Deployment of a National Research Data Grid Powered by iRODS

Ilari Korhonen
PDC Center for High Performance Computing, KTH

10th iRODS Users Group Meeting, June 6th 2018, Durham, NC
SNIC Storage and iRODS

- In the spring of 2017 SNIC (Swedish National Infrastructure for Computing) decided to fund the deployment of iRODS storage into its national distributed storage infrastructure (called Swestore)
- The previous generation of Swestore is based on dCache
- SNIC will be supporting both platforms in production
- Funding decisions for 2 x 1 PB of storage systems to be placed in PDC at KTH and NSC at Linköping University
- Procurements done, delivered and at PDC deployed and in production, and at NSC being deployed at this moment.
- High performance filesystem (GPFS) at PDC - landing zone
Software Stack

• iRODS (version 4.1.11) - upgrade to 4.2.3 imminent
• PostgreSQL 9.4 w/ streaming replication

• CentOS 7 (some older servers still running CentOS 6)
• Davrods for WebDAV and anonymous access via http
• MetaLnx for Web UI
• Kanki as an iRODS native client

• FreeIPA for IdM with LDAP and Kerberos V5 (also Heimdal at PDC)
• Python iRODS Client for integration scripts

• ZFS on Linux
• IBM Spectrum Scale (GPFS)
• IBM Spectrum Protect (TSM)

• Git and GitHub for repositories
• Sphinx for documentation
• Vagrant and Ansible for deployment and testing
Geographically Distributed Data Grid

- For cost-effective high availability and disaster recovery, the two supercomputing centers PDC and NSC are operating the data grid in collaboration with two administrative domains.
- The physical distance between the centers is ~ 130 miles.
- Our iRODS catalog services provider (a.k.a iCAT) is hot-standby and replicated across the two centers, via PostgreSQL streaming replication (async, there is latency).
- Also, the storage resources which iRODS manages are replicated as well, via iRODS (async).
- Swedish University Network (SUNET, SunetC) based on a 100 Gbit/s backbone (dual-ring topology).
Initial Results for (Long-Distance) Transfers

• PDC (Stockholm) <-> LUNARC (Lund) ~ 1.0 GB/s avg.
  • Physical distance ~ 370 miles, latency ~ 8.7 ms
  • 10 Gbit/s link speed at transfer node

• PDC (Stockholm) <-> NSC (Linköping) ~ 2.0 GB/s avg.
  • Physical distance ~ 130 miles, latency ~ 3.4 ms
  • 40 Gbit/s backbone at NSC

• Locally at PDC over 100 GbE (no routing) up to 8 GB/s
  • reading from GPFS, writing to GPFS (via iRODS)
  • GPFS via 100 Gbps EDR InfiniBand
Rollout into Production

- Had to be done in phases, wasn’t possible to do everything at once (not because of iRODS of course)
- Data migration from legacy systems, one at PDC and also another one at NSC

- Legacy data (migrated) - 96.6 TiB total
  1) 3,184,073 data objects (54.7 TiB) - NSC
  2) 2,551,581 data objects (41.9 TiB) - PDC

- One round of applications opened to researchers, more to come after summer holiday season
- New applications have been submitted and accepted
Data Migration via iRODS

- We had to migrate some research groups and users from our old EUDAT iRODS instance at PDC to the new SNIC iRODS - old PDC EUDAT instance decommissioned
- Since we are running federated zones, data migration can of course be done fully via iRODS native mechanisms

**EUDAT**

$ iadmin mkuser rods#snic.se rodsuser
$ ichmod -rvM own rods#snic.se /eudat.se/home
$ ichmod -rvM own rods#snic.se /eudat.se/projects

**SNIC**

$ irsync -Krv -R eudat-migration i:/eudat.se/projects
   i:/snic.se/migration/eudat.se/projects
Thank You

For more information, please do not hesitate to contact us!

Ilari Korhonen <ilarik@kth.se>

SNIC iRODS Team contact information (and my thanks go to the names below):

PDC
- Dejan Vitlacil <vitlacil@kth.se>
- Ilker Manap <manap@kth.se>

NSC
- Janos Nagy <fconagy@nscliu.se>
- Krishnaveni Chitrapu <krishnaveni@nscliu.se>