Prodiguer project : data burden, workflow and climate modeling group

Sebastien Denvil[†]; Mark Morgan; Ashish Bhardwaj; Jerome Raciazek
[†] CNRS/IPSL, France
Leading author: sebastien.denvil@ipsl.jussieu.fr

The Institute Pierre Simon Laplace (IPSL), like many other climate modelling groups, is involved in the development of a comprehensive Earth System Model (ESM) to study the interactions between chemical, physical, and biological processes. This work entails the coupling of different components (land, ocean, atmosphere, chemistry...etc) and requires an execution environment platform that can tackle the entire range of interdependent model configurations. Furthermore, the ever-increasing number of simulations, executed against model configurations within scientific computing centres, is generating a huge volume of data and meta-data that must be made readily available to researchers, modellers, students and general users. Each user group has a different set of information demands related to climate simulation data and metadata, and thus fulfilling the requirements of the entire community is highly challenging. This poster will focus upon the strategy adopted by IPSL to fulfilling the needs of the community whilst lowering the data distribution and data management burdens upon the climate modelling group due to the exploding interest in climate simulations data and information. To achieve these objectives we decided to leverage the efforts of international and European projects such as Earth System Grid Federation, METAFOR, IS-ENES within our execution environment platform. We will present the emerging workflow that will be in place to run not only CMIP5 simulations but also the "day-to-day" simulations that are run outside of the context of a large model inter-comparison project such as CMIP5. IPSL is implementing a French national project called Prodiguer whose objective is to ensure that data and meta-data can be delivered to the French & international communities in a timely and appropriate fashion, hence achieving the strategic goals outlined above. Prodiguer leverages, extends and builds upon the work of international projects, in particular the ESG-Federation software stack (ESG data node, security services, THREDDS ...etc.). The centrepiece of Prodiguer is an intuitive web portal (plus associated web services) that support use cases pertaining to simulation management and output data access. The user community has for the first time an integrated view across simulations executing across all IPSL compute nodes plus a search engine providing rapid access to output data held across all IPSL data nodes. Technically we will present both the physical infrastructure and software architecture being put in place to achieve the goals of Prodiguer. We will also present progress made in respect of intuitive user interface design, an often neglected aspect of climate modelling related software.