Introduction to CEDA-JASMIN

Malcolm Roberts – 29-06-2023 – EERIE science hour

(with thanks to Jon Seddon and the JASMIN team for their excellent documentation)

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

Funded by the European Union

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).

EERIE funding

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No. 101081383

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

CEDA-JASMIN platform

- One of EERIE's central places to do collaborative analysis on multi-model datasets
- One of the platforms we use for the EERIE hackathon in November during the GA
- UKRI fund CEDA-JASMIN, and they are one of our EERIE partners as well

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

Funded by the European Union

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).

Basic aspects – accounts etc

- Much of the basics are in the EERIE onboarding document via Jon's instructions
  - https://github.com/jonseddon/eerie_instructions
  - including how to get an account, how to get set up to access CMIP data etc
  - Home directory: /home/users/<username> (100 GB space)
- Group Work spaces (GWS)
  - These are the spaces to put large datasets that you want to keep
  - Note that EERIE only has ~250 TB GWS disk space in total for the whole project, as well as 50 TB object store
  - We also have access to the tape archive – not going to talk about this today
  - GWS path: /gws/nopw/j04/eerie/
  - To request permission to read/write to the EERIE GWS, visit: https://accounts.jasmin.ac.uk/services/group_workspaces/?query=eerie
- CMIP data available on disk
  - /badc/cmip6/data/ (need to link accounts, see eerie_instructions above)

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

Funded by the European Union

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).

JASMIN services

- Login via servers, e.g.
  - ssh -Y -A <username>@login2.jasmin.ac.uk
- Interactive analysis servers – sci[1-8].jasmin.ac.uk
  - for interactive work and testing scripts, not for long, intensive processing
  - different servers have smaller/larger memory and variable loading
- LOTUS batch cluster - 19,000 core batch computing system
  - used for long and intensive jobs, workflows
  - https://help.jasmin.ac.uk/category/4889-slurm
  - different queus for short/long jobs, all using slurm
- Experimental GPU cluster
  - need to register https://help.jasmin.ac.uk/article/4951-gpu-testing
- Directories for writing
  - As well as the GWS, there are various scratch directories to write intermediate files from workflows
  - https://help.jasmin.ac.uk/article/4700-understanding-new-jasmin-storage
  - depends on need for speed, file size, parallel write etc

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

Funded by the European Union

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).

Using JASMIN – software etc

- Software on JASMIN
  - https://help.jasmin.ac.uk/category/270-software-on-jasmin
- Environments
  - standard modules (module load) are
  - jaspy (python), jasmin-sci (extra software), jasr (R programming)
  - Can also choose particular variants of these e.g. python3.x (at least for some time)
- ESMValTool also available:
  - https://help.jasmin.ac.uk/article/4955-community-software-esmvaltool
- Own environments and software
  - Can also install your own environment and software, including via conda and virtual python environments
  - https://help.jasmin.ac.uk/article/5074-conda-environments-and-python-virtual-environments
  - https://help.jasmin.ac.uk/article/5075-creating-and-using-miniconda-environments

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

Funded by the European Union

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).

Hackathon resources

- JASMIN has hosted previous CMIP6 hackathons
  - https://cmip6moap.github.io/resources/
  - Variety of example Jupyter notebooks in github repos from these events
- Perhaps we can follow some similar ideas, e.g. at end of hackathon
  - commit any code, notebooks etc to our GitHub repo
  - Identify one or more figures that could be showcased
  - Store any data for future use in a GWS
- Notebooks
  - jupyter notebooks and service is available
  - https://help.jasmin.ac.uk/article/4851-jasmin-notebook-service

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

Funded by the European Union

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).

Potential future options

- Data management tool (DMT)
  - One implemented in PRIMAVERA
  - https://prima-dm1.jasmin.ac.uk/received_data_quick_query/
  - This is able to search PRIMAVERA-HighResMIP data in CEDA ESGF
- Plan to do something similar for EERIE
  - to make finding (and making available) data (across platforms this time) easier
  - This awaits Jon Seddon's return
- However, data can be searched in ESGF too (Pablo Ortega can demonstrate)

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract #22.00366.

Funded by the European Union

This work was funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant number 10040510).