A Framework for Performance Analysis of Computing Clouds

Mintu M. Ladani,Vinit Kumar Gupta

Abstract— Cloud Computing is widely spread technique which has emerged a new technology that provides large amounts of computing and data storage capacity to its users with a promise of increased scalability, better availability, and reduced administration and maintenance costs. Because the use of cloud computing environments increases, it becomes crucial to understand the performance of this environment. So, to assess the performance of computing clouds in terms of various metrics is of great importance, such as the overhead of acquiring and releasing the virtual computing resources, and other virtualization and network communications overheads. To address these issues, we have presented the architecture of the framework and discuss several cloud resource management alternatives. Then finally we have designed and implemented load balancing algorithm for performance increase. we state the requirements for defining such algorithm to assess the performance of computing clouds. Secondly, a new VM load balancing algorithm has been proposed and designed for an IaaS framework in cloud computing environment; i.e. 'Modified Weighted Active Monitoring Load Balancing Algorithm' on cloud, for the Datacenter to effectively load balance requests between the available virtual machines assigning a weight, in order to achieve better performance parameters such as response time and Data processing time.

Index Terms— About four key words or phrases in alphabetical order, separated by commas.

I. INTRODUCTION

Cloud computing has emerged as a new technology that lets users deploy their applications in an environment with a promise of good scalability, availability, and fault tolerance. As the use of cloud computing environments increases, it becomes crucial to understand the performance of these environments in order to facilitate the decision to adopt this new technology, and to understand and resolve any performance problems that may appear. In this paper, we present Framework, which is a framework for generating and submitting test overloads to computing clouds. By using Framework, users can assess the overhead of acquiring and releasing the virtual computing resources, they can Different configurations and they can evaluate different scheduling algorithms.

II. LITERATURE REVIEW

Modeling techniques for workloads and storage devices in order to do load balancing of workloads over a set of devices. The workloads were modeled independent of underlying device using the parameters inherent to a workload such as seek distances, read-write ratio, average IO size and outstanding IOs. For device modeling we used statistics such as IO latency with respect to outstanding IOs which is dependent on the storage device. Also described a load balancing engine that can do migrations in order to balance overall load on devices in proportion to their capabilities [18].

Research problem:

- How to characterize dynamic workloads for load balancing? Are percentile values for workload parameters good enough in real systems?
- Can we use static values of constants such as K1 to K4 for workloads running on different devices or do we need online estimation?
- How to measure and predict interference among various workloads accessing a device?
- How to suggest storage configuration changes to an administrator based on online workload monitoring. Ultimately the task of workload monitoring, device modeling and load balancing needs to happen in a feedback loop over time to handle churn in today's storage environments.

A systematic literature review is a means of Identifying, Evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomena of interest (Kitchen ham, 2004).[14]

Reasons for performing SLR

- It can identify gaps in current research for further investigation.
- It can be used as framework for new research activities.
- To generate a new hypothesis**.**

Importance of systematic literature review

- Literature review has less scientific value.
- Systematic literature review approach is fair.
- It summaries all existing information about some phenomenon in a fair and
- Unbiased manner.

Advantages

- Methodology is well defined and due to this results of literature will have less repeatability.
- New research activities can be developed.

Disadvantages

- More effort is required

Features of SLR

- Systematic review starts by defining review protocol.
- Based on search strategy systematic review is defined.
- Search strategy is documented.

Manuscript received on May, 2013.

Mintu M. Ladani, M.E. (Computer), Gujarat Technological University/ Hasmukh Goswami College of Engineering, Vahelal (Dist- Ahmedabad), India.

Vinitkumar Gupta, M.E. (Computer), Gujarat Technological University/ Hasmukh Goswami College Of Engineering, Vahelal(Dist-Ahmedabad) , India.

- Quality criteria for paper selection of primary studies are considered.
- Systematic review can be used as a prerequisite for quantitative Meta-analysis.

- **Conclusion Of Literature Review:**

By Conclude this all above paper how to characterize workload maintains in virtualization After Concluding all above paper how to characterize how to load-balance in Cloud Computing. Some virtualization security risks and problems create in dynamic workload in load balancing.so,I will do maintain load balancing in cloud computing.

III. PROBLEM STATEMENT

In earlier load balancing algorithms or in existing methods only how to balance load is stated but in those schemes only some factors are considered like, VM with least load etc. but other factors like process completion time, its total duration and other process related issues are not considered which can affect the performance of system greatly. So a new VM load balancing algorithm has been proposed and designed for an IaaS framework in cloud computing environment; i.e. 'Modified Weighted Active Monitoring Load Balancing Algorithm' on cloud, for the Datacenter to effectively load balance requests between the available virtual machines assigning a weight, in order to achieve better performance parameters such as response time and Data processing time

A. Proposed Algorithm

Modified 'Weighted Active Monitoring Load Balancing Algorithm' is designed the Active Monitoring Load Balancer by assigning a weight to each VM as discussed in Weighted Round Robin Algorithm of cloud computing in order to achieve better response time and processing time. In this proposed Load balancing algorithm using the concept of weights in active monitoring, the VM are assigned varying (different) amount of the available processing power of server/ physical host to the individual application services. To these VMs of different processing powers; the tasks/requests (application services) are assigned or allocated to the most powerful VM and then to the lowest and so on according to its weight and its availability. Hence optimizing the given performance parameters. After VM with least load and with least process duration is identified, it allocates requests to the most powerful VM according to the weight assigned. If there is more than one, then the one which was searched first is selected.

STEP 1: According to computing power of host/physical server in terms of its core processor, processing speed, memory, storage etc. Create VM's of different Datacenter.

STEP 2: According to the computing power of the VM's in Datacenter allocate weighted count.

Like if one VM is capable of having twice as much load as the other, the powerful server

Gets a weight of '2' or if it can take four times load then server gets a weight of '4' and so on.

STEP 3: An index table of VMs associated weighted count and the number of requests currently allocated to the VM is maintained by WeightedActiveVmLoadBalancer maintains. Initially at start time all VM's have 0 allocations.

STEP 4: When the Datacenter Controller arrives requests to allocate a new VM, WeightedActiveVmLoadBalancer parses the table and identifies the least loaded VM.

STEP 5: After VM with least load and with least process duration is identified, it allocates requests to the most powerful VM according to the weight assigned. If there are more than one, then the one which was searched first is selected.

Vm With least load and least process duration is identified, it allocates request to most powerful VM according to weight assigned. if result are more than one which was searched first is selected.

STEP 6: After that WeightedActiveVmLoadBalancer returns the VM id to the Datacenter Controller.

STEP 7: The Datacenter Controller then sends the request to the VM identified by the id which is sent by Balancer.

STEP 8: To notify the WeightedActiveVmLoadBalancer about the new allocation, the Datacenter Controller sends intimation to it.

STEP 9: After receiving notification, Weighted Active VmLoad Balancer updates the allocation table increasing the allocations count for that VM.

STEP 10: When the VM finishes processing the request, and the Datacenter Controller receives the response, it notifies the Weighted Active VmLoad Balancer of the VM de-allocation.

STEP 11: The WeightedActiveVmLoadBalancer updates the allocation table by decreasing the allocation count for the VM by one.

STEP 12: Continue from step 4.

IV. CONCLUSION AND FUTURE WORK

A. Conclusion

We have concluded that as the use of cloud computing environments increases, it becomes crucial to understand the performance of this environment. To address different issues, we presented the architecture of the framework and discussed several cloud resource management alternatives. Then we analyses and designed load balancing algorithm for performance increase. we stated the requirements for defining such algorithm to assess the performance of computing clouds. We designed 'Modified Weighted Active Monitoring Load Balancing Algorithm' on cloud, for the Datacenter to effectively load balance requests between the available virtual machines assigning a weight, in order to achieve better performance parameters such as response time and Data processing time. Hence we tried best to consider the most affecting factor (process duration) in performance increase.

B. Future Work

We have done some modification in previously stated load balancing algorithm i.e. here we have considered the extra factor, (process duration). Based on which the least loaded Vm is identified to balance the overall load of Datacenter. We will implement this algorithm in private cloud. After implementation, corresponding results and performance variance will be measured.

REFERENCES

1. The Economist, "A Special Report on Corporate IT," October2008,"http://www.economist.com/specialReports/showsurvey .cfm? Issue=20081025".
2. R. Buyya, C. S. Yeo, and S. Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities," CoRR, vol. abs/0808.3558, 2008.
3. T. Killalea, "Meet the virts," Queue, vol. 6, no. 1, pp. 14–18, 2008.
4. A. Iosup and D. Epema, "GrenchMark: A Framework for Analyzing, Testing, and Comparing Grids," in Proc. of the 6th IEEE/ACM Intl.Symposium on Cluster Computing and the Grid (CCGrid06). IEEE Computer Society, 2006, pp. 313–320, an extended version can be found as Technical Report TU Delft/PDS/2005-002, ISBN 1387-2109.
5. Amazon, "http://aws.amazon.com/ec2/," 2008.
6. Open Grid Forum, "Job Submission Description Language (JSDL) Specification, Version 1.0," 2005.
7. D. Feitelson, L. Rudolph, U. Schwiegelshohn, K. Sevcik, and P. Wong, "Theory and practice in parallel job scheduling," in 3rd Workshop on Job Scheduling Strategies for Parallel Processing, vol. LNCS 1291, 1997, pp. 1–34.
8. D. Tsafrir, Y. Etsion, and D. G. Feitelson, "Backfilling Using System-Generated Predictions Rather Than User Runtime Estimates," IEEE Trans. Parallel & Distributed Syst., vol. 18, no. 6, pp. 789–803, Jun 2007.
9. S. L. Garfinkel, "An Evaluation of Amazons Grid Computing Services: EC2, S3 and SQS," Center for Research on Computation and Society School for Engineering and Applied Sciences, Harvard University, Tech. Rep. TR-08-07, 2007.
10. S. Garfinkel, "Commodity Grid Computing with Amazon's S3 and EC2," ;login, vol. 32, no. 1, pp. 7–13, 2007.
11. E. Walker, "Benchmarking Amazon EC2 for High-Performance Scientific Computing," ;login, vol. 33, no. 5, pp. 18–23, 2008.
12. D. Nurmi, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman, L. Youseff, and D. Zagorodnov, "Eucalyptus : A Technical Report on an Elastic Utility Computing Architecture Linking Your Programs to Useful Systems," Department of Computer Science, University of California, Santa Barbara, Tech. Rep. 2008-10, 2008.
13. M. R. Palankar, A. Iamnitchi, M. Ripeanu, and S. Garfinkel, "Amazon S3 for science grids: a viable solution?" in DADC '08: Proceedings of the 2008 international workshop on Data-aware distributed computing. New York, NY, USA: ACM, 2008,.
14. Karthik Paladugu Sumanth Mukk "Systematic Literature Review and Survey on High Performance Computing in Cloud" Systematic literature review (SLR) September 2012
15. By Cloud Computing Ajith Singh."An Approach on Semi-Distributed Load Balancing Algorithm for Cloud Computing System" N Dept. of Computer Science Karpagam University M. Hemalatha Dept. of Computer Science Karpagam University,oct-2012
16. By Mariana Carroll1,, Paula Kotzé1,3,"Secure virtualization Benefits, Risks and Controls " Alta van der Merwe Secure Virtualization
17. By Rajkumar Buyya1,2, Chee Shin Yeo1, and Srikumar Venugopal" Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities "
18. By Ajay Gulati, Chethan Kumar, Irfan Ahmad "Modeling Workloads and Devices for IO Load Balancing in Virtualized Environments"