Future Generation Computer Systems 51 (2015) 36–44

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Large-scale climate simulations harnessing clusters, grid and cloud infrastructures

V. Fernández-Quiruelas *, C. Blanco, A.S. Cofiño, J. Fernández

Grupo de Meteorología, Dpto. Matemática Aplicada y CC. Computación, Universidad de Cantabria, Santander, Spain

HIGHLIGHTS

- Hybrid DCI (HDCI) is a common computing paradigm in earth science.
- The key problems found by standard applications to be run in HDCI are presented.
- The components of a new execution framework to run climate models are proposed.
- WRF4G is a successful implementation of the framework to run the WRF model in HDCI.

ARTICLE INFO

Article history:
Received 24 January 2014
Received in revised form
24 March 2015
Accepted 23 April 2015
Available online 5 May 2015

Keywords:
Grid computing
Cloud computing
HPC
Regional climate model
WRF
Hybrid distributed computing infrastructures

ABSTRACT

The current availability of a variety of computing infrastructures including HPC, Grid and Cloud resources provides great computer power for many fields of science, but their common profit to accomplish large scientific experiments is still a challenge. In this work, we use the paradigm of climate modeling to present the key problems found by standard applications to be run in hybrid distributed computing infrastructures and propose a framework to allow a climate model to take advantage of these resources in a transparent and user-friendly way. Furthermore, an implementation of this framework, using the Weather Research and Forecasting system, is presented as a working example. In order to illustrate the usefulness of this framework, a realistic climate experiment leveraging Cluster, Grid and Cloud resources simultaneously has been performed. This test experiment saved more than 75% of the execution time, compared to local resources. The framework and tools introduced in this work can be easily ported to other models and are probably useful in other scientific areas employing data- and CPU-intensive applications.

1. Introduction

The improvements achieved on commodity computers during the last two decades have changed the accessibility and availability of computing resources for research. Although supercomputers still play an important role for the research community, clusters and other infrastructures based on commodity computers such as Grid and Cloud infrastructures are widely used due to their low cost. This situation has promoted the spread of new computing facilities and, as a consequence, researchers can simulate in a wide range of computing resources. Today, most researchers have access to several clusters and Grid infrastructures and can rent on-demand Cloud resources to temporarily solve peak workloads [1–3]. The aggregation of these resources as a single Hybrid Distributed Computing Infrastructure (HDCI) can provide a great computing potential. This work introduces a framework for performing large experiments with climate models on HDCIs. The framework has been designed to access transparently these heterogeneous distributed environments providing a uniform interface to run simulations.

The heterogeneous and distributed nature of HDCIs poses new challenges to the applications willing to exploit them. Although there are several pieces of middleware that facilitate the use of HDCIs, there are still unsolved aspects. Moreover, applications with special requirements, such as data-intensive ones or those with long-term runs, require a workflow modification in order to make an efficient use of resources. Section 2 of this work describes the main issues a user finds when trying to port an application to an HDCI.

So far, some efforts have been made to run climate models on HDCIs. Several works have been devoted to adapt a climate model to perform a given experiment in a given Grid infrastructure [4–6]. These works do not solve the issues related to the heterogeneity of resources nor to the data management. Instead,

* Corresponding author.
E-mail address: valvanuz.fernandez@unican.es (V. Fernández-Quiruelas).

http://dx.doi.org/10.1016/j.future.2015.04.009
0167-739X/

they provide ad-hoc solutions that only work in a very limited set of computing resources. The main disadvantage of this approach is that the solution is not scalable nor reusable. In a previous work, Fernández-Quiruelas et al. [7] developed a framework for executing a climate model on a Grid infrastructure (the EGEE testbed). This first prototype was scalable and was able to exploit all the available Grid resources. However, it could not manage other computing infrastructures such as Clusters or Cloud resources and only allowed the execution of a given experiment. Blanco et al. [8] presented a scientific gateway focused on running climate workflows on Grid resources. This solution is useful to handle experiments with complex workflows among different Grid virtual organizations, but it lacks scalability and scheduling capabilities.

Other groups have explored new possibilities to run applications on HDCIs without a common middleware. Bretherton et al. [9] developed a framework based on web services that can be used to run different climate models by slightly modifying the job configuration. This solution provides some advanced features like the abstraction that allows the execution of different models. However, some important issues, such as job scheduling, efficient data management or recovery of failed simulations, were not handled.

The complexity of climate models poses challenges not only to the HDCI middleware, but also to the standard software used in a single cluster or supercomputer [10]. The execution of climate models usually involves managing complex workflows that produce large volumes of data and occasionally long-term runs lasting for weeks or even months. Furthermore, new trends in climate modeling employ ensemble prediction [11] to sample the uncertainties inherent to the simulations. This technique requires running multiple times the same simulation with varying parameters and thus complicates the experiment management even further. The growth in the number of independent simulations required to perform ensemble prediction has forced the community to find new sources of computing power. The independent nature of these simulations makes them suitable to run on HDCIs. Adapting these models to efficiently take advantage of HDCIs can enormously enlarge the computing power accessible for climate researchers.

In order to simplify the execution of climate and weather experiments, some institutions have created their own frameworks that allow the users to easily perform experiments on their computing facilities. These frameworks provide a set of commands and services that hide the complexity of configuring, running and monitoring all the simulations involved in an experiment. Among other features, these frameworks usually provide means for running the whole experiment workflow unattended and restarting part or the whole experiment in case of failure. The FMS Runtime environment (FRE), the ECMWF Supervisor monitor scheduler (SMS) or the IC3's Autosubmit[1] are some examples [10]. These frameworks have been designed to work with a given model and a single batch-queuing system (usually the one used in the developers institution). Adapting them to use different resource managers or computing configurations might involve a lot of work.

The framework shown in this paper encompasses many features already available in other institutional climate modeling frameworks (e.g. FRE, SMS and Autosubmit), facilitating the management and execution of climate experiments. The main contribution of our framework is the ability to combine heterogeneous, distributed resources to run the simulations. Additionally, its layered design allows to take advantage of most of the developments and easily port the framework to other climate models. Section 3, describes WRF4G, the framework created for running the WRF [12] atmospheric modeling system on HDCIs. This section describes the framework architecture and shows how other applications could also benefit from it to run on HDCIs. It is important to note that this work is not only useful for the climate community, but could also be of interest to other disciplines. Applications that require long running times, large data transfers or complex workflows could take advantage of this work.

In order to illustrate the usefulness of the framework, Section 4 shows how using WRF4G to access Cluster, Grid and Cloud resources simultaneously, the time spent in running a real climate experiment has been reduced 4 times. The experiment performed consisted of 365 independent simulations, which were executed using computing and data resources from our institutional Cluster, a remote Cloud and an EGI Grid infrastructure. Finally, Section 5 presents some conclusions and future work.

2. Hybrid distributed computing infrastructures challenges

Although the combination of Cluster, Grid and Cloud resources on a single HDCI can offer a great computing potential [3,13], leveraging these heterogeneous resources poses several challenges. The distributed nature of these infrastructures complicates tasks such as the monitoring and debugging of applications. Furthermore, combining Cluster, Grid and Cloud resources poses an additional challenge: providing a uniform interface that allows interoperability among different job managers. Interoperability of data resources is also an issue on these infrastructures. Below, we show the main difficulties users might find when executing their applications on an HDCI.

2.1. Application monitoring and debugging

In traditional computing infrastructures, users have direct access to the simulation working environment. They can track the simulation status or find errors inspecting the files generated by the application. Moreover, they can debug errors by re-running the simulation from the last checkpoint file or just stopping the job and running it again with debugging parameters. When the computing nodes belong to an HDCI, very often it is not possible to have direct access to the working directory as the simulations run or even when they have finished. The adaptation of the simulation environment to the policies of each computing resource (i.e. disk, memory and CPU quotas, scratch directory, interconnection among nodes) is another issue that has to be faced when running on HDCIs.

Therefore, a framework to run simulations on HDCIs must allow users to track and control their simulations, and to adapt the simulation environment to each site policy. To provide the framework with such capabilities, it is necessary to orchestrate and monitor the simulation workflow. One approach to application monitoring on HDCIs is the use of a wrapper that registers in a central database all the events produced by the simulation. This wrapper can also transfer the output and checkpoint datasets to the data repositories as they are being produced. Thanks to the checkpoint datasets, in case of crash, the simulations can continue from their last checkpoint, with minimal data loss. This wrapper will also prepare the execution environment (application paths, location of the parallel libraries or scratch file systems...).

Section 3 describes how the central database and WRF wrapper have been implemented in WRF4G.

2.2. Executing jobs on heterogeneous resources

Running computational jobs on heterogeneous infrastructures can be difficult due to the different middleware available. The variety of such middlewares is quite large, even on each HDCI: PBS [14], SGE [15], LSF [16], SLURM [17], LoadLeveler [18] and Condor [19] are examples of Cluster middlewares; gLite [20], Globus

[1] https://redmine.dkrz.de/collaboration/attachments/194/autosubmit.pdf.