



At Ellexus we recognise the unique challenges posed by HPC and scientific workloads.

“Without adequate telemetry, it is not possible to narrow down the many complex choices available when moving to hybrid cloud.”

To avoid the pitfalls of being overwhelmed by the complexity of the cloud, it is necessary to put in place adequate telemetry from day one. Here we look at how to build insight in from the beginning using commodity components in a way suited to HPC workloads and scientific compute.



Job steering

Which jobs should be run on prem and which jobs should be in the cloud? For greatest efficiency, fill the on-prem datacentre with predictable workloads and experiment with bursty or noisy behaved workloads in the cloud. Workloads need to be profiled to provide the insight for these decisions.



Cost management

The cloud can be expensive when not properly managed. Additionally, the task of allocating budget based on high-level business metrics to ensure that every team can meet their goals is complex and needs good telemetry and data pipelines.



Resource sizing and efficiency

The cloud gives engineers the opportunity to access any type of compute and storage capabilities that a workload may require, but this must be done in a cost-effective way to remain in budget.



Agility and innovation

The cloud provides the ability to allow you to scale your architecture up or down as needed, future proofing your storage needs. Take advantage of managed services to free up people's time to be innovative.

“Improving run time often doesn't require extensive rewrites. Knowing where to look is key.”

**Keiran Raine, Cancer Researcher,
Wellcome Sanger Institute**

Using the I/O profiling tools from Ellexus, the Wellcome Sanger Institute saved both time and money on the PanCancer Analysis of Whole Genomes.

The Sanger's project needed the cancer pipelines to be portable and to be tuned for cloud deployment.

As well as reducing run time from 32 hours to 22hrs, as a result of the work with Ellexus, the Institute discovered that it could save a significant 10% of project costs by choosing a cheaper storage option.

Read the full whitepaper, [“Accelerating cloud-based genomics pipelines through I/O profiling for analysis of more than 3,000 whole genome pairs on AWS”](#) at www.ellexus.com



On-prem first vs cloud native



As organisations start to design a hybrid cloud environment, many questions will arise. The best solution for each business depends on whether you intend to maintain an on-prem environment or fully move to a cloud environment. The tools from Ellexus can help you whichever route you choose.

On-prem first

- Understand the compute needs of today so you can plan for tomorrow
- Profile workloads to steer workloads and size cloud compute appropriately
- Use the cloud to isolate problem workloads and make better use of on-prem resources
- Plan to buy hardware for known compute needs going forward for efficiency savings
- Keep HPC expertise in house

Cloud native

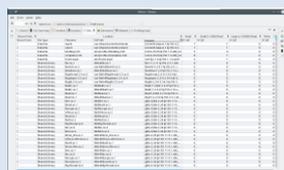
- Build telemetry in from the beginning
- Test and adjust new workflows in the cloud first then move them back on prem when they become stable
- Take advantage of burst capacity for short-lived projects and accelerated time to market
- Use managed services to free up engineers to innovate on what matters for your business
- Use data pipelines to converge efficiency, costs, budget and high-level business goals
- Experiment with instance feedback to find the setup that works best for your organisation

Ellexus data pipelines

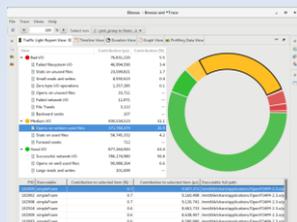
On-demand detailed profiling



Relational data lake
PostgreSQL/MySQL?



Job dependencies
for data hygiene,
regression testing
and migration



Detailed job snapshot
Detailed per-job
dashboard for debug,
optimization and user
education

USERS and DEVELOPERS

Always-on live telemetry



Time series data lake
Elk/Treasuredata/Kakfa/
Databricks?



Operational telemetry
with per job, per user, per
file system and system
overview reporting

Grafana dashboards?

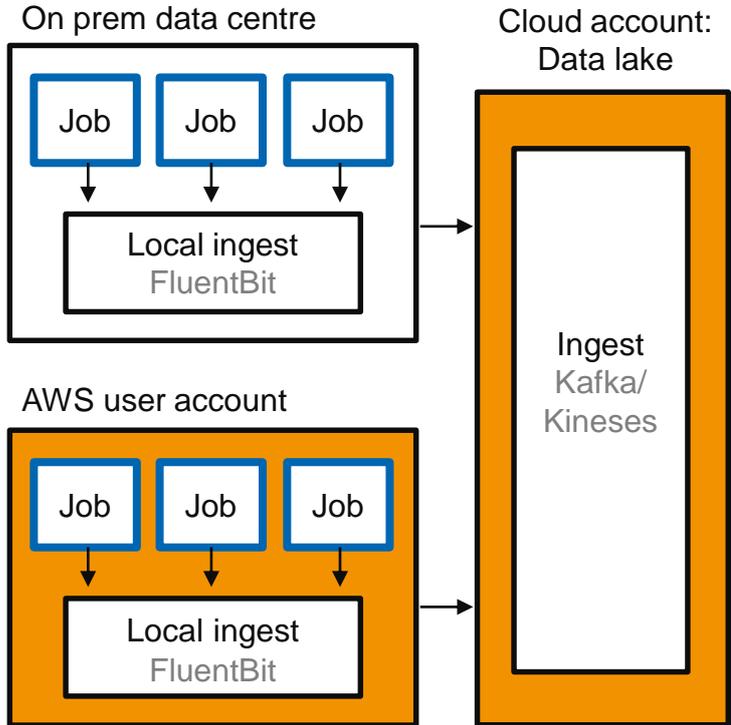
SYSTEM ADMINS



Business intelligence data
pipeline for
forecasting,
chargeback and high-
level optimization

SYSTEM MANAGERS

Ellexus hybrid cloud telemetry architecture



Every application is wrapped in the Ellexus Mistral monitoring technology giving a live feed of CPU, memory, I/O and performance.

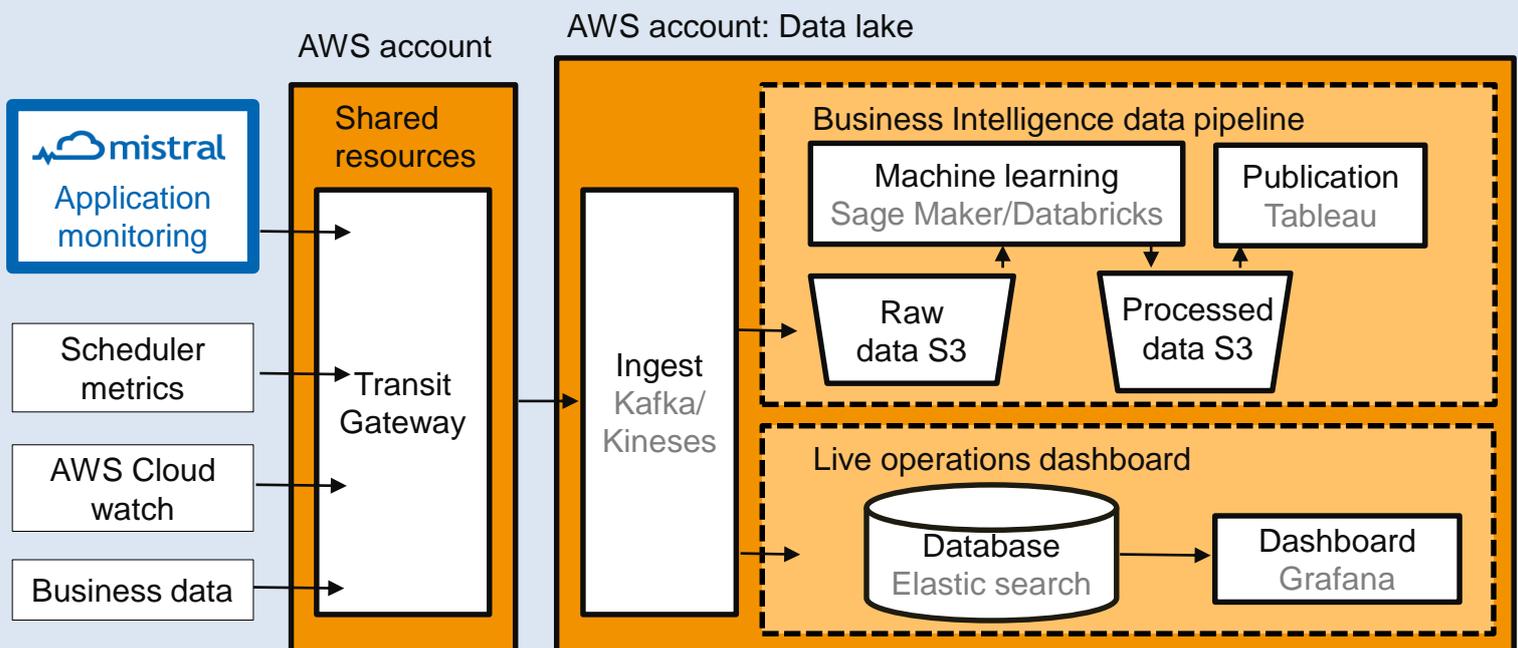
Local ingest can be a local agent such as FluentBit to minimise latency if the jobs are short. Mistral can also push data directly to a distributed Kafka cluster in long jobs where the overhead won't be noticed.

AWS also provides Kafka as a managed service called Kinesis.

The Kafka cluster can span on-prem and cloud infrastructure with multiple data sources and multiple stream subscribers.

Ellexus hybrid cloud telemetry architecture on AWS

Ellexus' solutions are cloud agnostic. The following pipeline can be built from open-source technology on any platform. AWS provides the components through managed services and partner solutions.



It is expected that the RAW data in S3 will be kept and archived, but the database for the live dashboard will only keep data short term.