

Report on the Earth System Grid Global Federation meeting 2009-10-05

Stephen Pascoe, Bryan Lawrence

Introduction

This document summarises the architecture and processes of the global federation assembled to provide data services for CMIP5, as they were discussed at a meeting in Hamburg on the 5th of October, 2009. The intended audience is interested parties who were unable to attend, especially those in collaborating projects such as IS-ENES.

The meeting involved wide ranging discussion on many aspects of the CMIP5 project and, in the interests of providing a high level overview, this account necessarily lacks some detail.

Attendees

Anne Chervenak (ISI), Pauline Mak (ARCS)

Bryan Lawrence, Stephen Pascoe, Phil Kershaw (BADC)

Luca Cinquini, Don Middleton, Eric Nienhouse, Nate Wilhelmi (NCAR)

Dean Williams (PCMDI), Ben Evans, Joseph Antony (ANU)

Sergai Nikonov (GFDL), Mark Elkington (UKMO)

Frank Toussaint, Michael Lautenschlager, Stephan Kinderman (DKRZ)

Jean-Yves Peterschmitt, Sebastian Denvil (IPSL)

Access Control

Access to the data in the global archive will be controlled by PCMDI authorisation, although individual data providers will be able to provide access to their own data via their own authorisation procedures.

There is still some controversy as to what it will mean to “open” the archive, and what criteria will be used when to control access, so the comments about data access criteria and timing below need to be treated as preliminary.

Initial timeline

The timeline is still in a state of flux, both from the technical point of view, and from the WGCM point of view. The following is our best guess:

Nov. 2009 - Jan. 2010: The first data will be published by CMIP5 data nodes. No version control will be in place and data is subject to change without notice. The data will be made available for testing purposes only. Authorisation for access will be via PCMDI, likely only to a very limited group.

Feb. 2010: Version control in place. A wider group is likely to be allowed access, possibly those authorised as being from the modelling centres themselves.



Summer 2010: All model runs completed. Analysis begins

Dec. 2010: Public release of CMIP5 (archive opens). Access will be open to all registered users.

May 2011: Journal articles accepted for the 2nd AR5 WG1 meeting

Spring 2013: IPCC AR5 published

ESG-CET Architecture

The CMIP5 archive will be built on the ESG-CET software infrastructure [ESG-CET]. To summarise, the main components are:

- **Data node:** A server running the ESG-CET data node software. A data node provides download and other services for the data it holds.
- **Gateway:** A server running the ESG-CET gateway software. A gateway provides a web interface to the data available on *data nodes* in the federation. Gateways can share metadata via inter-gateway metadata exchange.
- **ESG Publisher:** A tool that is run on a *data node* to inform a *gateway* of new data held at the data node. Metadata is sent to the *gateway* thus making that data discoverable via that *gateway*.

The network of gateways and data nodes serving the CMIP5 archive will be referred to here as the **Earth System Grid Federation (ESGF)** to be distinct from the “Earth System Grid” project (in its incarnation as ESG-CET) which is providing the software.

Within the ESGF there will be several types of Gateway and data node. PCMDI will operate a central Gateway that will register authorised users. Single sign-on will be implemented via OpenID. PCMDI will also coordinate replication of the core standard output (see replication below).

Some sites will deploy both a gateway and a data node. In this case data will be initially published to their local gateway and then federated via inter-gateway metadata exchange. NCAR and BADC are planning to operate in this configuration.

Some sites, typically those associated with a modelling centre, will only deploy a data node and are expected to publish to the PCMDI gateway, although any gateway within the federation could be used. GFDL is planning to operate in this configuration.

Data description

The CMIP5 experiments are divided into 3 tiers: Core, Tier 1 and Tier 2 [CMIP5]. A subset of the output from the core experiments, known as the standard output, will be replicated in a number of places: certainly PCMDI, BADC and DKRZ, and probably in ANU, ORNL and the University of Tokyo. The standard output is expected to be in the range 0.75-1.2 PB.

All data will be divided into atomic datasets identified using the Data Reference Syntax [DRS] that aggregates the data by modeling centre, experiment, scenario, variable, time resolution and ensemble member. Each atomic dataset is expected to comprise of a number of CF-NetCDF files conforming to the CMOR [CMOR2] standard.

Additional metadata describing the models used and how they conform to the CMIP5 experiments will



also be required, with WGCM having agreed that modelling centres will be required to provide this information via a detailed questionnaire being developed by the METAFOR team, *before* the data can be made available. The information from that questionnaire will be imported into the ESGF catalogues, amongst others.

Note that there is a conflict between the expectation that data will begin being published in Nov 2009 and this requirement to have metadata in place before publication. That conflict will be resolved shortly.

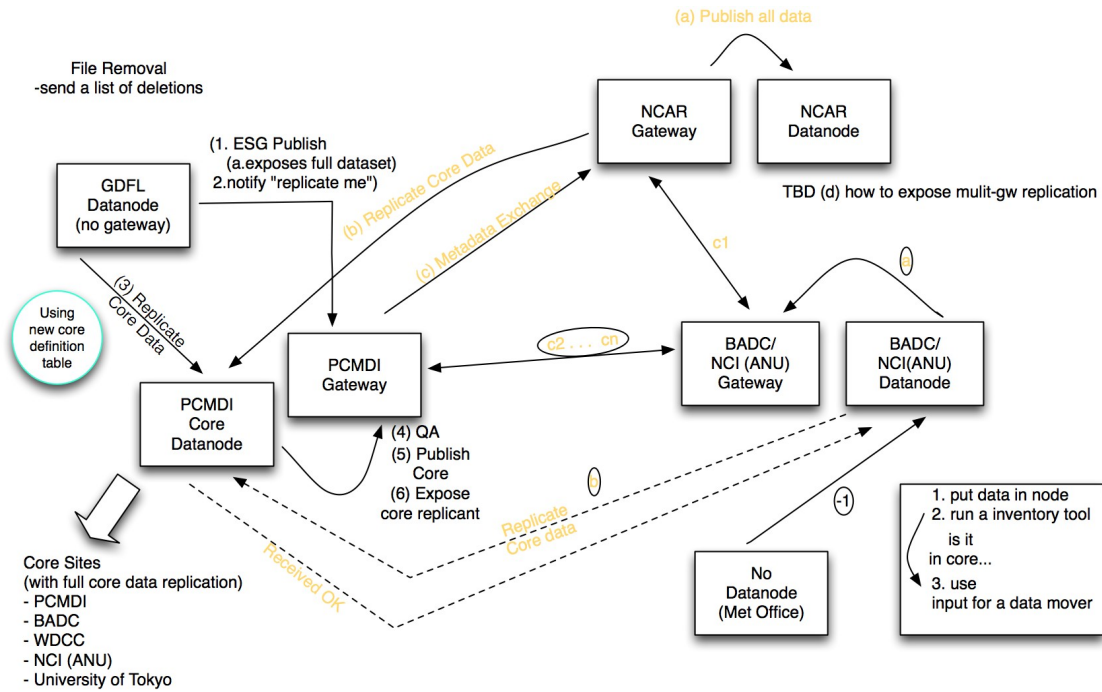


Figure : The identified data flow use cases for the CMIP5 archive (Courtesy of Joseph Antony and Ann Chervenak)

Data Replication

As noted above, standard data will be replicated amongst some “core” data centres.

Data will be initially published on data nodes.

Replication can be divided into 2 processes:

1. Movement of data. Pump-priming of the replicated data is likely to be done via physical disk. PCMDI are procuring a set of 8TB disks for this purpose, based on the estimate of 16TB of cores standard output per modeling centre. To minimise bottle necks, disks will be sent between replicating centres as a round robin. Updates to data will done over the network when feasible.
2. Re-publication at the replicating data node. ESG-Publisher will be run on the replicated data, publishing it to the target gateway. ESG-Publisher and the Gateway software are being adapted to flag replicated data.



How these processes are triggered has not been resolved and will probably occur manually in the first instance.

How the core standard output will be identified for replication has not been resolved. The core standard output is expected to be defined in a set of CMOR tables. Two scenarios were discussed:

1. The replicating data node (generally PCMDI) could identify the data to replicate from the metadata published to it's gateway and request that data from the data node.
2. The source data node could identify the core standard output and send it to PCMDI.

Where and when quality assurance will take place has not been resolved, although PCMDI is expected to do QA on the core standard output it receives.

The effect of replication will be that multiple copies of replicated data will be visible at some gateways. For instance in the case of GFDL it will publish data from its data node to the PCMDI gateway then send the core standard output to PCMDI who will then re-publish the replica to the PCMDI gateway, resulting in 2 copies being visible from the PCMDI gateway. If inter-gateway metadata exchange is considered gateways could receive metadata on multiple replicas. How this is managed within the gateway has not been resolved.

Versioning

We expect modelling centres to identify problems with published data and publish revised datasets. In a distributed, replicated archive it is vital that these revisions are recorded and made available to users.

Nate Nienhouse presented a plan for implementing version control of a data node's dataset hierarchy that includes cascading version numbers up the hierarchy [ESG-VCS]. It was agreed that this system was unnecessarily complex for the current requirements and timescale. The level of granularity of version control was discussed. If we record only changes to atomic datasets there will be a storage overhead when only 1 file has been changed in the dataset. There was some discussion on whether this case would be common in practice and whether the overhead was acceptable to reduce complexity in the versioning system. It was agreed that versions would be attached to individual files and atomic datasets would have version numbers calculated from the versions of their constituent files. Although the importance of fixing *version-number* as a component in the DRS was identified, the precise mechanism by which versions would be encoded was not agreed.

Gateways

BADC, DKRZ and ACU are planning to deploy gateways but are yet to deploy and test the software. Unlike the data node software, the gateway software has not yet been packaged for external deployment by ESG-CET and therefore some help will be required from the ESG-CET team to deploy it at the other sites. Due to scarce developer resources at NCAR and PCMDI it was agreed to start testing the gateway software in November. A new mailing list for Gateway installers will be created.

Process

The meeting identified a set of documents that are needed ASAP to clarify the effective running of the



federation including:

1. Replication procedure
2. Replication software description
3. Data node installation
4. Detailed milestones and time line
5. Resolution of conflict between data availability and satisfactory metadata

A series of regular targeted teleconferences are being arranged to address the technical and procedural issues identified in the meeting. Details of these telcos, and an updated list of documents, will be published on the go-essp-tech mailing list.

References

[CMIP5] Coupled Model Inter-comparison Project 5. <http://cmip-pcmdi.llnl.gov/cmip5>

[ESG-CET] Earth System Grid Centre for Enabling Technology. <http://esg-pcmdi.llnl.gov>

[DRS] CMIP5 Data Reference Syntax. http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax_v0-20_clean.pdf also see the latest version from http://cmip-pcmdi.llnl.gov/cmip5/data_documents.html

[CMOR2] Climate Model Output Rewriter. http://cmip-pcmdi.llnl.gov/cmip5/output_req.html

[ESG-VCS] Data Publication, ESG-CET NCAR wiki.
<https://wiki.ucar.edu/display/esgcet/Data+Publication>

