Computation Institute

Big process for big data

Ian Foster

foster@anl.gov

NASA Goddard, February 27, 2013

Argonne NATIONAL LABORATORY

computationinstitute.org

The Computation Institute

= UChicago + Argonne

= Cross-disciplinary nexus

= Home of the Research Cloud

High energy physics

Genetics

Molecular biology

Climate change

Cosmology

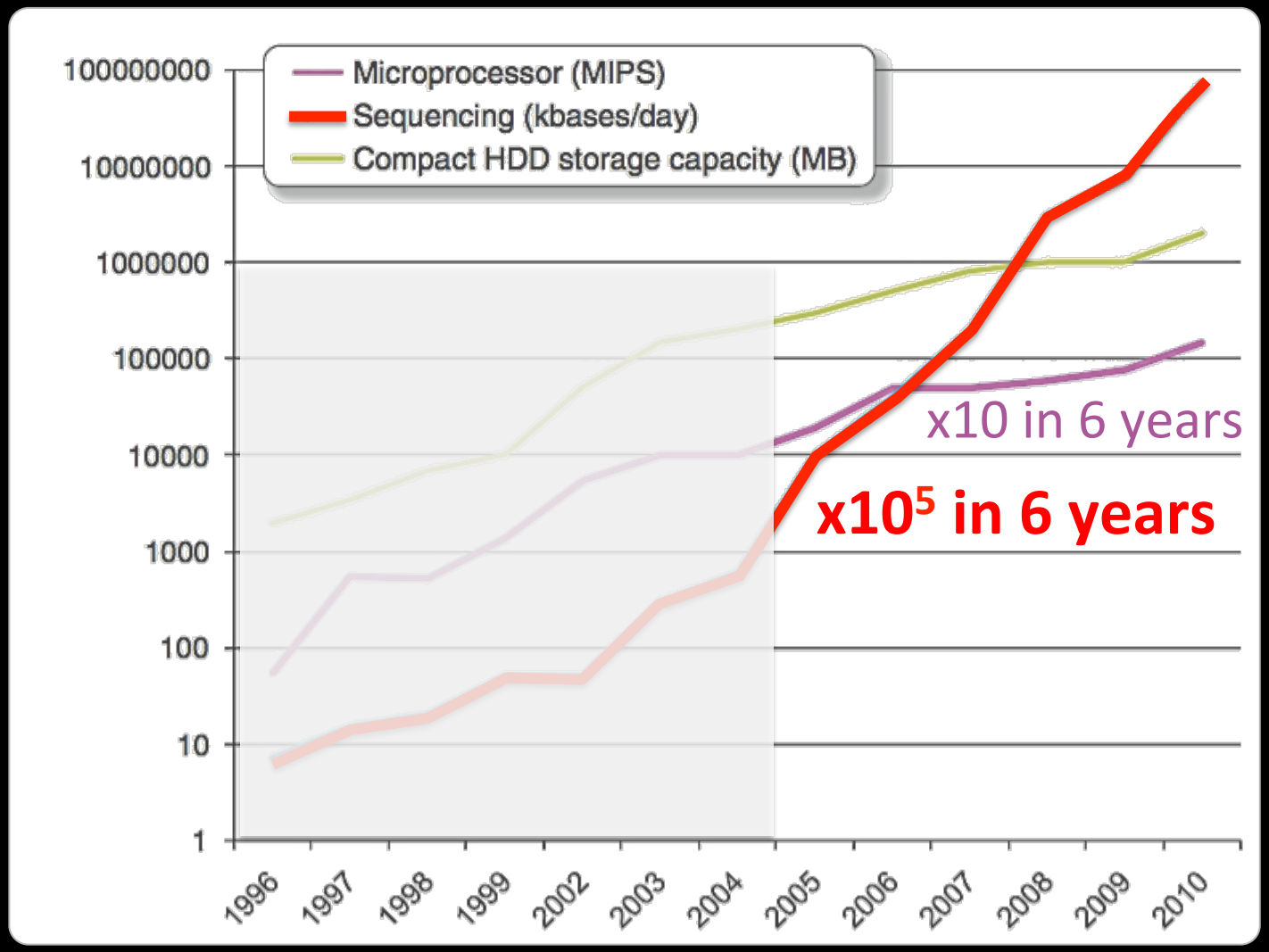
Linguistics

Metagenomics

Economics

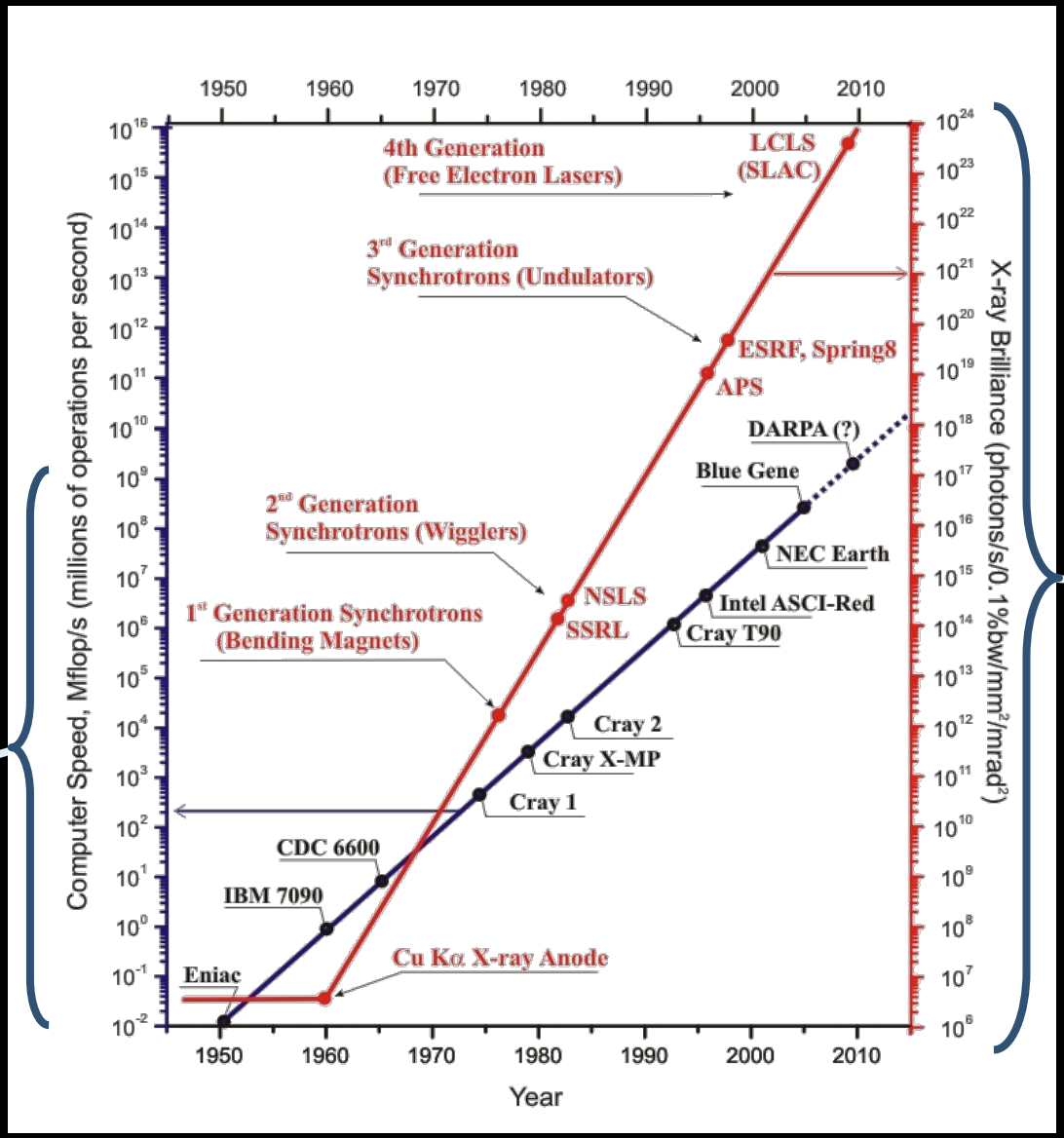
Visual arts

Will data kill genomics?



Kahn, Science, 331 (6018): 728-729

Moore's Law for X-Ray Sources



12 orders of magnitude In 6 decades!

18 orders of magnitude in 5 decades!

Earth System Grid Federation

ESGF

1.2 PB of climate data

Delivered to 23,000 users

We have exceptional infrastructure for the 1%

We have exceptional infrastructure for the 1%

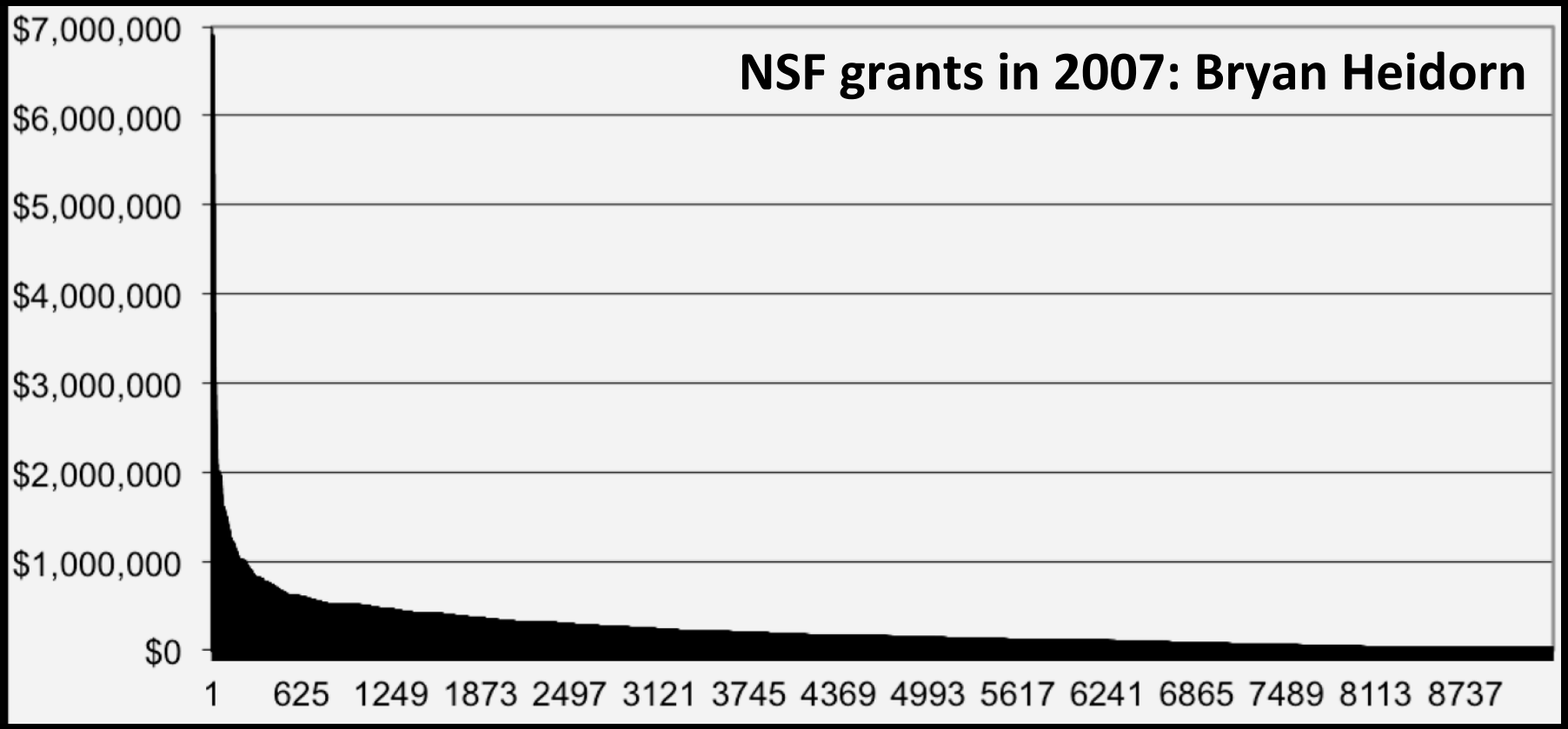
What about the 99%?

We have exceptional infrastructure for the 1%

What about the 99%?

Big science. Small labs.

The long tail of research



80% of awards and 50% of $$ are for grants < $350,0000

Can they remain competitive?

"Well, in our country," said Alice … "you'd generally get to somewhere else — if you run very fast for a long time, as we've been doing."

"A slow sort of country!" said the Queen. "Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must **run at least twice as fast as that**!"

Need: A new way to deliver research cyberinfrastructure

Frictionless
Affordable
Sustainable

We asked ourselves:

What if the research work flow could be managed as easily as…

flickr

…our pictures

…our e-mail

Gmail by Google

NETFLIX

…home entertainment

What makes these services great?

Great **User Experience**

+

High performance

(but invisible) **infrastructure**

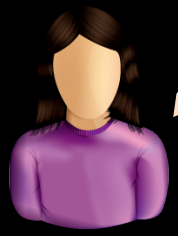
We aspire (initially) to create a great user experience for **research data management**

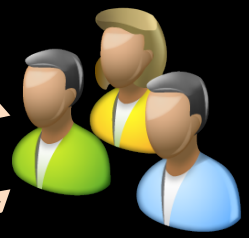
What would a “dropbox for science” look like?

- Collect
- Move
- Sync
- Share
- Analyze
- Annotate
- Publish
- Search
- Backup
- Archive

...for **BIG DATA**

A common work flow…





Staging Store

Ingest Store

Registry

Community Store

Analysis Store

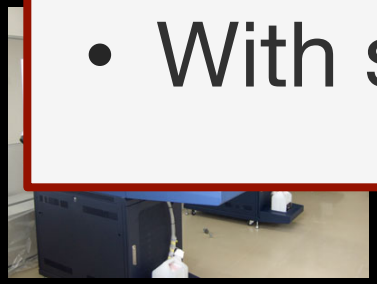
Archive

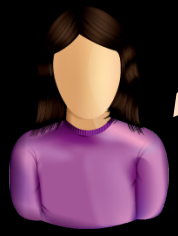
Mirror

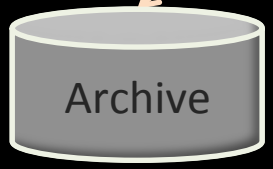
… with common challenges

Data movement, sync, and sharing

- Between facilities, archives, researchers
- Many files, large data volumes
- With security, reliability, performance



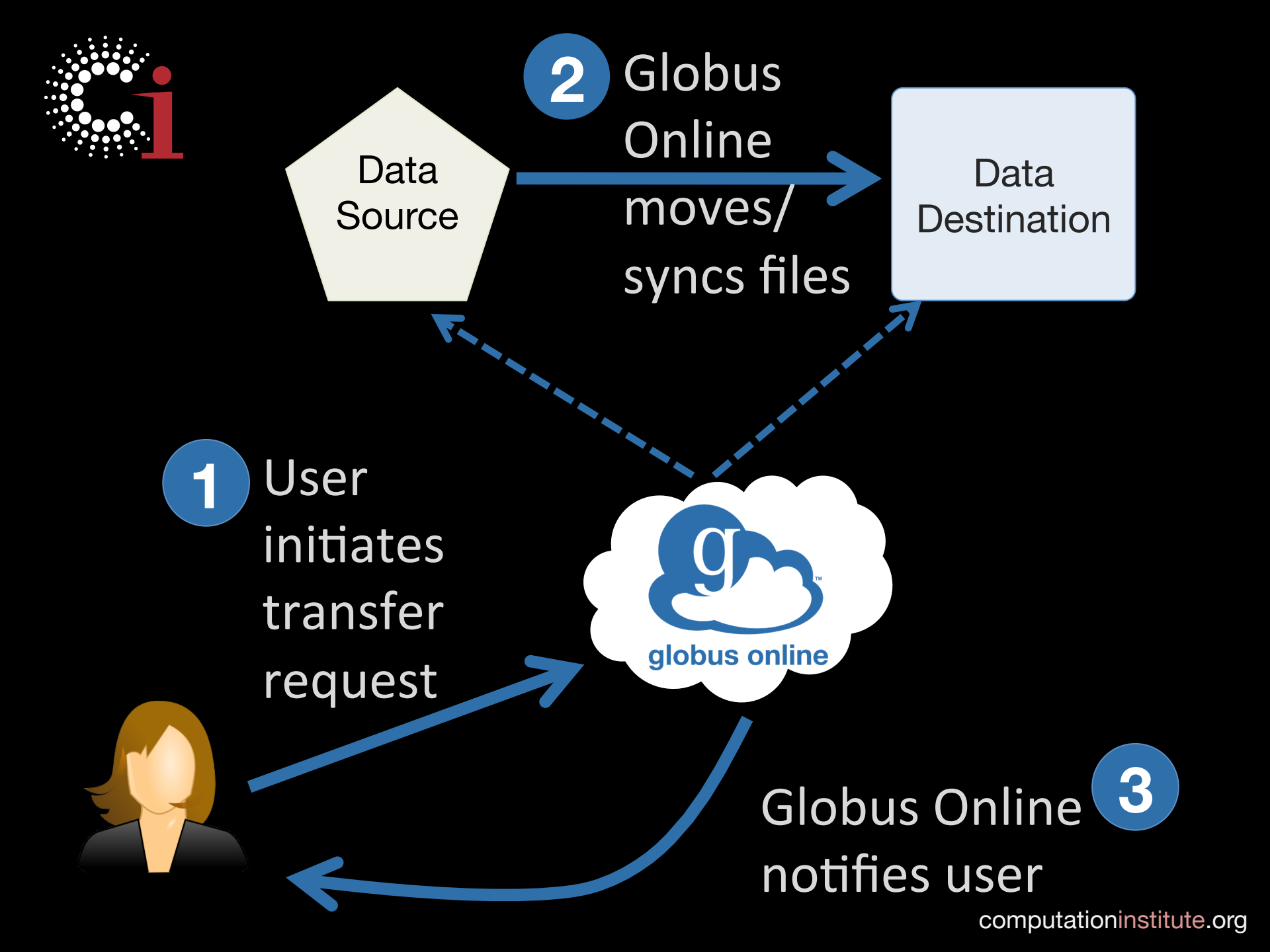


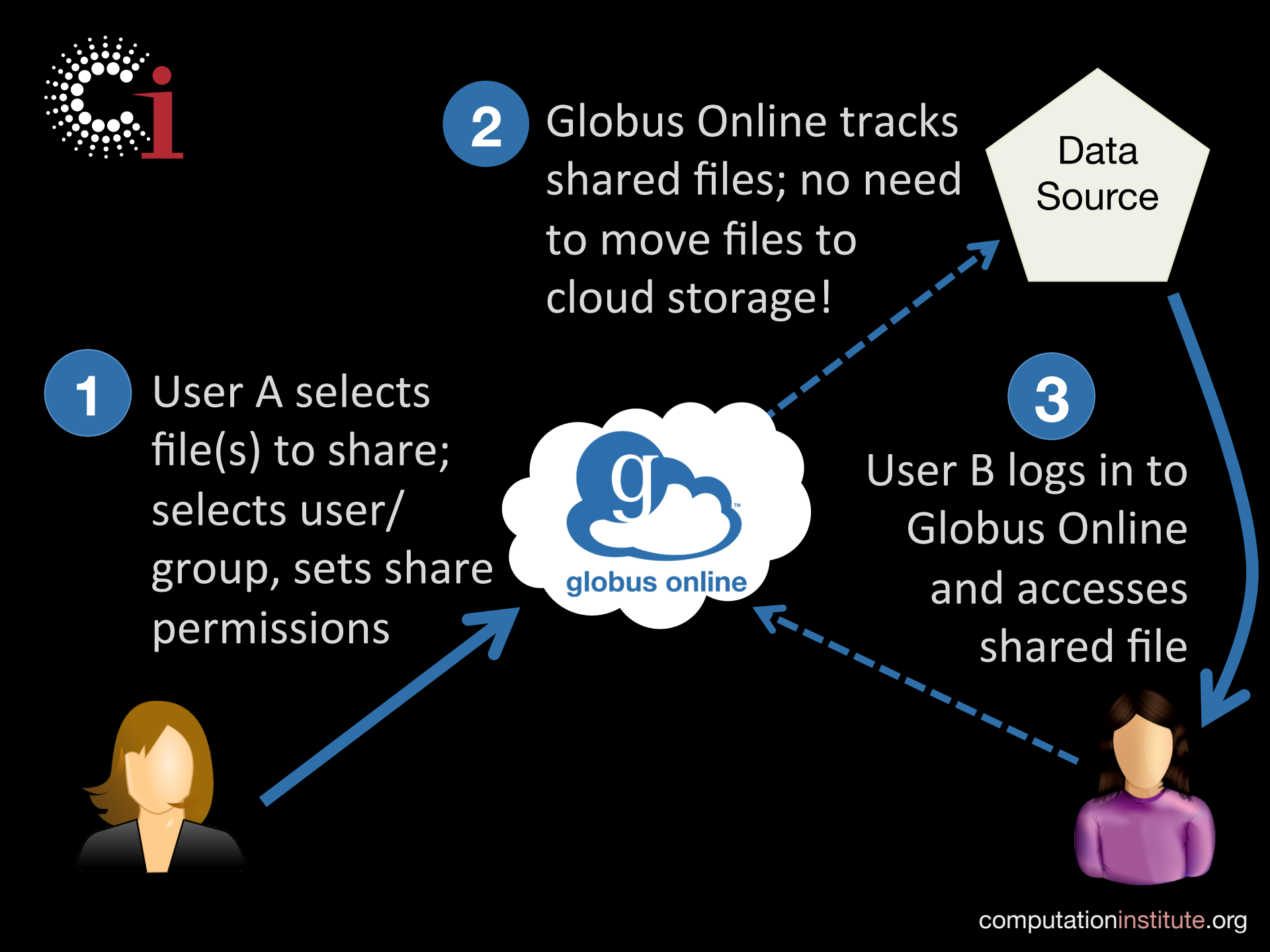


- Collect
- Annotate
- **Move**
- **Sync**
- **Share**
- Publish
- Search
- Backup
- Analyze
- Archive

globus online

Capabilities delivered using Software-as-Service (SaaS) model


2 Globus Online moves/ syncs files
Data Source
Data Destination
1 User initiates transfer request
globus online
Globus Online notifies user 3
computationinstitute.org


2 Globus Online tracks shared files; no need to move files to cloud storage!
Data Source
1 User A selects file(s) to share; selects user/ group, sets share permissions
globus online
3 User B logs in to Globus Online and accesses shared file
computationinstitute.org

Extreme ease of use

- InCommon, Oauth, OpenID, X.509, …
- Credential management
- Group definition and management
- Transfer management and optimization
- Reliability via transfer retries
- Web interface, REST API, command line
- One-click “Globus Connect” install
- 5-minute Globus Connect Multi User install

Early adoption is encouraging

XSEDE
Extreme Science and Engineering Discovery Environment

NERSC

HARVARD

MICHIGAN

UNIVERSITY OF EXETER

UNIVERSITY of WASHINGTON

NCSA

Carnegie Mellon University

NGS

APS physics

NORTHWESTERN UNIVERSITY 1851

THE UNIVERSITY OF AUCKLAND
NEW ZEALAND
Te Whare Wānanga o Tāmaki Makaurau

ISI
Information Sciences Institute

KSU

NASA

INDIANA UNIVERSITY

Fermilab

THE JOHNS HOPKINS UNIVERSITY BALTIMORE 1876

EMORY

ESnet
Energy Sciences Network

BERKELEY LAB
Lawrence Berkeley National Laboratory

Los Alamos
NATIONAL LABORATORY
EST. 1943

CERN

CORNELL UNIVERSITY

Cal

Ole Miss

THE UNIVERSITY OF CHICAGO

NEW YORK UNIVERSITY

Argonne
NATIONAL LABORATORY

Early adoption is encouraging

8,000 registered users; ~100 daily

~10 PB moved; ~1B files

10x (or better) performance vs. scp

99.9% availability

Entirely hosted on AWS

Delivering a great user experience relies on **high performance network infrastructure**

Science DMZ + optimizes performance

Bandwidth Requrements to move Y Bytes of data in Time X

Bits per Second Requirements

	1H	8H	24H	7Days	30Days
10PB	25,020.0 Gbps	3,127.5 Gbps	1,042.5 Gbps	148.9 Gbps	34.7 Gbps
1PB	2,502.0 Gbps	312.7 Gbps	104.2 Gbps	14.9 Gbps	3.5 Gbps
100TB	244.3 Gbps	30.5 Gbps	10.2 Gbps	1.5 Gbps	339.4 Mbps
10TB	24.4 Gbps	3.1 Gbps	1.0 Gbps	145.4 Mbps	33.9 Mbps
1TB	2.4 Gbps	305.4 Mbps	101.8 Mbps	14.5 Mbps	3.4 Mbps
100GB	238.6 Mbps	29.8 Mbps	9.9 Mbps	1.4 Mbps	331.4 Kbps
10GB	23.9 Mbps	3.0 Mbps	994.2 Kbps	142.0 Kbps	33.1 Kbps
1GB	2.4 Mbps	298.3 Kbps	99.4 Kbps	14.2 Kbps	3.3 Kbps
100MB	233.0 Kbps	29.1 Kbps	9.7 Kbps	1.4 Kbps	0.3 Kbps

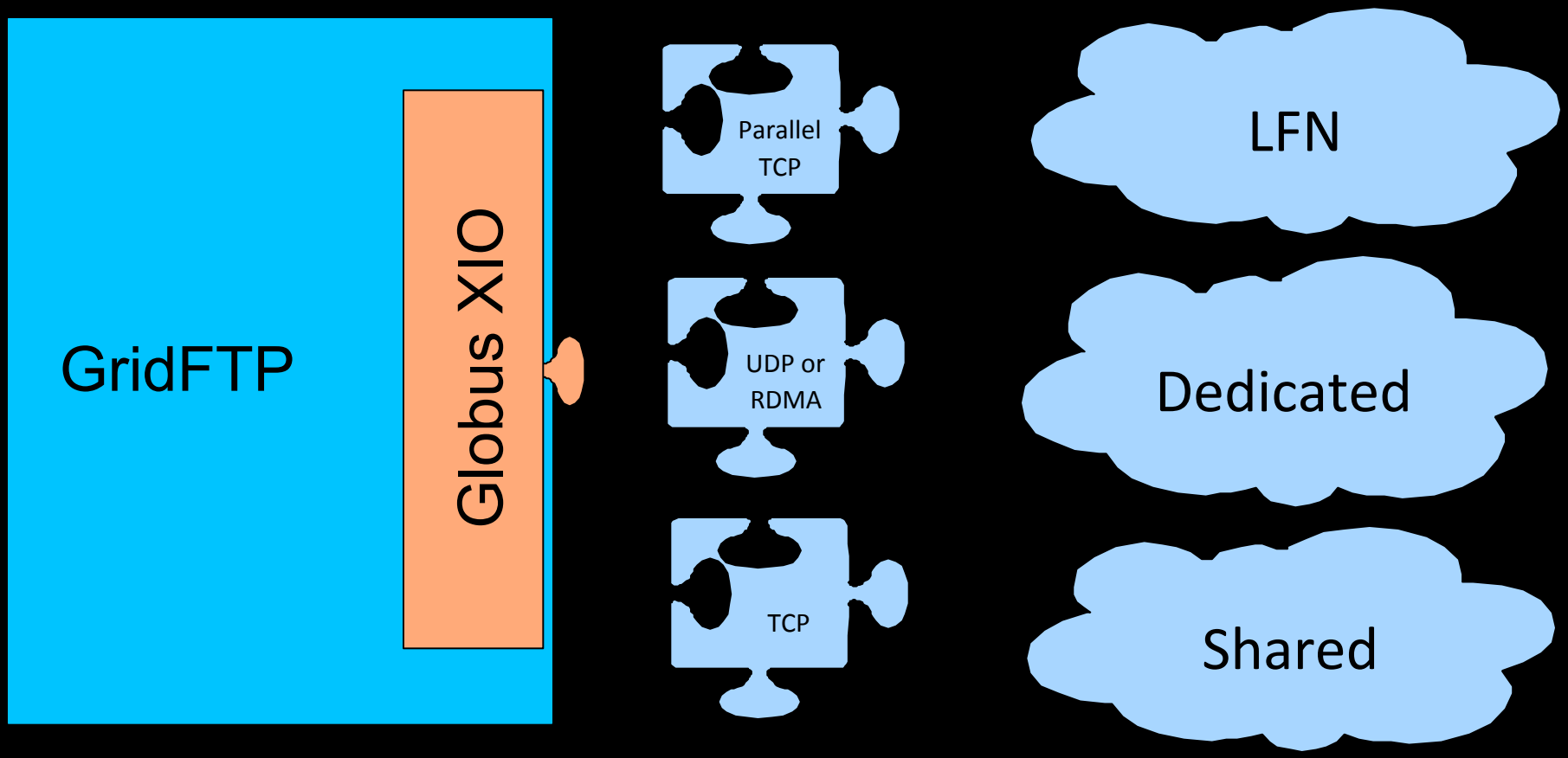
What is a Science DMZ?

Three key components, all required:

- "Friction free" network path
  - Highly capable network devices (wire-speed, deep queues)
  - Virtual circuit connectivity option
  - Security policy and enforcement specific to science workflows
  - Located at or near site perimeter if possible
- Dedicated, high-performance Data Transfer Nodes (DTNs)
  - Hardware, operating system, libraries optimized for transfer
  - Optimized data transfer tools: Globus Online, GridFTP
- Performance measurement/test node
  - perfSONAR

Details at http://fasterdata.es.net/science-dmz/

Globus GridFTP architecture



Internal layered XIO architecture allows alternative network and filesystem interfaces to be **plugged in** to the stack

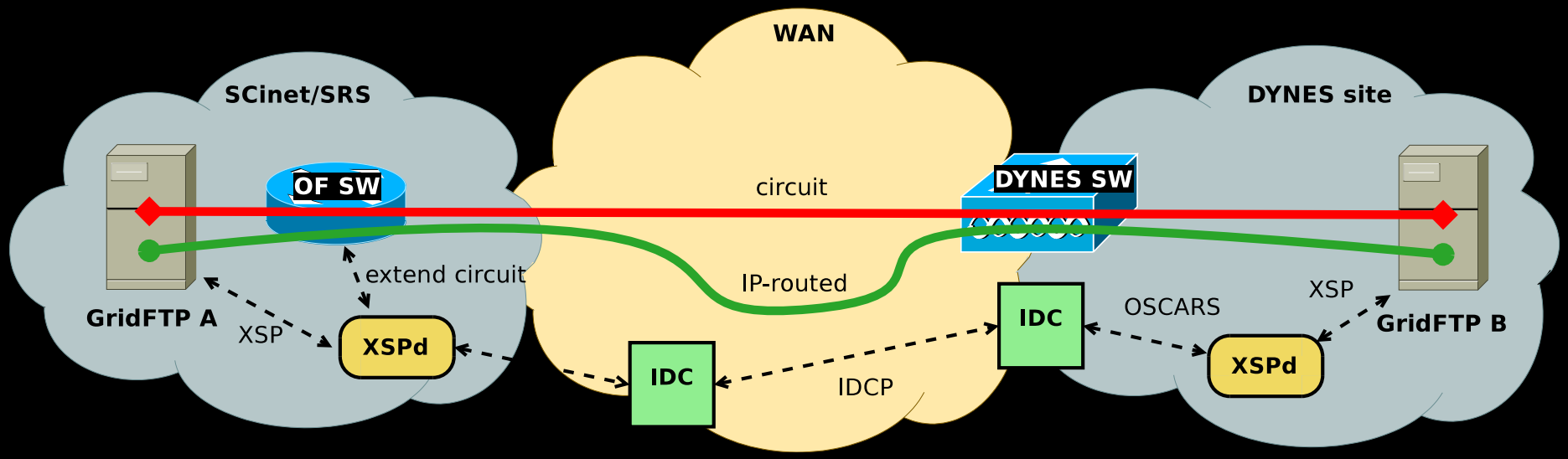
GridFTP performance options

- TCP configuration
- Concurrency: Multiple flows per node
- Parallelism: Multiple nodes
- Pipelining of requests to support small files
- Multiple cores for integrity, encryption
- Alternative protocol selection*
- Use of circuits and multiple paths*

Globus Online can configure these options based on what it knows about a transfer

* Experimental

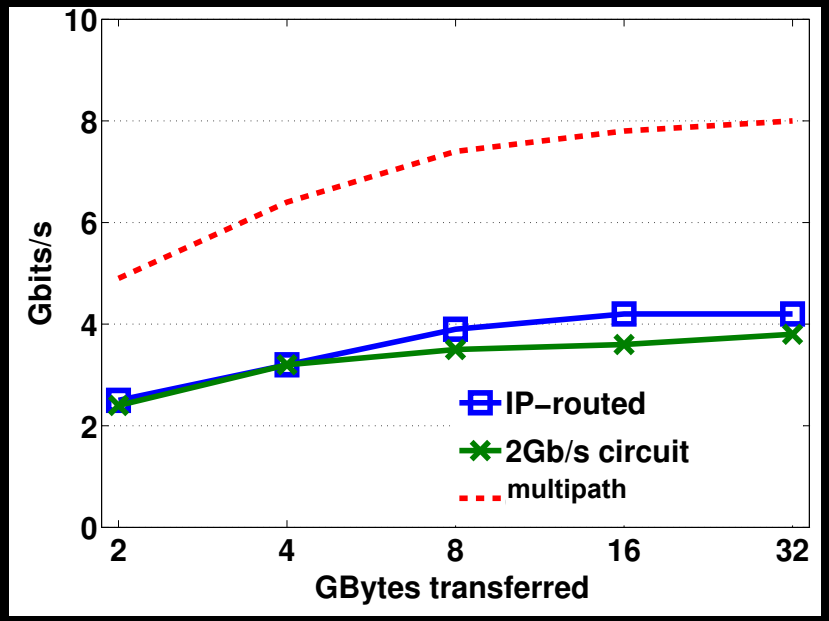
Exploiting multiple paths



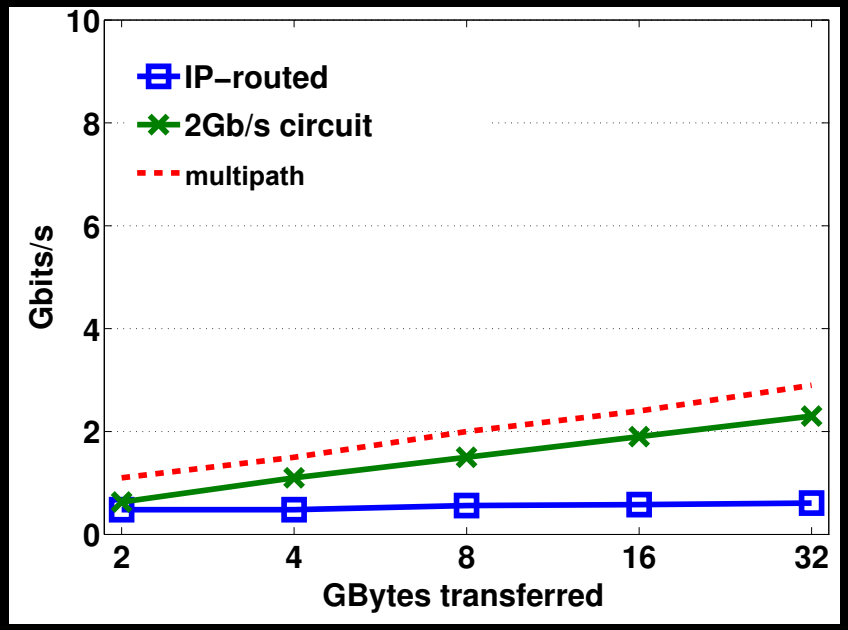
- Take advantage of multiple interfaces in multi-homed data transfer nodes
- Use circuit as well as production IP link
- Data will flow even while the circuit is being set up
- Once circuit is set up, use both paths to improve throughput

Exploiting multiple paths

Transfer between NERSC and ANL

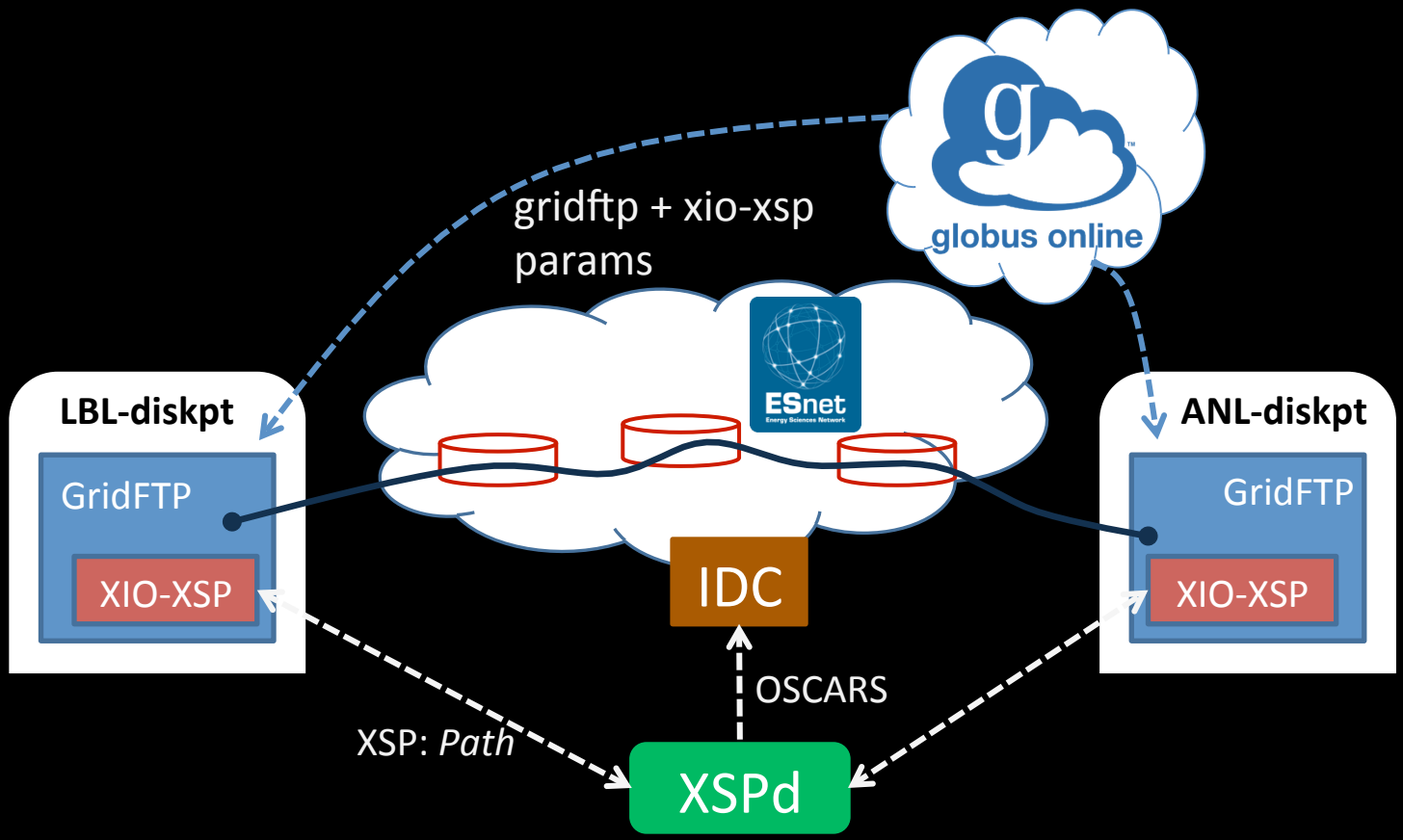


Transfer between UMich and Caltech



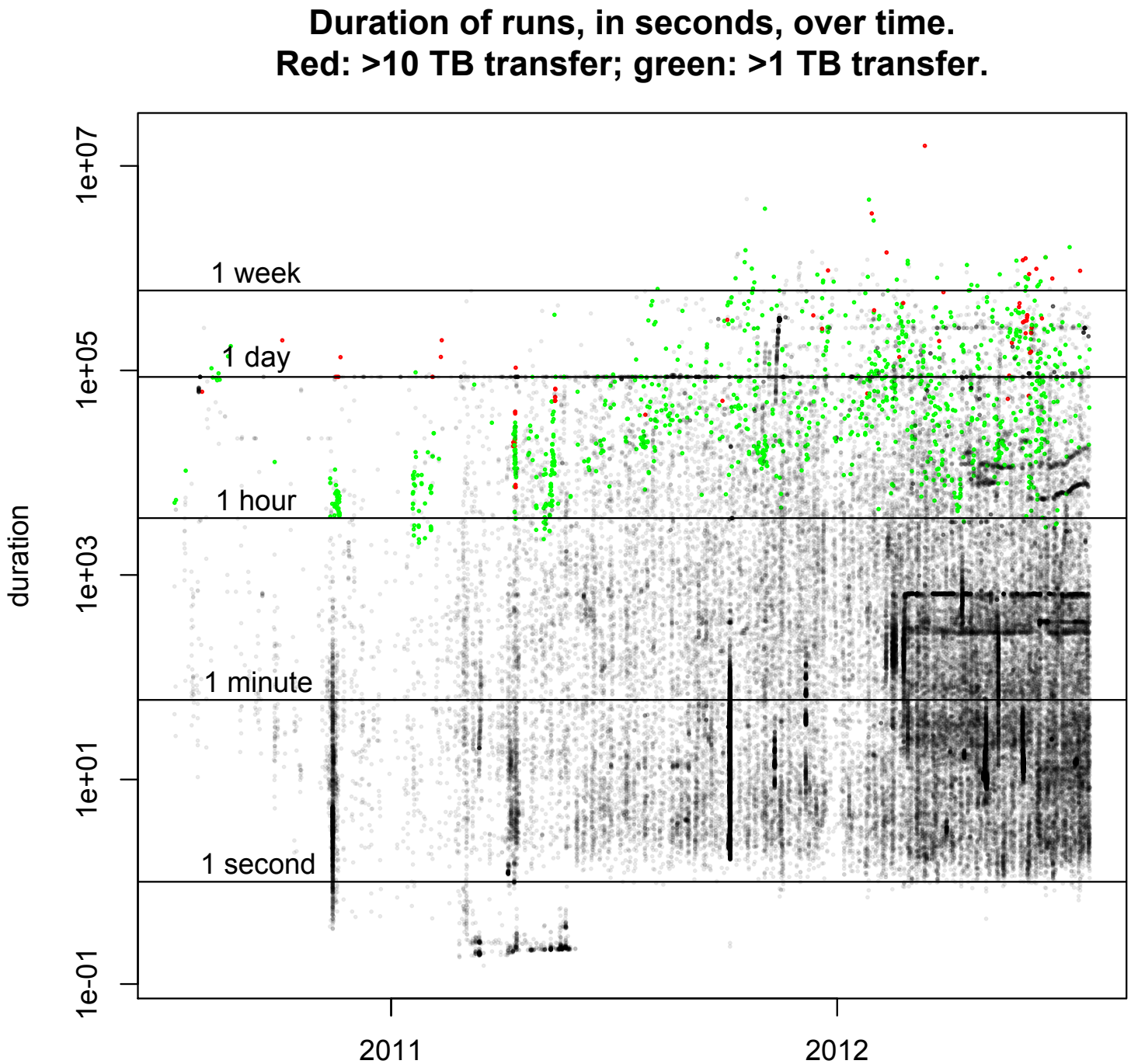
Default, commodity IP routes
+ Dedicated circuits
= Significant performance gains

Raj Kettimuthu, Ezra Kissel, Martin Swany, Jason Zurawski, Dan Gunter



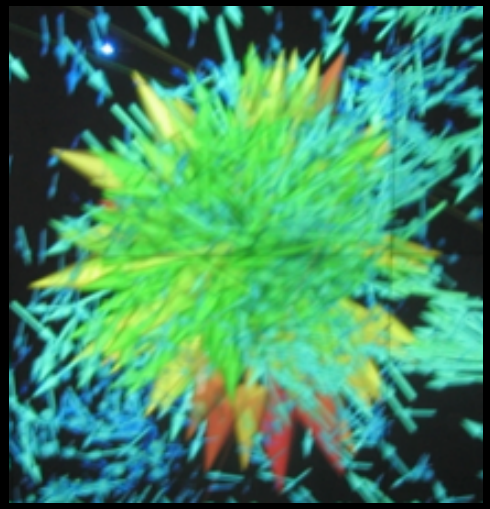
Provisioning OSCARS circuit for Globus Online transfers using Extensible Session Protocol (XSP)

Raj Kettimuthu, Ezra Kissel, Martin Swany

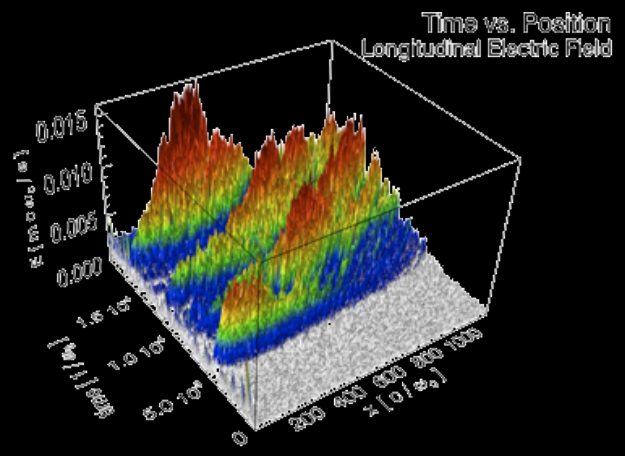

Duration of runs, in seconds, over time.
Red: >10 TB transfer; green: >1 TB transfer.
duration
1e+07
1e+05
1e+03
1e+01
1e-01
1 week
1 day
1 hour
1 minute
1 second
2011
2012

Dark Energy Survey

Receives 100,000 files each night in IL
Transmits files to Texas for analysis…
…then moves results back to IL
Globus Online makes this process reliable, routine, and efficient



K. Heitmann (Argonne) moves 22 TB of **cosmology** data LANL → ANL at 5 Gb/s



B. Winjum (UCLA) moves 900K-file **plasma physics** datasets UCLA →NERSC

Dan Kozak (Caltech)
replicates 1 PB LIGO
astronomy data for resilience

- Collect
- Move
- Sync
- Share
- Analyze
- Annotate
- Publish
- Search
- Backup
- Archive

...for BIG DATA

- Collect
- Move
- Sync
- Share
- Analyze
- Annotate
- Publish
- Search
- Backup
- Archive

...for **BIG DATA**

Many more capabilities planned …

Globus Online Research Data Management-as-a-Service

Ingest, Cataloging, Integration	**Sharing, Collaboration, Annotation**	**Backup, Archival, Retrieval**	**…**	**SaaS**
Globus Integrate (Globus Nexus, Globus Connect)				**PaaS**

A platform for integration

Catalog as a service

Approach

- Hosted user-defined catalogs
- Based on tag model

  <subject, name, value>
- Optional schema constraints
- Integrated with other Globus services

Three REST APIs

/query/

- Retrieve subjects

/tags/

- Create, delete, retrieve tags

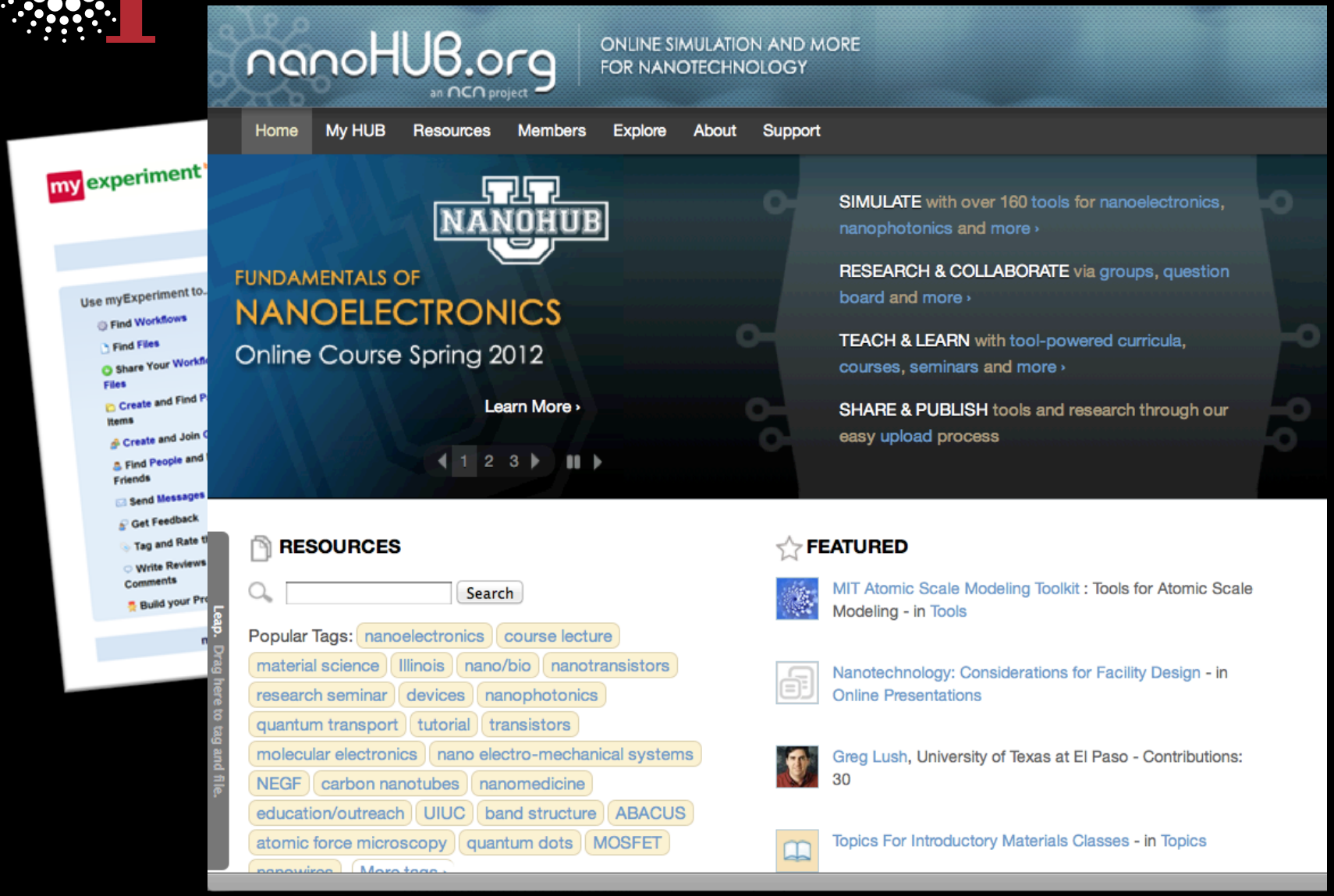
/tagdef/

- Create, delete, retrieve tag definitions

Builds on USC Tagfiler project (C. Kesselman et al.)

Other early successes in services for science…

myexperiment beta
myExperiment makes it really easy to find, use and share scientific workflows and other files, and to build communities.
All
Search
Use myExperiment to...
Find Workflows
Find Files
Share Your Workflows and Files
Create and Find Packs of Items
Create and Join Groups
Find People and Make Friends
Send Messages
Get Feedback
Tag and Rate things
Write Reviews and Comments
Build your Profile and
Explore
Find Workflows
About myExperiment
Join the Mailing List
Give us Feedback
For Developers
The myGrid Project
Taverna Workflow Workbench
The BioCatalogue Project
myExperiment Publications
Register
or Login:
Username or Email:
Password:
Remember me:
Or use OpenID:
(eg: name.myopenid.com)
Login
Forgot Password?
myExperiment has over 1250 users, 100 groups, 490 workflows, 130 files and 40 packs


nanoHUB.org
an ncn project
ONLINE SIMULATION AND MORE FOR NANOTECHNOLOGY
Home
My HUB
Resources
Members
Explore
About
Support
my experiment
Use myExperiment to...
Find Workflows
Find Files
Share Your Workflows Files
Create and Find Packs Items
Create and Join Groups
Find People and Friends
Send Messages
Get Feedback
Tag and Rate
Write Reviews Comments
Build your Profile
NANOHUB U
FUNDAMENTALS OF
NANOELECTRONICS
Online Course Spring 2012
Learn More ›
1 2 3
SIMULATE with over 160 tools for nanoelectronics, nanophotonics and more ›
RESEARCH & COLLABORATE via groups, question board and more ›
TEACH & LEARN with tool-powered curricula, courses, seminars and more ›
SHARE & PUBLISH tools and research through our easy upload process
Leap. Drag here to tag and file.
RESOURCES
Search
Popular Tags: nanoelectronics course lecture material science Illinois nano/bio nanotransistors research seminar devices nanophotonics quantum transport tutorial transistors molecular electronics nano electro-mechanical systems NEGF carbon nanotubes nanomedicine education/outreach UIUC band structure ABACUS atomic force microscopy quantum dots MOSFET
FEATURED
MIT Atomic Scale Modeling Toolkit : Tools for Atomic Scale Modeling - in Tools
Nanotechnology: Considerations for Facility Design - in Online Presentations
Greg Lush, University of Texas at El Paso - Contributions: 30
Topics For Introductory Materials Classes - in Topics

nanoHUB.org
ONLINE SIMULATION AND MORE FOR NANOTECHNOLOGY

MG-RAST

metagenomics analysis server

LOGIN REGISTER PASSWORD FORGOT?

login

Browse Metagenomes

search for metagenomes

Register Contact Help Upload* News

About

MG-RAST (the Metagenomics RAST) server is an automated analysis platform for metagenomes providing quantitative insights into microbial populations based on sequence data.

# of metagenomes	35,586
# base pairs	9.24 Tbp
# of sequences	85.21 billion
# of public metagenomes	7,167

The server provides web based upload, quality control, automated annotation and analysis for samples up to 10GBp. Comparison between large numbers of samples is enabled via pre-computed abundance profiles. MG-RAST was launched in 2007 and has over 5000 registered users and 35,586 data sets. The current server version is 3.1.2.

Updates | MG-RAST Version 3.1.2 released

Science EXCHANGE

Sign Up | Login | How It Works

Search for services, facilities...

The Scientific Services Marketplace

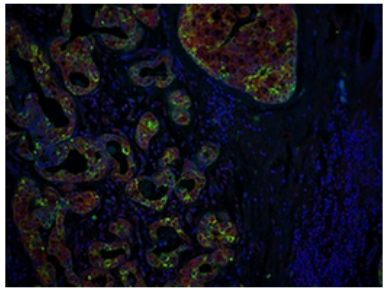
The easiest way to get experiments conducted by researchers in top core facilities and institutions.

Search All Services

Immunohistochemistry	$10.00 per Sample
Cloning Constructs	$162.00 per Sample
NGS Sequencing	$488.00 per Sample
HTS Screening	$0.10 per Sample
Mass Spectrometry	$10.00 per Sample
RNA microarray	$107.50 per Sample

Featured Immunohistochemistry Providers

Science Exchange has 22 verified Immunohistochemistry providers including the following featured providers.



Histopathology and Tissue Shared Resource

Georgetown Lombardi Comprehensive Cancer Center | Washington, DC, United States

The Histopathology and Tissue Shared Resource at Georgetown Lombardi Comprehensive Cancer Center offers a full complement of histology services through combined Human Tissue Bank, histology, and IHC departments.

$19.00 USD per Sample

REQUEST ESTIMATE

Georgetown | Lombardi · REVEAL · NYU School of Medicine · UAMS · UF University of Florida · +17 others · View →

Free

No cost for providers to list services • No cost for researchers to request services • Easy to pay any provider via payment platform

Trusted

Trusted by researchers from leading institutions including Stanford, USC, Princeton and UC Davis • Great customer service by real people

Safe

Secure access to the world's best research infrastructure • Leading SSL and site encryption technologies to protect site information

From The Blog

January 03, 2013

Bringing Clarity to the Jargon

VISIT BLOG →

my experiment beta

Browse Metag

About

MG-RAST (the Metage providing quantitative i

# of metagenomes
base pairs
of sequences
# of public metagenomes

Updates | MG-RAST Version 3.1.2 released

Our vision for a 21st century cyberinfrastructure

To provide **more** capability for **more** people at **substantially lower cost** by creatively **aggregating** (“cloud”) and **federating** (“grid”) resources —leveraging reliable, **high performance networks**

Our vision for a 21st century cyberinfrastructure

To provide **more** capability for **more** people at **substantially lower cost** by creatively **aggregating** (“cloud”) and **federating** (“grid”) resources

“Science as a service”

Thank you to our sponsors!

U.S. DEPARTMENT OF ENERGY

NSF

THE UNIVERSITY OF CHICAGO

NATIONAL INSTITUTES OF HEALTH

Argonne NATIONAL LABORATORY

computationinstitute.org