Towards a Globally Distributed Open Cyberinfrastructure

Frank Würthwein
Director, San Diego Supercomputer Center

October 25th 2024

Why
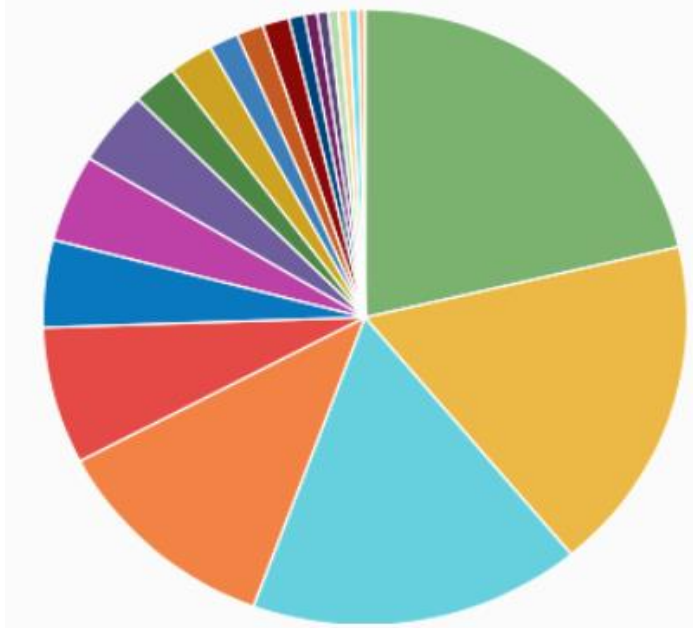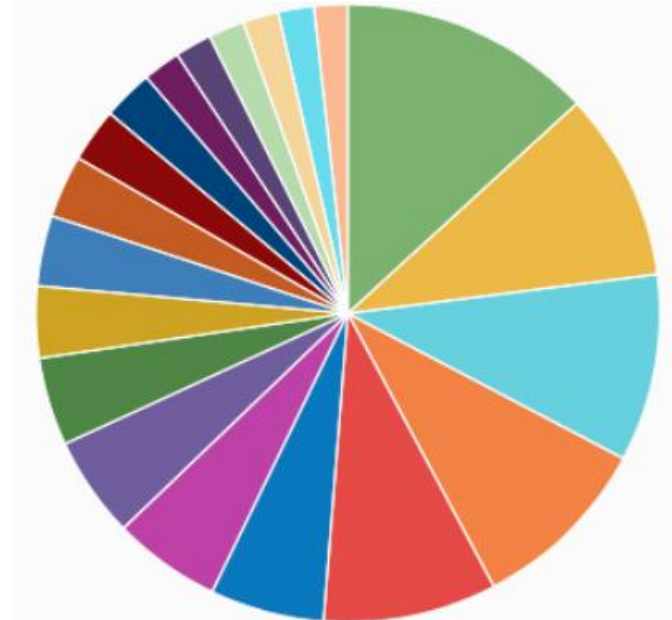
Big Science is a Team Sport

LIGO Scientific Collaboration
~ 1200 members ~ 100 institutions, 18 countries

Global Computing & Data Infrastructure

LIGO computing

Many funding agencies contribute hardware in many places.

IceCube computing

Distribution of computing matches distribution of researchers that collaborate on the science.

AI is an Experimental Science

The future AI workforce of the USA
is educated across
3900 Institutions of higher learning.

From data in
https://arxiv.org/abs/2202.05924
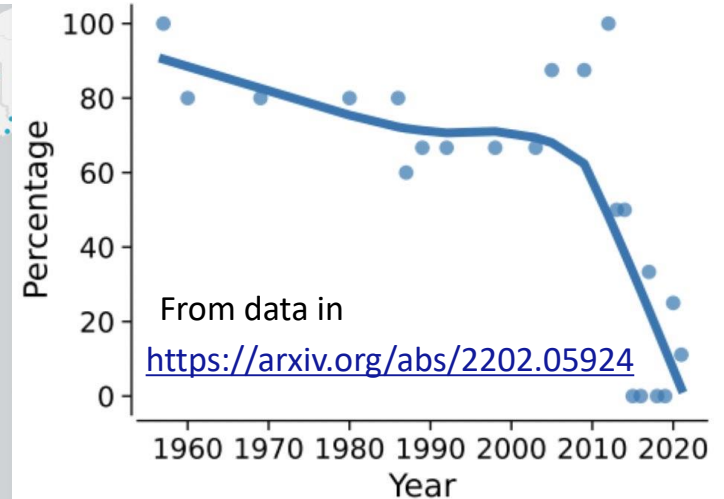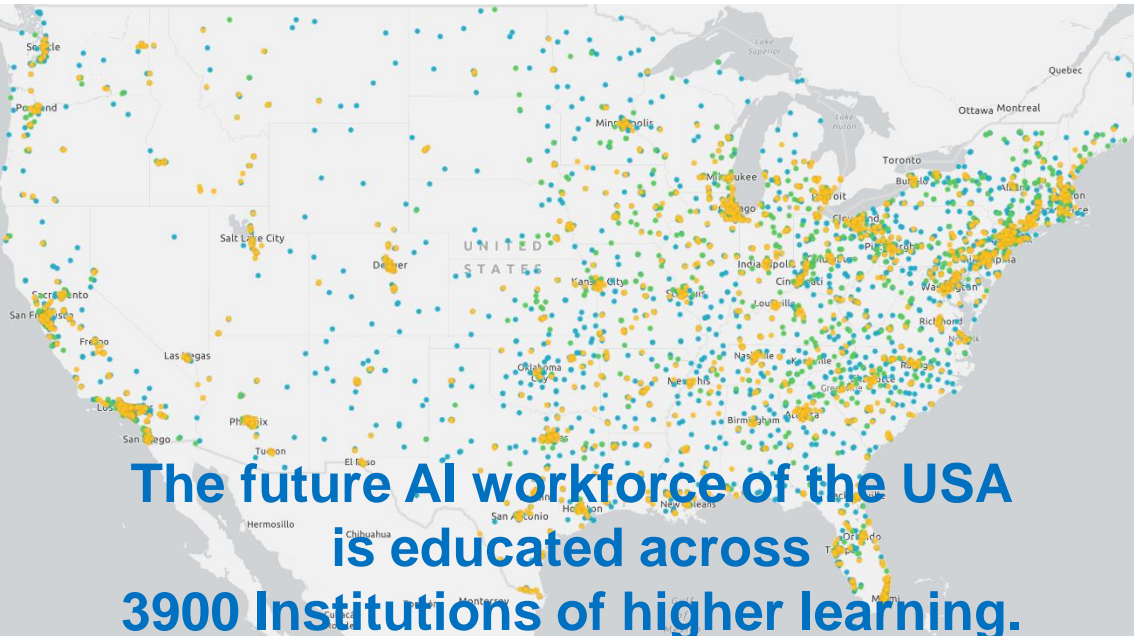
% of large-scale AI
results from academia
is basically zero

To guarantee access to AI education at these 3900 institutions
requires rethinking of how to provide cyberinfrastructure.

Democratize Access

The Minds We Need

- **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

6

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
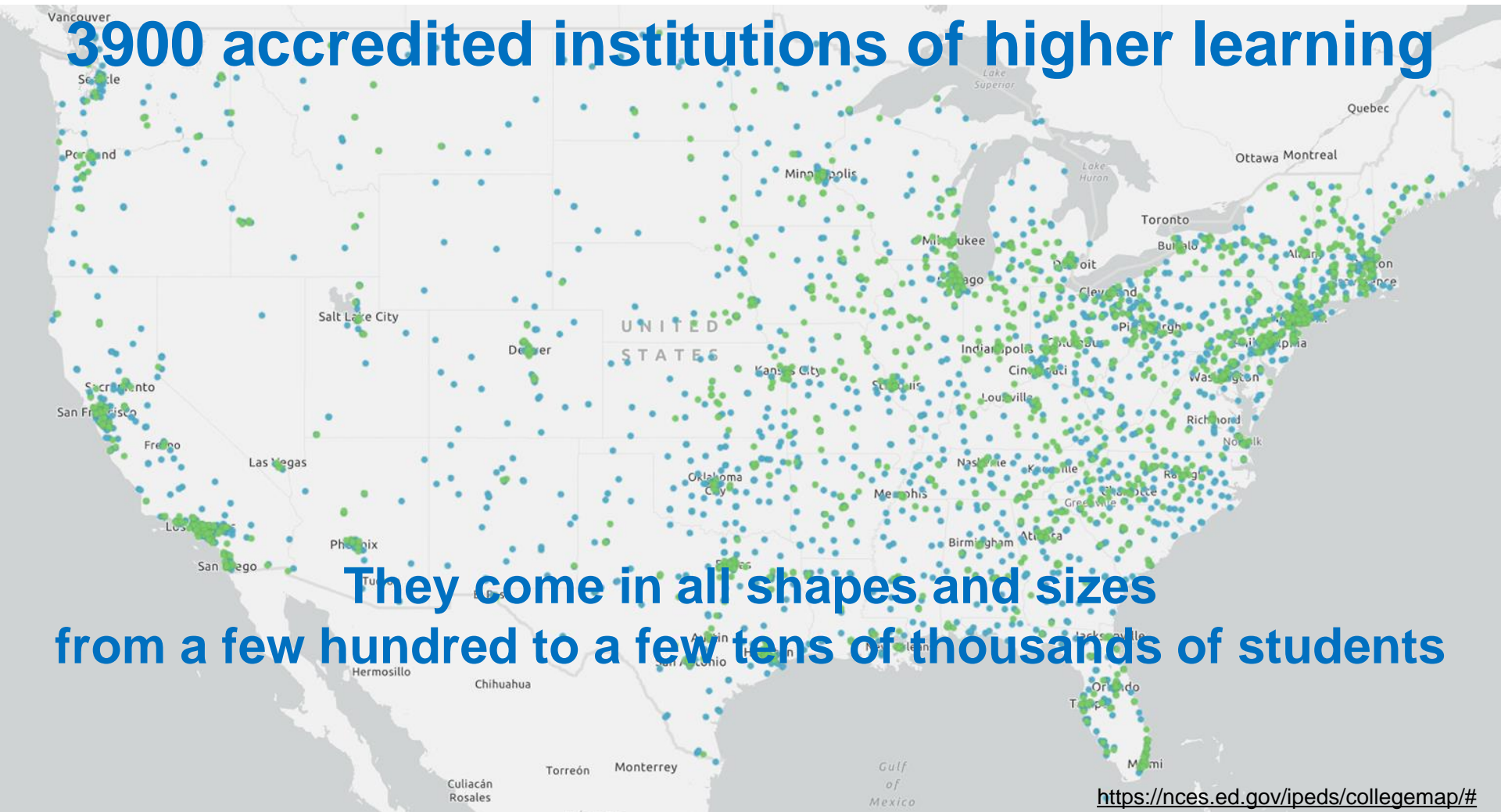
What should the role of NRAO be?

- Design, Develop & operate instruments
- Take data and make it available to the community
- Provide the computing resources to the community to
  - train AI models on the data
  - do inference on the data
  - analyze the data

This talk is relevant to you if NRAO wants to do the first 2 but not the 3rd.

About 2/3 of the Compute Resources operated by SDSC are not available via federal allocation mechanisms

SDSC operates roughly

3,000 GPUs

250,000 CPU cores

100PB of storage

And that's only 1/3 of what's in our data center

What about the Cloud?

- The fundamental difference for scientists between on-prem and cloud is what happens if they underestimate their computing needs.
  - On-prem: your career goals are delayed
  - Cloud: your career dies

Disclaimer:
SDSC is the cloud "reseller" and "support organization" for the NSF across Azure, AWS, GCP, NVIDIA DGX, IBM Quantum, …

What

Vision of an Open Infrastructure

Horizontally open => institutions can integrate their resources

Vertically open => projects can build on the infrastructure

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 ← Open Compute
    - ← Open Storage & CDN
    - ← Open devices/instruments/IoT, …?

Openness for an Open Society

How

Community vs Funded Projects

Community with Shared Vision
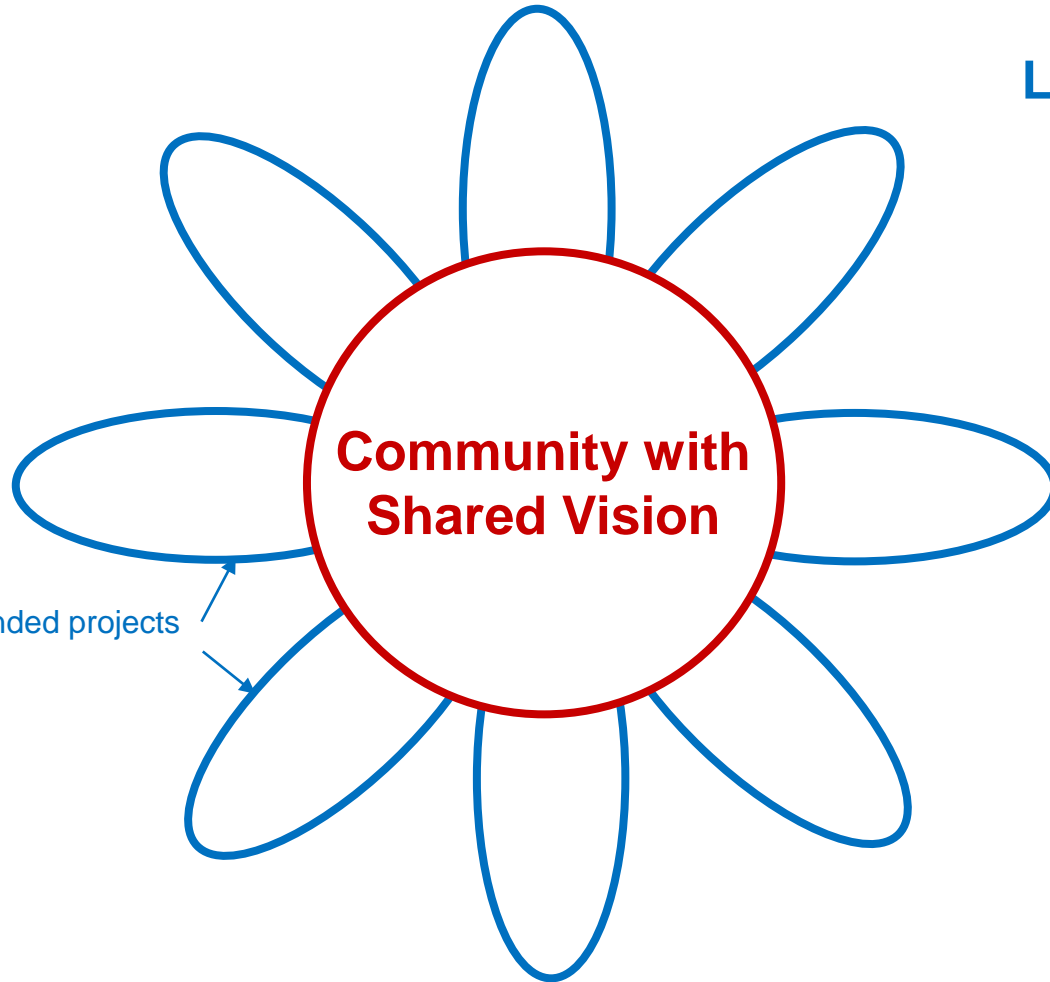
Funded projects

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

Hardware funded by NSF, DOD, DOE, …

<u>Petals of this flower include:</u>
Prototype National Research Platform
PATh
    Open Science Data Federation
    Open Science Compute Federation
National Data Platform
Pelican
Fusion Data Platform for AI
R&E platforms for multiple campuses
    GP-Engine
    TIDE
    …

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

NATIONAL RESEARCH PLATFORM (NRP)

HOW WE EXECUTE ON THIS VISION

Flexible Architecture to build on horizontally and vertically

- Depending on effort available and control desired, you can build on NRP both vertically and horizontally at different layers of the stack.

Federate your existing batch and storage cluster with OSG/PATh

Federate your existing Kubernetes infrastructure, including commercial cloud, with NRP

OSG/PATh	
SLURM HTCondor	Admiralty
Kubernetes	
IPMI, Firmware, BIOS	
Hardware	

Join & build on the existing NRP Kubernetes infrastructure

Have the NSF funded PNRP project operate your hardware via IPMI

- NRP is a non-local extendable container deployment platform, thus allowing many uses unthinkable for a SLURM cluster in a data center.

68 Institutions integrated in NRP
43 integrated at IPMI layer
25 integrated at Kubernetes layer
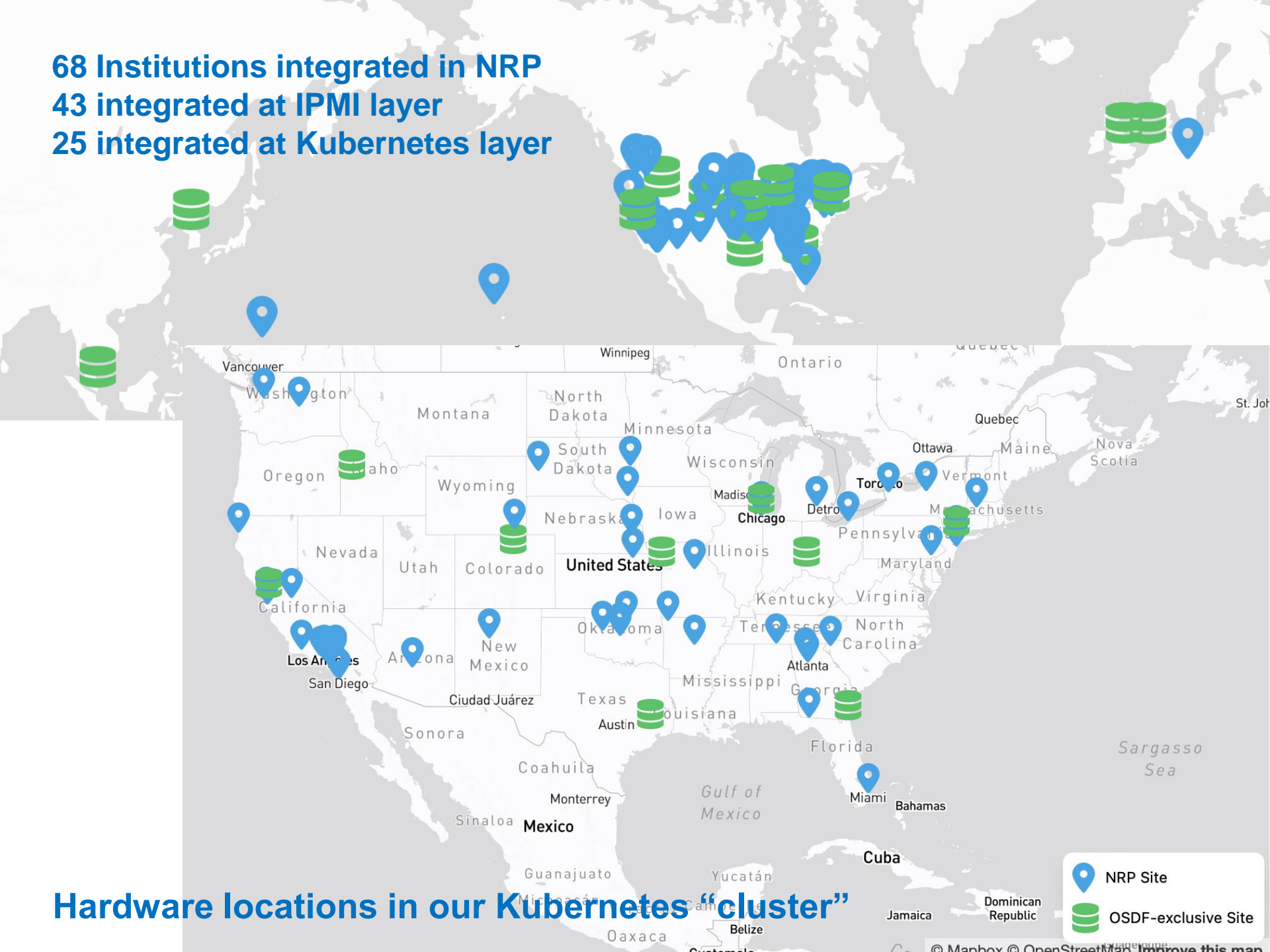
Hardware locations in our Kubernetes "cluster"

NRP Site

OSDF-exclusive Site

Combining all 3 layers, we integrate ~200 institutions across 5 continents

Open Science Data Federation (OSDF)

An example "application" that builds on NRP

Data stored on Origins
is accessed via caches

24.9 PB read in June 2024

on average:
10 Gigabytes/second

Origin
12 Sites, 10 Institutions

Cache
29 Sites, 20 Institutions

Cache and Origin
5 Sites, 5 Institutions

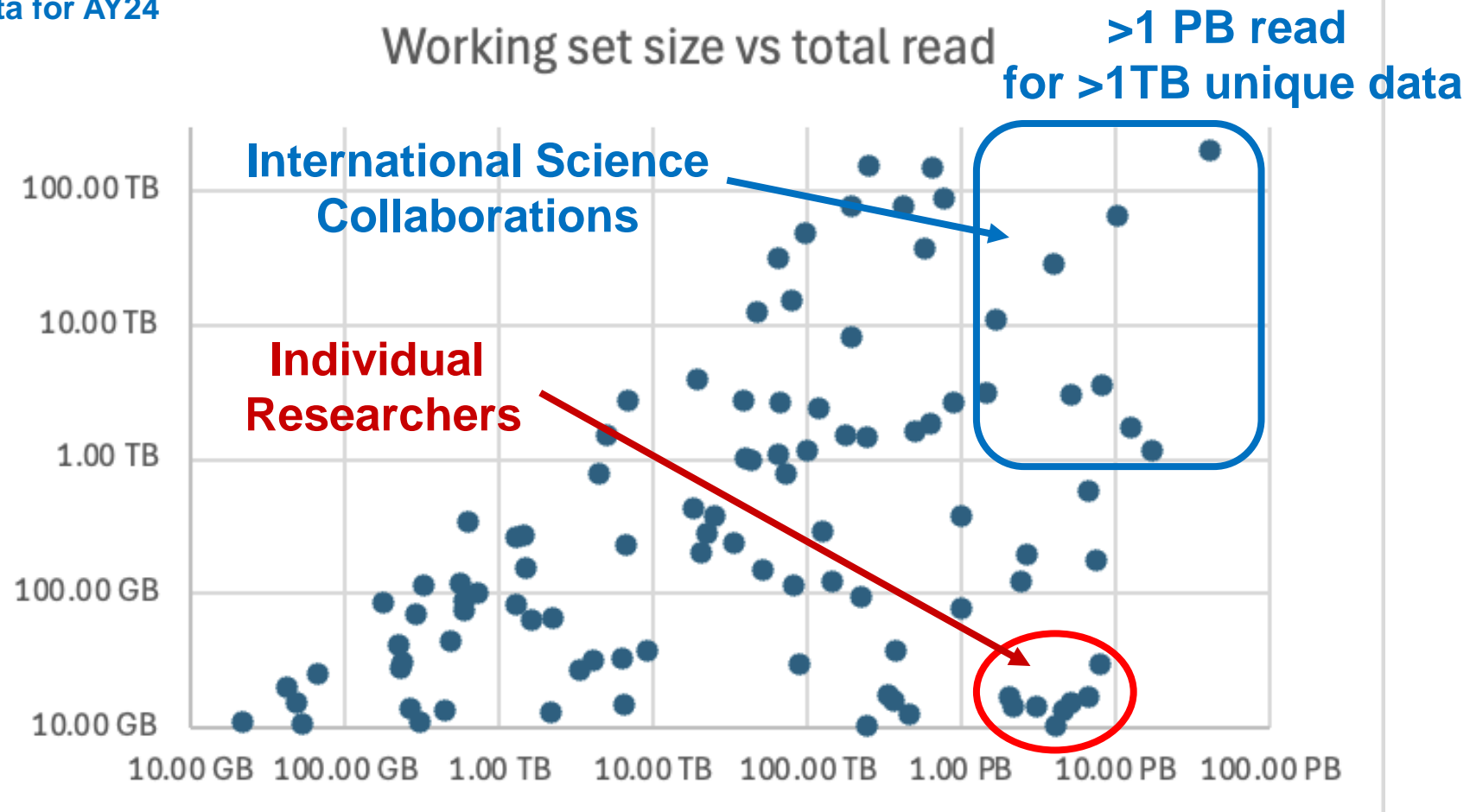Leaflet | Map data © OpenStreetMap contributors, Imagery © Mapbo

80 Gigabit per second … that's 80% of a 100G pipe
Observe <3% cache misses => OSDF caches save >75Gbps in network traffic

17 Origins and 34 caches across 5 continents

Let's look at two usage patterns

Usage Data for AY24

Working set size vs total read

>1 PB read for >1TB unique data

International Science Collaborations

Individual Researchers

>1 PB read for <50 GB unique data

Each of these patterns comprise ~1/3 of the namespaces with >1 PB read

Bridging Education & Research by having them co-exist on one platform

A lot of the smaller campuses care more about educational use than research use on NRP.

Building AI Education on NRP … Three Examples …

- [GP-Engine (NSF OAC 2322218) builds a regional compute cluster across 7 states in the Great Planes Regional Network.](#)
  - Hardware is mostly GPUs, with strong focus on supporting STEM education, especially AI.

- **UC San Diego**

- **CENIC AI Resource**
  - Collaboration between California's regional R&E network, NRP, and multiple California State Universities, and Community Colleges with strong focus on AI education
  - Includes "The California State University System Technology Infrastructure for Data Exploration (TIDE)" ([NSF OAC 2346701](#))
    - The CSU system includes 23 campuses, 21 of which are Hispanic-Serving Institutions.
      - "… most diverse higher education system in America … provides 50% of all bachelor degrees in California …"
    - " TIDE creates a pioneering computational core facility within the California State University (CSU) system, focused on ML and AI "

Jupyter & Chat

- **NRP Operates a Chat infrastructure**
  - All Educators across all institutions are signed up to a channel to share experiences.
  - Each course has its own chat, moderated by the educator(s) that teach the course
- **NRP supports Jupyter**
  - One JupyterHub for all newcomers
  - Experienced Institutions tend to clone their own, and build their own custom infrastructure on NRP
    - Join their own hardware & customize user experience for their users
- **We support interactive and batch access to GPUs**
    - Jupyter, cube-ctl (Unicorn as scheduler), HTCondor

GP-Engine
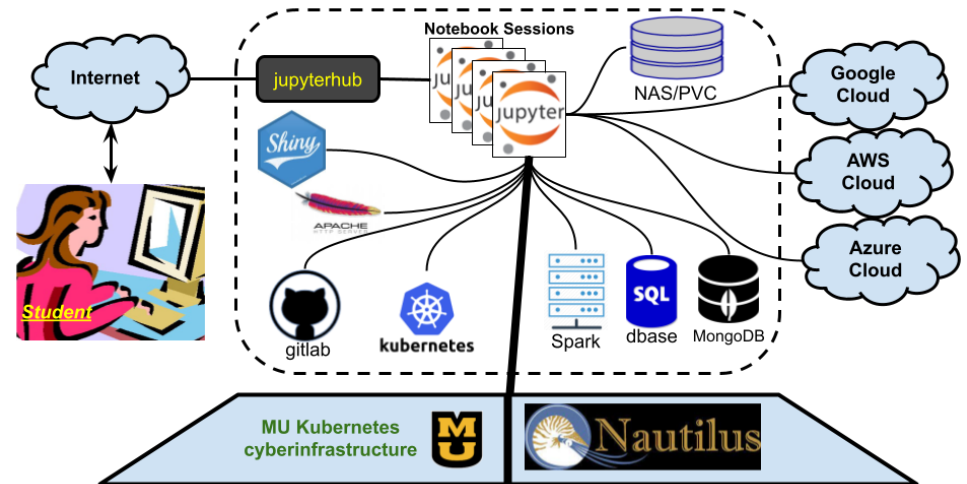
NSF OAC 2322218

- Computer Science Undergraduate Classes
- Computer Science HPC Classes
- Undergraduate Data Science
- HPC Emphasis Graduate Data Science
- State & Federal Government Training Programs

AI Education beyond STEM

- AI Filmmaking Hackathon
  - Using Generative AI to create film
  - Uses NRP to host the AI models and the inference, ie, standard diffusion

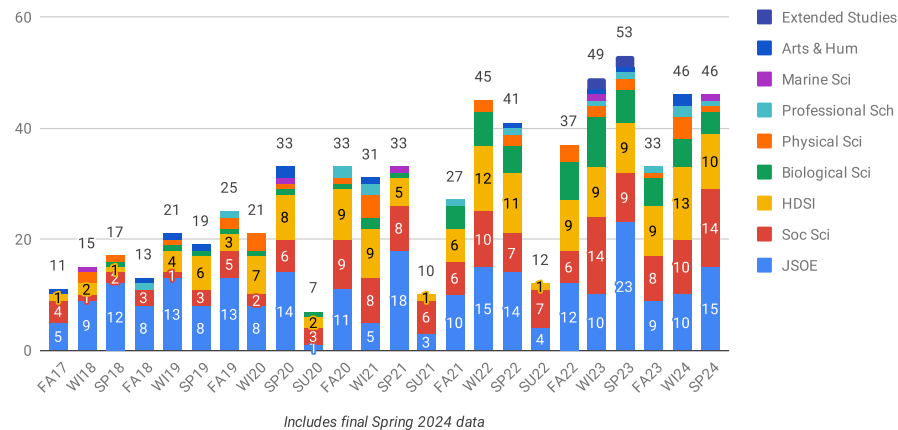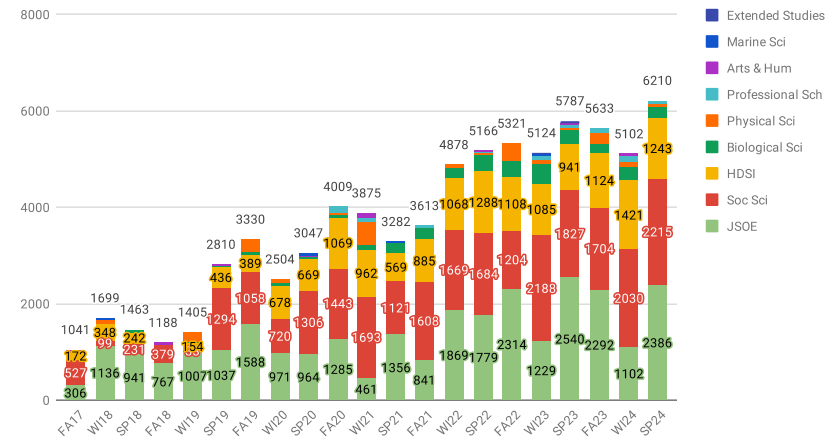**Twomey @ UNL's
Johnny Carson School of Theatre & Film**

Image generated with stable diffusion instance on the NRP

Educational Computing at UCSD

#Courses by School/Division, Quarter

Includes final Spring 2024 data

#Students by School/Division, Quarter

UCSD operates a modest size cluster (~140) 32-bit GPUs for use in the classroom

Total in AY24	9171
UG Students	7329
Grad/Prof Students	1842

~ 6,000 student enrollment per quarter
~ 50 courses across all schools
> 9,000 unique students per year

Selected UCSD Courses

- Advanced Computer Vision
- Bioinformatics for Immunologists
- Computational Physics: Probabilistic Models/Sim.
- Data Analysis/Design for Biologists
- Data Science/Spatial Analysis
- Deep Learning and Applications
- Intro to Causal Inference
- Neural Networks/Pattern Recognition
- Numerical Analysis for Multiscale Biology
- Robot Manipulation and Control

AI is accelerating a trend that already existed prior.

- CENIC serves the entire range of education in CA
  - K-12
  - Community College System (CCC)
  - California State University (CSU)
  - University of California
  - Public Libraries

CENIC AI Resource

**CENIC AIR Hosting Campuses:
CSUs & CCCs – The Next Generation**

CSU Chico

Sacramento State U

CSU San Bernardino

San Diego Community College Dist.

AI for Education via Jupyter is the "killer App"

Cal Poly Humboldt

Monterey Bay

CSU Fullerton

San Diego State U

Example: 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

UI at 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

AI/ML at San Bernardino

5 Departments teach courses on 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

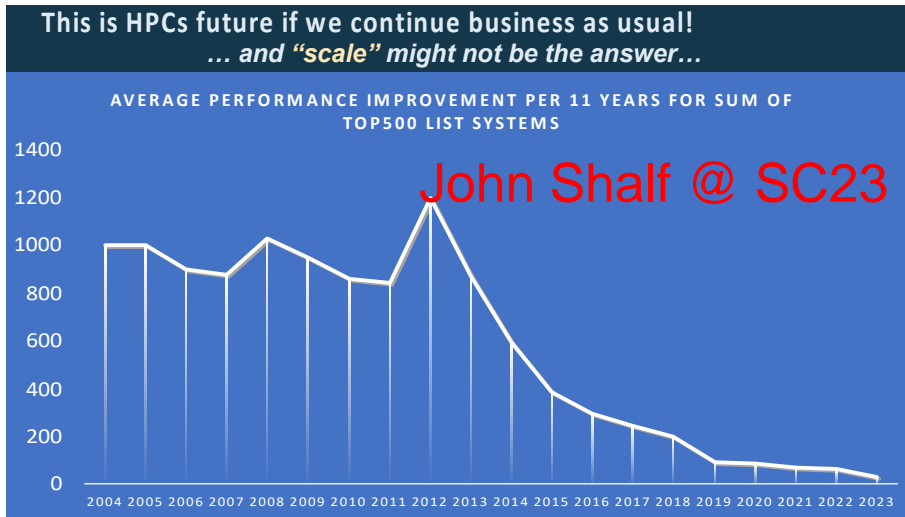MATH 3465 - Computational Statistics

14 AI/ML Projects on NRP

- Brain Tumor classification with Deep Learning
- Public Opinion Sentiment Analysis
- Medical Notes Labelling
- Cloud Removal from Satellite Imaging
- Wearable Technology for Assessing Autism
- GPT empowered personalized eLearning for programming
- ….

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

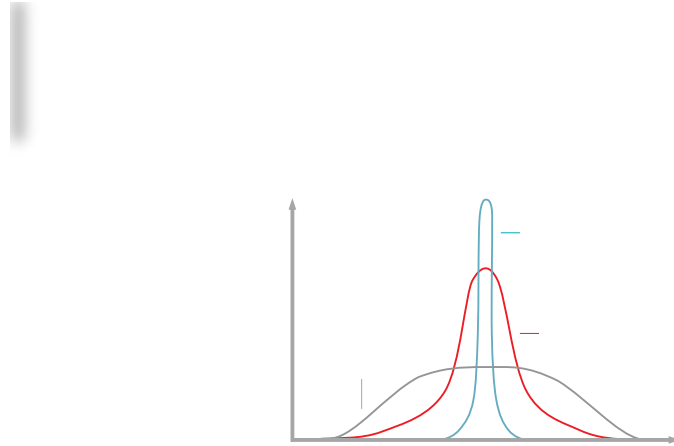NRP blurs the lines between "testbed" and "production" CI

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

"Domain Specific Architectures"

"end of Moore's law" motivates new architectures

This is HPCs future if we continue business as usual!
... and *"scale"* might not be the answer...

AVERAGE PERFORMANCE IMPROVEMENT PER 11 YEARS FOR SUM OF TOP500 LIST SYSTEMS

John Shalf @ SC23

Performance improvements vs time slowed down by O(100)

Mark Papermaster, CTO of AMD

PI, Tajana Rosing

PRISM, a Jump 2.0 project funded by SRC is early user of FPGAs@NRP

37

NRP supports FPGAs (Xilinx & Intel), P4 switches, NVIDIA DPUs & HGXs

Committed to be a "Playground" of technologies, easily deployed and operated.

Advanced Technology Laboratory on NRP

- **Programmable computational capabilities emerged in devices of all kinds**
  - Storage devices with embedded FPGAs => "Computational Storage"
  - GPUs on Network Interface Cards => "Data Flow Programming"
  - Programmable switches, down to individual ports => "Programmable Networks"

- We innovate nextGen systems in ATL to solve grand challenges of science
- Innovations made available to all of open science via our Open Infrastructure

Strategic Objective is to bring CS Research closer to Domain Research in the hope of decreasing time to adoption of new technologies & ideas
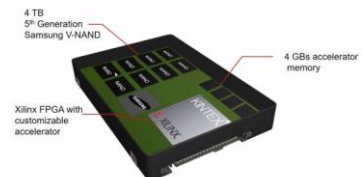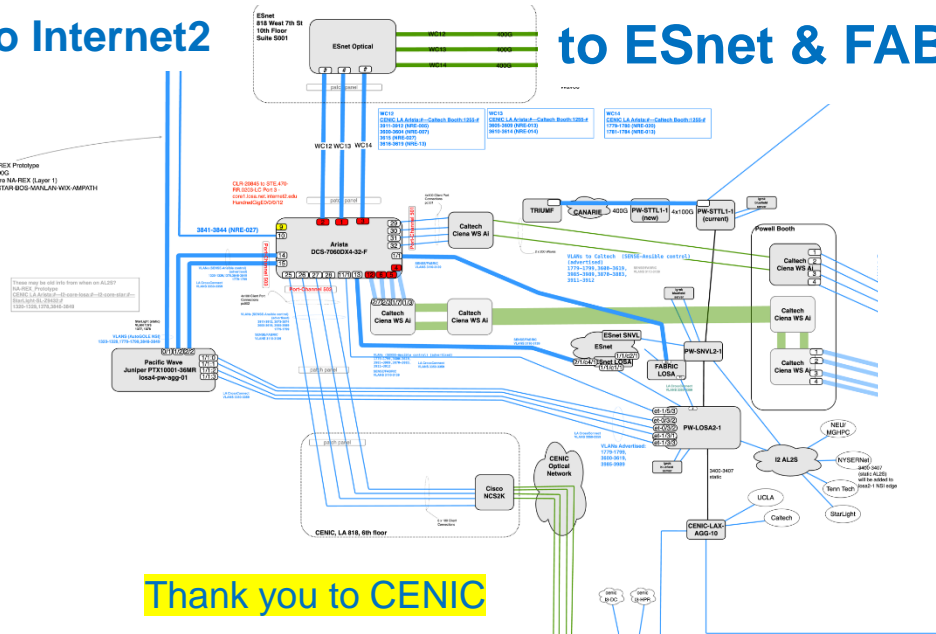
NVIDIA BlueField DPU
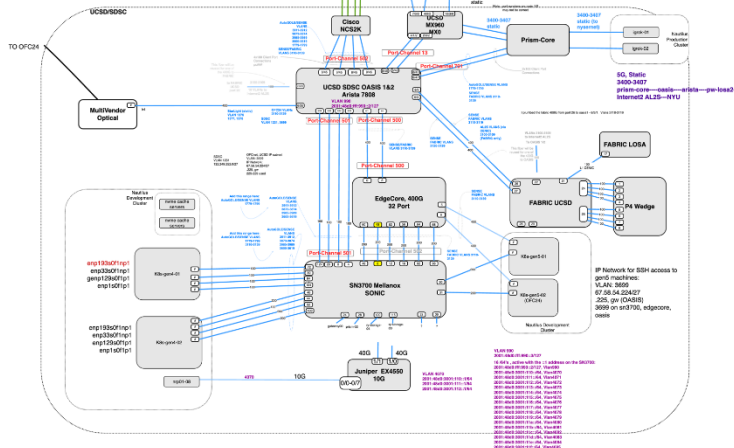
P4 programmable switches

Xilinx SmartSSD

4 TB
5th Generation
Samsung V-NAND

4 GBs accelerator memory

Xilinx FPGA with customizable accelerator

400G WAN Infrastructure

to Internet2

to ESnet & FABRIC

Infrastructure at SDSC:

- FPGAs: 32 U55C, 24 Bitware 520
- 400G P4 programmable switches
- 8 NVIDIA HGX w 8xA100 80G each
- 400TB of NVMe
- FABRIC node

Thank you to CENIC

**We peer at 400G in LA
with multiple networks
via our 400G Arista switch**

Summary & Conclusion

- **We have a very ambitious vision**
  - Horizontally open
    - Hardware across 68 institutions has joined **NRP**
    - Hardware across 145 institutions is available via **OSG**
  - Vertically Open
    - We have built the "Open Science Data Federation" on top of NRP,
    - … and are starting to build "Fusion Data Platform for AI" on top of NRP
    - … and are starting to build elements of the National Discovery Cloud for Climate on top of NRP (Pelican, National Data Platform, NCAR integration, …)
  - "Playground" for CS R&D on the same platform as "Production" system for Domain Scientists
- **Education increasingly requires significant CI**
- **NRP is owned & built by the community for the community**
  - We are always looking for new partners!

Acknowledgements