

Private Cloud Setup : A Survey

¹Rajalakshmi Vairamani, ²Sinduja Ravi,
³Sripriya Dhanraj, ⁴Sunesh Govindaraj,
^{1,2,3,4} UG Student

Department of Computer Science and Engineering,
Sri Ramakrishna Engineering College,
Coimbatore,INDIA

⁵Dr.M.Suresh Kumar
⁵ Professor

Department of Computer Science and Engineering,
Sri Ramakrishna Engineering College,
Coimbatore,INDIA

Abstract - Cloud computing is a phrase that gets a lot of play today, and for good reason. It represents a new paradigm for computing and the delivery of services, with the ability to scale from a single user to millions, almost instantly, while reducing costs and complexity. Private Cloud is a cloud infrastructure dedicated for only one customer. A private cloud can be located in the customer's own datacenter or computer room but it can also be hosted in a service provider's datacenter. A private cloud can be managed by the customer, by the hosting provider or by a third party. The customer however in need of private cloud software to maintain the cloud. This paper will explain the features of different types of clouds and private cloud setup tools.

Keywords: *private cloud, cloud computing, IaaS, PaaS, SaaS*

I INTRODUCTION

In the cloud computing business, there are different types of clouds: public and private, as well as hybrid clouds that utilize both. Cloud computing, at its simplest, is a service that is delivered to the user over a network (usually the Internet). The most basic type of cloud computing, while being the most flexible and granular in control, is Infrastructure as a Service (IaaS). It provides the basic, but critical, compute and storage services. This most closely mimics having your own physical machines, except it is a virtual machine along with many other virtual machines running on physical computers orchestrated by software. This makes the most efficient use of physical computing resources, by virtually slicing and dicing them based on customer needs and usage. This is superior to older methods of running one service on single piece of hardware, where average utilization rates were often below 10%. On the other hand, other services might be at or near capacity, and

can't expand for technical or organizational reasons.

A private cloud can offer the same services as a public cloud; however, its services are meant for in-house enterprise use only. [2] There are some main reasons why cloud computing has been on the rise : security, elasticity and scalability. This allows variable amounts of users on at the same time, or even enabled whenever it is necessary to alleviate bottlenecks. There is a downside to private clouds, their cost. [2]The capital required to build and operate a private cloud is substantially more than having a public cloud hosted. The need to purchase the hardware, as well as hiring an administrator to manage the private cloud, is the two main disadvantages.[3] In general, public clouds are highly structured and automated. Usually it's not possible for enterprises to get particular SLA's (Service Level Agreement) for their specific needs. Public clouds usually come with a standard SLA, so you just have the choice to take it or leave it.

The different cloud services are Software as a Service (SaaS), Platform as a service (PaaS) and Infrastructure as a Service (IaaS). [3]Software as a Service (SaaS)is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet.

Platform as a service (PaaS)[3] is a cloud computing model that delivers applications over the Internet. In a PaaS model, a cloud provider delivers hardware and software tools -- usually those needed for application development -- to its users as a service. A PaaS provider hosts the hardware and software on its own infrastructure.

Infrastructure as a Service (IaaS)[3,7] is a form of cloud computing that provides virtualized computing resources over the Internet. IaaS is one of three main categories of cloud computing

services, alongside Software as a Service (SaaS) and Platform as a Service (PaaS).

II OPEN SOURCE SOFTWARE FOR PRIVATE CLOUDS

Open source software for private clouds is widely available. OpenStack, Eucalyptus, Ganeti, OpenNebula are the main competitors in the private cloud area[4]. All of these except for Ganeti provide a clean user GUI, so we will just compare OpenStack, Eucalyptus and OpenNebula.

EUCALYPTUS

HPE Helion Eucalyptus is an open solution for building private and hybrid clouds compatible with Amazon Web Services (AWS) APIs[11]. It can dynamically scale up or down depending on application workloads and is well-suited for enterprise clouds. With HPE Helion Eucalyptus, organizations can deliver a public cloud experience using private IT resources to achieve better economics, stronger data governance, and faster application performance. The architecture of Eucalyptus is shown in figure 1.

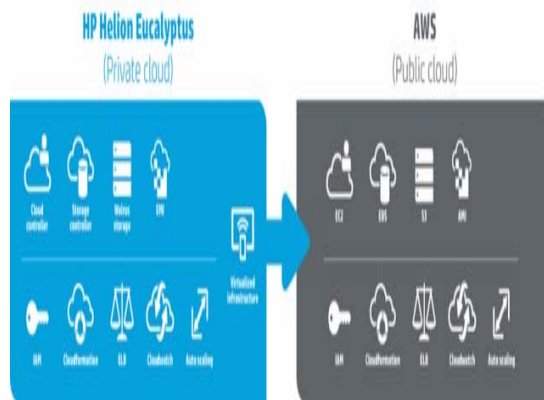


Fig 1: Eucalyptus Architecture

Each Eucalyptus service component exposes a well-defined language agnostic API in the form of a WSDL document containing both the operations that the service can perform and the input/output data structures. Inter-service authentication is handled via standard WS-Security mechanisms. The important service components of Eucalyptus are cloud controller, cluster controller, node controller, storage controller and walrus.

Cloud Controller (CLC) is the entry-point into the cloud for administrators, developers, project managers, and end-users[11]. The CLC is responsible for querying the node managers for information about resources, making high level scheduling decisions, and implementing them by making requests to cluster controllers. In essence, the CLC is responsible for exposing and managing the underlying virtualized resources (servers, network, and storage) via a well-defined industry standard API (Amazon EC2) well based user interface.

Cluster Controller (CC) generally executes on a cluster front-end machine or any machine that has network connectivity to both the nodes running NCs and to the machine running the CLC. CCs gather information about a set of VMs and schedules VM execution on specific NCs. The CC also manages the virtual instance network and participates in the enforcement of SLAs as directed by the CLC. All nodes served by a single CC must be in the same broadcast domain (Ethernet).

Node Controller (NC) is executed on every node that is designated for hosting VM instances. A UEC node is a VT-enabled server capable of running KVM as the hypervisor. UEC automatically installs KVM when the user chooses to install the UEC node. [11,12]The VMs running on the hypervisor and controlled by UEC are called instances. Eucalyptus supports other hypervisors like Xen apart from KVM, but Canonical has chosen KVM as the preferred hypervisor for UEC. The NC runs on each node and controls the life cycle of instances running on the node on one side and the CC on the other side. NC queries the operating system running on the node to discover the node's physical resources – the number of cores, the size of memory, and the available disk space. It also learns about the state of VM instances running on the node and propagates this data up to the CC.

Storage Controller (SC) implements block-accessed network storage (e.g., Amazon Elastic Block Storage -- EBS) and is capable of interfacing with various storage systems (NFS, iSCSI, etc.). An elastic block store is a Linux block device that can be attached to a virtual machine but sends disk traffic across the locally attached network to a

remote storage location.[12] An EBS volume cannot be shared across instances but does allow a snapshot to be created and stored in a central storage system such as Walrus, the Eucalyptus storage service.

Walrus (put/get storage) allows users to store persistent data, organized as eventually-consistent buckets and objects. It allows users to create, delete, list buckets, put, get, and delete objects, and set access control policies. Walrus is interface compatible with Amazon's S3, and supports the Amazon Machine Image (AMI) image-management interface, thus providing a mechanism for storing and accessing both the virtual machine images and user data. Using Walrus, users can store persistent data, which is organized as buckets and objects. WS3 is a file-level storage system, as compared to the block-level storage system of Storage Controller.

The Key Benefits of Eucalyptus are listed below:

- Build and manage self-service heterogeneous on-premise IaaS clouds using either existing infrastructure or dedicated compute, network and storage resources.
- Support high-availability IaaS for the most demanding cloud deployments.
- Gain precise control of private cloud resources via enterprise-ready user and group identity management along with resource quotas.
- Pool dynamic resources with built-in elasticity, allowing organizations to scale up and down virtual compute, network and storage resources.
- Integrate robust storage, enabling IT to easily connect and manage existing storage systems from within Eucalyptus clouds.
- Build hybrid clouds between on-premise Eucalyptus clouds and AWS and AWS-compatible public clouds.
- Run Eucalyptus or Amazon Machine Images as virtual cloud instances on Eucalyptus and AWS-compatible clouds.

- Leverage vibrant AWS ecosystem and management tools to manage Eucalyptus IaaS clouds.

OPEN NEBULA

OpenNebula is a cloud computing platform for managing heterogeneous distributed data center infrastructures. The OpenNebula platform manages a data center's virtual infrastructure to build private, public and hybrid implementations of infrastructure as a service and it is free and open-source software. OpenNebula combines existing virtualization technologies with advanced features for multi-tenancy, automated provisioning and elasticity. A built-in virtual network manager maps virtual networks to physical networks. OpenNebula is vendor neutral, as well as platform- and API-agnostic. It can use KVM, Xen or VMware hypervisors. OpenNebula architecture is shown in figure 2.

OpenNebula has some main components: Interfaces & API, Groups and Users, Networking, Storage, Hosts and Clusters.[4]The two main interfaces for interaction with OpenNebula are the Command Line Interface (CLI) and the Graphical User Interface (GUI). The Graphical User Interface is also known as "Sunstone". OpenNebula offers different APIs for Developers to extend the functionality or built on top of OpenNebula. These APIs are currently available as Amazon Web Services (AWS) APIs and OCCI Implementation.

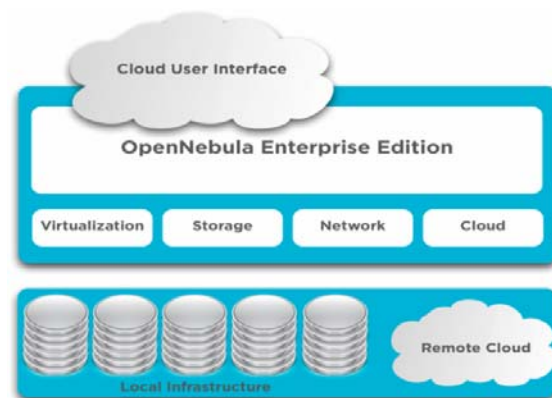


Fig 2: OpenNebula Architecture

OpenNebula allows different groups and users. It is also possible to integrate with different services such as LDAP and Microsoft's Active Directory. Multi-tenancy is possible by default, which eases

billing and accounting. OpenNebula comes with the following standard users: administrators, regular users, public users and service users. Administrators are in charge of administrative tasks within OpenNebula, regular users can use the functionality of OpenNebula in the self-service Portal. Public users are restricted users that may only use a subset of the functionality and service users are users that can use the APIs or Interfaces in OpenNebula.

The Networking interface in OpenNebula is fully extensible, which allows almost any integration in existing data centers. There is also support for VLAN and Open vSwitch. OpenNebula supports different storage systems such as the file system storage, distributed network storage or block storage.

OpenNebula supports the following hypervisors: Xen, KVM and VMware on the host. A host has three main components[4,9]:the host management, the cluster management and the host-monitoring component. The host management is implemented by the “onehost” and allows common operations on hosts such as initial setup or the machine lifecycle management. The cluster management allows placing a host in a specific cluster. This is implemented by the “onecluster” command. Host monitoring is done with the information driver (IM). Monitoring allows administrators to gather information about the health of a host.

Clusters are a pool of hosts that share networking and data stores[9]. A Cluster can be compared to a zone. Clusters are typically fulfilling different needs such as the production/testing differences. OpenNebula allows the grouping of different Hosts into a virtual data center (VDC) within a cluster. Different Hosts can also be grouped into zones that allow better administration for similar hosts.

The Key Features of OpenNebula are:

1. Automatic execution and scaling of multi-tiered applications
2. Automatic installation and configuration
3. Interfaces for Administrators and Advanced Users
4. Fine-grained ACLs for resource allocation
5. Resource Quota Management to track and limit computing, storage and networking resource utilization
6. Dynamic creation of Clusters as pools of hosts that share datastores and virtual networks for load balancing, high availability, and high performance computing.
7. High availability architecture
8. Persistent database backend with support for high availability configurations
9. Configurable behavior in the event of host or VM failure to provide easy to use and cost-effective failover solutions
10. fully platform independent

OPENSTACK

OpenStack is a global collaboration of developers and cloud computing technologists producing the ubiquitous open source cloud computing platform for public and private clouds.[5] “OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter, all managed through a dashboard that gives administrators control while empowering their users to provision resources through a web interface”. OpenStack has many features deploying and managing a private cloud simple. OpenStack is architected to provide flexibility as you design your cloud, with no proprietary hardware or software requirements and the ability to integrate with legacy systems and third party technologies[6,11]. We are able to run OpenStack on top of Ubuntu. OpenStack has the capabilities to make the transition from othercloud services with load balancing, which allows it to migrate the services without downtime to the end users. Openstack

Architecture is shown in figure 3.

Nova is the underlying cloud computing fabric controller for the OpenStack cloud. All activities needed to support the life cycle of instances within the OpenStack cloud are handled by Nova. It manages all the compute resources, networking, authorization, and scalability needs of the OpenStack cloud. Nova is a management platform and does not provide any virtualization capabilities

by itself instead, it uses libvirt APIs to interact with the supported hypervisors.

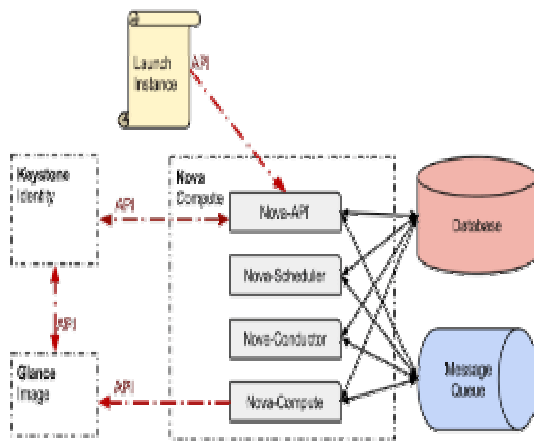


Fig 3: Openstack Architecture

Nova Cloud Fabric is composed of the following major components:

- **API Server (nova-api)**
API Server provides an interface to the outside world to interact with the cloud infrastructure. API server is the only component that outside world uses to manage the infrastructure.

The management is done through web services calls using EC2 API. API Server in turn communicates with the relevant components of the cloud infrastructure through the Message Queue.

- **Message Queue(rabbit-mq server)**

OpenStack Cloud Controller communicates with other nova components such as Scheduler, Network Controller, and Volume Controller via AMQP(Advanced Message Queue Protocol). Nova uses asynchronous calls for request response, with a call-back that gets triggered once a response is received. Since asynchronous communication is used, none of the end points get stuck for long in a waiting state. This is especially true since many actions expected by the API calls such as launching an instance or uploading an image are time consuming.

- **Compute Workers (nova-compute)**
Compute workers deal with instance management life cycle. they receive the requests for life cycle management via the Message Queue and carry out operations. There are several Compute Workers in a typical production cloud deployment. An instance is deployed on any of the available compute worker based on the scheduling algorithm used.
- **Network Controller (nova-network)**
Any valuable data that gets accumulated during the life cycle of an instance should be written to a volume, so that it can be accessed later. This typically applies to the storage needs of database servers, etc.
- **Compute Workers (nova-compute)**

The scheduler maps the nova-API calls to the appropriate openstack components. It runs as a daemon named nova-schedule and picks up a compute/network/volume server from a pool of available resources depending upon the scheduling algorithm in place[6]. A scheduler can base its decisions on various factors such as load, memory, physical distance of the availability zone, CPU architecture, etc. The nova scheduler implements a pluggable architecture.

Swift is an object store to store a large number of objects distributed across commodity hardware. Swift has built-in redundancy and failover management and features like backing up or archiving data, serving graphics or videos. It is scalable to multiple petabytes and to billions of objects. Swift provides elasticity and flexibility of cloud based storage for your web applications.

Functions and Features

- Storing the machine images
- Working as an independent data container
- Redundancy and Failover
- Backup and Archival
- Extremely scalable

OpenStack Imaging Service is a lookup and retrieval system for virtual machine images. Components of OpenStack Imaging Service are:

- Glance-api
- Glance-registryOpenStack

III PROPRIETARY SOFTWARE FOR PRIVATE CLOUDS

MICROSOFT AZURE

Microsoft Azure is a cloud computing platform and infrastructure created by Microsoft for building, deploying, and managing applications and services through a global network of Microsoft-managed datacenters. [1]It provides both PaaS and IaaS services and supports many different programming languages, tools and frameworks, including both Microsoft-specific and third-party software and systems.

The Windows Azure Pack integrates with System Center and Windows Server to provide a self-service portal for managing services such as websites, Virtual Machines, and Service Bus[10]. It serves for administrators to manage resource clouds and as a portal for multi-tenant services. Management Portal provides access to Cloud Service or hosted service and management tasks. It also provides at-a-glance status information about the overall health of your deployments and accounts.

Windows Azure provides four options – Virtual Machine, Web Sites, Cloud Service and Mobile Service; for creating and executing the applications. Each of them can be used separately or along with others. Data Service can be divided into two parts – Data Management and Business Analytics.

Data Management

Windows Azure provides three options – SQL Database, Tables and Blobs; for creating, querying and managing database for the applications. Each option addresses a different need; SQL Database for relational storage, Tables for fast access to potentially large amounts of simple typed data, and Blobs for unstructured binary storage.

Business Analytics

Windows Azure provides two options – SQL Reporting and HD Insight (Hadoop); for analyzing data and provides report. SQL Reporting is a subset of the reporting services included with SQL Server and allows you to build reporting into applications running on Windows Azure or on premises. The reports can be created in various formats like HTML, XML, PDF, Excel etc.

HDInsight is the name of the Windows Azure's Apache Hadoop-based service. HDInsight store data on the cluster and distribute it across multiple VMs. It also spreads the logic of a MapReduce job across those VMs. The architecture of Azure is shown in figure 4.

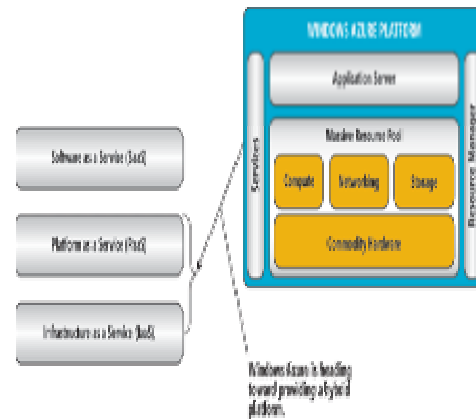


Fig 4: Azure Architecture

Windows Azure provides two options for networking – Virtual Network and Traffic Manager; for managing network traffic for the applications. Windows Azure Virtual Network is used to connect your own on-premises local network to a defined set of Windows Azure VMs. When your Windows Azure application is running in multiple data-centers, Windows Azure Traffic Manager is used to route requests from users intelligently across instances of the application.

App services can be divided into five parts – Identity, Performance, Messaging, Workflow Manager and Media Services.

Identity

Windows Azure Active Directory stores information about users and the organizations they belong to. It lets users log in and then supplies them with tokens they can present to applications to prove their identity. It also allows synchronizing user information with Windows Server Active Directory running on premises in your local network.

Performance

Windows Azure provides two options – in-memory Caching and Content Delivery Network (CDN); for caching frequent accessed data and improves your application performance. CDN is used to cache blob data that will be accessed fast by users around the world.

Messaging

Windows Azure provides two options – Queues and Service Bus; for handling complex interactions between two applications.

In Queue approach, one application places a message in a queue, and that message is eventually read by another application. [10]If your application needs just this straightforward service, Windows Azure Queues might be the best choice. Service Bus allows your applications to exchange data anywhere; whether your application is running in the cloud, in your data center, on a mobile device, or somewhere else. Moreover, Service Bus provides both queues (one-to-one) and publish-and-subscribe (one-to-many) mechanisms.

Work Flow Manager

It provides the capability to host workflows in a high scale, high density, and multi-tenant environment. Workflows is supported by using Microsoft .NET Framework 4.5 and Windows Workflow Foundation (WF45)

Media Services

Windows Azure media service process and serves the various media like video very fast to the client around the world.[10] Video created by using Media Services can be consumed by various client systems, including Windows, Macintosh, HTML 5, iOS, Android, Windows Phone, Flash, and Silverlight.

Windows Azure Marketplace and Windows Azure Store allows you to find and buy Windows Azure applications and commercial datasets. The difference between these two is that Marketplace is outside of the Windows Azure Management Portal, but the Store can be accessed from the portal.

Microsoft currently provides language-specific SDKs for .NET, Java, PHP, Node.js, Ruby, and Python. There's also a general Windows Azure SDK that supports any language such as C++.

The Windows Azure SDK for .NET includes the Windows Azure Tools for Microsoft Visual Studio, which extends Visual Studio to enable the creation, building, packaging, running, and debugging of scalable web applications and services on Windows Azure.

IV CONCLUSIONS

Focused on aspects such as architecture, characteristics, application and so on, a detailed comparison has been presented in this paper we took a closer look into three common Infrastructure as a Service (IaaS) framework implementations; Eucalyptus, OpenStack, OpenNebula and Microsoft Azure. For each framework, we illustrated its architecture and design as well as provided a feature comparison. OpenStack is the best choice for public cloud providers due to its scalability. Eucalyptus is a possible choice for private cloud providers and researchers due to its extensibility and quick installation. OpenNebula focuses tremendously on Data Center virtualization, and has fewer options for controlling what we want from a private cloud. Azure is the best choice for public as well as the private clouds due to its massive success in PaaS. This helps users to better understand and choose the appropriate cloud platforms based on service type, compatibility, deployment models, etc.

REFERENCES

- 1) Thilagavathi M* School of Information Technology and Engineering, VIT University, India, "CLOUD PLATFORMS – A COMPARISON", 2003
- 2) Private Cloud Development and Architecture
<https://www.simpletalk.com/cloud/develo>

- pment/private-cloud,-what-is-it-and-why-do-you-need-it/
- 3) Diversity Ltd “Understanding the cloud computing stack SaaS,PaaS,IaaS”
 - 4) OpenNebula Cloud Architecture <http://opennebula.org/new-open-cloud-reference-architecture/>
 - 5) Girish L S , Dr. H S Guruprasad “BUILDING PRIVATE CLOUD USING OPENSTACK”, June 2014 International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) ISSN 2278-6856
 - 6) Jilin Zhang, Jianhui Zhang, Hong Ding, Jian Wan, Yongjian Ren and Jue Wang “DESIGNING AND APPLYING AN EDUCATION IAAS SYSTEM BASED ON OPENSTACK” , 1 Feb 2013 An International Journal Natural Sciences Publishing Cor.
 - 7) Sushil Bhardwaj , Leena Jain , Sandeep Jain “CLOUD COMPUTING: A STUDY OF INFRASTRUCTURE AS A SERVICE (IAAS) “;Copyright© 2010 waves publishers ISSN 0975-5292 International Journal of Engineering and Information Technology Vol 2 , No. 1.
 - 8) Cloud Computing and Rackspace Cloud Files, <http://www.rackspacecloud.com>.
 - 9) Opennebula Components <http://cloudvane.com/2013/06/25/opennebula-components/>
 - 10) Microsoft Azure <https://azure.microsoft.com/en-in/documentation/articles/fundamentals-introduction-to-azure/>
 - 11) Openstack Setup https://community.spiceworks.com/how_to/72806-launch-your-private-cloud-with-openstack
 - 12) Eucalyptus Architecture <http://www.cloudbook.net/resources/stories/the-eucalyptus-open-source-private-cloud>
 - 13) Eucalyptus, <http://www8.hp.com/in/en/cloud/helion-eucalyptus.html>