Privacy Protected Documents On Openstack Cloud

Nithya A K
Assistant Professor
College of Engineering,Vadakara,Kerala,India
Email: ak.nithya@gmail.com

Dhannya A K
Junior Manager (Information Management)
Kerala State Civil Supplies Corporation
Email: dayathanal@gmail.com

Abstract—Cloud computing has become an extremely attractive area of research and practice over the last few years. However, there are many concerns over privacy of shared documents while adopting and using public cloud solutions. Private cloud solution is useful for organizations to implement their own privacy policies. Through this paper we propose a method to upload privacy preserving application on private cloud and later privacy protected documents can be publish in public cloud by using a hybrid cloud approach. This is a suitable solution for organizations those who want to share their valuable documents to outside world for data mining and research purpose. Applications running on virtual cloud environment are accessible to more users than a non cloud application. In this paper we discuss about various aspects in designing and setting up a private cloud infrastructure in academic as well as scientific environments by means of the open source software known as OpenStack and analyzing performance of applications on installed cloud infrastructure.

Keywords—*Cloud Computing, Privacy preservation, Private cloud, OpenStack*

I. INTRODUCTION

Cloud computing system usually offers services (DaaS, SaaS, IP-MaaS, PaaS, and so on) on the other side of the Internet in terms of its users[5].The secret information of individual users and business are stored and managed by the service providers, which consequently results in privacy concerns. Privacy issues exist for a long time in the computing literature, and many legal acts have been formulated and published to protect user's individual privacy as well as business secrets. Nevertheless, these acts are become outdated and inapplicable to the new scenarios, where an entirely different relationship has been evolved between users and providers [6].

Health research is central to the advancement of health care, which imperatively requires access to health data. However, health records are intrinsically identifiable, and control over the use of the data is necessary. When trying to publish health data, custodians of such data encounter inevitable hurdles in connection with privacy. To address such issues, a privacy rule was established in the United States which constitutes a comprehensive federal protection for individual's medical records and is known as Health Insurance Portability and Accountability Act (HIPAA). The legislation regulates health care groups, organizations or businesses, on how to use and disclose privacy critical data[4].

OpenStack is open source software designed to provision and manages large networks of virtual machines, creating a redundant and scalable private cloud computing environment [1].

The forthcoming sections of this paper discuss how to implement a privacy preserving solution using OpenStack to protect the private information, while publishing precious datasets to the outside world. For this, we propose an anonymization application as cloud computing service using OpenStack.

II. RELATED WORKS

Pierangela Samarati[7] addresses the problem of releasing micro data while safeguarding the anonymity of the respondents to which the data refer. The approach is based on the definition of k-anonymity. The paper introduces the concept of minimal generalization that captures the property of the release process not to distort the data more than needed to achieve k-anonymity, and present an algorithm for the computation of such a generalization.

Kabir et al. [8] presents a systematic clustering method based on k-anonymization technique to minimize the information loss while at the same time assuring data quality. Lin et al.[9] proposes a method to minimize the information loss due to anonymization,group similar data together and then anonymize each group individually.

L.Sweeney[10] propose k-anonymity protection model for protecting privacy. Zhang et al.[11]investigate and classify various privacy issues in cloud environment and point out potential key areas in cloud privacy protection and preservation.

Many anonymization applications already developed in non cloud applications. While using anonymizing application on cloud system, it is possible to share application as Software as a Service among users.

Authorized licensed use limited to: Kennesaw State University. Downloaded on October 13,2025 at 18:40:47 UTC from IEEE Xplore. Restrictions apply.

So here this paper discusses about private cloud implementation using OpenStack software, in which three node architecture is used for implementation. The rest of the paper is organized as follows; Section 3 describes the proposed method. The experimental results and discussion of them are presented in section 4. Then conclusion and future work are presented.

III. OPENSTACK

OpenStack is open source cloud computing software, released under the same terms as of apache license. It has the ability to control extensive processes, storage and resources all through the data centre, all supervised through a dashboard that gives the admin control while enabling the clients with procurement assets through a web interface. It is a cloud solution to provide an Infrastructure as a Service (IaaS). OpenStack is basically a collection of open source projects. Organizations can set up and run their cloud compute and storage infrastructure using this collection of open source projects.

OpenStack project is founded by NASA and Rackspace (RAX) and NASA provided a code for compute part and Rackspace provided hosting and storage infrastructure. OpenStack has become a non-profit organization entity which was established in September 2012.It manages and promotes distribution, development and adoption of the cloud computing software.

OpenStack is a global collaboration of developers and cloud computing technologists, producing a ubiquitous open source cloud computing platform for public and private clouds. The project aims to deliver solutions for all types of clouds by being simple to implement, massively scalable, feature rich.

A. Openstack Architecture

Open stack has a modular architecture which provides users the flexibility to design the cloud as required, without any proprietary issues related to hardware or software requirements. Further it has the ability to put together with existing as well as third party technologies.The architecture of the OpenStack is shown in fig 1.

B. Core Projects

OpenStack is defined as a collection of independently developed software projects, glued together. Each service has an HTTP REST interface to communicate with OpenStack clients.

- The identity service, all the services uses, since every single service has to validate the authentication token received from its clients. On the other hand, the Identity service also provides a catalog of REST endpoints, so every service has to register itself into the catalog to make itself known to OpenStack clients.

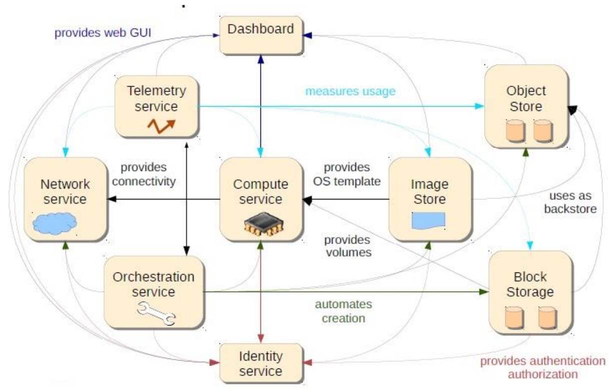


Fig. 1: OpenStack Architecture

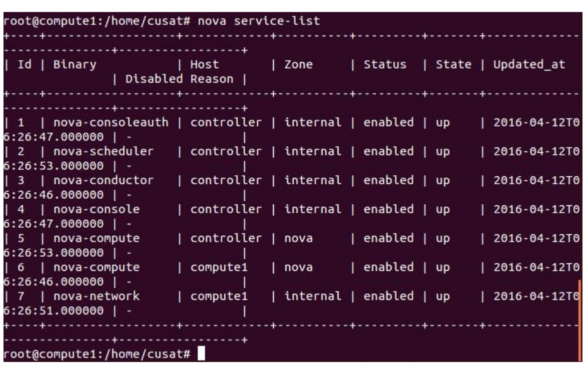
- The Dashboard,nicknamed Horizon, that is a Python/Django based Web GUI, with plug-ins for all other projects in OpenStack.
- Compute service (code name Nova) the heart of the whole OpenStack installation. This service is responsible for the creation of virtual machines on the top of a hypervisor (such as KVM, Xen, VMWare, etc.)
- The Image service (code name Glance) this is a simple storage for pre-installed OS images. Images can either be stored locally on the server where Glance is running, or uploaded to remote storage, such as the Object Store
- Block Storage service (code name Cinder) Besides OS images, we also need additional volumes to store data. These volumes are provided by Cinder. This service can also use the Object store to create backups of data volumes.
- Object Store (code name Swift) The Object Store is a petabyte scale storage built from up to hundreds of commodity servers and cheap disks. Files are up/downloaded via a simple REST/HTTP interface.
- The Network service (code name Neutron) it provides Software Defined Networking (SDN) functionality for OpenStack. With Neutron, users can build up complex network topologies to connect their VMs.
- The telemetry service (called Ceilometer) This service collects usage data for different resources, such a CPU cycles for virtual machines, storage space used out of the Block Storage. This telemetry data can be used to create bills for customers. It is also capable of raising alarms in case of an overuse of a service (such as running CPU hog applications).
- The orchestration service, nicknamed Heat It makes it easier to provision a larger set of dependent resources. With Heat you can download and launch templates (called stacks). A stack can start, for example, multiple running VMs provisioned with the same software stack. Using telemetry service alarms, it could also scale up/down a stack on demand.

Authorized licensed use limited to: Kennesaw State University. Downloaded on October 13,2025 at 18:40:47 UTC from IEEE Xplore. Restrictions apply.

IV. IMPLEMENTATION

An implementation of a prototype private cloud setup with open-source cloud operating environment on Linux platform. The nodes such as compute node and controller node, storage node are installed and configured [2][3] .

- Cloud Controller (Server1) Installation
  The API processes as well as the single-instance ones (MySQL, RabbitMQ, Keystone) run on the controller node. Beyond API processes, some services like Cinder, Glance, and Horizon have additional processes on the controller node.
- Compute Node (Server2) Installation
  This node runs a hypervisor to create virtual machines. The most important process that runs on the node is nova-compute responsible for
  - gathering all the necessary resources needed for a VM
  - creating the hypervisor related configuration for the VM
  - commanding the hypervisor to start the VM
- Storage Node (Server3) Installation
  Storage nodes usually provide distributed object store, or distribute block storage service.Installed OpenStack Block Storage service (cinder), which adds persistent storage to a virtual machine. Provides an infrastructure for managing volumes
  - Interacts with OpenStack Compute to provide volumes for instances.
  - Enables management of volume snapshots, and volume types.



```
root@compute1:/home/cusat# nova service-list
+----+------------------+------------+----------+---------+-------+------------
--------------+-----------------+
| Id | Binary           | Host       | Zone     | Status  | State | Updated_at
              | Disabled Reason |
+----+------------------+------------+----------+---------+-------+------------
--------------+-----------------+
| 1  | nova-consoleauth | controller | internal | enabled | up    | 2016-04-12T0
6:26:47.000000 | -               |
| 2  | nova-scheduler   | controller | internal | enabled | up    | 2016-04-12T0
6:26:53.000000 | -               |
| 3  | nova-conductor   | controller | internal | enabled | up    | 2016-04-12T0
6:26:46.000000 | -               |
| 4  | nova-console     | controller | internal | enabled | up    | 2016-04-12T0
6:26:47.000000 | -               |
| 5  | nova-compute     | controller | nova     | enabled | up    | 2016-04-12T0
6:26:53.000000 | -               |
| 6  | nova-compute     | compute1   | nova     | enabled | up    | 2016-04-12T0
6:26:46.000000 | -               |
| 7  | nova-network     | compute1   | internal | enabled | up    | 2016-04-12T0
6:26:51.000000 | -               |
+----+------------------+------------+----------+---------+-------+------------
--------------+-----------------+
root@compute1:/home/cusat#
```

Fig. 2: List of all nova services

In house development and deployment of the cloud involves in choosing hardware resources, configuration and deployment of the fundamental components OpenStack over the virtual infrastructure including compute, storage and network components. Necessary management services should be configured and customization of services through Application Programming Interface. The deployment should provide user a Graphical Interface to access, provision and automate cloud based resources. The OpenStack cloud is setup by using three node architecture, and creates a virtual

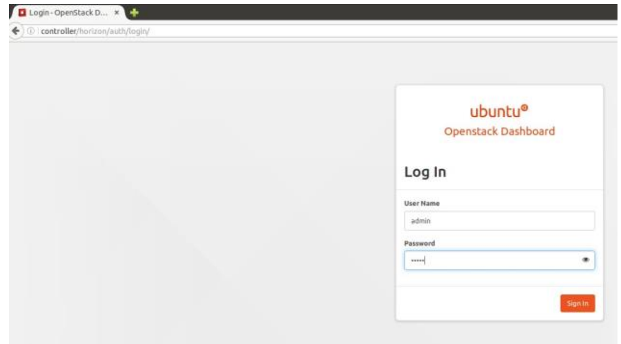


Fig. 3: Openstack login

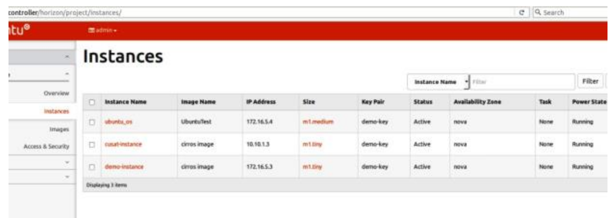


Fig. 4: Instances available in cloud

machine on it, which host anonymization application for privacy preservation.

V. CONCLUSION AND FUTURE WORK

In this experiment we created a private cloud infrastructure with minimum hardware requirements that can be used for small, mid sized and large scale organizations demanding high efficiency and security. This experimental setup of OpenStack found to be very much suitable for various communities, organizations, institutions, research centers to setup a cloud infrastructure in a cost effective way. Also implemented privacy preservation models for data management in cloud environment, which is highly beneficial to organizations, those who share cloud data with outside world.

The action plan for future aims to explore the scope of improving the performance of private cloud by adding more nodes to cloud infrastructure and to design and setup a hybrid cloud model for sharing of sensitive data with public cloud for publishing in outside world [12].

VI. ACKNOWLEDGMENT

The authors would like to thank Mr. Muralidharan K B, Department of Computer Science, Cochin University of Science and Technology for his guidance. This work is supported by the Technical Education Quality Improvement Program (TEQIP) fund for Basic Research.

Authorized licensed use limited to: Kennesaw State University. Downloaded on October 13,2025 at 18:40:47 UTC from IEEE Xplore. Restrictions apply.

REFERENCES

[1] Sefraoui, Omar, Mohammed Aissaoui, and Mohsine Eleuldj. "OpenStack: toward an open-source solution for cloud computing."International Journal of Computer Applications 55.3 (2012).

[2] Girish.L, Dr.H S GuruprasadBuilding a Private Cloud Using Open-Stack, International Journal of Emerging Trends and Technology In Computer Science Vol 3, Issue3, May-June 2014.

[3] Aruna P, L. Yamuna Devi, D. Sudha Devi, N. Priya, S. Vasantha, and K. Thilagavathy."Private Cloud for Organizations:An Implementation using OpenStack."International Journal of Scientific and Engineering Research,Volume4,Issue10,October-2013.

[4] Paul, Mithun, Christian Collberg, and Derek Bambauer. "A Possible Solution For Privacy-Preserving Cloud Data Storage" Cloud Engineering (IC2E), 2015 IEEE International Conference on. IEEE, 2015.

[5] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," National Institute of Standards and Technology, USA2009.

[6] Wang, Jian, et al. "Providing privacy preserving in cloud computing." 3rd International Conference on Human System Interaction. IEEE, 2010.

[7] Pierangela Samarati, Protecting Respondents Identities in Microdata Release ", IEEE Transactions On Knowledge And Data Engineering, VOL. 13,NO. 6, November/December, 2001.

[8] M.E. Kabir, H. Wang, E, Bertino, "Efficient systematic clustering method for k-anonymization ", Acta Informatica, Volume 48, Number 1,pp.51-66, 2011.

[9] Lin, Jun-Lin, and Meng-Cheng Wei. "An efficient clustering method for k-anonymization." Proceedings of the 2008 international workshop on Privacy and anonymity in information society. ACM, 2008.

[10] L.Sweeney, K-anonymity: a model for protecting privacy ", International Journal on Uncertainty Fuzziness and Knowledge-based Systems,10(5):557570(2002)

[11] Zhang, Gaofeng, et al. "Key research issues for privacy protection and preservation in cloud computing." Cloud and Green Computing (CGC), 2012 Second International Conference on. IEEE, 2012.

[12] Nepal, Surya, et al. "A secure storage service in the hybrid cloud." Utility and Cloud Computing (UCC), 2011 Fourth IEEE International Conference on. IEEE, 2011.

[13] Ken Pepple,Deploying OpenStack July 2011, First Edition, ISBN: 978-1-449-31105-6

[14] Open stack (website): http://www.openstack.org.

Authorized licensed use limited to: Kennesaw State University. Downloaded on October 13,2025 at 18:40:47 UTC from IEEE Xplore. Restrictions apply.