An Energy Efficient Enhanced Load Balance Technique Using Min Min Scheduling Algorithm (E3lbmin)

[1]S.Ananthi, [2]Dr.P.Vaishnavi

[1]Assistant Professor, [2]Assistant professor
[1]National college, [2]Anna University
Trichy, Tamilnadu

ABSTRACT
Cloud computing is considered as a great significance in recent technologies and also in IT sector. The tremendous growth in data produces heavy amount of carbon emission and also consume nearly half of the total energy consumption. Various kinds of scheduling algorithm are implementing to lessen the execution time but the ultimate issue energy consumption not yet considered. Energy aware scheduling algorithm is which is concentrate on both makespan and also in energy consumption. The Proposed methodology provides a novel scheduling algorithm based on the factors workload and job type to predict the makespan and also energy consumption. The motivation of this scheduling algorithm is to achieve the energy efficient green task scheduling and optimize the scheduler that using the sigmoid neural task predictor is the ultimate task for implementation

Keyword – Energy efficient, Load predictor, Job scheduling, carbon emission

1.INTRODUCTION

Nowadays Energy conservancy and performance are the top most fascination for the research community. The abundant energy utilization issue came to the spotlight owing to unintentional advancement of severe environmental changes across the world. The environment accomplished an unimaginable increase in the levels of Green House Gases (GHG) and carbon footprints. Information and Communication (ICT) sector has been detect as their prime emitter [1–6]. The abiding growth of aggregate and various data-intensive applications and services has embellished the energy concern within ICT. The ICT energy demand has provided the factor to not only fulfill the energy requirements but also to develop and implement effective energy efficiency mechanisms. The current research in this direction involves designing of energy efficient algorithms. Many of these algorithms can also optimize Quality of Service (QoS) parameters such as deadline, budget, reliability etc. [7,8]. A number of ICT firms have designed and implemented energy-aware resource scheduling and Load balancing strategies or algorithms that tend to improve the resource utilizations and also achieve energy efficiency. Recently, cloud computing, based on virtualization, has emerged as a prominent technology facilitating for energy efficiency [9,10]. The mechanism of minimizing the energy utilization and carbon emissions control through cloud is termed as Green cloud computing. Additionally, the fast processing of task facilitated by the current job scheduling architecture improvises the resource levels utilization. This paper proposes an “E3LBMin” model for achieving the energy efficiency through cloud computing by fulfilling the scheduler and controller specifications imposed by the users. It focuses on energy efficiency and thus given a prefix “green”. The underlying cloud system in enhanced green computing consists of workload on which tasks are scheduled for operating while optimizing energy utilizations. The result shows that the average carbon food print is reduced 8% from CEFF and 47% from other heuristic algorithms (Non-Carbon efficient). The proposed algorithm consumes less power while compare to other algorithms. The power consumption is achieved by efficient scheduling.

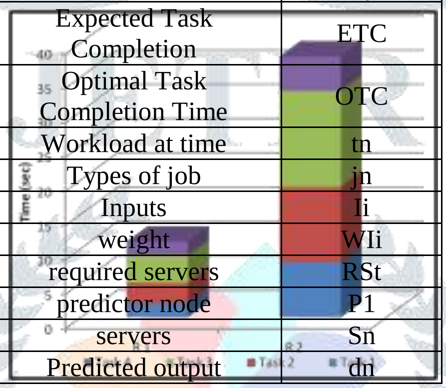
2.RELATED WORK

In this work the authors implement the technique to minimize power utilization and also concentrate on workload to use on/off terminology to nodes. The proposed algorithm monitors the load of resources and works on the basis of Single Point of Failure (SPF) system. It is also applied to the multi tier architecture [7] Chase et al. [8] proposed the technique for load balancing in efficient way. It is applied to all economic frameworks. Energy consumption is achieved by migrating the idle servers to power saving modes (e.g. sleep, hibernation). Moreover, likewise [7], diverse software configurations are not handled, which can be addressed by applying the virtualization technology. Elnozahy et al. analyzed the issue of power efficient management in a single web-application environment. Despite the variable nature of the workload, unlike [8], this scheme is concentrate on the fluctuation of resources.Nathuji and Schwan [10] is based on power consumption on VM Centers and also deals with both hard and soft scaling states on both local and global policies based on legacy. Raghavendra et al. [11] have investigated the problem of power consumption to merge different kind of policies to apply feedback control loop to controllers. It is claimed that, similarly to [10], the approach is different for various kind of workload and also with CPU management. It is mainly coordinate the resources based on power strategy and also with. Kusic et al. [12] have defined the problem of power control in heterogeneous environments using a optimization and Limited Look ahead Control (LLC). Its main aim is to maximize the resource provider’s profit. Kalman filter is applied to predict the future state of the system and perform necessary reallocations. Srikantaiah et al. [13] have studied the problem of scheduling in multi-tiered web-applications to reduce energy consumption. They have found that the energy consumption per transaction results in a ‘‘U’’-shaped curve, and it is possible to determine the optimal utilization point.

Cardosa et al. [14] have proposed an idea for the problem of power-efficient allocation of VMs in virtualized heterogeneous computing environments. They have leveraged the min, max and shares parameters of VMM. Verma et al. [15] have formulated the problem of power-aware dynamic placement of applications in virtualized heterogeneous systems as continuous optimization. Gandhi et al. Gupta et al. [17] have suggested network interfaces, links, switches and routers into sleep modes to save the energy consumed by the Internet. Based on the foundation laid by Gupta et al. [17], a number of research works have been done on the energy-efficient traffic routing by ISPs and applying sleep modes and performance scaling of network devices [18,19]. Chiaraviglio and Matta [20] have proposed cooperation between ISPs and content providers that allows the achievement of an efficient simultaneous allocation of resources and network paths that minimizes energy consumption under performance constraints.Tomas et al. [22] have analyze the problem of scheduling Message Passing Interface (MPI) jobs in Grids considering network data transfers satisfying the QoS requirements. Dodonov and de Mello [23] have proposed an approach to scheduling distributed applications in Grids based on predictions of communication events. They have proposed the migration of communicating processes if minimizing the total execution time. Gyarmati and Trinh [24] have investigated the energy consumption implications of data centers' network architectures. However, optimization of network architectures can be applied only at the data center design time and cannot be applied dynamically.

Table 1. Definitions and Notation

Notation	Description
unmapped tasks	Ui
resources	Rj
Expected Task Completion	ETC
Optimal Task Completion Time	OTC
Workload at time	tn
Types of job	jn
Inputs	Ii
weight	WIi
required servers	RSt
predictor node	P1
servers	Sn
Predicted output	dn



3.PROPOSED MODEL

Energy Efficient Enhanced load Balance Min Min (E3LBMin) Task Scheduling is the process of mapping the unmapped tasks to the appropriate resource . The task scheduling algorithms are exceuted to allocate the best server to achieve best performance like reduced makespan efficient resource utilization. The proposed scheduling algorithm outperforms Enhanced Load Balanced Min-Min algorithm (ELBMM) to achieve energy efficient green task scheduling . E3LBMin has two phases namely identification of minimum required task completion time and energy efficient load balancing the expected completion for a set of unmapped tasks (Ui) with available resources (Rj) Expected Task Completion (ETC) matrix is computed with Un rows Rm columns . The Optimial Task Completion Time (OTC) is the minimum completion time of the task Ui by the resource Rj. OTC matrix is extracted from ETC Matrix.

$$ETC = \begin{bmatrix} 2 X 10^6 & 64 X 10^5 \\ 3 X 10^6 & 88 X 10^5 \\ 4 X 10^6 & 112 X 10^5 \\ 2 X 10^6 & 4 X 10^6 \end{bmatrix} \qquad OTC = \begin{bmatrix} T1 & R1 \\ T2 & R2 \\ T3 & R1 \\ T4 & R1 \end{bmatrix}$$

	R1	R2
Task 1	2x10⁶	64x10⁵
Task 2	3x10⁶	88x10⁵
Task 3	4x10⁶	112x10⁵
Task 4	2x10⁶	4x10⁶

Task and capacity of resources in Milli seconds (R1 100 MIPS, R2 80 MIPS)

Fig.1 Execution time in seconds of tasks in recourses

3.1 ENHANCED GREEN COMPUTING OPTIMIZED MIN- MIN SCHEDULING

The cloud provider consists of scheduler and controller. The scheduler consist s of register and predictor. The register provides the information about total number of servers (n), the status of the server. The status of the server under goes 4 stages each and every server in the cloud energy to any one of the state. Namely ON/OFF, starting and shutting. The predictor system is used to predict the workload at time (t1), the time workload and type of the job are given as input to the predictor system. The workload (item) with j1, j2 …. Jn types of jobs are taken dataset at time

T1, T2, T3 the prediction model has four layers

Layer 1 – Input layer
Layer 2 – workload predictor
Layer 3 – Job type predictor
Layer 4 – output layer

The workload layer of the neural predictor accepts the Input from input layer. The total number of inputs is the sum of the total in time T1, T2 and T3.

3.2 MODEL FORMULATION

The predictor node (PN1, PN2, PN3) are present in layer 1 and (PN4, PN5) are used in layer 2. PN6 used for the output layer 3. The predictor node (PN4, PN5) receives Input from predictor node PN1, PN2 and PN3. The final output layer fives the number of total expected at time T4.

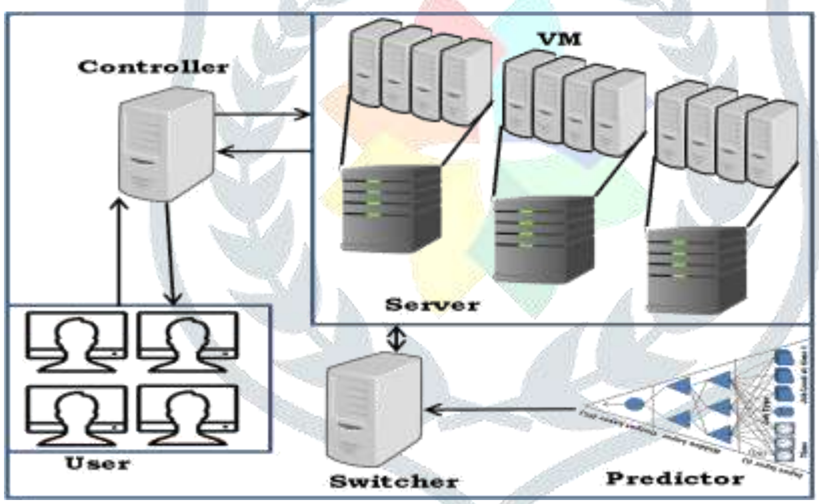


Fig.2 Proposed framework

Step 1: The Data controller plane updates the states of the servers at every time interval (t) to the Register plane.

Step 2: Whenever a scheduler receives resource request from the users (u1, u2,….. , un). It computes the available resources
Where,

Total servers (Sn)

$$S_n=S_{on}+S_{off}+S_{shutting}+S_{restarting}$$

Algorithm 1 Scheduler()

Description:

If working server < S_n
Then,
Get status of the server ()
$S_{on} = w$
$S_{off} = x$

$S_{starting}$ = y
$S_{restarting}$ = z
If (S_{start} > 0)
Then, Wait for the Server to begin
Else
Send request to the controller plane to ON the server Si Send the job to the dedicated servers Si after $TS_{starting}$

Algorithm 2 Data Controller()

Description:

If job arrival rate == server capacity (Sc)
Then, no change in server state
else if (job arrival rate>server capacity)
Job is wait state
else if (job arrival rate<server capacity)

Compute required server for current job arrival rate get job arrival rate at T from the PREDICTOR PLANE to compute required time

if required time < threshold (t)
then, don't change the state of the server
else
Change idle server to OFF.

Step 4: The 3D predictor model predicts the required servers at time (T4) using the history of availability at T1, T2, and T3.

3.3 SIGMOID NEURAL TASK PREDICTION

Cloud computing depends on the Data Centers. By increasing the number of data centers the Cloud Computing can provide efficient services to the Cloud Consumers. It consumes lot of time and energy with various process and the developers developed various algorithms to reduce the time for the efficient services, also lot of researches on process to control the time and cost for Cloud Services. Our proposed model is giving the solution to reduce the energy and save the cost for the services. The server is doing n number of job in a particular time, after finishes that particular job the server has to wait for the next job for that the existing model approached if the server is not in process then we can shut down the system. For this case our proposed predictive model predicts the waiting time of the server for the next job to be proceeded and it calculates the difference between idle time and off time. If an idle time < off time means our predictive model calculates the energy consumption for the idle time and it stops the server to shutting off.

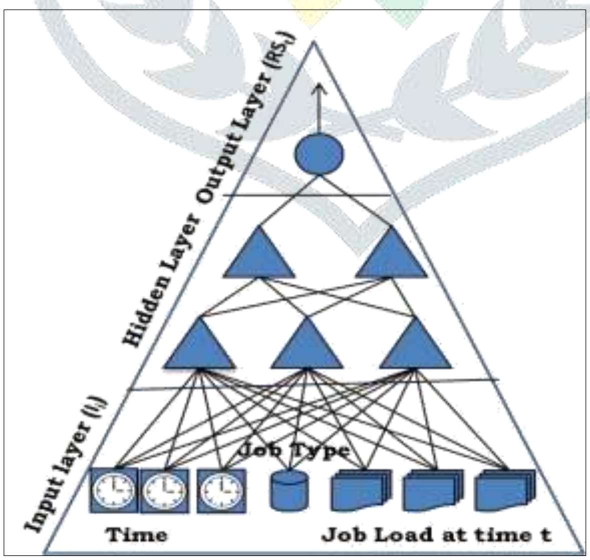


Fig.3. Green Job Predictor

3.3 Parameters

Input [I1 , I2 , I3 , ,,,,, In]

Weight [WI1, WI2, ,,,, WIm]

The structure of Input I i is prime mover to cause perception for output action .

$$F(x) = \frac{1}{1+e^{-BX}}$$

The task predication function executes using weighted sum of the inputs. The weight WIi is multiplied with sum of Inputs Ii .

$$TP = (I, W) = \sum_{i=0}^{n} Ii \;\; WIi$$

The activation function and output are equal when the neuron function is linear. The prediction function output is sigmoid function.

$$RSt = \frac{1}{1 + e^{TA(1\ WIi\ )}}$$

Were RSt is required servers.

The error in predicted output is the difference between the actual value and predicted output. The error rate is depends on the weight value obtained from Input (I i) . The error is minimized by adjusting the weight. The error function for the predicted output is (I , W i , d)

D is the difference in output and actual value.

$$Et = (I, Wi , d) = (RSt (I, WI) - Dt)^2$$

The error of the system is the sum of errors of all RS

$$E (I, WI , d) = \Sigma\, T(RSt\ (I, Wi\ ) - Dt\ )^2$$

The Adjustment of weight (ΔWIi) is obtained by gradient descendent

$$\Delta\, WI_i = -\eta \frac{\partial E}{\partial \mathrm{WIi}} \qquad \text{--------------------(1)}$$

The adjustment of weight is the product of negative of a constant eta (η) by the error which is derived with respect to WIi. The error rate with respect the output RSt is

$$\frac{\partial E}{\partial \mathrm{RSt}} = 2(RS_t - D_t) \qquad \text{-----------(2)}$$

The relation of error rate derived from respect to weight and output

$$\frac{\partial \mathrm{RSt}}{\partial\ \mathrm{WIi}} = \frac{\partial \mathrm{RS_t}}{\partial TPt} * \frac{TPt}{\mathrm{WIi}} = RS_t\,(1 - RS_t) \qquad \text{--------------(3)}$$

Thus the process of weight adjustment is carried out as many as the accuracy is improve .Since the time required the training the network is grows exponentially.

4. Experimental

The simulation is carried out using cloudSim. It is extended to simulate energy efficiency and carbon efficiency in task scheduling. The simulation is carried out using five data centers with 400 servers. It shows the carbon food print with varied number of task. The carbon food print of proposed algorithm is compared with Carbon Efficient First Fit (CEFF) and First Fit Power Efficient (EEPE). The result shows that the average carbon food print is reduced 8% from CEFF and 47% from other heuristic algorithms (Non-Carbon efficient). The proposed algorithm consumes less power while compare to other algorithms. The power consumption is achieved by efficient scheduling.

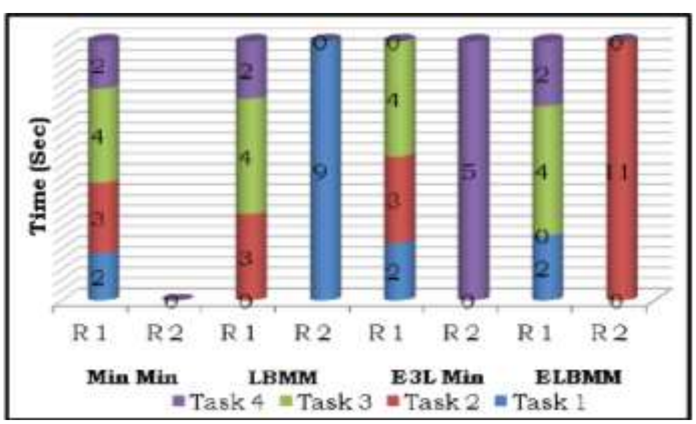


Fig.4. Makespan Comparision with Existing Algorithm

The Fig .4 show the makespan comparison with the existing Scheduling Algorithm MinMin, LBMM, E3L Min, ELBMM with the task and capacity of resources in Milli seconds (R1 100 MIPS, R2 80 MIPS). Fig.5 Show the power utilization of the various scheduling algorithm depends on R1, R2 In this way, the transmission parameters and energy consumption will also improve which can increase load traffic in a certain part of the cloud.

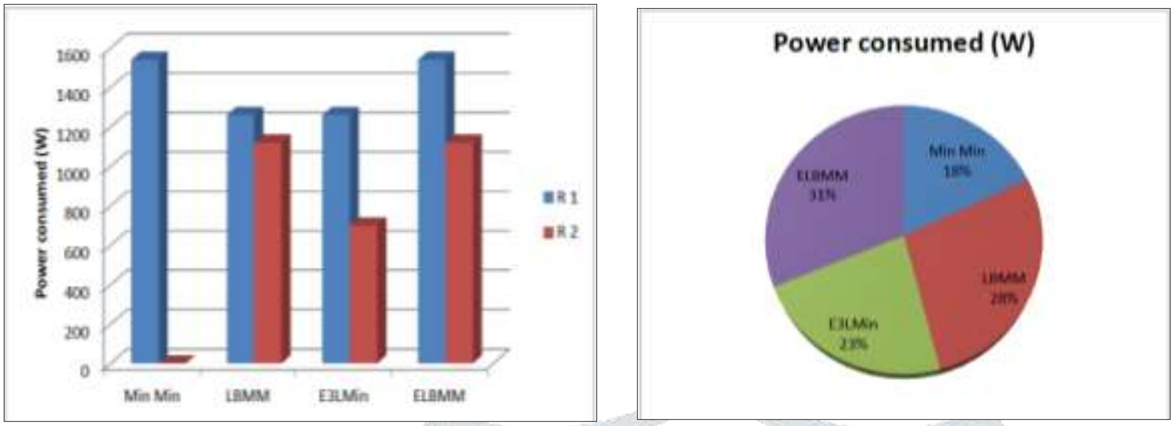


Fig.5.Power Consumption Comparision with Existing Algorithm

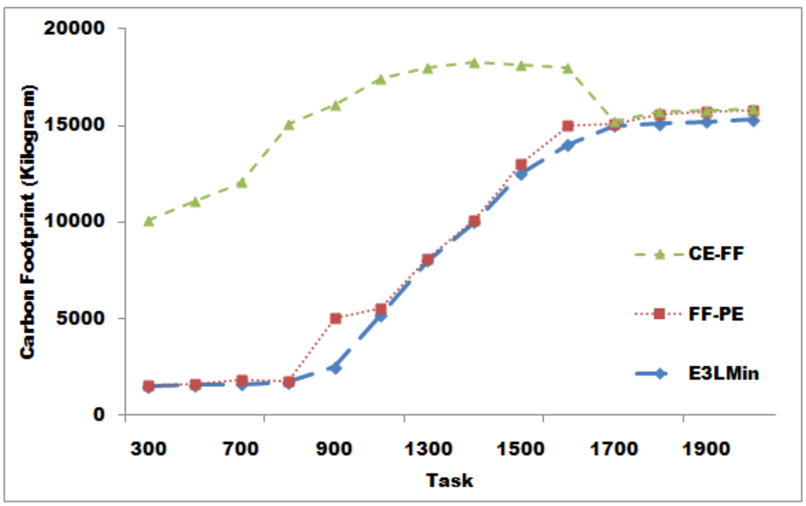


Fig. 6 Carbon foot print with varied number of task

Conclusion

In this paper task predictor scheduling concentrate on energy reduction was studied. In order to achieve the best result, i.e. reducing makespan and energy consumption, the most important parameters including the relationship between the tasks, scheduler required to run the tasks, VM status and the amount of energy utilization to allocate the tasks were determined. The scheduling algorithm chooses the best node for task execution by MinMin which actually has the highest influential among the parameters. The results show that the relationship between the tasks has the greatest focus on the overall makespan of the tasks and the overall energy consumption of the cloud. As a future work the advancement of task effects of the VMs' mode and the reduction of this effect will be examined on the resource allocation.

Reference

[1] K. Roth, F. Goldstein, J. Kleinman, Energy Consumption By Office and Telecommunications Equipment in Commercial Buildings–Volume I: Energy consumption Baseline. Washington, DC: Prepared by Arthur D. Little for the U.S. Department of Energy. A.D. Little Reference no. 72895 00. 2002 (accessed July, 2016). http://www.biblioite.ethz.ch/downloads/Roth_ADL_1.pdf.
[2] S. Albers, Energy-efficient algorithms, Communications of the ACM 53 (5) (2010) 86–96. http://doi.acm.org/10.1145/1735223.1735245.
[3] M. Jens, A. Moberg ˚, D. Lundén, G. Finnveden, N. Lövehagen, Greenhouse gas emissions and operational electricity use in the ICT and entertainment & media sectors, J. Indus. Ecol. 14 (5) (2010) 770–790.
[4] A. Ashraf, B. Byholm, I. Porres, Distributed Virtual Machine consolidation: A systematic mapping study, Turku Centre for Computer Science Technical Report No 1171, December, (2016).
[5] Trends in Global Emissions. https://www.epa.gov/ghgemissions/global-greenhouse-gas-emissions-data.
[6] ICT helping tackle climate change could help cut global emissions 20% by 2030, August, (2016). http://newsroom.unfccc.int/unfccc-newsroom/ ict-sector-helping-to-tackle-climate-change/ (accessed January, 2017).

[7] E. Pinheiro, R. Bianchini, E.V. Carrera, T. Heath, Load balancing and unbalancing for power and performancee in cluster-based systems, in: Proceedings of the Workshop on Compilers and Operating Systems for Low Power, 2001, pp. 182–195.

[8] J.S. Chase, D.C. Anderson, P.N. Thakar, A.M. Vahdat, R.P. Doyle, Managing energy and server resources in hosting centers, in: Proceedings of the 18th ACM Symposium on Operating Systems Principles, ACM, New York, NY, USA, 2001, pp. 103–116.

[9] E. Elnozahy, M. Kistler, R. Rajamony, Energy-efficient server clusters, PowerAware Computer Systems (2003) 179–197.

[10] R. Nathuji, K. Schwan, Virtualpower: coordinated power management in virtualized enterprise systems, ACM SIGOPS Operating Systems Review 41 (6) (2007) 265–278.

[11] R. Raghavendra, P. Ranganathan, V. Talwar, Z. Wang, X. Zhu, No ''power'' struggles: coordinated multi-level power management for the data center, SIGARCH Computer Architecture News 36 (1) (2008) 48–59.

[12] D. Kusic, J.O. Kephart, J.E. Hanson, N. Kandasamy, G. Jiang, Power and performancee management of virtualized computing environments via lookahead control, Cluster Computing 12 (1) (2009) 1–15. [13] S. Srikantaiah, A. Kansal, F. Zhao, Energy aware consolidation for cloud computing, Cluster Computing 12 (2009) 1–15.

[14] M. Cardosa, M. Korupolu, A. Singh, Shares and utilities based power consolidation in virtualized server environments, in: Proceedings of the 11th IFIP/IEEE Integrated Network Management, IM 2009, Long Island, NY, USA, 2009.

[15] A. Verma, P. Ahuja, A. Neogi, pMapper: power and migration cost aware application placement in virtualized systems, in: Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware, Springer, 2008, pp. 243–264.

[16] A. Gandhi, M. Harchol-Balter, R. Das, C. Lefurgy, Optimal power allocation in server farms, in: Proceedings of the 11th International Joint Conference on Measurement and Modeling of Computer Systems, ACM, New York, NY, USA, 2009, pp. 157–168.

[17] M. Gupta, S. Singh, Greening of the internet, in: Proceedings of the ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, SIGCOMM 2003, New York, NY, USA, 2003, pp. 19–26.

[18] N. Vasic, D. Kostic, Energy-aware traffic engineering, in: Proceedings of the 1st ACM International Conference on Energy-Efficient Computing and Networking, e-Energy 2010, Passau, Germany, 2010, pp. 169–178.

[19] C. Panarello, A. Lombardo, G. Schembra, L. Chiaraviglio, M. Mellia, Energy saving and network performance: a trade-off approach, in: Proceedings of the 1st ACM International Conference on Energy-Efficient Computing and Networking, e-Energy 2010, Passau, Germany, 2010, pp. 41–50.

[20] L. Chiaraviglio, I. Matta, GreenCoop: cooperative green routing with energyefficient servers, in: Proceedings of the 1st ACM International Conference on Energy-Efficient Computing and Networking, e-Energy 2010, Passau, Germany, 2010, pp. 191–194.

[21] M. Koseoglu, E. Karasan, Joint resource and network scheduling with adaptive offset determination for optical burst switched grids, Future Generation Computer Systems 26 (4) (2010) 576–589.

[22] L. Tomas, A. Caminero, C. Carrion, B. Caminero, Network-aware metascheduling in advance with autonomous self-tuning system, Future Generation Computer Systems 27 (5) (2010) 486–497.

[23] E. Dodonov, R. de Mell, A novel approach for distributed application scheduling based on prediction of communication events, Future Generation Computer Systems 26 (5) (2010) 740–752.

[24] L. Gyarmati, T. Trinh, How can architecture help to reduce energy consumption in data center networking? in: Proceedings of the 1st ACM International Conference on Energy-Efficient Computing and Networking, e-Energy 2010, Passau, Germany, 2010, pp. 183–186.

[25] C. Guo, G. Lu, H. Wang, S. Yang, C. Kong, P. Sun, W. Wu, Y. Zhang, Secondnet: a data center network virtualization architecture with bandwidth guarantees, in: Proceedings of the 6th International Conference on Emerging Networking EXperiments and Technologies, CoNEXT 2010, Philadelphia, USA, 2010.

[26] L. Rodero-Merino, L. Vaquero, V. Gil, F. Galan, J. Fontan, R. Montero, I. Llorente, From infrastructure delivery to service management in clouds, Future Generation Computer Systems 26 (8) (2010) 1226–1240.

[27] R.N. Calheiros, R. Buyya, C.A.F.D. Rose, A heuristic for mapping virtual machines and links in emulation testbeds, in: Proceedings of the 38th International Conference on Parallel Processing, Vienna, Austria, 2009.

[28] R.K. Sharma, C.E. Bash, C.D. Patel, R.J. Friedrich, J.S. Chase, Balance of power: dynamic thermal management for internet data centers, IEEE Internet Computing (2005) 42–49.

[29] J. Moore, J. Chase, P. Ranganathan, R. Sharma, Making scheduling ''cool'': temperature-aware workload placement in data centers, in: Proceedings of the Annual Conference on USENIX Annual Technical Conference, Anaheim, CA, USA, 2005.

[30] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, I. Brandic, Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility, Future Generation Computer Systems 25 (6) (2009) 599–616.

[31] A. Beloglazov, R. Buyya, Y.C. Lee, A. Zomaya, A taxonomy and survey of energy-efficient data centers and cloud computing systems, Adv. Comput. 82 (2) (2011) 47–111.

[32] N.J. Kansal, I. Chana, Artificial Bee Colony based energy-aware resource utilization technique for cloud computing, Concurrency Comput. 27 (5) (2015) 1207–1225. http://dx.doi.org/10.1002/cpe.3295.

[33] I. Chana, T. Kaur, Delivering IT as a utility-A systematic review, Int. J. Found. Comput. Sci. Technol. 3 (3) (2013). arXiv preprint arXiv:1306.1639http: //dx.doi.org/10.5121/ijfcst.2013.3302.

[34] T. Kaur, I. Chana, Energy efficient cloud: trends, challenges and future directions, in: International Conference on Next generation computing and communication technologies (ICNGCCT 14), 2014.