High performance data

21st ECMWF workshop on high performance computing in meteorology

2025-09-19

This is a discussion session, not the usual slide presentation

I’ve put some discussion questions on slides, but also happy to go off-topic if it’s useful/interesting

Put your hand up if you want to comment, people will bring a microphone to you

Reward: if you make a comment, you get a chocolate biscuit

HPC systems have traditionally had a parallel filesystem such as GPFS or Lustre, with other storage interfaces either layered on top in software or only used for slower external archives.

Object stores have some attractive features such as operating on whole objects, integration with notification/message systems and promises of huge scalability, but existing large applications usually expect to work with a filesystem.

Is the complexity of handling multiple storage interfaces worth it for the object store features, or better to stay with the simplicity of one HPC storage system?

Do you see a useful role for object storage with HPC?

How much structure is appropriate for long term, large volume stores such as tape archives?

ECMWF's MARS has a hypercube structure, Met Office's MASS is hierarchical with some structured data classes, some other agencies have just a big hierarchical filesystem with file/directory naming conventions.

What have you used?

Based on your experience, would you prefer the storage to be more structured or less?

AI model training prefers to work with multi-dimensional cube style data, particularly Zarr.

Traditional large scale storage systems have generally been GRIB records or NetCDF files with a naming convention.

Is it better to accept that transformation into a custom set of files will be needed before feeding AI model training, or make efforts to align the long term storage with AI model training access patterns?

Increasing use of package heavy languages (Python) leads to many small files and small IOs at task startup.

What have you found to work well to minimise the impact of small IO operations?

For example, filesystems that handle many files well, containers or unpacking the environment into node-local storage.