Advanced Scheduling in IaaS clouds

Joint Institute for Nuclear Research

N. Balashov, A. Baranov, S. Belov, I. Kadochnikov, V. Korenkov, N. Kutovskiy, A. Nechaevskiy, I. Pelevanyuk

CHEP`2018, Sofia, Bulgaria

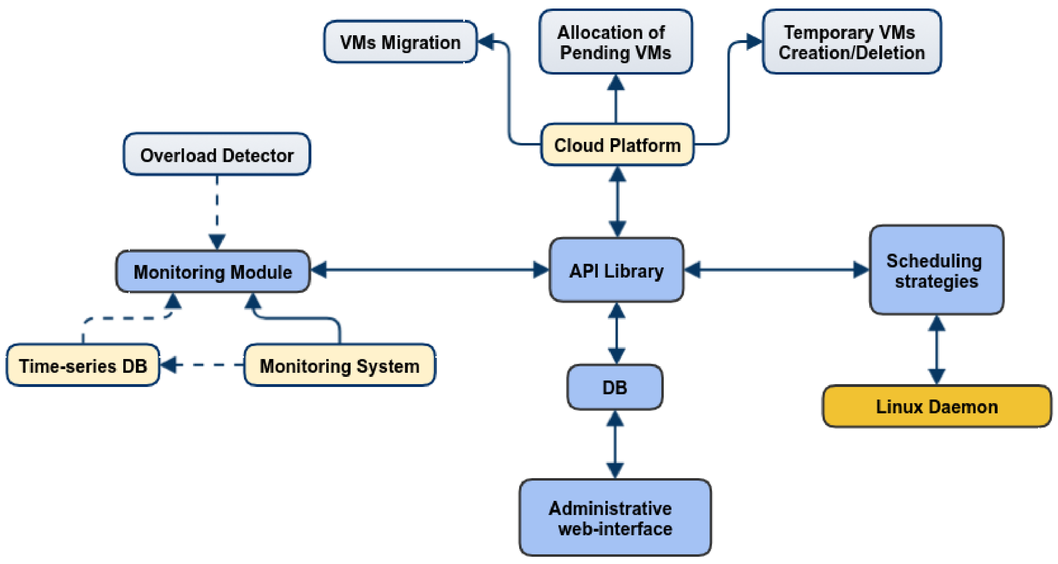
INTRODUCTION

The JINR cloud service is used as a universal computing resource supporting workload-intensive scientific computations as well as low-load activities. The complexity of modern software libraries and applications makes it hard to predict possible workloads generated by the software. For this reason cloud resources are often over-allocated leading to a high degree of underutilization of underlying equipment. To address this issue a smart scheduler for the JINR cloud is was developed. To increase the cloud resources utilization the smart scheduler uses consolidation of virtual machines (VMs) combined with a method for dynamic re-allocation of virtual machines.

System Architecture

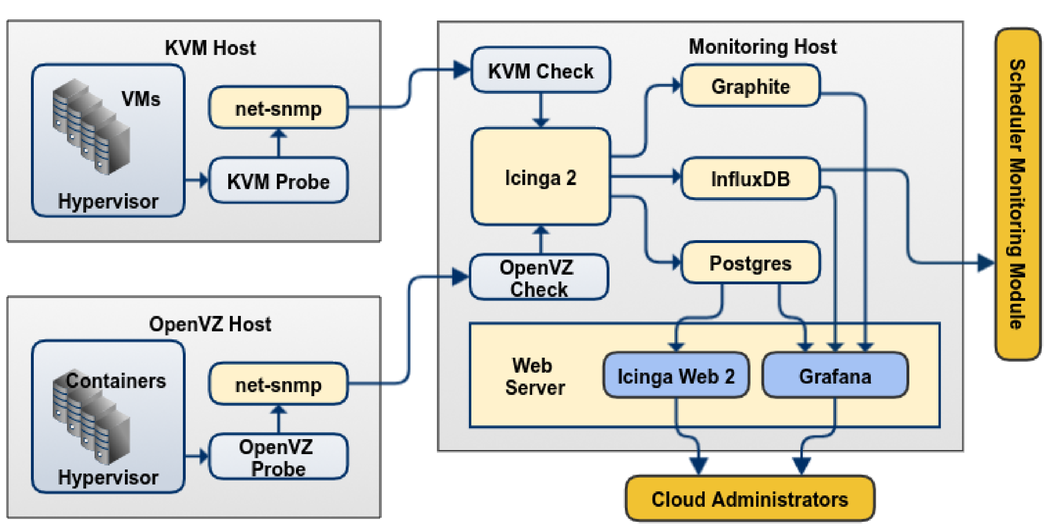
The system is designed to be modular with each module being replaceable:

- the API library is the system core that links all the modules;
- the Linux daemon starts scheduling strategies and keeps track of them;
- the monitoring module provides monitoring information and historical records;
- the integrated database can be used to store additional meta-information needed by the algorithms;
- the API includes implementations of different cloud platform interfaces (currently just OpenNebula XML-RPC).



The proposed framework allows to build a custom scheduling system with a specific module-set to suit the needs of a particular cloud service.

Monitoring and Data Storage

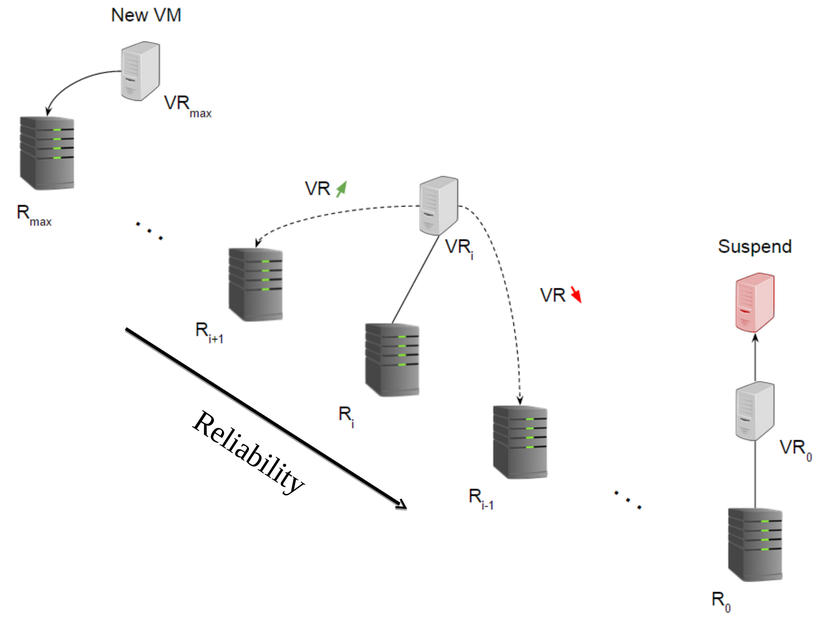


The impelmented monitoring system, the scheme of which is shown above, consists of many components giving the system all the necessary functionality:

- constant monitoring of the system state and instant reaction to failures;
- accumulation and storage of historical performance records for a long period of time;
- aggregation of the old data to reduce the database size;
- convenient visualization of the monitoring data;
- all components are open-source.

The project was funded by the RFBR grant 15-29-07027

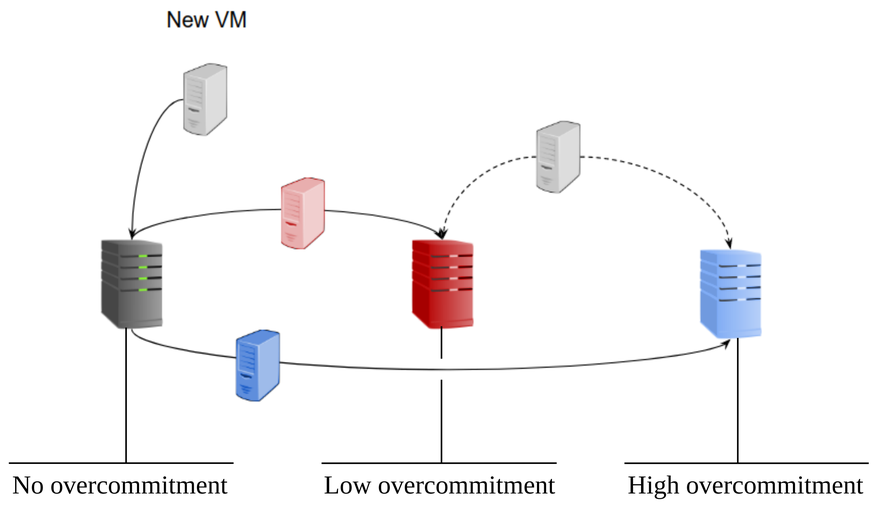
Basic Method



The proposed method of dynamic virtual machines reallocation in the cloud environment is based on ranking of hosts and VMs:

- Host ranks are statically defined and have different overcommitment ratios
- The lower the rank, the higher the overcommitment ratio
- Hosts can be grouped forming same-rank clusters
- VMs get ranks dynamically depending on the workload they generate
- Gives control over reliability of VMs by adjusting overcommitment ratios;
- Helps detect unused VMs.

Scheduling Strategy Used at JINR



The algorithm implemented at JINR is based on a two-rank strategy:

- VMs are classified as “hot”, generating high loads, and “cold”, low-active VMs
- Hosts are grouped into three clusters holding hot, cold and not classified yet VMs
- Free resources are occupied by short-living VMs processing Grid jobs

Conclusions

Cloud technologies represent an approach to providing convenient and universal way of performing scientific computations, but the experience of managing JINR cloud shows that it also leads to a lower degree of resources utilization efficiency compared to other computing models, such as Grid. The developed system and methods are intended to smooth out the drop in efficiency and give the community a framework for further developments and improvements of cloud scheduling approaches.