Infrastructure Systems for Advanced Computing in E-science applications

Olivier Terzo

Istituto Superiore Mario Boella, Via P. C. Boggio 61, Torino, Italy

In the e-science field are growing needs for having computing infrastructure more dynamic and customizable with a model of use "on demand" that follow the exact request in term of resources and storage capacities.

The integration of grid and cloud infrastructure solutions allows us to offer services that can adapt the availability in terms of up scaling and downscaling resources. The main challenges for e-sciences domains will on implement infrastructure solutions for scientific computing that allow to adapt dynamically the demands of computing resources with a strong emphasis on optimizing the use of computing resources for reducing costs of investments.

Instrumentation, data volumes, algorithms, analysis contribute to increase the complexity for applications who require high processing power and storage for a limited time and often exceeds the computational resources that equip the majority of laboratories, research Unit in an organization.

Very often it is necessary to adapt or even tweak rethink tools, algorithms, and consolidate existing applications through a phase of reverse engineering in order to adapt them to a deployment on Cloud infrastructure.

For example, in areas such as rainfall monitoring, meteorological analysis, Hydrometeorology, Climatology Bioinformatics Next Generation Sequencing, Computational Electromagnetic, Radio occultation, the complexity of the analysis raises several issues such as the processing time, the scheduling of tasks of processing, storage of results, a multi users environment.

For these reasons, it is necessary to rethink the writing model of E-Science applications in order to be already adapted to exploit the potentiality of cloud computing services through the uses of IaaS, PaaS and SaaS layer.

An other important focus is on create/use hybrid infrastructure typically a federation between Private and public cloud, in fact in this way when all resources owned by the organization are all used it will be easy with a federate cloud infrastructure to add some additional resources form the Public cloud for following the needs in term of computational and storage resources and release them where process are finished. Following the hybrid model, the scheduling approach is important for managing both cloud models.

Thanks to this model infrastructure every time resources are available for additional request in term of IT capacities that can used "on demand" for a limited time without having to proceed to purchase additional servers.