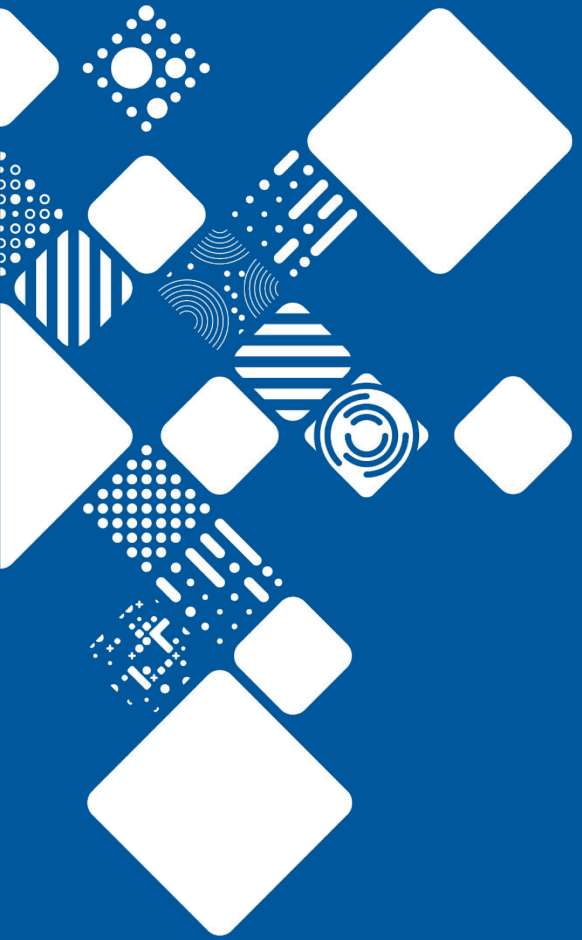


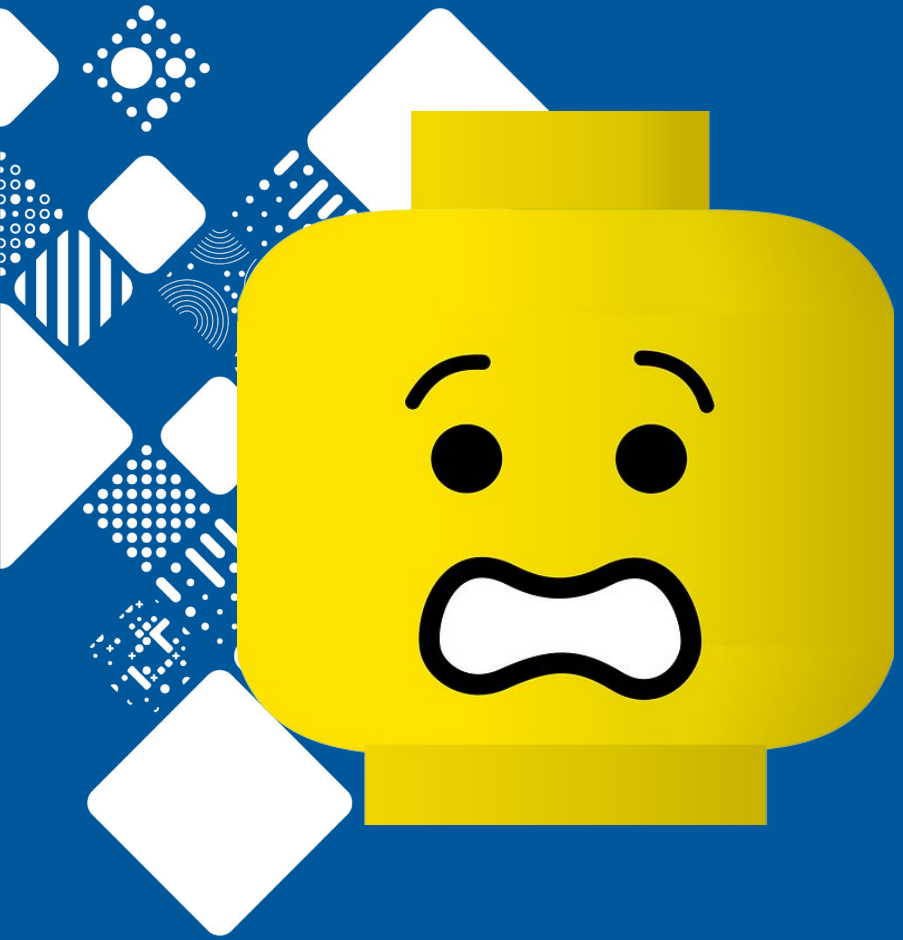
Wellcome Sanger Institute iRODS Deployment Seven Years On

John Constable
(Informatics Support Group)
https://www.sanger.ac.uk/science/groups/informatics-support-group

wellcome sanger institute

It’s been seven years since "Implementing a genomic data management system using iRODS in the Wellcome Trust Sanger Institute.” (https://www.ncbi.nlm.nih.gov/pubmed/21906284) was published.

wellcome sanger institute



Gosh

wellcome sanger institute

Recap

wellcome sanger institute

Image Credit: Pablo Gonzalez (Flikr)

“Increasingly large amounts of DNA sequencing data are being generated within the Wellcome Trust Sanger Institute (WTSI). The traditional file system struggles to handle these increasing amounts of sequence data. A good data management system therefore needs to be implemented and integrated into the current WTSI infrastructure. Such a system enables good management of the IT infrastructure of the sequencing pipeline and allows biologists to track their data”

wellcome sanger institute

First failed.
Second failed.
Third one stayed up!

wellcome sanger institute

So we installed iRODS 1.0

(the paper was written on 2.4.0)

wellcome sanger institute

It had seven servers!

Two Zones!

Two iCAT’s, federated.
Four iRES, replicated.

It authenticated against Active Directory.

We used Oracle as the Catalog Backend database, as was the fashion at the time.

wellcome sanger institute

Photo Credit: nolnet (Flikr)

We started by adding Storage via SAN.

First Nexsan, then DDN.

~400TB per server

It got used a lot so we added more zones. More capacity each year.

wellcome sanger institute

We (Pete) upgraded to 3.3.1

Photo Credit: Brian Rinker (Flikr)

wellcome sanger institute

We moved half of the storage to another data centre.

While the system was live.

With no one noticing.

On a lorry.

(You may have seen my colleague Jon Nicholson’s talk on this)

Photo Credit: Brickset (Flikr)

wellcome sanger institute

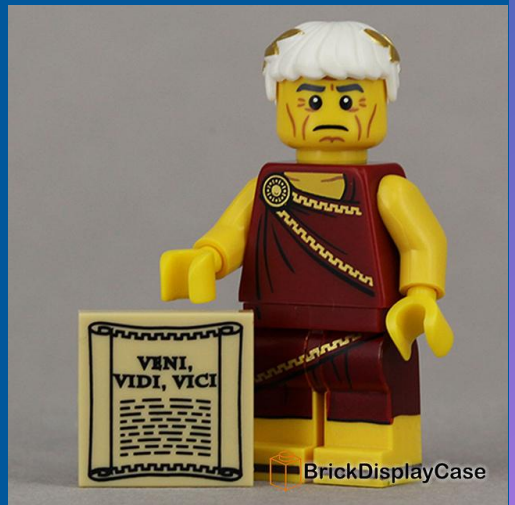


Photo Credit: brickdisplaycase.com (Flikr)

We upgraded to 4.1.8

(You may have seen my previous talk about this)

Took a year of prep.

Further upgrades took an hour, including prep.

Currently on 4.1.10, 4.1.11 on dev.
Hoping to jump to 4.1.12 soon

wellcome sanger institute

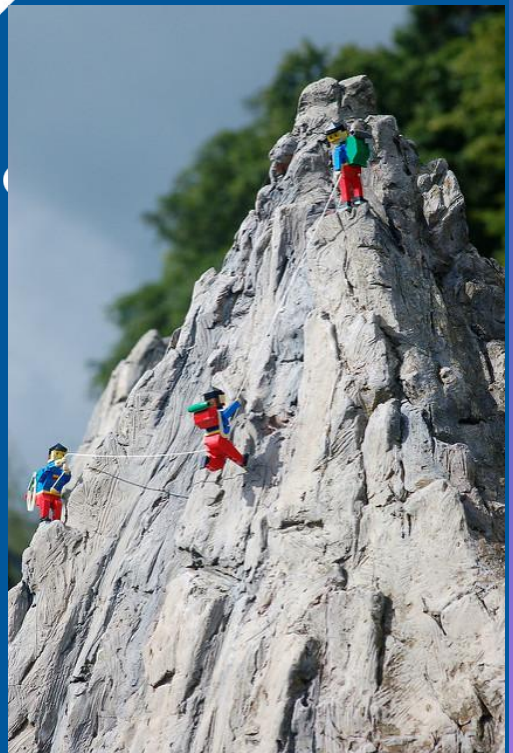


Photo Credit: Rob Young (Flikr)

We ran into scaling issues;

- One server could get its 10G overloaded.
- The number of multipath paths got to over 2k on each server!
- Could not readily make LUN's > 60TB due to fsck memory limits
- One server maintenance took a lot of storage offline

wellcome sanger institute

We switched to using 4U servers incorporating 10G networking and 60 disks.

Initially Ubuntu 12.04

Recently Red Hat 7

Photo Credit: Fred Dunn (Flikr)

wellcome sanger institute

This scaled very nicely.

wellcome sanger institute

One part time of one FTE to manage.

Today one full time FTE, plus others at times of high load

Photo Credit: Judy van der Velden (Flikr)

One Zone exports its resources via read-only NFS.

Allows researchers to compute across ‘all their data’, maintains the same workflow with migrating between file tracking platforms.

wellcome sanger institute

Almost all data is from automated pipelines, very few users upload their own data.

Getting the automated pipelines has been the key to ubiquity, for us.

wellcome sanger institute

Wrote our own tools and automation:

CFEngine and Ansible

Baton

Assorted Python maintenance scripts

Vagrant environment for testing (you have have seen my previous talk on this)

Scripts to recover a Resource from other replicas

Unit tests (not enough)

Photo Credit: noriart (Flikr)

wellcome sanger institute

Monitoring;

- Ganglia
- Collectd & Graphite for specific dashboards
- Quota dashboards & PDF monthly reports
- Capacity (this is by far the hardest)
- Access Usage (this has been by far the most valuable)
- Logging; Splunk and ElasticSearch
- Nagios

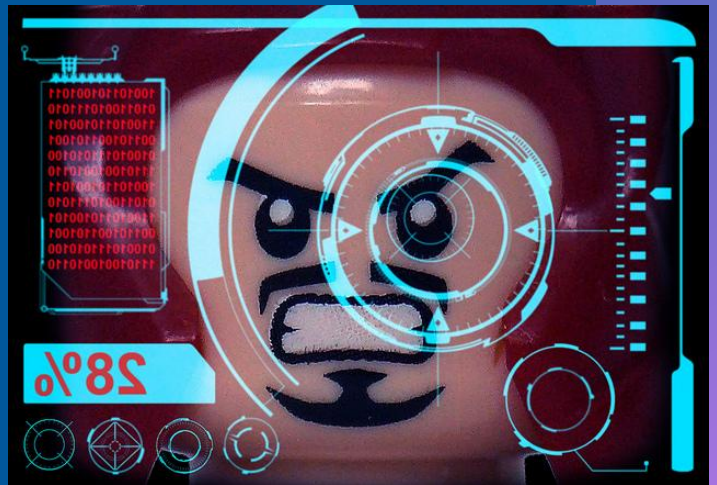


Photo Credit:
Okay Yaramanoglu (Flikr)

wellcome sanger institute

Current Infrastructure:

129 servers
~18PB (~9PB, replicated)

- Includes Dev zone that mirrors production (smaller resources)
- Six Zones (one not federated)
- One Zone HA (You may wish to see my upcoming talk about this)

Photo Credit: Paul Hartzog (Flikr)

wellcome sanger institute

Photo Credit: Benjamin Lim (Flikr)

Lessons learned

- Monitoring, logging and instrumentation (aka ‘observability’) still very early days
- Could really do with an Infrastructure As Code approach to spinning up dev environments on our Openstack Cloud
- When problems found resolution in months. We are not bleeding edge but scale brings its own challenges. So even community battle tested releases have edges unknown.

wellcome sanger institute

With Thanks to:

Dr Peter Clapham

Dr James Smith
(lego collector extraordinaire)

The lego community that make their work available via Creative Commons

wellcome sanger institute

Thank you for listening!

john.constable@sanger.ac.uk

@kript