ISSN No. 0976-5697
Volume 3, No. 2, March-April 2012
International Journal of Advanced Research in Computer Science
RESEARCH PAPER
Available Online at www.ijarcs.info

Energy consumption in cloud computing and power management

Prasanta Kumar Panda*
Dept. of Computer Science & Engineering
NMIET, BPUT, Bhubaneswar, India,
silu143.panda@gmail.com

Shrabanee swagatika
Dept. of Computer Science & Engineering
SOA University Bhubaneswar, India
shrabanee.swagatika@iter.ac.in

Abstract- "Cloud Computing" is becoming increasingly relevant, as it will enable companies involved in Spreading this technology. Network-based cloud computing is rapidly expanding as an alternative to conventional office-based computing. As cloud computing becomes more widespread, the energy consumption of the network and computing resources that underpin the cloud will grow. This is happening at a time when there is increasing attention being paid to the need to manage energy consumption across the entire information and communications technology. While data centre energy use has received much attention recently, there has been less attention paid to the energy consumption of the transmission and switching networks that are key to connecting users to the cloud. In this paper, we present an analysis of energy consumption in cloud computing and provide a solution for energy efficient cloud utilisation.

Keywords- Virtual machine, Qos, Computing resources, Green cloud computing, IAAS.

I. INTRODUCTION

In the context of getting dynamics computing services on demand the cloud computing plays a significant role. As the three different services provided by cloud computing like software as a service (SAAS), platform as a service (PAAS) & infrastructure as a service (IAAS) almost covers all applications. So gradually it becomes a very popular and on demand application approach to work with for. But it needs a huge amount of energy to work with which not only decreases the amount of static source of energy but also generates a huge amount of carbon foot prints which pollutes the environments. Thus we need green cloud computing as a solution which not only minimizes the operational cost but also reduce environmental impact [1]. In this paper we suggest an architecture framework for energy efficient cloud computing management.

Our proposed work suggest to reduce the energy by first searching for the virtual machine service(VMS) providing same kind of application for which their resource requirement performed by the data centers can also be minimized by keeping the extra sleep mode by distributing their workload to lower or lightly loaded systems. The suggested solution not only reduce the cost also energy and in spite of any negotiation with the quality of service in particular.

II. RELATED WORK

One of the first works, in which power management has been applied at the data center level, has been done by Pinheiro et al. [2]. In this work the authors have proposed a technique for minimization of power consumption in a heterogeneous cluster of computing nodes serving multiple web-applications. The main technique applied to minimize power consumption is concentrating the workload to the minimum of physical nodes and switching idle nodes off. This approach requires dealing with the power/performance trade-off, as performance of applications can be degraded due to the workload consolidation. Requirements to the throughput and execution time of applications are defined in SLAs to ensure reliable QoS. The proposed algorithm periodically monitors the load of resources (CPU, disk storage and network interface) and makes decisions on switching nodes on/off to minimize the overall power consumption, while providing the expected performance. The actual load balancing is not handled by the system and has to be managed by the applications. The algorithm runs on a master node, which creates a Single Point of Failure (SPF) and may become a performance bottleneck in a large system. In addition, the authors have pointed out that the reconfiguration operations are time-consuming, and the algorithm adds or removes only one node at a time, which may also be a reason for slow reaction in large-scale environments. The proposed approach can be applied to multi-application mixed-workload environments with fixed SLAs. Chase et al. [3][4] have considered the problem of energy-efficient management of homogeneous resources in Internet hosting centres. The main challenge is to determine the resource demand of each application at its current request load level and to allocate resources in the most efficient way. To deal with this problem the authors have applied an economic framework: services ''bid'' for resources in terms of volume and quality.

This enables negotiation of the SLAs according to the available budget and current QoS requirements, i.e. balancing the cost of resource usage (energy cost) and the benefit gained due to the usage of this resource. The system maintains an active set of servers selected to serve requests for each service. The network switches are dynamically reconfigured to change the active set of servers when necessary. Energy consumption is reduced by switching idle servers to power saving modes (e.g. sleep, hibernation). The system is targeted at the web workload, which leads to a ''noise'' in the load data. The authors have addressed this problem by applying the statistical ''flip-flop'' filter, which reduces the number of unproductive reallocations and leads to a more stable and efficient control.

The proposed approach is suitable for multi-application environments with variable SLAs and has created a foundation for numerous studies on power-efficient resource allocation at the data center level. However, in contrast to [5][6] the system deals only with the management of the CPU, but does not consider other system resources. The latency due to switching nodes on/off also is not taken into account. The authors have noted that the management algorithm is fast when the workload is stable, but turns out to be relatively expensive during significant changes in the workload [7]. Moreover, likewise diverse software configurations are not handled, which can be addressed by applying the virtualization technology.

A. Understanding Power Consumption:

Understanding the relationship between power consumption, CPU utilization and the transition delay between different server's states is essential to design efficient strategies for energy savings. We examined this relationship by measuring power consumption of typical machines in different states. We also measured power consumption of pre-defined low power states, including shutdown, hibernate, suspend-to disk, suspend-to-RAM, power-on-suspend with the Linux machine, and shutdown, hibernate, standby with the Windows machine. So now are aim is to our aim is to Lowering the energy usage by efficient processing and utilization of a computing infrastructure and also to minimize energy consumption. The system should satisfy QOS & fulfills service level agreement and also reduce energy consumption. But also be acceptable in the view of Commercial, scientific & technological aspect [8] [9].

a. Steps to fulfill our aim.
   a) Define architecture & framework
   b) Energy aware
   c) Must be able to self manage power obtained.
   d) Suitable mapping of VMS to cloud resources.
   e) Explore challenges in doing so.

B. Green Cloud Architecture:

To design next generation data centers by adding them with capabilities like network for virtual services and Virtual services includes hardware, database, user interface, application logic [10].

So that user can access and deploy any application from anywhere in the world on demand at a competitive price depending on the quality of service requirements. Main 4 entities involved

a. Consumer or brokers
b. Green service allocator
   a) Green service negotiator
   b) Service analyzer
   c) Consumer profile
   d) Pricing
   e) Energy monitor
   f) Service scheduler
   g) VM manager
   h) Accounting

c. VMS:

Dynamically migrate VMS across physical machines, workloads can be consolidated & unused resources can be switched to a low power mode.

d. Physical machines:

The services whose actual hardware helps for creating virtualized resources to meet service demand.

III. OUR PROPOSED WORK

We proposed techniques to inform power management at data centre level for those of homogeneous cluster of computing nodes serving kind of similar applications.

We proposed to do so via….

a. By concentrating to distribute the total workload to a minimum number of physical nodes so that the other idle computing nodes can be switch off or in sleep mode.
b. Here we proposed to save the amount of energy investment on power per performance tradeoff.
c. Before the distribution of application the internet hosting center must determine first the amount and type of resources demanded of each application at its current provided load level.
d. Then an efficient technique will be used to verify the type of resource held by different data centers or net host [11].
e. Then the VMs which serving the either moderate or lightly weighted application start distributing or transferring their workload to the other terminals for which we proposed an minimization of migration algorithm, this algorithm transfer the workload of some VMs to other depending on their resource type available & the kind of work they are performing with.
f. Then after the transfer of workload is finished the idle VMs & datacenter will be switched off or turned off or hibernate or can be move on to sleep mode.
g. Thus reducing energy consumption.

Among all active VMS need to find out their resources requirement for full filling the execution of their appropriate applications [11]. But we have to chosen such VMS which only serves same type of application so can they minimize their migration position to different net host provider.

Except that we need to go for a minimization of migration algorithm which will arrange the VMS in increasing order of their migration so that the system used for the less migration would provided with more work load and by doing so we can consume the energy utilization [12].

A. Algorithm:

Our algorithm includes basically two phases:-

The first VM should have the utilization higher than the threshold capacity previously fixed by the data centre. Second if at worst migration is being needed then the power consumption difference between two host changes must be minimum one. Here the complexity of the algorithm is

proportional to the product of the no of over utilization net host & allocated active VMS.
Steps:
Need to check every VM from the VM list:

a) For each VM from VM list, check its application & resources requirement.
b) Map those requirements with appropriate net host from net host list.
c) Then arrange the VM in decreasing order of their utilization & type of requirement.
d) With the maximum type of resource requirement VMs are the output VMs with maximum migration.

Add the migrated VMs to migrated list to check for potential growth policy. Here the highest potential growth policy migrates VMs that have the lowest usage or the CPU relatively to the CPU capacity defined by the VM parameters in order to minimize the potential increase of the hosts' utilization & prevent service level agreement violation.

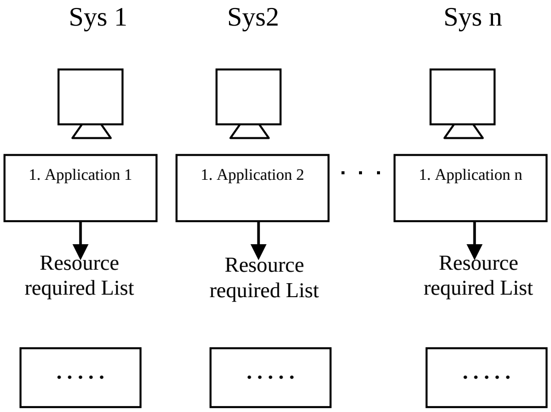

Fig. 1

The total work load can be provided on to the active working nodes on a specific application. We can just distribute the workload.

"Fig. 1" the system 3 can transfer its workload onto system 1 & 2 respectively and system 1 & 3 can transfer its workload on to system 1 & 2 respectively then system 1 & n will be free & can be switched off. Then the energy associated with system 3 & n be saved by turning them off.

IV. CHALLENGES

A. How to recognize what are the systems ready to transfer workload.
B. How to identify which system is running what applications.
C. Synchronization follows which algorithm?

a) Solution for challenge (A):

Every system must be provided with a timer which will tell us what amount of time each system is been worked for till. Then after the recognition of the application and associated systems are identified the system started transferring the load to other depending on the time period which is larger in comparing to that other system.

In this context two things will happen. First energy will be saved which is our objective and second the system working for a larger period will take rest so that better resource like CPU, I/O will get best utilization.

Note:

Before transforming the load the system capacity and functionality must be better checked.

b) Solution for challenge (B):

For knowing which system is running what application a register program must be run within the master processor. It will keep track of all information regarding which system is running what application.

At the time of different applications are coming for fetch into different system, they need to take following two steps. First the application will match their requirement with acquired resource list and type obtained by each system. Then the application which suited most appropriate system will fetch in.

c) Solution for challenge (C):

The last challenge for the synchronization purpose. Here the system involved need to go for a best synchronization algorithm for better handling of challenges A & B.

a. Depending on the programmer can go for any suited synchronization algorithm which will best fit on to their requirement and demand. Various synchronization algorithms are already available as processor pool methods.

V. CONCLUSION

In this paper we presented a Green solution for energy efficient cloud which makes good use of cloud resources results in savings in Cloud computing. The prediction here is to make turning off/on decisions to minimize the number of running servers. In order to demonstrate the solution, we have performed simulations with different parameters and running modes.

In future work, we plan to compare the presented solution with other power management schemes which employ different load prediction mechanisms. The system model for the solution should be extended to deal with a more diversity of workloads and application services, as well as architectures of data centres for a better simulation of cloud environments. A deployment of the solution in real server farms to show its efficiency in a real setting is also worth considering.

VI. REFERENCES

[1] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I.Pratt, A.Warfield, Xen and the art of virtualization, in: Proceedings of the 19th ACM Symposium on Operating Systems Principles, SOSP pp. 177, 2003, 2003.

[2] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, A view of cloud computing, Communications of the ACM 53 volume-4, 2009 pp 50–58.

[3] R. Buyya, A. Beloglazov, J. Abawajy, Energy-efficient management of data center resources for cloud computing: a vision, architectural elements, and open challenges, in: Proceedings of the 2010 International Conference on Parallel and Distributed Processing Techniques and Applications, PDPTA 2010,Las Vegas, USA, 2010.

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

[5] 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.

[6] E. Elnozahy, M. Kistler, R. Rajamony, Energy-efficient server clusters, Power- Aware Computer Systems (2003) pp 179–197.

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

[8] 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 volume 1, 2008 pp 48–59.

[9] D. Kusic, J.O. Kephart, J.E. Hanson, N. Kandasamy, G. Jiang, Power and performance management of virtualized computing environments via look ahead control, Cluster Computing 12 (1) (2009) pp 1–15.

[10] S. Srikantaiah, A. Kansal, F. Zhao, Energy aware consolidation for cloud Computing, Cluster Computing 12 (2009) pp 1–15.

[11] 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.

[12] Design issues of simple cloud based software as a service using 3G and 4G technology by Shrabanee Swagatika, Dr. Prasanta Kumar Pattnaik ,International Journal of Instrumentation, Control & Automation (IJICA), Volume 1, Issue 1, 2011, pp-11 - 14 .

