CLOUD COMPUTING: A NEW VISION OF THE DISTRIBUTED SYSTEM

Taha Chaabouni[1] and Maher Khemakhem[2]

[1]MIRACL Lab, FSEG, University of Sfax, Sfax, Tunisia
chaabounitaha@yahoo.fr
[2]MIRACL Lab, FSEG, University of Sfax, Sfax, Tunisia
maher.khemakhem@fsegs.rnu.tn

ABSTRACT

Cloud computing is a new emerging system which offers information technologies via Internet. Clients use services they need when they need and at the place they want and pay only for what they have consumed. So, cloud computing offers many advantages especially for business. A deep study and understanding of this emerging system and the inherent components help a lot in identifying what should we do in order to improve its performance. In this work, we present first cloud computing and its components then we describe an idea which attempts to optimize the management of cloud computing system that are composed of many data centers.

KEYWORDS

Cloud computing, distributed system, cloud technologies, data center, multi agent system, system management

1. INTRODUCTION

Users consume electricity without knowing how this power has been made. Cloud computing, this buzzword, is the adopting of this concept in the Information Technologies (IT) industries. So, cloud computing is a business model where IT services are offered to the consumer. The most important purpose of this system is to provide computing services easily like offering water, electricity, gas and telephone. So, the consumers use the resource and/or the service when they need according to the amounts of their requirements and consequently pay based on the rate of their use. Actually, cloud computing is the new way of informatics consuming. It represents a computational model in which a resource (software or hardware) is hosted, run, controlled and administrated through internet in large data centers. This resource is provided as a service. So, cloud computing offers a flexible and dynamic IT infrastructure.

Cloud computing represents the response to the new requirement in the IT. In fact, it can be said that cloud computing is an evolution of the concept "grid computing". The most important difference between them is in the method of management [8]. In grid computing, the user must manage the entire system (server, network element, operating system, software…). But, in cloud computing, the system is offered like a service. So, the user deals only with what he needs and doesn't be concerned with other services/issues. It means that cloud computing can be used friendly [10]. In fact, the cloud computing adopts the concept of utility computing. So, most people can use it without any specific knowledge of how the system operates or the need to mange anything. The grid computing is generally oriented to scientific researchers who have an important knowledge in computer sciences.

Sundarapandian et al. (Eds): ICAITA, SAI, SEAS, CDKP, CMCA, CS & IT 08,
pp. 245–252, 2012. © CS & IT-CSCP 2012 DOI : 10.5121/csit.2012.2521

Many projects are developed or under construction to respond to the increasing demand. Amazon with Elastic Compute cloud (EC2) was the leader in this domain. Then every actor-major in the IT follows and presents his own system like Microsoft and her system Azure, Google and App Engine, IBM and Blue Cloud...

This work is divided into two parts. The first one presents the notion of cloud computing, the corresponding architecture and the possible cases of use. The second part details an idea that attempt to optimize the management of cloud computing that is composed of more than one data center.

2. DEFINITION

The term "cloud" has been used many times ago to finally describe the use of information technology via internet. Cloud computing is not really a new technology. It represents a new way of using more than a really new technology. That is why we find now several proposed definitions. Everyone tries to find out a definition but it covers just certain aspects of the technology. The work [2] tries to summarize them. Finally, it proposes the following definition: clouds are a large pool of virtual and easy to reach resources (hardware and/or software). These resources can be dynamically adjusted depending to all the demand. In this model, we only pay what we consume.

The U.S. National Institute of Standards and Technology (NIST) proposes [3] the definition: "Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction".

This definition, that seems to have captured the commonly agreed aspects of cloud computing, was adopted by the most researchers. It is important to be agreeing on one definition to limit the scope of research and emphasize the potential business benefits.

In other ways, the main reason for the existence of these different definitions can be summarized in the fact that cloud computing is not really a new technology, but it is a new adoption of already existing technologies to run businesses in a new way. Cloud computing is based on several old concepts like [1 and 4]: Service-oriented architecture (SOA), distributed and grid computing (utility computing) and virtualization.

3. CLOUD COMPUTING ARCHITECTURE

In this section, we describe cloud computing in term of composition model, business model and deployment model.

3.1. Cloud computing composition model

Generally, cloud computing is divided into 4 layers [7]: the hardware layer, the infrastructure layer, the platform layer and finally the application layer. This layer architecture allows developing the system easily. In fact, every layer can be updated and/or changed without any knowledge and/or modification in other layers. This layer division is compared to OSI model [5] of the network protocol. So, with such architecture, the deployment of new software or the installation of a new hardware component does not affect other element of the system.

The layer architecture is composed of:

The hardware layer: This layer regroups all the hardware components of the cloud computing system. It concerns the management of the physical server, network component, power and controlling system. In this layer, we speak about the supervision of the data center. Usually, a cloud computing provider manages several data centers.

The infrastructure layer: It represents the virtualization layer. It allows creating the virtual resource that will be used by the upper layer. The most used virtualization technologies are Xen[1], KVM[2] and VMware[3].

The platform layer: This layer is dedicated to the operating system and application frameworks. It depends on the virtual machine created in the lower layer.

The application layer: This layer is the highest level of the hierarchy. All the cloud applications are combined in this layer. It represents the front office of the cloud computing system.

The business model of cloud computing is based on this layer architecture. Each offer in the business model corresponds to one or two layers in the architecture model [10].

3.2. Cloud computing business Model

As we have seen in the introduction, cloud computing represents a response to the new requirement in the Information Technologies. The businesses model combines the offer of cloud computing to the consumers. The authors in [11] try to describe the possible offers. They conclude that everything is a service represented as XaaS like SaaS (Software as a Service), PaaS (Platform as a Service), HaaS (Hardware as a Service), DaaS ([Development, Database, Desktop] as a Service), IaaS (Infrastructure as a Service)… More examples can be found in [12].The three services most used in this model are IaaS, PaaS, and SaaS [6, 7, 13, 14 and 15] (figure 1).

IaaS (Infrastructure as a Service): consumers use directly the IT infrastructure (computing power, networks, storage …). These resources are provided over virtualization technologies. The physical resources are integrated or decomposed to respond to the consumers demand. The virtualization strategy consists on creating virtual machines as many as the need. So in this service, the provider manages only the resources and it is up to the consumers to define the operating system and the application that will be used.

PaaS (Platform as a service): This service provides the software resource including operating system, development frameworks… So, in this type of service, the consumer has to develop and manage only his application. The service provider offers all the necessary tools to the consumer to allow him to run his application.

SaaS (Software as a service): It refers to provide on-demand application over the internet. So all the system, from the hardware layer to the final application is administered and controlled by the service provider. The consumer uses the application only when he needs and has nothing to manage or to create to perform his need.

[1] Xen : www.xensource.com
[2] KVM : Kernel Based Virtual Machine, www.linux-kvm.org
[3] VMWare : www.vmware.com/products/esx

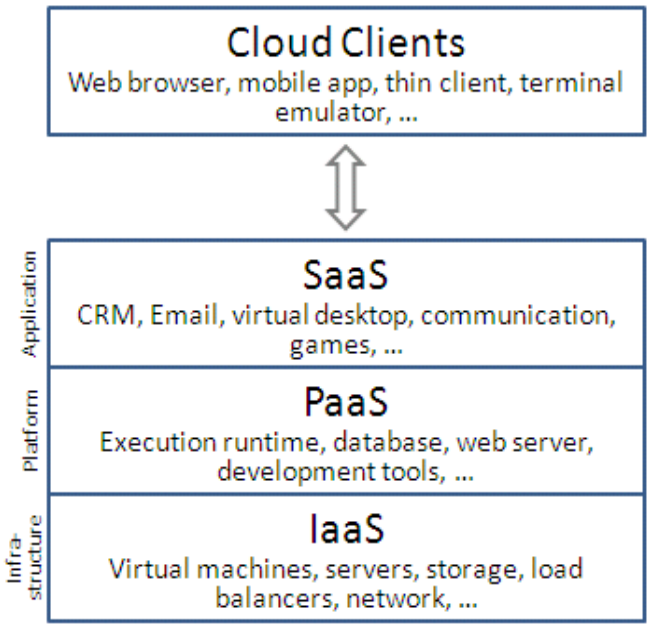

Figure 1. Business model of cloud computing

3.3. Cloud computing deployment model

The deployment model is composed essentially of 4 types defined in the cloud community [6, 7 and 16]:

Public Cloud infrastructure: The provider of this kind of system offers a set of resources (hardware or/and software) as a service to general public. The public clouds present many advantages like no initial capital investment on infrastructure. In exchange in this infrastructure, there is a lower control of the system by the user which hampers the efficiency in many business scenarios.

Private Cloud infrastructure: This type of deployment is operated for one user (organization). The management of this system can be performed by the organization itself or a third part. In private cloud, the user has more control in the system. That's why the use of this type is preferred in business especially at the first integration of cloud technology.

Hybrid cloud infrastructure: In this system, we have a combination of the other type of cloud computing deployment. It appropriates the business: the private cloud for the essential use and the public cloud when there is an increase of the need. So, hybrid cloud can be used in order to optimize users' resources depending on the actual activities (figure 2).

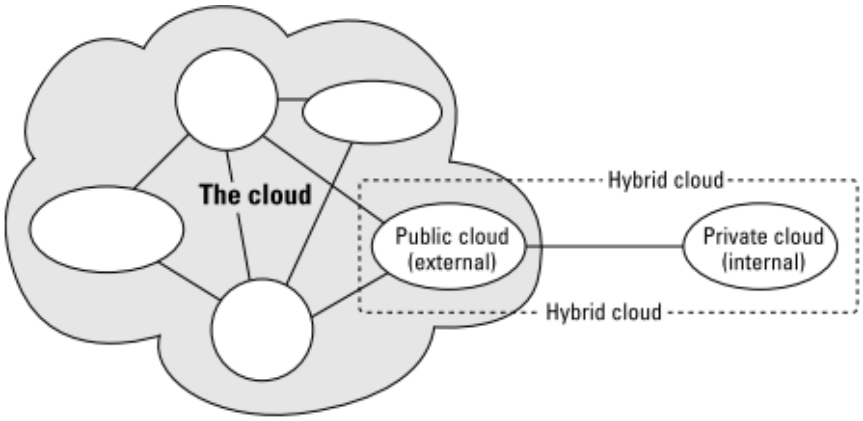

Figure 2. Example of hybrid cloud deployment

Community Cloud infrastructure: This system supports a specific community with a common function or purpose. It can be created by one organization or several organizations that have mutual concerns such as their mission, policies, security… The community cloud may be managed by the constituent organization(s) or by a third part.

4. INTEGRATION OF AN AGENT SYSTEM IN CLOUD COMPUTING

In this section, we present an idea of management optimization. Infrastructure as a Service (IaaS) offers resources as a service. Generally, these resources are presented in term of virtual machine (VM). The users demand a service (i.e. a set of resources (storage, computing power, bandwidth…)) and the provider of IaaS tries to respond to the need. The resources of the provider are usually organized in different data centers. Typically, a data center is composed of several physical servers which are interconnected and virtualized to optimize their use. The customers can be located in different places. So, the provider of the service has an interest to create data centers in different locations (figure 3)

.

Our idea is to integrate a multi-agent system to control the different data centers to optimize the management of the system and respond to the maximum of user demand.

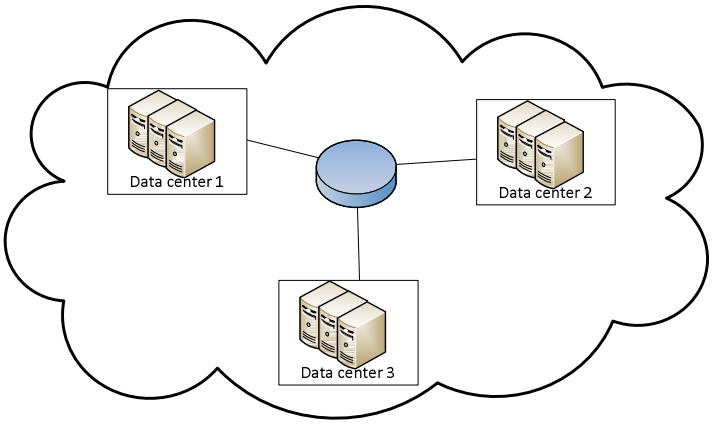

Figure 3. Exemple of Cloud Computing System

4.1. Related work

Cloud computing is still in the developing phase. Foster describes in [8] the current state of distributed systems and how to manage them. Furthermore, in cloud computing we can use the technologies that have been used in grid computing. The work [17] proposed an agent based framework for scalability in cloud computing. It focused on the communication inter agent to provide the user requirement even if we use a complementary service of another provider to satisfy the request. So, they attempt to use the agent technology to optimize the consumer work. For the best of our knowledge, there is no previous work which attempted to deal with multiple data centers belonging to the same provider. Consequently, we describe in the next section our proposal which aims to optimize resources utilization in case of having multiple data centers belonging to the same provider.

4.2. Proposed work

This work focuses on how to use efficiently resources in the different data center locations. The idea is to find out a strategy to respond to the max of client requests while maintaining a good quality of services. Typically, when a client demands a set of resources, the common idea is to affect this request to the nearest data center to him. But this solution can not be the best choice in all cases especially when there is network latency or the data center is overloaded. So, the provider needs a strategy to affect the client request to the best fit data center. In this case, the knowledge of the global state (forecast state) of cloud computing system is critical information that influences the choice of the best data center to use depending on the defined criteria. Actually, to obtain this result, an automatic communication system is needed to ensure the management of the cloud system. It will be responsible of all communications and negotiations between the different data centers to be able to choose the best one. So, a multi-agent system can be a solution to control the cloud computing system and to choose the best data center to the consumer request. We present two type of agent: Global Cloud Agent (GCA) and Local Agent (LA) (figure 4).

For each data center in the cloud system, we create a Local Agent. This agent is responsible for controlling and managing the local data center system. At any moment, it gives any information that the Global Cloud Agent needs like the free resources, the network state... Furthermore, LA keeps a history of the state of the data center. So, with this information it can predict and give the GCA more data that can help to manage better the cloud computing system. In this cloud computing system case, the local agent has an oversight role rather than a controlling (management) role.

In the cloud computing system, we found only one Global Cloud Agent. This agent will be responsible for controlling and managing the entire system. For each costumer's request, GCA collects information from the LA then decides which data center is the best for this client. This decision can be made according to the policy defined by the cloud computing provider.

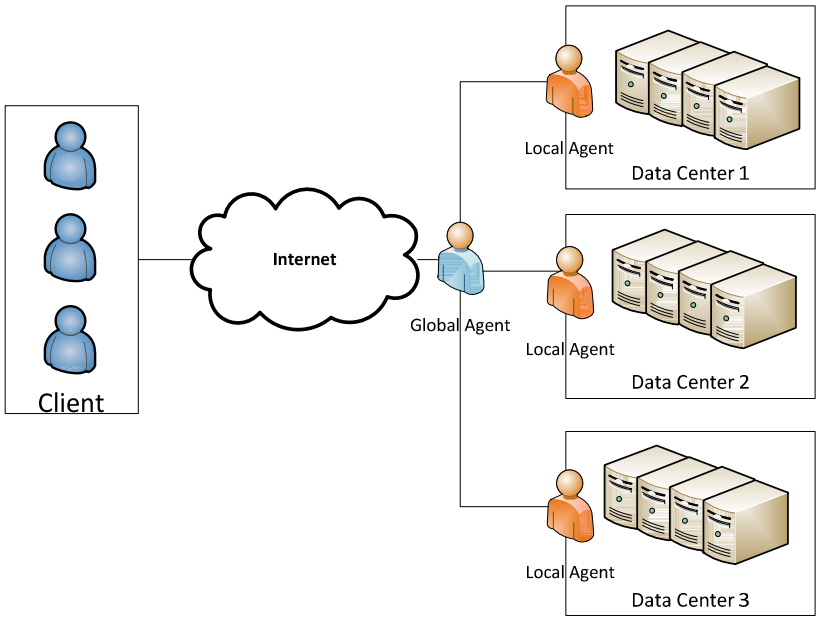

Figure 4. Integration of agent in cloud computing system

For example, we can choose to load balance between the data center or to use a data center only when the previous is fully loaded. Another strategy consists in using the lowers cost data center in the first place.

5. CONCLUSION

This work was divided into two parts: in first place, we have presented the cloud computing notion and architecture. Then, we have introduced an idea of using Multi-agent System to optimize the exploitation of cloud computing system. In future works, we attempt to define the method of communication between agents then the implementation of this system to test his real potential.

REFERENCES

[1] L. Youseff, M. Butrico and D. Da Silva, Toward a Unified Ontology of Cloud Computing, Grid Computing Environments Workshop (GCE '08), 2008.

[2] L. Vaquero, L. Rodero-Merino , J. Caceres and M. Lindner, A break in the clouds: towards a cloud definition. SIGCOMM Computer Communication Review, Volume 39, Number 1, pp. 50-55, January 2009.

[3] NIST definition of cloud computing, http://csrc.nist.gov/publications/drafts/800-145/Draft-SP-800-145_cloud-definition.pdf

[4] M. Vouk, Cloud computing- Issues, Research and Implementations. Journal of Computing and Information Technology (CIT) Volume 16, Number4, pp. 235–246, 2008.

[5] H. Zimmerman, OSI reference model — The ISO model of architecture for open systems interconnection. IEEE Transactions on Communications, 28, Apr. 1980, pp. 425–432.

[6] B. Sonisky, chapitre 1, cloud computing bible,wiley publishing inc 2011

[7] Q. Zhang, L. Cheng and R. Boutaba, Cloud computing : state of the art and research challenges. Journal of Internet Services and Applications, vol. 1, pp. 7–18, 2010.

[8] Foster, Y. Zhao, I. Raicu and S. Lu, "Cloud Computing and Grid Computing 360-Degree Compared," Grid Computing Environments Workshop (GCE '08), 2008.

[9] L. Wang and G. von Laszewski, Scientific Cloud Computing: Early Definition and Experience," in 10th IEEE International Conference on High Performance Computing and Communications, HPCC '08. , pp.825-830,2008

[10] R. Buyya, R. Ranjan and R. N. Calheiros, Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities. Proceedings of the Conference on High Performance Computing and Simulation (HPCS 2009), Leipzig, Germany. IEEE Press: New York, U.S.A., 21–24 June 2009; 1–11.

[11] B. P. Rimal, E. Choi and I. Lumb, A Taxonomy and Survey of Cloud Computing Systems. Fifth International Joint Conference on INC, IMS and IDC, pp 44–51, 2009.

[12] Gathering clouds of XaaS! https://www.ibm.com/developerworks/mydeveloperworks/blogs/sbose/entry/gathering_clouds_of_xaas?lang=en

[13] C.N. Hoefer and G. Karagiannis, Taxonomy of cloud computing services. In: Proceedings of the 4th IEEE workshop on enabling the future service-oriented Internet (EFSOI'10), Workshop of IEEE GLOBECOM 2010, pp 1345–1350, 2010

[14] C. N. Hofer and G. Karagiannis, Cloud computing services: taxonomy and comparison, Journal of Internet Services and Applications, Springer, Vol. 2, No. 2, pp. 69-79, September 2011.

[15] A. Lenk, M. Klems, J. Nimis, S. Tai and T. Sandholm. What's Inside the Cloud? An Architectural Map of the Cloud Landscape. In ICSE Workshop on Software Engineering Challenges of Cloud Computing, May 2009.

[16] T. Dillon, C. Wu and E. Chang, Cloud Computing: Issues and Challenges, in 24th IEEE International Conference on Advanced Information Networking and Applications, pp. 27-33, 2010.

[17] A. Snigh and M. Malhotra, Agent based framework for scalability in cloud computing, International Journal of Computer Science & Engineering Technology (IJCSET) Vol. 3, No. 4, April 2012.

Authors

Taha Chaabouni received his Master degree from the University of Sfax, Tunisia in 2010. He is currently preparing a Ph.D. at the University of Sfax, Tunisia. His research interests include cloud computing, distributed systems and performance analysis.
E-mail address: chaabounitaha@yahoo.fr

Maher Khemakhem received his Master of Science, his Ph.D. and Habilitation accreditation degrees from the University of Paris 11 (Orsay), France respectively in 1984, 1987 and the University of Sfax, Tunisia in 2008. He is currently Associate Professor in Computer Science at the Higher Institute of Business Administration at the University of Sfax, Tunisia. His research interests include distributed systems, performance analysis, Networks security and pattern recognition.
E-mail address: maher.khemakhem@fsegs.rnu.tn.