Volume 4, No. 4, March-April 2013

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

A Survey on Modeling and Simulation of Cloud Computing Environments

Garima Gupta*
Mtech
USICT, GGSIP University, Dwarka, Delhi
garimagupta2001@yahoo.com

Arunadha Chug
Assistant Professor
USICT, GGSIP University Dwarka, Delhi
a_chug@yahoo.co.in

Abstract: Cloud computing is a recent advancement in era of technologies wherein IT infrastructure, software, resources and applications are provided as "services" to the end-users under a pay-per-use based payment model. Based on the customer requirements, not only services can be offered to them but their dynamic needs can also be fulfilled. It basically provides an opportunity to dynamically scale the computing resources for applications. The application services which are hosted under Cloud computing model have various complex provisioning, composition, configuration, and deployment requirements. With the cloud computing ability to provide users dynamically scalable applications, providing the platform to share resources over the internet and avoid large expenses, cloud computing has recently emerged as a promising hosting platform. In this study an effort has been made to survey and explore various complexities and features in the field of cloud computing.

Keywords: Cloud Computing, CloudSim, Simulation, Virtualization, Data Centre, Performance modelling

I. INTRODUCTION

"Cloud computing is the collection of virtualized and scalable resources, capable of hosting application and providing required services to the users with the "pay only for use" strategy where the users pay only for the number of service units they consume". In simple terms, a cloud computing can be defined as a set of network enabled services which provides scalable, QoS guaranteed, normally personalized and inexpensive computing infrastructures on basis of the consumer's needs, which could be accessed in an efficient, simple and pervasive way throughout the world [10]. Cloud computing delivers infrastructure, platform, and software as services, which are available on subscription-based model services in a pay-per-you-use model to the consumers. These services are commonly referred to as IaaS (Infrastructure as a Service), PaaS (Platform as a Service), and SaaS (Software as a Service). On the other hand cloud computing requires the establishment of strong and complex networking connectivity and server configurations that are considered to be as self-monitoring and self-healing [18].

Clouds are made to be self-monitored so that they can handle and balance the workload effectively especially during peak hours and can decommission the resources which are not required during non peak hours. This way huge amount of money can be saved and can be utilized in some place. Clouds are made as self-healing in such a manner so that in case of any failure of physical hardware machine the backup of that machine can get created in as minimum time as possible. This results the rapid recovery in scenario when datacenter shuts down completely. Cloud technology is aiming to power up the newly created data centers as it will lead to huge advantages in terms of energy utilization. Also cloud aim to develop a consolidated environment having network of virtual services which may include the hardware, or business logic or a database or a user-interface under a single roof which will help users to access and deploy applications. The main advantage is that the resources can be accessible from all over the world just with the help of internet. The resources would be provided on the basis of the demand of the user or an enterprise at competitive costs. Developers and research associates with innovative ideas are not required any more to develop large capital outlays in both the hardware as well as software infrastructures. Also the cost of deployment of services can be reduced and human expenses would also get avoided. It offers remarkable and significant benefit to IT companies. So there is no need to set up basic hardware and software infrastructures.

II. TODAY'S BUSINESS NEEDS

Before we go further, the most important thing is to understand the today's business objectives. In terms of day by day advancement in technology with the growing competition and economics has raised a scenario where a business needs the following objectives when it comes to computing:-

a. ***Dynamism:*** - Dynamism means incorporating changes which are made at runtime. It is quite simple and very much similar to the way we use our mobile phone connection. If we want to talk more, we'll buy a top-up card. If we are a post-paid customer we'll change our plan to meet our changing requirements. Our needs are dynamic, so should be our infrastructure to support the changing needs. And therefore the dynamism becomes the basic need of our today's business requirements.

b. ***Abstraction:*** - This is the basic demand for today's business needs as with abstraction the user does not need to think about the operating system and hardware resources rather the time can be spent in building the application. This way the consumer can focus on main core part of the application instead worrying about the platform and infrastructure.

c. ***Resource Sharing:*** **-** The computing architecture should be implemented in such a manner that provides us the convenient way where applications and other hardware and software resources like network, datacenters and virtual machines can be shared. This will result in a need based flexible architecture which would help us achieving our business goals in an effective way with huge reduction in cost incurred.

III. PROPOSED SOLUTION

There's one model or style of computing which satisfies the above mentioned requirements to the satisfaction of the consumers. And therefore is becoming the technology trend of near future. It is well known as Cloud Computing. Cloud computing is an emerging technology which uses the internet and centrally designed remote servers to maintain data as well as the applications. Cloud computing allows consumers and industry associates to use various applications without installation on computer just with the help of an internet access [5]. This cloud technology allows much more then an inefficient computing by providing centralized access to storage, memory, processing and bandwidth and many other characteristics [1]. A simple example of cloud computing which is used on daily basis by almost everyone is Yahoo email, Gmail, or Hotmail etc. We do not need any software or any server or any other extra plug-in to use them. All what a consumer would require is just an internet connection and one can start communication through emails by sending and receiving. So in all we can say an internet is a cloud which is holding everything on it and providing services like email. These servers and email services are well managed by different service providers like Yahoo, Google etc. A cloud service has mainly three distinct characteristics that differentiate it significantly from traditional hosting. These services are sold on the basis of demand and on the basis of pay-per-use model; these services are purely dynamic in nature as a user can use as per their requirements just with internet connection. Significant innovations in the concept of virtualization and distributed computing with growing demand, an improved access to high-speed internet facilities and a weak economy, have increased business's interest in cloud computing [6].

IV. CHARACTERISTICS OF CLOUD COMPUTING

Cloud computing shows five important characteristics which are defined by NIST (National Institute of Standards and Technology) [11].

a. ***On-demand self-service:*** **-** A consumer can be given services at its discrete requirements.

b. ***Broad network access:*** - Capabilities are available throughout the network and can be accessed from any where in the world through standard mechanisms.

c. ***Resource pooling:*** - The resources are shared amongst multiple users on basis of their requirements. Different virtual resources are dynamically assigned and taken back according to the user needs and demands.

d. ***Rapid elasticity:*** **-** As the model is purely on demand basis so there is a flexibility to scale up or down the resources at any time.

e. ***Measured service:*** **-** Cloud systems have self monitored and self healing mechanism which would help in balancing load between servers and in case of failures the secondary server will take place.

V. CLOUD ARCHITECTURE

The basic architecture of cloud computing divide it into two sections: the front end and the back end. They connect to each other through a network, usually the Internet [4]. The front end is the side the computer user, or client. The back end is the "cloud" section of the system. Cloud architecture is the systems architecture of the software systems involved in the delivery of cloud computing. It basically involves multiple cloud components communicating amongst themselves over numerous application programming interfaces, usually web services. Basically cloud does provide everything as a service is it a software, platform, network or infrastructure. There are three different categories in which we can divide cloud services are infrastructure, platform, application etc [2].

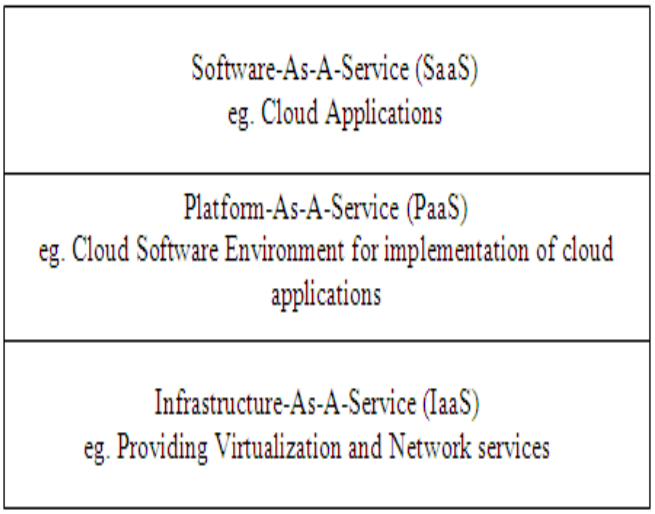

Figure. 1. Services provided by cloud computing.

These services are delivered and consumed in real-time scenario over the internet. We described them here as below:

a. ***Infrastructure as a Service (IaaS):-*** It involves various hardware related services using the fundamentals of cloud computing. These include some storage services like database or disk storage or virtualized machines. This is the base layer and acts as a foundation or basic building block for SaaS and PaaS. It includes the monitoring and management of virtualized servers and provides various virtualization and network services. The keyword behind this stack is Virtualization. Amazon [7] EC2 is an example of IaaS. In Amazon EC2, Elastic Compute Cloud the application gets executed on a virtual computer which is also called as an instance of a server. We have our own choice of virtual computer, like we have an option to select a configuration of CPU, virtual machines, memory and storage that is optimal for our

application. The IaaS provider supplies cloud infrastructure like application servers, routers, load-balancing, firewalls, database storage systems and other network equipment. The consumer buys these resources as a service on pay-per-usage basis.

b. ***Platform as a Service (PaaS):-*** It includes the cloud software environment for implementation of cloud applications. This is the middle layer i.e., PaaS (Platform as a Service). This middle layer of cloud is consumed mainly by developers for implementing cloud applications [20]. It also provides developers with a platform including systems and their end-to-end phases of developing, testing, deploying and hosting of web applications as a service delivered by a cloud based PaaS.

c. ***Software as a service (SaaS):-*** It is complete software offered as a service to the consumer. This is basically a cloud application which is being called as software here. And these applications are hosted by the cloud from where users can access them easily. Stand alone applications are quite costly and can only be managed by big organizations. Also there are various other requirements for making the application successful. The basic requirement for SaaS is a personal computer with a browser which is quite basic need. To access any cloud application there is a need of an internet connection. SaaS is a subscription based model which is delivered to the customer on demand – Pay as you use. Depending on the usage requirements the services can be subscribed. SaaS [12] [13] commonly referred to as the Application Service Provider model. Examples of key providers are SalesForce.com, Net Suite, IBM, Oracle and Microsoft etc.

VI. LAYERED DESIGN OF CLOUD COMPUTING ARCHITECTURES

Figure below shows the layered design of service-oriented Cloud computing architecture. This depicts that we can broadly classify the architecture in client browser, middleware tools, Web Application Servers and in the Database Storage System. IaaS comprises of physical cloud resources along with middleware features. The user-level middleware aims at providing Platform as service capabilities. The top most layers focuses on the application services or we can also call it as front-end which will directly interact with the user and will provide services to the layer beneath it. The Client browser sends an HTTP or SOAP request to the Middleware Tool which would perform validation, translation and transformation of the request file as per the requirement. This in turn sends the request ahead to the Application server or the backend system and waits for the reply to come back. Depending on the requirement again the servers can interact with the database storage systems. After processing the response is sent back to Middleware and which in turn again translates and transforms the response in the format which can be understood by the front end application. This is the basic structure we need to follow while developing any cloud application. As for the complete processing we need to have consolidated view of all three layers. All these three layers are interconnected for delivering the best performance. Below is the detailed diagram for the same and all these three layers have been individually explained here.

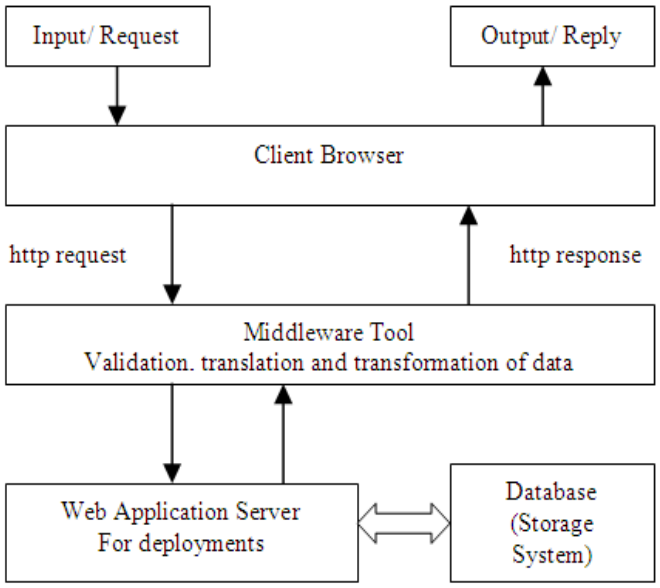

Figure. 2. Layered Design of Cloud Computing Architecture.

a. ***Application Level:*** - This layer acts like a frontend system and comprises of the applications that are directly available to the end-users or the customers. End users are the active entity which utilizes services throughout the world with the help of internet connection. These applications may be supplied by the Cloud provider (SaaS providers) and accessed by end-users either via a subscription model or a pay-per-use basis. This layer comprises of the services like social computing, enterprises etc. In this layer, there is a direct interaction from user and user it accepts the request. The input file used may be in xml format depending on the platform used.

b. ***Middleware Tools and Services:*** -This layer includes various middleware tools and offers various services to the enterprise. The layer also provides those programming environments that ease the creation, implementation, deployment, and execution of applications in clouds. In this layer there are also various application development technologies, such as JSP, Servlets, TIBCO, Enterprise Java Beans, Spring and hibernate, struts, velocity, Ruby on Rails which can be deployed to support applications running. This layer also implements the platform level services for hosting and managing User-Level application services in a runtime environment. Core services at this layer include Accounting, Billing, Execution monitoring and management. The well-known examples of services which are operating in this layer are Google App Engine [8] and Amazon EC2. The functionalities exposed by this layer are accessed by both SaaS and IaaS services. Basically it does provide the interaction between front

end applications to the back end applications by providing a common platform. Critical functionalities that are needed at this layer include messaging, service discovery, and load balancing. These functionalities are usually implemented by Cloud providers and offered to application developers at an additional premium. For example, Amazon offers a load balancer and a monitoring service (Cloud watch) for the Amazon EC2 developers/consumers.

c. ***Server Level:*** -This layer defines the concept of virtualization mainly. It provides us various features like virtual networking, virtual compute, virtual security and virtual storage. The power in different Cloud computing environments is supplied by data centers which are particularly installed with hundreds to thousands of hosts. At the Server Level layer there exist physical resources like storage servers and application servers that power up the data centers. These datacenters should be power efficient for the performance modeling. These servers are managed by the higher level virtualization policies and services and toolkits that allow sharing between its various virtual instances. These Virtual Machines are isolated from each other, thereby making fault tolerant behavior and isolated security context possible. This layer also interacts with the database via SQL requests and response.

VII. DESIGN AND IMPLEMENTATION

Our main goal is to design a simulator for academic and industrial purpose to model and to test new policies for better utilization of the cloud infrastructure and the services. Using this simulator, performance and utilization of resources will be measured for resource allocation policies and application scheduling [9]. Developers will be able to add the fault tolerance policy and placement and migration policy to test the overall performance of the system.

This simulator will help to model and analyze different components of cloud computing. Using this simulator, cloud service providers can measure the performance of the applications running on behalf of the customers. Here is the list of features of an ideal cloud computing simulator:

a. ***Physical machine:*** - Physical machine is a collection of resources. These resources are shared among multiple virtual machines, by using virtualization technology. Interface will allow creating, to start, to stop and to destroy virtual machines within a physical machine. All the machines can be monitored using monitoring tool as per the requirements of any organization. Also interface will provide different kind of options to configure and assign priorities among different virtual machines.

b. ***Resource allocation:*** - Resource allocation policies decide the amount of resource to be allocated to a particular or set of virtual machines. This policy can also update the resource allocation dynamically depending on the requirement. For implementing prioritization, we can provide more resources to a particular virtual machine, than other virtual machine. Resources can be interfering or sharing with each-other. Different technologies and prioritization schemes will help in opting out for the best methodology that matches with the industry requirements. And would be based on various factors like different seek latency, cpu utilization and overhead incurred. Performance of storage technology can significantly vary the system throughput.

c. ***Storage system:*** - Storage system is an important part of data center. Data centre is a server which can be divided into a) Storage area network (SAN) b) Direct area storage [19]. Most cases storage area network is used in a datacenter. Storage system contains virtual machine images, which are snapshot of virtual machines and can be managed effectively. On some Using storage is network creates some challenges for placement of virtual machines but also provides flexibility.

d. ***Placement Policies:*** - Placement policies both refer to virtual machine and application placement policies. Virtual machine placement policy decides the target physical host based on some parameter value. And application placement depends on, available resource and number of applications running on virtual machines. Placement policies can be classified as a) initial placement policy b) dynamic placement policy. Initial placement policy is executed at the start-up time of an application or virtual machine placing. Dynamic placement policies are executed on the fly depending upon different parameters like availability of different resources and violation of service level agreements.

e. ***Migration policies:*** - Migration policies decide when, which and where to migrate a virtual machine. This policy specifies the limiting condition which triggers a specific virtual machine to migrate. The target machine, which also to be manipulated should satisfy some condition. For example, a virtual machine should be migrated when the load produced by the running application is higher than 90% of virtual machines capacity. The selected target physical machine should have sufficient high resource available for hosting the migrated virtual machine. This resource availability condition can be twice of the previous resource allocation.

f. ***Virtualization techniques:*** - Different virtualization techniques implement resource allocation different way. In case of XEN, there exists Dom-0 (privileged domain) and Dom-U (virtual machines). But in KVM, there is no concept of domains. Also migration process implementation is also different [17].

g. ***Fault Tolerance:*** - Fault tolerant feature is for the backup of the servers. Cloud computing is an emerging technology which aims to deliver high performance applications and services which will be secured and reliable and sustainable. These services are further categorized as Software, Infrastructure, or Platform as services (SaaS, IaaS, PaaS). So, with this feature when primary server gets failed due to any circumstances the secondary server will take its place without hampering the normal functionality of the business.

VIII. MODELING AND SIMULATION UNDER VIRTUALIZED ENVIRONMENT

Many challenges are to be looked upon when researchers and industry professionals attempt to design and model the various complexities in virtualized environments so that they can assess and analyze performance and other issues. The complexity arises because of the virtual machines interaction as they not only interact with the physical machine but they also interact amongst themselves [14]. Therefore, the concept of virtualization needs to be realized so that effective utilization of resources can be made without incurring much cost.

a. CloudSim: - It is a recently created simulator which provides us a dynamic platform to simulate our datacenters. It has various features due to which it allows seamless modeling, simulation, and also experimentation of new emerging Cloud computing applications, services and infrastructures. The framework on which CloudSim is based upon consists of a datacenter node and a virtualization engine. This virtualization engine itself creates virtual machines on this datacenter node [15]. CloudSim is capable of handling all simulations of cloud computing services and applications and infrastructures under real time scenarios. It can also be used to determine the loop holes before the application gets deployed in real production environment. Developers can use this simulation framework effectively to determine performance issues. This tool gives us direction and opportunity for future work.

b. Features of CloudSim functionalities: -

a) It supports for modeling and simulation of various data centers.
b) It supports for simulation of virtualized servers and allows creating virtual machines and datacenters.
c) Also support for modeling and simulation of energy-efficient resources.
d) Application can be analyzed by entering the real data using CloudSim.
e) Support for modeling and simulation of interconnected clouds.
f) It also supports for dynamic addition of simulation elements, and for various allocation policies like allocation of hosts amongst various virtual machines.
g) Simulation can be stopped and started as per the user requirement.
h) Also support for the allocation of host resources [3].

IX. FUTURE SCOPE

Modeling and simulation of such environments that consist of providers encompassing multiple services and routing boundaries present unique challenges. Further, recent studies have revealed that data centers consume unprecedented amount of electrical power, hence they incur massive capital expenditure for day-to-day operation and management. To achieve simulation of the aforementioned Cloud computing environments, much of my future work would investigate new models and techniques for allocation of services to applications depending on energy efficiency and expenditure of service providers. Cloudsim aims to model and test, policies (placement, migration, resource allocation) to utilize cloud resources efficiently. Our future work is to develop other features based on these specific developed features like migration and fault tolerance specially. Because this feature helps many more features to be implemented in the simulator.

X. CONCLUSION

In this study we explored regarding the cloud computing, advantages of cloud computing, risks involved in cloud computing and various other approaches to solve those risks. However cloud computing is still struggling in its infancy, with all its pros and cons. IT technicians are working hard on the current challenges and pursuing research to improve on the lacking areas. Several communities and various computing groups have been formed with the common goal of exploring the various techniques and different concepts which are offered by cloud computing. So that the standards can be established this can commonly be used by different services providers. Cloud computing is also surrounded with several issues of security and accessibility. Its security deficiencies and benefits need to be considered very carefully before we go ahead towards its implementation [16]. With the rapid advances of Cloud technologies, there is a new demand for tools to study and analyze the benefits of the technology and the best practices to apply the technology to large-scaled applications. Therefore, the tools will evolve over the time, and the result of the processes will improve quality of the model and of the analysis it supports. In the long term through various simulation experiment testers would be able to identify new features and issues, model them, and develop and evaluate new mechanisms and algorithms for resource management, this way improving performance of emerging Cloud applications.

XI. REFERENCES

[1] R. Buyya, C. S. Yeo, and S. Venugopal, "Market oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities", published in the proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (2008)

[2] 2.R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility", presented in Future Generation Computer Systems, Elsevier (2009)

[3] 3.Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, César A. F. De Rose and Rajkumar Buyya, "CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms", presented by Cloud Computing and Distributed Systems (CLOUDS) Laboratory Department of Computer Science and Software Engineering, The University of Melbourne, Australia.

[4] 4.Amazon Web Services, "Amazon Elastic Compute Cloud (Amazon EC2)". Available: http://www.amazon.com/ec2/.

[5] 5. A. Weiss, "Computing in the Clouds", published in netWorker, ACM Press, New York, USA (2007)

[6] 6. R. Buyya, S. Pandey and C. Vecchiola, "Cloudbus Toolkit for Market - Oriented Cloud Computing", presented by Cloud Computing and Distributed Systems (CLOUDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Australia (2009).

[7] 7. Bhathiya Wickremasinghe, "CloudAnalyst: A CloudSim-based Tool for Modelling and Analysis of Large Scale Cloud Computing Environments" (2009)

[8] 8. Google App Engine, "Run your web apps on Google's Infrastrusture" Available at: http://code.google.com/appengine/

[9] 9. Bhathiya Wickremasinghe1, Rodrigo N. Calheiros, and Rajkumar Buyya, "Cloud Analyst: A CloudSim-based Visual Modeler for Analyzing Cloud Computing Environments and Applications", presented in Cloud Computing and Distributed Systems (CLOUDS) Laboratory, Department of Computer Science and Software Engineering, The University of Melbourne, Australia.

[10] 10. Wang et al., "Cloud computing: A Perspective study", published in Proceedings of the Grid Computing Environments (GCE) workshop (2008)

[11] 11. "The NIST Definition of Cloud Computing", National Institute of Standards and Technology (NIST), Information Technology Laboratory (www.csrc.nist.gov), (2009)

[12] 12. "Measuring, Modeling and Managing VM Shared Resources" published in The International Journal of Computer and Telecommunications Networking (2009)

[13] 13. "Probabilistic Performance Modeling of Virtualized Resource Allocation", published in ICAC 10th Proceeding of the 7th International conference on Autonomic Computing, pp.99-108, ACM, 2010. Available at: http://doi.acm.org/10.1145/1809049.1809067.

[14] 14. "Application Performance Modeling in a Virtualized Environment" , published in High Performance Computer Architecture (HPCA), IEEE 16th International Symposium on , pp.1-14 Jan. 2010.

[15] 15. "CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services", published in ICCP 2009. Available at: http://www.gridbus.org/reports/CloudSim-ICPP2009.pdf.

[16] 16. Traian Andrei and Raj Jain, "Cloud Computing Challenges and Related Security Issues". Available at: http://www.cse.wustl.edu/~jain/cse571-09/ftp/cloud.pdf

[17] 17. Tickoo et al., "Modeling Virtual Machine Performance: Challenges and Approaches", published in ACM SIGMETRICS Performance Evaluation Review (2009). http://portal.acm.org/citation.cfm?doid=1710115.1710126

[18] 18. Armbrust, M., et al., "Above the clouds: A Berkeley view of cloud computing". Published in Technical Report, UCB/EECS-2009-28, EECS Department, UC Berkeley. Available at: http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf (2009)

[19] 19. "On the Use of Fuzzy Modeling in Virtualized Data Center Management", published in Proceedings of the Fourth International Conference on Autonomic Computing (ICAC), IEEE Computer Society, (2007).

[20] 20. "A Taxonomy and survey of Cloud Computing Systems", published in Fifth International Joint Conference on INC, IMS and IDC (2009)