Power and Data Aware Best Fit Algorithm for Energy Saving in Cloud Computing

Pradeep Kumar[1], Dilbag Singh[2], Ankur Kaushik[3]

[1]Dept. of Computer Science and Applications, [2]Dept of Computer Science and Applications, [3]Dept. of Computer Science
[1]Chaudhary Devi Lal University, [2]Chudhary Devi Lal University, [3]YMCA Faridabad
[1]Budhakhera, Kaithal, India, [2]Cdlu Sirsa, India, [3]Old Faridabad, India

ABSTRACT- **Cloud computing is a computing advancement that provides a large number of computing and storage resources to different people. Management of huge amount of data in cloud computing is done by VM consolidation. Nowadays, energy saving has become a key aspect in designing VM consolidation algorithms. Algorithms for selecting next VM have never given consideration to both power and data at the same time in past. This paper proposes power and data aware best fit algorithm (PDABFD) for selecting next VM with power and data constraints. Proposed algorithm work in the direction of maximum resource utilization by VMs. Proper utilization of resources leads to conservation of energy. As a result, energy consumption using PDABFD algorithm is less than previous algorithms.**

KEYWORDS- **Cloud computing, CloudSim, Clouds, Simulator, PDABFD, Power Aware algorithm, Energy saving algorithm**

I. INTRODUCTION

Cloud computing is the evolution associated with a variety of technologies which may have come together to alter an organization's approach of constructing out an information technology infrastructure. There is nothing in essence completely new within the technologies comprising cloud computing, nonetheless they at the moment are built obtainable together to masses. In recent times software in addition to hardware has developed a lot and such is the advancement of this approach where resources that were isolated earlier are utilized in such a way that one may receive benefits from these as a single virtualized unit [1].

There are various definitions in order to explain cloud computing although not a single one is universally recognized [2]. These various definitions can be summarized as "A pool of virtual resources that are useful and accessible and can be used as resources on demand with or without nominal charges". Cloud computing is very large believed vision associated with providing computing as a service, it is compared to electricity a shared pool associated with pay on demand service. With the advancement of cloud computing, balancing of load between VMs and conservation of energy became key issues.

Load balancing is done in cloud computing in different ways, there are various load balancing algorithms proposed. Energy saving is a key aspect in cloud computing environment for better performance and efficiency of clouds. Whichever technique of load balancing is used saving energy is of utter importance, so a constraint of energy can be put on any of the load balancing algorithm. Process of consolidation of VMs is done for conserving energy in cloud computing environment. Consolidation of VMs is divided into overload detection, under load detection, selecting VM and next VM selection.

Clouds can be divided into Public cloud, Private cloud, Hybrid cloud and Community cloud on the basis of type of storage service they provide. In recent times parallel processing was used. Upcoming advancement in case of parallel processing was multiprogramming where multiple programs were submitted to the processors each taking some bite of processor time. Right after parallel, there was concept of multiprocessing using workload giving. Then vector processing was developed to improve processing in multitasking environment. Sharing of resources as well as management concerning processors is task of vector processing. Cloud computing use hardware virtualization and server virtualization. Hardware virtualization is essential when it comes to reducing and preservation of large number of data stores. To improve sharing and utilization of computers virtualization of memory, input, output and other resources is significant.

Hardware virtualization permits operating several operating systems and also software package on a single physical platform. Hardware virtualization deals with different virtual machines having user software on it utilizing various services, different operating systems with software named hypervisor called virtual machine monitor. With fresh technologies similar to hardware virtualization, live migration of VMs and multi-core chips, virtualization on

server systems is getting more popular. Earlier advantages of virtualization on server systems were rapid sharing, high reliability and superior usage. New benefits are superior migration, workload isolation and consolidation. Workload isolation is accomplished given that most program instructions are usually fully kept inside a VM, which leads to enhancement in security. Reliability is greater because if one particular machine fails other doesn't get impacted and also isolation is improved upon simply because one particular machine doesn't rely on some other while operating in concert.

Numerous platforms are usually integrated at one particular location leading to better utilization of resources. Migrations of workload targets at facilitating load balancing, hardware maintenance and disaster recovery. It is performed through an operating system state within virtual machine and also letting it to be fully serialized, migrated to a different platform, suspended and resumed immediately or preserved to store at a later date. Through the advancement with technology traditional computing has been substituted by cloud computing [3].

This paper proposes a modification of best fit data (BFD) algorithm for selecting next VM depicted as power and data aware best fit data (PDABFD) algorithm. This algorithm uses energy constraint on best fit data and sort all the VMs in the decreasing order of their current CPU utilization and allocates each VM to a host that provides the least increase of power by proper allocation of resources and data awareness.

II.   RELATED WORKS

As cloud computing is computing of resources over a distributed network with virtualization and service level agreements. There are different works done by various researchers which give a useful insight for selecting next VM in cloud computing. This work done is very helpful in understanding what has been done in the field of selecting next VM. Consolidation of VMs is done for conserving energy in cloud computing. Beloglazov et al. [4] proposed energy efficient resource management in virtualized cloud data centers. This is achieved by continuous consolidation of VMs according to current utilization of resources, virtual network topologies established between VMs and thermal state of computing nodes. In this architecture data about request for VMs and their migration, size is collected and used for purpose of energy saving. Placement of VMs and heuristic are only for simplified version of first stage, algorithms for other proposed optimization stages need to be developed and then combined for an algorithm for all stages.

Anandharajan et al. [5] presented co-operative scheduling energy aware load balancing technique for an efficient computational cloud. This works on the principle of grouping similar nodes together and working on these groups. In this paper, energy efficient load balancing was done through estimated ram frequency dynamic voltage scaling. It tells speed is cube root of power hence speed is scaled for in order to perform energy aware scheduling. In this paper load balancing and job scheduling is not done in real time environment which is more dynamic.

Nguyen Quang-Hung et al. proposed [6] genetic algorithm for power aware (GAPA) scheduling of resources for better VM allocation. This algorithm provides a way to reduce energy consumption in VM allocation and solve static VM allocation problem. GAPA is a genetic algorithm which uses tree structure to encode chromosome of an individual. This structure has three levels. First level is root node which doesn't have much meaning. Second level is collection of nodes which consists physical machines and third level is collection of nodes which consists virtual machines. It concluded that energy consumption of best fit algorithm is 130% more than GAPA algorithm. Ts'epoMofolo et al. [7] presented heuristic based VM placement. This VM placement was based upon modified best fit data (MBFD) algorithm. MBFD algorithm takes as input sorted list of VMs to be migrated in descending order of current CPU utilizations and allocate each to the selected host that provide the least number of remaining processing capacity caused\by allocation. This ensures high resource utilization as resource on target will not be idle.

Kyong Hoon Kim et al. [8] proposed power aware provisioning of cloud resources. They defined real time virtual machine which includes three parameters of real time application utilization, million instructions per second and lifetime of real time application. Dynamic voltage scaling model was designed that reduces power consumption of data centers and increase their profit.

Makris at al. [9] proposed networking and computing environment's integration: a novel mobile cloud resource provisioning approach. This explains that mobile network and cloud computing are revolutionary nowadays and newly found integration is mobile cloud computing (MCC). MCC resources are needed to be taken into consideration so this paper gives a scheme for MCC resource provisioning which handles radio and computing resources together rather than handling them individually. This paper does not deal with mobility management and seamless mobile services QoS provisioning challenges.

III.   PROPOSED METHODOLOGY

This model needs local and global managers. Local manager reside on each node of a virtual machine. The global manager resides on the master node and collects information from local managers to maintain overall view of system's resource utilization. Let us assume that at each local manager we have H as set of VMs where each VM V_k has its own processing powers. H can be defined as

$$H=V_k (C_k) \text{ where } k=1, 2 \ldots K \qquad (1.1)$$

Where K is the maximum number of virtual machines, V_k denotes VMs processing powers and C_k denotes their computation powers. Assume the set jobs J of i number where each job j_i has its own data size and processing power. Throughput is expressed as

$$Th=j_i (s_i, q_i) \text{ where } i=1. 2 \ldots.I \qquad (1.2)$$

Where j_i denotes i^{th} job, s denotes data file size of job i and I denotes number of jobs. Global manager maintains a list denoted as elist which keeps track of which physical machine is not capable of running a specific application.

The algorithm needs to fetch the average CPU load, network throughput and RAM. To facilitate the processing, three dimensions are normalized. The complementation of CPU load, network throughput and RAM is the principle and core part of algorithm. To the j-th physical machine, suppose that it can carry k virtual machines. The three load dimensions mentioned above are expressed as *cl, nt* and *mem* respectively. Therefore for the virtual machines carried on the j-th physical machine their load dimensions can be expressed as three arrays: $[cl_{j1}, cl_{j2}, cl_{jn}]$, $[nt_{j1}, nt_{j2}, nt_{jn}]$ and $[mem_{j1}, mem_{j2}, mem_{jn}]$. For the convenience of data process, all the values are normalized into interval of [0, 1].

$$\text{Avg CPU load } (cl_j) = cl_{ji} \text{ for } i=0 \text{ to } K \qquad (1.3)$$

$$\text{Avg net Th } (nt_j) = nt_{ji} \text{ for } i=0 \text{ to } K \qquad (1.4)$$

$$\text{Avg mem } (mem_j) = mem_{ji} \text{ for } i=0 \text{ to } K \qquad (1.5)$$

From equations (1.3), (1.4), (1.5) variance of average values on physical machine can be calculated as

$$Var_j = \{avg\ cl_j - (avg\ cl_j + avg\ nt_j + avg\ mem_j)^3\}^2 + \{avg\ nt_j - (avg\ cl_j + avg\ nt_j + avg\ mem_j)^3\}^2 + \{avg\ mem_j - (avg\ cl_j + avg\ nt_j + avg\ mam_j)^3\}^2 \qquad (1.6)$$

Equation (1.6) denotes the variance of average values of j-th physical machine. Now addition of variance of all physical machines will provide overall variance denoted by V_i.

$$V_i = \sum Var_j \qquad (1.7)$$

Obviously, the variance in equation (1.7) represents the deviation of the load dimensions represented by i-th VM placement scheme. If the value of variance is smaller, the average utilization of load is better. A large value of variance indicates that the loads are unbalanced

$$Th = [\sum (V - V_i)\ V] / [\sum (V - V_j)\ V] \qquad (1.8)$$

Proposed Algorithm

This algorithm presents modification of the best fit data (BFD) algorithm. Proposed algorithm put power and data constraint together on BFD algorithm for better energy conservation and resource utilization. This algorithm sorts all VMs according to their decreasing order of CPU utilization and allocates each VM to a host (next VM) that decreases power consumption by proper allocation of resources and data awareness.

Algorithm PDABFD

1. Sort VM (1, 2 …k) in order of their decreasing CPU utilization
2. **For** every V_i in V (1, 2 …k) perform
   minPower← Max
   allocatedhost← elist[]

3. **For** every M_j in M (1, 2 ...n) perform
   If M_j has enough resources for V_i **then**
   Power← estimatePower (M_j, V_i)
4. **If** Power< minPower **then**
   allocatehost← host
   minPower← power
5. **Elseif** allocatehost= NULL **then**
   Add (allocatehost, V_i) to NextVM
6. **return** NextVM

IV.   CONCLUSION

In this paper, a modification to BFD algorithm for selecting next VM is given. This modified algorithm emphasizes more on power conservation and data in data centers. This algorithm minimizes power consumption by providing a constraint for power and data. Simulation of algorithm in CloudSim simulator shows better resource utilization and thus less power consumption. Power saving is done by proposed algorithm by efficient consolidation of VM. We believe that this work can be further extended and additional functionalities can be added to it.

paper. Also, without the grace of GOD and our parents this work wouldn't have been possible.

REFERENCES

[1] Xueping lio, "Data security in Cloud Computing", International Conference on cybernetics and informatics, vol. 163, pp. 801-806, 2012.

[2] Lizhe Wang, Gregor von Laszewski, Andrew Younge, Xi He, Marcel Kunze, Jie Tao, Cheng Fu, "Cloud Computing a perspective study", vol. 28, Issue 2, pp. 137-146, April 2010.

[3] George Reese, *Cloud Application Architectures*, 2009.

[4] Anton Beloglazov, Rajkumar Buyya, "Energy Efficient Resource Management in Virtualized loud Data Centers", Cluster Cloud and Grid Computing (CCGrid), 10th IEEE ACM International Conference, pp. 826-831, May 2010

[5]   T.R.V. Anandharajan, M.A. Bhagyaveni, "Co-operative Scheduled Energy Aware Load-Balancing Technique for an Efficient Computational Cloud", IJCSI International Jouranl of Computer Science Issues, vol. 8, Issue 2, March 2011.

[6] Nguyen Quang-Hung, Pham Dac Nienz, Nguyen Hoai Namz, Nguyan Huynh Tuong, Nam Thoa, "A Genetic Algorithm for Power-Aware Virtual Machine Allocation in Private Cloud", Information and communication technology, vol. 7804, pp. 170, Feb 2013.

[7] Ts'epoMofolo, R. Suchithra, "Heuristic Based Resource Allocation Using Virtual Machine Migration: A Cloud Computing Perspective", International Journal of Computer Application Technology and Research, vol. 2, Issue 6, pp. 731-736, 2013.

[8] Kyong Hoon Kim, Anton Beloglazov, Rajkumar Buyya, "Power-Aware Provisioning of Cloud Resources for Real-time Services", International Workshop on Middleware for Grids, Clouds and e-services, ISBN: 978-160558-884-6, 2009.

[9] Prodromos Makris, Dimitrios N. Skoutas, Charalabos Skianis, "On networking and computing environments' integration: a novel mobile cloud resources provisioning approach", International conference on telecommunication and multimedia (TEMU), pp. 71-76, Aug 2012.

[10] Anton Beloglazov, Jemal Abawajy, Rajkumar Buyya, "Energy-Aware resource allocation heuristic for efficient management of data centers for cloud computing", Future generation computer systems, pp. 755-768, 2012.

[11] Esha Barlaskar, N. Ajith Singh, Y. Jayanta Singh, "Energy optimization methods for Virtual Machine Placement in Cloud Data Center", Journal of engineering technology, vol. 1, Issue 1, pp. 23, 2014.

[12] Jiankang Dong, Xing Jin, Hongbo Wang, Yangyang Li, "Energy Saving Virtual Machine Placement in Data Centers", Cluster, Cloud and Grid Computing (CCgrid), pp. 618-624, 2013.

[13] A. Singh, M. Hemalatha, "Virtual Machine Placement by Using Honeybee Forager Algorithm in Cloud Computing", Karpagam Journal of Computer Science, vol. 7, pp. 209, 2013.

[14] Sukhpal Singh, Inderveer Chana, "Energy based Efficient Resource Scheduling: A step towards Green Computing, International Journal of Energy, Information and Communications, vol. 5, Issue 2, pp. 35-52, 2014.

[15] Sukhpal Singh, Inderveer Chana, "Energy based Efficient Resource Scheduling: A Step towards Green Computing", International Journal of Energy, Information and Communications, vol. 5, Issue 2, pp. 35-52, 2014.