Data Challenges for UK k-scale modelling

Tools and Workflow

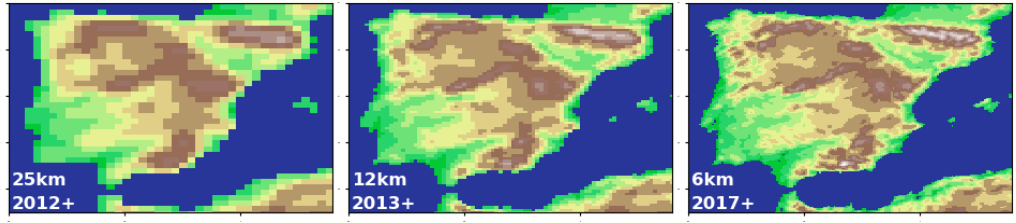
Bryan Lawrence

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Atmopheric Science
and the University of Reading

16th June 2023

Volume — the reality of global 1km grids



What about 1km?

Consider N13256 (1.01km, 26512x19884):

- 1 field, 1 year, 6 hourly, 180 levels
- 1 x 1440 x 180 x 26512 x 19884 = 1.09 PB
- 760 seconds to read one 760 GB (xy) grid at 1 GB/s
- but it's worse that that: 10 variables hourly, $>$ 220 TB/day!

Can no longer consider serial diagnostics, and even parallelised is a challenge for the I/O system!

Data Challenges

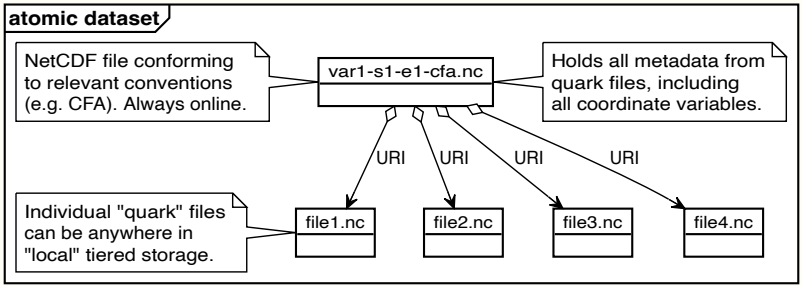
1. Volume (obviously) (io-servers, inline diagnostics, compression, ensemble statistics, reduced resolution output);
2. Native Grid manipulation (finite element support in conventions and tools);
3. Coupling (internal and external, in-fight diagnostics);
4. Intercomparison (common grids, healpix?);
5. Diagnostic Parallelisation (dask);
6. Diagnostic Tools (ESMValTool, underlying libraries, cf-python, IRIS);
7. Number and location of files (aggregation, tape, netzero, and tiered storage);
8. Distributed Data (active storage, dask gateways);

Standards based aggregation

Motivation

- We have lots of files, often lots of files for the same variable.
- Those files can be anywhere on tiered storage, but we tend to talk about “one variable, one-frequency, for one-simulation” as an atomic dataset.
- Can think of an *atomic dataset* being spread across storage. Individual files can be grouped into *quarks* each of which is on one tier of storage.

We can have an aggregate view of those files, stored in an *aggregation file*.

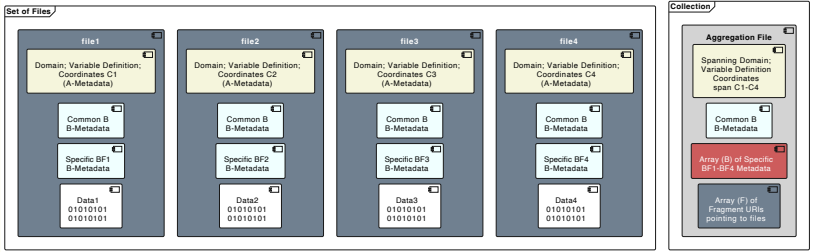


Standards based aggregation

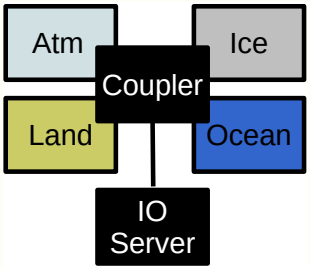
Motivation

- We have lots of files, often lots of files for the same variable.
- Those files can be anywhere on tiered storage, but we tend to talk about “one variable, one-frequency, for one-simulation” as an atomic dataset.
- Can think of an *atomic dataset* being spread across storage. Individual files can be grouped into *quarks* each of which is on one tier of storage.

We can have an aggregate view of those files, stored in an *aggregation file*.



Ensemble Diagnostics

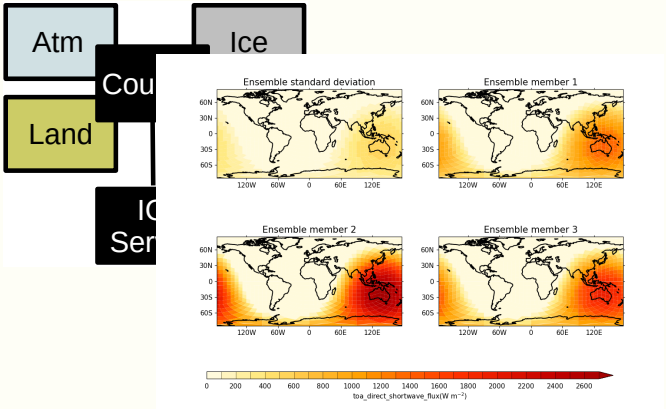


Use the IO server to control multiple ensemble members and carry out “in-flight” ensemble diagnostics.

Can do any kind of reductions “trivially”, more complicated options coming with future versions of XIOS (our I/O server)

(It is not trivial to make all the various MPI communicators play nicely.)

Ensemble Diagnostics

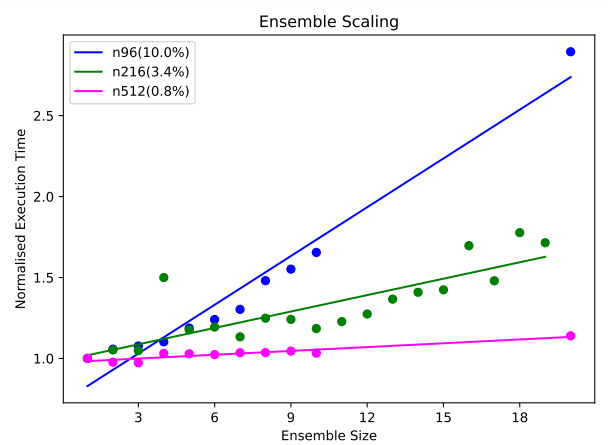


Use the IO server to control multiple ensemble members and carry out “in-flight” ensemble diagnostics.

Can do any kind of reductions “trivially”, more complicated options coming with future versions of XIOS (our I/O server)

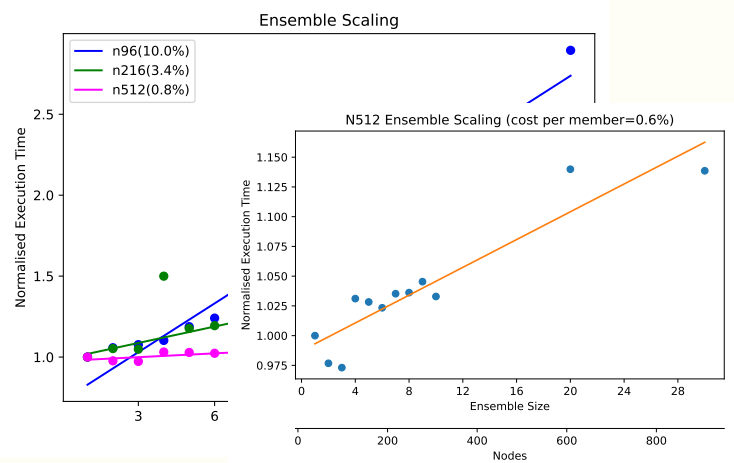
(It is not trivial to make all the various MPI communicators play nicely.)

Scaling Results on ARCHER2



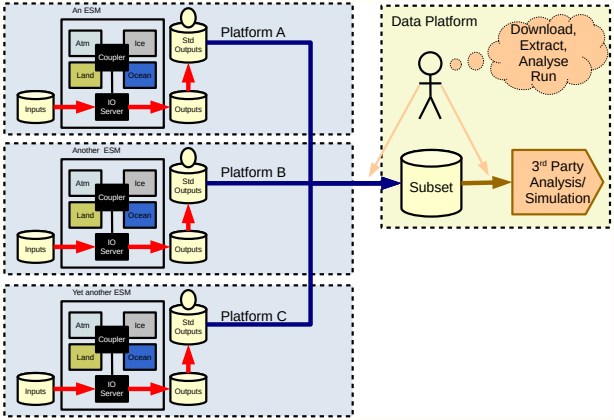
- Ensemble in-flight analysis becomes very cheap at high resolution and very valuable for large ensembles!
- Can consider in-flight ensemble output pruning (ie. avoid writing all ensemble members)

Scaling Results on ARCHER2



- Ensemble in-flight analysis becomes very cheap at high resolution and very valuable for large ensembles!
- Can consider in-flight ensemble output pruning (ie. avoid writing all ensemble members)

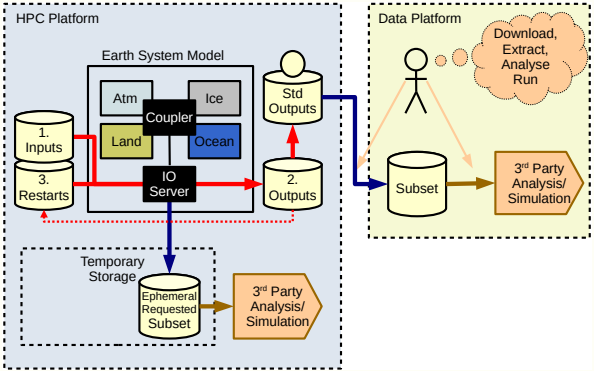
Data Workflows - CMIP6 era



- Outputs converted to standard outputs locally.
- Analysis of detailed outputs and full set of standard outputs is a local activity.
- Subsets of data brought from multiple sites to a common platform (e.g. JASMIN) for analysis.
- Considerable effort to standardise data.
- Data analysis separated from production over both space and time: issues for documentation and citation!

Data Workflows - into the future

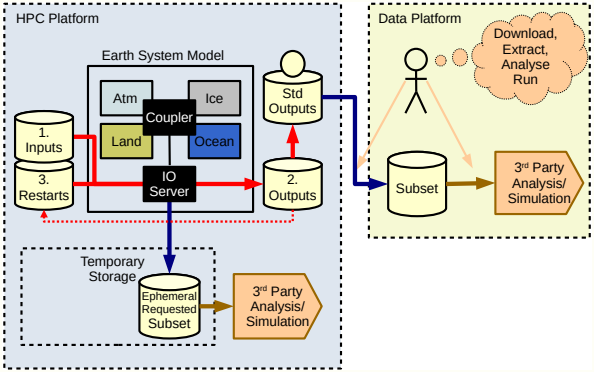
All of the issues from the CMIP era and some new ones!



- In-flight diagnostics supported by a coupler and/or IO-server (visualisations, ensemble diagnostics, downstream models using high frequency data).
- Data published to data platforms for wider sharing and analysis,
- Data platforms supporting co-located computation to support *bringing compute to the data*.

Data Workflows - into the future

All of the issues from the CMIP era and some new ones!

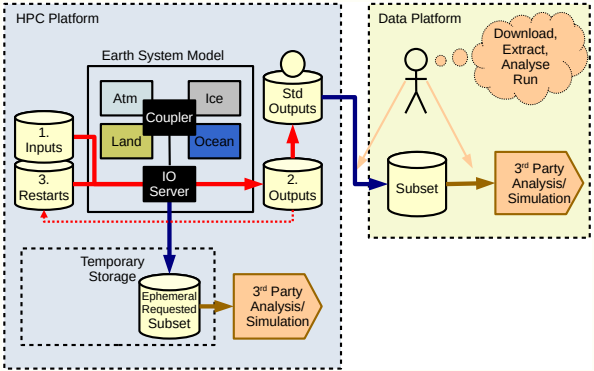


- In-fli… …d by a coup… (visu… diagn… …els using…
- Data p… …ms for wid…
- Data p… …pporting co-located computation to support *bringing compute to the data*.

Cloud has a place in data platforms, but it will not be the only solution.

Data Workflows - into the future

All of the issues from the CMIP era and some new ones!



- In-flight diagnostics supported by a coupler and/or IO-server (visualisations, ensemble diagnostics, downstream models using high frequency data).
- Data published to data platforms for wider sharing and analysis,
- Data platforms supporting co-located computation to support *bringing compute to the data*.

Our view of producers and consumers will have to change as we treat large modelling projects more like satellite missions:

- Well advertised in advance, community dicussions about what is important, well documented, etc.
- We will need to invest even more in our data systems and standards!

Data Challenges

1. Volume (obviously) (io-servers, inline diagnostics, compression, ensemble statistics, reduced resolution output);
2. Native Grid manipulation (finite element support in conventions and tools);
3. Coupling (internal and external, in-fight diagnostics);
4. Intercomparison (common grids, healpix?);
5. Diagnostic Parallelisation (dask);
6. Diagnostic Tools (ESMValTool, underlying libraries, cf-python, IRIS);
7. Number and location of files (aggregation, tape, netzero, and tiered storage);
8. Distributed Data (active storage, dask gateways);