Geophysical Research Abstracts
Vol. 19, EGU2017-16416-2, 2017
EGU General Assembly 2017

Use of several Cloud Computing approaches for climate modelling: performance, costs and opportunities

Diego A. Perez Montes (1), Juan A. Añel Cabanelas (1,2), David C. H. Wallom (3), Alberto Arribas (4), Peter Uhe (3,5), Pablo V. Caderno (6), and Tomas F. Pena (7)

(1) EPhysLab, Universidade de Vigo, Campus As Lagoas, 32004, Ourense, Spain (kabute@uvigo.es), (2) SSEE, University of Oxford, OX1 3QY, Oxford, UK (j.anhel@uvigo.es), (3) OeRC, University of Oxford, OX1 3QG, Oxford, UK (david.wallom@oerc.ox.ac.uk), (4) Met Office Informatics Lab, Exeter, UK (alberto.arribas@informaticslab.co.uk), (5) Environmental Change institute, University of Oxford, OX1 3QY, Oxford, UK (peter.uhe@ouce.ox.ac.uk), (6) mAdme Technologies Ltd, 32 Lower Lesson Street, Dublin 2, Dublin, Ireland (kaderno@gmail.com), (7) CITIUS, University of Santiago de Compostela, Santiago de Compostela, 15782, Spain (tf.pena@usc.es)

Cloud Computing is a technological option that offers great possibilities for modelling in geosciences. We have studied how two different climate models, HadAM3P-HadRM3P and CESM-WACCM, can be adapted in two different ways to run on Cloud Computing Environments from three different vendors: Amazon, Google and Microsoft. Also, we have evaluated qualitatively how the use of Cloud Computing can affect the allocation of resources by funding bodies and issues related to computing security, including scientific reproducibility.

Our first experiments were developed using the well known ClimatePrediction.net (CPDN), that uses BOINC, over the infrastructure from two cloud providers, namely Microsoft Azure and Amazon Web Services (hereafter AWS). For this comparison we ran a set of thirteen month climate simulations for CPDN in Azure and AWS using a range of different virtual machines (VMs) for HadRM3P (50 km resolution over South America CORDEX region) nested in the global atmosphere-only model HadAM3P. These simulations were run on a single processor and took between 3 and 5 days to compute depending on the VM type.

The last part of our simulation experiments was running WACCM over different VMS on the Google Compute Engine (GCE) and make a comparison with the supercomputer (SC) Finisterrae1 from the Centro de Supercomputacion de Galicia. It was shown that GCE gives better performance than the SC for smaller number of cores/MPI tasks but the model throughput shows clearly how the SC performance is better after approximately 100 cores (related with network speed and latency differences).

From a cost point of view, Cloud Computing moves researchers from a traditional approach where experiments were limited by the available hardware resources to monetary resources (how many resources can be afforded). As there is an increasing movement and recommendation for budgeting HPC projects on this technology (budgets can be calculated in a more realistic way) we could see a shift on the trends over the next years to consolidate Cloud as the preferred solution.