

A Solution to Optimize the Resources Present in the Cloud via Green Cloud Meeting Service Level Agreement

P.Elango, PhD.
Department of Information Technology
PKIET, Karaikal – 609 603.

M. Jannani
M. Tech. Scholar, Department of CSE
PEC, Puducherry – 605 014.

ABSTRACT

The Green cloud meeting the SLA (Service level agreement) in Cloud Computing presents an interesting and economical solution for optimizing the resources present in the cloud. The idea of this paper is to implement a cloud architecture that optimizes the energy consumed in the cloud without compromising the performance of high profile users, taking the energy consumption of the processor of a node controller under consideration. The Cloud controller optimizes the virtual machines accordingly to make the processor in low power state. The basics of Cloud Computing is briefly discussed. The architecture of the Green cloud system and its modules are discussed in detail. The Pseudo code is discussed for the Energy save calculator. The test cases with their expected results are listed.

General Terms

Cloud Computing, Resource Optimization, High Speed networks.

Keywords

Green Cloud, Virtual machines, Service Level Agreement and Energy Efficiency

1. INTRODUCTION

In the current data-centric world, all the data are uploaded into the cloud and the energy consumption of data centers has been a major problem that creates the carbon print in the environment. Hence it is high time that we take the necessary steps to optimize the resources present in the cloud. It is at this juncture that the term 'Green cloud' is coined. It is an environment-friendly approach to bring out a new paradigm shift to the cloud.

Cloud computing is a general term for anything that involves delivering hosted services over the Internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS). A cloud service has three distinct characteristics that differentiate it from traditional hosting. It is sold on demand, typically by the minute or the hour; it is elastic -- a user can have as much or as little of a service as they want at any given time; and the service is fully managed by the provider (the consumer needs nothing but a personal computer and Internet access). Significant innovations in virtualization and distributed computing, as well as improved access to high-speed Internet and a weak economy, have accelerated interest in cloud computing.

2. LITERATURE REVIEW

To implement a cloud architecture that optimizes the energy consumed in the cloud without compromising the performance of high profile users. Modern high-end computing can provide high performance computing solutions for scientific and engineering applications. However, today's high performance computers consume tremendous amounts of energy. For example, a 360-Tflops supercomputer (such as IBM Blue Gene/L) with conventional processors requires 20 MW to operate, which is approximately equal to the sum of 22,000 US households power consumption [1], [2]. Furthermore, it is estimated that the servers consume 0.5 percent of the world's total electricity usage [3], of which if current demand continues, is projected to quadruple by 2020. Unfortunately, the computing system temperature may increase rapidly due to inefficient cooling equipments. Based on Arrhenius time-to-fail model [4], every 100 degree C increase of temperature leads to a doubling of the system failure rate. Hence, it is obvious that power-aware resource management for high end computing is highly desirable. Dynamic voltage and frequency scaling (DVFS) [5] is an efficient technology to control the processor power consumption. With aid of supporting technologies such as Intel Speed Step and AMD Power now modern processors can be operated in several frequencies with different supply voltages.

Virtual machine technology is adopted for high end computing to achieve efficient computing resource usage. Some technical work had reported that virtual machine could be used for scientific applications with tolerable performance punishment [6], [7]. A Virtual Machine (VM) is a software based machine emulation technique to provide a desirable, on demand computing environments for users. As documented in [8], [6], [9], virtual machine provisioning is a popular part of cluster deployments. The normal process of a cluster operating with the use of virtual machines for executing jobs is shown as follows:

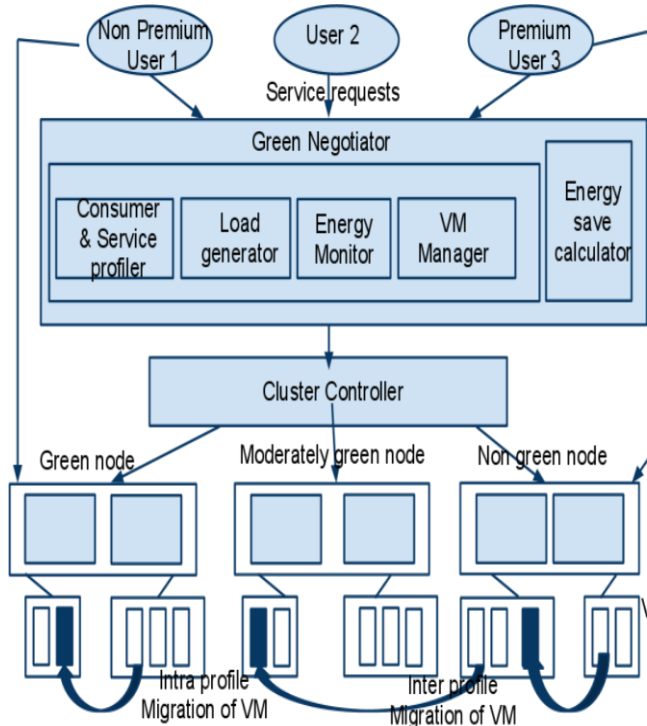
- 1) A compute cluster provides various virtual machine templates.
- 2) When a job arrives at the cluster, the cluster scheduler allocates the job with a preconfigured virtual machine and then starts it on proper compute nodes.
- 3) The job is executed in the virtual machine.
- 4) After the job is executed, the virtual machine is shutdown.

3. DETAILED DESCRIPTION

The idea of this paper is to aim at the integration of green characteristics of a cloud along with maintaining service level agreements between the customer and the cloud service provider. For this, three profiles of users are created: Premium, Average and

Non-premium users. Premium users are users with highest priority in providing service. Hence all jobs with comparatively lower priority have to be preempted and the new service has to be allocated with a new virtual machine.

3.1. ARCHITECTURAL DIAGRAM OF THE PROPOSED GREEN CLOUD SYSTEM:



- The parameter that is considered for handling different test cases is CPU usage. CPU usage is the percentage of CPU currently under utilization.
- The user requests are handled by a web interface. The user logs in into the system with his/her own username and password. Then he provides the service he needs i.e. whether web browsing or graphics intensive application.
- This request reaches the green negotiator. Green negotiator is at the heart of the project that is implemented in cloud or cluster controller.
- Under cluster controller, a predefined number of node controllers are present. The services are handled by the virtual machines that are created under these node controllers. Each node controller can have any number of virtual machines thus handling any number of service requests. The node controllers are defined according to their characteristics as follows:
 - Green nodes
 - Moderately green nodes
 - Non-green nodes
- Every node controller is in two states according to their energy usage:
 - Low power state: Every node controller will be allocated with a separate processor running in its maximum frequency. It has been found in research that a processor with CPU usage of 0% consumes 70% of energy as consumed when its CPU usage is

100%. Hence energy is wasted when the CPU is not in usage. Hence when the node controller is not allocated with any VM, then its CPU is brought to low energy consumption state by reducing the frequency of CPU. That is, if a CPU runs at 2.2 GHz of CPU cycle speed, then reducing its frequency to 1.2 GHz makes the CPU to use less amount of energy and brings it to low power state.

- High power state: A node controller in turn a CPU is in high power state if the frequency at which the processor runs is at its maximum level. There are many tools that help reduce the frequency cycles of CPU. Eg. CPU frequency
- Green nodes are a group of physical nodes i.e. node controllers in which virtual machines are created for the services from non-premium users. The special characteristic of green node is that when the CPU usage of a green physical node reaches a particular lower threshold because of the addition of more no. of virtual machines into it, then all virtual machines running under this node controller are suitably migrated and distributed to other node controllers under the green profile. Since there is no virtual machine running under the previous node controller, this node controller is put in low power state (refer low power state).
- In green node, maximum numbers of virtual machines are made to run in minimum number of node controllers. But non green nodes are created for the purpose of bringing into picture the service level agreement, i.e. performance plays much more important role as compared to green characteristics. Hence in non-green nodes, the virtual machines are distributed among non-green node controllers so that every virtual machine can use maximum CPU usage.

3.2. MODULES DESCRIPTION

The following are the modules in this project:

- Consumer and Services profiler
- Load generator
- Energy monitor
- VM manager
- Energy save calculator

3.2.1. CONSUMER AND SERVICES PROFILER:

This maintains the details of the users and their profiles. There are three profiles namely Premium user, average user and non-premium user.

1. Add new user: Get the username and desired profile and save the details.
2. Delete a user
3. Maintain services available and the CPU usage needed by the service.

3.2.2. LOAD GENERATOR:

Input: Username, Service Requested, Service time.

Output: If the service can be accommodated, a new virtual machine is generated for the service in a physical node. Otherwise the service is rejected.

Description:

This module creates a dummy load into the virtual machine. For example, if a green user selects web browsing, then it gets the suitable physical node from VM manager module, and then creates a dummy load, say 20% of CPU usage, in a newly allocated virtual machine under that physical node. These dummy loads can be created using Faban load generator.

3.2.3. ENERGY MONITOR:

Description:

The module checks the average CPU usage of physical machines for every 5 seconds. If they reach the upper threshold, some VMs are migrated to the other physical nodes else if they reach the lower threshold, all VMs are migrated to other physical nodes under same profile if possible and then the machine is run in low power state.

3.2.4. VM MANAGER:

Description:

1. Maintains the details about the VMs currently under different physical machines and the average CPU usage of each VM for previous 5 seconds.

2. Selects the physical node suitable for a new service request or to accommodate a VM already allocated in another physical node.

Input: Customer profile, CPU usage

Output: A physical node n for new VM

3.2.5. ENERGY SAVE CALCULATOR:

Description:

The actual energy used is directly proportional to frequency of processors present in physical nodes. Hence reducing frequency reduces energy used. The energy saved by the optimized cloud is calculated by the decrease in frequency of physical nodes at free time.

Pseudo code:

For every 1 time unit

```
{ For every physical node n
  {i=i+1
```

Totalfrequency(n)=Totalfrequency(n)

+ Current frequency(n)

```
}
}
```

Average frequency(n)=Total frequency(n)/i

Frequency saved=Maximum Frequency(n) * i – Average frequency(n)

3.3. TEST CASES:

TEST CASE ID	INPUT	EXPECTED RESULT
1.	<p>Current state of physical nodes in Green profile</p> <p>Physical node 1: High</p> <p>VMs with CPU usage: 20%,30%,40%</p> <p>Total CPU usage: 90%</p> <p>Physical node 2: High</p> <p>VMs with CPU usage: 1a. 30% 1b. 55%</p> <p>Total CPU usage: 1a. 30% 1b. 55%</p> <p>Action: VM with 40% usage completes its job in PN 1</p>	<p>1a. VM in PN 2 is checked whether it can be accommodated in PN 1. If so, the VM is migrated to PN1. PN2 is brought to low-power state.</p> <p>1b. After VM with 40% completes, CPU usage of PN1=50%. Now VM in PN2 with 55% usage cannot be accommodated. So it remains in the same PN.</p>
2.	<p>Current state of non-green nodes</p> <p>PN1: CPU usage 80% with upper threshold=70%.</p>	<p>select the optimum VM so that usage non green nodes is within the threshold</p> <p>and then select the optimum physical node in green profile to migrate the VM</p>
3.	<p>Current state of green nodes:</p> <p>PN1 : 90%</p> <p>Action: A new service with expected CPU usage of 20%</p>	<p>The service is put in wait queue.</p>

4. CONCLUSION

The idea of this paper is to take the energy consumption of the processor of a node controller under consideration. The Cloud controller optimizes the virtual machines accordingly to make the processor in low power state. Virtual Machine (VM) technology has been widely applied in data center environments due to its seminal features, including reliability, flexibility, and the ease of management. We present the GreenCloud architecture, which aims to reduce data center power consumption, while guarantee the performance from users' perspective. Green Cloud architecture enables comprehensive online-monitoring, live virtual machine migration, and VM placement optimization. This can be extended to be used in various other resources like memory, storage, etc. In the future, there are still a number of research activities that we plan to carry out, which could improve the performance of GreenCloud and bring solid value to users to achieve their business goals and their social responsibility in Green IT.

5. REFERENCES

- [1] A. Gara, M. and and .A. Blumrich "Overview of the Blue Gene/L system architecture," *IBM Journal of Research and Development*, vol. 49, no. 2-3, pp. 195–212, 2005.
- [2] K. Li, "Performance Analysis of Power-Aware Task Scheduling Algorithms on Multiprocessor Computers with Dynamic Voltage and Speed," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 11, pp. 1484–1497, 2008.
- [3] W. Forrest, "How to cut data centre carbon emissions?" Website, December 2008. [Online]. Available: <http://www.computerweekly.com/Articles/2008/12/05/233748/how-to-cut-data-centre-carbon-issions.htm>.
- [4] P. W. Hale, "Acceleration and time to fail," *Quality and Reliability Engineering International*, vol. 2, no. 4, pp. 259–262, 1986.
- [5] G. Magklis, G. Semeraro, D. Albonesi, S. Dropsho, S. Dwarkadas, and M. Scott, "Dynamic frequency and voltage scaling for a multiple-clock domain microprocessor," *IEEE Micro*, vol. 23, no. 6, pp. 62–68, 2003.
- [6] L. Wang, M. Kunze, and J. Tao, "Performance evaluation of virtual machine based Grid workflow system," *Concurrency and Computation: Practice and Experience*, vol. 20, no. 15, pp. 1759–1771, 2008.
- [7] W. Huang, J. Liu, B. Abali, and D. K. Panda, "A case for high performance computing with virtual machines," in *Proceedings of the 20th Annual International Conference on Supercomputing*, 2006, pp. 125–134.
- [8] V. Buge, Y. Kemp, M. Kunze, and G. Quast, "Application of Virtualisation Techniques at a University Grid Center," in *e-Science*, 2006.
- [9] N. Fallenbeck, H. Picht, M. Smith, and B. Freisleben, "Xen and the art of cluster scheduling," in *First International Workshop on Virtualization Technology in Distributed Computing*, 2006, pp. 4–4.
- [10] G. Magklis, G. Semeraro, D. Albonesi, S. Dropsho, S. Dwarkadas, and M. Scott, "Dynamic frequency and voltage scaling for a multiple-clock domain microprocessor," *IEEE Micro*, vol. 23, no. 6, pp. 62–68, 2003.
- [11] L. Wang, M. Kunze, and J. Tao, "Performance evaluation of virtual machine based Grid workflow system," *Concurrency and Computation: Practice and Experience*, vol. 20, no. 15, pp. 1759–1771, 2008.
- [12] W. Huang, J. Liu, B. Abali, and D. K. Panda, "A case for high performance computing with virtual machines," in *Proceedings of the 20th Annual International Conference on Supercomputing*, 2006, pp. 125–134.
- [13] V. Buge, Y. Kemp, M. Kunze, and G. Quast, "Application of Virtualisation Techniques at a University Grid Center," in *e-Science*, 2006.
- [14] N. Fallenbeck, H. Picht, M. Smith, and B. Freisleben, "Xen and the art of cluster scheduling," in *First International Workshop on Virtualization Technology in Distributed Computing*, 2006, pp. 4–4.
- [15] Y. Wiseman and D. Feitelson, "Paired gang scheduling," *IEEE Transactions on Parallel and Distributed Systems*, vol. 14, no. 6, pp. 581–592, 2003.
- [16] D. Marr, F. Binns, D. Hill, G. Hinton, D. Koufaty, J. Miller, and M. Upton, "Hyper-threading technology architecture and microarchitecture," *Intel Technology Journal*, vol. 6, no. 1, pp. 4–15, 2002.
- [17] K. Yu, "Xen PM," Website, March 2009. [Online]. Available: <http://wiki.xensource.com/xenwiki/xenpm>.
- [18] J. Fontan, T. Vazquez, L. Gonzalez, R. S. Montero, and I. M. Llorente, "OpenNEBula: The Open Source Virtual Machine Manager for Cluster Computing," in *Open Source Grid and Cluster Software Conference, San Francisco, CA, USA, May 2008*.