

# RESEARCH COMPUTING



# COMPUTING AND DATA SERVICES

Fall 2019

# RESEARCH COMPUTING AT PURDUE UNIVERSITY

## ABOUT THE COMMUNITY CLUSTERS

Information Technology at Purdue (ITaP) operates a significant shared cluster computing infrastructure developed over several years through focused acquisitions using funds from grants, faculty startup packages, and institutional sources.

These “community clusters” are now the foundation of Purdue’s research cyberinfrastructure.



We welcome any Purdue faculty or staff with computational needs to join this growing community and enjoy the enormous benefits this shared infrastructure provides:

### PEACE OF MIND    LOW OVERHEAD    COST EFFECTIVE    FLEXIBLE

<b>Peace of Mind</b>	ITaP system administrators take care of security patches, software installation, operating system upgrades, and hardware repair so faculty and graduate students can concentrate on research. Research support staff are available to support your research by providing consultation and software support.
<b>Low Overhead</b>	The ITaP provides all infrastructure: racks, floor space, cooling, power, networking and storage are added value included with the Community Cluster Program at no charge. In addition, each cluster is built with a lifespan of five (5) years, with free support the life of the cluster.
<b>Cost Effective</b>	ITaP works with vendors to obtain the best price for computing resources by pooling funds from different disciplines to leverage greater group purchasing power and provide far more computing capability at less cost than would be possible with individual purchases. Through the Community Cluster Program, partners have invested several million dollars in computational and storage resources since 2006.
<b>Flexible</b>	Partners in a community cluster always have ready access to the capacity they purchase and potentially to much more, if they need it. The Community Cluster Program shares compute nodes among cluster partners when the nodes are idle. This allows each partner to get more computational value per dollar than could be on his or her own.

## ADDITIONAL BENEFITS

- **Parallel Filesystem:** Access to large-scale, high-performance, parallel scratch for running jobs
- **Archive:** Access to the high-performance HPSS Archive system “Fortress”, for long-term storage of research data
- **Research Data Depot:** High-performance, expandable space is available to any research group to:
  - Share data and results among your group, or with collaborators, using Globus transfer service
  - Centrally install and manage the group’s applications
  - Define and manage access to custom UNIX groups for easy project-based collaboration
- **Cloud Lab Folders:** Centralize your lab’s documents and collaborate in a managed folder in Box.com.
- **Version Control:** Self-managed Purdue-hosted Github repositories for documents and source code
- **Remote Desktops:** Access community cluster systems via user-friendly Thinlinc Remote Desktop connections.
- **Notebooks:** Work in Python notebooks on cluster resources, for reproducible, shareable data analysis.
- **Cluster Science Gateway:** Access clusters, files, and applications from your browser using Open OnDemand.

# RESEARCH COMPUTING AT PURDUE UNIVERSITY

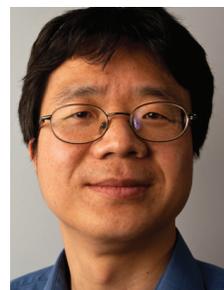


**Jeffrey Greeley**

Professor of  
Chemical Engineering

“Knowing there was a good group of experienced professionals I could rely on for support and establishing the computational infrastructure that I needed was very comforting when I was considering coming to Purdue. It frees up my time and the time of my graduate students and post-docs. We can focus on the scientific problems, which are our primary interest.”

“We want to solve the virus structures at high resolution, which means more details, and faster. The high-performance computing is key. I can’t imagine what we would do without the community clusters.”



**Wen Jiang**

Professor of  
Biological Sciences



**Gerhard Klimeck**

Professor of Electrical and  
Computer Engineering, director of the Network for  
Computational Nanotechnology and nanoHUB.org

“We do extremely intensive calculations requiring computation at a magnitude that has been impossible because of the lack of sufficiently powerful computers. The availability of such large-scale machines and the codes that can utilize them enables us to move nano science to nano engineering.”

## FACULTY PARTNERS BY CLUSTER

### STEELE

18 departments  
62 faculty

### COATES

25 departments  
81 faculty

### ROSSMANN

16 departments  
36 faculty

### HANSEN

11 departments  
22 faculty

### CARTER

26 departments  
60 faculty

### CONTE

26 departments  
62 faculty

### SNYDER

26 departments  
28 faculty

### RICE

38 departments  
73 faculty

### HALSTEAD

33 departments  
75 faculty

### BROWN

40 departments  
89 faculty

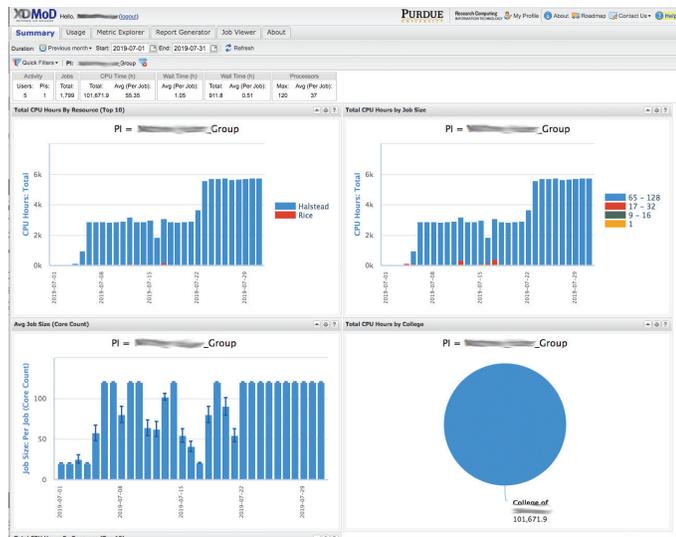
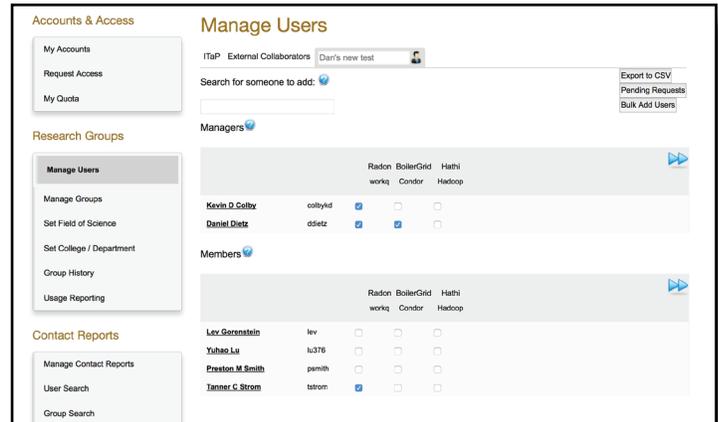
# RESEARCH COMPUTING AT PURDUE UNIVERSITY

## SELF-SERVICE TOOLS

### MANAGE YOUR GROUP

You or your delegate can enable or remove access for your students, staff, or collaborators on any cluster queue that you own.

Create and define UNIX groups for students and collaborators to work with group storage.



### TRACK YOUR USAGE

Track which students use the most computing, generate reports for sponsors, and monitor trends in your group's resource usage.

### ADD NEW RESOURCES

Easily purchase cluster nodes or Research Data Depot storage space for your research group.

#### Community Clusters

ITaP is currently offering purchase of access to nodes in several clusters serving distinct communities. Read the description of each to see which community may best suit your needs, or contact us at [rcac-help@purdue.edu](mailto:rcac-help@purdue.edu) if you have any questions.

Item	Quantity	Price	Sub-total
<b>Brown High-Performance Cluster (through Nov 2023)</b>	<input type="checkbox"/>		\$0.00
96GB memory, Two 12-core Intel Xeon Gold 6126 "Sky Lake" processors (24 cores total), EDR Infiniband interconnect. Access through this purchase will end November 2023. Original cost \$5599			
Brown is optimized for a wide variety of highly-parallel and HPC applications.			
<b>Research Data Depot - 1TB (1024 GB) storage</b>	<input type="checkbox"/>		\$0.00
per TB/year			

The Research Data Depot consists of over 2 petabytes of storage, available to any Purdue research group or campus unit in need of a high-capacity central solution for storing large, active research data sets at a competitive price. Research Data Depot storage is available for use on all research clusters, and lab computers.  
1 Terabyte (TB) of Research Data Depot Storage. This is billed on an annual basis.

## COMPUTATIONAL SCIENCE EXPERTISE

In addition to the peace of mind gained from professional systems engineering staff, community cluster partners can draw from the expertise of ITaP's experienced computational scientists, software engineers, and visualization experts.

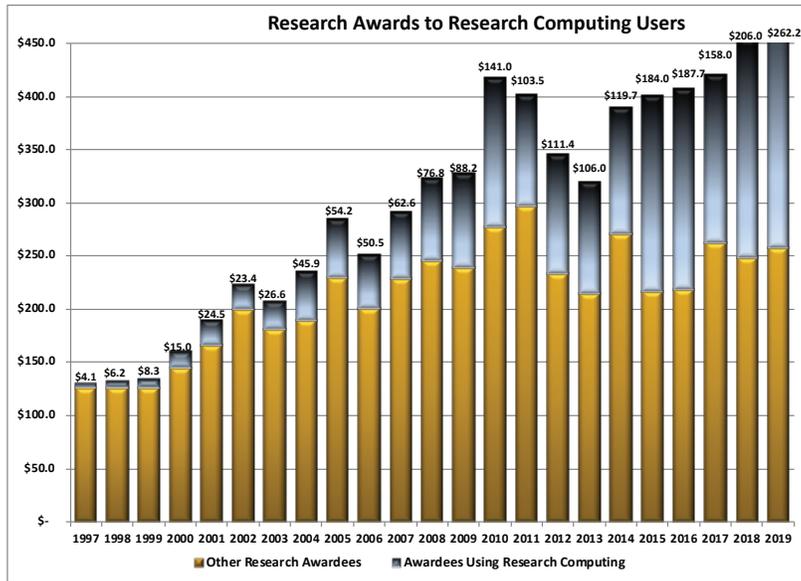
ITaP computational scientists are experienced users of computational resources, with advanced degrees in Engineering, Big Data, Bioinformatics, Biology, Chemistry, and Physics. Computational science staff can help with a wide range of issues: from answering user questions and providing training, code development, software installation, designing effective workflows, and performance analysis. Additionally, research solutions engineers are available to consult on applying new technology solutions for science problems.

# RESEARCH COMPUTING AT PURDUE UNIVERSITY

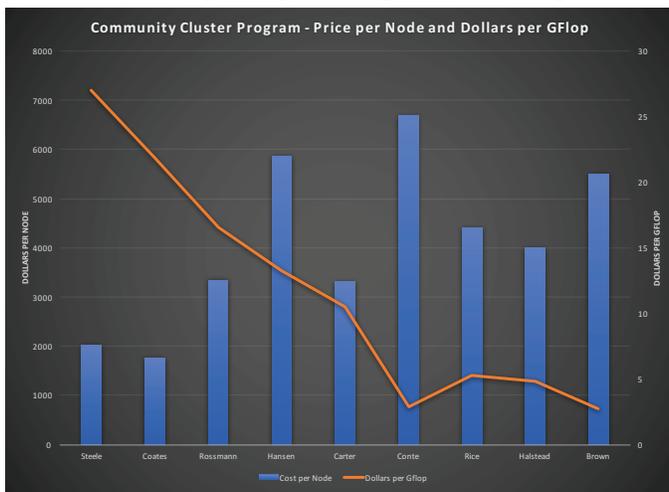
**PURDUE**  
UNIVERSITY®

Information Technology

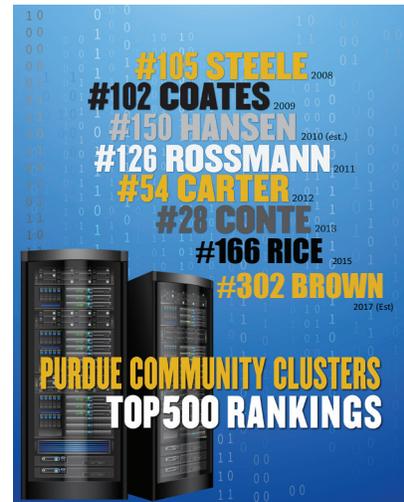
## POWERING RESEARCH AT PURDUE



## HIGH PERFORMANCE COMPUTING AT THE HIGHEST PROVEN VALUE



## WORLD-CLASS HPC



## DIVERSE COMMUNITY

College	Partners
Engineering	265
Science	227
Agriculture	111
Health and Human Sciences	28
Management	24
Polytechnic	21
Liberal Arts	9
Pharmacy	9
Education	7
Veterinary Medicine	5

Computing and Storage Partners

## A CLUSTER FOR ALL COMMUNITIES

### GILBRETH

- GPU Acceleration
- Machine Learning/AI
- Compute in Jupyter Notebooks
- Flash-based storage

### BROWN

- Traditional HPC
- High-speed interconnect
- Parallel computations
- High-performance, parallel scratch

### DATA WORKBENCH

- Interactive Computing
- Data Analytics
- Windows Virtual Desktops
- Non-Batch users

# RESEARCH COMPUTING AT PURDUE UNIVERSITY

**PURDUE**  
UNIVERSITY®

Information Technology

## RESEARCH DATA

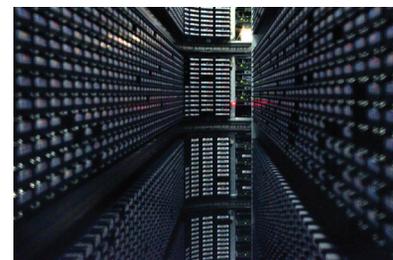


### RESEARCH DATA DEPOT

The Data Depot is a high-capacity, high-performance, reliable and secure data storage service designed, configured and operated for the day-to-day storage needs of Purdue researchers. The Data Depot is ideal for sharing data and collaborating with researchers on or off-campus. With spaces centered around a researcher's lab, the Data Depot provides storage at a competitive annual rate.

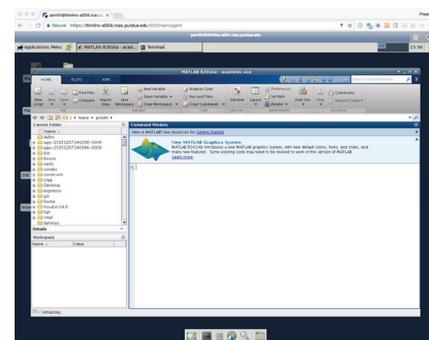
### FORTRESS ARCHIVE

The Fortress HPSS archive is a large, long-term, multi-tiered file caching and storage system utilizing both online disk and robotic tape drives. Fortress is ideal for permanent storage of raw data, results, or other critical research data. Access to Fortress is available to any researcher free of charge.



### DATA WORKBENCH

The Data Workbench provides access to advanced research storage systems, applications, and powerful hardware for interactive computing and data analytics. Access statistical packages, remote desktops, Python notebooks, and Windows virtual machines for computation, all from the convenience of your browser.



## PURR: INSTITUTIONAL REPOSITORY

The Purdue University Research Repository (PURR) is a research collaboration and data management solution for Purdue researchers and their collaborators. PURR allows you to collaborate on research and publish datasets online. Sharing the data that supports your research allows other scholars to reuse and cite your data as well as to reproduce your research.

To meet the requirements from funding agencies, such as the National Science Foundation and the National Institutes of Health, that a grant proposal include a Data Management Plan, PURR is available to any Purdue faculty, staff, or student.

# RESEARCH COMPUTING AT PURDUE UNIVERSITY

**PURDUE**  
UNIVERSITY®

Information Technology

## ADDITIONAL RESOURCES

Computing systems are available for more than just traditional high-performance computing communities. ITaP operates resources appropriate for many different types of computing requirements.

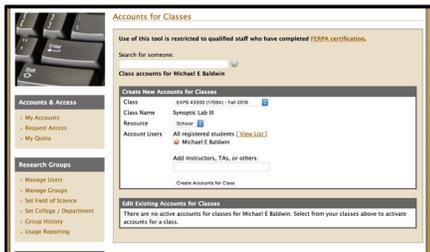
### GILBRETH: GPU ACCELERATION

The Gilbreth Community Cluster provides over 1 PetaFLOP of GPU capacity to solve the most challenging problems in AI, Machine Learning, data science, or other accelerated applications.



### SCHOLAR CLUSTER

The Scholar cluster is open to Purdue instructors from any field whose classes include assignments that could make use of supercomputing for modeling or data science, from high-end graphics rendering, weather modeling, simulation of millions of molecules, and exploring masses of data to understand the dynamics of social networks.



### SECURE COMPUTING

ITaP Research Computing provides resources for data and computation in support of projects with heightened security requirements. Research requiring protection for human subjects data, Export Control (EAR, ITAR), or Controlled Unclassified Information can all be performed on ITaP computing systems.



## COFFEE HOUR CONSULTATIONS

Coffee Hour Consultations are excellent opportunities in a casual setting to consult and discuss computing questions with ITaP computational scientists.

Researchers can attend to learn more about ITaP's high-performance computing, data storage and other research computing resources, ask general research computing questions, or get help with a specific research problem. Sessions are held three times a week at various coffee shops around campus.

[rcac.purdue.edu/coffee](http://rcac.purdue.edu/coffee)

# RESEARCH COMPUTING AT PURDUE UNIVERSITY

**PURDUE**  
UNIVERSITY®

Information Technology

## WANT TO LEARN MORE?

### CONTACT US:

#### **PURCHASE QUESTIONS:**

[rcac-cluster-purchase@lists.purdue.edu](mailto:rcac-cluster-purchase@lists.purdue.edu)

#### **SUPPORT QUESTIONS:**

[rcac-help@purdue.edu](mailto:rcac-help@purdue.edu)

#### **BUSINESS ADDRESS:**

ITaP - Young Hall  
155 S Grant Street  
West Lafayette, IN 47907-2114



Purdue University is an equal access, equal opportunity University.

## **FIND US ONLINE:**

**ITaP:** <http://www.itap.purdue.edu>

**ITaP Research Computing (RCAC):**  
<http://www.rcac.purdue.edu>

**Community Cluster Program:**

<https://www.rcac.purdue.edu/services/communityclusters/>

[facebook.com/PurdueRCAC](https://facebook.com/PurdueRCAC)

Twitter: [@PurdueRCAC](https://twitter.com/PurdueRCAC)

[youtube.com/PurdueRCAC](https://youtube.com/PurdueRCAC)

Photo Credits - Front: Purdue University photo/Andrew Hancock Page 2: Patrick Finnegan  
Back: Purdue University photo (top), Patrick Finnegan (bottom)

