

Investigation of Improving Load Balancing and Data Fragmentation in Cloud Computing Performance

Mahboobeh Najmi*

IT Network Engineering, Islamic Azad University of DR Hesabi, Tafresh, Iran; mahboobe.najmi@gmail.com

Abstract

Objective: In this study we have investigated the affecting parameters in cloud computing in information technology. To this end, security and load balancing in cloud computing is still considered to be one of the most important views owe to the sensitive and confidential information stored in the cloud by the users. **Methods:** This study has as a prime focus on the matters related to the security and load balancing aspects of CC. Furthermore, we have tried to review the researches about cloud computing, improving load balancing and data fragmentation in cloud computing and related matters in regard of its performance. **Findings:** As well, this research paper has shown all of aspect of investigations in recent years. In fact, this study has located novel ways for finding new paths in this field for future researches. **Application:** In the latest part of the paper we have presented some future themes as follows: Load balancing with novel load balancing algorithms for better efficiency, and also security issues considering some aspects such as data fragmentation and encrypted in CC are unknown for research.

Keywords: Cloud Computing, Data Fragmentation, Data, Load Balancing, Security Challenges

1. Introduction

During last years, Cloud Computing (CC) has become very popular in the industry in around the world. So CC has been defined as a computational model and an embedded model for managing and providing services over the internet. As well, it is a suitable model for demand and network access with a common space of configurable computing resources. For instance, networks and services that can be quickly deployed to customers with the least endeavor of manage or interact provide to customer. The evolution of the calculations is that after the need for elements such as electricity¹, water, gas and telephone and so on, it is supposed to be a critical element. In this case, users try to get the service, without concentration where the service is located, and based on needs. The world of computing continues towards that evolve of softwares so it is available to everyone instead of running on separate computers as a service. The computing systems were designed to allow users to allocate their systems for gen-

eral computing tasks at ideal times. Therefore, a large number of small and large computing resources have joined together in a grid as called voluntary computing and created a huge processing power. Also some of the old and novel computing systems include cluster computing, optical computing and recent cloud computing.

Cloud computing concept is referred to the 21st century, but its foundation dates backs to the early 1950s, when there were server rooms with gigantic ultra-powerful computers. These supercomputers shared between multiple users and through connections that are made by shared terminals, while most of the processing was carried out on the supercomputer itself².

The CC is like a repository of resources and services that is accessible through the internet. In fact, CC services are delivered through a data center around the world, and they cloud achieve users with a virtual platform for services and resource utilization³.

Generally, resources are located on the server instead of on the user's side, which reduces costs. The provided services

*Author for correspondence

through CC are according to the user's sample, and the user pays based on the type of resources using⁴, similar to other public services according to the level of human usage.

The widespread presence of large companies such as Microsoft, Google, Amazon, etc. in the competitive landscape of CC, is indicator of the rapid development and dominance of such computing in the world of information technology.

One of the important profits of CC provided for its developers is the ability to build cloud bank in places of the globe that are cost-effective in terms of cost and energy efficiency. So, this means by cause of security provision for natural disasters that may occur in a particular area, and assigns for the moderation of the traffic and the system load.

Cloud computing models are usually divided into two types, one according to place and another one based on the type of service.

The cloud computing types (based on service type) are as follows:

Three basic services that include:

- Platform as a Service (PaaS).
- Infrastructure as a Service (IaaS).
- Software as a Service (SaaS).

Also, the CC types (based on position) are as follows:

- Private
- Public
- Hybrid

As we mentioned in the Figure 1, we observed difference between public and private clouds based on their restrictions.

As well as, we could mention to four main advantages of clouds as follows:

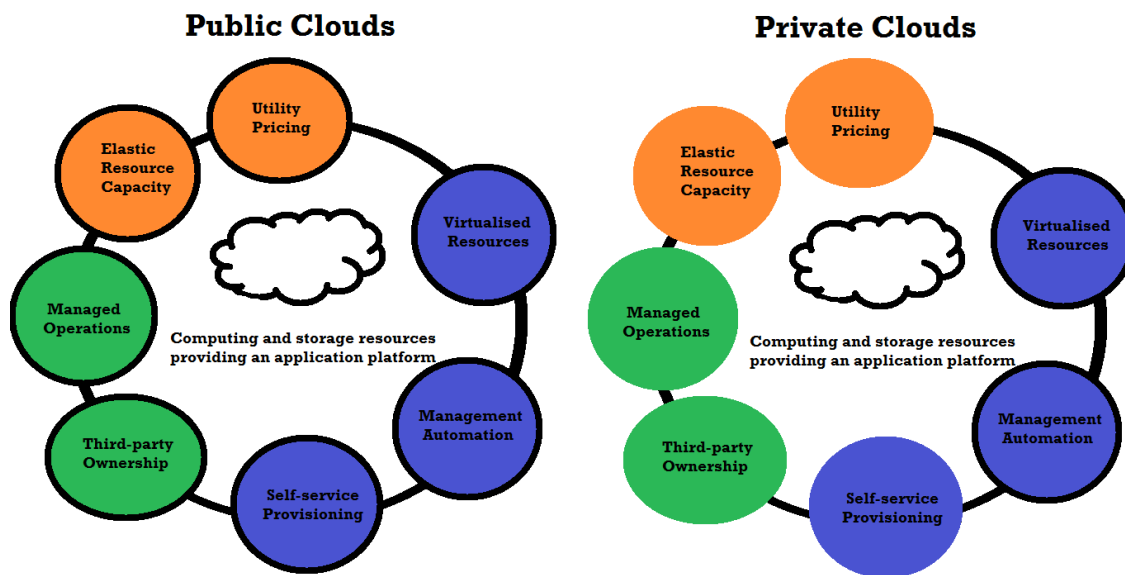


Figure 1. A comparison picture of Public clouds and Private clouds²⁷.

- Cost reduction.
- Storage space increment.
- Flexibility.
- Office management.

As well as, the most challenges of cloud using include:

- Unemployment time of apparatus.
- Maintenance of private environs⁵.

The study has been categorized as follows: review on load balancing, data security and data slipping on cloud com-

puting are presented in part two. The analysis of the paper's research has been presented in part three. The conclusion and future work has been presented in part four.

2. Review on Security, Load Balancing and Data Fragmentation on Cloud Computing Scope

In this section, we investigate previous reports on the kinds of security, load balancing and data slipping aspects.

2.1 Security

The cloud is a next generation model that makes dynamic resource pools, virtualization, and high accessibility. Today, we have the ability to utilize scalable, distributed computing spaces within the confines of the Internet, a practice known as cloud computing⁶. So, CC is becoming a well-known buzzword nowadays. Many companies, such as Google, Facebook and etc. accelerate their paces in developing CC systems and enhancing their services to provide for a larger amount of users. Although, security and privacy issues present a strong barrier for users to adapt into CC. The authors in⁷ find those concerns are not adequate and more should be added in terms of five aspects (i.e., availability, confidentiality, data integrity, control) for security. Moreover, the authors claim that the prosperity in Cloud Computing literature is to be coming after those security and privacy issues having been resolved⁷.

There is an important need to securely store, manage, share and investigate massive amounts of complex (e.g., semi-structured and unstructured) data to assign patterns and trends for amend to the quality of healthcare, better safeguard the nation and research alternative energy. Because of the important nature of the applications, it is important that clouds be secure. The most security problem with clouds is that the owner of the data may not have control of where the data is allocated. This means that if one wants to operate the profits of using CC, one must also use the resource placement and scheduling provided by clouds. Therefore, we require guarding the data in the midst of untrusted processes⁸.

The CC, a rapidly development information technology, has aroused the concern of around the world. The CC is Internet-based computing, whereby shared resources, software and information, are achieved to computers and devices on-demand, such as the electricity grid. Also, the CC is the product of the fusion of traditional computing technology and network technology like grid computing, distributed computing parallel computing and etc. It purposes to build a complete system with powerful computing capability through a large number of relatively low-cost computing entity, and using the advanced business models like SaaS (Software as a Service)⁹. With the development of parallel computing, distributed computing, grid computing, a new computing model appeared. The notion of computing comes from the grid, public computing and SaaS.

It is a novel approach that shares basic framework. The basic rules of the CC are to build the computing be

assigned in a great number of distributed computers, rather than local computer or remoter server. The running of the enterprise's data center is just like Internet. This builds the enterprise applicate the resource in the application that is needed, and access computer and storage system based on the requirement. The authors in the study¹⁰ recognize the background and rule of the CC, the character, style and actuality. This research paper also introduces the application field the merit of the CC. It achieves secure and dependable data storage center, therefore user need not do the awful things such storing data and killing virus, this kind of task can be done by professionals. It can realize data share through different equipment. It analyses some questions and hidden troubles, and puts forward some solutions, and discusses the future of the CC. The CC is a computing style that provides power referenced with IT as a service. Users can enjoy the service even he knows nothing about the technology of cloud computing and the enhancement for data security in CC environment professional knowledge in this field and the power to control it¹⁰.

2.2 Load Balancing

One important matter associated with this scope is dynamic load balancing or task scheduling. The load balancing technique used to make sure that none of the node is in idle state while other nodes are being utilized.

Load balancing algorithms were reviewed heavily in kinds of environments. The CC is the main concerns involve efficiently assigning tasks to the cloud nodes such that the effort and request processing is done as efficiently as possible, while being able to tolerate the various affecting constraints such as heterogeneity and high communication delays¹¹⁻¹⁴.

One of the technologies used in the CC is virtualization; virtualization is a runtime unit that acts as the basis for cloud computing technology. Along with the benefits of virtualization, it requires load management by load balancing algorithms to maintain stability and efficiency in the CC. Therefore, an algorithm that can improve virtual machines is essential. In fact, the units of task and virtual machines are in measured.

Load balancing algorithms:

The load balancing algorithms are divided into two static and dynamic general categories.

Static load balancing algorithm: These types of algorithms cannot adapt themselves to changes in the environment. The principle goal of these algorithms is to reduce overall execution time.

Dynamic load balancing algorithm:

These kinds of algorithms have higher accuracy and can produce better load balancing results. An important advantage of dynamic load balancing algorithms is decision-making based on the current state of the system, which improves system performance.

Also, in the study¹⁵ the authors arouse study is to understand the current challenges in CC, primarily in cloud load balancing using static algorithms and finding gaps to bridge for more efficient static cloud load balancing in the future.

In the following, we mentioned related algorithms:

- Round Robin Algorithm (RAA):

In this algorithm, requests are assigned to a list of virtual machines rotated. Initially, the first request is assigned to a randomly selected virtual machine, and then the next request is assigned to another virtual machine with a rotational arrangement.

- Weighted Round Robin Algorithm:

The revised version of the rotational turning algorithm is considered by each machine based on its characteristics. If a virtual machine has the power of processing two requests, it takes two weights, so each time the virtual machine arrives it allocates two requests. The assignment function is like rotational algorithm in this algorithm.

- Dynamic Round Robin Algorithm:

This algorithm randomly assigns one selected node to one accessible virtual machine. This algorithm is very simple and cannot recognize the high and low load of the virtual machine^{16,17}.

2.3 Data Fragmentation

This part states the related works in data fragmentation in cloud computing.

Data Fragmentation in Cloud for Optimal function and Security that collectively methods the security and function matters.

For respond to the question “How can we ensure that the use of unreliable cloud computing resources, it will ensure the confidentiality of information?” We describe this question and related conception in CC in the following paragraphs:

In¹⁸ have proposed a fragmentation plot, which is a column-based partition in server side; there are only encrypted partitions.

Each piece uses a unique identifier to support the query. Questions are processed twice. Initially, the whole part is transmitted to the client based on item identifier. When the client acquires that piece, we will decrypt it over the fragmentary piece for a re-query request. Finally, the result is returned after the second query is executed. While this method can be appropriate for a small data set, a significant amount of overload is needed to collect the whole database. In this approach, there is one case in which the entire database should be transmitted to the client side, which poses questions about the benefits of using CC. In addition, query processing reduces system performance twice. Hence, this approach is not the right solution for this problem; it eliminates the main advantage of cloud computing in terms of storage delivery. It also requires a lot of amount of overload time to run the query. Other studies have taken this approach¹⁹⁻²¹. They are mainly focus in the query optimization method as means for the attempt to address the constraint¹⁸, which is related to performance.

Fragmentation method:

The fragmentation technique consists of two aspects:

A primary cloud and some general cloud. In an initial configuration, the entire database is encrypted using a highly secure encryption algorithm and stored in the main cloud. While not disclosing the encryption key to the main cloud provider.

The main cloud is the core section of our plan so that the whole connection is maintained in one situation. Furthermore, the encrypted algorithm operated on the original cloud which is a block that should be enough to protect the secure data²²⁻²⁴.

3. Analysis

In this regard, we suppose set m include from set of virtual machines $V_m = \{V_{m1}, V_{m2}, \dots, V_{mn}\}$. Which n process task there are in $T = \{T_1, T_2, \dots, T_n\}$. All running

virtual machines are parallel and unrelated, and each virtual machine runs on its source. There is no sharing of resources from other virtual machines. We schedule non-exclusive affiliates for these virtual machines. Each “ n ” assigned task to “ m ” which is presented to the virtual machine as the LP model from the formulas (1 and 2). In these formulas, j is defined the virtual machine and i to the task²⁵.

Process time (PT_{ij}) is assigned to j and i .

If we assign “ i ” function, otherwise:

$$x_{ij} \begin{cases} 1 \\ 0 \end{cases} x_{ij} = 0 \text{ or } 1 \quad (1)$$

Then the Linear Programming (LP) model (by using LP can achieve the best result (for example, the highest profit or the lowest cost) in certain circumstances and with specific constraints) is presented as follows:

$$\left\{ \begin{array}{l} \text{Minimize } Z = \sum_{i=1}^n \sum_{j=1}^m PT_{ij} x_{ij} \\ \text{Subject to: } \sum_{i=1}^n x_{ij} = 1, j = 1, 2, \dots, m \\ x_{ij} = 0 \text{ or } 1. \end{array} \right. \quad (2)$$

Resource using rate: The maximum amount of resource using isother important subjects which is derived from Formulas (3) and (4).

$$\text{Average using rate} = \frac{\sum_{j \in VM_s} CT_j}{\text{Makespan} \times \text{Number of } VM_s} \quad (3)$$

Makespan is minimum of work completion time. Whereas makespan can state as follows:

$$\text{Makespan} = \max \left\{ \frac{CT_j}{j \in VM_s} \right\} \quad (4)$$

The virtual machine capacity is as follows:

$$C_{VM} = Pe_{num} \times Pe_{mips} \quad (5)$$

Which C_{VM} is virtual machine capacity, Pe_{num} is processing components number in virtual machine and Pe_{mips} is millions instructions.

All virtual machine capacity is equal to:

$$C = \sum_{i=1}^m C_{VM_j} \quad (6)$$

Which C is total all virtual machine capacity and assigned capacity to application program/environment.

As well as, the task length is that:

$$TL = T_{mips} \times T_{pe} \quad (7)$$

And also, work length is that:

$$JL = \sum_{k=1}^p TL_i \quad (8)$$

In the above formula, p is tasks number which is related for work.

Task load relation: The task load relation in Formulas (9) and (10) is calculated. By introducing and assigning task to virtual machines, it defines as follows:

$$TLR_{ij} = \frac{TL_i}{C_{VM_j}}, i = 1, 2, \dots, n \text{ task}, j = 1, 2, \dots, m \text{ virtual machines} \quad (9)$$

Which TL_i is task length that it estimates in starting of every action.

$$\text{If } TLT_{ji} \begin{cases} \text{assigning task to virtual machine is performed} \\ \text{Otherwise it is not assign} \end{cases} \quad (10)$$

Often, two schedules used in a non-exclusive system are Round Robin (RR) and Weighted Round Robin (WRR) Algorithms. The proposed algorithm is Improved Weighted Round Robin (IWRR).

It is mentioned that the IWRR algorithm is the most optimal algorithm and it allocates the jobs to the most suitable VMs based on the VM's information like its processing capacity, load on the VMs, and length of the arrived tasks with its priority. The proposed improved weighted round robin algorithm is the most optimal algorithm for assigning tasks to the most suitable virtual machines based on virtual machine information such as processing capacity, virtual machine load, and the length of tasks entered with priority. For example, in the study²⁶ the proposed WRR service broker policy is a modified version of the service proximity policy and considers the processing capacities of DCs in terms of two important parameters, namely number of processors and processor speed. In the proposed WRR policy, an initialization algorithm maintains a weighted list of all DCs in a single region.

Load balancing in supposed algorithm: The load balancing by collecting of the suspended execution time of each created virtual machine is initialized, and then it arranges in the ascending order, and then identifying the

number of tasks in each virtual machine, and finally organizes it in an incremental order.

After mathematic modeling and some definition in this regard, we describe about the analyzing of the result. The function of IWRR algorithm based on simulation result in cloud analyzed. Response time is analyzed in Round Robin (RR), Weighted Round Robin (WRR) Algorithms and Improved Weighted Round Robin (IWRR) under the combination of homogeneous and heterogeneous work lengths with heterogeneous source status.

Homogeneous tasks in non-homogeneous sources (virtual machines):

At first, we investigate total runtime comparison in three above algorithms. It is mentioned that, the data center is divided into two homogeneous and heterogeneous groups, which are equal in terms of processor power homogeneity, while in heterogeneous mode, the processors are not equal in terms of execution power. The analysis of the following cases is the highest and lowest performance algorithms of presented homogeneous work in heterogeneous environment.

Figures 2 and 3 provides an IWRR with a working length of a faster completion time, then demonstrate load balancing algorithms (WRR and RR) in non-homogeneous sources (virtual machines) and homogeneous work? IWRR scheduling algorithms considered length of work along with the processing capacity of non-homogeneous virtual machines for assigning a time. Thus, the huge number of tasks assign to virtual machines with high-capacity in homogeneous work on non-environment homogeneous. This helps to complete the task in a short time.

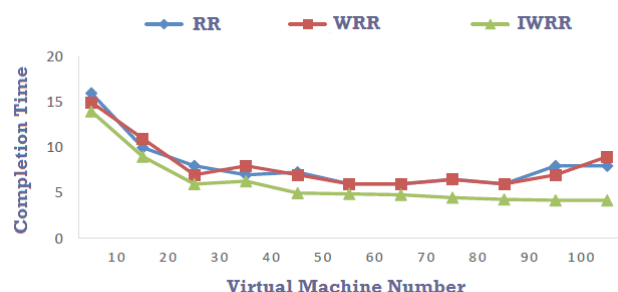


Figure 2. Completion execution time (in common space).

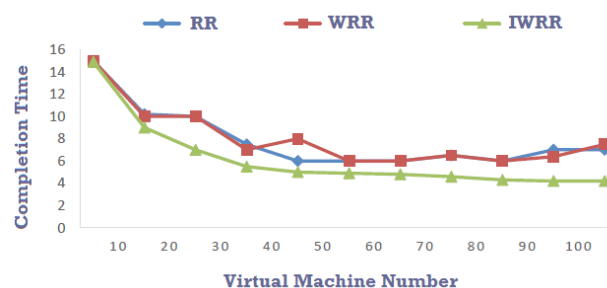


Figure 3. Completion execution time (in common time).

The load balancing in the IWRR runs at end of each tasks completion, if the load balances of each virtual machine were observed the dedicated tasks which is completed, so the virtual machine determines heavy loads from group of virtual machines and the it calculates probability completion time of current works in high-load and low-load or non-load virtual machines is more likely to work in, with little or no load. If low-capacity virtual machine can all the current work finished in as short time as possible in high-load virtual machine, so that work moves toward the low-load virtual machine.

The WRR algorithm considers the virtual machine's capacity ratio to the total capacity of the virtual machine, and allocates the corresponding number of works entered into the virtual machine. Thus, it runs on the next level.

But if all the long tasks are assigned to the low capacity of the virtual machine based on the above calculations, then it will delay the execution completion time.

RR algorithm did not take into account any varies in the space, with the capacity of virtual machines and the work length. It simply put the things in list one by one alternately on a regular basis. So its execution time is higher than the other two algorithms.

It is mentioned that, load balancing and data fragmentation are related to each other for increasing the efficiency of the system. And we can prolong the following result on data fragmentation.

4. Conclusion and Future Work

Cloud computing is relatively a new subject and it must have to overcome the load balancing and data fragmentation issues in order to be more and more tangible technology of the future. We suggest future themes in this regard as follows:

- Load balancing with novel load balancing algorithms for better efficiency.
- Security issues concerned with data fragmentation and encryption in CC are unknown for research.

5. References

1. Moghimi SM, Sarlak G. Energy management optimizing in multi carrier energy systems considering net zero emission and CHP temperature effects. *Physical Science International Journal*. 2016; 10(3):1-2. <https://doi.org/10.9734/PSIJ/2016/24779>
2. Feng DG, Zhang M, Zhang Y, Xu Z. Study on cloud computing security. *Journal of Software*. 2011; 22(1):71-83. <https://doi.org/10.3724/SP.J.1001.2011.03958>
3. Nurmi D, Wolski R, Grzegorzczak C, Obertelli G, Soman S, Youseff L, Zagorodnov D. The eucalyptus open-source cloud-computing system. *IEEE/ACM International Symposium on Cluster Computing and the Grid*; 2009. p. 124-31. <https://doi.org/10.1109/CCGRID.2009.93>
4. Shaikh FB, Haider S. Security threats in cloud computing. *International Conference for Internet Technology and Secured Transactions*; 2011. p. 214-9.
5. Surcel T, Alecu F. Applications of cloud computing. In *International Conference of Science and Technology in the Context of the Sustainable Development*; 2008. p. 177-80.
6. Somani U, Lakhani K, Mundra M. Implementing digital signature with RSA encryption algorithm to enhance the data security of cloud in cloud computing. *International Conference on Parallel, Distributed and Grid Computing (PDGC - 2010)*; 2010. p. 211-6. <https://doi.org/10.1109/PDGC.2010.5679895>
7. Zhou M, Zhang R, Xie W, Qian W, Zhou A. Security and privacy in cloud computing: a survey. *International Conference on Semantics, Knowledge and Grids*; 2010. p. 105-12. <https://doi.org/10.1109/SKG.2010.19>
8. Hamlen K, Kantarcioglu M, Khan L, Thuraisingham B. Security issues for cloud computing. *International Journal of Information Security and Privacy*. 2010; 4(2):39-51. <https://doi.org/10.4018/jisp.2010040103>
9. Yang J, Chen Z. Cloud computing research and security issues. *International Conference on Computational Intelligence and Software Engineering*; 2010. p. 1-3.
10. Zhang S, Zhang S. Cloud computing research and development trend. *International Conference on Future Networks*; 2010. p. 93-7.
11. Aceto G, Botta A, Donato WD, Pescapè A. Cloud monitoring: A survey. *Computer Networks*. 2013; 57(9):2093-115. <https://doi.org/10.1016/j.comnet.2013.04.001>
12. Deng Y, Lau RYH. On delay adjustment for dynamic load balancing in distributed virtual environments. *IEEE Transactions on Visualization and Computer Graphics*. 2012; 18(4):529-37. <https://doi.org/10.1109/TVCG.2012.52> PMID:22402679
13. Alam M, Khan ZA. Issues and challenges of load balancing algorithm in cloud computing environment. *Indian Journal of Science and Technology*. 2017; 10(25):1-12. <https://doi.org/10.17485/ijst/2017/v10i25/105688>
14. Alam M, Shahid M. A load balancing strategy with migration cost for independent Batch of Tasks (BoT) on heterogeneous multiprocessor interconnection networks. *International Journal of Applied Evolutionary Computation (IJAEC)*. 2017; 8(3):74-92. <https://doi.org/10.4018/IJAEC.2017070104>
15. Shah N, Farik M. Static load balancing algorithms in cloud computing: challenges and solutions. *International Journal of Scientific and Technology Research*. 2015; 4(10):366-7.
16. Abdulraza A, Aliyu S, Mustapha AM, Abdullahi SE. An Additional Improvement in Round Robin (AAIRR) CPU scheduling algorithm. *International Journal of Advanced Research in Computer Science and Software Engineering*. 2014; 4(2):601-10.
17. Sangwan P, Sharma M, Kumar A. Improved round robin scheduling in cloud computing. *Advances in Computational Sciences and Technology*. 2017; 10(4):639-42.
18. Executing SQL over Encrypted Data in the Database-Service-Provider Model. Available from: <https://www.ics.uci.edu/~chenli/pub/sigmod02.pdf>
19. Hore B, Mehrotra S, Canim M, Kantarcioglu M. Secure multidimensional range queries over outsourced data. *The Very Large Data Bases (VLDB) Journal*. 2018; 21(3):333-58. <https://doi.org/10.1007/s00778-011-0245-7>
20. Hore B, Mehrotra S, Tsudik G. A privacy-preserving index for range queries. *VLDB'04 Proceedings of the 13th International Conference on Very large data bases*. 2004; 30:720-31. <https://doi.org/10.1016/B978-012088469-8.50064-4>
21. Hacg H. Query optimization in encrypted database. *International Conference on Database Systems for Advanced Applications*; 2005. p. 43-55.
22. Mohamed EM. Enhanced data security model for cloud computing. *International Conference on Informatics and Systems (INFOS)*; 2012. p. 12-7. PMID:29403715
23. Arasu A, Blanas S, Eguro K, Joglekar M, Kaushik R, Kossmann D, Ramamurthy R, Upadhyaya P, Venkatesan R. Secure data bases-a-service with Cipherbase. *Proceedings of the International conference on Management of data - SIGMOD*; 2013. p. 1-1033.

24. Nalinipriya G, Aswin Kumar R. Extensive medical data storage with prominent symmetric algorithms on cloud- A protected framework. International Conference on Smart Structures and Systems - Icsss'13; 2013. p. 171–7.
25. Chitra Devi D, Uthariaraj VR. Load balancing in cloud computing environment using improved weighted round robin algorithm for nonpreemptive dependent tasks. The Scientific World Journal. 2015; 1–14.
26. Radi M. Weighted round robin policy for service brokers in a cloud environment. The International Arab Conference on Information Technology (ACIT2014); 2014. p. 46–8.
27. Ghanbari H, Simmons B, Litoiu M, Iszlai G. Feedback-based optimization of a private cloud. Future Generation Computer Systems. 2012; 28(1):104–11. <https://doi.org/10.1016/j.future.2011.05.019>