

# Globus 2016: From Science Software to Science Platform

GlobusWorld 2016

Keynote



globus



# Research data management simplified.

**155,466,716,596** MB  
TRANSFERRED

## Researchers

Focus on your research, not IT problems. We make it easy to move, manage, and share big data.

[LEARN MORE](#) >

[GET STARTED](#) >



## Resource Providers

Globus gives you more control over your data infrastructure, while providing excellent ease-of-use for your researchers.

[LEARN MORE](#) >

[GLOBUS PROVIDER PLANS](#) >



## Our Users

Researchers and resource providers are our greatest inspiration and we love it when they say nice things about Globus.

[USER QUOTES](#) >

[CASE STUDIES](#) >



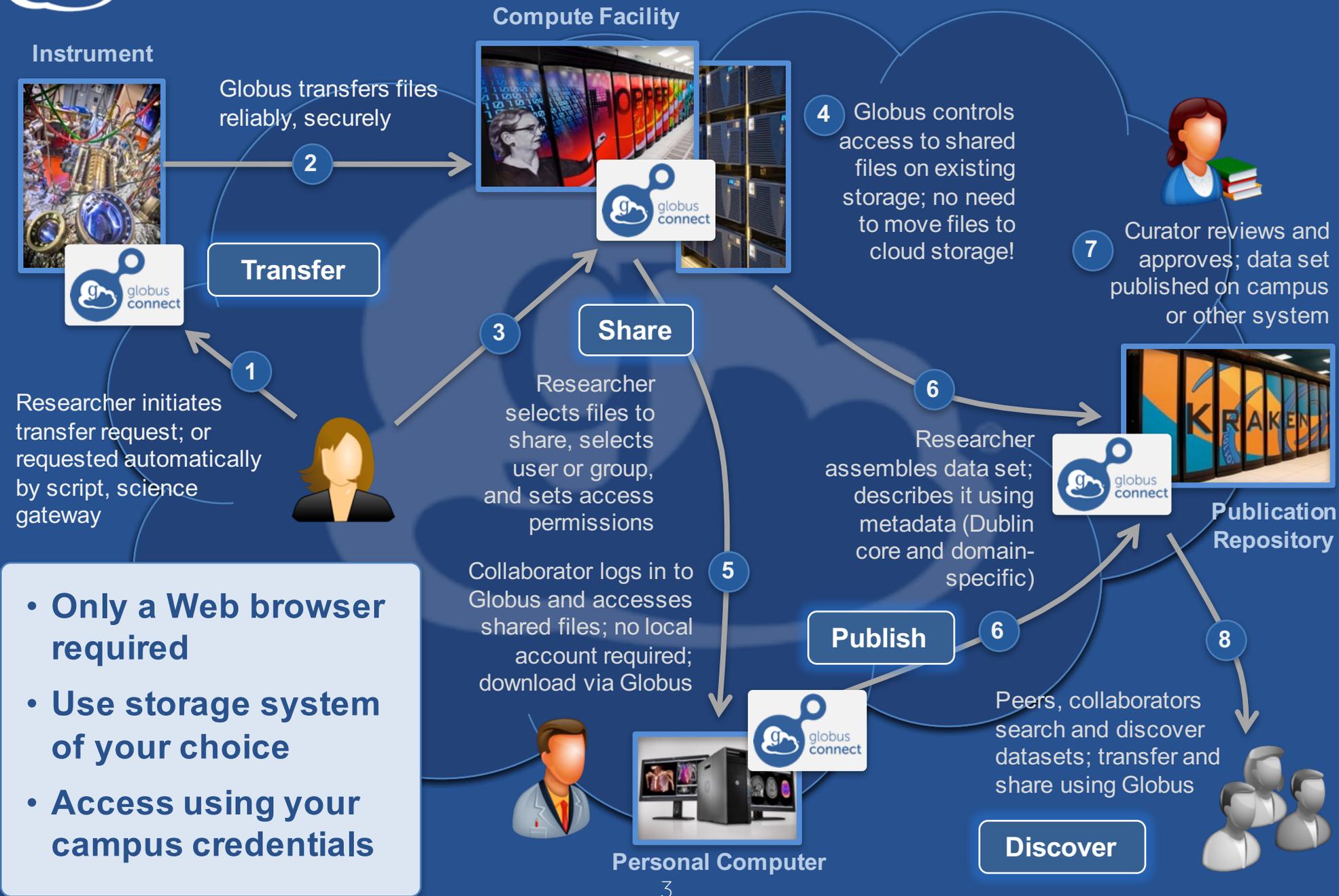
Fast, Reliable, Secure File Transfer



UPCOMING EVENTS



# Globus SaaS: Research data lifecycle





# Globus by the numbers

**4**

major services

**155 PB**  
transferred

**25 billion**  
files processed

**38,000**  
registered users

**13**

national labs  
use Globus

**2,500**  
active endpoints

**450+**  
active daily users

**99.9%**  
uptime

**35+**

institutional  
subscribers

**1 PB**  
largest single  
transfer to date

**3 months**  
longest  
continuously  
managed transfer

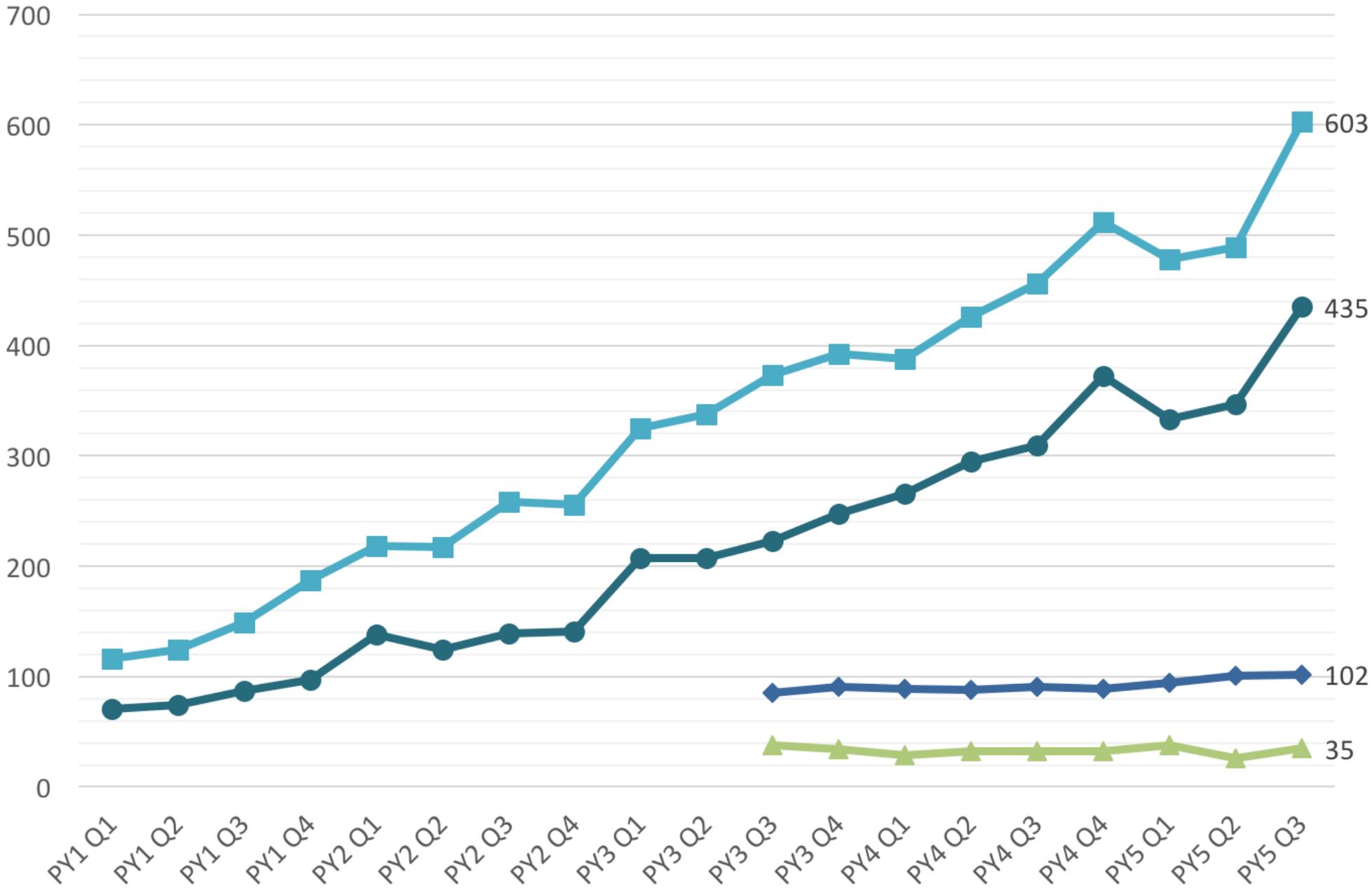
**130**  
federated  
campus identities



# Who uses it and for what?

- **Researchers**
  - Data sharing, transfer, archiving, upload, download, ...
- **Research data portal operators**
  - Data upload, download
- **Operators of experimental facilities**
  - Data delivery, data replication
- **Campus computing administrators**
  - Research acceleration, resource management

Users (to/from XSEDE)      Users (to/from XSEDE via GCP)  
Campuses (to/from XSEDE)      Campuses (all)

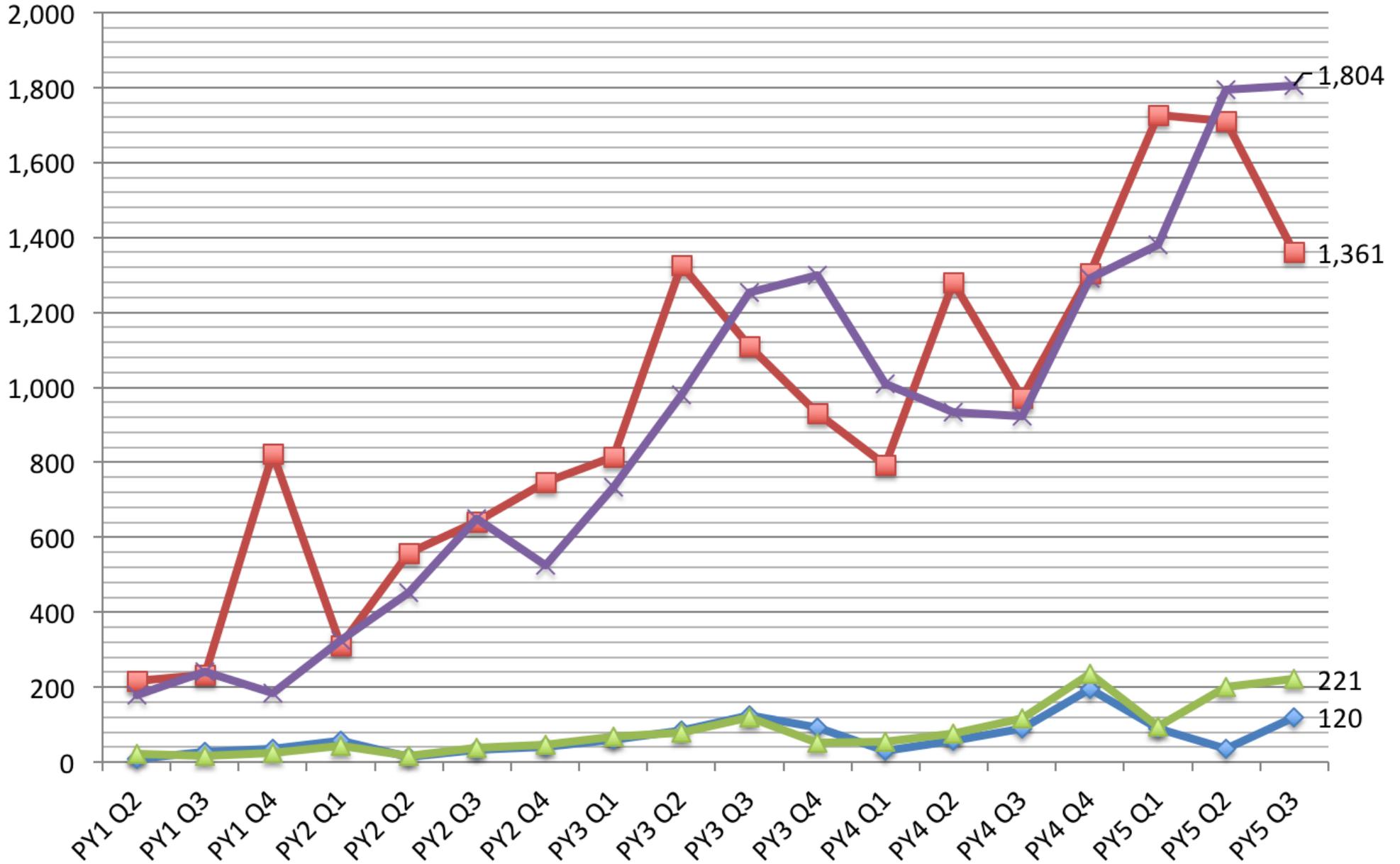


Files into XSEDE endpoints (millions)

Files from XSEDE (milions)

TB into XSEDE endpoints

TB from XSEDE endpoints



# Sanger Imputation Service

This is a free genotype **imputation** and **phasing** service provided by the [Wellcome Trust Sanger Institute](#). You can upload GWAS data in VCF or 23andMe format and receive imputed and phased genomes back. Click [here](#) to learn more and [follow us on Twitter](#).

## Before you start

Be sure to [read through the instructions](#).

You will need to set up a free account with [Globus](#) and have [Globus Connect](#) running at your institute or on your computer to transfer files to and from the service.

## Ready to start?

If you are ready to upload your data, please fill in the details below to **register an imputation and/or phasing job**. If you need more information, see the [about](#) page.

What is this [?](#)

## News

[@sangerimpute](#)

**11/05/2016**

Thanks to [EAGLE](#), we can now return **phased data**. The HRC panel has been updated to r1.1 to fix a [known issue](#). See [ChangeLog](#) for more details.

**15/02/2016**

Globus API changed, please see [updated instructions](#).

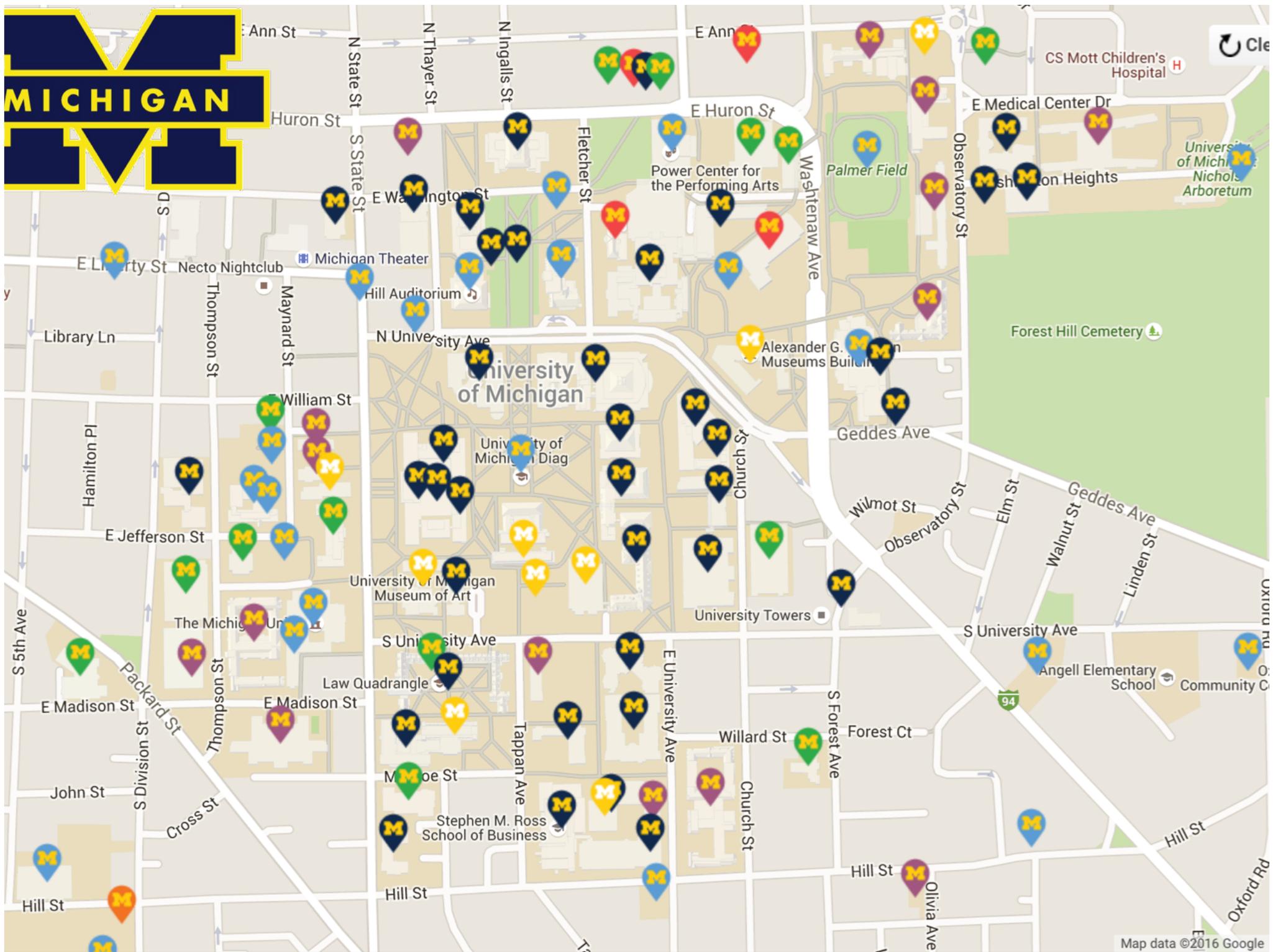
**17/12/2015**

New status page and reworked internals. See [ChangeLog](#).

**09/11/2015**

Pipeline updated to add some features requested by users. See [ChangeLog](#).

[See older news...](#)





# The full value of Globus

Is realized when you have it on  
**all** storage systems across  
campus, not just in the  
research computing center

**Don't stop at the Science DMZ**

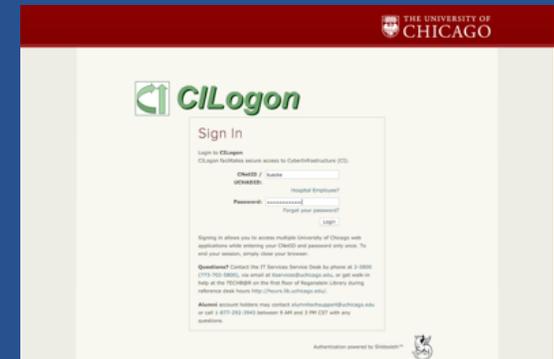


# New features and enhancements



# No Globus usernames required!

- **Globus username & password no longer needed!**
  - Old Globus usernames moved to optional “Globus ID” identity provider
- **Any identity recognized by Globus is now sufficient to access Globus**
- **Globus Account is a primary identity plus a set of linked identities**
  - Verified email address can be a linked identity
- **Sharing & groups can use any identity**





- **Globus requires consent to use your identity**

The screenshot shows a consent dialog for the Globus Web App. At the top left is the Globus logo and the text 'globus'. At the top right is an 'Account' dropdown menu. The main content area is white with a blue border. It starts with the heading 'Globus Web App would like to:'. Below this are three items, each with a green checkmark icon and an information icon: 'View your identities on Globus Auth', 'Manage your Globus Groups', and 'Transfer files using Globus Transfer'. This is followed by the heading 'To work, the above will need to:' and two more items with green checkmarks and information icons: 'Manage your Globus Groups' and 'View your identities on Globus Auth'. Below these items is a paragraph of text: 'By clicking "Allow", you allow **Globus Web App**, in accordance with its [terms of service](#) and [privacy policy](#), to use the above listed information and services. You can rescind this and other [consents](#) at any time.' At the bottom are two buttons: a blue 'Allow' button and a grey 'Deny' button.

- **Standard practice, e.g. when logging into another service with Facebook or Google**



# New Globus account management

- Revamped account management system
- Primary identity is created when you first access Globus
- Can link multiple identities

The screenshot shows the Globus account management page. At the top, there is a navigation bar with the Globus logo, 'globus', and links for 'Manage Data', 'Publish', 'Groups', 'Support', and 'Account'. The main content area is titled 'Account' and includes a 'Manage Your Consents' link and a 'Logout' button. Below this is the 'Identities' section, which lists two identities: 'Steve Tuecke via Globus ID' and 'Steven Tuecke via University of Chicago'. Each identity entry shows the name, username, email, and organization. The first identity is marked as the 'primary identity'. Below the identities section is the 'Globus Plus' section, which provides information about enhanced features and lists the user's current groups: 'Globus Team Plus Sponsor' and 'Tutorial Users'. There is also a dropdown menu to 'Select another organization' and a 'Didn't find your organization?' section with a note about contacting support.



- Share with any identity
- Invite collaborators via email (akin to Dropbox)
- If they have not used Globus, will need to provide some info, e.g. name and organization
- Then Globus account provisioned automatically
- Same is true for Globus group membership



# XSEDE has adopted Globus identity management

The screenshot shows the XSEDE website header with the logo and tagline 'Extreme Science and Engineering Discovery Environment'. The Globus logo and 'Account' dropdown are in the top right. The main content area displays a consent request from 'Jetstream Web App' to 'Access all Jetstream resources'. Below this is a paragraph explaining that clicking 'Allow' grants access to resources without terms of service or privacy policy provided, and that users can rescind consent at any time. At the bottom are two buttons: 'Allow' (blue) and 'Deny' (grey).

**XSEDE**  
Extreme Science and Engineering  
Discovery Environment

globus Account ▾

Jetstream Web App would like to:

- ✓ Access all Jetstream resources ⓘ

By clicking "Allow", you allow **Jetstream Web App** (no terms of service or privacy policy provided) to use the above listed information and services. You can rescind this and other [consents](#) at any time.

Allow Deny



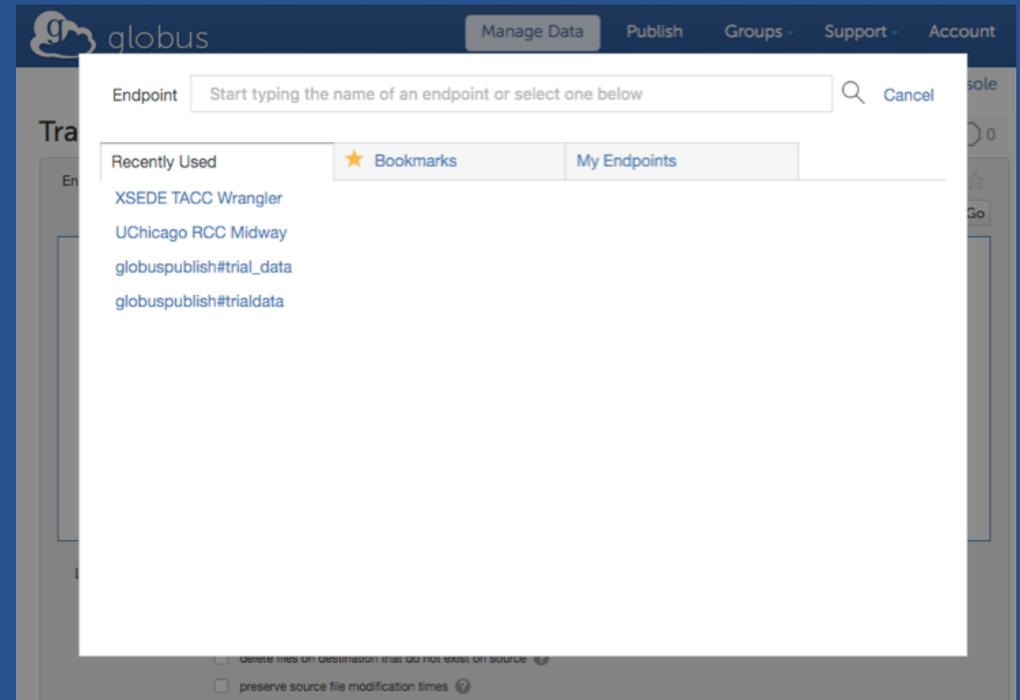
# Adding your campus identity provider to Globus

- **InCommon identity providers that release research & scholarship attributes to CILogon *(free)***
- **OpenID Connect identity provider supported by Globus Auth *(subscription)***



# New endpoint search

- We have >40,000 endpoints so need a way to discover
- Added multiple attributes to endpoint definition so that users can validate that they have the right one
- Endpoint searches across all attributes





# Endpoint search



Manage Data

Publish

Groups ▾

Support ▾

Account

Endpoint

xsede|



Cancel

**XSEDE Testing Endpoint 7 at PSC**

owner: xsede@globusid.org

no description provided

**XSEDE Trestles**

owner: xsede@globusid.org

Trestles is a dedicated XSEDE cluster at SDSC consisting of 324 compute nodes. This endpoint ...

**XSEDE Comet**

owner: xsede@globusid.org

Comet is a dedicated XSEDE cluster. This endpoint can be used to access data stored on the C...

**XSEDE Nautilus**

owner: xsede@globusid.org

This endpoint can be used to access data stored on the Nautilus file system.

**XSEDE Mason**

owner: xsede@globusid.org

Mason at Indiana University is a large memory computer cluster configured to support data-inte...

**XSEDE Lonestar**

owner: xsede@globusid.org

This endpoint can be used to access data stored on the Lonestar file system.

**XSEDE Keeneland**

owner: xsede@globusid.org

Keeneland is a hybrid CPU/GPGPU system for use with codes that can take advantage of GPU ...

delete files on destination that do not exist on source

preserve source file modification times



# Bookmarks

- Add frequently used endpoints to your bookmarks
- Globus also maintains list of recently used endpoints for quick access

The screenshot shows the 'Manage Bookmarks' page in the Globus interface. The page has a dark blue header with the Globus logo and navigation links: 'Manage Data', 'Publish', 'Groups', 'Support', and 'Account'. Below the header, there are tabs for 'Transfer Files', 'Activity', 'Endpoints', 'Bookmarks', and 'Console'. The main content area is titled 'Manage Bookmarks' and displays a list of four bookmarks. Each bookmark entry includes its Name, Endpoint, and Path, along with icons for editing, deleting, and refreshing.

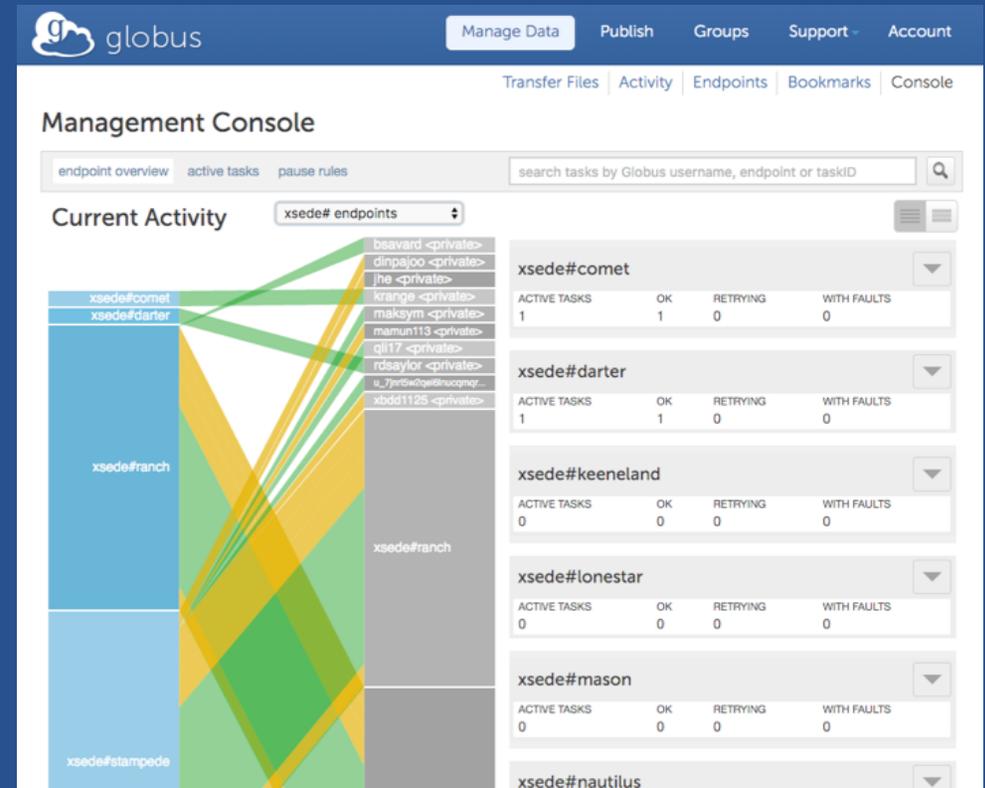
Name	Endpoint	Path	Actions
Gordon Project X	Tuecke Share on XSEDE Gordon	/ProjectX/	[Edit] [Delete] [Refresh]
Macbook tmp	Macbook	/~/tmp/	[Edit] [Delete] [Refresh]
midway/~myproject	UChicago RCC Midway	/~/myproject/	[Edit] [Delete] [Refresh]
UC Midway	UChicago RCC Midway	/~/	[Edit] [Delete] [Refresh]

© 2010-2016 The University of Chicago legal



# Management Console

- Real-time view into endpoint activity
- Useful for troubleshooting errors
- Ability to pause, resume, and cancel transfers facilitates endpoint management





# Data Publication and Discovery

- Announced at GW15, generally available for ~9 months
- Example: Compute Canada uses for national repository

g globus Log In Sign Up

To submit a dataset or view datasets that have restricted access, please log in.

Search

## Globus Data Publication (Trial)

Welcome to the Globus data publication "trial sandbox". The [User Guide](#) explains how to publish data. Please [contact us](#) with questions or feedback. When you are ready to create a publication community for your institution, complete [this form](#).

**Communities**  
Choose a community to browse its collections.

- Compute Canada
- Globus Demo Community
- Kansas City Federal Reserve
- Open Trial  
Open to all Globus users.
- Pilot Communities  
Groups performing trials

**Discover**

**Author**

- Pruyne, Jim 1
- Smith, John 2
- User, Jane 2
- smith, asd 1

**Subject**

- multiple text files 1
- Test 1
- test 1
- Transient 1
- Trial 1
- unverified dataset 1

**Issue Date**

- 2200 - 2222 1
- 2000 - 2099 5

© 2010-2015 Computation Institute, University of Chicago, Argonne National Laboratory [legal](#)

g globus Manage Data - Publish Groups Support - janesmith -

Browse & Discover | Data Publication Dashboard | Communities & Collections

License Describe Describe Describe Assemble Verify Complete

## Submit: Describe Dataset

All fields with an asterisk (\*) are required

A name or title by which the data is known

**Title \*** Jane's First Dataset to Curated Trial Collection

The main researchers involved in producing the data

**Authors \*** Smith Jane + Add More

The year when the data was or will be made publicly available

**Publication Year \*** Month: March Day: 15 Year: 2015

Select the language of the main content of the item. If the language does not appear in the list below, please select 'Other'. If the content does not really have a language (for example, if it is a dataset or an image) please select 'N/A':

**Language** English (United States)

The name of the entity that holds, archives, publishes prints, distributes, releases, issues, or produces the data

**Publisher \*** Trial Sandbox Holdings

< Back Save & Exit Continue >

© 2010-2015 Computation Institute, University of Chicago, Argonne National Laboratory [legal](#)

# docs.globus.org

- The reference for all things Globus
- Consolidated all documentation from support site and other sources
- Guides for administrators
- FAQs for all users
- API docs for developers



The screenshot shows the top navigation bar of the docs.globus.org website. It includes the Globus logo and the text "globus docs" on the left, and a search bar and navigation links for "APIs", "How To", "Guides", "Toolkit", and "Support" on the right. Below the navigation bar is a large banner with a grid of colorful icons representing various research fields. The text on the banner reads: "Harness the power of the Globus research data management cloud." Below the banner are three circular icons with labels: "Transfer API" (with a code symbol), "Resource Providers" (with a server rack icon), and "Toolkit" (with a 'T' icon).



# New mailing lists

- **User, admin, developer**
- **Hosted on Google Groups**
- **Public forum**
- **[www.globus.org/mailing-lists](http://www.globus.org/mailing-lists)**

The screenshot shows the 'Globus Email Lists' page. At the top, there is a navigation bar with the Globus logo, the word 'globus', and links for 'Products', 'News', 'About', 'Support', 'Log In', and 'Sign Up'. Below the navigation bar, the page title 'Globus Email Lists' is displayed. The main content area contains an introductory paragraph: 'If you use Globus, then participating in one or more of the public email lists is an excellent way to keep in touch with your peers in the Globus Community. There are three lists you can join – all are public forums:'. Below this, there are three columns, each representing a mailing list. The first column is for the 'User-discuss List', described as 'For general questions about using the Globus service.', with 'SUBSCRIBE' and 'VIEW DISCUSSIONS' links. The second column is for the 'Admin-discuss List', described as 'For questions about managing your Globus deployment, e.g. installing software for a Globus endpoint, configuring your firewall, and integrating your institution's identity system.', with 'SUBSCRIBE' and 'VIEW DISCUSSIONS' links. The third column is for the 'Developer-discuss List', described as 'For questions about integrating Globus services into your science gateway, portal, or other web application via the Globus REST APIs.', with 'SUBSCRIBE' and 'VIEW DISCUSSIONS' links.



# Broadening our reach

# Spectra Logic Company Overview

- Celebrating Over 3 Decades of Innovative Technology

- Privately held corporation founded in 1979
- Debt free from since conception
- 100% employee owned

- Storage Industry Leadership

- Founding member: Active Archive Alliance
- Founding member: Tape Council
- Standard member: SNIA

- Worldwide Presence

- HQ's located in Boulder, CO – Bracknell, UK – Melbourne, AUS
- Over 450 employees in 11 countries
- 20,000 libraries installed in over 50 countries

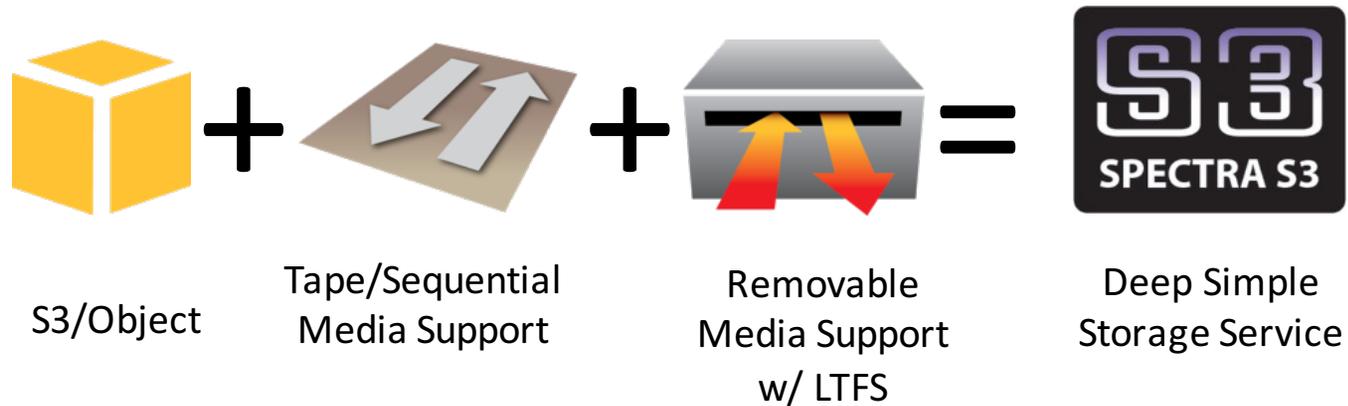


# Black Pearl



Store Massive Data Forever

# Applying Object Storage to the Archive

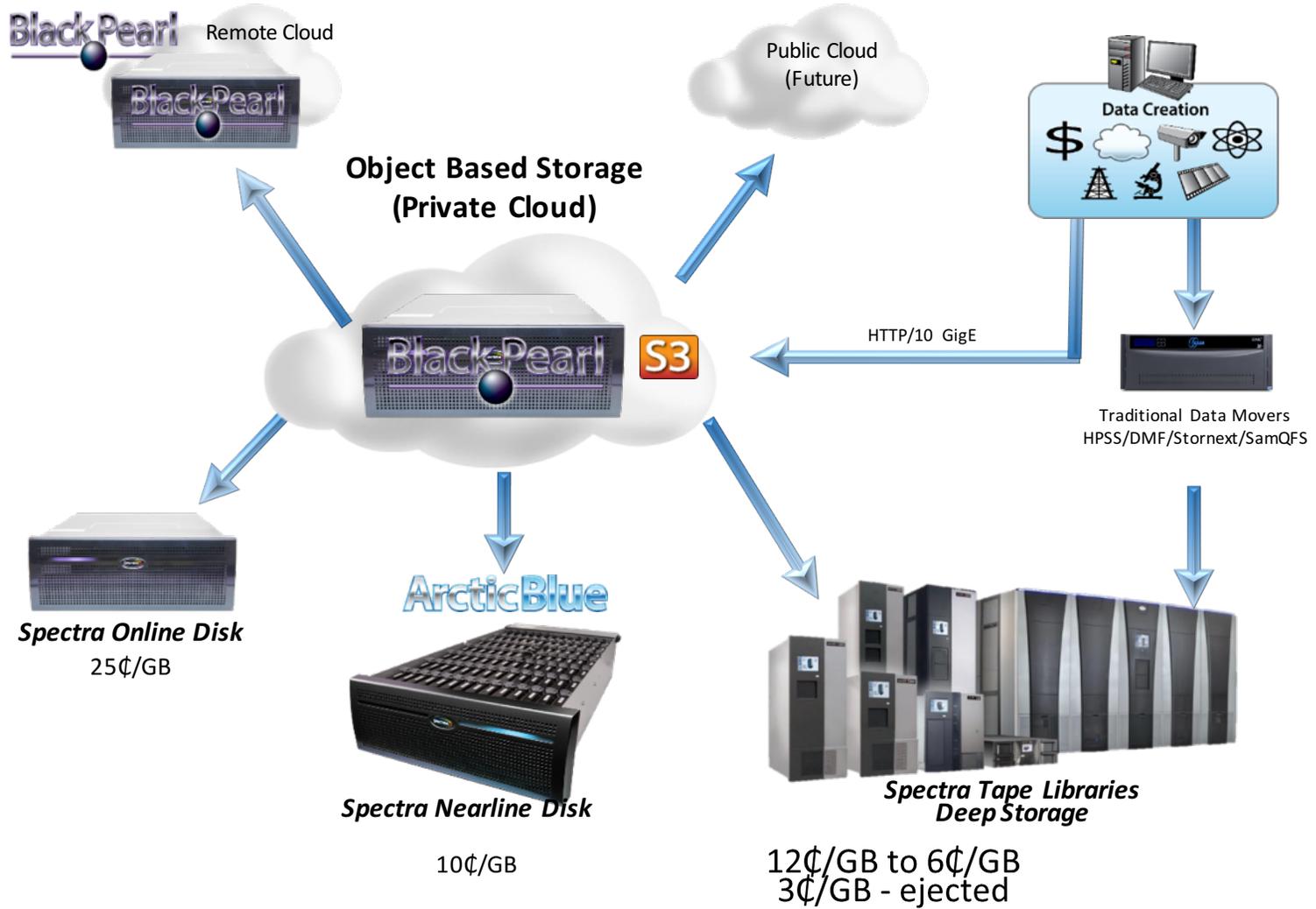


Start with Object, S3 and Rest

Use Archive media as the target

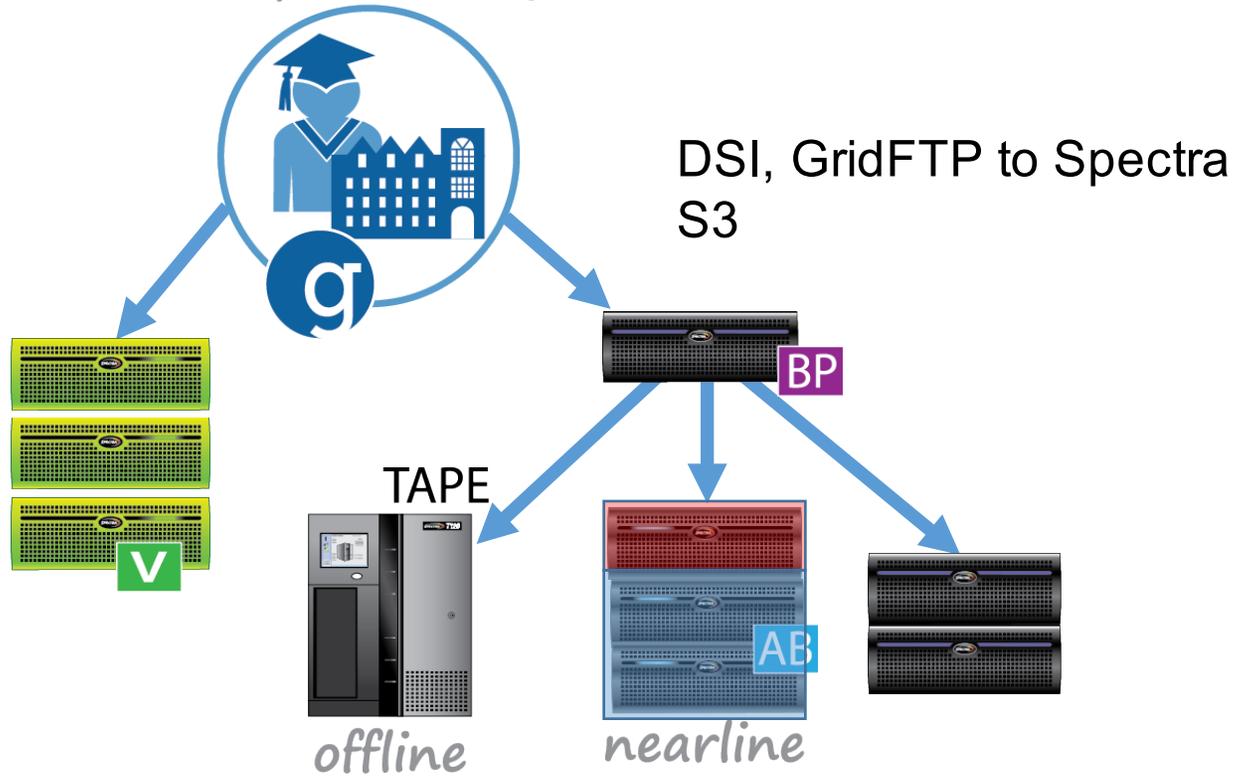
Spectra S3 facilitates an Objected based Archive

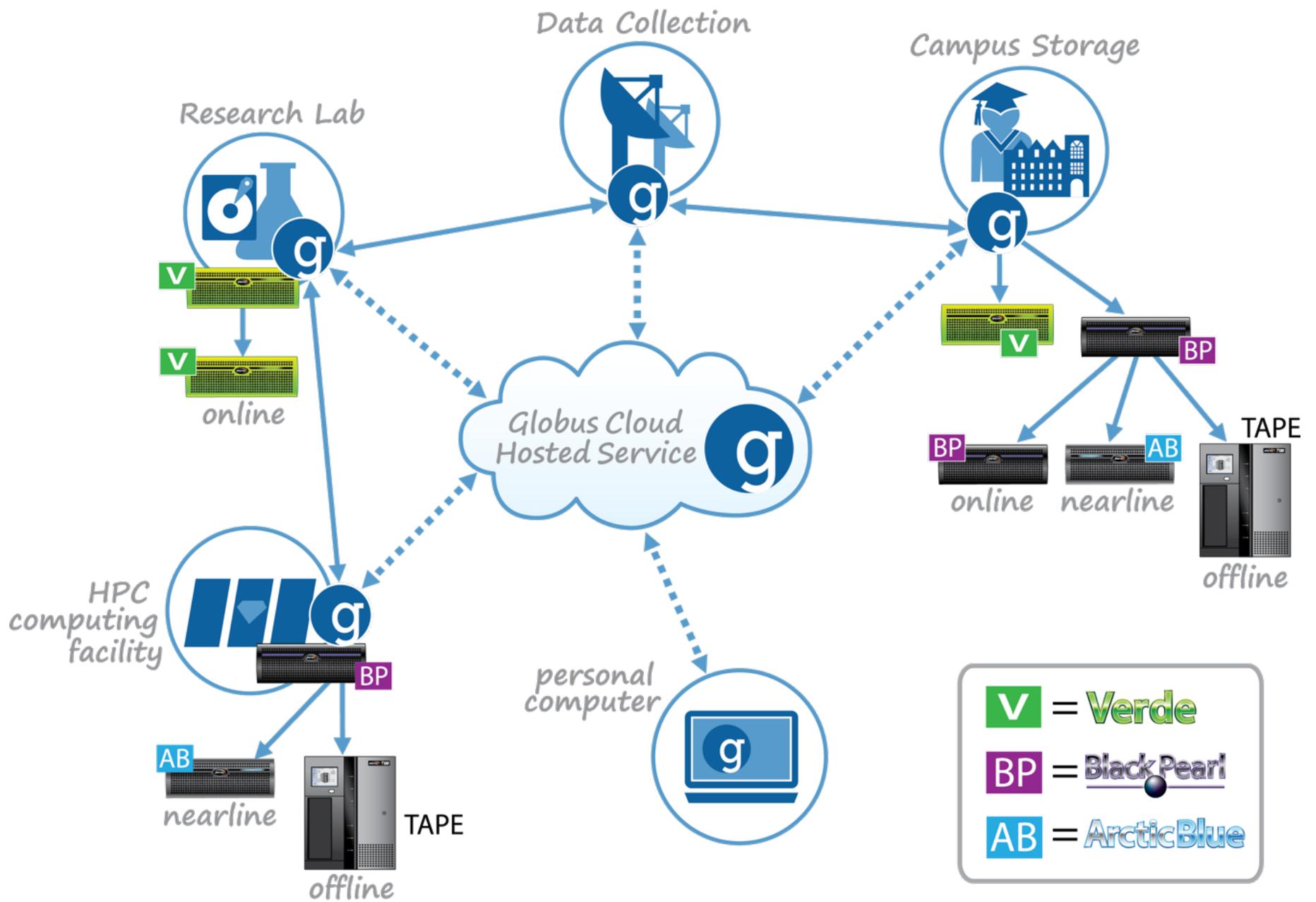
# BlackPearl Ecosystem





Campus Storage







# Other storage integrations

- **HDFS**
- **HPSS**
- **Ceph**
- **S3 improvements**
- **Google Drive (Fall 2016)**



# New CLI ⚡

- **Client (not ssh)**
- **Uses Python SDK**
- **Open source**



# HTTPS Access to Endpoints (coming soon) (demo)

- **Enhanced use of research storage**
  - Asynchronous, bulk transfer via GridFTP
  - (soon) Synchronous remote access via HTTPS
- **Enhanced Globus web app**
  - Browser-based upload/download
  - Inline file viewer
- **Integration with clients and web apps**



# HTTPS server demo

“A single global information space”





# HTTPS access to endpoints



Making  
research data storage  
a first-class entity on the web



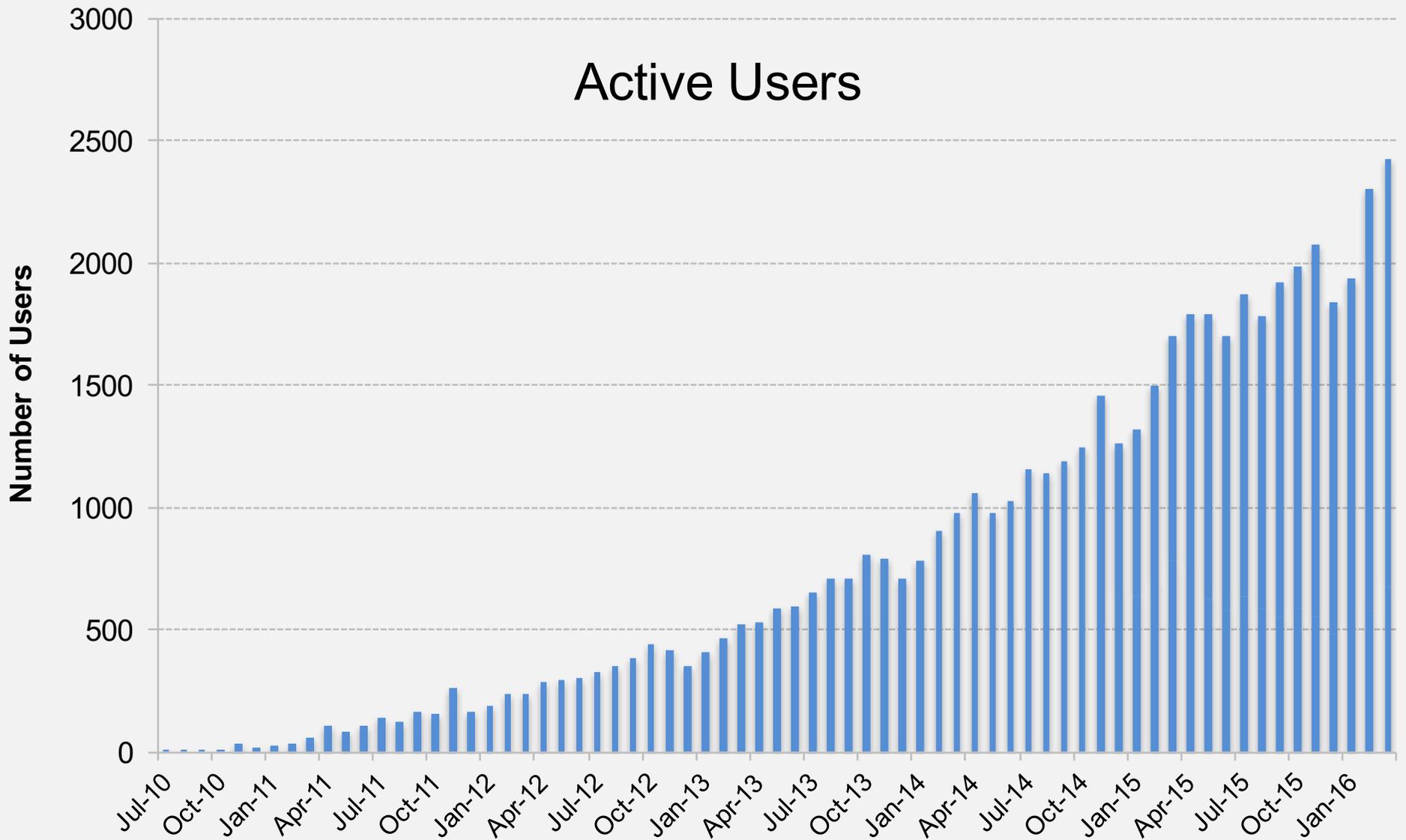
# Sustainability Update

Vas Vasiliadis



# Users, usage continue steady growth...

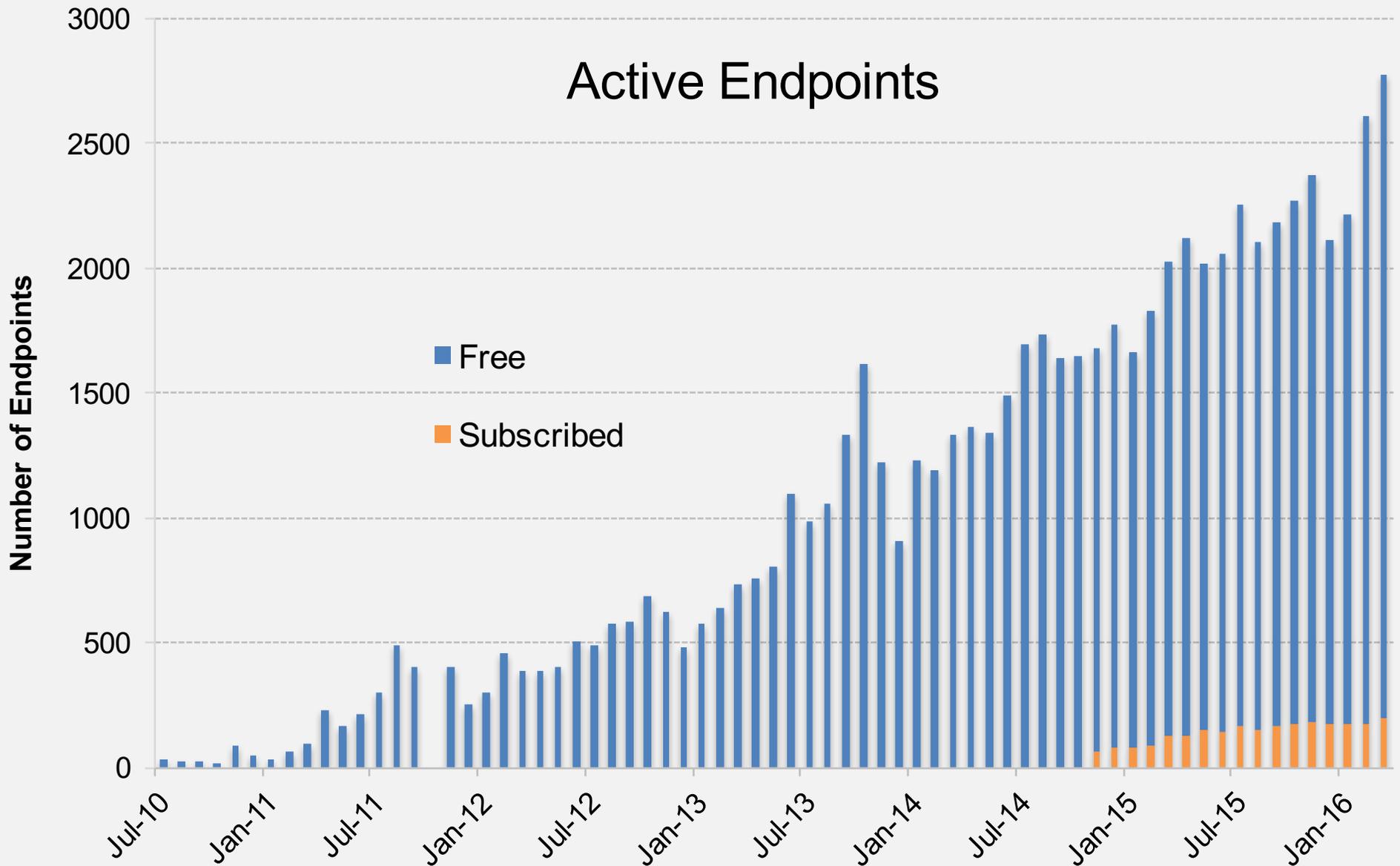
## Active Users





# ...but freemium gap is widening

## Active Endpoints





\$2MM



## Institutional adoption lags subscription

- **300+ institutions using**
- **64 institutions subscribing**
- **Subscriptions: \$2.0MM**
- **Annual operating costs: \$4.0MM**



# Planned changes to subscriptions

- **Encourage campus wide adoption (usage based model discourages it)**
- **Usage based tiers → Unlimited tiers**
- **Tiers based on criteria to be determined**
- **Multi-year phased payment plan**



# Tough choices

- **NSF will not be single-payer for the community**
- **Limit free use, e.g. 1TB/mo.**
- **Remove free tier**
- **...**



# Ask what you can do for Globus!

- **Make the business case to leadership**
  - We can help with usage data, rationale
- **Be a Globus champion on campus**
  - Actively promote usage
  - We can help with material, webinars, etc.
- **Bring a friend**
- **Consider the value Globus delivers ....is it worth \$27.26/day?**
  - What can we do to help you understand the value and sell it to your institution?

# From Science Software to Science Platform



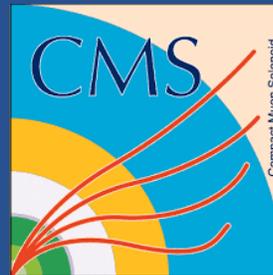
globus



# Custom cyberinfrastructure



ATLAS  
PERIMENT



Open Science Grid



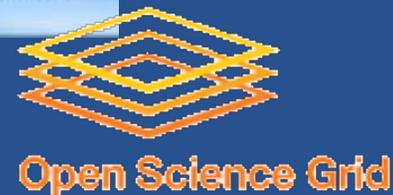
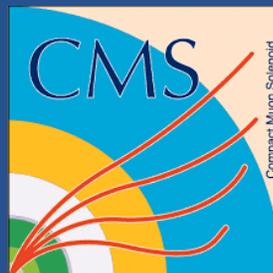
National Ecological Observatory Network

- Expensive
- Lack of skilled staff
- Quality challenges
- Non-interoperable
- Technical debt
- Support burden
- Slows down research!





# Omni-disciplinary CI

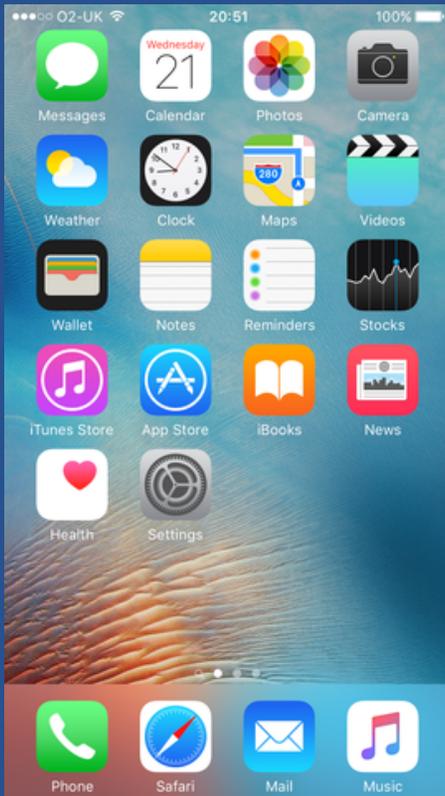


Omni-disciplinary CI software:  
platform and apps



# The importance of platforms

A good platform reduces costs and promotes usability, extensibility, interoperability



Apple iOS and app store:

- Rapid app development
- Common user experience
- Global distribution and monetization

How do we do the same for science and engineering?

iOS



# Cloud has transformed platforms and how software is delivered

Software as a service: **SaaS**  
(web & mobile apps)



**NETFLIX**



Platform as a service: **PaaS**



Microsoft Azure



Infrastructure as a service: **IaaS**



Microsoft Azure

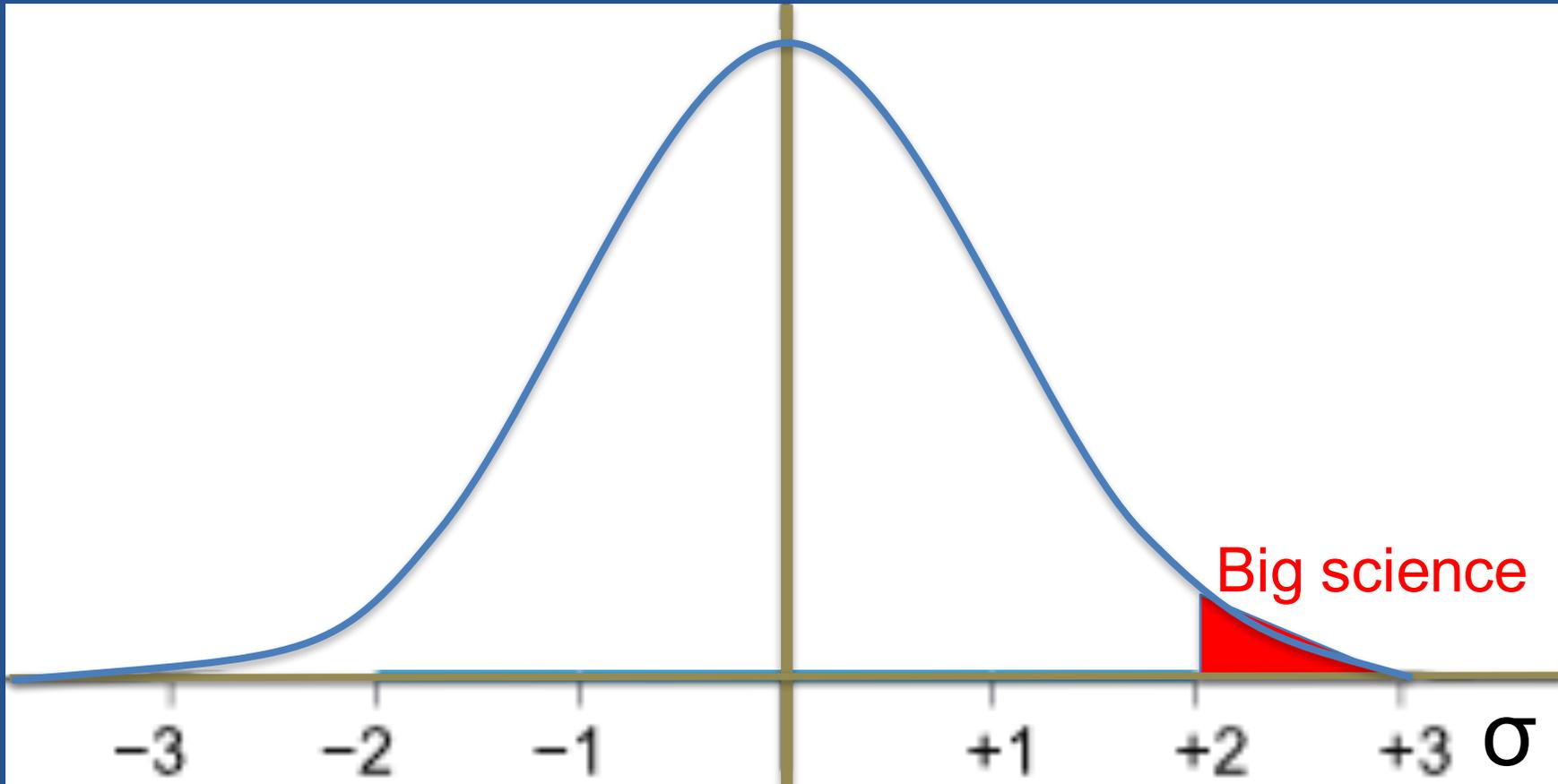


Google Compute Engine

PaaS enables more rapid, cheap, and scalable delivery of powerful apps—as SaaS



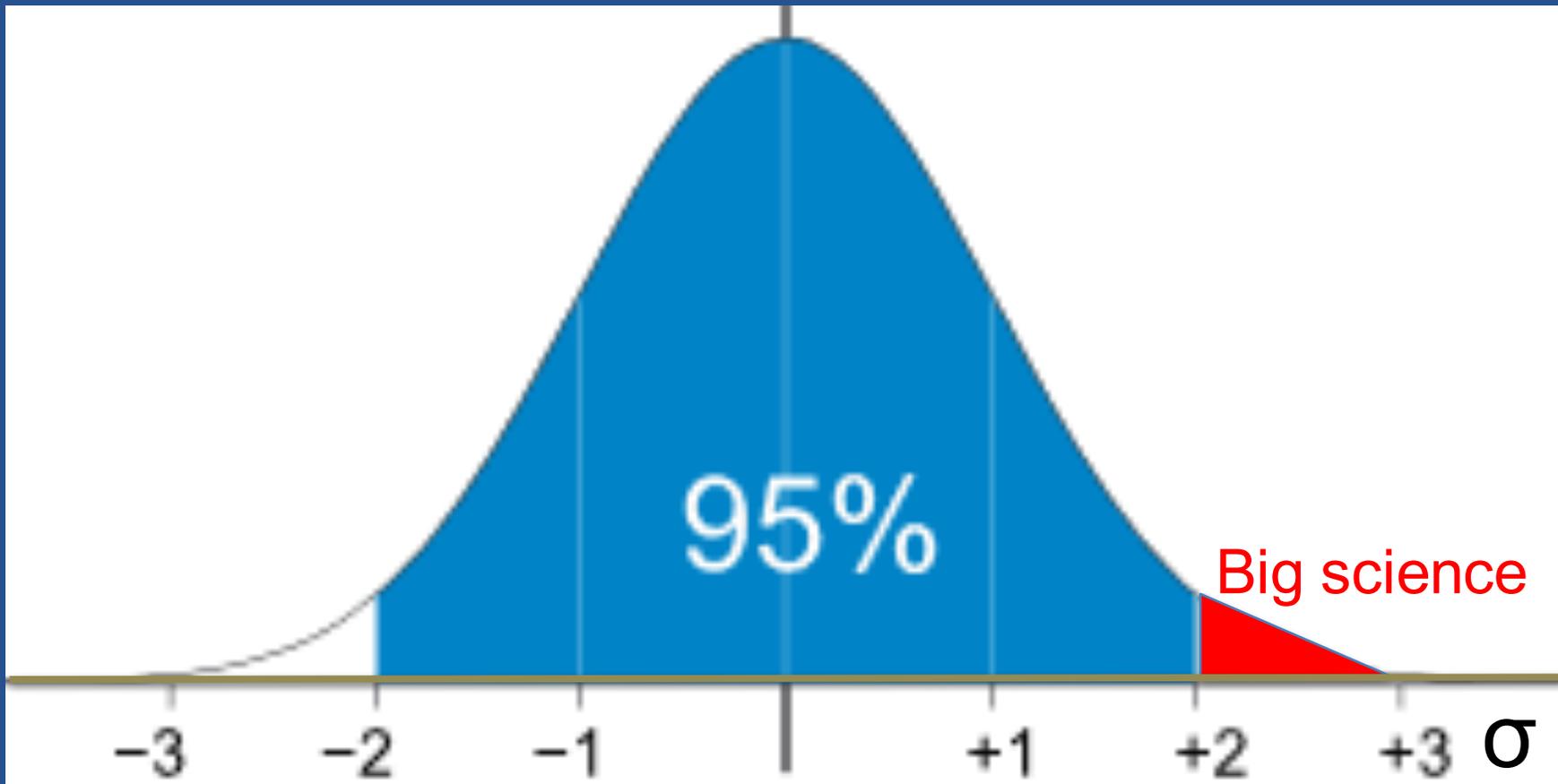
# PaaS can do the same for science



We can leverage cloud to provide solutions that span the vast majority of researcher needs



# PaaS can do the same for science



We can leverage cloud to provide solutions that span the vast majority of researcher needs



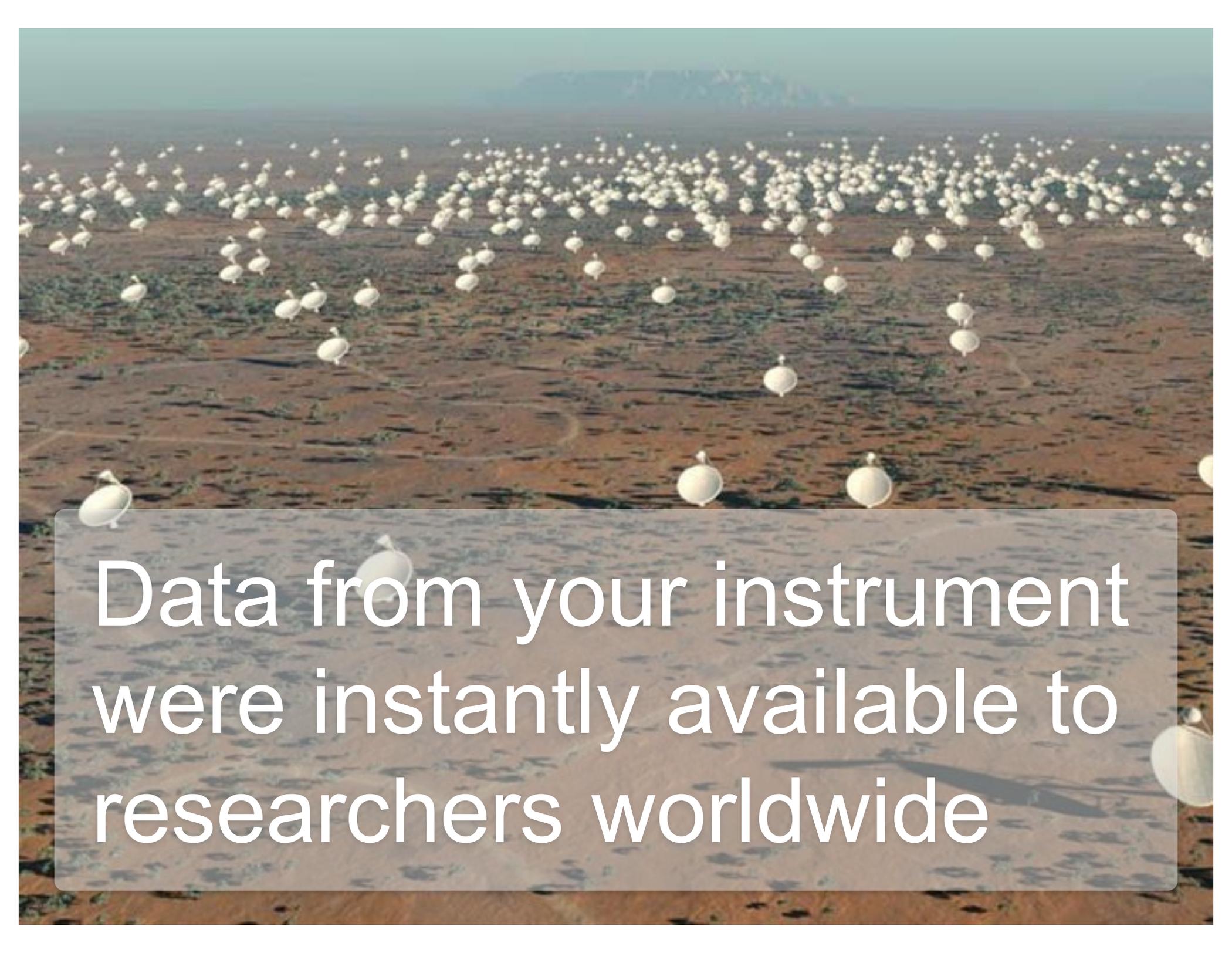
Why does this matter?

Consider the various infrastructures to which researchers have access ...

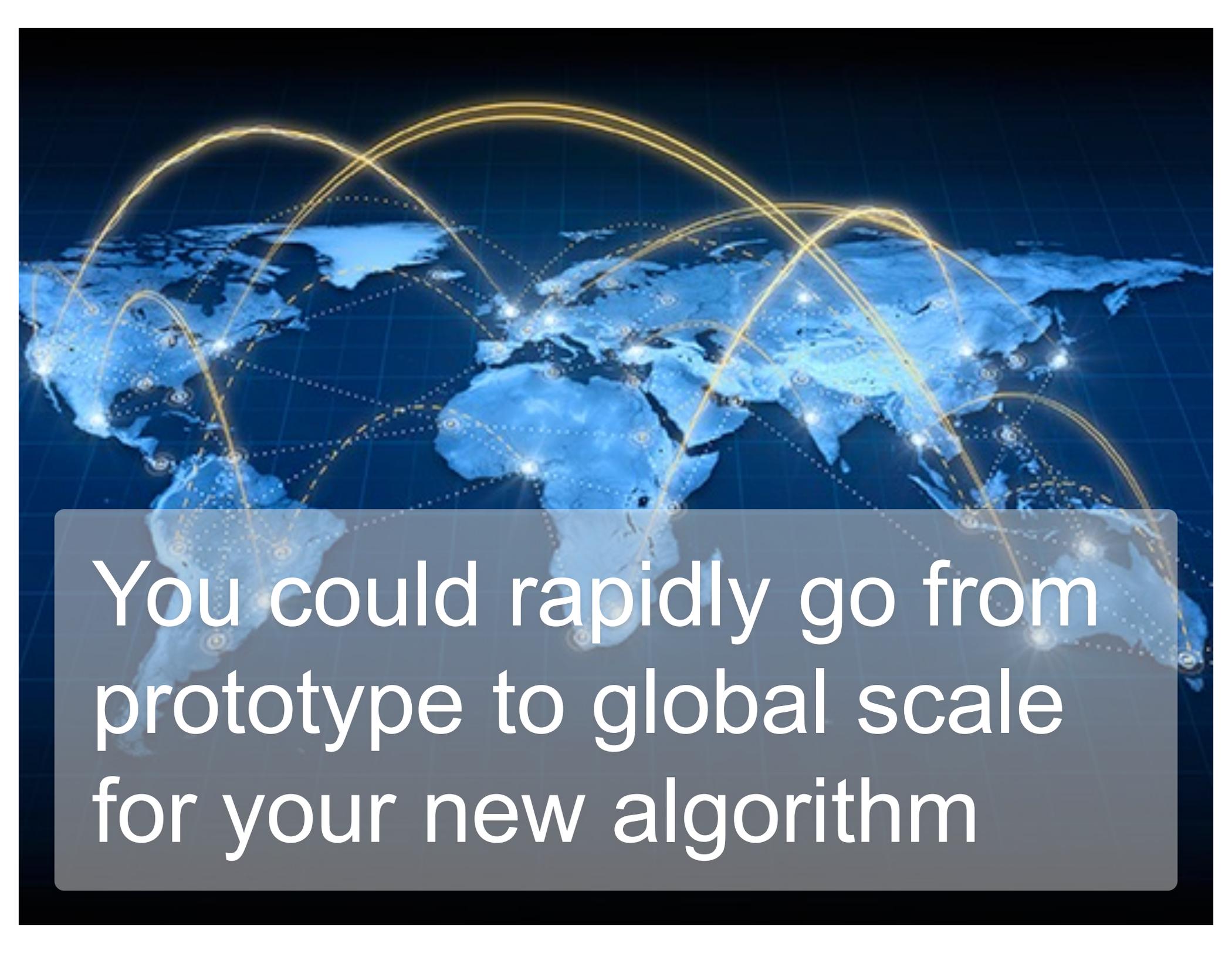
What could they do if they had this platform?



Imagine if...

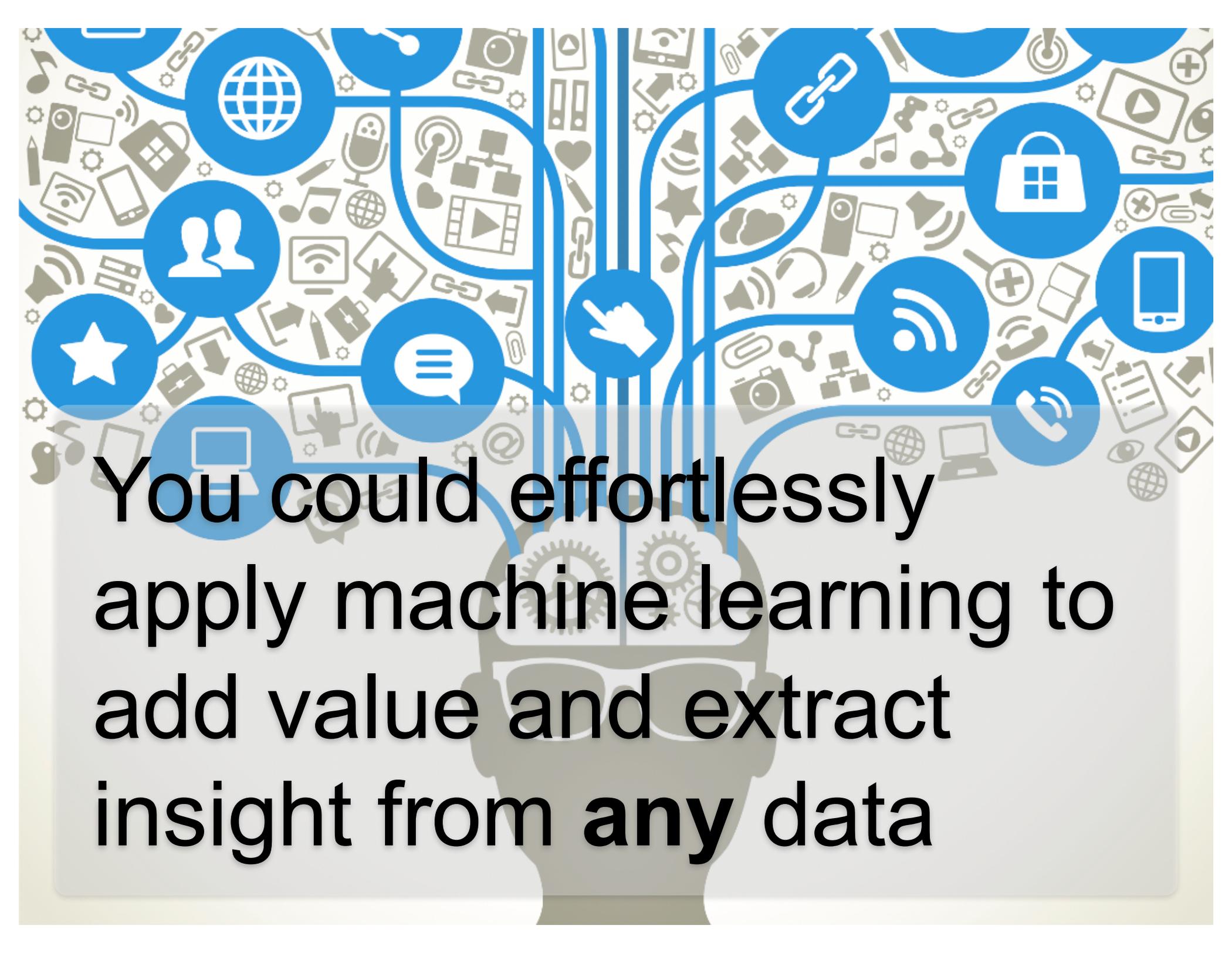
A wide-angle photograph showing a vast, flat landscape of reddish-brown earth with sparse green vegetation. In the foreground and middle ground, hundreds of white birds, likely terns, are scattered across the terrain. The birds are mostly facing away from the camera, and their shadows are cast on the ground. In the far distance, a range of low mountains or hills is visible under a clear, light blue sky. A semi-transparent grey box with white text is overlaid on the bottom half of the image.

Data from your instrument  
were instantly available to  
researchers worldwide



You could rapidly go from  
prototype to global scale  
for your new algorithm





You could effortlessly  
apply machine learning to  
add value and extract  
insight from **any** data

Your research computing center became an innovation center



A photograph of a wooden bridge with a metal truss railing, spanning a stream in a lush green forest. The bridge is made of light-colored wooden planks and leads towards the background. The surrounding trees are dense and vibrant green. A semi-transparent dark grey rectangular box is overlaid on the upper portion of the image, containing white text.

Your campus infrastructure  
linked seamlessly with  
global cyberinfrastructure



# A science CI platform can spur creation of a science CI ecosystem

## Software as a service: **SaaS**



(web and mobile apps)

**NETFLIX**



## Platform as a service: **PaaS**



Microsoft Azure



## Infrastructure as a service: **IaaS**



Microsoft Azure



Google Compute Engine

In so doing, we can slash costs, improve quality, and accelerate discovery across the sciences



# A science CI platform can spur creation of a science CI ecosystem

## Software as a service: SaaS



(web and mobile apps)

**NETFLIX**



2010-

## Platform as a service: PaaS



Microsoft Azure



2014-

## Infrastructure as a service: IaaS



**EC2**



Microsoft Azure



**S3**



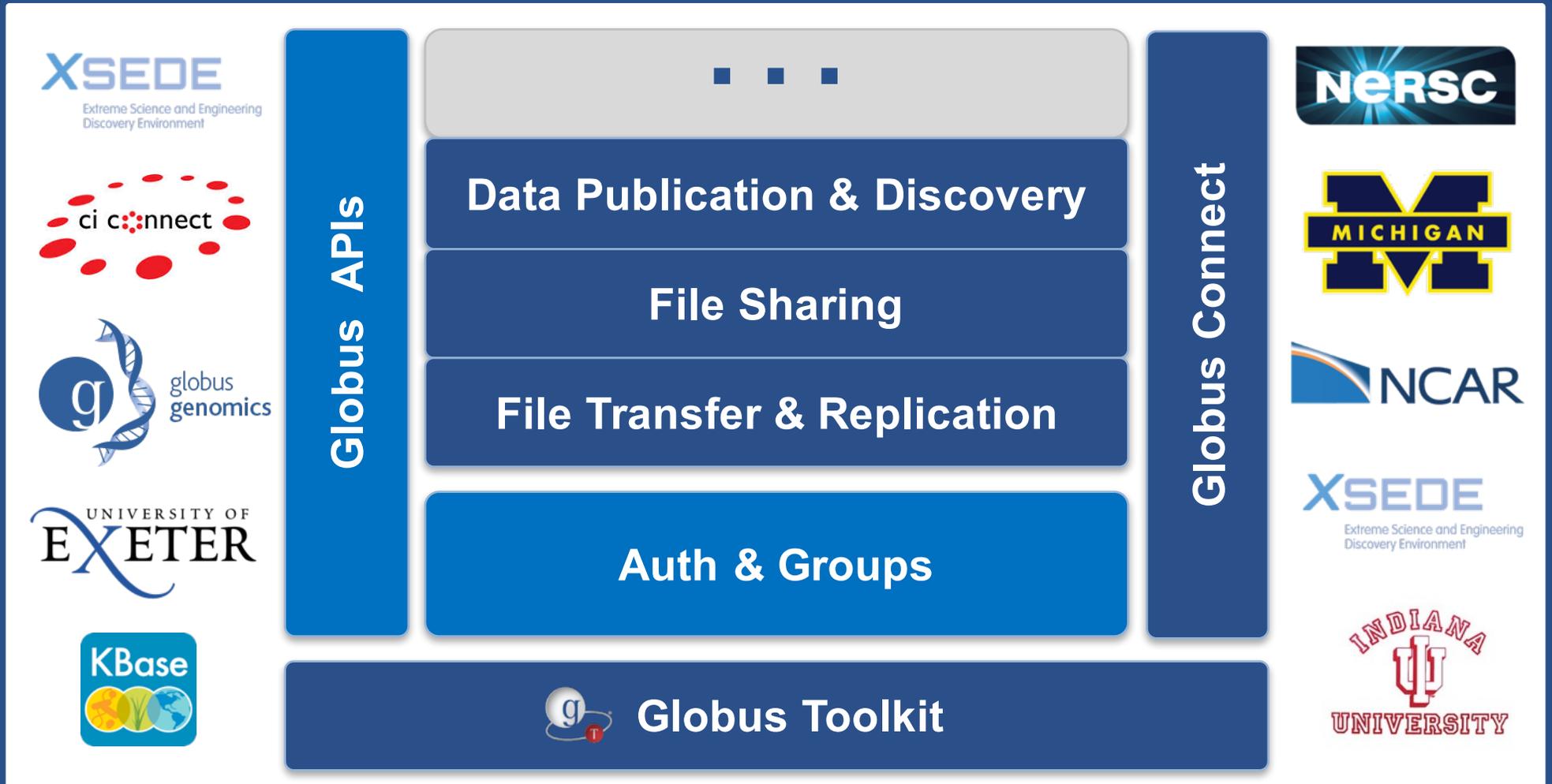
Google Compute Engine



In so doing, we can slash costs, improve quality, and accelerate discovery across the sciences



# Globus PaaS: Ecosystem enabler





# Globus PaaS at NCAR

- Research Data Archive at NCAR
- Integrate Globus for data downloads
- Shared endpoint with subfolder per request
- Single sign on via streamlined account provisioning

Find Data | Ancillary Services | About/Contact

All Datasets | Recently Added/Updated | Browse the RDA

- GCMD Topic:
  - Agriculture • Atmosphere • Biosphere • Climate
  - Oceans • Paleoclimate • Solid Earth • Spectral/En
- Atmospheric Reanalysis Data:
  - All Reanalysis Datasets • BPRC Arctic System Reanal
  - ECMWF ERA15 Reanalysis (ERA15) • ECMWF ERA40
  - ECMWF Interim Reanalysis (ERA-I) • JMA Japanese 2
  - JMA Japanese 55-year Reanalysis (JRA55) • NCEP Cl
  - NCEP North American Regional Reanalysis (NARR) •
  - NCEP/NCAR Reanalysis Project (NNRP) • NOAA-CIRE
- Station Observations:
  - Land Surface Air Temperature: Hourly, Monthly

Find Platform Observations datasets

**CISL Research Data Archive**

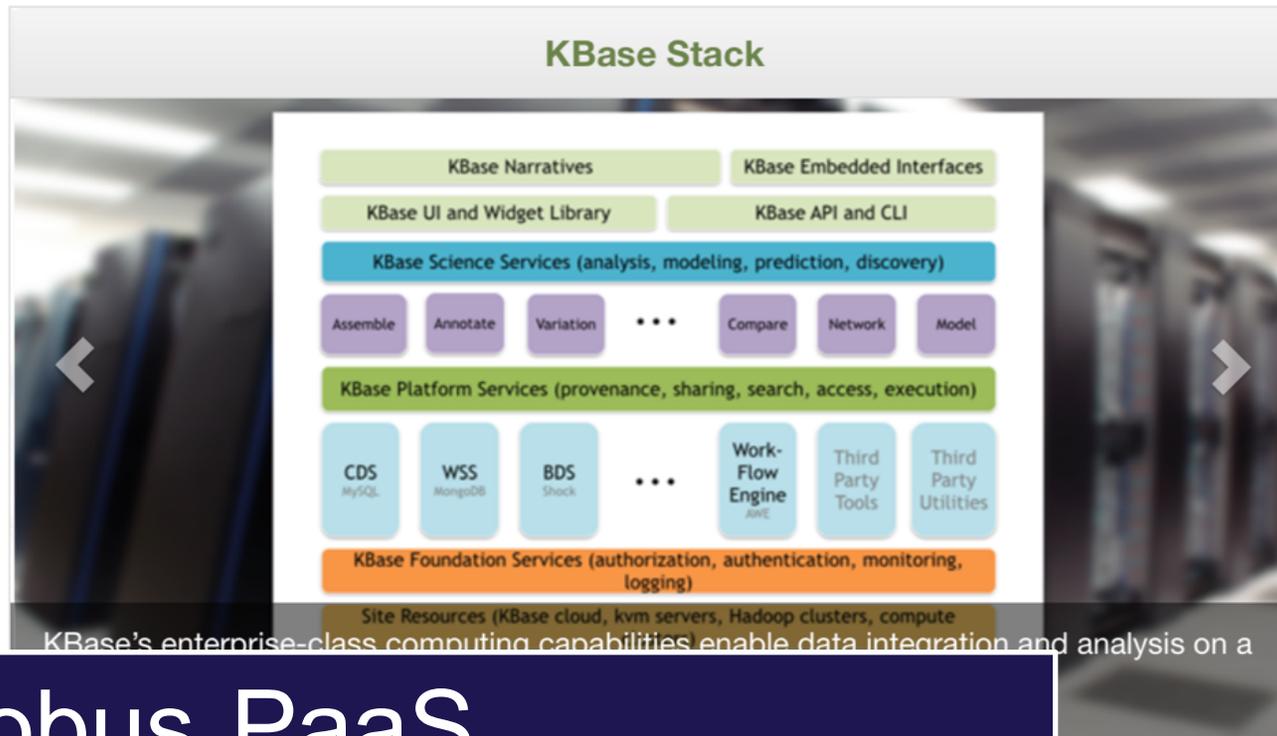
Managed by NCAR's Data Support Section  
Data for Atmospheric and Geosciences Research

**RDA**



**KBase Maintenance Window - 5/27/15** in 13 hours Wed May 27 from 11:00am to 3:00pm

## KBase: The Department of Energy Systems Biology Knowledgebase



 [New to KBase?](#)

 [Search Data](#)

 [Sign In](#)

**Globus PaaS accelerates development of new kBase services**

...cs and systems biology for microbes, ...hods with other scientists.

Society of Microbiology (ASM) 2015 Annual Meeting



# Globus PaaS and Open Science Grid

Show the bookmarks in this folder

https portal.osgconnect.net/SignIn

Reader

Contacts Bonjour MWT2 ATLAS OSG Coding IB OSX Mac PM FENIX Media GO Biz Props



Support

Resources

OSG Connect

Sign In / Register

## Efficiently connect your science to cycles and data



OSG Connect offers users simple access to distributed high throughput computing resources, and reliable, high-performance file transfer services.

### Sign In

[Sign Up with Globus Online](#)

Using your [InCommon / CILogon](#) login.

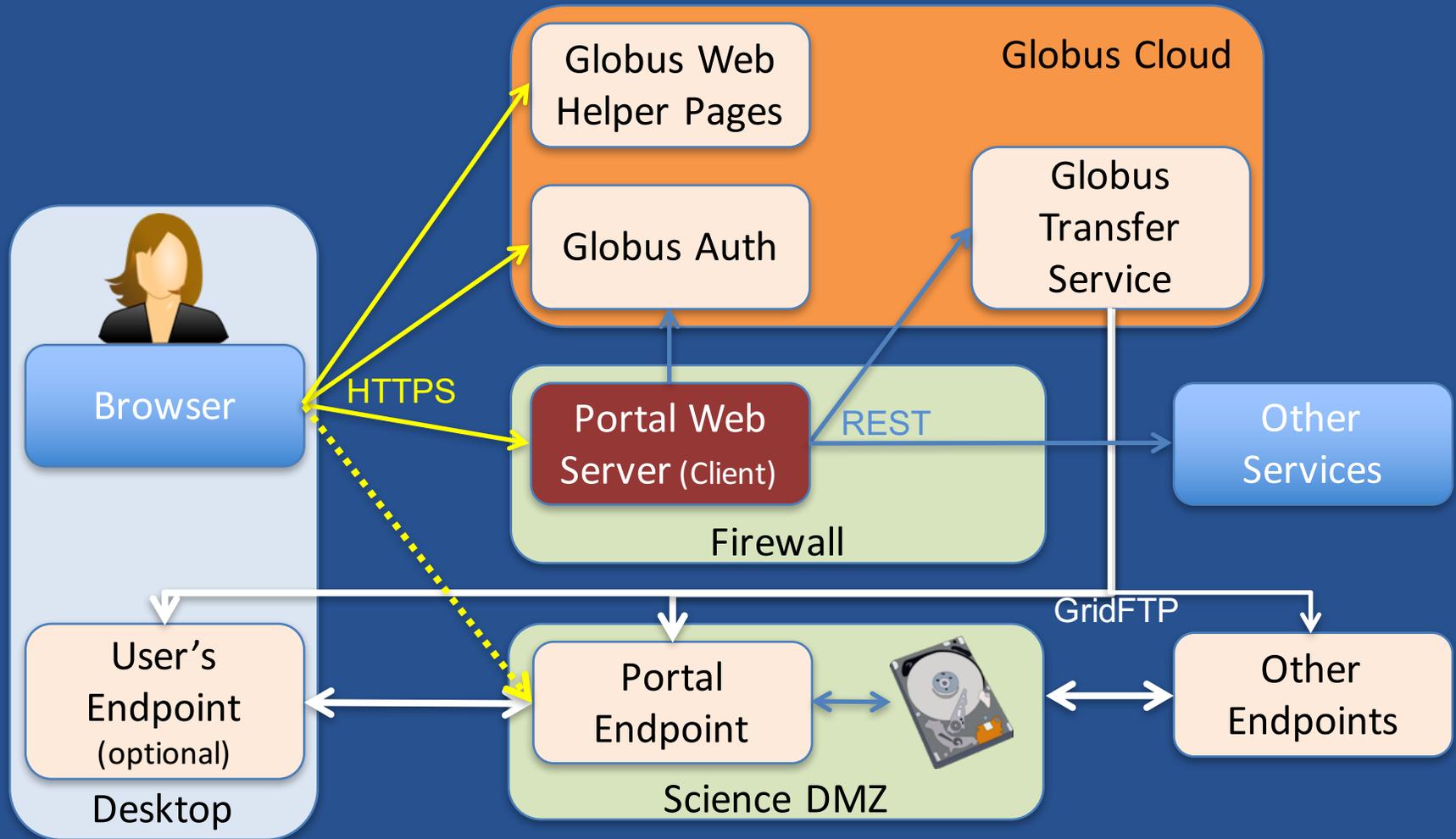
[alternate login](#)

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

Proceed



# Prototypical research data portal



- Move portal storage into Science DMZ, with Globus endpoint
- Leave Portal Web server behind firewall
- Globus handles the security and data heavy lifting



Modern Research Data Portal

[TRANSFER](#) | [GRAPH](#) | [PROFILE](#) | [LOGOUT](#) | [IAN@GLOBUSID.ORG](mailto:IAN@GLOBUSID.ORG)

# Modern Research Data Portal

It's how research data management is done!

## **Globus Transfer API**

API reference for transfer and sharing functions.

---

## **Globus Auth API**

API reference for authentication and authorization.

---

## **Frequently Asked Questions**

When all else fails...



# Repository

Select some dataset(s) to transfer **or** click on a dataset name to browse its files.

Dataset Name	Select
 Atlanta International Airport Climate Data	<input checked="" type="checkbox"/>
 Boston Logan International Airport Climate Data	<input type="checkbox"/>
 Baltimore-Washington International Airport Climate Data	<input type="checkbox"/>
 Dallas-Fort Worth International Airport Climate Data	<input type="checkbox"/>
 Jacksonville International Airport Climate Data	<input type="checkbox"/>
 Las Vegas McCarran International Airport Climate Data	<input type="checkbox"/>
 LaGuardia Airport Climate Data	<input type="checkbox"/>
 Miami International Airport Climate Data	<input type="checkbox"/>
 Milwaukee Mitchell International Airport Climate Data	<input type="checkbox"/>
 Philadelphia International Airport Climate Data	<input type="checkbox"/>
 San Francisco International Airport Climate Data	<input type="checkbox"/>

**Transfer**

## Browse Endpoint

Endpoint



Path

Go

Start by selecting an endpoint.

Label This Transfer

This will be displayed in your transfer activity.

Submit

# Repository

Select some dataset(s) to graph **or** click on a dataset name to browse its files.

Dataset Name	Select
 Atlanta International Airport Climate Data	<input checked="" type="checkbox"/>
 Boston Logan International Airport Climate Data	<input type="checkbox"/>
 Baltimore-Washington International Airport Climate Data	<input type="checkbox"/>
 Dallas-Fort Worth International Airport Climate Data	<input type="checkbox"/>
 Jacksonville International Airport Climate Data	<input type="checkbox"/>
 Las Vegas McCarran International Airport Climate Data	<input type="checkbox"/>
 LaGuardia Airport Climate Data	<input type="checkbox"/>
 Miami International Airport Climate Data	<input type="checkbox"/>
 Milwaukee Mitchell International Airport Climate Data	<input type="checkbox"/>
 Philadelphia International Airport Climate Data	<input type="checkbox"/>
 San Francisco International Airport Climate Data	<input type="checkbox"/>

Year

Generate Graphs



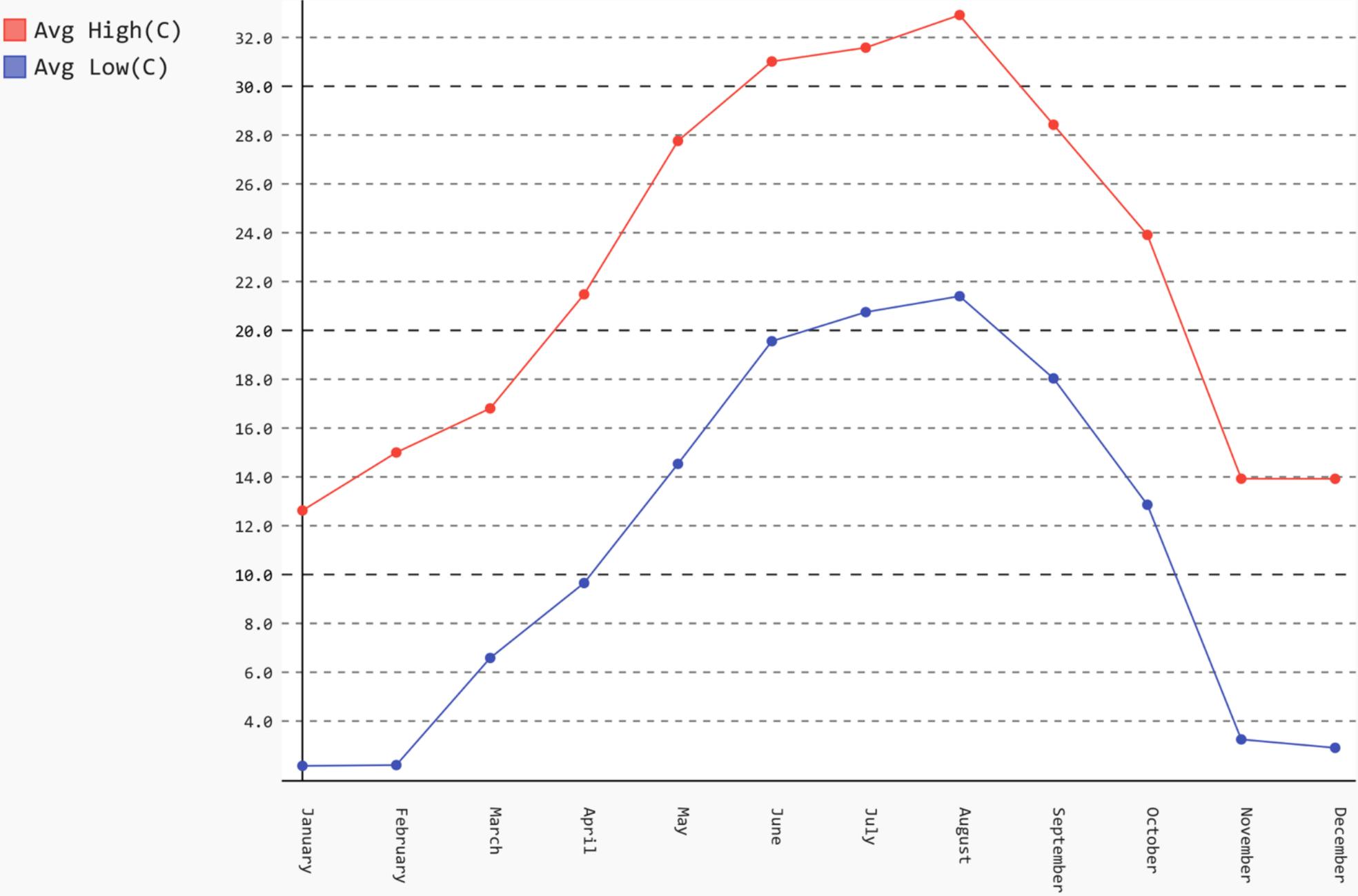
3-file SVG upload to /portal/processed/Graphs for ian@globusid.org/ on Globus Tutorial HTTPS Endpoint Server completed!

## Browse Endpoint

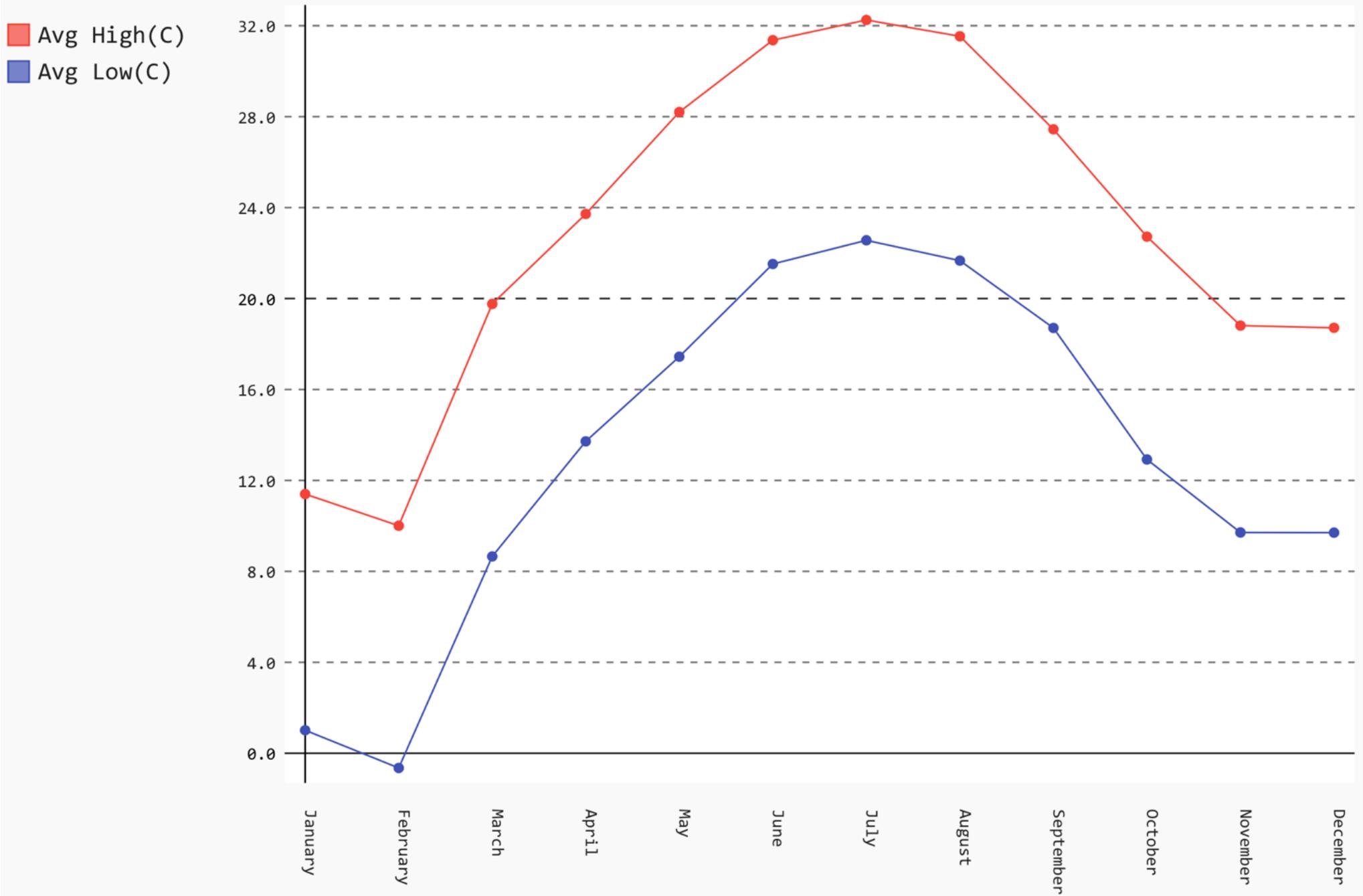
Displaying a list of files in endpoint **Globus Tutorial HTTPS Endpoint Server**.  
Click on a file name to display the contents in your browser or download via HTTP.  
You can also [transfer files using the Globus Webapp](#).

File Name	Size
 Average High-Low Temperature from Atlanta International Airport Climate Data for 1951.svg	25757
 High Temperature Ranges from Atlanta International Airport Climate Data for 1951.svg	30610
 Precipitation from Atlanta International Airport Climate Data for 1951.svg	20406

# Average High-Low Temperature from Atlanta International Airport Climate Data for 1951

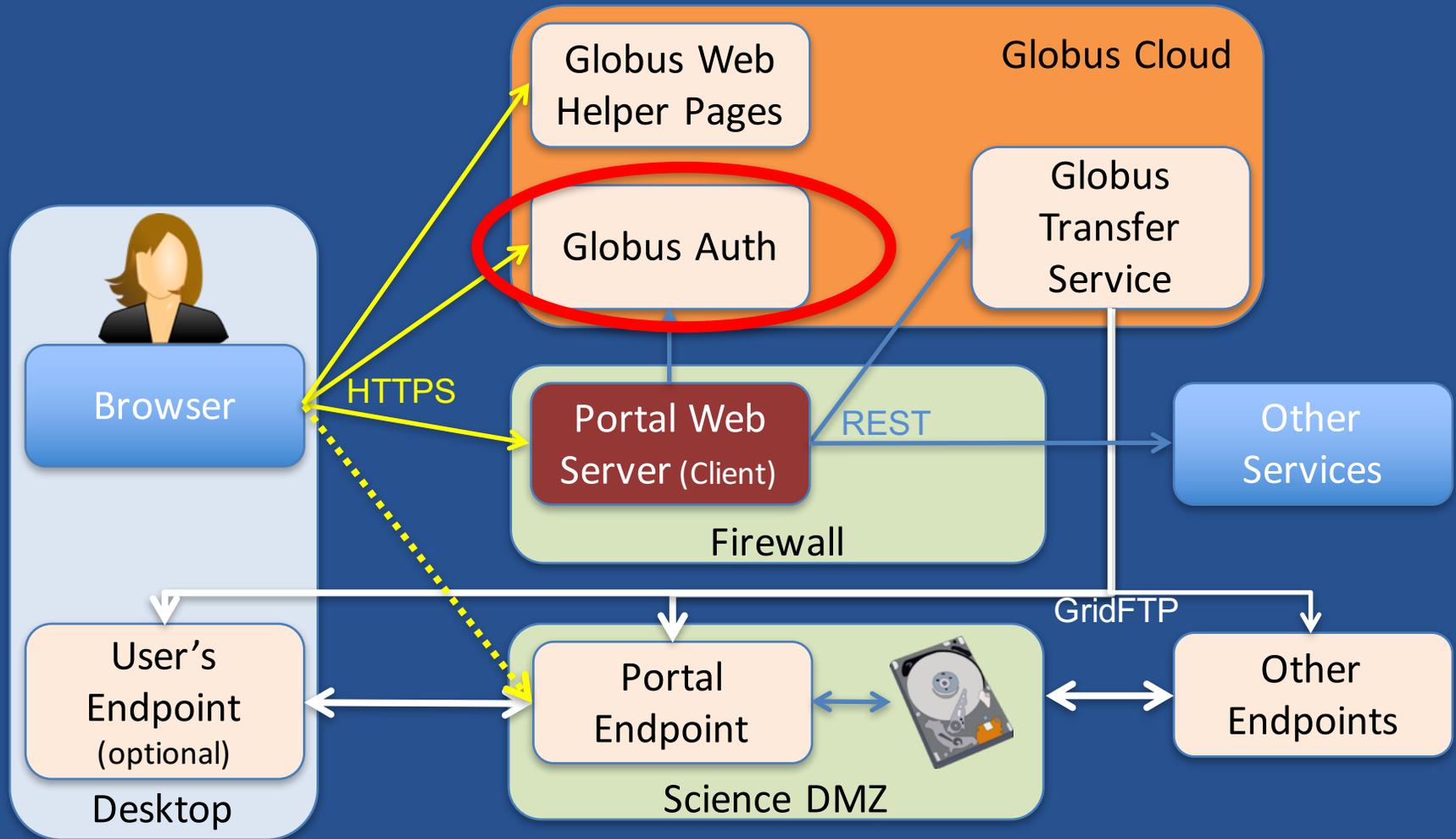


# Average High-Low Temperature from Atlanta International Airport Climate Data for 2015





# Prototypical research data portal



- Move portal storage into Science DMZ, with Globus endpoint
- Leave Portal Web server behind firewall
- Globus handles the security and data heavy lifting



# Globus Auth

- **Foundational identity and access management (IAM) platform service**
- **Simplify creation and integration of advanced apps and services**
- **Brokers authentication and authorization interactions between:**
  - End-users
  - Identity providers: XSEDE, InCommon, web apps
  - Resource servers: services with REST APIs
  - Clients: web, mobile, desktop, command line apps
  - Resource servers acting as clients to other resource servers

***<https://docs.globus.org/api/auth>***



Based on widely used web standards

- **OAuth 2.0 Authorization Framework**
  - aka OAuth2
- **OpenID Connect Core 1.0**
  - aka OIDC
- **Allows use of standard OAuth2 and OIDC libraries**
  - E.g., Google OAuth Client Libraries (Java, Python, etc.), Apache mod\_auth\_openidc

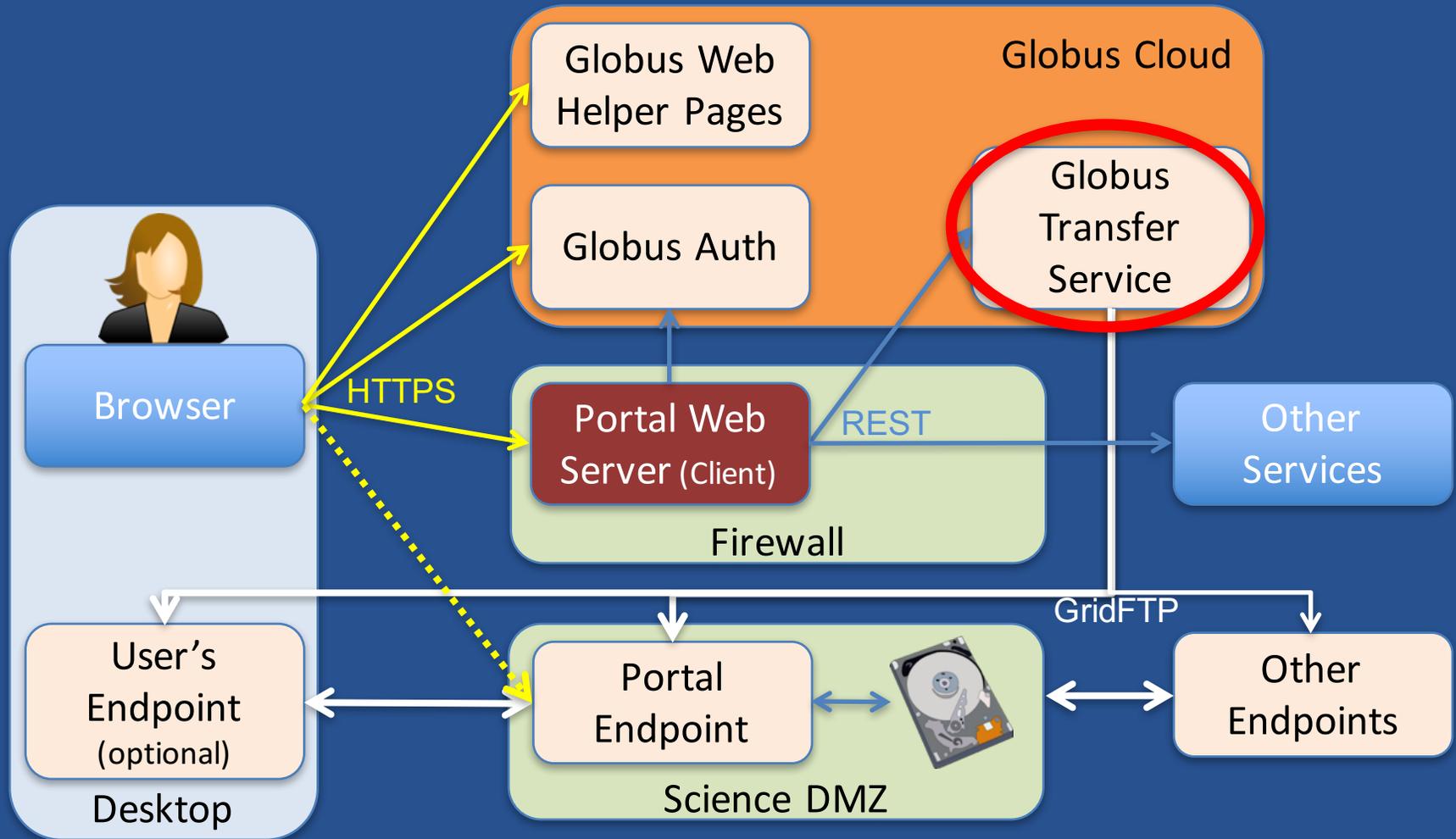


# Globus Auth uses

- **Login to web app**
  - “Log in with Globus”
  - Mobile, desktop, command line apps coming
- **Protect all REST API communications**
  - App → Globus service
  - App → non-Globus service
  - Service → service



# Prototypical research data portal



- Move portal storage into Science DMZ, with Globus endpoint
- Leave Portal Web server behind firewall
- Globus handles the security and data heavy lifting



# Introduction to REST APIs

- **Remote operations on resources via HTTPS**
  - POST ~ = Create (or other operations)
  - GET ~ = Read
  - PUT ~ = Update
  - DELETE ~ = Delete
- **Globus APIs use JSON for documents and resource representations**
- **Resource named by URL**
  - Query params allow refinement (e.g., subset of fields)
- **Requests authorized via OAuth2 access token**
  - Authorization: Bearer asdfkqhafsdafeawk



# Globus transfer API

**Nearly all Globus Web App functionality implemented via public Transfer API**

***<https://docs.globus.org/api/transfer/>***

[HOME](#) / [GLOBUS APIS](#)

## Transfer API Documentation

This API provides a REST-style interface to the [Globus](#) reliable file transfer service. The Transfer API supports monitoring the progress of a user's file transfer tasks, managing file transfer endpoints, listing remote directories, and submitting new transfer and delete tasks. The API is ideal for integration into

Transfer API  
Documentation

API Overview



# Globus Python SDK

- Python client library for the Globus Auth and Transfer REST APIs

<http://globus.github.io/globus-sdk-python/>

globus-sdk-python 0.2.3 documentation »

[next](#) | [modules](#) | [index](#)

## Table Of Contents

- Globus SDK for Python (Beta)
- Installation
- Basic Usage
- API Documentation
- License

## Next topic

- High Level API

## This Page

- Show Source

## Quick search

Enter search terms or a module, class or function name.

## Globus SDK for Python (Beta)

This SDK provides a convenient Pythonic interface to [Globus](#) REST APIs, including the Transfer API and the Globus Auth API. Documentation for the REST APIs is available at <https://docs.globus.org>.

Two interfaces are provided - a low level interface, supporting only `GET`, `PUT`, `POST`, and `DELETE` operations, and a high level interface providing helper methods for common API resources.

Source code is available at <https://github.com/globus/globus-sdk-python>.

## Installation

The Globus SDK requires [Python 2.6+](#) or [3.2+](#). If a supported version of Python is not already installed on your system, see this [Python installation guide](#) .

The simplest way to install the Globus SDK is using the `pip` package manager (<https://pypi.python.org/pypi/pip>), which is included in most Python installations:

```
pip install globus-sdk
```



# Jupyter (iPython) notebook

<https://github.com/globus/globus-jupyter-notebooks>

## Globus SDK

<https://github.com/globus/globus-sdk-python>

## Globus SDK Docs

<http://globus.github.io/globus-sdk-python/>

## Requirements

- You need to be in the tutorial users group for sharing: <https://www.globus.org/app/groups/50b6a29c-63ac-11e4-8062-22000ab68755>
- Installed Globus Python SDK

```
In [15]: from __future__ import print_function # for python 2
tutorial_endpoint_1 = "ddb59aef-6d04-11e5-ba46-22000b92c6ec" # endpoint "Globus Tutorial Endpoint 1"
tutorial_endpoint_2 = "ddb59af0-6d04-11e5-ba46-22000b92c6ec" # endpoint "Globus Tutorial Endpoint 2"
tutorial_users_group = "50b6a29c-63ac-11e4-8062-22000ab68755" # group "Tutorial Users"
```

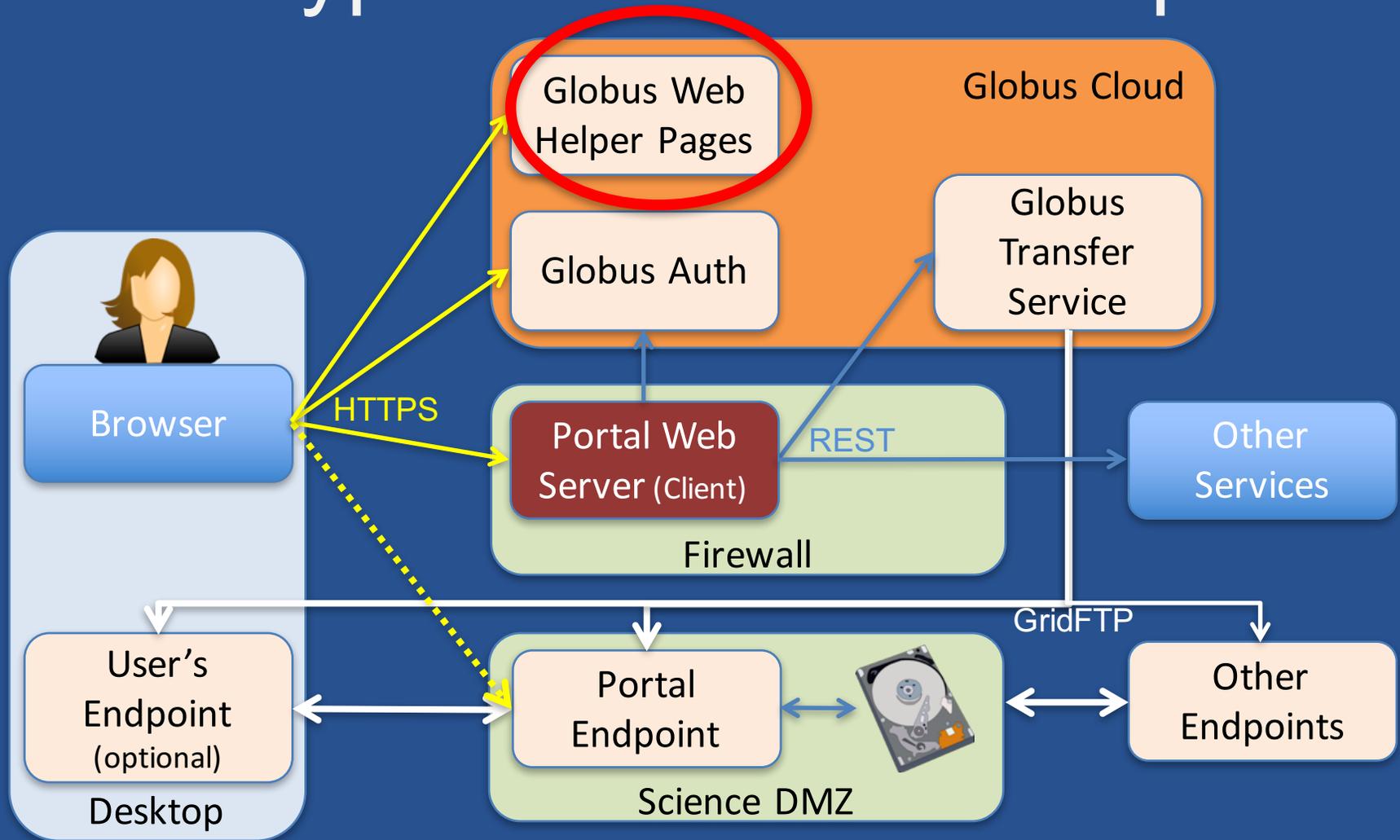
## Configuration

First you will need to configure the client with an OAuth2 access token. For the purpose of this tutorial, you can obtain access tokens via the [tokens.globus.org](https://tokens.globus.org) website. Click the "Jupyter Notebook" option and copy the resulting text below, or click on "Globus CLI" and copy the resulting text into `~/.globus.cfg`.

```
In [16]: transfer_token = None # if None, tries to get token from ~/.globus.cfg file
```



# Prototypical research data portal

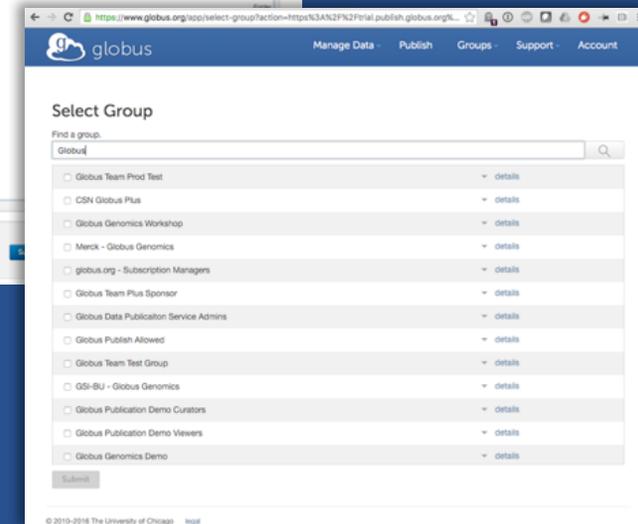
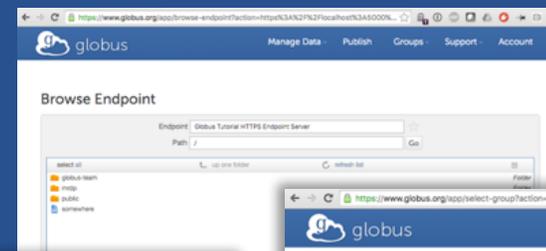
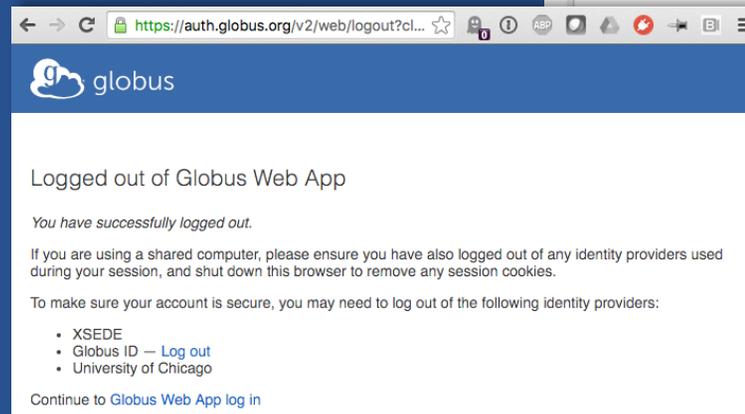


- Move portal storage into Science DMZ, with Globus endpoint
- Leave Portal Web server behind firewall
- Globus handles the security and data heavy lifting



# Globus Helper Pages

- Globus provided web pages designed for use by your web apps
  - Browse Endpoint
  - Select Group
  - Logout



<https://docs.globus.org/api/helper-pages/>



# Globus Helper Pages



Manage Data ▾

Publish

Groups

Support ▾

tuecke ▾

Browse & Discover

Data Publication Dashboard

Communities & Collections

...knowledge and accept  
that datasets you submit to this trial collection: (1)  
will be

globus Manage Data ▾ Publish Groups Support ▾ tuecke ▾

### Select Group

Find a group.

Input Form\*

Datacite Mandatory + R ▾

Submission Workflow\*

Default ▾

Curation Type\*

Edit Metadata ▾

Collection Permissions

Submitters

- All Users
- Restricted to Group...

Access to Data

- All Users
- Restricted to Group...

Curation Group

null



# Branding

## Can skin Globus Auth pages

Header

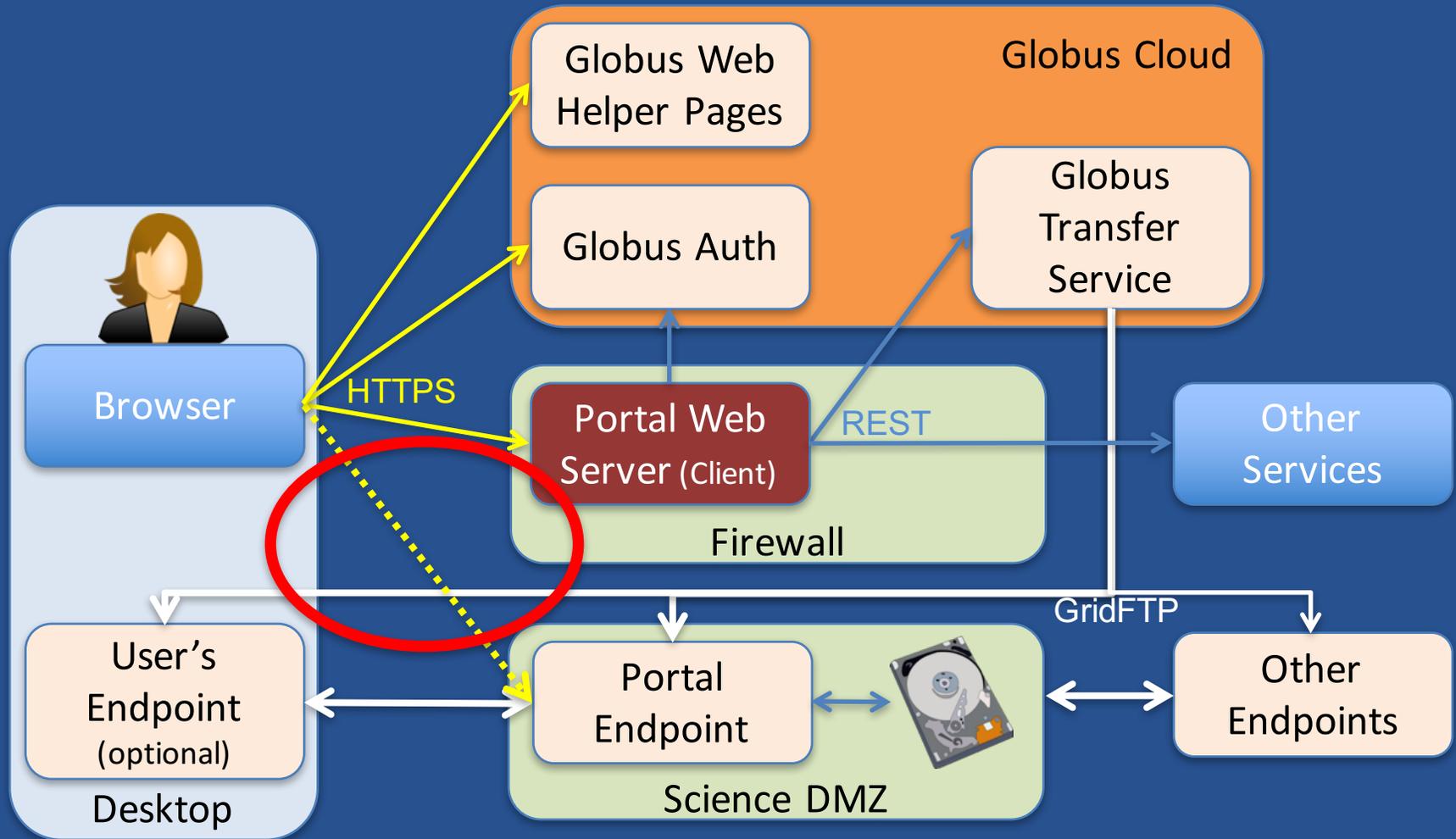
Text

Default IDP

The screenshot shows a web browser window with the URL `https://auth.globus.org/p/login?client_name=globus_webapp_computecanada&state=e...`. The page features a header with the Globus logo and the text "Globus Account Log In". The main content area is branded for "compute canada" and "calcul canada". It includes a colorful geometric graphic on the left and a login form on the right. The form contains the text "Log in to use Compute Canada Globus Web App", "Use your existing organizational login", and a dropdown menu currently showing "Compute Canada". A blue "Continue" button is below the dropdown. A link "Why has this page changed?" is also present. At the bottom, there is a note: "Didn't find your organization? Then use Globus ID to sign up." Three orange arrows point from the text labels on the left to specific elements: "Header" points to the top navigation bar, "Text" points to the main body text, and "Default IDP" points to the dropdown menu.



# Prototypical research data portal



- Move portal storage into Science DMZ, with Globus endpoint
- Leave Portal Web server behind firewall
- Globus handles the security and data heavy lifting

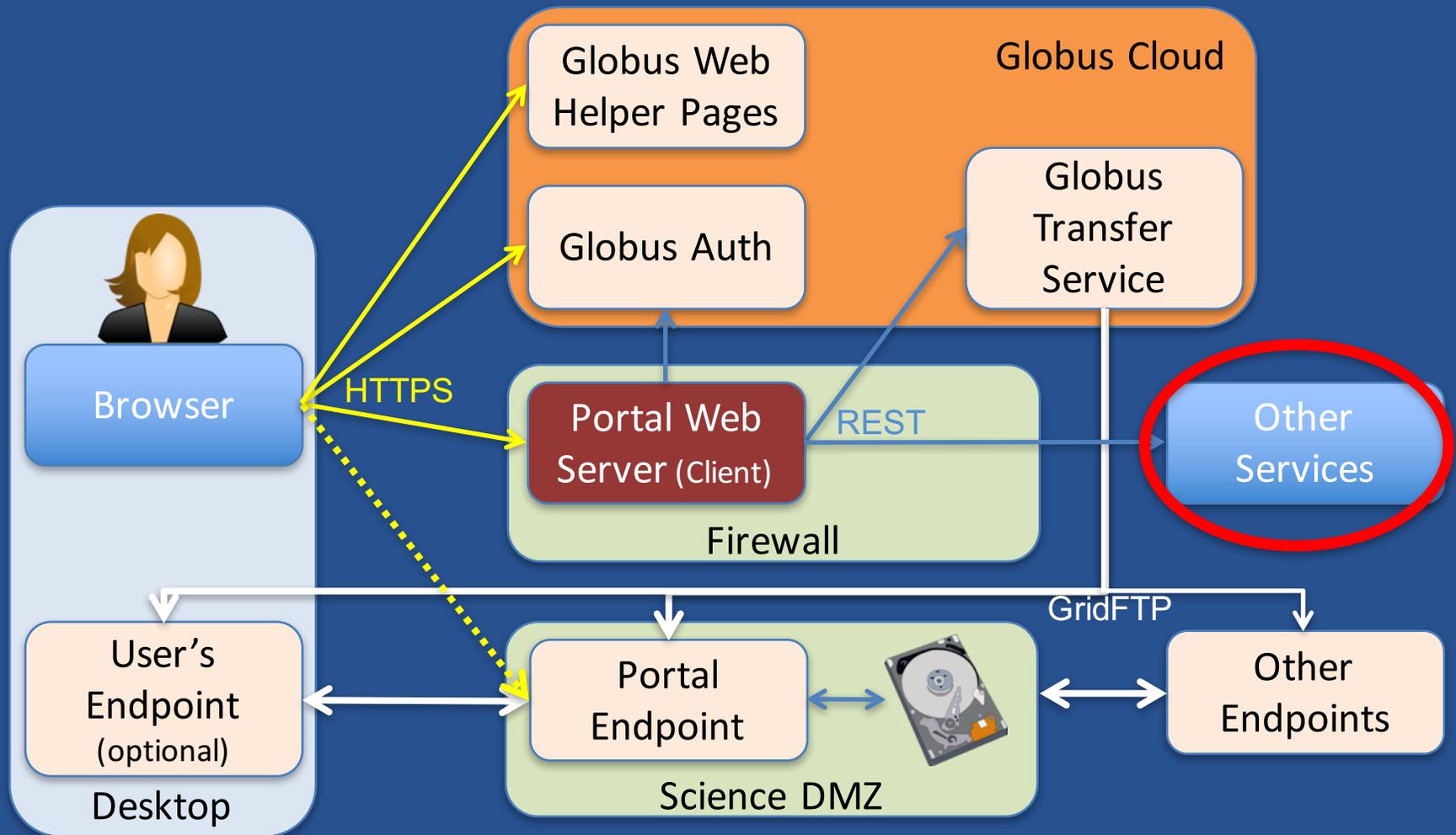


# Globus Connect HTTPS

- **The future of research CI is ... the web**
- **Globus Connect HTTPS unlocks all research storage to the web**
- **Globus Auth provides security glue using standard web security**
- **GridFTP doesn't go away – async, bulk data transfer is important, but its not the end-all, be-all**



# Prototypical research data portal



- Move portal storage into Science DMZ, with Globus endpoint
- Leave Portal Web server behind firewall
- Globus handles the security and data heavy lifting



# Why create your own services?

- **Front-end / back-end within your portal**
  - Remote backend for portal
  - Backend for pure Javascript browser apps
- **Extend your portal with a public REST API, so that other app and service developers can integrate with and extend your portal**



# Why Globus Auth for your service?

- **Outsource all identity management and authentication**
  - Federated identity with InCommon, Google, etc.
- **Outsource your REST API security**
  - Consent, token issuance, validation, revocation
  - You provide service-specific authorization
- **Apps use your service like all others**
  - Its standard OAuth2 and OIDC
- **Your service can seamlessly leverage other services**
- **Other services can leverage your service**

*Add your service to the national cyberinfrastructure platform*



# An extensible platform for CI

**Apps**

**Domain-independent and domain-specific services**

**AWS, Google, Azure services**

**Foundation services: Globus Auth, Groups, etc.**

- Globus provides foundation and other services
- Community can extend to meet domain-specific and domain-independent needs
- Organized for curation and sustainability



# Questions for you

- **What else do we need in a foundational CI platform?**
- **What would you want to build on top of such a platform?**
- **How can we encourage a vibrant community of app developers?**
- **Will you join us in answering these questions?**



Together we can  
create an **integrated ecosystem**  
of **services and applications**  
for the research and education  
community



# Enhancing Globus SaaS



Transforming  
research data storage



Thank you to our supporters!



U.S. DEPARTMENT OF  
**ENERGY**



THE UNIVERSITY OF  
**CHICAGO**



**NIST**



**Argonne**  
NATIONAL LABORATORY



powered by  
**amazon**  
web services



# Thanks!

- **Our 38,000 registered users**
- **The many campuses who provide Globus to their users**
- **Developers of innovative applications that leverage Globus PaaS**
- **Our subscribers**
- **You!**





# Program Preview

- **Today**
  - Lightning talks
  - Cocktail reception and poster viewing
- **Tomorrow**
  - Guest keynote
  - Lightning talks
  - Roundtable discussions



#globus2016

@globusonline

Tell me more about...  
[outreach@globus.org](mailto:outreach@globus.org)