iRODS speaks SFTP: More ways to securely transfer your data

Illyoung Choi  
CyVerse / University of Arizona  
iychoi@arizona.edu

Edwin Skidmore  
CyVerse / University of Arizona  
edwin@cyverse.org

Nirav Merchant  
CyVerse / University of Arizona  
nirav@arizona.edu

July 6, 2022  
iRODS User Group Meeting
Problem statement

Science domains working with confidential data require:

• Encrypted data transfer
• Secure user authentication
• Well known & validated client software
• Easy to integrate into existing research workflows
SFTP is a solution

Meet the requirements:

- Based on SSH Protocol, supporting various encryption algorithms
  - AES (128/256/192, GCM/CTR/CBC), 3DES CBC, ARCFOUR (128/256) ...
- Public-key authentication
- Well known, validated SFTP clients
  - OS built-in SFTP client, Curl, FileZilla, Cyberduck, ETC...
- Users & administrators are already familiar with SFTP
SFTPGo

• An open-source SFTP server written in pure Go

• Various backend storage
  • Local disk, S3, Azure Blob Storage, GCS, other SFTP, ETC...

• Built-in defender
  • IP filter, blocking brute-force login trial

• Additional data transfer protocols
  • FTP, SCP, WebDAV

• Easy integration to other services
  • Admin REST API
  • External authentication module
SFTPGo for iRODS

1. Login attempt
   - User

2. User authentication
   - via password: perform iRODS auth.
   - via public-key: login via proxy mode, check authorized_keys. (/iplant/home/username/.ssh/authorized_keys)

3. Create user & VFolder mappings
   - iRODS user home
   - Community shared

4. User authentication via proxy mode

5. Data access

Data Flow

Authentication Flow
“/zone/home/username/.ssh/authorized_keys”

- A text file containing a list of public keys, located in “.ssh” dir. under user home
- Options in a comma-separated string
  - “expiry-time=YYMMDD[HHMM]”: Set an expiration time
  - “from=98.165.125.0/24,!98.165.125.240”: Set an IP filter
  - “home=/zone/home/username/subdir”: Set a home directory (collection)

expiry-time="20221231",home="/iplant/home/iychoi/device1",from="98.165.125.0/24" ssh-rsa AAAAB3N...
from="98.165.125.0/24" ssh-rsa AAAE2Vj...
I/O Performance of backend storage modules

iRODS backend storage module requires data transfer over LAN, while local fs module doesn’t.
I/O Performance of iRODS clients

- iCommands: 58.87 MB/s (Upload), 132.93 MB/s (Download)
- WebDAV (Davrods): 68.70 MB/s (Upload), 28.43 MB/s (Download)
- SFTPGo: 52.80 MB/s (Upload), 48.07 MB/s (Download)
Quick Demo: Password-based authentication
Quick Demo: Public-key authentication
Applications

Health Science Projects (CyVerse Health)
- Encrypted data transfer between sites and services
- Integration to existing workflows

Agriculture Project
- Assign different public-key auth. to each Raspberry Pi, generating sensor data
- Set a home directory for each device (cannot access other directory when bleached)
- SFTP is already supported in Python or via SFTP CLI
Deployment plan

Alpha release (May 18th)
• Only accessible from UA network or via VPN
• May undergo maintenance any time

(Open) Beta release (August 2nd, Subject to change)
• Domain name change to “data.cyverse.org”
• No access restriction
• May undergo maintenance any time when issues occur

Release (To be determined)
• Scheduled maintenance (same day as Data Store maintenance day)
Open-source

Backend storage module for iRODS

- [https://github.com/cyverse/sftpgo](https://github.com/cyverse/sftpgo)
- Forked from [https://github.com/drakkan/sftpgo](https://github.com/drakkan/sftpgo)
- Added iRODS backend storage module

External authentication module for iRODS

- [https://github.com/cyverse/sftpgo-auth-irods](https://github.com/cyverse/sftpgo-auth-irods)
- A simple command-line program, working with JSON data via STDIN/STDOUT

Deployment scripts & configuration

- [https://github.com/cyverse/sftpgo-deploy](https://github.com/cyverse/sftpgo-deploy)
- Deploy using docker-compose
Conclusion

Provide SFTP access to iRODS
  • The service runs separately from iRODS (no configuration change on iRODS)
  • Secure authentication and data access

Implemented public-key authentication
  • Implemented on top of iRODS proxy authentication
  • Provide “options” for access control

To be used in science projects
  • Health science projects
  • Agriculture project
  • ...
Questions?