OPTIMIZATION OF THE JINR CLOUD'S EFFICIENCY

N. Balashov[1,a], A. Baranov[1], S. Belov[1,2], I. Kadochnikov[1,2], V. Korenkov[1,2], N. Kutovskiy[1], A. Nechaevskiy[1], I. Pelevanyuk[1]

[1] *Laboratory of Information Technologies, Joint Institute for Nuclear Research, 6 Joliot-Curie, Dubna, Moscow region, 141980, Russia*

[2] *Plekhanov Russian University of Economics, 36 Stremyanny per., Moscow, 117997, Russia*

E-mail: [a] balashov@jinr.ru

Clouds built on Infrastructure-as-a-Service (IaaS) model (such as JINR Cloud) gave us new universal and flexible tools and ways to use computing resources. These new tools may help scientists speed up their research work, though at the cost of a significant drop (compared to the more traditional systems in science such as grid) of the overall utilization efficiency of an underlying infrastructure. The talk covers Smart Cloud Scheduler project aimed at optimizing the performance of the IaaS-based clouds, including its architecture, development status and plans. The project includes development of a software framework that would allow one to implement custom schemes of dynamic reallocation and consolidation of virtual machines. The resulting system will give a possibility to dynamically rebalance the cloud workload in an automated fashion in order to increase the overall infrastructure utilization efficiency.

Keywords: cloud computing, virtualization, self-organization, intelligent control, datacenters, VM consolidation

1. Introduction

Cloud technologies gave us a convenient and universal way to organize and use large computing infrastructures. The JINR Cloud [1] is an example of such an infrastructure which is built on OpenNebula platform and is based on Infrastructure as a Service model. The JINR Cloud was initially born in the Laboratory of Information Technologies as a small project with the main goal to facilitate the deployment of virtual infrastructures which were used to conduct trainings for Grid users and system administrators. Eventually, the project evolved into a general-purpose cloud service which was made available for all JINR employees and which is now widely used by more than 150 users (including students, researchers, system administrators and software developers) for a variety of different activities ranging from hosting production services to using virtual machines as personal computers.

In our previous works [2-4] we showed the possibility of optimizing JINR Cloud efficiency and suggested a possible solution in the form of a software framework – SmartScheduler. In this paper we give a status report of this project.

2. SmartScheduler overview

The SmartSchedulder project contains a software framework which can be extended with custom interfaces and algorithms, and it also contains example interfaces and algorithms developed specifically for the JINR Cloud environment. The main components of the system are depicted on figure 1 and include:

1. Administrative interface based on Django web framework which helps control and configure custom scheduling policies and algorithms.
2. Relational database which can be any database supported by Django and is used to store configuration parameters and additional data.
3. A library providing interfaces for interacting with the cloud. Right now it is highly dependent on the OpenNebula data format but we will work out a more general approach to ease the development of other cloud platforms support.
4. Monitoring and statistics collection system based on Icinga2 with custom sensors, InfluxDB and Grafana.
5. A simple daemon used to execute the scheduling strategies and keep them running.

The light yellow boxes on figure 1 show the components that are not included in the framework and can be represented by different solutions. In the next two sections we give an example of how these components are implemented for the JINR cloud environment.

3. Cloud platform

The cloud platform could be any platform implementing IaaS model. The minimal requirements to the cloud platform are the following: it must provide API to create and delete VMs, to migrate a VM from one host to another and it also has to expose interfaces for getting monitoring information (which is covered in more details in the next section).

The testing of the SmartScheduler framework will be conducted on the testbeds of the JINR Cloud and will eventually be put into operation in the JINR Cloud production infrastructure. Since this infrastructure is based on OpenNebula platform the framework will include interfaces for interacting with OpenNebula-based clouds. These interfaces then could be used as an example to start with when developing other platforms support, e.g. OpenStack.

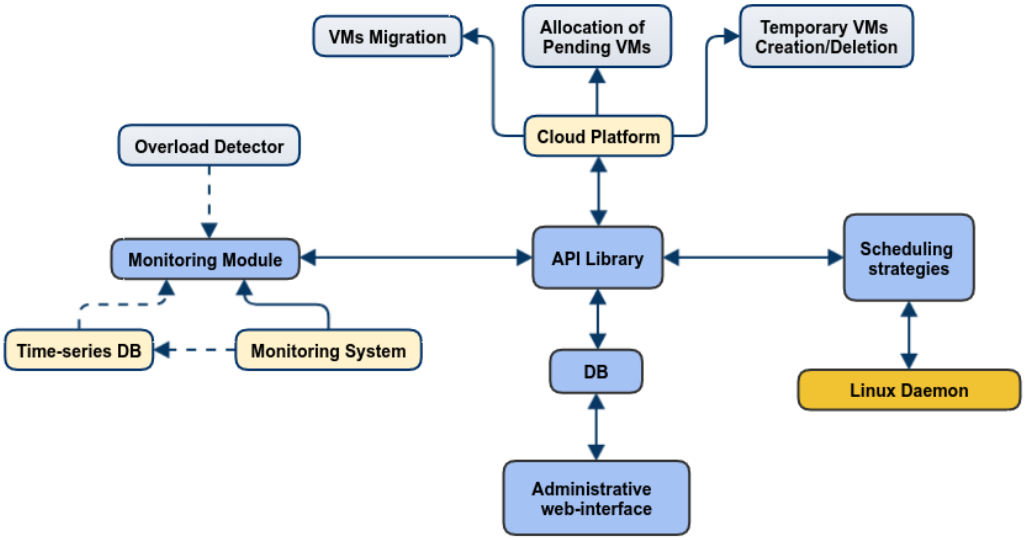

Figure 1. Main components of the SmartScheduler framework

One of the features of the JINR Cloud is that it supports two types of hypervisors: KVM and OpenVZ. Taking this into account, the SmartScheduler must be able to distinguish different hypervisors and to treat them independently.

4. Monitoring module

Monitoring and statistics collection is one of the most important parts of any complex computing infrastructure including clouds. Nevertheless, there is no "one size fits all" solution to build monitoring systems for complex computing infrastructures. Moreover, the monitoring systems for complex computing infrastructures in most cases are based on a set of different software solutions, each solving one particular task: data collection, issue detection, creation of reports, visualization, etc. The JINR cloud monitoring system is also comprised of a number of different software solutions and its scheme is shown on figure 2.

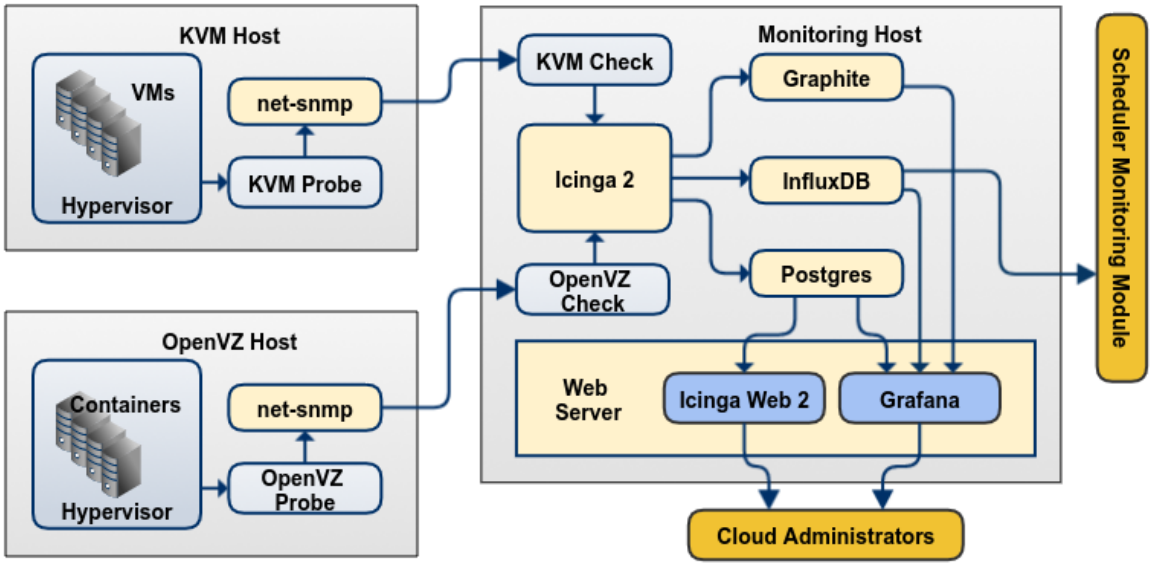

Figure 2. JINR Cloud monitoring system

The JINR Cloud monitoring system is composed of the following software solutions:

- Icinga 2 which is used for collecting performance metrics and can be used for alarming.
- InfluxDB and Graphite are used for storing those metrics collected by Icinga sensors.
- Grafana is used for visualizing the JINR Cloud operation.

The SmartScheduler project will include all necessary interfaces to interact with the monitoring system described above and a set of guides and recommendations on how to deploy such a system. Nevertheless, it is possible to use other monitoring solutions and the SmartScheduler can be adjusted to fit with other possible architectures of monitoring systems.

5. Conclusion

Presently, the basic software framework of the SmartScheduler project is designed, a functional prototype is fully implemented and is being tested on the JINR Cloud's testbeds. In the next stage the scheduling policies will be developed based on algorithms suggested in our previous works [2-4].

In this paper we gave an overview of the SmartScheduler project. To get more technical details, refer to our GitHub repositories [5] for all the source-code and manuals.

6. Acknowledgements

The project was supported by the RFBR grant 15-29-07027.

References

[1] Baranov A.V. et al. JINR cloud infrastructure evolution // Physics of Particles and Nuclei Letters, ISSN 1547-4771, eISSN: 1531-8567, 2016, vol. 13, No. 5, pp. 672–675. DOI: 10.1134/S1547477116050071.

[2] N. Balashov, A. Baranov, V. Korenkov "Optimization of over-provisioned clouds" // Physics of Particles and Nuclei Letters. — 2016. — Vol. 13, No. 5. — P. 609–612.

[3] Kadochnikov I., Balashov N., Baranov A., Pelevanyuk I., Kutovskiy N., Korenkov V., Nechaevskiy A. Evaluation of monitoring systems for metric collection in intelligent cloud scheduling // CEUR Workshop Proceedings, Vol. 1787, Pages 279-283, 2016.

[4] N.A. Balashov, A.V. Baranov, I.S. Kadochnikov, V.V. Korenkov, N.A. Kutovskiy, A.V. Nechaevskiy, I.S. Pelevanyuk Software complex for intelligent scheduling and adaptive self-organization of virtual computing resources based in LIT JINR cloud center // Izvestiya SFedU. Engineering Sciences, Vol. 12 (185), pp. 92-103, 2016.

[5] "Homepage of SmartScheduler project's code repositories", https://github.com/jinr-lit, accessed on: 26.10.2017