NRP NATIONAL RESEARCH PLATFORM

UCSD

Status of the National Research Platform (NRP)

Frank Würthwein
Director, San Diego Supercomputer Center

March 20th 2024

Challenges NRP addresses

- The gap between those who have and those who can't afford is becoming wider.
- Cyberinfrastructure needs are growing for Education
- The end of Moore's Law is leading to a proliferation of "architectures" … domain science adoption is at risk.

NRP NATIONAL RESEARCH PLATFORM

UCSD

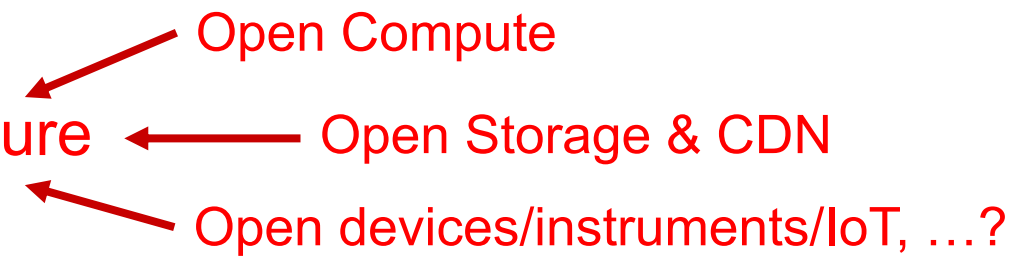
Vision of an Open Infrastructure

Horizontally open => institutions can integrate their resources

Vertically open => projects can build on the infrastructure

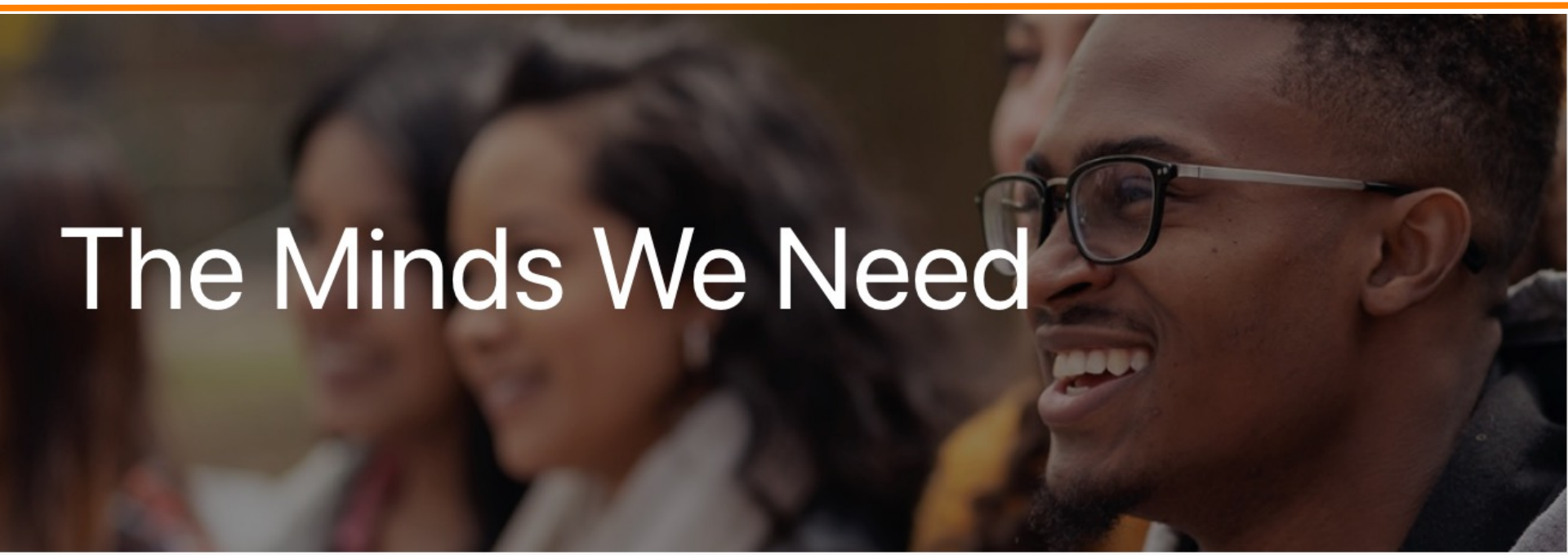
Long Term Vision

- 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



Openness for an Open Society

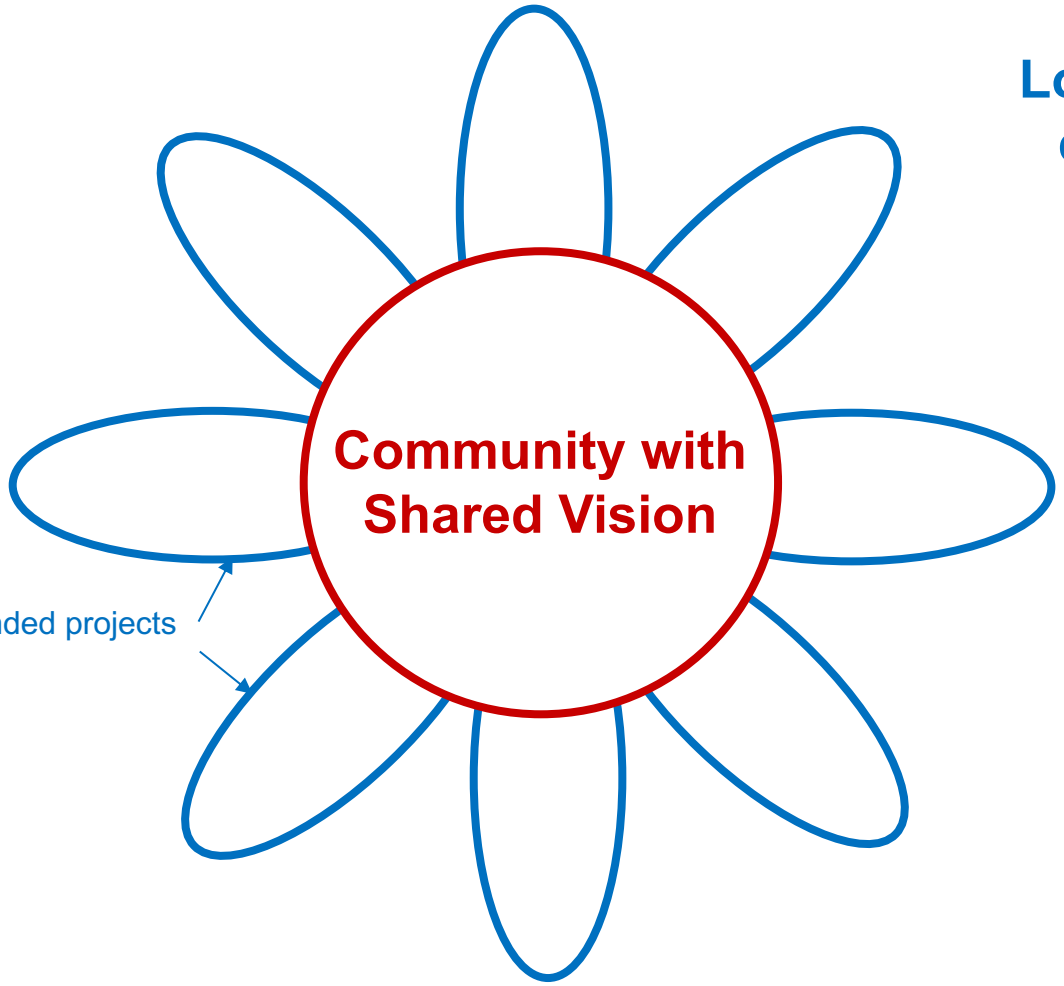
Democratize Access



- **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

Community vs Funded Projects



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

Hardware funded by NSF, DOD, DOE, …

Petals of this flower include:

- 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

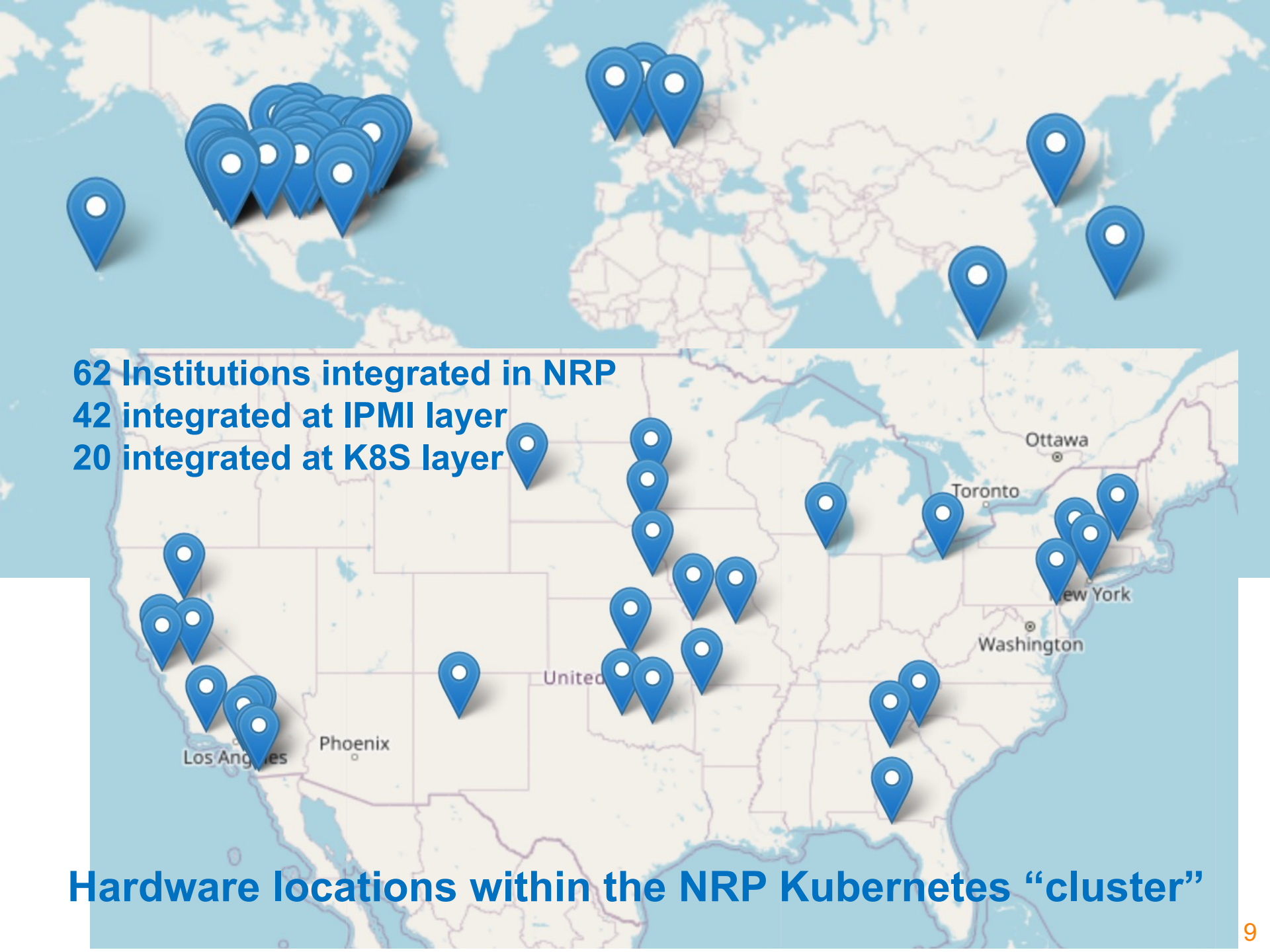
NATIONAL RESEARCH PLATFORM (NRP)

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.

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

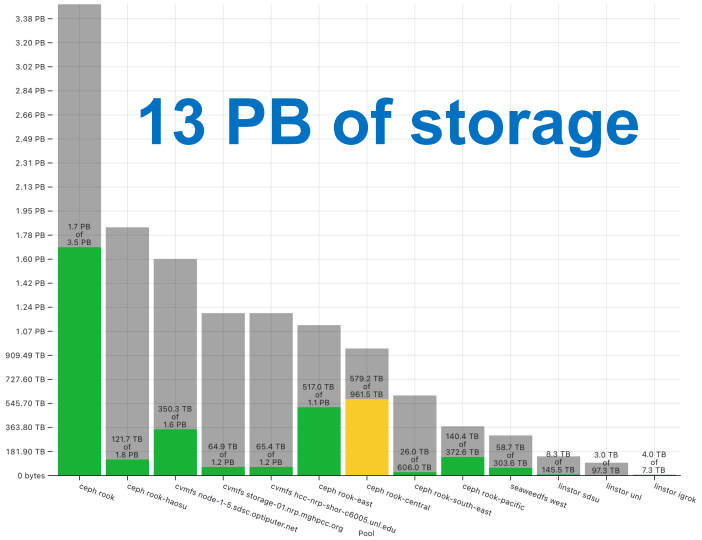


Hardware locations within the NRP Kubernetes “cluster”

NRP as an AI platform

Type of Accelerator	# of that type in NRP
A100 80G	224
A100 40G	2
A10	288
A40	10
A6000	48
A5000	6
A4000	32

Type of Accelerator	# of that type in NRP
4090	10
3090	247
2080Ti	152
1080Ti	139
Xilinx U55C	32
Other GPUs	104
Total	**1262**

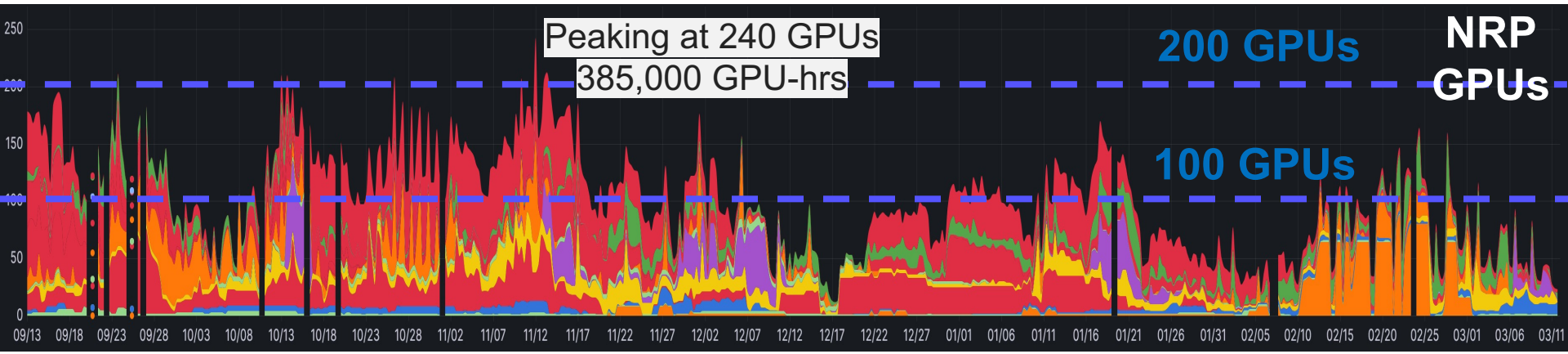


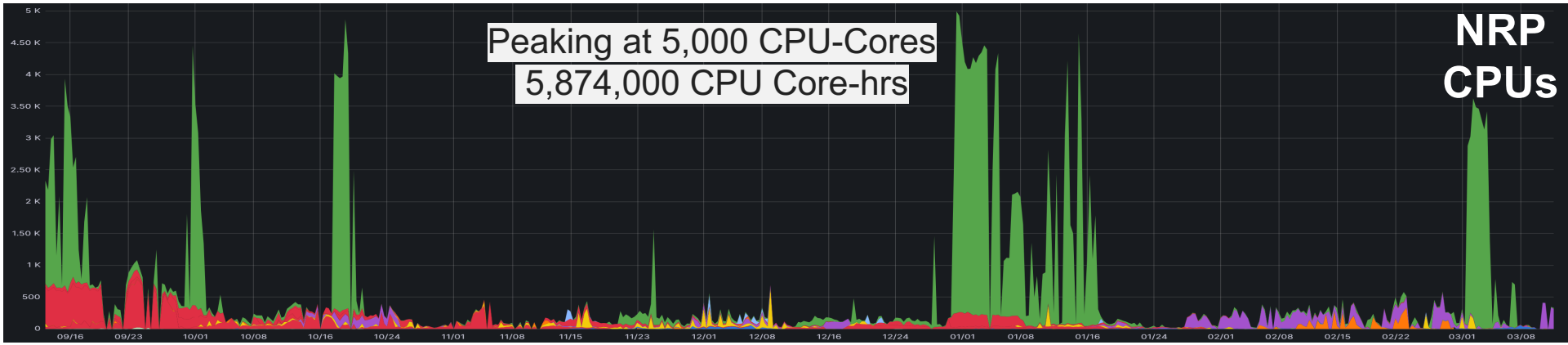
In addition, there are 21,070 x86 CPU cores, 7,406 are in nodes that have no accelerators.

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

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

Use of NRP by the 8 science speakers





The eight 5NRP science speakers presenting today & tomorrow used up to 240 GPUs and 5,000 CPU-Cores daily over the last 6 months

More details on Science@NRP in Larry Smarr's talk on Thursday.

Open Science Data Federation

an "application" deployed on NRP

143 Institutions joining at the batch or storage system layer

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

NRP NATIONAL RESEARCH PLATFORM

UCSD

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.

We made education on NRP a special focus of 5NRP with sessions on Thursday and Friday

Building AI Education on NRP … Two 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.

More details on Friday
Alex Hurt & Derek Weitzel

- **CENIC AI Resource**
  - Collaboration between California's regional R&E network, NRP, and multiple California State Universities 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 "

More details on Friday
Chris Bruton & Jerry Sheehan

Select UCSD Courses requiring CI Spring 2023

UCSD operates a modest size cluster (~140) 32-bit GPUs for use in the classroom
6,502 undergraduates & 1,847 graduate students used this in AY22/23

- 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

More details on Thursday
Michael Farley & Valerie Polichar

AI is accelerating a trend that already existed prior.

1/3 of use is easily identifiable AI/ML containers

Training events yesterday

We had a registration limit of 20 people

Significant fraction of the audience crashed the party.

Clearly, there is much more demand than we prepared for.

We will prepare better next year.

NRP NATIONAL RESEARCH PLATFORM

UCSD

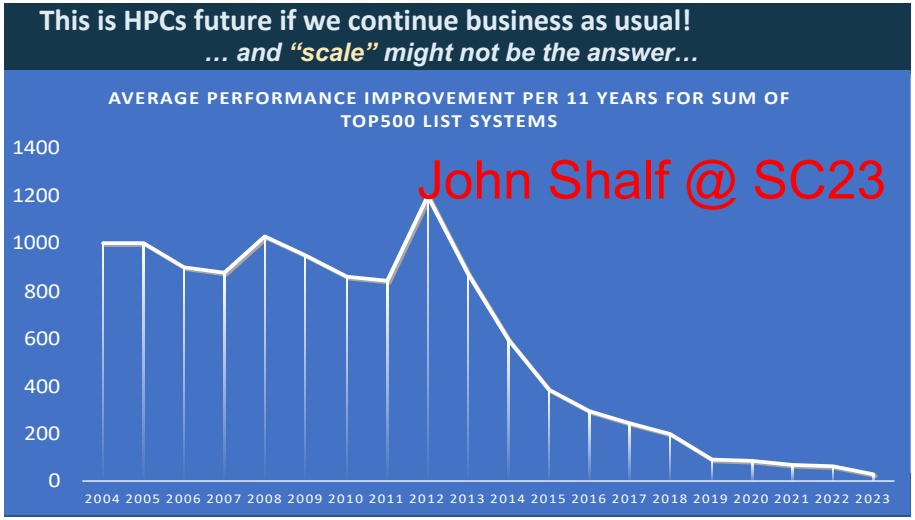
NRP brings CS R&D and Domain R&D onto the same platform

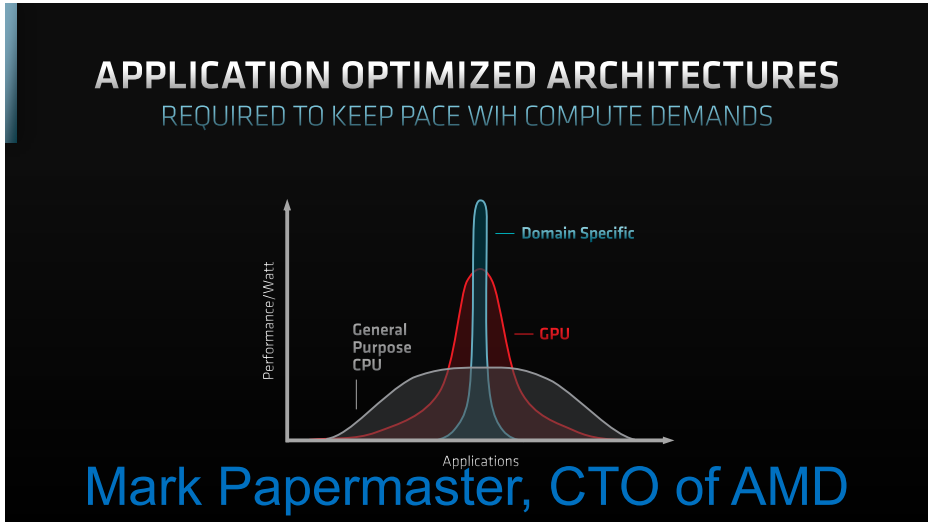
NRP blurs the lines between ”testbed” and “production” CI

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

"Domain Specific Architectures"

"end of Moore's law" motivates new architectures





Performance improvements vs time slowed down by O(100)

PI, Tajana Rosing

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

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

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

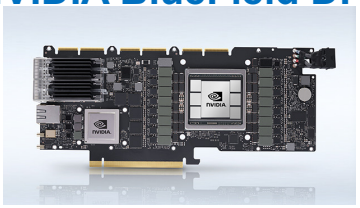
Advanced Technology Laboratory on NRP

- **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

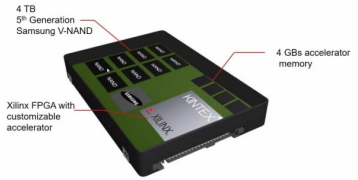
Strategic Objective is to bring CS Research closer to Domain Research in the hope of decreasing time to adoption of new technologies & ideas

NVIDIA BlueField DPU

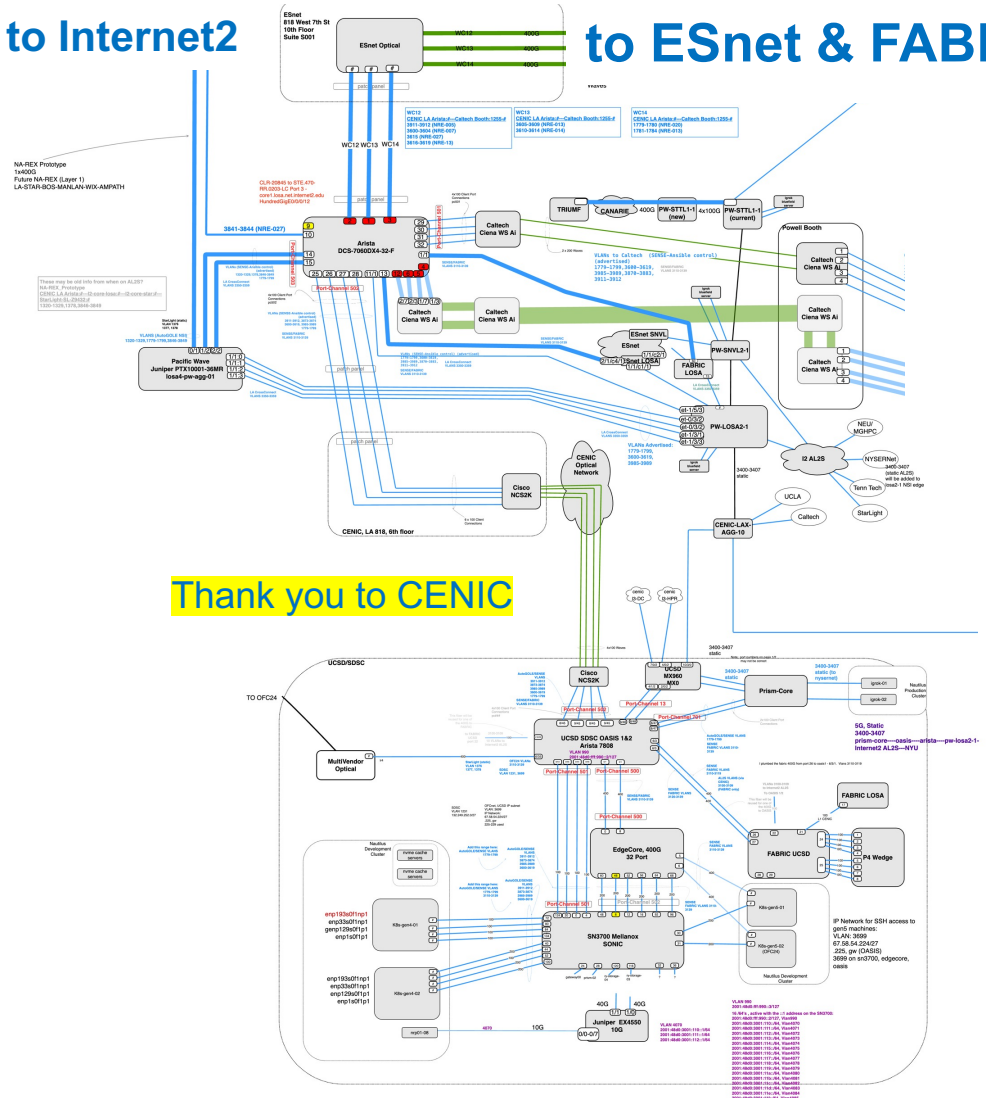


P4 programmable switches

Xilinx SmartSSD



400G WAN Infrastructure



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

Would like to offer this Infrastructure to FABRIC community via FABRIC-NRP co-allocation

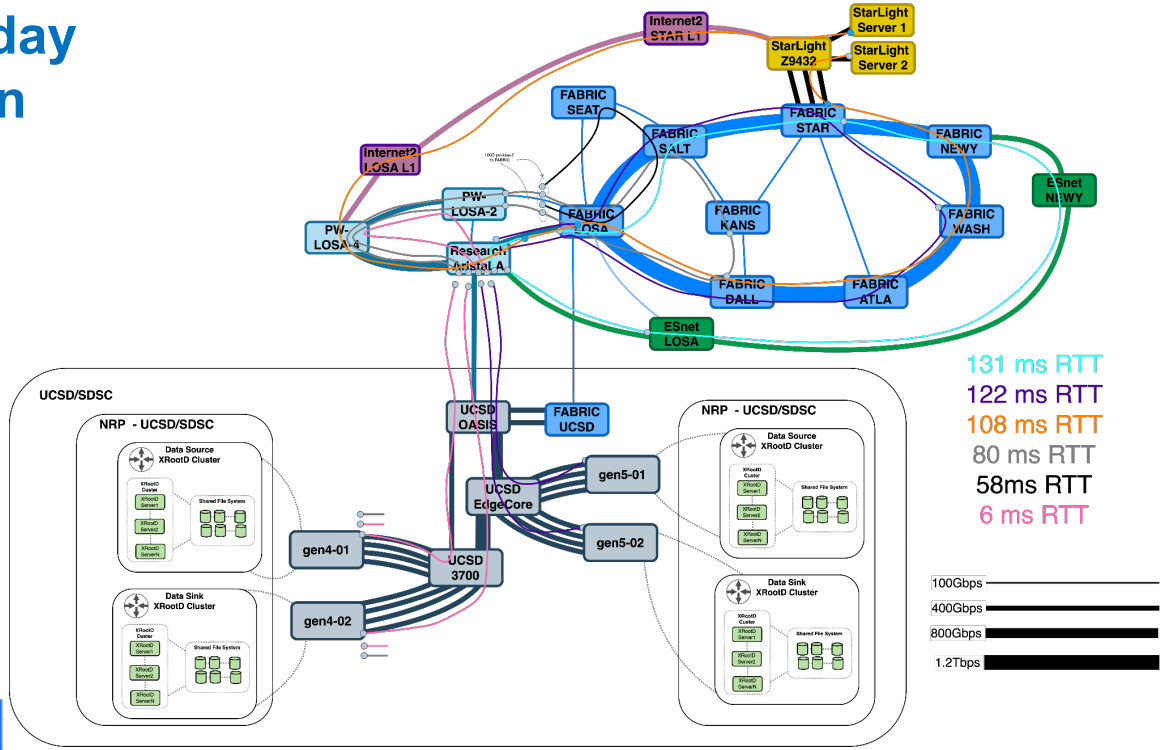
We peer at 400G in LA with multiple networks via our 400G Arista switch

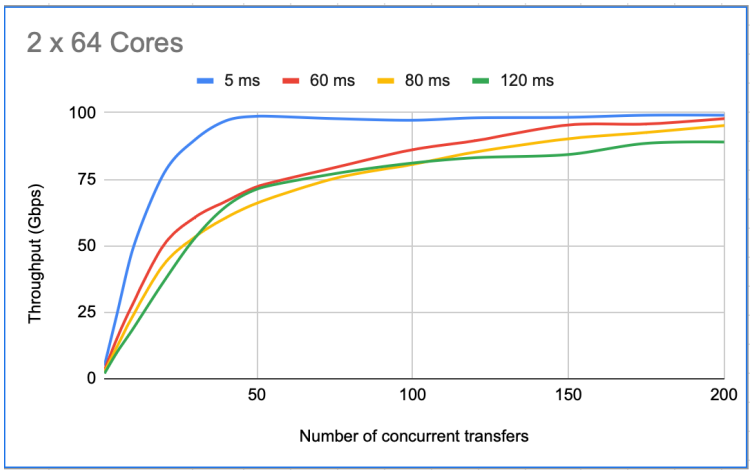
Example FABRIC-NRP Integration Exercise

See more details in Wednesday KNIT8 talk by Tom Lehman

Create various RTT loops across FABRIC, NRP, ESnet, CENIC, StarLight, Internet2

All begin & end on NRP@SDSC





Started using this testbed for measurements yesterday.

Summary & Conclusion

- **NRP has a very ambitious vision**
  - Horizontally open
    - Today about 3x # of GPUs total than what was part of Cat-II PNRP award
      - PNRP award started testbed operations phase on 3/27/23
  - 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**
- We are off to an excellent start … but there is lot’s more to come over the course of the next 5 years.

Acknowledgements

- This work was partially supported by the NSF grants OAC-1541349, OAC-1826967, OAC-2030508, OAC-1841530, OAC-2005369, OAC-21121167, CISE-1713149, CISE-2100237, CISE-2120019, OAC-2112167