EC-Earth and ESGF

Kim Serradell
Pierre-Antoine Bretonniere
(Computational Earth Sciences, BSC)

Prashanth Dwarakanath
( Linkoping University/NSC)

Co-funded by the European Union

03/12/2017

7th Annual ESGF F2F 2017
Outline

• Barcelona Supercomputing Center
• EC-Earth climate model
• BSC Experience
• LIU
• EC-Earth and ESGF
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

BSC-CNS objectives

- Supercomputing services to Spanish and EU researchers
- R&D in Computer, Life, Earth and Engineering Sciences
- PhD program, technology transfer, public engagement

BSC-CNS is a consortium that includes:

- Spanish Government 60%
- Catalonian Government 30%
- Univ. Politècnica de Catalunya (UPC) 10%
Earth Sciences Department

What
Environmental forecasting

Why
Our strength …
... research …
... operations …
... services …
... high resolution …

How
Develop a capability to model air quality processes from urban to global and the impacts on weather, health and ecosystems
Implement climate prediction system for subseasonal-to-decadal climate prediction
Develop user-oriented services that favour both technology transfer and adaptation
Use cutting-edge HPC and Big Data technologies for the efficiency and user-friendliness of Earth system models

Earth system services
Climate prediction
Atmospheric composition
Computational Earth sciences

BSC
Barcelona Supercomputing Center
Centro Nacional de Supercomputación
EC-Earth
EC-Earth Model

- The Integrated Forecasting System (IFS) as atmosphere model
- The Nucleus for European Modelling of the Ocean (NEMO) as ocean model
- The OASIS3-MCT coupler
- The Louvain-la-Neuve sea-Ice Model 3 (LIM3) as sea ice model
Community driven development

Schematic outline of the EC-Earth consortium and its members as of May 2014.
PRIMAVERA H2020

• a European Union Horizon2020 project,
• a new generation of advanced and well-evaluated high-resolution global climate models,
• simulations and predictions of regional climate with unprecedented fidelity,
• for the benefit of governments, business and society in general.

Ocean surface currents from HadGEM3-based global coupled (atmosphere-ocean/sea-ice) models at three different resolutions - (left) 25km-1/12 degree, (middle) 60km-1/4 degree, (right) 130km-1 degree (courtesy of Malcolm Roberts)
BSC Experience
BSC Data Node

- ESGF Tier 2 “data” node: publishing BSC and AEMET PRIMAVERA and CMIP6 data
- Index node at STFC (partner involved in PRIMAVERA)
- PRIMAVERA Stream 1 data still uploaded to Jasmin but published from BSC node
- PRIMAVERA Stream 2 data will be only stored and published from BSC
- Estimate of ~600TB of PRIMAVERA data
ESGF back-end index node for publication by Tier 2 sites at CEDA

This is a back-end index node at CEDA. This node is provided to allow publication by Tier 2 sites, and to allow search indexes (at CEDA and externally) to search those records via SoLr.

The current list of Tier 2 sites publishing to this index is as follows:

- Barcelona Supercomputing Centre

END USERS

It is not recommended that you use this node directly. Any searches performed on this CoG site will only see the data that is hosted by the 2 sites that publish to this index, and will not see either any records published by CEDA or any published to other Tier 1 indexes outside CEDA.

The main CEDA ESGF search is at https://esgf-index1.ceda.ac.uk/

TIER 1 SITES

You may do automated searching or replication from this Solr shard on this node via /solr URLs on port 80 or 443. As for end users, use of the CoG UI is not supported.

SELECTED TIER 2 SITES

If you appear on the above pre-arranged list, you may use publication service URLs on this server. (This is not a general service for all Tier 2 sites.)
BSC cmorization

- Part of PRIMAVERA project
  - ece2cmor3 is trying to standardize the cmorization within EC-Earth
  - Developed by a set of partners (KNMI/e-Science Center and BSC)
  - But each partner cmorizes its own data (lots of coordination needed)
    - online (on HPC as the simulation runs to only store cmorized data)
    - offline (adding time to publish)
BSC Data Node (installer)

• March 2017:
  • started installing version 2.3 to publish CMIP5/SPECS data as test

=> impossible due to dependencies pointing to updated versions of packages breaking compatibility of the overall installation

• August 2017:
  • Successfully ran version 2.5 of the installer, even if still under development
BSC Data Node (publisher)

Once we had the node up and running, updating:

- updating the esg.[primavera-cmip6].ini
- esgcet_models_table.txt
- certificates and PIDs landing pages

} PRIMAVERA specific

} EC-Earth specific

} BSC specific
BSC recommendations

• Fresh start installation testing

• Semantic versioning of
  • dependencies (git submodules?)
  • releases

• Allow users to install dependencies by hand within the installer run

• Identify who (person or institution) is responsible for what (PID, installation, publication,...)

• Wiki/documentation for “beginners”
  • How to’s?
  • How to deploy a production and a testing server
LiU on ESGF and CDNOT

• LiU operates a Tier-1 ESGF node, offering all ESGF services, including index peering to other datanodes, funded by SMHI and IS-ENES projects.
• Hosts CORDEX, CMIP5, SPECS, and CLIPC projects, from SMHI.
• Manages attribute services for CORDEX data access.
• Mailing list maintainer for esgf-cordex@lists.nsc.liu.se, for users registered to access CORDEX data on ESGF.
• LiU Manages one of the ESGF Federation CAs and helped develop CA policy for ESGF.
• Prashanth Dwarakanath (LiU) co-leads the ESGF Installation Working Team, and is a representative for EC-EARTH in CDNOT.
• Kai Lu (LiU) is an active contributor to user support on ESGF Users and IWT mailing lists.
Ongoing and future activities on CDNOT

- LiU has contributed to achieving better documentation by authoring manuals for data node administrators, and looks to take an active role in documenting procedures for CMIP6 data publication and node operations.

- Has conducted training workshops and code sprints and is working on developing a suite of tools to reduce uncertainties with infrastructure setups for workshops and collaborative development events.

- Actively working on setting up of an Incident-response team, and to disseminate training and documentation to ensure adoption of security best-practices.
EC-Earth and ESGF
Previous experiences

• CMIP5 and SPECS
  • Centralized at BADC
    • Only a reduced set of partners was involved in the public
  • Data Node in LIU
  • Partners without data node had to:
    • Upload lots of data
    • Needing assistance
Future experiences

• CMIP6 (PRIMAVERA)
  • More data
    • spatial resolution increasing
    • spatial dissemination increasing (more partners)

• More partners
  • more ESGF/EC-Earth data nodes
  • more “ESGF beginners”

• How to deal with such a project?
  • at the EC-Earth level
  • at the ESGF level
## EC-Earth’s contributions to CMIP6

<table>
<thead>
<tr>
<th>Institute</th>
<th>Contact</th>
<th>Contributions</th>
<th>HPC platform</th>
<th>ESGF node</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. AEMET</td>
<td>J. A. Parodi Perdomo</td>
<td>DECK (EC-Earth3-CC)</td>
<td>Cray XC30@ECMWF + BULL@AEMET</td>
<td>? (BSC?)</td>
</tr>
<tr>
<td>2. BSC/IC3</td>
<td>F. Doblas-Reyes</td>
<td>DECK, DCPP, HighResMIP (?)</td>
<td>Marenostrum@BSC</td>
<td>BSC</td>
</tr>
<tr>
<td>3. CNR</td>
<td>J. von Hardenberg</td>
<td>DECK (EC-Earth-CC, EC-Earth-CC-LR)</td>
<td>CNR</td>
<td></td>
</tr>
<tr>
<td>4. DMI</td>
<td>S. Yang</td>
<td>DECK, DCPP, GeoMIP, HighResMIP, ISMIP6, LS3MIP, SenarioMIP</td>
<td>Cray XC30@DMI/IMO</td>
<td>DMI</td>
</tr>
<tr>
<td>5. ENER</td>
<td>A. Alessandri</td>
<td>LS3MIP</td>
<td>Cray XC30@ECMWF</td>
<td></td>
</tr>
<tr>
<td>6. FMI</td>
<td>H. Korhonen</td>
<td>AeroChemMIP</td>
<td>Cray XC30@FMI</td>
<td>FMI/UHEL</td>
</tr>
<tr>
<td>7. IPMA</td>
<td>P. Viterbo</td>
<td>??</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. KIT</td>
<td>A. Arneth</td>
<td>C4MIP, LUMIP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. KNMI</td>
<td>R. Bintanja</td>
<td>DECK, AeroChemMIP, CFMIP(?), HighResMIP, SenarioMIP</td>
<td>Cray XC30@ECMWF + BULL@KNMI</td>
<td>KNMI</td>
</tr>
<tr>
<td>10. MetEirre/ICHEC</td>
<td>R. McGrath</td>
<td>DECK, HighResMIP, SenarioMIP</td>
<td>Cray XC30@ECMWF + Cray XC40@PRACE</td>
<td>ICHEC</td>
</tr>
<tr>
<td>11. LU</td>
<td>P. Miller</td>
<td>C4MIP, LUMIP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. MISU</td>
<td>L. Brodeau</td>
<td>DECK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. NGSU</td>
<td>Q. Zhang</td>
<td>PMIP</td>
<td></td>
<td>NSC</td>
</tr>
<tr>
<td>14. SMHI</td>
<td>K. Wyser</td>
<td>DECK, C4MIP, HighResMIP, C4MIP, SenarioMIP</td>
<td></td>
<td>LIU</td>
</tr>
<tr>
<td>15. UHEL</td>
<td>R. Makkonen</td>
<td>AeroChemMIP</td>
<td>Cray XC30@CSC</td>
<td>FMI/UHEL</td>
</tr>
</tbody>
</table>
Thank you

kim.serradell@bsc.es