Dynamic Resource Allocation using Virtualization Technology in Cloud Computing

Miss. Gawali Anita D. Prof. S. K. Sonkar

Abstract— **Cloud computing is ubiquitous and promise a cost-effective realization of the utility computing principle, allowing users and providers easy access to resources in a self-service, pay-as-you-go fashion, thus decreasing cost for system administration and improving resource utilization and accounting. It allows business users to scale up and down the resource usage according to their need of use. In this paper we give the dynamic Resource allocation strategies and green computing technology to improve the performance.**

Index Terms— **Cloud Computing, Resource Management, Virtualization, Green Computing, Hot a d Cold Spots, .**

I. Introduction

CLOUDs are dynamic environment that guarantee availability, reliability and related quality aspects through automated, elastic management of the hosted services. The services provided by the cloud are- Infrastructure as a Service (IaaS), Platform as a Service (PaaS) or Software as a Service(SaaS) .The automated management thereby aims at optimizing the overall resource utilization whilst maintaining the quality constraints. Cloud Computing consist of the Distributed processing, parallel processing and grid computing [3]. The basic principle of cloud computing is that user data is not stored locally but is stored in the data center of internet. So the user can access the data anywhere at any time by using API which is provided by the cloud. The most popular definition is probably the one provided by the National Institute of Standards and Technology (NIST) "Cloud computing is a model for enabling ubiquitous, convenient, on demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [3]. Resource allocation is a process of assigning the resource based on the application demands by using resource provisioning over the internet.

In this paper we are trying to achieve two goals-Overload avoidance and green computing.

In overload avoidance is achieved by minimizing the number of servers used and Green computing by shutting down underutilized servers.

Manuscript received June, 2015.

Miss. Gawali Anita D*, Departmant of Computer Engineering, Amrutvahini College of Engineering Sangamner, India,*

Prof. S. K. Sonkar*, Departmant of Computer Engineering, Amrutvahini College of Engineering.,Sangamner, India.*

II. Literature Survey

Cloud systems with virtualization open a new opportunity to widen the scope of contention-aware scheduling, as virtual machines can cross legacy system boundaries with live migration. In this paper, we use live VM migration to dynamically schedule VMs for minimizing the contention on shared caches and memory controllers[4]. Furthermore, this study considers the effects of non-uniform memory accesses (NUMA) in multi socket systems commonly used in cloud servers. Contention-aware cloud scheduling techniques is used for cache sharing and NUMA affinity. The techniques identify the cache behaviors of VMs on-line, and dynamically migrate VMs, if the current placements of VMs are causing excessive shared cache conflicts or wrong NUMA affinity. Since the techniques identify the VM behaviors dynamically and resolve conflicts with live migration it will not required any prior knowledge on the behaviors of VMs[5]. The first technique, cache-aware cloud scheduling minimizes the overall last-level cache (LLC) misses in a cloud system. The second technique, NUMA aware cloud scheduling extends the first technique by considering NUMA affinity.

Zhen Xiao [1] gives the strategy for dynamic resource allocation with Skewness and load prediction algorithm. He uses Xen hypervisor Usher controller. The merits in his system are no overheads, high performance. It requires less number of migrations and residual resource is friendly to virtual machines. It improves the scheduling effectiveness. The demerit of the system is it is not cost effective.

T.Wood [2] gives the Black and Grey box strategies with BG algorithm. He used Xen hypervisor and finds with Nucleous and monitoring engine, Grey-box enables proactive decision making. While it has the limitation as, Black-box is limited to reactive decision making and BG algo requires more number of migrations.

III. System Architecture

The system architecture is described in the following figure.VM scheduler consists of four components- Predictor, Hot spot solver, cold spot solver and migration list[1].

Predictor collects the historical statistics and predicts current and future resource needs of VMs and ultimately the load on PMs. Hot Spot Solver observes the PMs and compare its utilization with the hot threshold temperature. If it is above the hot threshold then it migrate it away by sending this case to migration list otherwise to the cold spot solver for further observations[6]. Cold spot solver does the same with PM resource utilization by comparing it with cold spot threshold or green computing threshold and if it found to be less than it that PM is migrated to the efficient one PM and original PM is potentially turns off to achieve green computing. Migration list is compiled and result is sent to the Usher CTRL for further execution.

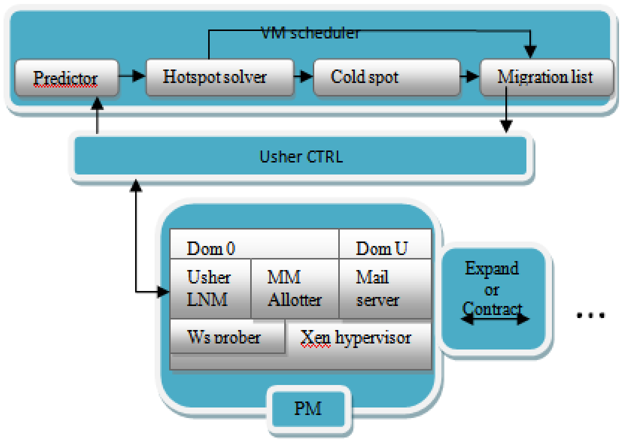

Fig. System Architecture

Then each PM runs Xen hypervisor which is a Virtual Machine Monitor. It has privileged domain 0 and more than one domain U. Domain U can run more applications like web server, remote desktop, mail, Dns, etc.

In the Usher framework, the multiplexing of VMs to PMs is well managed. On domain 0 of Xen hypervisor, each PM runs a Usher local node manager (LMN). This LNM aggregates all the statistics of resource on particular node. For example, the CPU and networking resource utilization of a PM is calculated by monitoring scheduling events on it[7].

This statistics is then sent to the Usher CTRL (Usher Central Controller) where scheduler is running. The scheduler is invoked after some period of time and these observed statistics are compared and further managed by several Scheduling components.

Our system consist of following modules-
1. Cloud server
2. Cloud Computing
3. Authentication
4. Resource management
5. Virtualization
6. Green Computing

Cloud Server is the entity which provides the services to the users. It manages the resources so as to fulfill the needs of the user requests. Cloud Computing refers to the applications offered over the internet. This is the technology which provides the services to the users on the pay as u go basis. Various services like Software Platform and Applications are provided by the Cloud Authentication gives the Access restrictions for the user as well as to the Cloud service Administrator[6].

Resource allocation is the technology where we are allocate the resources to the users dynamically. User is not using all reserve memory for the current use so we are allocating that unused memory storage for the another user and return it to the original user when he require it. In Virtualization [4] we are using the Virtual machines to Increase the efficiency of the system. Green Computing is the technology we are adapting to save the electricity as well as increase system performance.

IV. Mathematical Model

Let S be the system that Provide Dynamic resource allocation in a efficient way with green
computing, such that S = I, F, O
Where,
I Represents the set of Inputs.
I = I1, I2,I3
I1 - Cloud user registration and Authentication.
I2 - Cloud service provider registration and Authentication.
I3 - Resource allocation.
F Represents the set of Functions.
F = F1, F2, F3, F4
F1 - Cluster detailing.
F2 - Site creation.
F3 - Site hosting.
F4 - Resource allocation status.
O Represents the set of Outputs.
O = O, Where O - Valid Efficient resource allocation And green computing.

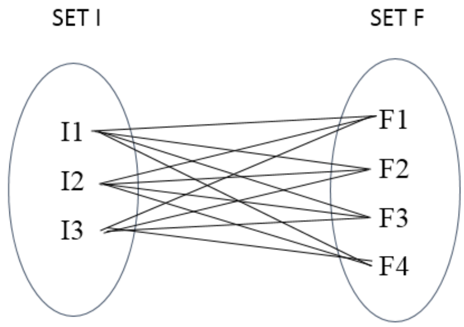

Fig. Venn diagram 1

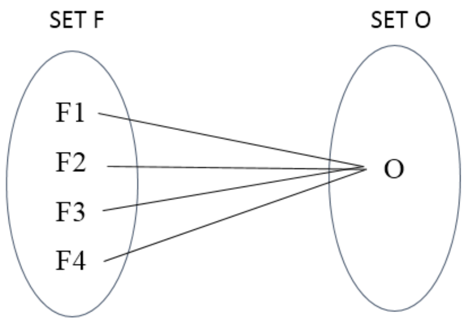

Fig. Venn diagram 2

V. *Mathematical Formulation*

The mathematical formulation [9] is designed to minimize the Skewness i.e. uneven utilization of the resources. Consider i=1 to m denoting the requests for resources. Then j=1 to n as servers which fullfils the need for resources, k=1 to l is the virtual machine with the given processing capacity, a_j is the total CPU capacity of the server j and b_i is the total CPU capacity required for request i. C_{ijk} is the per unit cost of utilization of resource by i^{th} request when it is executed on k^{th} VM on ^{jth} server.X_{ijk} is the amount of the CPU capacity being utilized by i^{th} request when it is executed on k^{th} VM on jth server. Thus we got the expression as,

$$Minimize\ F= \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{l} (Cijk\ Xijk)$$

$$Subjected\ to\ \sum_{j=1}^{n} \sum_{k=1}^{l} (Xijk) = ai, i = 1,2,\dots,m.$$

We calculate the Skewness i. e. Uneven utilization of the resources by using the following formula;

$$\text{Skewness (P)} = \sqrt{\sum_{i=1}^{n} (\frac{ri}{\bar{r}} - 1)2}$$

Where, n is the number of resources, ri be the i^{th} Resource utilization. $\bar{r}$ is the average utilization of the resources on Server P. minimizing the skewness we can reduce the overall workload on the system improve the utilization of server resources.

VI. HOT AND COLD SPOTS

We can say the server as a Hot spot if its temperature is above the hot threshold this indicates that the server is overloaded and load should be migrated away. Cold spot if its temperature is below cold or green computing threshold. It indicates that this server is in the condition that the load on this server may get transferred to another server which is not overloaded and that server should get switched off.
Now the question is what is threshold Temperature? That can be calculated using[1,8]

$$\text{Temperature(P)} = \sum_{r \in R} (r - rt)2$$

Where R is the se of overloaded resources in server p and rt is the hot threshold for resource r.

Migrations- our goal is to eliminate all hot spots or keep the temperature of them as low aspossible, the aim is tomigrate that VM in the server which result in low temperature of the server. In the hot to warm migrations we migrate the VM to the warm spot i. e. not cold nor hot.

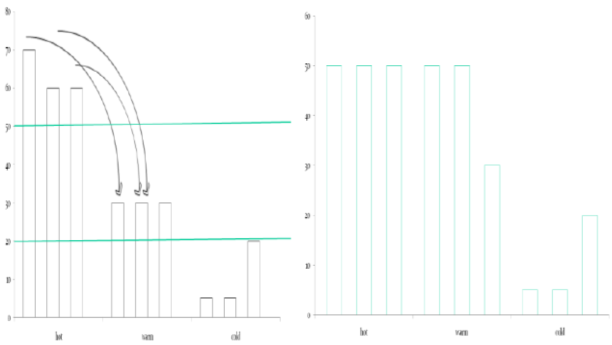

Fig. Hot to warm spot migration

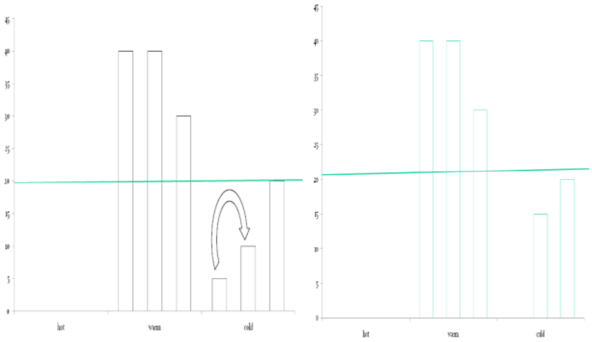

Fig. cold spot migrations

VII. CONCLUSION

Described paper gives the overview if the Dynamic resource allocation system with two goals- overload avoidance and green computing. Both of these goals are achieved by the VM Migrations. It increases the system performance as well as saves the electricity.3

System uses the skewness metric to combine VMs with different resource characteristics appropriately so that the capacities of servers are well utilized we achieve green computing by migrating the loads on the servers.

ACKNOWLEDGMENT

F.A. want to thank Amrutvahini College of engineering for facilitating the opportunity for such Application development. She also want to thank the Staff of department of computer engineering for their valuable guidance. .

REFERENCES

[1] Zhen Xiao, Senior Member, IEEE, Weijia Song, and Qi Chen *"Dynamic* resource allocation using virtual machine in cloud computing environment*"* ieee transaction on parallel and distributed systems ,year 2013.
[2] T. Wood, P. Shenoy, A. Venkataramani, and M. Yousif, *Black-box and gray-box strategies for virtual machine migration*, in Proc. of the Symposium on Networked Systems Design and Implementation (NSDI07), Apr. 2007..
[3] ZhixiongChen, JongP.Yoon,"International Conference on P2P, Parallel, Grid, Cloud and Internet Computing", 2010 IEEE: pp 250-257.
[4] Lutz schubert, Keith Jeffery "Advances in CLOUDS" the expert group report public version 1.0, European commission, 2012.
[5] Amol Selokar, Prof.S.D.Zade, Prof.C.U.Chavan "Survey on Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment" International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 5, May 2014
[6] L. Dhivya, Ms. K. Padmaveni, *"Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment"*, IJREAT International Journal of Research in Engineering Advanced Technology, Volume 2, Issue 1, Feb-Mar, 2014 ISSN: 2320 – 8791.
[7] SAROJA V, SIVAGAMI V M, *"Dynamic Resource Management Using Skewness and Load Prediction Algorithm for Cloud Computing"*, International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169, Volume: 2 Issue: 4, 798 804.
[8] Pooja.Kulkarni, Patil.Yogita, *Allocation of Resources Dynamically using Virtual Machines for Cloud Computing*, 2014 IJEDR –Volume 2, Issue 2 — ISSN: 2321-9939.
[9] Kumar N. and Agarwal S. Department of Computer Science, Babasaheb Bhimrao Ambedkar University, *An Analytical Model for Dynamic Resource Allocation Framework in Cloud Environment*, Research Journal of Recent Sciences ISSN 2277-2502 Vol. 3(IVC-2014), 1-6 (2014)

Authors Profile

Ms. A. D. Gawali received the B.E. degree in Information Technology from University of Pune, in 2013. Currently he is pursuing Master's degree in Computer Engineering from Amrutvahini college of Engineering, Sangamner under University of Pune. His areas of interest are network Security and cloud computing. He is currently working in the field of Network security and Cloud computing.

Prof. S. K. Sonkar received the Master degree in Computer science and Engineering from SRTMU Nanded. He is currently pursuing the Ph.D. degree in computer science from University of Pune. He is presently working as Assistant Professor in Dept. of Computer Engineering in Amrutvahini college of Engineering, Sangamner, India. His current research interests include network security and Cloud Computing.

