Geophysical Research Abstracts
Vol. 16, EGU2014-4565, 2014
EGU General Assembly 2014
© Author(s) 2014. CC Attribution 3.0 License.

Flexible server-side processing of climate archives

Martin Juckes, Ag Stephens, and Eduardo Damasio da Costa
STFC, BADC, Chilton, Didcot, Oxon, United Kingdom (martin.juckes@stfc.ac.uk)

The flexibility and interoperability of OGC Web Processing Services are combined with an extensive range of data processing operations supported by the Climate Data Operators (CDO) library to facilitate processing of the CMIP5 climate data archive. The challenges posed by this peta-scale archive allow us to test and develop systems which will help us to deal with approaching exa-scale challenges. The CEDA WPS package allows users to manipulate data in the archive and export the results without first downloading the data – in some cases this can drastically reduce the data volumes which need to be transferred and greatly reduce the time needed for the scientists to get their results. Reductions in data transfer are achieved at the expense of an additional computational load imposed on the archive (or near-archive) infrastructure. This is managed with a load balancing system. Short jobs may be run in near real-time, longer jobs will be queued. When jobs are queued the user is provided with a web dashboard displaying job status. A clean split between the data manipulation software and the request management software is achieved by exploiting the extensive CDO library. This library has a long history of development to support the needs of the climate science community. Use of the library ensures that operations run on data by the system can be reproduced by users using the same operators installed on their own computers. Examples using the system deployed for the CMIP5 archive will be shown and issues which need to be addressed as archive volumes expand into the exa-scale will be discussed.