



Hérons, Yaks and Technical Debt

iRODS at Wellcome Sanger Institute in 2020

John Constable

Who Am I?

Why Am I Here?

Five Zones, four of which Federated

One hundred and thirty servers (ish)

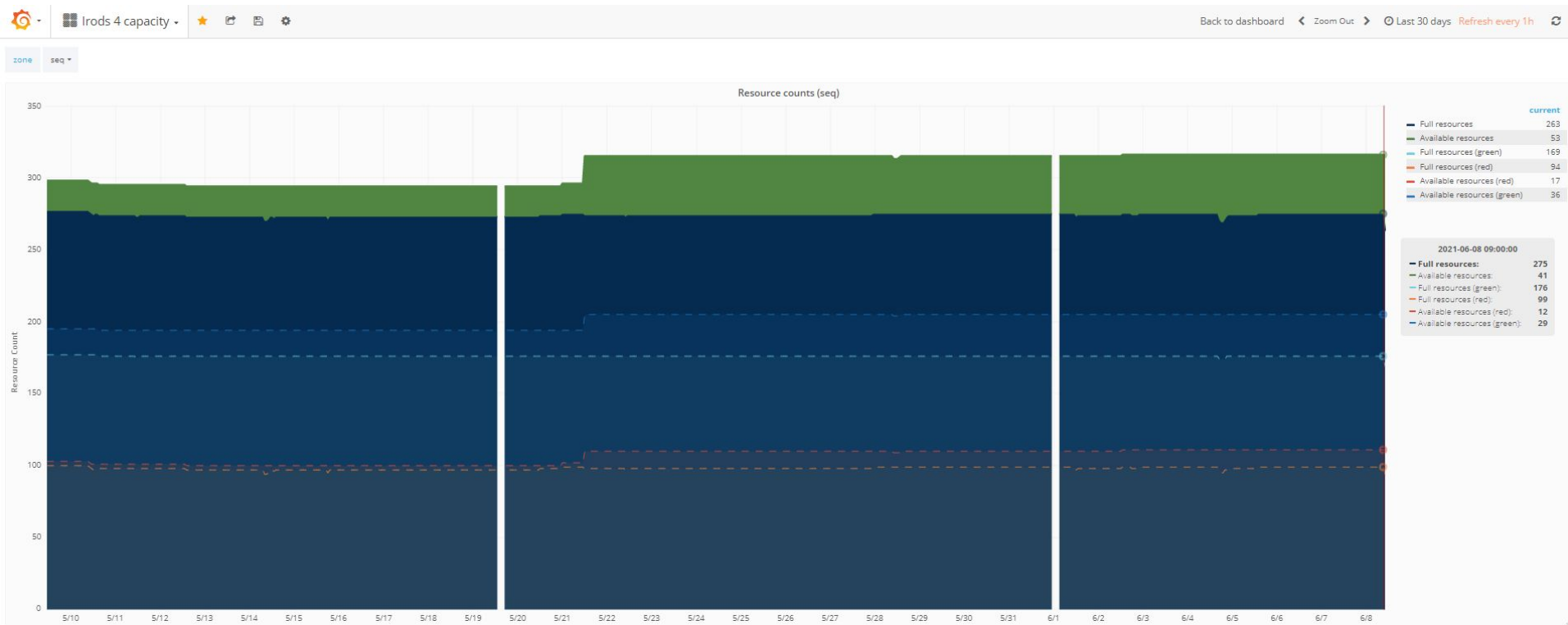
30+PB managed by iRODS

329,274,949 objects

270,813,869 metadata items

The complexities of a large installation

This is just one of our Zones (OK, the biggest)



Some Terminology



Technical Debt!



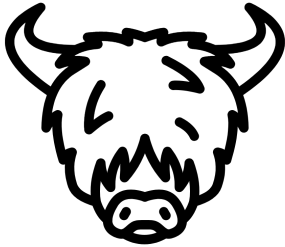
Yak Shaving!



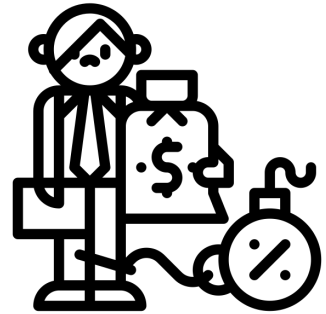
Herons are long-legged, long-necked, freshwater and coastal birds.

Credit: [Wikipedia](#)

How many can you count in this talk?



Created by Yu luck
from Noun Project



Created by Chaowalit Koetchuea
from Noun Project

The Operating System Upgrades



Technical Debt



Ubuntu 12.04 /
Red Hat 7

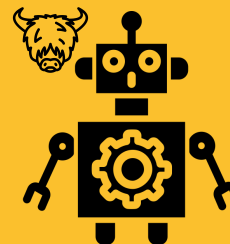


iRODS 4.1.12



CFEngine

CFEngine



Created by SBTS
from Noun Project

Semi-automated
install.

Security Challenge!

Update iRODS
and its OS's
before an audit
in August.

The challenge



Ubuntu 12.04 EOL



4.2 was significantly
different



CFEngine

CFEngine



Created by Wilson Joseph
from Noun Project

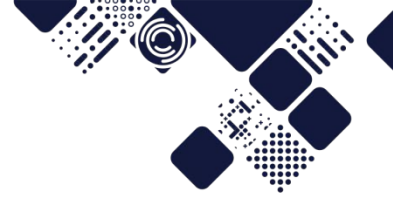
Team size



The Solution!

Created by Marek Polakovic
from Noun Project

Operating System Upgrade



We upgraded the OS in-situ to 18.04

Operating System Upgrade



We installed 18.04 manually *over the top* of Red Hat, being careful to use the same OS partitions.

Provider OS's



Most providers are on VMWare, so we build new ones on Ubuntu 18.04 and switched over at upgrade.

The large zones were physical, so we built new servers to switch to.

The possibility of
significant data loss didn't
worry me at all.



Created by WARPAINT Media Inc.
from Noun Project

OS Upgrade Issues



14.04 upgrade

[Debian Bug 953201](#)



16.04 Upgrade

UDEV rules to
needed updating.



18.04 Upgrade

SSSD cache rebuild



18.04 Upgrade

Remove old kernels to
make space in /boot

4.2.7

The 4.2.7 Upgrade Itself

- Planned to do the smaller zones with the largest and the Federation Master last.
- Terraform created zone on our Openstack.
- Informatics teams needed to test and fix any issues found.
- A dev zone was upgraded first, along with a test system.



The scary bit



Two database operations;

1. [Deprecate no longer used values](#)
2. Convert the Resource Names to ID's for Every Single. Object.



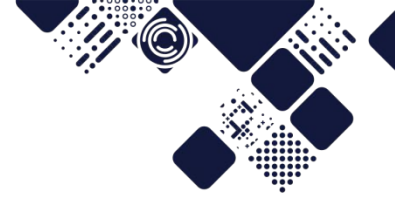
Guess how long
this took?

The cgp zone

- ~5 hours to run the update.
- ~4 million objects.

So how long do
you think the
largest zone
took?

Around 250 *million* objects...



It took a week!

**We did have one
unintended
outage**

**While testing
the upgrade, we
found a few
bugs.**



`msiSysChksumDataObj`
deprecated in favour of
`msiDataObjChksum`



Github Issues: [4590](#) / [4659](#)



Support Oracle TNS via unixODBC



[3569](#)



setup_irods.py
is no longer
idempotent

[4602](#)



Upgrading on Oracle sets the database plugin config wrong



[4603](#)



update_json.py is
not present in 4.2.X



[4604](#)



Attempting to
upgrade an IRES
from 4.1.12 to
4.2.7 first attempts
to delete all the local
resources



[4612](#)

Systemd. Oh systemd....

[2774](#) / [Pull Request Comments 3999](#)



Wait, why is systemd so terrible?

Well, yes, quite. Snark aside, the standard init.d emulator didn't work for us because;

- We have local filesystems as Resources and we want the service to come up **after** them so that it can serve the files.
- We use sssd for uid/guid caching and it needs to start **before** the irods service.
- We have customer users and groups per zone (and sometimes per server because we're terrible people), so the service needs to run as the right user.
- We use Oracle, and so the LD_LIBRARY_PATH and ORACLE_HOME variables need setting because bionic doesn't set LD_LIBRARY_PATH in the environment of a service because that would be *helpful*...
- On Ubuntu 18.04 the path needed setting manually because (to quote our git commit) "systemd doesn't *actually* honour the \$PATH, so we need to hand off responsibility to something with fewer stoats, like /bin/bash".
- The default emulation didn't know when/how to restart the service ([4946](#)) **and** we needed to set KillMode to control-group to ensure both/all processes restarted ([4947](#)).
- We needed two config files while upgrading; one for 4.1.x and one for 4.2.x, because the server binary had moved (of *course* it had!)



4.2.7 icommands don't work on server pre-4.2.7



[4636](#)



ShowCollAcls query
needs to be added
as our iRODS was
initially installed
before this query
was created



4636



iRODS packages
should declare
configuration files



4722

imeta qu ignores
the -z option



4426

Baton, the tool our
pipelines use, needs
to be compiled
against iRODS.



It's deployed via conda, which
turned out to be a lot of work to get
working in 4.2.7, in part due to
the assorted dependencies.

[Pull Request 4953](#)

[Pull Request 4952](#)



So...
What's with the
heron?



Created by Yu luck
from Noun Project

UK launches whole genome sequence alliance to map spread of coronavirus

News article

23 Mar 2020

The Wellcome Sanger Institute will collaborate with expert groups across the country to analyse the genetic code of COVID-19 samples circulating in the UK, providing public health agencies with a unique tool to combat the virus



Created by Yu luck
from Noun Project

No pressure
then!



Created by Nithinan Tatah
from Noun Project



But that's not all!

- SSL enabling
- PAM Authentication

PAM token
extension.. doesn't



pam_password_max_time does
not allow pass_expiry_ts above
1209600 #5096



Our users often
forgot to use
`iinit` rather than
`kinit`



[pep_api_auth_response_pre](#)
[can be triggered without _post](#)
[or _except #5201](#)





**We did a few other
things too.**

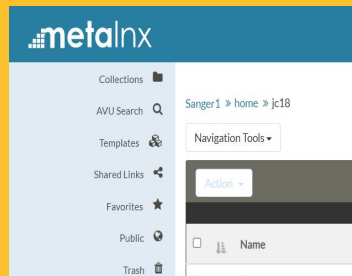
Other projects



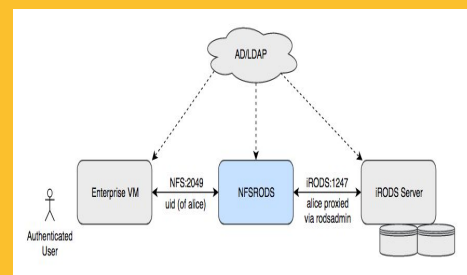
**Migrating 8PB of data to
new hardware**



**Proof Of Concept
Indexing
Capability**



**Proof Of Concept
MetalINX**



**Proof Of Concept
NFSRODS**

Migrating data: an update



[wtisi-ssg/irods_migrate:](#)
[tooling for bulk migration](#)
[of data in iRODS using](#)
[parallel iphymv](#)

Indexing Capability Plugin



- [Indexing should maintain a stable number of delayed rule executions #45](#)
- [buffer space / available inode exhaustion #58](#)

Metalnx web GUI

- [The warning message for attempting to download files larger than metalnx.download.limit should reflect the actual size #220](#)



NFSRODS

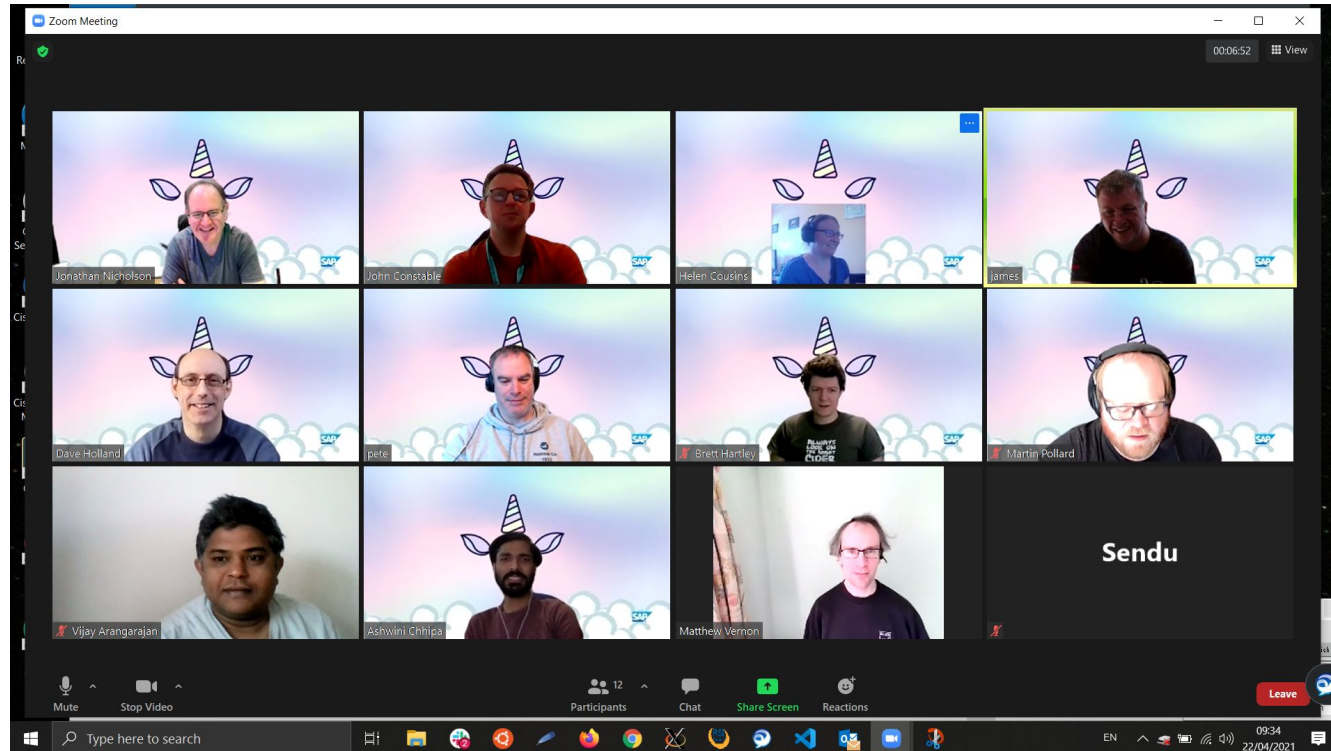
- Large file collections cause long lookup issues
- Doesn't do parallel transfer, so can be quite slow.



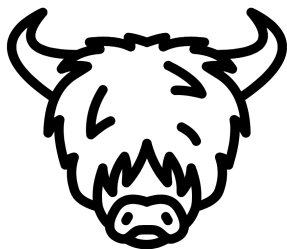


So what now?

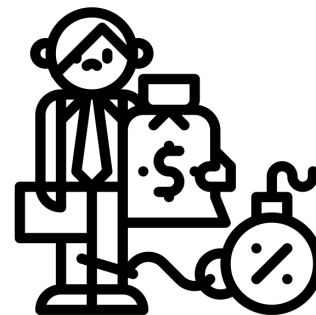
Thanks to my colleagues



How many did you count?



Created by Yu luck
from Noun Project



Created by Chaowalit Koetchuea
from Noun Project

Photo Credits

Technical Debt Photo by [Martijn Baudoin](#) on [Unsplash](#)

Yak Photo by [Manav Agarwal](#) on [Unsplash](#)

Heron By I, Acarpentier, CC BY 3.0, <https://commons.wikimedia.org/w/index.php?curid=2945358>

Filing Cabinet Photo by [Maksym Kaharlytskyi](#) on [Unsplash](#)



John Constable

System Administrator specialising in iRODS. Shaves
Yaks as a profession.



Questions?

LinkedIn Profile via QR on the left

Twitter: @kript