

UITS Research Technologies Cyberinfrastructure for IU Research and Academics

Robert Ping

UITS Research Technologies, Cyberinfrastructure and Service Center
Indiana University Pervasive Technology Institute



Science Gateways



Storage



Computation



Analysis/
Software



Bio/Health



Visualization



Campus
Bridging



Education/
Outreach



Grant
Support



Cyber risk
Mitigation

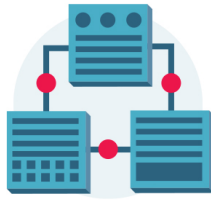
IT-28

Overview

- UITS Research Technologies (RT)
 - Changing the way IU uses Cyberinfrastructure
- RT Service Areas
 - From Analysis to Visualization and everything in between



Analysis and software
delivery and support



Visualization facilities
and services

(Inter)National Cyberinfrastructure

A SPECIAL REPORT

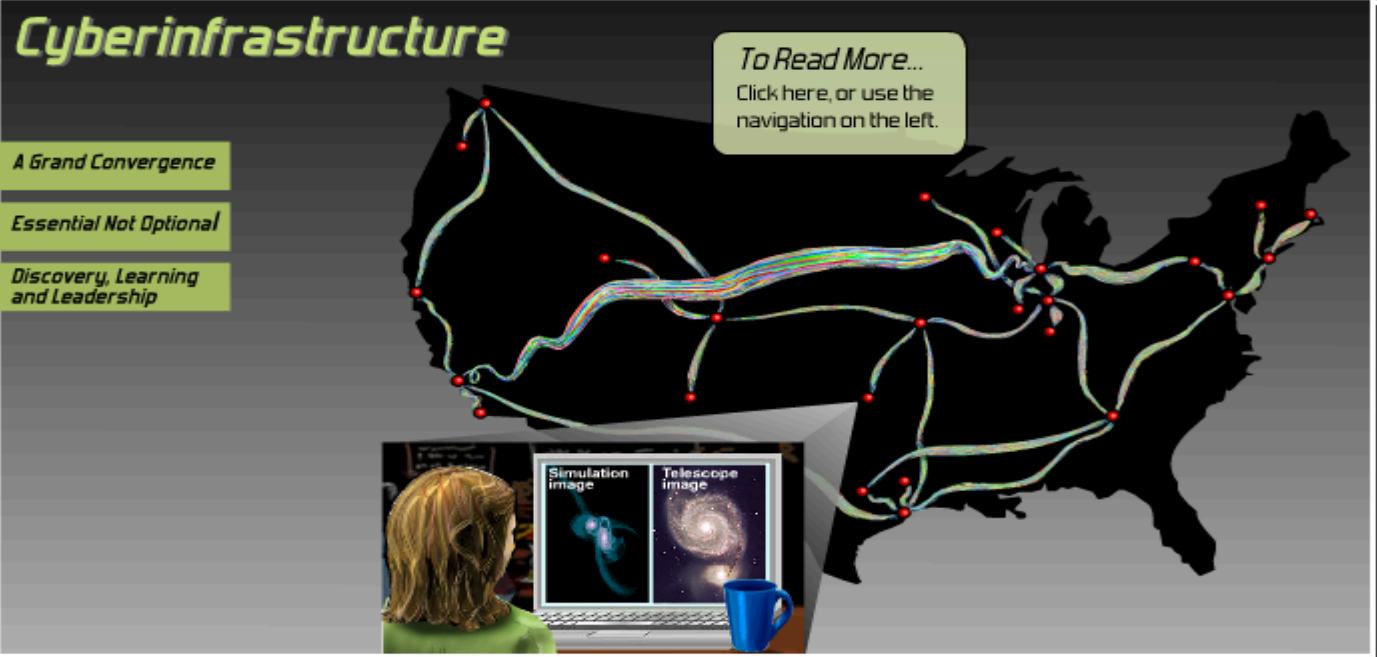
Cyberinfrastructure

To Read More...
Click here, or use the navigation on the left.

A Grand Convergence

Essential Not Optional

Discovery, Learning and Leadership



The graphic features a dark map of the United States with a complex network of multi-colored lines (red, green, blue, yellow) representing data or network paths. A callout box in the upper right of the map area contains the text: "To Read More... Click here, or use the navigation on the left." On the left side of the map, there are three stacked green boxes with white text: "A Grand Convergence", "Essential Not Optional", and "Discovery, Learning and Leadership". Below the map is an inset image showing a person with long blonde hair looking at a laptop. The laptop screen is split into two panels: "Simulation image" showing a blue and white molecular or biological structure, and "Telescope image" showing a spiral galaxy. A blue mug is on the desk next to the laptop.

*National Science Foundation, "A Special Report". http://www.nsf.gov/news/special_reports/cyber/.

Accessed March 31, 2014.



Margaret Dolinsky, Associate Professor and Research Scientist
H.R. Hope School of Fine Arts, Indiana University Bloomington

RT Services – Science Gateways

Web interfaces to supercomputers, massive storage, and shared workflows

- RT works to simplify the usage experience of computational resources for users through nationally funded grid projects. The Science Gateways Group does this in particular by implementing open source software to support science gateway projects.*



HOME TOOLS DATA NEWS ABOUT

Latest Analysis

Forecast Station Monitor Scorecard

Map Satellite

Google

Status	Color
no status change:	
status changes on selected date:	
status changed in last 30 days before selected date:	
no data on selected date:	

2014-3-15 Go to latest

March 2014

Su	Mo	Tu	We	Th	Fr	Sa
23	24	25	26	27	28	1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22

RT Services – Science Gateways

- IU Cyberinfrastructure Gateway – the “one stop shopping” portal for all IU central cyberinfrastructure as well as external resources such as Penguin on Demand, Amazon Web Services, and Microsoft Azure.
- URL available at <http://researchtech.iu.edu> and in RT printed overview brochure



INDIANA UNIVERSITY Cyberinfrastructure Gateway Home News Software Contact Links Log in

Portal to IU computing resources

- Get updates on Big Red II, Quarry, and Mason status.
- Move files from your desktop to IU's Scholarly Data Archive and Data Capacitor.
- Monitor and manage running jobs on Big Red II, Quarry, and Mason.
- Find information on available software.

[Login](#)

Chat With a Consultant
[ITHelpLive](#) (IU login required)
 Available 24 hours a day, 7 days a week

ITS Assistance
 Ask RT For Assistance
[ITS Support Center](#)

IT Notices
 Access Management site maintenance scheduled
 Other Announcement
 Last updated: Thu 3/27/2014 3:59 PM EDT
[...Read More](#)
 IU Address Book maintenance scheduled
 IU Address Book (people.iu.edu) Announcement
 Last updated: Thu 3/20/2014 9:34 AM EDT
[...Read More](#)

Latest Updates
 10th annual Polar Technology Conference coming to Bloomington
 IUTS Research Technologies enables study of complete RNA collection of fruit fly uncovers unprecedented complexity

System Status

Resource	Status	Utilization	Job
Big Red II	Healthy	Total Nodes: 1020 Running Nodes: 3 Busy Nodes: 314 Idle Nodes: 658 Drained Nodes: 3 Other Nodes: 42	Total: 784 Running: 31 Idle: 753 Not Queued: 0 Completed: 0 Other: 0
Mason	Healthy	Total Nodes: 16 Running Nodes: 4 Busy Nodes: 0 Idle Nodes: 10 Drained Nodes: 0 Other Nodes: 2	Total: 111 Running: 12 Idle: 99 Not Queued: 0 Completed: 0 Other: 0
Quarry	Healthy	Total Nodes: 273 Running Nodes: 26 Busy Nodes: 20 Idle Nodes: 224 Drained Nodes: 3 Other Nodes: 0	Total: 1051 Running: 97 Idle: 952 Not Queued: 0 Completed: 2 Other: 0

Jan 2014 System Summary

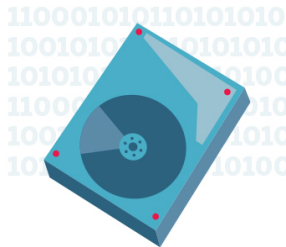
Cluster	Jobs	Users	Runtimes (hrs)	Resources (hrs)	Median Queue (hrs)	Median Runtime (hrs)
Big Red II	36359	149	149087	7932979	0.23	0.42
Mason	3516	94	29548	281768	0.53	0.73
Quarry	127083	88	493710	682613	0.30	0.23

RT Services – Data Storage

High-quality storage services for fast access or large data sets

- **Data Capacitor II**

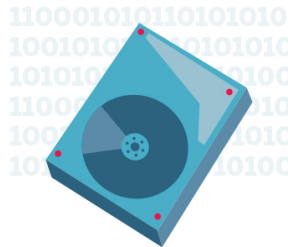
- Data storage on disk, not backed up
- Temporary storage of research data – purged regularly
- Large-capacity, high-throughput, high-bandwidth
 - Lustre-based file system
 - Serves all IU campuses
 - Mounted on Big Red II, Quarry, and Mason research computing systems



RT Services – Data Storage

High-quality storage services for fast access or large data sets

- **Data Capacitor Wide-Area Network (WAN) II**
 - Data storage on disk, not backed up
 - Temporary storage of research data – purged regularly
 - Allows researchers to access remote data as if stored locally
 - Serves all IU campuses and several research centers throughout the US
 - Makes it easy to share large amount of data among multiple research sites

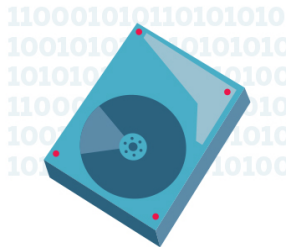


RT Services – Data Storage

High-quality storage services for fast access or large data sets

- **Research File System (RFS)**

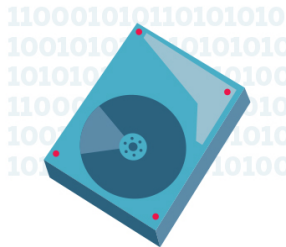
- Centralized disk storage designed to support researchers
- Backed up regularly – 2 copies (IUB and IUPUI)
- Open to IU community – undergrads/non-IU must have sponsor
 - 100 GB of disk storage
 - Can utilize project areas in file system
 - HIPAA-aligned (does not encrypt stored data – individual users responsible)



RT Services – Data Storage

High-quality storage services for fast access or large data sets

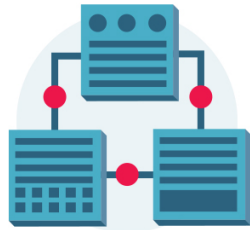
- **Scholarly Data Archive (SDA)**
 - Distributed tape storage for large-scale archival/near-line storage
 - Mirrored– 2 copies (IUB and IUPUI)
 - Open to IU community – undergrads/non-IU must have sponsor
 - 42 PB of tape storage
 - Supports SFTP, HSI, HPSS API
 - HIPAA-aligned (does not encrypt stored data – individual users responsible)



RT Services – Computation

Fast calculations, advanced simulations, and big data analysis

- **Big Red II** – IU's main system for high-performance parallel computing
 - Owned solely by IU
 - 1 petaFLOPS
 - CPUs/GPUs
 - 3.5 PB of scratch



- Available through itaccounts.iu.edu
- Support/consulting available
- Tours available

RT Services – Computation

Fast calculations, advanced simulations, and big data analysis

- **Research Database Complex** – research databases and data applications

- Supports oracle and MySQL Databases
- Strictly devoted to research
- Runs Red Hat Enterprise Linux

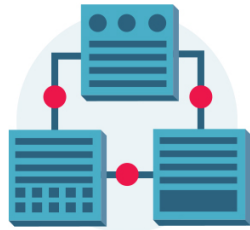


- Examples of research projects currently supported
 - Collaborative Initiative on Fetal Alcohol Spectrum Disorders
 - Indiana Spatial Data Service
 - National Gene Vector Biorepository

RT Services – Computation

Fast calculations, advanced simulations, and big data analysis

- **Rockhopper** – Penguin Computing’s Penguin-On-Demand (POD) supercomputing cloud appliance
 - Hosted by Indiana University
 - Burst capacity
 - Collaborative effort between Penguin Computing and several US institutions

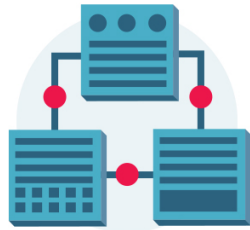


- Users log into the cluster using SSH

RT Services – Computation

Fast calculations, advanced simulations, and big data analysis

- **Quarry** – IU's primary distributed Linux cluster computing environment
 - For research and research instruction use
 - Serves as condominium cluster
 - Runs Red Hat Enterprise Linux
 - Accounts available via the Account Management Service (AMS)



RT Services – Computation

Fast calculations, advanced simulations, and big data analysis

- **Mason** – large memory computer cluster
 - Supports data-intensive, high-performance computing tasks
 - Specifically for genome researchers
 - Uses include genome assembly software, large-scale phylogenetic software, and other genome applications



RT Services Spotlight – Computation 2/2014

Cluster	Jobs	Users	Core Hours	Median Queue (hours)	Median Runtime (hours)
Big Red II	36359	145	7,932,979	0.2	0.4
Mason	3516	27	281,768	0.5	0.7
Quarry	127083	22	682,613	0.3	0.2
Total	166958		8,897,360		

RT Services – Analysis and software delivery and support

Tools for analytics and big data research

- Software and application consulting – long term and short term
- Technical support
- Software packages like GIS and remote sensing software
- Purchase software
- Software licensing
- Benchmark and evaluate HPC systems
- Tutorials/Documentation
- Public software availability
- Application tuning



RT Services – Services for biomedical, biological, and health-related research

Clinical, genomic, and HIPAA-aligned services

- Examples include:
 - The Advanced Biomedical IT Core (ABITC)
 - Open Science Grid, Grid Operations Center
 - Indiana CTSI HUB
 - INResearch
 - I2iConnect



RT Services – Services for biomedical, biological, and health-related research

- The Advanced Biomedical IT Core (ABITC)
 - REDCAP - a HIPAA aligned service to enable Research Electronic Data Capture by allowing research staff to design and deploy research study forms quickly and easily. REDCap may also be used to conduct electronic surveys for research.
 - Alfresco Share - a HIPAA aligned document sharing website meant to empower collaboration. The service is typically used by members of extended research groups to share information, and by groups of researchers during grant-writing.



RT Services – Services for biomedical, biological, and health-related research

- National Center for Genome Analysis Support
 - Consulting for biologists undertaking genome analysis and genomic researchers
 - Genome analysis software
 - Optimization
 - Hardening
 - Enhancement
 - Data storage
 - Results archiving
 - Delivery and assistance in use of supercomputer clusters




Funded by the
National Science
Foundation



RT Services – Visualization facilities and services

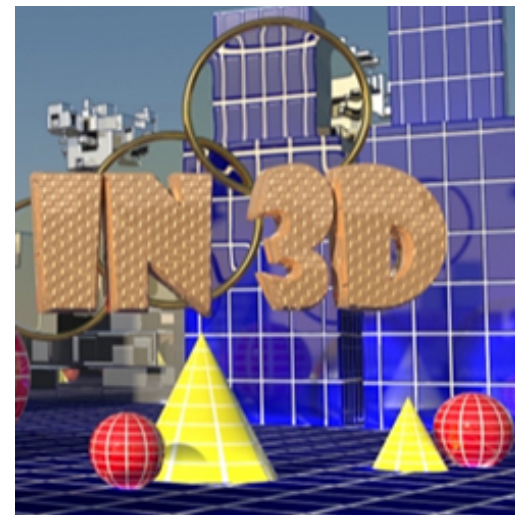
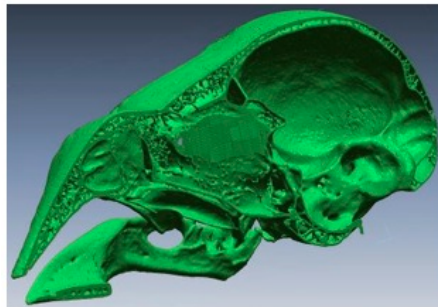
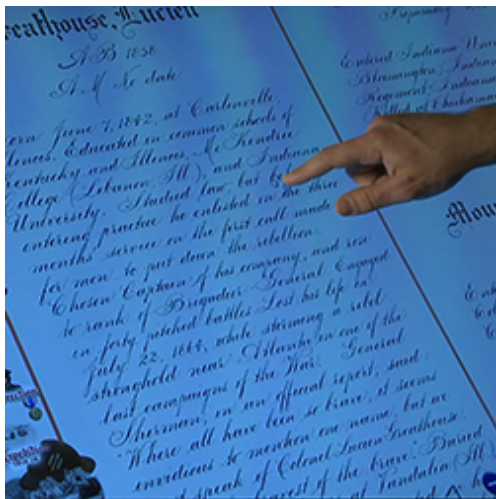
Understanding advanced simulations and big data

- Project consultation
 - Test and demonstrate new technologies and procedures
 - Tours of IU visualization facilities
 - Facility reservations
 - Vis & Collaboration Theater
 - Virtual Reality Theater
 - Science on a Sphere
- 
- IQ-Series - Wall/Tilt/Table
 - Digitization & 3D
 - Advanced Media Capture
 - Haptic Interface Systems

RT Services – Visualization facilities and services

Examples:

- *Digital Golden Book*
- *3D scanning, analysis, and printing*
- *Generating Interest in STEM using 3D*





**UITS Research
Technologies**
September 17, 2014



Science
Gateways



Storage



Computation



Analysis/
Software



Bio/Health



Visualization



Campus
Bridging



Education/
Outreach



Grant
Support



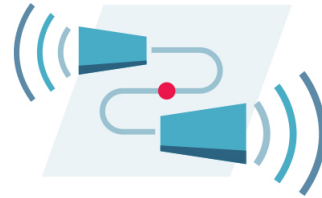
Cyber risk
Mitigation

IT-28

RT Services – Campus Bridging: connecting local and national cyberinfrastructure

Powerful resources as easy to use as your computer's peripheral devices

- Examples include:
 - IUPUI researchers who are using DC-WAN II to transfer data that's local to their laboratory to the HPC systems at IUB
 - Research Database Complex Web Services
 - GitHub Enterprise
 - Open Source Mirror
 - Red Hat Enterprise Linux



RT Services – Education and Outreach

Can help you figure out which RT services might benefit your research and will help with planning events regarding IU Cyberinfrastructure

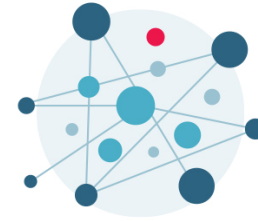
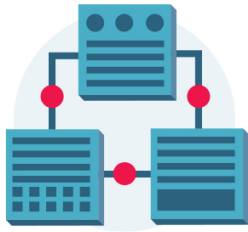
- K-12 outreach activities [Ready, Set, Robots!](https://www.youtube.com/watch?v=g7nDId4Ndtw)
<https://www.youtube.com/watch?v=g7nDId4Ndtw>
- In-person classes and workshops
- InfoShares tailored to your campus, department, class
- Tours of facilities like the Data Center or Science on a Sphere
- On-line tutorials and workshops
- Represent IU at conferences and within HPC organizations



RT Services – Digital Humanities

Provides support to all service areas with an emphasis on digital humanities.

- Recently added brand new service!



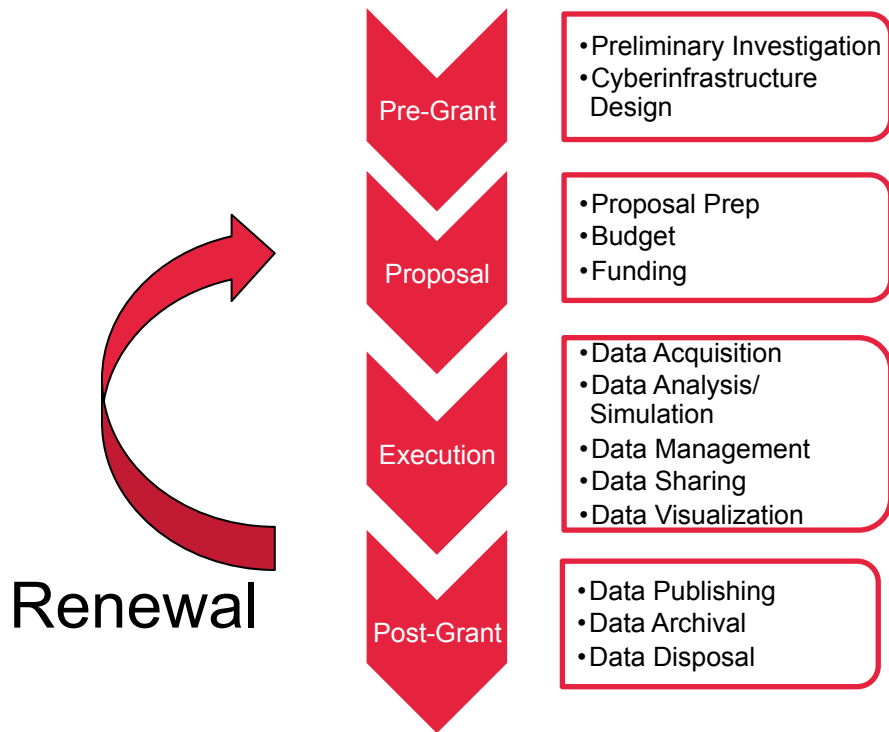
RT Services – Grant support and custom fee-based services

Helps with proposals that require cyberinfrastructure.
Application development and custom programming for a fee.

- Data Management Plans
- Facilities Statements
- Letters of Support
- Application Development
- Custom programming



The Grant/Data Life Cycle



RT Services – Services in support of cyber risk mitigation and IT-28 implementation

Data Center and Cluster services to secure research data

- II-Basic and IUFS
- Scholarly Data Archive (SDA)
- Condominium computing
- Cluster as a service – Penguin Computing



The Health Insurance Portability and Accountability Act of 1996 (HIPAA)

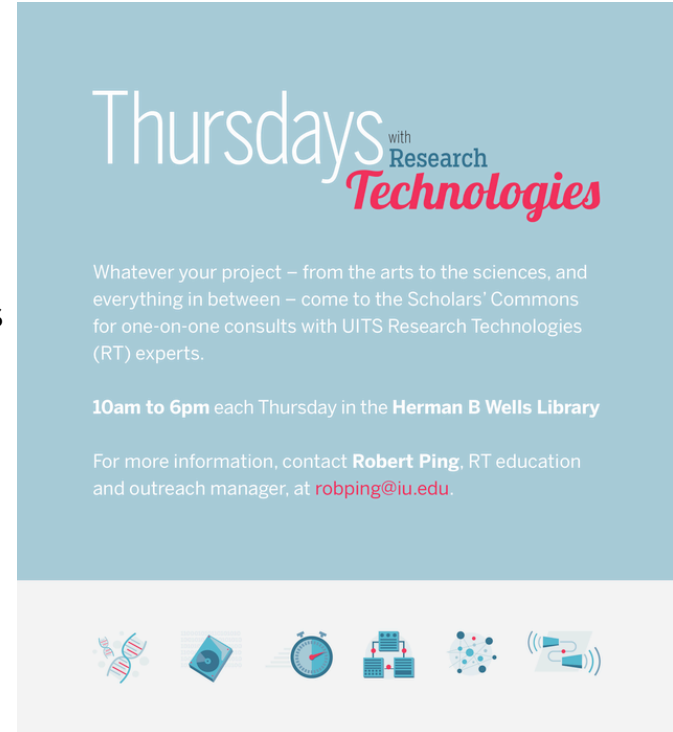
RT helps you be proactive and preventative

- Data Storage
 - Scholarly Data Archive (SDA)
 - Research File System (RFS)
- Shared and project scratch space
 - Data Capacitor II
- Research databases
 - Research Database Complex (RDC)
- Data Management
 - REDCap
- Collaboration and document sharing
 - Alfresco Share

Contact us ...

Right here in the library -

- Every Thursday
 - 1st – Storage and Campus Bridging
 - 2nd – Vis, Grant Support, and Science Gateways
 - 3rd – Medical/bio research, and EOT
 - 4th – Computation and Analysis
- Sign-up for the CI Newsletter



Thursdays with Research **Technologies**

Whatever your project – from the arts to the sciences, and everything in between – come to the Scholars' Commons for one-on-one consults with UITS Research Technologies (RT) experts.

10am to 6pm each Thursday in the **Herman B Wells Library**

For more information, contact **Robert Ping**, RT education and outreach manager, at robping@iu.edu.

Icons: DNA, storage, clock, server, network, speaker

Contact Information

Email: researchtech@iu.edu

Contact:

Robert Ping

Education and Outreach Manager

robbing@iu.edu



Please help us help you: InfoShare Survey <http://go.iu.edu/alj>