

How will CMIP5 Model Output be Archived and Served?

Dean Williams and Karl E. Taylor (PCMDI)

Bryan Lawrence (BADC), and Michael Lautenschlager (WDCC)

Please note that the information in this document may become dated rather quickly as the ESG Federation is established. If the current date differs from the document version date, check the CMIP5 website to ensure that a later version has not been produced.

1. Introduction

As with CMIP3, PCMDI is charged with facilitating access to a unified database of model output from CMIP5 simulations. PCMDI will collaborate with partners (at data centers elsewhere who have been funded to help support CMIP5 and AR5 data delivery) and will also rely on the climate modeling centers to contribute substantially to the effort. Under the leadership of PCMDI and with the help of these other groups, the data management needs of CMIP5 will be met by an internationally federated, distributed data archival and retrieval system, referred to as the Earth System Grid Federation (ESGF). The federation will exploit software primarily developed by the Earth System Grid Center for Enabling Technologies (ESG-CET) consortium in the U.S.

A single CMIP5 model will typically produce tens of terabytes of the requested standard output (http://cmip-pcmdi.llnl.gov/cmip5/data_description.html), which will all be made available to the scientific community by the ESG Federation. A substantial fraction of the requested standard output will comprise what is called the “core data”¹, and this core data, because it is expected to be in relatively high demand, will be replicated, at least at five sites (PCMDI, BADC, WDCC, ANU, and University of Tokyo/JAMSTEC), referred to here as “Core Data Nodes”. The final list of core data is in preparation.

The core data, because it has been replicated, will provide some system-wide redundancy, which will increase data availability when there is a disruption in a single site’s services. It also will circumvent the transfer of data from multiple data nodes (and speed server-side processing) when requests are received to return, for example, a multi-model mean field, since the data from all the models will all reside locally.

The ESG Federation data delivery system will consist of two major components (see figure 1 below):

¹ The “core data” will be extracted from all the simulations performed as part of CMIP5, not just output from the so-called “core experiments” referred to in Taylor (2008; http://cmip-pcmdi.llnl.gov/cmip5/docs/Taylor_CMIP5_design.pdf).

1) Data Nodes, where the data actually reside, and 2) Gateways, which support portal services and serve as interfaces to end-users who can search, discover, and request data and data products. The model output is “published” to the Data Node, which makes the data visible to a Gateway and enables its delivery to end-users. It is hoped that most modeling centers will install the Data Node software through which they can serve their CMIP5 model output. The Gateway software, on the other hand, is expected to be installed only at a small number of sites already devoted to serving data (e.g., PCMDI, BADC, WDC, ANU, and University of Tokyo/JAMSTEC).

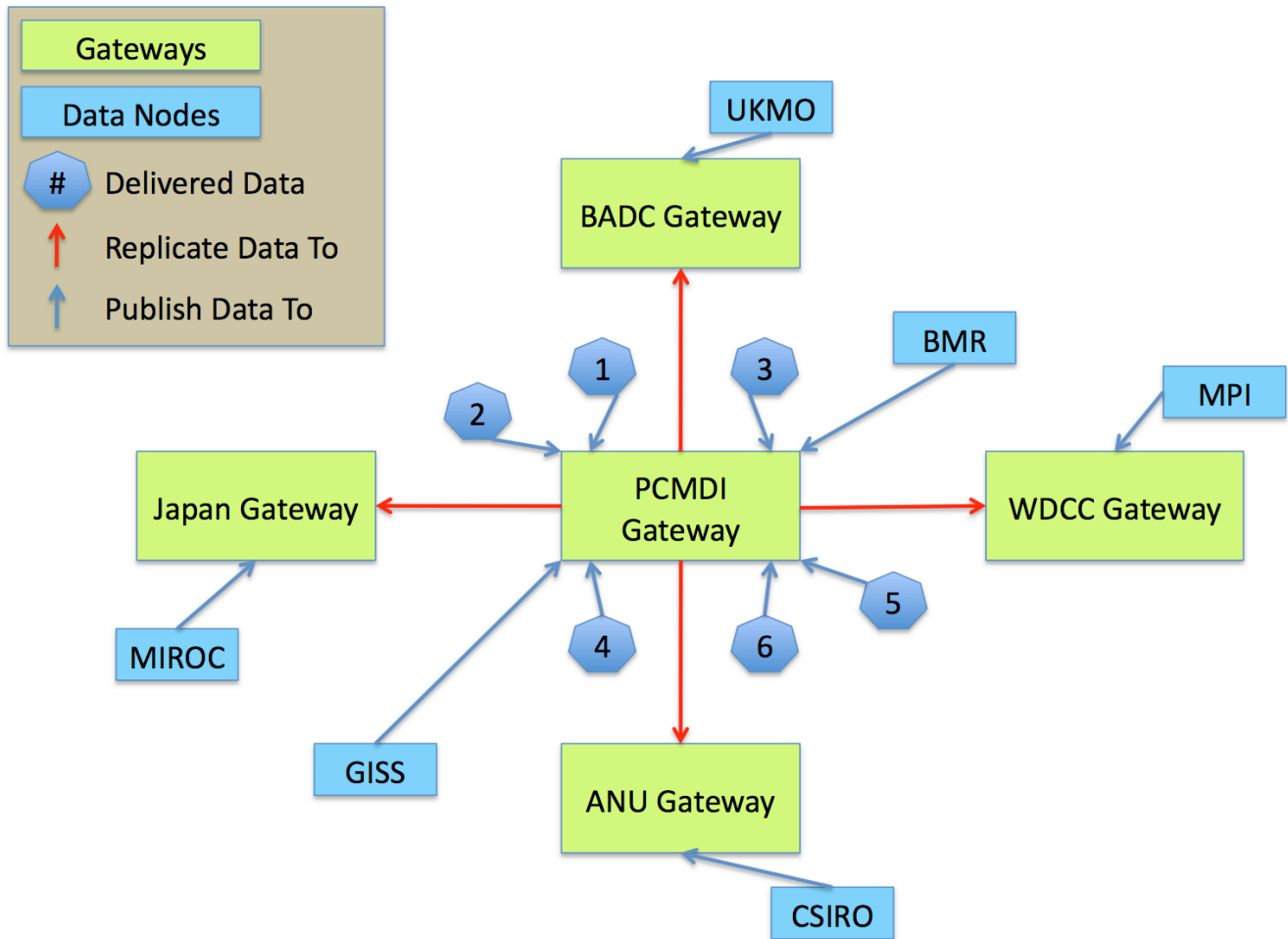


Figure 1: The ESG Federation includes both Gateways and Data Nodes, with complete copies of the CMIP5 core data archived in “Core Data Nodes” at some of the Gateway sites. If modeling centers choose not to standup a Gateway or a Data Node, then their data will be delivered to PCMDI or BADC or WDC.

This document briefly summarizes how modeling centers can make their standard output available either by directly placing it on their own ESG Data Node (the preferred option; section 2) or by sending it to PCMDI or a designated core data node (section 3). It also describes the procedure by which data will be sent to a designated core data node (section 4), how flawed data will be corrected (section 5), and how the core data will be replicated at the “Core Data Nodes” (section 6).

Special CMIP5 Data Considerations

The complete requirements for preparing CMIP5 output are not included in this document, but notably:

- Model output will not be made available until the CMIP5 metadata questionnaire has been completed. The questionnaire will be used to obtain model documentation and also documentation of each simulation performed by the model. The questionnaire will be linked from the CMIP5 web page at <http://cmip-pcmdi.llnl.gov/cmip5/>. (The questionnaire is scheduled to be available by December 2009).
- The file and directory structure and names must be constructed consistent with the requirements of the “Data Reference Syntax” (DRS). (See http://cmip-pcmdi.llnl.gov/cmip5/docs/cmip5_data_reference_syntax_v0-20_clean.pdf)
- The data must be written in conformance with CF-metadata standards and the requirements for metadata imposed by CMOR2; it is highly recommended that CMOR2 be used to rewrite model output. (See http://cmip-pcmdi.llnl.gov/cmip5/output_req.html?submenuheader=2#cmor)

2. Option for providing data through an ESG Data Node (preferred option)

CMIP5 modeling groups are encouraged to install an ESG Data Node and archive their output on it, thereby making it accessible via the ESG Gateway to end users. The output archived on the node should include all the requested CMIP5 standard output. It may also include from the same simulations additional output not called for by CMIP5 but nevertheless of interest.

Deploying an ESG Data Node is not difficult and has been completed in one case (involving an amateur user no less) in less than an hour. The difficulty in installing a Data Node will depend, however, on whether the prerequisite operating system software and support tools have already been installed. There are also minimum requirements for system hardware capabilities which must be met. A full description of the procedure and requirements for deploying an ESG data node can be found in a companion document, “Taking Part in ESG: Data Node Installation.” This accompanying document can be on the CMIP5 web page at <http://cmip-pcmdi.llnl.gov/cmip5/>.

As model output is produced for each CMIP5 simulation, the modeling group will store it, and after checking it for errors, they will “publish” it to their ESG Data Node, thereby making it visible to the CMIP5 Gateways and enabling its delivery to end-users. The “publishing” procedure is straightforward and can be scripted, as described in “Publishing Data to the ESG Federation”. (See <http://cmip-pcmdi.llnl.gov/cmip5/>)

Once the “core data” from each simulation has been “published” on the group’s own Data Node, it should be transferred to PCMDI (or to the alternative Core Data Node designated for certain groups), which will provide a back-up and build redundancy into the data delivery system. The usual procedure for doing this is described in section 4 and will rely on multi-terabyte portable hard disks sent to PCMDI via a parcel shipping service (e.g., FedEx). In unusual circumstances, alternative means of transferring data (e.g., via the internet) might, by special arrangement with PCMDI (or with the alternative data node), be possible, unless the alternative proves too slow. In any case, the transfer of the “core data” to PCMDI (or to an alternative core data node) will not delay access to a modeling group’s own Data Node, once the data has been published.

It should be noted that groups electing to deploy their own ESG Data Node will be able to make their data available for analysis much sooner than those who cannot host a Data Node (probably at least a few weeks sooner). Note also that a group maintaining its own Data Node will be able to immediately correct or withdraw data if mistakes are found.

3. Option for groups unable to host an ESG Data Node

Groups unable to host a Data Node may transfer their standard output via physical disks to PCMDI (or, in designated cases, to the BADC or WDCC², or perhaps another “Core Data Node”). It is estimated that the transfer of multiple terabytes of output to a Core Data Node would take much longer via the internet (by means of normal network hardware and conduits) than via physical disks, which is why the data will usually be transferred using portable hard disks. There might be special cases, however, when transfer through a different means would make sense and would therefore be considered. The procedure by which the model output should normally be transferred to the designated “Core Data Center” is described in the next section.

4. Sending model output to PCMDI (or the designated “Core Data Center”)

The first step in sending data to a Core Data Node is to email a request for one or more multi-TB disks to PCMDI (or a designated “Core Data Center”). A disk or disks will then be sent to the modeling center. Once the modeling center has transferred its CMIP5 model output to the disks, the disks will be shipped to the Core Data Node assigned to that center. The data will then be uploaded, archived and published as part of the CMIP5 archive. This procedure is similar to what was done in CMIP3. We are in the process of investigating what size and type of disks to purchase for this purpose, with a very important consideration being the reliability of the disk hardware and firmware.

If all the model output were to arrive at a Core Data Node at the same time, resources would be severely stretched and there would be considerable delays in making the data available for use. We therefore ask modeling groups to send their data in several chunks, as different sets of CMIP5 experiments are completed. It will be inefficient to make the “chunks” too small, as this will require many disks to be sent back and forth, but waiting until all simulations have been completed will also cause a delay, since the last simulation might be completed many months or more after the first is ready for distribution. It is estimated that the time between a request for a disk and the “publishing” of the data will be a minimum of three weeks, so this can be used as a guide in deciding how to optimally send output to a Core Data Center.

When requesting a disk from PCMDI, send an email to Tony Hoang (hoang1@llnl.gov) and include the following information:

1. Contact (name and email)
2. Address (to which the disk(s) should be sent; not a P.O box number)

² For all UK modeling groups, for IPSL, and for EC-Earth, data will be sent to BADC. For all German modeling groups, WDCC will take the data. BADC and WDCC will ensure the data is then ingested into the ESG federated delivery system.

3. Institution I.D.
4. Model I.D.
5. How many terabytes of data are ready to be transferred?

5. Correcting or replacing archived model output

It is certain that errors will be found in model output initially “published” to the ESGF. When errors are found, the first step should be to withdraw the flawed data from the CMIP5 archive. Groups with their own ESG Data Node can immediately do this, as described elsewhere. Groups who have sent their data to a “Core Data Center” must send an email request for the data to be withdrawn to PCMDI (hoang1@llnl.gov). Once corrected data becomes available, it can either be served on the modeling group’s own ESG Data Node (if one exists), or sent to the appropriate Core Data Center. The method of transferring data (either via physical disk or via the internet) will depend on the volume of data.

6. Replication of data at the “Core Data Centers”

A subset of the data hosted at each Data Node (i.e., the “core data”), will be replicated across all the Core Data Centers. How this will be accomplished may be of interest only to those hosting a Core Data Center.

Once a Core Data Center receives model output (directly) from a modeling group, it should inspect the data for flaws and then pass it along to the other Core Data Centers, the first priority being to replicate the data at PCMDI. Depending on how much data is involved, the data will either be sent via hard disk(s) or via the internet.