Energy Efficient Resource Management in Cloud Data Centers

< Gowtham.P>[1], <Sathishkumar.R>[2],<S.S.G.Krishnan>[3]

[1]Student, Department of Information Technology,
SKP Engineering College, Tiruvannamalai, Tamil Nadu
gowthamkarathe@gmail.com

[2]Student, Department of Information Technology,
SKP Engineering College, Tiruvannamalai, Tamil Nadu
shocksathish@gmail.com

[3] Assistant Professor, Department of Information Technology,
SKP Engineering College, Tiruvannamalai, Tamil Nadu
ssgkrishnan@skpec.in

Abstract: Rapid growth of the demand for computational power by scientific, business and web-applications has led to the creation of large-scale data centers consuming enormous amounts of electrical power. We propose an energy efficient resource management system for virtualized Cloud data centers that reduces operational costs and provides required Quality of Service (QoS). Energy savings are achieved by contin-uous consolidation of VMs according to current utilization of resources, virtual network topologies established between VMs and thermal state of computing nodes. We present first results of simulation-driven evaluation of heuristics for dynamic reallocation of VMs using live migration according to current requirements for CPU performance. The results show that the proposed technique brings substantial energy savings, while ensuring reliable QoS. This justifies further investigation and development of the proposed resource management system.

Keywords: *energy-efficient, data center, cloud computing, Virtual Machine.*

1. Introduction

Data centers are becoming increasingly popular for the provisioning of computing resources. The cost and operational expenses of data centers have skyrocketed with the increase in computing capacity.

Energy consumption is a growing concern for data centers operators. It is becoming one of the main entries on a data center operational expenses (OPEX) bill. The Gartner Group estimates energy consumptions to account for up to 10% of the current OPEX, and this estimate is projected to rise to 50% in the next few years.

The slice of roughly 40% is related to the energy consumed by information technology (IT) equipment, which includes energy consumed by the computing servers as well as data center network hardware used for interconnection. In fact, about one-third of the total IT energy is consumed by communication links, switching, and aggregation elements, while the remaining two-thirds are allocated to computing servers. Other systems contributing to the data center energy consumption are cooling and power distribution systems that account for 45% and 15% of total energy consumption, respectively.

The first data center energy saving solutions operated on a distributed basis and focused on making the data center hardware energy efficient. There are two popular techniques for power savings in computing systems.

The Dynamic Voltage and Frequency Scaling (DVFS) technology, adjusts hardware power consumption according to the applied computing load and the Dynamic Power Management (DPM), achieves most of energy savings by powering down devices at runtime. To make DPM scheme efficient, a scheduler must consolidate data center jobs on a minimum set of computing resources to maximize the amount of unloaded servers that can be powered down (or put to sleep). Because the average data center workload often stays around 30%, the portion of unloaded servers can be as high as 70%.

Most of the existing approaches for job scheduling in data centers focus exclusively on the job distribution between computing servers [1] targeting energy-efficient or thermal-aware scheduling. To the best of our knowledge, only a few approaches have considered data center network and traffic characteristics for developing energy-efficient data center schedulers.

Ref. [2] identifies the problem associated with existing multi-path routing protocols in typical fat tree network topologies. Two large traffic flows may be assigned to share the same path if their hash values collide leaving other paths under-loaded. The problem is solved with the introduction of a complex central scheduler that performs flow differentiation and analysis of flow traffic demands across the data center network. Traffic-aware virtual machine placement is proposed in [3]. Relying on the knowledge about network topology, virtual machines are placed to optimize traffic flows inside a data center network. The approach presented in [4], also allows job migration control during runtime with a specifically designed network-aware scheduler. The migration scheduler is aware of the migration delays and bandwidth resources required. As we may see, most of the existing solutions leave the networking aspect unaccounted for in an energy-efficient optimization setting.

This paper presents a data center scheduling methodology that combines energy efficiency and network awareness. The methodology is termed DENS, which is an acronym for **d**ata center **e**nergy-efficient **n**etwork-aware **s**cheduling. The DENS methodology aims to achieve the balance between individual job performances, job QoS requirements, traffic demands, and energy consumed by the data center. Data intensive jobs require low computational load, but produce heavy data streams directed out of the data center as well as to the neighboring nodes. Such data intensive jobs are typically

A. Research scope

The focus of this work is on energy-efficient resource management strategies that can be applied on a virtualized data center by a Cloud provider (e.g. Amazon EC2). The main instrument that we leverage is live migration of VMs. The ability to migrate VMs between physical hosts with low overhead gives flexibility to a resource provider as VMs can be dynamically reallocated according to current resource requirements and the allocation policy. Idle physical nodes can be switched off to minimize energy consumption.

In this paper we present a decentralized architecture of the resource management system for Cloud data centers and propose the development of the following policies for continuous optimization of VM placement:

- Optimization over multiple system resources – at each time frame VMs are reallocated according to current CPU, RAM and network bandwidth utilization.
- Network optimization – optimization of virtual network topologies created by intercommunicating VMs. Network communication between VMs should be observed and considered in reallocation decisions in order to reduce data transfer overhead and network devices load.
- Thermal optimization – current temperature of physical nodes is considered in reallocation decisions. The aim is to avoid "hot spots" by reducing workload of the overheated nodes and thus decrease error-proneness and cooling system load.

B. Research challenges

The key challenges that have to be addressed are:

1) How to optimally solve the trade-off between energy savings and delivered performance?
2) How to determine when, which VMs, and where to migrate in order to minimize energy consumption by the system, while minimizing
3) migration overhead and ensuring SLA?
4) How to develop efficient decentralized and scalable algorithms for resource allocation?
5) How to develop comprehensive solution by combining several allocation policies with different objectives?

The remainder of the paper is organized as follows. In the next section we discuss related work followed by the proposed system architecture in Section III. In Sections IV and V we present the allocation policies and evaluation of

III. SYSTEM ARCHITECTURE

In this work the underlying infrastructure is represented by a large-scale Cloud data center comprising *n* heterogeneous physical nodes. Each node has a CPU, which can be multi-

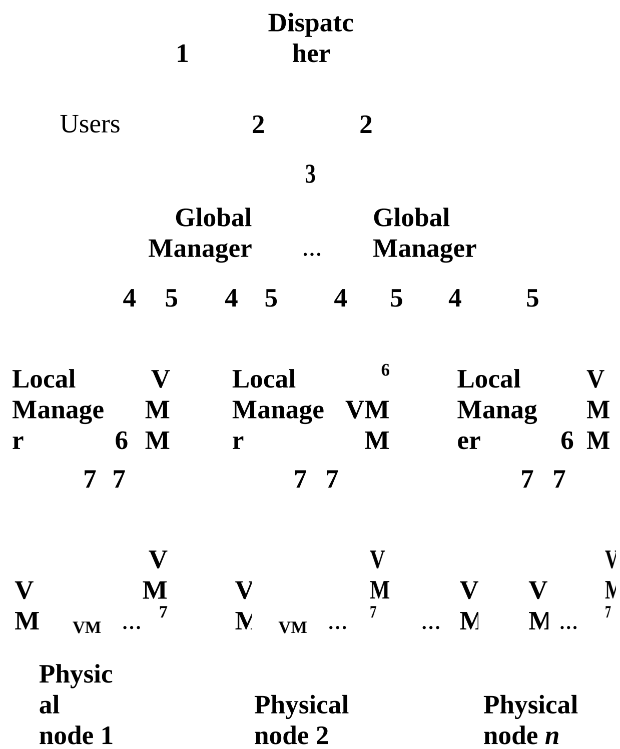

Fig 1. System Architecture

core, with performance defined in Millions Instructions Per Second (MIPS). Besides that, a node is characterized by the amount of RAM and network bandwidth. Users submit requests for provisioning of *m* heterogeneous VMs with resource requirements defined in MIPS, amount of RAM and network bandwidth. SLA violation occurs when a VM cannot get the requested amount of resource, which may happen due to VM consolidation.

The software system architecture is tiered comprising a dispatcher, global and local managers. The local managers reside on each physical node as a part of a Virtual Machine Monitor (VMM). They are responsible for observing current utilization of the node's resources and its thermal state. The local managers choose VMs that have to be migrated to another node in the following cases:

- The utilization of some resource is close to 100% that creates a risk of SLA violation.
- The utilization of resources is low, therefore, all the VMs should be reallocated to another node and the idle node should be turned off.
- A VM has intensive network communication with another VM allocated to a different physical host.
- The temperature exceeds some limit and VMs have to be migrated in order to reduce load on the cooling system and allow the node to cool down naturally.

The local managers send to the global managers the information about the utilization of resources and VMs chosen to migrate. Besides that, they issue commands for

VM resizing, application of DVFS and turning on / off idle nodes. Each global manager is attached to a set of nodes and processes data obtained from their local managers. The global managers continuously apply distributed version of a heuristic for semi-online multidimensional bin-packing, where bins represent physical nodes and items are VMs that have to be allocated. The decentralization removes a Single Point of Failure (SPF) and improves scalability. Each dimension of an item represents the utilization of a particular resource. After obtaining allocation decision, the global managers issue commands for live migration of VMs.

As shown in Figure 1, the system operation consists of the following steps:

1) New requests for VM provisioning. Users submit requests for provisioning of VMs.
2) Dispatching requests for VM provisioning. The dispatcher distributes requests among global managers.
3) Intercommunication between global managers. The global managers exchange information about utilization of resources and VMs that have to be allocated.
4) Data about utilization of resources and VMs chosen to migrate. The local managers propagate information about resource utilization and VMs chosen to migrate to the global managers.
5) Migration commands. The global managers issue VM migration commands in order to optimize current allocation.
6) Commands for VM resizing and adjusting of power states. The local managers monitor their host nodes and issue commands for VM resizing and changes in power states of nodes.
7) VM resizing, scheduling and migration actions. According to the received commands, VMM performs actual resizing and migration of VMs as well as resource scheduling.

IV. ALLOCATION POLICIES

We propose three stages of VM placement optimization: reallocation according to current utilization of multiple system resources, optimization of virtual network topologies established between VMs and VM reallocation considering thermal state of the resources. Each of these stages is planned to be investigated separately and then combined in an overall solution. The developed algorithms have to meet the following requirements:

- Decentralization and parallelism – to eliminate SPF and provide scalability.
- High performance – the system has to be able to quickly respond to changes in the workload.
- Guaranteed QoS – the algorithms have to provide reliable QoS by meeting SLA.
- Independence of the workload type – the algorithms have to be able to perform efficiently in mixed-application environments.

The VM reallocation problem can be divided in two: selection of VMs to migrate and determining new placement of these VMs on physical hosts. The first part has to be considered separately for each optimization stage. The second part is solved by application of a heuristic for semi-online multidimensional bin-packing problem.

At the first optimization stage, the utilization of resources is monitored and VMs are reallocated to minimize the number of physical nodes in use and thus minimize energy consumption by the system. However, aggressive consolidation of VMs may lead to violation of performance requirements. We have proposed several heuristics for selection of VMs to migrate and investigated the trade-off between performance and energy savings. To simplify the problem for the first step we considered only utilization of CPU. The main idea of the policies is to set upper and lower utilization thresholds and keep total utilization of CPU created by VMs sharing the same node between these thresholds. If the utilization exceeds the upper thresholds, some VMs have to be migrated from the node to reduce the risk of SLA violation. If the utilization goes bellow the lower thresholds, all VMs have to be migrated and the node has to be switched off to save the energy consumed be the idle node. Another problem is to determine particular values of the utilization thresholds. The results of the proposed algorithms evaluation are presented in Section V.

Due to continuous reallocation, some intensively communicating VMs can be placed inefficiently leading to excessive load on the network facilities. Therefore, it is crucial to consider network communication behavior of VMs in reallocation decisions. The aim of the second proposed optimization stage is to place communicating VMs in a way minimizing the overhead of data transfer over network.

A cooling system of a data center consumes a significant amount of energy, therefore, the third proposed optimization stage is aimed at optimization of cooling system operation. Due to consolidation, some computing nodes experience high load leading to overheating and thus require extensive cooling. Monitoring of the nodes' thermal state using sensors gives an opportunity to recognize overheating and reallocate workload from the overheated node to allow the natural cooling. The network and temperature optimizations are subjects for the ongoing research work.

V. EVALUATION

As the proposed system is targeted on a large-scale Cloud data center, it is necessary to conduct large-scale experiments to evaluate the algorithms. However, it is difficult to run large-scale experiments on a real-world infrastructure, especially when the experiments have to be repeated for different policies with the same conditions [18]. Therefore, simulation has been chosen as a way to evaluate the proposed heuristics. We have chosen CloudSim toolkit [18] as a simulation framework, as it is built for simulation of Cloud computing environments. In comparison to alternative simulation toolkits (e.g. SimGrid, GangSim), CloudSim supports modeling of on-demand virtualization enabled resource application management. We have extended the framework in order to enable energy aware simulations as the core framework does not provide this capability. In addition, we have incorporated the abilities to account SLA violations and to simulate dynamic workloads that correspond to web-applications and online services.

The simulated data center consists of 100 heterogeneous physical nodes. Each node is modeled to have one CPU core with performance equivalent to 1000, 2000 or 3000 MIPS, 8 Gb of RAM and 1 TB of storage. Users submit requests for provisioning of 290 heterogeneous VMs that fill the full capacity of the data center. For the borderline policies we simulated a Non Power Aware policy (NPA) and DVFS that adjusts the voltage and frequency of CPU according to current utilization. We simulated a Single Threshold policy (ST) and two-threshold policy aimed at Minimization of Migrations (MM). Besides that, the policies have been evaluated with different values of the thresholds.

Table I
SIMULATION RESULTS

Policy	Energy	SLA	Migr.	Avg. SLA
NPA	9.15 KWh	-	-	-
DVFS	4.40 KWh	-	-	-
ST 50%	2.03 KWh	5.41%	35 226	81%
ST 60%	1.50 KWh	9.04%	34 231	89%
MM 30-70%	1.48 KWh	1.11%	3 359	56%
MM 40-80%	1.27 KWh	2.75%	3 241	65%
MM 50-90%	1.14 KWh	6.69%	3 120	76%

The simulation results are presented in Table I. The results show that dynamic reallocation of VMs according to current utilization of CPU can bring higher energy savings comparing to static allocation policies. MM policy allows to achieve the best energy savings: by 83%, 66% and 23% less energy consumption relatively to NPA, DVFS and ST policies respectively with thresholds 30-70% and ensuring percentage of SLA violations of 1.1%; and by 87%, 74% and 43% with thresholds 50-90% and 6.7% of SLA violations. MM policy leads to more than 10 times fewer VM migrations than ST. The results show the flexibility of the algorithm, as the thresholds can be adjusted according to SLA requirements. Strict SLA (1.11%) allow achievement of the energy consumption of 1.48 KWh. However, if SLA are relaxed (6.69%), the energy consumption is further reduced to 1.14 KWh.

VI. CONCLUSION AND FUTURE WORK

In this paper have presented a decentralized architecture of the energy aware resource management system for Cloud data centers. We have defined the problem of minimizing the energy consumption while meeting QoS requirements and stated the requirements for VM allocation policies. Moreover, we have proposed three stages of continuous

REFERENCES

[1] R. Brown *et al.*, “Report to congress on server and data center energy efficiency: Public law 109-431,” *Lawrence Berkeley National Laboratory*, 2008.

[2] G. Semeraro, G. Magklis, R. Balasubramonian, D. H. Al-bonesi, S. Dwarkadas, and M. L. Scott, “Energy-efficient processor design using multiple clock domains with dynamic voltage and frequency scaling,” in *Proceedings of the 8th International Symposium on High-Performance Computer Architecture*, 2002, pp. 29–42.

[3] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, “Xen and the art of virtualization,” in *Proceedings of the 19th ACM symposium on Operating systems principles*, 2003, p. 177.

[4] R. Buyya, C. S. Yeo, and S. Venugopal, “Market-oriented
cloud computing: Vision, hype, and reality for delivering it services as computing utilities,” in *Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC’08)*. IEEE CS Press, Los Alamitos, CA, USA, 2008.

[5] R. Neugebauer and D. McAuley, “Energy is just another re-source: Energy accounting and energy pricing in the nemesis OS,” in *Proceedings of the 8th IEEE Workshop on Hot Topics in Operating Systems*, 2001, pp. 59–64.

[6] H. Zeng, C. S. Ellis, A. R. Lebeck, and A. Vahdat, “ECOSys-tem: managing energy as a first class operating system resource,” *ACM SIGPLAN Notices*, vol. 37, no. 10, p. 132, 2002.

[7] E. Pinheiro, R. Bianchini, E. V. Carrera, and T. Heath, “Load balancing and unbalancing for power and performance in cluster-based systems,” in *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, and 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] R. Nathuji and K. Schwan, “Virtualpower: Coordinated power management in virtualized enterprise systems,” *ACM SIGOPS Operating Systems Review*, vol. 41, no. 6, pp. 265–278, 2007.

[10] D. Kusic, J. O. Kephart, J. E. Hanson, N. Kandasamy, and G. Jiang, “Power and performance management of virtual-ized computing environments via lookahead control,” *Cluster Computing*, vol. 12, no. 1, pp. 1–15, 2009.

[11] S. Srikantaiah, A. Kansal, and F. Zhao, “Energy aware con-solidation for cloud computing,” *Cluster Computing*, vol. 12, pp. 1–15, 2009.

[12] Y. Song, H. Wang, Y. Li, B. Feng, and Y. Sun, "Multi-Tiered On-Demand resource scheduling for VM-Based data center," in *Proceedings of the 2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid-Volume 00*, 2009, pp. 148–155.

[13] M. Cardosa, M. Korupolu, and A. Singh, "Shares and util-ities based power consolidation in virtualized server envi-ronments," in *Proceedings of IFIP/IEEE Integrated Network Management (IM)*, 2009.

[14] A. Verma, P. Ahuja, and A. Neogi, "pMapper: power and migration cost aware application placement in virtualized systems," in *Proceedings of the 9th ACM/IFIP/USENIX Inter-national Conference on Middleware*. Springer-Verlag New York, Inc., 2008, pp. 243–264.

[15] R. N. Calheiros, R. Buyya, and C. A. F. D. Rose, "A heuristic for mapping virtual machines and links in emulation testbeds," 2009.

