

Cloud Computing – Research Issues, Challenges, Architecture, Platforms and Applications: A Survey

Santosh Kumar and R. H. Goudar

Abstract—Cloud computing is the development of parallel computing, distributed computing, grid computing and virtualization technologies which define the shape of a new era. Cloud computing is an emerging model of business computing. In this paper, we explore the concept of cloud architecture and compares cloud computing with grid computing. We also address the characteristics and applications of several popular cloud computing platforms. In this paper, we aim to pinpoint the challenges and issues of cloud computing. We identified several challenges from the cloud computing adoption perspective and we also highlighted the cloud interoperability issue that deserves substantial further research and development. However, security and privacy issues present a strong barrier for users to adapt into cloud computing systems. In this paper, we investigate several cloud computing system providers about their concerns on security and privacy issues.

Index Terms— Cloud computing, architecture, challenges, cloud platforms, research issues.

I. INTRODUCTION

Cloud computing is a complete new technology. It is the development of parallel computing, distributed computing grid computing, and is the combination and evolution of Virtualization, Utility computing, Software-as-a-Service (SaaS), Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS). Cloud is a metaphor to describe web as a space where computing has been pre installed and exist as a service; data, operating systems, applications, storage and processing power exist on the web ready to be shared. To users, cloud computing is a Pay-per-Use-On-Demand mode that can conveniently access shared IT resources through the Internet. Where the IT resources include network, server, storage, application, service and so on and they can be deployed with much quick and easy manner and least management and also interactions with service providers. Cloud computing can much improve the availability of IT resources and owns many advantages over other computing techniques. Users can use the IT infrastructure with Pay-per-Use-On-Demand mode; this would benefit and save the cost to buy the physical resources that may be vacant.

Organization. The rest of the paper is organized as follows: In Section II, we define architectural components such as Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS) and Data as a Service (DaaS). Then, we compare cloud and grid computing

in Section III and explain some popular cloud computing platforms in Section IV. In Section V, we include few applications of cloud computing. We further explained about issues and challenges of cloud computing in Section VI, VII and VIII. Finally, we conclude in Section IX.

II. ARCHITECTURAL COMPONENTS

Cloud service models are commonly divided into SaaS, PaaS, and IaaS that exhibited by a given cloud infrastructure. It's helpful to add more structure to the service model stacks: Fig. 1 shows a cloud reference architecture [13] that makes the most important security-relevant cloud components explicit and provides an abstract overview of cloud computing for security issue analysis.

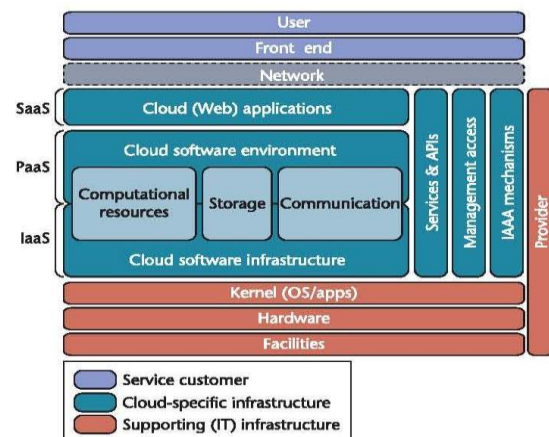


Fig. 1. The cloud reference architecture.

A. Software as a Service (SaaS)

Cloud consumers release their applications in a hosting environment, which can be accessed through networks from various clients (e.g. Web browser, PDA, etc.) by application users. Cloud consumers do not have control over the cloud infrastructure that often employs multi-tenancy system architecture, namely, different cloud consumers' applications are organized in a single logical environment in the SaaS cloud to achieve economies of scale and optimization in terms of speed, security, availability, disaster recovery and maintenance. Examples of SaaS include Salesforce.com, Google Mail, Google Docs, and so forth.

B. Platform as a Service (PaaS)

PaaS is a development platform supporting the full "Software Lifecycle" which allows cloud consumers to develop cloud services and applications (e.g. SaaS) directly on the PaaS cloud. Hence, the difference between SaaS and PaaS is that SaaS only hosts completed cloud applications whereas PaaS offers a development platform that hosts both

completed and in-progress cloud applications. This requires PaaS, in addition to supporting application hosting environment, to possess development infrastructure including programming environment, tools, configuration management, and so forth. An example of PaaS is Google AppEngine.

C. Infrastructure as a Service (IaaS)

Cloud consumers directly use IT infrastructures (processing, storage, networks and other fundamental computing resources) provided in the IaaS cloud. Virtualization is extensively used in IaaS cloud in order to integrate/decompose physical resources in an ad-hoc manner to meet growing or shrinking resource demand from cloud consumers. The basic strategy of virtualization is to set up independent virtual machines (VM) that are isolated from both the underlying hardware and other VMs. Notice that this strategy is different from the multi-tenancy model, which aims to transform the application software architecture so that multiple instances (from multiple cloud consumers) can run on a single application (i.e. the same logic machine). An example of IaaS is Amazon's EC2.

D. Data as a Service (DaaS)

The delivery of virtualized storage on demand becomes a separate Cloud service - data storage service. Notice that DaaS could be seen as a special type IaaS. The motivation is that on-premise enterprise database systems are often tied in a prohibitive upfront cost in dedicated server, software license, post-delivery services and in-house IT maintenance. DaaS allows consumers to pay for what they are actually using rather than the site license for the entire database. In addition to traditional storage interfaces such as RDBMS and file systems, some DaaS offerings provide table-style abstractions that are designed to scale out to store and retrieve a huge amount of data within a very compressed timeframe, often too large, too expensive or too slow for most commercial RDBMS to cope with. Examples of this kind of DaaS include Amazon S3, Google BigTable, and Apache HBase, etc.

III. COMPARISON BETWEEN CLOUD AND GRID COMPUTING

A comparison [6] can be summaries as follows:

- 1) Construction of the grid is to complete a specified task, such as biology grid, Geography grid, national educational grid, while Cloud computing is designed to meet general application and there are not grid for a special field.
- 2) Grid emphasizes the "resource sharing" to form a virtual organization. Cloud is often owned by a single physical organization (except the community Cloud, in this case, it is owned by the community), who allocates resources to different running instances.
- 3) Grid aims to provide the maximum computing capacity for a huge task through resource sharing. Cloud aims to suffice as many small-to-medium tasks as possible based on users' real-time requirements. Therefore, multi-tenancy is a very important concept for Cloud computing.

- 4) Grid trades re-usability for (scientific) high performance computing. Cloud computing is directly pulled by immediate user needs driven by various business requirements.
- 5) Grid strives to achieve maximum computing. Cloud is after on-demand computing – Scale up and down, in and out at the same time optimizing the overall computing capacity.

IV. POPULAR CLOUD COMPUTING PLATFORMS

A. AbiCloud

Abicloud [5] is a cloud computing platform, It can be used to build, integrate and manage public as well as private cloud in the homogeneous environments. Using Abicloud, user can easily and automatically deploy and manage the server, storage system, network, virtual devices and applications and so on. The main difference between Abicloud and other cloud computing platforms is its powerful web-based management function and its core encapsulation manner. Using the Abicloud, user can finish deploying a new service by just dragging a virtual machine with mouse. This is much easier and flexible than other cloud computing platforms that deploy new services through command lines.

Abicloud can be used to deploy and implement private cloud as well as hybrid cloud according to the cloud providers' request and configuration. It can also manage EC2 according to the rules of protocol. Besides, apply the Abicloud, a whole cloud platform based on Abicloud can be packed and redeployed at any other Abicloud platform. This is much helpful for the transformation of the working environment and will make the cloud deployment process much easier and flexible.

B. Eucalyptus

Eucalyptus (Elastic Utility Computing Architecture for Linking Your Programs to Useful Systems) [5] mainly was used to build open-source private cloud platform. Eucalyptus is an elastic computing structure that can be used to connect the users' programs to the useful systems, it is an open-source infrastructure using clusters or workstation implementation of elastic, utility, cloud computing and a popular computing standard based on a service level protocol that permit users lease network for computing capability.

Currently, Eucalyptus is compatible with EC2 from Amazon, and may support more other kinds of clients with minimum modification and extension.

C. Nimbus

Nimbus [5] is an open tool set and also a cloud computing solution providing IaaS. It permits users lease remote resources and build the required computing environment through the deployment of virtual machines.

Generally, all these functional components can be classified as three kinds. One kind is client- supported modules which are used to support all kinds of cloud clients. Context client module, cloud client module, reference client module and EC2 client module are all belonging to this kind of component. The second kind of component is mainly

service-supported modules of cloud platform, providing all kinds of cloud services. It includes a context agent module, web service resource framework module, EC2 WSDL module and a remote interface module. The third kind of component is the background resource management modules which are mainly used to manage all kinds of physical resources on the cloud computing platform, including work service management module, IaaS gateway module, EC2 and other cloud platform support module, workspace pilot module, workspace resource management module and workspace controller.

D. OpenNebula

OpenNebula [5] is also an open source cloud service framework. It allows user deploy and manage virtual

machines on physical resources and it can set user's data centers or clusters to flexible virtual infrastructure that can automatically adapt to the change of the service load. The main difference of OpenNebula and nimbus is that nimbus implements remote interface based on EC2 or WSRF through which user can process all security related issues, while OpenNebula does not. OpenNebula is also an open and flexible virtual infrastructure management tool, which can use to synchronize the storage, network and virtual techniques and let users dynamically deploy services on the distributed infrastructure according to the allocation strategies for data center and remote cloud resources. Through the interior interfaces and OpenNebula data center environment, users can easily deploy any types of clouds.

TABLE I: THE COMPARISON OF SERVER CLOUD COMPUTING PLATFORMS [5]

	Abicloud	Eucalyptus	Nimbus	OpenNebula
Cloud Character	Public/private	Public	Public	Private
Scalability	Scalable	Scalable	Scalable	Dynamic, Scalable
Clouds form	IaaS	IaaS	IaaS	IaaS
Compatibility	Not support EC2	Support EC2, S3	Support EC2	Open, multi-platform
Deployment	Pack and redeploy	Dynamical deployment	Dynamical deployment	Dynamical deployment
Deployment Manner	Web interface drags	Command line	Command line	Command line
Transplant-ability	Easy	Common	Common	Common
VM support	Virtual Box, Xen, VMware, VM	Xen, VMware, KVM	Xen	Xen, VMware
Web interface	Libvirt	Web service	EC2, WSDL, WSRF	libvirt, OCCI, EC2, API
Structure	Open platform encapsulates core	Module	Lightweight components	Module
Reliability	-	-	-	Rollback host and VM
OS support	Linux	Linux	Linux	Linux
Development language	Ruby, c++, python	Java	Java, python	Java

V. APPLICATIONS

There are a few applications of cloud computing [4] as follows:

- 1) Cloud computing provides dependable and secure data storage center.
- 2) Cloud computing can realize data sharing between different equipments.
- 3) The cloud provides nearly infinite possibility for users to use the internet.
- 4) Cloud computing does not need high quality equipment for the user and it is easy to use.

VI. ISSUES IN CLOUD COMPUTING

More and more information on individuals and companies is placed in the cloud; concerns are beginning to grow about just how safe an environment it is? Issues of cloud computing [3] can summarize as follows:

A. Privacy

Cloud computing utilizes the virtual computing technology, users' personal data may be scattered in various virtual data centers rather than stay in the same physical

location, users may leak hidden information when they are accessed cloud computing services. Attackers can analyze the critical task depend on the computing task submitted by the users.

B. Reliability

The cloud servers also experience downtimes and slowdowns as our local server.

C. Legal Issues

Worries stick with safety measures and confidentiality of individual all the way through legislative levels.

D. Compliance

Numerous regulations pertain to the storage and use of data requires regular reporting and audit trails. In addition to the requirements to which customers are subject, the data centers maintained by cloud providers may also be subject to compliance requirements.

E. Freedom

Cloud computing does not allow users to physically possess the storage of the data, leaving the data storage and control in the hands of cloud providers.

F. Long-Term Viability

You should be sure that the data you put into the cloud will never become invalid even your cloud computing provider go broke or get acquired and swallowed up by a larger company.

G. Issues in Cloud Interoperability

1) Intermediary Layer

A number of recent works address the interoperability issue by providing an intermediary layer between the cloud consumers and the cloud-specific resources (e.g. VM).

2) Open Standard

Standardization appears to be a good solution to address the interoperability issue. However, as cloud computing just starts to take off, the interoperability problem has not appeared on the pressing agenda of major industry cloud vendors.

3) Open API

SUN has recently launched the Sun Open Cloud Platform [10] under the Creative Commons license. A major contribution of this platform is the proposed (in-progress) the cloud API. It defines a set of clear and easy-to-understand RESTful Web services interfaces, through which cloud consumers are able to create and manage cloud resources, including compute, storage, and networking components in a unified way.

4) SaaS and PaaS Interoperability

While the aforementioned solutions generally tackle with IaaS interoperability problems, SaaS interoperability often involves different application domains such as ERP, CRM, etc. A group of experts in the field of data mining raises the issue of establishing a data mining standard on the cloud, with a particular focus on “the practical use of statistical algorithms, reliable production deployment of models and the integration of predictive analytics” across different data mining-based SaaS clouds.

PaaS interoperability not yet discovered Since PaaS involves the entire software development life-cycle on the cloud, it would be more difficult to reach the uniformity with regards to the way consumers develop and deploy cloud applications.

VII. CHALLENGES ON CLOUD ADOPTION PERSPECTIVE

Based on a survey conducted by IDC in 2008[1]

A. Security

Well-known security issues such as data loss, phishing, botnet (running remotely on a collection of machines) pose serious threats to an organization's data and software. The multi-tenancy model and the pooled computing resources on cloud computing has introduced new security challenges such as shared resources (hard disk, data, VM) on the same physical machine invites unexpected side channels between a malicious resource and a regular resource. And, the issue of “reputation fate-sharing” will severely damage the reputation of many good Cloud “citizens” who happen to, unfortunately, share the computing resources with their fellow tenant - a notorious user with a criminal mind. Since they may share the same network address, any bad conduct will be attributed to

all the users without differentiating real subverters from normal users.

B. Costing Model

Cloud consumers must consider the tradeoffs amongst computation, communication, and integration. While migrating to the Cloud can significantly reduce the infrastructure cost, it does raise the cost of data communication.

C. Charging Model

From a cloud provider's perspective, the elastic resource pool (through either virtualization or multi-tenancy) has made the cost analysis a lot more complicated than regular data centers, which often calculates their cost based on consumptions on static computing.

D. Service Level Agreement

It is vital for consumers to obtain guarantees from providers on service delivery. Typically, these are provided through Service Level Agreements (SLAs) negotiated between the providers and consumers.

VIII. SECURITY AND PRIVACY ISSUE

Cloud computing can provide infinite computing resources on demand due to its high scalability in nature, which eliminates the needs for Cloud service providers to plan far ahead on hardware provisioning. Many companies, such as Amazon, Google, Microsoft and so on, accelerate their paces in developing cloud computing systems and enhancing its services providing to a larger amount of users.

In this paper, we investigate the security and privacy concerns of current cloud computing systems provided by an amount of companies. As cloud computing refers to both the applications delivered as services over the Internet and the infrastructures (i.e., the hardware and systems software in the data centers) that provide those services.

Based on the investigation security and privacy concerns provided by companies nowadays are not adequate, and consequently result in a big obstacle for users to adapt into the cloud computing systems. Hence, more concerns on security issues, such as availability, confidentiality, data integrity, control, audit and so on, should be taken into account.

A. Security on Demand

Cloud services are applications running somewhere in the cloud computing infrastructures through internal network or Internet. Cloud computing allows providers to develop, deploy and run applications that can easily grow in capacity (scalability), work rapidly (performance), and never (or at least rarely) fail (reliability), without any concerns on the properties and the locations of the underlying infrastructures.

Cloud computing systems can achieve the following five goals together [2]:

1) Availability

The goal of availability for cloud computing systems (including applications and its infrastructures) is to ensure its users can use them at any time, at any place. As its web-native nature, cloud computing system enables its users

to access the system (e.g., applications, services) from anywhere. This is true for all the cloud computing systems (e.g., DaaS, SaaS, PaaS, IaaS, and etc.). Required to be accessed at any time, the cloud computing system should be severing all the time for all the users (say it is scalable for any number of users). Two strategies, say hardening and redundancy, are mainly used to enhance the availability of the cloud system or applications hosted on it.

2) Confidentiality

It means keeping users' data secret in the cloud systems. There are two basic approaches (i.e., physical isolation and cryptography) to achieve such confidentiality, which are extensively adopted by the cloud computing vendors.

3) Data integrity

In the cloud system means to preserve information integrity (i.e., not lost or modified by unauthorized users). As data are the base for providing cloud computing services, such as Data as a Service, Software as a Service, Platform as a Service, keeping data integrity is a fundamental task.

4) Control

In the cloud system means to regulate the use of the system, including the applications, its infrastructure and the data.

5) Audit

It means to watch what happened in the cloud system. Auditability could be added as an additional layer in the virtualized operation system (or virtualized application environment) hosted on the virtual machine to provide facilities watching what happened in the system. It is much more secure than that is built into the applications or into the software themselves, since it is able watch the entire access duration.

IX. CONCLUSION

This paper discussed the architecture and popular platforms of cloud computing. It also addressed challenges and issues of cloud computing in detail. In spite of the several limitations and the need for better methodologies processes, cloud computing is becoming a hugely attractive paradigm, especially for large enterprises.

Cloud Computing initiatives could affect the enterprises within two to three years as it has the potential to significantly change IT.

REFERENCES

[1] T. Dillon, C. Wu, and E. Chang, "Cloud Computing: Issues and Challenges," *2010 24th IEEE International Conference on Advanced Information Networking and Applications(AINA)*, pp. 27-33, DOI= 20-23 April 2010

[2] M. Q. Zhou, R. Zhang, W. Xie, W. N. Qian, and A. Zhou, "Security and Privacy in Cloud Computing: A Survey," *2010 Sixth International*

Conference on Semantics, Knowledge and Grids(SKG), pp.105-112, DOI= 1-3 Nov. 2010

[3] J. F. Yang and Z. B. Chen, "Cloud Computing Research and Security Issues," *2010 IEEE International Conference on Computational Intelligence and Software Engineering (CiSE)*, Wuhan pp. 1-3, DOI= 10-12 Dec. 2010.

[4] S. Zhang, S. F. Zhang, X. B. Chen, and X. Z. Huo, "Cloud Computing Research and Development Trend," In *Proceedings of the 2010 Second International Conference on Future Networks (ICFN '10)*. IEEE Computer Society, Washington, DC, USA, pp. 93-97. DOI=10.1109/ICFN.2010. 58.

[5] J. J. Peng, X. J. Zhang, Z. Lei, B. F. Zhang, W. Zhang, and Q. Li, "Comparison of Several Cloud Computing Platforms," *2009 Second International Symposium on Information Science and Engineering (ISISE '09)*. IEEE Computer Society, Washington, DC, USA, pp. 23-27, DOI=10.1109/ISISE.2009.94.

[6] S. Zhang, S. F. Zhang, X. B. Chen, and X. Z. Huo, "The Comparison between Cloud Computing and Grid Computing," *2010 International Conference on Computer Application and System Modeling (ICCSAM)*, pp. V11-72 - V11-75, DOI= 22-24 Oct. 2010.

[7] M. M. Alabbadi, "Cloud Computing for Education and Learning: Education and Learning as a Service (ELaaS)," *2011 14th International Conference on Interactive Collaborative Learning (ICL)*, pp. 589 – 594, DOI=21-23 Sept. 2011.

[8] P. Kalagiakos "Cloud Computing Learning," *2011 5th International Conference on Application of Information and Communication Technologies (AICT)*, Baku pp. 1 - 4, DOI=12-14 Oct. 2011.

[9] P. Mell and T. Grance, "Draft nist working definition of cloud computing - vol. 21, Aug 2009, 2009.

[10] "Sun Microsystems Unveils Open Cloud Platform," [Online]. Available: <http://www.sun.com/aboutsun/pr/2009-03/sunflash.20090318.2.xml,2009>.

[11] W. Dawoud, I. Takouna, and C. Meinel, "Infrastructure as a Service Security: Challenges and Solutions," *2010 7th International Conference on Informatics and System*, pp. 1-8, March 2010.

[12] W. Itani, A. Kayssi, and A. Chehab, "Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures," *2009 8th IEEE International Conference on Dependable, Autonomic and Secure Computing*, 2009, pp. 711-716.

[13] B. Grobauer, T. Walloschek, and E. Stöcker, "Understanding Cloud Computing Vulnerabilities," *2011 IEEE Security and Privacy*, pp. 50-57, DOI= March/April 2011.

[14] W. A. Jansen, "Cloud Hooks: Security and Privacy Issues in Cloud Computing," *Proceedings of the 44th Hawaii International Conference on System Sciences*, 2011.



Santosh Kumar received the BE degree in Computer Science and Engineering from Visveswaraiah Technological University, Karnataka, India in 2007. He is currently working toward the M.Tech degree in the Computer Science from Graphic Era University, Dehradun, India. His research interest includes cloud computing security.



Dr. R H Goudar, currently working as an Associate Professor, Dept. of CSE, Graphic Era University, Dehradun. He also worked as Faculty at International Institute of Information Technology, Pune for 4 years and Indian National Satellite Master Control Facility, Hassan, India. His Subjects of Interest include Semantic Web, Network Security and Wireless Sensor Networks.