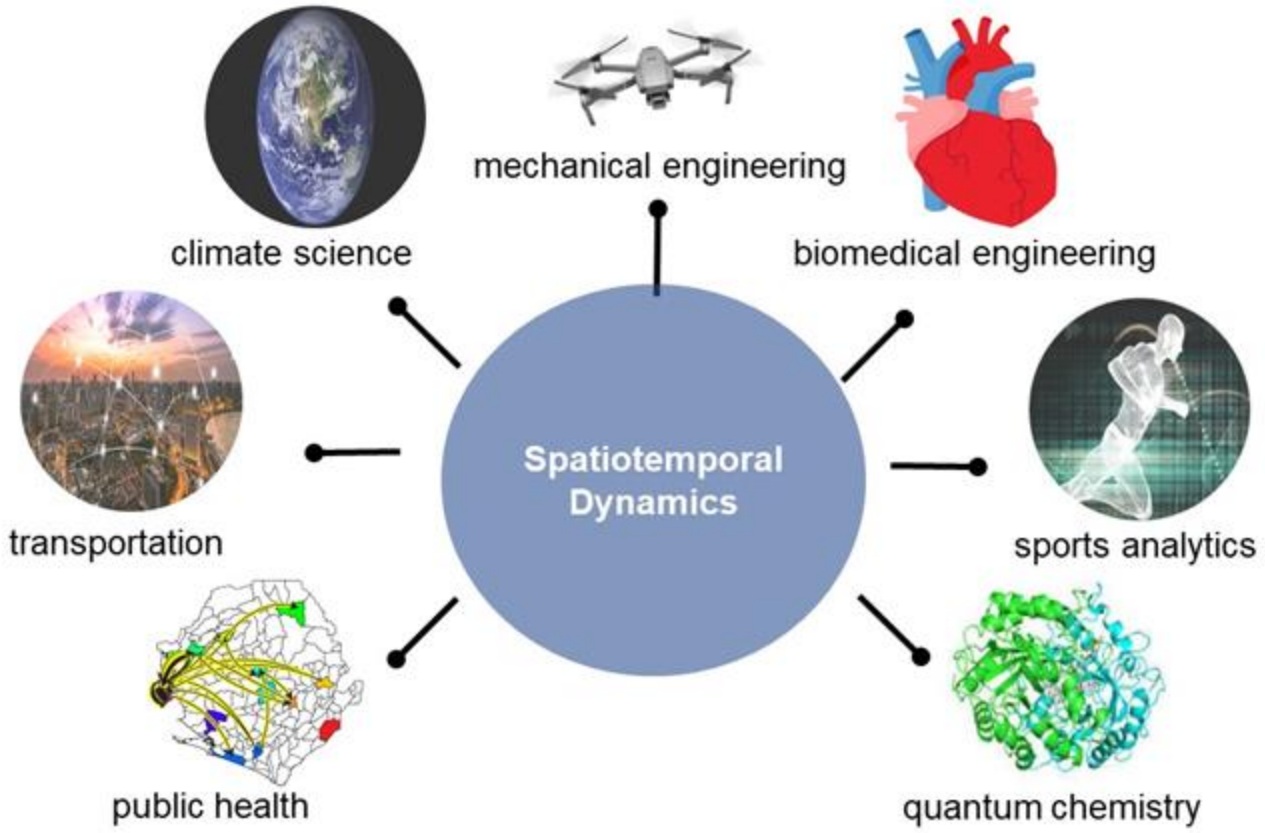
ChatGPT

Meta

PNRP Usage Last 9 Months:
6.4k GPU-hrs & 37k CPU Core-hrs
Namespaces cesmii-scw, nsf-maica

Source: Bingbing Li

Physics-Guided AI for Large-Scale Spatiotemporal Data: Learning Spatiotemporal Dynamics



Rose Yu uses I5 in DOE-FES award

Rose Yu owns 144 GPUs & a CEPH server in NRP

PNRP Usage Last 9 Months:
25k GPU-hrs & 146k CPU Core-hrs
Namespaces *deep-forecast, spatiotemporal-decision-making, climate-ml, ai-md*

Source: Rose Yu
UC San Diego

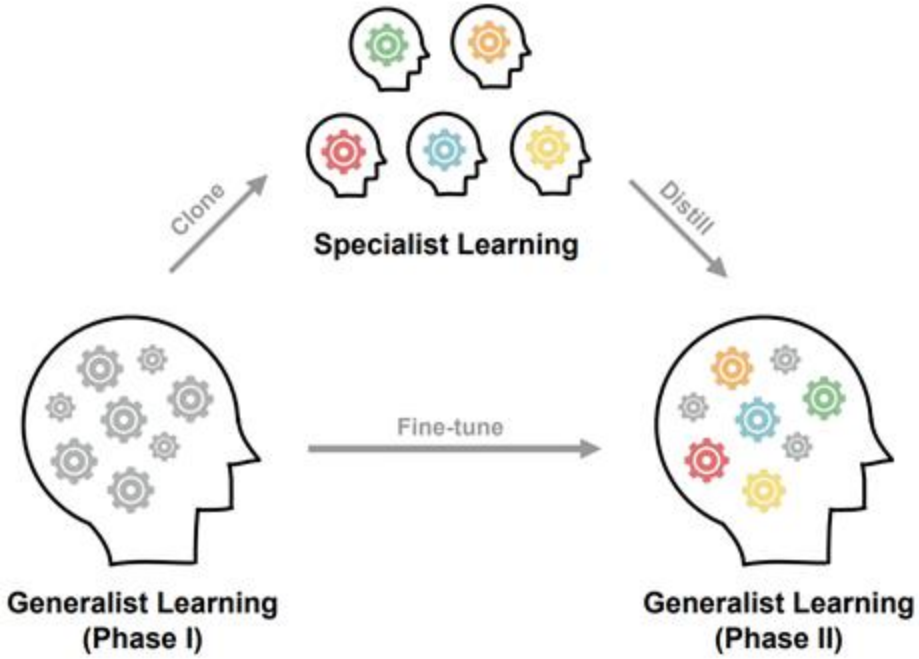
A Major Project in UCSD's Hao Su Lab is Large-Scale Robot Learning

Hao Su owns 192 GPUs & Ceph in NRP

- 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

PNRP Usage Last 9 Months:
117k GPU-hrs & 4.3M CPU Core-hrs
Namespace *ucsd-haosulab*



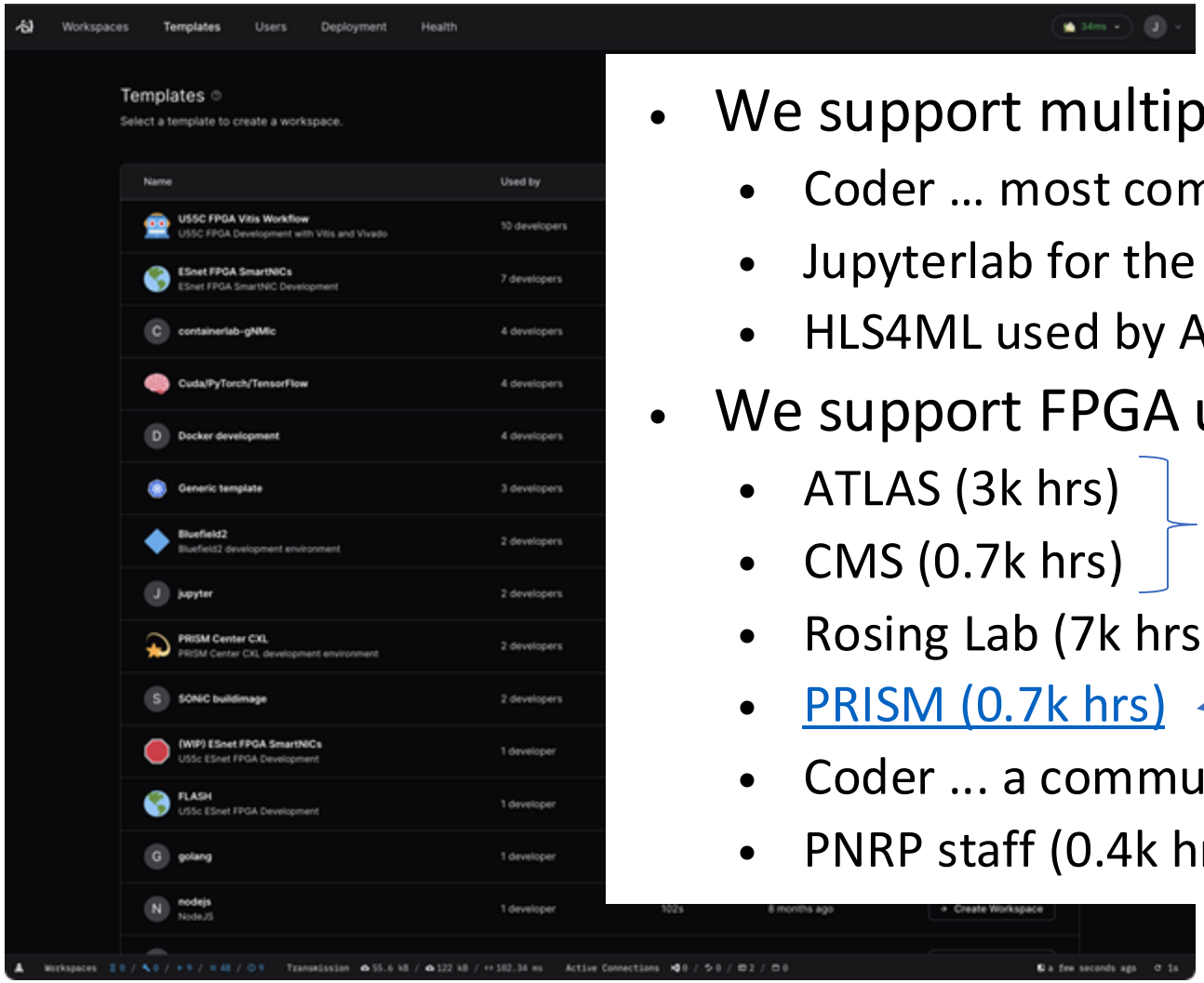
Source: Prof. Hao Su, UCSD

NRP brings CS R&D and Domain R&D onto the same platform

NRP blurs the lines between ”testbed” and “production” CI

Create social cohesion to accelerate domain science adoption of new programming paradigms & architectures

Use of FPGAs within the 9 months 10.1.2023 to 6.1.2024



- We support multiple software environments: (22k hrs)
  - Coder … most common … supporting multiple templates
  - Jupyterlab for the PRISM project
  - HLS4ML used by ATLAS and CMS
- We support FPGA use across 5 major namespaces
  - ATLAS (3k hrs)
  - CMS (0.7k hrs)

  LHC trigger R&D
  - Rosing Lab (7k hrs) — Genomics, data analytics, cybersecurity, …
  - PRISM (0.7k hrs) — Precision medicine & data analytics
  - Coder … a community namespace … (11k hrs)
  - PNRP staff (0.4k hrs)

Users from 10 institutions

Ongoing Projects

- Making PNRP FPGAs, DGX, and NVMe available to FABRIC community
  - LHC Data Management Stack Overhead Investigation
  - Network bandwidth accountability via SENSE/AutoGOLE
  - ESnet SmartNIC and DPDK
- Packet header inspection via ESnet High Touch
- Wireless to Wireless via PNRP
- Microbiome gene sequencing on FPGAs
- PRISM Center R&D on PNRP (SRC JumpStart 2.0 project)

Summary & Conclusions

- **NRP addresses 3 challenges:**
  - The gap between those who have and those who can't afford is getting wider
  - CI needs for Education are growing
  - The "end of Moore's law" is leading to a proliferation of "architectures" … domain science adoption is at risk …
- **Enabled by NSF funded PNRP Cat II project which does so via 5 Innovation thrusts. *NRP core services (systems and people) are supported by this project.***
- **Exceptionally strong science impact enabled by NRP, dominated by:**
  - Big Science in Astronomy & Physics
  - AI across multiple areas of engineering
  - Health and life sciences
- **BYOR program is a significant driver of Broader Impact, and vice versa.**
- **Deployed several development environments to support NRP's innovative features (Gitlab, Coder, CI/CD integration between gitlab and Kubernetes cluster)**

Project team has excelled in deploying an innovative, distributed system and supported the system in Operations since March 2023

SDSC: User support; Lead Nautilus Operations; Operations of hardware at SDSC

- Fabio Andrijauskas
- Tom DeFanti (Co-PI)
- John Graham
- Christopher Irving
- Tom Hutton
- Elham Khoda
- Dima Mishin
- Nick Patience
- Tajana Rosing (Co-PI)
- Mohammad Firas Sada
- Scott Sakai
- Igor Sfiligoi
- Fernando Silva
- Robert Sinkovits
- Shawn Strande
- Mahidhar Tatineni (Co-PI)
- Nicole Wolter

Thank you to this outstanding team, whose expertise and commitment has made all this possible!

We are grateful for **student contributions** to the project

- Aashay Arora, Isaac Neely, Kevin Vo Khai Vu

UNL: Nautilus Operations; Operations of hardware at UNL & MGHPCC & BYOR

- Derek Weitzel (Co-PI), Garhan Attebury, Ashton Graves, Huijun Zhu

MGHPCC: FP32 system hosting and remote support

- John Goodhue, Greg Shomo, Jim Culbert

Internet2 (in-kind contribution): Remote hands for 5 caches at Internet2 POPs

- Chad Sorrell, Matt Mullins