2017 COMMUNITY CLUSTER

PURDUE UNIVERSITY

Preston Smith

Director of Research Computing Services

Alex Younts

Senior Research Engineer

8/29/2017

HPC FACULTY MEETING

COMPUTATION

RESEARCH COMPUTING STRENGTH

Since Steele in 2008, Research Computing has deployed many world-class offerings in computation

PURDUE UNIVERSITY

9 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

26 departments, 60 faculty investors

#54 on June 2012 Top 500

CONTE

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

Installed August 2013

26 departments, 62 faculty investors

#28 on June 2013 Top 500

DATA DEPOT

2.5 PB of disk storage
Installed Nov. 2014

400+ faculty investors from every academic college

RICE

13,200 cores

Installed May 2015

23 departments

69 faculty investors

HALSTEAD

10,160 cores

Installed December 2016

25 departments

62 faculty investors

Halstead-GPU Expansion
May 2017

PURDUE
UNIVERSITY

PARTNERS

287M hours delivered in 2016

Almost 200 investors from 36 departments, from every College, and 3 Purdue campuses

The gold standard for condo-style computing

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

A selling point to attract people to Purdue!

PURDUE UNIVERSITY

FACULTY DEMOGRAPHICS

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

BID PROCESS

Open bid process requesting:

- Quantity 500 nodes
- Included various interconnects (EDR/HDR Infiniband, OmniPath)
- “Broadwell” or “Sky Lake” processors, 2.4 GHz or better
- At least 128GB memory per node
- SSD boot drive, 250G or better
- Optional uplift for one Nvidia P100 GPU per node

Responses ranged from $4,400-4,800 for a node like Halstead, or $6,000-7,700 per node for the latest Intel CPU.

Vendor	Cores/Node	CPU	GB RAM/Node	$/node	RDMA Network	$ per GF
Vendor 1	24	Sky Lake	192	$6,050.00	EDR IB	$3.03
Vendor 2	24	Sky Lake	192	$7,675.13	EDR IB	$3.88
Vendor 3	24	Sky Lake	192	$7,754.82	EDR IB	$3.84
Vendor 4	20	Broadwell	128	$4,756.70	EDR IB	$6.19
Vendor 5	20	Broadwell	128	$4,456.84	EDR IB	$5.80
Halstead	***20***	***Haswell***	***128***	***$3,600.00***	***EDR IB***	***$4.80***

PURDUE UNIVERSITY

FINAL NODE SPECS

BID RESULTS

Base node: Dell R640

- 24-core node, 2.6 GHz Intel Xeon Gold “Sky Lake” processors (Xeon Gold 6126)
    - Higher node price, but 50% faster processors
    - 32 Flops per cycle!
- 96 GB DDR4 memory
    - 384 GB, 768 GB & 1536 GB options
    - Memory prices are high world-wide
    - (More on this later!)
- EDR Infiniband interconnect
    - 100 Gbps, 3:1 fat tree – very similar in speed to Halstead
    - Converged fabric – IP traffic uses Infiniband rail

Also bid systems for provost equipment program, leading to substantial savings!

PURDUE UNIVERSITY

STORAGE

CPU IS ONLY PART OF THE PICTURE

Halstead

- 1 PB GPFS
- 22 GB/sec bandwidth
- 100k IOPS

Brown

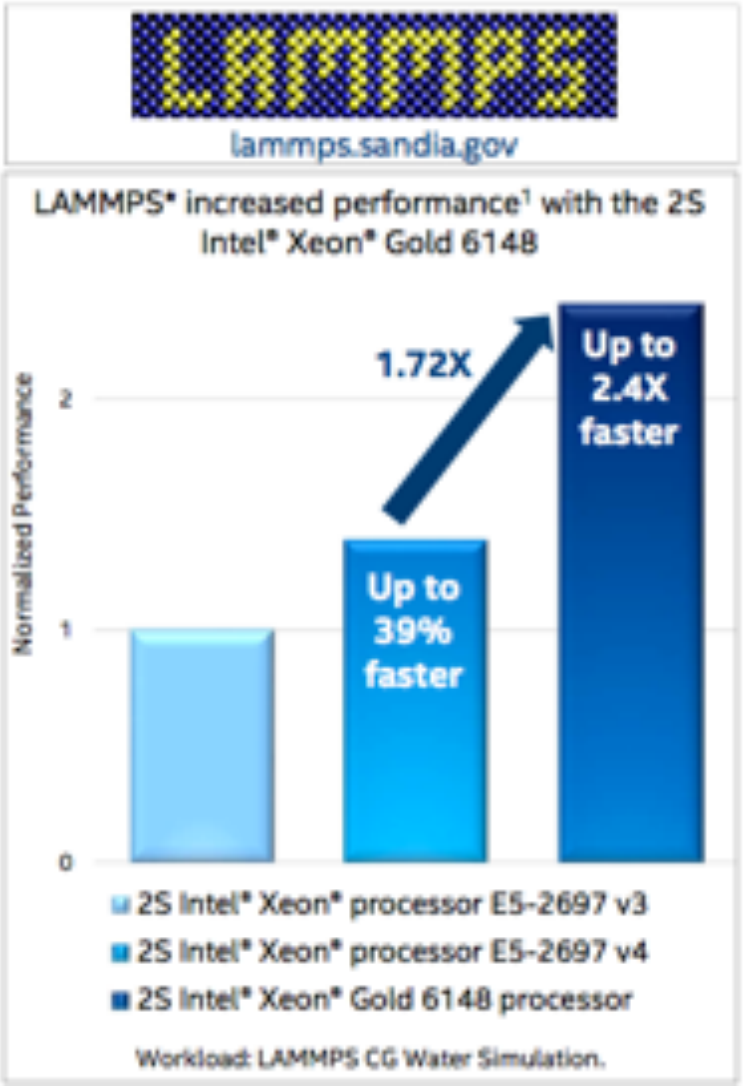
- 3 PB Lustre
- 40 GB/sec bandwidth
- 400k IOPS

THE BOTTOM LINE

XEON GOLD VS HASWELL (HALSTEAD)

For codes used by community cluster partners

- Ansys Fluent: 1.6x faster
- Converge CFD: up to 1.29x faster
- Gaussian: 1.25x faster
- LAMMPS: Up to 2.4x faster
- GROMACS: Up to 2x faster
- VASP: Up to 1.9x faster
- AMBER: Up to 1.73x faster
- NAMD: 1.67x faster
- HOMME: Up to 1.67x faster
- WRF: Up to 1.41x faster
- HS06: 675.27



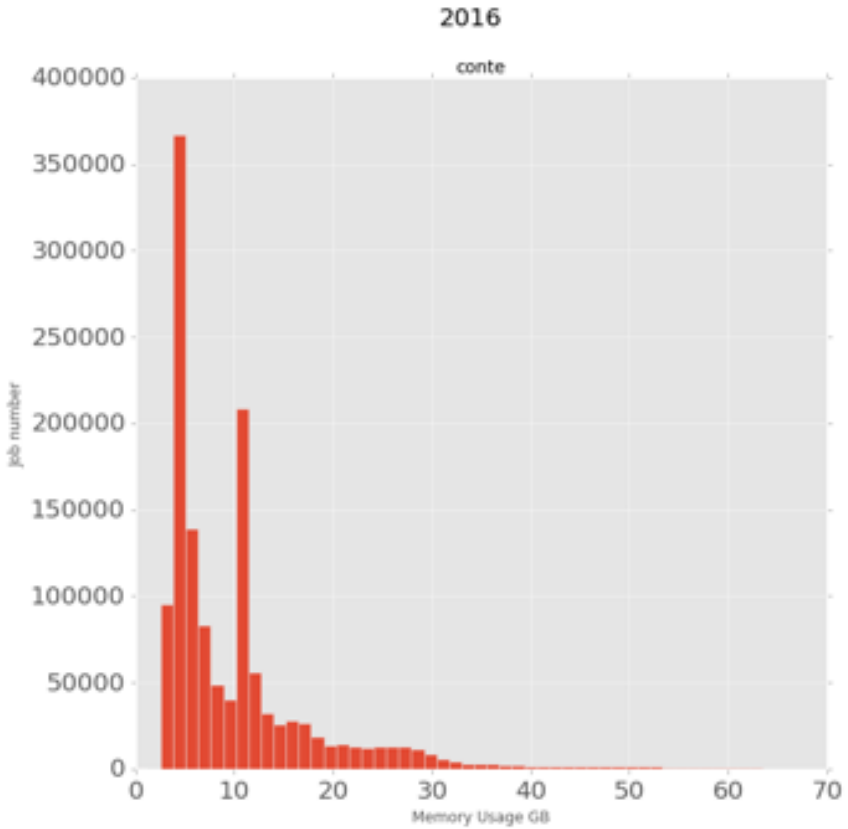
PURDUE UNIVERSITY

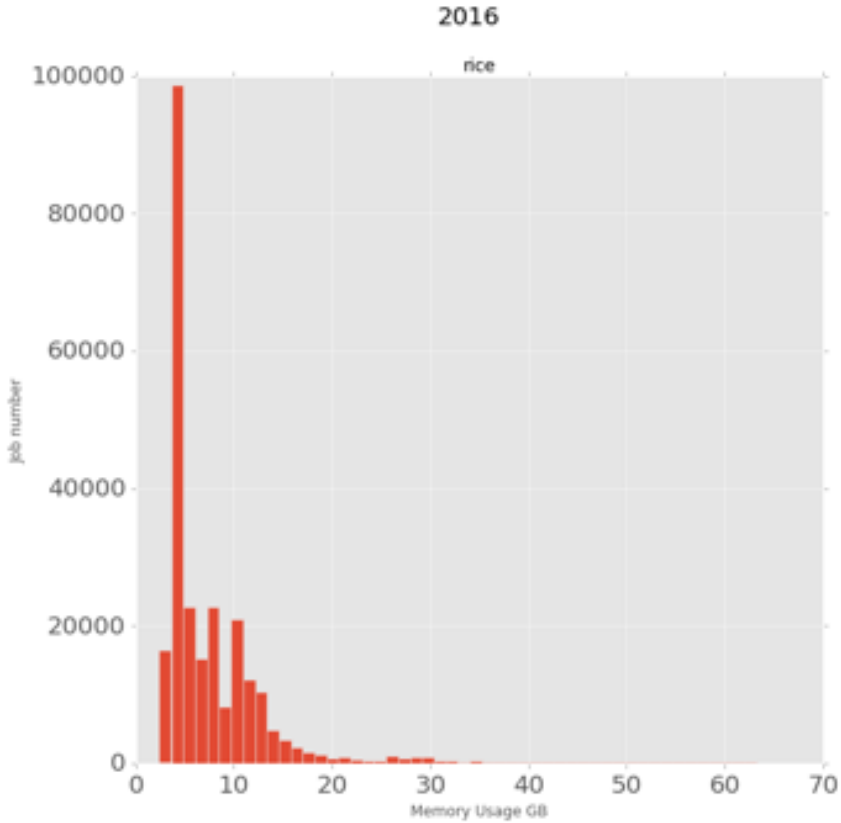
MEMORY

WHAT IS THE IDEAL AMOUNT?

Halstead had 128G RAM per node -

Why 96GB for Brown?





Cluster	95th Percentile	99th Percentile
Rice	17.2	36.6
Conte	27.3	44.7
Snyder	184	510.4

GB/node per job

PURDUE UNIVERSITY

FOR THE MAJORITY-

Brown: A traditional HPC system

The same, familiar model:

- New cluster acquisition every year
- Service through October of 2022

Improved floating point performance
Vastly improved I/O subsystem

Nobel Laureate
Herbert C. Brown

Great for most science and engineering codes

550 Nodes of Brown:
1.1 PF peak

Equivalent to all of Conte, including Xeon Phi Accelerators!

PURDUE
UNIVERSITY

OPTIONS

NODE PRICES

Base node option, **plus**

- Node with 3 Nvidia GPUs
- Large memory options, also with 768 GB and 1.5 TB configurations available.

96G Brown Node	384G Snyder Node	3-GPU Brown Node
$5,599.00	$9,500.00	$18,325.00

PURDUE UNIVERSITY

GPUS

GPU PARTITION

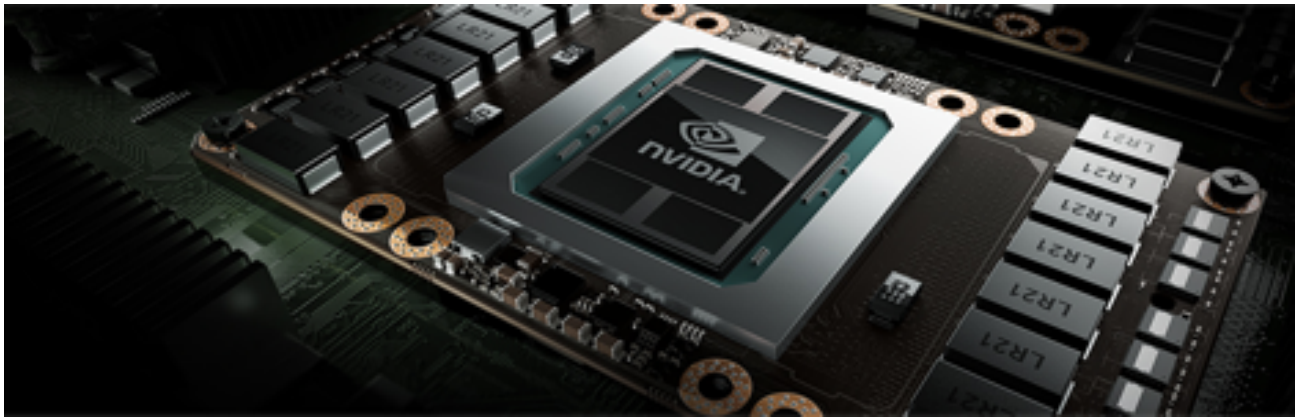
Requests for GPU for accelerating simulations, computer architecture research, or deep learning.

12 researchers are currently evaluating Halstead's GPU partition

With Brown:

Plans are for a 16-node GPU partition, with 3 Nvidia P100 GPUs per node.

250 TF of GPU!

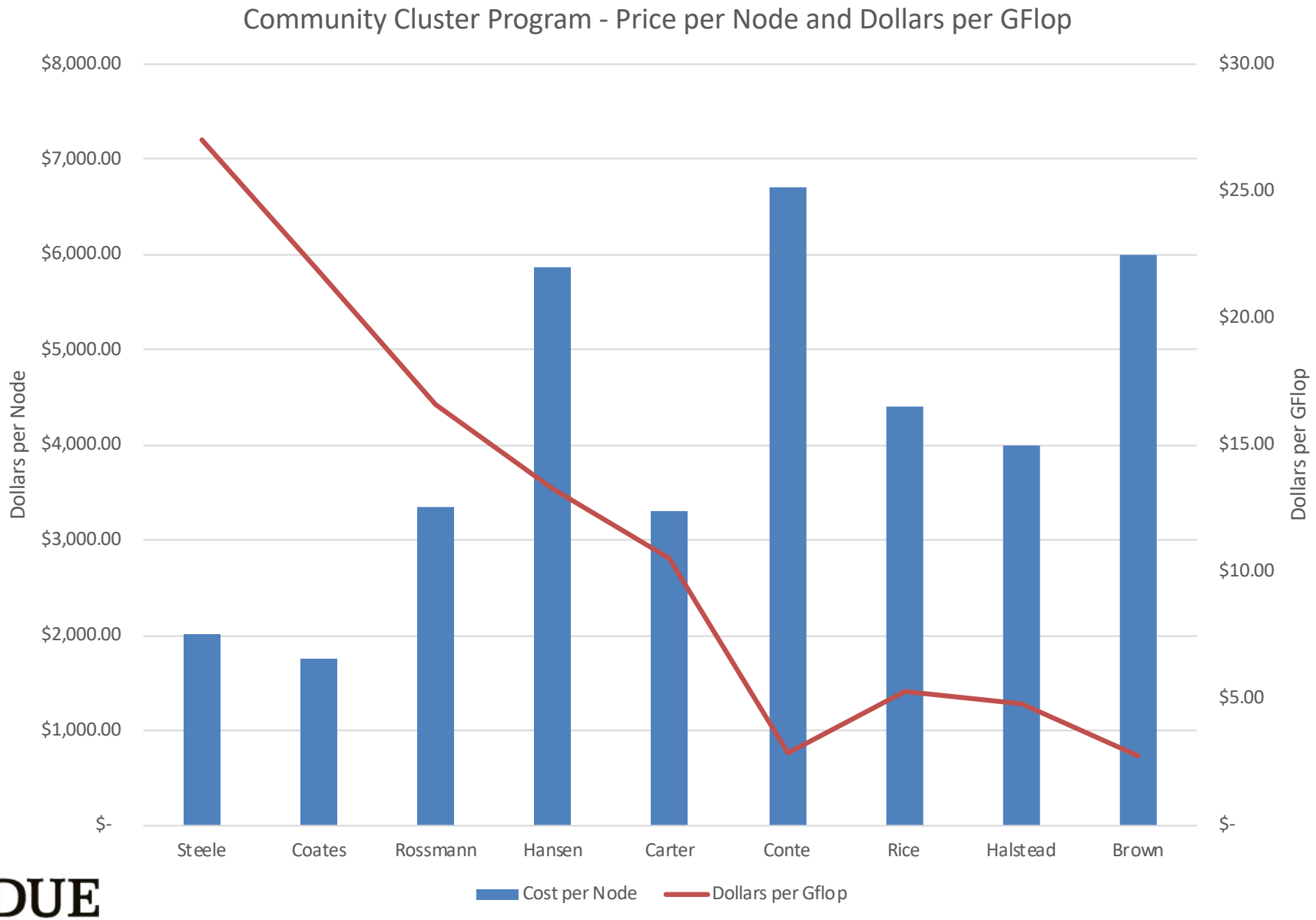


PURDUE UNIVERSITY

Annual subscription to access to GPU nodes:
$2,500 per year

INCREASING VALUE

TRENDS



PURDUE UNIVERSITY

IMPLEMENTATION

General implementation schedule:

- Facilities preparation already in progress - in the POD modular datacenter.
- Plan is for early access and burn-in early October
- Top 500 benchmark run week of Oct 16

- General availability Oct 30

- ***No install event this year.***

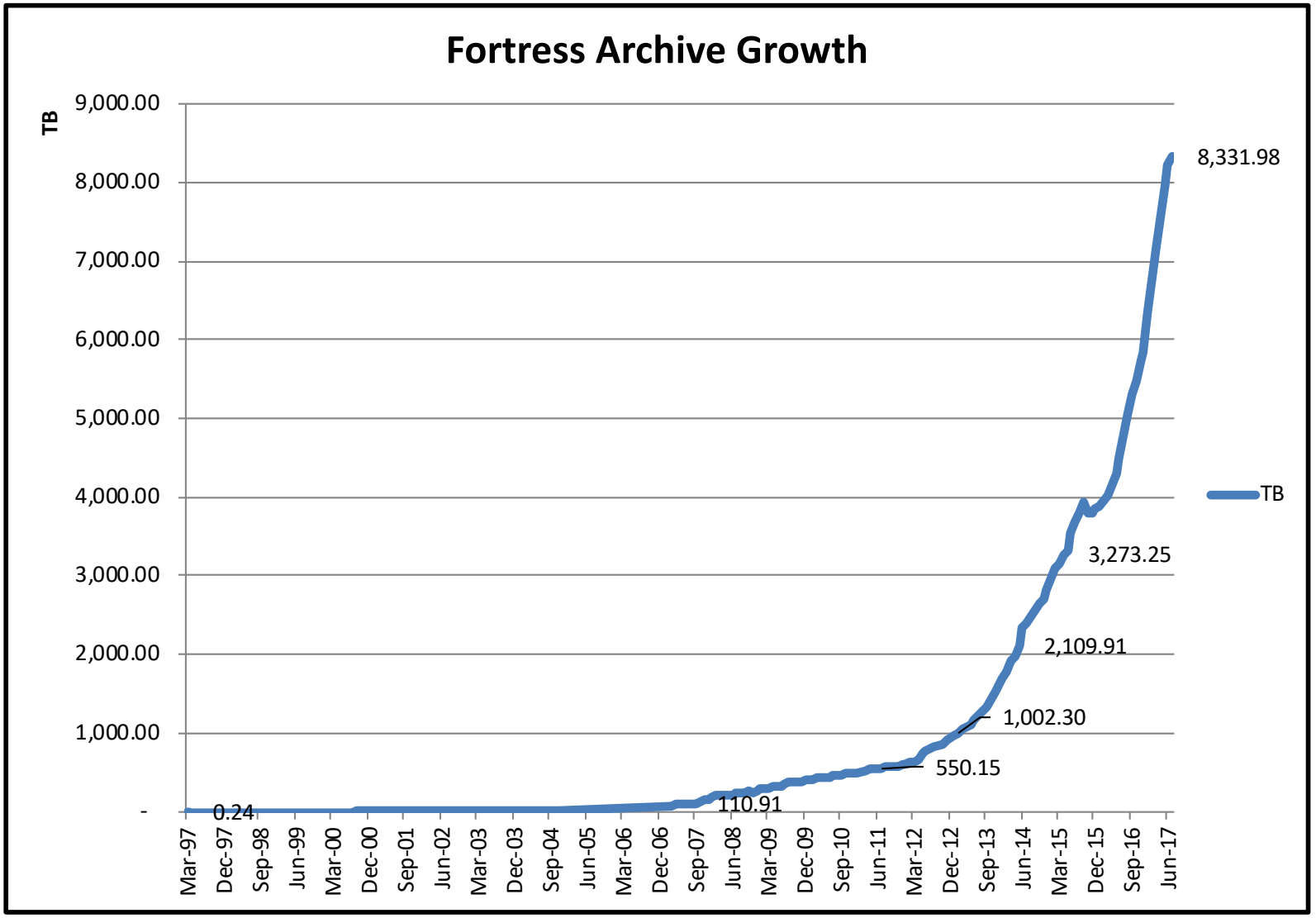
PURDUE UNIVERSITY

PURDUE
UNIVERSITY

RESEARCH DATA STORAGE

DATA IS HUGE!

AND GROWING



PURDUE UNIVERSITY

RESEARCH DATA DEPOT

THE 2014 HPC SYSTEM

New price!

At $75/TB per year:

- Storage oriented around your research lab, with
  - Snapshots
  - Multi-site copies of your data
  - Disaster protection
  - A scalable, expandable storage resource optimized for HPC
- Access to Globus data transfer service, and endpoint sharing

PURDUE UNIVERSITY

GET ACCESS

GIVE IT A TRY!

To buy 1 or more TB of space,

Or to set up a trial for your lab

Order online:

https://www.rcac.purdue.edu/purchase/

PURDUE UNIVERSITY

A hit!

- Over 400 research labs are Depot partners!
  - ***Many are not HPC users!***
  - *Thousands of individual users*
- 1.25 PB sold
- A research group purchasing space has purchased, on average, nearly 10 TB.
- Other institutions looking to Purdue for leadership from our Depot storage service

GLOBUS

STATISTICS

Transfer and share large datasets….

…. With dropbox-like characteristics ….

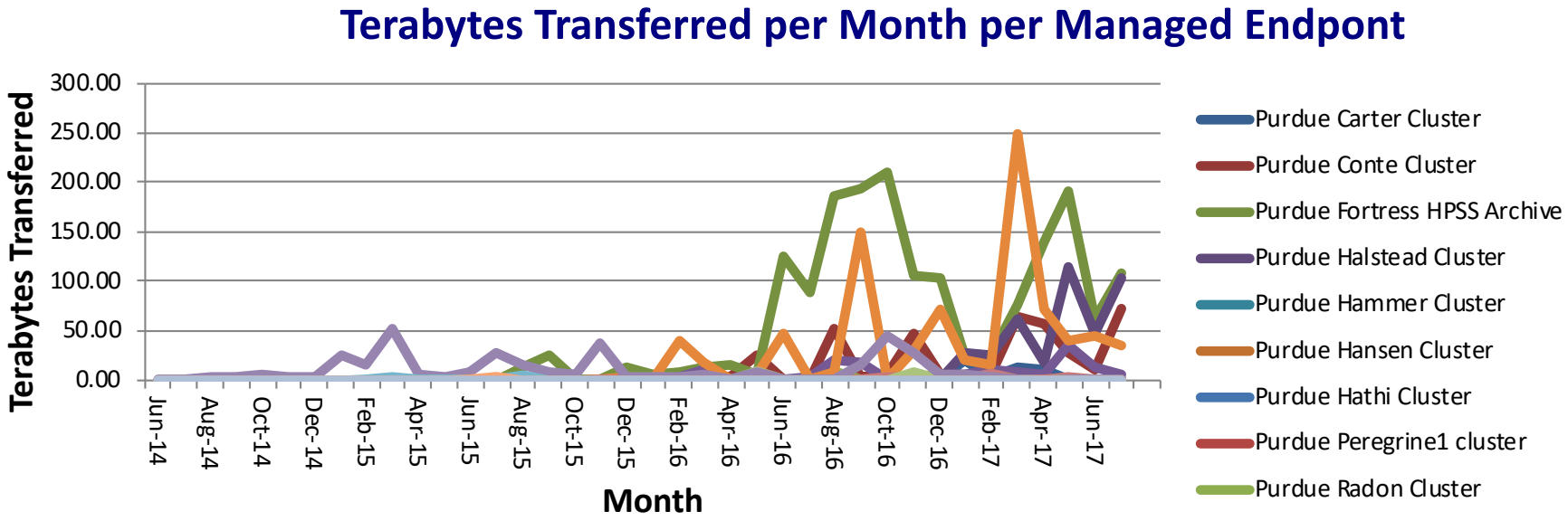
https://transfer.rcac.purdue.edu

…. *Directly from your own storage system!*

PURDUE UNIVERSITY

GLOBUS

ADOPTION



Fortress: 1.74PB since Aug ‘15!

- Last 12 months:
- 2.4 PB transferred (388TB in March!)
- Average of 185 TB, 84 unique users per month

PURDUE UNIVERSITY

PURDUE
UNIVERSITY

RESEARCH SERVICES

OTHER SERVICES YOU MIGHT BE INTERESTED IN

GROWTH

NON-TRADITIONAL HPC RESEARCHERS

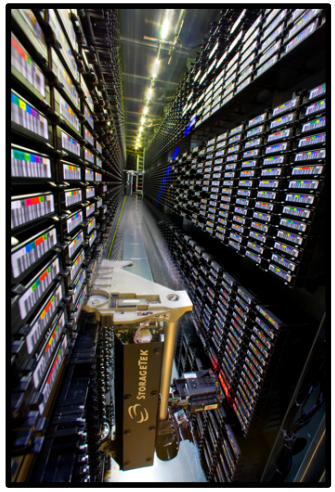
College	2014	2017	Growth 2014-2017	Growth Rate 2014-17
Agriculture	48	111	63	131%
Engineering	161	265	104	65%
Science	199	227	28	14%
Education	1	7	6	600%
Liberal Arts	1	9	8	800%
Management	20	24	4	20%
Pharmacy	5	9	4	80%
Polytechnic	13	21	8	62%
Heath and Human Sciences	14	28	14	100%
Veterinary Medicine	0	5	5	

PURDUE UNIVERSITY

DATA WORKBENCH

WE HAVE THE PARTS

But not wrapped up in a package



globus

git

PURDUE UNIVERSITY

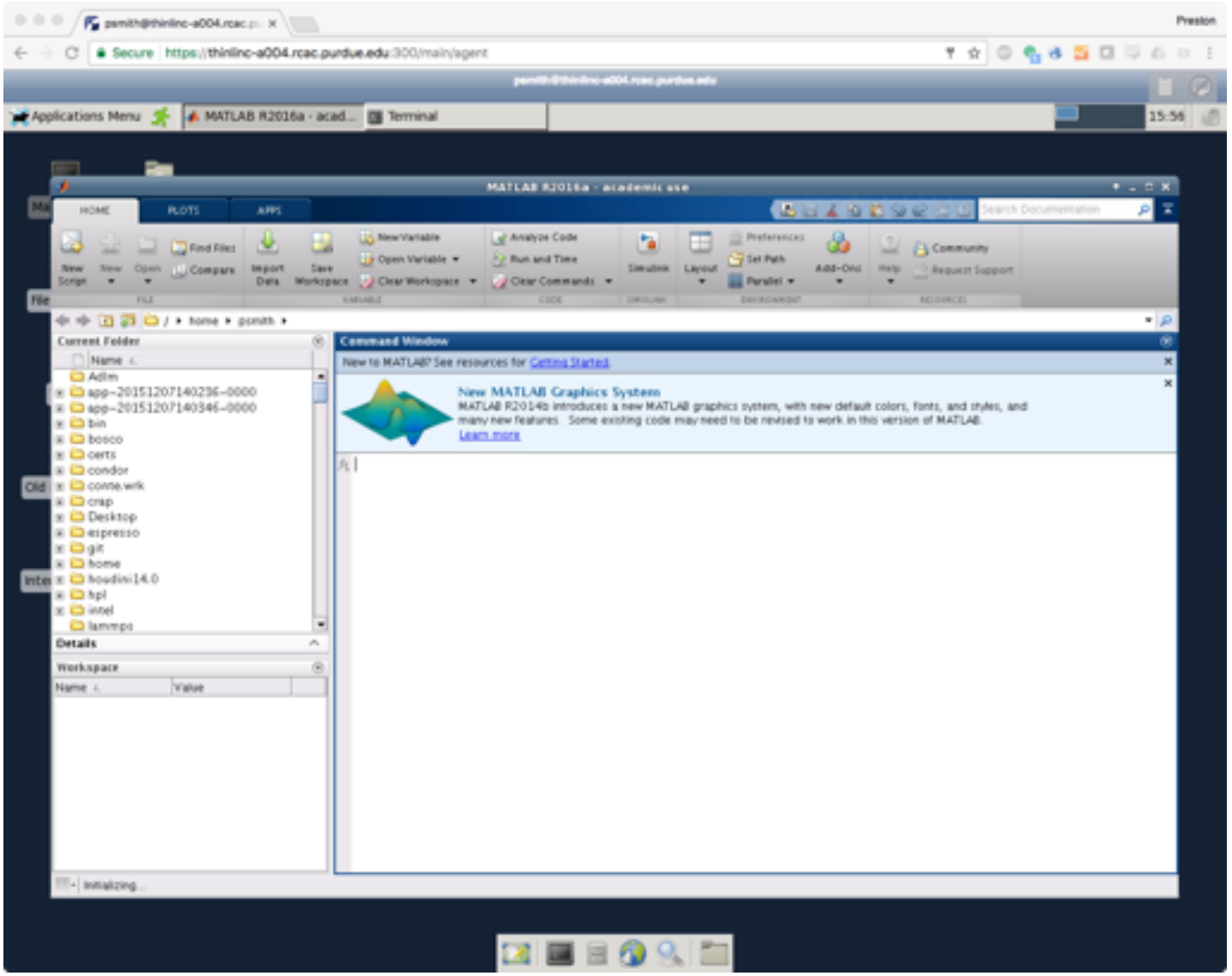
MySQL

APACHE Spark

hadoop

COMPUTATIONAL DESKTOPS

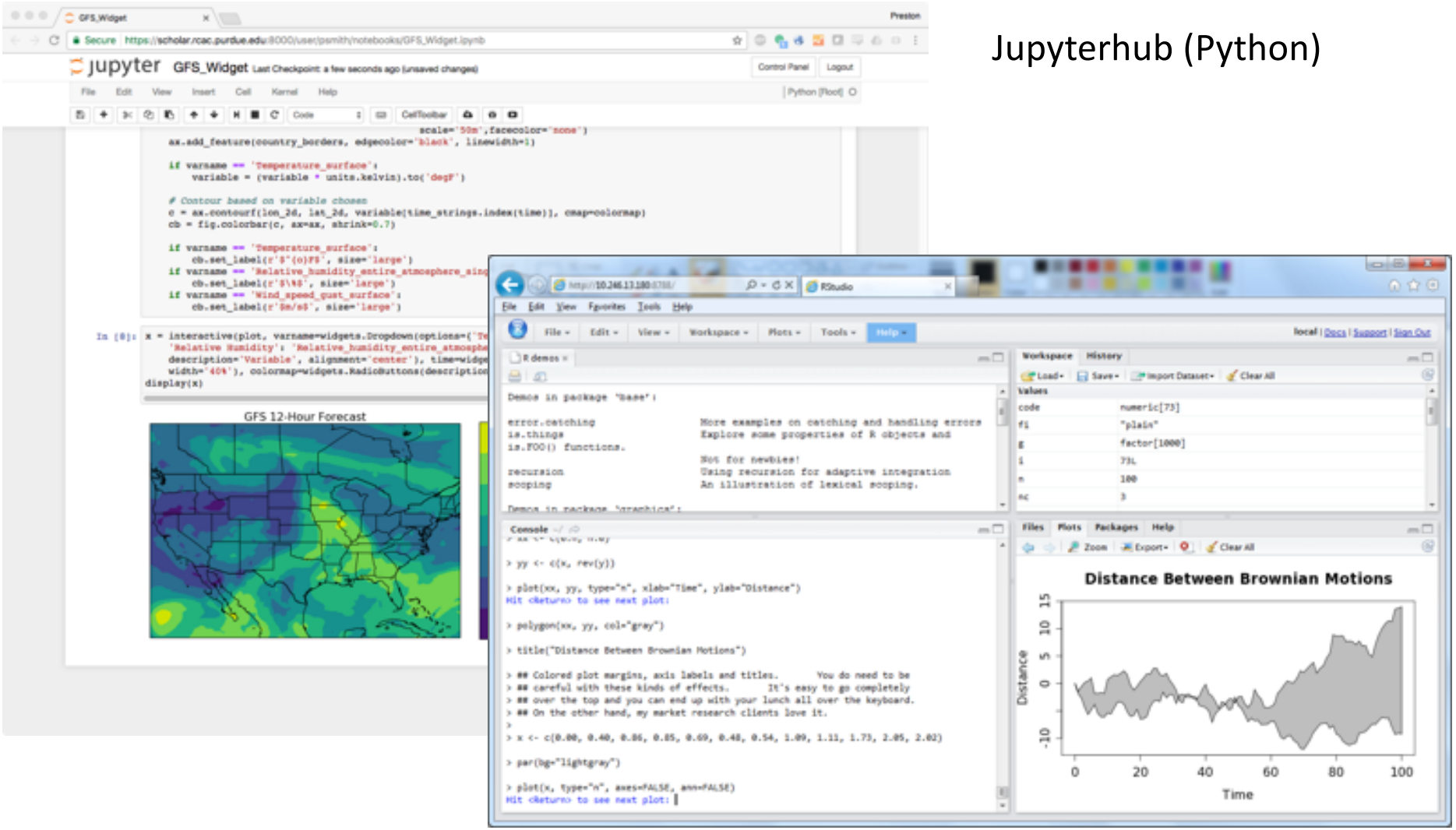
THINLINC



PURDUE UNIVERSITY

DATA ANALYTICS

IN THE BROWSER



Jupyterhub (Python)

Rstudio

PURDUE UNIVERSITY

DATA WORKBENCH

FEATURES

- Approx. $150 annual charge for access
- Easy access to data analysis tools
- Run virtual private Windows desktops
- Run virtual private Linux images in containers
- Integrate with RCAC services
  - Depot, Fortress, Globus, Github, Self Service, etc.
- Grow to batch HPC as your needs grow!
- Add same interactive capabilities to community cluster frontends

PURDUE UNIVERSITY

COMPUTATIONAL SUPPORT

SCIENCE EXPERTISE

DOMAINS

Chemistry

Physics

Astrophysics

Earth and Atmospheric Sciences

Computer Science

Chemical Engineering

Electrical and Computer Engineering

Cell and Molecular Biology

Agriculture

APPLICATION SPACES

Molecular Dynamics

Image Processing

Quantum Chemistry

Weather Modeling

Machine Learning

Big Data

Computer Architecture

Finite Element Analysis

Statistics

Bioinformatics

Geospatial

Remote Sensing

Visualization

PURDUE UNIVERSITY

NEED HELP?

COFFEE BREAK CONSULTATIONS

- Hard to solve problems with HPC?
- Need help building your software or optimizing your workflow?
- Need to learn what resources are available?

PURDUE UNIVERSITY

SCHOLAR

HPC FOR INSTRUCTION

- Need to teach students to use HPC or work with big data in a course?
- Scholar cluster is available to any instructor at no cost.

Spring 2016:	EAPS
CS	AAE
STAT	ANSC
CHEM	ME

Just add your CRN

PURDUE UNIVERSITY

VERSION CONTROL

COLLABORATE ON CODE OR DOCUMENTATION

Local-to-Purdue Github repositories for your lab, managed by you!

PURDUE UNIVERSITY

WOMEN IN HPC

WHPC WOMEN IN HIGH PERFORMANCE COMPUTING
PURDUE UNIVERSITY

Fall kick-off meeting Wednesday, August 30th 12:00-1:00PM in RAWL 2082.

Our invited speaker: Dr. Beth M. Holloway, Director of the Women in Engineering program and Assistant Dean of Undergraduate Education, College of Engineering.

Join WHPC for computing scholarship, networking, and mentorship opportunities!

PURDUE UNIVERSITY

TRAINING

COMING EVENTS

Big data research with Hadoop and Spark XSEDE Workshop

September 12 & 13, 2017 11:00am – 5:00pm

Unix 101

September 19 & 21, 2017

Jetstream cloud resource for science and engineering

September 20, 2017 1:30pm – 4:30pm

Unix 201

October 3 & 5, 2017

XSEDE HPC Monthly Workshop Series - MPI Day 1

October 3& 4, 2017 11:00am – 5:00pm

Clusters 101

October 17& 19, 2017

XSEDE HPC Monthly Workshop Series - GPU Programming Using OpenACC

November 7, 2017 11:00am – 5:00pm

Upcoming but not yet scheduled:

NVIDIA GPU - Deep learning and CUDA programming
Software Carpentry – Programming with Python
Software Carpentry - Programming with R
Software Carpentry - R for Reproducible Scientific Analysis
Software Carpentry - Programming with MATLAB

PURDUE UNIVERSITY

THE END

QUESTIONS?

rcac.purdue.edu/purchase

PURDUE UNIVERSITY

SPARE SLIDES

EVERYTHING PAST HERE IS SPARE

PURDUE
UNIVERSITY