

NRP for the next 10 years

Frank Würthwein
Director, San Diego Supercomputer Center

December 8th 2022



Vision

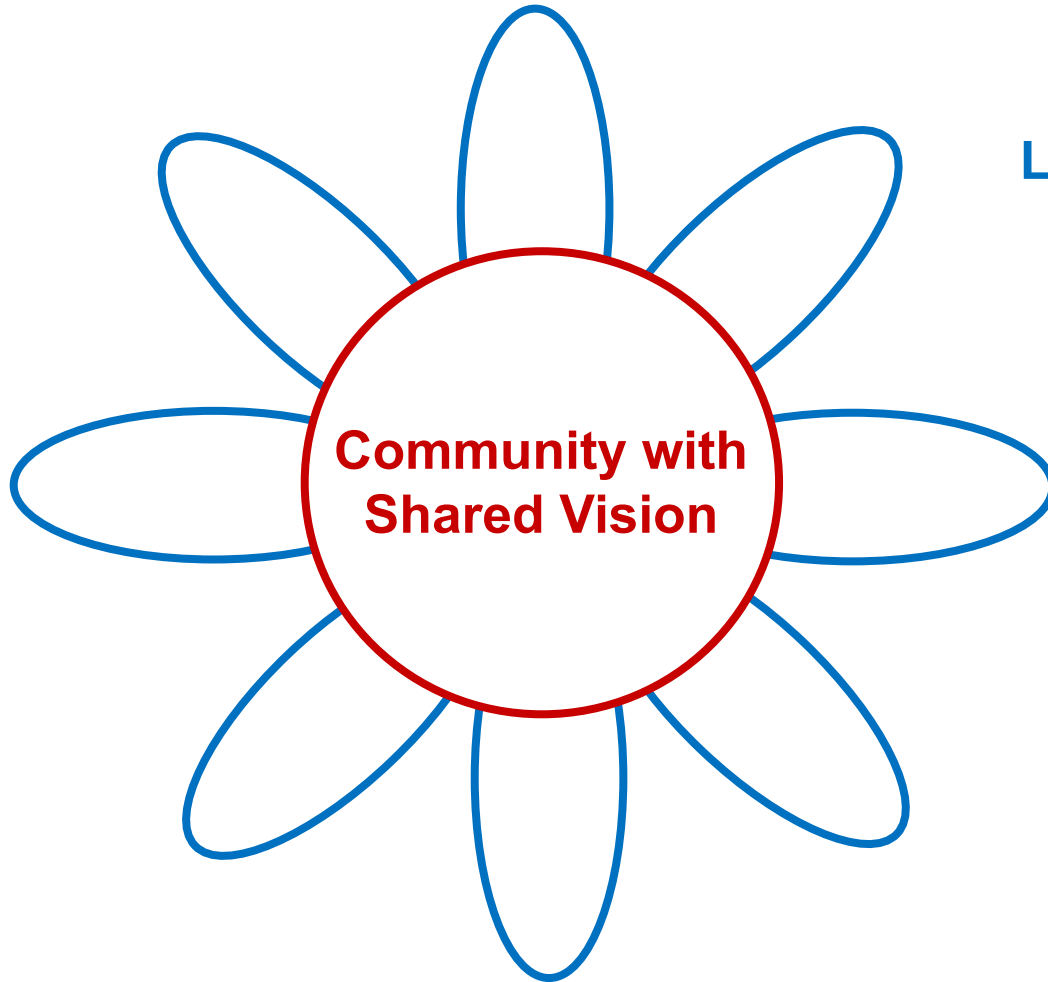


The Minds We Need

- **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

- 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



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

3 Ways to build Open Infrastructure Federation(s)

- At the container orchestration layer, e.g. K8S as implemented in Pacific Research Platform (PRP).
- Federating independent container orchestration frameworks, e.g. via SLATE or via Admiralty on K8S.
- Federating storage and compute clusters at the storage and batch system layer., e.g. via OSG.

**More (cybersecurity) control
implies more effort to join & operate.**

Some Institutions will never be able to join because of a mismatch between the effort they have and the control they desire.

Complementarity in Implementation of “Bring Your Own Resource” model

OSG focused on campus cluster integration.

PRP focused on individual node integration instead of clusters.

National Research Platform (NRP) supersedes PRP in Fall 2022.

Supporting Nautilus for the next decade

Nautilus = K8S infrastructure of PRP for the last 5+ years

Nautilus = K8S of NRP for the next 10 years

- NSF supports via the Cat-II program novel systems ideas.
 - 3 year “testbed” phase
 - The PI owns the resource, and has (some) freedom regarding who uses it.
 - No requirements for making it available via any specific allocation mechanism.
 - It is expected that not all features work on day 1.
 - 3 years of experimentation & development of features
 - 2 year “allocation” phase
 - The resource is made available via an NSF supported allocation mechanism.
 - The solicitation mentions the possibility of an additional 5 year renewal without re-competition if system is successful.
- **We decided that this is an ideal program to try and secure NRP core operations funding for the next 10 years**
 - And thus provide the stability necessary for growth of NRP.

NATIONAL RESEARCH PLATFORM

Designed for Growth & Inclusion

HPC/HTC Resource

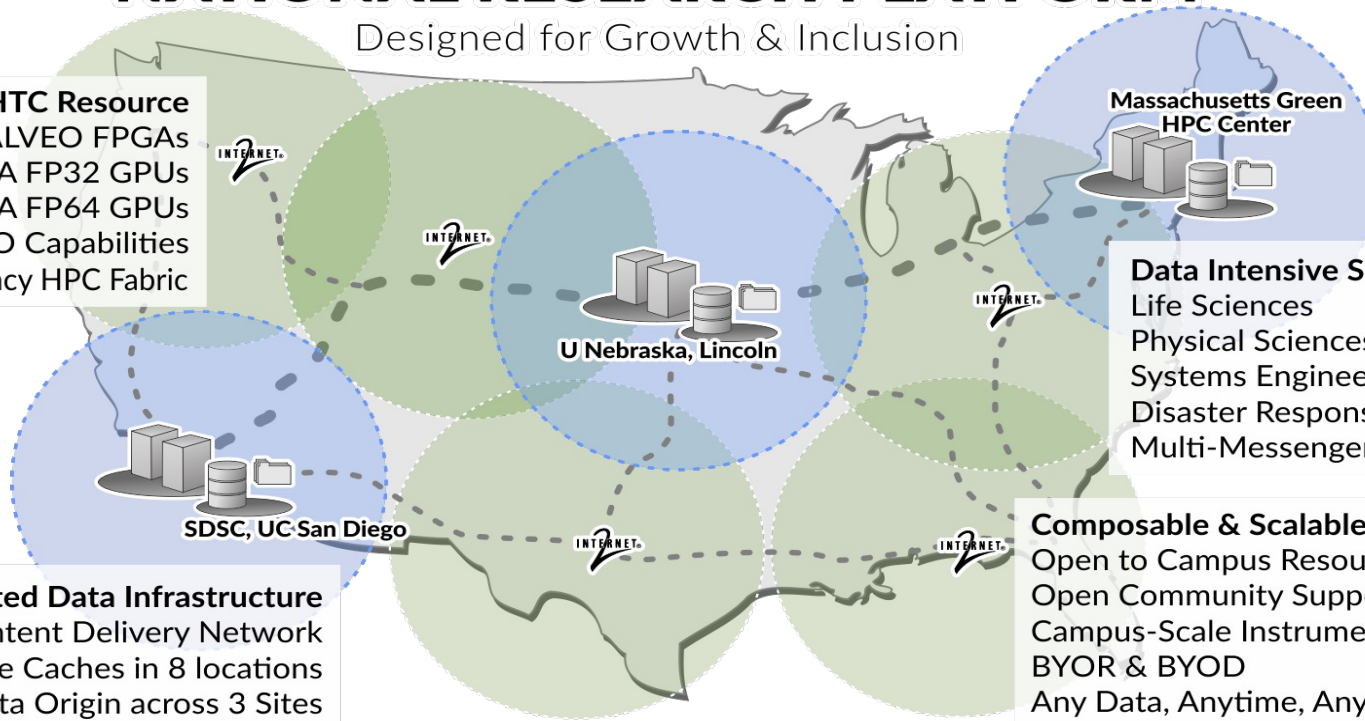
32 ALVEO FPGAs

A10 288 NVIDIA FP32 GPUs

80GB A100 64 NVIDIA FP64 GPUs

Tbps WAN IO Capabilities

Configurable Low Latency HPC Fabric



Massachusetts Green HPC Center

Data Intensive S&E

- Life Sciences
- Physical Sciences
- Systems Engineering
- Disaster Response
- Multi-Messenger Astrophysics

Distributed Data Infrastructure

National Scale Content Delivery Network

50TB 100Gbps NVMe Caches in 8 locations

4.5PB Distributed Data Origin across 3 Sites

Composable & Scalable Innovation

- Open to Campus Resource Integration
- Open Community Support Model
- Campus-Scale Instrument integration
- BYOR & BYOD
- Any Data, Anytime, Anywhere

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

Supports Nautilus, and thus the core NRP infrastructure

Promises to go beyond “PRP” functionality

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

- 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.**

- 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 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)

Want Others to build higher level data services on top

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, Nikon 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

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

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.

- **Growth of NRP infrastructure**
 - 1,000++ GPUs end of 2022
 - 50 PB storage end of 2024
 - Growth in diversity of community
 - # and types of campuses and their researchers
- **Introduce new capabilities into NRP**
 - Machine learning at 100TB scale
 - Experiment with new types of “nodes” on NRP
 - Expand NRP into Wireless, Edge, IoT
- **New Directions initiated by the Community**

See you at 4th NRP Workshop, San Diego, Feb. 8-10th 2023

“Bring Your Own Resource” as Strategy for Meeting these Goals

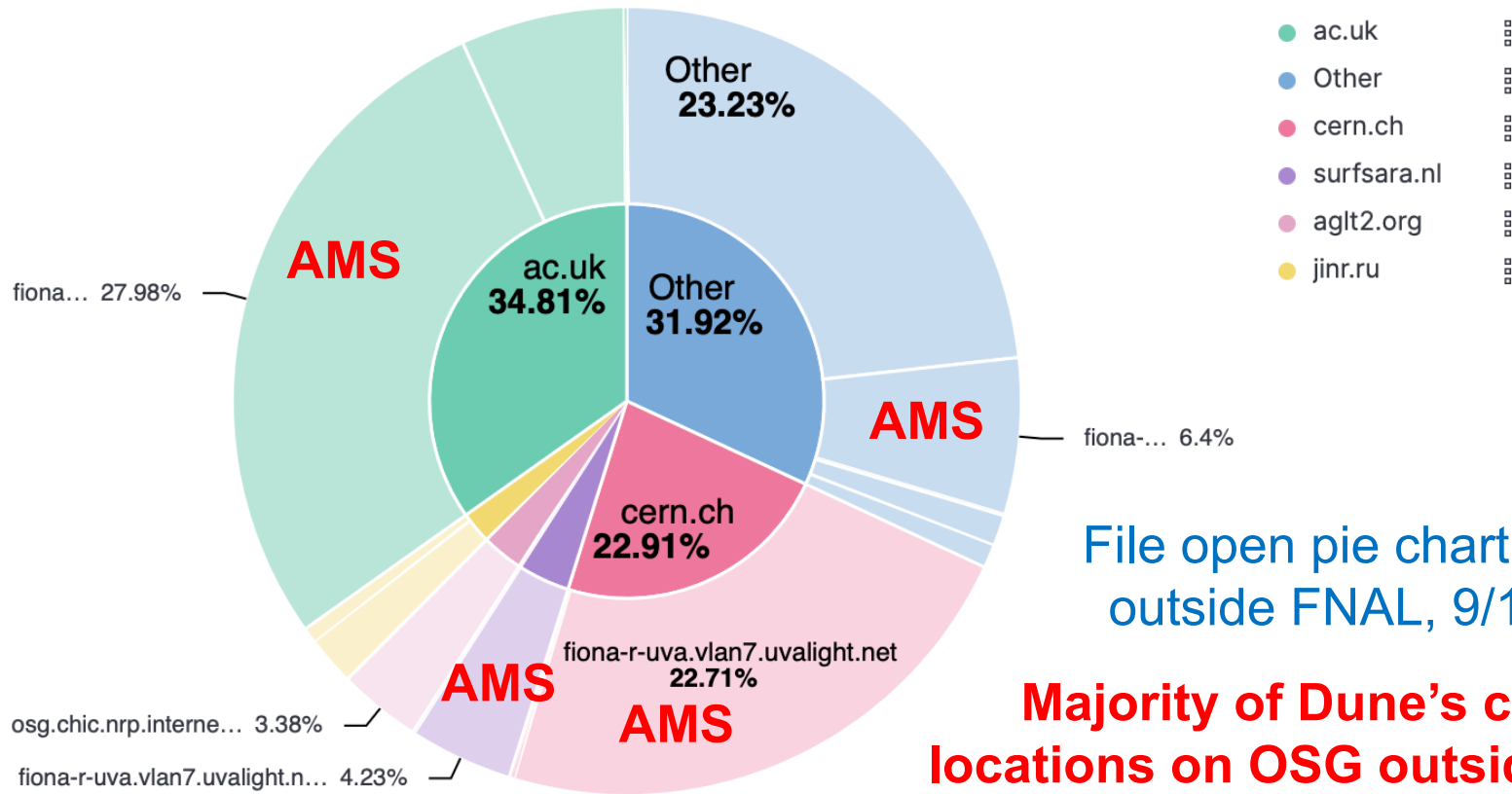
GPU goal will be met based on new hardware from half a dozen new projects that are committed to work with us.

What about storage & data ?



20 Caches ... 6 of which are in R&E network backbone
10 Data Origins ... incl. one in Europe

- 8 from NRP to cover the continental USA with 500 mile circles.
 - 9th cache in Miami as part of PATH facility
 - 10th cache in Hawaii
- 2 from ESnet at London & Amsterdam POPs
 - Amsterdam is currently one of our oldest & most heavily used caches.
 - It needs an upgrade !
 - London cache will take on significant load from Amsterdam.



File open pie chart for Dune outside FNAL, 9/1 to 9/29.

Majority of Dune's compute locations on OSG outside FNAL hit the Amsterdam cache.

Compute resources in EU comprise 1/4 of global Dune resources on OSG

- The NSF CC* program awarded 9 campuses with \$500k storage system awards each.
 - We guess this pays for 5PB of storage each.
- Some of these campuses may decide to integrate their CC* storage into the OSDF.
- Some of these campuses have storage from other projects that they may integrate with the OSDF in addition.
- Remains to be seen how much total data will become available in OSDF as a result.

- The NSF announced the 2023 CC* program ... **NSF 23-526**
- ... and it again includes opportunities for campuses to propose Data Origins.

=> More opportunities to grow the OSDF total data volume beyond 50 Petabytes

- **PRP ended, and was replaced by NRP**
 - Responsibility for the bulk of K8S operations moved from San Diego to Lincoln, Nebraska.
 - # of GPUs available double in 2022.
 - new GPUs (A10, 3080, 3090, A100) much more powerful than older GPUs
 - # of FPGAs increase from a few to a few dozen in 2022.
 - # 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.**

Comments & Questions ?