Geophysical Research Abstracts
Vol. 16, EGU2014-12369, 2014
EGU General Assembly 2014

Convergence in France facing Big Data era and Exascale challenges for Climate Sciences

Sébastien Denvil (1), Jean-Louis Dufresne (1), David Salas (2), Yann Meurdesoif (1), Sophie Valcke (3), Arnaud Caubel (1), Marie-Alice Foujols (1), Jérôme Servonnat (1), Stéphane Sénési (2), Julien Derouillat (4), and Pascal Voury (5)

(1) Institut Pierre Simon Laplace (IPSL), Climate modeling group, Paris, France, (2) CNRS-GAME (Météo-France / CNRS), Toulouse France, (3) Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), Toulouse, France, (4) Maison De La Simulation, Orsay, France, (5) Institute for Development and Resources in Intensive Scientific Computing (IDRIS), Orsay, France

The presentation will introduce a french national project : CONVERGENCE that has been funded for four years. This project will tackle big data and computational challenges faced by climate modeling community in HPC context.

Model simulations are central to the study of complex mechanisms and feedbacks in the climate system and to provide estimates of future and past climate changes. Recent trends in climate modelling are to add more physical components in the modelled system, increasing the resolution of each individual component and the more systematic use of large suites of simulations to address many scientific questions. Climate simulations may therefore differ in their initial state, parameter values, representation of physical processes, spatial resolution, model complexity, and degree of realism or degree of idealisation. In addition, there is a strong need for evaluating, improving and monitoring the performance of climate models using a large ensemble of diagnostics and better integration of model outputs and observational data.

High performance computing is currently reaching the exascale and has the potential to produce this exponential increase of size and numbers of simulations. However, post-processing, analysis, and exploration of the generated
data have stalled and there is a strong need for new tools to cope with the growing size and complexity of the underlying simulations and datasets. Exascale simulations require new scalable software tools to generate, manage and mine those simulations ,and data to extract the relevant information and to take the correct decision.

The primary purpose of this project is to develop a platform capable of running large ensembles of simulations with a suite of models, to handle the complex and voluminous datasets generated, to facilitate the evaluation and validation of the models and the use of higher resolution models. We propose to gather interdisciplinary skills to design, using a component-based approach, a specific programming environment for scalable scientific simulations and analytics, integrating new and efficient ways of deploying and analysing the
applications on High Performance Computing (HPC) system. CONVERGENCE, gathering HPC and informatics expertise that cuts across the individual partners and the broader HPC community, will allow the national climate community to leverage information technology (IT) innovations to address its specific needs.

Our methodology consists in developing an ensemble of generic elements needed to run the French climate models with different grids and different resolution, ensuring efficient and reliable execution of these models, managing large volume and number of data and allowing analysis of the results and precise evaluation of the models. These elements include data structure definition and input-output (IO), code coupling and interpolation, as well as runtime and pre/post-processing environments. A common data and metadata structure will allow transferring consistent information between the various elements.

All these generic elements will be open source and publicly available. The IPSL-CM and CNRM-CM climate models will make use of these elements that will constitute a national platform for climate modelling. This platform will be used, in its entirety, to optimise and tune the next version of the IPSL-CM model and to develop a global coupled climate model with a regional grid refinement. It will also be used, at least partially, to run ensem-

bles of the CNRM-CM model at relatively high resolution and to run a very-high resolution prototype of this model.

The climate models we developed are already involved in many international projects. For instance we participate to the CMIP (Coupled Model Intercomparison Project) project that is very demanding but has a high visibility: its results are widely used and are in particular synthesised in the IPCC (Intergovernmental Panel on Climate Change) assessment reports. The CONVERGENCE project will constitute an invaluable step for the French climate community to prepare and better contribute to the next phase of the CMIP project.