

USING GLOBUS AT UCSD



Research Computing & Data (RCD) Community Meeting

June 21, 2022



Agenda

- Meeting Basics
- Welcome to the UCSD RCD Community!
- What is Globus?
Vas Vasiliadis, Globus
- UCSD-Specific Globus Resources
Rick Wagner, UCSD Research IT
- Advanced Examples from Campus
 - GenePattern, Ted Liefeld, Mesirov Lab
 - CReD Portal, Brian Yee & Bill West, Yeo Lab
- Community Discussion

Meeting Basics



Google Doc notes page in chat



Meeting will be recorded



Use the Zoom chat for questions, ideas, and suggestions



Or, Raise Hand
(Reactions > Raise Hand)



Live Transcript (auto-generated closed captions) is available



Meeting content (recording, notes, discussion, slides, chat log) will posted on meeting page





WELCOME TO THE UCSD RCD COMMUNITY!

It's about us!



Research Computing and Data (RCD), also known as Cyberinfrastructure (CI), refers to foundational research support areas. This community includes researchers and their collaboration partners, those RCD professionals who include data scientists, systems engineers, software developers, research librarians, data wranglers, research facilitators, and more.

Meetings



Our bi-monthly meetings are a chance to come together to discuss issues important to the UCSD research community, learn about resources, software, tools, and services relevant to our work.

<https://ucsd-rcd.github.io/meetings/>

Mailing List

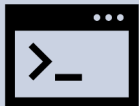


Our mailing list is a way to stay updated on the UCSD RCD Community, ask questions, and contribute your knowledge. We'd also love to hear your ideas for this community.

Google Groups [ucsd-rcd](https://groups.google.com/a/ucsd.edu/g/ucsd-rcd-l)
<https://groups.google.com/a/ucsd.edu/g/ucsd-rcd-l>

Website & GitHub Repos

Our website and GitHub repos are places to share resources, experiences, and examples in a way that's widely accessible.



<https://ucsd-rcd.github.io/>
<https://github.com/ucsd-rcd>

WHAT IS GLOBUS?

Vas Vasiliadis, Globus

UCSD-Specific Globus Resources

research-it@ucsd.edu

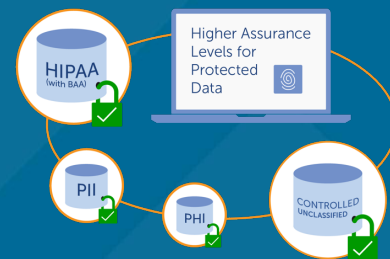
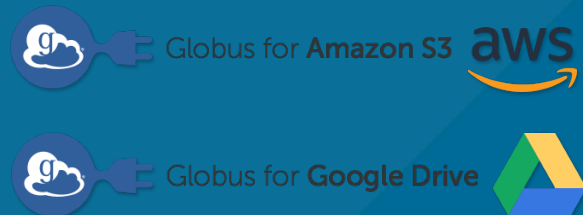
- Our Globus Subscription Features
- Existing GCS Instances
- Research IT DTN
- Project Guest Collections POC
- Cloud Storage Migration Testing



OUR GLOBUS SUBSCRIPTION FEATURES

research-it@ucsd.edu

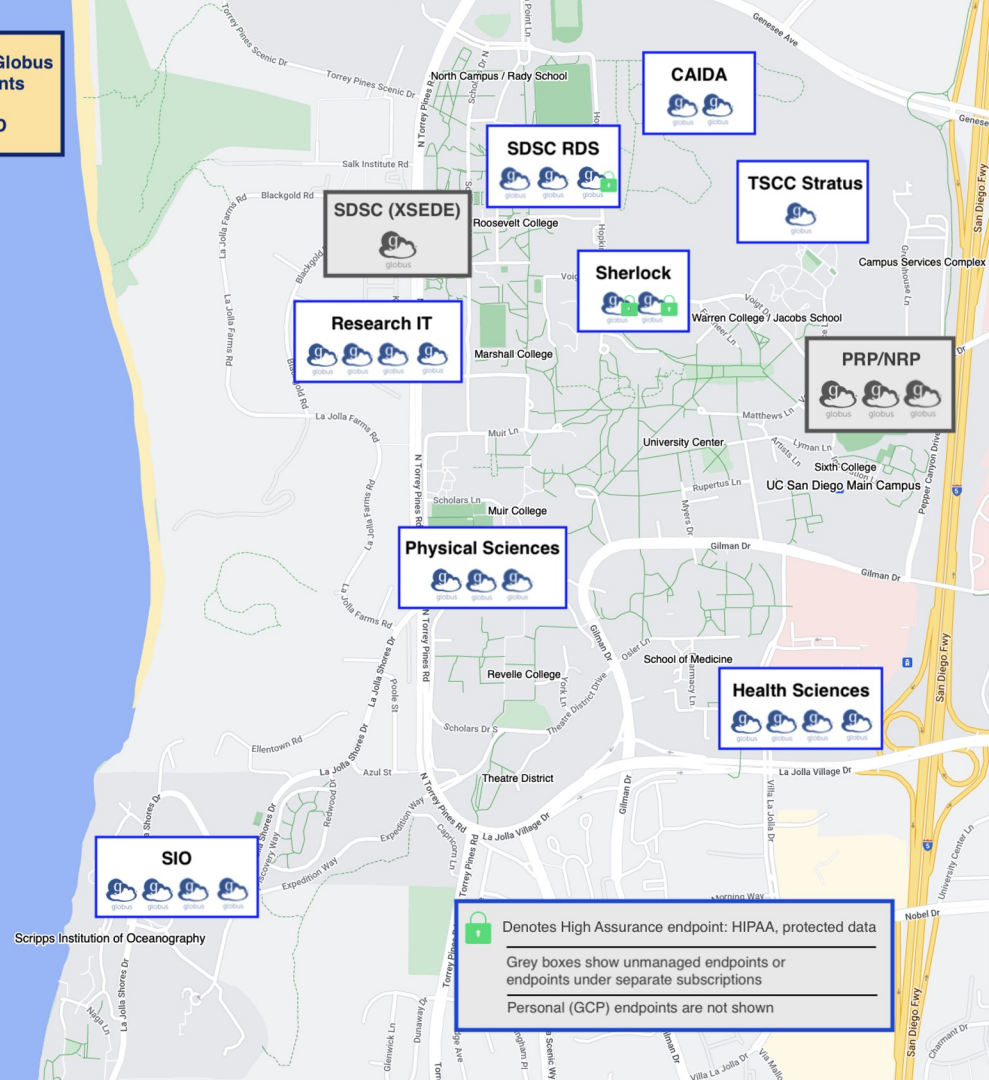
- Support! support@globus.org
- Managed GCS Instances
 - Guest Collections
 - HTTPS access to collections
 - Monitoring
- High Assurance
 - Use Globus for protected data
 - Globus Connect Server or Personal
 - Includes HIPAA
- S3 & Google Drive Connectors
 - <https://www.globus.org/connectors/s3>
 - <https://www.globus.org/connectors/google-drive>
- Globus Plus
 - Request to join “University of California San Diego Globus Plus” Group
<https://app.globus.org/groups/986ab968-a844-11e7-a508-22000b100078/about>
 - Transfer between Personal endpoints
 - Host Guest Collections on a Personal endpoint
 - Separate Globus Plus group for High Assurance



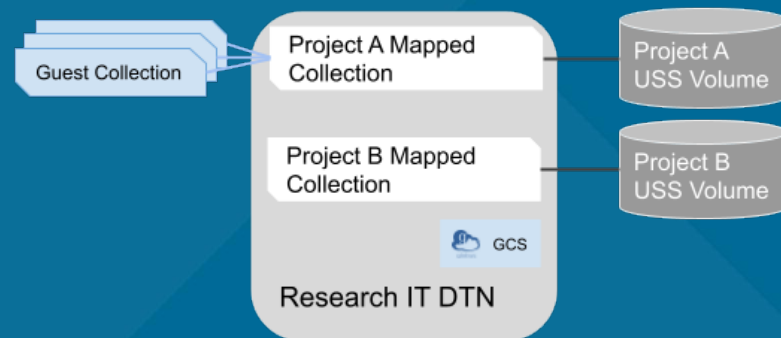
EXISTING GCS INSTANCES

- Several groups & departments already run GCS
- Try reaching out to your IT staff
 - There may already be an endpoint you can use
 - Or they may have experience setting up endpoints
- Contact us if you need help
research-it@ucsd.edu

Managed Globus Endpoints at UCSD



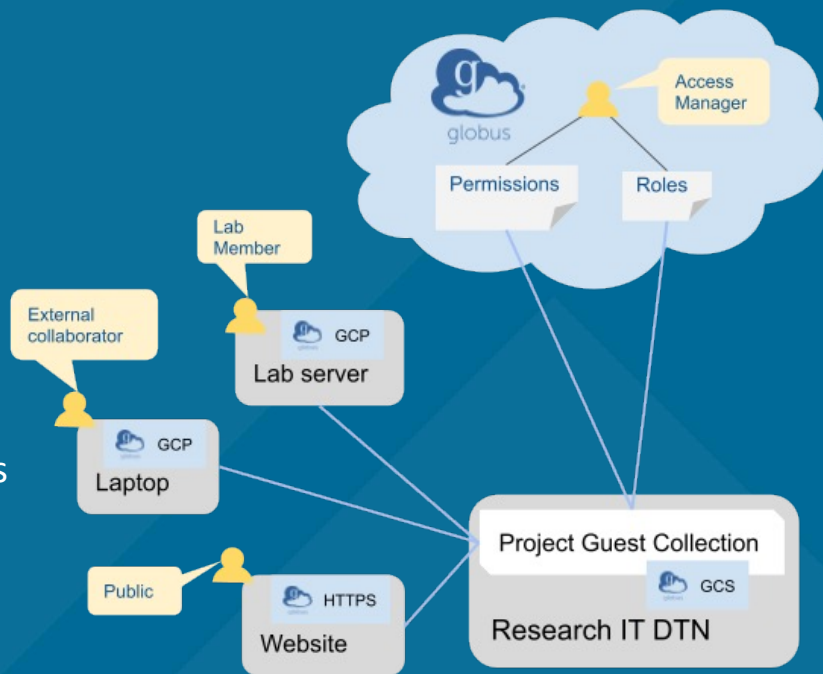
- Research IT operated Globus Connect Server
- Extends features of SDSC's Universal Scale Storage (Qumulo)
 - Includes chunks brokered through Research IT
- We work with you to mount your volume and set up
 - Identities
 - Permissions
 - Mapped Collections
 - Guest Collections
- No additional cost
- Convenient way to share data on USS with collaborators



PROJECT GUEST COLLECTIONS POC

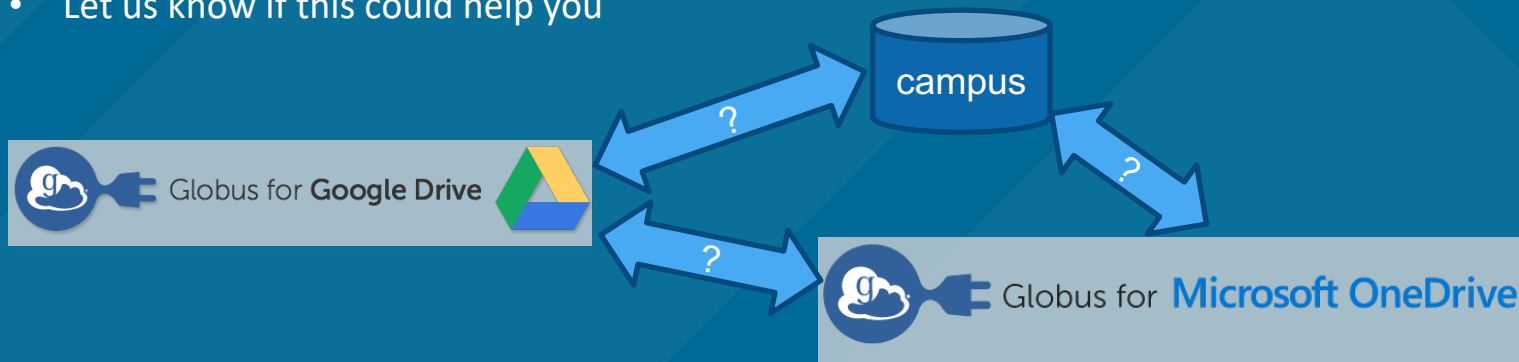
research-it@ucsd.edu

- Globus Guest Collections intended to promote data access
 - Lab members
 - Collaborators
 - Public
- We give you the space, you manage the permissions
- Supports HTTPS access to files
 - *Store data in collection, link to it from your site*
- Project PIs define Access Managers to manage permissions
- Request only require short project and usage descriptions
 - Default 1-year allocations of 500GB (may be expanded)
 - Priority for annual renewal will be given to active projects
 - Not intended to be backups or secondary copies of data
- Hosted by Research IT
- We can help projects if they start outgrowing the space



<https://collab.ucsd.edu/display/RESUP/Project+Guest+Collections+Proof-of-Concept>

- S3 connector is in use by several groups
- We're in the early phases of testing the Google Drive connector
- We have a trial subscription for the Microsoft OneDrive connector
- Want to identify critical use cases and limitations
 - Google Drive to campus systems?
 - Google Drive to OneDrive?
 - Vice versa?
 - One GCS instance with Google Drive for campus?
 - Separate ones for various departments or groups?
- Let us know if this could help you



ADVANCED EXAMPLES FROM CAMPUS

GenePattern
Ted Liefeld
Mesirov Lab

CRoD Portal
Brian Yee & Bill West
Yeo Lab

COMMUNITY DISCUSSION

Questions, ideas, suggestions!

Globus?

UCSD RCD Community?

Meeting Topics?

Needs?



Mailing List Reminder

Google Groups [ucsd-rcd](https://groups.google.com/a/ucsd.edu/g/ucsd-rcd-l)

<https://groups.google.com/a/ucsd.edu/g/ucsd-rcd-l>