

JINR MEMBER STATES CLOUD INFRASTRUCTURE

**N.A. Balashov¹, A.V. Baranov¹, N.A. Kutovskiy^{1,a}, Ye. M. Mazhitova^{1,2},
R.N. Semenov^{1,3}**

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

² *Institute of Nuclear Physics, 1 Ibragimova street, Almaty, 050032, Kazakhstan*

³ *Plekhanov Russian University of Economics, 36 Stremyanny per., Moscow, 117997, Russia*

E-mail: ^a kut@jinr.ru

One of the possible ways to speed up scientific research projects which JINR and its Member State organizations participate in is to join computational resources. It can be done in a few ways one of which is to build distributed cloud infrastructures integrating local private clouds of JINR and organizations from its Member States. To implement such scenario, a cloud bursting based approach was chosen since it provides more flexibility in resources management of each cloud integrated into such distributed infrastructure. For the time being a few private clouds from some organizations from the JINR Member States have been already integrated with the JINR one with help of custom cloud bursting driver developed by the JINR team. Various aspects of that activity are covered in more details in the article as well as implemented and planned changes in the JINR cloud such as a smart scheduler deployment in production, migration from DRBD high-availability (HA) to native OpenNebula HA, Ceph storage deployment and VMs' disks migration to it and others.

Keywords: cloud computing, OpenNebula, clouds integration, cloud bursting, ceph

© 2017 Nikita A. Balashov, Alexandr V. Baranov, Yelena M. Mazhitova,
Nikolay A. Kutovskiy, Roman N. Semenov

1. Clouds integration

To achieve scientific results faster, the research organizations join own resources. There are few ways to do that. One of them is to join one of the worldwide grid infrastructure (e.g. WLCG). However, the grid site deployment and maintenance can be difficult for small research institutions in contrast to a local private cloud. For that purpose a custom driver [1] for resources integration of the private clouds from the JINR Member State organizations was developed by the JINR cloud team. That driver implements the so call “cloud bursting” integration model.

At the moment of writing the article a progress with clouds integration of each of the listed in the Table 1 organizations is at the various stage.

Table 1. List of the JINR Member state parnter organizations involved into clouds integration activity

Organization	Location	Share of CPU cores	Share of RAM, GB	Share of disk capacity, TB	Status
JINR	Dubna, Russia	600	2000	14	integrated
Plekhanov Russian University of Economics	Moscow, Russia	10	20	5	integrated
Institute of Physics	Baku, Azerbaijan	24	256	16	integrated
Bogolyubov Institute for Theoretical Physics	Kiev, Ukraine	10	20	4	integrated
Institute for Nuclear Problems	Minsk, Belarus	10	20	10	integrated
Nazarbaev University	Astana, Kazakhstan	-	-	-	deployment
Institute of Nuclear Physics	Astana, Kazakhstan	-	-	-	deployment
Yerevan Physics Institute	Yerevan, Armenia	-	-	-	deployment
Georgian Technical University	Tbilisi, Georgia	-	-	-	deployment
Institute for Nuclear Research and Nuclear Energy	Sofia, Bulgaria	-	-	-	deployment
Sofia University	Sofia, Bulgaria	-	-	-	deployment
University “St. Kliment Ohridski”	Bitola, Macedonia	-	-	-	negotiations

Figure 1 shows on the map the locations of such distributed cloud infrastructure participants.

The cloud bursting driver allows one to link each cloud with another one following “peer-to-peer” (or it’s better to say “cloud-to-cloud”) model. A complexity of the maintenance of such a distributed cloud infrastructure increases sufficiently with the growth of the number of its participants.

That’s why a research work has been started recently to evaluate an architecture in which all clouds are interconnected with each other and with the help of an additional layer based on grid middleware with grid services managed centrally at JINR forming such way an united computing

facility (see Figure 2). Possible candidates for workload management system (WMS) of such a glue layer can be PanDA, DIRAC WMS, Vac/VCycle, VPManger/VCondor.



Figure 1. Map with the distributed cloud infrastructure participants

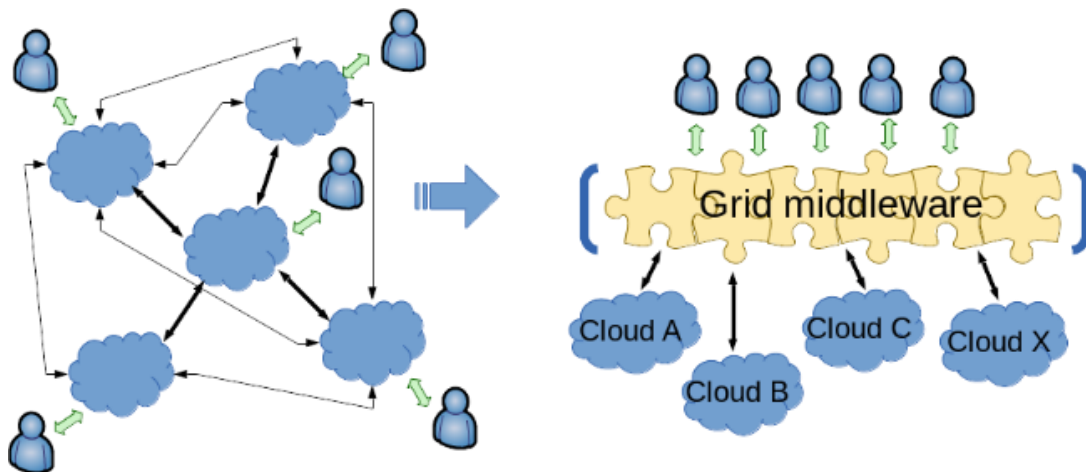


Figure 2. Possible schemes of clouds integration

2. Upcoming high availability setup based on a Raft consensus algorithm

The JINR cloud [2] is based on the OpenNebula platform [3]. Since its version 5.4 it became possible to build a high availability (HA) setup of front-end nodes (FNs) using just built-in tools implementing a Raft consensus algorithm [4]. It provides a cheap and easy way to make an OpenNebula master node replication and eliminates a need in MySQL clustering. A testbed based on the OpenNebula v5.4 was deployed to evaluate all its new features and get an experience in migrating from 4.12 version which our production cloud is running on now.

Apart from that a Ceph-based [5] software-defined storage (SDS) was deployed as well. Various benchmarks are running now against it to evaluate its performance and reliability.

All these at the end should lead to the new HA OpenNebula JINR cloud architecture shown on the Figure 3.

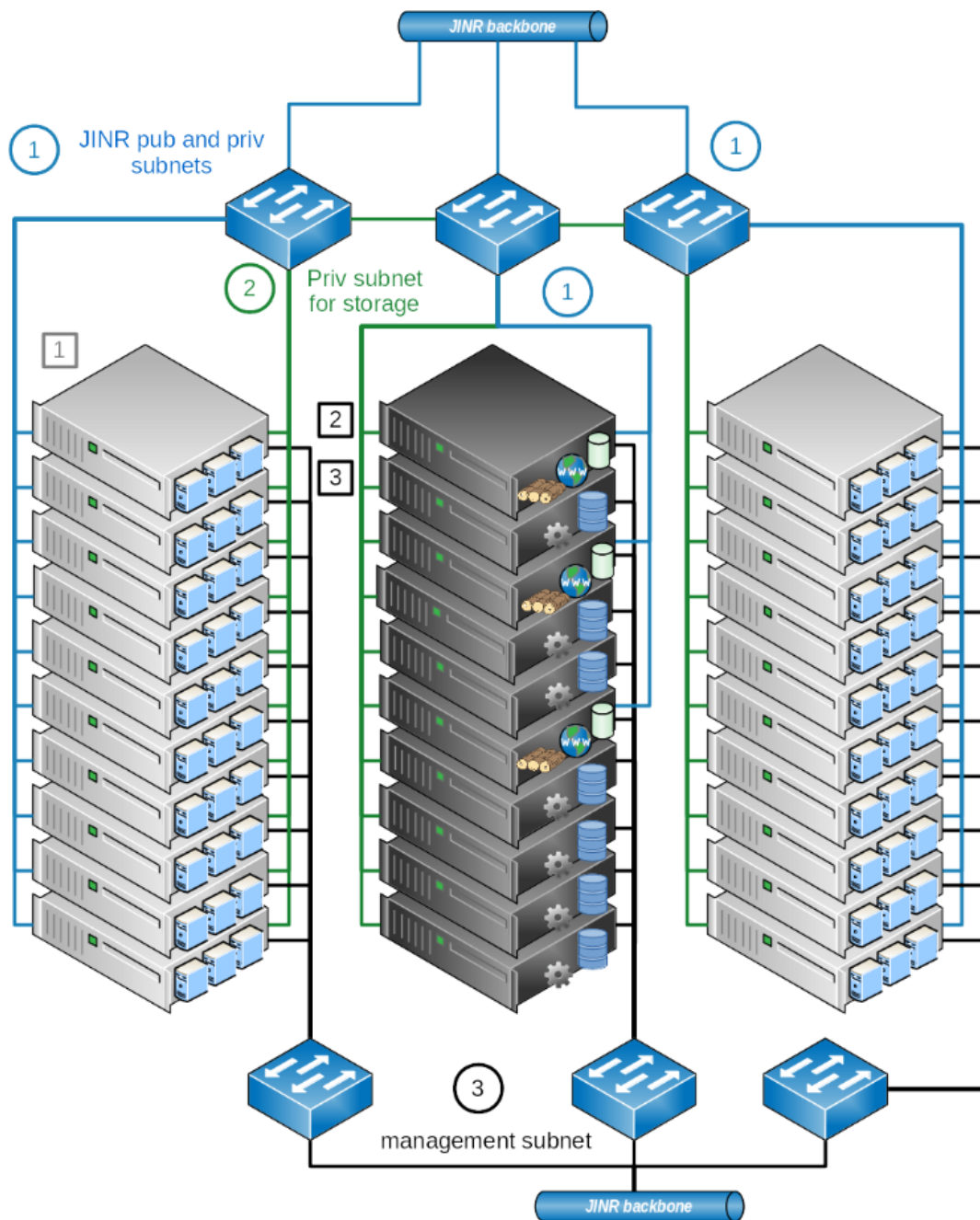


Figure 3. A schema of the new HA JINR cloud setup based on Raft consensus algorithm and Ceph SDS

There are the following types of the OpenNebula 5.4 testbed servers shown on the Figure 3

- Cloud worker nodes (CWNs) which host virtual machines (VMs) and containers (CTs) are marked on the Figure 3 by numeral “1” in a grey square;
- Cloud front-end nodes (CFNs) where all core OpenNebula services including database, scheduler and some other ones are deployed (such hosts are marked on the Figure 3 by black numeral “2” inside the same color square);
- Cloud storage nodes (CSNs) based on Ceph SDS for keeping VMs’ and CTs’ images as well as users’ data (marked by black numeral “3” inside the same color square).

All these servers are connected to the same set of networks:

- JINR public and private subnets (they are marked on the Figure 3 by blue lines near which the numeral “1” is placed in the same color circle and signed as “JINR pub and priv subnets”);

- An isolated private network dedicated for a SDS traffic (dark green lines with numeral “2” in a circle of the same color and signed as “Priv subnet for storage”);
- Management network (black lines with numeral “3” in a circle of the same color signed as “management subnet”).

All network switches excluding those which are for the management network have 48 ports with 10 GbE each as well as four 40 Gbps SPF-ports for uplinks.

All CSNs apart from HDDs disks for data have SSD disks for caching.

In addition to increasing a common availability and reliability of the JINR cloud operation it is planned that the JINR cloud in such a new configuration will sufficiently improve users' experience in terms of VMs/CTs deployment and migration times due to hosting its images on the fast SDS shared across CWNs as well as by providing users with the additional disk space to keep their data on the same SDS permanently and even more reliably.

3. Conclusion

The JINR cloud is actively developing: the new HA setup based on Raft consensus algorithm and Ceph-based SDS are under evaluation, new hardware is bought and put into production as well as private clouds of partner organizations from the JINR Member States are integrated into the united distributed cloud infrastructure. A research work on possible alternatives for clouds interconnection is in progress. More experiments show interest in the JINR cloud resources usage. Their needs and workflows are under evaluation to implement in our cloud environment.

References

- [1] A. V. Baranov et al. Approaches to cloud infrastructures integration // *Computer Research and Modeling*, vol. 8, no. 3, pp. 583–590, 2016
- [2] 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
- [3] "OpenNebula web-portal" [Online]. Available: <http://opennebula.org>. [Accessed on 30.10.2017]
- [4] "The Raft consensus algorithm web-site" [Online]. Available: <https://raft.github.io/>. [Accessed on 30.10. 2017]
- [5] "Ceph software defined storage web-portal" [Online]. Available: <http://ceph.com/>. [Accessed on 30.10.2017].