

Open Cyberinfrastructure for an Open Society

Frank Würthwein
Director, San Diego Supercomputer Center

December 20th 2023



- The gap between those who have and those who can't afford is becoming wider.
- How do we bridge the gap between classroom and research scale of any concept we teach?
- The end of Moore's Law is leading to a proliferation of "architectures" ... domain science adoption is at risk.

- I will switch back and forth between
 - Vision & Ideas Idea
 - R&D R&D
 - Existing “production” systems Production

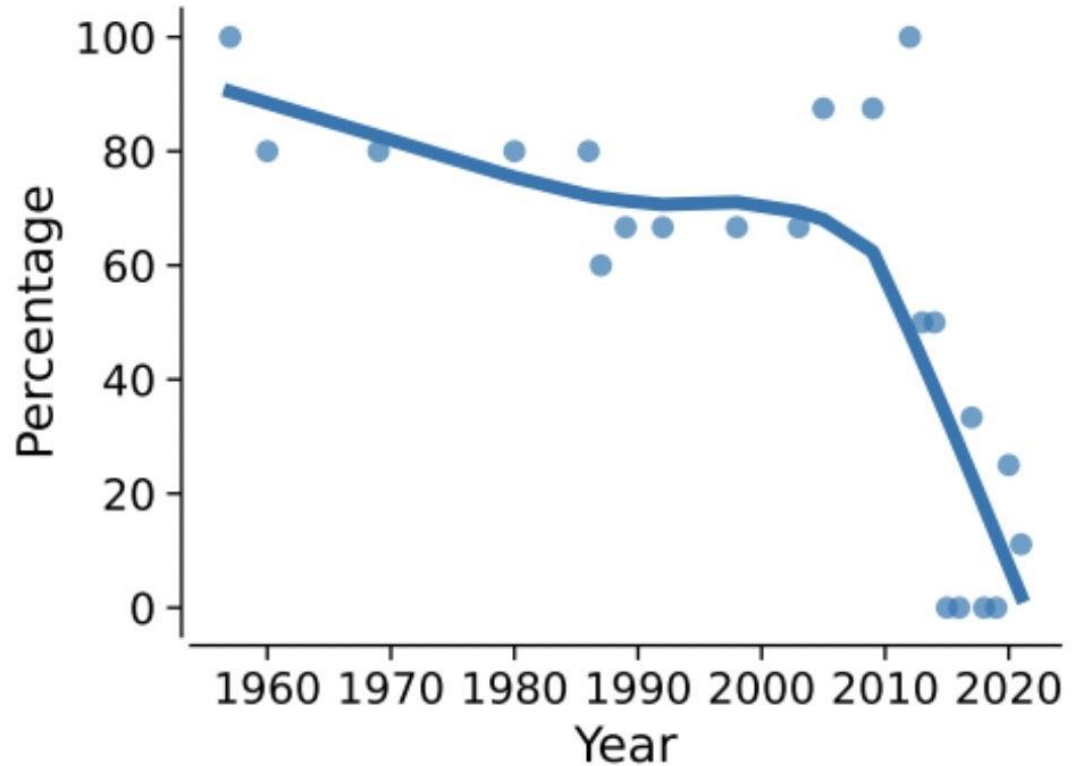
To minimize confusion, I’ll tag my slides with these labels as much as possible

**We will look at these challenges using AI as a
lense but the challenges are far more
general than that**

% of large-scale AI
results from academia

From data in

<https://arxiv.org/abs/2202.05924>



**Academia is no longer competitive
for large-scale AI model training.**

GPT Assistant training pipeline



Towards an Open Infrastructure

Horizontally open => institutions can integrate their resources

Vertically open => projects can build on the infrastructure

Production

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

3900 accredited institutions of higher learning

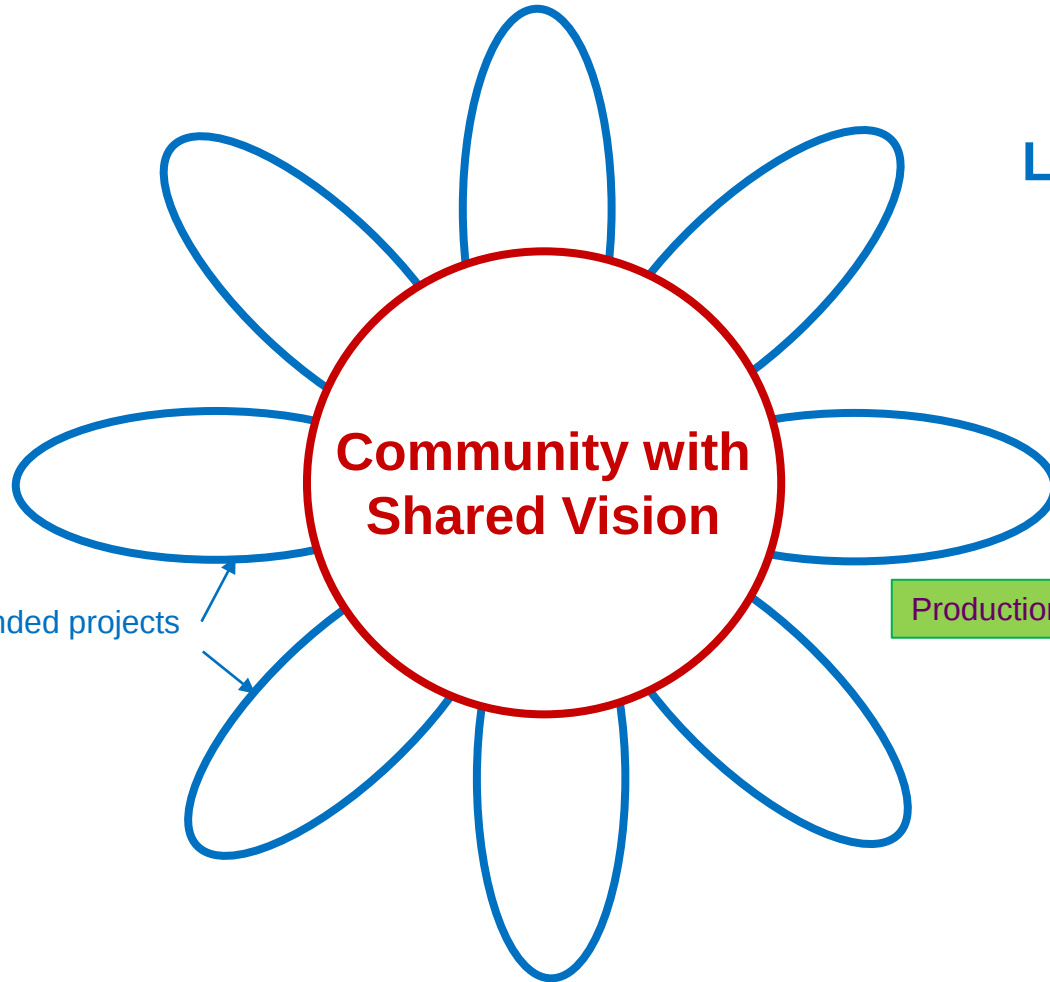


**They come in all shapes and sizes
from a few hundred to a few tens of thousands of students**



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



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

R&D

Production

Open Infrastructure is “owned” and “built” by the community for the community

- **Majority of colleges in USA are too small to afford a data center.**
 - A good number are still struggling with broadband connectivity
 - Even if they invested the funds, they don't have the human infrastructure to operate an HPC cluster and distributed storage system to even support AI education, leave alone AI research.
- **Most colleges that have a data center, aren't prepared to host modern hardware in that data center.**
 - All modern HPC or AI racks require liquid cooling.
 - Even when vendors don't support it (e.g. NVIDIA HGX)
 - Power density has gone from 20-30kW/rack to 50-150kW/rack within the last few years.
- **Even if they have power and cooling, the price per rack has exploded**
 - Expanse HPC system @ SDSC was ~\$10M for 13 racks in 2020, or ~\$0.75M per rack
 - Today, we are looking at ~\$4M per rack for high end AI equipment

We need R&D focus on reducing TCO of CI for AI

NATIONAL RESEARCH PLATFORM (NRP)

OUR ATTEMPT TO EXECUTE ON THIS VISION

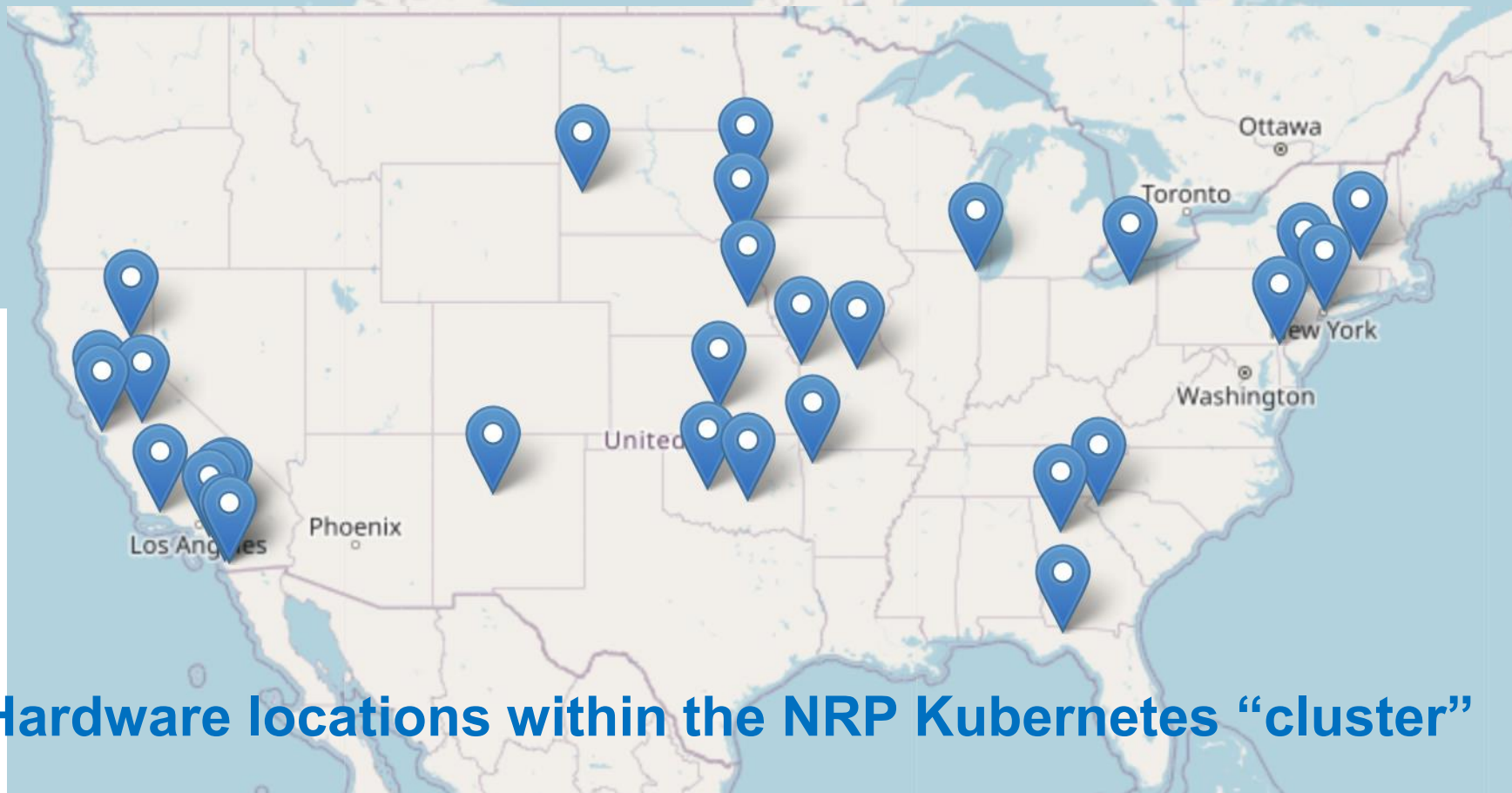
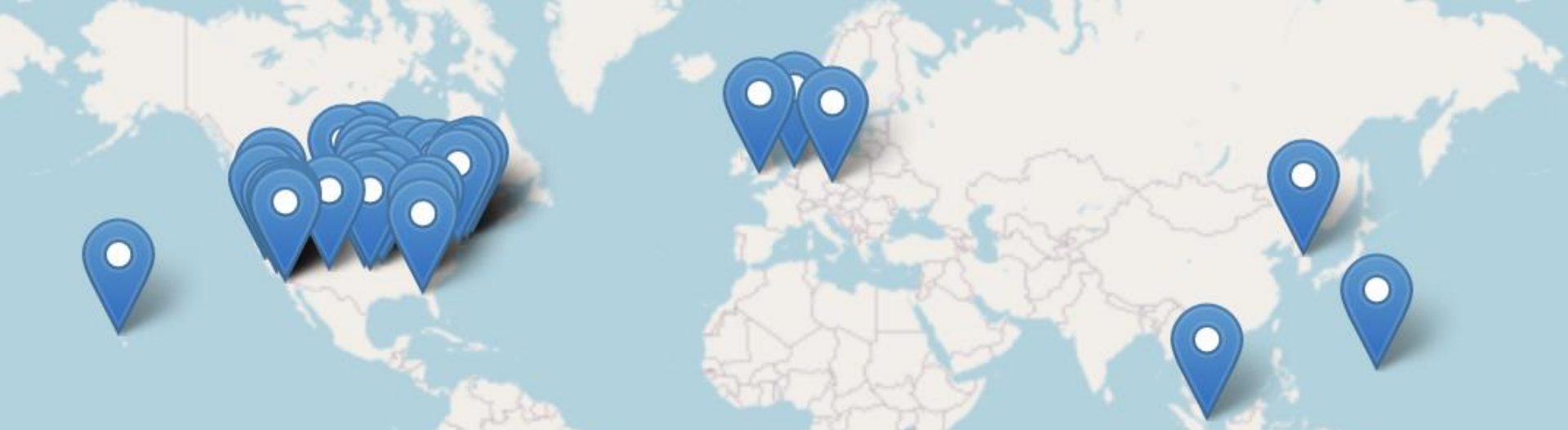
- Depending on effort available and control desired, you can build on NRP both vertically and horizontally at different layers of the stack.



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

Data is stored at origins
and accessed via caches





Hardware locations within the NRP Kubernetes “cluster”

Institutions joining at the batch or storage system layer



Combining both layers, we integrate >150 institutions across 5 continents

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

A lot of the smaller colleges care more about educational
use than research use.

All colleges struggle with applying classroom concepts to
research scale problems.

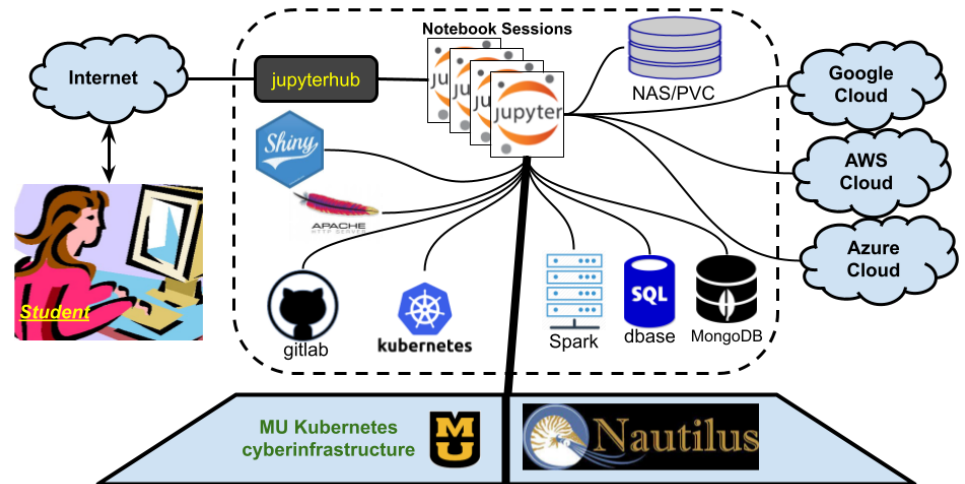
- University of Oklahoma Libraries presented at CNI 2022 how they built a R&E platform on NRP
- GP-Engine (NSF OAC 2322218) builds a regional compute cluster across 7 states in the Great Planes Regional Network.
 - Hardware is mostly GPUs, with strong focus on supporting STEM education, especially AI.
- **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 "

Grant Scott

<https://scottgs.mufaculty.umsystem.edu/>

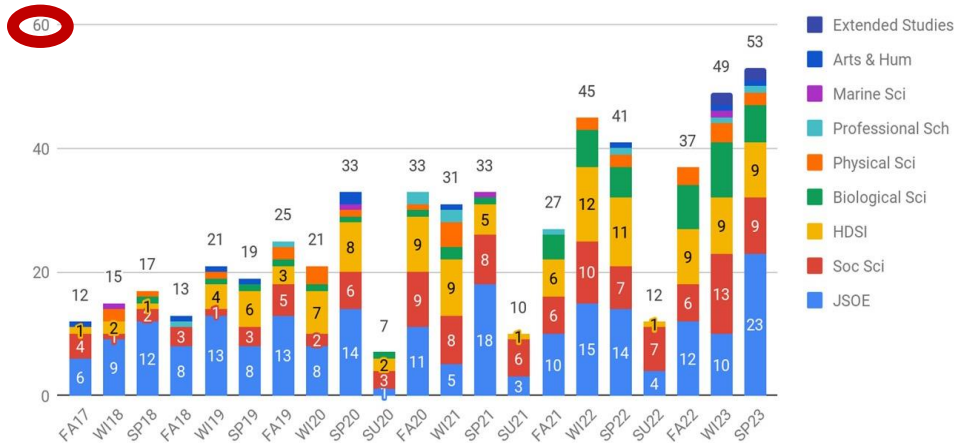


NSF OAC 2322218

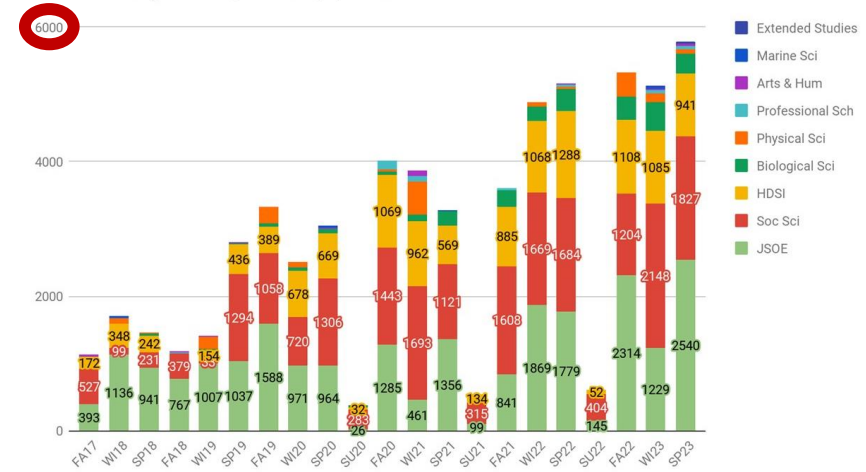


- Computer Science Undergraduate Classes
- Computer Science HPC Classes
- Undergraduate Data Science
- HPC Emphasis Graduate Data Science
- State & Federal Government Training Programs

#Courses by School/Division, Quarter



#Students by School/Division, Quarter



Source: Adam Tilghman, UCSD ITS
Data Science/ML Platform - 5/2023 Update

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

During FY23, this was used by 6,502 undergraduates & 1847 graduate students

33% of 1.5M pod-hours per year are used by AI/ML containers

In Spring Quarter 2023, 53 courses across 8 schools were taught on this system

- 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.

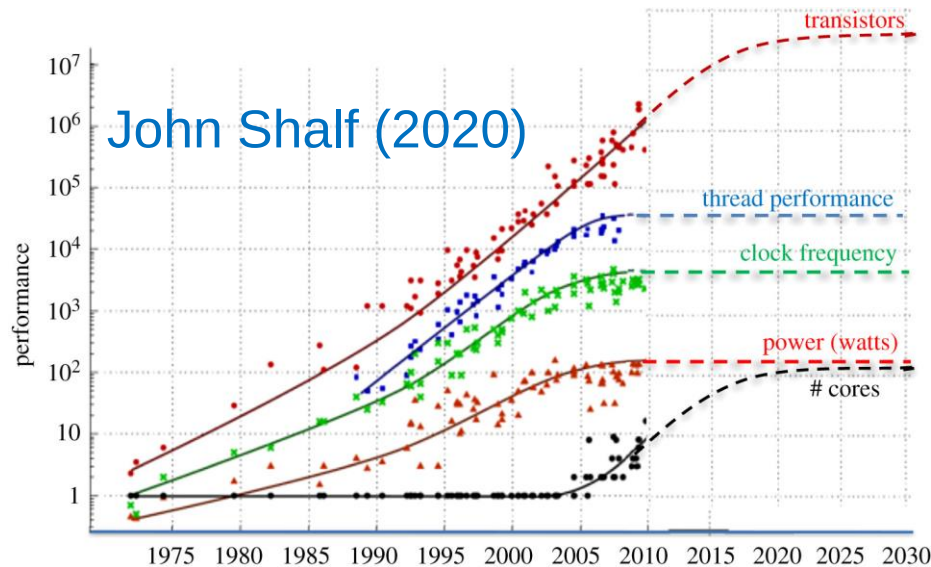
- SDSC did a keyword search for AI,ML across project descriptions on our systems to find AI projects active in 2023 with PIs from UC San Diego.
- We find >60 “faculty”, across **23 departments engaged in AI/ML @ SDSC.**
- The departments with the largest number of PIs engaged in AI/ML@SDSC are Computer Science (9), SDSC (7), Oceanography (6), Data Science (6), Cellular & Molecular Medicine (5), and Radiation Medicine (4).
- In terms of Schools, **Engineering (6 departments), Health Sciences (7 units), Physical Sciences (3 departments), Social Sciences, Biological Sciences, Pharmacy, and Oceanography** are engaged in AI@SDSC.
- We find that **Engineering and Health Sciences account for roughly ¼ of the AI/ML researchers each.**

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

“end of Moore’s law” motivates new architectures



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

R&D

PI, Tajana Rosing

Mark Papermaster, CTO of AMD

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

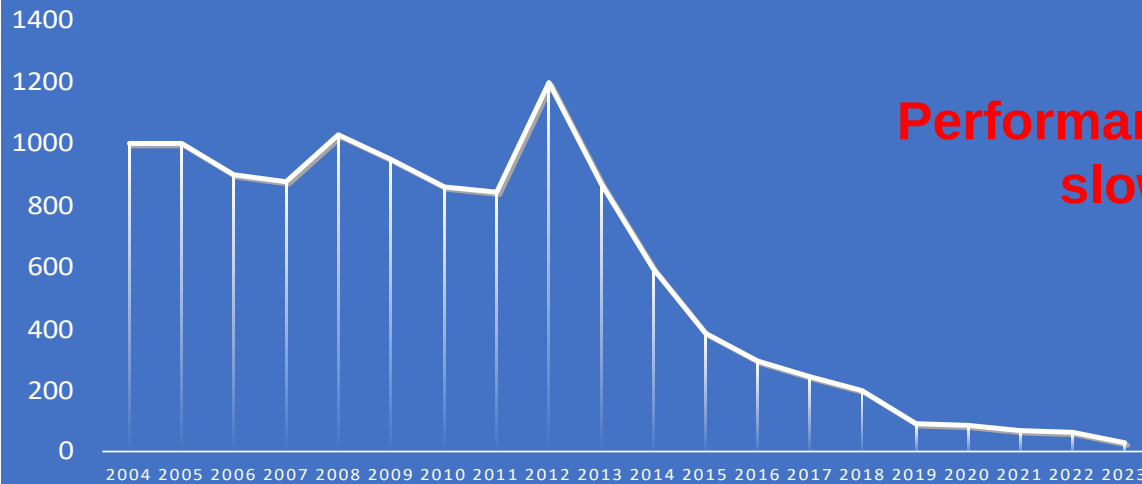
25

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

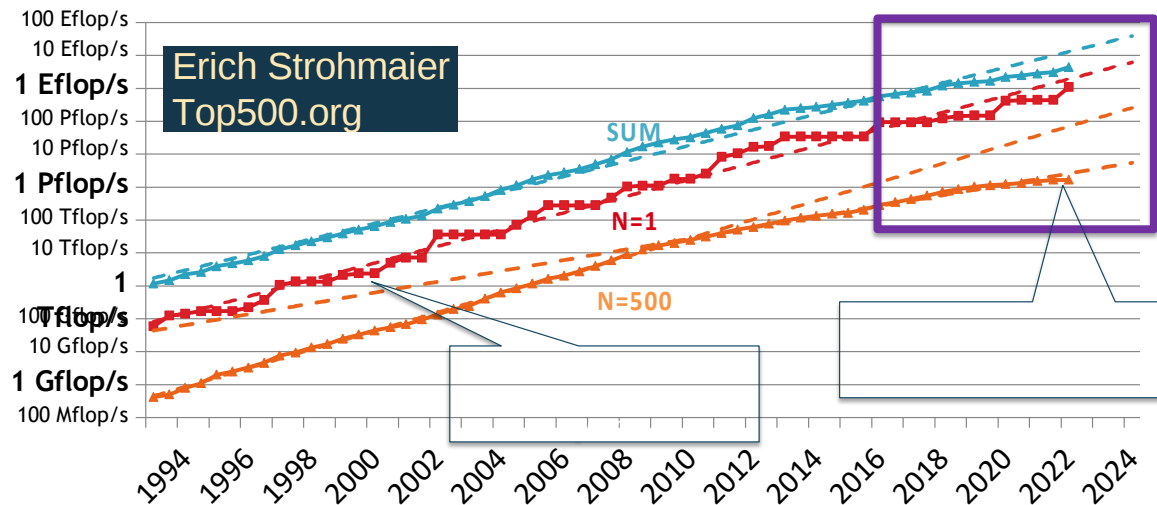
Committed to be a “Playground” of technologies, easily deployed and operated.

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



Performance improvements vs time
slowed down by $O(100)$



- **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

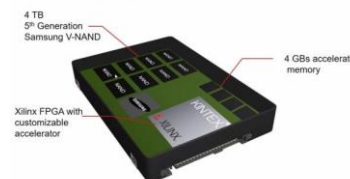
NVIDIA BlueField DPU



P4 programmable switches



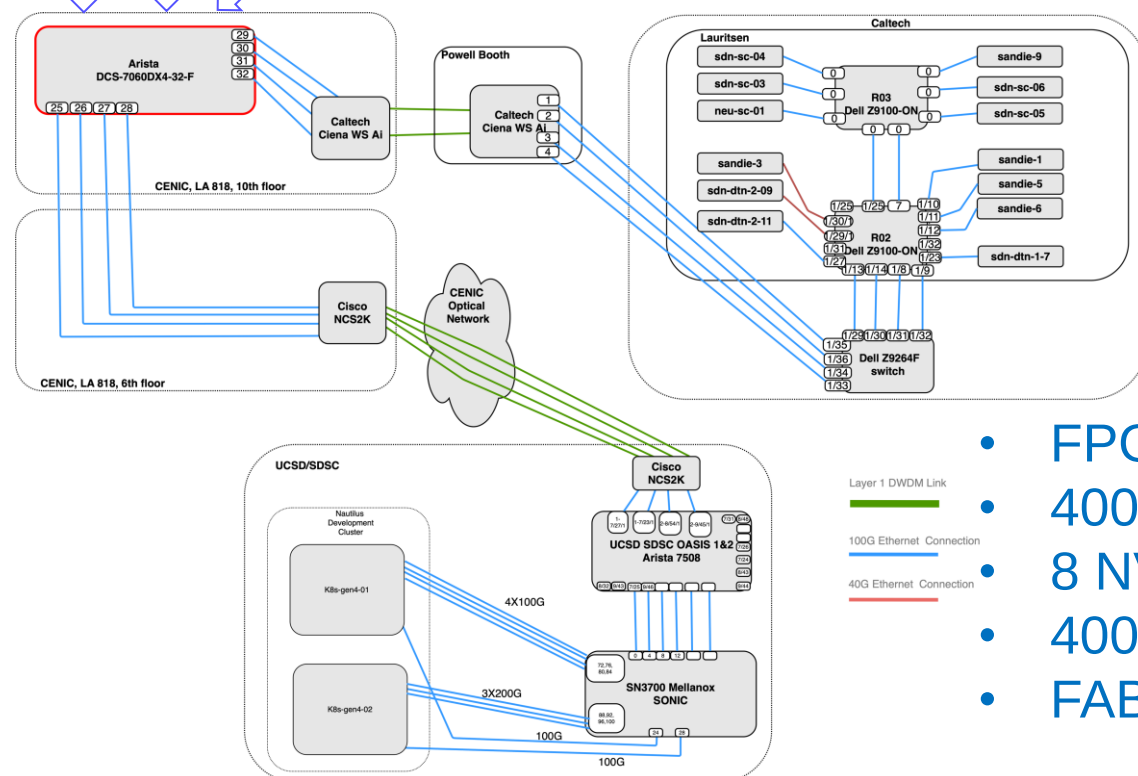
Xilinx SmartSSD



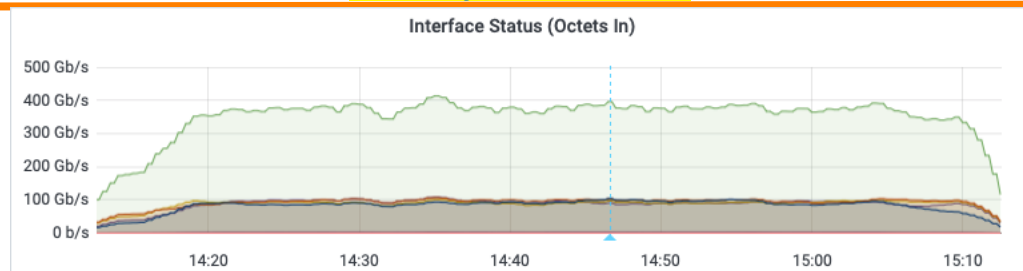
Thank you to CENIC

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

ESnet FABRIC Pacific Wave



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 8x80G A100s each
- 400TB of NVMe
- FABRIC node

Idea

Reference: [2209.13714](#)



Fundamental philosophy is to build a system where "the network" makes bandwidth promises that higher layers in the stack adjust to in order to maximize network bandwidth use given priorities of stakeholders inside an organization and between organizations.

This creates end-to-end accountability of network utilization (vertical and horizontal), and allows for prioritization of available bw by the organization's higher layers of the transfer stack.

Network may use dynamic routing to maximize aggregate bandwidth used by managing large long lasting flows.

R components that are affected by new higher priority request. R thus can adjust strategy if warranted.

- **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
- **Recent obsession with AI accelerates prior trends**
- We are off to an excellent start ... but there is lot's more to come over the course of the next 5 years.

- 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

