Introduction to grid and cloud technologies. Practical usage of EMI and OpenNebula middleware

Sergey Belov, Nikolay Kutovskiy
Laboratory of Information Technologies (LIT)

Introduction

Nowadays it's impossible to imagine a scientific research without computer assistance. In many fields scientists have a need to simulate complex processes as well as to store, manage, process and analyze huge amount of data. Grid and cloud technologies let solve such tasks.

The first one allowed to integrate resources of hundred of centers into common platform to be used by thousands of users all over the world as a unified tool for solving similar tasks in different scientific domains.

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

While 1) grid computing is more focused on openness, collaboration and sharing between different sites over a prolonged period of time, cloud computing is designed more for short-term, on-demand provision of computing power and storage and 2) clouds and grids are based on different paradigms with clouds being based on virtualization of resources, and grids being based on the sharing of resources across boundaries, scientists can benefit from both of these two approaches. That's why a modern trend is a synthesis of them both.

JINR actively participates in different international projects which are relied on grid and cloud technologies. JINR has Tier-1 and Tier-2 grid sites what are parts of Worldwide LHC computing grid (http://wlcg.web.cern.ch) infrastructure. Along with these production grid sites there is a training grid infrastructure aimed for learning, development and research activities. It is deployed on facilities of JINR cloud service (http://cloud.jinr.ru) which is also used for a wide range of tasks including trainings in cloud technologies.

Project aims

1. To discuss basic grid and cloud concepts, use cases and benefits of both technologies for science.
2. To provide theoretical knowledge on grid and cloud infrastructures, middlewares' architectures and services (mostly focusing on EMI and OpenNebula).
3. To obtaining practical skills to start unassisted work with EMI and OpenNebula, ones of the most widespread grid and cloud middlewares in the world:
   - security infrastructure (requesting a digital certificate, making a temporary proxy certificate, different manipulations with digital certificates);
   - job management (job preparation, requirements description, running it in the Grid, controlling, obtaining results);
   - data management (copying data to and from grid, other common operations);
   - operations with metadata catalogs (making file's replicas (exact copies), registration data in special file catalogs, assigning additional access rights and metadata descriptions);
   - planning productions using grid information systems;
   - basic skills on clouds (creating virtual machine image, making VM templates, VM creation and deployment, accessing VM, cloud testbed deployment):

This project is to teach applicants for basic knowledge in grid and cloud technologies, to give essential skills needed for further work in these environments and for development of grid/cloud-

enabled applications or deploying and administration of grid/cloud infrastructures. It could be useful for scientists interested in deeper understanding of how grids and clouds could improve their work and for IT specialists supporting computational infrastructure for scientific research. The project is mainly based on half-year educational course conducted in Dubna University and JINR University Centre.

Entry requirements

Good knowledge in Linux user-level skills (including command line operations) as well as administration and programming (including text editors and shell scripting) are expected from applicants.

Project supervisor

Prof. Vladimir Korenkov.