Geophysical Research Abstracts
Vol. 14, EGU2012-12913, 2012
EGU General Assembly 2012

Scientific workflow and support for high resolution global climate modeling at the Oak Ridge Leadership Computing Facility

V. Anantharaj (1), B. Mayer (2), F. Wang (1), J. Hack (1), D. McKenna (2), and R. Hartman-Baker (1)

(1) Oak Ridge Leadership Computing Facility, Oak Ridge National Laboratory, Oak Ridge, United States (anantharajvg@ornl.gov), (2) Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, United States

The Oak Ridge Leadership Computing Facility (OLCF) facilitates the execution of computational experiments that require tens of millions of CPU hours (typically using thousands of processors simultaneously) while generating hundreds of terabytes of data. A set of ultra high resolution climate experiments in progress, using the Community Earth System Model (CESM), will produce over 35,000 files, ranging in sizes from 21 MB to 110 GB each. The execution of the experiments will require nearly 70 Million CPU hours on the Jaguar and Titan supercomputers at OLCF. The total volume of the output from these climate modeling experiments will be in excess of 300 TB.

This model output must then be archived, analyzed, distributed to the project partners in a timely manner, and also made available more broadly. Meeting this challenge would require efficient movement of the data, staging the simulation output to a large and fast file system that provides high volume access to other computational systems used to analyze the data and synthesize results. This file system also needs to be accessible via high speed networks to an archival system that can provide long term reliable storage. Ideally this archival system is itself directly available to other systems that can be used to host services making the data and analysis available to the participants in the distributed research project and to the broader climate community.

The various resources available at the OLCF now support this workflow. The available systems include the new Jaguar Cray XK6 2.63 petaflops (estimated) supercomputer, the 10 PB Spider center-wide parallel file system, the Lens/EVEREST analysis and visualization system, the HPSS archival storage system, the Earth System Grid (ESG), and the ORNL Climate Data Server (CDS). The ESG features federated services, search & discovery, extensive data handling capabilities, deep storage access, and Live Access Server (LAS) integration.

The scientific workflow enabled on these systems, and developed as part of the Ultra-High Resolution Climate Modeling Project, allows users of OLCF resources to efficiently share simulated data, often multi-terabyte in volume, as well as the results from the modeling experiments and various synthesized products derived from these simulations. The final objective in the exercise is to ensure that the simulation results and the enhanced understanding will serve the needs of a diverse group of stakeholders across the world, including our research partners in U.S. Department of Energy laboratories & universities, domain scientists, students (K-12 as well as higher education), resource managers, decision makers, and the general public.