PSC
Carnegie Mellon University
University of Pittsburgh

XSEDE Resources @ PSC

Nick Nystrom
Chief Scientist
nystrom@psc.edu

Petascale Computing Institute 2019 · Pittsburgh & Remote · August 23, 2019

Outline

1. Bridges
2. Bridges-AI
3. Bridges-2
4. Summary

Outline

1. Bridges
2. Bridges-AI
3. Bridges-2
4. Summary

PSC
Carnegie Mellon University
University of Pittsburgh

NSF

Bridges Overview

Nick Nystrom
PI & Project Director
AD for Acquisition & Ops
nystrom@psc.edu

Paola Buitrago
Co-PI &
AD for AI & Data Science
paola@psc.edu

John Urbanic
Co-PI &
AD for EOT
urbanic@psc.edu

Sergiu Sanielevici
AD for User Support

Robin Scibek
Project Manager

Jim Marsteller
Security Manager

David Moses
Compliance Manager

Petascale Computing Institute 2019 · Pittsburgh & Remote · August 23, 2019

- ***Available at no cost for open research and courses and by arrangement to industry***
- *Easier access for CMU and Pitt faculty through the Pittsburgh Research Computing Initiative*
- 29,036 Intel Xeon CPU cores
- 216 NVIDIA GPUs: 64 K80, 64 P100, **88 V100**
- 17 PB storage (10 PB persistent, 7.3 PB local)
- 277 TB memory (RAM), up to 12 TB per node
- 44M core-hours, 173k GPU-AI-hours, 442k GPU-hours, and 343k TB-hours allocated quarterly
- Serving >2,100 projects and 16,000 users at 800 institutions, spanning 122 fields of study

- ***Bridges-AI:*** NVIDIA DGX-2 Enterprise AI system + 9 HPE 8-Volta Apollo 6500 Gen10 servers: total of 88 V100 GPUs

Bridges converges HPC, AI, and Big Data to empower new research communities, bring desktop convenience to advanced computing, expand remote access, and help researchers to work more intuitively.

- Funded by NSF award #OAC-1445606 ($20.9M), Bridges emphasizes usability, flexibility, and interactivity
- Available at no charge for open research and coursework and by arrangement to industry
- Popular programming languages and applications: Python, Jupyter, R, MATLAB, Java, Spark, Hadoop, …
- 856 compute nodes containing Intel Xeon CPUs and 128GB (800), 3TB (42), and 12TB (4) of RAM each
- 216 NVIDIA Tesla GPUs: 64 K80, 64 P100, **(new) 88 V100 configured to balance capability & capacity**
- Dedicated nodes for persistent databases, gateways, and distributed services
- The world's first deployment of the Intel Omni-Path Architecture fabric

NSF

Hewlett Packard Enterprise

intel NVIDIA

Carnegie Mellon University PSC University of Pittsburgh

Bridges Makes Advanced Computing Easy

Make HPC accessible to all research communities

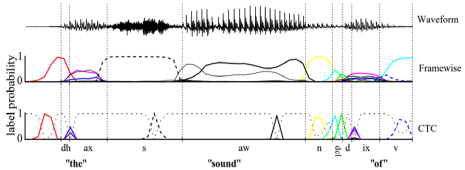
Converge HPC, AI, and Big Data

Support the widest range of science with an extremely rich computing environment

- 3 tiers of memory: 12 TB, 3 TB, and 128 GB
- Powerful, flexible CPUs and GPUs
- Familiar, easy-to-use user environment:

Elements not available in traditional supercomputers

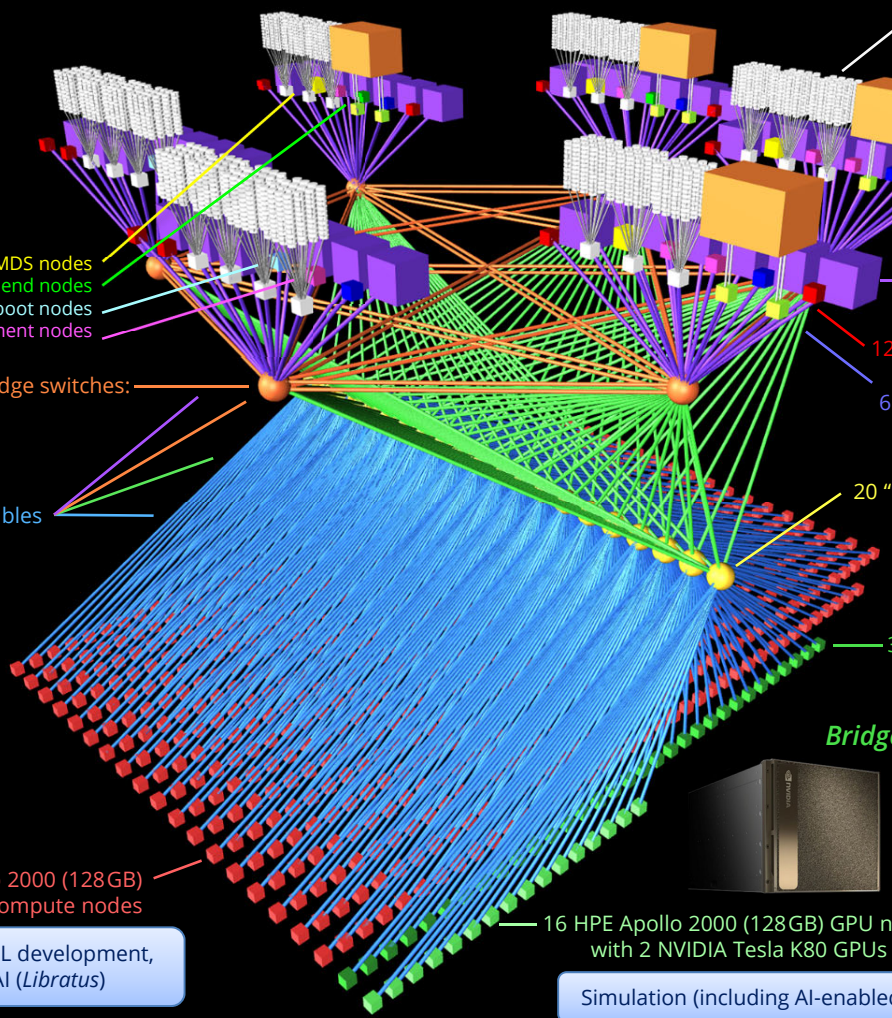
  - Interactivity
  - Popular languages and frameworks: Python, Anaconda, R, MATLAB, Java, Spark, Hadoop
  - AI frameworks: TensorFlow, Caffe2, PyTorch, etc.
  - Containers (e.g., NGC) and virtual machines (VMs)
  - Databases
  - Gateways and distributed (web) services
  - Large collection of applications and libraries



Carnegie Mellon University PSC University of Pittsburgh

BRIDGES

A PITTSBURGH SUPERCOMPUTING CENTER RESOURCE



20 Storage Building Blocks, implementing the parallel Pylon storage system (10 PB usable)

Project & community datasets

4 HPE Integrity Superdome X (12TB) compute nodes ...

... each with 2 gateway nodes

Large-memory Java & Python

4 MDS nodes
2 front-end nodes
2 boot nodes
8 management nodes

42 HPE ProLiant DL580 (3 TB) compute nodes

12 HPE ProLiant DL380 database nodes
6 HPE ProLiant DL360 web server nodes

User interfaces for AIaaS, BDaaS

6 "core" Intel® OPA edge switches: fully interconnected, 2 links per switch

Intel® OPA cables

Robust paths to parallel storage

20 "leaf" Intel® OPA edge switches

Distributed training, Spark, etc.

Representative uses for AI

Purpose-built Intel® Omni-Path Architecture topology for data-intensive HPC

32 HPE Apollo 2000 (128GB) GPU nodes with 2 NVIDIA Tesla P100 GPUs each

Deep Learning

Bridges-AI

Maximum-Scale Deep Learning

NVIDIA DGX-2 and 9 HPE Apollo 6500 Gen10 nodes: 88 NVIDIA Tesla V100 GPUs

748 HPE Apollo 2000 (128GB) compute nodes

16 HPE Apollo 2000 (128GB) GPU nodes with 2 NVIDIA Tesla K80 GPUs each

ML, inferencing, DL development, Spark, HPC AI (*Libratus*)

Simulation (including AI-enabled)

Bridges Virtual Tour: https://psc.edu/bvt

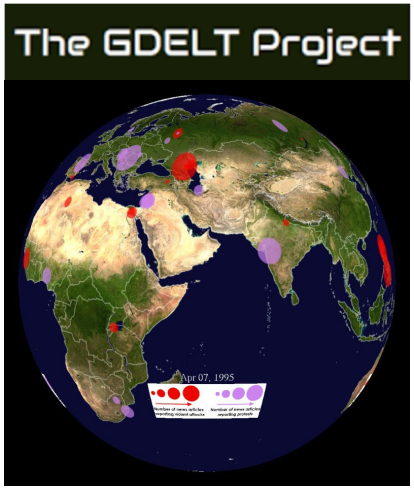
Community Datasets

- Hosting mature corpus of data and data tools for an open science community
  - Accessible by multiple users, multiple groups.
  - Provision of reusable data management tools
  - Facilitate collaboration
  - Offload data management
- Interoperable with HPC capabilities
  - High speed data transfer
  - High performance compute capabilities
- Support copies, maintenance, guarantee integrity
- Data resource not subject to project limitations

Some unique, others with local caching for efficiency and to drive interdisciplinary research

IMAGENET

Common Crawl



Carnegie Mellon University PSC University of Pittsburgh

The Expanding Ecosystem of *Bridges*

10s of PB

Hybrid on-prem data/AI/HPDA + Cloud

Human BioMolecular Atlas

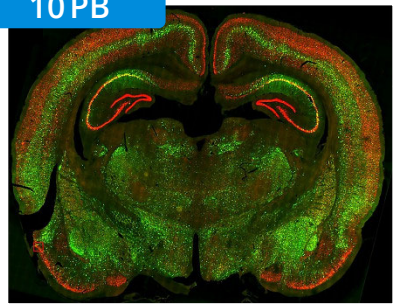
2.2 PB

Big Data for Better Health

BRIDGES

A PITTSBURGH SUPERCOMPUTING CENTER RESOURCE

10 PB



Brain Image Library

Dedicated resources + cloud use of Bridges

Campus Clusters

Carnegie Mellon University PSC University of Pittsburgh

Accessing Bridges: No Cost for Research & Education and Cost-Recovery Rates for Corporate Use

The following annual allocations are renewable and extendable, also at no cost for research and education.

	Open Research			Industry
	XSEDE			PSC Corporate Program
	Startup	Research	Education	
Cost	No charge	No charge	No charge	Cost recovery rates
CPU-hours	50k	Up to ~10^7	Up to ~10^6	Up to ~18M
GPU-hours	2500	Up to ~10^5	Up to ~10^4	Up to ~180k
GPU-AI hours	1500	Up to ~10^5	Up to ~10^4	Up to ~69k
TB-hours	1000	Up to ~10^4	Up to ~10^4	Up to ~137k
Developer	Yes	Yes	(Yes)	Yes
Accepted	Any time	Quarterly	Any time	Any time
Awarded	~1-2 days	Quarterly	~1-3 days	ASAP

Outline

1. Bridges
2. Bridges-AI
3. Bridges-2
4. Summary

PSC
Carnegie Mellon University
University of Pittsburgh

NSF

Bridges-AI Overview

Nick Nystrom
nystrom@psc.edu

Paola Buitrago
paola@psc.edu

Petascale Computing Institute 2019 · Pittsburgh & Remote · August 23, 2019

Bridges-AI: Overview

- *1 NVIDIA DGX-2*
  Tightly couples 16 NVIDIA Tesla V100 (Volta) GPUs at 2.4 TB/s bisection bandwidth, to provide maximum capability for the most demanding of AI challenges
- *9 Hewlett Packard Enterprise Apollo 6500 Gen10 servers*
  Each with 8 NVIDIA Tesla V100 GPUs connected by NVLink 2.0, to balance great AI capability and capacity
- *Bridges-AI is integrated with Bridges and allocated through XSEDE as resource "Bridges GPU-AI"*, analogous to *Bridges GPU*, *RM*, *LM*, and *Pylon*

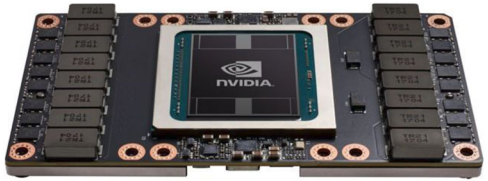
Volta introduces *Tensor Cores* to accelerate neural networks, yielding extremely high peak performance for appropriate applications.

Bridges-AI providea massive aggregate performance:

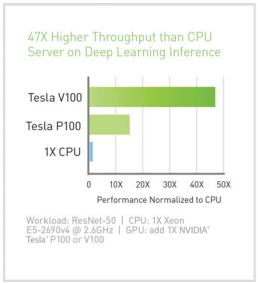
- *9.9 Pf/s mixed-precision tensor*
- 251 Tf/s 32-bit
- 125 Tf/s 64-bit

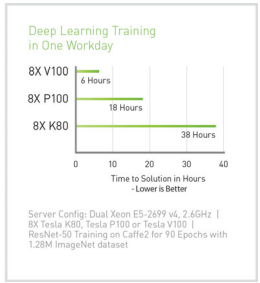
- *Bridges-AI* adds 9.9 Pf/s of mixed-precision tensor, 1.24 Pf/s of fp32, and 0.62 Pf/s of fp64. (Totals: 9.9 Pf/s tensor, 3.93 Pf/s fp32, 1.97 Pf/s fp64).
- *The $1.786M supplement includes additional staffing* to support solutions and scaling
- *Deployment: Bridges-AI* deployed on time. PSC ran an Early User Program from November-December 2018, and production operations began January 1, 2019.

The Heart of Bridges-AI: NVIDIA Volta



NVIDIA Tesla V100 SXM2 Module
with Volta GV100 GPU





Training ResNet-50 with ImageNet:
V100 : 1075 images/s[a]
P100 : 219 images/s[b]
K80 : 52 images/s[b]

a. https://devblogs.nvidia.com/tensor-core-ai-performance-milestones/
b. https://www.tensorflow.org/performance/benchmarks

New Streaming Multiprocessor (SM) architecture, introducing Tensor Cores, independent thread scheduling, combined L1 data cache and shared memory unit, and 50% higher energy efficiency over Pascal.

Tensor Cores accelerate deep learning training and inference, providing up to 12× and 6× higher peak flops respectively over the P100 GPUs currently available in XSEDE.

NVLink 2.0 delivering 300 GB/s total bandwidth per GV100, nearly 2× higher than P100.

HBM2 bandwidth and capacity increases: 900 GB/s and up to 32GB.

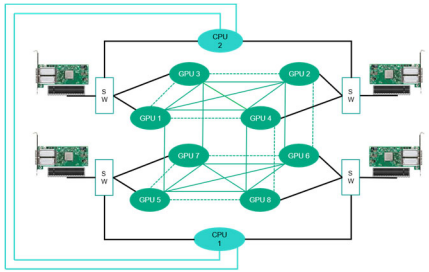
Enhanced Unified Memory and Address Translation Services improve accuracy of memory page migration by providing new access counters.

Cooperative Groups and New Cooperative Launch APIs expand the programming model to allow organizing groups of communicating threads.

Volta-Optimized Software includes new versions of frameworks and libraries optimized to take advantage of the Volta architecture: TensorFlow, Caffe2, MXNet, CNTK, cuDNN, cuBLAS, TensorRT, etc.

Balancing AI Capability & Capacity: HPE Apollo 6500

HPE Apollo 6500 Gen10 Server



HPE Apollo 6500 Gen10
hybrid cube-mesh topology

Bridges-AI adds 9 HPE Apollo 6500 Gen10 servers

Each HPE Apollo 6500 couples 8 NVIDIA Tesla V100 SXM2 GPUs

- 40,960 CUDA cores and 5,120 tensor cores

Performance: 1 Pf/s mixed-precision tensor, 125 Tf/s 32b, 64 Tf/s 64b

Memory: 128 GB HBM2, 7.2 TB/s aggregate memory bandwidth

2 × Intel Xeon Gold 6148 CPUs and *192 GB of DDR4-2666 RAM*

- 20c, 2.4–3.7 GHz, 27.5 MB L3, 3 UPI links

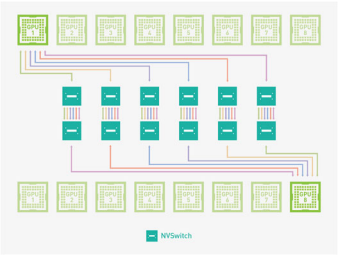
4 × 2 TB NVMe SSDs for user and system data

1 × Intel Omni-Path host channel adapter

Hybrid cube-mesh topology connecting the 8 V100 GPUs and 2 Xeon CPUs, using NVLink 2.0 between the GPUs and PCIe3 to the CPUs

Maximum DL Capability: NVIDIA DGX-2

NVIDIA DGX-2



NVIDIA DGX-2 with NVSwitch internal topology

Couples 16 NVIDIA Tesla V100 SXM2 GPUs
- 81,920 CUDA cores and 10,240 tensor cores

Performance: 2 Pf/s mixed-precision tensor, 251 Tf/s 32b, 125 Tf/s 64b

Memory: 512 GB HBM2, 14.4 TB/s aggregate memory bandwidth

2 × Intel Xeon Platinum 8168 CPUs and *1.5 TB of DDR4-2666 RAM*
- 24c, 2.7–3.7 GHz, 33 MB L3, 3 UPI links

2 × 960 GB NVMe SSDs host the Ubuntu Linux OS

8 × 3.84 TB NVMe SSDs (aggregate ~30 TB)

8 × Mellanox ConnectX adapters for EDR InfiniBand & 100 Gb/s Ethernet

The *NVSwitch* tightly couples the 16 V100 GPUs for capability & scaling
- Each of the 12 NVSwitch chips is an 18×18-port, fully-connected crossbar
- 50 GB/s/port and 900 GB/s/chip bidirectional bandwidths
- 2.4 TB/s system bisection bandwidth

Deep Learning Frameworks on *Bridges*

Outline

1. Bridges
2. Bridges-AI
3. Bridges-2
4. Summary

PSC
Carnegie Mellon University
University of Pittsburgh

NSF

Introducing *Bridges-2*

Nick Nystrom
PI & Project Director
AD for Acquisition & Ops
nystrom@psc.edu

Paola Buitrago
Co-PI &
AD for AI & Data Science
paola@psc.edu

Sergiu Sanielevici
Co-PI &
AD for User Support
sergiu@psc.edu

Robin Scibek
Co-PI &
AD for Broader Impact
flaus@psc.edu

Deb Nigra
Project Manager

Ken Goodwin
Networking Manager

Ken Hackworth
Allocations Manager

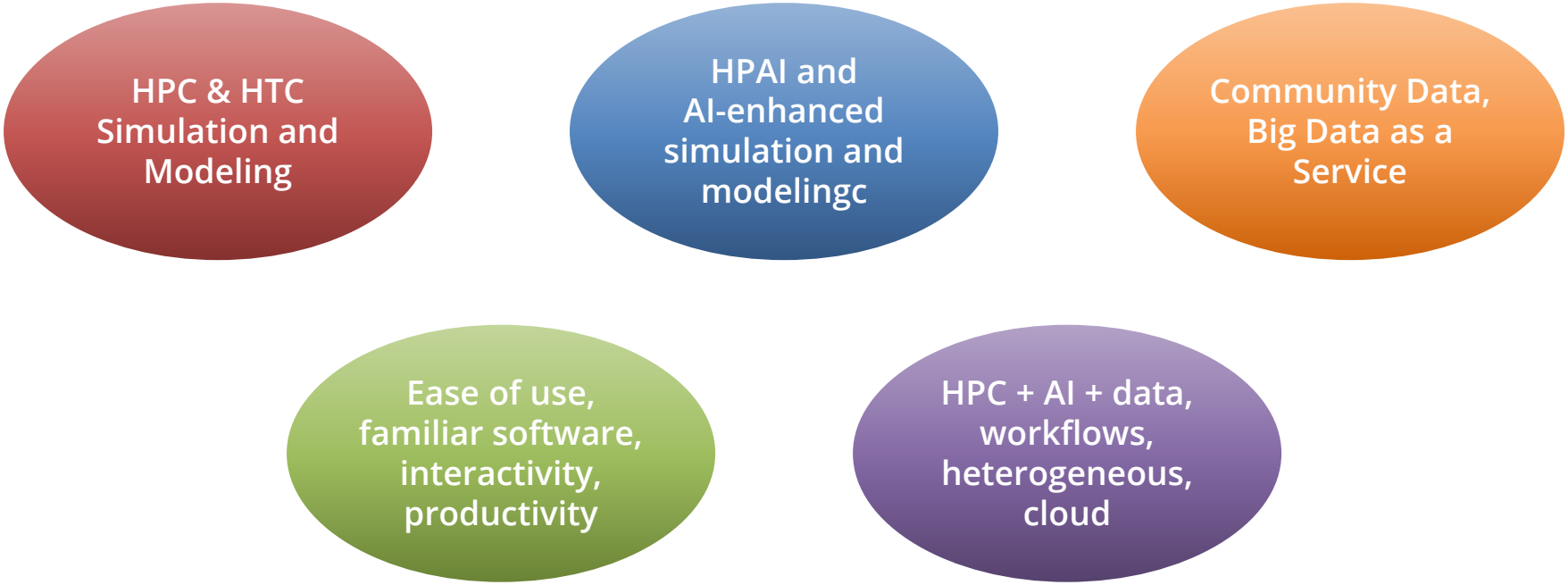
Jim Marsteller
Security Manager

David Moses
Compliance Manager

Philip Blood
XSEDE Liaison

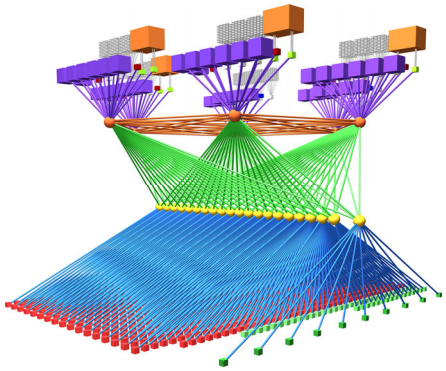
PEARC19 Plenary C · Chicago · August 1, 2019

Driving Rapidly Evolving Science and Engineering



An Ecosystem for Rapidly Evolving, Data-Intensive Science & Engineering

BRIDGES
Converged HPC, AI, and Big Data
A PSC RESOURCE

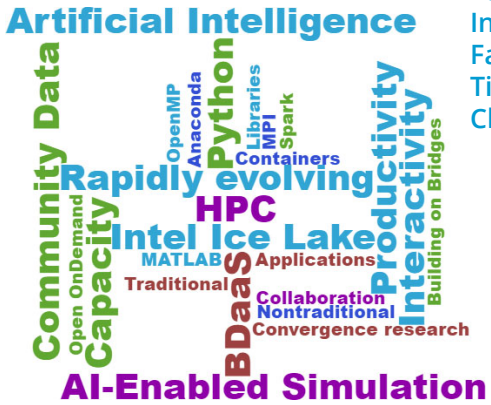


Pioneered HPC+AI+Big Data
Bridges-AI expansion
Intel OPA first installation
Has become an ecosystem

2072 projects
16,000 users
800 institutions
122 fields of study
130 education allocations

NSF Award OAC-1445606

BRIDGES-2
Converged HPC, AI & Big Data for Rapidly-Evolving Research
A PSC RESOURCE



Full-system HPAI
Intel Ice Lake
Fast flash array
Tiered data management
Cloud interoperability

NSF Award OAC-1928147

Carnegie Mellon University PSC University of Pittsburgh

AI @ PSC: An Evolution (1)

Prof. Tuomas Sandholm

Noam Brown

Blacklight

2010

TartarianN

2015

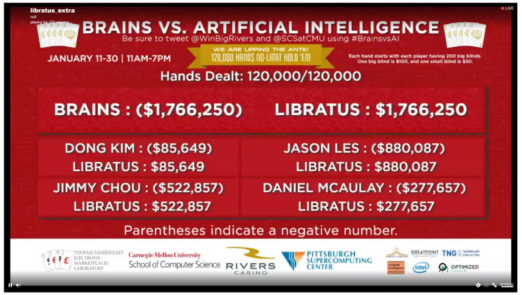
Claudico

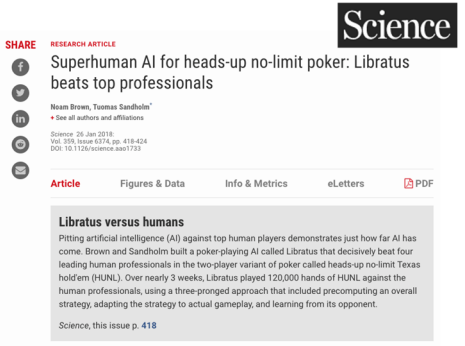
AI @ PSC: An Evolution (2)

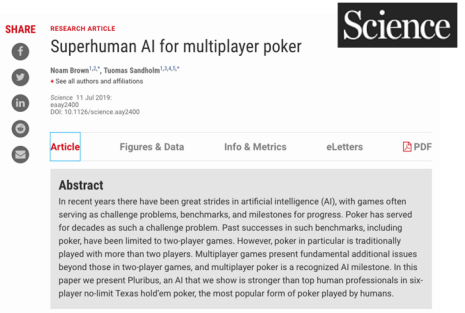
BRIDGES

2017

2019







Carnegie Mellon University PSC University of Pittsburgh

AI @ PSC: An Evolution (3)

BRIDGES

2019

Prof. Tuomas Sandholm

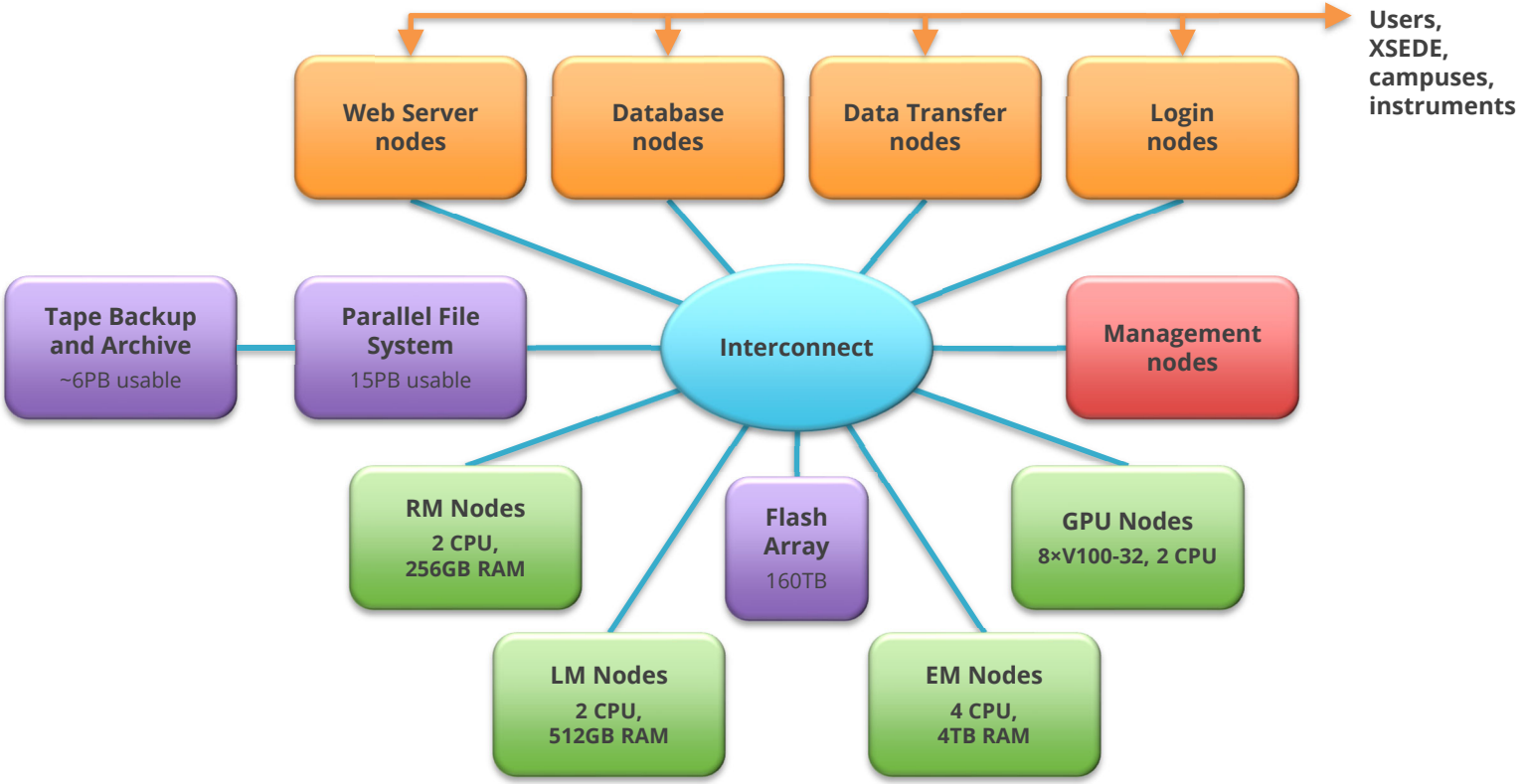
Noam Brown

2019 Marvin Minsky Medal for Outstanding Achievements in AI

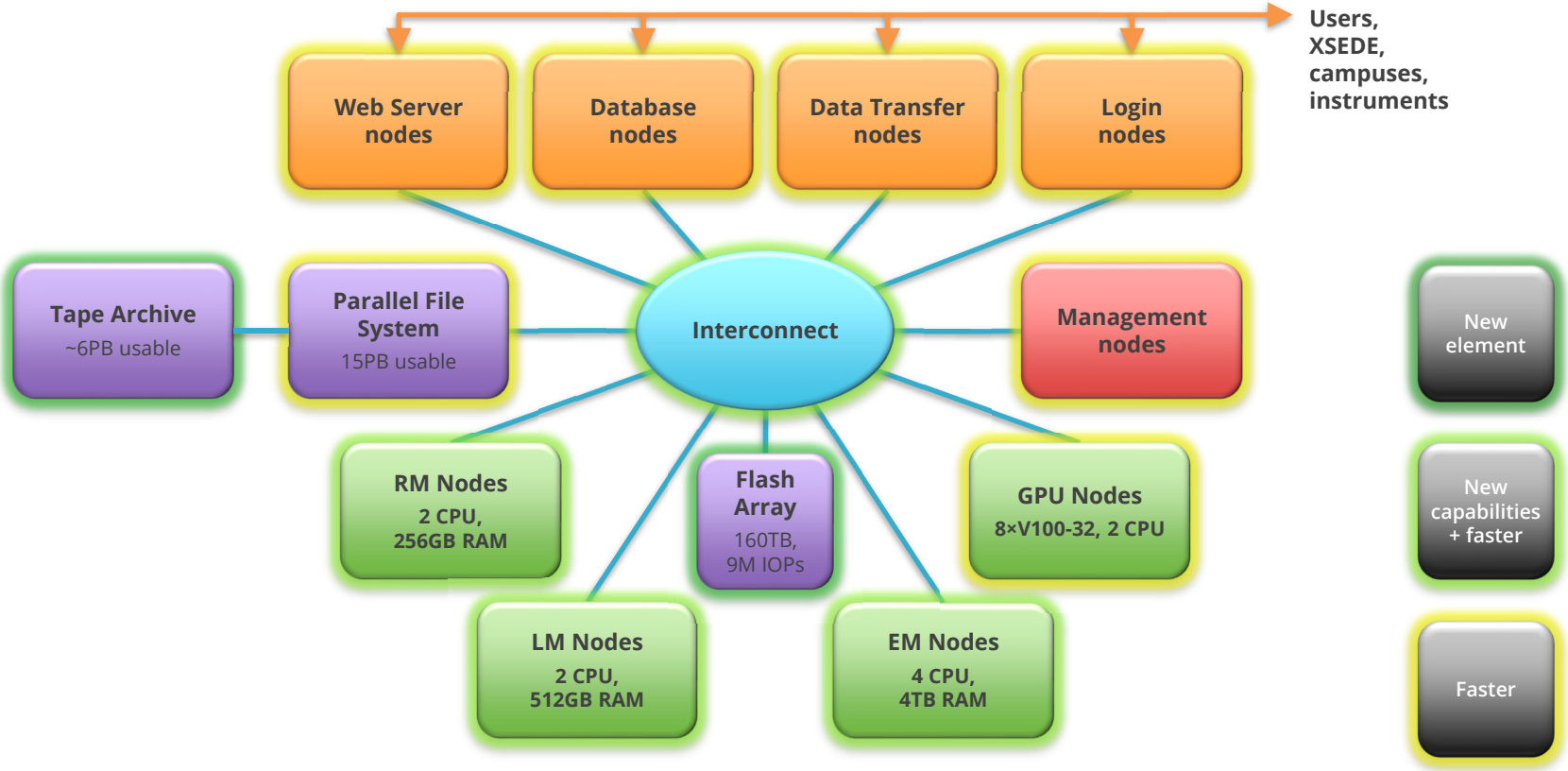
"*Poker is an important challenge for AI because any poker player has to deal with incomplete information. Incomplete information makes the computational challenge orders of magnitude harder. Libratus used fundamentally new techniques for dealing with incomplete information, which have exciting potential applications far beyond games.*"

— Professor Michael Wooldridge,
Chair of the IJCAI Awards Committee

Bridges-2: High-Level Architecture



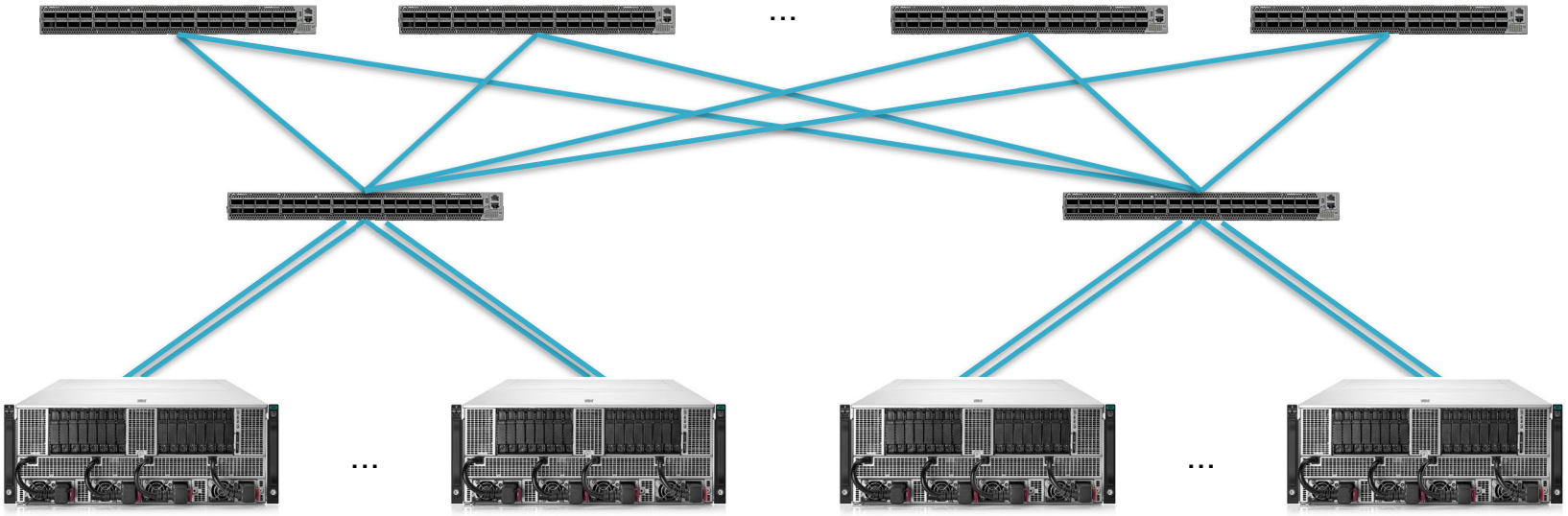
Bridges-2: High-Level Architecture



Carnegie Mellon University PSC University of Pittsburgh

GPU Infrastructure

12 Mellanox Quantum Spine Switches



12 × HPE Apollo 6500 Gen10 Server
Each: 1Pf/s tensor, 125 Tf/s fp32, 64 Tf/s fp64

12 × HPE Apollo 6500 Gen10 Server
Each: 1Pf/s tensor, 125 Tf/s fp32, 64 Tf/s fp64

Data Infrastructure

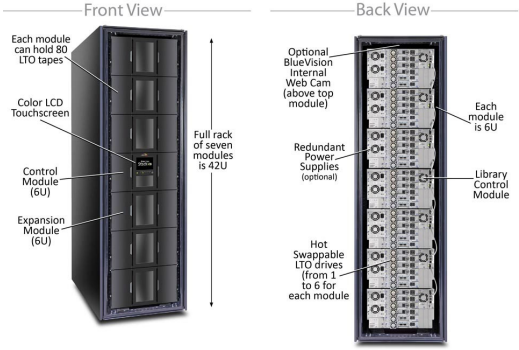
Bridges-2 filesystem (b2fs): Managed by HPE Data Management Framework (DMF) to provide a single namespace and user-friendly, rule-based migration.

5 HPE TierZero NVMe servers

- 160 TB, 9M IOPs, 180 GB/s
- Use cases: training on large data, genomics, databases
- To be implemented as a scheduled resource

HPE disk-based Lustre filesystem

- 15 PB usable, 21 PB raw
- 130 GB/s read, 140 GB/s write
- RAIDZ2
- 10 data server pairs, each serving 2.1 PB (raw)
- To be allocated through XSEDE



HPE StoreEver MSL6480 Tape Library

- 7 modules, scalable; 80 LTO tape slots per module
- 5 PB uncompressed, ~6 PB compressed
- 36 TB/hour
- Use cases: archiving, disaster recovery
- To be implemented as a resource to be allocated through XSEDE
- Option for external groups to fund project-specific expansion

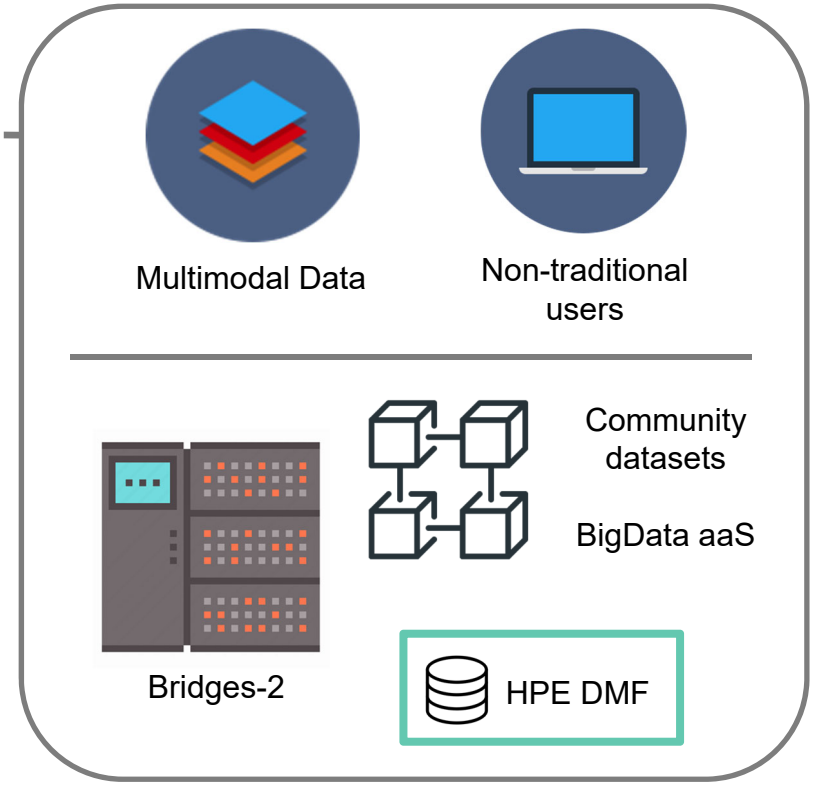
Carnegie Mellon University PSC University of Pittsburgh

AI and Data @ *Bridges-2*

Productivity

Usability

Interoperability



AI and Data @ *Bridges-2*

Productivity

Usability

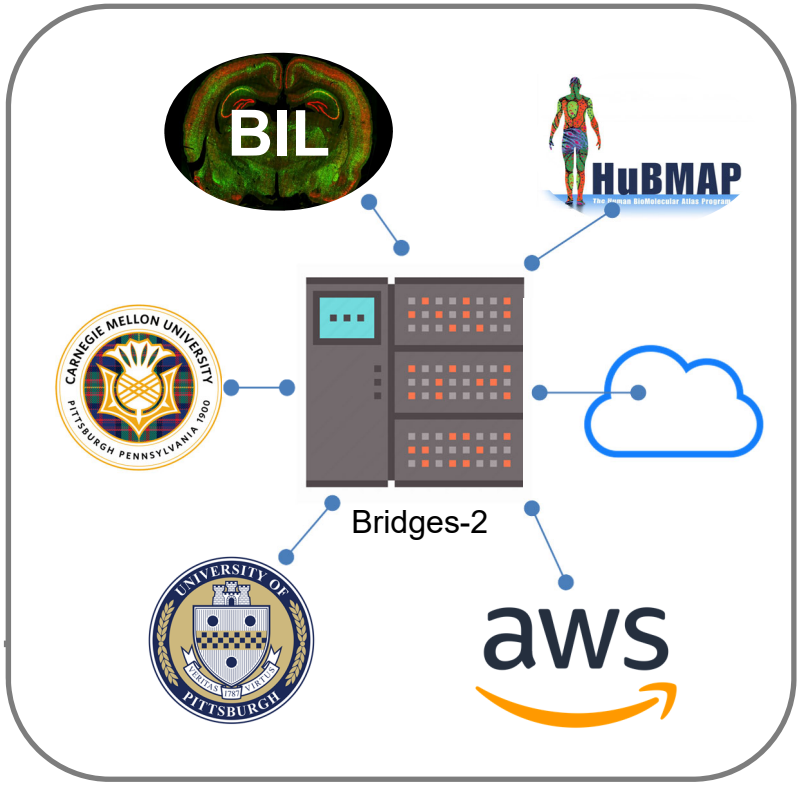
Interoperability

AI and Data @ *Bridges-2*

Productivity

Usability

Interoperability



Carnegie Mellon University PSC University of Pittsburgh

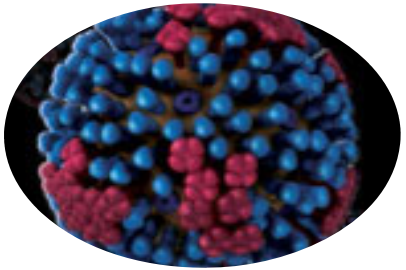
User Support

- Very strong focus on support for the full *Bridges-2* user community
  - Emphasis on AI & Data Science, nontraditional users and applications
- Leveraging XSEDE Novel & Innovative Projects
- Early Operations focus on the user environment, documentation, and content development

Broadening Impact

Building STEM Talent

Innovating for the Future



Improving Our Society

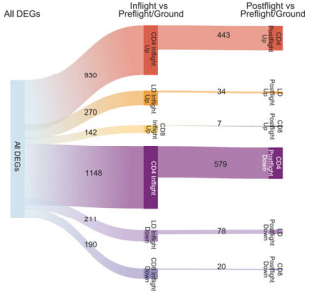
Reaching Beyond Borders

Engaging a Wider Audience

Carnegie Mellon University PSC University of Pittsburgh

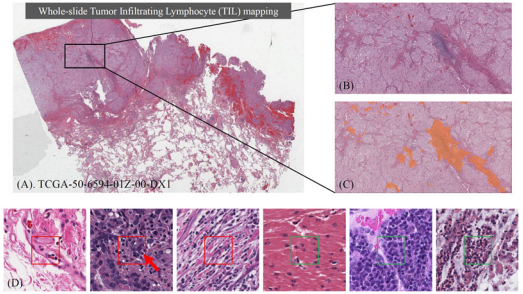
Bridges-2 Application Areas: Examples

Gene Expression



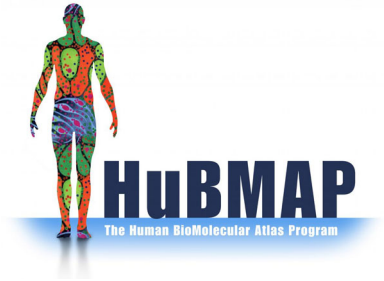
F. E. Garrett-Bakelman et al., *The NASA Twins Study: A Multidimensional Analysis of a Year-Long Human Spaceflight*, Science, 2019. DOI: 10.3847/1538-4357/aac329.

Advancing Digital Pathology with AI



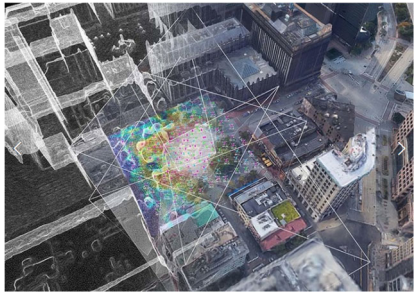
S. Abousamra et al., *Learning from Thresholds: Fully Automated Classification of Tumor Infiltrating Lymphocytes for Multiple Cancer Types*, 2019. ArXiv 1907.03960v1.

Mapping the Human Body at Cellular Resolution



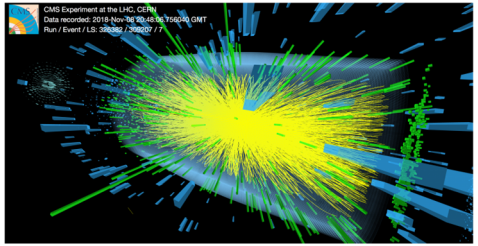
M. Snyder et al., *Mapping the Human Body at Cellular Resolution -- The NIH Common Fund Human BioMolecular Atlas Program*, Nature, to appear.

Developing Smart Cities



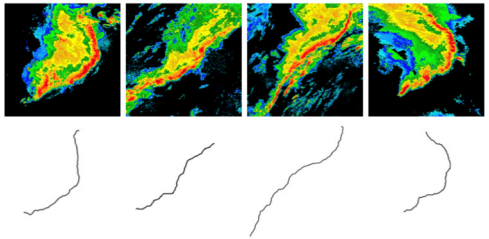
J. Argota and D. Cardoso Llach, *Using Computation to Understand How Pedestrians Use Market Square*, 2018. https://soa.cmu.edu/news-archive/2018/9/5/using-computation-to-understand-how-pedestrians-use-market-square.

Workflows for CMS @ HL-LHC



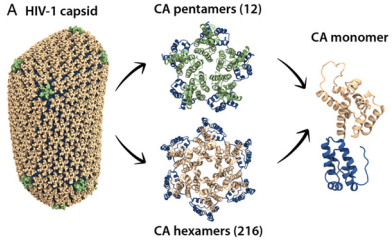
Event display of heavy-ion collision registered at the CMS detector on Nov. 8, 2018 (image: Thomas McCauley). From https://cms.cern/news/2018-heavy-ion-collision-run-has-started.

Improving Severe Storm Prediction



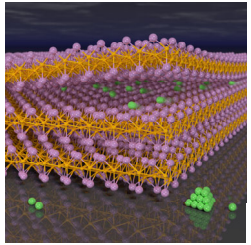
X. Zheng et al., *Detecting Comma-shaped Clouds for Severe Weather Forecasting using Shape and Motion*, IEEE Transactions on Geosciences and Remote Sensing, 2019. DOI: 10.1109/TGRS.2018.2887206.

Understanding Immunity



C. M. Quinn et al., *Dynamic regulation of HIV-1 capsid interaction with the restriction factor TRIM5a identified by magic-angle spinning NMR and molecular dynamics simulations*, PNAS, 2018. DOI: 10.1073/pnas.1800796115.

New Materials



M. Amsler et al., *Cubine, A Quasi Two-Dimensional Copper-Bismuth Nanosheet*, Chem. Mater. 2017. DOI: 10.1021/acs.chemmater.7b03997

Carnegie Mellon University PSC University of Pittsburgh

Target Timeline

October 1, 2019	Award start date; preparatory activities begin – System and user environment, documentation, content, dissemination, etc. – Broadly invite applications for the Early User Program
June-July, 2020	Accept initial round of XRAC proposals
July, 2020	Delivery, installation, initial testing
August, 2020	Early User Program
October 1, 2020	Start of Production Operations

Outline

1. Bridges
2. Bridges-AI
3. Bridges-2
4. Summary

Summary

- *Bridges* pioneered AI, HPC, and Big Data, and through its heterogeneous, very flexible architecture, created a large community of nontraditional users and interoperating cyberinfrastructure.
- *Bridges-AI* adds tremendous capability for scalable AI.
- *Bridges-2* will greatly extend those proven concepts with full-system HPAI, a new all-flash component for fast data, tiered data management, Intel Ice Lake CPUs, and enhanced cloud interoperability.

Bridges and Bridges-AI are available now!

The Bridges-2 Early User Program is planned for August 2020. Join us!