DYNAMIC LOAD BALANCING IN CLOUD BASED MULTIMEDIA SYSTEM WITH GENETIC ALGORITHM

Vinza V Suthan
Department of CSE
SCT College of Engineering
Trivandrum, India
vinuvinza@gmail.com

Kavitha K V
Department of CSE
SCT College of Engineering
Trivandrum, India
kavitha279@yahoo.co.in

***Abstract*—Nowadays cloud computing is the most advanced paradigm promising to show the vision of computing applicability into reality. It provides a versatile and straightforward way to store and retrieve immense information left-out concern the hardware required. Cloud based Storage (CS) incorporates a resource manager, cluster head and server clusters. Inside the resource manager the client's request for information service tasks to server clusters in line with the assignment features. And each cluster head distributes the assigned task to the servers inside its server cluster. In Cloud based Storage every server clusters completely handles a particular kind of information service and receives the client's request dynamically in different time steps. It is a research challenge to design an efficient load balancing algorithm which can assign the multimedia message jobs with minimum cost between server clusters and clients while not overloading the server clusters. Differing from previous works, this paper takes into account a more practical dynamic services scenario in which each server clusters handles only a particular type of multimedia task and each client request different type of multimedia services at different time. Such scenario can be handled as an integer linear programming problem which is computationally in feasible in general .In this paper an attempt to solve the problem by an efficient GA in immigrant scheme Simulation results exhibit that the proposed genetic algorithm can efficiently cope with dynamic multi-service load balancing in CMS**

Keywords— Genetic algorithm, metaheuristic, load balancing, cloud computing, multimedia system.

I. INTRODUCTION

Cloud computing is an emerging technology providing various computing and storage services over the Internet. It generally integrates infrastructure, platform, and software as a service. In the fast developing era of technology and communications, Cloud-based multimedia systems (CMS) are advancing to meet a large number of client services for different types of multimedia content processing and storage operations through the internet similarly [1]. It also has to support various protocols and devices to access the data. In general, all cloud-based multimedia systems and its platform software's are contributing connectivity for a large number of customers at a time for keeping and processing media data in a clustered fashion with set of expected Quality of Services though internet. With these highly configured cloud-based systems, clients will benefit to minimize the cost of IT infrastructure and maintenance by paying only for resources that are consumed by them over given period of time [2].

In dynamic load equalization the lightest server within the whole system is searched and most well-liked for equalization the load. Wireless system captures the advantage of IEEE802.3ad to dynamically alter the usage of information measure. Basically the dynamic load equalization is tough to simulate however best suited in heterogeneous atmosphere. Load equalization is required to issues the surplus dynamic work fairly to entire node within the whole cloud to realize the high resource utilization. Load equalization could also be static or dynamic. In static load equalization all data concerning equalization selections are celebrated earlier and ignore the present state of the system. However dynamic load equalization algorithms gather state data and react to system state if it modified. Migration time is outlined because the total time needed in leave the resources from one node to a different and fault tolerance allows the rule to continue operative accurately within the event of some failure. Migration is the rule capable to decide to create a decision to come to a arrangement to make your mind up to choose to determine that it ought to make changes of load distribution throughout beheading of method or not. The migration time is taken for the analysis of the performance of each algorithm [3].

Here centralized hierarchical CMS (as shown in Figure 1) composed of a resource manager and a number of server clusters, each of which is coordinated by a cluster head, and assume the servers in different server clusters to provide different services. Such a CMS is operated as follows. Every time when the CMS receives clients' requests for multimedia service tasks, the resource manager of the CMS assigns those task requests to different server clusters according to the characteristics of the requested tasks. Subsequently, the cluster head of each server cluster distributes the assigned task to some server within the server cluster. It is not hard to observe that the load of each server cluster significantly affects the

performance of the whole CMS. In general, the resource manager of the CMS is in pursuit of fairly distributing the task load across server clusters, and hence, it is of importance and interest to be able to cope with load balancing in the CMS.

II. RELATED WORKS

The template More research studies has been done on load balancing mechanisms, especially over wireless networks based on different factors, scheme, and theories. Existing works of load balancing on CMS to enhance the multimedia process over clustered servers make less transmitting payload information between clients and CMS clustered servers without losing QoS and overloading on the servers. In general, CMS provides many services to the clients to operate different tasks like keeping, organizing, sharing, searching the multimedia content like images, animations, audio, video, etc. It also supports different devices like laptops, tablets, smart phones etc for performing these operations. With lots of innovation in communications and technologies, some of the vendors are also delivering advanced features like camera-to-cloud to keep the live collecting data into cloud directly, multi-user concurrent operations (for team collaboration), live streaming, single & multi clients 3D graphical gaming etc. There is a lot of variation in utilization of server resources based on type of multimedia assignment for e.g., server need more CPU, RAM and storage space for processing gaming or videos when analyze to images or web pages. The earlier works which are on time based are won't be considerable for this research work.

To encounter above dynamic real time load balancing essential, we reinforced each CMS cluster group with particular multimedia process jobs so that each customer request of specific type always redirected to that particular cluster group based on the job type or characteristic. Thereafter, each customer requests multiple types of services at specific times and seizes load balancing issues for every trial and modeled them into mathematical optimizing program models for analyzing the problem in mathematical approach. This model helps to learn or discover the problem solution with practical methods and it also observed that the results of this proposal have shown significant improvement in load balancing of CMS by issuing the load equally across clustered servers. This approach may or may not impact other problem solving techniques of cloud distributed techniques.

To respond to the empirical requirements mentioned above, we assume that in the CMS, each server cluster can only handle a particular type of multimedia service task, and each client requests a different type of multimedia service at various times. At each specific time step, such a problem can be modeled as an integer linear programming formulation, which is computationally intractable in general [4]. Conventionally, intractable problems are usually solved by metaheuristic approaches, e.g. for simulated annealing [5], genetic algorithm [6], particle swarm optimization [7], [8], etc. Genetic algorithm (GA) for the agitated dynamic load balancing problem for CMSs. GA has already found applications in a variety of areas in computer science and engineering, such as quick co-variance matching [9], aircraft ground service scheduling problem [10], optimal electric network design [11], between others. In our setting of GA, elite immigrants and random immigrants are appended to new population, because they are suitable for solving the problems in dynamic environments [12]. The exploratory results show that to a certain extent, our approach is capable of dynamically spreading the multimedia task load evenly.

III. MATERIALS AND METHODS

Before The figure 1 showing the infrastructure of cloud storage model. Cloud based mostly Storage (CS) accommodates a resource manager, cluster lead and server clusters. The resource manager nominate client's request for knowledge service tasks to server clusters in line with the task characteristics. Every cluster head assigns the appointed task to the servers inside its server cluster. In cloud storage every server clusters solely handles a chosen sort of knowledge service and dynamically receives the client's request from resource manager on fully different time steps. The information is also any range of formats like pdf, html, txt, doc, and java. Effective genetic formula nearly new for optimizing CS – dynMLB.CS–dynMLB considers a centralized system composed of a resource manager and collection of server clusters, every of this coordinated by a cluster lead, and also the servers are fully different in different numerous server clusters to produce different services. Such a CS is operated as follows. Whenever the CS receives customer's requests for knowledge service tasks, the resource manager of the CS assigns those task requests to absolutely different server clusters in line with the characteristics of the requested tasks. Subsequent, the cluster lead of every server cluster distributes the appointed task to some server inside the server cluster.

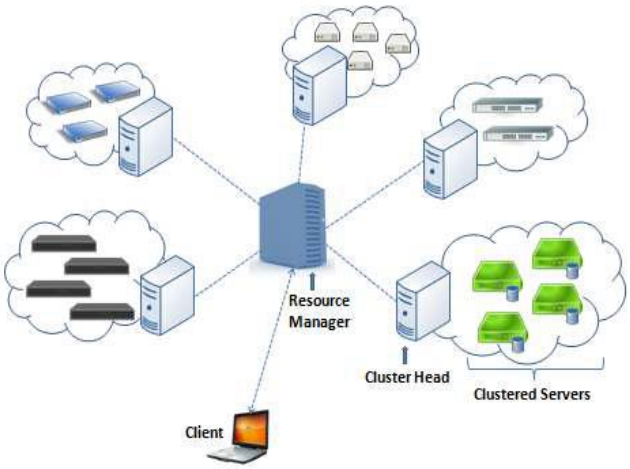


Figure 1 CMS Architecture [14]

IV. FUNCTIONS OF CS MODEL

A. Client's Requests and Request assignation

Customers request the service to the server cluster for storing the info in cloud. Request passed to the resource manager for more method. When the system receives customer's requests for storage services it assigns customers requests for cache service tasks to server clusters consistent with the task characteristics, so every cluster lead distributes the assigned task to the servers at intervals to its.

B. Resource manager

Resource manager monitors the entire, request from the clients. It pass this request to the server cluster and the resource manager of the system reserve, the service task load data collected from server clusters, and resolves the number of customer's requests allotted to every server cluster. The decision of assignment is predicated upon the characteristics of different service requests and also the data collected from server clusters.

C. Complete bipartite graph

Client and server clusters square compute thought of as a node for a whole weighted bipartite graph. At this step, the system will be viewed of as an whole weighted bipartite graph. Set of vertices that represent shoppers and edges are formed between shoppers and server clusters.

D. Weight calculation

Proximity between server cluster and shopper is required to be measured at any time step owing to dynamic modification of topology. Genetic algorithmic program is utilized to find a domestically optimum load charge answer[13]. Server utilization ratio is magnitude that removes the violating links, to find an optimal solution and also to compute the final load charge solution. For each available server cluster in the set that includes the server clusters connected to client, and deduct the landmark distance of server cluster.

V. OPERATIONS OF GENETIC ALGORITHM

A. Selection

Choose collection of parent people from the population in step with the fitness price.

B. Cross over

Crossover operation is done by selecting 2 parent people then making a replacement individual by alternating and reforming the components of these oldsters.

C. Mutation

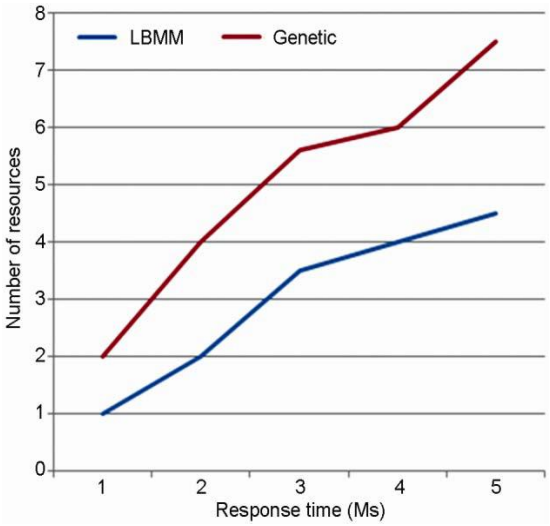
Mutation operator wants to conserve genetic diversity from one generation of a population of genetic formula chromosomes to future.

D. Fitness Function

Fitness perform is to decide the chromosomes area unit higher or worse. This performance may manage the trend of population development, therefore it's a crucial a part of GA. In this work fittest solution has minimum execution time, minimum communication cost, maximum processor utilization and maximum through put.

VI. COMPARISON

Taken response and migration time for the contrast of both algorithm performance. Migration time makes the algorithm skilled to resolve that it should make changes of load distribution during execution of process or not. Response time means amount of time taken to respond by a load balancing algorithm in a distributed system. For better performance, this parameter should be reduced. In Genetic algorithmic regulation the response time of the 2for 8time and response time depends on the amount of data to be stored in the cloud at the same time. If the number of data grows the processing time to complete the storage increases that make the migration time to increase. In LBMM rule, number of data increase gives greater response time compared to GA. This is compared to GA response time is within 3ms for 5reduced in GA compared to LBMM rule. Moving means the time to resolve or move the resource from one node to another. To increase the performance, time should be reduced. Genetic and LBMM algorithm has better results for the execution of cloud based storage model. But, genetic algorithm had much better when differentiate, to LBMM. The figure 2 shows the comparison between genetic algorithm and LBMM algorithm.



VII. RESULTS AND DISCUSSION

Genetic algorithm mimics the procedure of natural selection which is used as a problem solving technique. At first level of genetic algorithm, individuality is randomly selected from the population to see the closeness between them to solve the problem. An individuality who is fit than others of the population, are allowed to produce the next generation gives the best solution of the problem. Genetic algorithm (GA) is developed to find the most optimized solution for a given job. This algorithm is relevant to search for the solution of high degree of complexity that often involves attributes that are large, non-linear and discrete in nature.

VIII. CONCLUSION

Occurring Load Balancing techniques Algorithms that have been considered largely focus on reducing overhead, reducing the moving time and improving performance etc., but the response to request ratio is rarely considered. It is a difficulty of every engineer to build up the cloud platforms that can raise the throughput. A genetic algorithm approach for perfect the CMSdynMLB was proposed and implemented. The main difference in CMS model from previous models is that CMS

look at a practical multiservice dynamic scenario in which at different time steps, clients can change their locations, and every server cluster only handled a specific type of multimedia task so that two performance objectives were optimized at the same time. The main characteristic of this paper included not only the proposal of a mathematical establishment of the CMS-dynMLB problem but also a theoretical analysis for the algorithm convergence. A detailed simulation was also conducted to show the performance of GA approach.

REFERENCES

[1] W. Zhu, C. Luo, J. Wang, and S. Li, "Multimedia cloud computing: An emerging technology for providing multimedia services and applications," IEEE Signal Process. Mag., vol. 28, no. 3, pp. 59–69, May 2011.

[2] W. Hui, H. Zhao, C. Lin, and Y. Yang, "Effective load balancing for cloud-based multimedia system," in Proc. Int. Conf. Electron. Mech. Eng. Inform. Technol., 2011, pp. 165–168.

[3] H. Cheng and S. Yang, "Genetic algorithms with immigrants schemes for dynamic multicast problems in mobile ad hoc networks," Eng. Appl. Artif. Intell., vol. 23, no. 5, pp. 806–819, 2010

[4] M. Garey and D. Johnson, Computers and Intractability—A Guide to the Theory of NP-Completeness. San Francisco, CA, USA: Freeman, 1979Y.

[5] J. H. Holland, Adaptation in Natural and Artificial Systems. Ann Arbor, MI, USA: Univ. Michigan Press, 1975.

[6] J. Kennedy and R. Eberhart, "Particle swarm optimization," in Proc. IEEE Int. Conf. Neural Netw., Nov.–Dec. 1995, pp. 1942–1948.

[7] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in Proc. IEEE Int. Conf. Evol. Comput. May 1998, pp. 69–73.

[8] X. Zhang, S. Hu, D. Chen, and X. Li, "Fast covariance matching with fuzzy genetic algorithm," IEEE Trans. Ind. Eng., vol. 8, no. 1, pp. 148–157, Feb. 2012.

[9] W. Ip, D. Wang, and V. Cho, "Aircraft ground service scheduling problems and their genetic algorithm with hybrid assignment and sequence encoding scheme," IEEE Syst. J., vol. PP, no. 99, p. 111.

[10] F. Gonzalez-Longatt, P. Wall, P. Regulski, and V. Terzija, "Optimal electric network design for a large offshore wind farm based on a modified genetic algorithm approach," IEEE Syst. J., vol. 6, no. 1, pp. 164–172, Mar. 2012

[11] H. Cheng and S. Yang, "Genetic algorithms with immigrants schemes for dynamic multicast problems in mobile ad hoc networks," Eng. Appl. Artif. Intell., vol. 23, no. 5, pp. 806–819, 2010

[12] A. Pathan and R. Buyya, "A taxonomy and survey of content delivery networks," Grid Computing and Distributed Systems (GRIDS) Laboratory, University of Melbourne, Parkville, Australia, vol. 148,2006.

[13] D.E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning. Reading, Mass.: Addison-Wesley, 1989.M.

[14] Suresh Babu Kuntumalla and Lakshumaiah Maddigalla, "Enhanced Load Balancing in Clustered Cloud-based Multimedia System", International Journal of Engineering Trends and Technology (IJETT) – Volume 29 Number 4 - November 2015.