Distributed Cloud Computing and Distributed Parallel Computing: A Review

Zryan Najat Rashid
Computer Science Institute
Sulaimani Polytechnic University
Sulemani, Iraq
zryan.rashid@spu.edu.iq

Subhi R. M. Zebari
Akre Technical Institute
Duhok Polytechnic University
Akre, Iraq
subhi.rafeeq@dpu.edu.krd

Karzan Hussein Sharif
Computer Science Dept.
University of Human Development
Sulemani, Iraq
karzan.hussein@uhd.edu.iq

Karwan Jacksi
Computer Science Dept.
University of Zakho
Zakho, Iraq
karwan.jacksi@uoz.edu.krd

Abstract-- In this paper, we present a discussion panel of two of the hottest topics in this area namely distributed parallel processing and distributed cloud computing. Various aspects have been discussed in this review paper such as concentrating on whether these topics are discussed simultaneously in any previous works. Other aspects that have been reviewed in this paper include the algorithms, which simulated in both distributed parallel computing and distributed cloud computing. The goal is to process the tasks over resources then readjusted the calculation among the servers for the sake of optimization. These help us to improve the system performance with the desired rates. During our review, we presented some articles which explain the designing of applications in distributed cloud computing while some others introduced the concept of decreasing the response time in distributed parallel computing.

Keywords—Distributed Cloud Computing, Distributed Parallel Processing, Load Balancing, Scheduling, Measuring Throughput, Response Time.

I. INTRODUCTION

A recently growing technology, that is Cloud computing, entailing scheduling mechanism as a vital part, allows for the processing of huge amounts of data [1].

One field where we increasingly see cloud computing is high energy physics (HEP). In this study, we explore the reasons why they are integrated in the applications of HEP and how it is gradually becoming more common [2].

Cloud computing is a broad term generally referring to hosted services. It is the virtualization of physical hardware where data is organized in specified centres. Yet as a new technology with several merits, it is not without its issues, a major challenge is load balancing [3].

This article proposes the implementation of robot SLAM architecture in order to fulfil the Real-time requirement of practical robot systems, which is essential. The robot SLAM adopts two paralleled threads processing in order to fulfil this role. The computational complexity is dominantly determined by the particle number employed, two distributed threads of variable particle sizes are simultaneously executed [4].

In cloud computing services, Virtual clusters allows for the allocation of visualized resources energetically. User management and virtual storage is one of the commonest applications of cloud environment by IT and business companies. In cloud services, the input position is job organization, and its arrangement is vital for the competence of the whole cloud services. The range of appropriate property for job implementation is the mechanism for job implementation [5].

Large scale industrial processing of big data entails monitoring and modelling issues, and in order to deal with this, the approach of distributed and parallel designed principal component analysis is proposed. The large-scale process is decomposed initially into distributed blocks based on a priori process knowledge; this solves the problem of high dimensional process variables [6].

Some nodes can be lightly loaded while others heavily loaded in a cloud system [7], resulting in a poor performance. In cloud environment, distributing the load between the nodes is the function of load balancing, which is the spotlight of problems in cloud computing [8].

Even load distribution in the cloud system leads to a better resource utilization and is much desired [7].

In order to balance the total load on the system in attempt to ensure a good overall performance relative to some specific metric of system performance, a load balancing algorithm [9] [10] [11] transparently transfers the workload from heavily loaded nodes to lightly loaded nodes. The response time of the processes is what determines performance when the metric is involved. Yet the metric is total system throughput when performance is considered from the resource point of view [12]. Is the function of throughput to ensure fair treatment of all the users and that they make progress [9] [10] [12] [11] [8] [7], this in contrast to response time [10].

II. DISTRIBUTED CLOUD COMPUTING

Scientific applications utilizing clouds is seemingly a rising interest among researchers, at the same time many large corporations are contemplating switching over to hybrid clouds. Parallel processing is needed for effective execution of jobs by complex applications. In parallel process the presence of communication and synchronization allows for more effective use of CPU resources. Thus overall, maintaining the level of responsiveness of parallel jobs while achieving the effective utilization of nodes is mandatory for a data center [1].

978-1-5386-6696-8/18/$31.00 ©2018 IEEE

Through cloud computing, a client can request information, shared resources, software and other services at any time, according to his specifications. It is an on-demand service, and the term is commonly seen across the internet. You can view the whole internet as a cloud. Not to mention, utilizing cloud decreases the capital and operational costs. However, a major challenge in cloud computing is load balancing, a distributed solution to this issue is always required. Because of the complexity of cloud and widespread distribution of its component, it is difficult to have efficient load balancing by assigning jobs to appropriate servers and clients individually, and it is not cost effective or practical fulfilling the required demands by maintaining one or more idle services. While jobs are assigned, some uncertainty is attached [3].

A protocol is thus proposed and designed, the purpose of which is enhancing server throughput and performance, enhancing resource utilization and switching time. In this protocol, it is in the cloud where the jobs are scheduled and within the existing protocols the drawbacks are solved. In order to minimize the waiting and switching time, the job that offers better performance to the computer are given priority. Solving drawbacks of existing protocols by managing the scheduling of job has taken considerable effort, along with improvising the throughput and efficiency of the server [1].

Scientific applications utilizing systems based on cloud computing allows for high throughput computing (HTC). Applications in particle physics have systems with a unified infrastructure that utilize a number of distinct IaaS clouds. There are a number of criteria that the system of our cloud computing is based on. The applications of embarrassingly parallel single HEP need to run in a batch environment in the system. No inter-node or inter-process communications are required; However, the memory footprint that sharing process memory produces are reduced by using multi-process jobs [2].

Topology island formation and fast network topology processing is realized by CIM parallel topological processing, which the power network is based on and is discussed in this paper. High throughput, high reliability and high availability storages for cloud storage platform are designed by them. Also, the design idea of MySQL-CIM model was introduced by them as well and efficient tube MySQL-CIM model in data is realized by Ogma development. The development of power network topology processing (NTP) application is done [13].

Cloud computing is a broad term generally referring to hosted services. It is the virtualization of physical hardware where data is organized in specified centres. Yet as a new technology with several merits, it is not without its issues, a major challenge is load balancing, and an important topic in cloud computing requiring many researches and studies. Since many systems are involved in the structure of data centre, it is difficult to perform load balancing, especially in case of cloud computing. Majority of researches on the subject were have been done in distributed environment, yet using semi-distributed load balancing has been the target of so little research. Load balancing in semi-distributed way would allow the design of new algorithm for cloud computing [3].

The merits of decentralized computing away from data centres, along with the consideration of using infrastructure obtained from multiple providers and changing the infrastructure of cloud has been discussed in this paper. A new architecture for computing is demanded by these new trends and need to be fulfilled by cloud infrastructure in the future. Self-learning systems, service space, data intensive computing and people – device connections are areas expected to be most influenced by these new architectures. Finally, in order to realize the next generation cloud system's potential, a roadmap of the challenges that need to be considered has been proposed. [14]

Creating ad hoc clouds and harnessing computing for online applications and mobile both at the edge of the network, using computing models based on voluntarily providing resources has been discussed in this paper. The idea of paying for a cloud VM even if the server executing on the VM is idle is a traditional notion, and we have presented a computing model in order to replace it. An uprising computing model of cloud has been mentioned in this paper that integrates resilience and is software defined. Some areas will be influenced by newly forming computing architecture and changing cloud infrastructure. The internet-of-Things paradigm will be eased further enhancing the connectivity between people and devices, and new architectures will play a vital role. The volume of data provides a challenge in data intensive computing and novel techniques are needed to address this. There will be rising interest in new services such as acceleration, containers and function. Self-learning system will be realized when there will be convergence of search areas with cloud systems. The academia and the industry are leading forces in these changes, yet many challenges need to be solved in the future. Development of next generation cloud computing with sustainable systems and efficient management, expressing applications and improving security have been discussed in this paper, along with their approach and direction [14].

Grid computing distributed computing, and parallel computing developments are part of cloud computing. Firstly, two traditional parallel programming models are introduced in this paper, along with expounding the concept of cloud computing. Secondly, it analyses and studies the principles, advantages and disadvantages of OpenMP, MPI and Map Reduce, respectively. Finally, it discusses and compares MPI, OpenMP models and Map Reduce from the angle of cloud computing. The results of this paper are intended to provide a reference for the development of parallel computing. [15]

Great emphasis has been made on distributed computing in this paper. The differences between distributed and parallel computing has been studied as well, along with terminologies, task allocation, performance parameters, the advantages and scope of distributed computing, as well as parallel distributed algorithm models [16].

III. Distributed Parallel Processing

Some nodes can be lightly loaded while others heavily loaded in a cloud system, resulting in a poor performance. In cloud environment, distributing the load between the nodes is the function of load balancing, which is the spotlight of problems in cloud computing [7].

The Quality of service needs to meet the standards regarding the job arrangement to make the virtual users happy. Decreasing source price, minimizing makespan and safeguarding error

acceptance along with quality of service are used to improve resource allotment and job arrangement. Cloudsim toolkit has been used through the existing scheduling policies to evaluate the proposed algorithms. Early results of experiments have shown better results in terms of user response, execution time and cost, and time on different cloud workloads for the proposed framework compared to the already existing algorithms. Different virtual machine scheduling algorithms along with their performance that are based on various quality metrics has been discussed in this paper. The approach depends on CPU performance, network consumption, scheduling success grade, standard implementation time and so on.

We have shed some light on processing user request utilizing caching in the network itself. By making use of parallel processing a new catching strategy was introduced and its performance was heavily evaluated in terms of reducing redundant traffic and data access delay in different catching scenarios.

The cost of implementing the aforementioned catching network has also been studied. The improved performance by having decreased delay comes at the price of increased cost, and when implementing the proposed parallel processing, the trade-off needs to be carefully considered. The new strategy has shown marked effectiveness in the simulation results [17].

In order to improve the real time performance of PF based robot SLAM, an effective parallel implementation has been proposed in this paper. The acceleration of the overall SLAM algorithm can be increased by the discussed distributed parallel idea, because in the scenario when a keyfram laser can is grabbed, only a large number of particles is used. The results we have obtained from our experiments have verified that the temporal cost can be cut off effectively using this distributed architecture [4].

P. Srinivasa Rao et al. in [18] referred to the elements to take a balanced approach to effective, the information nodes of all other nodes. When a node receives a job, it has to query the status of the other buttons to find out which node has less usage to forward the work, and when all the nodes are query phenomenon overload happens. With the nodes broadcast a statement informed about its status will also cause huge load on the network. Next is the issue of time wasted at each node to perform query status. Besides, the current state of the network is also a factor affecting the performance load balancing. This is because, in a complex network, with multiple subnets, the network node configured to locate all the other node is a fairly complex task. Thus, querying the status of nodes in the cloud will affect the performance load balancing. [18]

Reference [19] showed that factors like response time will greatly affect the performance load balancing on the cloud. Another study outlined two outstanding issues of the previous algorithm: i) load balancing occurs only when the server is overloaded; ii) continuous information retrieval resources available lead to increased computational cost and bandwidth consumption So the authors based on the response time of the request have proposed an algorithm to assign the required decisions for servers appropriately, the algorithm approach has reduced the query information on available resources, reduced communication and computation on each server.

According to the algorithm, Min – Min [20] minimises the time to complete the work in each network node; however, the algorithm has yet to consider the workload of each resource. Therefore, the authors proposed the algorithm of Load Balance Improved Min-Min (LBIMM) to overcome this weakness. Failure to consider the workload of each resource leading to a number of resources are overloaded, some resources are idle; therefore, the work done in each resource is a factor affecting the performance load balancing on the cloud. There is a simple traditional Min - Min Algorithm and the current scheduling algorithm in cloud computing is based upon it.

Kapur in [21] proposed the algorithm of LBRS (Load Balanced Resource Scheduling Algorithm) to consider the importance of resource scheduling policies and load balancing for resources in cloud. The main goals are to maximise the CPU utilisation, maximise the throughput, minimise the response time, minimise the waiting time, minimise the turnaround time, minimise the resource cost and obey the Fairness Principle. In here, the parameters mentioned are QoS: throughput, response time, and waiting time. We have simulation and analysis of data on the impact of the parameters on cloud to perform load balancing. From there, we discovered that, the parameter of makespan (runtime) is of great significance for the data centre cloud. So, the task of the researchers is to study that the algorithms have effective load balancing to reduce the time makespan of virtual machines.

To achieve utilization of resources in an optimal way, the dynamic workload is distributed by the load balancing across multiple resources and this prevents single from being underutilized or overwhelmed; this however is a considerable optimization problem. A strategy for load balancing originating from Simulated Annealing (SA) has been proposed in this paper, and balancing the cloud infrastructure load is its primary function. A traditional Cloud Analyst simulator is modified and the effectiveness of the algorithm is measured. In comparison to existing approaches, like First Come First Serve (FCFS), local search algorithms i.e. Stochastic Hill climbing (SHC) and Round Robing (RR), the proposed algorithm has shown a better overall performance [22].

We tried to maximise the utilisation of resources to keep working resources available for tasks that are yet to come and also concentrate on the reliability of cloud services. We propose a new scheduling algorithm called Dabbawala cloud scheduling Algorithm based on Mumbai Dabbawala delivery system. In our proposed system, the tasks are grouped according to its cost required to complete in a Cluster and its VM resources. We found the lowest cost cluster and its VM for each task requested and group it together for getting services as in the Hadoop Map Reduce model. We have two phases called mapping the tasks and reduce the mapped tasks. The algorithm consists of four Dabbawala for each task to get serviced. From this algorithm, some available scheduling algorithms are compared. We achieve considerable gain in time and resources utilisation. [23].

IV. Existing Load Balancing Technique in Cloud

A. VectorDot

VectorDot is an innovative algorithm proposed by A. Singh et al. [18] for load balancing. It utilizes a flexible data centre with

technologies of storage virtualization and integrated server, to handle the multidimensional resource loads that are distributed across network switches, servers and storages and the hierarchical complexity of the data centre. VectorDot helps in improving overloads on storage nodes, switches and servers, at the same time based on item requirements distinguishes nodes using dot product.

B. LB of VM resources scheduling strategy

A scheduling strategy that utilizes the current state of the system and historical data for load balancing of VM resources, was proposed by J. Hu et al [9]. By implementing a genetic algorithm, this strategy reduces dynamic migration and accomplishes the best load balancing. It achieves a better resource utilization by dealing with the issues of high cost of migration and load imbalances.

C. Task Scheduling Based on LB

A mechanism to gain high resource utilization and satisfy the users' dynamic requirements based on load balancing with a two-level task scheduling, is discussed by Y. Fang et al. [11]. It maps tasks to virtual machines, then host resources, accomplishing load balancing, resulting in a cloud computing environment with better resource utilization, improved task response time and an overall gain in performance.

D. Active Clustering

Optimizing job assignments by using local re-wiring to connect similar services, was the self-aggregating technique for load balancing that M. Randles et al. [9] investigated. Using resources effectively in a high resource system lead to an increased throughput bettering the performance of the system. However, it is degraded with increase in system diversity.

E. Cloud Load Balancing Metrics

Various metrics considered in existing load balancing techniques in cloud computing are discussed below:

1- The number of the executed tasks is measured using throughput, and a high number indicates a good system performance.
2- When applying an algorithm for load balancing, the involvement of overhead is measured by overhead associates. The inter-process, inter-processor and task mobilization composes overhead and the more efficient a load balancing technique is, the less overhead is involved.
3- Load balancing needs to have a good technique for fault toleration. And it is the ability to do uniform load balancing by an algorithm, despite of link or arbitrary node failure.
4- Good performance systems have minimized migration times, and it is the time needed for migration of resources or jobs between individual nods, the less the better
5- Response time is another parameter, which if minimized leads to a better system performance and it is the time needed for a particular load balancing algorithm to respond in a distributed system.
6- Effective resource utilization is mandatory for an efficient load balancing and optimization should be done.
7- Scalability of algorithm is determined by its ability to perform load balancing for any finite number of nodes in a system. Enhanced scalability is desired.
8- The effectiveness of a system is measured by its performance, yet the cost effectiveness needs to considered and kept reasonable. An example would be keeping acceptable delays while decreasing task response times [24].

V. DISCUSSION

Distributed cloud computing is a new technology to interconnect data and applications served from different locations. In information technology, the term ‘distributed’ means that something is shared among multiple users or systems that are geographically different. As shown in Table I, we have some features that we can get from using distributed cloud computing, and each feature has an effect in using cloud technology.

An important feature used in more than one resources is multi-process job feature because an important aim of cloud computing is processing multi jobs at the same time via more than one server if the servers are in different locations. When we have a large amount of data to process, we can divide the big amount of data into small pieces, and each part may be processed by a different server. The aim of this process is to decrease the CPU usage, minimise switching time, minimise waiting time for processing data, improve server throughput and improve performance of the communication and computing of data.

Another feature is designing an application for cloud structure that helps users to easily use any cloud application for make communication with different users in different locations. One more feature is reducing the usage of memory because one problem in the last is using memory, and now after using cloud, user can use memory as they need by contacting with cloud application manage for expanding the memory in this time users only use cloud memory and reducing his storage.

The last important feature is improving server performance because the performance of communication is important. After reviewing all the references mentioned in Table I, we decided that the reference [1] is best works on distributed cloud computing because they cover great number of features that we discussed in above details.

Improving performance of digital computers and its other attributes (like cost effectiveness, reliability and so on) by means of various forms of concurrency is the concern of parallel processing, and achieves this through various algorithmic and architectural methods. There are three types of parallel processing approaches: distributed, shared and hybrid memory systems. In this review, we focused on distributed parallel processing and determined some important features as shown in Table II.

TABLE I. DISTRIBUTED CLOUD COMPUTING SUMMARY

Features	Ref 1	Ref 2	Ref 3	Ref 13	Ref 14	Ref 15	Ref 16
Decrease of CPU Utilization	✓						
Multi process jobs	✓	✓		✓		✓	✓
Minimize switching time	✓						
Minimize waiting time	✓						
Improve resource utilization	✓						
Improve server performance	✓						✓
Improve server throughput	✓						
Load balancing			✓				
Cost efficiency			✓				
Reduce memory usage		✓		✓			
Design an application for cloud structure				✓	✓		✓
Improve connectivity between peoples					✓		
Enhancing security					✓		
Managing efficiency and developing system					✓		
Expand the concept of cloud computing						✓	
Advantages and disadvantages of MPI, oprnMPI and MapReduce						✓	
Compare MPI, oprnMPI and MapReduce						✓	

The important feature mentioned in more than one resource is improving performance using load balancing technique. By using load balancing among servers, we can distribute the process and make balance between servers for processing the jobs and improve performance of our distributed system. Another feature is minimising resource cost because when we divide the load among servers, we can minimise the resource cost such as CPU, memory and storage. All of references implement here idea by proposing an algorithm for using distributed parallel processing based on response time of responding user requests because when the system have minimum response time it is better to user for using that system for responding user requests. After reviewing the references in this paper, we decided that reference [21] is better because it covers a great number of features including load balancing for improving system performance and minimising both response time and resource cost.

TABLE II. DISTRIBUTED PARALLEL PROCESSING SUMMARY

Features	Ref 4	Ref 7	Ref 17	Ref 18	Ref 19	Ref 20	Ref 21	Ref 22	Ref 23
Improve performance using load balancing	✓	✓	✓	✓		✓	✓	✓	✓
Querying the status of each nodes				✓					
Proposed an algorithm based on response time			✓		✓	✓			
Reduce query information on available resource					✓				
Reduce communication and computation on each server					✓				
Consider the workload of each resource						✓			
Maximize CPU utilization							✓		
Minimize throughput							✓		
Minimize response time							✓		
Minimize waiting time							✓		
Minimize resource cost			✓				✓		✓
Safeguard error acceptance and QoS							✓		
Mapping the tasks and reduce mapped tasks									✓
Efficient parallel implementation	✓								
Method to enhanced job arrangement			✓						
Enhance resource allotment			✓						
Better result in execution time			✓						

VI. CONCLUSION

This review paper has covered many ideas in distributed cloud computing and distributed parallel computing. Subjects such as the combination of both algorithms have been devoted in

this review paper. The main goal of this paper relates to the process of distributing workloads over servers and then process them among master and slave's nodes. However, articles have been discussed in this paper include the methodology of designing applications in distributed cloud computing and introducing a concept of optimizing the level of response times while executing user's images.

REFERENCES

[1] L. Tripathy and R. R. Patra "SCHEDULING IN CLOUD COMPUTING", in International Journal on Cloud Computing: Services and Architecture (IJCCSA), Volume: 4, No. 5, October 2014

[2] R. J. Sobie "Distributed Cloud Computing in High Energy Physics", in DCC '14 Proceedings of the 2014 ACM SIGCOMM workshop on Distributed cloud computing, Pages 17–22, 2014, Chicago, IL, USA, August 2014

[3] P. A. Pawade and Prof. V. T. Gaikwad "Semi-Distributed Cloud Computing System with Load Balancing Algorithm", 2014

[4] Xiuzhi Li, Songmin Jia, Ke Wang and Xiaolin Yin "DISTRIBUTED PARALLEL PROCESSING OF MOBILE ROBOT PF-SLAM", in International Conference on Automatic Control and Artificial Intelligence (ACAI 2012), Xiamen, China, April 2013.

[5] A. K.Indira, B. M. K. Devi "Effective Integrated Parallel Distributed Processing Approach in Optimized Multi-cloud computing Environment", in Sixth International Conference on Advanced Computing (lCoAC), pages 17-19, Chennai, India, Dec. 2014

[6] J. Zhu, Z. Ge and Z. Song "Distributed Parallel PCA for Modeling and Monitoring of Large-scale Plant-wide Processes with Big Data", in IEEE Transactions on Industrial Informatics, pages 1877 – 1885, Volume: 13, Issue: 4, Aug. 2017

[7] A. Khiyati, M. Zbakh, H. El Bakkali, D. El Kettani "Load Balancing Cloud Computing: State Of Art", in 2012 National Days of Network Security and Systems, pages 20-21, Marrakech Morocco, April 2012

[8] M. A. Vouk, "Cloud Computing – Issues, Research and Implementations", in ITI 2008 - 30th International Conference on Information Technology Interfaces, pages 23-26, Dubrovnik, Croatia, June 2008

[9] C. Lin, H. Chin, D. Deng, "Dynamic Multiservice Load Balancing in Cloud-Based Multimedia System", in IEEE Systems Journal , pages 225 – 234, Volume: 8, March 2014

[10] Y. Deng, Rynson W.H. Lau, "On Delay Adjustment for Dynamic Load Balancing in Distributed Virtual Environments", in IEEE transaction on visualization and computer graphics, Volume: 18, NO. 4, April 2012

[11] L.D. D. Babua and P. V. Krishna, "Honey bee behavior inspired load balancing of tasks in cloud computing environments", in Applied Soft Computing, Pages 2292-2303, Volume: 13, Issue 5, Amsterdam, The Netherlands, May 2013.

[12] D. Warneke and O. Kao, "Exploiting Dynamic Resource Allocation for Efficient Parallel Data Processing in the Cloud", in IEEE transaction on parallel and distributed systems, Volume: 22, NO. 6, NJ, USA, JUNE 2011

[13] X. L. Xingong and X. Lv "Distributed Cloud Storage and Parallel Topology Processing of Power Network", in Third International Conference on Trustworthy Systems and Their Applications, pages 18-22, Wuhan, China, Sept. 2016

[14] B. Varghese and R. Buyya "Next Generation Cloud Computing: New Trends and Research Directions1", in Future Generation Computer Systems, 07 September 2017.

[15] Z. Peng, Q. Gong , Y. Duan and Y. Wang "The Research of the Parallel Computing Development from the Angle of Cloud Computing", in IOP Conf. Series: Journal of Physics: Conf. Series 910, 2017

[16] Md. F. Ali and R. Zaman Khan "Distributed Computing: AnOverview", in International Journal of Advanced Networking and Applications, Pages: 2630-2635, Volume: 07 Issue: 01, 2015

[17] Y. Sun, Z. Zhu and Z. Fan "Distributed Caching in Wireless Cellular Networks Incorporating Parallel Processing" , in IEEE Internet Computing, pages 52-61, Volume: 22, Issue: 1, Feb. 2018

[18] P. Srinivasa Rao, V.P.C Rao and A. Govardhan,"Dynamic Load Balancing With Central Monitoring of Distributed Job Processing System", in International Journal of Computer Applications, Volume 65– No.21, March 2013.

[19] A. Sharma and S. K. Peddoju, "Response Time Based Load Balancing in Cloud Computing", In 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT). July 2014, Kanyakumari, India

[20] H. Chen, F. Wang, N. Helian, and G. Akanmu, "User-Priority Guided Min-Min Scheduling Algorithm For Load Balancing in Cloud Computing", in 2013 National Conference on Parallel Computing Technologies (PARCOMPTECH), Bangalore, India, Feb 2013

[21] R. Kapur, "A Workload Balanced Approach for Resource Scheduling in Cloud Computing", in 2015 Eighth International Conference on Contemporary Computing (IC3), page 20-22, Noida, India, Aug 2015

[22] B. Mondal and A. Choudhury "Simulated Annealing (SA) based Load Balancing Strategy for Cloud Computing", in International Journal of Computer Science and Information Technologies, pages 3307-3312, Volume: 6, Kolkata-700053 India, 2015

[23] S.K.S. Kumar and P. Balasubramanie "Cloud Scheduling Using Mumbai Dabbawala", in International Journal of Computer Science and Mobile Computing, Volume: 4, Issue. 10, October 2015

[24] N. J. Kansal and I. Chana "Existing Load balancing techniques in cloud computing: A SYSTEMATIC REVIEW", in Journal of Information Systems and Communication, pages 87-91, Volume: 3, Issue 1, 2012.