

Globus: What's New

Rick Wagner
The University of Chicago





What is Globus?

Globus delivers...

Big data transfer, sharing,
publication, and discovery...

...directly from your own
storage systems...

...via software-as-a-service



Globus by the numbers

4

major services

218PB

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



No Globus usernames required

- **Globus users no longer require a Globus username & password**
 - Old Globus usernames moved to separate, 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 use any identity**
- **Used by XSEDE and others**



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*)



Account

 [Manage Your Consents](#) [Logout](#)

Identities

 [Manage These Identities](#) [Add Linked Identity](#)

Steven Tuecke via University of Chicago

Name: Steven Tuecke

Username: tuecke@uchicago.edu

E-mail: tuecke@uchicago.edu

Organization: University of Chicago

This is your **primary** identity.

Steve Tuecke via XSEDE

Name: Steve Tuecke

Username: tuecke@xsede.org

E-mail: tuecke@uchicago.edu

Organization: University of Chicago

This is a **linked** identity.

Globus Plus

Globus Plus gives you [enhanced features](#) on your Globus Connect Personal endpoints.

You do not currently have Globus Plus.

You can get Globus Plus as part of your organization's [Globus Provider subscription](#), by joining their Globus Plus group.

Select your organization:

[Didn't find your organization?](#)





Used by XSEDE and others

The screenshot shows the XSEDE homepage. The XSEDE logo is prominently displayed on the left. Below it, the text "Extreme Science and Engineering Discovery Environment" is visible. In the top right corner, there is a "globus" logo consisting of a stylized 'g' inside a cloud-like shape.

Jetstream Web App would like to:

Access all Jetstream resources

By clicking "Allow", you allow **Jetstream Web App** (no term information and services. You can rescind this and other co

Allow

Deny

The screenshot shows the Compute Canada Globus Web App login page. At the top, there is a decorative graphic of colorful arrows pointing upwards. The Compute Canada and calcul canada logos are displayed. The page includes a "Log in to use Compute Canada Globus Web App" section and a "Use your existing organizational login" section. A dropdown menu is set to "WestGrid". A "Continue" button is at the bottom, and a note at the very bottom says "Didn't find your organization? Then use Globus ID to [sign up](#)".



Share with any identity

Manage Endpoints

[◀ Back to Transfer Files](#) [Endpoint List](#)

[+ add Globus Connect Personal endpoint](#)

[+ add Globus Connect Server endpoint](#)

Tuecke Share on XSEDE Gordon
Shared Endpoint on xsede#gordon:/

ready
never expires

Overview

Permissions (3)

Roles

Manage Permissions for Tuecke Share on XSEDE Gordon

Host Endpoint: xsede#gordon

Host Path: /

principal	read	write	
view link for sharing			
Globus Team	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
Steve Tuecke (tuecke@globusid.org)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Steven Tuecke (tuecke@uchicago.edu)	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
Rachana Ananthakrishnan (rachana@nersc.gov)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	

Add New Permission for Tuecke Share on XSEDE Gordon

Path *

/

Share With *

user group all users public

Identity/E-mail *

Permissions

read write



Group membership with any identity

[Manage Data](#)[Publish](#)[Groups](#)[Support](#)[Account](#)[My Groups](#) | [Search for Groups](#) | [Create New Group](#)

My Groups (30)

▼ MANAGER OF (10)

My Admin Queue

Globus Team

- Globus Team - Development
- Globus Team - User Services
- Globus Team Plus Sponsor
- go#s3 access
- SciNet HPSS Admin Group
- Tuecke Lab
- Tuecke lab managers
- tuecke test 1
- UChicago Computation Institute

▼ MEMBER OF (20)

- BIRN Community
- CNM Contributors
- confluence-users-birn
- connect
- Globus Data Publication Service Admins
- Globus Publication Demo Submitters
- Globus Publish Allowed

[Create New Group »](#)

Globus Team

About Members (16) Subgroups (2) Settings

← Members Invite People to Globus Team

Search View Invites

enter username or search for user, and then click Add

Username: tuecke@uchicago.edu

Email invitation to: tuecke@uchicago.edu



Endpoint search



globus

Manage Data

Publish

Groups ▾

Support ▾

Account

Endpoint

Start typing the name of an endpoint or select one below



Cancel

Recently Used

★ Bookmarks

My Endpoints

[XSEDE TACC Wrangler](#)

[UChicago RCC Midway](#)

[globuspublish#trial_data](#)

[globuspublish#trialdata](#)

delete files on destination that do not exist on source

preserve source file modification times



Endpoint search



globus

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



Management Console

[Manage Data](#)[Publish](#)[Groups](#)[Support ▾](#)[Account](#)[Transfer Files](#) | [Activity](#) | [Endpoints](#) | [Bookmarks](#) | [Console](#)

Management Console

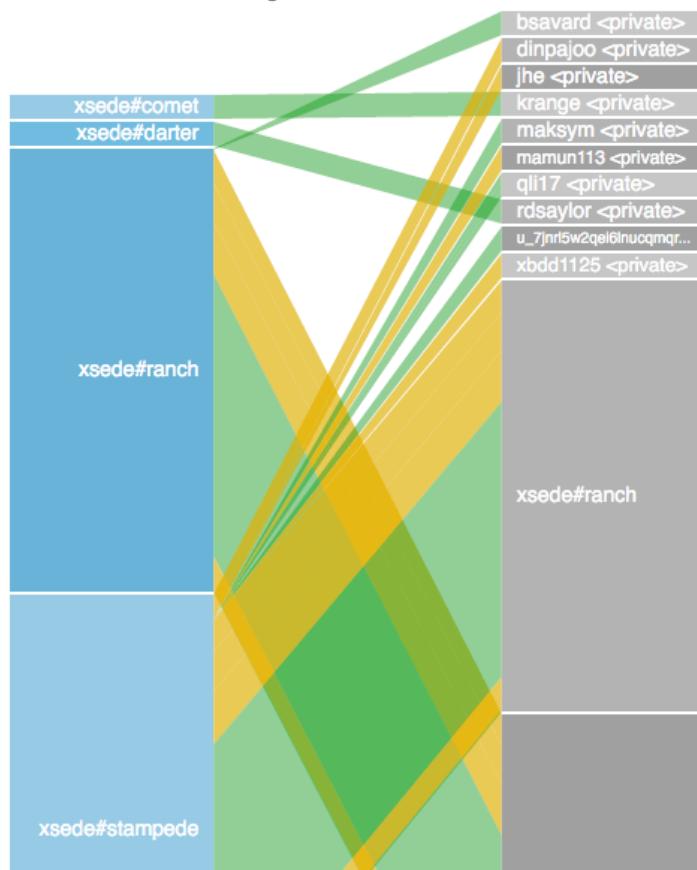
[endpoint overview](#)[active tasks](#)[pause rules](#)

search tasks by Globus username, endpoint or taskID



Current Activity

xsede# endpoints



xsede#comet

ACTIVE TASKS	OK	RETRYING	WITH FAULTS
1	1	0	0

xsede#darter

ACTIVE TASKS	OK	RETRYING	WITH FAULTS
1	1	0	0

xsede#keeneland

ACTIVE TASKS	OK	RETRYING	WITH FAULTS
0	0	0	0

xsede#lonestar

ACTIVE TASKS	OK	RETRYING	WITH FAULTS
0	0	0	0

xsede#mason

ACTIVE TASKS	OK	RETRYING	WITH FAULTS
0	0	0	0

xsede#nautilus



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

Alpha version: <http://globus.github.io/globus-cli/>

Issue tracker: <https://github.com/globus/globus-cli/issues>

```
[$ 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



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**



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: nnn.n

[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	
P	R	Web Server Holdings	Globus Transfer Service (GridFTP)	Subsetting	THREDDS Data Server	Central File System (GLADE) Holdings	Tape Archive (HPSS) Holdings
		Web File Listing	Request Globus Invitation	Get a Subset	TDS Access	GLADE File Listing	HPSS File Listing
		Web File Listing		Get a Subset		GLADE File Listing	HPSS File Listing
		Web File Listing		Get a Subset		GLADE File Listing	HPSS File Listing
		Web File Listing		Get a Subset		GLADE File Listing	HPSS File Listing

Union of Available Products

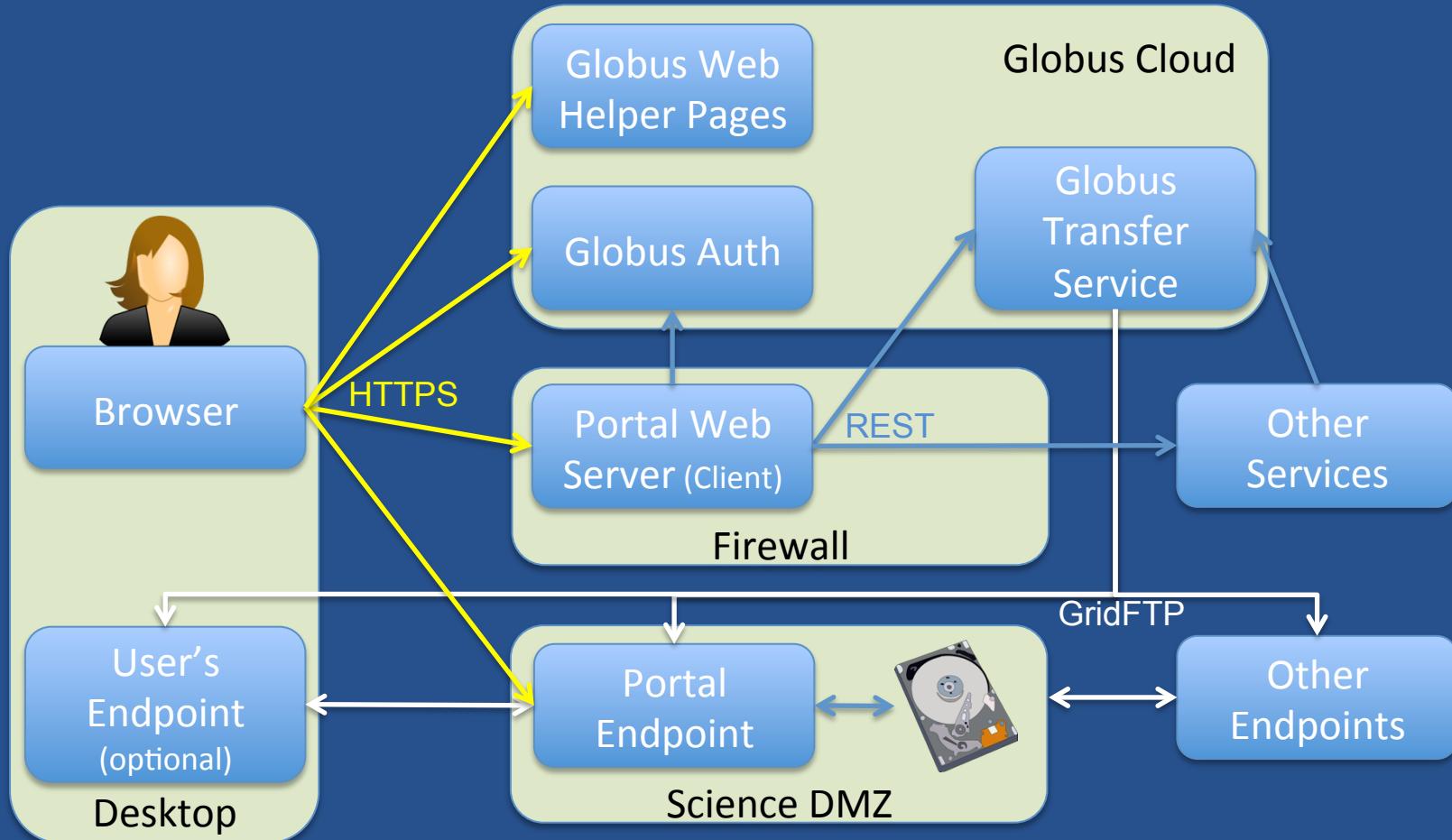
P Diurnal monthly means

R Regular monthly means

D Selected Parameter/Level Time Series



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 Transfer API

Nearly all Globus Web App functionality implemented via public Transfer API

The screenshot shows a web browser window with the following details:

- Header:** globus docs
- Top navigation bar:** APIs ▾, How To, Guides ▾, Support ▾, and a search icon.
- Breadcrumbs:** HOME / GLOBUS APIS / TRANSFER API DOCUMENTATION
- Title:** Transfer API Documentation
- Content Area:**
 - Last Updated: July 18, 2016
 - A detailed description of the Transfer API's REST-style interface and its capabilities.
 - Contents:** A list of API endpoints:
 - API Overview
 - Endpoint Activation
 - Task Submission
 - Task Monitoring
 - File Operations
 - Endpoint Management
 - Endpoint Search
 - Endpoint Roles
 - Endpoint Bookmarks
 - Endpoint ACL
- Right Sidebar:** A vertical sidebar with a list of API endpoints, where "Transfer API Documentation" is highlighted with a blue background.

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



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: InCommon, XSEDE, Google, portals
 - services: resource servers with REST APIs
 - apps: web, mobile, desktop, command line clients
 - services acting as clients to other services
- **Based on OAuth2 and OpenID Connect**
- **IAM platform for XSEDE (e.g., XUP, Jetstream, Jira)**

“A single global information space”



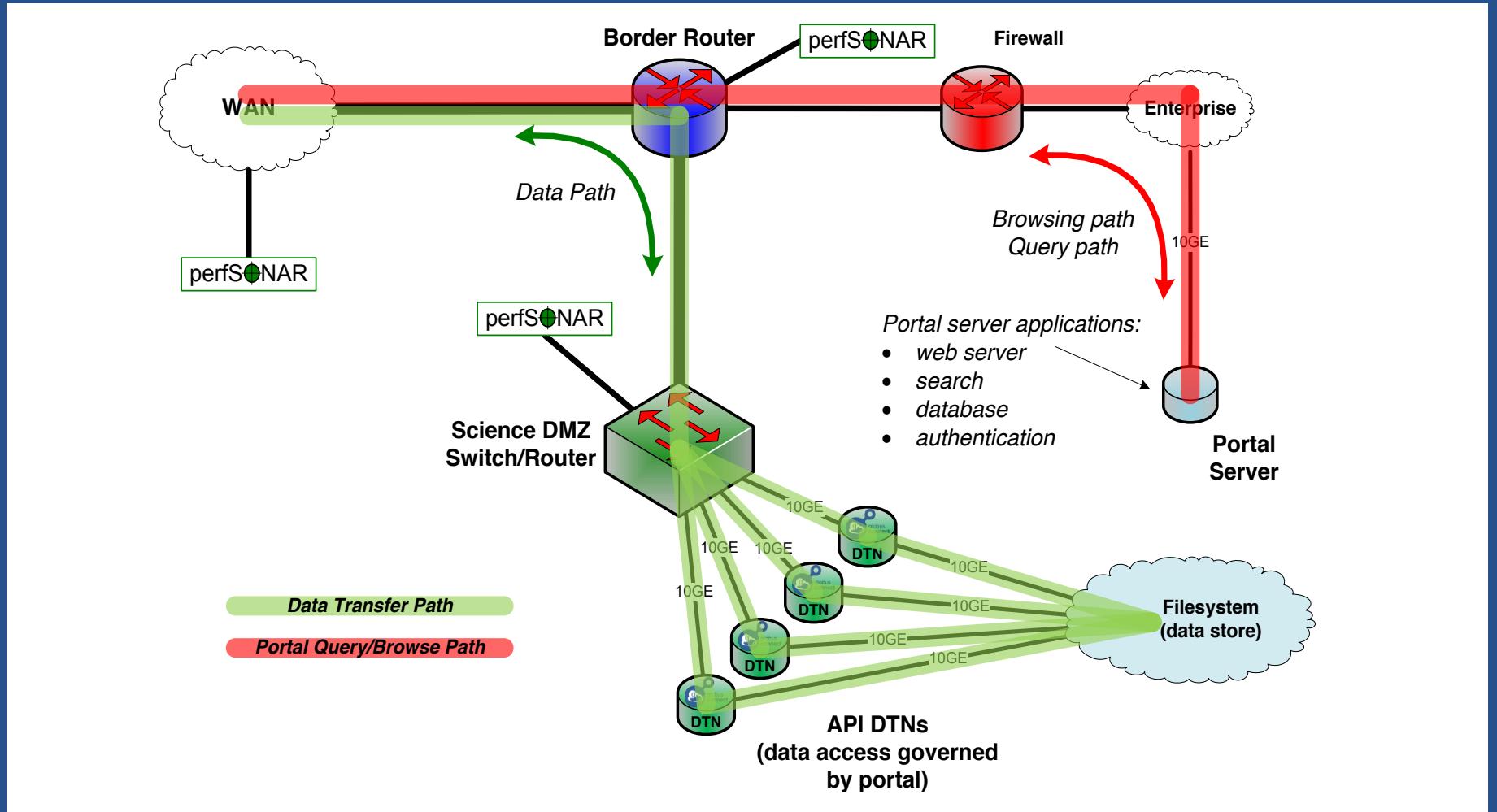


HTTPS Access to Endpoints (coming soon)

- **Enhanced use of research storage:**
 - asynchronous, bulk transfer via GridFTP
 - synchronous remote access via HTTPS
- **Enhanced Globus web app**
 - Browser-based upload/download
 - Inline file viewer
- **Integration with clients and web apps**



Next-Generation Portal Leverages Science DMZ



<https://fasterdata.es.net/>



docs.globus.org

[APIs](#)[How To](#)[Guides](#)[Support](#)

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](#)

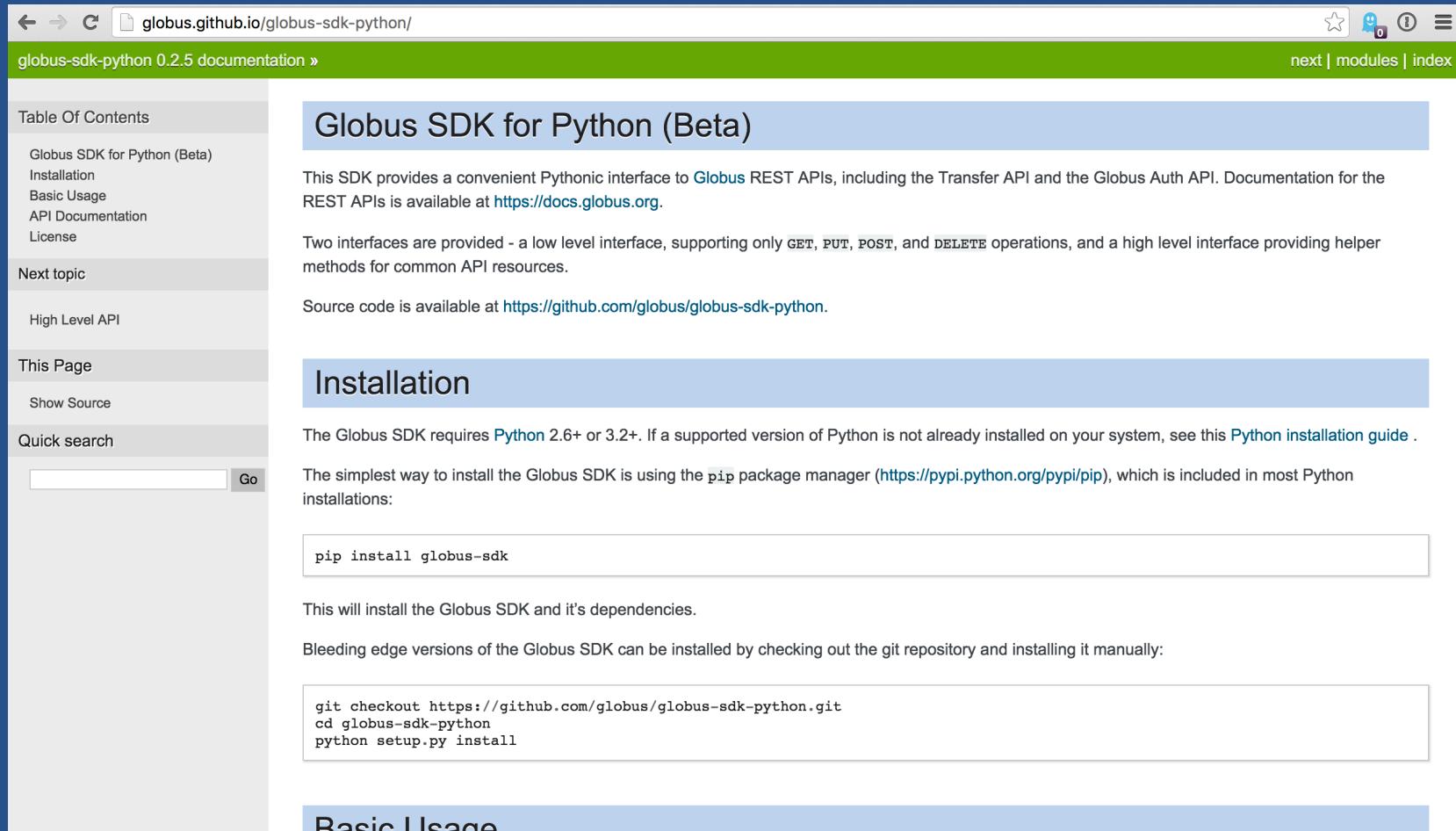
Toolkit

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 SDK Python



The screenshot shows a web browser displaying the [globus-sdk-python 0.2.5 documentation](https://globus.github.io/globus-sdk-python/). The page has a green header bar with the title and navigation links for 'next | modules | index'. On the left, there's a sidebar with a 'Table Of Contents' section containing links to 'Globus SDK for Python (Beta)', 'Installation', 'Basic Usage', 'API Documentation', and 'License'. Below that is a 'Next topic' section with a link to 'High Level API'. Further down are 'This Page' and 'Quick search' sections, along with a search bar and a 'Go' button.

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>



Sample python data portal



The screenshot shows the homepage of the Modern Research Data Portal. At the top left is the Globus logo. At the top right are 'LOGIN' and 'SIGN UP' buttons. The main title 'Modern Research Data Portal' is centered over a background image of blue DNA helixes. Below the title is the tagline 'It's how research data management is done!'. The page content area has a white background and contains three sections: 'Globus Transfer API' (API reference for transfer and sharing functions), 'Globus Auth API' (API reference for authentication and authorization), and 'Frequently Asked Questions' (When all else fails...).

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

The screenshot shows the homepage of the GlobusWorld Tour website. At the top, there's a decorative graphic with a blue circle containing a stylized 'g' logo, followed by the text 'The globusworld TOUR' with colorful dots and lines. Below this are four buttons: 'Home' (blue), 'Register' (grey), 'Program' (grey), and 'Contact Us' (grey). The main content area has a white background and features a large 'Welcome to the GlobusWorld Tour!' heading. It includes a paragraph about the workshops, a list of scheduled events, and a note for potential hosts. A 'Motivation' section follows. To the right, a 'Why Attend?' section lists benefits, and a blue callout box at the bottom right states: 'Workshops are free to attend and open to all, but we do require registration since'.

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**
- **Date TBD - NCAR, Boulder, CO**

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 and share research data at tremendous speeds and scales—but have also proved challenging to use in practice. Two new technologies now allow us to translate this potential into reality: Science DM² architectures provide frictionless end-

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, but we do require registration since

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



Summary

- Globus SaaS continues to improve its capabilities for researchers and research computing centers
- Globus PaaS opens new opportunities for developers creating web and mobile applications for researchers



Thank you to our sponsors!



U.S. DEPARTMENT OF
ENERGY



THE UNIVERSITY OF
CHICAGO



Argonne
NATIONAL LABORATORY



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



powered by
amazon
web services