LSST:UK Consortium

SOMERVILLE CLOUD

Somerville – Edinburgh High Performance Database Service

George Beckett

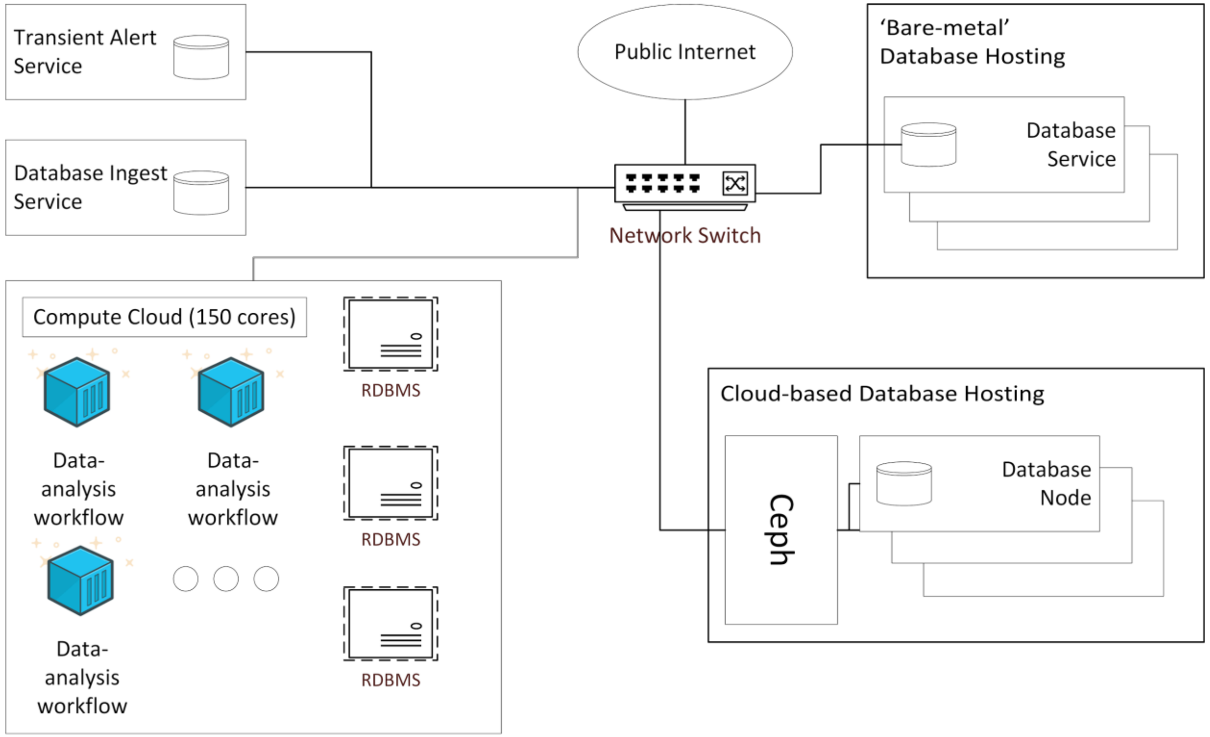
10th May 2022

Background

- In 2017, IRIS (at the time, known as UK Tier 0) submitted proposal to STFC e-Infrastructure Pilot call
  - £1.5M bid titled “A Common Cloud Platform for STFC Science”
  - Infrastructure for RAL, Manchester, Cambridge and Edinburgh
    - Plus consultancy from StackHPC re cloud federation strategies
  - Software infrastructure
    - Data Movement Services
    - VM Manager
- Edinburgh node had in-kind contribution from LSST:UK

Edinburgh Component

- High Performance Database Testbed
  - Target next generation astronomy surveys, which could not be accommodated on traditional database servers
  - Noting Euclid, LSST, SKA, Virgo and Higgs Centre for Innovation
  - Databases of O(10—100 TB)



Phase 1 Infrastructure

- 8×200TB storage nodes, configured as 2×500TB-usable Ceph clusters
  - One production; one experimental
- 4×24-core (w/ hyperthreading) hypervisor nodes (256GB RAM)
- 40 Gbps ‘data’ network
- 2×10Gbps uplink to JANET
- Exposed to users via OpenStack ‘Rocky’ cloud
- 1×24-core Service Node (with 256GB RAM)
- 1 Gbps ‘management’ network

Early Activity

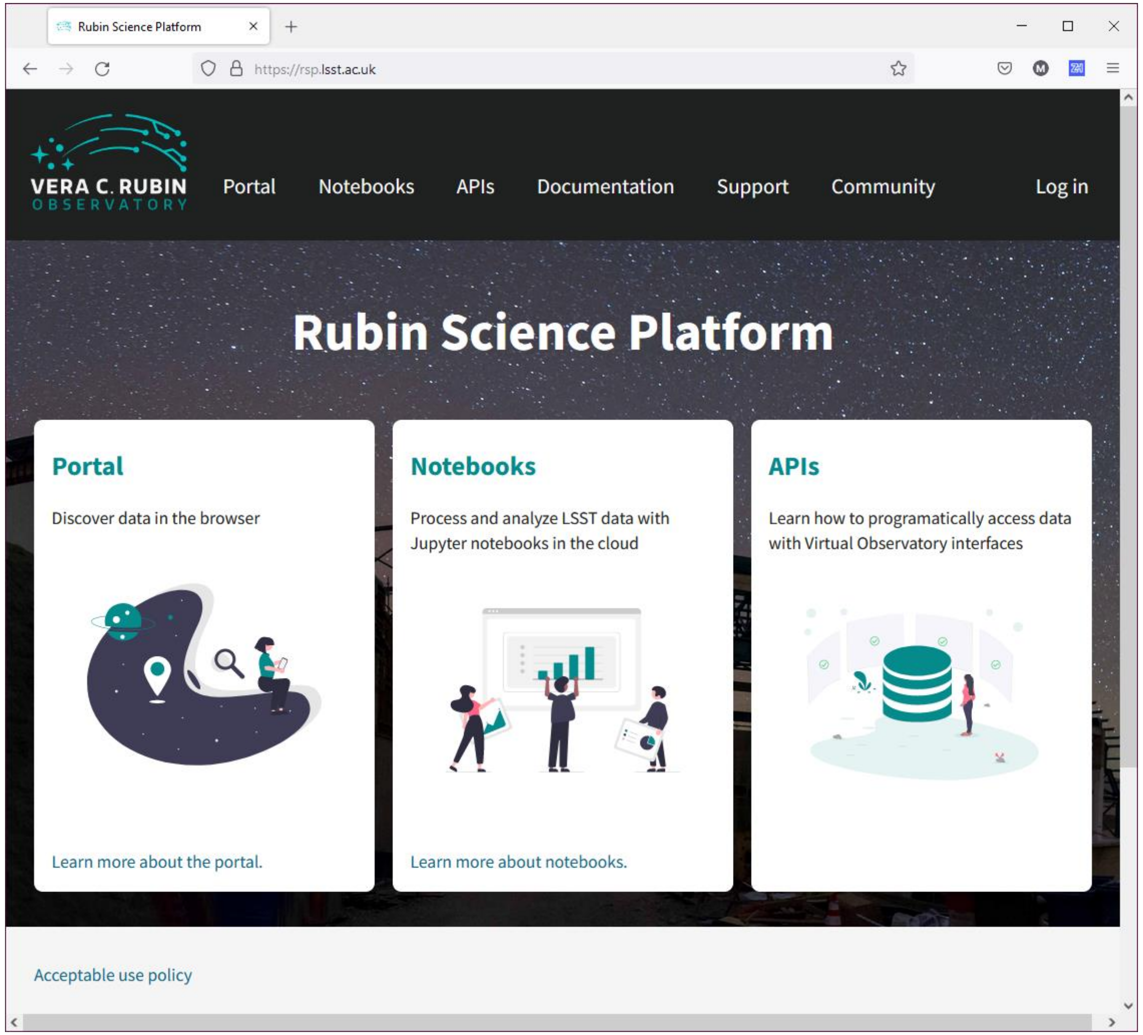
- Database Hosting (WFAU and LSST:UK)
  - WFAU VVV
  - Lasair ZTF
  - PanSTARSS PS1
  - UKIDSS
  - Initial discussions with VIRGO consortium
- Experiments
  - CephFS evaluation
  - Ceph appliance configuration options (block size, RAID configuration, etc.)
  - Database performance – based on WFAU standard server

Lessons Learned

- Support effort (0.3 FTE) insufficient
  - Maintenance difficult to schedule
  - Problem resolution time-consuming
  - Unable to upgrade from Rocky
- Insufficient compute resources (cores)
  - LSST:UK services exhausted compute-node provision
- Need better container solution (e.g., Magnum)
- IP address provision insufficient
- Assignment of Ceph clusters to production and experiment not clearcut
- Difficult to maintain consistency with other IRIS O/S deployments

Successes

- Built up good experience of OpenStack and Ceph
- Successfully deployed large-scale astronomy databases
- Ran production (and pre-production) services for
  - Lasair
  - Rubin Science Platform
- Initiated cloud-computing (people) network
  - StackHPC
  - Edinburgh Notable
  - STFC Research Cloud



Rubin Science Platform

Portal

Discover data in the browser

Learn more about the portal.

Notebooks

Process and analyze LSST data with Jupyter notebooks in the cloud

Learn more about notebooks.

APIs

Learn how to programatically access data with Virtual Observatory interfaces

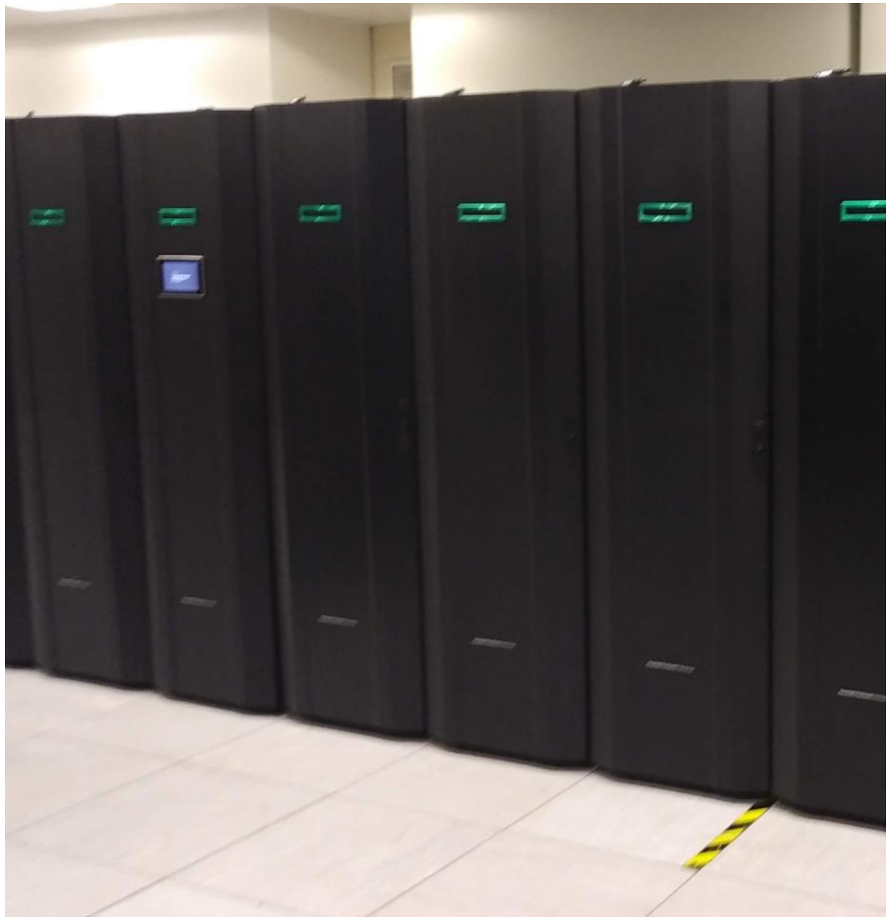
Acceptable use policy

Phase 2 Infrastructure

- In 2021, LSST:UK received infrastructure funding
  - to support in-kind contribution to Rubin Observatory
- Focus on Rubin pre-operations and commissioning
  - Running proto-DAC
  - Hosting and serving Data Previews to early users
  - Supporting Lasair scale-out experiments
  - … plus maintain previous projects
- Engage StackHPC to design and deploy revised Edinburgh OpenStack
  - Plus train staff and upgrade existing infrastructure

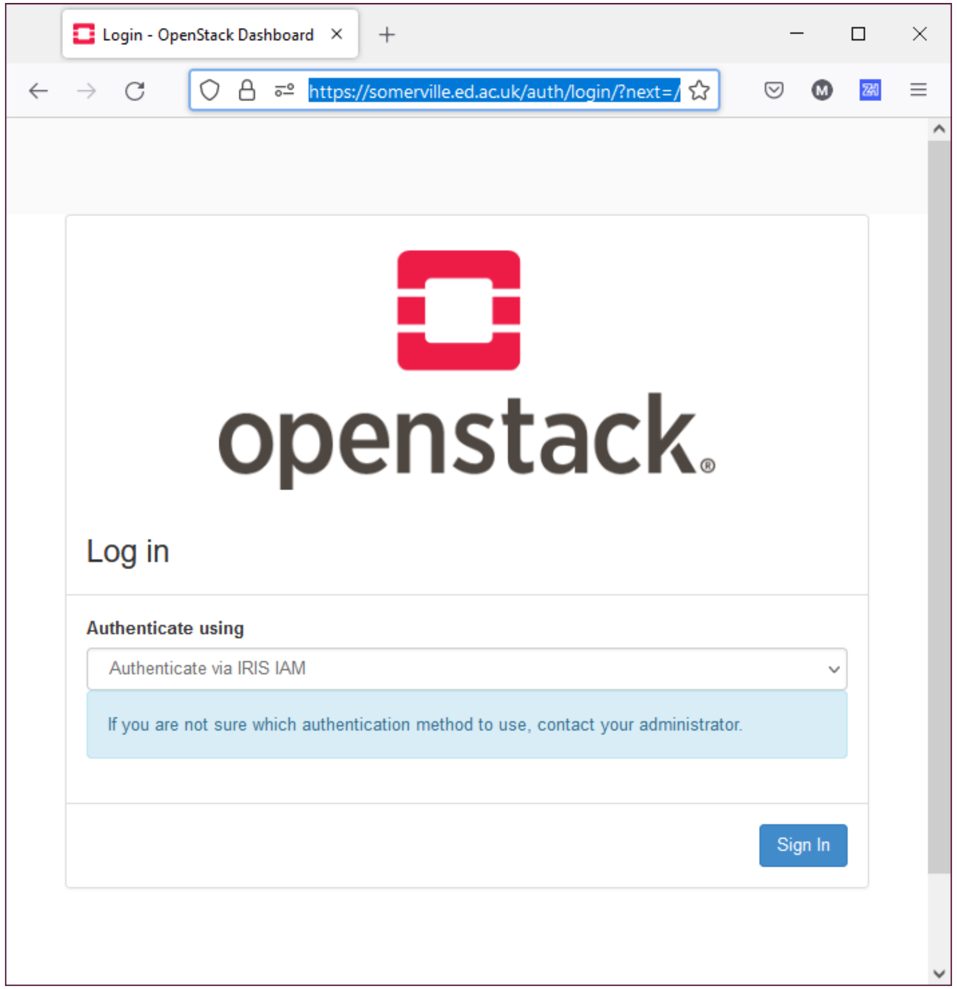
Somerville Cloud

- 7×288TB storage nodes
  - configured as 1 PB (usable) Ceph cluster
- 8×20TB NVMe storage
  - configured as 100 TB (usable) Ceph cluster
- 7×64-core hypervisor nodes (1TB RAM)
- 100 Gbps ‘data’ network
- 2×100 Gbps uplink to JANET
- Exposed to users via OpenStack ‘Wallaby’ cloud
- 1×16-core service node (96GB RAM)
- 10 Gbps and 1 Gbps OpenStack networks



Somerville Service

- Running Scientific OpenStack (StackHPC)
  - Virtualised compute
  - Storage (Posix, block, Object Store, ephemeral)
  - Containerisation (Magnum/ Kubernetes)
  - High-speed Internet access (100 Gbps)
- Covered by 1.2 FTE of support effort,
  - Initially with in-depth support from StackHPC
- Has Class /24 subnet of public IP addresses
- User authn and authz via IRIS IAM
  - IAM groups mapped to Somerville projects



Login - OpenStack Dashboard
https://somerville.ed.ac.uk/auth/login/?next=/
openstack.

Log in

Authenticate using

Authenticate via IRIS IAM

If you are not sure which authentication method to use, contact your administrator.

Performance (DB Queries)

- Aim to achieve similar performance to standard WFAU database
- Benchmark suite
  - Popular WFAU queries
  - LSST exemplar queries
  - Non-indexed trawl (classic, challenging query)
- Tuned instance of SQL Server
  - hosting UKIDSS DR8 detections catalogue
  - ~5 Billion rows (2.1 TB)

Performance (milliseconds)

Query	Phase 1	Phase 2	Phase 2 SSD	WFAU Standard
1	100	54	5	56
2	530	524	405	886
3	103	56	6	50
4	342,323	480,175	103,714	277,760
5	12,176	6,754	4,109	4,956
6	604,000	981,087	178,475	501,573
7	102,940	128,048	35,488	60,720
8	33,720	47,176	12,334	36,776
9	32,720	46,032	12,334	36,503
10	12,890	8,674	4,166	14,910
11	340	243	23	226
12	796,736	1,318,590	232,143	563,860

Performance (normalized)

Query	Phase 1	Phase 2 (*)	Phase 2 SSD	WFAU Standard
1	1.8	1.0	0.1	1.0
2	0.6	0.6	0.5	1.0
3	2.1	1.1	0.1	1.0
4	1.2	1.7	0.4	1.0
5	2.5	1.4	0.8	1.0
6	1.2	2.0	0.4	1.0
7	1.7	2.1	0.6	1.0
8	0.9	1.3	0.3	1.0
9	0.9	1.3	0.3	1.0
10	0.9	0.6	0.3	1.0
11	1.5	1.1	0.1	1.0
12	1.4	2.3	0.4	1.0

Future Plans

- Test and Development system
  - Old hypervisor nodes (out-of-warranty) will be provisioned as a TDS
  - Allow experiments with OpenStack configs/ upgrades without impacting live service
  - Potential to share this service with other sites, if interesting
- Migrate Phase 1 storage as lower-spec storage resource
- Documentation
  - Probably based on Materials for MkDocs with option for community contributions
- Upgrade of RAID controllers on Ceph cluster

Future Plans

- Better integrate into IRIS resources
    - Provide access to other IRIS experiments
    - Potential to join RSAP resource pool
- Contribute to IRIS roadmap
    - Maturation of Scientific OpenStack
    - Federation with other cloud sites
    - Host Scientific OpenStack testbed
- LSST early operations
    - ~0.5 PB of Commissioning and Data Previews catalogues in 2023
    - Rising to ~4 PB by end of 2024
    - …