iRODS®

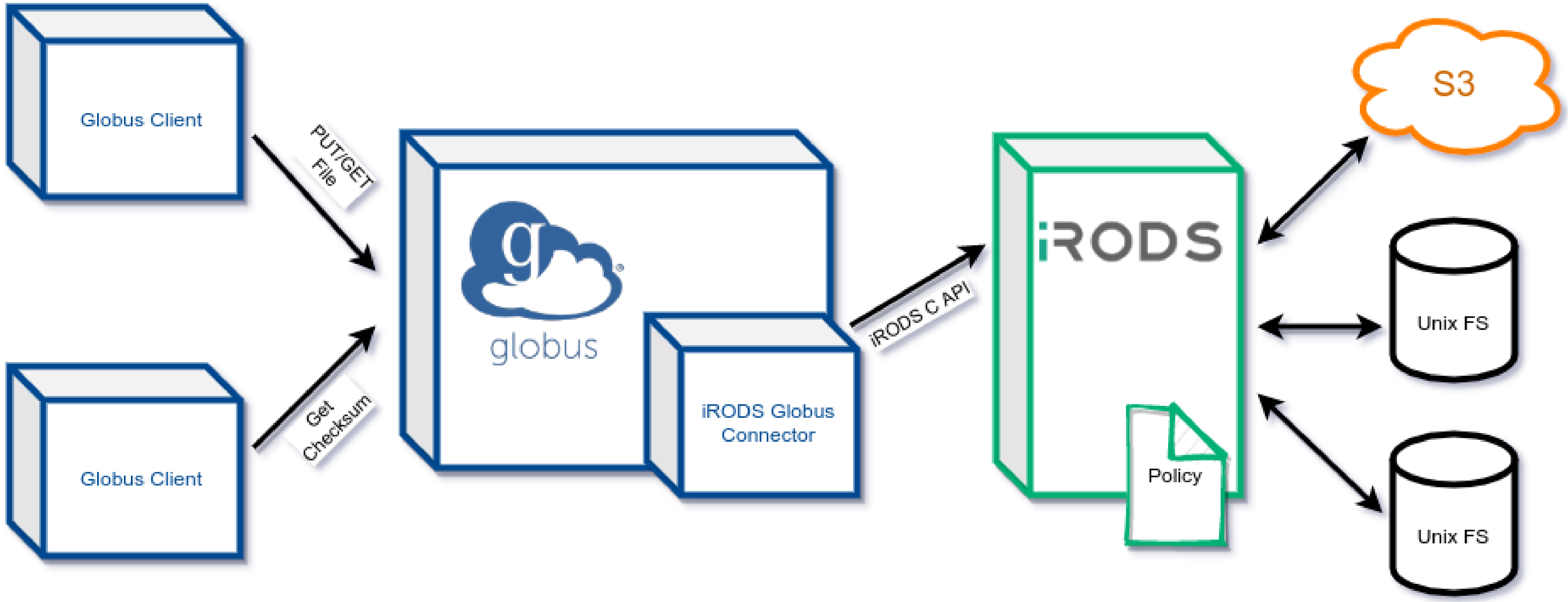
iRODS Globus Connector

Justin James
Applications Engineer
iRODS Consortium

June 8-11, 2021
iRODS User Group Meeting 2021
Virtual Event

iRODS Globus Connector - Overview

- The iRODS Globus Connector provides access to iRODS storage resources from a Globus Connect Server.
- The connector is a **plugin** for the Globus Connect Server and a **client** to the iRODS servers.
- All Globus requests use the iRODS C++ client APIs to get information and transfer data to iRODS



iRODS Globus Connector - Overview

This is built on the work from EUDAT:

https://github.com/EUDAT-B2STAGE/B2STAGE-GridFTP

We worked with Globus to implement a few enhancements and bug fixes.

iRODS Globus Connector - Scope of Talk

The majority of this talk is from the iRODS perspective

- Implementation details
- Plugin limitations
- Improvements
- iRODS APIs used by the plugin

iRODS Globus Connector - Globus Plugin Interface

The iRODS Globus Connector implements the **globus_gfs_storage_iface_t** interface.

The connector implements the following interface functions:

- INIT - Called when a new session is initiated. Reads the user environment and calls *clientLogin()*.
- DESTROY - Called at the end of the session. Cleans up and calls *rcDisconnect()*.
- SEND - Called when client requests to receive a file. Calls *rcDataObjOpen()/rcDataObjcLseek(),rcDataObjRead()/rcDataObjClose()*.
- RECEIVE - Called when client requests to transfer a file to the server. Calls *rcDataObjectOpen()/rcDataObjectCreate()/rcDataObjLseek()/rcDataObjWrite()/rcDataObjClose().*
- COMMAND - Called when a client sends a command to the server. (See next slide).
- STAT - Called when the server needs information about the object or collection. Calls *rcObjStat() or rclOpenColleciton/rclReadCollection()/rclCloseCollection()*.
- REALPATH

iRODS Globus Connector - Globus Client Commands

The following client commands are implemented.

- GLOBUS_GFS_CMD_MKD - Creates a collection in iRODS.   Calls *rcCollCreate()*.
- GLOBUS_GFS_CMD_RMD - Removes a collection. Calls *rcRmColl()*.
- GLOBUS_GFS_CMD_DELE - Deletes an object. Calls *rcDataObjUnlink().*
- GLOBUS_GFS_CMD_CKSM - Gets the checksum for an object. See next slide.

iRODS Globus Connector - Checksum/Hashing

Globus clients can request specific hashing algorithms. Originally the hashing algorithm that was calculated in iRODS was returned to the client. This did not necessarily match the algorithm the client had requested.

To support client requested hashing, the hash files are now calculated by the iRODS Globus Connector and stored in metadata as follows:

- AVU Name - Globus::<algorithm>
- AVU Value - <checksum value>
- AVU Units - POSIX time that the checksum was calculated

iRODS Globus Connector - Checksum/Hashing

The following pseudo-code describes roughly how a checksum request is served by the iRODS Globus Connector:

```
Receive request for hash for <alg1>
Query metadata for key(attr) = "Globus::<alg1>"
If exists:
    Compare AVU unit with update time.
    If update time <= unit value:
        Return checksum from metadata
    Else:
        Remove existing AVU for <alg1>
Download file and calculate checksum.
Write checksum to AVU.
Return checksum.
```

iRODS Globus Connector - Recent Enhancements / Fixes

- Support incremental directory listings for very large directories.
  - After either X number of entries are encountered or Y seconds have passed since the last partial listing, send additional entries via *globus_gridftp_server_finished_stat_partial()*.
  - Implemented heartbeats for long running checksum operations.
  - Implemented the realpath feature so that alternate paths don't allow users to bypass path restrictions.
  - Fixed some memory leaks in existing code.

iRODS Globus Connector - Future Enhancements

- Add the PEP to calculate checksums on data object updates.
  - In the policy the administrator would specify which checksums should be calculated.

- Investigate and fix other performance issues.

iRODS Globus Connector - Configuring Globus

Refer to the following page for more information on configuring Globus to be used with the iRODS Globus Connector.

https://docs.globus.org/premium-storage-connectors/v5/irods/

Questions?