Volume 2, No. 3, May-June 2011

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

CASE STUDY AND REPORT

Available Online at www.ijarcs.info

A Modern Architecture for Open Source Cloud Computing

S. Ramani
School of Computing Science and Engineering
VIT University,India
ramani.s@vit.ac.in

Abstract: With the advent of Web 2.0 and Software as a Service, cloud computing has come into vogue. Cloud computing has become synonymous with providing services anywhere anytime with the basic requirement being access to the internet. As a new model, cloud computing promises to make any online service available without a large upfront investment in infrastructure. The economics of running a full infrastructure changes dramatically since you only pay for what you use. And you can provision capacity for only what you require at any given time. Many companies who have supported their own needs for large-scale, distributed computing are beginning to export their technologies and services to support others. Open source is important in all aspects of cloud computing. It is used to build the core of the cloud and its services. In this paper we proposed architecture for open source service cloud computing which will avoid vendor lock-in and scalability. We detail the functional architecture and the core structure implementing such paradigm. The model divides the cloud computing system with various components in open source 3-layer hierarchy called infrastructure, platform and application**.**

Keywords: cloud computing, open source, vendor lock-in, scalability

I. INTRODUCTION

Everyone has an opinion on what is cloud computing. It can be the ability to rent a server or a thousand servers and run a geophysical modeling application on the most powerful systems available anywhere. It can be the ability to rent a virtual server, load software on it, turn it on and off at will, or clone it ten times to meet a sudden workload demand. It can be storing and securing immense amounts of data that is accessible only by authorized applications and users. It can be supported by a cloud provider that sets up a platform that includes the OS, Apache, MySQL™ database, Perl, Python, and PHP with the ability to scale automatically in response to changing workloads.

Cloud computing can be the ability to use applications on the Internet that store and protect data while providing a service anything including email, sales force automation and tax preparation. It can be using a storage cloud to hold application, business, and personal data. And it can be the ability to use a handful of Web services to integrate photos, maps, and GPS information to create a mashup in customer Web browsers.

While the first revolution of the Internet saw the three-tier (or n-tier) model emerge as a general architecture, the use of virtualization in clouds has created a new set of layers: applications, services, and infrastructure. These layers don't just encapsulate on-demand resources; they also define a new application development model. And within each layer of abstraction there are myriad business opportunities for defining services that can be offered on a pay-per-use basis.

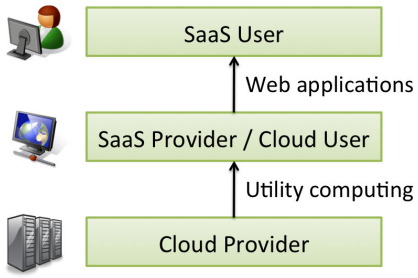

Figure 1: Users and Providers of Cloud Computing.

Cloud computing brings a new level of efficiency and economy to delivering IT resources on demand and in the process it opens up new business models and market opportunities.

II. RELATED WORK

Cloud computing systems fundamentally provide access to large pools of data and computational resources through a variety of interfaces similar in spirit to existing grid and HPC resource management and programming systems.

EUCALYPTUS – an open source software framework for cloud computing was proposed that implements what is commonly referred to as Infrastructure as a Service (IaaS); systems that give users the ability to run and control entire virtual machine instances deployed across a variety physical resources [5].A reference model divides the cloud computing system with various components in a 3-layer hierarchy called infrastructure, platform and application. The details of the components are presented for its functionality assumed. In the model, the cloud system is clarified with a 3-layer hierarchy as follow:

- Cloud computing infrastructure layer provides a cluster of hardware resource such as CPU, memory, bandwidth and storage.

- The platform layer includes the components such as kernel, distributed file system; cloud IO, computing driver/engine, management and UI interface.
- The application layer host business domain specific application.

The model is showed in the following diagram:

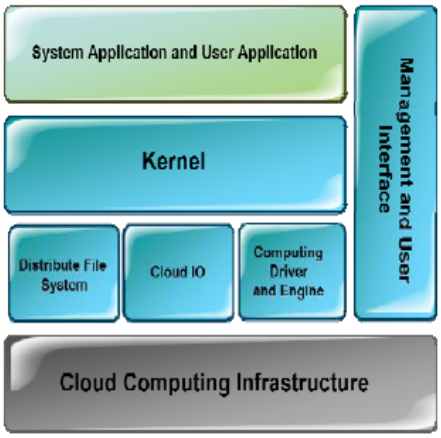

Figure 2: Reference model

Mapping to the reference model proposed, some components like kernel, distributed file system already have many implementations with commercial or open source software. For application developers, they don't need to care much detail of the service such as task management, memory, storage, communication any more. Contrarily, they only need to focus on application design including the application specific IO, distribute task design and domain use cases implementation. With the model proposed in this paper, the application developers can reuse the existed experience from traditional computing OS such as UNIX/LINUX and get a faster learning curve as the cloud utility users – which also make the communication more effective with the same context.

Also the open source software implementation for the components in the model is addressed[17].RESERVOIR project [3] is developing an advanced system and service management approach that will serve as the infrastructure for Cloud Computing and Communications and Future Internet of Services by creative coupling of service virtualization, grid computing, networking and service management techniques.

Globus, a system that they are developing to address these challenges. The Globus system is intended to achieve a vertically integrated treatment of application, middleware, and network. A low-level toolkit provides basic mechanisms such as communication, authentication, network information, and data access[17].

III.BUILDING A PRIVATE CLOUD

We can build a private cloud using UEC.
Ubuntu Enterprise Cloud (UEC) is a private cloud that embeds Eucalyptus cloud on Ubuntu server. The current release of UEC runs on Ubuntu 9.04 Server running Eucalyptus 1.5.

UEC is made up of three components: Cloud Controller (eucalyptus-cloud), Cluster Controller (eucalyptus-cc), and one or more Node Controllers (eucalyptus-nc). The Cloud Controller is the Web-services interface and the WEBUI server. The Cloud Controller also provides resource scheduling and S3 and EBS computable storage interfaces.

A cluster in UEC is synonymous with an availability zone in AWS. In this release of UEC the Cluster Controller has to run on the same machine as the Cloud Controller. The Cluster Controller provides network control for the defined cluster and manages resources within the cluster (i.e., resources on the node). The Cloud Controller and the Cluster Controller are sometimes referred to as the Front End. Typically the Node Controller runs on a separate box from the Front End box.

In a production environment there will be multiple Node Controllers making up a larger cluster (i.e., your cloud). Each Node Controller runs as a KVM hypervisor and all the Node Controllers in the cluster make up the cloud environment. In the current release, running multiple clusters is really not supported. In future releases of UEC, you will be able to run multiple clusters in one environment. Each cluster acts like an availability zone in the UEC environment.

IV. OPEN SOURCE ARCHITECTURE

Open source is an approach to the design, development, and distribution of software, offering practical accessibility to software's source code. Some consider open source as one of various possible design approaches, while others consider it a critical strategic element of their operations.

Before open source became widely adopted, developers and producers used a variety of phrases to describe the concept; the term open source gained popularity with the rise of the Internet, which provided access to diverse production models, communication paths, and interactive communities.
These are the key dimensions for Open source architecture.

- Licenses that permit and encourage redistribution, modification, and even forking
- An architecture that enables programs to be used as components where-ever possible, and extended rather than replaced to provide new functionality
- Low barriers for new users to try the software
- Low barriers for developers to build new applications and share them with the world

The architecture for open source architecture is shown in the figure.

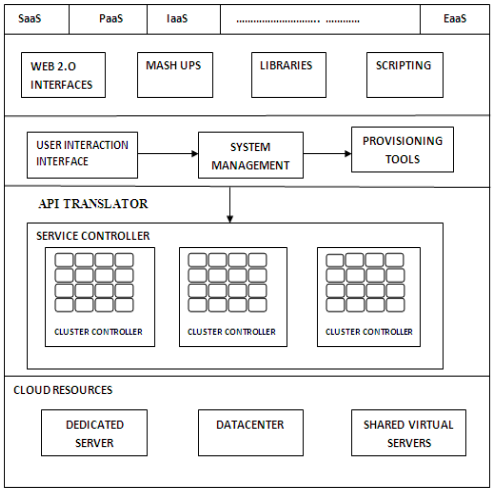

Figure 3: Open source Cloud Architecture

Virtualization and service integration plays a wide role here .The service integration perform the decision control and allows all services to the user for performing desire operations.

A. Virtualization

The first objective is to provide quality of service to virtual machines that is consistent with what we have been able to provide in manually configured, single-application environments. The second objective is to provide persistence of connectivity so that when a virtual machine moves between compute nodes, the connections and resources within the network will automatically follow the virtual machine to its new compute location. And the third objective is to provide tools for monitoring the performance levels being delivered and reporting back to end users so they can be sure they're actually receiving what they're paying for.

Virtualization technologies are transforming the way organizations deliver data to employees and customers, enabling a dynamic and agile business strategy while optimizing IT resources and reducing costs. Organizations can extend virtualization benefits by implementing a more holistic data center strategy, a process that begins and ends with creating dynamic pools of IT resources—servers, storage, applications, networks, and the interconnect layers—that can be allocated automatically or on demand as business requirements dictate.

The average disk bandwidth for 75 EC2 instances each writing 1 GB files to local disk. The mean disk write bandwidth is nearly 55 MBytes per second with a standard deviation of a little over 9 MBytes/sec, more than 16% of the mean. This demonstrates the problem of I/O interference between virtual machines.

B. Service Controller

Service controller controls the services and other control. It directed the users to access the required service. It has included the following controllers.

- Cluster Controller gathers information about and Schedules VMexecution on specific node controllers,as well as manages virtual instance network.
- Storage Controller (Walrus) is a put/get storage service that implements Amazon's S3 interface, providing a mechanism for storing and accessing virtual machine images and user data.
- Cloud Controller is the entry-point into the cloud for users and administrators. It queries node managers for information about resources, makes high level scheduling decisions, and implements them by making requests to cluster controllers.

V.SERVICE INTEGRATION

The future of cloud computing will be permeated with the notion of brokers negotiating relationships between providers of cloud services and the service customers. In this context, a broker might be software, appliances, platforms or suites of technologies that enhance the base services available through the cloud. Enhancement will include managing access to these services, providing greater security or even creating completely new services.

There will be three types of CSBs:

1. Cloud service intermediation brokers. Companies such as AT&T, Verizon, Telestra or Virgin Media will sell you services, with add-ons such as identity and access management tools. You may buy directly from these companies, or you may get their services as part of another SaaS offering.
2. Aggregation brokers. As the name implies, these companies will combine multiple services into one or more new services. They'll handle the data integration and other details for end users.
3. Cloud service arbitrages. These will be similar to aggregation brokers, but the services won't be fixed, so users will have more options and flexibility.

VI.TRENDS AND FEATURES

While many people think of current cloud computing offerings as purely "pay by the drink" compute platforms, they're really a convergence of two major interdependent IT trends:

IT Efficiency - Minimize costs where companies are converting their IT costs from capital expenses to operating expenses through technologies such as virtualization. Cloud computing begins as a way to improve infrastructure resource deployment and utilization, but fully exploiting this infrastructure eventually leads to a new application development model.

Business Agility-Maximize return using IT as a competitive weapon through rapid time to market, integrated application stacks, instant machine image deployment, and petascale parallel programming. Cloud computing is embraced as a critical way to revolutionize time to service. But inevitably these services must be built on equally innovative rapid-eployment-infrastructure models.

To be sure, these trends have existed in the IT industry for years. However, the recent emergence of massive network bandwidth and virtualization technologies has enabled this transformation to a new services-oriented infrastructure.

Cloud Computing has the following features.

Reduce capital expenditures - Cloud computing makes it possible for companies to convert IT costs from capital expense to operating expense through technologies such as virtualization.

Cut the cost of running a datacenter - Cloud computing improves infrastructure utilization rates and streamlines resource management. For example, clouds allow for self-service provisioning through APIs, bringing a higher level of automation to the datacenter and reducing management costs.

V. CONCLUSION

Cloud Computing holds a lot of promise and we believe that it is likely to be a major influence on hosting and application development. Cloud computing is the next big wave in computing. It has many benefits, such as better hardware management, since all the computers are the same and run the same hardware. It also provides for better and easier management of data security, since all the data is located on a central server, so administrators can control who has and doesn't have access to the files.

VI.REFERENCES

[1] Ali Arsanjani, Liang-Jie Zhang, Michael Ellis, Abdul Allam, Kishore Channabasavaiah, "S3: A Service-Oriented Reference Architecture," IT Professional, vol. 9, no. 3, pp. 10-17, May/Jun, 2007

[2] The Future of Cloud Integration http://www.itbusinessedge.com/cm/blogs/lawson/the-future-of-cloud-integration-what-model-will-win/?cs=34121

[3] B.Rochwerger, A.Galis, E.Levy, J. A. Cáceres, D.Breitgand, Y.Wolfsthal, I.M.Llorente, M.Wusthoff, R.S.Montero, E.Elmroth "RESERVOIR: ManagementTechnologies and Requirements for Next Generation Service Oriented Infrastructures" 2009 IFIP/IEEE International Symposium on Integrated Network Management (IM 2009).

[4] Enomalism elastic computing infrastructure. http://www.enomaly.com.

[5] Eucalyptus Public Cloud (EPC) http://eucalyptus.cs.ucsb.edu/wiki/EucalyptusPublicCloud/

[6] G. J. Popek and R. P. Goldberg, "Formal requirements for virtualizable third generation architectures" - Communications of ACM, vol. 17, no.7, pp. 412– 421, 1974

[7] Hadoop Open Source Project by apache, http://hadoop.apache.org/core/, 2009

[8] ISO/IEC 10040, 1998, "Information technology - Open Systems Interconnection - Systems management overview"

[9] L. Erlanger, "Virtualization: Doing More with Less", Inforworld Report,Sept 06, www.infoworld.com/article/06/09/11/37FEvirtcaseserv_1.html

[10] Liang-Jie Zhang; Qun Zhou; "CCOA: Cloud Computing Open Architecture", WebPage(s):607 – 616, 6-10 July 2009

[11] Mathias Dalheimer, Franz-Josef Pfreundt," GenLM: License Management for Grid and Cloud Computing Environments", 9th IEEE/ACM International Symposium on Cluster Computing and the Grid

[12] Nurmi, D.; Wolski, R.; Grzegorczyk, C.; Obertelli, G.; Soman, S.; Youseff, L.; Zagorodnov, D.;" The Eucalyptus Open-Source Cloud-Computing System", Cluster Computing and the Grid, 2009. CCGRID '09. 9th IEEE/ACM International Symposium Page(s):124 – 131, 18-21 May 2009

[13] OSI Model, http://www.osi.org, 2009

[14] P. Wallis, Keystones and Rivets, "Understanding Cloud Computing", Feb08 www.keystonesandrivets.com/kar/2008/02/cloud-computing.html

[15] Red Hat Global File System, http://www.redhat.com/whitepapers/rha/gfs/GFS_INS0032US.pdf.

[16] Reservoir project www.reservoir-fp7.eu

[17] Wenke Ji, Nanjing, P.R.C, Jiangbo Ma, Xiaoyong Ji," A Reference Model of Cloud Operating and Open Source Software Implementation Mapping", 18th IEEE International Workshops on EnablingTechnologies: Infrastructures for Collaborative Enterprises,2009