National Research Platform

Frank Würthwein
Director, San Diego Supercomputer Center

November 13th 2023

NRP is many things …

- **Open Infrastructure for an Open Society**
  - **Built by the Community for the Community**
- Infrastructure for Education
  - More details on Wednesday at 3:30pm in SDSC booth.
- **Testbed that brings Computer Science R&D and Domain Science R&D onto the same platform.**
  - Accelerate translation of innovation into practice

NRP NATIONAL RESEARCH PLATFORM

UCSD

Towards an Open Infrastructure

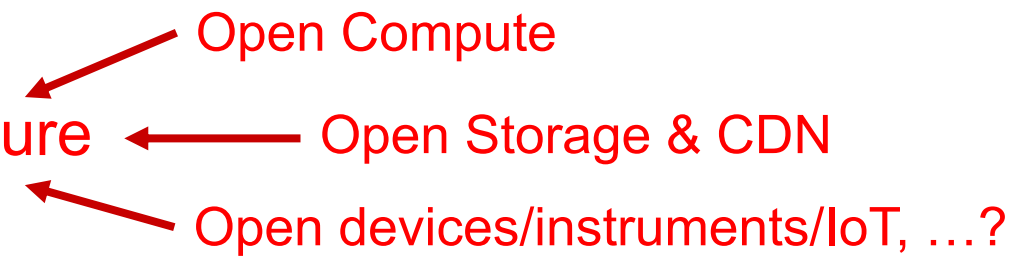
Horizontally open => institutions can integrate their resources

Vertically open => projects can build on the infrastructure

Long Term Vision

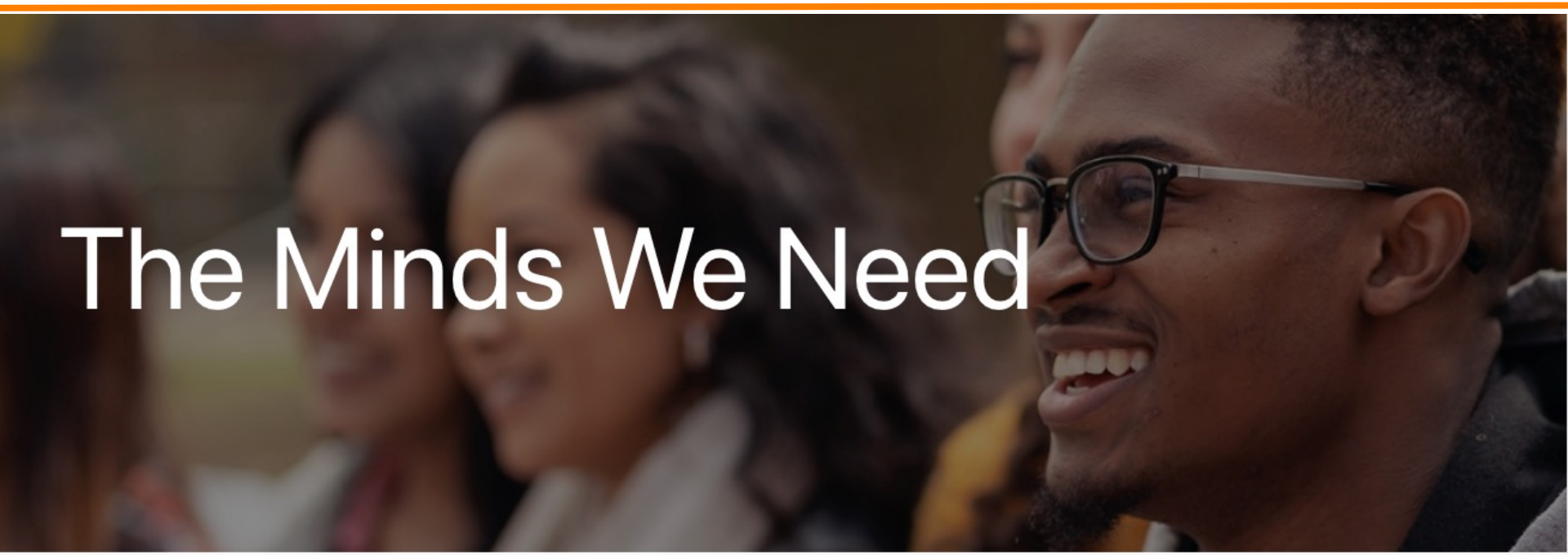
Idea

- 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

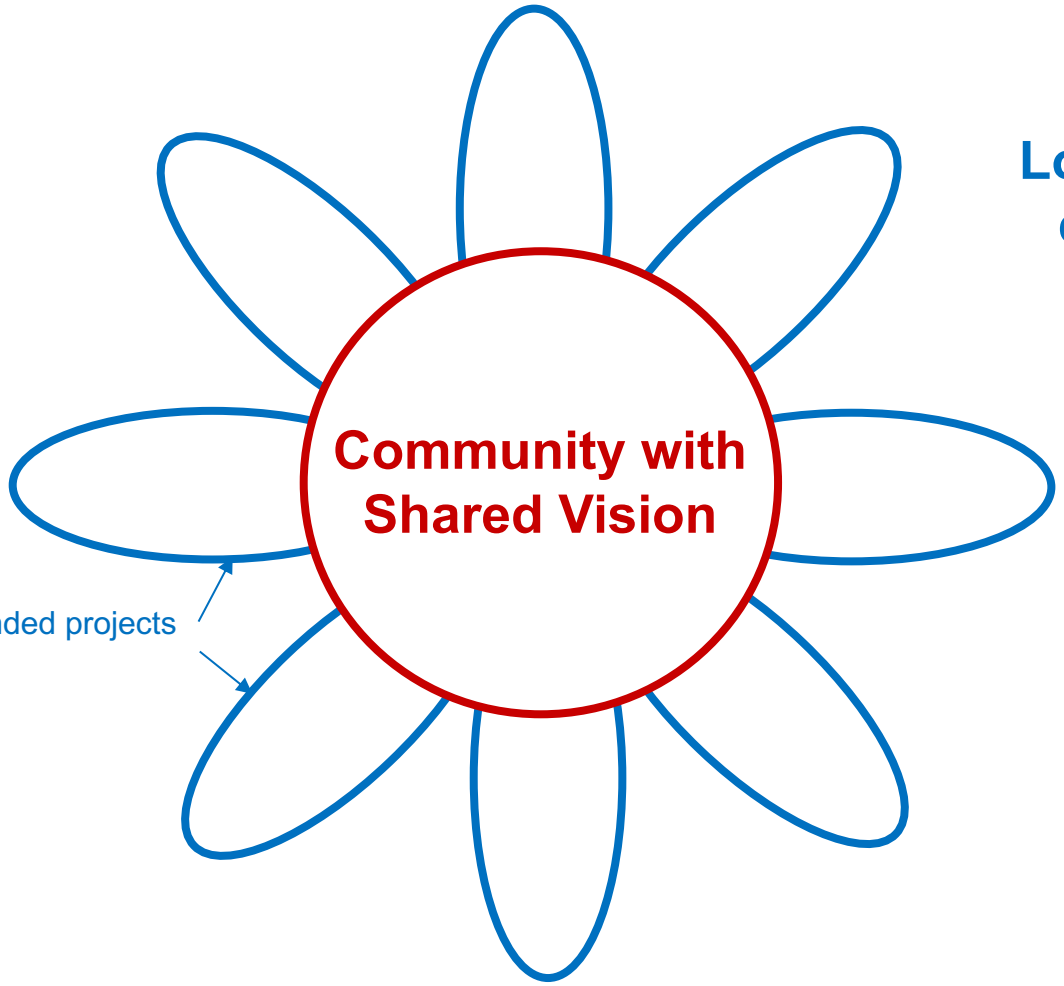
Democratize Access



- **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, …

Services building on top include:
Open Science Data Federation
Open Science Compute Federation
National Data Platform
Fusion Data Platform
R&E platforms for multiple campuses

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

NATIONAL RESEARCH PLATFORM (NRP)

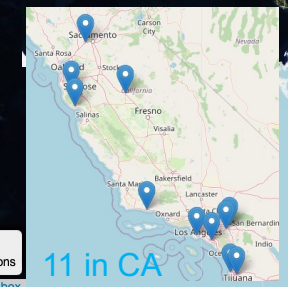
OUR ATTEMPT TO EXECUTE ON THIS VISION

Flexible Architecture to build on horizontally and vertically

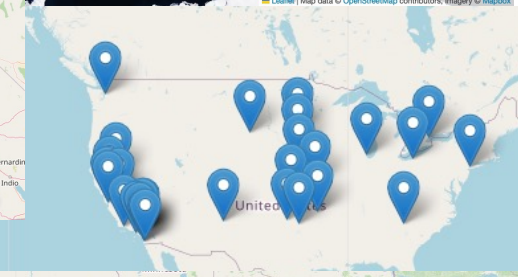
- Depending on effort available and control desired, you can build on NRP both vertically and horizontally at different layers of the stack.

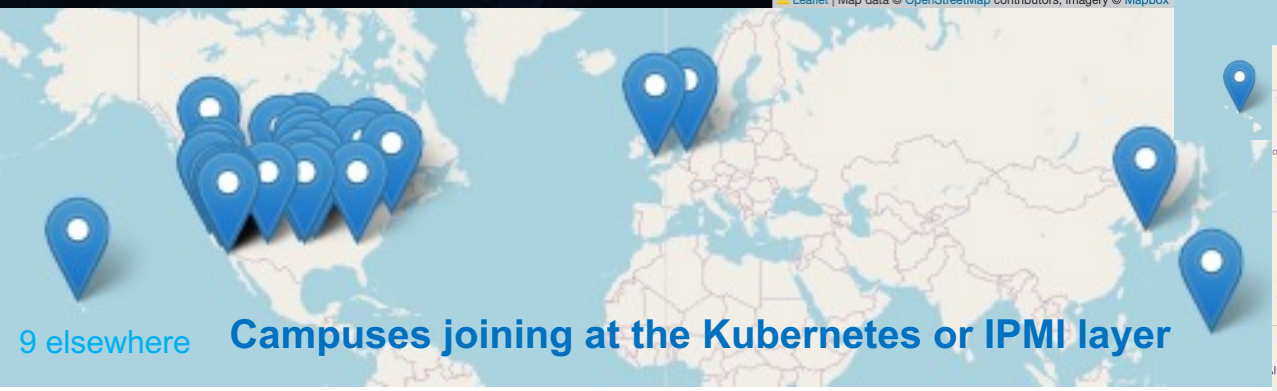
Campuses joining at the batch or storage layer

Campuses joining at the storage layer

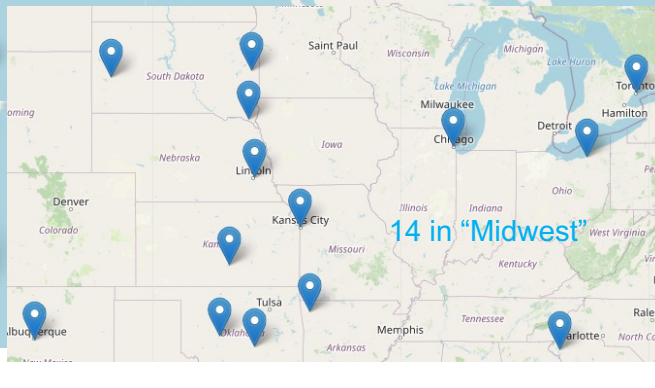


11 in CA




9 elsewhere

Campuses joining at the Kubernetes or IPMI layer



14 in “Midwest”

More than 150 Institutions across 5 continents are presently integrating resources this way

NRP NATIONAL RESEARCH PLATFORM

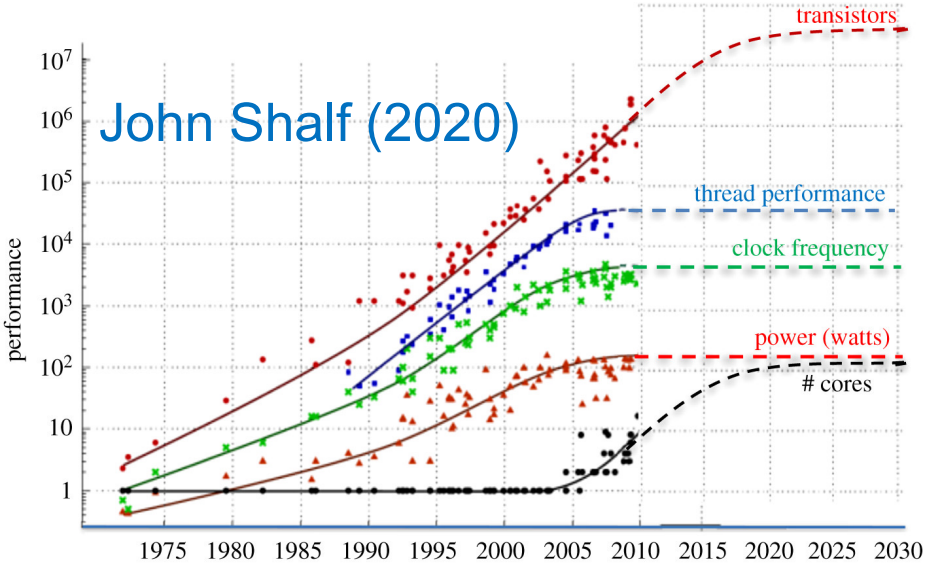
UCSD

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

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

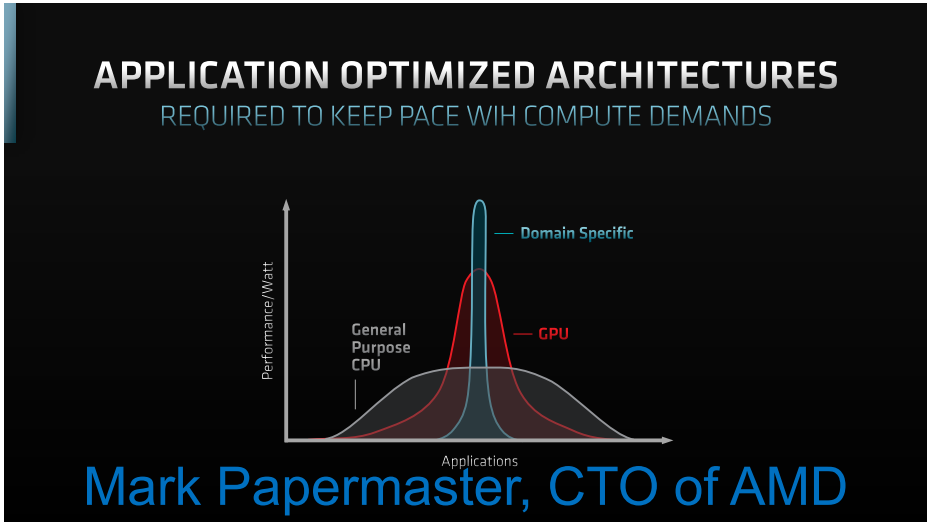
"Domain Specific Architectures"

"end of Moore's law" motivates new architectures



John Shalf (2020)

https://doi.org/10.1098/rsta.2019.0061



Mark Papermaster, CTO of AMD

R&D

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”

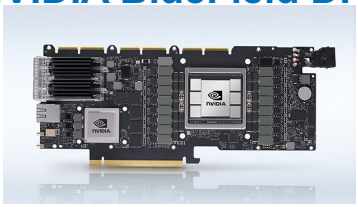
- We innovate nextGen systems in ATL to solve grand challenges of science
- Innovations made available to all of open science via our Open Infrastructure

Idea

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

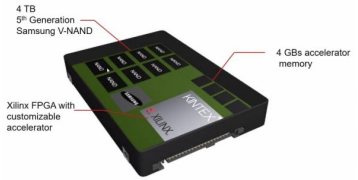
R&D

NVIDIA BlueField DPU



P4 programmable switches

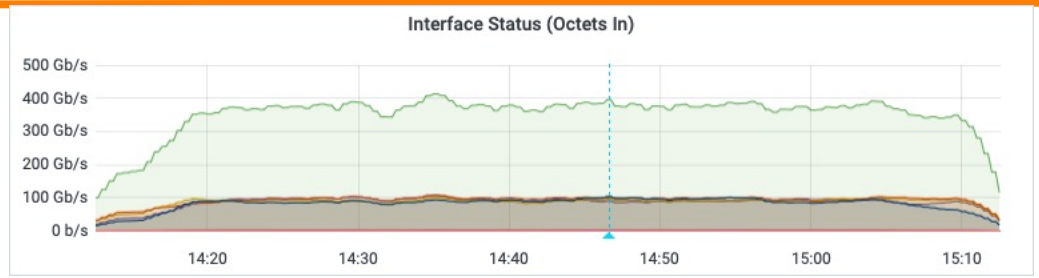
Xilinx SmartSSD



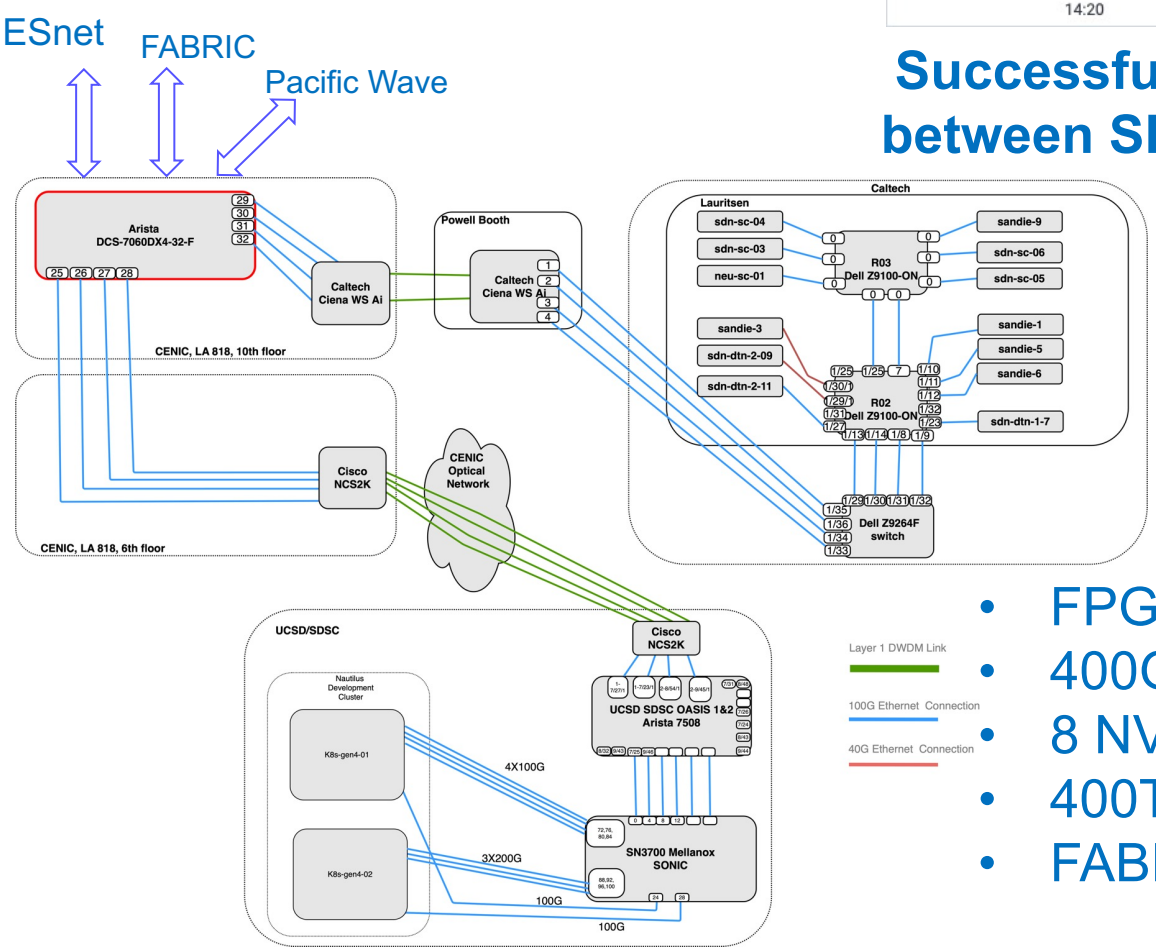
400G WAN Infrastructure

Thank you to CENIC

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



Successfully sustained 400G transfers between SDSC & Caltech using XRootD



Infrastructure at SDSC:

- FPGAs: 32 U55C, 24 Bitware 520
- 400G P4 programmable switches
- 8 NVIDIA HGX w 8 80G A100s each
- 400TB of NVMe
- FABRIC node

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
      - More details on Fusion on Wednesday at 10:30am in SDSC booth
    - … and are starting to build elements of the National Discovery Cloud for Climate on top of NRP (Pelican, National Data Platform, NCAR integration, …)
      - More details on Wednesday 3:30pm SDSC booth
  - “Playground” for CS R&D on the same platform as “Production” system for Domain Scientists
- 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