Performance Optimization in Multi-cloud Computing: A Review

[1]Vidhya R. Vasekar, [2] Prajakta V. Kshirsagar, [3]Aditi G. Ingawale
[1]Assistant Professor, [2]Student, [3]Student
[1]Computer Engineering Department,
[1]ZCOER, Pune, India

Abstract*:* Many users can use cloud computing technology to access the virtualized hardware and software infrastructure that is scattered and scalable over the Internet. Multi - cloud is a notion of assigning workload across several computers or other resources over network connections in order to achieve optimal use of resources, maximum throughput in minimum response time by ignoring overload. This needs load balancing and task scheduling algorithms and techniques on cloud computing. We have discussed papers that describe systems for load balancing to solve the problem of overall performance of the cloud computing environment which needs efficient task scheduling mechanism to improve resource utilization.

Index Terms - Performance Optimization, load balancing techniques

I. INTRODUCTION

The latest vision of large distributed computing is “Cloud”. Cloud based multi-media system (CMS) gained momentums as there are large number of users. Cloud computing is internet computing and provides demand computers and other devices with shared resources, software and information, such as a public utility. Large distributed computing latest vision is "Cloud."Multi - media cloud - based system (CMS) has gained momentum as there are many users. Cloud computing is internet - based technology, providing computers and other devices with common resources, software and information on demand, such as a public entity. Cloud storage for the IT user is the most burning topic recently. So when a user utilizes a personal or financial computer for a higher purpose, then they also must have some valuable documents that the user is willing to invest more to defend the file. Cloud computing service providers need to handle gigantic requests as the cloud scale increases. So, despite Cloud Computing glorious future, many real issues still need to be explored for their perfect awareness. Load balancing is one of these concerns. Organization uses high - level cloud. As the request for the user increases rapidly, the request must be answered in less time. The one of the biggest problems is load balancing in cloud. The cloud is distributed so that dividing the work load efficiently becomes a vital task. Many solutions have been proposed in the conventional cloud load balancing system

Conventional algorithms have the other big advantage, making them ineffective in cloud computing to keep up with demand. The primary task of load balancing is to improve system performance and efficiency. The better use should be made of the resources. The number of user’s increases, the QoS metric should remain the same. Despite an increase in user, the cloud service should not be affected. The response time when providing the service should be lower. The system should also be fault tolerant so that the system can change to another node or alternative path for distributing load whenever one node fails. Whenever a conflict occurs, due to one - point failure, the system should not be disturbed

By examining the difficulties of conventional algorithms, the genetic algorithm is used to deal with the problems actually faced by distributing the load in the cloud. It seeks the strongest and fastest efficient path and reduces the make span to distribute the work load as well. This algorithm seems to use positive response or acts like the real ant colony to discover the food and search one another with assistance of the pheromone they leave on their journey. The algorithm solves the cloud computing virtualization placement problem. Genetic algorithm is renowned for finding huge spaces and discovering the appropriate integration of things as well. This algorithm is not looking for the best possible solution, but is looking for the best and rigorous result especially compared to fitness criteria. And so, while measuring the fitness value, rather than dividing the appropriate local search, it looks for global. The genetic algorithm is based purely on Darwin's natural selection principle. In short, to work accordingly, the genetic algorithm uses search and scalable technique. In three stages, the genetic algorithm is processed. The first step includes population selection to enhance fit survival.

1.1 Cloud Computing

Apart from storing files on a primary hard drive or local storage device, cloud storage enables remote databases to save files. It is not necessary for the user to have access to information in a specific location as information to be accessed can be found in the "cloud". This particular system enables staff to work from anywhere. Companies allow users to store their database on remote servers, and then provide the data accessibility via the internet [1].

1.2 Multi-Cloud computing

Using multiple cloud computing and storage services in a single architecture which can be heterogeneous is known as Multi-cloud. An organization makes use of two or more cloud computing platforms to perform various tasks, this strategy is known as Multi-cloud. Organizations that don't want to rely on a single cloud provider may choose to use multiple providers ' resources to get the most out of each unique service. Combining software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service model (IaaS) is a multi - cloud solution. It may also refer to the use of several private and public cloud solutions [2].

1.3 Load Balancing

However, Cloud computing is cost - effective and interoperable to maintain processing flexibility so a huge number of tasks in the cloud computing environment it lead to an appalling disadvantage with load balancing receiving a great deal of research attention. In the cloud computing environment, load balancing has a very significant performance effect. Sensitive load balancing makes cloud computing cost - effective and enhances user requirements. The method which equally distributes the work dynamically across all nodes is load balancing, ensuring that no single node is weak, thereby increasing the overall performance and efficiency of the system. Load balancing can be a relatively recent technique by providing maximum response time with minimum response time to facilitate networks and resources. Proper load balancing will promote the ideal use of accessible resources, thereby minimizing resource consumption. There are various types of algorithms that help to load traffic between accessible servers. Websites are associated with a fundamental example of load balancing in our lifestyle.In the absence of load balancing , there will be experience of delays, timeouts and late replies to the problems, also wrong responces.

II. LITERATURE REVIEW

2.1 Load Balancing Task Scheduling

Wang, T. et.al. Mentioned regarding the scheduling and load balancing. A scheduling algorithm carried by double - fitness adaptive algorithm - job spanning time and load balancing genetic algorithm (JLGA) is constructed to solve the issue considering the new features of cloud computing and original adaptive genetic algorithm program (AGA).Then compare JLGA's performance with AGA by simulating [3].

Shalini Joshi et al. [8], Load balancing in cloud computing a major problem. The key objective of load balancing is to satisfy the needs of the user by distributing workload across multiple nodes in the system and maximizing the use of resources and improving system performance. The key objective of load balancing is to satisfy the needs of the user by distributing workload across multiple nodes in the system and maximizing the use of resources and improving system performance. Efficient load balancing is therefore crucial for system performance, resource utilization, stability, maximizing throughput and minimizing response time, which are the main goals of paper.

2.2 Hybrid Job Scheduling Algorithm

Mohammad Shojafar et al. [4], described a hybrid job planning approach, with the assistance of genetic formula and fuzzy theory, taking into account the loading feature of the system and reducing the total execution time and execution value The main priority of the analysis is to allocate the roles to the resources, taking into consideration the measuring unit of the VM and the time - span of the work. The experimental results highlight the importance of the planned approach in terms of time, execution value and average degree of inequity.

2.3 Host Scheduling Algorithm using Genetic Algorithm

Danilo Amendola et al.[5], represents that Scheduling is the major issue in the cloud atmosphere. Scheduling is responsible for the economic use of resources. A minimum network delay - based scheduling model using Suffrages Heuristic coupled Genetic algorithm for scheduling sets of freelance job algorithms is projected with the aim of reducing the make span.

Harshdeep Sharma and Gianetan Singh Sekhon [12], Genetic algorithm is based on biological thinking about population generation, a rapidly growing Artificial Intelligence area. The area unit of GA influenced by Darwin's evolution theory. If that theory is considered,"the fittest's survival". It is also applied for scheduling method in which the jobs are allocated resources in accordance with the schedules in the scheduling situation, which inform us which resource is to be allocated to which task. The program of genetic algorithms is based on the biological idea of population generation. The initial population is considered as the set of all the people used in the genetic algorithm and then finally it is used find the optimal solution. It is determine of an individual's superiority in the community. Good selection mechanism is produced to select an intermediate resolution that supports Darwin law's survival for consecutive generation. Hybrid operation can be accomplished by selecting 2 parents and making an individual tree replacement by alternating and reforming their ancestors ' elements. Mutation occurs when the population tends to be consistent due to continuous use of copy and crossover operator. They Propose a Load Balancing Enhanced Genetic (EGA)-based cloud scheduling policy and main contribution is to maintain the balance in the load of whole system while attempting to minimize the Make span of a set of tasks.

2.4 Efficient Load Balancing

Vlad BUCUR et.al [7], proposed the system which attempt. Which analyses and reviews the offerings of current cloud in the market, also the storage solutions which are provided on commercial as well as on theoretical level, also challenges in MultiCloud and any future research .They conclude in this paper that multi cloud is an emerging as a powerful solution and it offers safer, and reliable storage for many interconnected devices, that can be of any kind for example smart phones, computer or company servers.

2.5 Dynamic Load Balancing

Master - slave design is an intelligent design which has a main server which monitors all the tasks of slave servers and few slave servers, majorly used in Google's cloud computing such as MapReduce and Hadoop. The single task is then breaked into some executable tasks in the main node and the tasks generated are distributed to various slave nodes. Afterwards, with the guidance of the master node, the tasks are executed separately in the slave nodes and the outputs are returned to the master node. Finally, in the master node, the distributed results are combined and sent to the receiving user. In addition, it is the responsibility of the master node to monitor the entire process and again execute the previously failed tasks. The uneven distribution of tasks during this process can cause some slave nodes to be in low load conditions while others are in heavy load conditions. Load balancing procedure for the cloud platform should be performed dynamically to keep the platform stable and efficiently operating.

Prof. Kamal Mistry et.al proposed the concept in which there is assignment of the load on different servers according to the type of content. There are different platforms which consist of linode and cloudsim emulator. Cloudsim emulator is an open source and free of cost tool. Different parameters are checked like performance, efficiency, control and security in cloud etc. Dynamic load balancing will be useful for different community, peoples, organizations in number of ways. Because of load balancing dynamically , so user will have more time and that time can be utilized to do the another useful work. With the help of the load balancing, data management and loading of the data on the cloud becomes easier. As the data get increased so the accessing speed of the data from the cloud should be increased. So in future, existing algorithms should be improvised [9].

How and when to allocate and arrange tasks, to optimize the use of resources and avoid overload is a challenging. Ren Gao and Juebo Wu [13], They introduce a new load balancing approach through ant colony optimization (ACO) to dynamically balance the workload on a cloud computing software platform. Two methodologies are adopted, forward - looking ant mechanism and max - min rules, to quickly determine the candidate nodes for load balancing. They devise initialization of pheromones and update pheromones according to physical resources, including evaporation of pheromones, rules of incentive and punishment, etc. In combination with the the the prediction of task execution, they described the moving likelihood of ants in two ways, i.e. whether either forward ant meets the backward ant or not in the neighboring node, with the aim of speeding up the search process.

2.6 Performance Optimization In Multi-clouds

Huiyan Cao et al. [11], constructed a mathematical model for analysis of cloud execution of workflow and construct a budget constrained workflow mapping problem for optimization of network performance in Hadoop system in multi-cloud environments. The superiority of performance of the suggested mapping system compared to already estimated systems, they defined it by extensive simulations, also verified those by real life workflow experiments which were deployed in public clouds. 15% of lack of compatibility and variation were noticed between the theoretical approximate calculations and practical experimental measurements, which has validated the accuracy of the cost models and confirmed the accurate workflow mapping in real system. The three layer workflow execution architecture were used for MapReduce workflow execution in multi-cloud along with time and financial cost models where practically in real time networks there is a possibility of failure of servers under certain circumstances and the actual workload of information and workflow module is always dependent on real time changes.

Table 1 A Three layer Model to check the performance in multi-cloud

Layer	Description
Top	Used to define workflow structure with various MapReduce based modules. This module shows the dependency with intra- and inter-cloud data transfer and execution.
Middle	Used to define cloud-.model or network of virtual machines. These virtual machines occupied on physical machines. The physical machines are located at different data centers.
Bottom	It consists of number of distributed data centers. The data centers includes group of physical machines connected with speedy network.

III. OPEN CHALLENGES IN MULTI-CLOUD

Multi - cloud became the new normal for most firms. The 2018 State of the Cloud Report of RightScale found that 81 percent of firms do use multiple public or private clouds. They actually use 4.8 different clouds on average. This increased dependence on the cloud especially the public cloud, has made things easier for IT professionals in some contexts. It lowers the need for research, selection, deployment and maintenance of their own IT infrastructure. But multi - cloud has created new biz IT challenges in other ways. As complexity tends to increase, as it eventually does with Multi-cloud, it becomes much more difficult to monitor what is happening in all the different cloud environments. Some of the challenges in multi-cloud are discussed below [14].

3.1 Constant change

Because of the flexibility and scalability that this approach promises, many companies decide to pursue multi - cloud strategies. The cloud's greatest advantage, however, is also one of its greatest disadvantages. Dave Anderson, Digital Performance Expert at Management and Monitoring Vendor Dynatrace, explained in an email interview that the biggest challenge in multi - cloud is to "keep up with what's going on, where and how we're going to do this on Earth when the environment continues to change every second."

3.2 Siloed cloud vendor tools

Many organizations depend on the free or nearly free - cost tools provided by the vendors themselves to keep tabs on what's happening in their cloud environments. In reality, 37 percent of respondents said they rely on the monitoring services provided by their public cloud vendors in the Interop ITX 2018 State of the Cloud report. Unsurprisingly, these tools have a major drawback: maybe they just can only scan and handle workloads that run on the infrastructure of that one vendor. If IT monitors workloads — and they are — for multiple vendors, they must use multiple tools to get the job done. This also does not include workloads in their own data centers running in private clouds or traditional infrastructure. Anderson told a customer's story, "We've got a hybrid cloud monster made up of multiple clouds, and yet we still have a mainframe." Public cloud vendor monitoring tools ca n't provide a consolidated view of all these different environments.

3.3 No cloud support in legacy tools

For their in - house data centers, most organizations already have monitoring tools that they use. They may want to extend their use of these tools into the cloud in a perfect world. Most legacy tools, however, were not designed to monitor the cloud. While some legacy vendors revised their monitoring software to provide some cloud support, most fail to provide the amount of detail and detailed capabilities many businesses are looking for.

3.4 Challenges in selecting multiplatform monitoring tools

Companies often start shopping for a hybrid or multi - cloud monitoring and management tool in response to all these challenges. But it can be expensive for these tools. They also offer such varied and wide - ranging feature sets that it can be difficult to compare. For example, Gartner's Selection criteria for Systems and Tools for Cloud Management identify 215 different IT d criteria.

3.5 Lack of skills

When a new multiplatform monitoring tool is deployed by companies, they often must retrain staff. The rapidly evolving technology also make it difficult for companies to find employees with the multi - cloud skills they need. Whereas the lack of skills and other challenges may present obstacles to multi - cloud strategies enterprises, these challenges are not intractable.

IV. CONCLUSION

Multi-cloud provides solution to many industry problems, it is necessary to optimize its performance further to improve its use and make it efficient. Hence with the help of job scheduling algorithms also load balancing algorithms it is possible to enhance the performance to solve the problems like load balancing ,task scheduling ,also security challenges.

REFERENCES

[1] CLOUD COMPUTING. HTTPS://WWW.INVESTOPEDIA.COM/TERMS/C/CLOUD-COMPUTING.ASP

[2] Multicloud. https://en.wikipedia.org/wiki/Multicloud

[3]Wang, T., Liu, Z., Chen, Y., Xu, Y., & Dai, X. Load Balancing Task Scheduling Based on Genetic Algorithm in Cloud Computing. 2014 IEEE 12th International Conference on Dependable, Autonomic and Secure Computing.

[4]Mohammad Shojafar,Danilo Amendola,Nicola Cordeschi,Hongbo Liu, Ajith Abraham . Hybrid Job Scheduling Algorithm for Cloud Computing Environment. Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014.

[5] SaeedJavanmardi, Mohammad Shojafar, DaniloAmendola, Nicola Cordeschi,Hongbo Liu, and Ajith Abraham. Hybrid Job scheduling Algorithm for Cloud computingEnvironment .2014.

[6] Simranjit Kaur, Tejinder Sharma .Efficient Load Balancing using Improved Central Load Balancing Technique. Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018).

[7] Vlad BUCUR, Liviu MICLEA, Catalin DEHELEAN. Object storage in the cloud and multicloud :state of the art and the research challenges .978-1-5386-2205-6/18/$31.00 ©2018 IEEE.

[8] Shalini Joshi , Uma Kumari .Load Balancing in Cloud Computing: Challenges & Issues. 2nd International Conference on Contemporary Computing and Informatics (IC3I) 2016.

[9] Prof. Kamal Mistry, Mr. Rayan Dasoriya , Ms. Purvi Kotadiya. Dynamic Load Balancing in Cloud. 2017 International Conference on Networks & Advances in Computational Technologies (NetACT) |20-22 July 2017| Trivandrum.

[10] Prof.Jignesh Vania, Foram F Kherani. Load balancing in Cloud Computing. © 2014 IJEDR | Volume 2, Issue 1 | ISSN: 2321-9939.

[11] Huiyan Cao, Chase Q. Wu. Perfomance optimization of budget constrained MapReduce workflows in multiclouds.2018 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing

[12] Harshdeep Sharma, Gianetan Singh Sekhon. Load Balancing in Cloud Using Enhanced Genetic Algorithm. International Journal of Innovations & Advancement in Computer Science, ISSN 2347 – 8616 Volume 6, Issue 1 January 2017.

[13] Ren Gao, Juebo Wu. Dynamic Load Balancing Strategy for Cloud Computing with Ant Colony Optimization. Future Internet 2015, 7, 465-483; doi:10.3390/fi7040465.

[14] Multi-cloud challenges. https://www.networkcomputing.com/cloud-infrastructure/top-5-challenges-monitoring-multi-cloud-environments.