Report: BoF: NSF innovative computing technology testbed community exchange

Amit Majumdar[1], Tim Boerner[2], David Hancock[3], Honggao Liu[4], Dan Stanzione[5], Sergiu Sanielevici[6], Robert Harrison[7], and Eva Siegmann[7]

[1]San Diego Supercomputing Center
[2]University of Illinois Urbana-Champaign
[3]Indiana University
[4]Texas A&M High Performance Research Computing
[5]Texas Advanced Computing Center
[6]Pittsburgh Supercomputing Center
[7]Stony Brook University

PEARC Conference 2022

The session took place at PEARC'22 on July 12th in Boston. It was chaired by Robert Harrison and brought together the leadership, technical staff, and user communities of several NSF-funded innovative computing testbeds. The symposium lasted for one hour, with presentations from the individual testbeds (each 5-8 minutes), followed by an open discussion between the panelists and the audience. There were around 60 people present at the BoF session.

Figure 1: Panelists at the BoF session.

These diverse testbeds examine novel technologies, architectures, usage modes, etc., and explore new target applications, methods, and paradigms for discovery in science and engineering. Despite the great diversity of activities and resources that span networking, novel and advanced processor and accelerator technologies, advancing HPC/AI and its applications, these testbeds have many common interests, communities, and potential synergies.

This report includes a short description of all testbeds represented in the BoF, as well as a summary of the discussion during the session. The appendix includes the presented slides.

1 Description of the testbeds

- **Honggao Liu: ACES**

  ACES (Accelerating Computing for Emerging Sciences), an innovative composable computational prototype to be developed by Texas A&M University, tries to answer a fundamental question: how does one effectively offer a holistic computing platform that can simultaneously meet the needs of a continuum of users in diverse research communities with varying levels of computing adoption? ACES will allow researchers to creatively develop new programming models and workflows that utilize these architectures while simultaneously advancing HPC and data science projects. ACES is a dynamically composable high-performance data analysis and computing platform. ACES leverage Liqid's composable framework on Intel's Sapphire Rapid processors to offer a rich of accelerators testbed containing Intel Ponte Vecchio GPUs (Graphics Processing Units) , Intel FPGAs (Field Programmable Gate Arrays), NEC Vector Engines, NextSilicon co-processors, Graphcore IPUs (Intelligence Processing Units) coupled with Intel Optane memory and DDN Lustre storage interconnected with Mellanox NDR 400Gbps (gigabit-per-second) InfiniBand to support workflows that benefit from optimized devices.

The ACES platform removes significant bottlenecks in advanced computing by introducing the flexibility to aggregate various components (i.e., processors, accelerators and memory) on an as-needed basis to solve problems that were previously not addressable. By letting researchers switch and run on accelerators best suited for their workflows, ACES will benefit many research and development projects in the fields of artificial intelligence and machine learning (AI/ML), cybersecurity, health population informatics, genomics and bioinformatics, human and agricultural life sciences, oil & gas simulations, de novo materials design, climate modeling, molecular dynamics, quantum computing architectures, imaging, smart and connected societies, geosciences, and quantum chemistry. The initial phase of ACES with new accelerators such as Graphcore IPUs and Intel FPGAs, and composable Optane memory is available to the national research community via the NSF-funded FASTER platform since July 8, 2022.

- **Tim Boerner: Delta**

  The largest GPU resource in the NSF ecosystem when it launches, Delta will contain:

  - 124 CPU compute nodes and 8 utility nodes featuring AMD Epyc 7763 64-core Milan processors
  - 100 each of 64- and 32-bit quad-GPU nodes with Nvidia A100 and A40 GPUs, respectively
  - Five 8-way Nvidia A100 GPU nodes and one 8-way AMD MI100 GPU node, all of which do double-duty as high-memory nodes
  - 3 PB relaxed-POSIX SSD-based storage and 7 PB of Lustre-based storage

  Delta will support:

  - Traditional HPC, data-intensive, and AI/ML workloads, with multiple compilers and runtimes
  - Third-party software installation via Spack or Easybuild
  - A modules utility for modifying programming environments
  - Containers, including Singularity and vendor-supplied containers from Nvidia and AMD
  - Persistent user tasks, supported by the utility nodes, for databases, workflow orchestration, etc.
  - Gateways, through a key partnership with the Science Gateway Community Institute and engagement with multiple gateway communities

- **Amit Majumdar: Voyager**

  Voyager is an innovative resource for exploring AI processors in science and engineering. It has specialized training and inference processors, and optimized implementations of standard machine learning frameworks. The high speed interconnect between training nodes provides a unique capability to scale up training models – something the research community is keen to explore. Voyager will offer:

  - 42 training nodes, each with eight Habana Gaudi processors, Ice Lake host CPUs and 6.4 TB of node-local storage
  - Two inference nodes with eight Habana Goya processors and 3.2 TB of local storage
  - 36 Intel x86 CPU compute nodes
  - An on-chip, 400-GbE Arista Gaudi network

- A 3-PB high-performance storage system, accessible via a 25-GbE connection
- A 324-TB home file system, also connected via 25 GbE
- Deep-learning frameworks such as TensorFlow and PyTorch
- Software tools and libraries specific for Voyager's architecture to enable user to develop AI techniques.

- **Dan Stanzione: Chameleon**

  Chameleon is a deeply reconfigurable experimental testbed supporting Computer Science (CS) systems experimentation. The platform consists of two sites, University of Chicago (UC) and Texas Advanced Computing Center (TACC), along with support functionalities from Northwestern University and the University of North Carolina. the platform balances investment in large-scale hardware to support Big Compute and Big Data experimentation with diversity reflected by smaller clusters of graphic processing units (GPUs), field programmable gate arrays (FPGAs), specialized architectures, and innovative networking hardware. Users can reconfigure this hardware at bare-metal level, boot from custom kernel, get access to serial console, or provision and reconfigure Software Defined Networking (SDN)-enabled switches. Since its public availability in July 2015, Chameleon has attracted a community of over 4,000 users working on over 600 education and research projects.

- **Sergiu Sanielevici: Neocortex**

  Neocortex is a highly innovative resource that targets the acceleration of AI-powered scientific discovery by vastly shortening the time required for deep learning training, fostering greater integration of artificial deep learning with scientific workflows, and providing revolutionary new hardware for the development of more efficient algorithms for artificial intelligence and graph analytics.

  Neocortex democratizes access to game-changing compute power otherwise only available to tech giants for students, postdocs, faculty, and others, who require faster turnaround on training to analyze data and integrate AI with simulation. It also inspires the research community to scale their AI-based research and integrate AI advances into their research workflows.

- **David Hancock: Jetstream2**

  Jetstream2 capabilities will include:

  - An enhanced IaaS model with improved orchestration support, elastic virtual clusters, federated JupyterHubs, and improved storage sharing
  - A commitment to more than 99 percent uptime to better support science gateways and hybrid-cloud computation
  - A revamped user interface with unified instance management and multi-instance launch
  - More than 57,000 next-gen AMD EPYC processor cores
  - Over 360 Nvidia A100 GPUs
  - Greater than 17 PB storage in an NVMe/disk hybrid
  - A 100-GbE Mellanox network

- **Eva Siegmann: Ookami**

  Ookami provides researchers with access to the A64FX processor developed by Riken and Fujitsu for the Japanese path to exascale computing and is currently deployed in the until June 2022 fastest computer in the world, Fugaku. It is the first such computer outside of Japan. By focusing on crucial architectural details, the ARM-based, multi-core, 512-bit SIMD-vector processor with ultrahigh-bandwidth memory promises to retain familiar and successful programming models while achieving very high performance for a wide range of applications. It supports a wide range of data types and enables both HPC and big data applications.

  The Ookami HPE (formerly Cray) Apollo 80 system has 174 A64FX compute nodes each with 32GB of high-bandwidth memory and a 512 Gbyte SSD. This amounts to about 1.5M node hours per year. A high-performance Lustre filesystem provides about 800 TB storage. It also includes a few other nodes to facilitate users exploring current computer technologies and contrasting performance and programmability with the A64FX.

2 Discussion

- A user requested more information about the memory component of ACES system

  ACES provides 3TB composable Intel Optane SSDs addressable as memory using MemVerge Memory Machine. The compute node has standard DRAM and can compose optane SSDs as memory. If the application requests for more than 140GB memory, then the optane memory is used.

- Why is the testing phase time larger than production?

  This is the same on all testbeds. The initial testing phase lasts for up to 3 years. However, the systems are available to everyone during this initial period, not just staff.

- What are the gaps/barriers to broader deployment of your testbed's technologies within the NSF cyberinfrastructure or academic computing in general and how to address them?

  - SDSC – work closely with users, reaching out to and engaging with the broader community in general, hosting workshops. Habana is very hands on (with the staff@sdsc).
  - Neocortex – Cerebras works closely with neocortex staff. Neocortex staff hosts meetings, tutorials/training with collaborators. Identifying successful projects/use cases and presenting it to the community.
  - Chameleon – barriers complexity is wide. Pioneers want to mess with new hardware, but not users. One way is to prove that you can do it including the pain involved in the process (keep it realistic). Show successes to the community to address adoption issues. Flexibility of testbed systems for university users (not production use cases).
  - Ookami – The team is working on porting and tuning commonly used science for Fujitsu A64FX. They offer various support opportunities for users (slack channel, office hours, webinars, etc.). It's very important to work with the community to help them porting their applications. The project also works directly with students from different research groups to make respective codes performant.
  - Jetstream2 – all systems have different challenges. With Jetstream challenges were with OpenStack at scale and the rapidly evolving landscape of orchestration (e.g. Mesos / Swarm / Kubernetes) and containerization. Taking need-driven proof-of-concepts allow different techniques and then individuals can adapt and readily use them. It's an evolving path.
  - Delta – had a different transformation path. They already use commodity hardware – the new thing on their side is the non-POSIX file system. They show the benefits of new technology and how others can take advantage of it without putting in too much effort or work.
  - ACES – it's the latest testbed system which provides a common platform with several new accelerators. Employs Liqid technology for composable software/hardware. Having a common platform for cross-architecture of CPUs, GPUs and FPGAs is difficult. Dealing with innovation vs common/familiar/general use.

- What are the plans for providing trainings and workshops for researchers and general use cases when platform enters production?

  These are not general purpose systems, so nearly all users need more support and guidance than on conventional systems. User support plays a crucial role for the success of the testbeds. All testbeds emphasize that they have plans for various online trainings, publishing use cases, and documentation.

NSF innovative computing technology testbed community exchange

Name	Testbed	Institution
Robert Harrison (Chair)	Ookami	SBU
Honggao Liu	ACES	TEXAS A & M
Tim Boerner	Delta	NCSA
Amit Majumdar	Voyager	SDSC
Dan Stanzione	Chameleon	TACC
Sergiu Sanielevici	Neocortex	PSC
David Hancock	Jetstream2	Indiana University
Eva Siegmann	Ookami	SBU

ACES - Next Generation Composability

Honggao Liu, Ph.D

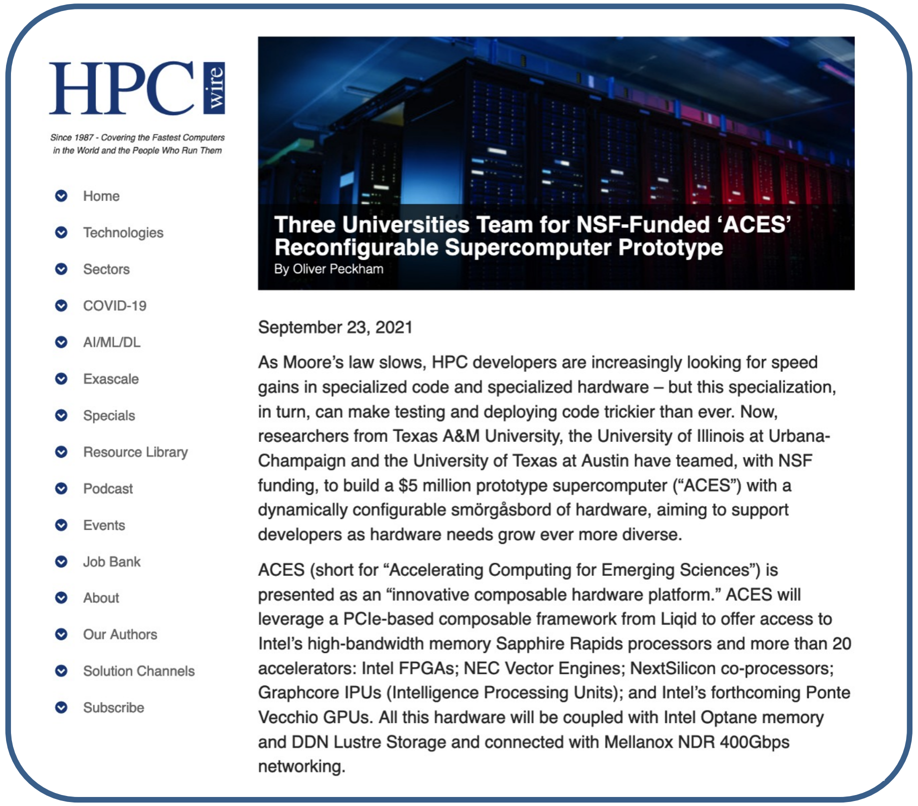
PEARC22 BOF

July 12, 2022

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

High Performance Research Computing
DIVISION OF RESEARCH

TACC
TEXAS ADVANCED COMPUTING CENTER



HPCwire

Since 1987 - Covering the Fastest Computers in the World and the People Who Run Them

- Home
- Technologies
- Sectors
- COVID-19
- AI/ML/DL
- Exascale
- Specials
- Resource Library
- Podcast
- Events
- Job Bank
- About
- Our Authors
- Solution Channels
- Subscribe

Three Universities Team for NSF-Funded 'ACES' Reconfigurable Supercomputer Prototype

By Oliver Peckham

September 23, 2021

As Moore's law slows, HPC developers are increasingly looking for speed gains in specialized code and specialized hardware – but this specialization, in turn, can make testing and deploying code trickier than ever. Now, researchers from Texas A&M University, the University of Illinois at Urbana-Champaign and the University of Texas at Austin have teamed, with NSF funding, to build a $5 million prototype supercomputer ("ACES") with a dynamically configurable smörgåsbord of hardware, aiming to support developers as hardware needs grow ever more diverse.

ACES (short for "Accelerating Computing for Emerging Sciences") is presented as an "innovative composable hardware platform." ACES will leverage a PCIe-based composable framework from Liqid to offer access to Intel's high-bandwidth memory Sapphire Rapids processors and more than 20 accelerators: Intel FPGAs; NEC Vector Engines; NextSilicon co-processors; Graphcore IPUs (Intelligence Processing Units); and Intel's forthcoming Ponte Vecchio GPUs. All this hardware will be coupled with Intel Optane memory and DDN Lustre Storage and connected with Mellanox NDR 400Gbps networking.

ACES - Accelerating Computing for Emerging Sciences

ACES
ACCELERATING COMPUTING FOR EMERGING SCIENCES

"ACES will enable applications and workflows to dynamically integrate the different accelerators, memory, and in-network computing protocols to glean new insights by rapidly processing large volumes of data," the NSF grant reads, "and provide researchers with a unique platform to produce complex hybrid programming models that effectively supports calculations that were not feasible before."

https://www.hpcwire.com/2021/09/23/three-universities-team-for-nsf-funded-aces-reconfigurable-supercomputer-prototype/

ACES

Accelerating Computing for Emerging Sciences

Our Mission:

- Offer an accelerator testbed for numerical simulations and **AI/ML workloads**
- Provide consulting, technical guidance, and training to researchers
- Collaborate on computational and data-enabled research

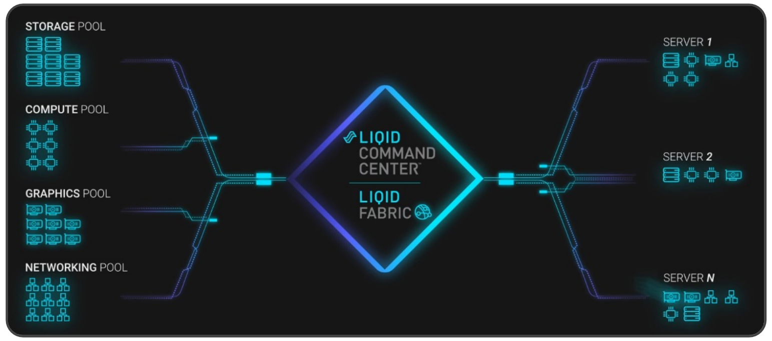
High Performance Computing (HPC) Architecture Comparison

Legacy HPC

- Built on Converged HW
- Static Hardware Design
- Fixed GPUs/Accelerators
- Fixed Memory
- Legacy Storage: SATA and SAS

Modern HPC

- Built on Disaggregated HW
- Composable Hardware Platform
- Composable GPUs/Accelerators
- Composable Memory - Optane
- Modern Storage: NVMe-oF



A Modern HPC Platform Supporting Composable GPUs/Accelerators and Memory

ACES - Accelerating Computing for Emerging Sciences (Phase I)

Component	Quantity	Description
Graphcore IPU	16	16 Colossus GC200 IPUs and dual AMD Rome CPU server on a 100 GbE RoCE fabric
Intel FPGA PAC D5005	2	FPGA SOC with Intel Stratix 10 SX FPGAs, 64 bit quad-core Arm Cortex-A53 processors, and 32GB DDR4
Intel Optane SSDs	8	3 TB of Intel Optane SSDs addressable as memory using MemVerge Memory Machine.

Available through FASTER (NSF Award #2019129)

ACES - Accelerating Computing for Emerging Sciences (Phase II)

Component	Quantity*	Description
Graphcore IPU	32	16 Colossus GC200 IPUs, 16 Bow IPUs, and a dual AMD Rome CPU server on a 100 GbE RoCE fabric
Intel FPGA PAC D5005	2	FPGA SOC with Intel Stratix 10 SX FPGAs, 64 bit quad core Arm Cortex-A53 processors, and 32GB DDR4
Bittware IA-840F FPGA	2	Accelerator based on Intel Agilex FPGA
NextSilicon coprocessor	20	Reconfigurable accelerator with an optimizer continuously evaluating application behavior.
NEC Vector Engine	24	Vector computing card (8 cores and HBM2 memory)
Intel Ponte Vecchio GPU	100	Intel GPUs for HPC, DL Training, AI Inference
Intel Optane SSDs	48	18 TB of Intel Optane SSDs addressable as memory w/ MemVerge Memory Machine.

****Estimated quantities***

ACES System Description (Phase II)

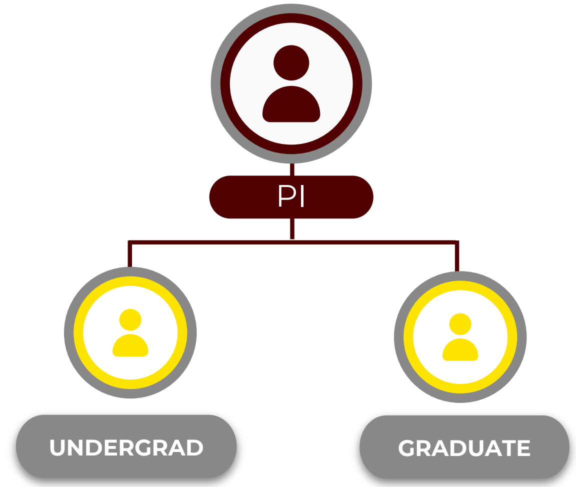
Component	Quantity	Description
Allocatable resources		Total cores: 11,520
CPU-centric computing with variable memory requirements	120 nodes (11,520 cores)	Dual Intel Sapphire Rapids 2.1 GHz 48 core processors 96 cores per node, 512 GB memory, 1.6 TB NVMe storage (PCIe 5.0), NVIDIA Mellanox NDR 200 Gbps InfiniBand
Composable infrastructure	120 nodes	Dynamically reconfigurable infrastructure that allows up to 20 PCIe cards (GPU, FPGA, VE, etc.) per compute node
Data transfer nodes	2 nodes	Same as compute nodes, 100 Gbps network adapter

Research Workflows - Accelerators (Phases I and II)

Hardware Profile	Applications Supported	
NEC Vector Engines	• AI/ML (Statistical Machine Learning, Data Frame) • Chemistry (VASP, Quantum ESPRESSO) • Earth Sciences • NumPy Acceleration	• Oil & Gas (Seismic Imaging, Reservoir Simulation) • Plasma Simulation • Weather/Climate Simulation
Graphcore IPUs	• Graph Data • LSTM Neural Networks	• Markov Chain Monte Carlo • Natural Language Processing (Deep Learning)
Intel/Bittware FPGA	• AI Models for Embedded Use Cases • Big Data • CXL Memory Interface • Deep Learning Inference • Genomics	• MD Codes • Microcontroller Emulation for Autonomy Simulations • Streaming Data Analysis
Intel Optane SSDs	• Bioinformatics • Computational Fluid Dynamics (OpenFOAM)	• MD Codes • R • WRF
NextSilicon	• Biosciences (BLAST) • Computational Fluid Dynamics (OpenFOAM) • Cosmology (HACC) • Graph Search (Pathfinder)	• Molecular Dynamics (NAMD, AMBER, LAMMPS) • Quantum ChromoDynamics (MILC) • Weather/Environment modeling (WRF)

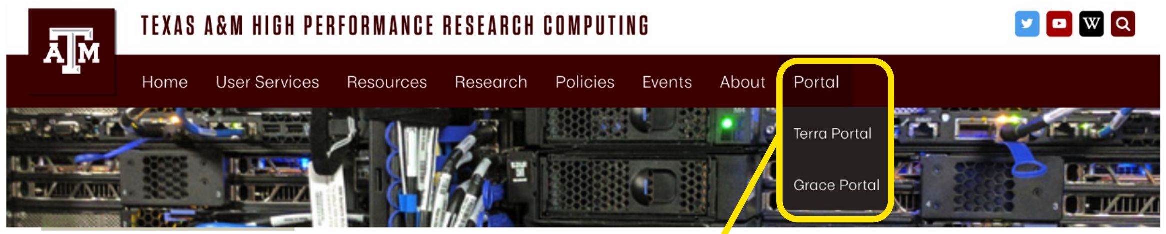
Getting on ACES Phase I

- Allocation is upon special request during this phase of deployment.
- You must have an XSEDE account!
- Applications are available at hprc.tamu.edu/aces/
- Email us at help@hprc.tamu.edu for questions, comments, and concerns



PIs can apply for an account and sponsor accounts for their researchers.

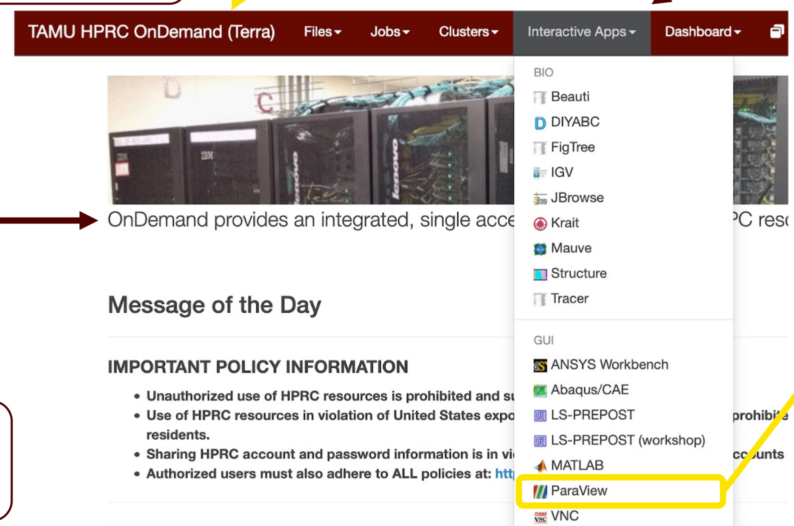
ACES Portal - OOD

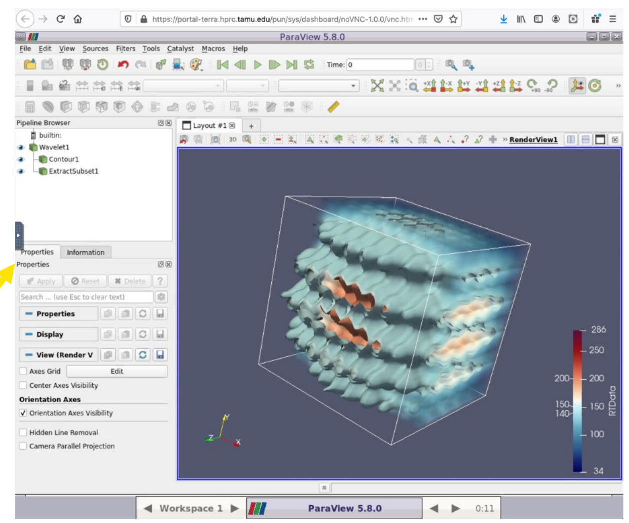


https://portal.hprc.tamu.edu

Interactive Apps: launch a software window right in your browser.

Open OnDemand (OOD) Portal is an advanced web-based graphical interface for HPC users.





HPRC Portal YouTube tutorials

Knowledge Base

Graphcore IPUs [edit]

From one of FASTER login nodes, ssh into poplar1 system.

```
[username@faster1 ~]$ ssh poplar1
```

1.Set up the Poplar SDK environment [edit]

In this step, set up several environment variables to use the Graphcore tools and Poplar graph programming framework.

```
[username@poplar1 ~]$ source /opt/gc/poplar/poplar_sdk-ubuntu_18_04-[ver]/poplar-ubuntu_18_04-[ver]/enable.sh
[username@poplar1 ~]$ source /opt/gc/poplar/poplar_sdk-ubuntu_18_04-[ver]/popart-ubuntu_18_04-[ver]/enable.sh
```

[ver] indicates the version number of the package.

hprc.tamu.edu/wiki/ACES

ACES Project Team

Project Management Board:

- Honggao Liu (PI, Texas A&M University)
- Lisa Perez (Co-PI, Texas A&M University)
- Dhruva Chakravorty (Co-PI, Texas A&M University)
- Shaowen Wang (Co-PI, University of Illinois Urbana-Champaign)
- Timothy Cockerill (Co-PI, University of Texas Austin)
- Francis Dang (SI, Texas A&M University)
- Costas Georghiades (SI, Texas A&M University)
- Edwin Pierson (SI, Texas A&M University)

Advisory Council:
Gabrielle Allen (U. Wyoming), Richard Gerber (NERSC), John Goodhue (MGHPCC), Dan Katz(NCSA), Victor Hazlewood (U. Tennessee), Anita Nikolich (UIUC), Barry Schneider (NIST), Carol Song (Purdue), and Dan Stanzione (TACC)

NSF Program Officer: Robert Chadduck (NSF OAC Program Director)

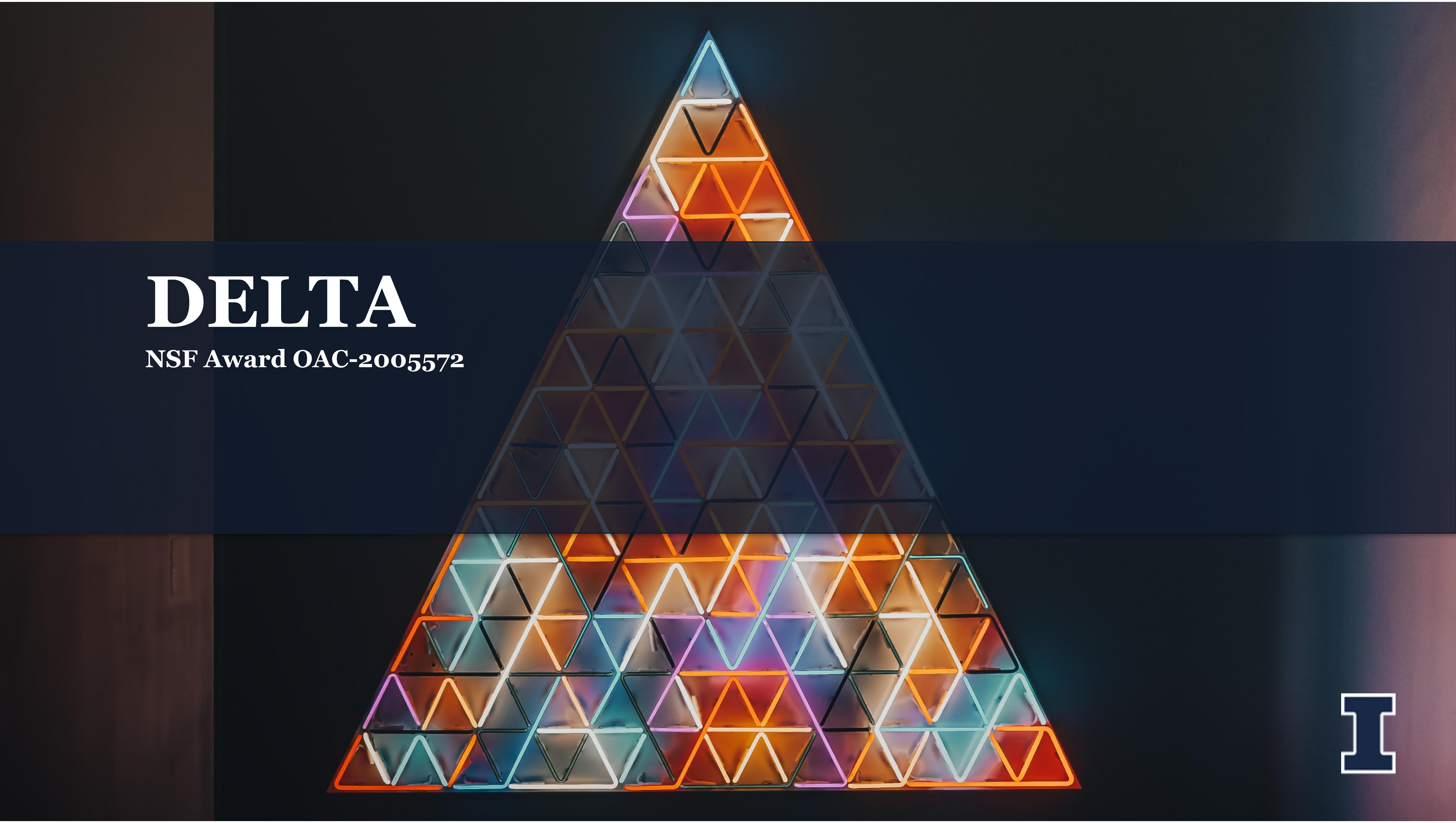
ACES

ACCELERATING COMPUTING FOR EMERGING SCIENCES

hprc.tamu.edu

HPRC Helpdesk:

help@hprc.tamu.edu
Phone: 979-845-0219



DELTA

NSF Award OAC-2005572

COMPUTE

Furthering GPU Adoption

DATA

Moving Beyond POSIX

INTERFACE

Improving Usability and Accessibility

Hardware Overview

RESOURCE COUNTS

CPU NODES	124 x CPU Compute Nodes 8 x CPU Utility Nodes
GPU NODES	100 x 64-Bit GPU 100 x 32-Bit GPU 5 + 1 x Dense GPU & High Mem
STORAGE	7 PB HDD (Lustre) 3 PB SSD (non-POSIX)

SYSTEM TOTALS

CPUs	476 x AMD EPYC 7763 64 core “Milan”
GPUs	440 x NVIDIA A100 400 x NVIDIA A40 8 x AMD MI100
PERF	10 PF double-precision 100 PF single-precision 200 PF tensor

Hewlett Packard Enterprise ddn NVIDIA AMD

BoF: NSF Innovative Computing Technology Testbed Community Exchange

Overview of the Habana Processor-based Voyager Project : Exploring Neural Network Processors for AI in Science and Engineering

PEARC2022

Amit Majumdar
Principal Investigator
SDSC Division Director

Co-PIs: Rommie Amaro, Javier Duarte, Mai Nguyen, Robert Sinkovits, UCSD
Project Manager: Shawn Strande, UCSD

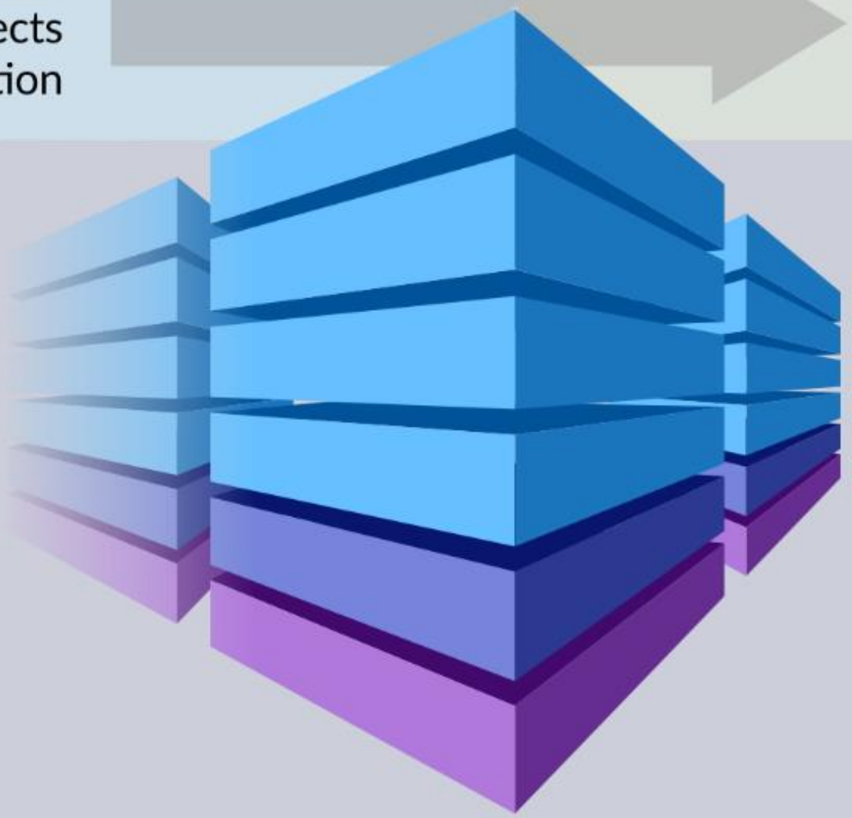
NSF Award 2005369

VOYAGER

EXPLORING AI PROCESSORS in SCIENCE and ENGINEERING

3-YEAR TESTBED PHASE
Focused Select Projects
Workshops, Industry Interaction

2-YEAR ALLOCATIONS PHASE
NSF Allocations to the Broader Community
User Workshops



INNOVATIVE AI RESOURCE
Specialized Training Processors
Specialized Inference Processors
High-Performance Interconnect
X86 Standard Compute nodes
Rich Storage Hierarchy

OPTIMIZED AI SOFTWARE
Community Frameworks
Custom user-developed AI Applications
PyTorch, Tensorflow

IMPACT & ENGAGEMENT
Large-Scale Models
AI Architecture Advancement
Improved Performance of AI Applications
External Advisory Board of AI & HPC Experts
Wide Science & Engineering Community
Advanced Project Support & Training
Accelerating Scientific Discovery
Industrial Engagement

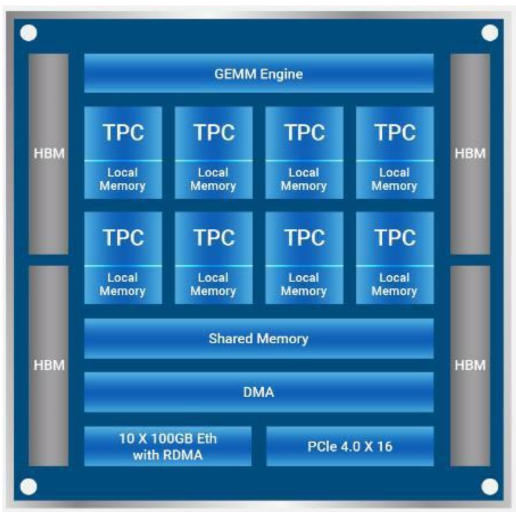
Voyager hardware and software innovations will lead to performance gains and ease model development

Gaudi training processor

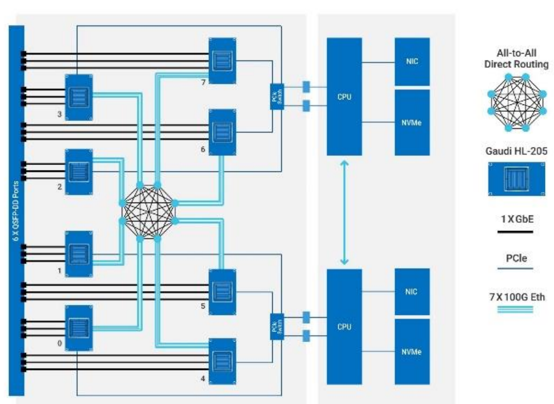
- Fully-programmable Tensor Processing Cores (TPC)
- Integrated 10x 100GbE RoCE ports on each chip
- Standard Ethernet within and across nodes
- Non-blocking connectivity within nodes
- HBM2 similar to GPU memory, so existing models fit within memory
- 24x100GbE for scale out

Synapse AI Library

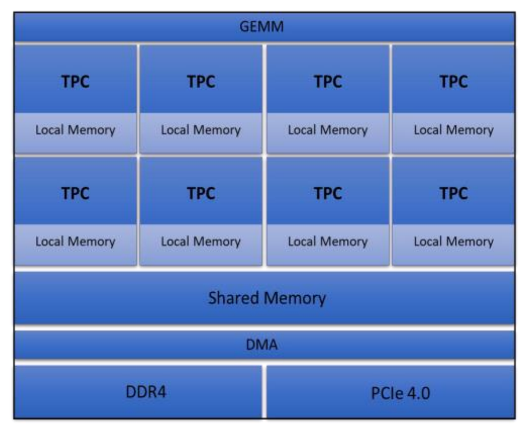
- Integrated with TensorFlow and PyTorch
- Minimal code changes to get started
- Maps model topology onto Gaudi devices
- Facilitates development of custom kernels
- Habana Developer Site & GitHub



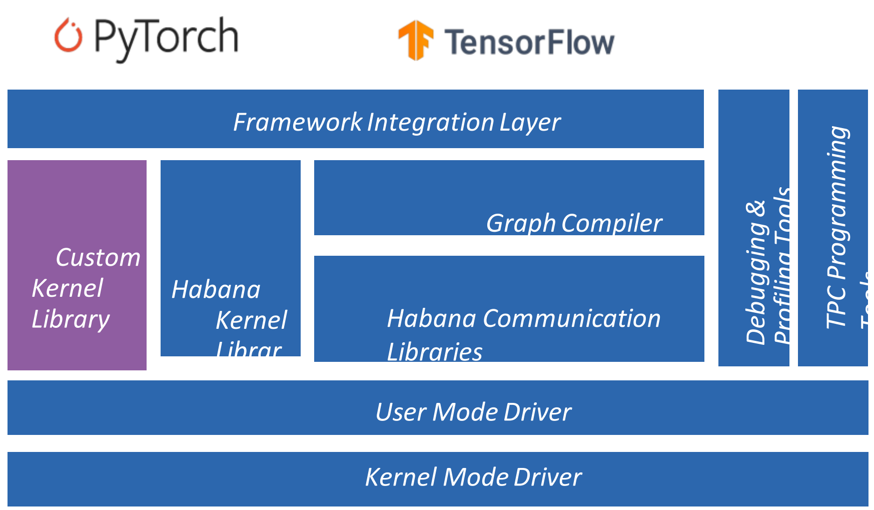
Gaudi training processor



Gaudi networking



Goya inference processor



Synapse AI Library

Example Early Users Applications on *Voyager*

Project	Model
Data Driven Weather Prediction	U-Net
High energy physics	GNN
Cardiac image analysis	U-Net
Ultrasound computed tomography for structural, medical and material applications	U-Net
Systems biology	Dense neural network with learned embeddings
Atmospheric sciences	VAE model
Human microbiome research	Categorical VAE
Astronomy	NN
Cognitive Neuroscience	CNN
Natural language processing	Transformers
Biochemistry – Molecular Dynamics	VAE, AAE, ANCA-AE

***Key**: CNN=Convolutional Neural Network, GAN=Generative Adversarial Network, GNN=Graph Neural Network; NN=Dense Neural Network, VAE=Variational Auto-Encoder, AAE=Adversarial Auto-Encoder, ANCA-AE=Anharmonic Conformational Analysis enabled Autoencoders.*

Voyager as a Testbed system for AI Research

- *Voyager's* AI-focused architecture will be an excellent resource for exploring AI for science and engineering
- The system was designed to support exploration in multiple dimensions (Gaudi and Goya processors, RoCE interconnect, 400 GbE switch, Kubernetes, Ceph, cnvrg.io, Slurm, et al)
- Excellent collaboration with our technology partners at Supermicro and Habana allowed us to bring this innovative architecture to the community
- Considering the amount of innovation, the acquisition, deployment and installation was remarkably smooth with issues resolved jointly with Supermicro and Habana experts
- In most cases, measured performance is better than projected, due in large part due to software improvements by Habana and the collaboration
- Excellent 3-way partnership between SDSC, Habana, and the research community to gain a deep understanding of how AI/ML applications perform and what work is needed to optimize them for these kind of specialized architectures
- We are gaining important knowledge and lessons in areas such as Kubernetes, AI model porting and performance, and scaling
- There are new opportunities to engage the broader community (academia, national labs, industry) via training and workshops
- Completed the NSF review, and Voyager is in Testbed phase since May 2022
- Please email Amit Majumdar (majumdar@sdsc.edu) if you are interested in gaining access

TEXAS ADVANCED COMPUTING CENTER

CHAMELEON

July 12, 2022

CHAMELEON – A NON-OAC TESTBED

- Questions for the panelists to address during their presentation:
- 1) Introduce yourself and the testbed project you are representing including its motivating technology opportunities and challenges.
- 2) Which disciplines/software/research should be especially interested in your system, and why?
- 3) What are the major findings/outcomes/impacts so far from your project?
- 4) What does the future hold for this technology and its science impact? Should we be excited? How to realize this future?

CHAMELEON IN A NUTSHELL

- Chameleons like to change: a testbed that adapts itself to your experimental needs
  - **Bare metal reconfigurability/isolation** + (alternative) KVM cloud (different cost/isolation trade-off)
  - Capabilities: power on/off, reboot, custom kernel, serial console access, etc.
- Balance: diversity and scale – from large to small
  - **Large to small:** from **supercomputing datacenters** (UC, TACC) with 100G network to **edge devices**
  - **Diverse:** FPGAs, GPUs, NVMe, NVDIMMs, Corsa switches, edge devices via CHI@Edge, etc.
  - **Distributed:** CHI-in-a-Box sites at **NCAR, Northwestern, and UIC**
- Cloud++: CHameleon Infrastructure (CHI) via mainstream cloud tech
  - Powered by **OpenStack** with bare metal reconfiguration (Ironic) + "special sauce" (50/50 split)
  - Blazar contribution recognized as official OpenStack component
- Reproducibility, repeatability, and sharing
  - **Jupyter integration** for imperative and non-transactional experiment packaging, **Chameleon daypass** for easy access, **Trovi** for sharing and finding experiments, integration with **Zenodo** for publishing

OPEN TESTBED – BY THE NUMBERS

450+
Papers published

45
Countries

6,000+
Users

800+
Unique projects

~3 more years to grow!

CHAMELEON HARDWARE

CHI@UC
Skylake, CascadeLake, IceLake, and AMD nodes
GPUs (A100, V100, RTX6000), FPGAs (Xilinx Alveo U280)
Storage (NVMe SSD, NVDIMM)
Network (Corsa SDN, 25G Ethernet, 200G InfiniBand)

CHI@Edge
Raspberry Pi, Jetson Nano, Jetson Xavier NX

CHI@TACC
Haswell, SkyLake, CascadeLake, and AMD nodes
GPU (K80, M40, P100, MI100), FPGAs (Altera)
Storage (NVMe, SSD, NVDIMM)
Network (Corsa SDN, 25G Ethernet, 200G InfiniBand)

Storage
~0.5 PB

Network
100Gbps public network

Storage
~3.5PB

FABRIC
and other partners

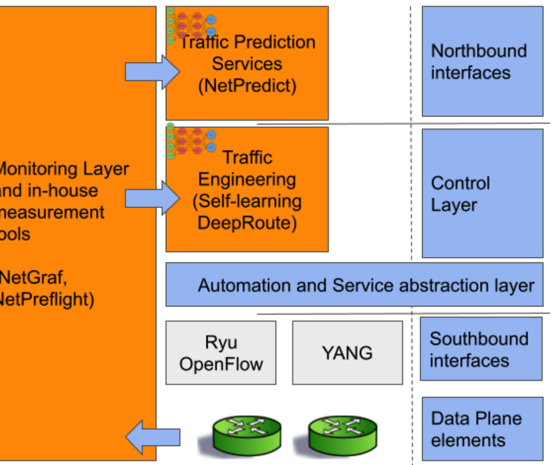
Commercial Clouds
via CloudBank

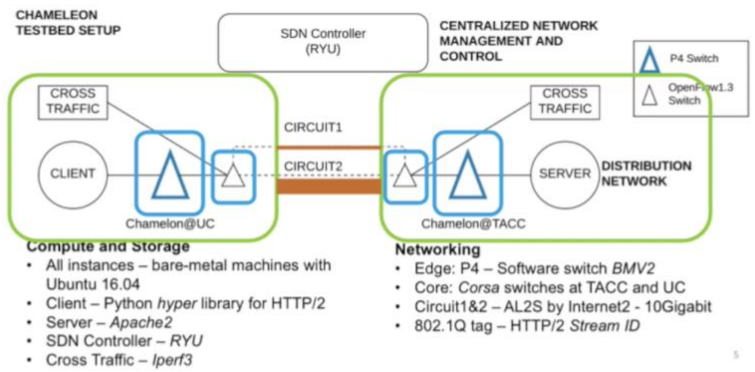
Chameleon
Associate Sites:
NCAR,
Northwestern.
UIC
and other partners

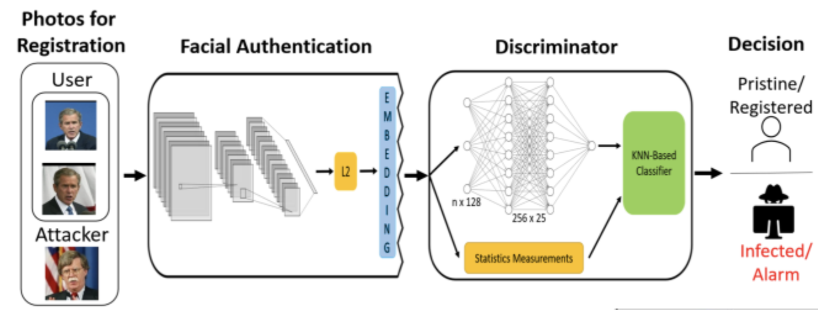
Chicago

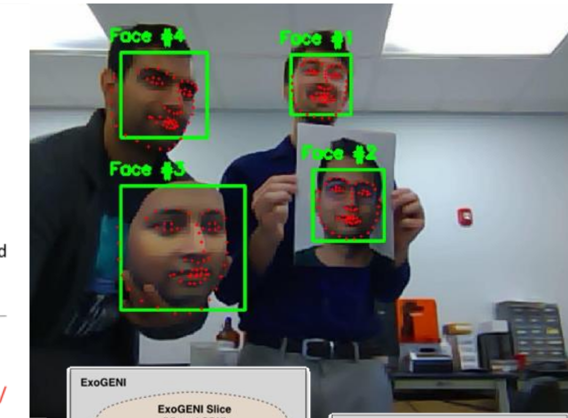
Austin

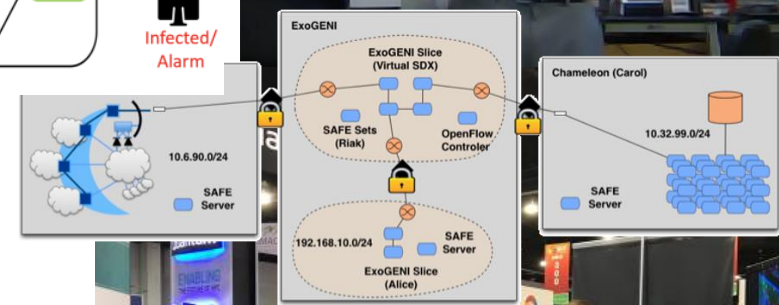
LEAVING NO EXPERIMENT BEHIND



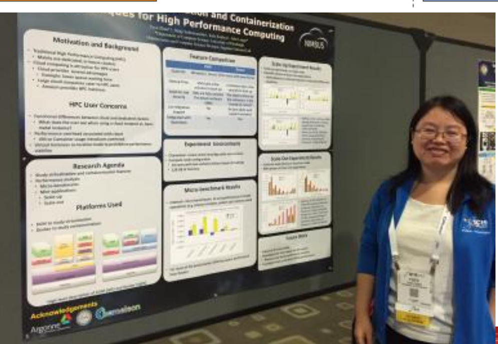


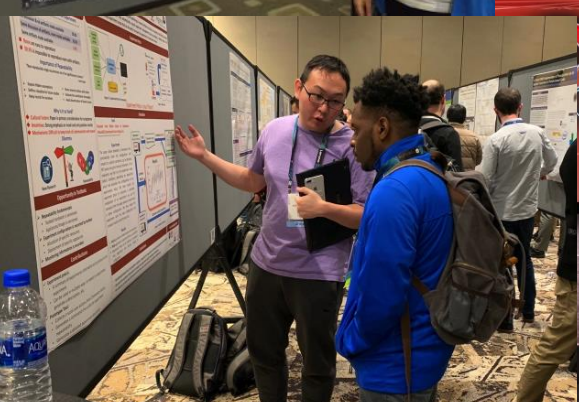


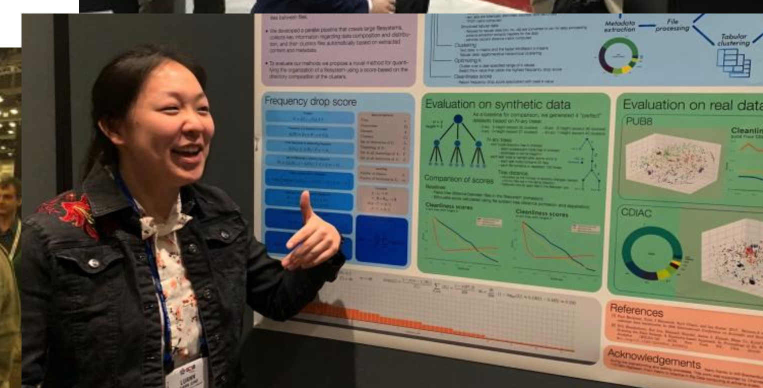




Supporting research projects in architecture, operating systems design, virtualization, power management, real-time analysis, security, storage systems, databases, networking, machine learning, neural networks, data science, and many others.







Chameleon www. chameleoncloud.org

Check out user experiment stories on our blog:
https://www.chameleoncloud.org/blog/category/user-experiment

CHAMELEON KEEPS EVOLVING!

- New hardware is being provided both by the Chameleon grant (e.g. Liqid composable hardware) and for TACC test hardware to find a home (e.g. AMD GPUs).

- Chameleon is a place to experiment – try out new hardware, new software, new ideas!

- Allocations are managed by the project – you can even get one day tickets they day you apply. Get access to Chameleon at www.chameleoncloud.org

NSF Testbed Project Update: Neocortex

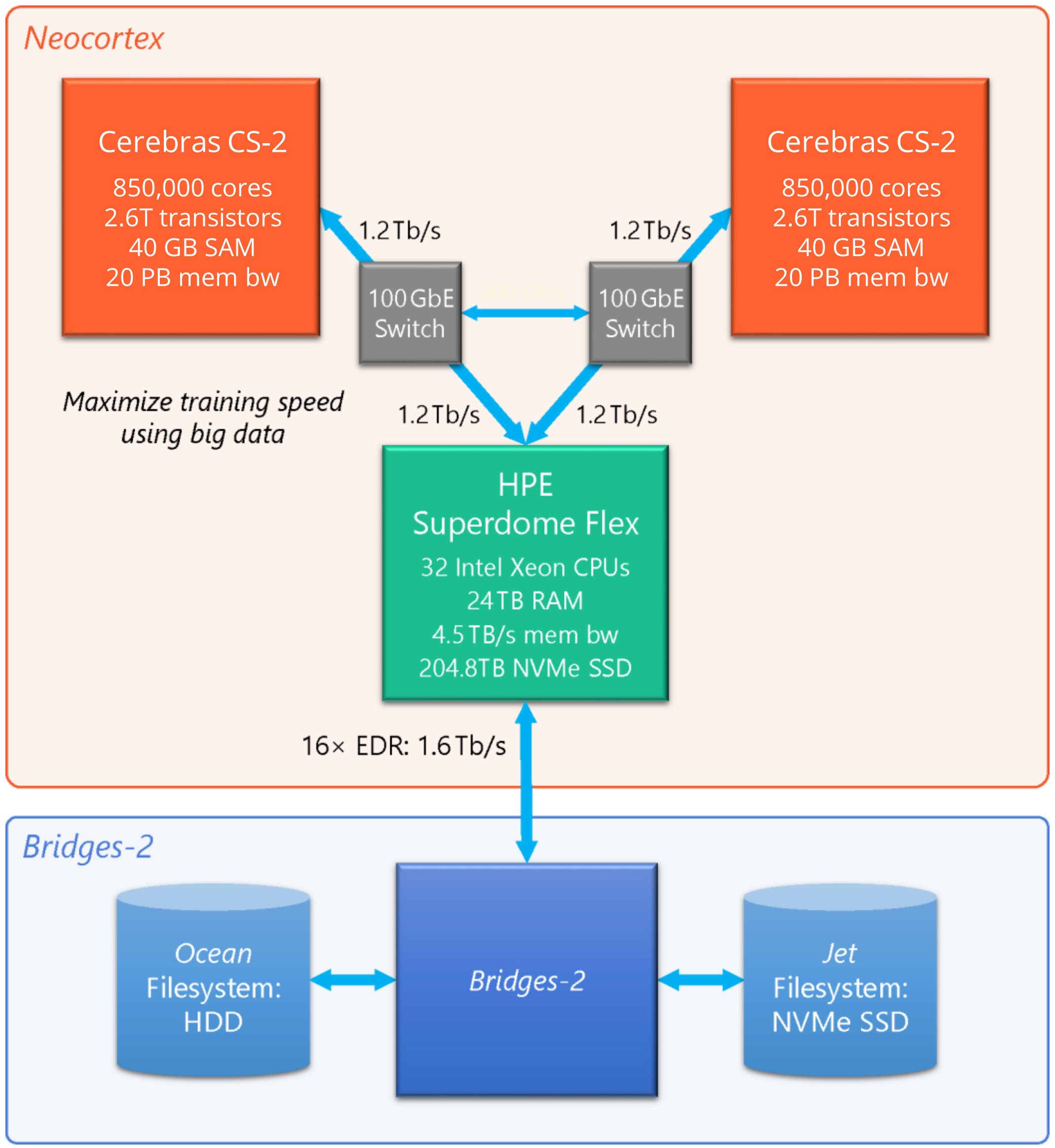
Sergiu Sanielevici
Neocortex co-Principal Investigator & AD for User Support
Director, User Support for Scientific Applications, Pittsburgh Supercomputing Center

7/12/2022

The Neocortex System

NEOCORTEX

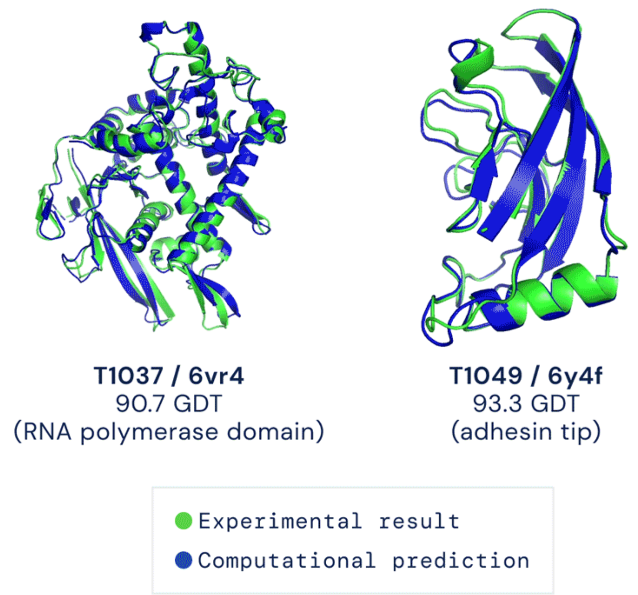
Neocortex System Overview



NEOCORTEX

Driving Use-Cases

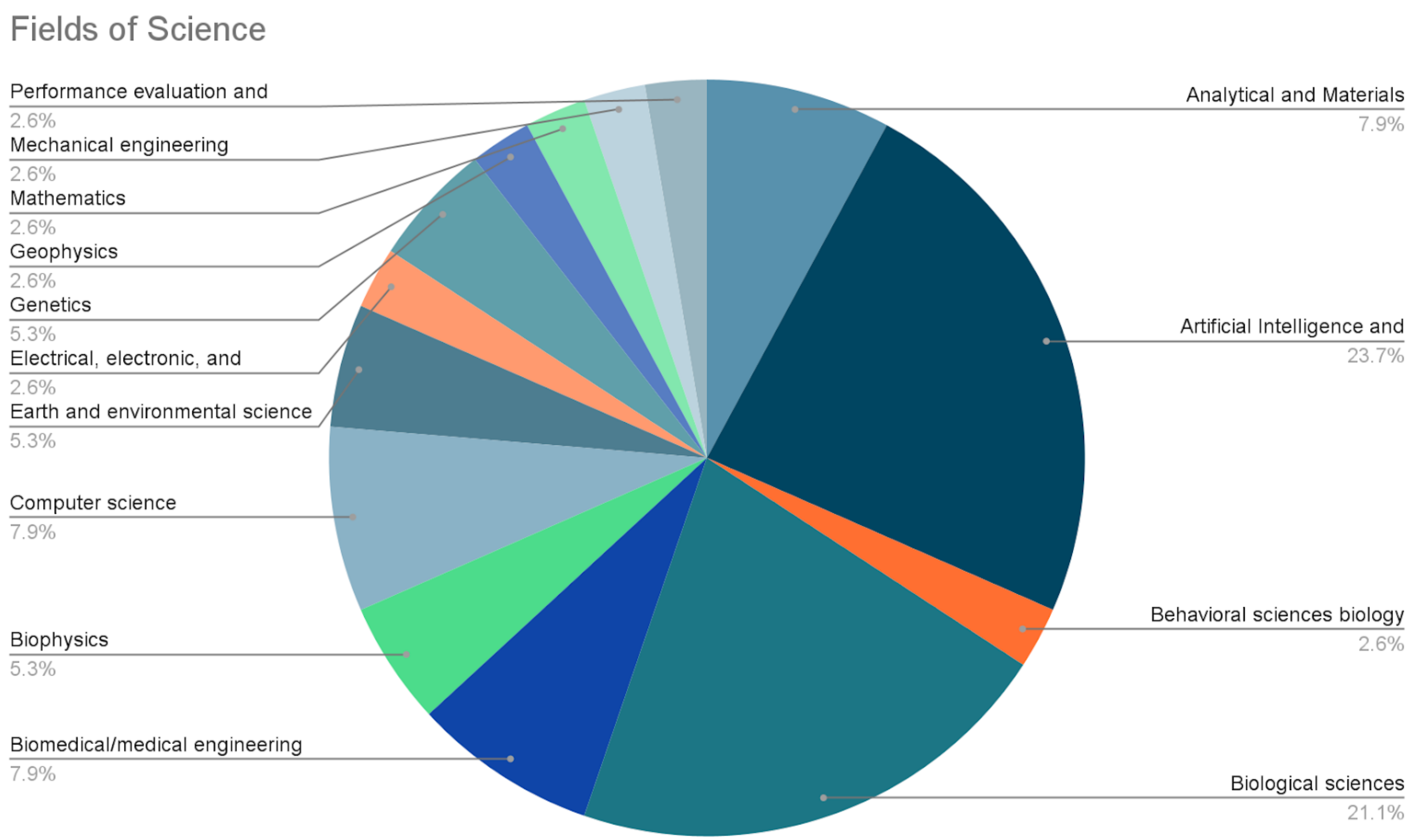
- Transform and accelerate AI-enabled research
- Development of new and more efficient AI algorithms and graph analytics
- Foster greater integration of artificial deep learning with scientific workflows
- Democratize access to game changing compute power
- Explore the potential of a groundbreaking new hardware architecture
- Support research needing large-scale memory (genomics, brain imaging, simulation modeling)
- Augmenting traditional computational science with rapidly-evolving methodologies and technologies
- User-centric and interactive computing modalities



Animation from https://deepmind.com/blog/article/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology. Retrieved on August 2021.

NEOCORTEX

Fields of Science for 38 Projects Active on Neocortex, May 2022



NEOCORTEX

Summer 2022 Call for Proposals

- All details available in the official webpage: https://www.cmu.edu/psc/aibd/neocortex/2022-06-cfp-summer-2022.html

Neocortex Summer 2022 Allocation Submissions	
Name	**Date (ET)**
Application begins	June 8, 2022
Application ends	July 15, 2022
Response ends	August 15, 2022

To Learn More and Participate

Watch the Neocortex website for updates!	https://www.cmu.edu/psc/aibd/neocortex/
Information about active projects	Neocortex Active Projects - HPC AI and Big Data Group - Pittsburgh Supercomputing Center - Carnegie Mellon University (cmu.edu)
Summer 2022 CFP Information	https://www.cmu.edu/psc/aibd/neocortex/2022-06-cfp-summer-2022.html
Contact us with questions, input, or requests	neocortex@psc.edu

NEOCORTEX

PERVASIVE TECHNOLOGY INSTITUTE

RESEARCH TECHNOLOGIES
UNIVERSITY INFORMATION TECHNOLOGY SERVICES

Jetstream2: Accelerating science and engineering on-demand

David Y. Hancock – Indiana University

Director for Advanced Cyberinfrastructure

Jetstream2 Primary Investigator

NSF innovative computing technology testbed community exchange

PEARC22 – 12 July 2022

Jetstream2

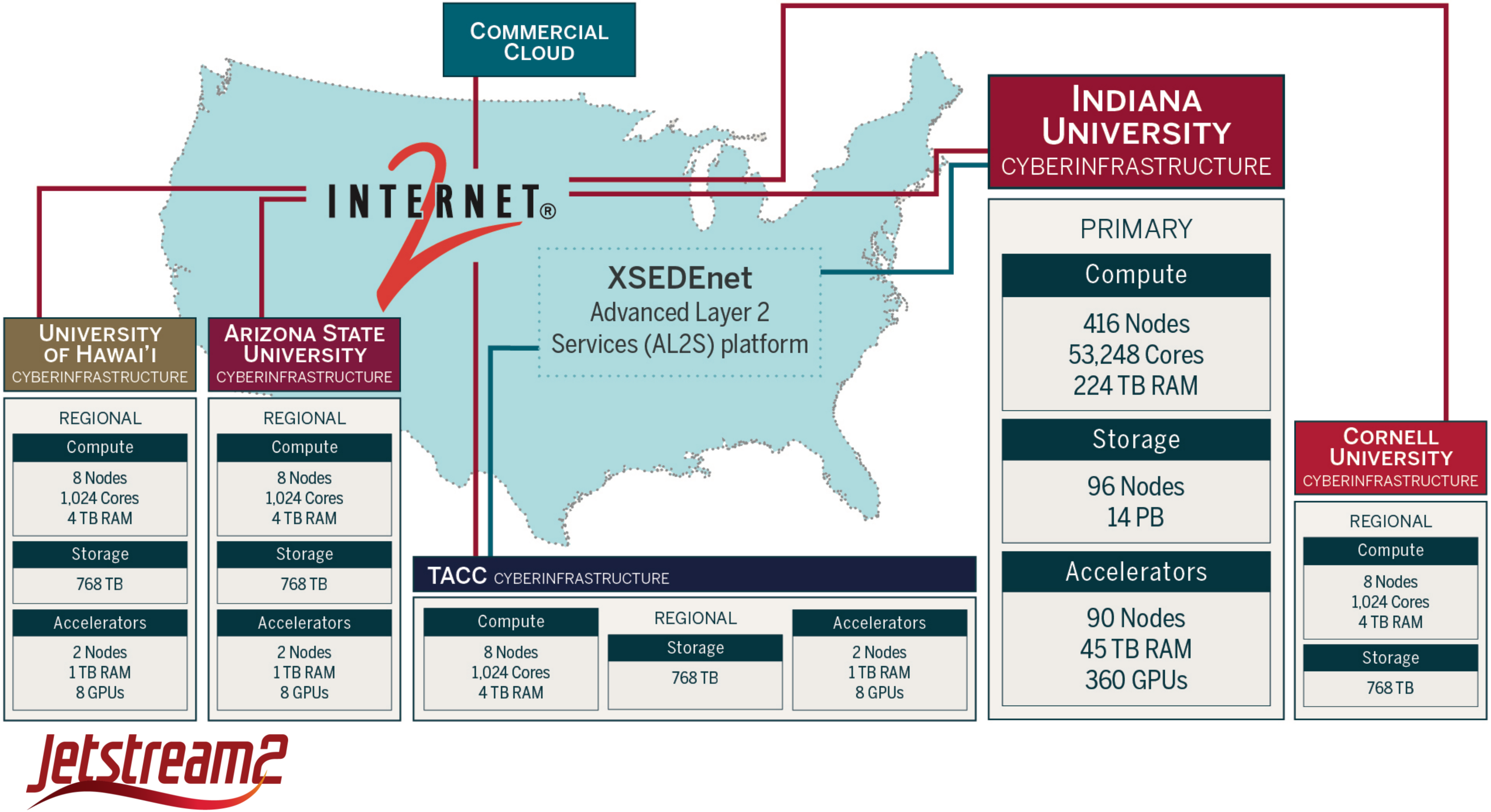
Is Jetstream2 a “testbed?”

- No, but it’s a great place to test!
- Ease-of-use focus, rapid on-ramp to XSEDE/ACCESS
- On-demand interactive computing and persistent services for science gateways
- Enables configurable environments; programmable cyberinfrastructure

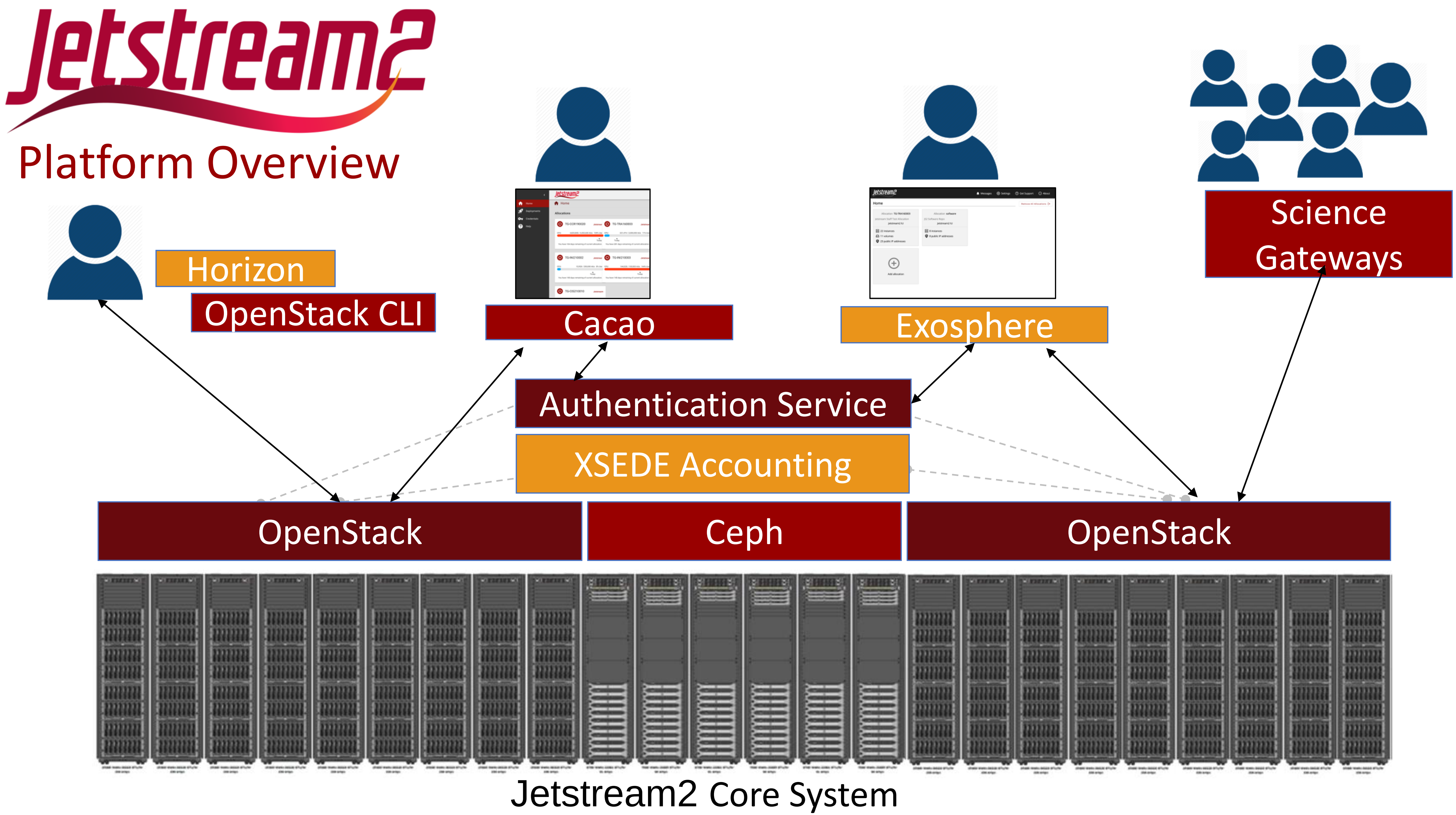
By Maria Morris: JS2 rear doors (lower) Banksy adaptation [non-commercial] (right)

Now with GPUs,
large-memory,
more faster PB!

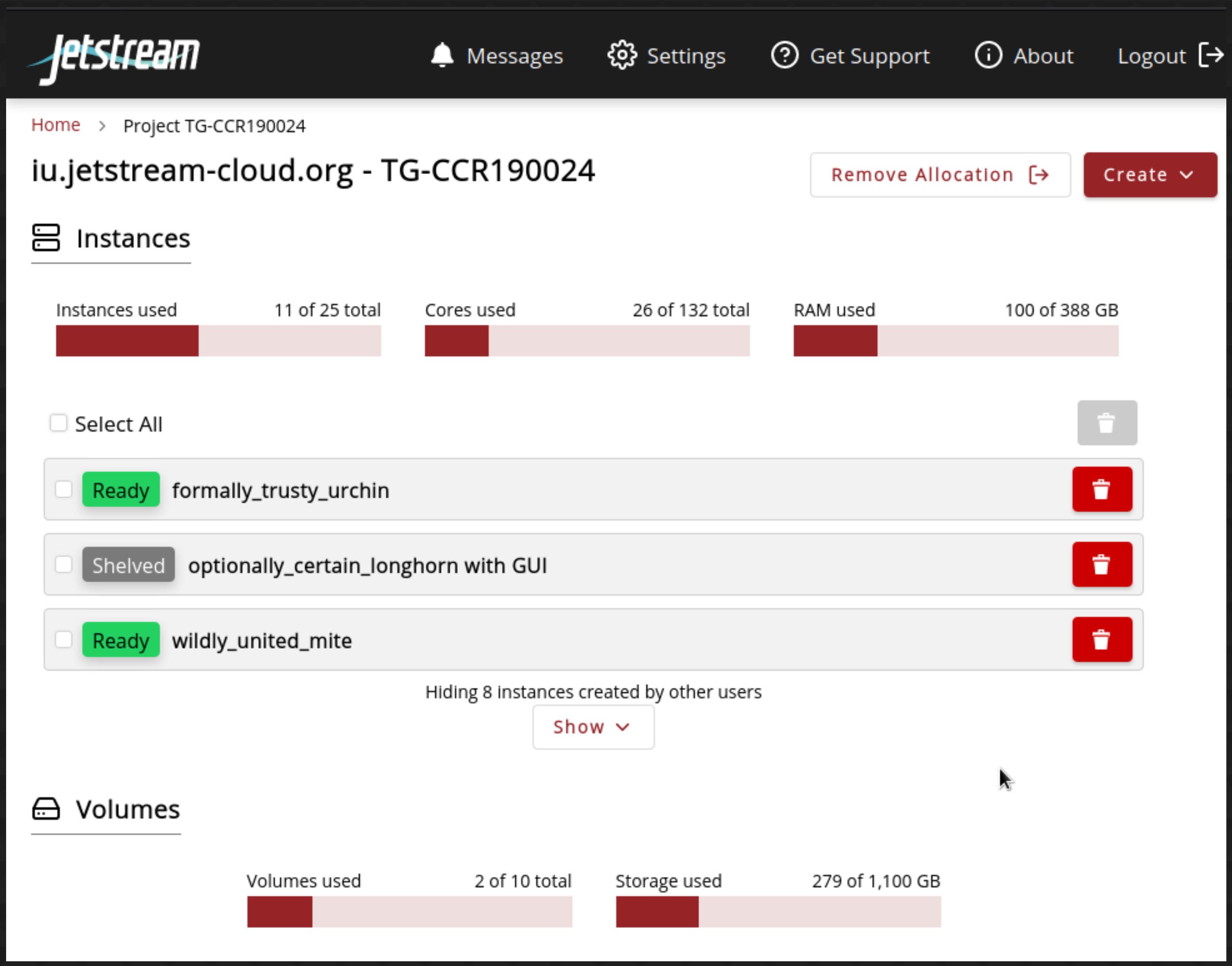
Jetstream2

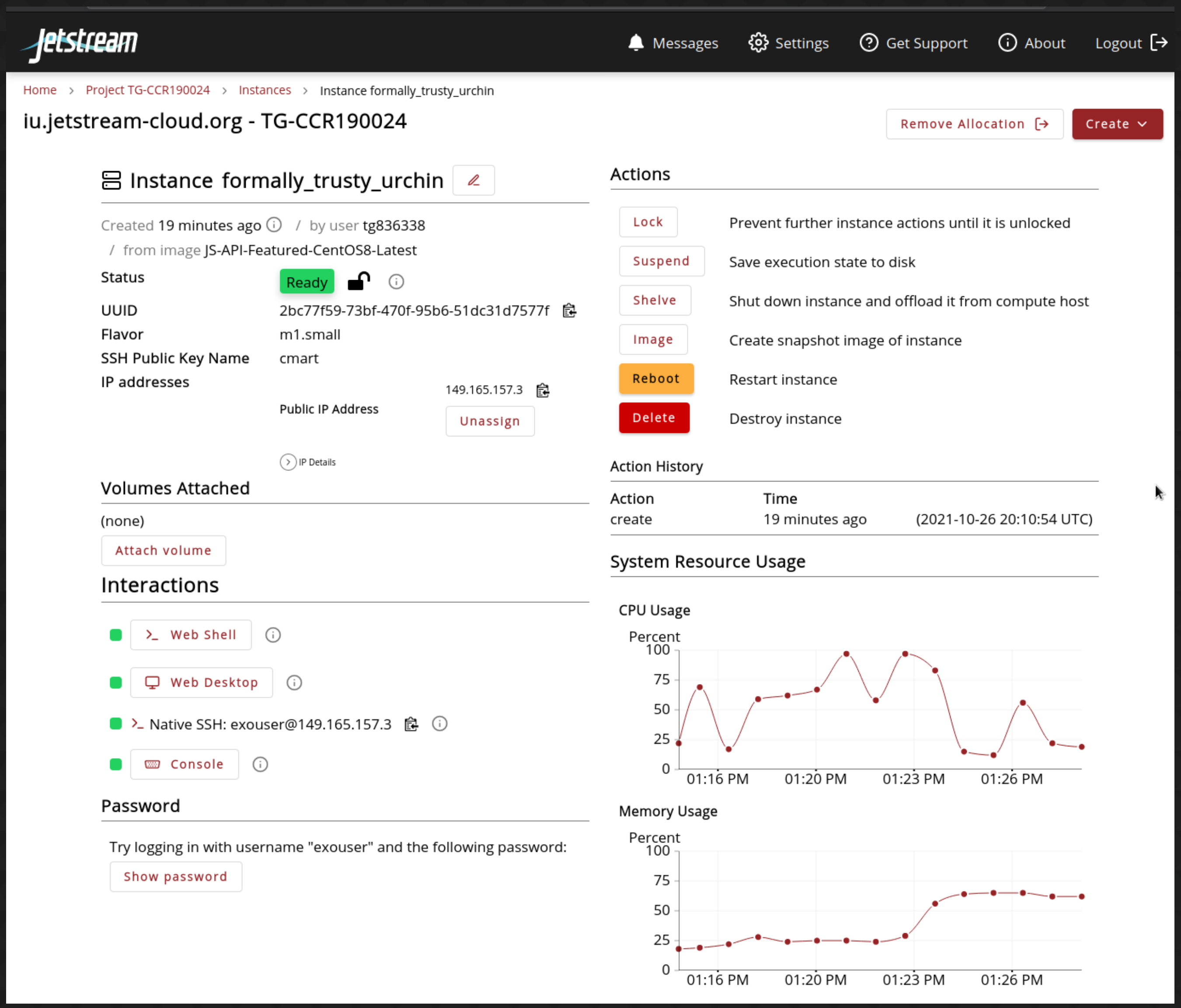


COMMERCIAL CLOUD
INTERNET2
XSEDEnet
Advanced Layer 2 Services (AL2S) platform
INDIANA UNIVERSITY
CYBERINFRASTRUCTURE
PRIMARY
Compute
416 Nodes
53,248 Cores
224 TB RAM
Storage
96 Nodes
14 PB
Accelerators
90 Nodes
45 TB RAM
360 GPUs
UNIVERSITY OF HAWAI'I
CYBERINFRASTRUCTURE
REGIONAL
Compute
8 Nodes
1,024 Cores
4 TB RAM
Storage
768 TB
Accelerators
2 Nodes
1 TB RAM
8 GPUs
ARIZONA STATE UNIVERSITY
CYBERINFRASTRUCTURE
REGIONAL
Compute
8 Nodes
1,024 Cores
4 TB RAM
Storage
768 TB
Accelerators
2 Nodes
1 TB RAM
8 GPUs
TACC CYBERINFRASTRUCTURE
REGIONAL
Compute
8 Nodes
1,024 Cores
4 TB RAM
Storage
768 TB
Accelerators
2 Nodes
1 TB RAM
8 GPUs
CORNELL UNIVERSITY
CYBERINFRASTRUCTURE
REGIONAL
Compute
8 Nodes
1,024 Cores
4 TB RAM
Storage
768 TB
Jetstream2


Jetstream2
Platform Overview
Horizon
OpenStack CLI
Cacao
Exosphere
Science Gateways
Authentication Service
XSEDE Accounting
OpenStack
Ceph
OpenStack
Jetstream2 Core System

Exosphere





Jetstream2

https://exosphere.Jetstream-cloud.org or try.exosphere.app

Early Operations Projects & Activity

- First PI invitations and projects added in early February 2022
- Followed push for all projects to migrate in May – July.
- June 2022: 245 projects and 1,144 individuals (391 students)
- Includes multiple science gateways and education/training allocations
- Entering "production" upon NSF approval
- Goal to retire Jetstream[1] this month

"Bike Exchange - 2009 IU Women's Little 500" by Indiana Public Media

Jetstream2

PERVASIVE TECHNOLOGY INSTITUTE

RESEARCH TECHNOLOGIES
UNIVERSITY INFORMATION TECHNOLOGY SERVICES

Acknowledgements

NSF Awards 1053575 & 1548562 (XSEDE), 1445604 (Jetstream) and 2005506 (Jetstream2)

This document was developed with support from the National Science Foundation. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

Special thanks to contributors & Jetstream2 partners

- Jeremy Fischer, J. Michael Lowe, Therese Miller, Maria Morris, Winona Snapp-Childs, George Turner, and Chris Martin.
- Vendors, particularly Dell and NVIDIA, also deserve recognition for their efforts

PERVASIVE TECHNOLOGY INSTITUTE

RESEARCH TECHNOLOGIES
UNIVERSITY INFORMATION TECHNOLOGY SERVICES

Jetstream2 partners

Arizona State University

The University of Arizona

Johns Hopkins University

TACC

UCAR

Jetstream2

http://jetstream-cloud.org/
National Science Foundation
Award #ACI-2005506

OOKAMI

- **Fujitsu A64FX** testbed at Stony Brook University in cooperation with the University at Buffalo
  - Same processor as in Fugaku
  - HPE/Cray Apollo 80
  - Arm V8-64bit
  - Support high calculation performance and low power consumption
  - 32 (4x8) GB HBM @ 1TB/s
  - Supports Scalable Vector Extensions (SVE) with 512-bit vector length
  - Suited for a wide range of applications
  - Current status: 83 projects, 250 users

OOKAMI

www.stonybrook.edu/ookami

Node	
Processor	A64FX
#Cores	48
Peak DP	2.76 TOP/s
Memory	32GB@**1TB/s**
System	
#Nodes	176
Peak DP	486 TOP/s
Peak INT8	3886 TOP/s
Memory	5.6 TB
Disk	0.8 PB Lustre
Comms	IB HDR-100

OOKAMI

2019 2022 2025

Procurement

Installation

System burn-in

Early user access

System open for testbed projects

System open for testbed & production projects

XSEDE integration
Level 2 Service / Research Provider

XSEDE

ADVANCE TO ACCESS

OOKAMI

www.stonybrook.edu/ookami

ookami_computer@stonybrook.edu