

Globus Introduction

Steve Tuecke
tuecke@globus.org





Thank you to our sponsors!



U.S. DEPARTMENT OF
ENERGY



THE UNIVERSITY OF
CHICAGO

NIST

**National Institute of
Standards and Technology**
U.S. Department of Commerce



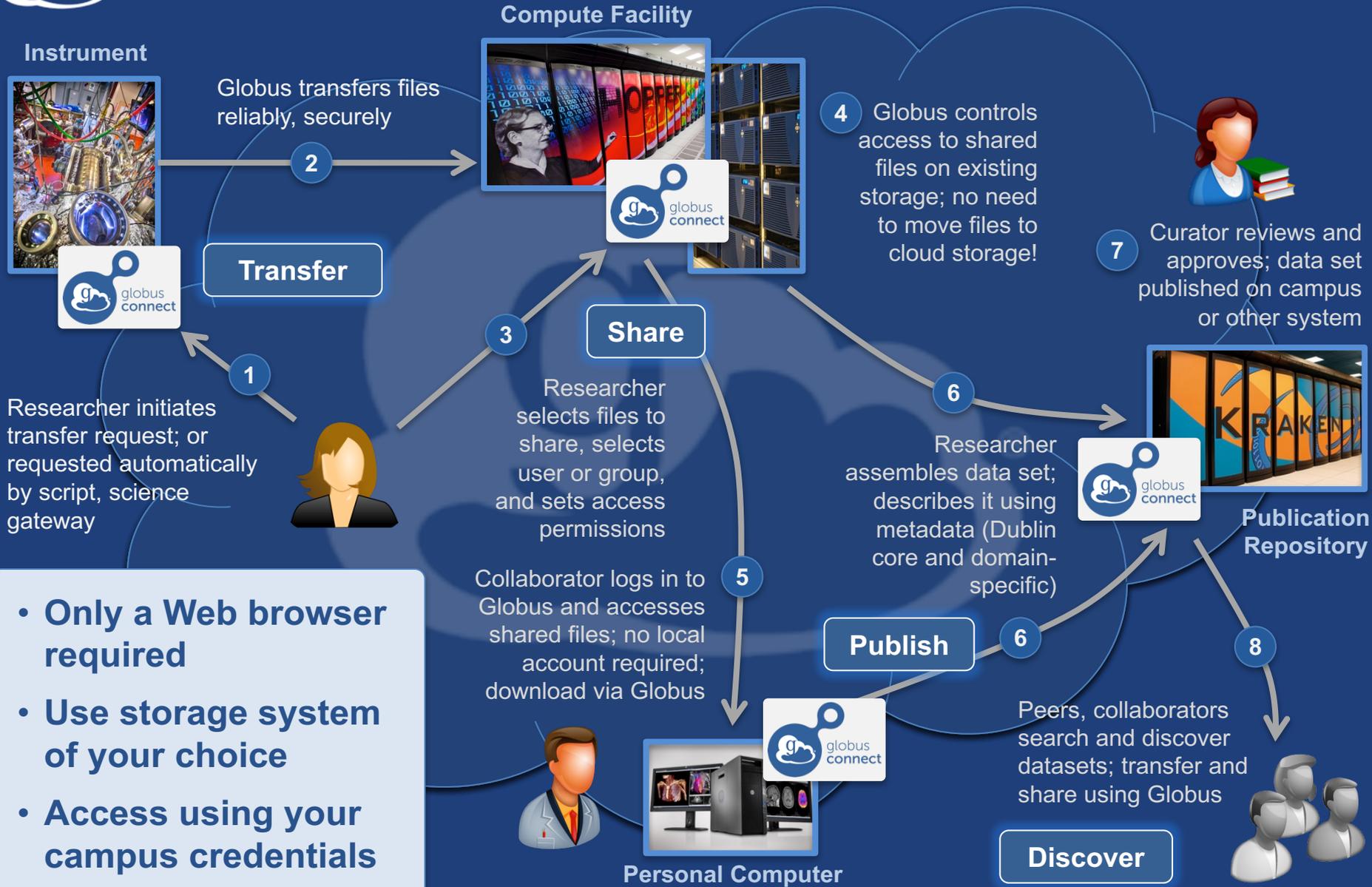
Argonne
NATIONAL LABORATORY



powered by
amazon
web services



Globus and the research data lifecycle



- Only a Web browser required
- Use storage system of your choice
- Access using your campus credentials



Benefits of SaaS

- **Users do not need to deploy software**
 - Easy to access via Web browser
 - Command line, REST interfaces for flexible automation and integration
- **New features automatically available**
- **Reduced IT operational costs**
 - Small local footprint (Globus Connect)
 - Consolidated support and troubleshooting



Globus by the numbers

4

major services

180 PB

transferred

20 billion

files processed

31,000

registered users

13

national labs
use Globus

3,000

active endpoints
per month

~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



Demo

- **Login using existing identity**
- **Identity linking**
- **Endpoint search**
- **Transfer**
- **Globus Connect Personal**
- **Bookmarks**
- **Sharing**
- **Groups**
- **Management console**



Storage connectors

- **Standard storage connectors (Posix)**
 - Linux, Windows, MacOS
 - Lustre, GPFS, OrangeFS, etc.
- **Premium storage connectors**
 - HPSS
 - HDFS
 - S3
 - Ceph RadosGW (S3 API)
 - Spectra Logic BlackPearl
 - Google Drive



New Globus CLI

(coming soon)

```
$ globus transfer
```

```
Usage: globus transfer [OPTIONS] COMMAND [ARGS]...
```

Interact with Globus Transfer API. Transfer, Delete, List, and Rename files on Endpoints, manage your Endpoints and Shares, and monitor your ongoing Transfer Tasks

Options:

```
-F, --format [json|text]  Output format for stdout. Defaults to text
-h, --help                Show this message and exit.
--version                 Show the version and exit.
```

Commands:

```
acl          Manage Endpoint Access Control Lists
async-delete Submit a Delete Task
async-transfer Submit a Transfer Task
bookmark     Manage Endpoint Bookmarks
endpoint     Manage Globus Endpoint definitions
ls           List Endpoint directory contents
mkdir        Make a directory on an Endpoint
rename       Rename a file or directory on an Endpoint
task         Manage asynchronous Tasks
```



Data Publication and Discovery

← → ↻ <https://publish.globus.org/jspui/handle/11466/1> ☆ 👤 ⓘ ☰

 globus Log In Sign Up

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

🔍

Materials Data Facility Community home page



**MATERIALS
DATA
FACILITY**

The Materials Data Facility (MDF) is a scalable repository where materials scientists can publish, preserve, and share research data. The repository provides a focal point for the materials community, enabling publication and discovery of materials data of all sizes.

MDF is a pilot project funded by NIST, and serves as the first pilot community of the [National Data Service](#).

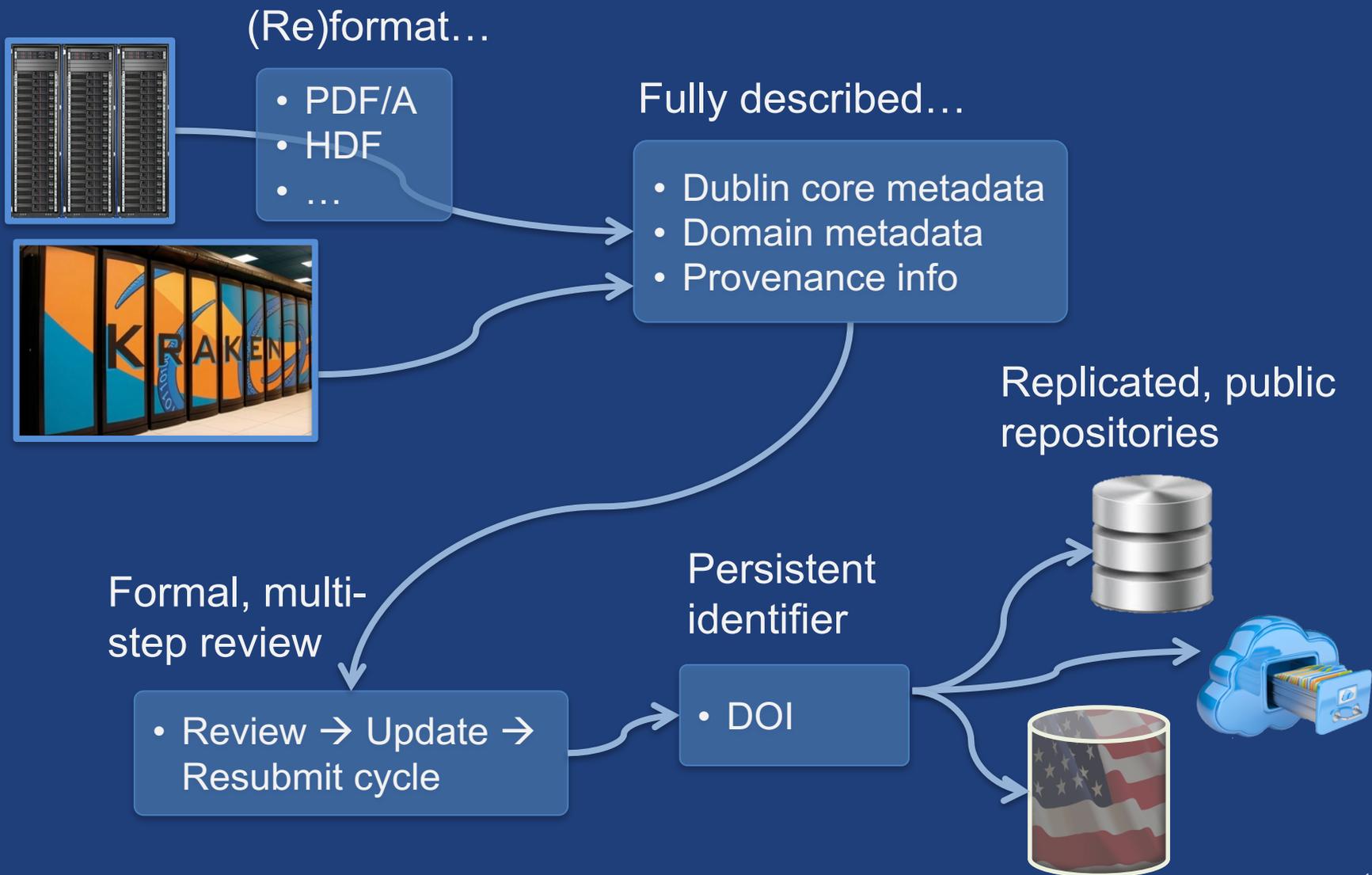
[Contact Ben Blaiszik \(blaiszik@uchicago.edu\) to begin publishing your data](#)

Browse

Issue Date	Author	Title	Subject
------------	--------	-------	---------



Peer reviewed paper data





Globus collection policies

Identifier



Description



Curation



Access



Preservation





Raw NGS output



Minimal metadata...

- Source environment
 - Instrument, timestamp,...
- Unique ID

No curation

- Automated dataset acceptance

Identify...

- Handle

High durability,
low cost store





Analysis results



Optional metadata...

- “Implicit” metadata
- Description through organization

Team review

- Any collaborator may approve

Identify...

- Globus share

Widely accessible stores





Demo

- **Publication**
- **Discovery**

“A single global information space”





HTTPS access to endpoints



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



HTTPS access to endpoints

(coming soon)

- **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**



Cloud has transformed how software and platforms are delivered

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



NETFLIX



Platform as a service: **PaaS**



Microsoft Azure



Infrastructure as a service: **IaaS**



EC2



S3



Microsoft Azure



Google Compute Engine

PaaS enables more rapid, cheap, and scalable delivery of powerful (SaaS) apps



Platform Questions

- **How do you leverage Globus services in your own applications?**
- **How do you extend Globus with your own services?**
- **How do we empower the research community to create an integrated ecosystem of services and applications?**



Research data portal

UCAR NCAR

Closures/Emergencies

Locations/Directions

Find People

Hello tuecke@uchicago.edu [dashboard](#) [sign out](#)

NCAR
UCAR



Research Data Archive
Computational & Information Systems Lab

weather • data • climate

Go to Dataset:

Home

Find Data

Ancillary Services

About/Contact

Data Citation

Web Services

For Staff



NCEP Climate Forecast System Version 2 (CFSv2) Monthly Products

ds094.2

For assistance, contact [Bob Dattore](#) (303-497-1825).

Description

Data Access

Mouse over the table headings for detailed descriptions

Data Description		Data File Downloads		Customizable Data Requests	Other Access Methods	NCAR-Only Access	
		Web Server Holdings	Globus Transfer Service (GridFTP)	Subsetting	THREDDS Data Server	Central File System (GLADE) Holdings	Tape Archive (HPSS) Holdings
Union of Available Products		Web File Listing	Request Globus Invitation	Get a Subset	TDS Access	GLADE File Listing	HPSS File Listing
P R O D U C	Diurnal monthly means	Web File Listing		Get a Subset		GLADE File Listing	HPSS File Listing
	Regular monthly means	Web File Listing		Get a Subset		GLADE File Listing	HPSS File Listing
	Selected Parameter/Level Time Series	Web File Listing		Get a Subset	TDS Access	GLADE File Listing	HPSS File Listing



Globus PaaS





Harness the power of the Globus research data management cloud.



Transfer API

The Transfer API provides a REST-style interface to the Globus reliable file transfer service. The API can be used to monitor the progress of file transfers, manage file transfer endpoints, list remote directories, and submit new transfer and delete tasks.

[LEARN MORE](#)



Resource Providers

Globus allows you, as a resource provider, to easily offer reliable, secure, high-performance research data management capabilities to your users and their collaborators, directly from your own storage infrastructure.

[LEARN MORE](#)



Toolkit

The open source Globus Toolkit is a fundamental enabling technology for the "Grid," allowing users to access high-performance computing resources securely across corporate, institutional, and geographic boundaries without sacrificing local autonomy.

[LEARN MORE](#)



Globus Transfer API

Nearly all Globus Web App functionality implemented via public Transfer API

The screenshot shows the 'globus docs' website. The navigation bar includes 'APIs', 'How To', 'Guides', and 'Support'. The breadcrumb trail is 'HOME / GLOBUS APIS / TRANSFER API DOCUMENTATION'. The main heading is 'Transfer API Documentation'. Below this, it states 'Last Updated: July 18, 2016'. The main text describes the API as a REST-style interface for the Globus file transfer service, supporting monitoring, endpoint management, and task submission. A 'Contents' section lists various API features with links. On the right, a table of contents lists: API Overview, Transfer API Documentation (highlighted), Endpoint Activation, Task Submission, Task Monitoring, File Operations, Endpoint Management, Endpoint Search, Endpoint Roles, Endpoint Bookmarks, and Endpoint ACL.

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



Globus SDK Python

← → ↻ globus.github.io/globus-sdk-python/ ☆ 👤 ⓘ ☰

globus-sdk-python 0.2.5 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

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
```

This will install the Globus SDK and its dependencies.

Bleeding edge versions of the Globus SDK can be installed by checking out the git repository and installing it manually:

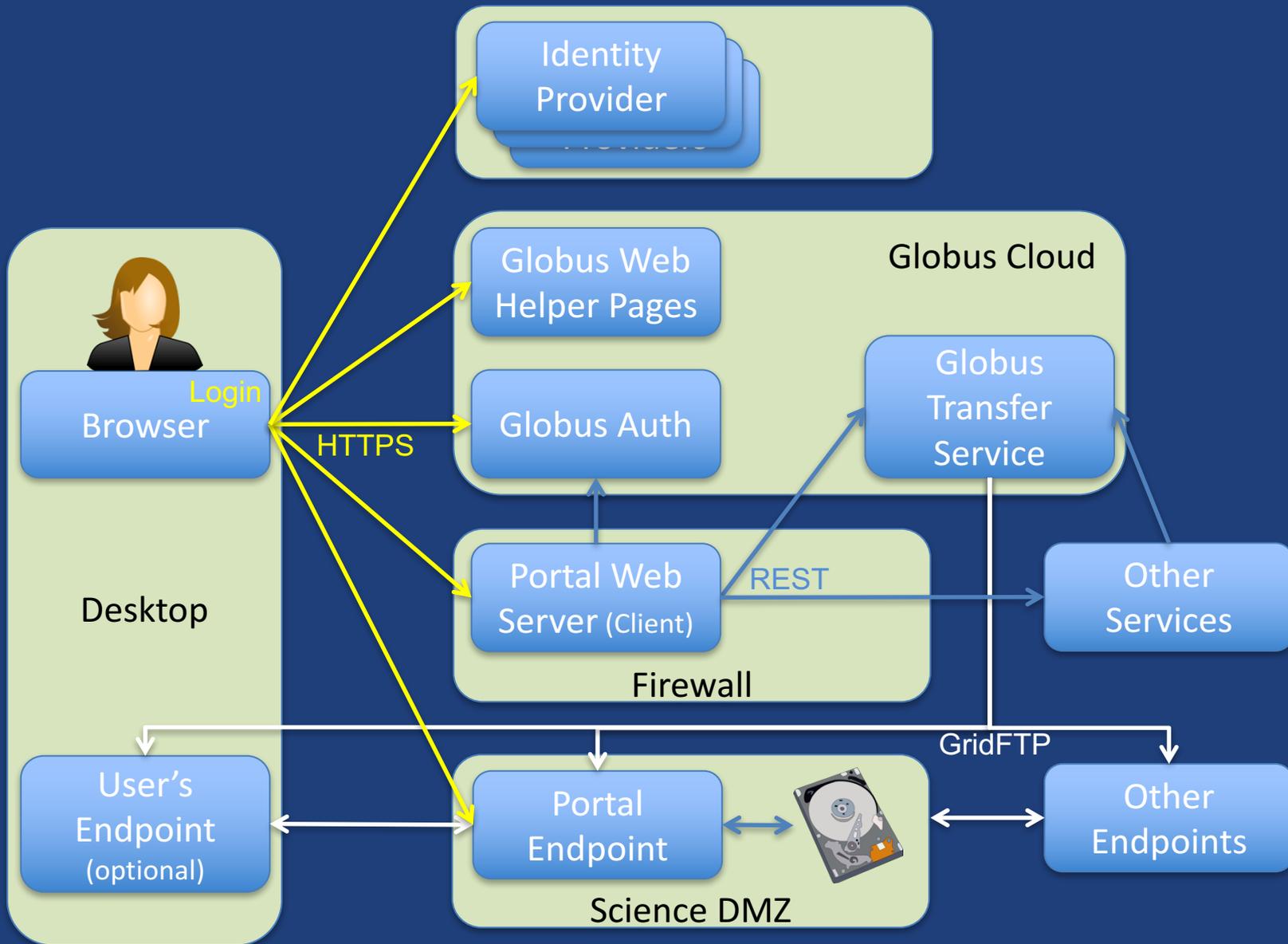
```
git checkout https://github.com/globus/globus-sdk-python.git
cd globus-sdk-python
python setup.py install
```

Basic Usage

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



Prototypical research data portal





Sample python data portal

Modern Research Data Portal LOGIN | SIGN UP

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...

<https://github.com/globus/globus-sample-data-portal>



GlobusWorld Developer Workshops

[Home](#)[Register](#)[Program](#)[Contact Us](#)

Welcome to the GlobusWorld Tour!

We're presenting a series of Globus tutorials and developer workshops across the US, building on the success of the [workshop held at GlobusWorld 2016](#). These workshops are made possible by the various hosting institutions that generously provide meeting space and other financial support.

The following workshops are currently scheduled:

- **September 13-14, 2016 - LBNL, Berkeley, CA**
- **October 12-13, 2016 - Yale University, New Haven, CT**
- **October 25-26, 2016 - NCAR, Boulder, CO**
- **Dates TBD - NIH, Washington D.C.**

If you would like to host a workshop at your institution please contact us.

Motivation: New high-speed networks make it possible, in principle, to transfer

Why Attend?

- Learn how the Globus platform simplifies development of web applications for researchers
- Experiment with new Globus services and APIs
- Exchange ideas with peers on ways to apply Globus technologies
- Expand your knowledge of Globus administration features

Workshops are free to attend and open to all,

<https://www.globusworld.org/tour/>



Globus Subscriptions

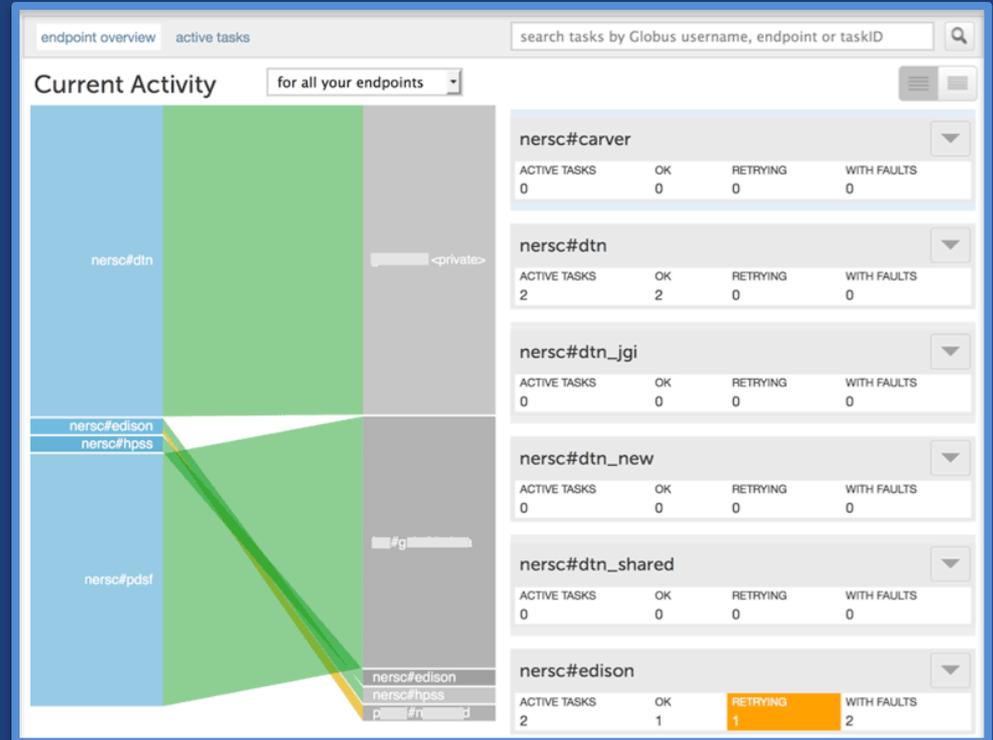
- **Globus Subscription**

- Shared endpoints
- Data publication
- Management console
- Usage reporting
- Priority support
- Application integration
- HTTPS endpoints

- **Branded Web Site**

- **Alternate Identity Provider (InCommon is standard)**

- **Premium Storage Connectors**



globus.org/subscriptions



Summary

- Globus SaaS delivers advanced capabilities for researchers and research computing centers
- Globus PaaS opens new opportunities for developers creating applications for researchers