ISSN (Online) : 2348 - 2001

International Refereed Journal of Reviews and Research

Volume 3 Issue 2 March 2015

International Manuscript ID : 23482001V3I2032015-10

(Approved and Registered with Govt. of India)

DYNAMIC LOAD BALANCING ALGORITHM FOR EFFECTIVE COMMUNICATION IN DATA CENTERS OF CLOUD INFORMATION

Anuradha

ABSTRACT

Cloud computing is the result of evolution and adoption of existing technologies and paradigms. The goal of cloud computing is to allow users to take the benefits from all of these technologies and computing resources, without the need for deep knowledge about or expertise with each one of them. The cloud aims to cut costs, and helps the users focus on their core business instead of being impeded by IT obstacles. In cloud infrastructure, the customers do not have to invest capital to purchase, manage, maintain and scale the physical infrastructure. The customers can take required resources on demand from the cloud providers and pay for it as they use. In this paper, the proposed work focus on the dynamic load balancing that is compared based on the parameters cost and execution time with the classical approach. The parameters and results in the proposed algorithmic approach is giving the optimal and effective results in multiple cloud scenarios.

Keywords – Cloud Security, Secured Load Balancing, Cloud Computing

ISSN (Online) : 2348 - 2001
International Refereed Journal of Reviews and Research
Volume 3 Issue 2 March 2015
International Manuscript ID : 23482001V3I2032015-10
(Approved and Registered with Govt. of India)

INTRODUCTION

Cloud computing refers the paradigm 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). The name cloud was inspired by the symbol that's often used to represent the Internet in flowcharts and diagrams.

The services that are provided by the cloud providers are broadly classified into three categories:

Infrastructure-as-a-Service (IaaS): In Infrastructure as a Service model, the service provider owns the equipments including storage, hardware, servers and networking components and is provided as services to the clients. The client typically pays on per-use basis. Amazon elastic Compute (EC2) and Simple Storage Service (S3) are typical examples for IaaS.

Platform-as-a-Service (PaaS): In Platform as a Service model, the service provider provides virtualized server, operating system and development tools as service. Using these services, users can develop, test, deploy and manage new applications in a cloud environment or run existing applications. These applications are delivered to users via the internet. Google App Engine is a typical example for PaaS.

Software-as-a-Service (SaaS): In Software as a Service model, the service provider provides software as a service over the Internet, eliminating the need to buy, install, maintain, update and run the application on the customer's own computers. Google Docs is a typical example.

A cloud service has four distinct characteristics as follows:

i. It is elastic: A user can dynamically scale up and scale down resources as they want at any given time.
ii. Pay per use: Usage is metered and user pays only for what they consume.

ISSN (Online) : 2348 - 2001

International Refereed Journal of Reviews and Research

Volume 3 Issue 2 March 2015

International Manuscript ID : 23482001V3I2032015-10

(Approved and Registered with Govt. of India)

iii. Operation: The service is fully managed by the provider.

iv. Self-service: Users can add a new CPU, a server instance or extra storage using the console offered by the cloud provider.

Cloud Computing research addresses the challenges of meeting the requirements of next generation private, public and hybrid cloud computing architectures, also the challenges of allowing applications and development platforms to take advantage of the benefits of cloud computing. The research on cloud computing is still at an early stage. Many existing issues have not been fully addressed, while new challenges keep emerging from industry applications. Some of the challenging research issues in cloud computing are given below.

- Service Level Agreements (SLA's)
- Cloud Data Management And Security
- Data Encryption
- Migration of Virtual Machines
- Interoperability
- Access Controls
- Energy Management
- Multitenancy
- Server Consolidation
- Reliability and Availability of Service
- Common Cloud Standards
- Platform Management

Service Level Agreements (SLA's) : Cloud is administrated by service level agreements that allow several instances of one application to be replicated on multiple servers if need arises;

dependent on a priority scheme, the cloud may minimize or shut down a lower level application. A big challenge for the Cloud customers is to evaluate SLAs of Cloud vendors. Most vendors create SLAs to make a defensive shield against legal action, while offering minimal assurances to customers.

Cloud Data Management : Cloud data Can be very large (e.g. text-based or scientific applications), unstructured or semi-structured, and typically append-only with rare updates Cloud data management an important research topic in cloud computing. Since service providers typically do not have access to the physical security system of data centers, they must rely on the infrastructure provider to achieve full data security.

Data Encryption : Encryption is a key technology for data security. Understand data in motion and data at rest encryption. Remember, security can range from simple (easy to manage, low cost and quite frankly, not very secure) all the way to highly secure (very complex, expensive to manage, and quite limiting in terms of access). User and the provider of the Cloud computing solution have many decisions and options to consider.

Migration of virtual Machines : Applications are not hardware specific; various programs may run on one machine using virtualization or many machines may run one program. Virtualization can provide significant benefits in cloud computing by enabling virtual machine migration to balance load across the data center. In addition, virtual machine migration enables robust and highly responsive provisioning in data centers.

Interoperability : This is the ability of two or more systems work together in order to exchange information and use that exchanged information. Many public cloud networks are configured as closed systems and are not designed to interact with each other. The lack of integration between these networks makes it difficult for organizations to combine their IT systems in the cloud and realize productivity gains and cost savings. To overcome this challenge, industry standards must be developed to help cloud service providers design interoperable platforms and enable data portability.

Energy Resource Management : Significant saving in the energy of a cloud data center without sacrificing SLA are an excellent economic incentive for data center operators and would also make a significant contribution to greater environmental sustainability.

Multi-tenancy : There are multiple types of cloud applications that users can access through the Internet, from small Internet-based widgets to large enterprise software applications that have increased security requirements based on the type of data being stored on the software vendor's infrastructure. These application requests require multi-tenancy for many reasons, the most important is cost. Multiple customers accessing the same hardware, application servers, and databases may affect response times and performance for other customers.

Server consolidation : The increased resource utilization and reduction in power and cooling requirements achieved by server consolidation are now being expanded into the cloud. Server consolidation is an effective approach to maximize resource utilization while minimizing energy consumption in a cloud computing environment. Live VM migration technology is often used to consolidate VMs residing on multiple under-utilized servers onto a single server, so that the remaining servers can be set to an energy-saving state. The problem of optimally consolidating servers in a data center is often formulated as a variant of the vector bin-packing problem, which is an NP-hard optimization problem.

Reliability & Availability of Service : The challenge of reliability comes into the picture when a cloud provider delivers on-demand software as a service. The software needs to have a network conditions (such as during slow network connections). There are a few cases identified due to the unreliability of on-demand software.

Common Cloud Standards : Security based accreditation for Cloud Computing would cover three main areas which are technology, personnel and operations. Technical standards are likely to be driven by organizations, such as, Jericho Forum1 before being ratified by established bodies, e.g., ISO2 (International Standard Organization).

ISSN (Online) : 2348 - 2001

International Refereed Journal of Reviews and Research

Volume 3 Issue 2 March 2015

International Manuscript ID : 23482001V3I2032015-10

(Approved and Registered with Govt. of India)

LITERATURE SURVEY

Randles, Martin, David Lamb, and A. Taleb-Bendiab.[31] "A comparative study into distributed load balancing algorithms for cloud computing." In *Advanced Information Networking and Applications Workshops (WAINA), 2010 IEEE 24th International Conference on*, pp. 551-556. IEEE, 2010 - This paper investigates three possible distributed solutions proposed for load balancing; approaches inspired by Honeybee Foraging Behaviour, Biased Random Sampling and Active Clustering.

Wang, Shu-Ching, Kuo-Qin Yan, Wen-Pin Liao, and Shun-Sheng Wang[43]. "Towards a load balancing in a three-level cloud computing network." In *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, vol. 1, pp. 108-113. IEEE, 2010 - In this study, a two-phase scheduling algorithm under a three-level cloud computing network is advanced. The proposed scheduling algorithm combines OLB (Opportunistic Load Balancing) and LBMM (Load Balance Min-Min) scheduling algorithms that can utilize more better executing efficiency and maintain the load balancing of system.

Kansal, Nidhi Jain, and Inderveer Chana[36]. "Cloud load balancing techniques: A step towards green computing." *IJCSI International Journal of Computer Science Issues* 9, no. 1 (2012): 1694-0814 - This paper discusses the existing load balancing techniques in cloud computing and further compares them based on various parameters like performance, scalability, associated overhead etc. that are considered in different techniques. It further discusses these techniques from energy consumption and carbon emission perspective.

Yi Lu et al. [60] proposed a mechanism to allocate computing resource to workloads with various mixes of best-effort and advance reservation requests. Incoming advance reservation leases are always scheduled right away, where best-effort leases are put on a queue. The scheduling function periodically evaluates the queue. It uses an aggressive backfilling algorithm [49, 23] to decide whether any best-effort leases can be scheduled. When an advance reservation lease comes, the scheduler will try to choose nodes that will not need preempting another lease.

Motivated by low resource use, Amazon EC2 introduced the spot instance mechanism to allow customers to bid for unused Amazon EC2 capacity [41]. Amazon EC2 at each availability zone has fixed number of virtual machine types and fixed number of instances of each VM type. Amazon EC2 runs one spot market for each VM type in each availability zone. A customer submits a request that specifies the type, the number of instances, the region desired and the biding price per instance-hour. The provider assigns resources to bidders in decreasing order of their bids until all available resources have been allocated or all resource requests have been satisfied. The selling price (i.e. the spot price) is equal to the lowest winning bid.

T.Kokilavani et.al defined a new class of games called Cloud Resource Allocation Games (CRAGs). CRAGs solve the resource allocation problem in clouds using game-theoretic mechanism. At the start of each round, all the clients submit the jobs for that round to the cloud. The provider first calculates his resource allocation vector for jobs that are running on the cloud.

Divya et.al presented a scheduling strategy on load balance of VM resources based on genetic algorithm. According to historical data and current state of the system and through genetic algorithm, this strategy computes ahead the influence on the system after the deployment of the needed VM resources. It then chooses the least-affective solution, through which it achieves the best load balance and reduces or avoids dynamic migration. This strategy solves the problem of load imbalance and high migration cost by traditional algorithms after scheduling.

Abhay Bhadani et.al proposed a power aware load balance algorithm. The PALB algorithm first gathers the utilization percentage of each active compute node. If all compute nodes n are above 75% utilization, PALB instantiates a new virtual machine on the compute node with the lowest utilization. Otherwise, the new virtual machine (VM) is booted on the compute node with the highest utilization (if it can accommodate the size of the VM). If all currently active compute nodes have utilization over 75%, PALB sends turning on command to power on additional compute nodes (as long as there are more available compute nodes). If the compute node is using less than 25% of its resources, PALB sends a shutdown command to that node.

PROPOSED WORK AND SIMULATION

The proposed algorithmic approach is simulated in the following environment -

- CloudSim
- GridSim
- Eclipse Juno IDE
- Notepad++

The proposed dynamic load balancing approach is implemented using the metaheuristic approach and the execution is done in parallel aspects.

Table 1 - Comparison of Efficiency of the Existing and Proposed Approach

Existing Approach	Improved Algorithm
41	68
47	59
36	88
42	52

ISSN (Online) : 2348 - 2001

International Refereed Journal of Reviews and Research

Volume 3 Issue 2 March 2015

International Manuscript ID : 23482001V3I2032015-10

(Approved and Registered with Govt. of India)

43	60
34	72
37	79
34	90
46	55
33	82

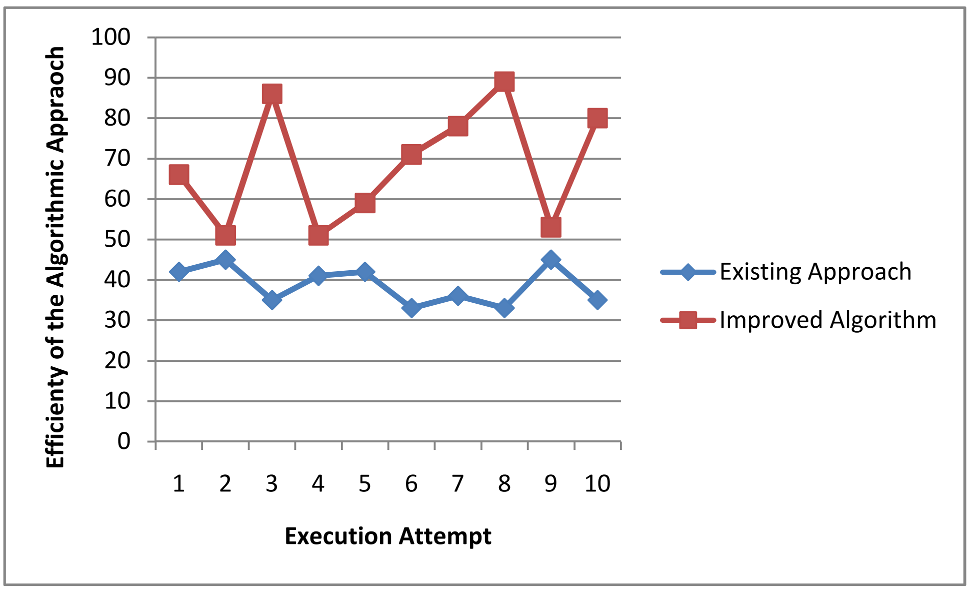

Figure 1 - Comparison between classical and proposed approach

The figure shows the comparative analysis between classical and proposed approach. It is evident from the results that the proposed approach is giving better results as compared to the classical in terms of efficiency and integrity.

ISSN (Online) : 2348 - 2001

International Refereed Journal of Reviews and Research

Volume 3 Issue 2 March 2015

International Manuscript ID : 23482001V3I2032015-10

(Approved and Registered with Govt. of India)

Table 2 – Comparison of Execution Time

Existing Approach	Improved Algorithm
2.1	1.6
3.1	2.2
3.2	2.2
3.1	2.3
4.9	3.6
4.2	3.7
2.9	2.3
3.8	2.5
3.5	2.7
4.2	3.5

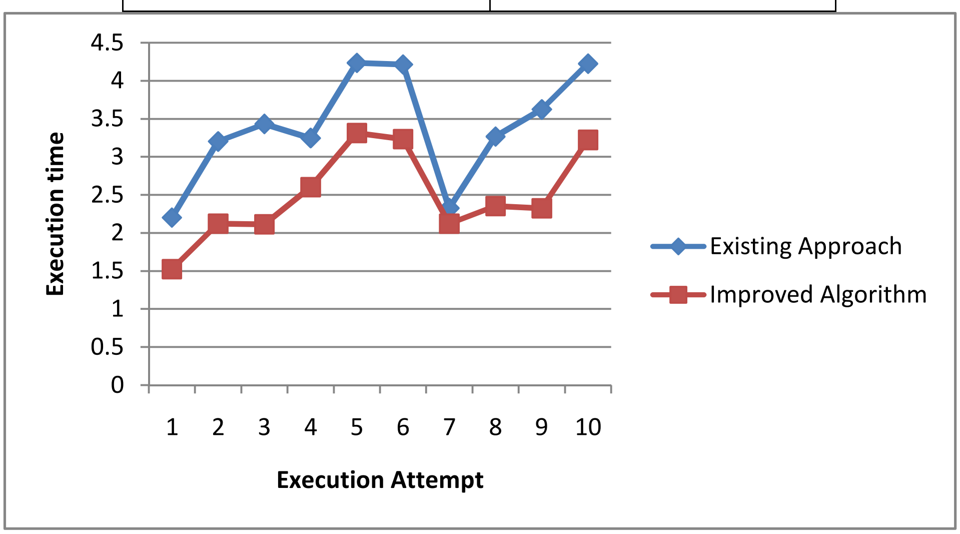

Figure 2 - Comparison between existing and proposed approach

CONCLUSION AND FUTURE SCOPE

This work mainly highlights the load balancing aspects in the cloud computing infrastructure. The existing algorithmic approaches are not effective by which the security can be enhanced. In the proposed approach, the improvements in the classical algorithms are implemented and effective results are found. The proposed approach gives optimized cost as well as the execution time that is directly proportional and related. Cloud Computing embodies the as-a-Service paradigm and allows for services to be provided en masse to consumers. As there is storage and accessing of data from cloud servers, the concerns about data confidentiality, authentication and integrity are being increased. Data stored within the Cloud is more vulnerable and open to attacks than before; its protection is paramount. Beside these issues there would also be chance of using a part of data or whole by cloud server for their financial gain which results the economic losses to data owner. The main reason behind the above defined issues is that the cloud servers are very likely to be the outside from trusted boundaries of data owner. One of the related issues is also the unwanted exposure of data as result of a software malfunction or malicious CSP. When entrusting data to the cloud the data creators i.e. service users need assurances over access to their data. In essence data creators need to regain control over this access i.e. data creators need to become empowered. A data owner centric three- tier privacy aware cloud computing model is also proposed to implement the algorithm in real-time cloud environment.

The power of proposed algorithmic approach and model may be checked by implementing and penetrating various attacks. After the penetration test, we may come to a conclusion about its robustness in terms of confidentiality, integrity and authenticity.

- The combination of access control techniques and cryptographic techniques may be used to maintain more privacy and security of data within the cloud.

- The QOS (quality of service) of proposed model may be determined in terms of availability, throughput and delay.
- The actual physical implementation and testing of proposed model required to detect its fault tolerant power and Data availability.
- The work can be enhanced using metaheuristics techniques including
  - Genetic Algorithms
  - Ant Colony Optimization
  - Neural Networks
  - HoneyBee Algorithm

REFRENCES

[1] Abhay Bhadani , Sanjay Chaudhary, "Performance Evaluation of Web Servers using Central Load Balancing Policy over Virtual Machines on Cloud", Proceedings of the Third Annual ACM Bangalore Conference, Article No. 16, ISBN: 978-1-4503-0001-8, January 2010, DOI: 10.1145/1754288.1754304.

[2] Anthony T.Velte, Toby J.Velte, Robert Eisenpeter, "Cloud Computing: A Practical Approach", Tata McGraw-Hill Publishers, 1st Edition, 2009, ISBN: 0071626948.

[3] Argha Roy, Diptam Dutta, "Dynamic Load Balancing: Improve efficiency in Cloud Computing", International Journal of Emerging Research in Management Technology, pp 78-82, Vol 2, Issue 4, ISSN:2278-9359, April 2013.

[4] Arzuaga, E., and Kaeli, D. R., 2010, Quantifying load imbalance on virtualized enterprise servers, Proceedings of the first joint WOSP/SIPEW international conference on Performance engineering, pp. 235–242.

[5] Barrie Sosinsky, "Cloud Computing Bible", Wiley publishers, 1st edition, December 2010, ISBN: 978-0-470-90356-8.

[6] Bellur, U., Rao C., and Kumar, M., 2010, Optimal Placement Algorithms for Virtual Machines, Proceedings of CoRR, pp.103-110.

[7] Bobroff, N., Kochut, A., and Beaty, K., 2007, Dynamic Placement of Virtual Machines for Managing SLA Violations, Proceedings of the 10th IFIP/IEEE Symposium on Integrated Network Management, pp. 119-128.

[8] Che-Lun Hung, Hsiao-hsi Wang, Yu-Chen Hu, "Efficient Load Balancing Algorithm for Cloud Computing Network", International Conference on Information Science and Technology (IST 2012), pp 251-253, April 28-30, 2012.

[9] Chandrasekaran, B., Purush, R., Douglas, B., and Schmidt, D., 2007, Virtualization Management Using Microsoft System Center and Dell OpenManage, Dell Power Solutions, pp. 40-44.

[10] Uddalak Chatterjee, "A Study on Efficient Load Balancing Algorithms in Cloud Computing Environment", International Journal of Current Engineering and Technology, pp 1767-1770, Vol 3, ISSN 2277-4106, December 2013.

[11] Vaidehi. M, Rashmi. K. S, Suma. V, "Enhanced Load Balancing to Avoid Deadlock in Cloud", International Journal of Computer Applications on Advanced Computing and Communication Technologies for HPC Applications, pp 31-35, June 2012.

[12] Van, H., and Tran, F., 2009, Autonomic resource management for service host platforms, Proceedings of Workshop on Software Engineering Challenges in Cloud Computing, pp. 1-8.

[13] Verma, A., Ahuja, P., and Neogi, A., 2008, pMapper: Power and Migration Cost Aware Application Placement in Virtualized Systems, Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware, pp. 243-264.

[14] W Kleiminger, E Kalyvianaki, P Pietzuch, "Balancing Load in Stream Processing with the Cloud", 27th IEEE International Conference on Data Engineering

ISSN (Online) : 2348 - 2001
International Refereed Journal of Reviews and Research
Volume 3 Issue 2 March 2015
International Manuscript ID : 23482001V3I2032015-10
(Approved and Registered with Govt. of India)

Workshops[ICDEW], Zurich, Switzerland, pp 16-21, 11-16 April 2011, DOI: 10.1109/ ICDEW.2011.5767653.

[15] Wood, T., Shenoy, P., and Arun, 2007, Black-box and gray-box strategies for virtual machine migration, NSDI 2007, pp. 229–242.

[16] Xu, J., and Fortes, J., 2010, Multi-objective Virtual Machine Placement in Virtualized Data Center Environments, Proceedings of the 2010 IEEE/ACM Conference on Green Computing and Communications, pp. 179-188.

[17] Yi Lu, Qiaomin Xie, Gabriel Kliot, Alan Geller, James R Larus, Albert Greenberg, “Join-Idle-Queue: A Novel Load Balancing Algorithm for Dynamically Scalable Web Services”, Elsevier, Journal on Performance Evaluation, pp 1056-1071, Vol 68, Issue 11, Nov 2011, DOI: 10.1016/j.peva.2011.07. 015.

[18] Zhao, Y., and Huang, W., 2009, Adaptive Distributed Load Balancing Algorithm based on Live Migration of Virtual Machines in Cloud, Proceedings of 5th IEEE International Joint Conference on INC, IMS and IDC, pp. 170-175.

[19] Gallery Images For (Basic Cloud Architecture) URL - http://imgarcade.com/1/basic-cloud-architecture

[20] Tziritas, Nikos, Cheng-Zhong Xu, Thanasis Loukopoulos, Samee Ullah Khan, and Zhibin Yu. "Application-aware Workload Consolidation to Minimize both Energy Consumption and Network Load in Cloud Environments." In Parallel Processing (ICPP), 2013 42nd International Conference on, pp. 449-457. IEEE, 2013.