

EU-Brazil Cloud Connect: Integrating services for heterogeneous infrastructures

Tuesday, 20 May 2014 14:30 (20 minutes)

EUBrazil Cloud Connect (EUBrazilCC), a project from the 2nd coordinated call of research projects of the European Commission and the Brazilian Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) aims at developing a heterogeneous integrated e-infrastructure for research.

EUBrazilCC leverages a set of components for the use of supercomputing, private cloud and cloud opportunistic resources in desktops. EUBrazilCC will expose these resources through programming frameworks and scientific gateways, easing the adaptation and deployment of the applications that use data and computing resources in both sides of the Atlantic.

EUBrazilCC has 3 use cases demonstrating the feasibility of using distributed and heterogeneous resources: epidemiology, advanced medical systems simulation and climate change.

The integration of the components take into account existing standards to maximize interoperability with provisioning systems and existing infrastructures.

This abstract describes the components and access models for the different types of resources. Tools for transparent access to supercomputers such as CSGrid are combined with execution frameworks to deploy applications on multiple private and public cloud resources (COMPSs) and with complex workflow managers that run on top of different infrastructures through eScienceCentral. EUBrazilCC services include a Parallel Data Analysis Service (PDAS) for big data analytics and the mc2 framework for scientific gateways.

Wider impact and conclusions

EUBrazilCC provides a combination of services addressing different types of resources for computing and big data analytics. The use of OCCI interfaces facilitates the integration of opportunistic resources and private clouds, and the use of execution frameworks hiding the specificities of different resources enables the interoperability required for the applications.

Components are being used in many different contexts, so the impact can extend beyond the use cases of the project. The project defines two tasks to achieve interoperability with major cloud infrastructures such as EGI and HelixNebula.

The project will improve international cooperation between Europe and Brazil, as it was started in the previous EUBrazilOpenBio project, and aims at setting up the basis for future collaborations.

URL(s) for further info

www.eubrazilcloudconnect.eu (under construction)

Description of work

Resources in EUBrazilCC come from three different flavours:

- * Supercomputing resources which accessible through a queue system.
- * On-premises cloud infrastructures, exposing an OCCI interface.
- * Federated cloud opportunistic desktop resources, exposing an OCCI interface.

These infrastructures are made available through three different execution frameworks:

- * COMPSs, which enables the execution of applications automatically exploiting the parallelism without requiring a programmatic adaptation. COMPSs can deploy the required number of VMs to complete the execution of the tasks composing the program in an OCCI-compliant infrastructure.
- * eScienceCentral, which executes complex workflows comprising heterogeneous components on cloud or supercomputer resources.
- * CSGrid, which hides the specificities of different local resource management systems, providing additional data life-cycle management.

Data-intensive applications can benefit from a specific component called Parallel Data Analysis Service (PDAS) to provide near real-time post-processing of data cubes in the order of Terabytes. The PDAS exposes itself as a service and will support the enactment of complex massive jobs on a dedicated infrastructure, as well as on a supercomputer.

Therefore, an application could combine the execution of an MPI-based application on a supercomputer while executing a high-throughput workload on an opportunistic or a dedicated on-premises cloud. One example of this is the whole arterial system simulation, which is achieved by combining the simulation of the heart with the Alya simulator in the BSC's MareNostrum supercomputer with the parameter-sweep coordinated execution of the ADAN simulator of the arterial system, in the SINAPAD resources in Brazil.

Primary author: Dr BLANQUER, Ignacio (UPVLC)

Co-authors: Prof. GIOVANNI, Aloisio (University of Salento); Dr GOMES, Antonio Tadeu A. (Laboratório Nacional de Computação Científica); Dr JACEK, Cala (University of Newcastle); LEZZI, Daniele (Barcelona Supercomputing Center); VILAR BRASILEIRO, Francisco (Universidade Federal de Campina Grande); Dr LIMA, Maria Julia (Pontifical Catholic University of Rio de Janeiro)

Presenter: Dr BLANQUER, Ignacio (UPVLC)

Session Classification: Porting new applications to EGI

Track Classification: Porting applications to the grid and cloud platform (Track Leaders: G. Sipos, D. Wallom)