CASU @ IRIS

Ground based Astronomical Spectroscopic Surveys Data Processing

Cambridge Centre of Excellence for Astronomical Data (CamCEAD)

Advantages & Capabilities

Scalability & Flexibility

- Dynamic resource allocation via Kubernetes
- Support for multiple concurrent projects
- Elastic scaling for reprocessing campaigns

Global Collaboration

- JupyterHub with configurable resources
- Remote desktop for GUI applications
- Data-proximate analysis for QC workflows

Distributed Processing

- Parallel pipeline execution across nodes
- GPU acceleration for ML/visualization
- 200 GB/day data ingestion capability

Operational Continuity

- Automated backups and redundancy
- Version-controlled data products
- 24/7 availability for survey operations

Platform accessible to CASU staff, collaborators, and approved IRIS users

Architecture & Deployment

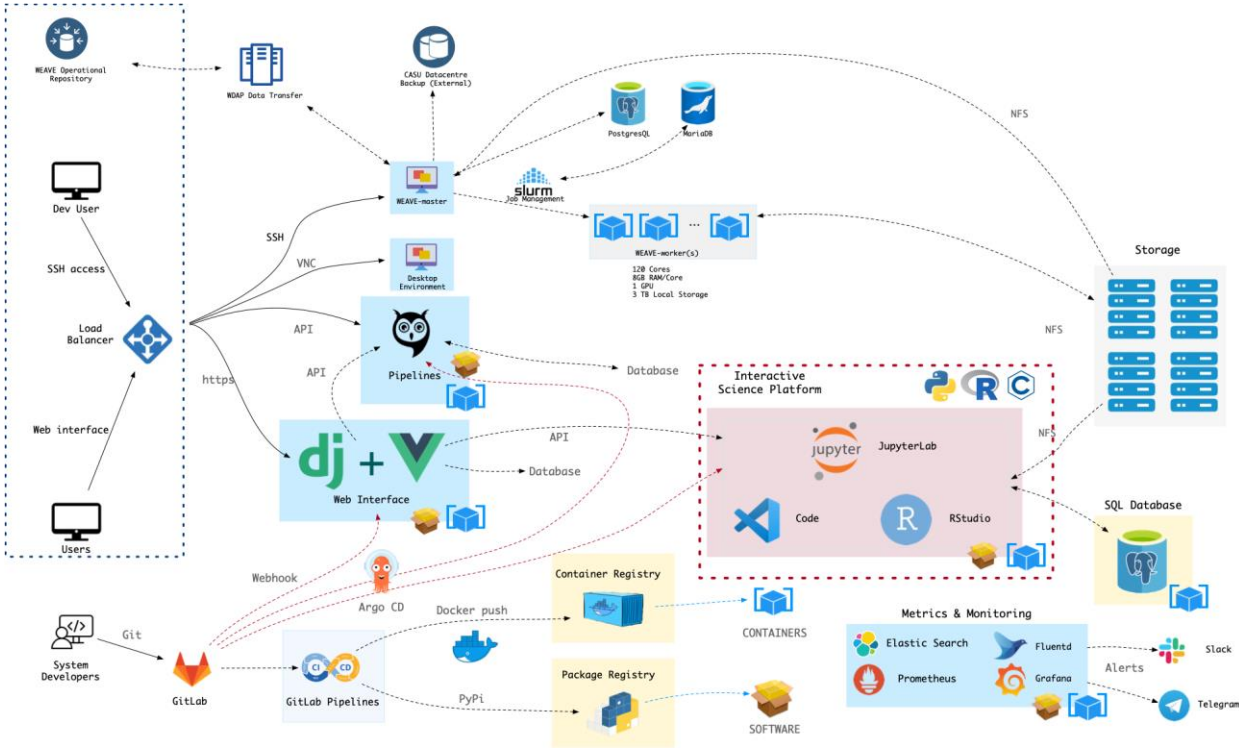
- **Kubernetes** cluster deployed via Kubespray on IRIS OpenStack
- 18 worker + 3 control-plane nodes at Cambridge CSD3
- **2500 vCPU cores, 5 GB RAM/core (12.5 TB total memory) CephFS storage via Manila CSI: 3 PB capacity with SSD tiers**

Compute Access Models:

- **Direct SSH**-accessible instances with **SLURM** job scheduler integration
- **JupyterHub** platform with web interface for interactive analysis
- **Remote desktop** environment (noVNC) for GUI applications
- S3-compatible object storage access via MinIO

Authentication & Management:

- RBAC-based Kubernetes API access with certificate authentication
- Dual-factor authentication for enhanced security
- Prometheus/Grafana monitoring stack



Analysis Platform And Compute Access Models

Direct SSH-accessible instances with SLURM:

- Traditional HPC-style SSH access for batch processing.
- SLURM workload manager for job scheduling and resource allocation.
- Automated pipeline execution and large-scale processing tasks.

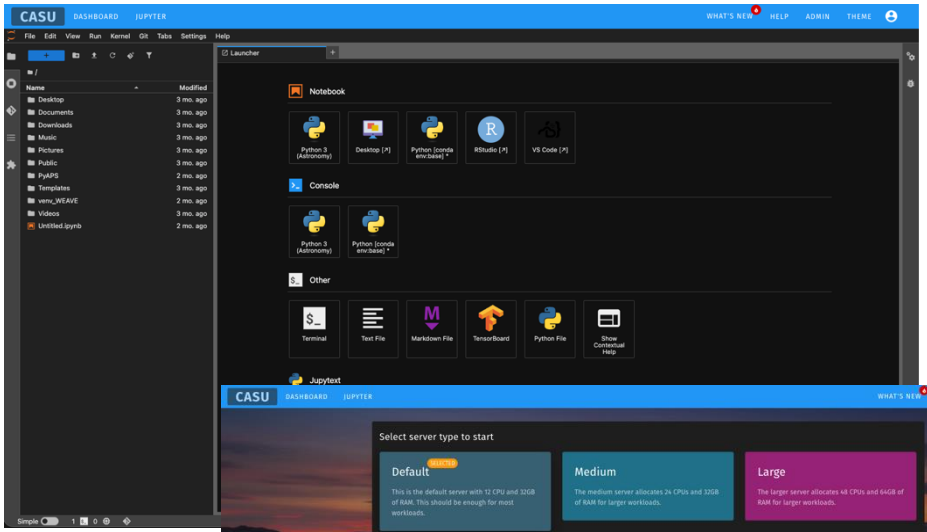
JupyterHub platform with web interface:

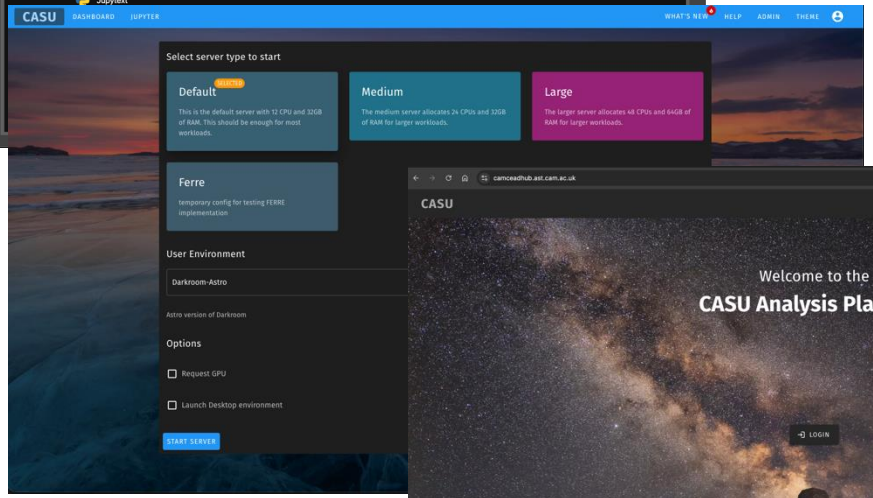
- Browser-based JupyterLab with configurable CPU, RAM, GPU resources.
- Pre-installed scientific stack (AstroPy, NumPy, TensorFlow, PyTorch).
- Supports up to 30 concurrent users with dynamic allocation.

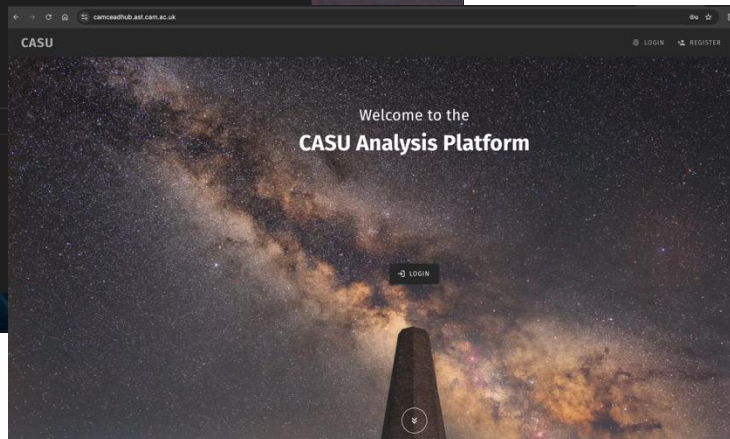
Remote desktop environment (noVNC):

- Web-based graphical desktop, with all essentials for astronomy visualization tools (DS9, TOPCAT, Napari)
- GPU-accelerated rendering for large datasets

https://camceadhub.ast.cam.ac.uk







Current Projects & Resources

Delivering science ready spectroscopic and imaging data

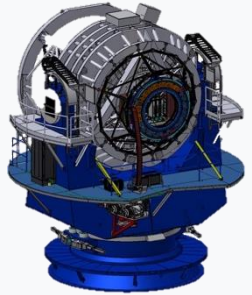
1. **Develop and deploy the data analysis systems** to generate the science data products required to deliver science from fibre fed highly multi-flexed optical and near IR spectroscopy facility instruments: **WEAVE, MOONS** and **4MOST**.

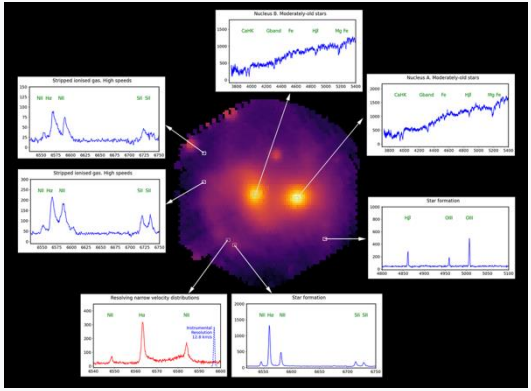
2. **Operate the WEAVE, MOONS and 4MOST data analysis systems**, providing the optimal fully calibrated science data products to the community.

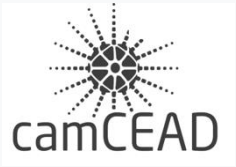
3. **Deploy and operate Data Processing Centre**, providing the capability to run the pipelines.

4. **Builds on heritage**, e.g. pipelines for ESO/VISTA, Gaia-ESO …

5. **Development team embedded in CamCEAD** with wider development for Gaia,
6. PLATO, etc + AI and HTC/HPC







Current Projects & Resources

WEAVE

Advanced Processing System (APS) | L1/L2 pipeline + operational repository

50% resources

1200 vCPU • 12 GPU • 1 PB

MOONS

VLT multi-object spectrograph | Full data-analysis platform

25% resources

800 vCPU • 8 GPU • 200 TB

WST (Pilot)

Wide-field Spectroscopic Telescope | Early workflow development

~12% resources

300 vCPU

QMOST (Fallback)

Redundant backup system | Database snapshots & continuity

5% resources

100 vCPU

Total Allocation — **2500 vCPU • 18 GPU (A100) • 3 PB storage**

Current usage (Oct 2024-Nov 2025): 19.24M vCPU-hours • 69.7B RAM-hours

Cost Comparison: IRIS vs Commercial Providers

Facility	Mode	RAW/night (GB)	L1/night (GB)	L2/night (GB)	AWS Compute cost/day *	AWS Storage cost/day *
WST (estimated)	MOS-LR	30	150–300	600	9880 USD	1500-3K USD
WST (estimated)	MOS-HR	19.5	97.5–195	390		
WST (estimated)	IFS	480	2400–4800	24000–48000		
WEAVE	MOS-LR	1.65	8.25–16.5	33	1440 USD	1500 USD
WEAVE	MOS-HR	3.0	15–30	60		
WEAVE	LIFU LR	4.5	22.5–45	100–130		
WEAVE	LIFU HR	18.0	90–180	300–550		
4MOST	LRS	3.6	18–36	72	3360 USD	1500 USD
4MOST	HRS	2.1	10.5–21	42		

*Cost estimates based on October 2025 pricing

Current usage (Oct 2024-Nov 2025): 19.24M vCPU-hours • 69.7B RAM-hours

Challenges & Future Plans

Current Challenges

I/O Performance

SSD storage critical for pipeline throughput and inter-component communication latency

Infrastructure Stability

Ongoing work with IRIS team to ensure OpenStack changes don't affect deployed services

Future Plans (2026-2030)

2027-2028	2028-2029	2029-2030
2800 vCPU	3200 vCPU	3600 vCPU
24 GPU (Blackwell)	32 GPU (Blackwell)	48 GPU (Blackwell)
3.5 PB storage	4.0 PB storage	4.8 PB storage

Key Drivers: Bulk reprocessing campaigns, MOONS operations (5+ years), 4MOST integration, WST scaling

Collaboration with IRIS/RCS essential for long-term sustainability