2018 COMMUNITY CLUSTER PARTNER ANNUAL MEETING

PRESTON SMITH
DIRECTOR OF RESEARCH COMPUTING SERVICES
PSMITH@PURDUE.EDU

ALEXANDER YOUNTS
PRINCIPAL RESEARCH COMPUTING ENGINEER
AY@PURDUE.EDU

PURDUE UNIVERSITY

COMMUNITY CLUSTERS

PURDUE UNIVERSITY®

Community Cluster Program

The Rules

- **We build a cluster each year**
  - Cluster is a service, buying in provides 5 years of access.
  - Since 2008, faculty invest approx. $2.5M of compute each year

- You get out at least what you put in
  - Buy 1 node or 100, you get a queue that guarantees access up to that many CPUs

PURDUE UNIVERSITY

Community Cluster Program

The Rules

- But wait, there's more!!
  - What if your neighbor isn't using his queue?
    - You can use it, but your job has to run in 4-hour chunks if he wants to run.
- You don't have to do the work
  - Your grad student gets to do research rather than run your cluster.
    - Nor do you have to provide space in your lab for computers.
  - ITaP provides data center space, systems administration, application support.
  - Just submit jobs!

PURDUE UNIVERSITY

10 HPC SYSTEMS

STEELE

7,216 cores, Installed May 2008

Retired Nov. 2013

COATES

8,032 cores, Installed May 2008

24 departments, 61 faculty investors

Retired Sep. 2014

ROSSMANN

11,088 cores, Installed Sept. 2010

17 departments, 37 faculty investors

Retired Sep. 2015

HANSEN

9,120 cores, Installed Sept. 2011

13 departments, 26 faculty investors

Retiring Oct. 2016

CARTER

10,368 cores

Installed April 2012 – Retired 2017

26 departments, 60 faculty investors

#54 on June 2012 Top 500

CONTE

9,280 Xeon cores (69,900 Xeon Phi)

Installed August 2013 – Retired 2018

26 departments, 62 faculty investors

#28 on June 2013 Top 500

DATA DEPOT

2.5 PB of disk storage
Installed Nov. 2014

500+ faculty investors from every academic college

RICE

13,200 cores, Installed May 2015

33 departments, 69 faculty investors

HALSTEAD

10,160 cores, Installed December 2016

39 departments, 62 faculty investors

BROWN

13,200 cores
Installed October 2017
36 departments
70 faculty investors

$5039 for 5 years of service

#302 on Nov 2017 Top 500


#105 STEELE 2008
#102 COATES 2009
#150 HANSEN 2010 (est.)
#126 ROSSMANN 2011
#54 CARTER 2012
#28 CONTE 2013
#166 RICE 2015
#302 BROWN 2017
PURDUE COMMUNITY CLUSTERS
TOP500 RANKINGS

Cluster Program Partners

302M hours delivered in 2017
177 active (over 200 all-time) investors from 50 departments, from every College, and 3 Purdue campuses

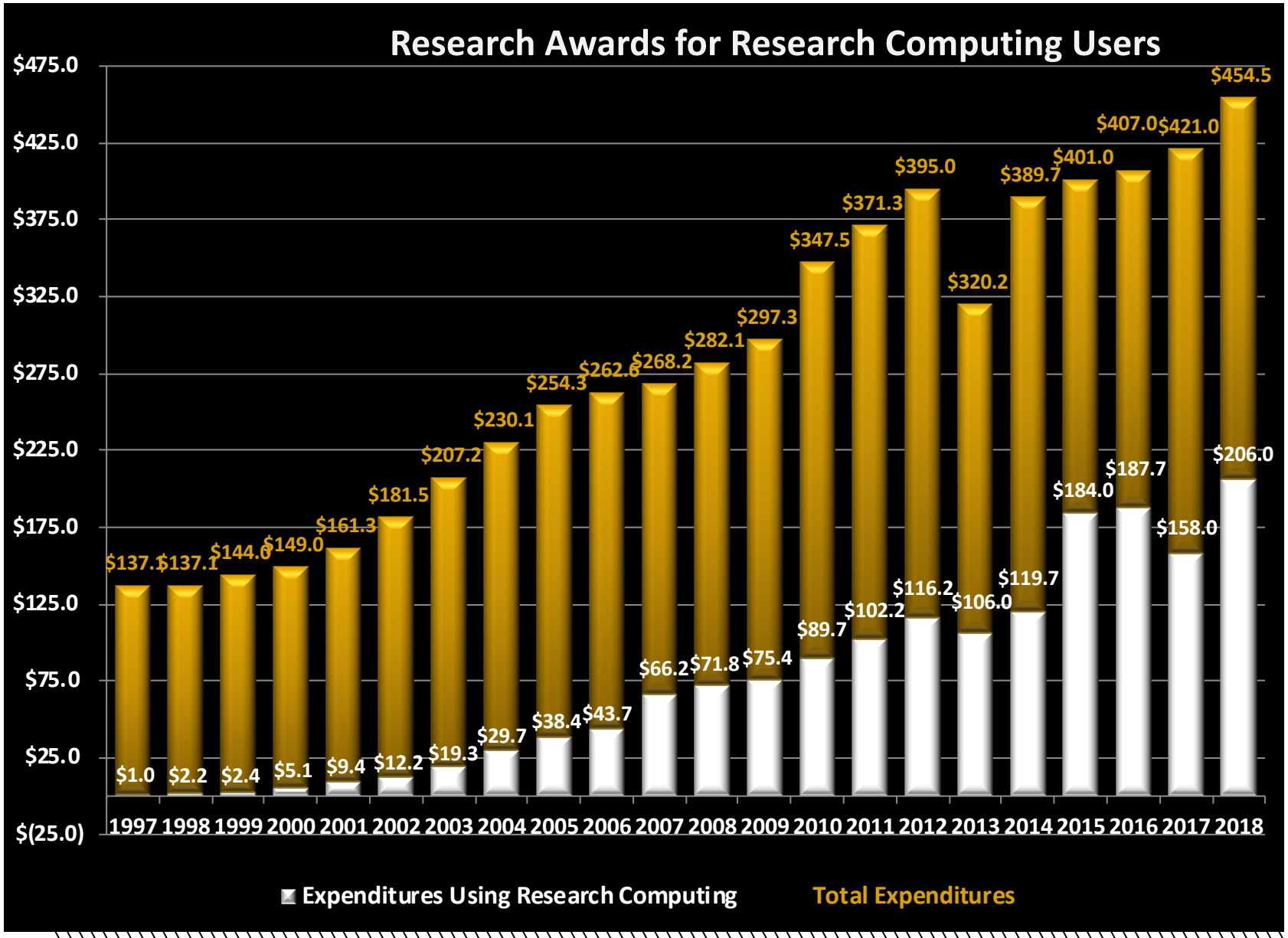
Today, the program is part of many departments' faculty recruiting process.

A selling point to attract people to Purdue!

Department	Cores
Aeronautics and Astronautics	5740
Mechanical Engineering	5556
CMS Tier2	5440
Electrical and Computer Engineering	4344
Earth, Atmospheric, and Planetary Sciences	2540
Materials Engineering	2064
Nuclear Engineering	1564
Other College of Engineering	980
Chemistry	824
Physics and Astronomy	820
Biomedical Engineering	640
Other Executive Vice President for Research and Partnerships	600
Statistics	512
Chemical Engineering	424
Agricultural and Biological Engineering (Biological Engineering)	368
Biological Sciences	356
Industrial Engineering	296
Civil Engineering	276
Computer and Information Technology	248
Medicinal Chemistry and Molecular Pharmacology	248
Mathematics	232
Bioinformatics Core	200
Agronomy	180
ITaP	176
Computer Science	156
Horticulture and Landscape Architecture	156
Cancer Center	96
Forestry and Natural Resources	96
Biochemistry	40
Botany and Plant Pathology	40
Industrial and Physical Pharmacy	40
Brian Lamb School of Communication	32
Agricultural Economics	20
Animal Sciences	20
Food Science	20
Health Sciences	20
Other College of Pharmacy	20
Agricultural and Biological Engineering (Agricultural Systems Mgmt)	16

PURDUE UNIVERSITY

Research Awards to HPC Partners



Brown – on the Floor Today

Specifications

If we were to build a CPU cluster today, it would look exactly like Brown!

Base node: Dell R640

- 24-core node, 2.6 GHz Intel Xeon Gold "Sky Lake" processors (Xeon Gold 6126)
  - 32 Flops per cycle!
- 96 GB DDR4 memory
  - 384 GB, 768 GB & 1536 GB options
- EDR Infiniband interconnect
  - 100 Gbps, 3:1 fat tree – very similar in speed to Halstead
  - Converged fabric – IP traffic uses Infiniband rail

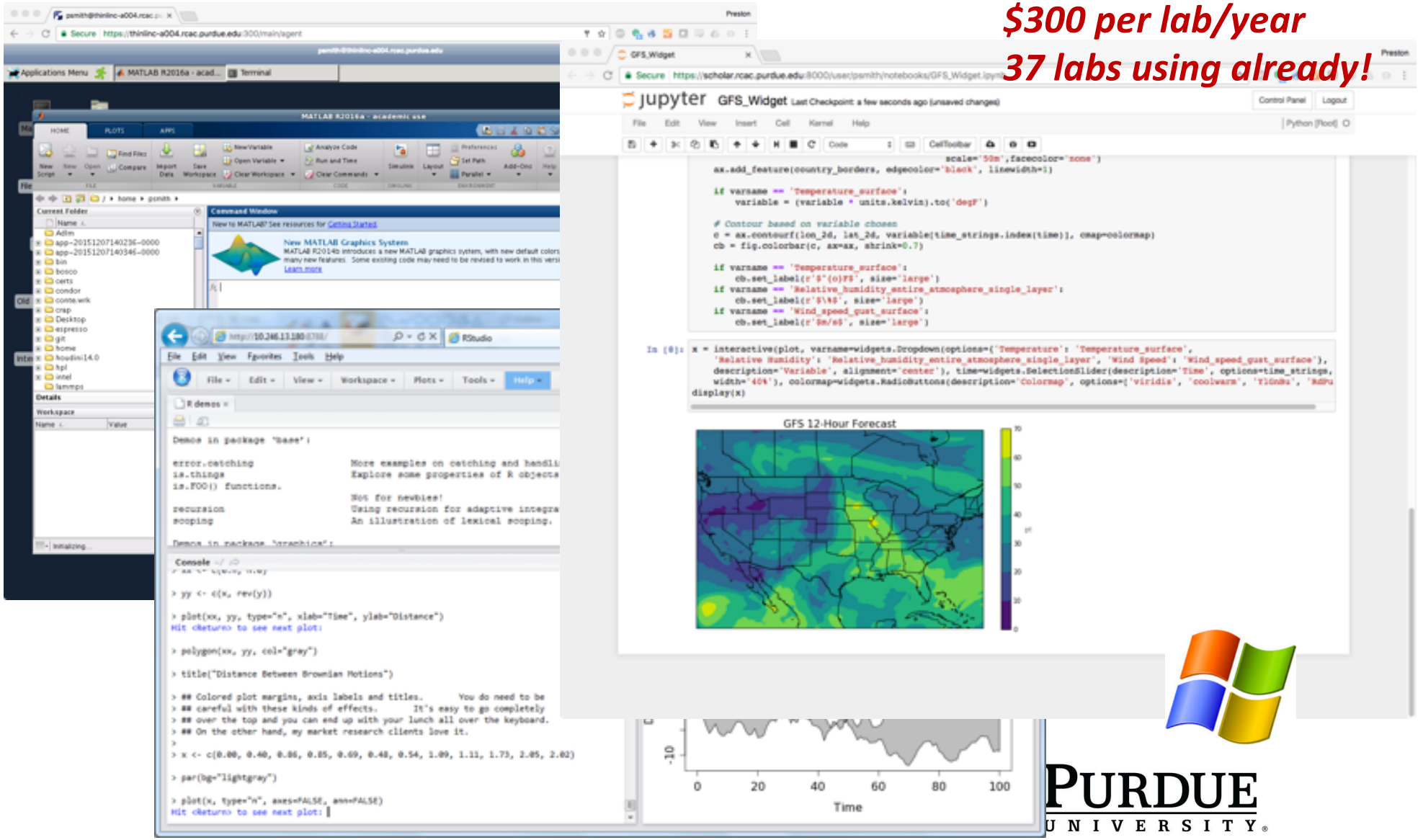
There are still Brown nodes available –**at ~~$5,039 per node~~** (through 2022)

–**now $4,480 per node**

Next spring: Look for next CPU cluster based on "**Cascade Lake**" or AMD "**Rome**"

PURDUE UNIVERSITY

Interactive Data Science – Data Workbench



$300 per lab/year
37 labs using already!

PURDUE UNIVERSITY

ACCELERATED COMPUTING

FPGAS, MICs and GPUS

PURDUE UNIVERSITY®

Past Forays into Accelerators

- FPGA (Brutus)
- Small number of K20-era GPUs (Carter)
- Xeon Phi (Conte)
- Halstead-GPU PoC (2016-2017)

PURDUE UNIVERSITY

GPU Cards around Campus

It all comes full circle!

- Like with clusters, networks of GPU workstations are everywhere.
- For individual PIs, consumer GPUs are cost-effective for quite a bit of capability.

- **However – Nvidia GeForce license doesn't allow for data center use!**

This is the 2018 version of "building a cluster in your office" and having your grad student run it!

PURDUE UNIVERSITY

GPU Usage and Requirement Survey

Faculty Survey

- May 2018
- Over 50 invited to participate

- 20 responses

PURDUE UNIVERSITY

Survey Data

Top Challenges facing the research group

68	*Not enough resources (GPUs) for the group*
42	Power/Cooling issues
42	Software or system installation and maintenance
35	System Performance
31	Hardware maintenance
30	Cost-effectiveness
20	Access to Storage Capacity
0	Storage Performance

Top Priorities when Selecting GPU servers or workstations

32	Per-GPU performance
31	Total system performance
28	Lowest total price per system
26	Memory per GPU
17	Single Precision FLOPs
13	Double Precision FLOPs

GPU numbers, by GPU type

Tesla P100	40
Tesla V100	12
Tesla K80	4
Tesla K40	8
Tesla K40	0
GTX 10XX	89
Titan	124
Other	16
	293

GPUs per server

Mean of 4.4, stdev of 2.2

PURDUE UNIVERSITY

Survey Data

Server Counts

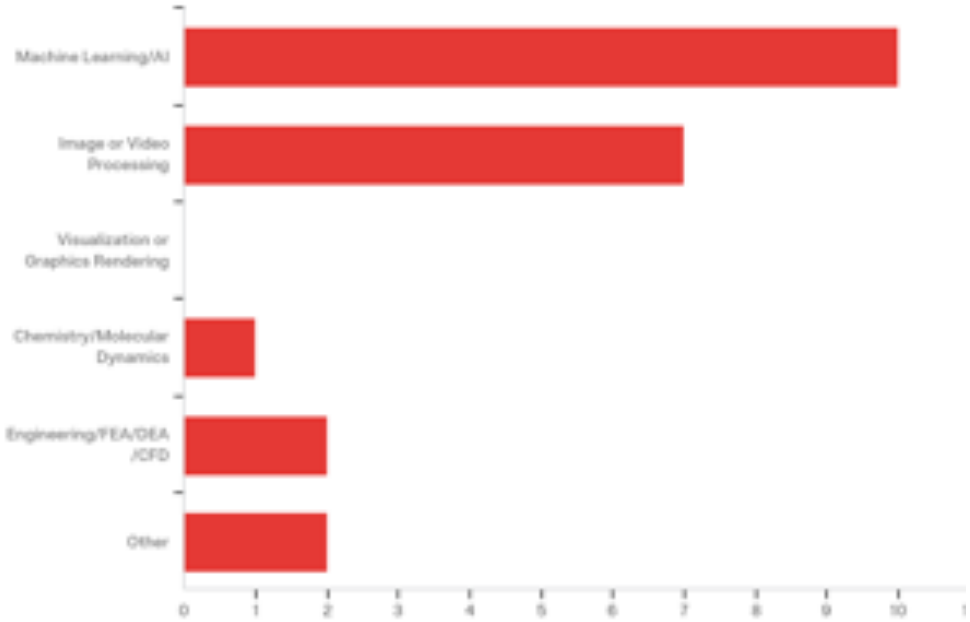
Lab-run: avg of 8.7 per PI, stdev of 7

IT-run: avg of 18 servers, serving 40 PIs

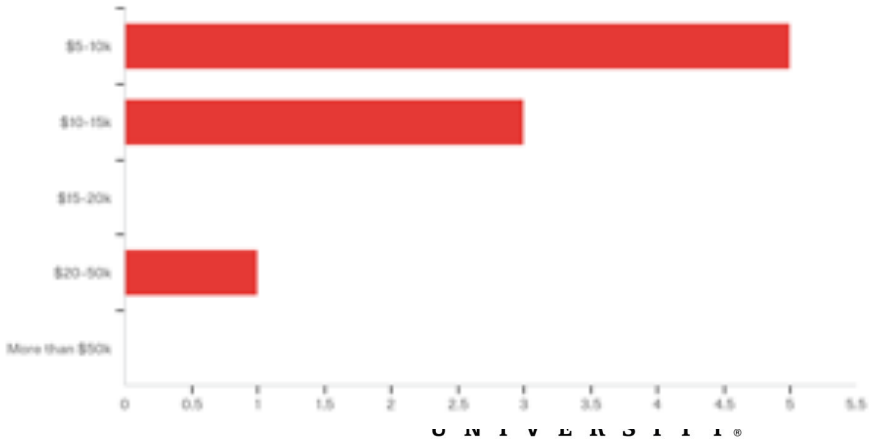
Memory Config per workstation/server

Answer	%
32 GB or less	9.09%
33-96 GB	27.27%
97-192 GB	18.18%
193-256 GB	0.00%
257-512 GB	9.09%
> 512 GB	36.36%
Total	100%

Apps run on GPU workstations/servers



Cost per Node on existing equipment



Key Takeaway – Campus Needs More GPUs

A Faculty Quote

"What is the point of having extensive studies of our strategy about Data Science when we don't even have [enough] computers???

The Biggest Challenge, from Survey:

Having sufficient GPUs for the lab

Also, facility problems, and
"taking care of all the computers" problems

PURDUE UNIVERSITY

2018 Community Cluster

Prof. Lillian Moller Gilbreth

Gilbreth

- GPU-based system ideal for machine learning, AI, **big data science** – as well as FEA, Chemistry, MD
  - ***.74 PF of computing!***
- Annual subscription for the lab's access to the entire cluster.

Hardware currently being received!

PURDUE UNIVERSITY

Bid Process

Open Bid Process – Requesting:

- Mellanox SwitchX-2 EDR for GPUDirect RDMA
- Intel Xeon Skylake Silver Processors
- 192GB of DDR4 Memory
- SSD boot drive, 250G or better
- Two Nvidia P100 16GB GPU per node

- Responses ranged from $14,100-14,700

PURDUE UNIVERSITY

Speeds and Feeds and Specs

Compute Node

- Base node: Dell R740
- 16-core node, 2.1-3.0 GHz Intel Xeon Sky Lake processors (Xeon Silver 4110)
  - 85w vs 125w
- 2x 16GB P100 GPUs per node
- 192 GB DDR4 memory
- Mellanox SwitchX-2 and ConnectX-4 EDR Infiniband interconnect
  - 100 Gbps
  - Converged fabric – IP traffic uses Infiniband rail

PURDUE UNIVERSITY

Combined with..

Existing GPU Hardware

- HalsteadGPU and BrownGPU nodes and owners will be merged into Gilbreth
  - (We'll reach out and work with you directly to migrate)

- .. PLUS 3 new dedicated model-training nodes, funded in partnership with OVPR
  - 4x 32GB V100 GPUs per node, with NVLink
  - 8TB local NVMe flash

- Caffe, TensorFlow, and CNTK are up to 3x faster with Tesla V100 compared to P100
- Up to 125 TFLOPS of TensorFlow operations per V100 GPU

PURDUE UNIVERSITY

All Told

GPU Resource Breakdown

- 2 GPU-enabled frontend nodes (1 P100 GPU each)
- 4 interactive queue nodes (8 P100 GPUs)
- 40 batch queue nodes (80 P100 GPUs)
- 3 Training Nodes (12 32GB V100 GPUs)
  - NVLink, NVMe flash

88 P100 GPUs
12 V100 GPUs
.74 PFLOPS

PURDUE UNIVERSITY

Speeds and Feeds and Specs

Storage

Brown (550 nodes)
- 3 PB Lustre
- 40 GB/sec bandwidth
- 400k IOPS

Gilbreth (~50 nodes)
- 2 PB Lustre
- 30 GB/sec bandwidth
- 200k IOPS
- Expandable to 10PB and 72GB/s

Gilbreth Flash Burst Buffer
- ~100 TB NVMe
- 60 GB/sec bandwidth
- 1M IOPS

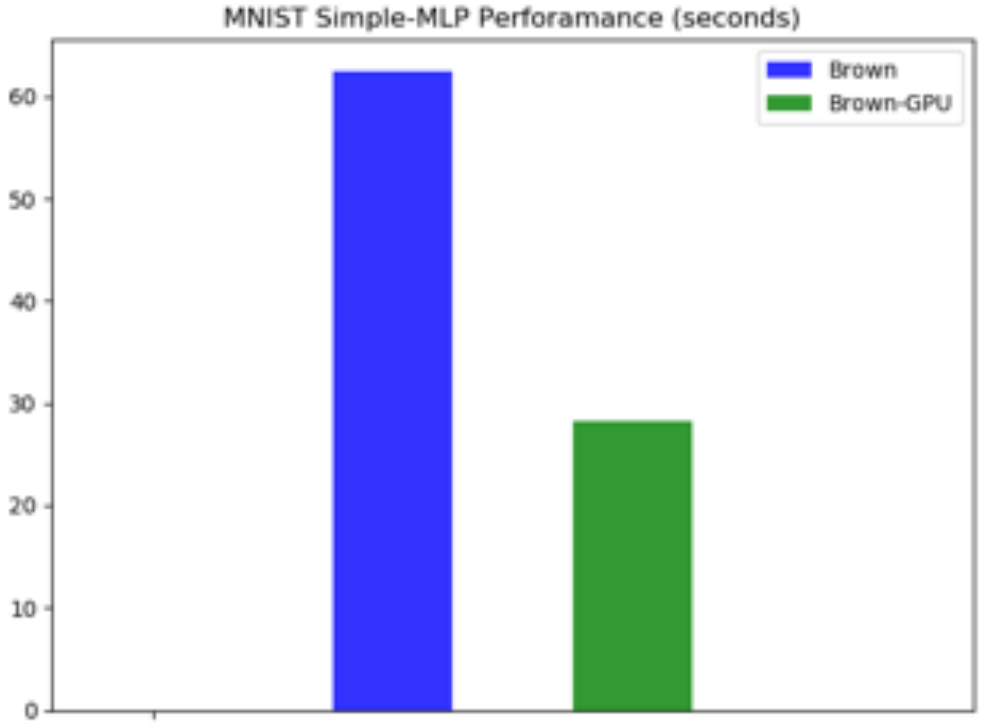
PURDUE UNIVERSITY

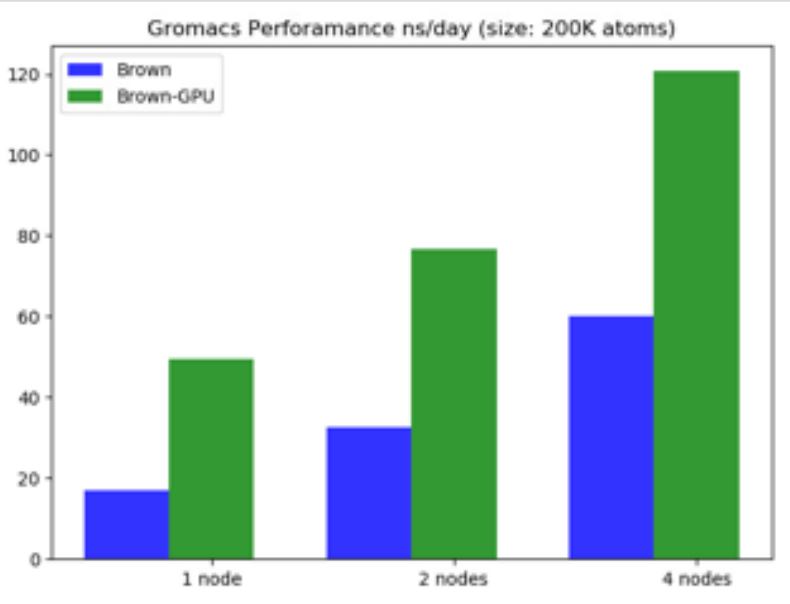
Application Performance

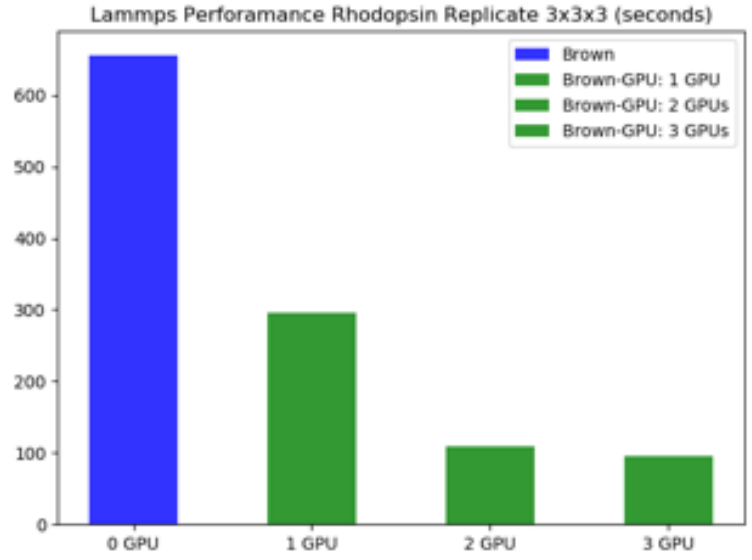
LAMMPS, GROMACS, Tensorflow, vs CPU

Remaining top code to benchmark:

- VASP







Gilbreth - Pricing

Costs

- Three ways to buy:

Buy Node (5y)	Buy Node (annual)	Shared Queue Subscription
$14,144	$2,829/year	$1,599/year

Shared queue model provides cheapest annual cost to get access to high-end GPUs! A proven model with Halstead and Brown.

Subscription is not a “node” purchase – access to a large pool of nodes!

PURDUE UNIVERSITY

Gilbreth - the Value Proposition

Community Cluster Guiding Principles

- You don't have to do the work
- You get to benefit from the pool of unused resources (get back more than what you put in)
- ITaP provides high-end networking, storage sufficient to drive these powerful systems
- Your lab doesn't have to **power, cool**, and **listen to** your GPU systems.
- ITaP manages the compute environment and software stacks so your students can get PhDs and write papers.

PURDUE UNIVERSITY

Gilbreth - The Value Proposition

Gilbreth Capabilities

- The big benefit is scale, for throughput!
  - You and your students will have access to a bigger pool of resources than you can do by yourself.
  - If you've used Brown-GPU or Halstead-GPU, please give us your feedback on scheduling set-up!

- Gilbreth's Pascal/Volta cards have more RAM than consumer cards.
- 200+ TB of Lustre scratch
- Additional shared flash
- Volta + nvlink training nodes
  - Local flash on training nodes
- GPU-enabled Jupyter notebooks
- CUDA 10
- Nodes dedicated to interactive use

Not a "node" purchase – it's access to a large pool of nodes!

PURDUE UNIVERSITY

Managed Software Stacks

ML Toolkit

- ml-toolkit-gpu/all/all
  ml-toolkit-gpu/caffe/1.0.0
  ml-toolkit-gpu/cntk/2.3
  ml-toolkit-gpu/gym/0.10.5
  ml-toolkit-gpu/keras/2.1.5
  ml-toolkit-gpu/opencv/3.3.1
  ml-toolkit-gpu/pytorch/0.4.0
  ml-toolkit-gpu/tensorflow/1.4.0
  ml-toolkit-gpu/tflearn/0.3.2
  ml-toolkit-gpu/theano/1.0.2

Non-CentOS Userland

- Bring your own Linux userland with Singularity containers!

  (Ubuntu, Slackware, Fedora, other)

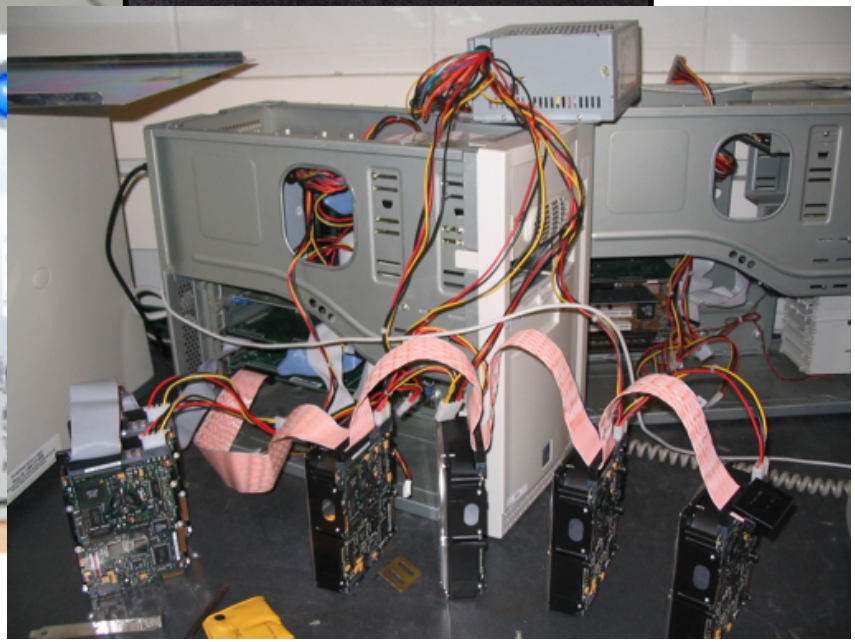
PURDUE UNIVERSITY

RESEARCH DATA

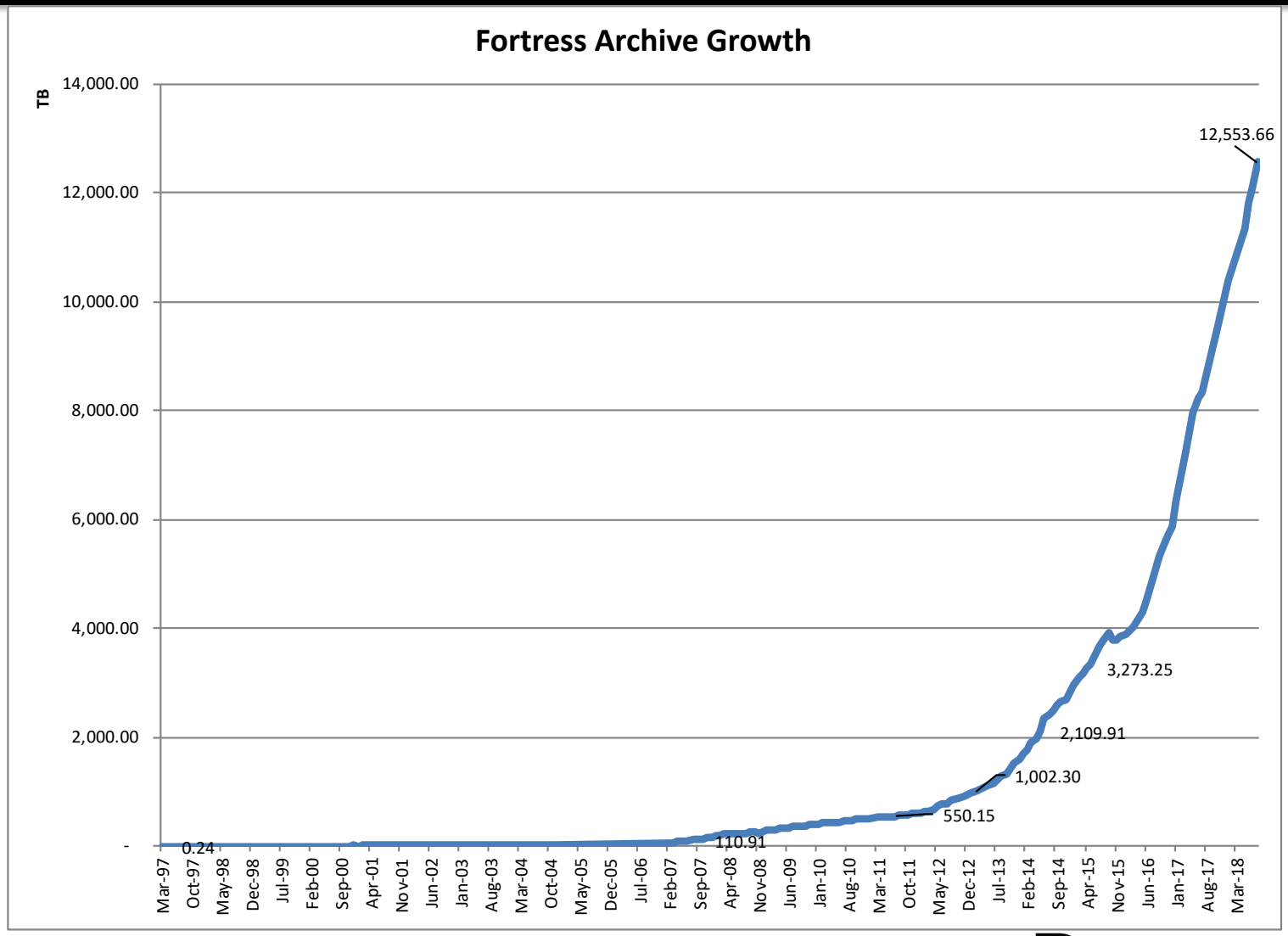
PURDUE
UNIVERSITY®

On-Campus Research Storage Solutions

Left to fend for themselves, researchers will find a way



Growth in Research Data Storage



PURDUE UNIVERSITY

The Research Data Depot

Impact

- Over 550 research labs are Depot partners!
  - ***60% are not HPC users!***
  - *Thousands of individual users*
- Over 2 PB sold
- A research group purchasing space has purchased, on average, nearly 10 TB.

Refresh/expansion underway to extend Depot for 5 more years!
Further enhancement (performance, scalability) to non-HPC users.

PURDUE UNIVERSITY

The Research Data Depot

$70

At ~~$75~~ per TB/year

- Storage oriented around the PI's research lab, with
  - Snapshots
  - Multi-site, active copies of data
  - Easy ways to do common research data management patterns
  - A scalable, expandable storage resource optimized for HPC
- Access to Globus data transfer service, and endpoint sharing
- ***Our goal: enabling the frictionless use and movement of data***

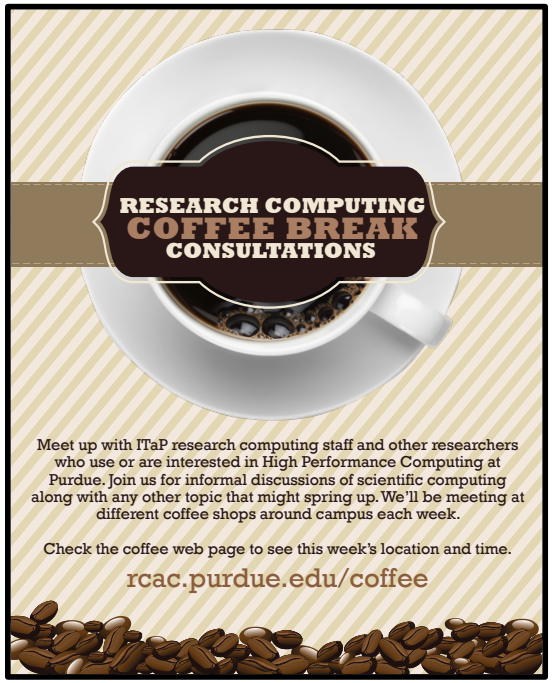
Expansion: now with a tertiary copy of data for near real-time disaster protection, for less $$$ per TB!

PURDUE UNIVERSITY

Education, Training, Expertise

Our key value – computational experts

- New faculty orientations
- One-on-one consultations
- UNIX, MPI, HPC, Python, R training offerings
- Cyberinfrastructure seminars



software carpentry

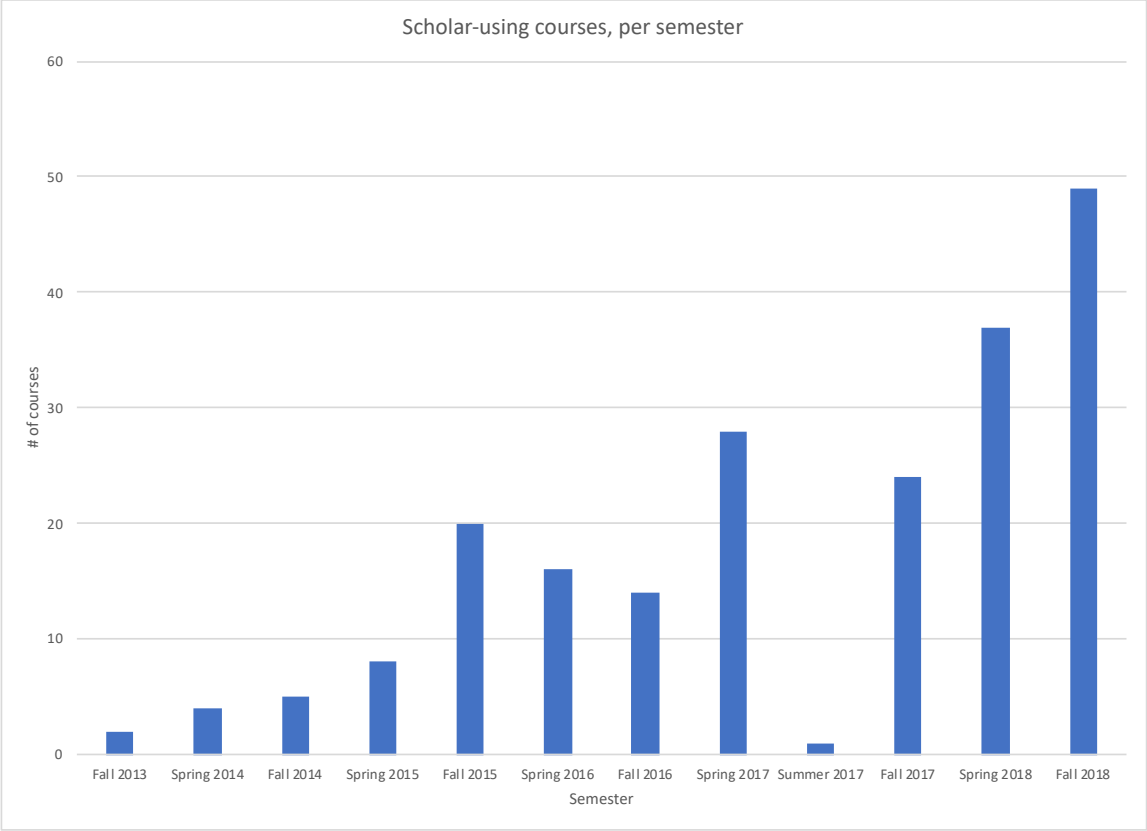
ADVANCED DOMAIN EXPERTISE

Chemistry
Physics
Astrophysics
Earth and Atmospheric Sciences
Computer Science
Chemical Engineering
Electrical and Computer Engineering
Cell and Molecular Biology
Entomology

PURDUE UNIVERSITY

Scholar – HPC and Data Science Education

Dedicated Cluster for HPC and Data Science



Dept	# of Courses
AAE	6
ABE	2
AGRY	3
ANSC	1
BIOL	4
BME	1
CGT	1
CHM	5
CNIT	3
CS	6
EAPS	17
ECE	3
EEE	1
FS	1
HORT	1
IE	1
LIBR	1
ME	5
MGMT	9
NUTR	1
PHYS	1
STAT	9

2035 students using Scholar to learn HPC and data science this semester!

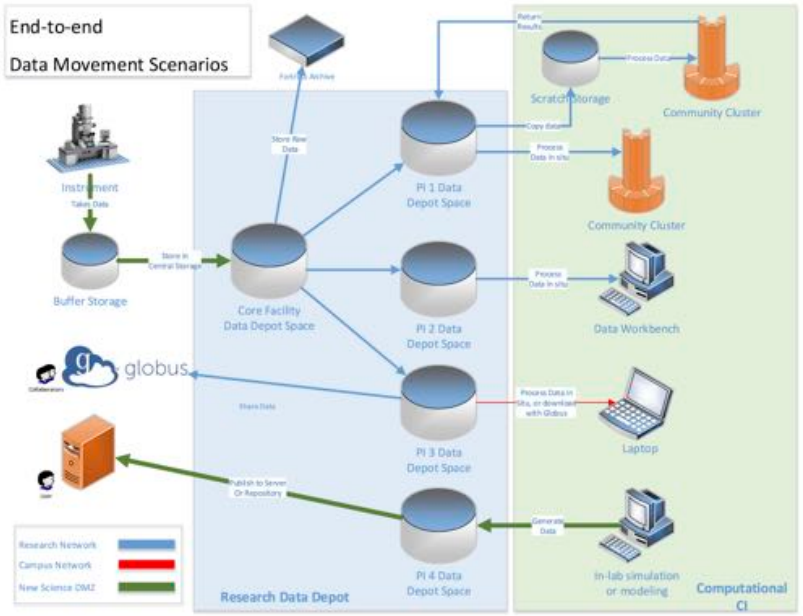
Interactive GPUs now available for instructional use!

PURDUE UNIVERSITY

UPCOMING ENHANCEMENTS

PURDUE UNIVERSITY®

Science DMZ to Data-Intensive Instruments



NSF CC* Grant

- Extend science DMZ to big data instrument “hot-spots”
- Create a cost-effective, sustainable architecture for research data networks.
- Develop research computing professionals.

The network is one part of all the tools we provide for supporting data-intensive instrumentation!

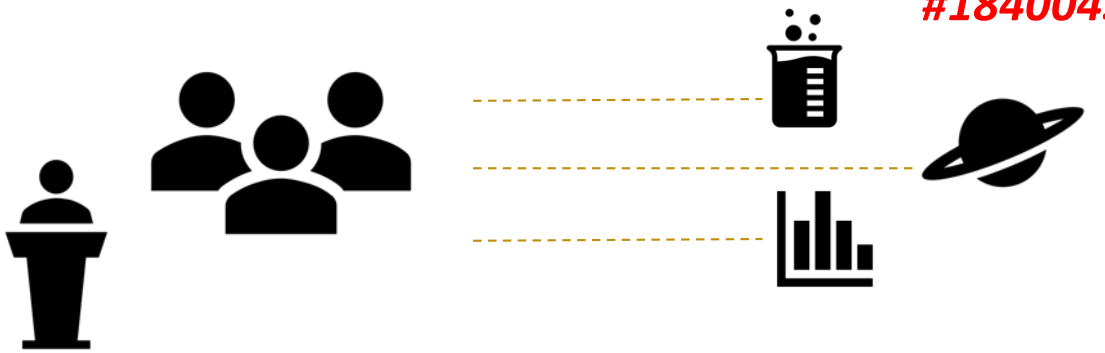
PURDUE UNIVERSITY

Cybersecurity Framework for Regulated Data

REED+

A managed research ecosystem with sufficient storage, high speed computing capability and security to efficiently and cost effectively handle Purdue's controlled research data and processing needs in a manner compliant with the highest level of cybersecurity applicable to Unclassified data possessed by Purdue University and Purdue University researchers.

2 year, $600k NSF Award #1840043 to build framework



PURDUE UNIVERSITY

Changes for Off-Campus Access to Clusters

BOILERKEY
TWO-FACTOR AUTHENTICATION
DUO

VPN Required for Off-Campus

- SSH
- Thinlinc remote desktops
- CIFS access to cluster scratch

PURDUE
UNIVERSITY

THANK YOU

www.rcac.purdue.edu
psmith@purdue.edu