Scientific Journal of Impact Factor (SJIF): 4.72
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406
International Journal of Advance Engineering and Research Development
Emerging Trends and Innovations in Electronics and Communication Engineering - ETIECE-2017
Volume 5, Special Issue 01, Jan.-2018 (UGC Approved)

Green Scheduling Algorithm for Cloud Centric Data Centers

Idris Afzal Shah [1], Syed Arshid Ahmad Simnani [2], Hamid Hussain Haqani [3], Faisal Rasheed Lone [4]

[1]*Department of CSE, University of Kashmir*
[2]*Assistant Programmer NIC*
[3]*Department of Computer Science, University of Kashmir*
[4]*Department of CSE, University of Kashmir*

ABSTRACT- *In cloud computing large number of data centers encompassing many servers are geographically distributed on the globe and connected in a network. These provide round the clock access to almost whole population on the globe.To keep all the servers of a cloud operational a considerable amount of energy is consumed , which presently accounts for 10% of the total operational cost and is expected to rise to 50% in the coming years.Statistically it has been found that on an average only 30% of the resources of a cloud are utilized at any given instant of time. Most of the resources therefore remain underutilized / unutilized. These resources which remain underutilized / unutilized consume a lot of energy as an idle server consumes about two-third of the energy of the peak load. This is because of the fact that the servers must manage modules, disks, I/O resources and other peripherals in acceptable state. Therefore, a lot of energy gets wasted.In this thesis a green scheduling algorithm is proposed which penalizes low server utilization and favors high server utilization. It concentrates the workload on a minimum set of servers and maximizes the number of servers that can remain in idle state. Finally servers that are in the idle state are shut down and put in sleep mode using Dynamic Power Management. As a result, significant reduction in energy consumption is achieved.*

Keywords- *Cloud Computing, Data Center, Green Scheduling, DPM.*

I. CLOUD COMPUTING

The cloud computing is being defined in numerous ways, below are given some precise definitions [1][2]:Cloud computing is a computing paradigm shift where computing is moved away from personal computers or an individual server to a 'cloud' of computers.Cloud computing is Internet-based ("cloud") development and use of computer technology ("computing"). The cloud is a metaphor for the Internet, based on how it is depicted in computer network diagrams, and is an abstraction for the complex infrastructure it conceals. computing is an example of computing in which dynamically scalable and often virtualized resources are provided as a service over the Internet.US National Institute of Standards and Technology (NIST) put forward a definition that has now become widely accepted as the closest that the industry has a definitive answer. The NIST definition is as follows [3]."Cloud computing is a model for enabling, convenient, on-demand network access to a shared pool of configurable computing resources (eg. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction".

The following are some of the notable challenges associated with cloud computing:

a) Security and Privacy
b) Lack of Standards
c) Continuously Evolving
d) Compliance Concerns
e) Lack of interoperability and Portability
f) Energy Concerns[4][5][6]

II. THREE TIER ARCHITECTURE

Three-tier data center architectures are the most common nowadays. They include: (a) access, (b) aggregation, and (c) core layers as presented in Fig.4.2. The availability of the aggregation layer facilitates the increase in the number of server nodes (to over 10,000 servers) 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 (only four are presented in Fig. 4.2).Such architecture implements an 8-way ECMP that includes 10 GE Line Aggregation Groups (LAGs) [11], which allow a network client to address several links and network ports with a single MAC address. While the LAG technology is an excellent methodology to increase link capacities, its usage has several fundamental drawbacks that limit network flexibility and performance. LAGs make it difficult to plan the capacity for large flows and make unpredictable in it case of a link failure. In addition, several types of traffic patterns, such as ICMP and broadcast are usually routed through a single link only. Moreover, full mesh connectivity at the core of the network requires considerable amount of cablings.

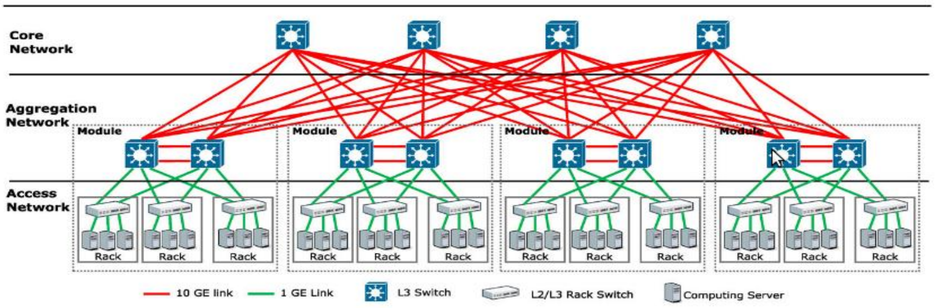

Figure 1. Three Tier Architecture

III. PROBLEM DEFINITION

The operation of large geographically distributed data centers requires considerable amount of energy that accounts for a large slice of the total operational costs for cloud data centers. Gartner group estimates energy consumptions to account for up to 10% of the current data center operational expenses (OPEX), and this estimate may rise to 50% in the next few years [6]. However, computing based energy consumption is not the only power-related portion of the OPEX bill. High power consumption generates heat and requires an accompanying cooling system that costs in a range of $2 to $5 million per year for classical data centers. The workload of a data center fluctuates on the weekly (and in some case on hourly basis), it is a common practice to overprovision computing and communicational resources to accommodate the peak (or expected maximum) load. In fact, the average load accounts only for 30% of data center resources [6]. An idle server consumes about 2/3 of the peak load [10].

IV. DYNAMIC POWER MANAGEMENT

Dynamic power management (DPM) is a design methodology that dynamically reconfigures an electronic system to provide the requested services and performance levels with a minimum number of active components or a minimum load on such components.DPM encompasses a set of techniques that achieve energy-efficient computation by selectively turning off(or reducing the performance of) system components when they are *idle* (or partially unexploited). DPM is used in various forms in most portable (and some stationary) electronic designs. The fundamental premise for the applicability of DPM is that systems (and their components) experience non uniform workloads during operation time. Such an assumption is valid for most systems, both when considered in isolation and when internetworked. A second assumption of DPM is that it is possible to predict, with a certain degree of confidence, the fluctuations of workload. Workload observation and prediction should no consume significant energy[7].Dynamic power managers can have different embodiments, according to the level (e.g., component, system, network) where DPM is applied and to the physical realization style(e.g., timer, hard-wired controller, software routine). Typically, a *power manager* (PM) implements a control procedure based on some observations and/or assumptions on the workload .The control procedure is often called *policy*. An example of a simple policy, ubiquitously used for laptops and palmtops, is the *timeout* policy, which shuts down a component after a fixed inactivity time, under the assumption that it is highly likely that a component remains idle if it has been idle for the timeout time[8[9][10].

V. PROPOSED METHODOLGY

As the average load accounts only for 30% of data center resources, it means that we can put the rest of the 70% of resources in sleep mode. In order to reduce achieve this we propose energy aware scheduling algorithm that

a) Concentrate the workload on a minimum number of computing resources
b) Maximize the amount of resources that remain in 'idle state'
c) Applying DPM/DNS on the resources in 'idle state' so as to put them in 'sleep mode'.

In order to concentrate the workload on a minimum set of computing resources/servers we propose a scheduler named 'Green Scheduler' .The scheduler is green in the sense that it will transfer the user workload on a minimum set of servers within their computing capacity. Then using DPM we will dynamically shut down the servers that remain idle and put them in a low power state to achieve efficiency in terms of energy consumption.

INTRODUCTION TO GREENCLOUD

GreenCloud is developed as an extension of a packet-level network simulator Ns2 [12]. Unlike the existing cloud computing simulator CloudSim [13], GreenCloud extracts, aggregates, and makes information about the energy consumed by computing and communication elements of the data center available on an unprecedented fashion. In particular, a special focus is devoted to accurately capture communication patterns of currently deployed and future data center architectures.

VI. RESULT AND ANALYSIS

Simulations for three tier and three tier high speed architectures are presented. For comparison purposes the numbers of servers are fixed to 1536 while number and interconnection of network switches is varied. Three scenarios are considered:

Round Robin Scheduler (distributes the workload evenly across all the servers)
Green Scheduler
Green Scheduler with DNS (dynamic network shutdown)

TABLE 1: SIMULATION SETUP PARAMETERS

Parameter	Three-Tier
Core nodes(C1)	8
Aggregation nodes(C2)	16
Access switches(C3)	512
Servers(S)	1536
Link(C1-C2)	10GE
Link(C2-C3)	1GE
Link(C3-S)	1GE

DATA CENTER

Parameter	Three-Tier
Core nodes(C1)	8
Aggregation nodes(C2)	16
Access switches(C3)	512
Servers(S)	1536
Link(C1-C2)	10GE
Link(C2-C3)	1GE
Link(C3-S)	1GE

TABLE 2: THREE TIER (ROUND ROBIN SCHEDULER)

Scheduler	**Architecture**	**Energy saving**
Round Robin	**Three tier**	**No**

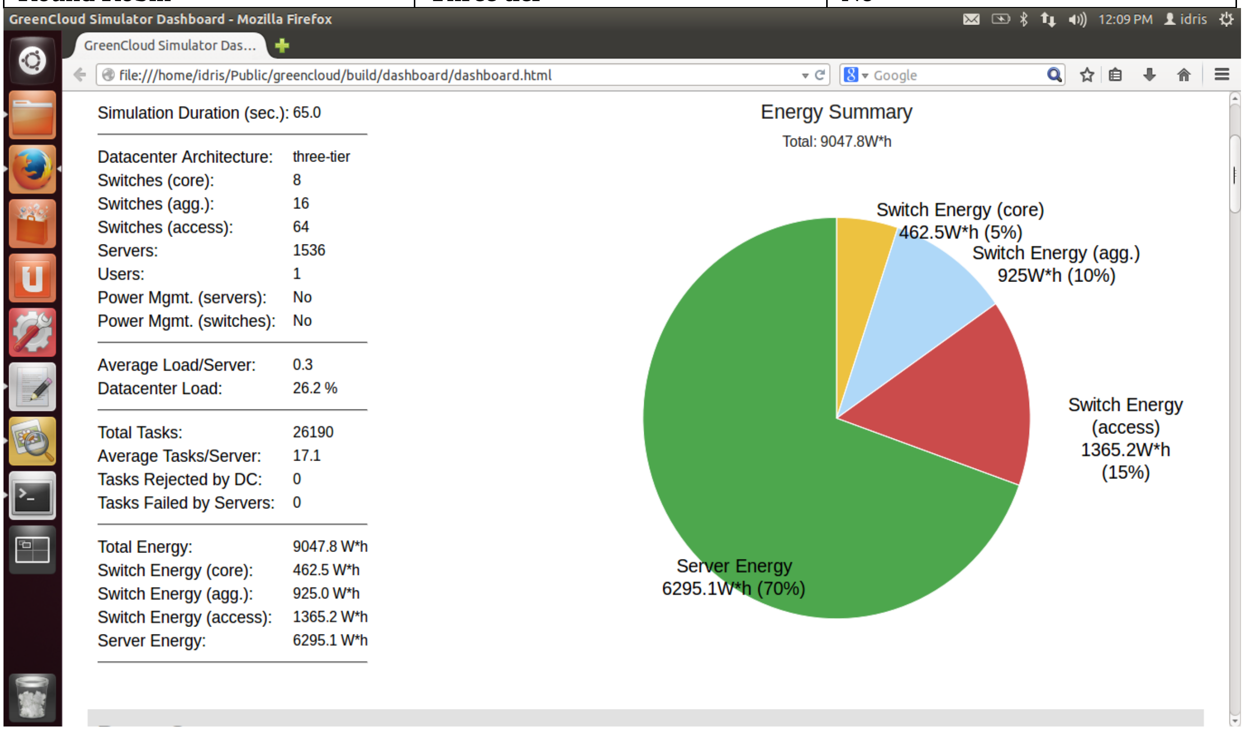

Figure 1. Energy Consumption(Round Robin)

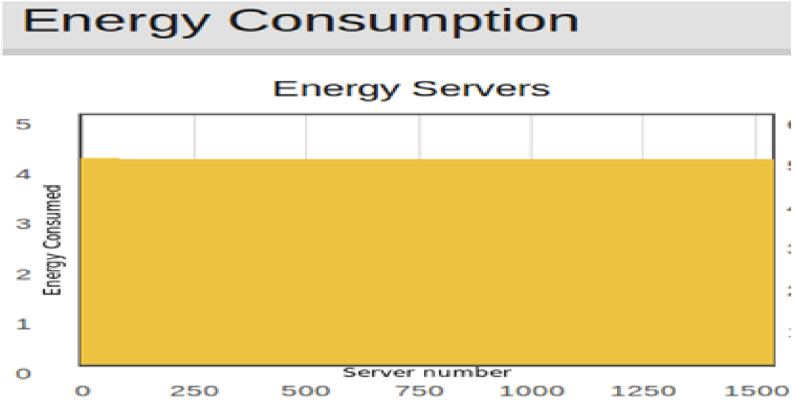

Figure 2. Energy Consumption(Servers)

TABLE 3 : THREE TIER (GREEN SCHEDULER)

Scheduler	Architecture	Energy saving
Green Scheduler	Three tier	No

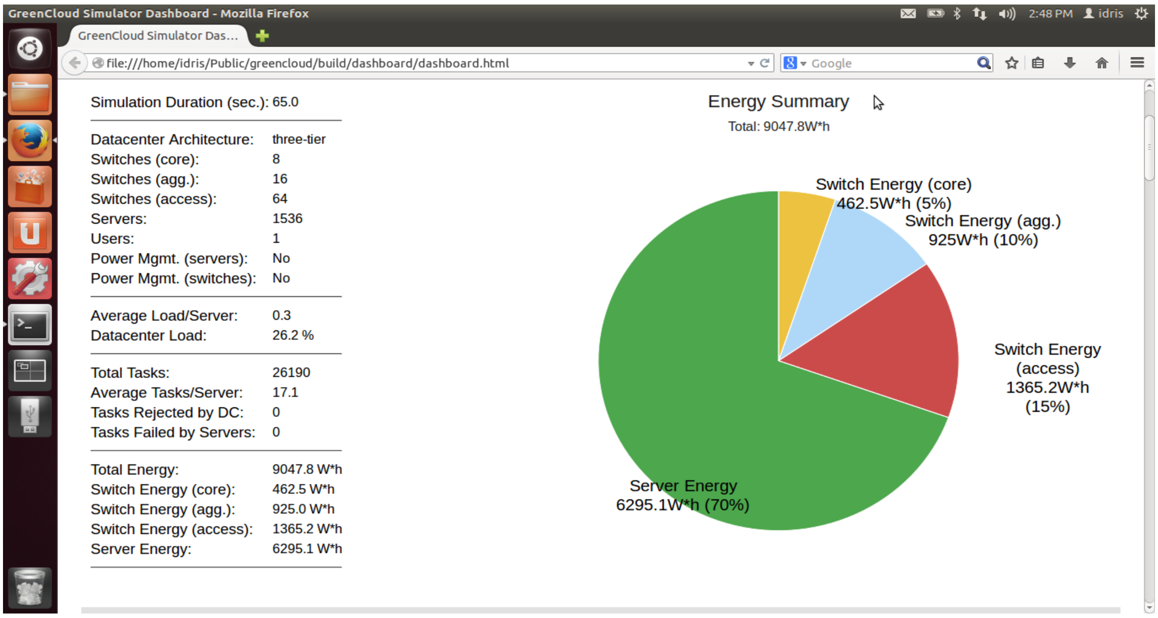

Figure 3. Energy Consumption(Green Scheduler)

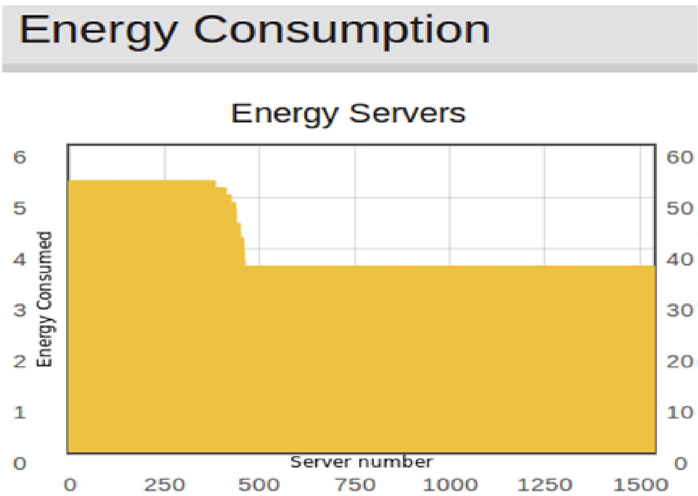

Figure 4. Energy Consumption(Servers)

TABLE 4: THREE TIER (GREEN SCHEDULER WITH DNS)

Scheduler	Architecture	Energy saving
Green Scheduler	Three tier	DNS

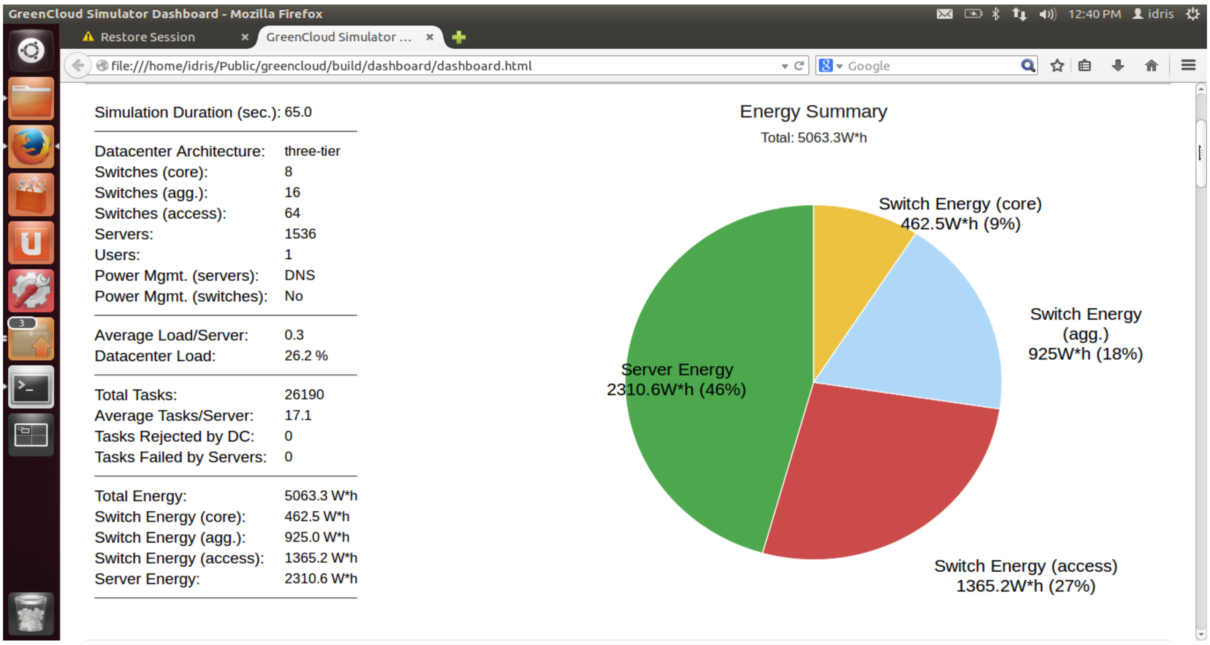

Figure 5. Energy Consumption(Green Scheduler with DNS)

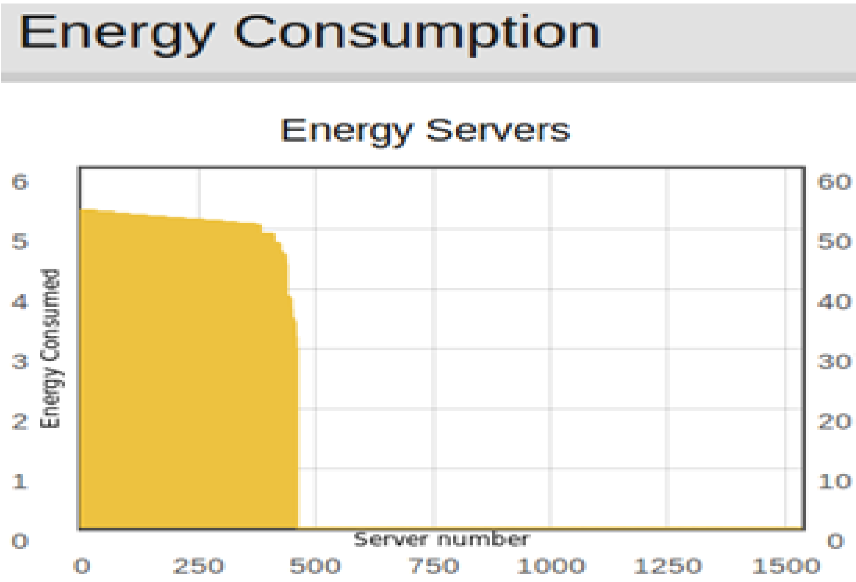

Figure 6. Energy Consumption(Servers)

VIII. SUMMARY

Parameter	Three Tier
Data Center	9047.8
Servers	6295.1
Switches(Core)	462.5
Switches(Aggregation)	925
Switches(Access)	1365.2

Parameter	Three Tier(no energy saving)	DNS
Data Center	9047.8	5063.3
Servers	6295.1(70%)	2310.6(46%)

IX. CONCLUSION AND FUTURE WORK

The overall energy consumption drops by 44% by using the proposed scheme.
The future work will focus on further refinement of energy model and also modelling and designing virtual machine migration processes. Also the effects of variable task size and dense traffic loads in the interconnection network will be explored.

REFERENCES

[1] www.dialogic.com, White Paper on ‘Introduction to Cloud Computing’
[2] http://en.wikipedia.org/wiki/Cloud_computing
[3] csrc.**nist**.gov/publications/**nist**pubs/800-145/SP800-145
[4] http://techankit.com/what-is-cloud-computing
[5] T. Dillon, Wu Chen,E Chang, “Cloud Computing: Issues and Challenges” 24th IEEE International conference on Advanced Information Networking and Applications(AINA),2010 ,Perth,WA.
[6] Gartner Group, available at: http://www.gartner.com/
[7] Benini L, Bogliolo A, De Micheli G (2000) A survey of design techniques for system-level dynamic power management. In: IEEE transactions on very large scale integration (VLSI) systems, 8(3), pp 299–316
[8] Mastroleon L, Bambos N, Kozyrakis C, Economou D (2005) Automatic power management schemes for internet servers and data centers. In: IEEE global telecommunications conference, 2005, GLOBECOM ’05, p 5
[9] B. Khargharia, S. Hariri, F. Szidarovszky, M. Houri, H. El-Rewini, S. U. Khan, I. Ahmad, and M. S. Yousif, “Autonomic Power and Performance Management for Large-Scale Data Centers,” IEEE International Parallel and Distributed Processing Symposium (IPDPS), Long Beach, CA, USA, March 2007.
[10] G. Chen, W. He, J. Liu, S. Nath, L. Rigas, L. Xiao, and F. Zhao,“Energy-aware server provisioning and load dispatching for connection intensive internet services,” the 5th USENIX Symposium on Networked Systems Design and Implementation, Berkeley, CA, USA, 2008.
[11] IEEE Std. 802.3ad-2000 (2000) Amendment to carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications-aggregation of multiple link segments. IEEE Press, New York
[12] The Network Simulator Ns2, avialable at: http://www.isi.edu/nsnam/ns/
[13] R. Buyya, R. Ranjan, and R. N. Calheiros, “Modeling and Simulation of Scalable Cloud Computing Environments and the CloudSim Toolkit: Challenges and Opportunities,” 7th High Performance Computing and Simulation Conference, Leipzig, Germany, June, 2009.

