

iRODS

Technology Update

Terrell Russell, Ph.D.
@terrellrussell
Chief Technologist, iRODS Consortium

June 5-7, 2018
iRODS User Group Meeting 2018
Durham, NC

iRODS Release	Issues Closed
4.1.11	92
4.2.2	39
4.2.3	76 + 18

Plugins

- Python Rule Engine Plugin
- Storage Tiering Rule Engine Plugin
- Auditing (AMQP) Rule Engine Plugin
- S3 Resource Plugin
- GSI Authentication Plugin
- Kerberos Authentication Plugin
- Curl Microservice Plugin

Clients

- Python iRODS Client
- Automated Ingest Framework
- Cloud Browser

In The Last Year

```
~/irods $ git shortlog --summary --numbered 4.1.10..4-1-stable
```

```
31 Ben Keller
20 Alan King
20 Terrell Russell
7 Hao Xu
7 Rick Skarbez
7 jkgill
6 Zoey Greer
5 Jason Coposky
2 Matthew Vernon
1 Brian N. Makin
```

```
~/irods $ git shortlog --summary --numbered 4.2.1..4-2-stable
```

```
51 Alan King
27 Jason Coposky
22 Hao Xu
18 Terrell Russell
17 Ben Keller
17 Zoey Greer
15 Kory Draughn
9 jkgill
8 Justin James
5 Andrew Kelly
5 Kyle Ferriter
2 d-w-moore
2 xu-hao
1 Felix A. Croes
1 Marco Grossi
1 Sam Morris
```

- iRODS 4.1.12
- iRODS 4.2.4
- iRODS 4.3.0
- Python iRODS Client (PRC)
- Indexing Framework
- Lustre Integration
- Metalnx
- CockroachDB Database Plugin
- Ceph RADOS Resource Plugin
- QueryArrow Database Plugin
- R Client Library (R-iRODS)
- Multipart Transfer, v5 API
- NFS4J Integration
- Metadata Templates
- Testing Infrastructure

4.1.12 Update

Replicas in a complex tree (across 2 zones)...

Not enough replicas! [#3746](#)

Problem: 1 copy == BAD

Cause(s): Network, storage, OS...

Solution: Try harder

[Replication retry](#) (tested with [mungefs](#))

Too many replicas! [#3665](#)

Problem: Filling up storage

Cause: Surprise concurrency? Repl resc?

Solution: ???

["Enhanced logging"](#) with PID/user info

To be continued...

```

root:passthru
└─ replicate:replication
   └─ green-rand:random
      │   └─ green0:random
      │       │   └─ green0ufs0
      │       │   └─ green0ufs1
      │       └─ green1:random
      │           │   └─ green1ufs0
      │           └─ green1ufs1
      └─ red-rand:random
          └─ red0:random
              │   └─ red0ufs0
              │   └─ red0ufs1
              └─ red1:random
                  │   └─ red1ufs0
                  └─ red1ufs1
  
```

- irodsMonitor process
- Checksums moving down into resource plugins
- Containerize installation
- Logging overhaul

4.3.0 Logging Update

Today

- Inconsistent
- Incomplete (syslog support)
- Not very helpful in tracking what caused an error
- Not very helpful when multiple servers are involved

Design Goals

- Leverage existing logging library (spdlog, etc.)
- Enable admins to easily capture, process, and analyze activity
- Consistent
- Easily track errors across multiple servers (hostname, timestamp, PID, plugin, etc.)
- Tie in existing infrastructure
- Provide more options for controlling output

Join Us!

- What would you like to see accomplished with the logging refactor?
- RFC: https://github.com/irods/irods_rfcs/blob/master/0001_logging.md
- Have questions/suggestions? Let's talk!

July 2011

- Python → Node.js → RabbitMQ → Celery → Eucalyptus

October 2012

- Python → Node.js → ssh → OpenStack

January 2013

- Hudson → Python → OpenStack

October 2013

- Hudson → Python → vSphere long-running VMs

Spring 2015

- Jenkins → Python → Ansible → vSphere dynamic VMs

Spring 2015 - onwards

- Jenkins → Python → Ansible → zone_bundles → vSphere dynamic VMs

Changes Since 2017

- Ended support for CentOS 6 and Ubuntu 12
- Moved iRODS build/test logic from Ansible to per-repository python modules
- Unified Jenkins workflow to test all supported plugins
- Consolidated to two parameterized Jenkins jobs
 - run-script-on-vm
 - run-script-on-irods-zone

iRODS Build and Test - Today

Jenkins

2

search

Jaspreet | log out

ENABLE AUTO REFRESH

add description

1. Main 2. Personal All UGM-2018 +

S	W	Name ↓	Last Success	Last Failure	Last Duration
		build_and_test_all_plugins_workflow	6 hr 25 min - #16	4 days 16 hr - #15	6 hr 25 min
		irods-build-and-test-workflow	9 hr 21 min - #1035	4 days 16 hr - #1028	7 hr 23 min
		plugin-auth-gsi-test-workflow	23 min - #23	1 hr 5 min - #22	9 min 2 sec
		plugin-auth-kerberos-test-workflow	1 hr 5 min - #20	10 days - #16	28 min
		plugin-microservice-curl-build-test-workflow	1 hr 5 min - #16	4 days 16 hr - #14	8 min 25 sec
		plugin-resource-s3-test-workflow	13 days - #16	1 hr 5 min - #26	2 hr 28 min
		python-irodscient-coverage-and-test-results	1 day 18 hr - #2973	N/A	0.19 sec
		python-irodscient-github-tracking-master	1 day 18 hr - #85	3 days 2 hr - #84	17 min
		python-irodscient-test-workflow	1 day 18 hr - #303	3 days 2 hr - #302	17 min
		rule-engine-plugin-audit-amqp-build-test-workflow	1 hr 15 min - #68	1 hr 5 min - #67	26 min
		rule-engine-plugin-audit-amqp-github-tracking-master	1 hr 15 min - #23	9 days 22 hr - #22	26 min
		rule-engine-plugin-python-build-test-workflow	4 days 16 hr - #55	5 days 16 hr - #51	6 hr 25 min
		rule-engine-plugin-python-github-tracking-master	1 day 0 hr - #16	11 days - #15	5 hr 28 min
		rule-engine-plugin-storage-tiering-build-test-workflow	1 hr 5 min - #180	11 days - #171	32 min
		rule-engine-plugin-storage-tiering-github-tracking-master	1 hr 36 min - #73	11 days - #71	32 min
		run-script-on-irods-zone	19 min - #5207	1 hr 1 min - #5193	4 min 55 sec
		run-script-on-ums	4 min 3 sec - #2137	4 days 16 hr - #2102	3 min 42 sec

Build Queue

No builds in the queue.

Build Executor Status

1 Idle
2 Idle
3 Idle
4 Idle
5 Idle
6 Idle
7 Idle
8 Idle
9 Idle
10 Idle
11 Idle
12 Idle
13 Idle
14 Idle

iRODS Build and Test - Future

- Increase coverage (more plugins in CI)
- Move pipeline scripts to GitHub (no logic in Jenkins)
- Address inconsistency (false reds / pyvmomi errors)
- Containerize Jenkins (easier to test / update / redeploy)
- Possibly move from VMs to containers (speed / fewer moving parts)

Testing a distributed system is hard.

Input and other ideas are welcome.

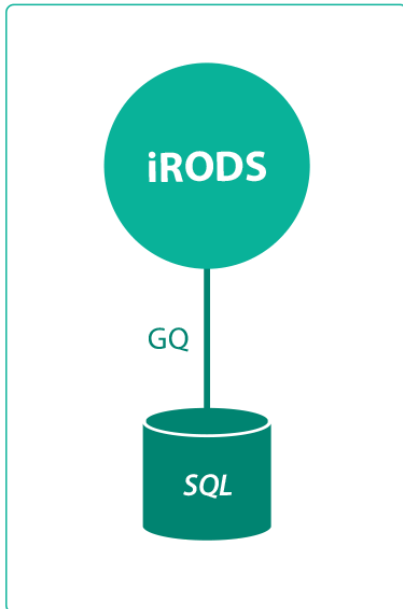
4.2 is about sending things out

- Dynamic PEPs
- Pluggable Rule Engine
- Enterprise Messaging

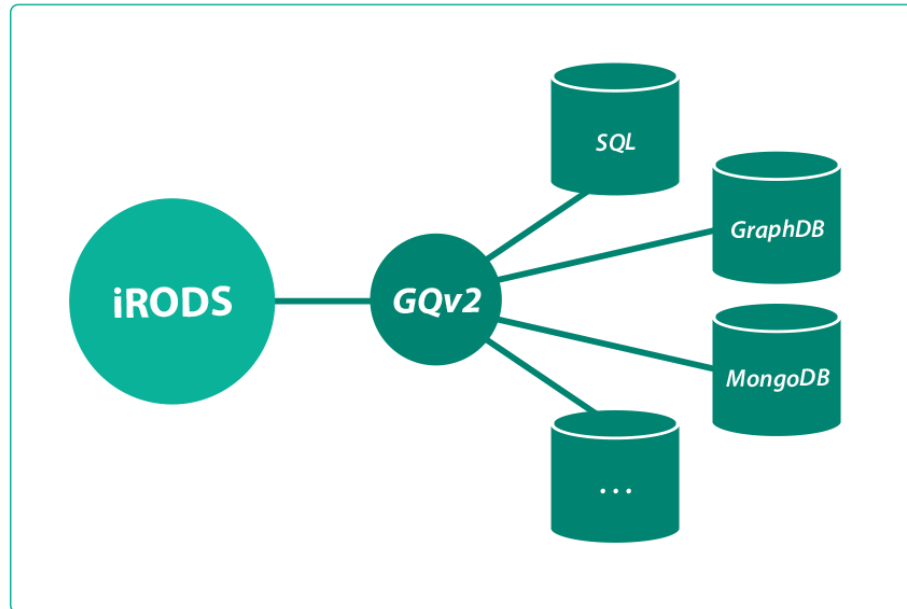
4.3 (and friends) will be about bringing them back

- Facilitated by GQv2

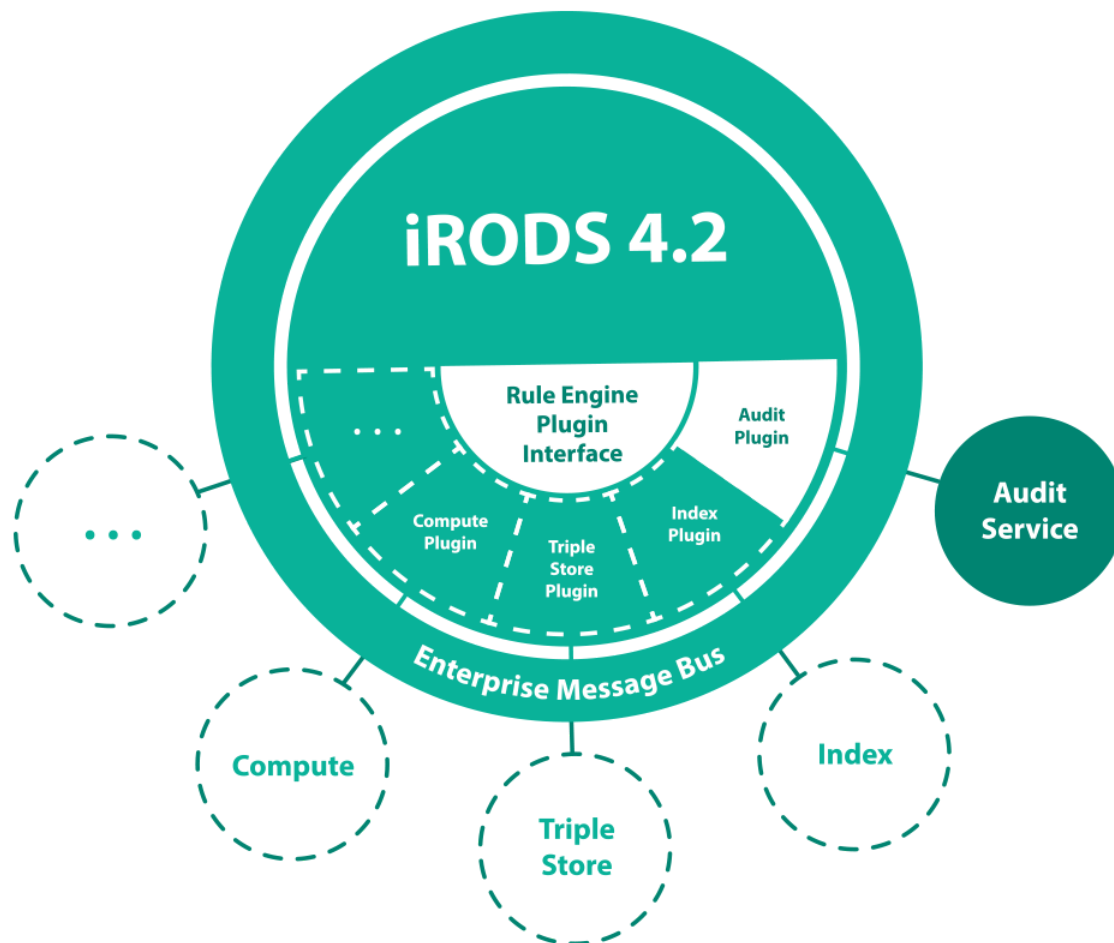
Current GQ



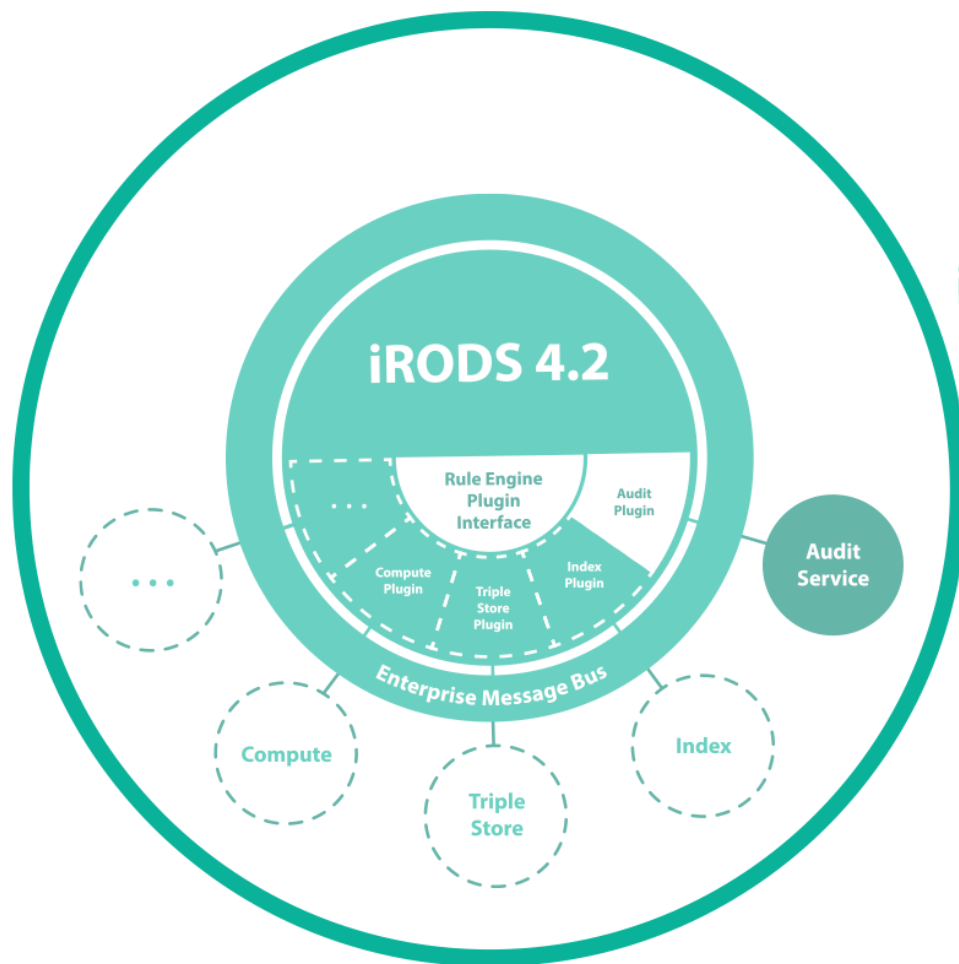
Future GQv2



Enterprise Messaging - Scatter

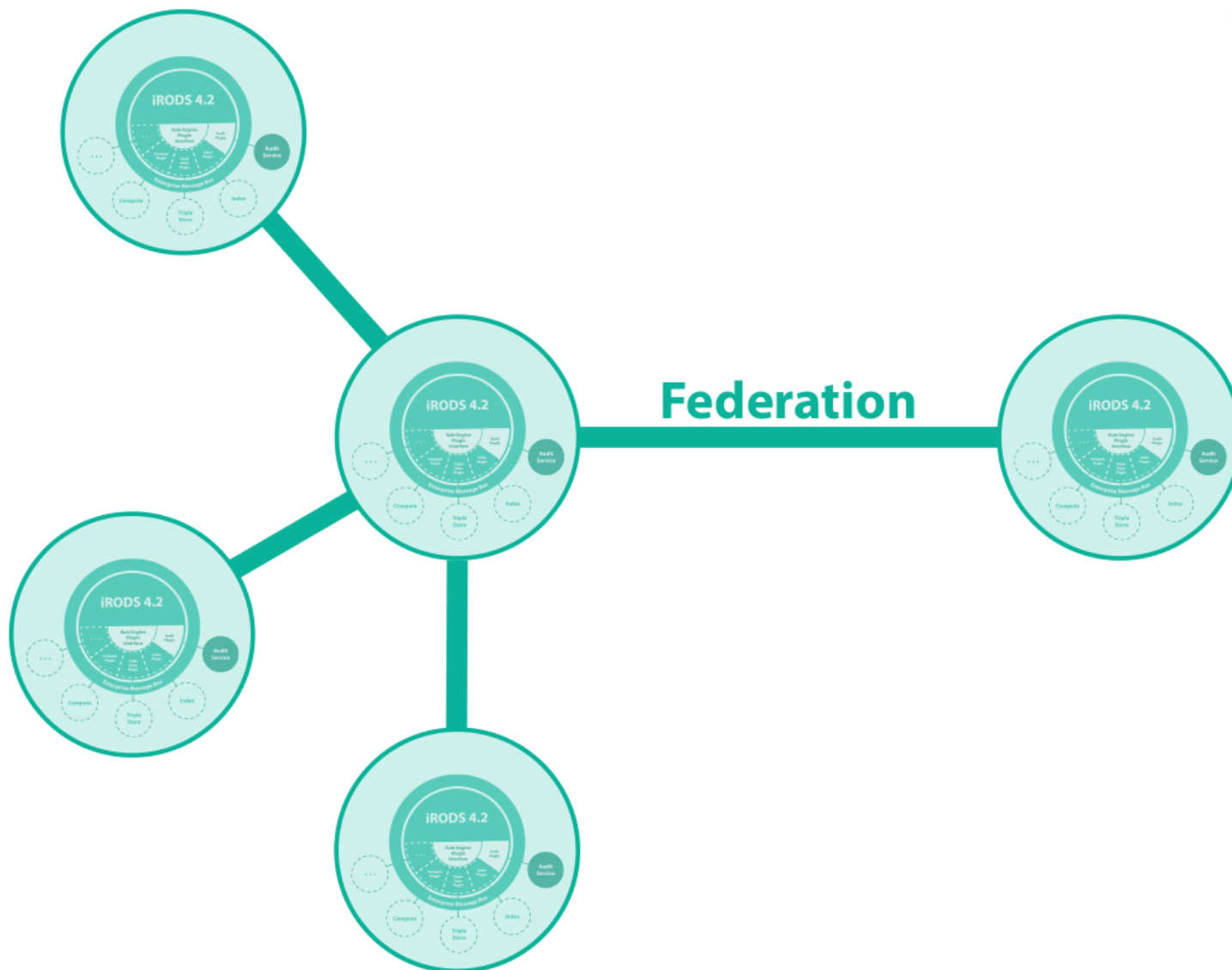


Metadata Virtualization - Gather



iRODS 4.3

GQv2 - Suite of Plugins
Virtualizing Metadata



- Technology Working Group
- Metadata Templates Working Group
- GitHub Issues
- Pull Requests
- Chat List
- Consortium Membership