Proceedings of the XXVI International Symposium on Nuclear Electronics & Computing (NEC'2017)
Becici, Budva, Montenegro, September 25 - 29, 2017

SERVICE RELIABILITY IN THE CLOUD OF DATA CENTERS UNDER OPENSTACK

P.V. Fedchenkov[1], S.E. Khoruzhnikov[1], V.A. Grudinin[1], O.L. Sadov[1], A.E. Shevel[1,2,a], A.B. Kairkanov[1], O.I. Lazo[1], A.A. Oreshkin[2]

[1] *Department of Network and Cloud Technologies, ITMO University, St.-Petersburg, 197101, Russia*

[2] *National Research Centre "Kurchatov Institute" PETERSBURG NUCLEAR PHYSICS INSTITUTE, Gatchina, 188300, Russia*

E-mail: [a] shevel.andrey@gmail.com

University ITMO (ifmo.ru) is developing the cloud of geographically distributed data centers under Openstack. The term “geographically distributed” in our proposal means data centers (DC) located in different places far from each other at least by many tens of kilometers. Authors follow the conception of “dark” DC, i.e. the DC has to perform normal operation without permanent maintainers even with minor problems (single machine or a number of disk drives went down). In such “dark” DC the staff might visit DC if required to fix appeared problems. Authors describe thoughts and experiments with service reliability for cloud of DCs under Openstack.

Keywords: Openstack, service reliability, cloud of data centers, SDN, CEPH

1. Cloud of Data Centers

The main cloud services to be provided in Data Centers (DCs) include Computing (VMs), Storage, Data transfer channels. Here it is considered several globally dispersed DCs which are functioning like united cloud of DCs. There are several advantages for such the architecture [1]: improved reliability of the data storage due to existing several replicas of data on several DCs, improved reliability of computing due to possibility to continue the user computing when one DC went down. We can eliminate a single point of failure and achieve High Availability (HA) deployment. HA is the ability for the cloud to continue functioning after failure of one or more of the hardware or software components [2].

All DCs are running under Openstack (ocata) [3]. It means that users have ability to be logged to Openstack panel Horizon and use computing and other Openstack services divided by several zones.

Obviously DCs need the data links in between each other. In the project it is distinguished two types of virtual channels: regular channels (encrypted) for data transfer and control channels (with OpenVPN). Also the cloud needs distributed storage which is implemented with CEPH [4]. Further discussion is concentrated to computing service reliability. The conditions where the reliability is planned to be increased in this project are following:

- User can be logged into any Openstack instance in the cloud of DCs and start up Virtual Machine (VM) on any DC in the cloud;
- If DC went down where the user uses Openstack and user's VM was running in the same DC, then user can log into another Openstack instance in different DC in the cloud and start new user's VM instance. Presumably all intermediate results from broken DC are available in distributed cloud storage.

In the table 1 several options shown when one DC in the cloud went down. From the table can be seen that even with cloud consisting of two DCs the reliability to get the computing service for user is significantly higher to compare with one DC.

Table 1. Possible problems and solutions in distributed DCs

Number	**Situation**	**Solution**
1	User was logged into Openstack instance in broken DC and user's VM was running in the same DC.	User might log on into another Openstack instance and start VM again in different DC of the cloud. It is assumed user's VM in broken DC is dead but intermediate data are kept in distributed storage which is available in different DCs.
2	User was logged into Openstack instance in broken DC but user's VM is running on different DC.	Nothing bad is expected for running user's VM. User might log on into another Openstack instance in different DC and watch his running VM.
3	User was logged into Openstack instance in running DC but user's VM was running in broken DC	It is assumed that user's VM is dead but intermediate results are saved in distributed storage and available in different DCs. User might start up new VM instance in different DC and use saved earlier data from distributed storage.
4	One of the DCs became broken but user is logged in Openstack instance in different DC and user's VM is running in running DC.	No affection to user and user's VM.

To increase computing service reliability it is used at least one independent Openstack deployment in each DC. To make united cloud with decentralized administration [5], which is available for users as integrated resource it is, required to build up the special interconnection of the all Openstack deployments in different DCs. With all measures being discussed further in the paper it is planned to increase significantly the computing service reliability in according to the above table.

2. Cloud Services

As a storage subsystem, the CEPH platform is chosen to provide long-term distributed storage. It is configured and deployed the CEPH storage cluster to geographically distributed servers using the Internet line with nominal capacity 1 Gbit. Servers are divided into two groups (two sites) which are far from each other at a distance of about 40 Km. In other words it has been configured the testbed for geographically distributed storage. The only disadvantage of this configuration is the relatively low speed of replication. At the same time this testbed is quite enough for reliable archiving purposes for relatively long time (years). Clients might use the virtual storage when part of data are located on one group of storage servers and another part is in another group in the manner shown on the fig.1. All the data transfers between remote sites are performed over network tunnel with OpenVPN.

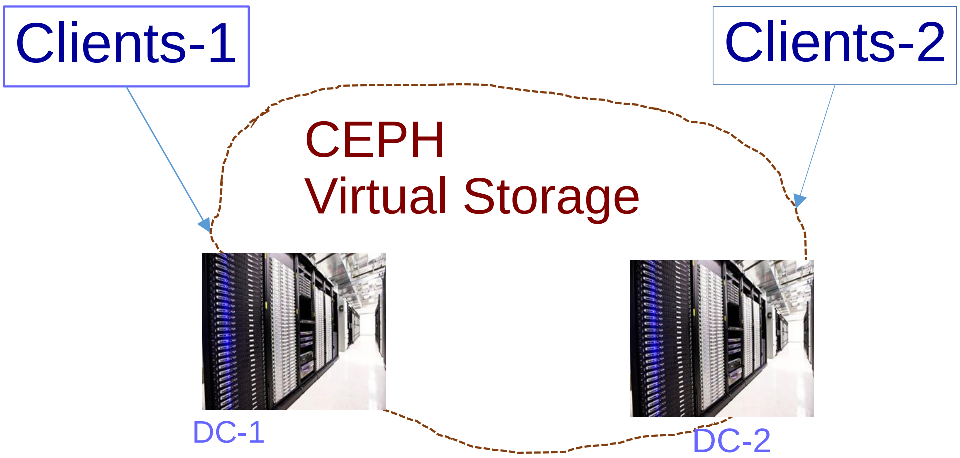

Figure 1. CEPH storage cluster distributed between DCs

Figure 1 represents an approach of using CEPH virtual storage. We can create a set of clusters, which may use servers of DC1, DC2 or both of them. In the case of using both DCs in the same CEPH cluster, we can achieve data service High Availability (HA), where clients (end users or modules of system), may request both of DCs to get block devices. It helps to start VM from another DC without data loss. Also it improve Active/Passive possibilities of clusters.

3. The computing reliability testbed

In the proposed testbed (fig.2), we focused on providing fault tolerance for computing. Components of the OpenStack do use common MySql database to store information about project settings, created virtual entities, provide message delivery. In order to do that database available on both DCs it were created MySql cluster in Master-Master mode. To make possible control interconnection by message quesues between sites it was created RabbitMQ. Both keystone database content are available in each DCs, and therefore you can make changes from any working DC.

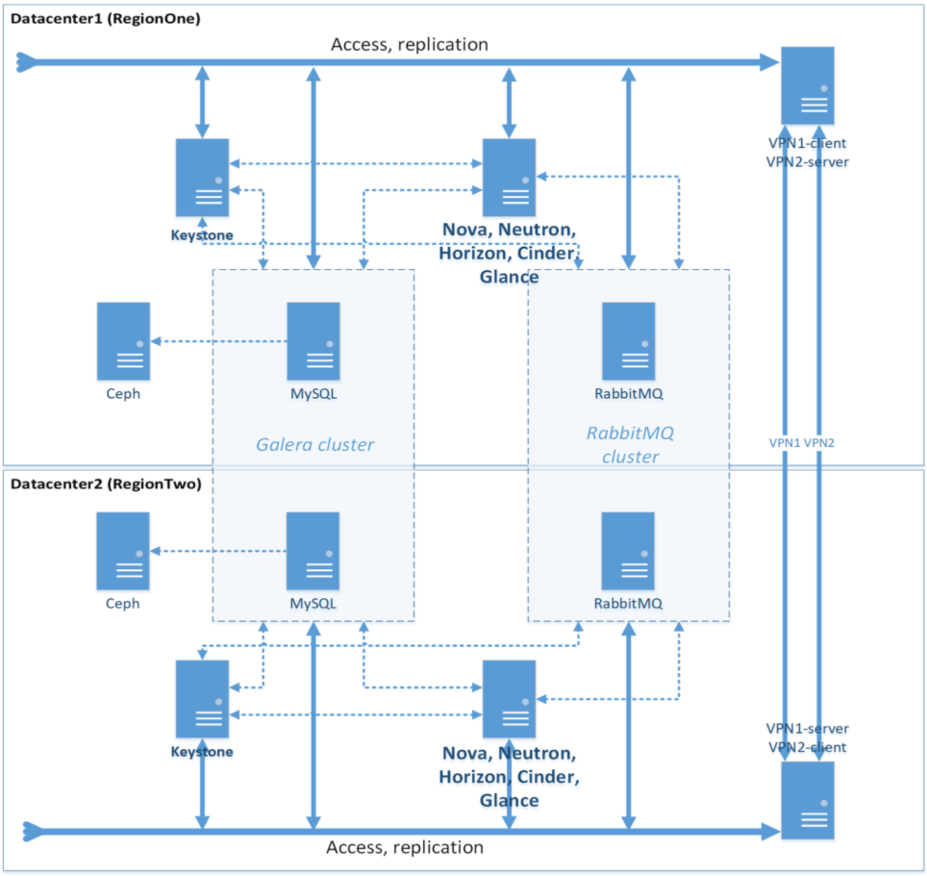

Figure 2. Two independent OpenStack setup on two DCs with common keystone database

Figure 2 shows the configuration of the OpenStack on two DCs. Each DC OpenStack represent the separate region. Each region has full independent OpenStack controller (Nova, Neutron, Horizon, Cinder, Glance, Keystone). Both DCs are in same IP address space interconnected with OpenVPN. This option allows to use local services in the event when one DC is down. The common keystone database (MySql cluster) and the message queue cluster are used. Thanks to the single common for both regions keystone database, each of the keystone can access the keystone data of another region.

In created testbed is possible to logon into Horizon DC1, start and see running VMs and other resources of both regions (DCs) and vice versa; shutdown DC1 and continue to use DC2 and vice versa; use (with light modification by adding third DC in form of VM).

Assuming that most common problem with DC is breaking the link between DCs in the cloud we consider the use of pacemaker and virtual IP to go to most appropriate DC at the concrete time. However, the use of Virtual IP (VIP) to enter the most fit at the time region (DC) must done with precautions. With loss of connectivity between regions and when DCs have the same weight (or importance of the DC expressed in numerical value) the cloud has been just destroyed. To prevent it the different weights were assigned to DCs. In this situation the DC with more weight will function in full scale. However second DC will have limited functionality – only use active virtual objects, but can't change state of any virtual object.

4. Near future testbed improvement

With number of regions more than two, the cloud configuration with Pacemaker and HAProxy might give more advantages. Pacemaker can make automation of Active/Passive cluster availability. HAProxy can be used in this case not only for load balancing, but also to ensure routing of packets.

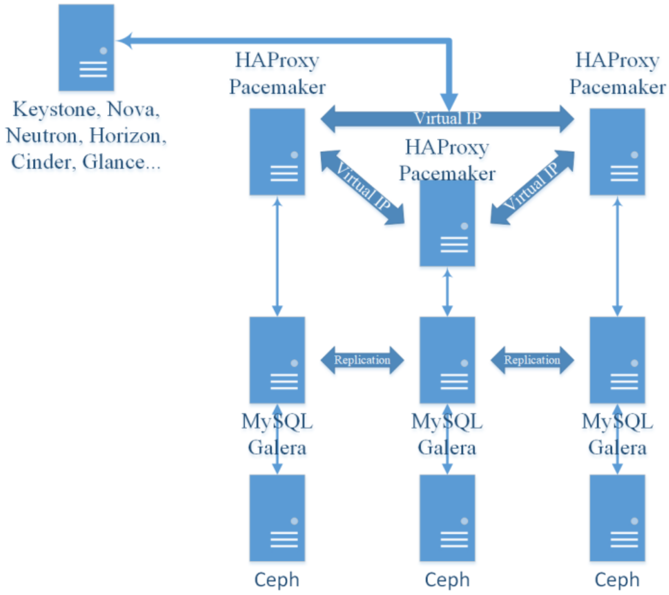

Figure 3. Future testbed configuration improvements

Figure 3 shows the plans of implementing HA feature in testbed. With HAProxy it will obtain balancing between deployed Openstacks in DCs of the cloud. Also an odd number of DCs is required for the full functionality of majority arbitration in case of equal weights of each DCs in the cloud.

5. Future project developments

As feature development the reliable data links between DCs are required. It is assumed that in between DCs more than one data link is to be used. The data links are planned to be configured with failover option with SDN approach.

To guarantee strong encryption of transferred data in this project is planned use Quantum Key Distribution [6]. The network side developments are running with extensive use of approach NFV.

6. Acknowledgement

This work was financially supported by the Ministry of Education and Science of Russian Federation (project No. 02.G25.31.0229).

7. Conclusion

It was described a range of tests on developed testbed when one DC is down to verify the reliability of computing services. It is shown that it is possible to increase cloud computing service reliability under specially configured Openstack in the cloud of geographically distributed data centers significantly.

References

[1] Justin Riley, John Noss, Wes Dillingham, James Cuff, Ignacio M. Llorente, "A High-Availability Cloud for Research Computing", Computer, vol. 50, issue 6 , pp. 92-95, 2017.

[2] M. Nabi, M. Toeroe, F. Khendek, "Availability in the Cloud", *J. Network and Computer Applications*, vol. 60, pp. 54-67, 2016.

[3] Sefraoui, O., Aissaoui, M., & Eleuldj, M. (2012). OpenStack: toward an open-source solution for cloud computing. *International Journal of Computer Applications*, *55*(3).

[4] Weil, S. A., Brandt, S. A., Miller, E. L., Long, D. D., & Maltzahn, C. (2006, November). Ceph: A scalable, high-performance distributed file system. In *Proceedings of the 7th symposium on Operating systems design and implementation* (pp. 307-320). USENIX Association.

[5] Almutairi, A., Sarfraz, M., Basalamah, S., Aref, W., & Ghafoor, A. (2012). A distributed access control architecture for cloud computing. *IEEE software*, *29*(2), 36-44.

[6] Chistyakov, V. V., Sadov, O. L., Vasiliev, A. B., Egorov, V. I., Kompaniets, M. V., Fedchenkov, P. V., ... & Khoruzhnikov, S. E. (2017). Software-defined subcarrier wave quantum networking operated by OpenFlow protocol. *arXiv preprint arXiv:1709.09081*.