

# Research Computing: A view from the trenches

George Turner, Chief Systems Architect  
Indiana University

Visualization Lab, 240 Braunstein Hall  
Geology-Math-Physics (GMP) Library

7-Nov-2019



DCS<sup>2</sup>

Data & Computational  
Science Series 2019



RESEARCH TECHNOLOGIES  
PERVASIVE TECHNOLOGY INSTITUTE



# University Information Technology Services

Popular searches: [Passphrase reset](#), [Email-eTexts](#), [Exchange](#), [IUware](#), [Gmail](#)



Box



Gmail at IU



One.IU



Exchange



Canvas

## IT Notices

▲ Difficulties accessing Scholarly Data Archive

[View all notices](#)

## IT Support



Bloomington

[How do we know?](#)

Knowledge Base: [kb.iu.edu](http://kb.iu.edu)

## Accounts & passphrases

Includes eligibility, quotas, and requests for access

## Email

Includes set-up instructions, forwarding, list management, policies, and access to other services

## Computers & printing

Includes student lab locations, hours, printing, and tools to find available stations

## Software & hardware

Includes recent software downloads, on-

## IU Sitehosting

## Help & support

Includes technology repair, contact information, hours, and locations

## Connection services

Includes Wi-Fi (IU Secure), VPN, videoconferencing, and set-up instructions

## Training

Technology training workshops, seminars,



# UITS Research Technologies

SERVICES   IMPACT   NEWS & EVENTS   ABOUT US   SUPPORT   HUF19   THE CAAV 19

## Innovation Award

Indiana University wins 2019 Citrix Innovation Award

2019 Citrix Innovation Award

### Impact Highlights



IU hosts two high-tech conferences—in the same week

IU SCA supports quality control for medical imaging

IU leads NSF-sponsored workshop for new application benchmarks



We're advancing the future of research, scholarship, and creative activity at IU. Not sure where to start? You're already halfway there.

[Contact us](#)



#### Expert Consulting

Connecting you with software, computing, tools, and other resources to advance your work



#### Compute & Storage Resources

Enabling fast calculations, advanced simulations, and massive secure storage



#### Research Software

Delivering apps for analytics and big data research; distributing stat/numerical and open-source software



#### Visualization & Data Services

Promoting interactive models, virtual and augmented reality, advanced digital arts and media as well as secure data analysis



#### Training & Outreach

Connecting you with the research services you need to reach discovery through targeted training, seminars, workshops, and tours



# Pervasive Technology Institute

[CONTACT](#) / [NEWSLETTERS](#) / [INTERNAL](#)

[ABOUT](#)

[CENTERS](#)

[SERVICES](#)

[IMPACT](#)

[MAJOR PROJECTS](#)

[OUTREACH & TRAINING](#)

[NEWS & EVENTS](#)



## SEVEN CENTERS. ONE MISSION.

### Latest news about PTI



IU hosts two high-tech conferences—in the same week

Indiana University to help secure Indiana's 2020 elections

IU SCA supports quality control for medical imaging





# Pervasive Technology Institute

[CONTACT](#) / [NEWSLETTERS](#) / [INTERNAL](#)

[ABOUT](#)

[CENTERS](#)

[SERVICES](#)

[IMPACT](#)

[MAJOR PROJECTS](#)

[OUTREACH & TRAINING](#)

[NEWS & EVENTS](#)



- Center for Applied Cybersecurity Research (CACR)
- Data to Insight Center (D2I)
- Digital Science Center (DSC)
- HathiTrust Research Center (HTRC)
- National Center for Genome Analysis Support (NCGAS)
- Science Gateways Research Center (SGRC)
- UITS Research Technologies (RT)
- Retired centers

## Latest news about PTI



IU hosts two high-tech conferences—in the same week

Indiana University to help secure Indiana's 2020 elections

IU SCA supports quality control for medical imaging



## OUR EXPERTISE

The GlobalNOC is unique among network operation centers. We are a nonprofit focused exclusively on the R&E community and the research and development of advanced network services.

INTERNET

- IP Network Layer 1 Services
- Address Layer 2 Service (Layer 2 Network)
- Physical Access Layer 1 Services

By R&E, For R&E

About GlobalNOC



# XSEDE: Resources for Science and Engineering

**XSEDE**

Extreme Science and Engineering  
Discovery Environment

# The origin of research computing at IU

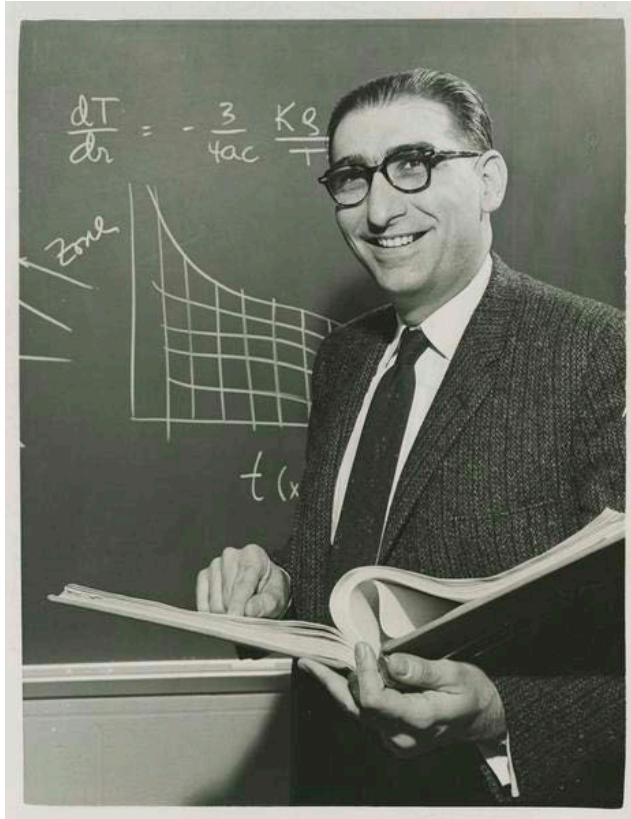
- 1949: Office of Naval Research grant for an IBM 602A Calculating Punch
  - Korean War, selective service required GPA, this machine could do division!
  - 8:00am-5:00pm used by the Registrar's Office
  - 5:00pm-8:00am used by Profs Merritt (Chem), Gucker (Chem), Wrubel (Astro)
  - President Wells would spend evenings with the researchers in the computing center
- 1953: President Herman B. Wells calls for the creation of a research computing center in his state of the university address
- 1954: committee formed to with the goal of proposing a Scientific Computing Center



# The origin of research computing at IU (cont.)

- Two recommendations of the committee
  - Research computing equipment should be available on an “open shop” basis; i.e. researchers should be able to use the equipment themselves with support from the staff
  - Computing services should be available to all researchers, including those who could not personally fund their use.

# The origin of research computing at IU (cont.)



According to the legend handed down by Astronomy & Astrophysics students, Dr. Wrubel was reported to have stated:

Everybody engaged in the university's business at Indiana University has a valid need for access to research computing systems. *Free, open, and equal* access will be afforded to all faculty, staff, and students ...including my undergraduate astronomy students.

Marshall Wrubel,  
first official director of Indiana University's  
Research Computing Center

“*For years, [IU] was one of the few universities where computing was free. **That required a lot of support from campus administrators!***”

Dr. E Wainwright Martin  
School of Business



IU 's first real computer, the IBM 650, was delivered in the fall of 1956. It rented for \$31,200 a year, included no software or operating system, had only 2,000 words of memory, occupied 1.000 square feet of floor space, and contained approximately 1.000 vacuum tubes, which generated a great deal of heat. In 1958 the University remodeled and air conditioned a special room for it in the basement of Swain Hall East, at a cost of \$72,000.



IU's first real computer, the IBM 650, was delivered in the fall of 1956. It rented for \$31,200 a year, included no software or operating system, had only 2 occupied ,000 words of memory, 1,000 square feet of floor space, and contained approximately 1,000 vacuum tubes, which generated a great deal of heat. In 1958 the University remodeled and air conditioned a special room for it in the basement of Swain Hall East, at a cost of \$72,000.



The Research Computing Center was in the basement of what is today Swain Hall West. This image appears on page 10 of the October 1958 edition of the Indiana Alumni Magazine. The caption reads, in part: "New magnetic tape reading machines (center) are speeding up scientific research computations...The tape machines and the Center's 'I.B.M. Tape 650 magnetic drum data-processing machine' (right) for the first combination of its kind in a university. The result is up to 60 per cent increases in speed of the Center's equipment, which had already cut some computations from months to minutes. The magnetic drum machine...'remembers' instructions by 'storing' them on a rapidly revolving magnetic drum, on which up to 20,000 digits can be stored. The new tape devices will increase the storage capacity to two and one-half million digits.

# IU-Bloomington Data Center

<https://dcops.iu.edu>



**DCS<sup>2</sup>**

Data & Computational  
Science Series **2019**



**RESEARCH TECHNOLOGIES**  
PERVASIVE TECHNOLOGY INSTITUTE

# Distributed Processing Compute Resources

<https://kb.iu.edu/d/alde>

- **Big Red II** – Cray XE6/XK7 hybrid with 344 CPU and 676 GPU nodes (NVIDIA k20) for 1 petaFLOPS of capacity
- **Big Red II+** – Cray XC30 with 552 CPU nodes and 286 teraFLOPS of capacity. Restricted to the IU Grand Challenge projects and dedicated large-scale distributed computing researchers.

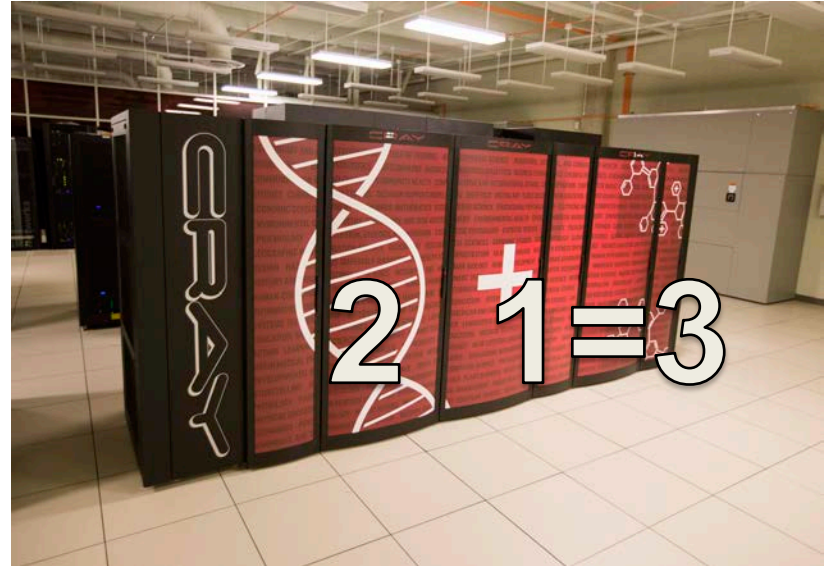




# Future [near term] Compute Resources

<https://kb.iu.edu/d/aoku>

- **Big Red III (3)**
  - An upgrade of Big Red II+ to a Cray XC40 (Intel Haswell processors)
  - 3X the capacity at nearly 1 petaFLOPS from its 936 dual-processor nodes.
  - Open for all users of Big Red II (faculty, staff, and graduate students, undergraduates with sponsorship)
  - Availability planned for October 14



# Big Red 200

- IU's next supercomputer
- ~6 petaFLOPS
- Installation in early 2020
- Cray Shasta supercomputer
  - AMD "Rome" CPUs
  - NVIDIA V100 GPUs
  - Cray Slingshot fabric

## How big is Big Red 200?

The new supercomputer is the latest major milestone in IU's decades-long leadership in pushing the boundaries of computing to advance world-class research.

Big Red 200 can process **53 times more data** in memory than Big Red and **10 times more** than Big Red II. For example, it can process **71,000** 3 GB brain scans in RAM simultaneously.

Big Red: 1,300 brain scans processed simultaneously



Big Red II: 7,100



Big Red 200: 71,000



 OFFICE OF THE  
VICE PRESIDENT FOR  
INFORMATION TECHNOLOGY  
and Chief Information Officer

Big Red 200 is almost **300 times faster** than Big Red and **6 times faster** than Big Red II..

Big Red: 20.48 teraFLOPS



Big Red II: 1 petaFLOPS



Big Red 200: 5.9 petaFLOPS



... which is almost **6 million times faster** than an iPhone XS.

# Current Compute Resources

<https://kb.iu.edu/d/alde>

- **Karst** – Lenovo NeXt Scale system with ~300 CPU nodes and 100 teraFLOPS of capacity. Used by many condo and colocation customers.
- **Carbonate** – Lenovo NeXt Scale system with 80 larger memory CPU nodes (256-512GB). Newer condo and colocation nodes installed here.
- **Carbonate Deep Learning (DL)** – Lenovo SD530 system with 12 GPU nodes (NVIDIA P100 & V100).



# Research Desktop (RED): a friendly gateway to HPC

<https://kb.iu.edu/d/apum>

- Making supercomputers more user friendly
  - RED provides a new way to login and interact with HPC
  - A GUI/desktop instead of a terminal
- **New:** Web browser access and a dedicated service for apps like Jupyter and Rstudio
- Research Desktop in the browser:
  - <https://red.uits.iu.edu>
- IUWare (Thinlinc) client download:
  - <https://iuware.iu.edu/search?q=thinlinc>



# AI and Deep Learning (DL) applications

<https://kb.iu.edu/d/avjk>

- Currently 40+ projects on Carbonate Deep Learning test bed
- Support a full DL stack on Python 2 & 3 including: Tensorflow, Pytorch, Theano, Keras, Mxnet, Caffe2, CNTK, Matlab
- Consultation on performance and throughput optimization
- URL for requests: <https://rtstats.uits.indiana.edu/DLprojects>
  - Future: <https://hpcprojects.uits.iu.edu/>

# Current Storage Resources

<https://kb.iu.edu/d/avkm>

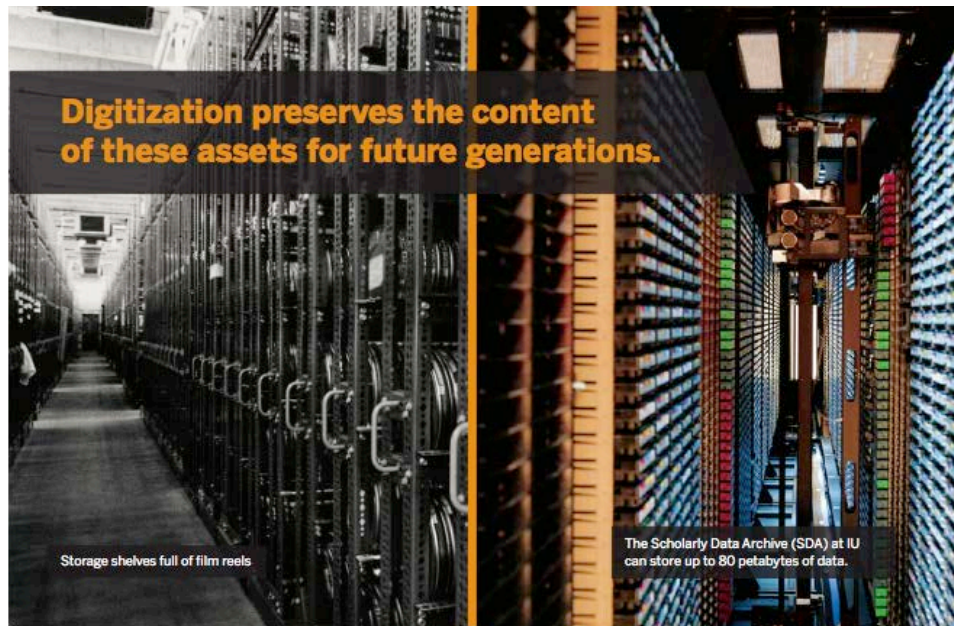
- **Research Home Directories (Geode)**
  - Allocated with accounts on central supercomputing systems.
  - Replicated between campuses with snapshot capability, 100 gigabyte default quota.
- **Scholarly Data Archive**
  - Long-term replicated archive with 79 petabytes of capacity.
  - 50 terabytes available to faculty, staff, and graduate students.



# Future Storage Resources

<https://kb.iu.edu/d/avkm>

- **Scholarly Data Archive**
  - Limits of 25K files and 50TB per user to be enforced.
  - Additional capacity available for \$0.025 per GB per year.
  - Tape drive upgrades in plan (20TB capacity) for next fiscal year.
  - Storage maintained for active accounts.



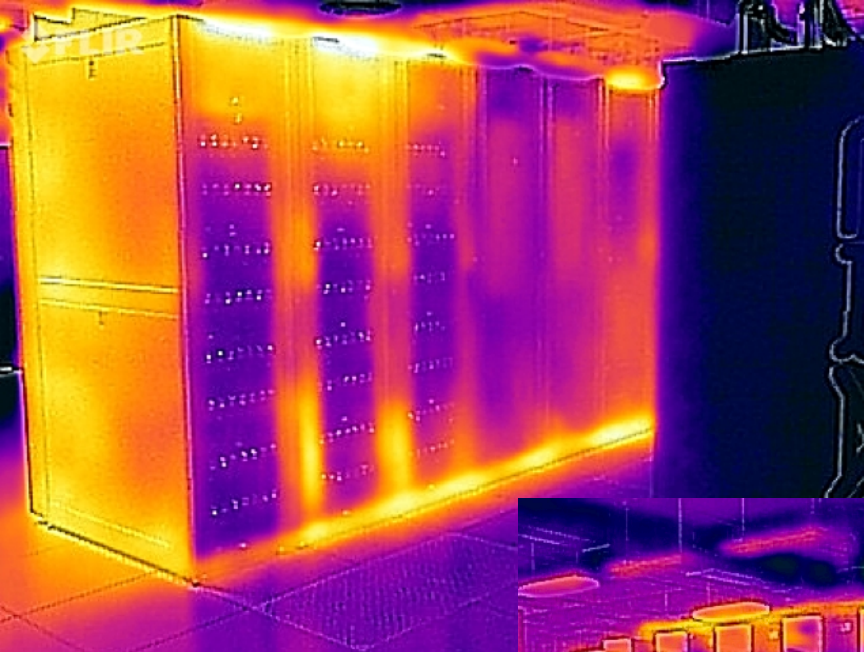
# Jetstream, a national resource

<https://jetstream-cloud.org>

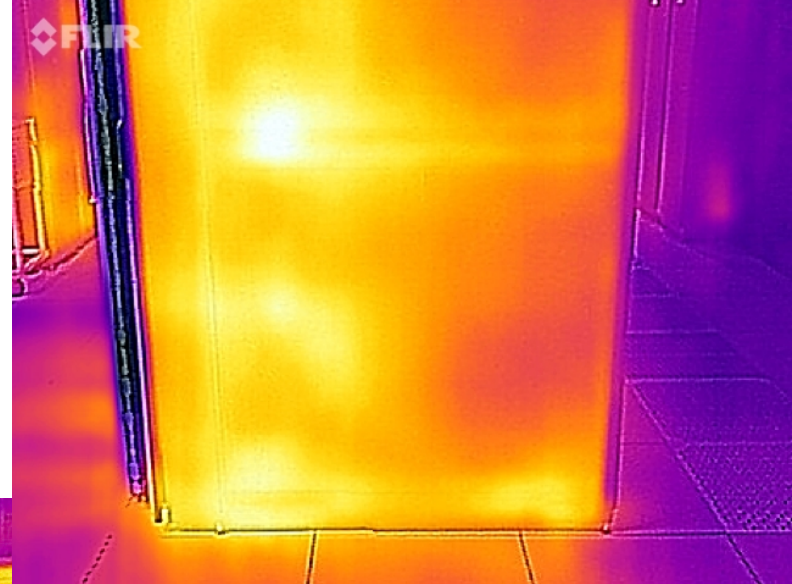
- Jetstream – Dell M1000 Blade System, 640 nodes distributed between IU and TACC
- NSF's first production cloud environment, includes a modest IU expansion.
- On-demand, interactive computing environment using virtual machines
- Supports 25 science gateways with over 24,000 users.
- Allocated through XSEDE; discretionary time also available including for piloting GPU services on cloud infrastructure.



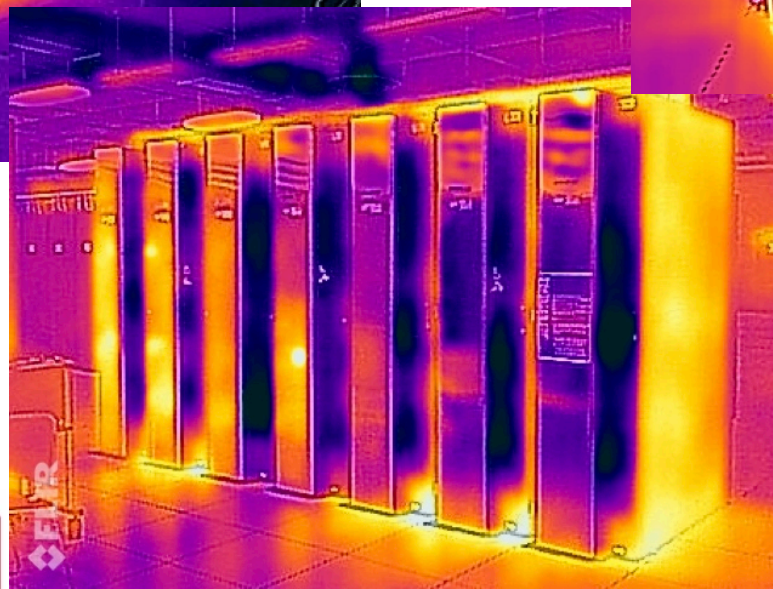




Front



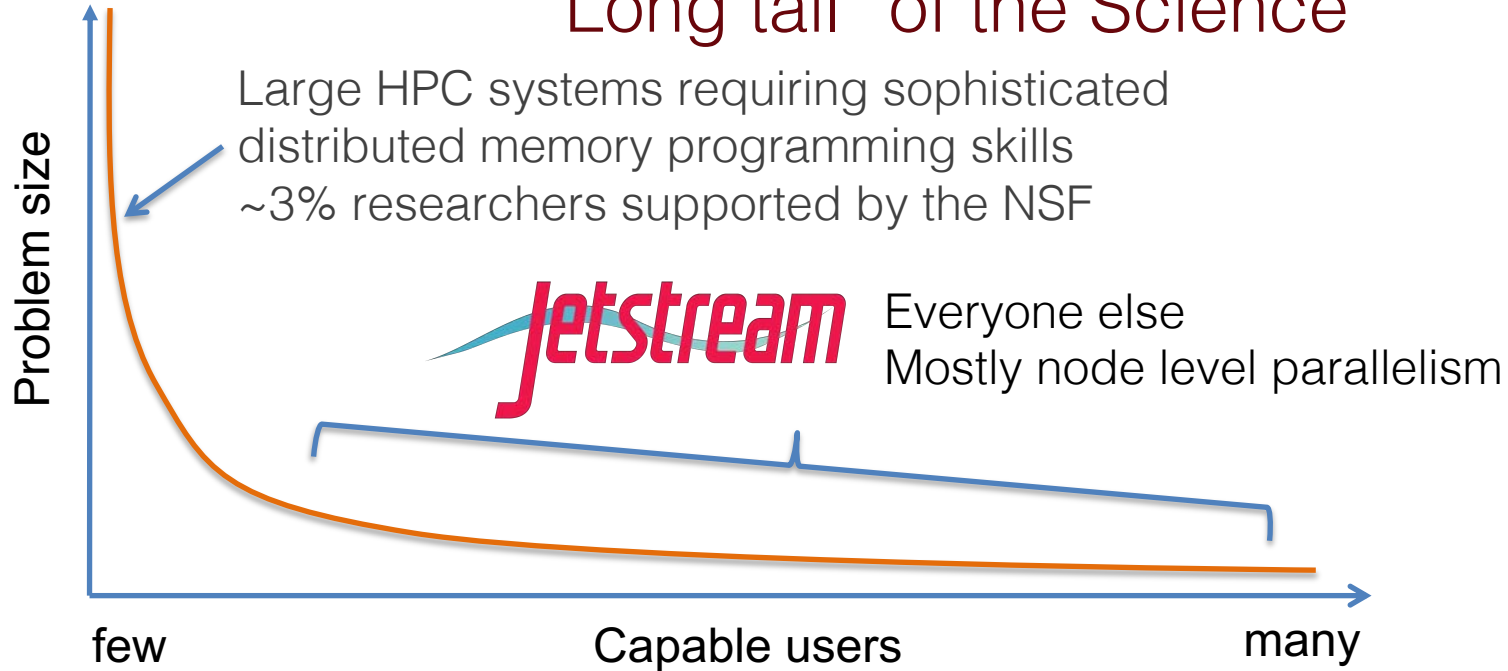
Side



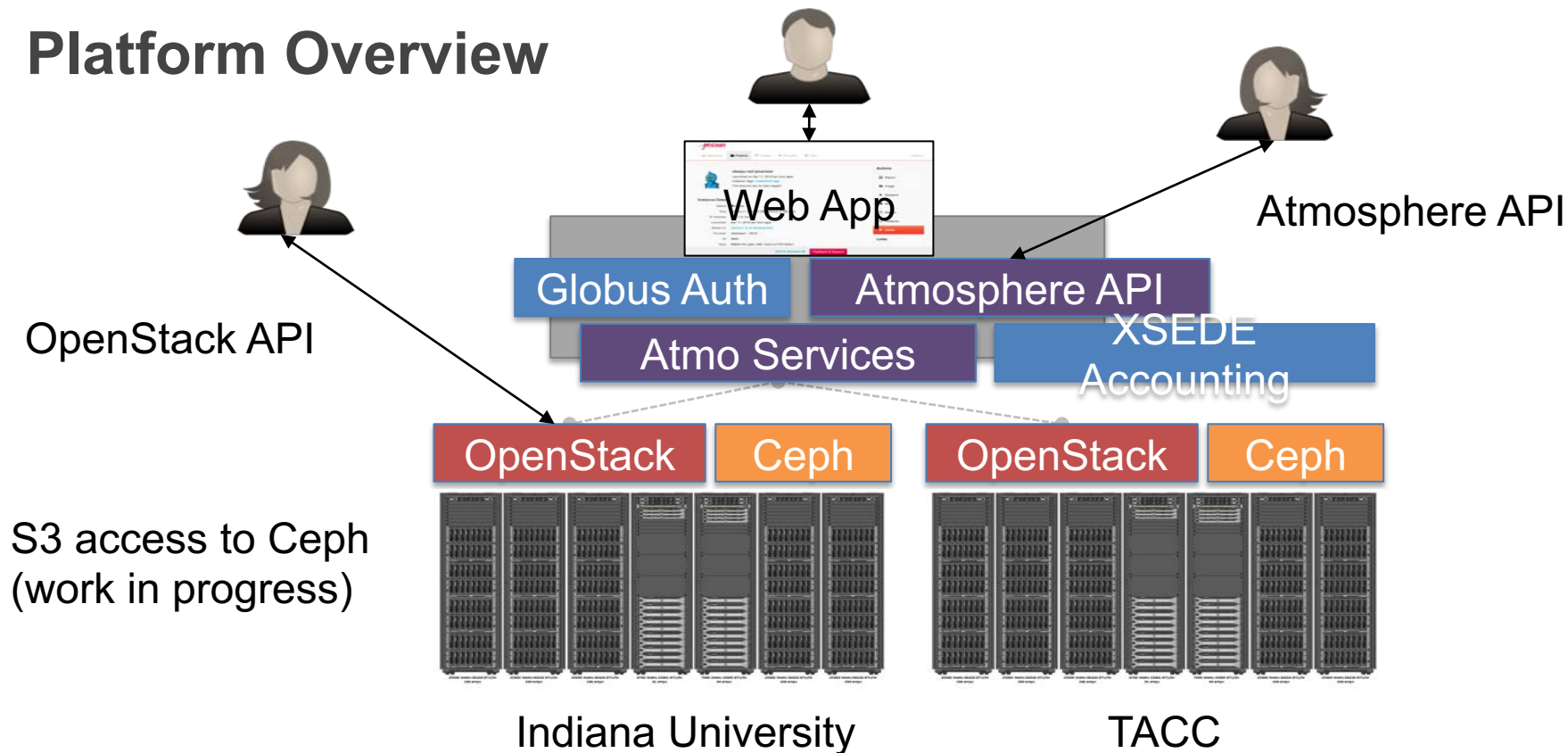
Back

# What is Jetstream? (cont)

## “Long tail” of the Science

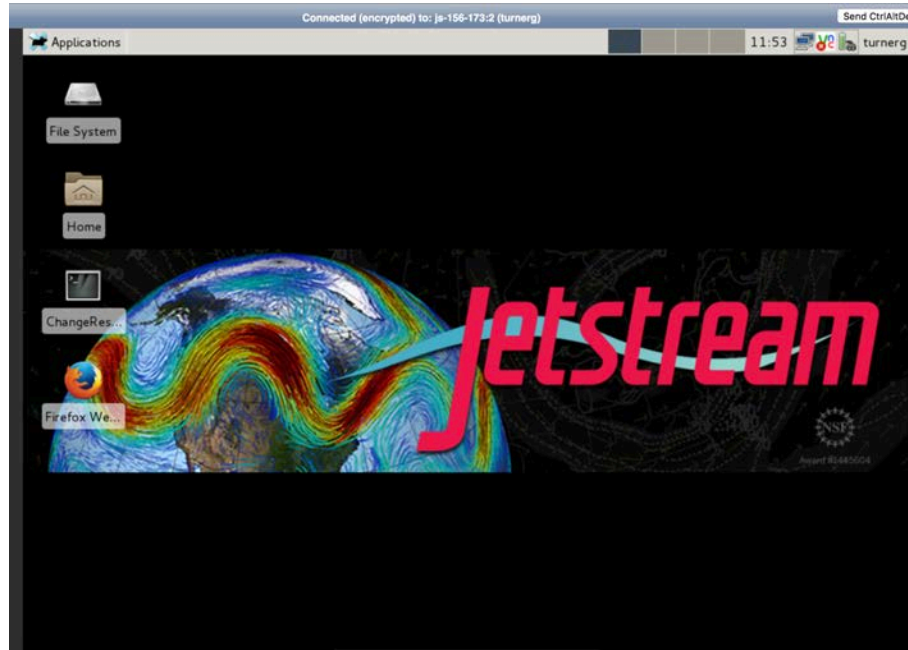


# Platform Overview



# Jetstream's Atmosphere Interface

Web desktop access to instance  
VNC access also available



# Research Technologies

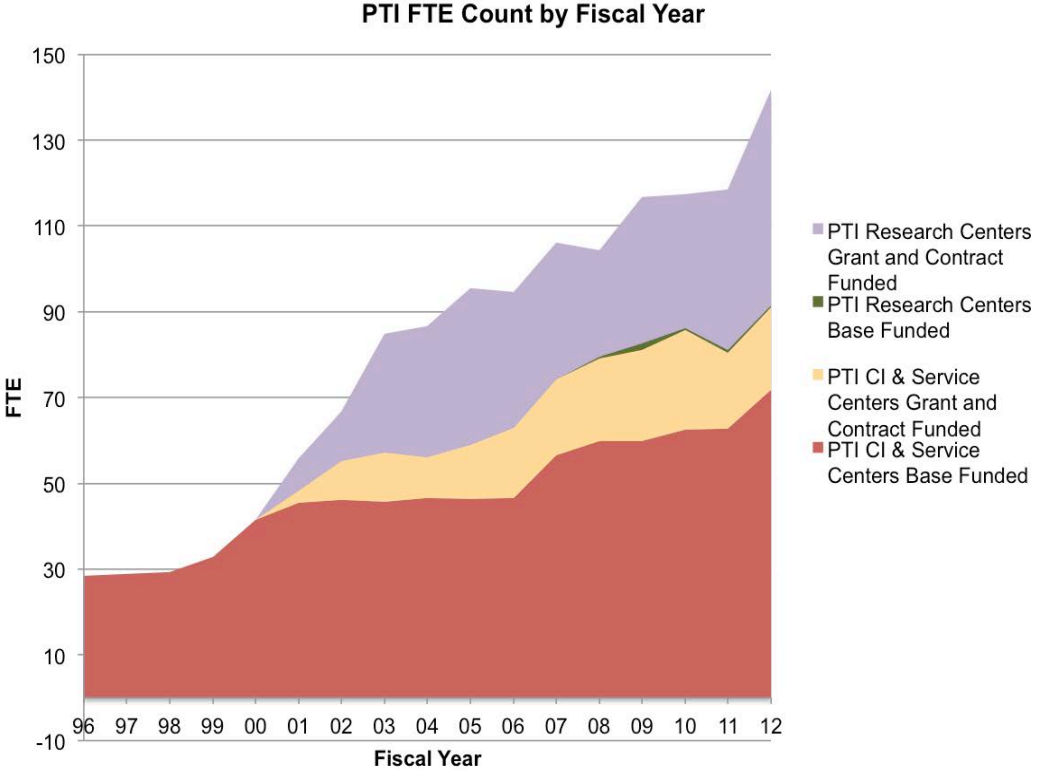
- Mission is to support research.
- Deliver a wide range of services.
  - Advanced Visualization
  - High Performance Computing
  - High Performance Storage
  - Research Software & Solutions
  - Education and Outreach
  - HIPAA & FISMA/CMS
- Contact <https://rt.iu.edu>
- Our impact <https://rt.iu.edu/impact/>
- Partner with faculty on grants
  - 52 grants totaling \$51M over the past 10 years
- Partner to solve problems



# RT Areas of Engagement

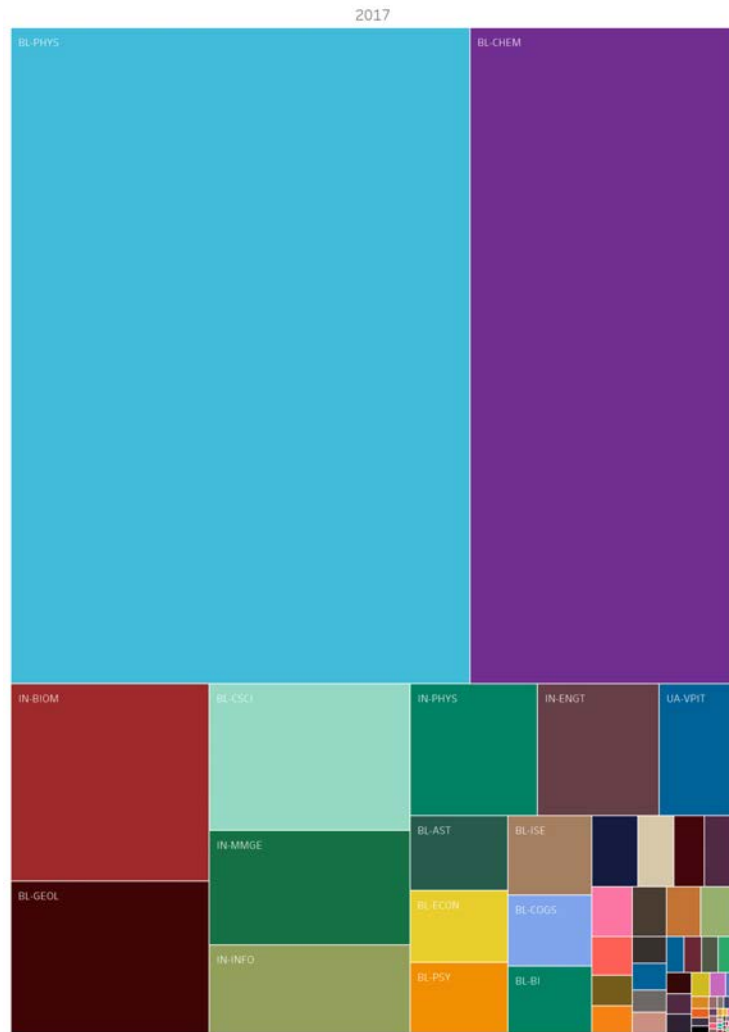
1. AVP - Matt Link
  - Senior Technical Advisor – Tony Walker
  - Chief Systems Architect - George Turner
2. Advanced Cyberinfrastructure – Dave Hancock, Dir.
3. Visualization & Data Services – Eric Wernert, Dir.
4. Community Engagement – Therese Miller, Dir
5. Research Software & Solutions – Robert Henschel, Dir.

# Continued job growth... PTI as an example



# RT Supports Everyone at IU

- 153 different departments and over 330 disciplines use IU's advanced research cyberinfrastructure provided and supported by RT
- Breadth of impact on IU research, scholarship, and creative communities.
- Over 1,000 IU researchers, who use IU supercomputers or other HPC tools, received nearly \$313 million in grant awards in 2018.





# RT Fun Facts

- Enabled \$308M in grants and contracts in FY19 (~45% of all funded research).
- Worked with 153 different IU departments
- Supported 331 academic disciplines
- Big Red 200 system will be the first of Cray's revolutionary new Shasta supercomputers installed at a U.S. university.
- Research Desktop provides a graphical interface to supercomputers for over 1,000 researchers.
- IU's use of Lustre started with an NSF-funded major research instrumentation grant in 2006 and sustained with IU funding since.

# Where research computing at UC is heading....

[research.uc.edu/arc](http://research.uc.edu/arc)

**UC CINCINNATI** OFFICE OF RESEARCH

News Facts & Figures For Researchers Signature Programs & Events Leadership & Offices

## INITIATIVE

The advanced research computing (ARC) initiative establishes a high-performance computing (HPC) infrastructure that supports and accelerates computational and data-enabled research, scholarship and education at the University of Cincinnati.

### ABOUT

The University of Cincinnati's Advanced Research Computing (ARC) initiative offers a readily accessible hybrid CPU/GPU computing cluster, supporting the next generation of computational and data science researchers while developing a highly competitive workforce.

We will partner with you to utilize the core suite of HPC services and resources. With the ARC resources, researchers can advance theoretical knowledge and expand the realm of discovery, generating leading edge research and applications suitable for innovation and commercialization in line with UC's Next Lives Here strategic direction.

This sustainable high-performance computing (HPC) infrastructure with technical support services, accelerates the time to discovery and enables sophisticated and increasingly realistic modeling, simulation and data analysis and will help to bridge users to the local, regional and national HPC ecosystem.

### IMPACT

ARC resources support all disciplines, including healthcare, sciences, engineering and social sciences/humanities, in their quest to harness big data analytics, modeling and simulation, visualization, artificial intelligence and machine learning.

**DCS<sup>2</sup>**

Data & Computational  
Science Series **2019**



**RESEARCH TECHNOLOGIES**  
PERSVASIVE TECHNOLOGY INSTITUTE

Where research computing at UC is heading....



# Where research computing at UC is heading....



**Someone's always ready to light this place up.**

Jane Combs

# Where research computing at UC is heading....



**Someone's always ready to  
light this place up.**

Jane Combs                      Amy Latessak  
Larry Schartman                Adam Steele

# Where research computing at UC is heading....



**Someone's always ready to  
light this place up.**

Jane Combs                      Amy Latessak  
Larry Schartman                Adam Steele

Faculty Advisory Committee  
Prashant Khare,                Chair  
Thomas Beck                    Emily Kang  
Mario Medvedivic              Chuck Sox  
Philip Taylor                    Ad Hoc

Project website: <http://rt.iu.edu/>

Project email: [researchtechnologies@iu.edu](mailto:researchtechnologies@iu.edu)

Direct email: [turnerg@iu.edu](mailto:turnerg@iu.edu)

### License Terms

- Turner, G.. 2019. Research Computing; A view from the Tranches, Data & Computational Science Series 2019; Cincinnati, OH. Also available at: <http://jetstream-cloud.org/publications.php>
- Jetstream is supported by NSF award 1445604 (David Hancock, IU, PI)
- XSEDE is supported by NSF award 1053575 (John Towns, UIUC, PI)
- This research was supported in part by the Indiana University Pervasive Technology Institute, which was established with the assistance of a major award from the Lilly Endowment, Inc. Opinions presented here are those of the author(s) and do not necessarily represent the views of the NSF, IUPTI, IU, or the Lilly Endowment, Inc.
- Items indicated with a © are under copyright and used here with permission. Such items may not be reused without permission from the holder of copyright except where license terms noted on a slide permit reuse.
- Except where otherwise noted, contents of this presentation are copyright 2015 by the Trustees of Indiana University.
- This document is released under the Creative Commons Attribution 3.0 Unported license (<http://creativecommons.org/licenses/by/3.0/>). This license includes the following terms: You are free to share – to copy, distribute and transmit the work and to remix – to adapt the work under the following conditions: attribution – you must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work). For any reuse or distribution, you must make clear to others the license terms of this work.