VM Migrations and Consolidation In Cloud Computing

Weqar Fatima Sayed Abedi[1],
[1]Student,
TYBSc IT, Department of Information Technology Jai Hind College (Autonomous), Mumbai
weqarfatima@gmail.com

ABSTRACT

Cloud computing, cloud users, and their workload are increasingly rising, whereas QoS (quality of service) is decreasing proportionally. Cloud computing is a recent development in the field of information technology that allows for more efficient service delivery and versatility. While dynamic consolidation of virtual machines (VMs) in a data centre is an effective way to reduce energy consumption and increase physical utilization, CM consolidation in Cloud Computing is a huge challenge. Since it directly impacts end-user programmers, there must be a balance between power usage and performance degradation. The most important thing to consider is consistency, which should not be compromised even if the amount of user requests on the cloud increases. Only an effective live VM migration technique can handle a large number of cloud users while also reducing time and energy expended on cloud data centers and delivering better computing services. The data centre is at the heart of cloud computing. The pressure on data centers increases as the demand for cloud storage services expands. Some data-center resources can become overloaded or under loaded over time, resulting in increased energy consumption, as well as decreased functionality and resource waste. [10]. there are two forms of virtual machine migration: pre-copy and post-copy. The purpose of this literature survey paper is to reduce power consumption while making maximum use of resources powered on physical servers and reviewing all the existing live VM migration techniques with their advantages and disadvantages. Since the number of data centers and applications is rising at an exponential pace, load balancing and server consolidation can help to optimize resource efficiency and increase Quality of Service (QoS) metrics. Pre-copy and post-copy VM migration strategies have been contrasted using simulations with CPU, Memory, and Network as parameters.

KEYWORDS

Cloud Computing, VM Migration, Energy Consumption, Virtualization, Virtual Machines, Consolidation,

1. INTRODUCTION

Many physical computers in cloud data centers use virtualization. The user's requests are packaged as virtual machines (VMs), which are then assigned to various hosts based on relevant parameters, such as optimizing resource usage, reducing the number of VM migrations, and meeting the cloud providers' and users' Service Level Agreement (SLA) specifications. The growing number of cloud data centers, as well as the high demand for internet services, has boosted the idea of building large-scale data centers. The number of cloud data centers capable of supporting large-scale Internet services is rapidly growing, and the cost of operation is rising due to these data centers' rising energy consumption. [1].

Cloud computing is a form of computer technology that enables companies to offer services to customers on demand. This technology makes knowledge more accessible across a variety of devices, such as Smartphone's, PDAs, PCs, and Tablets.

Infrastructure as a Service (IaaS), Software as a Service (SaaS), and Application as a Service (PaaS) are three service models in cloud computing that provide various benefits (PaaS). Everyone needs the cloud nowadays, from a single individual to a big corporation. Before its inception, people required proper hardware and software installations to run one of the techniques known as virtualization, which consolidates the number of virtual machines on one physical server and dramatically reduces power consumption. This concept entails running virtual machines on a smaller number of physical servers. For example, if we have two physical servers, each with one virtual machine (VM), running both VMs on one (same) physical server will be much more effective because the power consumed by both servers would be cut in half when they are consolidated on one physical server. We can save more power without hurting the data centers if we do it this way.

Load balancing is the process of allocating and reassigning load among available resources to maximize throughput while lowering costs, response times, and energy consumption, and thus improving resource utilization and efficiency [2], [3]. According to the survey, GOOGLE's datacenter used nearly 260 million megawatts of energy in 2013, accounting for 0.01 percent of global energy consumption. This immense amount of energy is more than enough to fuel 200,000 households [4, 5]. The purpose of this literature survey is to reduce power consumption while making maximum use of resources powered on physical servers and reviewing all the existing live VM migration techniques with their advantages and disadvantages. VM migration techniques such as pre-copy and post-copy have been compared by simulation corresponding to CPU uses, Memory and Network as parameters.

A. Power Consumption and Management:

Virtualization architecture is regarded as a critical component of data centre power management.

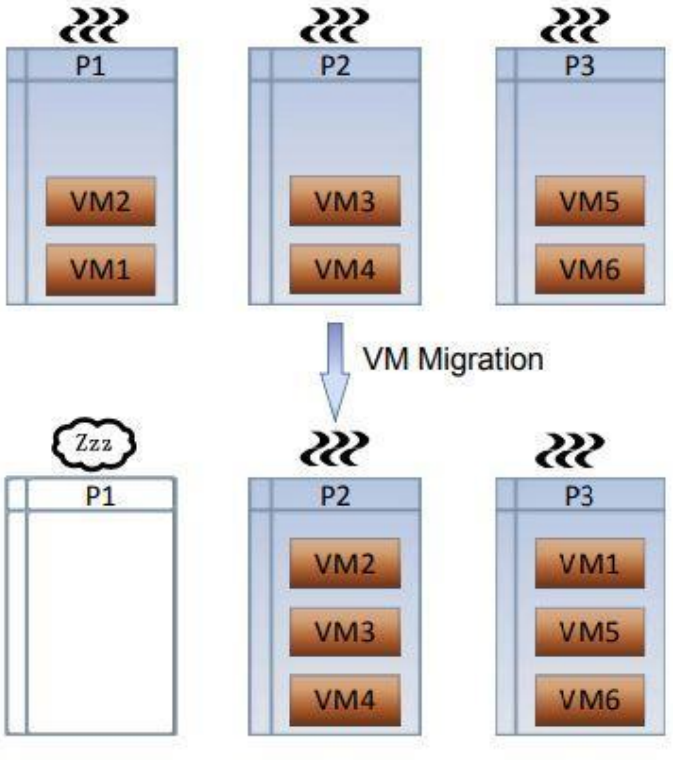

Fig.1 Resource Consolidation in Cluster. The diagram shows six virtual machines (VM1-VM6) operating on three physical servers (P1-P3). VM1 and VM2 are consolidated on P2 and P3 with the aid of VM migration. P1 is also switched off, lowering the cluster's power consumption.

Previously, before the concept of virtualization, computing services were delivered by grid computing, which was extremely time consuming and only used by a small number of people. In order to increase the performance of modern data centers, virtualization technology is being implemented and used more frequently. The migration of virtual machines, on the other hand, results in increased power usage. According to the EPA [8,] data centers in the United States would cost 100 billion kWh, or $7.4 billion per year, by 2011, rendering data centre power management a critical problem in most

countries. Consolidation is a well-known technique for dynamically reducing the number of nodes used within a running cluster by freeing nodes that aren't necessary for the current step of the computation [9]. Fig. 1 In the case of the use of VM migration technology, the data center's power consumption is reduced. However, there are some drawbacks to restructuring, such as failure to meet the Service Level Agreement (SLA), and increased power usage during the migration phase. As a result, in order to formulate an effective consolidation plan, you must first consider the effects of consolidating applications.

B. Cloud Computing Goals:

Task scheduling and resource distribution should be the two priorities of cloud computing; as a result, the outcome of these goals should be [6]:

1. High resource availability
2. Increasing resource utilization
3. Reduction in resource cost
4. Preserving the elasticity of cloud computing
5. Reduction of carbon emissions
6. Energy savings.

Within a computer structure, resources are sets of physical or virtual components with limited availability; any computer connected to the system is considered a resource, as is any internal part of the system. as listed in Table 1 [7].

Physical resources	Logical resources
Storage Memory CPU Workstations Network elements Sensors actuators	Bandwidth (BW) Energy Operating system Information security, protocols APIs Network loads, delays

TABLE 1. Physical and logical resources [7].

2. VIRTUAL MACHINE MIGRATION

By moving a virtual machine through physical servers, VM migration has become the primary method for improving data centre performance. Thanks to server consolidation via VM migration, more computers can be turned off. VM migration is essential for keeping the server's workload and isolating the data centre. Virtual machine migration transfers the state of the processor (I/O, CPU, and memory) between physical machines. It moves the processor's state (I/O, CPU, and memory) from one physical machine to another. Using resource-rich servers to replace underutilised servers and provide better support as the number of users grows In the cloud data centre, the VM migration technique is used. There are basically two migration patterns for VM migration: live migration and non-live/regular migration. Service provisioning is stopped during a non-live migration. System provisioning is not disrupted during a live migration. The second approach pauses the current server, copies its memory contents, and then restarts it on the new host, moving a virtual machine from one host to another. The first performs the same logical function as the second, but without requiring the server domain to be paused. Live migration, as opposed to traditional migration, shows a major potential for using virtual machines and virtual machine migration technologies to efficiently handle workload consolidation and, as a result, improve overall data centre power use.

A. Non-Live VM Migration

When a virtual machine migrates from a source server to a destination server, it stops serving customers. The VM state at the target server is not resumed until the entire VM has been moved. The base is process migration

Advantages: The VM migration takes a predictable amount of time, and memory pages are only transferred once.

Disadvantages: Application efficiency is not optimal due to service disconnection during VM migration.

B. Live VM Migration

It ensures that the service is not interrupted during the VM migration. (1) Server performance optimization during VM migration, (2) Efficient bandwidth usage, and (3) Reduce high migration time and downtime during migration are all goals of live VM migration. The advantages and drawbacks of live migration are categorically discussed in the live migration scheme subsection.

1. ***Pre-copy VM Migration:***
   In this migration, memory pages are copied iteratively until a suitable state is reached, which adds to the migration time. Furthermore, dirty memory pages were migrated using system resources and the network. This migration process has many phases, including memory migration and virtual machine migration.
   The CPU aggressively transfers dirty pages during the memory migration phase, after which VM migration from the source server to the destination server occurs.
   a. ***Warm- up phase:***
      The CPU copies the required memory pages from the source node to the destination node during this step. And all of this was accomplished without putting the virtual machine on hold at the source node [10]. Before the entire fault pages are copied to the destination, any duplicate pages that shift during the copying process will be reduplicated.
   b. ***Stop-and-Copy phase:***
      After the warm-up phase is done, the stop-and-copy phases begin. The VM is stopped at the source, the changed data is copied to the destination, and the VM resumes processing at the destination node during this stage. The entire pre-copy VM migration process is broken down into six distinct stages, which include.
      (1) Target host selection, (2) Resource reservation, (3) Iterative pre-copying rounds, (4) Stop and copy phase, 5) Commitment, and (6) VM activation at the target serve.

Advantage: It performs better because there is less memory transfer. Migration mode is more optimal if there are fewer page faults during VM migration.

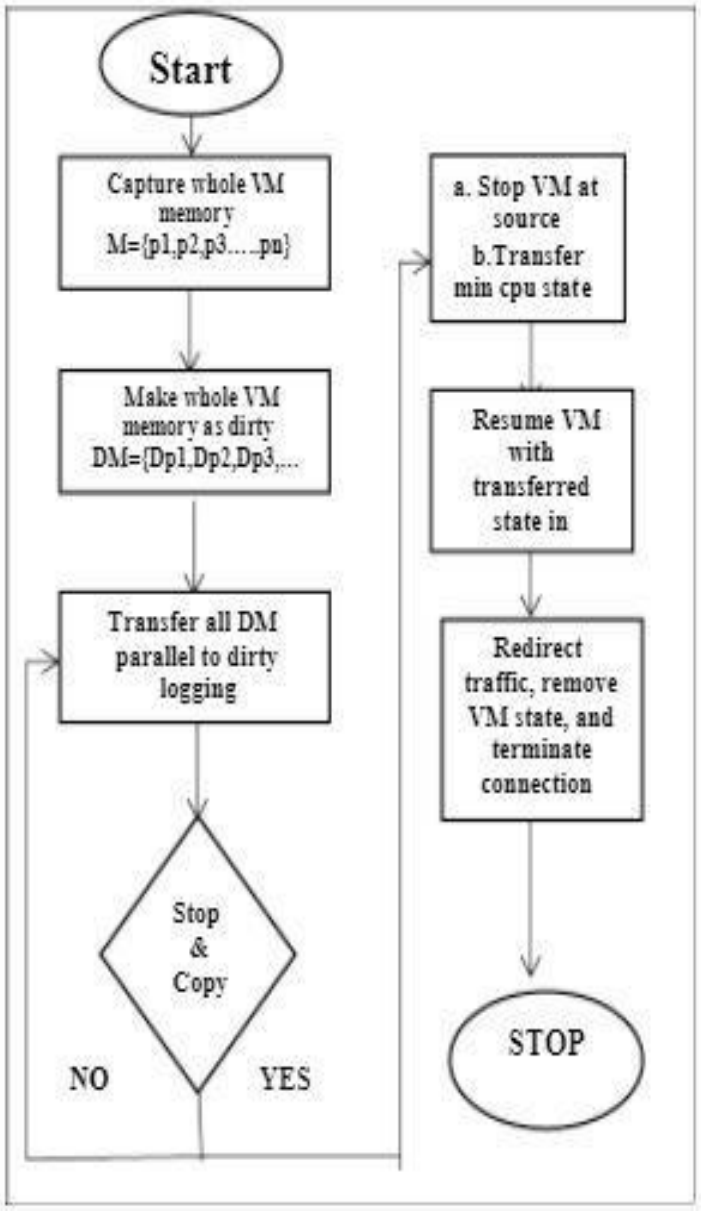

Fig.2 Pre-copy migration flow chart

Disadvantage: Output is not optimal when a virtual machine has many cores. In comparison to migration time, there is more downtime [10].

2. Post-copy VM Migration

The VM is transferred to the destination first, followed by the memory pages in this migration. Pre-copy migration is the polar opposite of this. A small portion of the virtual machine's running state (CPU Registers) is sent to the destination after a portion of the virtual machine pauses for a brief period of time on the source host. When started, this virtual machine will start running at the destination. In response to these network faults, the source hosts submit fault pages, but this degrades application performance [11].

Advantages: This approach solves the post-copy problem of more downtime than migration time.

Disadvantage: Output suffers as there are more page faults and network faults.

3 LOAD BALANCING

The market for cloud storage varies depending on user requirements and resource availability. Load balancing is one of the most important issues in this area, and it should not be ignored. It is the process of allocating and reassigning load among available resources to maximize usage and reduce costs, energy consumption, and response time [12]. Load balancing distributes workload evenly among all resources in order to maximize resource performance, customer loyalty, equal resource distribution, increased scalability, and avoid over-provisioning and bottlenecks, among other things. Fig. 4 shows an overview of the load balancing model. The model presented here incorporates certain data-center elements, such as physical components. (Servers), as well as virtualized components (i.e VMs). We have the opportunity to ensure that the tasks load balancer receives the demands of the clients and implements a task load balancing algorithm to spread demands among the VMs. The load balancer decides which VM should be assigned to the next task. A demand Project management is the responsibility of the data centre controller. Administration As a result, tasks were allocated to the load. The load balancing algorithm is applied by the balancer, It selects the

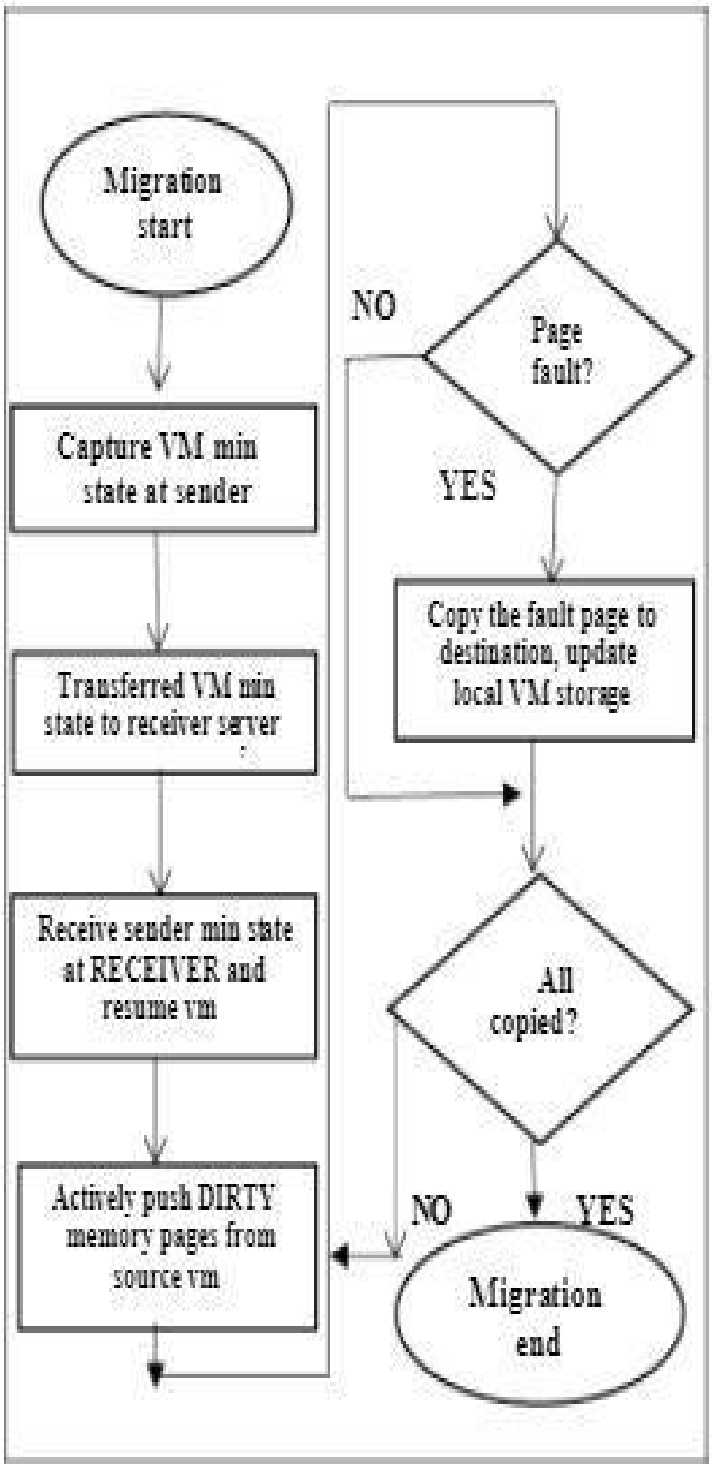

Fig.3 Post-copy migration flowchart

appropriate virtual machine (VM) to handle the task the PMs' equilibrium at all times. The VM manager is in charge of virtual machines. Virtualization, which aims to disperse costly hardware among VMs, is the

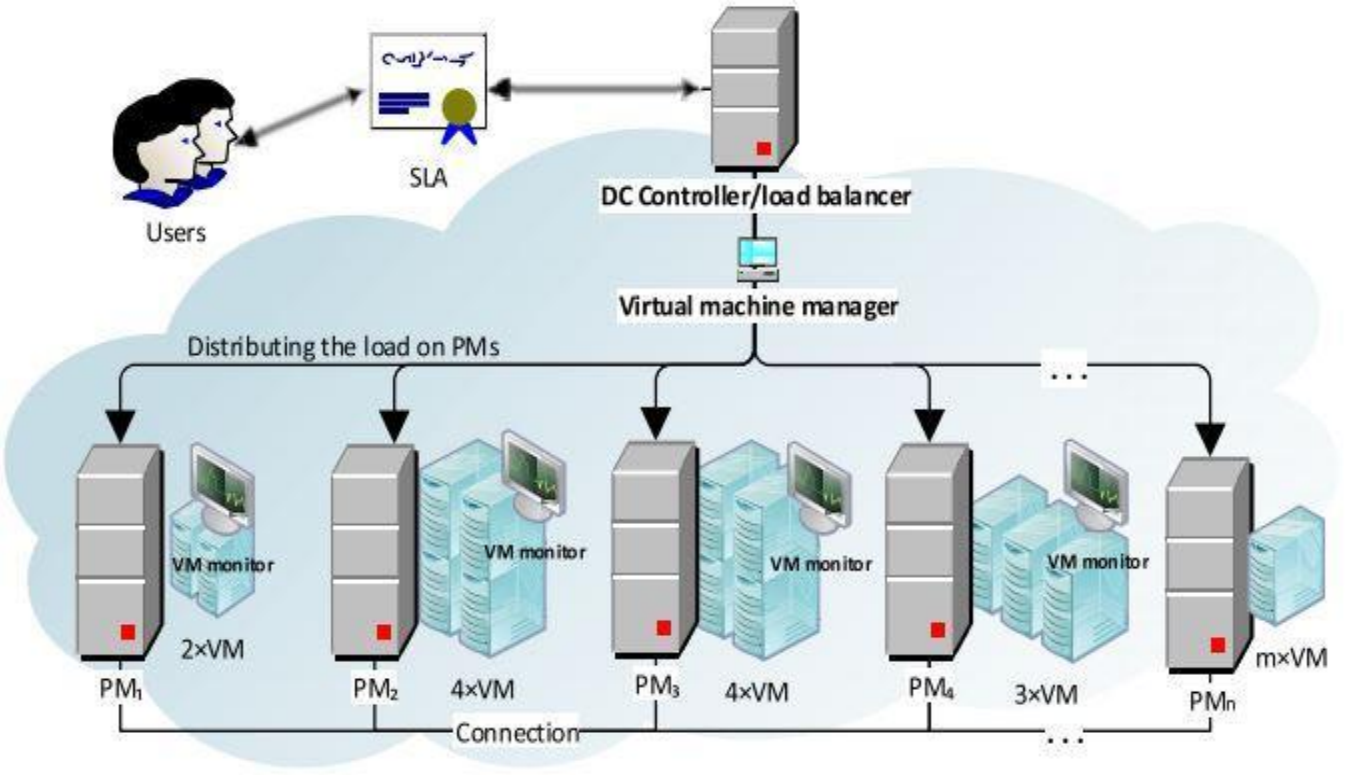

Fig 5. Load Balancing

dominant technology in the cloud computing world. A virtual machine is a computer software application which is in charge of the systems that allow applications to run. Users of cloud computing are located all over the world and send their requests to the VMs at random. As a result, role assignment is one of the most critical issues in cloud computing, and it should be considered in order to maintain service quality. When certain VMs are idle, overloaded, or have few tasks to perform, the QoS is reduced, resulting in customer frustration and attempts to move work to another service provider. The Virtual Machine Monitor (VMM), also known as the "hypervisor," is used to control and build virtual machines [14]. Provision (resume), suspension (storage), multiplexing, and live migration are the four procedures provided by VMM [15]. These procedures are essential for load balancing.

4 SERVER CONSOLIDATION

Data-center services spread at an exponential rate. Virtualized PMs in active virtual machines are used by cloud providers to present their services. Clients must be convinced of this by promising high performance and a large data repository volume. Meanwhile, data-centre virtualization technology is commonly used to make handling PMs or ‘‘servers" simpler. However, if done incorrectly, assigning PMs to VMs may have an effect on data-center performance. On the one hand, this technology eliminates data-center sprawl, energy usage, and a high carbon footprint; on the other hand, it provides several advantages such as resource distribution, virtual machine resizing, live migration, and server consolidation. . Server consolidation is commonly used to reduce overall energy consumption and carbon emissions in data centers. Furthermore, resource waste is at the root of cloud computing popularity. The virtual machine live migration is the key feature that has

made server consolidation a hot topic among researchers. The best way to minimize energy usage by reducing the number of active servers in the data centre is to migrate live.

5 CONCLUSION

In this paper literature survey has been done to reduce power consumption while making maximum use of resources powered on physical servers and reviewing all the existing live VM migration techniques with their advantages and disadvantages. VM migration techniques such as pre-copy and post-copy have been compared by simulation corresponding to CPU uses, Memory and Network as parameters. VM migration is key to realize VM-based resource reservation and power reduction. And understanding its impact is important to make power-efficient deployment in data centers. Less power usage and increased processing performance has always been a major concern for cloud owners. Virtual machine consolidation is key to achieving the objectives of powered-on physical servers and VM migration. The VM consolidation manager works with virtual machines to move them, power on/off physical servers, and assign users to the most powerful physical server available at the time. Thus, In virtualized data centers, it is important to reduce energy consumption while maintaining SLA compliance.

6 REFERENCES

[1] Zhu X, Young D, Watson B J, Wang Z, Rolia J, Singhal S, McKee B, Hyser C, Gmach D, Gardner R, Christian T, Cherkasova L. 1000 islands: an integrated approach to resource management for virtualized data centers. Cluster Computing, 2009, 12(1): 45–57

[2] P. Singh, P. Baaga, and S. Gupta, ‘‘Assorted load balancing algorithms in cloud computing: A survey,’’ Int. J. Comput. Appl., vol. 143, no. 7, pp. 34–40, 2016.

[3] S. Goyal and M. K. Verma, ‘‘Load balancing techniques in cloud computing environment—A review,’’ Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 6, no. 4, pp. 583–588, Apr. 2016.

[4] M. Uddin, A. Shah, R. Alsaqour and J. Memon, “Measuring efficiency of tier level data centers to implement green energy efficient data centers,” Middle East J Sci Res 15:200–207, 2013.

[5] A. Beloglazov and R. Buyya, “Energy efficient resource management in virtualized cloud data centers. In: Proceedings of the 2010 10th IEEE/ACM international conference on cluster, cloud and grid computing,” IEEE Computer Society, New York, pp 826–831, 2010.

[6] I. N. Ivanisenko and T. A. Radivilova, ‘‘Survey of major load balancing algorithms in distributed system,’’ in Proc. Inf. Technol. Innov. Bus. Conf. (ITIB), 2015, pp. 89–92.

[7] S. S. Manvi and G. K. Shyam, ‘‘Resource management for Infrastructure as a Service (IaaS) in Cloud computing: A survey,’’ J. Netw. Comput. Appl., vol. 41, pp. 424–440, May 2014

[8] D. Meisner, B. T. Gold, and T. F. Wenisch, “Powernap: eliminating server idle power,” in ASPLOS, 2009, pp. 205–216

[9] F. Hermenier, N. Loriant, and J.-M. Menaud, “Power management in grid computing with xen,” in ISPA Workshops, 2006, pp. 407–416. http://csc.csudh.edu/btang/seminar/papers/paper5.pdf

[10] M. Uddin, A. Shah, R. Alsaqour and J. Memon, “Measuring efficiency of tier level data centers to implement green energy efficient data centers,” Middle East J Sci Res 15:200–207, 2013. K. Ibrahim and Hofmeyr, “Optimized pre-copy live migration for memory intensive applications,” Seattle, Washington, USA Copyright 2011 ACM 978-1-4503-0771, November 12-18, 2011.

[11] R. Singh and K. Khalon, “Comparative Study of virtual machine migration techniques and challenges in Post Copy Live Virtual Machine Migration,” International Journal of

Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value, 2013.

[12] P. Singh, P. Baaga, and S. Gupta, ‘‘Assorted load balancing algorithms in cloud computing: A survey,’’ Int. J. Comput. Appl., vol. 143, no. 7, pp. 34–40, 2016. S. Goyal and M. K. Verma, ‘‘Load balancing techniques in cloud computing environment—A review,’’ Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 6, no. 4, pp. 583–588, Apr. 2016.

[14] H. Gupta and K. Sahu, ‘‘Honey bee behavior based load balancing of tasks in cloud computing,’’ Int. J. Sci. Res., vol. 3, no. 6, pp. 842–846, Jun. 2014.

[15] K. Hwang, J. Dongarra, and G. C. Fox, Distributed and Cloud Computing: From Parallel Processing to the Internet of Things. San Mateo, CA, USA: Morgan Kaufmann, 2013.