Some Techniques for Increasing Energy Efficiency in a Cloud Data Center

Ezmerina Kotobelli¹, Ylli Seitaj², Blerina Zanaj³, Igli Tafaj⁴ and Elma Zanaj⁵

¹,²,⁴,⁵Faculty of Information Technology, Polytechnic University of Tirana, Tirana, Albania
¹ekotobelli@fti.edu.al, ²ylliseitaj@gmail.com, ⁴itafaj@fti.edu.al, ⁵ezanaj@fti.edu.al
³Agriculture University of Tirana, Tirana, Albania
bzanaj@ubt.edu.al

Abstract
Today the cost of consumed energy from computer systems occupies a considerable part of the total operating cost of a corporation. Consequently there is a need for finding and implementing techniques for saving consumed energy by these systems as much as possible. In this paper will be shown the efficiency of Dynamic Voltage and Frequency Scaling (DVFS) and VM migration techniques for energy saving in a cloud data center.

Keywords: Energy saving, Cloud computing, Techniques, Virtualization, DVFS.

1. Introduction
One of the challenges of today is not only to find ways to produce green energy or building technologies to realize this but also to find methods for saving as much as possible the energy consumed. In this paper will be described two methods for energy saving in a cloud data center, by using a well-known simulator called CloudSim. Through CloudSim, will be implemented a set of experiments in order to show clearly what amount of energy can be saved using each of the techniques and also will be done a comparison between them in terms of saved energy.

2. Cloud Computing
Cloud Computing is often described as a stack, as a response to the broad range of services built on top of one another under the moniker “Cloud.” The generally accepted definition of Cloud Computing comes from the National Institute of Standards and Technology (NIST). Cloud computing is defined as a model for enabling ubiquitous, 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 cloud model is composed of: five essential characteristics, three service models, and four deployment models.

2.1 Essential Characteristics
NIST also offers up several characteristics that it sees as essential for a service to be considered “Cloud” [1].

- **On-demand self-service.** A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

- **Broad network access.** Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g., mobile phones, tablets, laptops, and workstations).

- **Resource pooling.** The provider’s computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of
resources include storage, processing, memory, and network bandwidth.

- **Rapid elasticity.** Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

- **Measured service.** Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

## 2.2 The cloud infrastructure

A cloud infrastructure is the collection of hardware and software that enables the five essential characteristics of cloud computing. The cloud infrastructure can be viewed as containing both a physical layer and an abstraction layer. The physical layer consists of the hardware resources that are necessary to support the cloud services being provided, and typically includes server, storage and network components. The abstraction layer consists of the software deployed across the physical layer, which manifests the essential cloud characteristics. Conceptually the abstraction layer sits above the physical layer.

## 2.3 Service Models

The diagram show in Fig.1 is obtained from [2] and depicts the Cloud Computing stack. It shows three distinct categories within Cloud Computing: Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS).

A very simplified way of differentiating this service model of Cloud Computing is as follows:

- **SaaS** applications are designed for end-users, delivered over the web. The capability provided to the consumer is to use the provider’s applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user specific application configuration settings.

- **PaaS** is the set of tools and services designed to make coding and deploying those applications quick and efficient. The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

- **IaaS** is the hardware and software that powers it all – servers, storage, networks, operating systems. The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls).

## 2.4 Deployment Models

**Private cloud.** The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers.
(e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.

**Community cloud.** The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises.

**Public cloud.** The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

**Hybrid cloud.** The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds) [1].

### 3. Virtualization

Virtualization is a vital technology of Cloud Computing which offers two important features – abstraction and encapsulation [3]. It is about creating an abstract layer between hardware and software. Usually, the virtualization layer is set above the physical layer of the Cloud’s architecture. Virtualization technology is used widely in Cloud Computing data centers owing to the benefits offered, such as utilizing resources, lowering costs, easier management of servers, server consolidation, and live migration of virtual machines [4]. Virtualization is mostly used in Cloud Computing platforms as means to optimize resource usage [5]. Through virtualization, the number of hardware resources used in Clouds can be reduced to minimize the capital cost as well as the cost of power consumption and cooling systems [6]. For instance, through server consolidation, multiple (virtual) servers can be allowed to run simultaneously on a single physical server. Also, live migration of the virtual machine to the not fully utilized physical servers would allow more and more physical servers to be turned off, which would lead to better achievement of energy efficiency for data centers. Furthermore, visualization in Cloud Computing can offer dynamic configurations for different application’s resource requirements, and aggregate these resources for different needs. Additionally, it can improve responsiveness by monitoring, maintaining and provisioning resources automatically [7].

### 4. Some Techniques for Saving Energy

In this section will be described two techniques for saving energy. These techniques are VM migration and DVFS.

VM migration has to do with the migration of virtual machines in as little as possible physical machines. Smaller number of physical machines in operation led to smaller amount of consumed energy.

DVFS is an abbreviation for Dynamic Voltage and Frequency Scaling. In order to give a simple explanation will separate it in two parts: Dynamic Voltage Scaling and Dynamic Frequency Scaling. In Dynamic Voltage Scaling, the voltage is increased (overvolting) or decreased (undervolting) depending on the circumstances. Undervolting is used in order to conserve power. Overvolting is used in order to increase computer performance. In Dynamic Frequency Scaling, the frequency of a microprocessor can be adjusted automatically on the fly for energy saving.

### 5. Implementation and Evaluation of Experiments

There are three types of experiment scenarios that will be implemented. In the first scenario will not be used any of the techniques for saving energy. In the second one, each experiment will be executed by using the DVFS technique. In the third scenario, for each experiment will be implemented another techniques known as VM migration. Each scenario will be executed as a set of experiments where each single one of them will have different values of host, VM and Cloudlet parameters. Initially they will have the values 10, 20 and 20 respectively. These will be
the minimal values which will be increased. The maximal values for each scenario will change from one experiment to another.

Scenario 1: without any technique for saving energy. In this scenario are implemented 15 experiments, where each one is executed 7 times and then is calculated the average value of them. It is not used any technique for saving energy. Fig.2 summarizes the entire scenario.

Scenario 2: using DVFS technique. In this scenario will be used DVFS technique for saving energy. The amount of experiments, the number of tests for each experiment and the respective values of Hosts, VMs and Cloudlets will be equal with the selected values of the first scenario. Fig.3 summarizes the entire scenario.

Scenario 3: With VM migration. In this scenario will be used VM migration as a technique to save energy. CPU utilization will not exceed 80% in order to elude the performance degradation of the system. Fig.4 shows the results for this scenario. Experiment 6 to 15, cannot be executed with this technique due to the amount of hosts bandwidth.

A comparison between three scenarios. In order to give a clear view of contribution of each scenario in saving energy, the first five experiments for each scenario are presented at Fig. 5.

6. Conclusions
As it is already shown, the implementation of saving energy techniques contributes a lot in the reduction of consumed energy in a cloud data center. VM migration technique is more efficient in saving energy compared with DVFS.
On the other end, VM migration is limited in the number of experiments that may be executed compared with DVFS technique. As is shown, in each technique the first five experiments are executed but when it comes to the execution of experiment 6 (number of Hosts = 320, number of VMs = 640 and number of Cloudlets = 640) and so on, where the number of Hosts, VMs and Cloudlets is increased experiment after experiment, with the VM Migration, the experiments (6 and up) cannot be executed due to Hosts bandwidth. This problem is not shown using DVFS technique.

References