NRP NATIONAL RESEARCH PLATFORM

SDSC SAN DIEGO SUPERCOMPUTER CENTER

National Research Platform (NRP)

Integration of PNRP resources

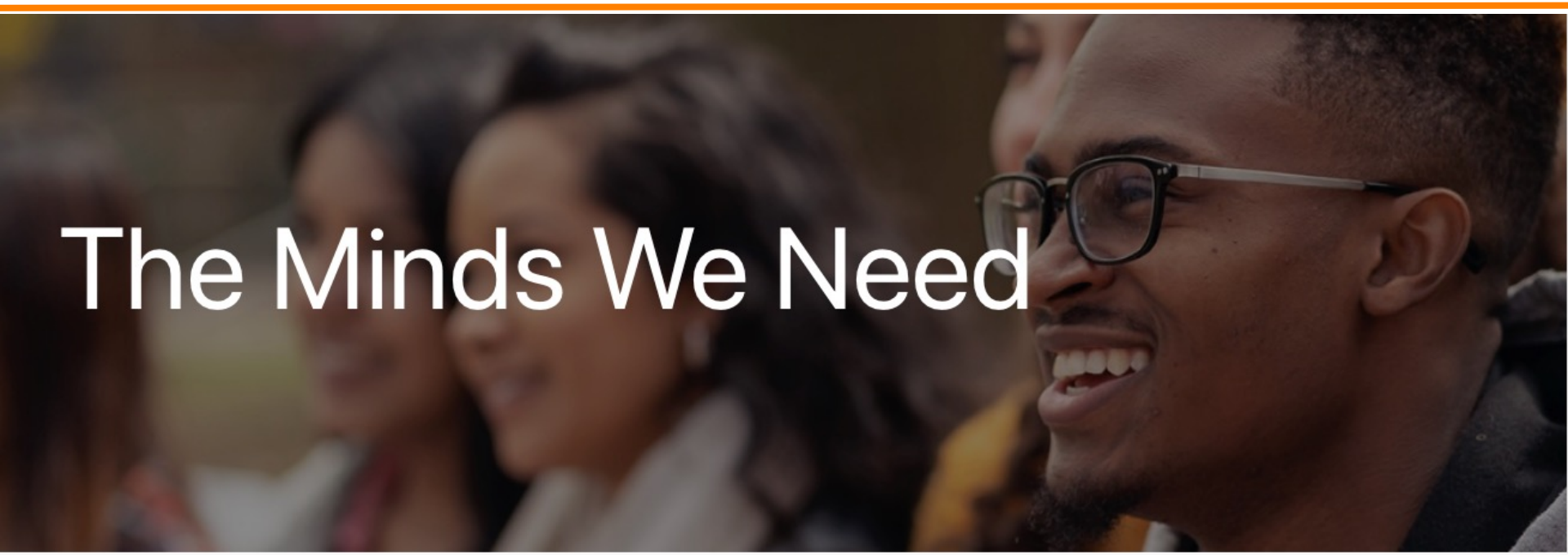
Mahidhar Tatineni
Director, User Services, SDSC

4th GLOBAL RESEARCH PLATFORM WORKSHOP
Oct 9-10, 2023

Acknowledgement: Frank Würthwein
Director, SDSC

NRP Vision

Democratize Access

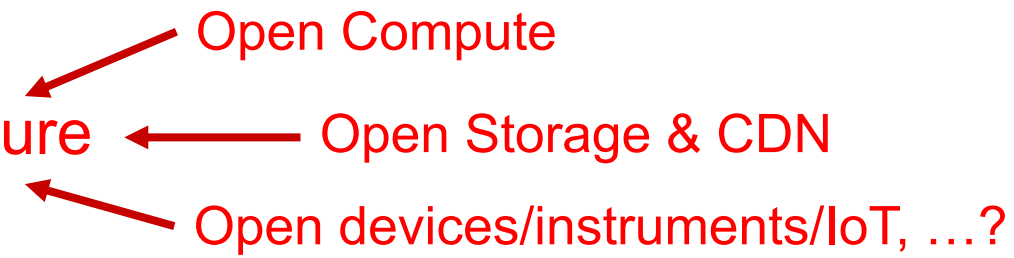


- **Connect every community college, every minority serving institution, and every college and university, including all urban, rural, and tribal institutions** to a world-class and secure R&E infrastructure, with particular attention to institutions that have been chronically underserved;
- **Engage 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

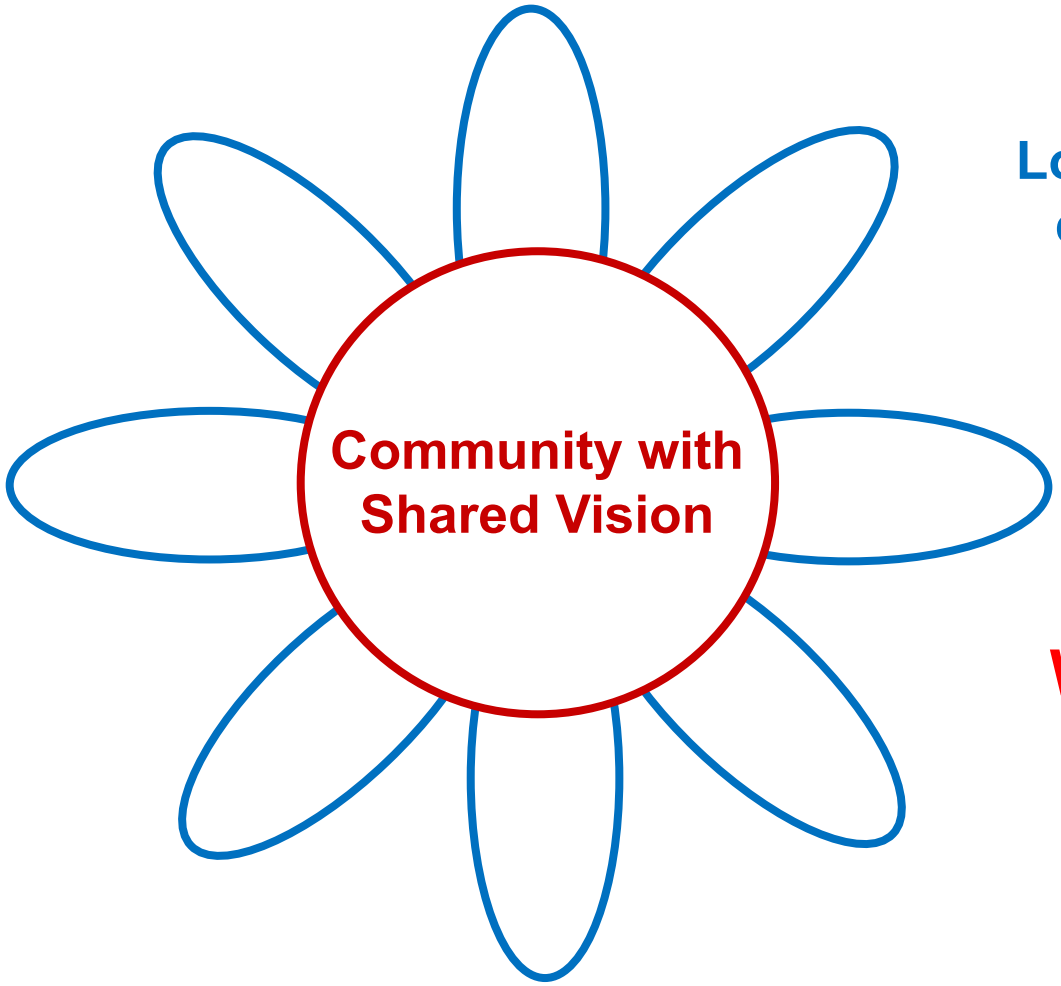
Long Term Vision

- 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.

We want you to …
… grow NRP.
… build on NRP.

NRP is "owned" and "built" by the community for the community

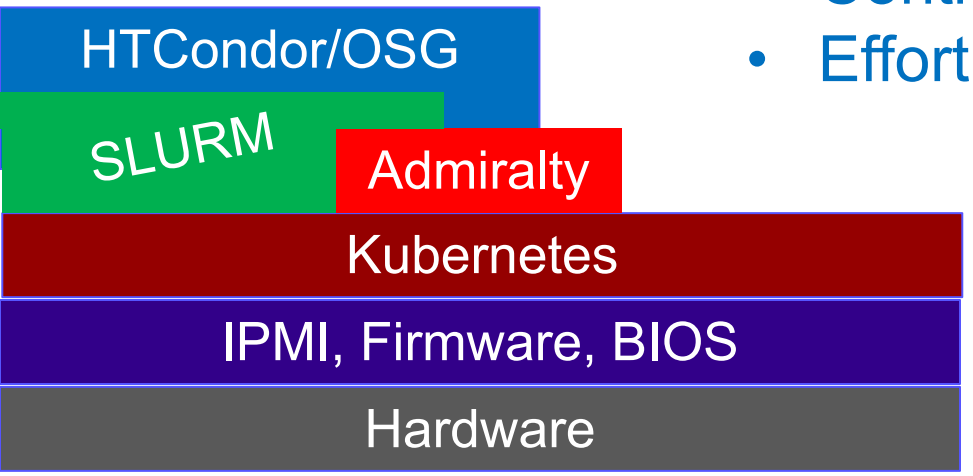
Cyberinfrastructure Stack

NRP operates at all layers of the stack, from IPMI up

- IPMI reduces TCO and lower threshold to entry
- Kubernetes allows service deployments
  - Also the natural layer for application container deployment
- Admiralty allows K8S federation with folks who want control
  - Including cloud integration to access TPUs & other cloud only architectures
- HTCondor allows NRP to show up as a “site” in OSG

The layer you integrate at depends on

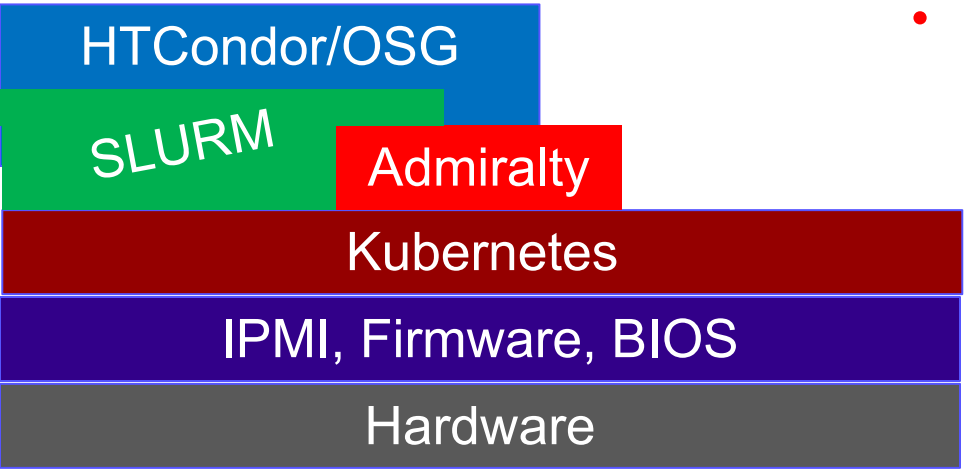
- Control you want
- Effort you can afford



Cyberinfrastructure Stack

NRP operates at all layers of the stack, from IPMI up

- **IPMI reduces TCO and lowers threshold to entry**
- Kubernetes allows service deployments
  - Also the natural layer for application container deployment
- Admiralty allows K8S federation with folks who want control
  - Including cloud integration to access TPUs & other cloud only architectures
- HTCondor allows NRP to show up as site in OSG



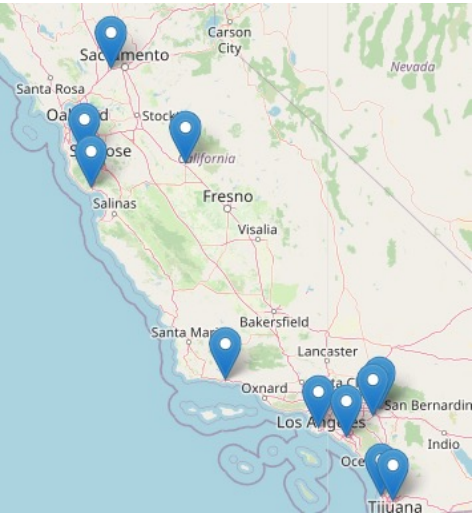
- **Under-resourced institutions**
- **Network providers and their POPs**
- **CS & ECE faculty specialized on:**
  - **AI/ML => gaming GPUs**
  - **systems R&D**

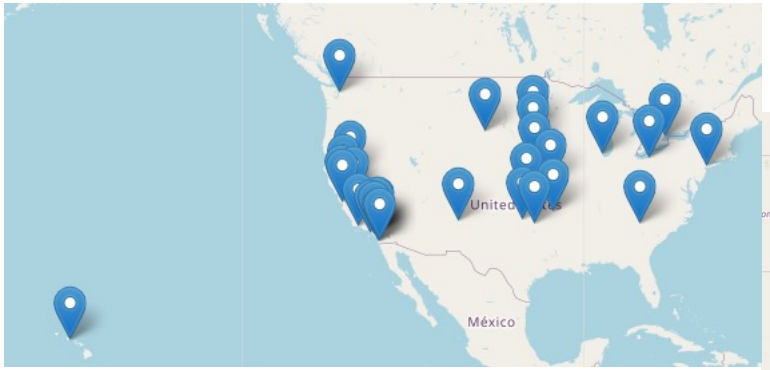
All of these find it difficult to justify staff to support all layers

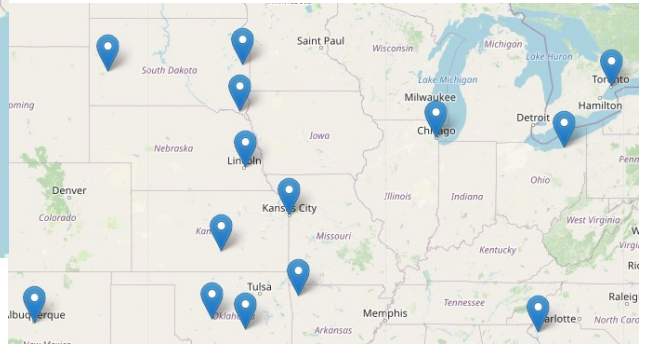
Hardware on NRP is Global

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

NRP integrates hardware in USA, EU, and Asia







Cyberinfrastructure Stack

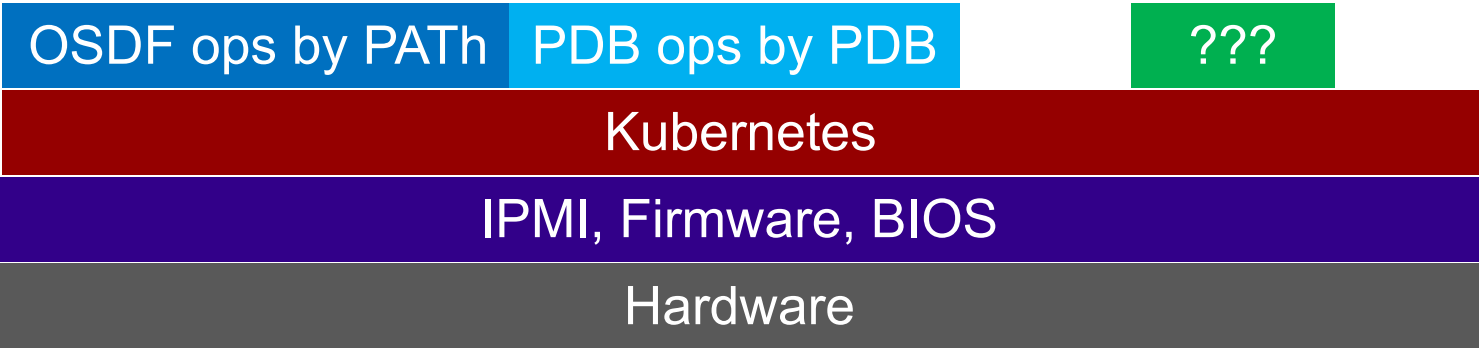
SDSC SAN DIEGO SUPERCOMPUTER CENTER

NRP operates at all layers of the stack, from IPMI up

- IPMI reduces TCO and lowers threshold to entry
- **Kubernetes allows service deployments**
  - Also the natural layer for application container deployment
- Admiralty allows K8S federation with folks who want control
  - Including cloud integration to access TPUs & other cloud only architectures
- HTCondor allows NRP to show up as site in OSG

NRP is unique in its support of global service deployments

- **Open Science Data Federation**
  - Origins & Caches in US, EU, Asia
- **Protein Data Bank**
  - (Future) Replicas in EU & Asia

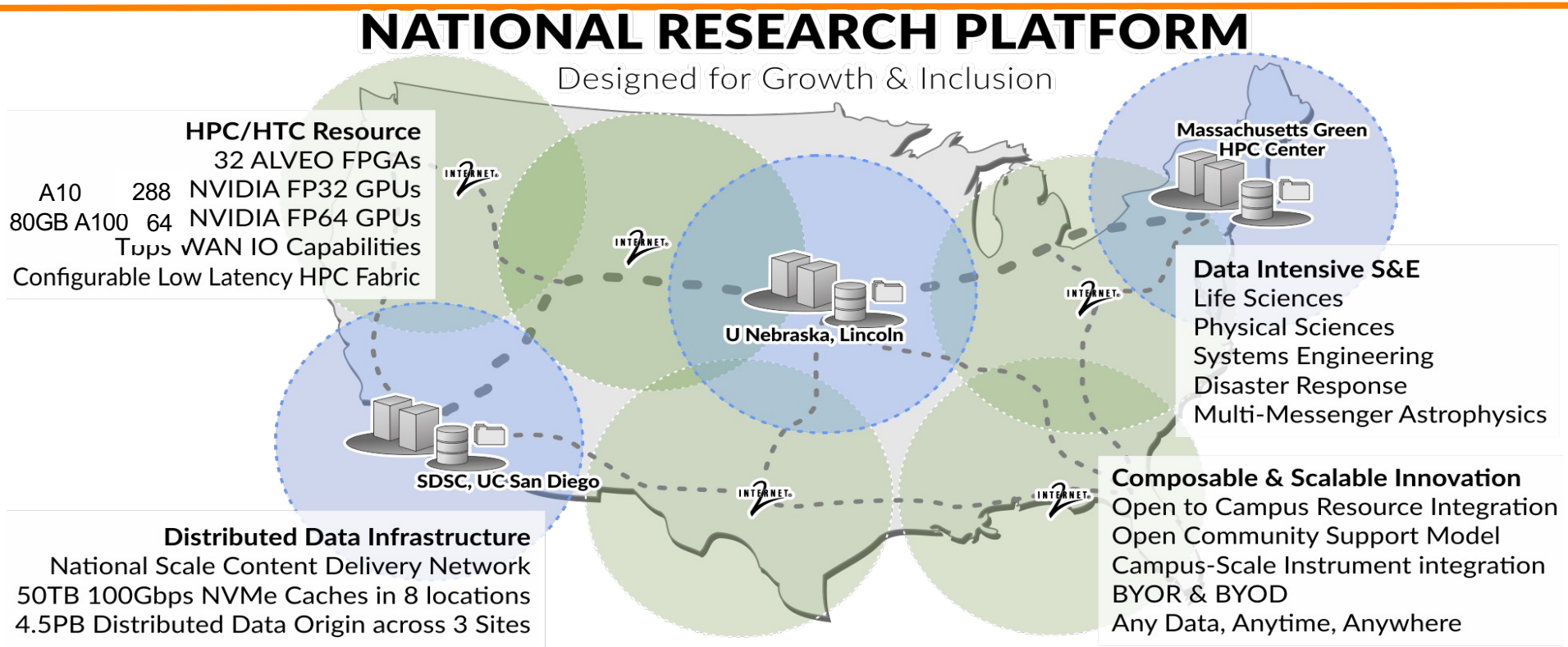


- **NSF Funded PNRP added significant resources to NRP**
- **PNRP project has operational funding and Nautilus will be the integrated resource.**

=> Nautilus transitioned from a PRP k8s cluster to NRP k8s cluster with operational funding for 5 years (with possibility of another 5-year renewal).

Cat-II: Prototype National Research Platform (PNRP)

SDSC
SAN DIEGO SUPERCOMPUTER CENTER



5 year project with $5M hardware & $6.45M people

Supports Nautilus, and thus the core NRP infrastructure

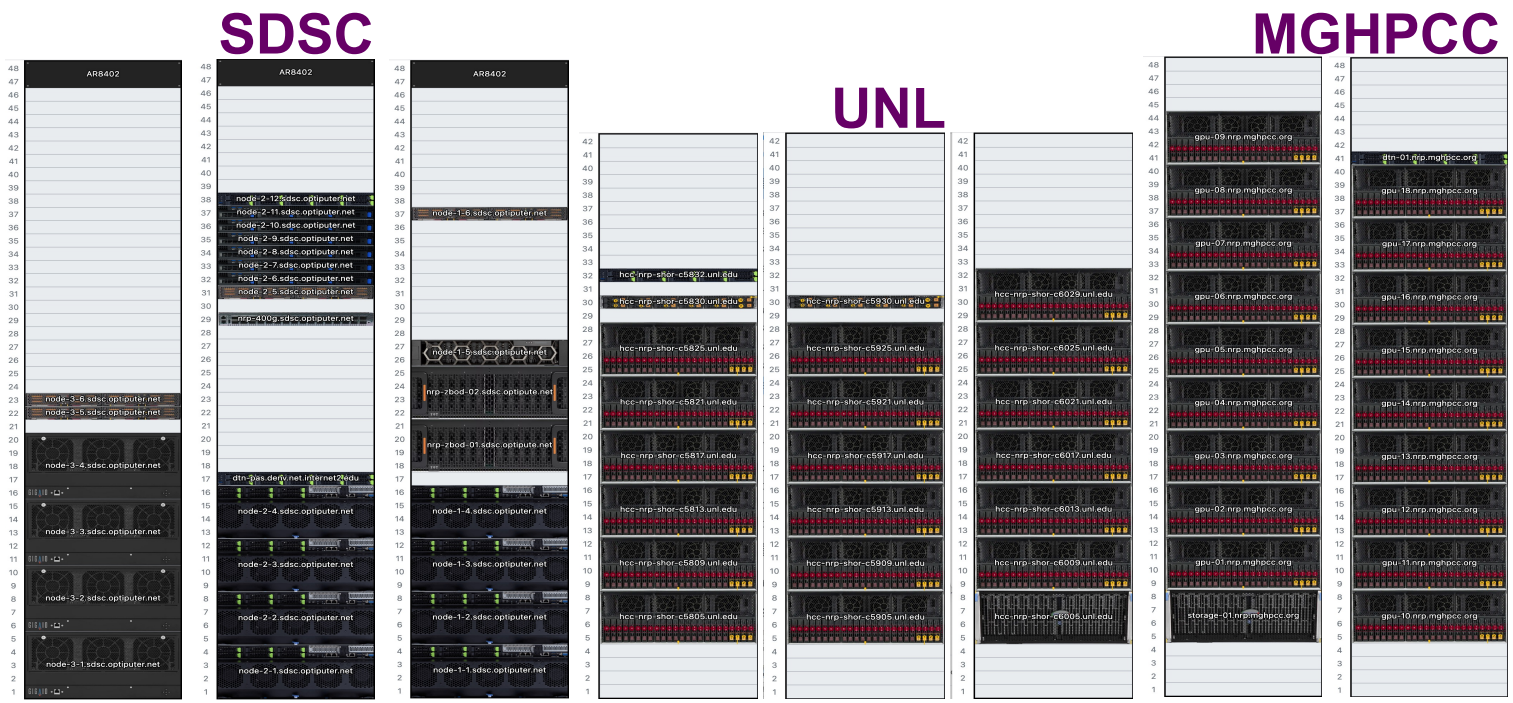
Promises to build on “PRP” functionality, and go beyond

NSF Acceptance Review completed, System in Testbed Phase

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

Funded as NSF 2112167

PNRP contributions



PNRP contributions

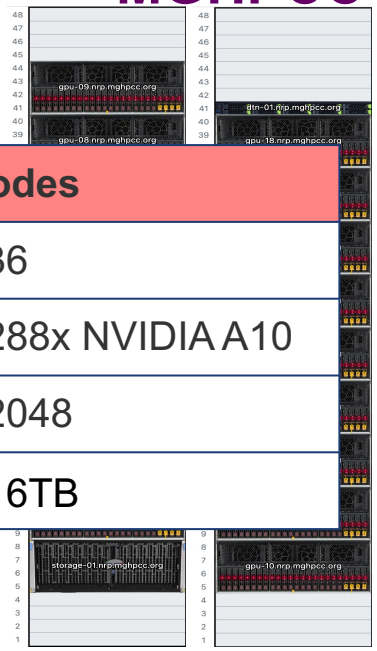
SDSC

Composable System	
Total Nodes	**Composable Each CPU node**
GPUs	64x NVIDIA A100
FPGAs	32x XILINX U55C
Cores	1792
Memory	10TB

UNL

FP32 Nodes	
Total Nodes	36
GPUs	288x NVIDIA A10
Cores	2048
Memory	16TB

MGHPCC



PNRP Project Innovations

SDSC SAN DIEGO SUPERCOMPUTER CENTER

- I1: Innovative network fabric that allows “rack” of hardware 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.

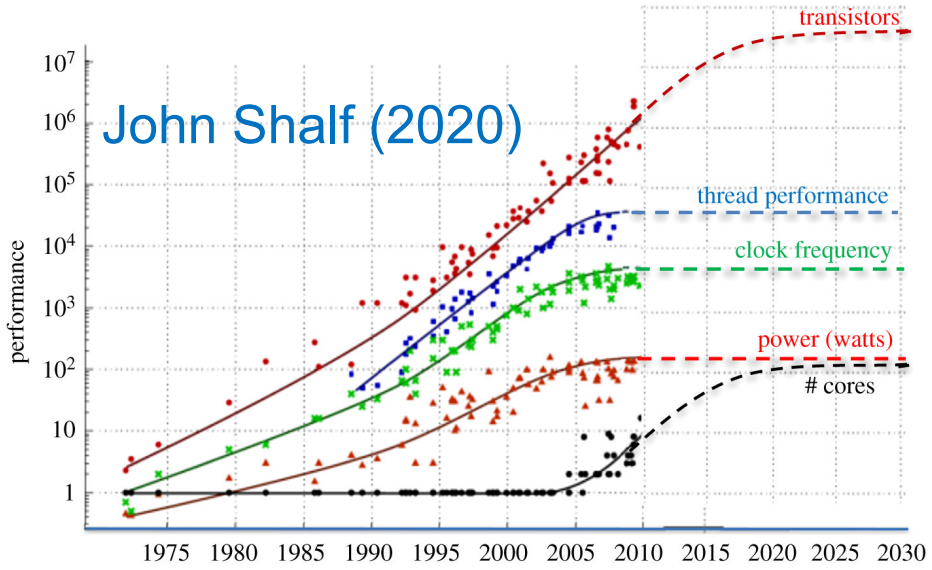
I3 & I4 & I5 turn ”PRP” into “NRP” and sustains it into the future.

I1 & I2 are totally new.

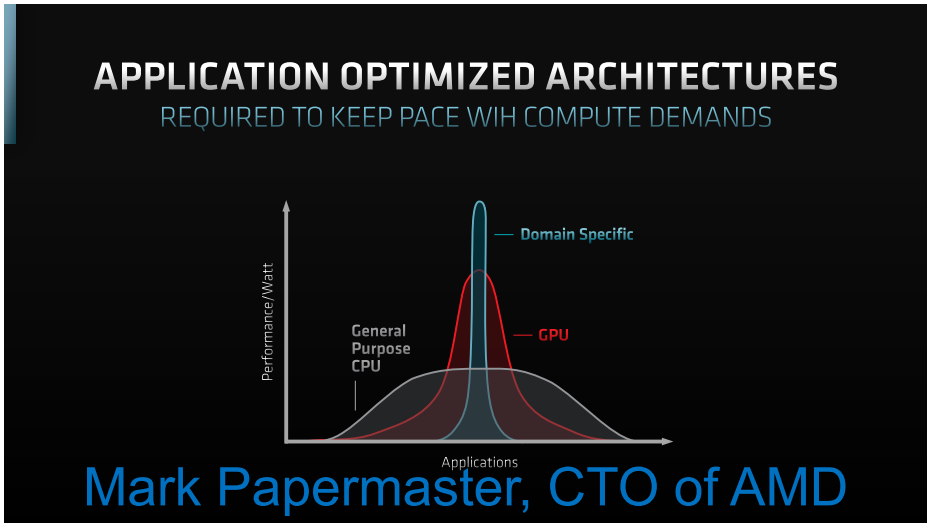
Domain Specific Architectures

- I1: Innovative network fabric allowing “composable hardware”.
- I2: Innovative application libraries allowing “domain optimized architectures” on FPGAs

“end of Moore’s law” motivates new architectures



John Shalf (2020)



Mark Papermaster, CTO of AMD

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

PI, Tajana Rosing

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

Bring Your Own Resources and Devices (BYOR/BYOD)

- "Bring Your Own Resources" (BYOR)
  - Typically, nodes host 8 GPUs in Science DMZs, can be CPU-only nodes
  - Campuses pay for networking, space, power, and hardware maintenance
  - PNRP and allied grants supply highly-automated sysadmin support
  - The Nautilus community provides training and onboarding help
  - Half of the nodes in Nautilus today are BYOR nodes.
- "Bring Your Own Devices" (BYOD)
  - Internet of Things (IoT) devices hanging off lab nodes in Science DMZs
  - Useful for 5G experimentation, for example.

Data Infrastructure Model

- Support regional Ceph storage systems across the USA.
  - Campuses can join individual storage hosts to the Ceph system in their region.
  - All regional storage systems are Origins in OSG Data Federation (OSDF)
  - Deploy replication system such that researchers can decide what part of their namespace should be in which regional storage.
- Deploy caches in Internet2 backbone such that no campus nationwide is more than 500 miles from a cache.

NRP data infrastructure model combines best of PRP & OSG

From PRP we take the regional Ceph storage concept
From OSG/PATh we take the data origin & caching concepts

And then we add as a totally new feature:
User controlled replication of partial namespaces across regions.
(We will develop this during 3-year "testbed" phase of PNRP Project)

Want Others to build higher level data services on top

Matrix of Science x Innovations

SDSC SAN DIEGO SUPERCOMPUTER CENTER

Table 3.1 Representative Science and Engineering Use Cases

Application domain	Lead researcher & Institution	Science Driver Themes	NRP Innovations
LIGO	Peter Couvares, LIGO Lab; Erik Katsavounidis, MIT	BGS, UC, AI	I2, I3, I4, I5
IceCube	Benedikt Riedel, UW Madison	BGS, UC, AI	I3, I4
Astronomy (DKIST & Sky Surveys)	Curt Dodds, U. Hawai'i	BGS, AI	I3, I5,
Campus Scale Instrument Facilities	Mark Ellisman, NCMIR; Samara Reck-Peterson, Nicon Imaging Center; Johannes Schoeneberg, Adaptive Optics Lightsheet Microscopy; Kristen Jepsen, Institute for Genomic Medicine; Tami Brown-Brandl, Precision Animal Management	SD, UC, H	I1, I2, I3, I4, I5
Molecular Dynamics	Rommie Amaro, UCSD; Andreas Goetz, SDSC; Jonathan Allen, LLNL	MD, AI, H	I1, I2, I3
Human microbiome	Rob Knight, UCSD	G, AI, H	I1, I2, I3
Genomics & Bioinformatics	Alex Feltus, Clemson	G, AI, H	I3, I4, I5
Fluid Dynamics	Rose Yu, UCSD	AI	I1, I2, I3
Experimental Particle Physics, IAIFI	Phil Harris, MIT	AI, BGS, SD	I1, I2
Computer Vision	Nuno Vasconcelos, UCSD	AI, CV	I3
Computer Graphics	Robert Twomey, UNL	CV, AI	I3
Programmable Storage	Carlos Malzahn, UCSC	SD	I1, I2, I5
AI systems software stack for FPGAs	Hadi Esmaeilzadeh , UCSD	SD	I1, I2
WildFire Analysis & Prediction	Ilkay Altintas, UCSD	UC, AI, CV	I3, I4

Key: The NRP innovations column lists those innovations among I1 through I5 listed in Section 2.1 that a given science driver most benefits from.

Lot's of AI …
but so much more …

NSF MREFCs

Incl. 4 campus scale instrument facilities

Incl. a very diverse set of sciences and engineering

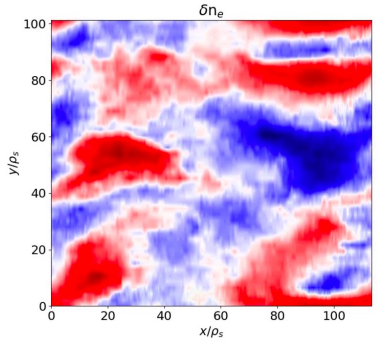
Status of innovative PNRP additions

Innovation	Status of deployment and testbed activities
I1: Network fabric	Deployed and configured GigaIO FabreX PCIe fabric for composition of CPU servers, FPGAs via pooling appliance, NVMe storage, HGX A100 systems, A100s via pooling appliance. Running multiple user applications over this.
I2: FPGA libraries	Deployed Xilinx U55C FPGAs and libraries for multiple research teams
I3: BYO Resource	More than twice as much FP32 GPU capacity available to PNRP user community than expected. More than twice as much data infrastructure hardware available to PNRP user community than expected
I4: Scheduling	Innovative scheduling via Kubernetes and composed units, and user interfaces, including Jupyter. Integrated scheduling across all PNRP hardware via Nautilus access. Integrated scheduling from OSG.
I5: Data Infrastructure	Global scale Content Delivery System already used as a production system for multiple research groups NSF CC* program will use PNRP for its storage awards. 9 awards made in 2022. Engagement with some of the awards already in full swing.

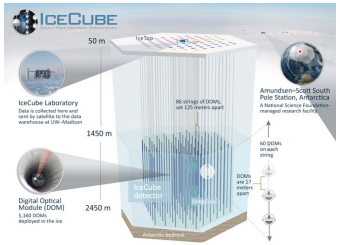
Status of PNRP additions

SDSC
SAN DIEGO SUPERCOMPUTER CENTER

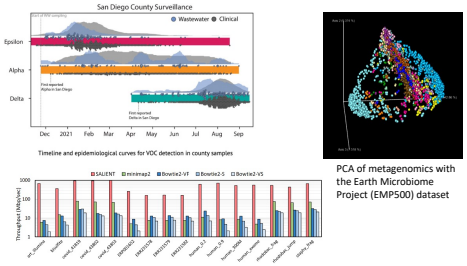
Early Users on PNRP tested/utilized the innovative system features



Fusion energy research (A100)

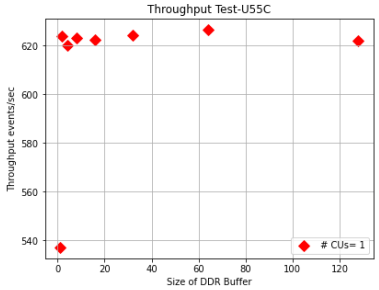


IceCube Neutrino Observatory (A10, Data)

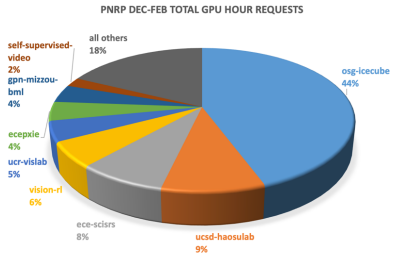


PCA of metagenomics with the Earth Microbiome Project (EMP500) dataset

Genomics pipeline (FPGA)

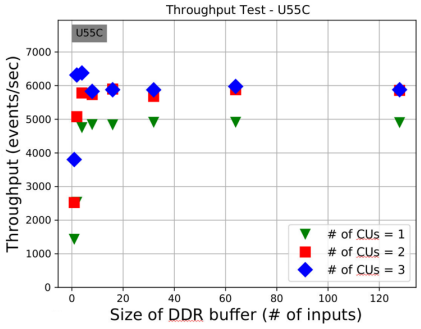


LIGO data analysis (FPGAs, Data)

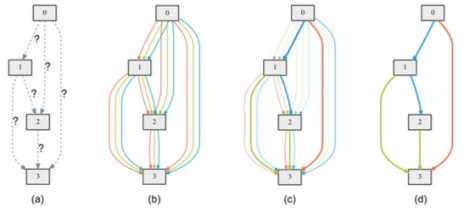


Namespace	PNRP Dec-FebTotal GPU hour requests
osg-icecube	234,859
ucsd-haosulab	49,380
ece-scisrs	42,436
vision-rl	30,060
ucr-vislab	24,182
ecepxie	22,705
gpn-mizzou-bml	21,267
self-supervised-video	10,922
all others	93,938
total	**529,749**

Broad community adoption



Large Hadron Collider (FPGAs)



Machine learning for NLP (A10, A100)

Summary & Conclusions

- **PRP ended, and was replaced by NRP**
  - Significant new capabilities via Cat-II system “PNRP”
    - PNRP project includes funding of ops that will support Nautilus cluster for the future
  - Major increase in # of GPUs in the past couple of years including significant additions via PNRP project (288 A10s, 64 A100s)
  - # of FPGAs also increased with PNRP adding 32 Xilinx U55C FPGAs.
  - # of caches grow by 50% in 22/23
    => more consistent coverage across USA
  - Data volume served expected to grow substantially in 23/24/25.
    - How much? As yet too hard to predict.
- Hoping to recruit new partners to build **FAIR capabilities on top of OSDF within the next 5 years**.
- Hoping to expand NRP into **sensor networks using 5G & 6G in the next 10 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