

Cloud Bursting, Brokerage and Aggregation with Load Balancing

Snehlata Mishra¹, Ritu Tondon²

¹RGPV University, TRUBA College of Engineering and Technology, Indore-Rao Bypass Road, India

²TRUBA College of Engineering and Technology, RGPV University, Indore-Rao Bypass Road, India

Abstract: Today through Cloud Computing, the customer and business can access their personal files at any computer and with the help of centralizing data storage, processing and bandwidth, this technology permits much more efficient computing. The cloud bursting model in cloud computing helps to share the application's resources between private and public cloud. The problem related to cloud bursting arise from incompatibility between the different platforms and the limited availability of management tools. To manage hybrid environments cloud computing service providers have developed tools to share workloads on cloud, but they often require all environments to be based on same platform. Cloud Bursting is the limited availability of management tools that are cross compatible over multiple platforms. This paper introduce the concept of Load Balancing or Scheduling in Cloud Computing to share the dynamic workload across multiple nodes to ensure that no single node is overloaded and able to satisfy over-capacity needed by private cloud with very short advance warning. It helps in proper utilization and improve the performance of resources and able to load balance the work load in such a way that nobody will suffer any service deterioration to existing to other customers.

Keywords: Cloud Scheduling, Cloud Bursting, Private Cloud, Load Balancing, Cloud performance.

1. Introduction

Cloud Computing is a model for enabling ubiquitous, convenient, on demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort. It providing online resources and online storage to the user's it provide all the data at a lower cost. In cloud computing users can access resources all the time through internet. They need to pay only for those resources as much they use. In Cloud computing cloud provider outsourced all the resources to their client. There are many existing issues in cloud computing. The main problem is load balancing in cloud computing. Load balancing helps to distribute all loads between all the nodes. It also ensures that every computing resource is distributed efficiently and fairly. It helps in preventing bottlenecks of the system which may occur due to load imbalance. It provides high satisfaction to the users. Load balancing is a relatively new technique that provides high resource utilization and better response time. [1] [2] [3] [4] Cloud computing provide many advantages to the users.

A. Cloud computing consist of several characteritestic: [20] [6].

- **On demand self service:**-The computing capabilities like server time and network storage can be provisioned by the users without performing any human interaction with each and every service provider.
- **Geographic Distribution:**-Having Large number of geographical distributed datacenter sites allow flexibility of cloud provider to allocate workload to resources close to end user.
- **Homogeneity:**- In order to maximize the operational efficiencies, successful cloud deployments will limit the range of different types hardware ,infrastructure ,software platforms and policies they support.

- **Measured Service:**-In cloud computing resource usage can be monitored, controlled for both the provider and consumer of the all service.
- **Broad Network Access:**-In cloud computing capabilities are available over the network .All the capabilities are accessed through different mechanisms.
- **Resource Pooling:**-Different models are used to pooled the resources which provide by the providers to their consumers. All the resources dynamically assigned and reassigned according to consumer demand.
- **Rapid Elasticity:**-Quantity of resources is increase at any time according to the customer's requirements.

B. Challenges in Cloud Computing

There are many challenges in cloud computing:-

- 1) Security
- 2) Efficient load balancing
- 3) Performance Monitoring
- 4) Consistent and Robust Service abstractions
- 5) Resource Scheduling
- 6) Scale and QoS management
- 7) Requires a fast speed Internet connection.

C. Layers of Services

All the services have number of layers which manage by the users and providers.

Public Cloud: The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization .Anyone can use public cloud as they want without restriction

Private Cloud: The cloud infrastructure is used by a single organization. Private cloud is only managed by the

organization or a third party. General Public not able to use the private cloud directly.

Community Cloud: The cloud infrastructure is shared by many organizations. Community cloud supports a specific community that has shared concerns. Ex:-security requirements, policy, compliance considerations. It may be managed by the organizations or a third party.

Hybrid Cloud: Hybrid cloud is a combination of two or more clouds (private, community, or public). That remains unique entities but is bound together by standardized technology that enables data and application portability. Ex:- **cloud bursting for load-balancing between clouds. This capability enables hybrid clouds to cloud bursting for scaling across clouds. A primary advantage of cloud Bursting is that an organization only pays for extra compute resources when they are needed**

2.Literature Survey

The technique of extending capabilities of enterprises resources by leasing public cloud capabilities is introduced in 2011[24], but there is only a marginal improvement at the original scale of trace, whereas for smaller task sizes the behavior is similar.

In 2012 [17],[19], various security algorithms like RSA, MD5, AES etc are proposed for secure cloud bursting but the solutions are not stable and revolution always concern with the new problem.

In 2013[7] to seamlessly burst applications between public and private cloud, the cloud bursting model was suggested, but the cloud capabilities have not yet been developed and researched to level that allows their exploitation to a full degree.

Cloud services are currently among the top-ranked high growth areas in computer services and seeing an acceleration in enterprise adoption with the worldwide market predicted to reach more than \$140b in 2014[14],[3]. For secure cloud bursting and aggregation, the author uses encryption scheme of 64-bit cipher. But they fail to apply this concept in real environment check for the real time simulations on different platforms.

In 2014[5], the problem of load balancing in cloud bursting was discovered and various scheduling algorithms are proposed to solve the problem but the practical implementation of that solutions is still remaining.

In 2015[19], automatic cloud bursting was suggested which allows computer resources to be dynamically reconfigured to meet users' demands but the launching of virtual machines on commercial cloud takes much more time than on fermicloud. Hence further more tests in virtual machines' performance on commercial cloud needed to be done in the future.

3.Problem Definition

Cloud bursting is the idea that an application normally runs in one type of cloud and is capable of utilizing additional resources of another cloud type during peak periods, or „bursting.“ The most common example of this type of utilization would be a retail company utilizing a private cloud for day-to-day operations bursting to the public cloud for peak periods such as a holiday season. Experts recommend cloud bursting for high performance, on-critical applications that handle non-sensitive information. In cloud bursting the designated public cloud platform should be fully compatible with the private cloud to successfully run the bursting applications. Cloud Bursting is the limited availability of management tools.



Figure 2: (Basic Idea Cloud Bursting)

When the demands of capacity spikes in an organization who is owning their own cloud infrastructure and willing to use resources from public cloud provider for certain time intervals and given certain circumstances triggering this use. In this situation, the service providers and virtualization vendors have developed tools to send workloads to the cloud and manage hybrid environment, but they often require all environment to be based on the same platform and the drawback in satisfy the load balance in such a way that nobody suffers any service denial to existing to other customer is a major problem.

Benefits of cloud bursting

Cloud bursting can be advantageous and utilizing it within a hybrid cloud deployment model can have further benefits. Benefits of cloud bursting are:

Ability to expand and retract services based on changes in capacity needs

- Cost reduction, enterprises only pay for extra computing resources on demand
- Increased performance with spikes in workloads
- Bursting new instances of applications to another cloud
- Ability to burst scalable application tiers to improve performance at peak periods
- Reduced capital expense

Further benefits of utilizing a hybrid cloud include:

- Fail over or disaster recovery
- Increased availability

4. Proposed Solution

To solve the above stated problem of cloud bursting, the concept of Load Balancing or Scheduling is proposed so that traffic can dynamically burst out to a private cloud and take the advantage of the capacity and when the peak is over, the business can return that capacity back to pool and shed the cost.

Load balancing is used to distributing a larger processing load to smaller processing nodes for enhancing the overall performance of system. In cloud computing environment load balancing is required to distribute the dynamic local workload evenly between all the nodes. . [10][11][12][13]

- Load balancing helps in fair allocation of computing resource to achieve a high User satisfaction and proper Resource utilization High resource utilization and Proper load balancing helps in minimizing resource consumption. It helps in implementing fail over, scalability, and avoiding bottlenecks.
- Load balancing is a techniques that helped networks and resources by providing a maximum throughput with minimum response time. Load balancing is dividing the traffic between all servers, so data can be sent and received without any delay with load balancing.
- In cloud environment many algorithms are available that helps in proper traffic loaded between all available servers. Most of them can be applied in the cloud environment with suitable verifications. In cloud computing environment load balancing algorithms can be divided into two main groups : first algorithm type is Batch mode heuristic scheduling algorithms (BMHA) and second is online mode heuristic algorithms. In BMHA Jobs are combined together when they are arriving in the system. The BMHA scheduling algorithm will start after a fixed time period.
- In one vision of the future, the shifting of load is automated to enable organizations to configure clouds and cloud balancing and then turn their attention to other issues, trusting that the infrastructure will perform as designed.

Technical Goals of Cloud Balancing

From a purely technical perspective, the goals of cloud balancing are similar to those associated with traditional GSLB: ensure the availability of applications while simultaneously maximizing performance, regardless of the location or device from which users are accessing the application. Whether that access point is within an organization's data center utilizing private cloud resources or via a cloud provider, DNS requests are sent to the most appropriate location.

These technical goals are met through a combination of application and network awareness and collaboration between the global application delivery solution and local load balancing solutions. By coordinating across application deployments in multiple data centers, whether in the cloud or traditionally based, organizations can, through careful monitoring of capacity and performance-related variables, achieve optimal application performance while ensuring availability.

Another goals

- **Cost effectiveness:** Load balancing help in provide better system performance at lower cost.
- **Scalability and flexibility:** The system for which load balancing algorithms are implemented may be change in size after some time. So the algorithm must handle these types'' situations. So algorithm must be flexible and scalable.
- **Priority:** Prioritization of the resources or jobs needs to be done. So higher priority jobs get better chance to execute.

To distribute the dynamic local workload evenly between all nodes in public and private cloud to achieve better throughput and improve the response time in cloud environment we can use any one of the following algorithms mentioned below:-

- **Task Scheduling based on LB:** This algorithm mainly consist two level task scheduling mechanism which are based on load balancing to meet dynamic requirements of users . It obtains high resource utilization. This algorithm achieves load balancing by first mapping tasks to virtual machines and then all virtual machines to host resources .It is improving the task response time .It also provide better resource utilization .
- **Opportunistic Load Balancing:** OLB is to attempt each node keep busy, therefore does not consider the present workload of each computer. OLB assigns each task in free order to present node of useful .The advantage is quite simple and reach load balance but its short coming is not consider each expectation execution time of task, therefore the whole completion time (Make span) is very poor.
- **Round Robin:** - In this algorithm all the processes are divided between all processors. In this each process is assigned to the processor in a round robin order. The work load distributions between processors are equal. Different processes have not same job processing time. At many point of time some nodes may be heavily loaded and others remain idle In web servers where http requests are of similar nature and distribute equally then RR algorithm is used . In Round Robin Scheduling the time quantum play a important role. When time quantum is very large then RR Scheduling Algorithm is same as the FCFS Scheduling. and when time quantum is too small then Round Robin Scheduling is known as Processor Sharing Algorithm.
- **Randomized:** This algorithm is static in nature. In this algorithm a process can be handled by a particular node n with a probability p. When all the processes are of equal loaded then this algorithm work well. Problem arises when loads are of different computational complexities. This algorithm is not maintaining deterministic approach.
- **Min-Min Algorithm:** It starts with a set of all unassigned tasks .In this minimum completion time for all tasks is found. Then after that among these minimum times the minimum value is selected. Then task with minimum time schedule on machine. After that the execution time for all other tasks is updated on that machine then again the same procedure is followed until all the tasks are assigned on the resources. The main problem of this algorithm is has a starvation.
- **Shortest Response Time First:** The idea of this algorithm is straight forward. In this each process is assigned a priority which is allowed to run. In this equal priority processes are

scheduled in FCFS order. The (SJF) algorithm is a special case of general priority Scheduling algorithm. In SJF algorithm is priority is the inverse of the next CPU burst. It means, if longer the CPU burst then lower the priority. The SJF polices selects the job with the shortest (expected) processing time first. In this algorithm shorter jobs are executed before long jobs. In SJF, it is very important to know or estimate the processing time of each job which is major problem of SJF.

5. Conclusion and Future Work

In this paper we plan to implement the cloud bursting seamlessly by using any one of the suggested scheduling algorithm with the aim of effortlessly burst application between private and public cloud.

References

- [1] Mell, Peter and Grance, Tim, "The NIST definition of cloud computing", National Institute of Standards and Technology, 2009, vol53, pages50, Mell2009
- [2] Y. Hu, R. Blake, D. Emerson, "An optimal migration algorithm for dynamic load balancing", *Concurrency: Practice and Experience* 10 (6)(1998) pages 467–483.
- [3] Sidhu, S. Kinger, "Analysis of load balancing techniques in cloud computing", *INTERNATIONAL JOURNAL OF COMPUTERS & TECHNOLOGY* 4 (2) (2013) pages737–741.
- [4] EU Expert Group, The Future of Cloud Computing, <http://cordis.europa.eu/fp7/ict/ssai/docs/cloud-report-final.pdf>
- [5] M. HSU (ED.), "Special Issue on Workflow and Extended Transaction Systems", *IEEE Data Engineering Transaction Systems*, *IEEE Data Engineering*, Vol. 16(2), June 1993
- [6] "Security Guidance for Critical Areas of Focus in Cloud Computing". Cloud Security Alliance. 2011. Retrieved 2011-05-04
- [7] Gabriel Mateescu, Wolfgang Gentzsch, Calvin J Ribbens "Hybrid Computing – Where HPC meets grid and Cloud Computing" Science Direct November 2010
- [8] Indraveer Chana and Anju Bala, "A Survey of Various Workflow Scheduling Algorithm in Cloud Environment", NCICT 2011.
- [9] K. Jackson, L. Ramakrishnan, K. Muriki, S. Canon, S. Cholia, J. Shalf, H. Wasserman, and N. Wright. Performance Analysis of High Performance Computing Applications on the Amazon Web Services Cloud. IEEE International Conference on Cloud Computing Technology and Science (CloudCom 2010), Nov. 30-Dec. 1, 2010, Bloomington, Ind
- [10] Yung-Terng Wang and Robert J. T. Morris. Load sharing in distributed systems. *IEEE Transactions on Computers*, C-34(3):204–217, March 1985
- [11] <http://www.personal.kent.edu/~rmuhamma/OpSystems/Myos/prioritySchedule.htm>, "Priority Scheduling - Operating Systems Notes."
- [12] S. S. MOHARANA, R. D. RAMESH, D. POWAR, "Analysis of load balancers in cloud computing" *International Journal of Computer Science and Engineering*, 2013, vol 2, pages 101-108
- [13] O. Elzeki, M. Reshad, M. Elsoud, "Improved max-min algorithm in cloud computing, *International Journal of Computer Applications*" vol 50 (12) (2012) pages 22–27..
- [14] Hai Zhong, Kun Tao, Xuejie Zhang, "An Approach to Optimized Resource Scheduling Algorithm for Open-source Cloud Systems," 2010 Fifth Annual China Grid Conference (China Grid), July 16-18, 2010, IEEE, pp.124-129
- [15] Information Society and Media, "An EGEE Comparative Study: Grids and Clouds - Evolution or Revolution," Technical Report Enabling Grids for E-science, v1.1, Jun 11, 2008, pp. 1-33
- [16] H. Chen, F. Wang, N. Helian, G. Akanmu, "User-priority guided min-min scheduling algorithm for load balancing in cloud computing", *Parallel Computing Technologies (PARCOMPTECH)*, 2013 National Conference on, 2013
- [17] Mayur R. Palankar, Adriana Iamnitchi, Matei Ripeanu, and Simson Garfinkel. Amazon S3 for science Grids: a viable solution? In *International Workshop on Data-aware Distributed Computing (DADC'08)* in conjunction with HPDC 2008, pages 55–64, New York, NY, USA, 2008. ACM.
- [18] <http://letslearncloud.wordpress.com/> "Cloud Computing | Learn Cloud and its tips."
- [19] Iosup, H. Li, M. Jan, S. Anoep, C. Dumitrescu, L. Wolters, and D.H.J. Epema. The grid workloads archive. *Future Generation Computer Systems*, 2008..
- [20] NIST Definition of Cloud Computing v15, <http://csrc.nist.gov/groups/SNS/cloud-computing/cloudef-v15.doc>
- [22] Ahuva W. Mu'alem and Dror G. Feitelson. Utilization, predictability, workloads, and user runtime estimates in scheduling the IBM SP2 with backfilling. *IEEE Transactions on Parallel and Distributed Systems*, 12(6):529–543, 2001
- [23] Chaczk, Zenon and Mahadevan, Venkatesh and Aslanzadeh, Shahrzad and Mcdermid, Christopher, "Availability and load balancing in cloud computing", *International Conference on Computer and Software Modeling*, Singapore, chaczko2011availability
- [24] Gansen Z; Chunming R; Jin L; Feng Z; Yong T; (2010), "Trusted Data Sharing over Untrusted Cloud Storage Providers", 2010 IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom), pp 97, Nov. 30 2010-Dec. 3 2010.
- [25] Grobauer, B.; Walloschek, T.; Stocker, E.; (2011), "Understanding Cloud Computing Vulnerabilities", 5487489 searchabstr Security & Privacy, IEEE, Vol 9, pp 50.
- [26] Tian L.Q; NI Y, LING; (2010), "Evolution of user Behavior Trust in Cloud Computing", 2010 International Conference on Computer Application and System Modeling (ICASM 2010), Vol. 7, pp V7-567, 22-24 Oct. 2010.

Author Profile

Snehlata Mishra completed the BE in Computer Science from RGPV University, Indore in 2013. Currently she is doing dissertation of M.E. (CS) final semester on same topic as of paper, from RGPV University.

Ritu Tondon is working as an Asst. Professor (CS) in TRUBA College of Engineering and Technology, Indore under RGPV University. She is the guide of dissertation work of Snehlata Mishra.