Deployment of the Scientific Cloud Computing Infrastructure in Moldova

Nichita Degteariov, Petru Bogatencov, Nicolai Iliuha
RENAM Association
Chisinau, Moldova
nichita.degteariov@renam.md

Grigorii Horos
Institute of Mathematics and Computer Science of ASM
Chisinau, Moldova
grigorii.horos@gmail.com

***Abstract* — Development of scientific computing infrastructure significantly influences on various fields of science that require complex computational problems solving and unified sets of commonly used services providing. In Moldova over the last years such computing infrastructures as Grid and HPC have started deploying, but new technologies and new applications domains have led to new approaches of providing computational resources and services. Service oriented infrastructures based on Cloud paradigm simplify the use of distributed computing infrastructure, expand users community and applications areas. In the paper the results of federated Scientific Cloud infrastructure deployment in Moldova are described.**

***Keywords* — Federated Cloud infrastructure, on-demand services, integrated Grid and Cloud infrastructure.**

I. Introduction

Cloud is data processing technology in which distributed computing, networking and storage resources are provided to end-users as a service accessible over the Internet. Since its introduction in 2006, Cloud computing penetrates deeply into the various sectors of providing IT services and constantly becomes the more and more demanded approach of offering computational resources. Although there are many commercial realizations of Cloud infrastructures that are widely used for providing various services of common use, unified scientific Cloud infrastructures for supporting research and educational activities are only at the initial stage of development. This can be explained by: 1) specific requirements of complex applications used for research purposes; 2) relatively large amount of testing and debugging procedures; 3) necessity to permanently adopt computing resources to the needs of complex applications; and in many cases 4) a short lifetime of scientific applications that cannot be used as reusable services.

The mentioned above pushed away new initiatives in this area that are mainly based on open source software platforms.

II. First Results of Testing Cloud Infrastructure Deployment

Positive results on deploying open source Cloud computing infrastructures for research and educational area were already obtained during RENAM Association participation in the regional project "Experimental Deployment of an Integrated Grid and Cloud Enabled Environment in BSEC Countries on the Base of g-Eclipse (BSEC gEclipseGrid)" [1]. The main purpose of the Project was to deploy a regional integrated Grid and Cloud enabled environment based on g-Eclipse for the South-East Europe region including Armenia, Georgia, Moldova and Romania [2].

The main approach proposed for implementation in the BSEC gEclipseGrid project is the ability of realization in perspective joint computational environment that will combine Grid and Cloud resources to offer the united enhanced computational power that can adaptively, on demand allocate computational resources depending on workflow requirements (Fig. 1). As an example, if the user requires parallel computational resources, he will submit a job on the Grid; but if the user needs any specific software or environment to solve some special problem, to create new distributed service, he can use a dedicated Cloud service or virtual images for that purpose.

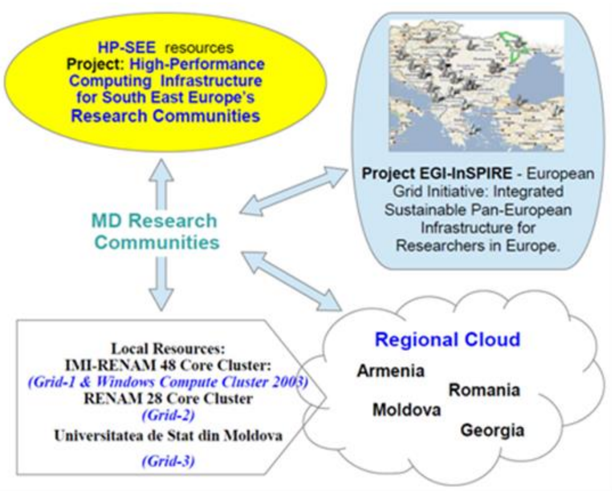

Fig.1. Integration of HPC, Grid and Cloud infrastructures.

In the "BSEC gEclipseGrid" project the distributed Cloud computing platform was deployed using resources provided by each partner country. As a result in Moldova the Cloud testing infrastructure based on OpenNebula 4.4.1 middleware had been deployed. The national Cloud resources have been joined together using the OpenNebula Zones (oZones), which allows centralized management of multiple instances of OpenNebula and especially designed to create federated Cloud infrastructure. Some disk images with preinstalled operation system and main open source applications for solving computational problems in different scientific domains such as environment, meteorology, seismology and astrophysics [3] were integrated in this computing Cloud infrastructure as it is shown in Fig. 2.

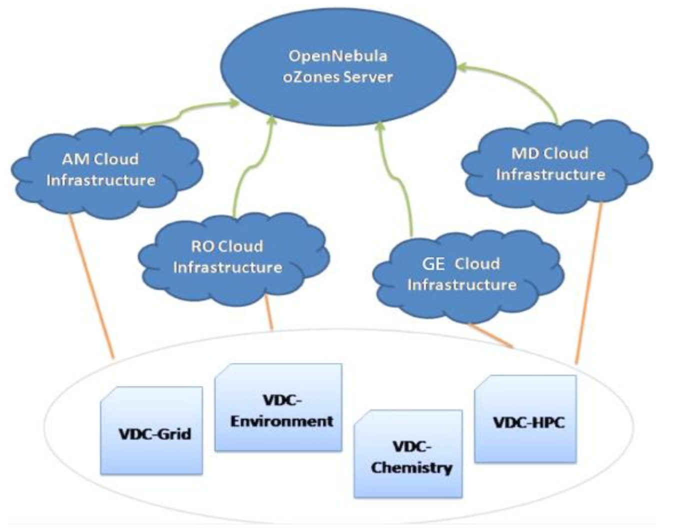

Fig.2. Realization of Federalized access to Cloud.

Organization of federated access to Cloud computing resources is a very important component of every federated cloud computing infrastructure operation. At the moment of cloud infrastructure putting in testing operation there were no unified solutions that could provide unified access to cloud infrastructures and be integrated in the creating Research & Educational identity management federations operated within eduGAIN inter-federation authorization & authentication mechanism (AAI). The first practical results in the area of implementation of unified federated access including various Cloud infrastructures we obtained due to participation in EGI-Inspire AAI Cloud Pilot project "Federated Authentication and Authorization Infrastructure (AAI) for NREN services". Useful experience for the federated AAI deployment and use was also got during administration of OpenStack cloud infrastructure, which stands in a basement of GN3plus project service activity SA2: Testbed as a Service – GEANT Testbed Service (GTS). Work of the team from Moldova in the above-mentioned initiatives pushed us to deploy additional Cloud computing infrastructure on the base of OpenStack middleware.

OpenStack cloud middleware is highly configurable to meet different needs with various computing, networking, and storage options. OpenStack in general has more built-in features and is considered as a leading open source Cloud realization. These middleware main components provide at least the same functionality in common sense that any other open source cloud solutions. It is actively being developed and funded by around 180 organizations all over the world, including such big names as Rackspace, Red Hat, Dell, HP, IBM, Cisco, VMware, Canonical and others.

OpenStack middleware can be used for building any types of Clouds: private, public, mixed and those Clouds can be of any complexity and scale.

III. OPENSTACK CLOUD INFRASTRUCTURE DEPLOYMENT

After thorough acquaintance with OpenStack configuration documentation [4], we elaborated the implementation plan, which specifies ways of computational resources allocation for the second experimental Scientific Cloud segment. Our strategy meant initial deploying of OpenStack with minimal recommended hardware requirements and future enlargement of resources and additional components (services) after initial "test-drive" procedure.

For easier and faster configuration of this Cloud and lacking of idle resources, we decided to allocate one physical server with 4 CPU cores, 8Gb RAM and 250Gb disk space for Compute Node and two virtual machines with 1 CPU core, 1Gb RAM and 30Gb disk space, which will act as Controller Node and Network Node. This approach gave us the possibility to create snapshots and revert them if something goes wrong during configuration, which in general saves deployment time. Although this setup doesn't reach the recommended minimal configuration (see Fig. 3) but it is more than enough for testing infrastructure and can be easily scaled up in the future by adding if required more worker nodes, networks, data storage devices, etc. to the existing infrastructure.

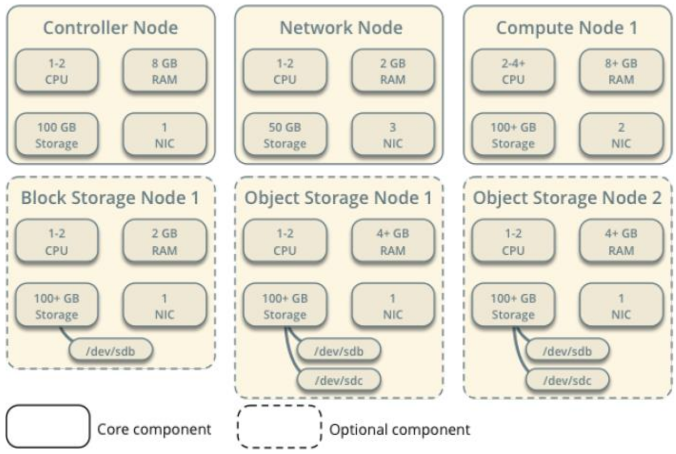

Fig. 3. Minimal OpenStack architecture example with networking (neutron).

We have used the latest at the moment of deployment Ubuntu Server 14.04 LTS as a base operation system for all nodes and the latest version of OpenStack release "Juno". This infrastructure is interconnected via two dedicated Gbit switches - one for management and one for data network. Internet connectivity and internal networking for virtual machines provided via the Network Node. The Network Node runs SDN (Software Defined Network) technology software – Open Virtual Switch. It creates virtual network infrastructure for virtual machines and segregates different network slices using GRE (Generic Routing Encapsulation) tunneling protocol. It also supports many networking protocols including OpenFlow and is a very flexible and powerful instrument that can be compared to proprietary switch facility used in the licensed private VMware Cloud products.

Network layout for OpenStack Cloud testing infrastructure is presented in Fig.4.

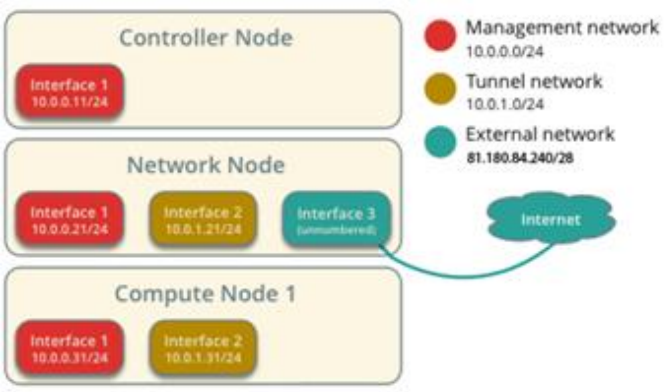

Fig. 4. Network layout for OpenStack Cloud testing infrastructure.

This network model provides connectivity between all VMs in a tenant network, regardless of which compute node the VMs reside in, and segregates VMs, which are in different tenant networks. In this sense, GRE tunnels are an alternative to VLANs. Open Virtual Switch based network infrastructure in our second Cloud testing infrastructure is a big step forward compared to previous one built using standard Linux bridge module.

Another serious enhancement compared to our first OpenNebula based Cloud testing infrastructure is that virtual machine's disk images use deltas of the initial images, which in turn are already stored on the Compute Node, so they should not be copied from the outside before deployment. This significantly reduces VM deployment time and economizes storage space.

For organization of unified access to the both deployed cloud segments, we consider important the implementation of two existing AAI solutions for cloud infrastructures that are being developed by European Grid Initiative within AAI Cloud Pilot project "Federated Authentication and Authorization Infrastructure (AAI) for NREN services". After practical deployment of the National Identity Management federation finalizing in Moldova [5] there will be elaborated technical profiles and installed appropriate software for ensuring federated access to the deployed scientific cloud infrastructures.

IV. Conclusion

The created regional cloud testing infrastructure, although it has limited computational resources, is the successful example of adaptation of new technologies and open source software platforms for providing computational resources to scientific community.

There are perspectives to continue development of the scientific Cloud infrastructure and technologies at national and regional levels. We intend together with BSEC gEclipseGrid project partners to continue work focused on adaptation and implementation of new open source tools for extension of the created federated cloud infrastructures. Other perspective direction that can get new experience for national cloud infrastructure and services development is cooperation with partners within new projects initiating by European Grid Initiative. It comprises integration of the national Grid and Cloud computing infrastructures.

Acknowledgment

This work is significantly supported by the regional project "Experimental Deployment of an Integrated Grid and Cloud Enabled Environment in BSEC Countries on the Base of g-Eclipse (BSEC gEclipseGrid)" (sponsored by the Organization of the Black Sea Economic Cooperation) and by the European Commission under EU FP7 project EGI-Inspire (contract number 261323).

References

[1] P. Bogatencov, N. Degteariov, N. Iliuha, and P. Vaseanovici "Implementation of Scientific Cloud Testing Infrastructure in Moldova," Proceeding of The Third Conference of Mathematical Society of the Republic of Moldova 2014, Chisinau, Moldova, pp. 463–466, August 2014.

[2] Experimental Deployment of an Integrated Grid and Cloud Enabled Environment in BSEC Countries on the Base of g-Eclipse (BSEC gEclipseGrid) project: http://blacksea-cloud.net/

[3] Hrachya Astsatryan, Andranik Hayrapetyan, Wahi Narsesian, Peter Bogatencov, Nicolai Iliuha, Ramaz Kvatadze, Nugzar Gaamtsemlidze, Vladimir Florian, Gabriel Neagu, Alexandru Stanciu. "Deployment of a Federated Cloud Infrastructure in the Black Sea Region". Computer Science and Information Technologies. Proceedings of the CSIT Conference, September 23-27, 2013, Erevan, Armenia, pp.283-285. ISBN 978-5-8080-0797-0.

[4] OpenStack documentation and materials: http://docs.openstack.org/

[5] BOGATENCOV P., POCOTILENCO V. Implementation of national IdP Management Systems for Access to Resources of European R&E E-Infrastructures. "Networking în Education and Research", Proceedings of the 11th RoEduNet IEEE International Conference, Sinaia, Romania, January 17-19, 2013, pp. 96-100. ISSN-L 2068-1038.