

Hosted services for managing shared cyberinfrastructure

Ian Foster

Argonne National Laboratory & The University of Chicago

Joint work with Rachana Ananthakrishnan, Josh Bryan,
Kyle Chard, Mattias Lidman, Steven Tuecke, and others

GENI Engineering Conference, NYC, October 28, 2013

Using **cloud** services to accelerate **discovery**

Ian Foster

Argonne National Laboratory & The University of Chicago

Joint work with Rachana Ananthakrishnan, Josh Bryan,
Kyle Chard, Mattias Lidman, Steven Tuecke, and others

GENI Engineering Conference, NYC, October 28, 2013



- “a technological and sociological solution to the problem of efficiently connecting laboratories, data, computers, and people with the goal of enabling derivation of novel scientific theories and knowledge” [Wikipedia]
- AKA eScience, eResearch, Computer Supported Collaborative Work, Grid, ...

“The Anatomy of the Grid,” 2001



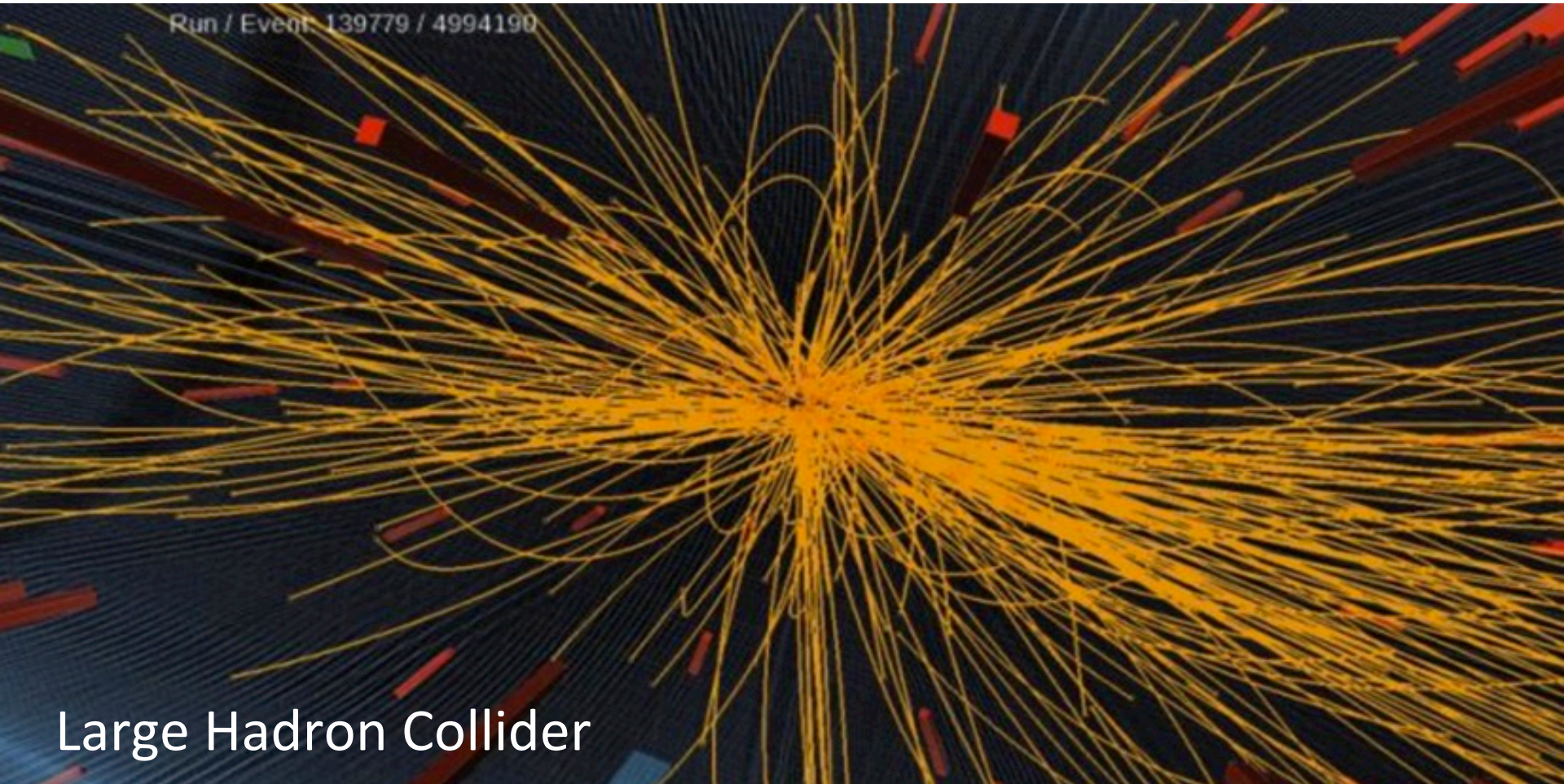
The ... problem that underlies the Grid concept is coordinated resource sharing and problem solving in **dynamic, multi-institutional virtual organizations**. The sharing that we are concerned with is not primarily file exchange but rather direct access to computers, software, data, and other resources, as is required by a range of collaborative problem-solving and resource-brokering strategies emerging in industry, science, and engineering. This sharing is, necessarily, highly controlled, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs. **A set of individuals and/or institutions defined by such sharing rules form what we call a virtual organization (VO).**

Grid technology accelerates discovery



Higgs discovery “only possible because of the **extraordinary achievements of ... grid computing**” —Rolf Heuer, CERN DG

Run / Event: 139779 / 4994190



Large Hadron Collider

LHC Computing Grid “virtual organizations”

Service View

VO View

Go to another VO

- astrogrid
- astron
- astrop
- ✓ atlas
- atlas-se
- atlas.ac.il
- atlas.o
- atlast3
- atlas_test
- auger
- augerint
- auvergrid
- babar
- balticgrid
- bauvogrid
- bbmri.nl
- beapps
- becms
- belle
- belle2
- belle2.org
- bes
- betest
- bfg
- bg-edu.grid.acad.bg
- bgtut
- biit
- bing.vo.ibergrid.eu
- bio
- bio.ndgf.org
- biocomp
- bioif
- biomed
- biotech.grid.acad.bg
- bisarid

Viewing Statistics or Graphs

CE Status

Show 10 entries



ID	TotalJobs	RunningJobs	WaitingJobs	FreeJobSlots	ERT	WRT	
/atlas/Role_lcgadmin	0	0	0	1	0	0	lcgceatlas pbs-atlas
/atlas/Role_pilot	144	19	125	0	17525	44236800	lcgceatlas pbs-anala
/atlas/Role_pilot	3	0	3	0	17525	44236800	lcgceatlas pbs-atlas
/atlas/Role_production	398	279	119	0	17770	41126400	lcgceatlas pbs-atlas
atlas	0	0	0	2	0	0	t2-ce-01.f
atlas	0	0	0	2	0	0	t2-ce-01.f
atlas	0	0	0	0	0	0	gridce2.p
atlas	309	170	139	0	83028	1317788	creamce
atlas	0	0	444444	0	2146660842	2146660842	glite.univ.
atlas	718	673	45	0	417543	81000000	node74.d atlasana
Total	162,407	111,907	14,274,109	27,721			

Showing 1 to 10 of 289 entries

<http://gstat2.grid.sinica.edu.tw/gstat/vo/atlas/>

3,761

GridFTP servers
reporting usage

169 million GRAM jobs submitted

by the Open Science Grid in 2012

300,000 jobs/day reported

382 million operations

29 petabytes transferred

during February 2013



Open Science Grid



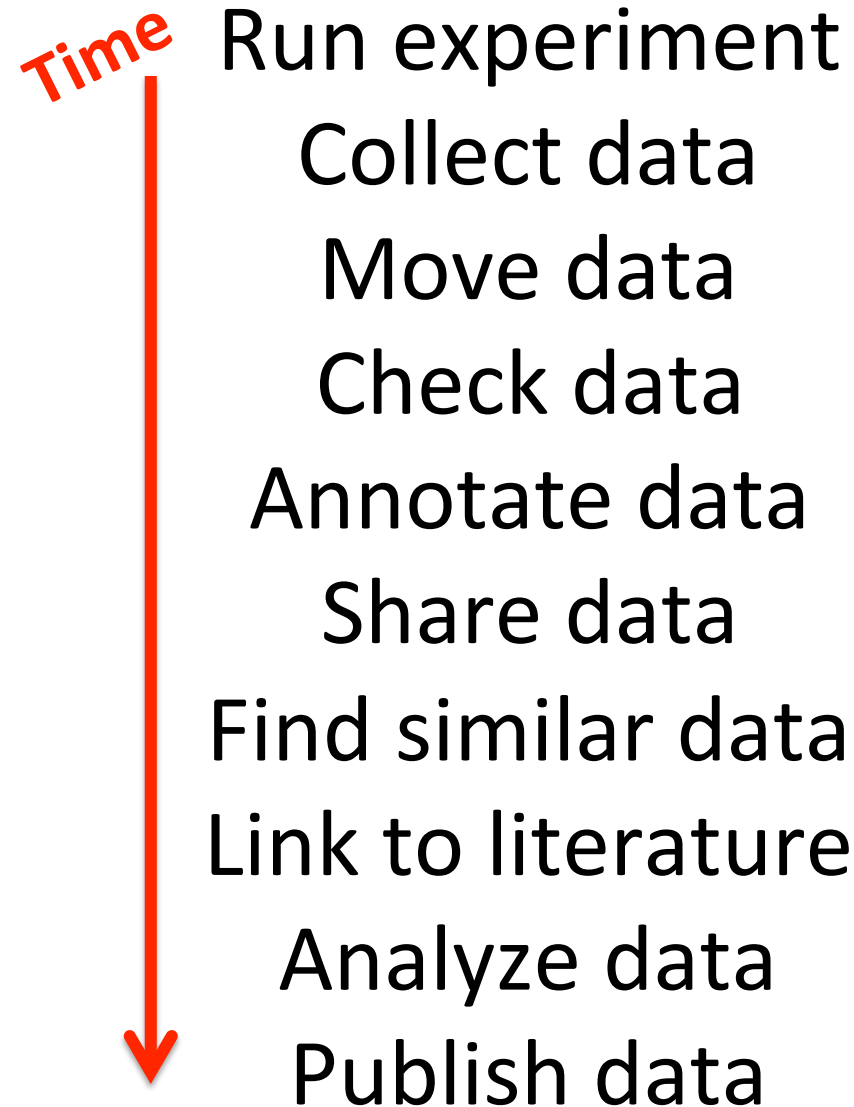
100 MyProxy servers

2,000,000 requests per week

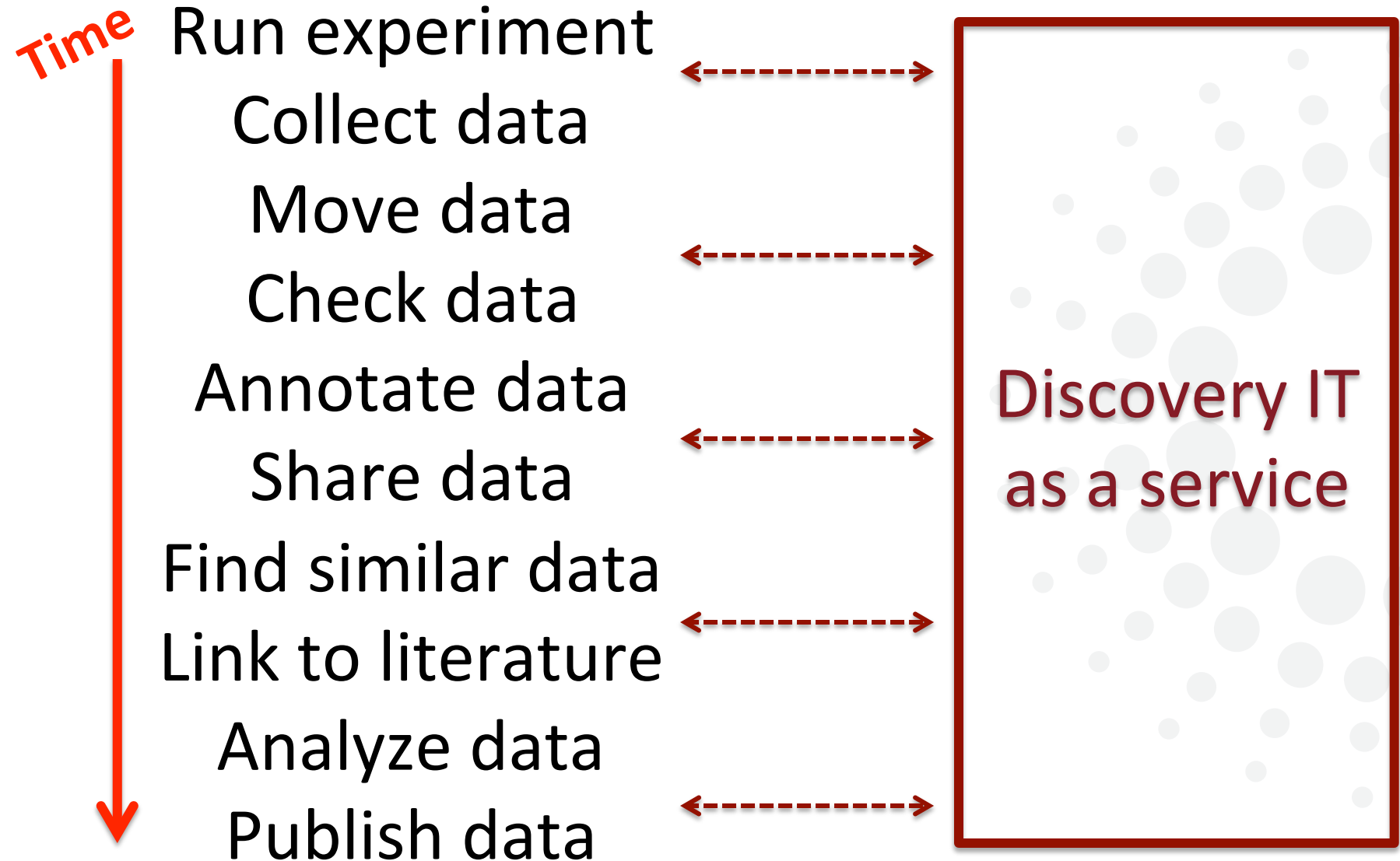
1,000 GSI-OpenSSH servers

1,000,000 login requests per week

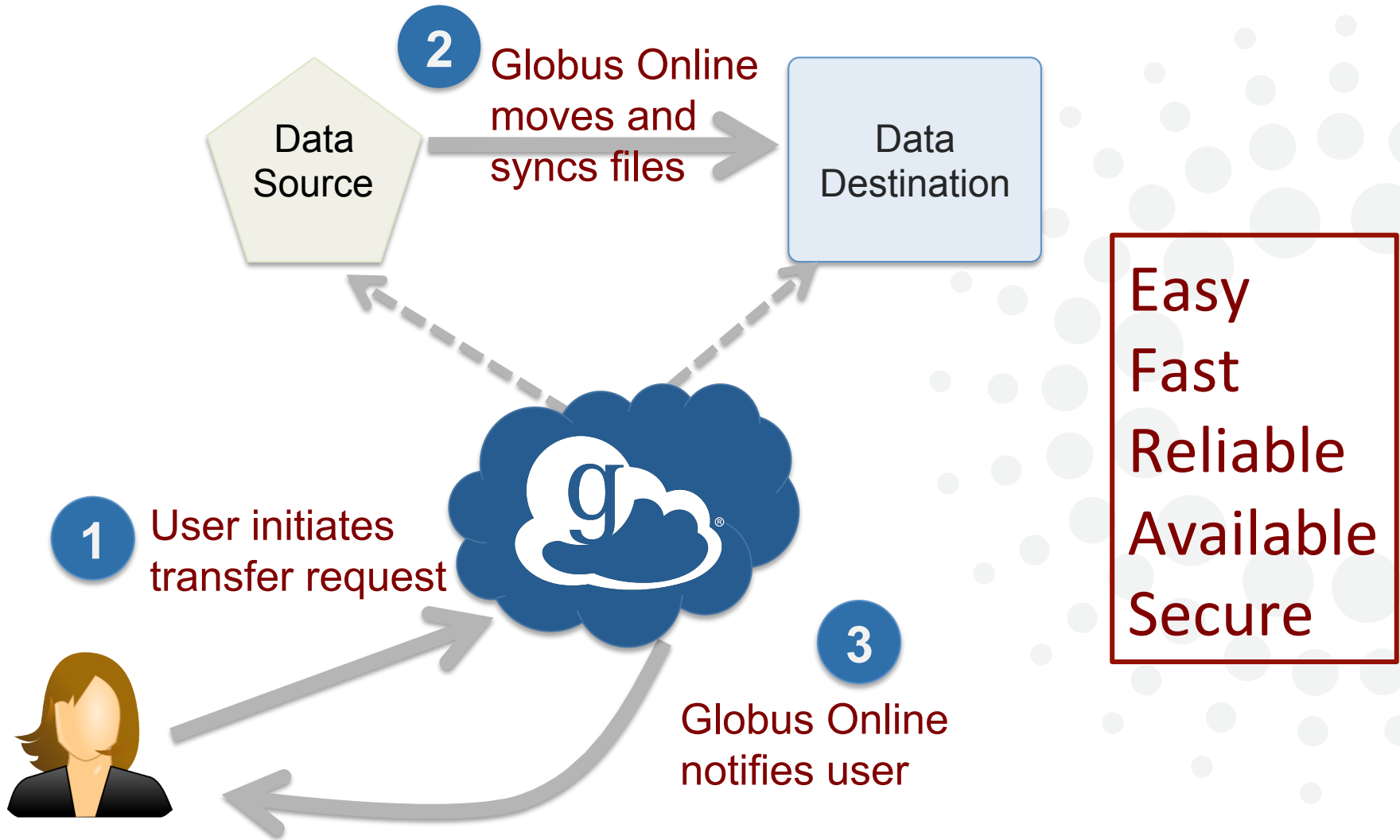
Complexity in research is large and growing



Process automation for discovery



First: File transfer as a service





26,312,788,281 MB
TRANSFERRED

researcher

Globus Online enables you to move, sync, and share your data using just a web browser. We take care of time consuming, error prone IT tasks so you can focus on your research.

 [sign up now](#)

You've got BIG data.

provider

Globus Online helps you deliver robust data management solutions to your researchers. We provide a secure, reliable, high performance service backed by great support.

 [learn more](#)

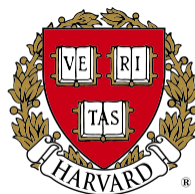
Early adoption is encouraging



XSEDE

Extreme Science and Engineering
Discovery Environment

NERSC



MICHIGAN

UNIVERSITY OF
EXETER

W
UNIVERSITY of
WASHINGTON

NCSA

**Carnegie
Mellon
University**

NGS

**APS
physics**



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

ISI
Information Sciences Institute

KSU



**INDIANA
UNIVERSITY**

Fermilab



EMORY

ESnet
Energy Sciences Network

BERKELEY LAB
Lawrence Berkeley National Laboratory

CERN

**Los Alamos
NATIONAL LABORATORY**
EST. 1943

**CORNELL
UNIVERSITY**

Cal

Ole Miss



THE UNIVERSITY OF
CHICAGO



NEW YORK UNIVERSITY

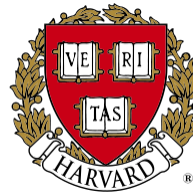
Argonne
NATIONAL LABORATORY

Early adoption is encouraging



XSEDE

Extreme Science and Engineering
Discovery Environment



Carnegie Mellon



12,000 registered users; >150 daily
>25 PB moved; >1B files
10x (or better) performance vs. scp
99.9% availability
Entirely hosted on Amazon



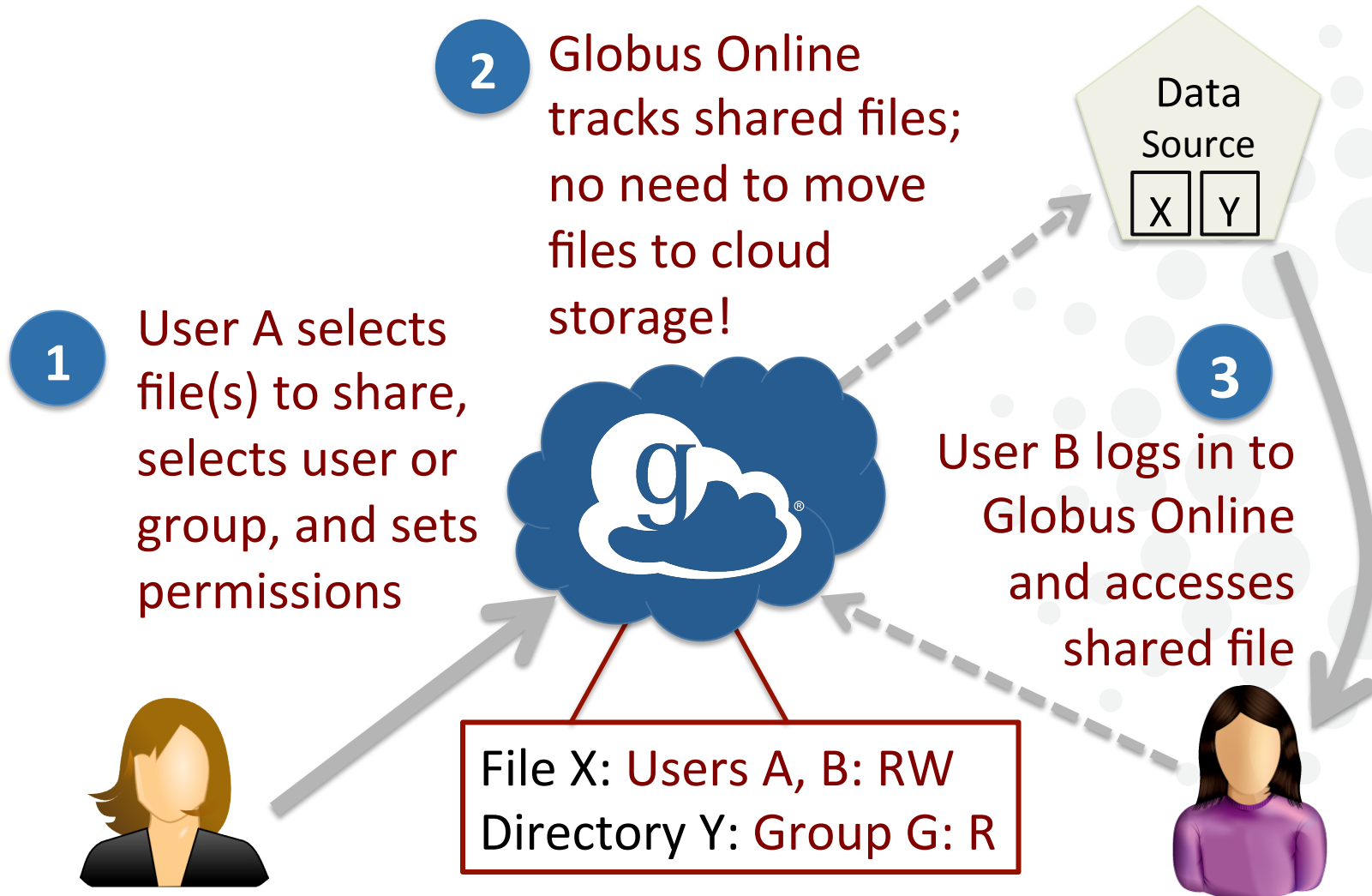
THE UNIVERSITY OF
CHICAGO



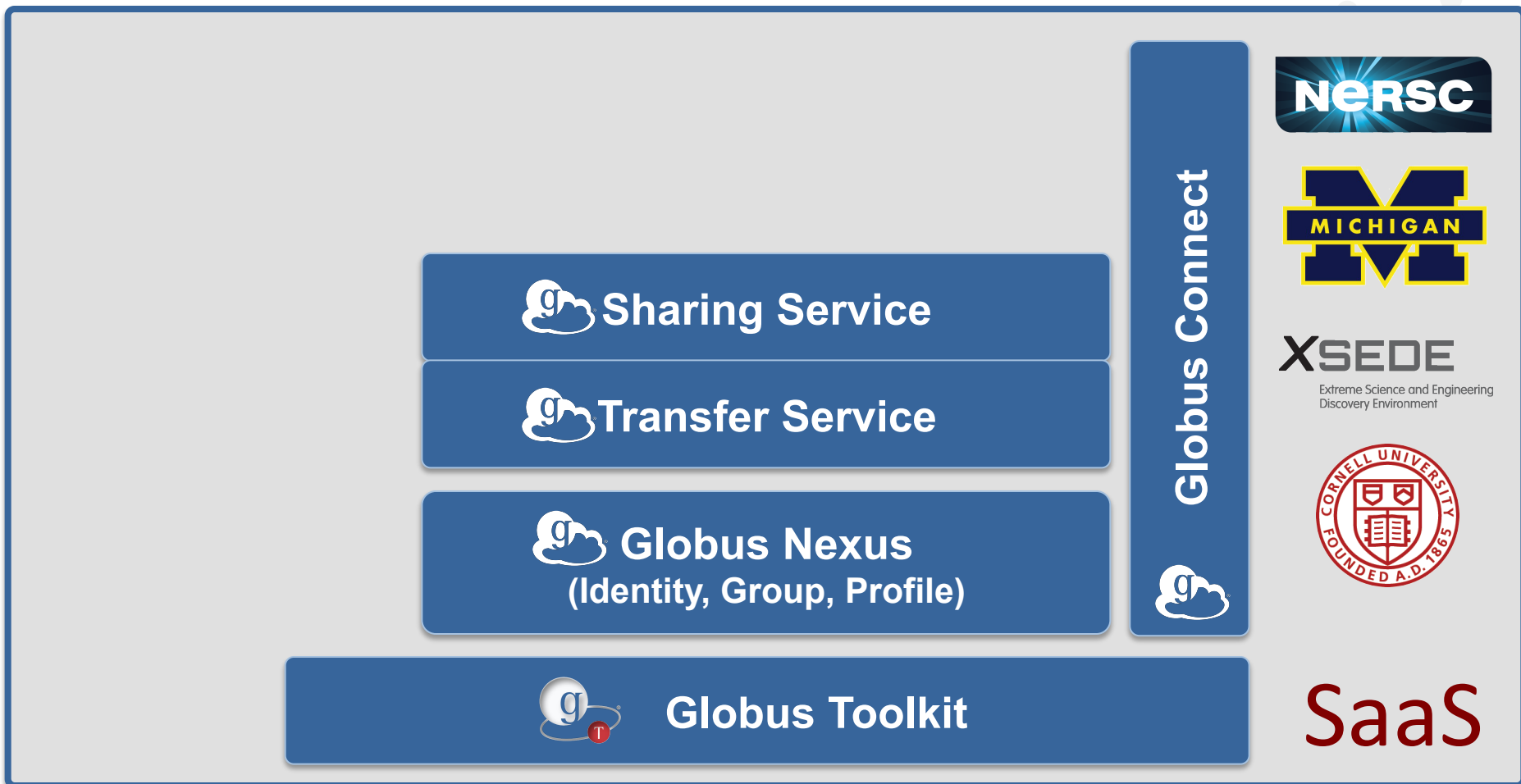
NEW YORK UNIVERSITY



Next: Share big data from existing storage



Globus Online is SaaS for science



We are now expanding to a platform



PaaS

Globus Online APIs



Sharing Service

Transfer Service

Globus Nexus
(Identity, Group, Profile)



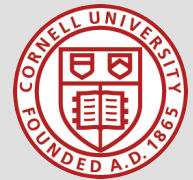
Globus Toolkit

Globus Connect



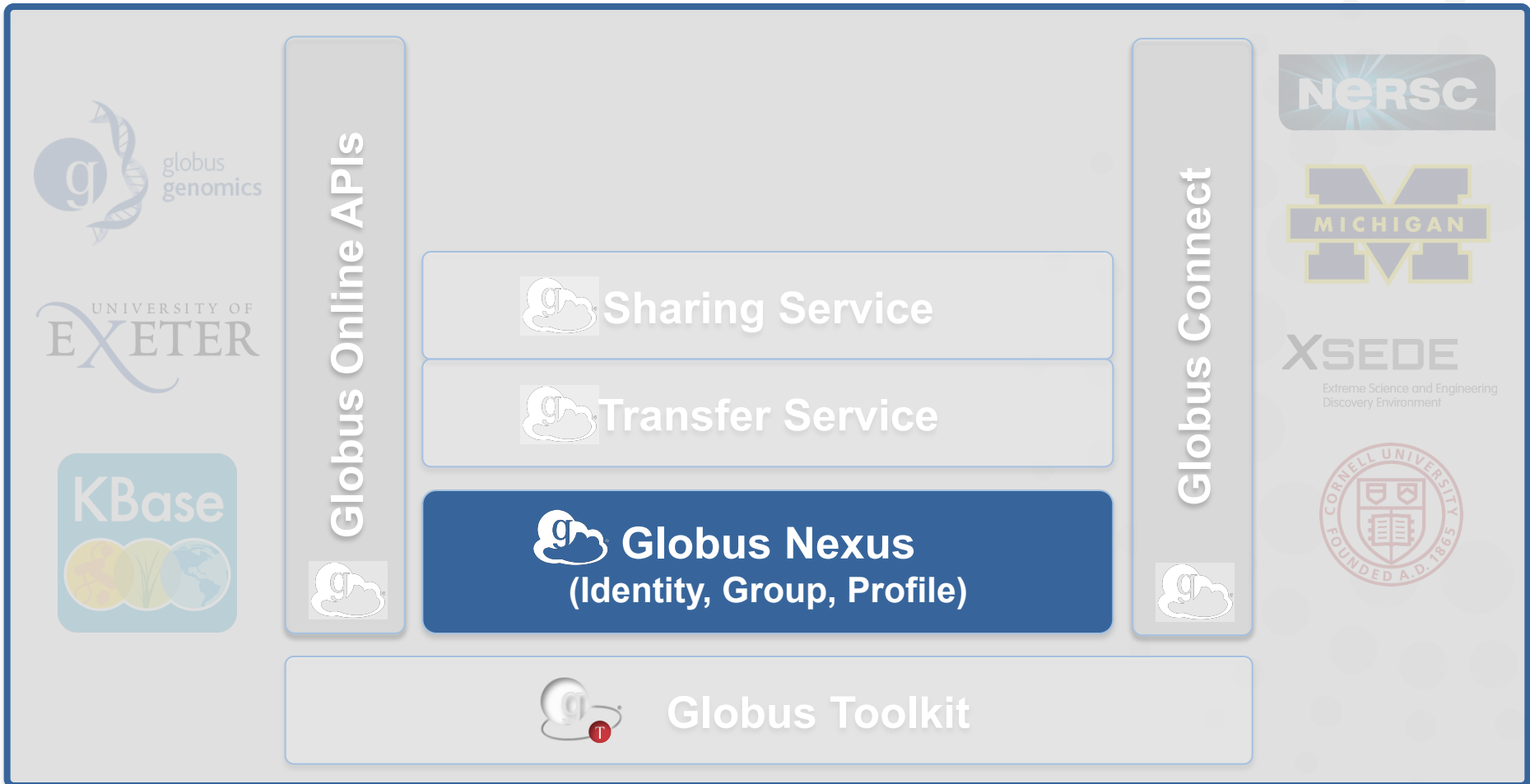
XSEDE

Extreme Science and Engineering
Discovery Environment



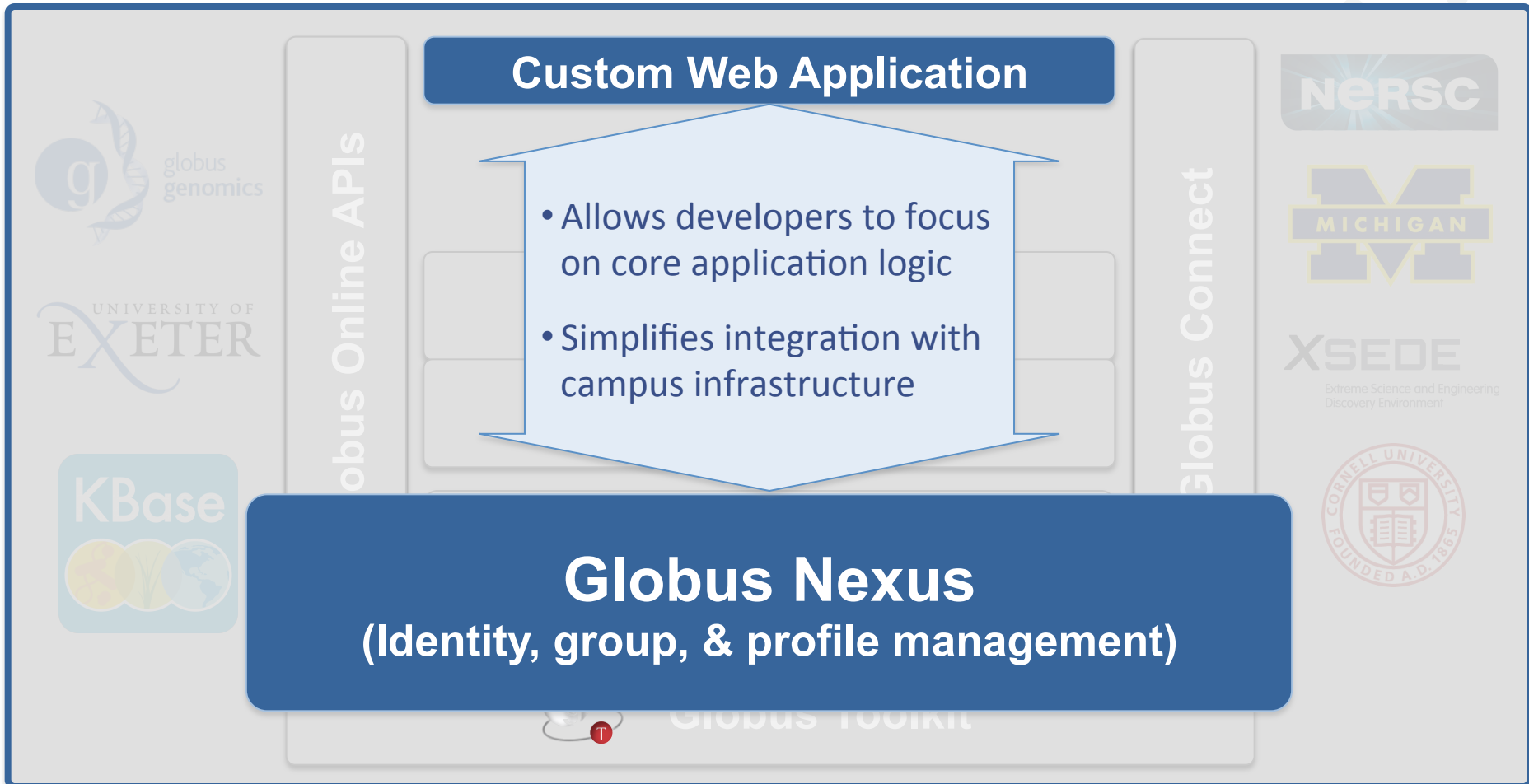
SaaS

Globus Online: Platform-as-a-Service





- Developers of collaborative science tools, applications, and cyberinfrastructures need to:
 - **Assign identities** to their users
 - Manage **user profiles**
 - Organize users into **groups** for authorization
- Providing **high-quality implementations** of such capabilities is challenging
 - Complexity of associated security protocols
 - Reliability, availability, scalability, security are all hard
- The result is **many identity ‘islands’** across science domains and projects—often poorly implemented



Nexus provides four key capabilities



Identity provisioning

- Create and manage Globus identities

Identity hub

- Link with other identities; use to authenticate to Nexus and other services

Group hub

- User-managed group creation, management; groups can be used for authorization

Profile management

- User-managed profile attributes and visibility; can be used in group admission

I Identity provisioning



- Globus Nexus can act as an identity provider (IDP) for a project
 - User management, email validation...
- DOE Systems Biology Knowledge Base (kBase) is an example of such a project. ~700 identities to date



DOE Systems Biology Knowledgebase



Log In Sign Up

Sign Up...

Already a member? [Sign In.](#)

KBase has partnered with Globus Online to provide a single account and group management system for applications across KBase.

Full Name

Email

Username

- **Link identities** from other federated IDP(s) with a Nexus identity
 - E.g., InCommon (SAML), Google (OpenID), XSEDE (OAuth MyProxy), IGTF-certified X.509 CA, SSH
- **Use linked identity** to authenticate to Nexus as the Nexus identity (e.g., use campus identity)
- **Leverage Nexus federated IDP** to 3rd-party services
 - Via Oauth or LDAP
 - E.g., to XSEDE, Jira, Zendesk, Drupal, Globus data management, Confluence
- Have Nexus **cache delegated credentials**
 - X.509, via CILogon, MyProxy

Identity management



Manage Identities

[Add External Identity](#) | [Add X.509 Certificate](#) | [Add SSH Public Key](#)

Alias	Type	Authentication Provider			
system-generated openid alias	openid	google.com	View Details	Update	Delete
CI	myproxy	grid.ci.uchicago.edu	View Details	Update	Delete
	myproxy	dev.esg.anl.gov:7512	View Details	Update	Delete
MBook	ssh2	SSH Public Key	View Details	Update	Delete
galaxy-climate	x509	/O=Grid/OU=DemoG.			Delete
Ian Laptop	x509	/C=US/O=Globus C...			Delete
InCommon / CILogon Sign In	oauth	cilogon.org			Delete

[Add External Identity](#) | [Add X.509](#)

Add External Identity

Alias

Account Provider

InCommon / CILogon

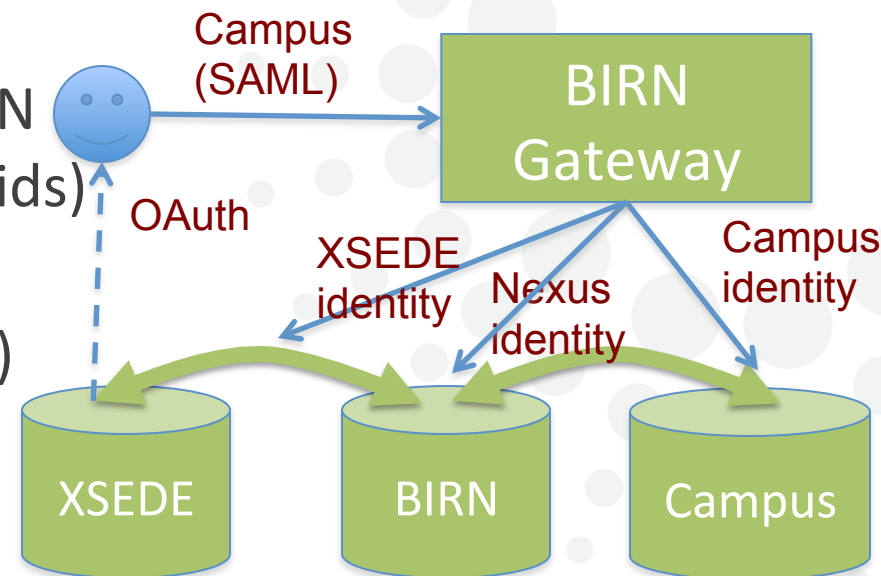
[Add Account](#) [Close](#)



- Dr. Smith creates a Nexus id, via BIRN project interface
- Dr. Smith links campus id and XSEDE id
- Dr. Smith can then:

Name: Dr. Smith
Email: smith@u.edu
Linked id: Campus
Linked id: XSEDE

- Authenticate to BIRN with campus id
- Query catalog (Nexus/BIRN id)
- Request data transfer from BIRN to campus (Nexus and campus ids)
- Request transfer from BIRN to XSEDE (Nexus and XSEDE ids)
- Repeat these tasks: use cached credentials



(BIRN=Biomedical Informatics Research Network)

Use linked identity



The screenshot shows the Globus Online 'Sign In' page. A modal window titled 'Select a different login' is open, displaying a list of login providers. The 'XSEDE' option is circled in red. The modal window contains the following list of providers:

GlobusOnline	InCommon / CILogon
Argonne LCF	LRZ
Argonne MCS & LCRC	NCSA
BIRN	NCSA Blue Waters
CLI Transition	NERSC
EGI	UChicago CI
ESG ANL	UChicago iBi
Exeter	UK NCS
Google	westGrid
	XSEDE

The screenshot shows the Globus Online 'Sign In' page. The text 'Use Your XSEDE login' is displayed, along with a 'Proceed' button. The page also includes a link for 'alternate log' and a message: 'You will now be redirected to XSEDE's authentication page.'

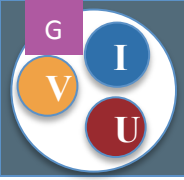
The screenshot shows the XSEDE User Portal Authorization Page. The page title is 'Welcome to the XSEDE User Portal Authorization Page'. It contains the following information:

Science Gateway Access
The XSEDE Science Gateway or Service below is requesting access to your XSEDE account. If you approve, please sign in with your XSEDE username and password.

Note: Only members of active XSEDE project allocations will be able to sign in on this page.

SCIENCE GATEWAY INFORMATION	SIGN IN
The XSEDE Science Gateway listed below is requesting access to your XSEDE account. If you approve, please sign in.	Username: <input type="text" value="demodoc"/>
Name: Globus Online URL: https://www.globusonline.org/	Password: <input type="password" value="*****"/>
	<input type="button" value="SIGN IN"/> <input type="button" value="CANCEL"/>

Please send any questions or comments about this site to help@xsede.org.



- User-managed group creation, management
- Flexible control over admission policies and visibility
- Groups can be used in authorization decisions

Example: kBase

- Every kBase user added to **kbase_users**
- Subgroups also created
- Groups used for access control

The screenshot displays the kBase user interface. At the top, a navigation bar includes 'Dashboard', 'Groups', 'ian's account', and 'Log Out'. Below this, the 'My Groups' section shows a list of groups, with 'kbase_users' highlighted. To the right, there are three main menu items: 'Browse Groups' (Browse and join groups), 'My Account' (View and change your account settings, including contact information and security credentials), and 'KBase Site' (Return to the main KBase website). Below the 'My Groups' section, a detailed view of the 'kbase_users' group is shown, featuring buttons for 'Home', 'Members', and 'SubGroups'. Underneath these buttons, a list of subgroups is displayed: 'kbase_staff', 'kbase_microbes', and 'ENIGMA'.

Group membership interface



Group Membership

Group Name	Status	
Cornell demo	Pending Admin Approval	✕
Data Summit Slides	Administrator	✕
Globus Team	Active Member	✕
Ian's secret stuff group	Administrator	✕
Ag-GRID	Active Member	✕

Kyle's Group

Home Members SubGroups Settings

1 active 2 invited 1 rejected

[+ Invite people to this group](#)

Invitation Sent - Awaiting Response

name	email	username	request sent
-	chard@uchicago.edu	-	today
josh	josh.bryan@gmail.com	jbryan	today

Active Members

name	email	username	member since
Kyle Chard	kyle@ci.uchicago.edu	kyle	today

Kyle's Group

Home Members SubGroups Settings

Policies Membership Fields Advanced

edit

- Group is visible to** logged in users (visible to all Globus Online members).
- Member names are visible to** this group's administrators only.
- Subgroups for this group may be created by** admins only
- Users may request membership if they are** invited to join this group.
- Membership requests are approved** automatically for invitations sent by admins & managers

Branded sites



BLUE WATERS
SUSTAINED PETASCALE COMPUTING

Reliable, high-performance, secure file transfer by Globus Online.

Blue Waters has partnered with the Globus Online file transfer service.

You may access this service by entering your Blue Waters username and password.

Sign In

Use Your NCSA Blue Waters login alternate login

You will now be redirected to NCSA Blue Waters authentication page.

Proceed

XSEDE

osg connect

Support Resources OSG Connect Sign In / Register

Efficiently connect your science to cycles and data

osg connect

Sign In Sign Up with Globus Online

Using your InCommon / CILogon login alternate login

You will now be redirected to InCommon / CILogon's authentication page.

Proceed

Open Science Grid

Research Computing Center

THE UNIVERSITY OF CHICAGO

Log In Sign Up

Sign In

Use Your InCommon / CILogon login alternate login

You will now be redirected to InCommon / CILogon's authentication page.

Proceed

University of Chicago

KBase
DOE Systems Biology Knowledgebase

Log In Sign Up

Sign In Sign Up with Globus Online

Use Your GlobusOnline login alternate login

Username

Password

Sign In Forgot Password?

DOE kBase

INDIANA UNIVERSITY Cyberinfrastructure Gateway

Sign In

Sign In

Use Your InCommon / CILogon login alternate login

You will now be redirected to InCommon / CILogon's authentication page.

Proceed

Indiana University

UNIVERSITY OF EXETER

Log In Sign Up

Sign In

Use Your Exeter login alternate login

You will now be redirected to Exeter's authentication page.

Proceed

University of Exeter

globus online

Manage Data Groups News & Events About Support Log In Sign Up

Sign In

Use Your EG login alternate login

You will now be redirected to EG's authentication page.

Proceed

Globus Online

NERSC

Sign Up Sign In

Sign In

Use Your NERSC login alternate login

Username

Password

Sign In

NERSC

BIRN ACCESS Group Management

Manage Data Groups Log In Sign Up

ACCESS is a user and group management system used by the BIRN community to allow editing access to the BIRN WGL.

Sign in with your Globus Online ID.

Are you working with a BIRN Working Group or user community and need editing access to the BIRN WGL? Sign up here.

Username

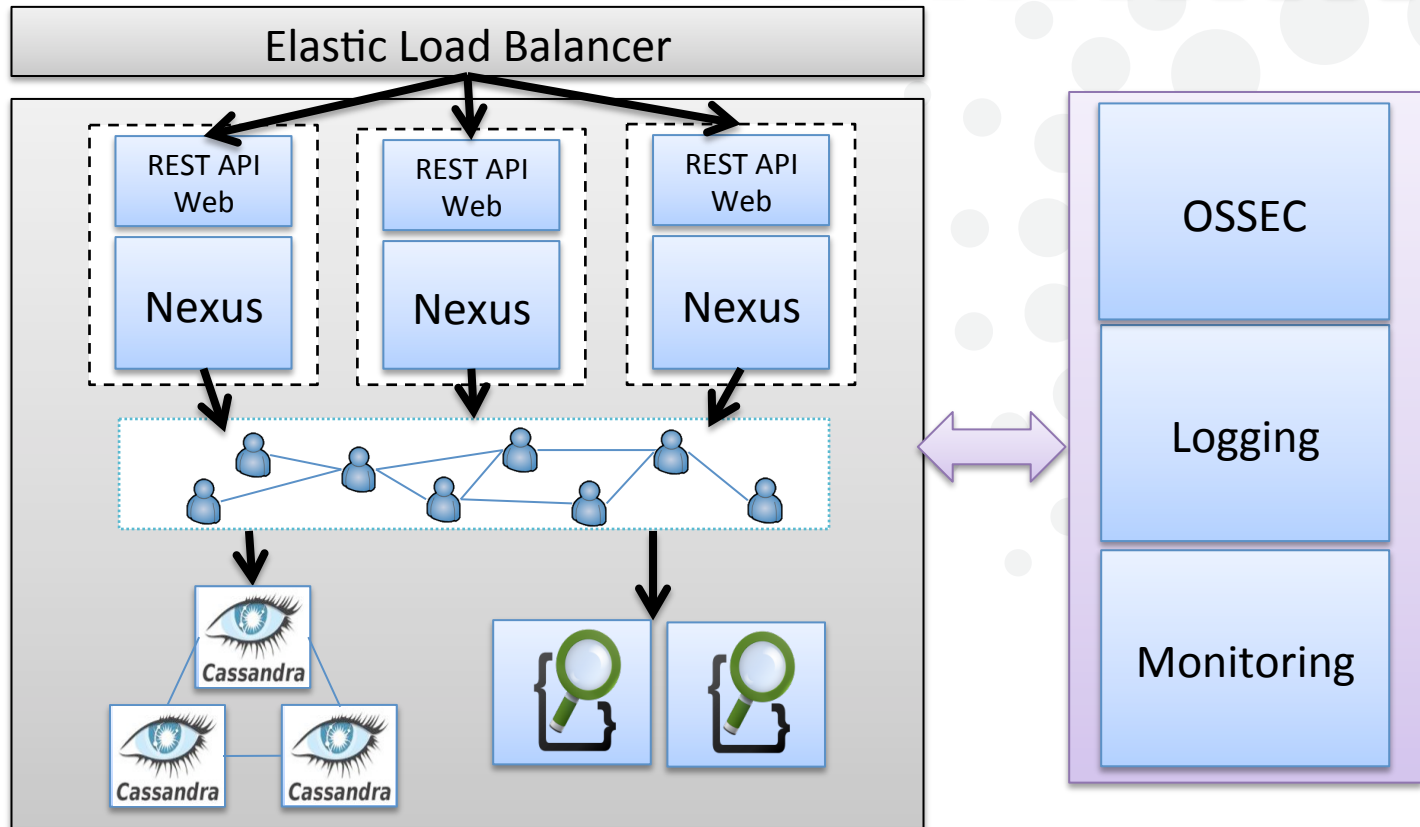
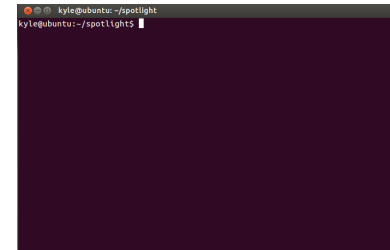
Password

Sign In

Forgot Password? Sign Up Here

NIH BIRN

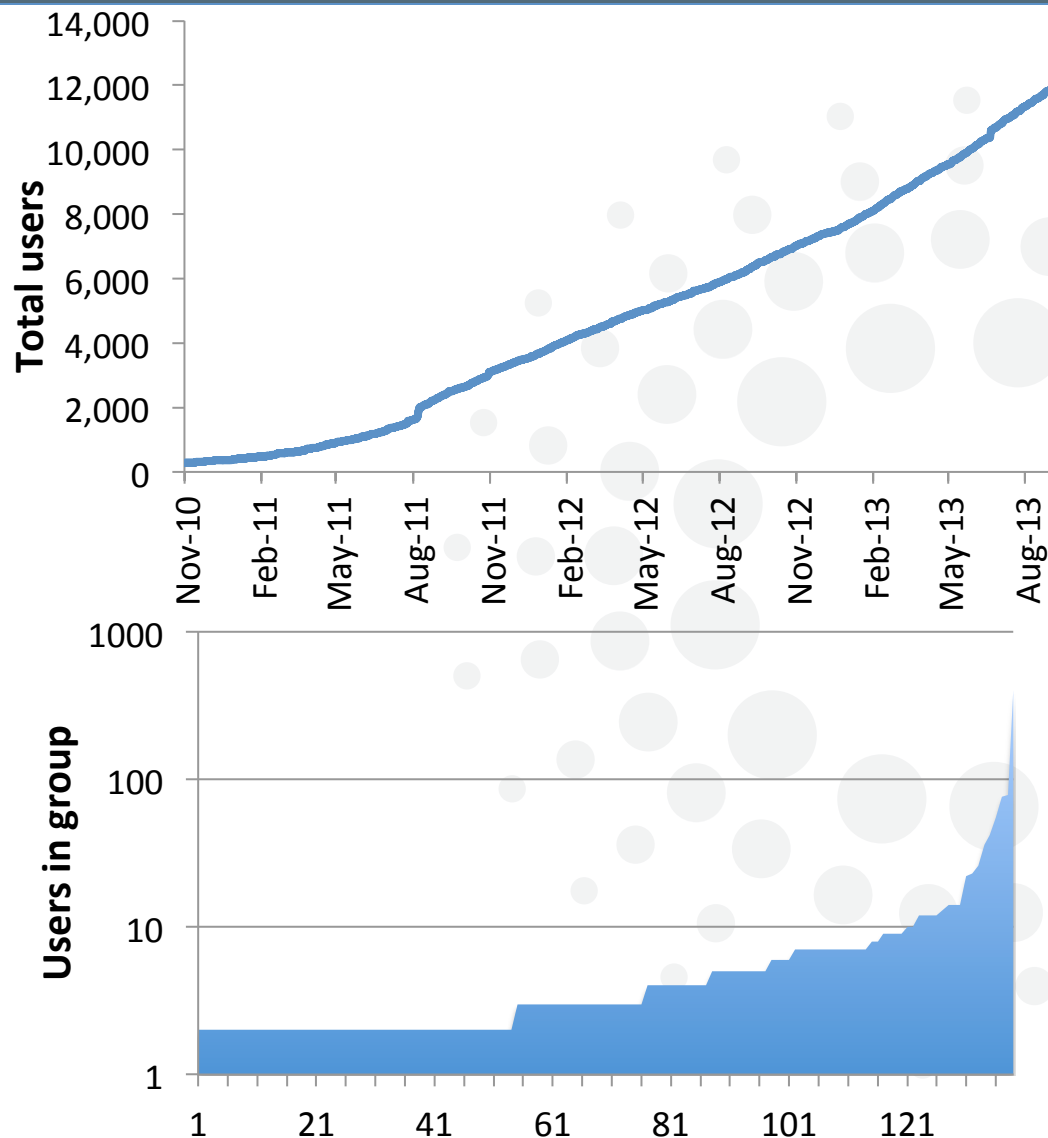
Implementation and deployment



Globus Nexus usage as of 9/13



- **>12,000** users and **4977** linked identities
- **557** groups totaling:
 - **1638** active members
 - **229** pending or invited members
 - **162** rejected or suspended members
- Largest group (kbase) has **402** members





- Proposal: Replace current ad-hoc systems with Globus Nexus identity and group service
 - Reduce complexity, reduce cost, increase capability
- Careful process of documentation and review
 - “Architecture and development requirements: User and identity management”
 - “User management proposal: Affected use cases”
 - “User management proposal: Motivating stories”
 - “Proposal: Refactoring XSEDE identity and group capabilities”
- Hope to reach closure by end of 2013

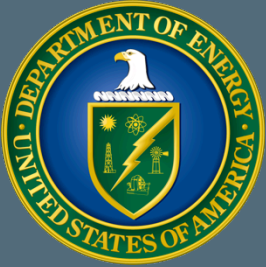


Accelerate discovery and innovation worldwide by providing **research IT as a service**

Leverage software-as-a-service to

- provide millions of researchers with unprecedented access to powerful tools;
- enable a massive shortening of cycle times in time-consuming research processes; and
- reduce research IT costs dramatically via economies of scale

Thanks to ...



U.S. DEPARTMENT OF
ENERGY



THE UNIVERSITY OF
CHICAGO



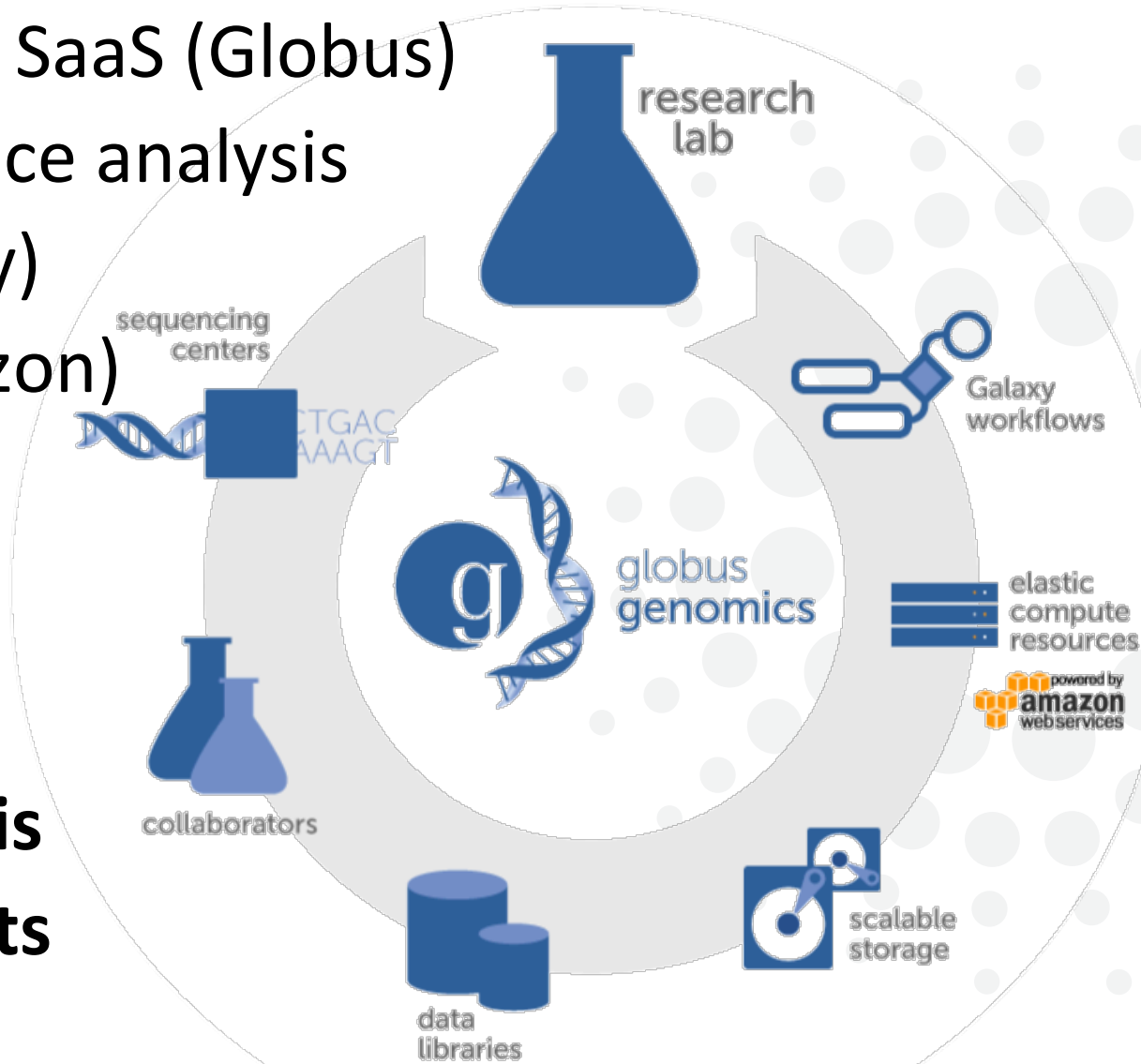
Thank you! Questions?

foster@anl.gov

foster@uchicago.edu

www.globusonline.org

Data management SaaS (Globus)
+ Next-gen sequence analysis
pipelines (Galaxy)
+ Cloud IaaS (Amazon)
**= Flexible,
scalable,
easy-to-use
genomics analysis
... for all biologists**



Globus Online: Platform-as-a-Service




Globus Online APIs



 Catalog Services

 Sharing Service

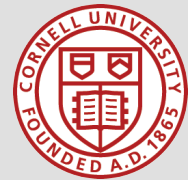
 Transfer Service

 Globus Nexus
(Identity, Group, Profile)



Globus Toolkit

Globus Connect



Globus Online: Platform-as-a-Service

