iRODS at CC-IN2P3: overview

Jean-Yves Nief

Talk overview

- **CC-IN2P3 activity.**
- **iRODS in production:**
  - Hardware setup.
  - Usage.
  - Prospects.
- **iRODS developpements in Lyon:**
  - Scripts.
  - Micro-services.
  - Drivers.
  - Resource Monitoring System.
  - iCommand.
- **SRB to iRODS migration.**

CC-IN2P3 activities

- Federate computing needs of the french scientific community in:
    - Nuclear and particle physics.
    - Astrophysics and astroparticles.
- Computing services to international collaborations:
    - CERN (LHC), Fermilab, SLAC, ….
- Opened now to biology, Arts & Humanities.

iRODS setup @ CC-IN2P3

- 9 servers:
  - 3 iCAT servers (metacatalog): Linux SL4, Linux SL5
  - 6 data servers (200 TB): Sun Thor x4540, Solaris 10.
- Metacatalog on a dedicated Oracle 11g cluster.
- HPSS interface: rfio server (using universal MSS driver).
- Use of fuse-iRODS:
  - For Fedora-Commons.
  - For legacy web applications.
- TSM: backup of some stored data.
- Monitoring and restart of the services fully automated (*crontab* + *Nagios* + *SMURF*).
- Automatic weekly reindexing of the iCAT databases.
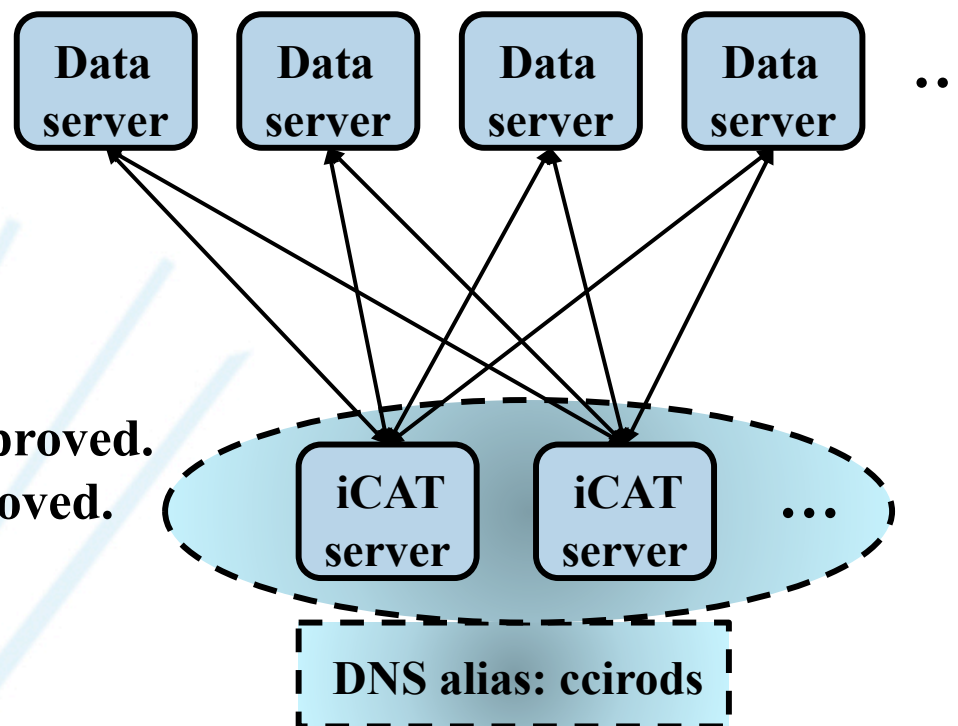- Accounting: daily report on our web site.

DNS alias:
- load balanced.
- redundancy improved.
- scalability improved.

Data server   Data server   Data server   Data server   …

iCAT server   iCAT server   …

DNS alias: ccirods

iRODS usage at CC-IN2P3: TIDRA

- TIDRA: Rhône-Alpes area data grid.
- Used by:
  - Biology.
  - Biomedical applications:
    - Animal imagery, Human data.
    - → Automatic bulk metadata registration in iRODS based on DICOM files content (Yonny).
  - Coming soon: synchrotron data (ESRF – Grenoble).
- Already 3 millions of files registered.
- Up to 60000 connections per day on iRODS.
- Authentication: using password or grid certificate.
- Expecting growth: + 20 TBs or more.
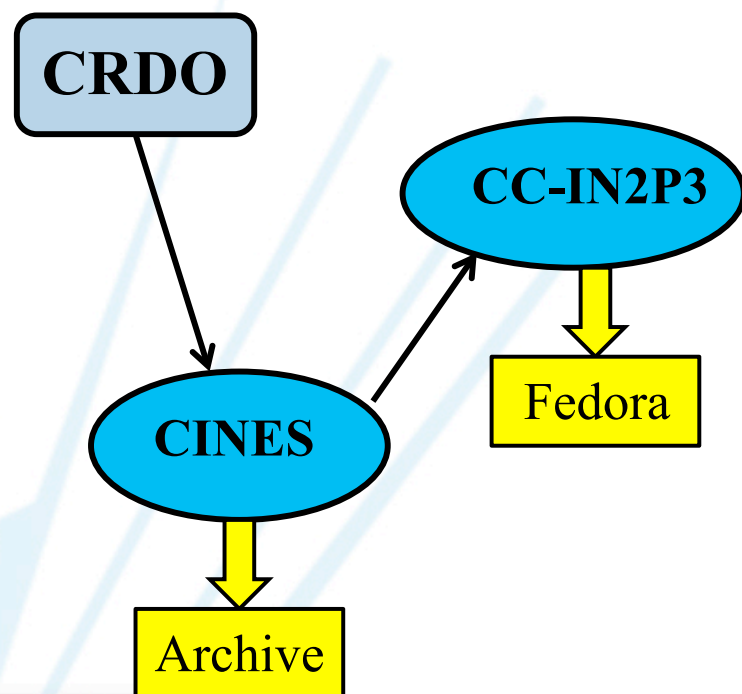
iRODS usage at CC-IN2P3: Adonis

- Adonis: french academic platform for Arts and Humanities.
- Various projects:
  - Data archives.
  - Data online access through:
    - Fedora-commons.
    - Web site.
  - Data access and processing by batch jobs.
- iRODS at the heart of Adonis.
- Already 20 TB of data (2 M files), more than 100 TB at the end of this year.

- Ex: archival and data publication of audio files (CRDO).

1. Data transfer: CRDO → CINES (Montpellier).
2. Archived at CINES.
3. iRODS transfer to CC-IN2P3: *iput file.tar*
4. Automatic untar at Lyon.
5. Automatic registration in Fedora commons.

iRODS usage: prospects

- Starting:
  - Neuroscience: ~60 TB.
  - IMXGAM: ~ 15 TB (X and gamma ray imagery).
  - dChooz (neutrino experiment): ~ 15 TB / year.
- Coming soon: LSST (astro):
  - For the IN2P3 electronic test-bed: ~ 10 TB.
  - For the DC3b data challenge: 100 TB ?
- Thinking about a replacement of light weight transfer tool (bbftp).

➔ communities: High Energy physics, astrophysics, biology, biomedical, Arts and Humanities.

iRODS contributions

- Scripts:
  - Test of icommands functionnalities.
- icommand:
  - iscan (release 2.3): admin command.
- Micro-services:
  - Access control: flexible firewall.
  - Msi to tar/untar files and register them in iRODS.
  - Msi to set ACLs on objects/collections.
- Universal Mass Storage driver.
- Miscealeneous (related to the Resource Monitoring System):
  - Choose best resource based on the load.
  - Automatic setup of status for a server (up or down).
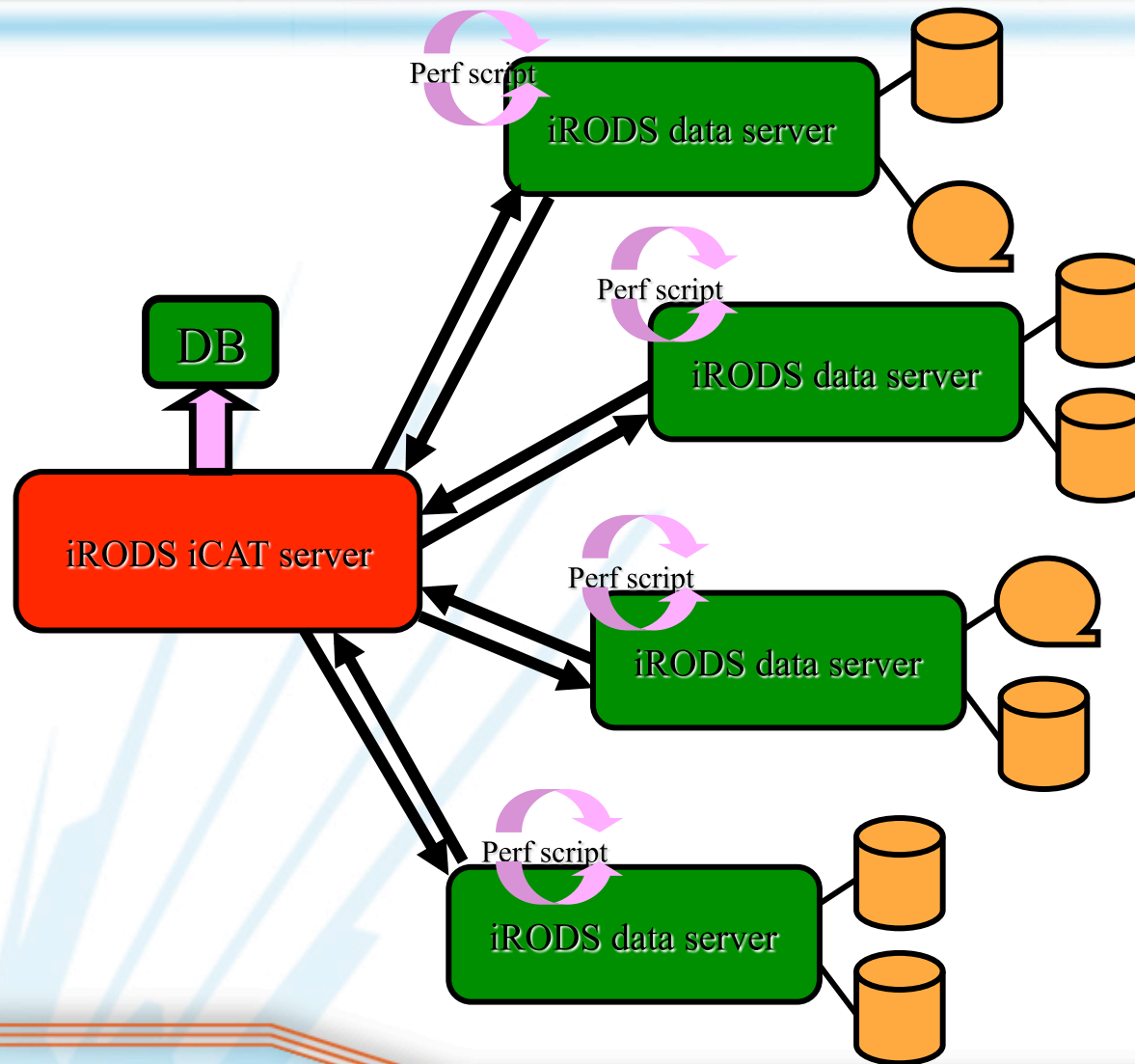
1. Ask each server for its metrics: rule engine cron task (msi).

2. Performance script launched on each server.

3. Results sent back to the iCAT.

4. Store metrics into iCAT.

5. Compute a «quality factor» for each server stored in an other table: r.e. cron task (msi).

JUX: Java Universal eXplorer

- Provide a single GUI for accessing the data on the GRID.
- JUX tries to be intuitive and easy to use for non-expert users:
  - use context menus, drag-and-drop…
  - close to widely used explorer (i.e. Windows explorer)
- Written in Java by **Pascal Calvat**.
- Based on the JSAGA API developed at ccin2p3 by Sylvain Reynaud.
- JSAGA provides the data management layer:
  - Protocols: srb, irods, gsiftp, srm, http, file, sftp, zip…
  - SRB and iRODS plugins are using Jargon.
  - Can add a plugin easily for a new protocol.
- JSAGA provides security mechanisms:
  - Globus proxy, VOMS proxy, Login/Password, X509

JUX: Java Universal eXplorer

- Download: https://forge.in2p3.fr/wiki/jux
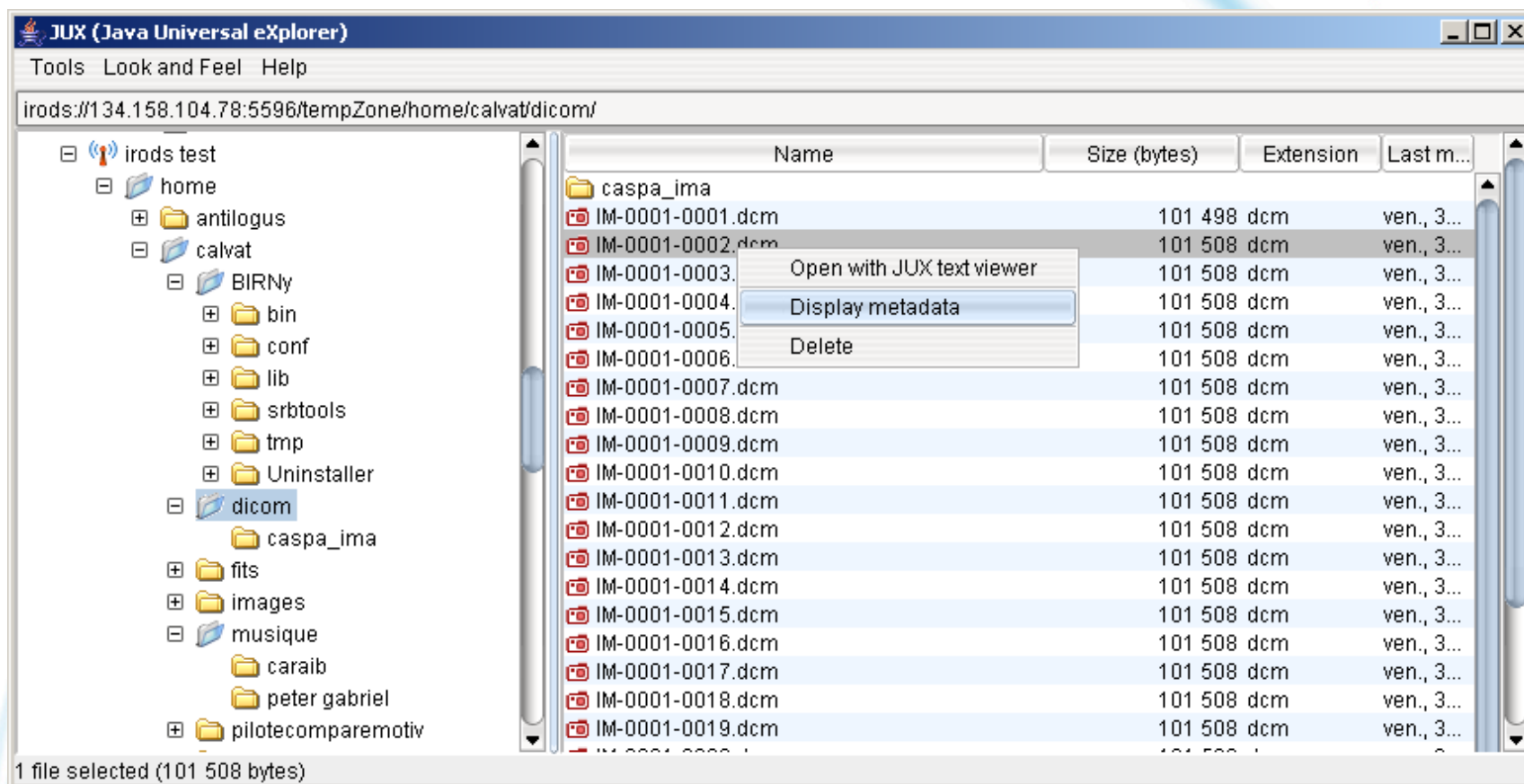
SRB migration to iRODS

- SRB as of today:
  - More than 10 experiments using it.
  - Key component for all of them.
  - **More than 2 PBs** handled by the SRB.
  - Max network traffic 15 TB / day.
  - 500,000 connections per day, even more.
  - OS: Windows, MAC, Linux (SL, Debian, Suse), Solaris, AIX.
  - Clients: from laptop to supercomputers.
  - Connections as far as Australia and Hawaii.
- Still growing.

SRB migration to iRODS

- Tricky to do it in production:
  - Java part ok: almost transparent.
  - Shell commands: need to adapt client side scripts.
  - Will provide a simple tool to help converting scripts from Scommands to icommands.
- Will start this year (2 years process) with BioEmergence (extra 60 TBs in iRODS).
- Won't do migration for ended projects (BaBar, SNFactory etc…).

iRODS overall assessement

- iRODS is becoming more and more popular in IN2P3 community and beyond.

- Very flexible, large amount of functionnalities.

- As SRB, it allows to go way above simple data transfer.

- Can be interfaced with many different technologies (no limit):

  - Cloud, Mass Storage, web services, databases, ….

    → Able to answer a vast amount of needs.

- Lot of projects = lot of work for us !

- Goal for this year: ~ x00 TB (guess: > 300 TBs).

- Should reach PB scale very quickly.

iRODS issues

- Filenames and metadata with accentuated letters.
- Need better control on servers' activity:
  - Connections can come from anywhere (laptops, batch farms …).
  - Limit the number of processed requests at the same time.

Acknowledgement