THE UNIVERSITY OF QUEENSLAND
AUSTRALIA
CREATE CHANGE

Site update: UQ's Storage Scale Journey.

...a site update for 2023.

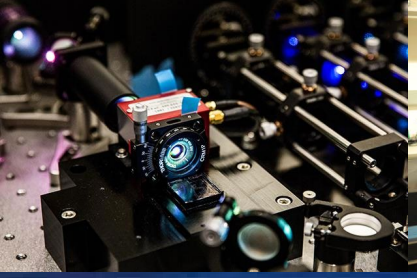
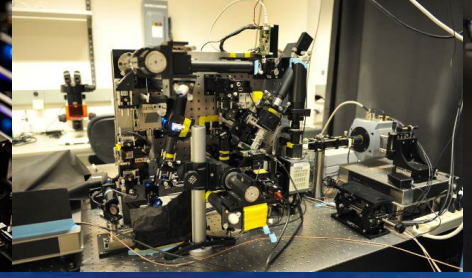
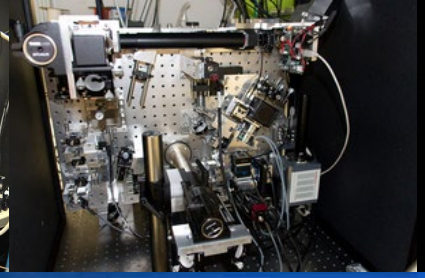
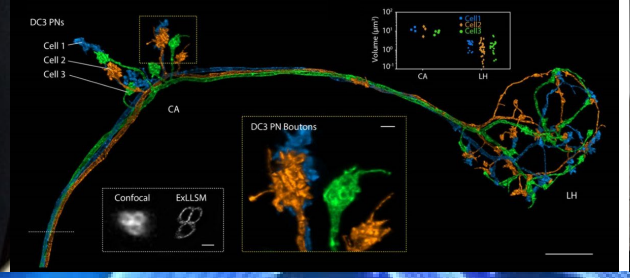
Jake Carroll, Chief Technology Officer, Research Computing Centre, The University of Queensland, Australia.

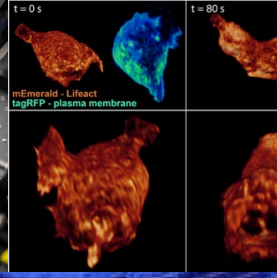
jake.carroll@uq.edu.au

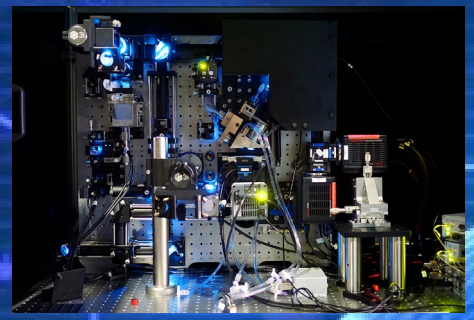
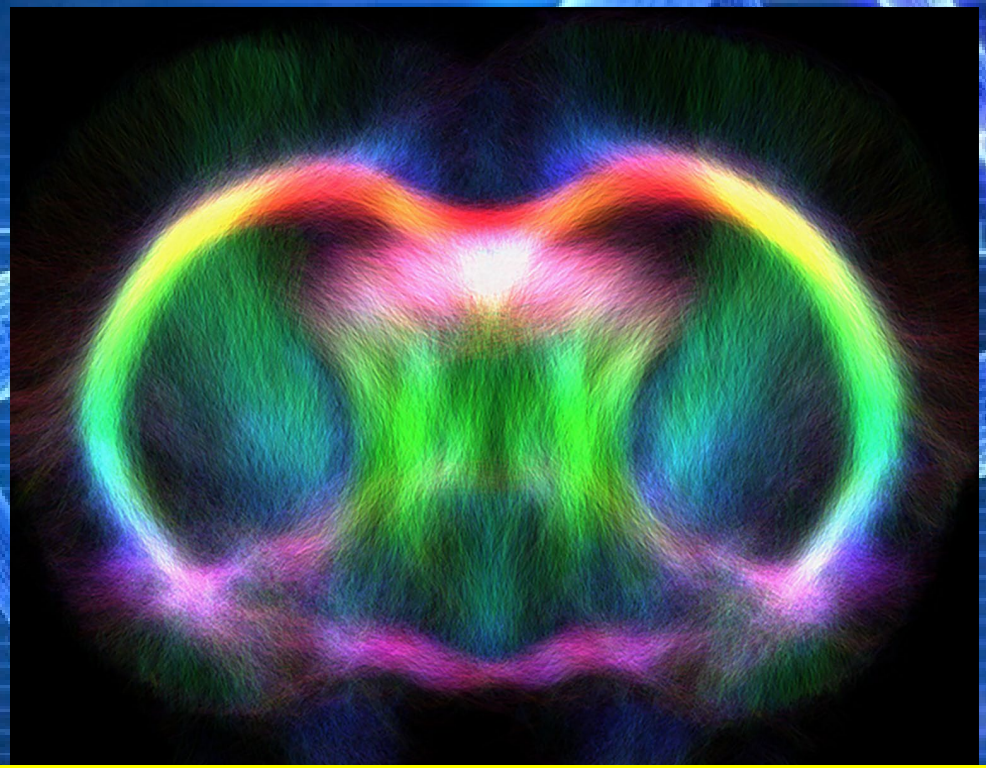
Context

UQ is:

- Six major faculties
- Eight institutes
- Fifteen sites
- ~$441.5 HERDEC income (2021)
- ~55,400 students (2022)
- ~7,410 full time staff (2022)
- ~25,000 endpoints
- 2 supercomputers, more than 10,000 CPU cores
- ~100 PB of research data

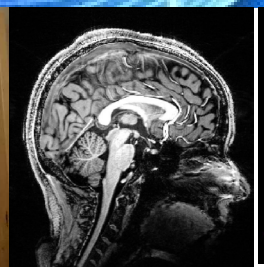
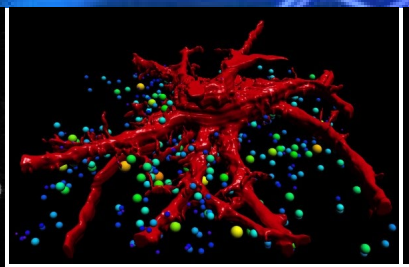
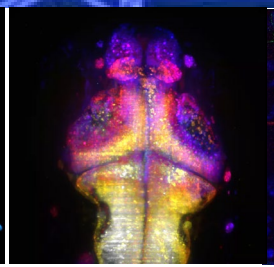
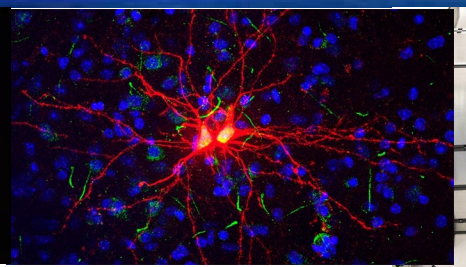
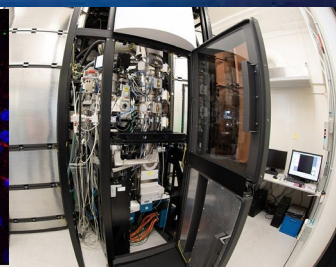





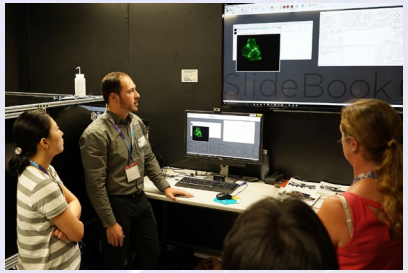




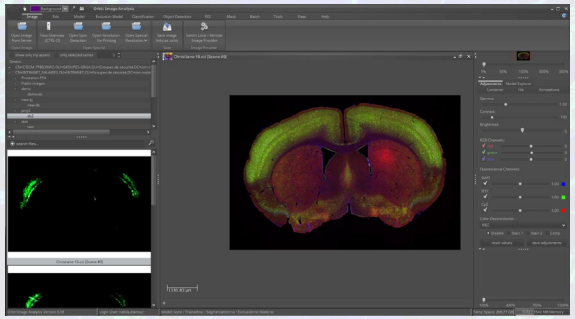


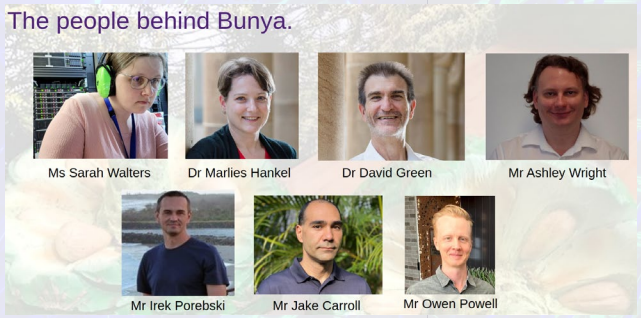
100’s of terabytes generated per day.

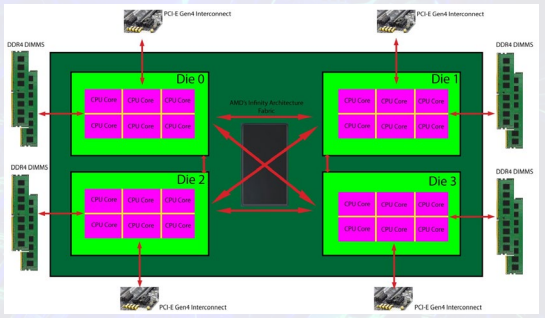






UQ Research Computing Centre










LATER
NEXT
NOW

*"**Authenticity**. I'd rather tell you the whole story. Not just the bits that make it sound great…"*

--me.

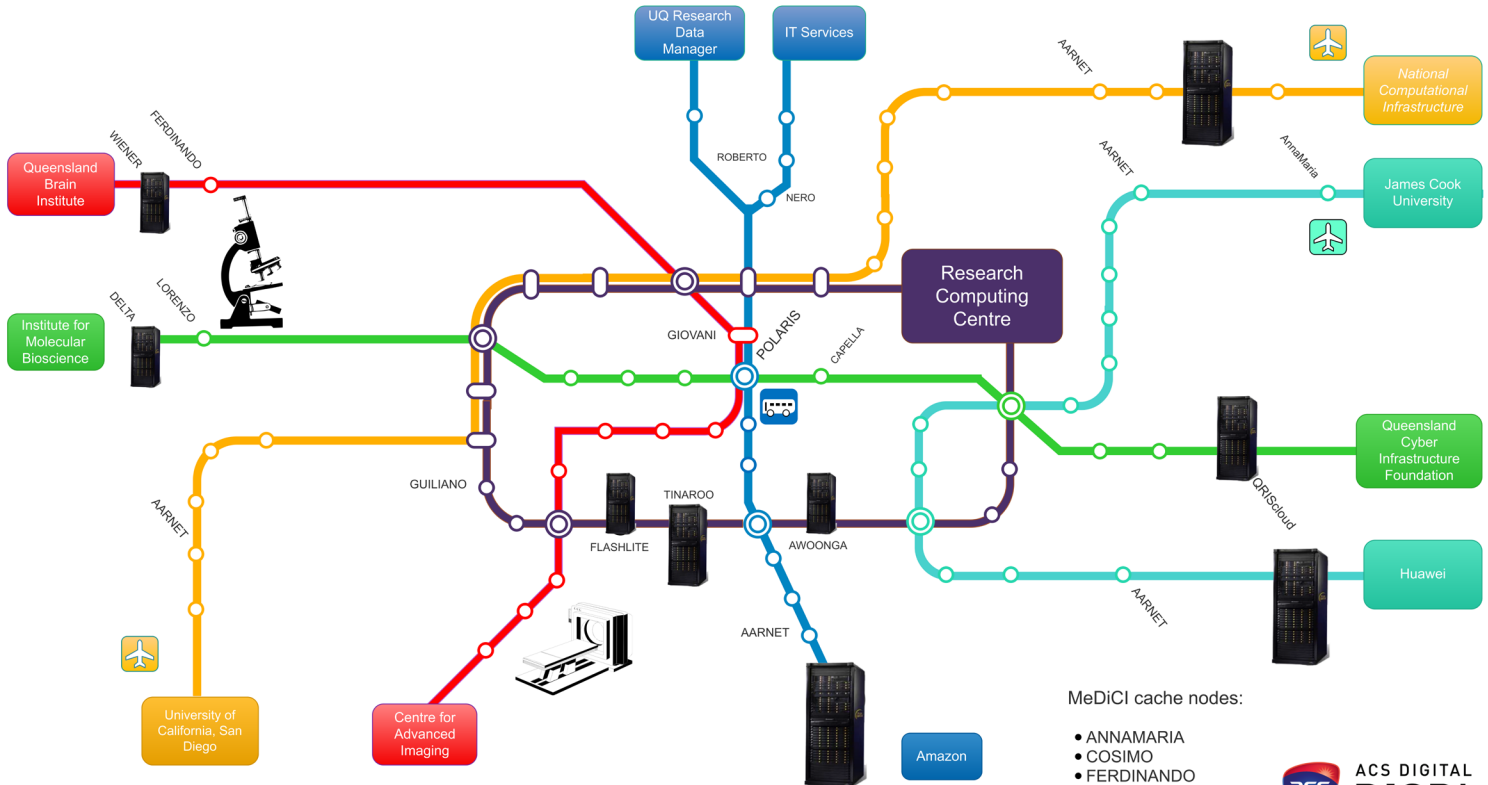
What do we do with Storage Scale?

- Since 2015 we've been building a unified data fabric using Storage Scale, AFM and ESS.
- IBM took a journey with us and didn’t run away when things got difficult, complex and entered “breaking point” territory.
- Australians are a bit crazy but we do cool things with filesystems. Blame our toxic wildlife?
- We took scientific research data infrastructure to better places in scale, performance and capability for the scientific research and higher education sector.

MeDiCI (2019): UQ's Data Fabric of Choice

The University of Queensland's Metropolitan Data Caching Infrastructure (MeDiCI) provides seamless access to data regardless of where it is created, manipulated and archived.

Developed by the Research Computing Centre (RCC), MeDiCI holds copies of data on campus until it is not required for some time. Data is moved between on and off-campus storage on demand without user involvement. MeDiCI is underpinned by HPE DMF, DDN Grid Scalar storage, and IBM Spectrum Scale technologies.



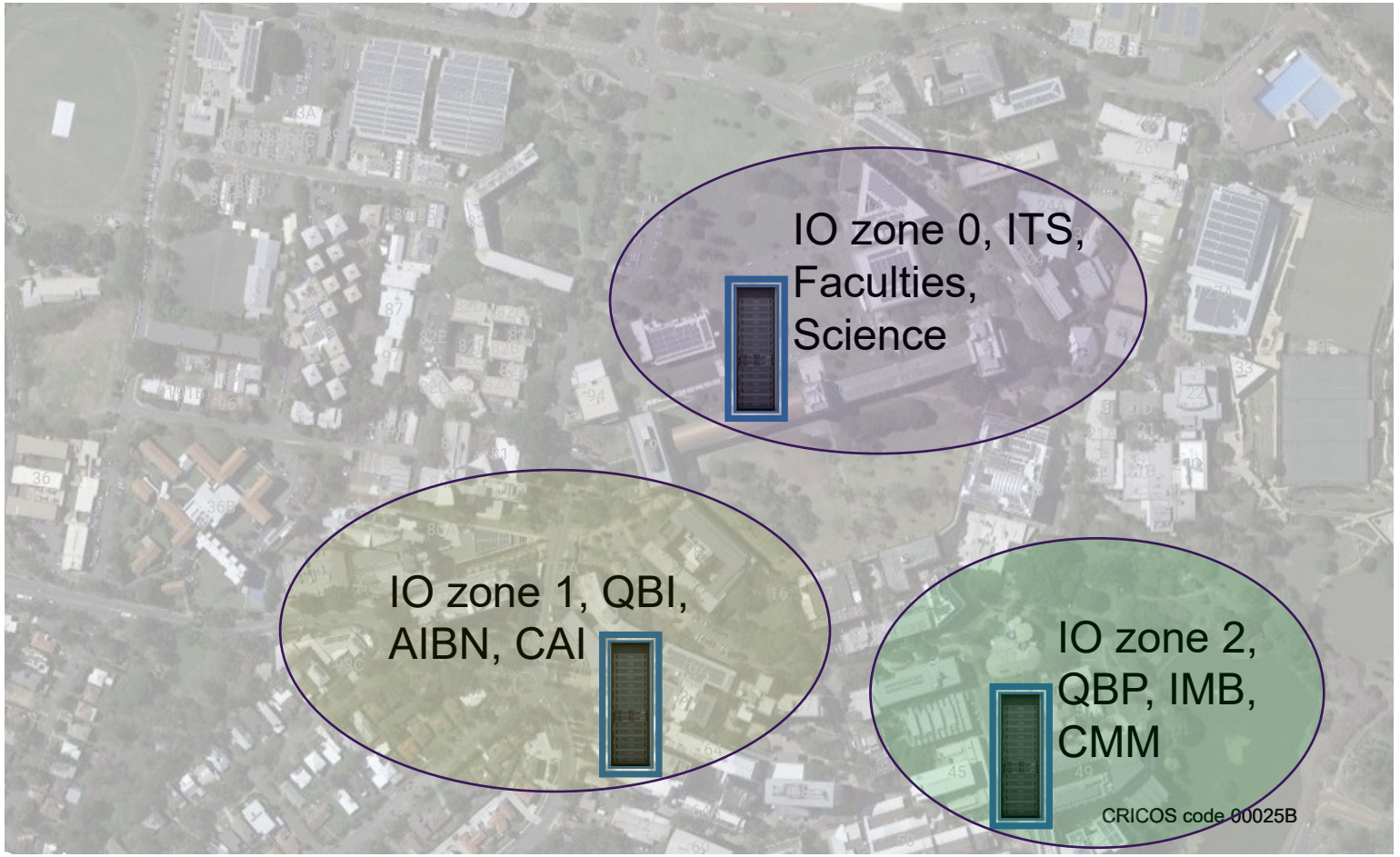
MeDiCI cache nodes:

- ANNAMARIA
- COSIMO
- FERDINANDO
- GUILIANO
- GIOVANI
- CAPELLA
- NERO
- ROBERTO
- PIERO

ACS DIGITAL DISRUPTORS AWARDS

Distributed IO zones around campus.

Cache and IO zone distribution around campus – putting caches near our instruments and IO intensive locations, using AFM, protocols and an SMB, to “knit” the fabric together…

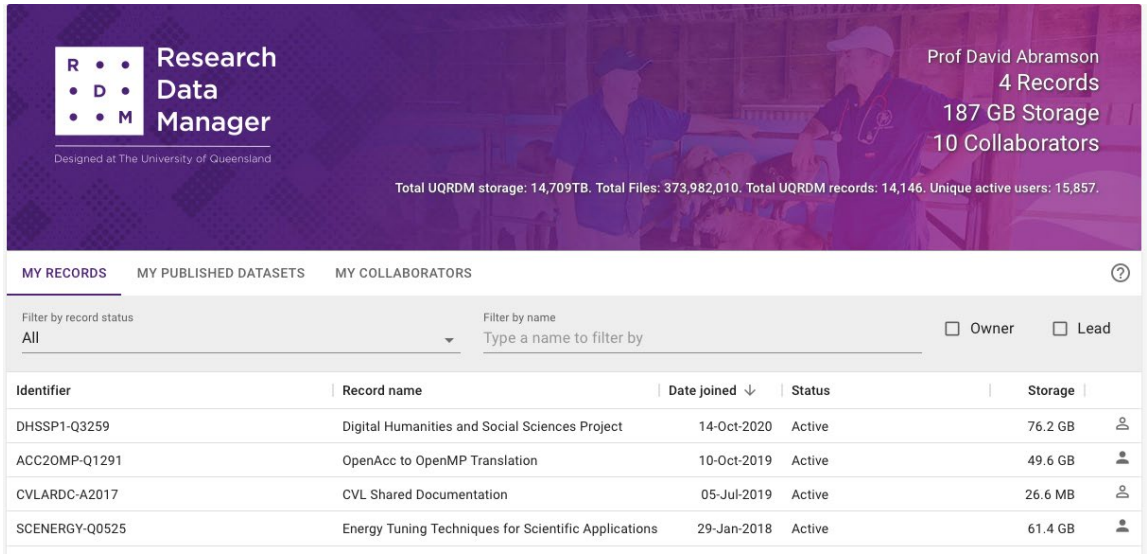


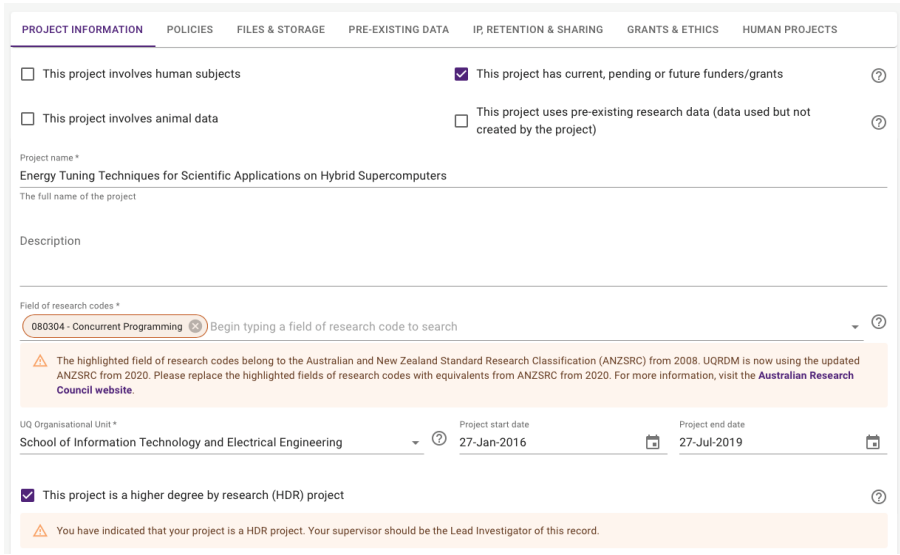
We put a layer on top to provision storage

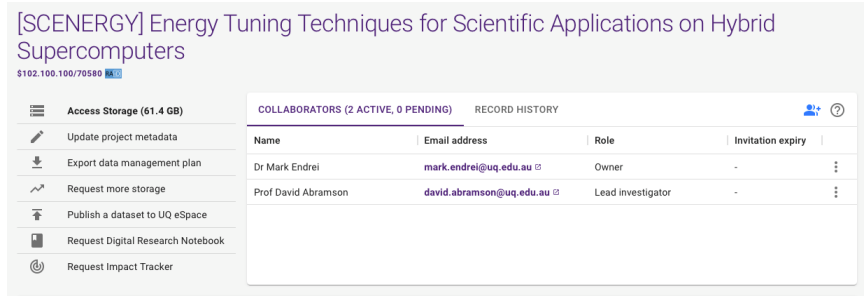
Research Data Manager

Designed at The University of Queensland

- Designed by our Research Systems team.
- Came into existence about a year or so after we launched the data fabric.



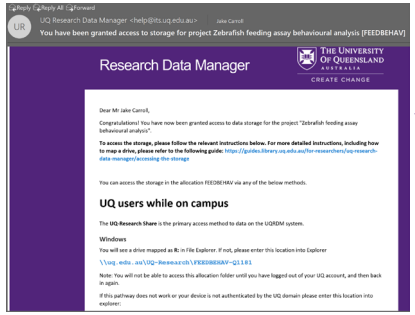




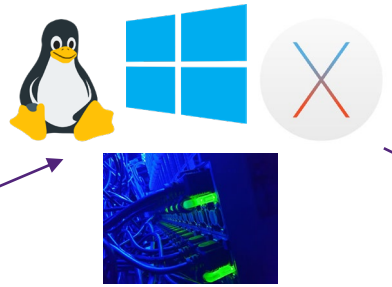
A typical workflow

Researcher obtains a collection. Q1234
[This is a GPFS AFM fileset!]

Research Data Manager
Designed at The University of Queensland

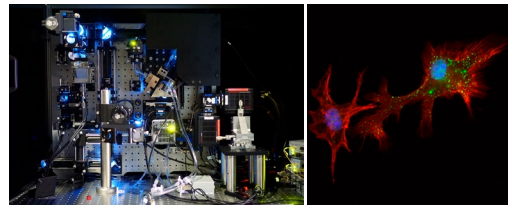


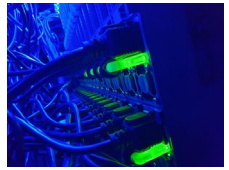
Researcher is sent an email explaining access instructions.

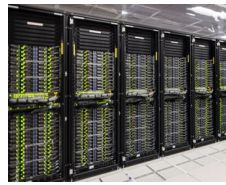


Q1234 then mounted on the fabric.

Researcher then acquires data and stores in Q1234

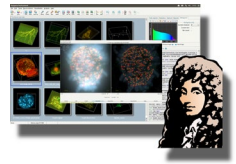


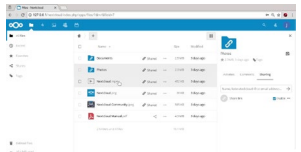


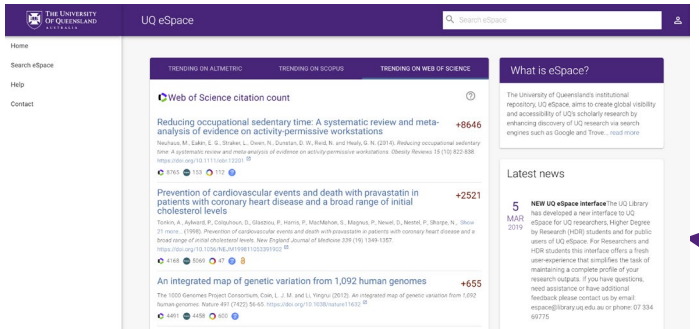


PYTORCH

nectar





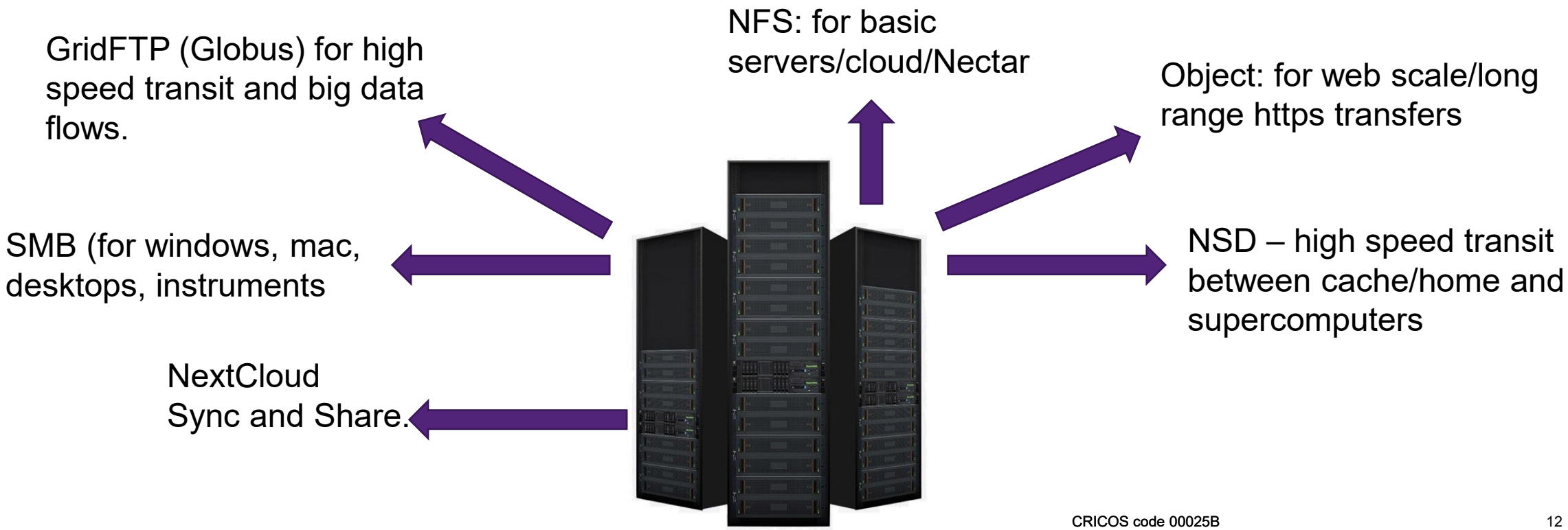


DOIs can be minted, published to UQ eSpace.
RDM can facilitate data linking back for durable URL.

It is already on the fabric. Can be accessed via, supercomputers, desktops, laptops, services.

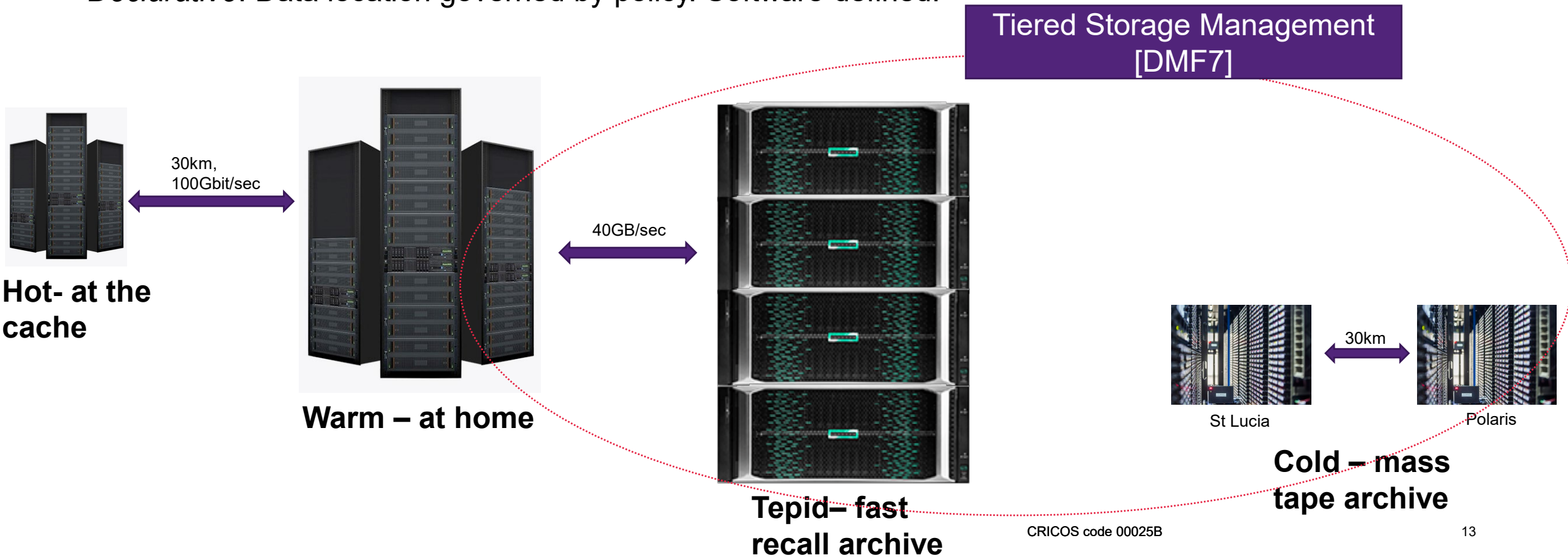
Offer data in a way that suites the use case. **Protocols.**

- Recognise that different users, workloads and communities have different ways they present their data.
- Use technology that can present a view of the data in many forms, from one namespace.



Data where it needs to be, when it needs to be there.

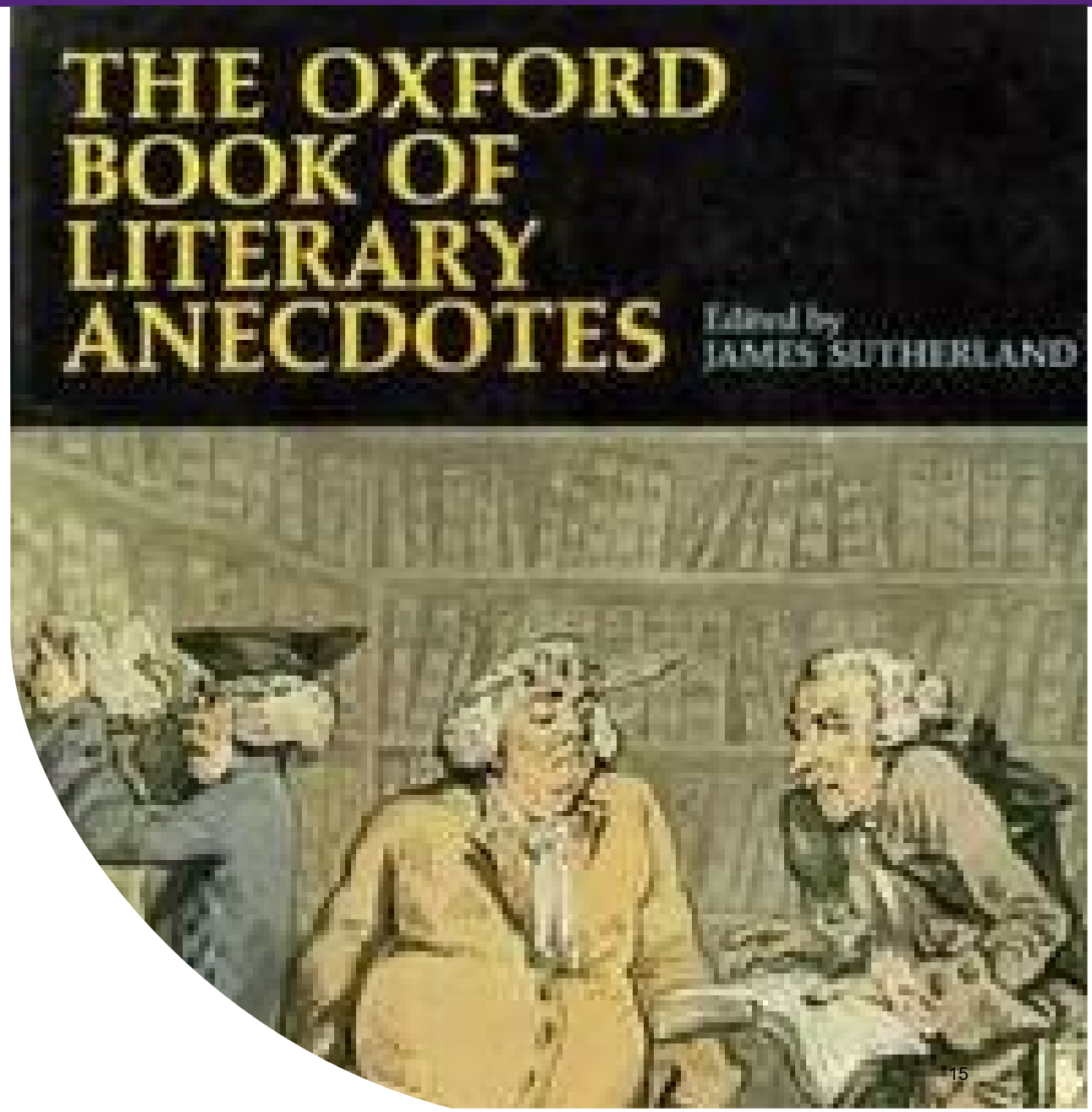
- Using tiering technology and the concept of caches so that data is where it needs to be *on-access* and not taking up space on expensive media, when it is cold and infrequently used.
- *Declarative*. Data location governed by policy. Software defined.




FAST
FORWARD

Anecdotes.

- Story 1: The ESS 3500 hybrid array experience.
- Story 2: Scalability limits of SMB+NFS tools, shares, admin issues
- Story 3: Difficult ESS Upgrades.
- Story 4: Storage Scale + DMF7
- Story 5: Storage Scale ESS + AFM at 400Gbit/sec over the wide area.
- Story 6: Fusion HCI and Discover adjacent to a Storage Scale environment – and the cool things one can do with that.
- Story 7: The support experience



Story 1: The ESS 3500 hybrid array experience.

- The implementation of one of the first ESS3500's [NVME+NL-SAS] in Australia was a bumpy experience.
- It appears that the OEM of the JBOD didn't consider regional power and electricity differences.
- The JBOD sensors didn't appreciate 240V.
- The JBOD sensors didn't like power fluctuating by tiny amounts.
- The JBOD had fan PWM control tolerance misreporting.
- The JBOD thought it was too cold.
- Resolved after two months with IBM's help and two controller firmware upgrades later. I guess the OEM/ODM got involved.

Take away message for IBM to consider:

- We are good at finding these problems pre-production *with* you.
- We've done this type of work with you before.
- If you can get silicon in front of us earlier – we (likely) won't end up with an "*oh….s@#$t*" moment days before production.
- Better for everyone. Create value with us early on (please).

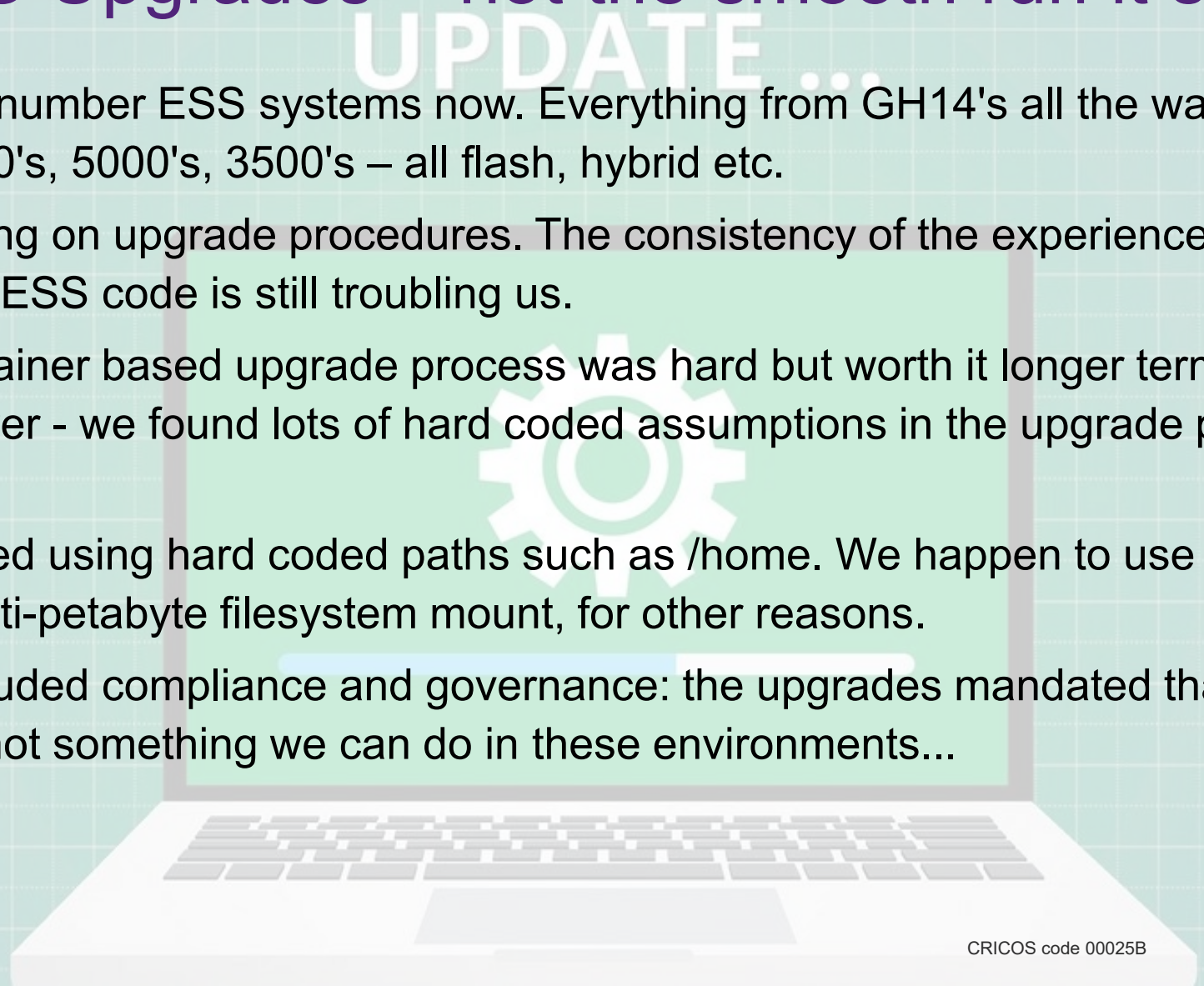
Story 2: SMB & NFS-Ganesha management scalability.

- We discovered that Storage Scale CES administration tools don't perform well with:
  - ~16,000 SMB shares.
  - ~34,000 NFS-Ganesha exports.
- We've reached scalability limits of the **admin** tools of these protocols on Storage Scale CES clusters.
  - Timeouts when updating SMB ACL's, reloading NFS, watching Ganesha hang and block IO on reload of very long export lists.
- To be clear: the protocols themselves perform well in terms of throughput at scale. Many GB/sec out of SMB, likewise out of NFS when reading or writing to Storage Scale filesystems.
- Breaking the clusters up into individual less populous CES clusters isn't a viable option.
- We've raised this with development teams at IBM to see what can be done.
  - Major issue for us, as our (thousands of) users rely on Storage Scale exports to function predictably.
- Tuxera Fusion File Share approach being contemplated if it can remove scalability problems of the above commands and things like ctdb.

Take away message for IBM to consider:

- Your filesystem is awesome and continues to get incrementally more awesome.
- Your use of SMB and NFS tools from community are good.
- Your reliance upon the admin tools that traditionally manipulate these protocols is failing at the large enterprise scale.
- Time to find a new way or risk losing business to the NetApp/Isilon “NAS is king” crowd - even if we all know they’d face similar issues.
- *Perception in market is key.*

Story 3: ESS Upgrades – not the smooth run it should be



- We have quite a number ESS systems now. Everything from GH14's all the way through to multiple ESS3000's, 5000's, 3500's – all flash, hybrid etc.
- We keep stumbling on upgrade procedures. The consistency of the experience to get between point releases of ESS code is still troubling us.
- Migration to container based upgrade process was hard but worth it longer term. Once that was ironed out however - we found lots of hard coded assumptions in the upgrade process that broke our environment.
- Examples included using hard coded paths such as /home. We happen to use /home as a location for a multi-petabyte filesystem mount, for other reasons.
- Other issues included compliance and governance: the upgrades mandated that we turn off firewalls. That's not something we can do in these environments...

Take away message for IBM to consider:

- Our take: It should be a better experience, hands down, compared to buying a white-box or OEM "GPFS" experience, or brewing our own.
- We need it to be a first class citizen compared to what we could cobble together by hand, ourselves.
- Diligence is being paid: IBM support has been first class in hearing out our concerns in post mortem meetings – and real attention has been paid. What happens next to turn it into positive change?

Story 4: DMF7 + Storage Scale: a world class combination.

- It took us several years working with HPE and IBM to perfect this, but we now have what we think of as a world class preservation and archiving facility that combine HPE's DM7 and IBM's Storage Scale.
- Storage Scale "cache" [which is an AFM home] natively HSM managed by DMF7.
- New and powerful workflows for data migration inside storage scale managed environments, using DMF7, NoSQL and Object Reflection Tables to do all the heavy lifting.
- Importantly: Storage Scale is a cohesive part of a world leading PRESERVATION system. Storage Scale needs to be thought of as an active contributor in the digital preservation market and high end archiving market as well as the "go fast" market – but you need to pair it with the right integrations and interactions to get the very best.

On the other side of the AFM home, we've natively integrated HPE's DMF7 to space manage Storage Scale in our ESS.

ESS 3500+5000
Storage Scale
Filesystems

/UQ00
/UQ01
/UQ0..
/UQ07

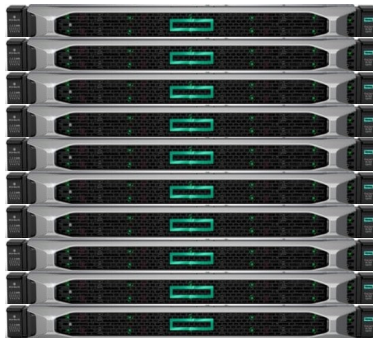
Light weight events via Ganesha
interface processed by DMF7
to events_filter for DMAPI processing

dmf7-gpfs-dm1
dmf7-gpfs-dm2

200G HDR fabric

MESOS

Apache Zookeeper



Cassandra Object Scale db, all nvme HDR200 connected event processing and reflection table generation cluster

Spark

MARATHON

A more holistic picture of the new architecture shape…

[Future?] Remote Preservation Copy

Research Data Manager

Designed at The University of Queensland

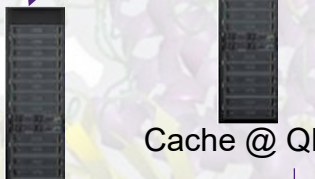
Q collection
Request from User

DMF7 engine

- Policy
- Placement
- Preservation
- Data movement
- Integrity

TS4500 + 10 *
TS1160's
Preservation Copy 2

/afm
DMAPI space
managed



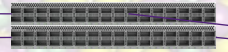
Cache @ IMB

Cache @ QBI

Cache @ ITS

UQ OneNet 100G
Core

25km of highway



Polaris 100G Edge

GPFS AFM “home”

ZWS Archive Disk
Preservation Copy1
@ Polaris

TS4500 + 10 *
TS1160's
Preservation Copy 1

- Back in DMF6 days: No DMAPI integration with GPFS – all cache chained with NFS. *Ugly.*
- A native DMAPI experience. ENOTREADY ability vs "*stale file handle, goodnight*" poor behaviours.
- Very high throughput over HDR fabric – line rate, RDMA + ZCO means very low CPU utilisation on all front ends but lots of low latency throughput.
- The HPE DMF7 backend can instantiate "tier0" representations of the data Storage Scale ingests
- Direct traffic from GPFS to any repository we like (s3, glacier-type, spin down disk technologies, our TS4500 tape libraries etc)
- We tier even further. IBM ESS → {DMF7} --> ZeroWattStorage --> IBM TS4500's
- DMAPI managed regions are understood by Spectum Scale in terms of `mmlsattr` (has DMAPI region outputs) – means Scale is aware of the "offlineness" and treats accordingly.
- We didn't have to "lift and shift" to other technologies, as we had DMF already.
- Just "directories" as AFM targets at home. No more dependent/independent filesets 1000 limit sorrows.
- RAO: When a user asks for lots of data that is offline, that request is passed through Storage Scale to DMF7 which understands RAO in the TS1160 tape drive.

“All my stuff is *mysteriously* deleted, Jake!”

How we recover with GPFS and DMF7 fused together:

1. From the DMF7 MDS/mgmt node, on the GPFS filesystem, we cd to the directory in question…
2. Find what we’re looking for…

```
dmf find -t dmfind \
  ‘object.path like “/gpfs/prodfs1/fset123/mydata*”’
```

Then, once we’re happy with it…

3. We construct a view of the entire object reflection table versions that ever existed as a result of the filesystem light weight event watcher stream GPFS sends to DMF7.

```
dmf stage -t dmfind \
  ‘object.path like “/gpfs/prodfs1/fset123/fset123/mydata*”’ /gpfs/prodfs1/fset123/
```

Which restores the entire tree and creates to /gpfs/prodfs1/fset123/

```
ie. /gpfs/prodfs1/fset123//gpfs/prodfs1/fset123/fset123/mydata/file.bam
```

Then it’s a matter of moving it back in to place.

Data migration tricks that saved our (data) lives.

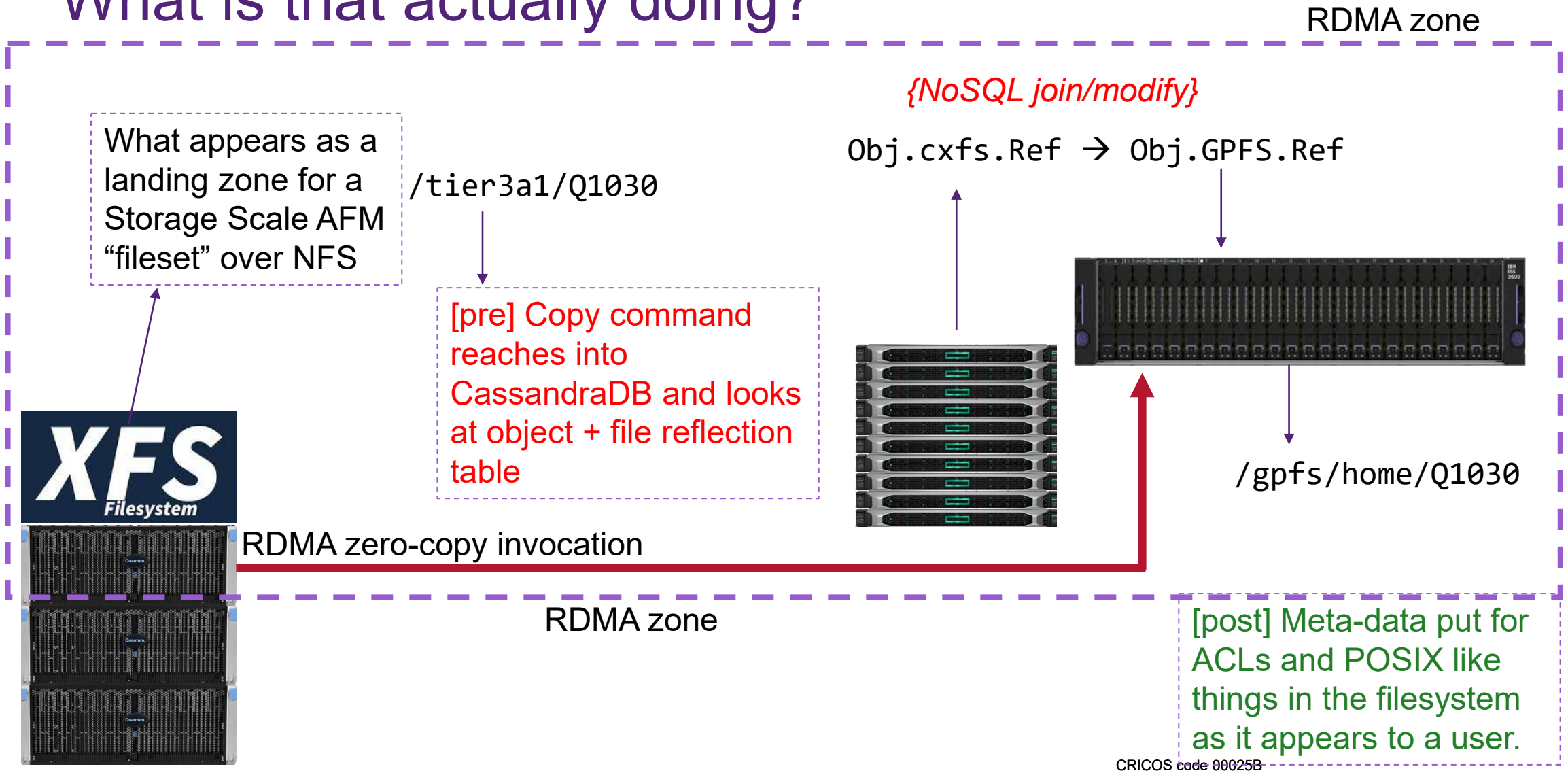
- Unbounded by the limits of traditional data movement techniques or any preconceived ideas – we built workflows with the DMF7 engineering team to move entire fileset backing stores out of Lustre and CXFS and into GPFS.
- Let's say all your data is trapped in a CXFS, or XFS...or another *FS and you want to get it into GPFS but you're doing that with the help of DMF7 behind the scenes:

```
/opt/hpe /dmf-file-copy/sbin/dmf-file-copy --source <source directory> --target <target directory>
```

- Example of us moving an entire fileset ZeroCopy using the power of the parallel filesystem and the NoSQL database cluster behind DMF7 as an object + file reflection table:

```
/opt/hpe/dmf-file-copy/sbin/dmf-file-copy --source /tier3a1/Q1030 --target /gpfs/UQ07/test/tier3c/Q1030 --log-level=debug --overwrite --verbose –dataset=Q1030 --log-target=/var/log/hpe/dmf/dmf-file-copy-Q1030.log
```

What is that actually doing?

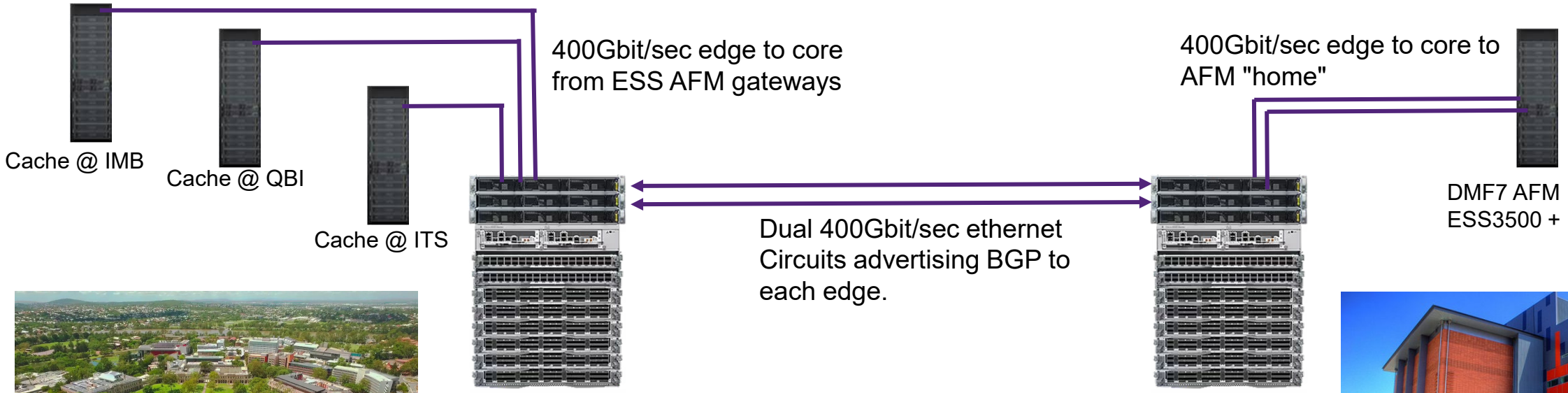


Take away message for IBM to consider:

- Storage Scale can be a cohesive part of a world leading digital preservation systems.
- Storage Scale needs to be thought of as an active contributor in the digital preservation market and high end archiving market as well as the “go fast” market. If you’re clever – you get both. Have cake, eat it too.
- You need to pair it with the right products, integrations and interactions to get the very best.
- Clients and customers like all of us need to be willing to dig in, get very involved and push for those “perfect” integrations to get the best for our organisations – rather than accepting what is necessarily on the table from one vendor. Don’t settle for the workflow they told you would fit and suffice. Push for your own. Build your own boat and work with your technology partners to realise its supportability for the **world**.

Story 5: Storage Scale AFM at 400Gbit/sec.

- We're building a 400Gbit/sec campus supercore.
- My IBM colleagues tell me, nobody has tried to run AFM at 400Gbit/sec before. All the ESS building blocks are now capable of more than 40GB/sec of sustained IO...



UQ St Lucia Campus

CRICOS code 00025B

Polaris Data Centre

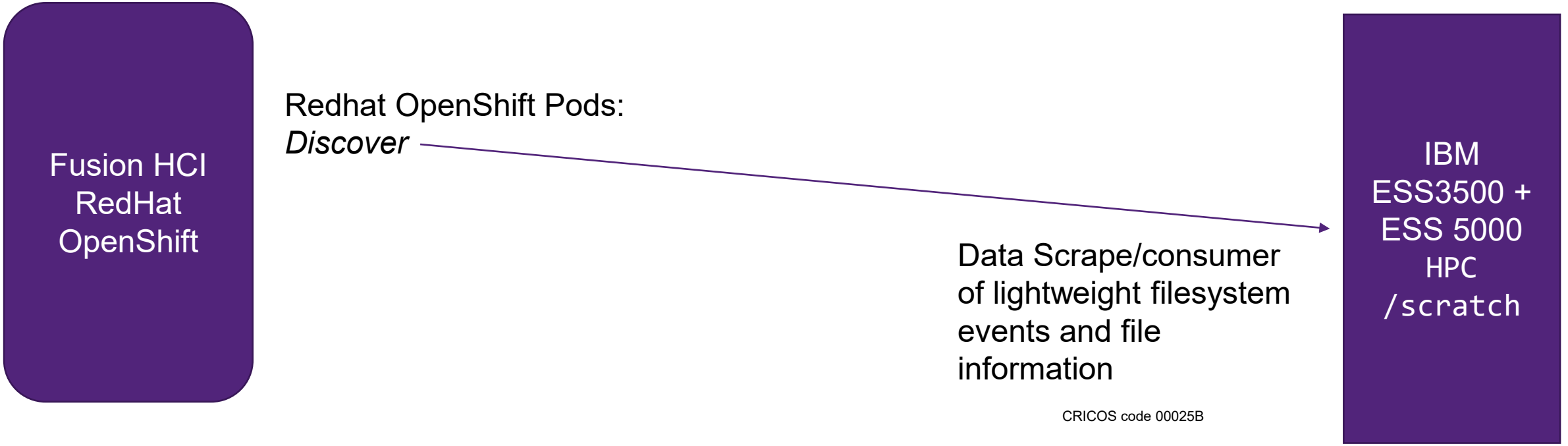
Take away message for IBM to consider:

- Early discussions from our heroic IBM unstructured storage specialist in Australia (Andrew Beattie) and our team frame this as more about our ability to recover a cache quickly in the event of a serious incident, a cache-re-warm-up, replacement or a recovery.

- This might be less about how fast each AFM stream runs from cache to home and more about aggregation of what we can do with that much bandwidth between cache and home. Is the NSD protocol, the AFM stack and all the things that sit in between ready for that?

Story 6: Storage Discover + IBM Fusion HCI

- In conjunction with IBM, we've deployed a new IBM HCI Fusion system.
- We're deploying IBM Storage Discover and targeting our ESS3500 HPC /scratch arrays.
- Why? We're interested in learning more about our data than we currently know about.
- After that, we're keen to then use this to automate workflows for users who can trigger data, steps and processes inside our supercomputers based on tooling inside Discover.



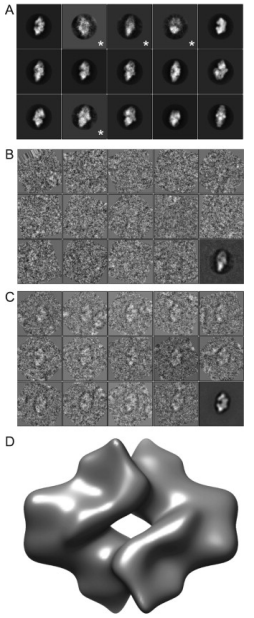
CryoTEM Beamline micrograph generation

100Gbit/sec connected AFM cache

AFM Home

IBM Spectrum Discover

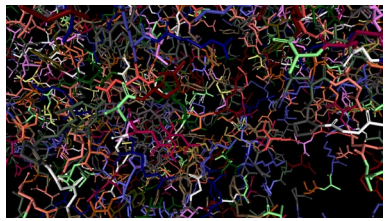
FS.watcher
Tagging file types as they appear on filesystem inside specific filesets in near real time.



Spawn Relion single particle picking workflow on HPC using SLURM

RELION

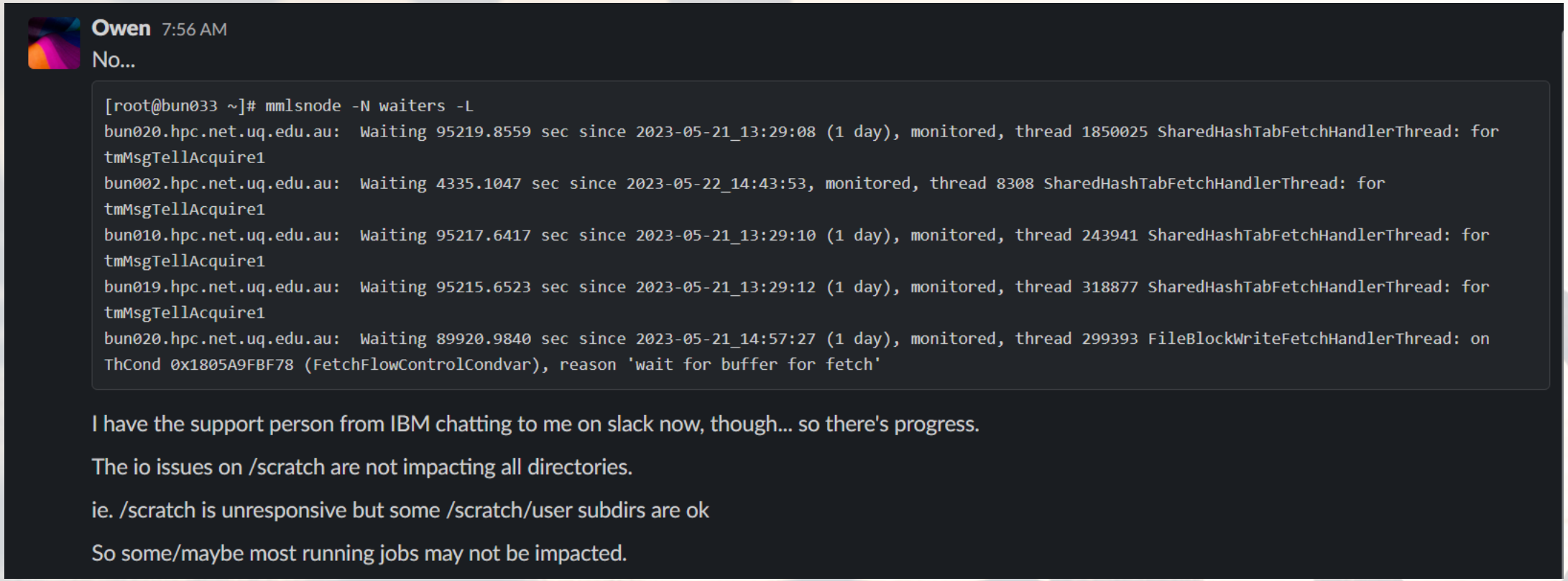
Structural output



Story 7: The support experience

- When the worst happens and disaster strikes, it is fair to say we've had some of the very best experiences with support and the comprehensiveness of practice they provide.
- Incredibly robust post incident review procedures and post mortem incident working groups.
- Attention being paid. A first class experience, honestly.

Story 7: The support experience



Owen 7:56 AM

No...

```
[root@bun033 ~]# mmlsnode -N waiters -L
bun020.hpc.net.uq.edu.au:  Waiting 95219.8559 sec since 2023-05-21_13:29:08 (1 day), monitored, thread 1850025 SharedHashTabFetchHandlerThread: for
tmMsgTellAcquire1
bun002.hpc.net.uq.edu.au:  Waiting 4335.1047 sec since 2023-05-22_14:43:53, monitored, thread 8308 SharedHashTabFetchHandlerThread: for
tmMsgTellAcquire1
bun010.hpc.net.uq.edu.au:  Waiting 95217.6417 sec since 2023-05-21_13:29:10 (1 day), monitored, thread 243941 SharedHashTabFetchHandlerThread: for
tmMsgTellAcquire1
bun019.hpc.net.uq.edu.au:  Waiting 95215.6523 sec since 2023-05-21_13:29:12 (1 day), monitored, thread 318877 SharedHashTabFetchHandlerThread: for
tmMsgTellAcquire1
bun020.hpc.net.uq.edu.au:  Waiting 89920.9840 sec since 2023-05-21_14:57:27 (1 day), monitored, thread 299393 FileBlockWriteFetchHandlerThread: on
ThCond 0x1805A9FBF78 (FetchFlowControlCondvar), reason 'wait for buffer for fetch'
```

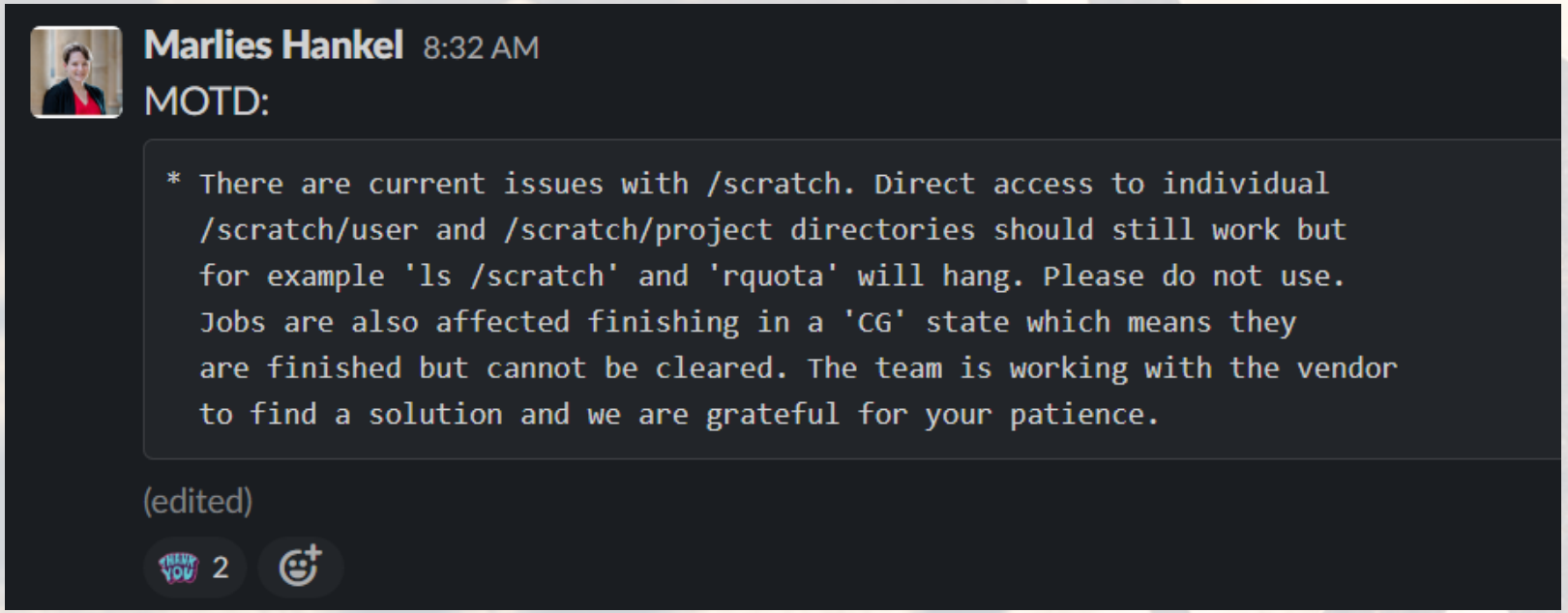
I have the support person from IBM chatting to me on slack now, though... so there's progress.

The io issues on /scratch are not impacting all directories.

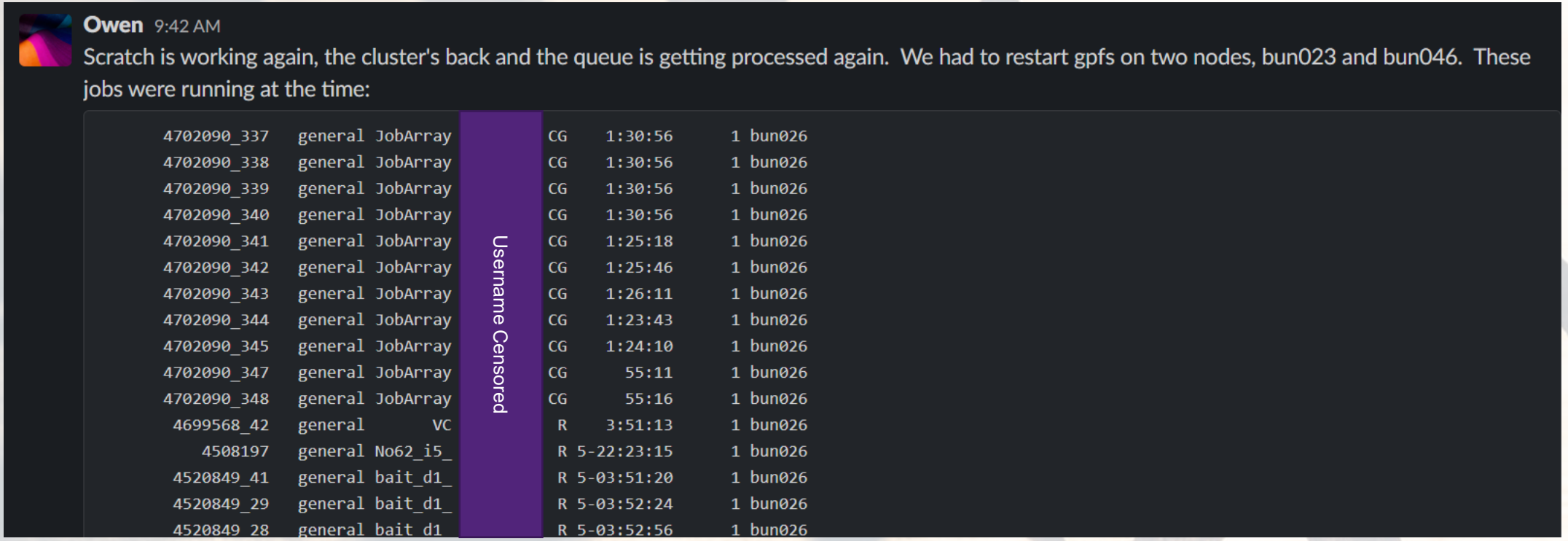
ie. /scratch is unresponsive but some /scratch/user subdirs are ok

So some/maybe most running jobs may not be impacted.

Story 7: The support experience



Story 7: The support experience



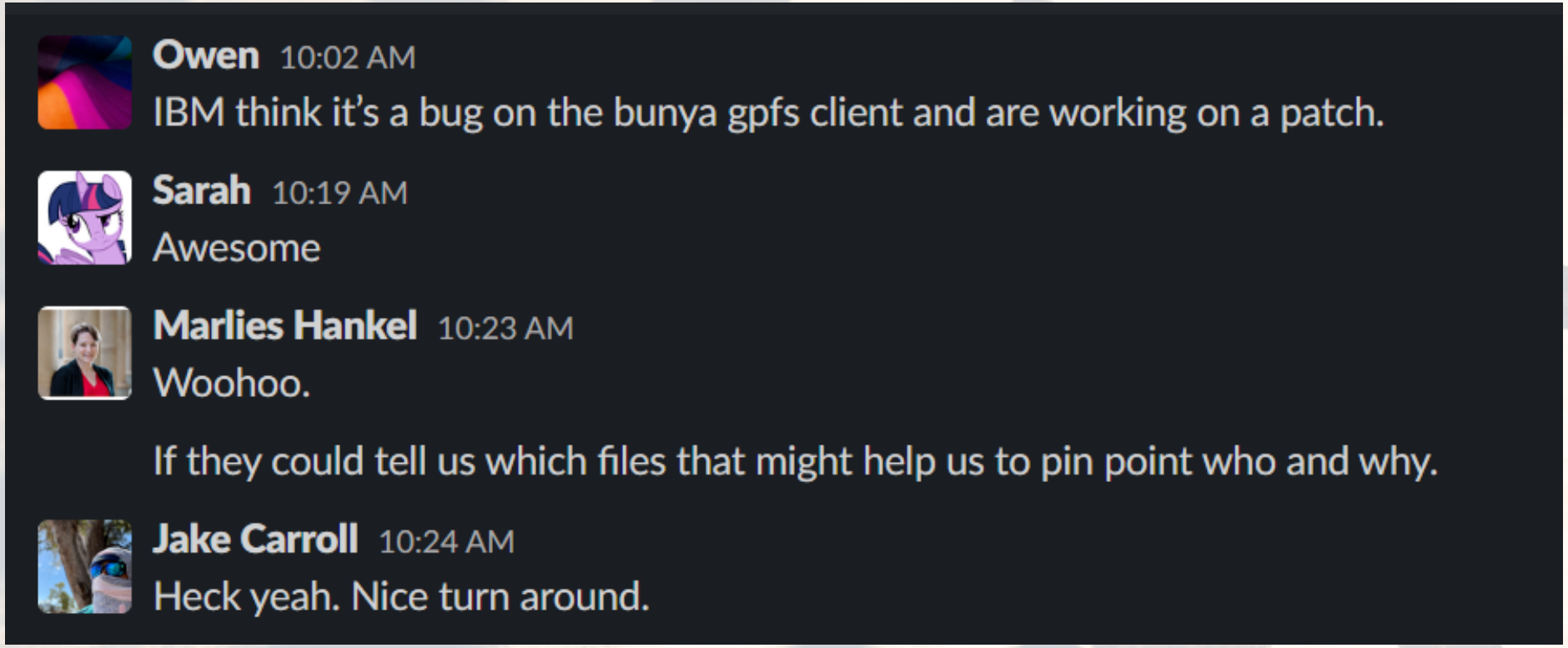
Owen 9:42 AM

Scratch is working again, the cluster's back and the queue is getting processed again. We had to restart gpfs on two nodes, bun023 and bun046. These jobs were running at the time:

4702090_337	general	JobArray	CG	1:30:56	1	bun026
4702090_338	general	JobArray	CG	1:30:56	1	bun026
4702090_339	general	JobArray	CG	1:30:56	1	bun026
4702090_340	general	JobArray	CG	1:30:56	1	bun026
4702090_341	general	JobArray	CG	1:25:18	1	bun026
4702090_342	general	JobArray	CG	1:25:46	1	bun026
4702090_343	general	JobArray	CG	1:26:11	1	bun026
4702090_344	general	JobArray	CG	1:23:43	1	bun026
4702090_345	general	JobArray	CG	1:24:10	1	bun026
4702090_347	general	JobArray	CG	55:11	1	bun026
4702090_348	general	JobArray	CG	55:16	1	bun026
4699568_42	general	VC	R	3:51:13	1	bun026
4508197	general	No62_i5_	R	5-22:23:15	1	bun026
4520849_41	general	bait_d1_	R	5-03:51:20	1	bun026
4520849_29	general	bait_d1_	R	5-03:52:24	1	bun026
4520849_28	general	bait_d1	R	5-03:52:56	1	bun026

Username Censored

Story 7: The support experience



Still a journey. Never truly complete, nor **should** it be.

Thank you for your time, SSUG Deutschland!

…question time.