The Earth System Grid Federation (ESGF) Project

Nicolas Carenton-Madiec, Sébastien Denvil, and Mark Greenslade
CNRS/IPSL, Climate Data Management, Paris, France

The Earth System Grid Federation (ESGF) Peer-to-Peer (P2P) enterprise system is a collaboration that develops, deploys and maintains software infrastructure for the management, dissemination, and analysis of model output and observational data. ESGF's primary goal is to facilitate advancements in Earth System Science. It is an interagency and international effort led by the US Department of Energy (DOE), and co-funded by National Aeronautics and Space Administration (NASA), National Oceanic and Atmospheric Administration (NOAA), National Science Foundation (NSF), Infrastructure for the European Network of Earth System Modelling (IS-ENES) and international laboratories such as the Max Planck Institute for Meteorology (MPI-M) german Climate Computing Centre (DKRZ), the Australian National University (ANU) National Computational Infrastructure (NCI), Institut Pierre-Simon Laplace (IPSL), and the British Atmospheric Data Center (BADC). Its main mission is to support current CMIP5 activities and prepare for future assesments.

The ESGF architecture is based on a system of autonomous and distributed nodes, which interoperate through common acceptance of federation protocols and trust agreements. Data is stored at multiple nodes around the world, and served through local data and metadata services. Nodes exchange information about their data holdings and services, trust each other for registering users and establishing access control decisions. The net result is that a user can use a web browser, connect to any node, and seamlessly find and access data throughout the federation.

This type of collaborative working organization and distributed architecture context en-lighted the need of integration and testing processes definition to ensure the quality of software releases and interoperability. This presentation will introduce the ESGF project and demonstrate the range of tools and processes that have been set up to support release management activities.