ANALYSIS OF ENERGY CONSUMPTION IN CLOUD DATA CENTER USING GREEN CLOUD SIMULATOR

***Vakula Rani Gumpina**
Associate Professor, CMR University, Bangalore

ABSTRACT

Cloud computing characterizes a novel and potential model for managing and providing ICT resources to remote users based on pay-as-you-go model. Data centers are popular for hosting the cloud computing applications which consumes tremendous amount of energy, contributing to high operational costs and carbon footprints to the environment[2]. The energy dearth, carbon emissions and global warming leads to concerns towards the power consumption of data centers and [3] the analysis of their energy consumption is extremely important. Therefore, we need Green Cloud Computing Solutions to increase energy savings and to reduce CO2 emission. This paper presents the analysis of cloud datacenter energy saving using virtualization and consolidation technologies . The green cloud simulator is used to estimate the energy consumed by data center core components such as servers, network switches, and access links. The simulated results obtained for various-tiers, and virtualization techniques such as DVFS (Dynamic voltage frequency scaling) ,dynamic power management (DPM) , and dynamic shutdown are presented with and without utilizing power reduction schema. This analysis is significant for users and organizations that are looking at Cloud computing as a solution for their administrative, infrastructural and management problems.

Keywords: *Energy efficiency, DVFS (Dynamic voltage frequency scaling), dynamic power management (DPM), Cloud computing, energy consumption.*

1. Introduction

Over the recent years there has been an alarming rise of high speed networks, resulting in massive

E-commerce transactions and web queries a day. This demand is handled through large-scale datacenters at Google, Amazon, eBay, and Yahoo etc. which host numerous servers, storage and network infrastructure. Cloud Computing has rapidly changed this proprietor - based system to payment - oriented system by providing on demand scalable infrastructure and services. Cloud computing benefits provide on-demand computing resources, faster and cheaper software development capabilities and offers compute power to organizations. Financial companies are able to maintain dynamic information about hundreds of clients. Clouds are viewed as potential market opportunity, according to IDC (International Data Corporation) report [4], the global IT Cloud services. Engrossed by growth prospects, Online businesses (Amazon, eBay, Salesforce.com), hardware vendors (HP, IBM, Cisco), telecom providers (AT&T, Verizon), software firms (EMC/VMware, Oracle/Sun, Microsoft) are investing in huge amount of capital for Cloud datacenters.

Cisco Systems forecasts that annual datacenter traffic will reach 6.6 zettabytes by the end of 2016 with a compound annual growth rate (CAGR) of 31% from 2011 to 2016 (Figure. 1). In 1992, daily Internet traffic was of the order of 100 Gbytes. In 2012, more than 12,000 Gbytes of data crossed the Internet every second. This number is projected to triple by 2017.

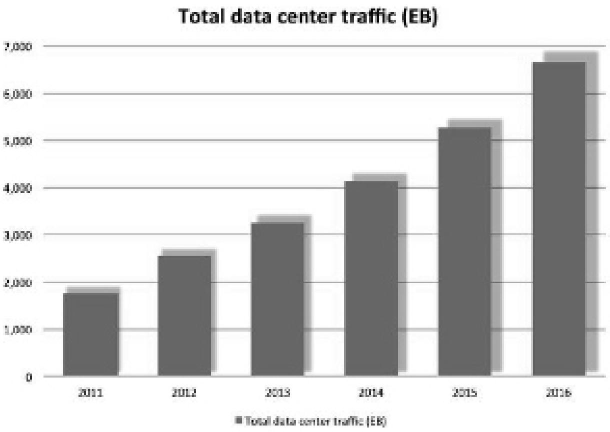

Fig 1: Data Center Traffic by the end of 2016[8]

Two-thirds of traffic in 2016 is forecasted to be cloud-based, displacing traditional datacenter traffic. While compute workloads on traditional servers will grow by approximately 33%, virtualized workloads on cloud servers are projected to increase by 200%, according to Cisco Systems [8]. The company's study also finds that nearly three-fourths of datacenter traffic is internal to the datacenter, inter-server, or between servers and storage.

Clouds are essentially virtualized datacenters and applications offered as services on a payment basis and require high energy usage for its operation. For a datacenter, the energy cost is a significant component of its operating and up-front costs. According to a report published by the European Union, a decrease in emission volume of 15%–30% is required before year 2020 to keep the global temperature increase below 20C. Thus, energy consumption and carbon emission by Cloud infrastructures has become a key environmental concern.

Studies also show that Cloud computing can actually make traditional datacenters more energy efficient by using technologies such as resource virtualization and workload consolidation. The traditional data centers running Web applications are often provisioned to handle irregular peak loads, which can result in low resource utilization and wastage of energy. Through the use of large shared virtualized datacenters cloud computing can offer large energy savings. The environmental sustainability of cloud computing by analyzing various technologies and mechanism that support this goal. Our analysis is important for users and organization that are looking at Cloud computing as a solution for their administrative, infrastructural and management problems. Finally, we also propose and recommend a Green Cloud framework for reducing its carbon footprint in wholesome manner without sacrificing the quality of service (performance, responsiveness and availability) offered by the multiple Cloud providers.

2. CLOUD COMPUTING

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. NIST defines cloud computing by describing five essential characteristics, three cloud service models, and three cloud deployment models. They are summarized in visual form in figure:2.

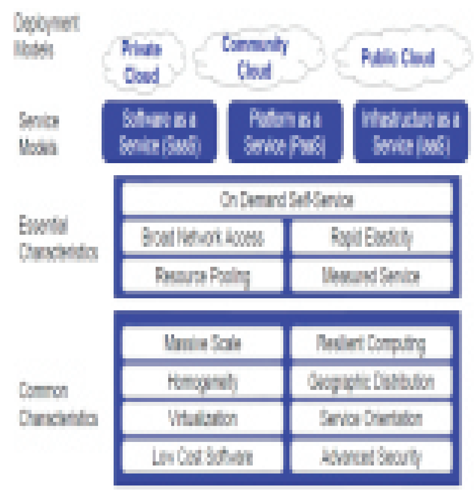

Fig2 :Working Definition of Cloud Computing -NIST

Cloud services exhibit five essential characteristics that demonstrate their relation to, and differences from, traditional computing approaches i) On-demand self-service ii) Broad network access iii) Resource pooling iv) Rapid elasticity v) Measured service. Cloud service classifications and are often referred to as the "SPI Model," where 'SPI' refers to Software, Platform or Infrastructure (as a Service), respectively. i) Software as a Service (SaaS) is the capability provided to use the provider's applications running on a cloud infrastructure. The Service Providers for SaaS are Google Docs, Salesforce.com. ii) Platform as a Service (PaaS) is the capability provided to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider. The Service Providers for PaaS are Google App Engine, Microsoft Azure Services, Amazon Elastic Map Reduce iii) Infrastructure as a Service (IaaS) is the capability provided 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 Service Providers for IaaS are Amazon EC2, Amazon S3, Linode and Rackspace.

When considering cloud computing options, good network connectivity is essential. Once the network is in place, the choices of how to create a cloud environment vary, but most fall into four different categories: private cloud, community cloud, public cloud, and hybrid cloud. Among the best known public cloud services are Amazon, Bing, Google, Yahoo, and others. However, these public clouds may not offer all of the security protections required by cloud users. A private cloud is one built for a specific customer either in the corporate sector or government with greater security protection for the customer's data. A community cloud allows several clients to share the same cloud to perform similar functions such as human resources or email. A hybrid cloud is a combination of two or more of the other three cloud infrastructure options. The benefits of Cloud Computing are i)Minimized capital expenditure ii) Location and device independence iii) Utilization and efficiency improvement iv) Very high scalability v) High computing power.

3. Virtualization

Clouds are essentially virtualized datacenters and uses a number of underlying technologies, services, and infrastructure-level configurations. Datacenter can reduce the energy consumed through server consolidation, whereby different workloads can share the same physical host using virtualization and unused servers can be switched off. Virtualization is a technique of masking the physical hardware and enabling multiple operating systems and multiple applications to run concurrently on a single or clustered physical machines. This is performed using a Hypervisor or Virtual Machine Monitor (VMM) which lies in between the hardware and the Operating System (OS). Then one or more virtualized OSs can be started concurrently, Virtualized servers allow for increased utilization on a common hardware platform through the use of hypervisors that permit multiple OS instances ("guest OSs") and their associated applications.

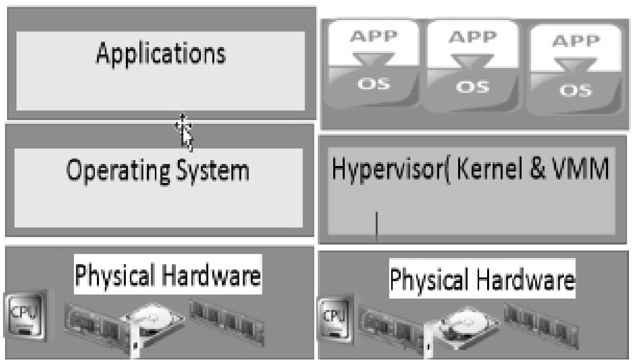

Fig 3: Server Architecture- Traditional VS Virtual

Virtual machines are deployed to execute incoming jobs. The file server provides virtual machine templates and all typical jobs are preconfigured in virtual machine templates. When a job arrives at the head node of the cluster, a correspondent virtual machine is dynamically started on a certain compute node within the cluster to execute the job.

4. Data Centre Architecture

A cloud datacenter could comprise of many hundreds or thousands of networked computers with their corresponding storage and networking subsystems, power distribution and conditioning equipment, and cooling infrastructures. Due to large number of equipment, datacenters can consume massive energy consumption and emit large amount of carbon. This creates a challenge in the design of interconnected network architecture and the set of communication protocols. Given the scale of a data center, the conventional hierarchical network infrastructure often becomes a bottleneck due to the physical and cost driven limitations of the used networking equipment. Three-tier data center architectures are the most common nowadays. They include: Access, Aggregation and Core layers. The availability of the aggregation layer facilitates the increase in the number of server nodes while keeping inexpensive Layer-2 (L2) switches in the access network, which provides a loop-free topology. Because the maximum number of ECMP paths allowed is eight, a typical three tier architecture consists of eight core switches. Such architecture implements an 8-way ECMP that includes 10 GE Line Aggregation Groups (LAGs), which allow a network address several links and network ports with a single MAC address

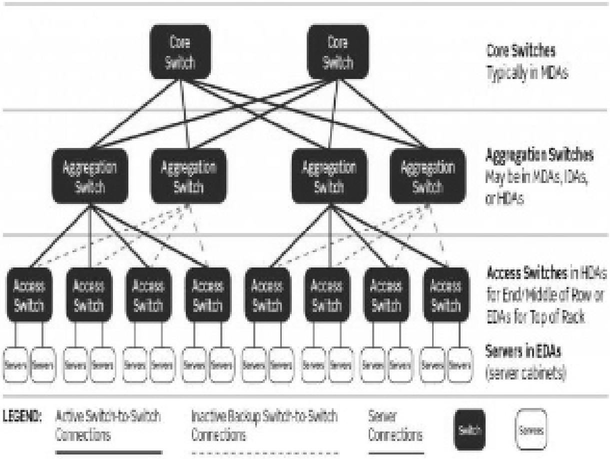

Fig 4: three-tier data center architecture[7]

The data center energy is utilized by i) computing servers ii) to maintain interconnection links and iii) network equipment operations iv) air-conditioning systems v) dissipates as heat energy. Part of the electricity is wasted in the power distribution system. The efficiency of a data center can be defined in terms of the performance delivered per watt, which may be quantified by the Power Usage Effectiveness (PUE), which describe which portion of the totally consumed energy gets delivered to the computing servers.

5. Green Cloud Simulator

Green Cloud simulator is a open source cloud computing environment for energy aware cloud computing data centers. It is an extension to the network simulator NS2, which is developed for the study of cloud computing environments. The Green Cloud offers users a detailed fine-grained modeling of i) the energy consumed by the elements of the data center, such as servers, switches, and links ii) study of workload distributions iii) the packet-level simulations of communications. Figure 5 presents the structure of the Green Cloud Architecture[15].

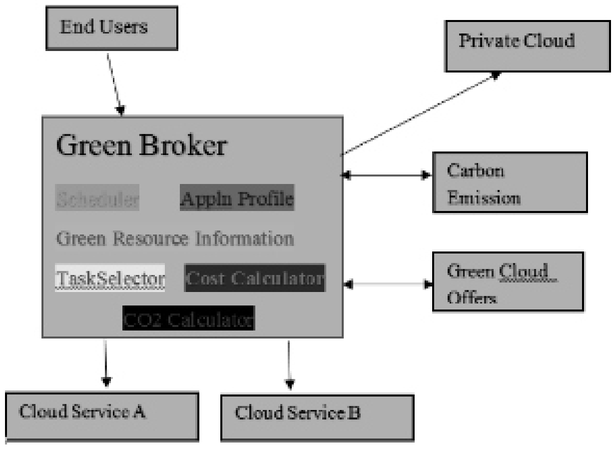

Fig 5: Green Cloud Architecture

Servers are the main components of a data center that are responsible for task execution. In Green Cloud, the server components implement single core nodes that have a preset on a processing power limit in MIPS (million instructions per second) or FLOPS (floating point operations per second), associated size of the memory/storage resources, and contain different task scheduling mechanisms ranging from the simple round-robin to the sophisticated DVFS- and DNS-enabled The servers are arranged into racks with a Top-of-Rack (ToR) switch connecting it to the access part of the network. The power model followed by server components is dependent on the server state and its CPU utilization. An idle server consumes about 66% of energy compared to its fully loaded configuration[7].This is due to the fact that servers must manage memory modules, disks, I/O resources, and other peripherals in an acceptable state. Then, the power consumption linearly increases with the level of CPU load. As a result, this model allows implementation of power saving in a centralized scheduler that can provision the consolidation of workloads in a minimum possible amount of the computing servers.

In Green Cloud, the energy models are implemented for every datacenter element (computing servers, core and rack switches). Moreover, due to the advantage in the simulation resolution, energy models can operate at the packet level as well. This allows updating the levels of energy consumption whenever a new packet leaves or arrives from the link, or whenever a new task execution is started or completed at the server.

6. Performance evaluation

This section presents case study simulations of an energy-aware data center for three-tier (3T), and three-tier high-speed (3Ths) architectures. For comparison reasons, we fixed the number of computing nodes to 1536 for all three topologies, while the number and interconnection of network switches varied. Table:1 Summarizes the main simulation setup parameters. The core switches are connected to the access network directly using 1 GE links (referred as C2–C3) and interconnected between them using 10 GE links (referred as C1–C2). The 3Ths architecture mainly improves the 3T architecture with providing more bandwidth in the core and aggregation parts of the network. The bandwidth of the C1–C2 and C2–C3 links in the 3Ths architecture is ten times of that in 3T and corresponds to 100 GE and 10 GE, respectively. The availability of 100 GE links allows keeping the number of core switches as well as the number of paths in the ECMP routing limited to 2 serving the same amount switches in the access. The propagation delay of all the links is set to 10 ns. The workload generation events and the size of the workloads are exponentially distributed. The average size of the workload and its computing requirement depends on the type of task. For CIW workloads, the relation between computing and data transfer parts is chosen to be 1/10, meaning that with a maximum load of the datacenter its servers will be occupied for 100% while the communication network will be loaded for 10% of its maximum capacity. For DIW workloads the relation is reverse.

The workloads arrived to the data center are scheduled for execution using energy aware Green scheduler and Round Robin scheduler. This green scheduler tends to group the workloads on a minimum possible amount of computing servers. In order to account for DIW workloads, the scheduler continuously tracks buffer occupancy of network switches on the path. In case of congestion, the scheduler avoids using congested routes even if they lead to the servers able to satisfy computational requirement of the workloads. The servers left idle are put into sleep mode (DNS scheme) while on the under loaded servers the supply voltage is reduced (DVFS scheme). The time required to change the power state in either mode is set to 100 ms.

Table: 1 : Simulation Setup parameter

DC Architecture	**Three Tier**		**Three Tier High Speed**	
Scheduling Algorithm	*Round Robin*	*Green Scheduling*	*Round Robin*	*Green Scheduling*
Core nodes (C1)	8	8	2	2
Aggregation nodes (C2)	16	16	4	4
Access switches (C3)	64	64	256	256
Servers (S)	1536	1536	1536	1536
Link (C1–C2)	10	10	100	100
Link (C2–C3)	1	1	10	10
Link (C3–S)	1	1	1	1
Avg. Load / Server	0.4	0.4	0.4	0.4
Data center average load	43.6	43.3	43.6	43.4
Simulation time	65.5	65.5	65.5	65.5
Total Tasks	581499	581499	581499	581499
Avg. Tasks / Server	378.6	378.6	378.6	378.6

Table:2 : Simulation Results

DC Architecture	**Three Tier**		**Three Tier High Speed**	
Scheduling Algorithm	Round Robin	Green Scheduling	Round Robin	Green Scheduling
Data Center Power Consumption ((kWh)				
Total Energy	**6808.4**	**6402.9**	**6902.1**	**6896.6**
Switch Energy (Core)	**466.1**	**466.1**	**1033.4**	**1033.4**
Switch Energy (Aggre.)	**932 .1**	**932.1**	**458.5**	**458. 5**
Switch Energy(Access)	**1375.7**	**1375.7**	**1375.7**	**1375.7**
Switch Energy	**4034.5**	**4029**	**4034.5**	**4029**

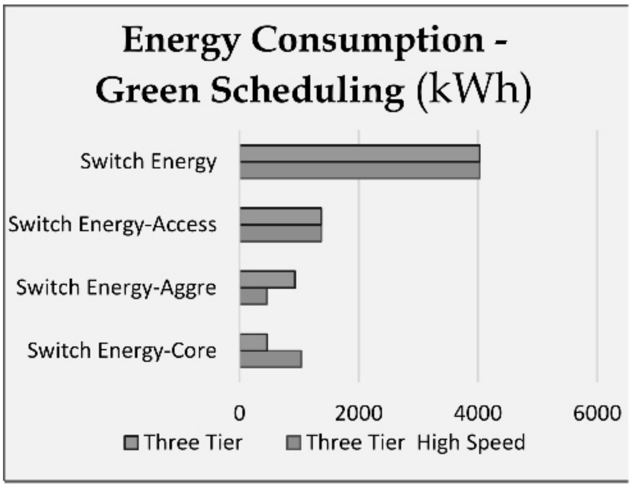

Fig:6- Energy Consumption Green Scheduling

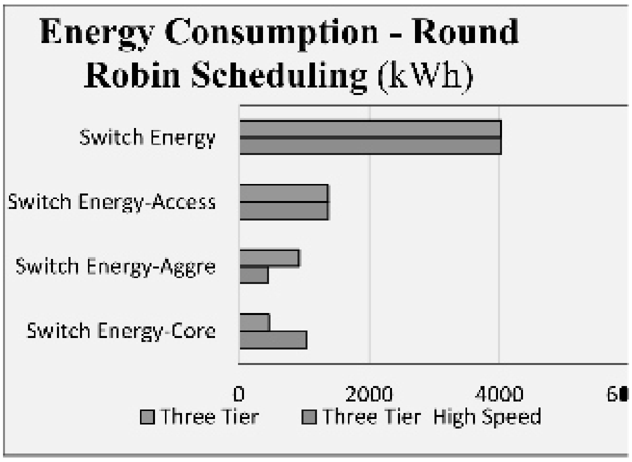

Figure:7- Energy Consumption – RR Scheduling

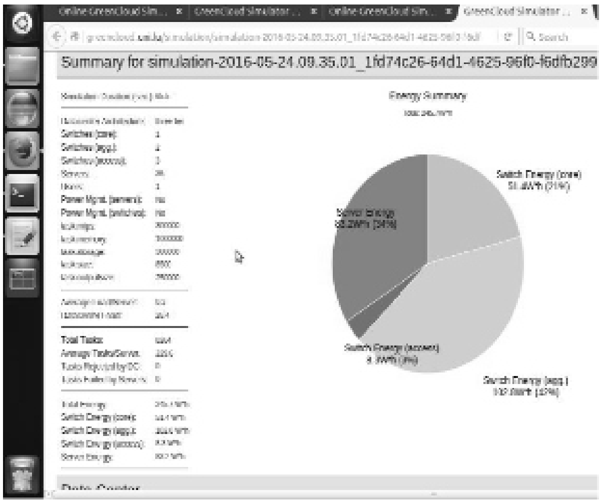

Fig :8 - Summary for Simulation

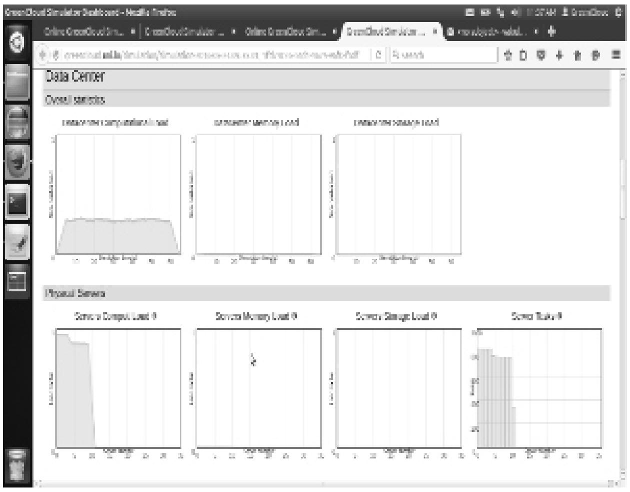

Figure9 : Data Center Overall Statistics

Fig:10- Data Center Network Statistics

Fig : 11a : Data Center Energy Consumption statistics

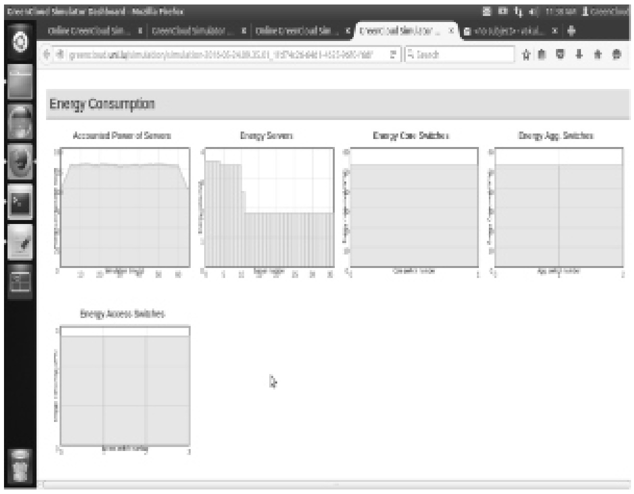

Fig: 11b - Data Center Energy Consumption Statistics

The switches' consumption is almost constant for different transmission rates as most of the power (85–97%) is consumed by their chassis and line cards and only a small portion (3–15%) is consumed by their port transceivers. Switch power consumption values are derived with a twisted pair cable connection considered for the rack switch (C3) and optical multimode fiber for the core (C1) and aggregation (C2) switches. Depending on the employed data center topology, the core and aggregation switches will consume differently. For the 3T topology where the fastest links are 10 G the core and aggregation switches consume a few kilowatts, while in the 3Ths topology where links are of 10 G speed faster switches are needed which consume tens of kilowatts. Table:2 presents simulation results obtained for three evaluated data center topologies with no energy saving management involved for an average load of the data center of 30%. The obtained numbers aim to estimate the scope of the energy-related spending components in modern data centers and define where the energy management schemes would be the most efficient. On average, the data center consumption is improved around 406 kWh using green scheduler when compared with Round Robin Scheduler.

The processing servers share around 70% of total data center energy consumption, while the communicational links and switches account for the rest 30%. The core and aggregation switches together account for 15% of total energy consumption. However, taking into account the requirements for network performance, load balancing, and communication robustness, the obvious choice is to keep core and aggregation switches constantly running possibly applying communication rate reduction in a distributed manner. The data center network accounts for the differences between power consumption levels of different data center architectures. With the respect to the 2T architecture, the 3T architecture adds around 25 kW for aggregation layer which enables the datacenter scale beyond 10 000 nodes. The 3Ths architecture contains fewer core and aggregation switches. However, the availability of 100 G links comes at a price of the Graphs increase per-switch energy consumption. As a result, a 3Ths network consumes more than a 3T network. Figure 7 reports an average distribution of energy consumption ina 3T data center. Figure:12 compares the impact on energy consumption of DVFS, DNS, and DVF Switch DNS schemes applied on both computing several and networking equipment. The results are obtained for balanced tasks loading both computing servers and interconnection network equally for an average system load of 30%.The DVFS scheme alone reduces power consumption to only 96% from the nominal level. Most of the power saving in servers comes from downshifting CPU voltage on the under-loaded servers. However, CPU accounts for 43% of server consumption only. The most effective results are obtained by DNS scheme. It is equally effective for both servers and switches as the most of their energy consumed shows no dependency on the operating frequency. However, in order to utilize DNS scheme effectively, its design should be coupled with the data center scheduler positioned to unload the maximum number of the servers.

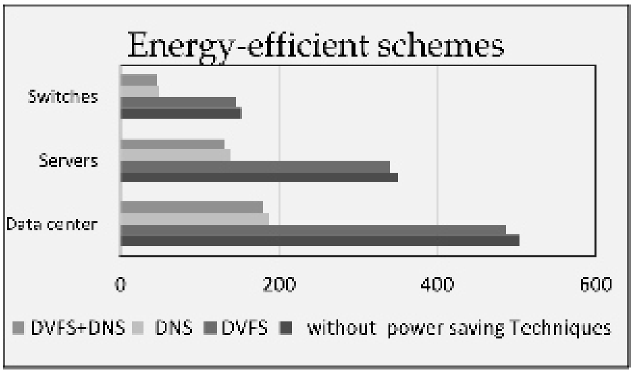

Fig :12 - Comparison of energy-efficient schemes

7. Conclusions

This paper presented a simulation environment for energy-aware cloud computing data centers. GreenCloud is designed to capture details of the energy consumed by data center components as well as packet-level communication patterns between them. The energy consumption analysis and simulation results are presented for, three-tier, and three-tier high-speed data center architectures with GS and RR schedulers. Also, demonstrated applicability and impact from the application of different power management schemes like DVPS and DSN applied on the computing as well as on the networking components.

References

1): "Energy efficient management of data center resources for cloud computing: A vision, architectural elements and Open Challenges" Raj Kumar Buyya, Anton Beloglazov, Jemal Abawajy Proc. of 9th IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2009), Rio De Janeiro, Brazil, May 2009.

2) AnkitaAtrey, Nikita Jain and Iyengar N.Ch. S.N"A Study on Green Cloud Computing " International Journal of Grid and Distributed ComputingVol.6, No.6 (2013) pp.93-102 http://dx.doi.org/10.14257/ijgdc.2013.6.6.08, ISSN: 2005-4262 IJGDC Copyright ⓒ 2 0SERSC .

3) Guérout, Tom, Thierry Monteil, Georges Da Costa, Rodrigo NevesCalheiros, Rajkumar Buyya, MihaiAlexandru, "Energy-aware simulation with DVFS", Simulation Modelling Practice and Theory 39, 2013, pp. 76-91.

4) Data Center Efficiency Assessment Scaling Up Energy Efficiency Across the Data Center Industry: Evaluating Key Drivers and Barriers Issue Paper August 2014 IP:14-08-a

5) Beant Singh Gill, Satinderjitkaur Gill, Prince Jain "Analysis of Energy Aware Data Center using GreenCloud Simulator in Cloud Computing" International Journal of Computer Trends and Technology (IJCTT) – volume 5 number 3 – Nov 2013,ISSN: 2231-2803 http://www.ijcttjournal.org Page 154

6) S. K. Garg and R. Buyya, "Green Cloud computing and Environmental Sustainability", Edited by S. Murugesan and G. R. Gangadharan, Wiley-IEEE Press Ebook (2012), Edition 1, no. 3, pp. 76-87.

7) http://www.datacenterjournal.com : Three-tier data center architecture.

8) http://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/global-cloud-index-gci : Cisco Global Cloud Index: Forecast and Methodology, 2015–2020 ,White Paper, Cisco Public.

9) "EMC IT's Journey to the Private Cloud: A Practitioner's Guide"

10) "Cloud Infrastructure and Services Student Guide", EMC Education Services, May 2, 2013.

11) A Platform Computing Whitepaper. "Enterprise Cloud Computing: Transforming IT." Platform Computing, pp6, 2010.

12) Uddin, M. and Rahman, A.A. (2012) 'Energy efficiency and low carbon enabler green IT framework for data centers considering green metrics', Renewable and Sustainable Energy Reviews, Vol. 16, pp.4078–4094.

13) Chen, F. F., Schneider, J., Yang, Y., Grundy, J., and He, Q. 2012 An energy consumption model and analysis tool for Cloud computing environments. Proceedings of the First International Workshop on Green and Sustainable Software 2012. on Software Engineering, June 2-9, Zurich, Switzerland, IEEE, Ed. IEEE Computer Society, 45-50.

14) G. Procaccianti, P. Lago, G. Lewis, Green architectural tactics for the cloud, in: Proceedings of the 11th Working IEEE/IFIP Conference on Software Architecture (WICSA), 2014.

15) Uddin M & Rahman A A, Techniques to Implement Green Data Center to achieve Energy Efficiency & Reduce Global Warming Effects, Inter J Global Warm, 3(4) (2011) 372- 389.