

ENES DTF Telco 17-01-11, 14 – 16:10

Attendees:

Sebastien Denvil (IPSL), Francesca Guglielmo (IPSL), Sylvie Joussaume (IPSL), Martin Juckes (CEDA), Stephan Kindermann (DKRZ), Michael Kolax (SMHI), Michael Lautenschlager (DKRZ, Minutes), Christian Page (CERFACS), Frank Toussaint (DKRZ),
Excused: Sandro Fiore (CMCC)

Topics that have been discussed:

CMIP6 European storage contributions

The estimated CMIP6 data volume is still not completely fixed and the discussion of a core data set of 2 PB size is ongoing with no clear direction. Therefore, for CMIP6 resource and replication planning within Europe we might face the full data set of 20 PB and more.

From the telco we collect the following CMIP6 disc storage contributions which is the status of January 17th, 2017. In the list we distinguish between primary data publication and replication together with national scientific requirements:

- CEDA: no numbers for primary data publication and 1 -2 PB are expected for replication
- CMCC: 1 PB for primary data publication, no replication planned
- DKRZ: 2 PB for primary data publication, 3 PB for replication and about 20 PB tape in addition
- IPSL: 2 PB for primary data publication, 3 PB for replication
- LIU: no numbers because the funding is yet not decided

CMIP6 European replication strategy

The discussion specified two objectives of data replication: data archive security and data accessibility enhancement. Archive security requires at least two full copies, preferably more of CMIP6 data throughout ESGF. For data accessibility enhancement, the global CMIP6 replication strategy and national scientific requirements have to be taken into account at the individual ESGF data nodes that participate in the CMIP6 data replication (CMIP6 tier-1 nodes). The balance between replication and scientific requirements defines data “hot spots” (priority 1) and “cold spots” (priority 2) for each data node in the ENES CDI. The synopsis of the CMIP6 data “hot spots” provides a clearer view of data that covered in the European replication and those CMIP6 data which are missing in the replication. From this picture we might be able to develop a replication strategy for storing a full copy of CMIP6 data at European CMIP6 tier-1 data nodes.

Action item: European ESGF data nodes that plan to participate in the CMIP-6 data replication identify national scientific requirements and transfer them into local data “hot spots”.

Institutional support for CMIP6 ESGF operation

The institutional support for CMIP6/ESGF as it stands for January 2017 is:

- DKRZ: 7 persons with 4 persons on soft money
- IPSL: 5 persons with 3 persons on soft money
- CEDA: 3 persons
- LIU: 1 person with 50% on soft money
- SMHI: not yet specified
- CMCC: not yet specified

The content related contributions in ESGf working groups and CMIP6 services are specified in the ENES DTF status report at the IS-ENES-2 3rd General Assembly in Paris January 16th – 18th, 2017.

ESGF user support after IS-ENES-2

The ESGF first level user support so far has been mainly covered by the EU-project IS-ENES-2. The project ends in March 2017 together with direct funding of the support person which is currently employed at DKRZ. DKRZ has no facility to finance the ESGF user support by its institutional budget. We agreed to look for alternative funding opportunities even nothing concrete could be indentified during the telco.

The ESGF support presently covers the first level support and in addition maintenance of web pages, FAQs and documentation. Even if the ESGF support tasks are shared between several partners we expect main drawbacks after end of IS-ENES-2. In addition the personal skills which have been accumulated in the run time of the project are lost for ESGF if the person moves to another position with new responsibilities.

ENES CDI after IS-ENES-2

It is agreed to continue the ENES CDI with institutional support and funding from national and European projects. The coordination will be performed by the ENES DTF. The European projects proposals are currently in COPERNICUS, E-INFRA12 and E-INFRA21.

In COPERNICUS we have two projects for providing climate model data, C3S_34a_CP4CDS coordinated by NCAS and C3S_34b_Regional-Climates-PIN coordinated by IPSL. ENES participation is agreed in the EGI-EUDAT-INDIGO proposal for the EOSC in the E-INFRA-12a call. In the parallel call E-INFRA-21 two ENES coordinated proposals are under development for C4I and the OPHIDIA data analysis platform. A continuation of the CLIPC developments might be possible in INFRADEV-01-2017.

A common strategy for the ENES CDI development is expected in the Update of the ENES Foresight document. This will be discussed at the IS-ENES-2 3rd General Assembly in Paris.

Other Business

The ENES HPC task force (HPCTF) suggests a closer interaction between HPC and CDI. Milestones should be coordinated between the ENES HPCTF and the ENES DTF