

SOMERVILLE CLOUD

**Research cloud infrastructure
for databasing and science
archive applications**

Presented by:

Mark Holliman

Wide Field Astronomy Unit
Institute for Astronomy
University of Edinburgh

Somerville Team:

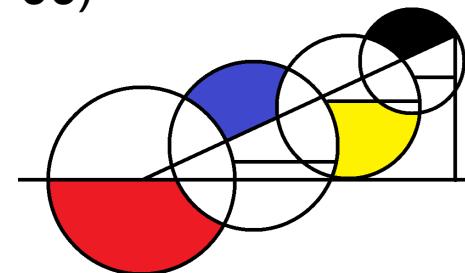
Mark Holliman (WFAU)

Greg Blow (EPCC)

George Beckett (WFAU/EPCC)

Terence Sloan (WFAU)

Daniel Pizarro (EPCC)



CLOUD FOR ARCHIVES

Somerville Cloud provides computing infrastructure for Astronomy and other IRIS experiments, designed to support the complex resource requirements of databasing and science archive applications.



- Scientific OpenStack cloud (<https://www.hpc.cam.ac.uk/scientific-openstack>)
- Housed at the University of Edinburgh's Advanced Compute Facility
- Operated by the Wide Field Astronomy Unit and EPCC of the University of Edinburgh
- Supported by StackHPC (<https://www.stackhpc.com/>)
- Resource allocations and funding coordinated through IRIS



ARCHIVE ON CLOUD

Science archive projects have complex resource requirements which can be difficult to fulfil on generic cloud infrastructure.

These often include specialized data use cases and software services which need to meet performance specifications and policy obligations derived from a variety of stakeholders.

Archives also require a degree of “permanence” not always associated with cloud deployments – data stored and served on them is expected to be accessible to end users in perpetuity and cannot easily be migrated to new locations/services.

ASTRONOMY ARCHIVES

ASTRONOMY ARCHIVES TYPICALLY HAVE TWO TIERS OF ACTIVITY:

- **Processing/Curation Activities** – the active tasks involved in ingesting raw data products and producing science data products for end users
 - **Tasks include image analysis, catalogue creation, database ingests/outgests, crossmatching related archives, etc**
 - **Typically these tasks are resource intensive with high IO patterns, complex workflows, and strict operational schedules**
 - **Curation activities have a known period of operation – only as long as the telescope/instrument is collecting new data**
- **Data Servicing/Hosting Activities** – the ongoing tasks of serving science data products and analysis services to end users
 - **Data products are usually served via web services, providing both GUI and API access, with varying levels of access for different community groups**
 - **Modern archives include value-added services like alert streams, Jupyter notebooks, and user defined bulk reprocessing of datasets too large to copy**
 - **Archive activities provide community value in perpetuity as astrophysics research often utilizes “old” datasets**

SOMERVILLE STRATEGY

BALANCING CLOUD FLEXIBILITY WITH PROJECT NEEDS

Somerville aims to meet the hosting needs of IRIS affiliated Science Archive projects on our OpenStack cloud. Our strategy involves:

- Discussing each project's technical requirements to determine hardware/software solutions
- Comparing requirements between projects to identify similarities which could be addressed with common systems/software, enabling flexible resource deployment and encouraging inter-project cooperation to leverage shared knowledge and experience
- Enabling data sharing capabilities between locally hosted projects via network and storage links
- Deploying tailored hardware/software solutions where necessary for meeting specific project requirements, and planning cloud growth and services to match predicted usage by the hosted archives

SOMERVILLE OFFERINGS

STORAGE SOLUTIONS

- High performance CephFS (backed by NVMe SSDs) – for use in large RDBMS database applications (10s-100s TB)
- High performance local disk (multiple NVMe SSDs on single hypervisors, single/multiple ephemeral disk instance flavours) – for sharded database and distributed dataset applications
- High reliability block/CephFS (backed by HDDs) – designed for bulk storage and flexible sharing of flat file data

INFRASTRUCTURE SERVICES

- Manila shares, backed by either HDD or SSD pools depending on use case
- Kubernetes clusters via Magnum, with ongoing investigation into Cluster API deployment
- Octavia load balancers for on-demand horizontal cluster scaling
- Swift storage for bulk object stores

CURRENT PROJECTS

LSST UK

- Lasair – Zwicky Transient Facility Event Broker
- Rubin Science Platform – User data processing tools and facilities for the Rubin Observatory
- Qserv – Sharded relational database of LSST survey observations
- TiDES - Time-Domain Extragalactic Survey



WFAU SCIENCE ARCHIVES

- WFCAM Science Archive – UKIRT Sky Surveys
- VISTA Science Archive – VISTA Sky Surveys
- OmegaCAM Science Archive
- SuperCOSMOS Science Archive



LSST: QSERV

LSST SURVEY DATABASE

- Qserv is an open source [MPP](#) SQL database system designed originally to host the astronomical catalogs of the [Legacy Survey of Space and Time](#) to be conducted by the [Rubin Observatory](#).
- LSST UK operates a Qserv cluster on Somerville that offers a database capacity of 70TB to external requests, primarily used for testing database capabilities and forward facing user services.
- Qserv uses HDD block storage on Somerville (and test environments with SSD blocks), with Kubernetes as orchestration and Terraform for provisioning. Tests have also been performed using local SSDs, and it is expected that the production implementation will utilize a combination of the two storage solutions.
- By 10th year of the survey, the catalog database is expected to run to approximately 60 trillion rows, requiring more than 10 petabytes of storage before considerations for replication, indices, or other overheads.



LSST: RUBIN SCIENCE PLATFORM

USER DATA TOOLS FOR LSST

- The Rubin Science Platform is a set of web-based tools that allow scientists to access and analyze Rubin LSST data.
- Consists of three main user-facing components, the Portal, a JupyterHub Service & a set of API for programmatic access.
- Deployed as a set of Kubernetes Applications using GitOps (ArgoCD, Helm, Github).
- The Kubernetes cluster is provisioned using Magnum and utilizes several Openstack features:
 - Storage:
 - Cinder Ephemeral Volumes provisioned automatically for Jupyter Sessions
 - NFS Server (backed by Cinder volumes) for persistent storage of user data
 - Object storage (Swift) for API query results
 - Networking:
 - Octavia Load Balancers, and more recently Amphora
- RSP:UK plans to support hundreds of current users, which will require growing number of resources, and addressing scalability & stability issues, i.e improved health monitoring, autoscaling. We also plan on moving from using Magnum to ClusterAPI for creating the Kubernetes clusters.



LSST: LASAIR

LASAIR TRANSIENT ALERT STREAM

- Lasair is a broker for astronomical transient alerts, providing an SQL-based platform for mining the alert stream to identify objects and events of interest. The event stream is generated by the Zwicky Transient Facility (ZTF), a wide angle telescope which images the full sky every 2 nights (<https://www.ztf.caltech.edu/>).
- The Lasair system is a proof-of-concept for the technology stack which will be used to process and serve the LSST event stream to users. It has tested a number of different storage and database solutions using the ZTF stream, with the aim of meeting LSST requirements which are a factor of 25 larger.
- Present deployment uses an Apache Kafka stream (<https://kafka.apache.org/>) and consists of 22 instances serving up 5 databases which are deployed as a mix of traditional RDBMS deployment on Ceph SSD volumes and Apache Cassandra NoSQL databases on ephemeral SSDs. Lasair has 370 user accounts and processes over 400,000 alerts per night (~30 GB per night).
- By 2025 Lasair will begin processing LSST events, growing to ~50 instances with over 10,000,000 alerts per night (~800 GB per night). It will continue to use a mix of Ceph SSD volumes and ephemeral disks for storing the database products, and will also use Swift storage for postage stamp images for each event.

<https://lasair.readthedocs.io/en/main/about.html>




WFAU SCIENCE ARCHIVES

WFAU GROUND BASED ASTRONOMY ARCHIVES

- WFCAM Science Archive: <http://wsa.roe.ac.uk/>
- VISTA Science Archive: <http://vsa.roe.ac.uk/>
- OmegaCAM Science Archive: <http://osa.roe.ac.uk/>
- SupeCOSMOS Science Archive: <http://ssa.roe.ac.uk/>



These archives consist of dozens of relational databases, large volumes of image/flat file data, and a series of GUI and API web services providing end user access.

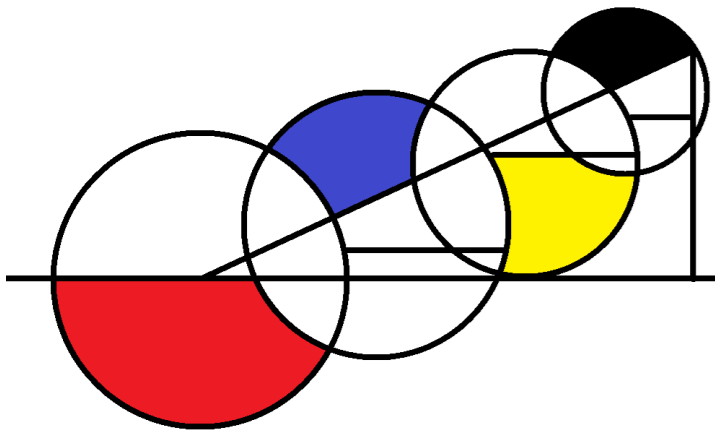
Relational databases range in size from 10's of GB to 140TB, with a total volume of ~.5PB (~200TB of data currently on Somerville, expanding in 2024). Database IO is random and most performant on SSDs.

Image/flat file data is shared across all four archives, with a total volume of 2.5PB (~500TB of data currently on Somerville, expanding in 2024). File IO is serial and performance is acceptable on HDDs.

BEYOND ASTRONOMY

IRIS AFFILIATED PROJECTS ARE WELCOME

Somerville's main projects are all astronomy based right now, with a particular focus on LSST UK activities. But we are available to other IRIS affiliated projects interested in archive hosting on cloud infrastructure. To discuss possibilities and understand your requirements, contact Mark Holliman (msh@roe.ac.uk). After that, resource requests would go through the standard IRIS RSAP process.



SOMERVILLE ONLINE:

[HTTPS://GITHUB.COM/LSST-UK/SOMERVILLE-OPERATIONS](https://github.com/LSST-UK/SOMERVILLE-OPERATIONS)

For questions or requests:

Mark Holliman

msh@roe.ac.uk