CLOUD TECHNOLOGIES - SERVICES, ARCHITECTURES AND SIMULATIONS

Georgi Petrov Georgiev

Abstract: Cloud communications are an evolving technology that requires attention to various issues such as quality of service, resource provisioning, security, energy management and reliability. The purpose of this paper is to discuss the development of cloud communications and to analyze and compare the various cloud networks researches and simulation tools, focusing on widely used programming simulation systems such as Green Cloud. Most cloud network studies rely on simulation, as the use of real cloud infrastructures for research is a complex and time-consuming task, has a very high cost, and requires a lot of resources to achieve realistic results. In this paper the main services provided by cloud computing network are presented. The models of services and general topologies are pointed.

Keywords: Cloud Computing, Simulation Systems, Green Cloud, Services, Energy Management

ITHEA Keywords: K.6 Management of Computing and Information Systems; K.6.2 Installation Management - Computing Equipment Management

Introduction

After the appearance of cloud technologies, they have become fundamental to IT operations worldwide, replacing traditional business models. Businesses now have access to a huge nomenclature of software and services online via a virtualized environment, avoiding the need for expensive IT infrastructure investments.

Cloud services are defined as a shared pool of on-demand computer and computing services available over the global Internet network. Their configuration is dynamic, to achieve as much as possible better optimization of their work. Cloud communications are based on a wide range of different technologies such as distributed systems, high performance computers, virtualization, information storage, networks, security, management and automation, Service-Oriented Architecture (SOA), Business Process Management (BPM), Service-Level Agreement (SLA), Quality of Service (QoS) etc. [Jaraweh et al, 2012].

Cloud communications are the result of the evolution of many existing technologies such as: Grid Computing, Utility Computing, Services Computing, and Distributed Systems in general. The
relationship between cloud communications and other distributed processing technologies is presented in Figure 1.

The Cloud overlaps with existing technologies like Web 2.0 that covers almost the whole spectrum of service-oriented applications. Cloud Computing lies at the large-scale side like Supercomputing and Cluster Computing that have been more focused on traditional non-service applications. Grid Computing overlaps with all these fields where it is generally considered of lesser scale than supercomputers and Clouds.

The generally accepted definition of Cloud Computing comes from the National Institute of Standards and Technology (NIST) [Mell and Grance, 2011]. The NIST definition essentially says that “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”. In plain terms, it is the ability for the end-users to utilize parts of bulk resources. These resources can be acquired quickly and easily. NIST also offers up several characteristics that it sees as essential for a service to be considered “Cloud-based”. These characteristics include:

- On-demand self-service. The ability for an end user to sign up and receive services without the long delays that have characterized traditional IT.
- Broad network access. Ability to access the service via standard platforms (desktop, laptop, mobile etc).

- Resource pooling. Resources are pooled across multiple customers.

- Rapid elasticity. Capability can be scaled to cope with demand peaks.

- Measured Service. Billing is metered and delivered as an utility service.

Cloud services are constantly evolving, and are also rapidly spreading computing systems that are of interest to both the industry and the academic community. Many experts believe that these services will very quickly become dominant in the IT sector. As a result, universities are actively introducing cloud systems into their curricula. The big problem faced by cloud services is the lack of programs to provide simulation and modeling of these systems [Jaraweh et al, 2012].

Cloud architecture and services

Three types of architectures are known for building cloud systems.

Two Layer Architecture: Works well for early (new emerging) data centers with a limited number of computing servers. Depending on the type of switches used in the access network, two-tier data centers can support up to 5500 servers. The number of main switches and the main links capacity determine the maximum network transmission rate. Figure 2 shows a two-layered architecture.

![Two Layer Architecture](image-url)
Three layer architecture: The three-layer data center (layer) architectures are the most common nowadays. These include: access layer, aggregation and a core layer. Availability of the aggregation layer makes it easier to increase the number of servers (to over 10,000 servers), at the same time they maintaining low cost Layer-2 (L2) switches. For my simulation studies, this type of architecture is used on a cloud with parameters that I have set. Its work has been studied (see Figure 6).

The three layer high-speed architectures (shown in Figure 3) of the data centers are designed to optimize the number of nodes, core capacity (backbone) and distribution networks, which are currently the ones that limit the maximum number of nodes in a data center or speed of transmission. It has possible connections of the order of 100 Gb / s (IEEE 802.3ba standard) and is standardized in June 2010.

It should be mentioned that there are three basic models for the delivery of cloud services:

- Public cloud - the provider of this type of service operates a shared services environment that is available to each user over the Internet. Data centers are the property of the vendor.
- Private cloud - in this model the cloud was created and operated only for the institution for which it was designed. It may be located on the premises of the establishment and provides a high level of data security.

- Hybrid cloud - offers a combination of the above two models. The institution may require to use a public cloud for temporary use while the cloud remains for the core activities [Garison et al., 2013].

Cloud computing providers offer their cloud services based on several basic models, such as [Calheiros et al., 2011]:

- Infrastructure as a Service (IaaS) – With this model, cloud service providers offer virtual machines, virtual memory, and other physical resources. The cloud user must have an operating system installed to deploy their cloud applications.

- Platform as a Service (PaaS) - With this model, the cloud provider provides a platform of resources including operating system, programming languages, libraries, services and tools supported by the provider [Vilaplana et al., 2015].

- Software as a Service (SaaS) - Cloud service providers in this case provides users with access to the installed software and databases.

Figure 4. Layered Cloud Network Architecture [Calheiros et al., 2011].
Figure 4 shows the layered design of cloud network architecture. It includes the following layers: User level; User-level middleware (SaaS); Core of middleware (PaaS); System level (IaaS) [Calheiros et al, 2011].

**A review of tools for modeling and simulating the operation of the cloud systems**

Simulations and modeling are used in the design and pre-testing of various devices. These simulation tools and environments can also be used to test cloud systems. Experimenting in a real cloud environment has the following problems [Devi and Sujan, 2014]:

- Expensive, because it takes a very large number of resources for a long time;
- Repeating of the experiments is impossible;
- The environment offered by the provider may not support developing applications.

When researching and simulating cloud systems, it is very important to choose a product that meets our specifications and can provide us with a wide range of cloud related information, data channel load, server load, power consumption, etc. It is also essential that the product is easy to work with and that it presents the information in a form that is accessible to us. Therefore, some of the most common simulators used to simulate cloud systems have been reviewed.

**CloudSim** - This is a set of tools (libraries) for simulating cloud environments. It is developed in the CLOUDS laboratory of the Department of Computer Science and Engineering at the University of Melbourne, Australia. CloudSim is not a tool that can be used immediately. The user cannot enter the selected data and the simulation starts immediately. Because CloudSim is a library, it is necessary to write a Java program using the built-in tools to create the user-defined scenario.

**Open Cloud Testbed (OCT)** - It is used to evaluate various cloud computing systems and their interoperability. It is used for the development of cloud computing programs for cloud infrastructure. Figure 6 shows the OCT architecture. The main features of the simulator are that various cloud computing systems and processes have been installed, such as Eucalyptos, Hadoop, CloudStore, Sector / Sphere and Thrift. This is to make OCTs easy to study interoperability, referral to other researches, network libraries development and monitoring systems. OCT architecture includes highly productive protocols, processes and infrastructure at all levels. It uses a highly efficient 10 Gb/s network based on extremely fast transport protocols (such as the UDP protocol) [Grossman et al, 2009].

**GreenCloud** - This is a complex packet level simulator for determining the energy efficiency of cloud data centers focusing on cloud communications. It offers very detailed modeling of the power consumption of data center and communication equipment such as computing servers, network devices and communication lines. It can be used to develop new solutions for monitoring, resource allocation,
load distribution, and optimization of unused protocols and network infrastructure. Figure 5 shows the architecture of GreenCloud [Kliazovich et al, 2015].

**Simulation program for Elastic Cloud infrastructure (SPECI)** is a simulation tool that allows you to study scaling aspects as well as presenting properties of future data centers. Given the size and middleware design policy, SPECI simulates the performance and behavior of data centers.

**Public IaaS Cloud Simulator (PICS)** is a cloud-based simulator designed to evaluate the cost and characteristics of the infrastructure provided as cloud services by identifying the necessary virtual machines and evaluating storage services, resource elasticity, task planning, and different load. The disadvantage of this simulator is the lack of support for the heterogeneous cloud characteristics and it does not include a cost model for communication models.

![Figure 5. GreenCloud architecture [Kliazovich et al, 2015]](image)
Also, many other simulators are reviewed in [Sinha and Shekhar, 2015].

After a review and in-depth analysis of the various cloud simulation tools, the GreenCloud simulator was selected. It was preferred because of the presence of a graphical interface that helps the program to be used by a wider range of experts and users without highly specialized knowledge. While CloudSim and OCTs that do not have such a graphical interface require more in-depth knowledge in some areas and, in particular, extensive JAVA programming skills. The second reason for choosing is accessibility. The other two simulators are not as readily available and can hardly be found without paying any fees or licenses, even if they are used for training and researchs. The third reason is that GreenCloud, in addition to the possibility of "standard" cloud simulations, also provides information on the energy efficiency of the entire system, which has been given great attention in the 21st century.

Simulation results and analyzes

Experimental simulations will be performed in a cloud with the following topology, consisting of 3 layers - access network, aggregation network and core network. This is the typical model that is used to create cloud communications with a data center. It is recommended and used by companies such as Cisco system, which mainly work with this established and proven topology. The core is made up of 1 switch. The aggregation network is made up of 2 switches, and two switches are used in the access network for the simulation. The total number of servers in the created system is 24. The structure of the simulated network is presented in Figure 6. The cloud simulation will be done at a low load, with the aim of assessing mainly its workability.

Figure 7 shows the system parameters as well as the energy consumed by the above-defined number of devices required to handle the tasks. According to the energy consumption per device, each layer shows that it is the largest for the aggregation network - 102.8 W*h then the core network - 51.4 W*h, the servers 23.7 W*h and the access network - 5.5 W*h. It can be seen that the aggregation consumes almost twice as much as the core network. Tasks submitted for processing are just over 16600.

The reference [Claudio et al, 2015] mentions that data centers require a huge amount of energy to work. In Europe, these data center values are projected to reach 93 TWh by 2020. Almost 75% of this consumption is attributable to IT and cooling equipment. Ideally, the power consumption of network devices should be proportional to the workload. In fact, however, their energy consumption consists of two parts, fixed (for switches) and variables (for active transmitters).
Figure 8 shows the traffic showing the data center load. As well as it shows the average load of the processors on different servers from the simulated cloud. Figure 9 shows the number of tasks performed by virtual machines in the cloud. As can be seen, the specified cloud (cloud) handles these tasks, and the load is far from the possible maximum. Also, Figure 9 shows the load of the individual lines connecting the switches from the distribution to the support network. As you can see, the connection line is one.

Figure 7. Consumed energy and parameters of the cloud system
In this simulation, cloud network equipment is more than necessary. There is a large reserve of equipment for future needs, but under such parameters we do not need a cloud of such computational power as this would lead to a significant increase in capital expenditure (CAPEX). It is seen that the cloud is dealing with the set tasks, with roughly a quarter of its total load, which means that it has a 75% reserve, which will be linked to huge investment costs. Also, let’s pay attention to the large total consumption, which is 183.4 W*h, and adding most equipment will result in Operating Expense (OPEX).
Because of the low data center load and low resource usability, a new cloud simulation is performed at a load close to the maximum. That's why we'll increase users three times.

Figure 10 shows the structure of the system as well as the energy consumed by the above-defined number of devices required to handle the tasks. According to the energy consumption per unit, each layer shows that it is the largest in the aggregation network - 102.8 W*h, exactly twice that of the upper level (core), which is 51.4 W*h. The access commutators consume 5.5 W*h, and servers 70.8 W*h, meaning the distribution network consumes the most energy from all devices. Total consumption is 229.7 W*h. Tasks submitted for processing are 16,503. The total number of servers in this system is 24. Consumers are increased three times.

Figure 10. Energy consumption and structure of the cloud system at high load

Figure 11 shows the traffic showing the data center load and shows the average load of the processors of the different servers from the simulated cloud. Figure 12 shows the number of tasks performed by the virtual machines in the cloud and the loading of the lines from the aggregation to core network. As can be seen, the set network (cloud) handles the tasks.

In order to simulate cloud network at high load it is decided to increase the users. As can be seen, the data center load is about 80%. Targeted Servers are just over 20 out of 24, which also leaves a similar reciprocal reserve of about 15%. The number of completed tasks from virtual machines is great. It was a well-balanced cloud and achieved very good ratio of price / quality.
Figure 11. Load the data center and load the processors on the computing servers

Figure 12. Number of tasks performed by virtual machines and loading of the lines from the distribution to the backbone

Now let's ask the following question, the cloud has the task, but are there dismissed or unprocessed tasks? The next figure (13) gives us an answer to this. As you can see, the average number of failures for a virtual machine is zero. We must note that the task is considered unsuccessful when it has not been processed for a specified time (5 seconds). Apart from the fact that all the tasks are processed, the time-window for processing each task (this was set by the software developers) was also followed.
After the simulation researches and analysis, it was identified that the cloud works perfectly and without losses. Green Cloud product provides sufficient information as to the work of the cloud, and the power consumption. Studied cloud system is suitable for building of a private cloud.

For future work should be focus on study the cloud structures with larger capacity, usage of different software to verify the results and to simulate the structures of the collision in them (cloud computing storms).

![Failed Tasks by VM](image)

**Figure. 13. Average value of failed tasks**

### Conclusion

An overview of the cloud-based services available to users, departments and organizations is made. Attention is also paid to the different cloud architectures. The following CloudSim, OCT, GreenCloud cloud simulation software and tools are presented. A short analysis and highlighting the main advantages and disadvantages is made. A program for simulations (GreenCloud) was selected and such were made: load the data center and CPU computing servers, the number of completed tasks of virtual machines and load lines of distribution to the backbone network, average failed tasks. The obtained results have been analyzed, which lead to the conclusion that the set cloud is dealing with the stated tasks and the load is far from the possible maximum, the distribution network is the busiest, there is a slight load on the connections and a large capacity in the reserve specific tasks for the particular cloud model. We have a significantly higher load simulation to determine the maximum capacity we can use the cloud to achieve maximum quality cost results.
Bibliography


Georgi Georgiev – Department of Telecommunications at University of Telecommunications and Post, Sofia, Bulgaria, tel.: +359889813894 e-mail: goshko.georgiev@gmail.com