AB-INITIO (INITIAL) RESOURCE ASSIGNMENT IN CLOUD COMPUTING: AN EXPERIMENTAL SURVEY

[1]Swapnil M Parikh, [2]Dr. Narendra M Patel
[1]PhD Scholar and Assistant Professor, [2]Associate Professor
[1]Computer Engineering,
[1]C U Shah University and Babaria Institute of Technology, Wadhwan and Vadodara, India

---

Abstract : Now a day, Cloud Computing is not new to researchers. It is basically epoch of remote computing where one can access their personal and public resources easily from any computer through Internet. Cloud Computing is a pay-per-use consumer-provider service model. Resource Management is always a foremost problem in any computing paradigm. Due to the availability of finite resources it is very challenging for cloud providers to provide all the requested resources and that is again in time. These cloud resources must be allocated in a fair and efficient manner. Ab-initio resource assignment is a basically an initial resource assignment process, in a manner that resources are requested by application (on behalf of cloud consumers) for the first time. In this research paper, various existing algorithms on Ab-initio resource assignment are implemented and compared for further analysis and research work.

IndexTerms **- Cloud Computing; Cloud Resources; Resource Management; Resource Assignment; Scheduling.**

I. INTRODUCTION

Because of the advancement in Information and Communication Technology (ICT) over the past few years, Computing has been considered as a utility like water, electricity, gas and telephony. These utilities are available to the consumers based on their requirement at any time. Consumers pay for these services to the service providers based on their usage [1][3][7].

Like all the other existing utilities, Computing utility is the basic computing service that meets the day to day needs of the general community. To deliver this vision, a number of computing paradigms have been proposed, of which the latest one is known as Cloud Computing.

Cloud is nothing but large pool of easily accessible and usable virtual resources. Cloud computing is a service provision model which provides various kinds of agile and effective services to the consumers where everything is considered as a service [1][3][7][10].

Resource management is always a foremost problem in many computing domains. In cloud computing various cloud consumers demand variety of services as per their dynamically changing needs. So it is the job of cloud computing providers to avail all the demanded services to the cloud consumers. But due to the availability of finite resources it is very difficult for cloud providers to provide all the demanded services in time. From the cloud providers perspective cloud resources must be allocated in a fair manner. So, it is a vital issue to meet cloud consumers QoS requirements and satisfaction [11].

Conventional resource management techniques are not sufficient for cloud computing because they are based on virtualization technology with distributed nature. Cloud computing introduces new challenges for resource management due to heterogeneity in hardware capabilities, on-demand service model, pay per use model and guarantee to meet QoS [2][5][9][14][15].

Ab-initio resource assignment is a basically an initial resource assignment process, in a manner that resources are requested by application (on behalf of cloud consumers) first time. In this research paper, various existing algorithms on Ab-initio resource assignment are implemented and compared for further analysis.

The rest of the paper is organized as follows: Section II discusses fundamentals of cloud computing. Section III introduces initial resource assignment algorithms. Section IV gives comparisons among all algorithms discussed in section III. Section V gives concluding remark and future work.

II. FUNDAMENTALS OF CLOUD COMPUTING

Cloud is like a big black box, nothing inside the cloud is visible to the cloud consumers. Cloud delivers computing as a utility as it is available to the cloud consumers on demand. Cloud Computing is a simple pay-per-use consumer-provider service model [1][3][7][10].

Below are the widely quoted definitions of Cloud Computing:

NIST [8]: Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Ian Foster [4]: A large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet.

Rajkumar Buyya [3]: A Cloud is a type of parallel and distributed system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resource(s) based on service-level agreements established through negotiation between the service provider and consumers.

Cloud computing is composed of three kinds of service models. These service models are based on the level and depth of the services provided by cloud computing [4][7][8][13][17][18].

Cloud Software as a Service (SaaS): In this service model, instead of using locally run applications the cloud consumer uses the cloud providers software services running on a cloud infrastructure. It is the job of cloud provider to maintain and manage the software services that are used by the cloud consumer. The cloud provider may charge according to software quantity and time usage. Salesforge.com and Customer Relationship Management (CRM) are the examples of such service model [8][11][12][16][17][18].

Cloud Platform as a Service (PaaS): In this service model, the cloud platform offers an environment on which developers create and deploy applications. It provides platform where applications and services can run. The consumers do not need to take care of underlying cloud infrastructure including network, servers, operating system or storage but has a control over deployed application. Google Application Engine, Microsoft Azure and RightScale are the example of such model [8][11][12][17][18].

Cloud Infrastructure as a Service (IaaS): In this service model, cloud providers manage large set of computing resources such as storing and processing capability. Cloud consumer can control operating system; storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls). Sometimes it is also called as a Hardware as a Service (HaaS). The cost of the Hardware can be greatly reduced here. Amazon Web Services, Open Stack, Eucalyptus, GoGrid and Flexiscale offers IaaS [8][11][12][17][18].

In cloud computing various deployment models have been adopted based on their variation in physical location and distribution. Regardless of the services, clouds can be classified among four models as mentioned below:

Private Cloud: It is private to the organization. All the cloud services are managed by the organization people themselves or any third party vendors. In private cloud services are not provided to the general public. Private cloud may exist on premise or off premise [6][8][17][18].

Public or Hosted Cloud: All the cloud services managed by the organization are made available as in pay as you go manner to the general public. The business people can adopt such cloud to save their hardware and/or software cost. Public cloud may raise number of issues like data security, data management, performance, level of control etc [8][17][18].

Community Cloud: Here cloud is available to specific group of people or community. All the cloud services are shared by all these community people. Community cloud may exist on premise or off premise [8][17][18].

Hybrid Cloud: It is a combination of two or more cloud models mentioned above [8][17][18].

III. INTRODUCTION TO AB-INITIO RESOURCE ASSIGNMENT ALGORITHMS

First Come First Serve: First Come First Serve (FCFS) is one of the simple and efficient resource management algorithms under Ab-initio resource assignment. It is also known as First-in First-out or First Come First Choice. Here cloud resources are assigned to various incoming cloudlets (tasks) in order in which they arrive. A queue is maintained for all incoming tasks. It is a non-preemptive resource scheduling algorithm. Figure 1 and Figure 2 shows FCFS algorithm for 10 cloudlets and 100 cloudlets respectively.

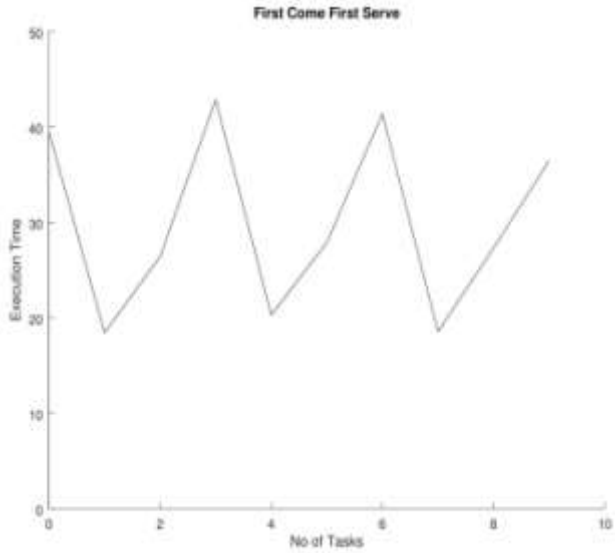

Figure 1 - FCFS (No of tasks: 10)

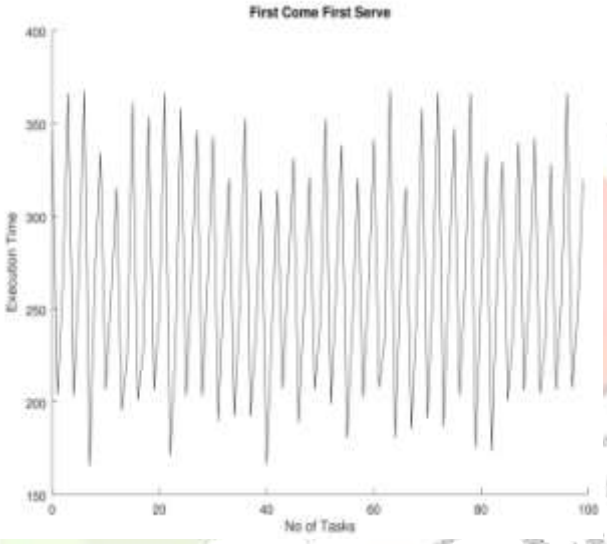

Figure 2 - FCFS (No of tasks: 100)

Shortest Job First: Shortest Job First (SJF) is a non-preemptive resource scheduling algorithm under Ab-initio resource assignment. Here cloud resources are assigned to various incoming cloudlets (tasks) that are having smallest CPU burst value. This algorithm requires knowledge of how long a task will run which is very difficult to measure in real time. Figure 3 and Figure 4 shows SJF algorithm for 10 cloudlets and 100 cloudlets respectively.

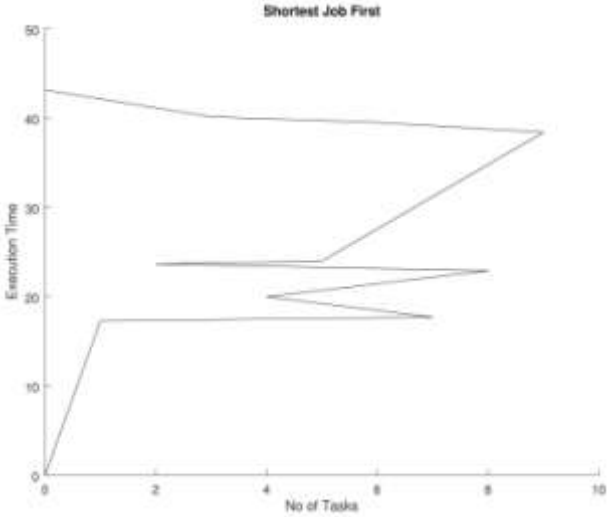

Figure 3 - SJF (No of tasks: 10)

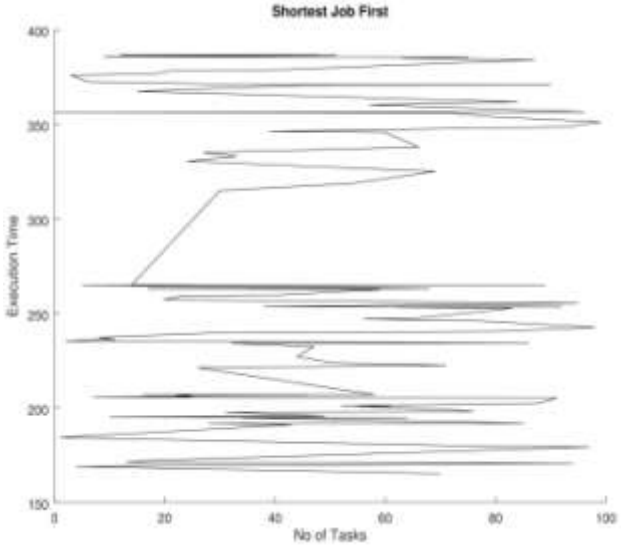

Figure 4 - SJF (No of tasks: 100)

Round Robin: Round Robin (RR) is a non-preemtive resource scheduling algorithm under Ab-initio resource assignment. Here cloud resources are assigned to various cloudlets (tasks) based on time intervals. This algorithm does not focus on length of the tasks and priority. Figure 5 and Figure 6 shows RR algorithm for 10 cloudlets and 100 cloudlets respectively.

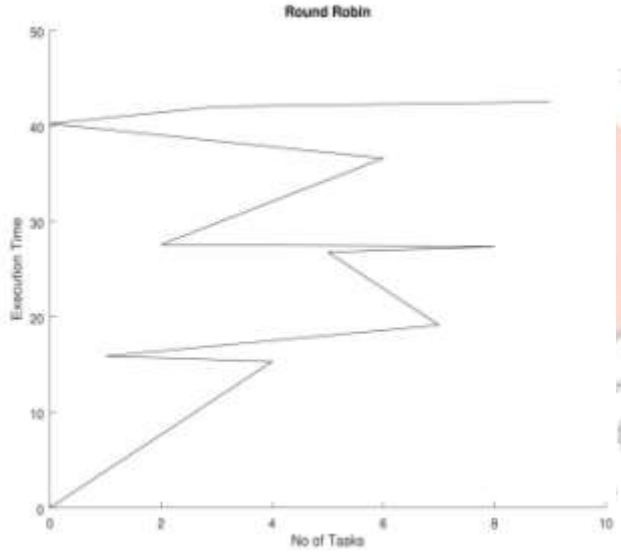

Figure 5 - RR (No of tasks: 10)

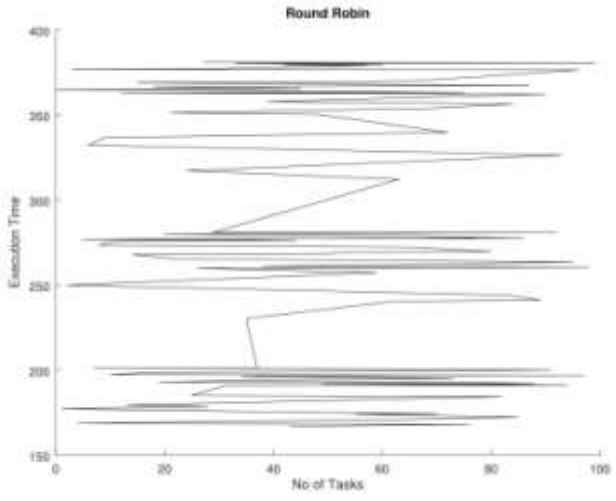

Figure 6 - RR (No of tasks: 100)

Priority Based: Priority based scheduling involves assigning priority to each and every tasks. Here cloud resources are assigned to incoming cloudlets (tasks) based on priority assigned. Tasks with higher priorities are executed first. Figure 7 and Figure 8 shows Priority based algorithm for 10 cloudlets and 100 cloudlets respectively.

Minmin: Minmin scheduling algorithm finds the execution time for all cloudlets. Then it chooses cloudlet with smallest execution time first and assigns cloud resources to it. Figure 9 and Figure 10 shows Minmin algorithm for 10 cloudlets and 100 cloudlets respectively.

Maxmin: Maxmin scheduling algorithm finds the execution time for all cloudlets. Then it chooses cloudlet with largest execution time first and assigns cloud resources to it. Figure 11 and Figure 12 shows Maxmin algorithm for 10 cloudlets and 100 cloudlets respectively.

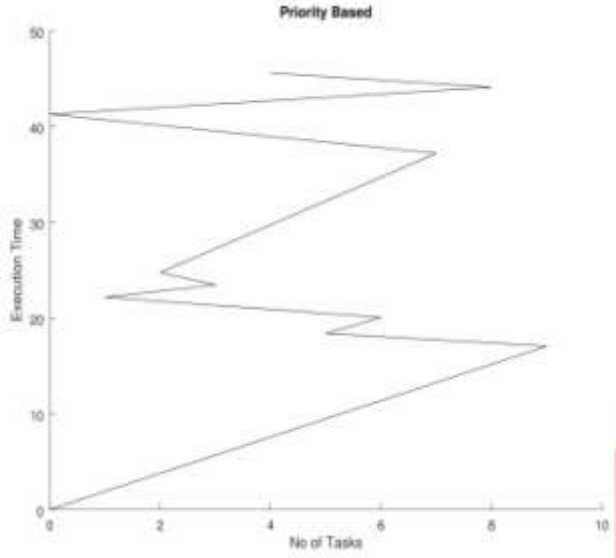

Figure 7 – Priority Based (No of tasks: 10)

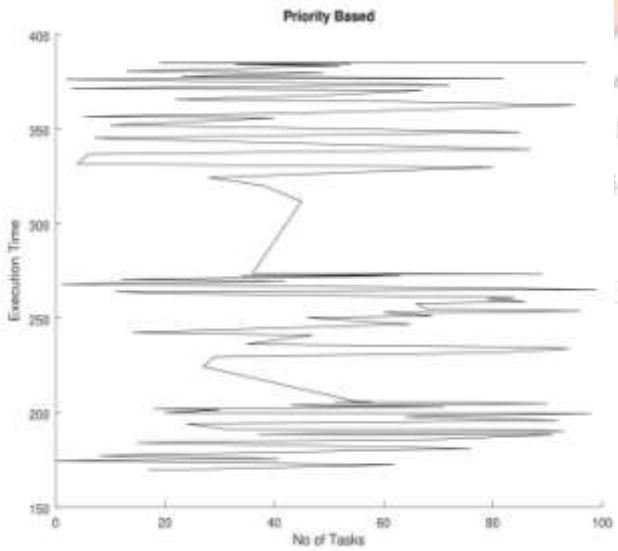

Figure 8 – Priority Based (No of tasks: 100)

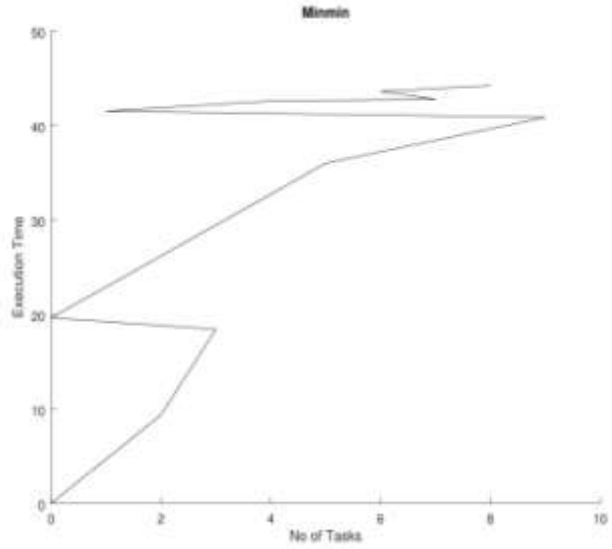

Figure 9 - Minmin (No of tasks: 10)

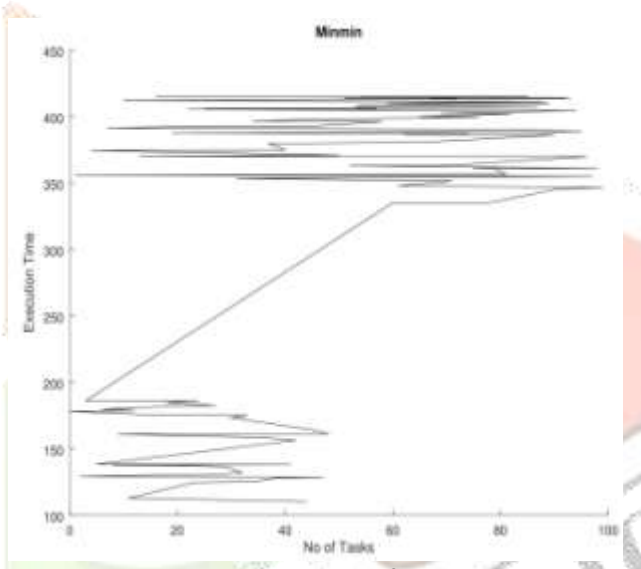

Figure 10 - Minmin (No of tasks: 100)

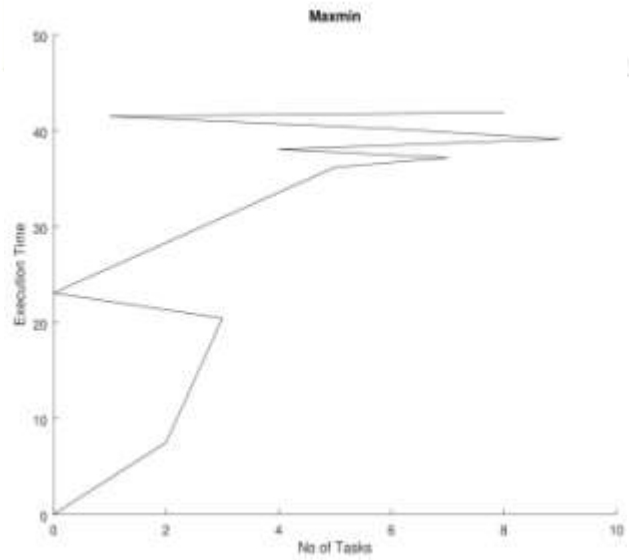

Figure 11 - Maxmin (No of tasks: 10)

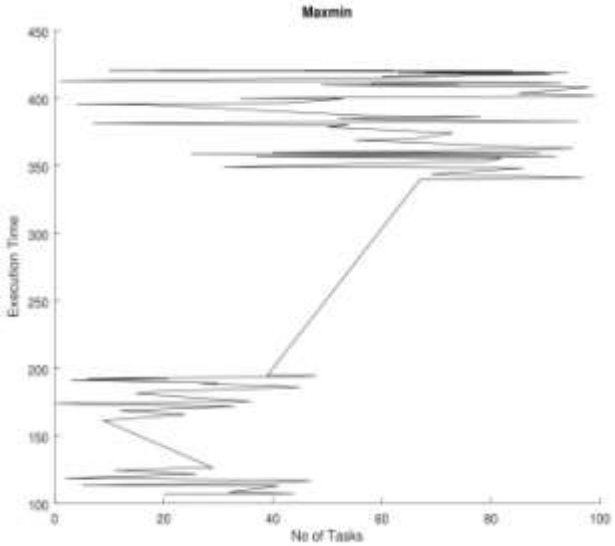

Figure 12 - Maxmin (No of tasks: 100)

IV. COMPARISONS AMONG AB-INITIO RESOURCE ASSIGNMENT ALGORITHMS

Figure 13 and Figure 14 shows Ab-initio resource assignment algorithms for 10 cloudlets and 100 cloudlets respectively.

Figure 15, Figure 16 and Figure 17 shows comparisons among Ab-initio resource assignment algorithms for 10 cloudlets, 50 cloudlets and 100 cloudlets respectively. The execution time in vertical line shows the total completion time (makespan) of all cloudlets. The horizontal line shows the various Ab-initio resource assignment algorithms.

Figure 15 shows that algorithms SJF and Minmin have the smallest makespan and algorithm FCFS has the largest makespan.

Figure 16 and Figure 17 shows that if we increase the no of cloudlets then algorithms Minmin and Maxmin gives unsatisfactory results. The makespan of both algorithms is much larger than other algorithms as it requires computing total execution time of all cloudlets at start of the algorithm. All other algorithms behave normally.

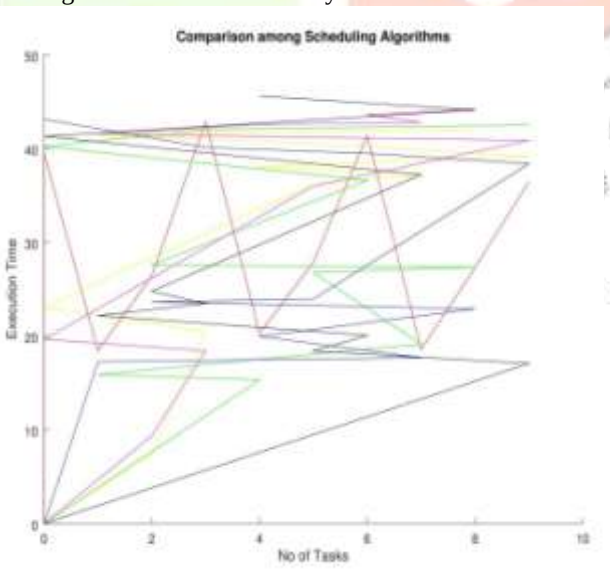

Figure 13 – Ab-initio resource assignment algorithms (No of tasks: 10)

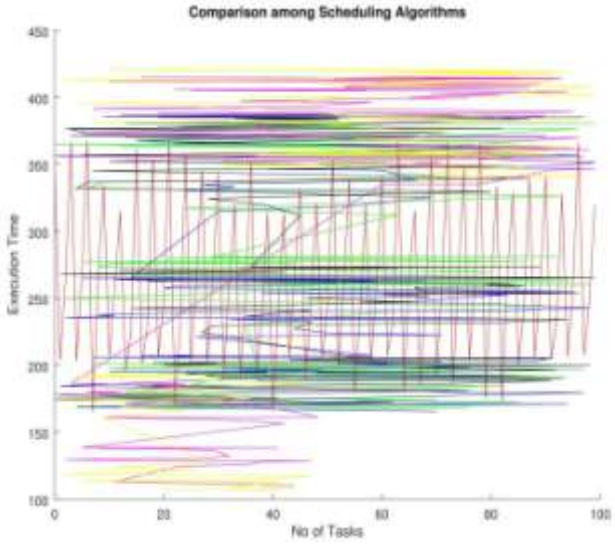

Figure 14 – Ab-initio resource assignment algorithms (No of tasks: 100)

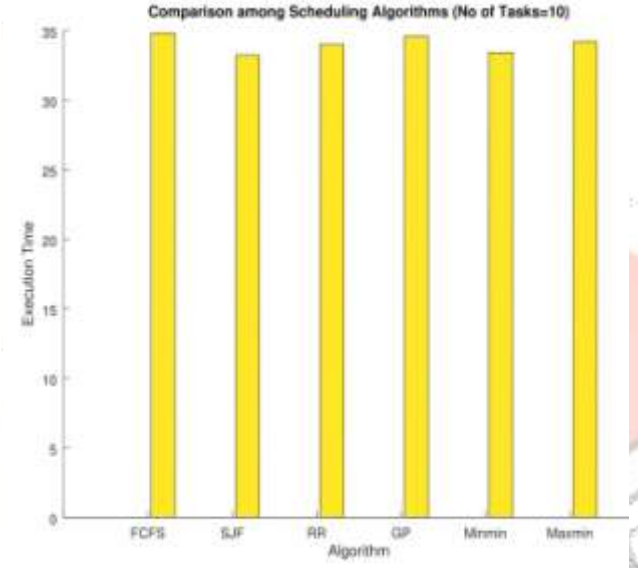

Figure 15 – Comparisons among Ab-initio resource assignment algorithms (No of tasks: 10)

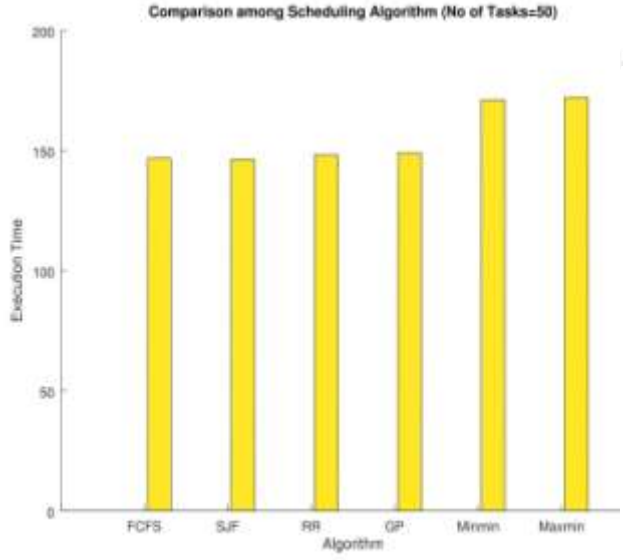

Figure 16 – Comparisons among Ab-initio resource assignment algorithms (No of tasks: 50)

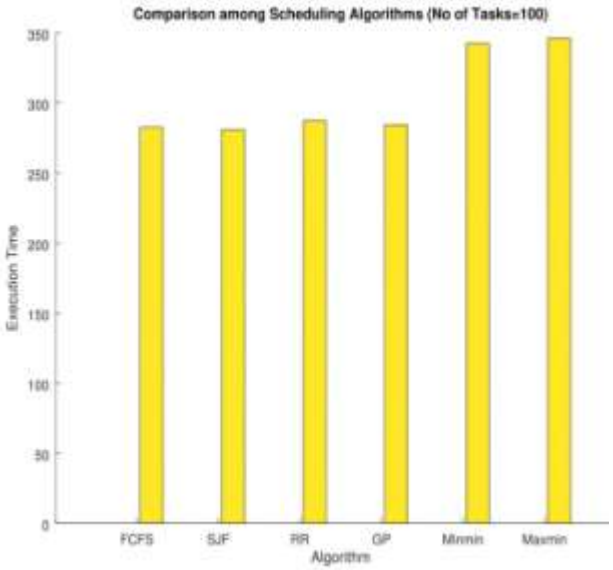

Figure 17 – Comparisons among Ab-initio resource assignment algorithms (No of tasks: 100)

V. CONCLUDING REMARK AND FUTURE WORK

All initial resource assignment algorithms are simple and single objective. The major drawback is that the current load on various machines has not been considered. The main idea is to propose new load balancing algorithm which is dynamic in nature with also includes other various criteria.

REFERENCES

[1] Armbrust M, Fox A, Griffith R, Joseph A, Katz R, Konwinski A, Lee G, Patterson D, Rabkin A, Stoica I, Zaharia M, 2010. A View of Cloud Computing (Draft). Communications of the ACM, Vol 53, No 4.
[2] Ben O, Ben M, Schuster A, Tsafrir D, 2012. The Resource-as-a-Service (RaaS) Cloud. 4th USENIX workshop on Cloud Computing.
[3] Buyya R, Yeo C, Venugopala S, Broberg J, Brandic I, 2009. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. ELSEVIER - Future Generation Computer Systems 25 pg 599-616.
[4] Foster I, Zhao Y, Raicu I, Lu S, 2008. Cloud Computing and Grid Computing 360-Degree Compared. Proceedings of Grid Computing Environments Workshop, pages 1–10.
[5] Gonçalves G, Endo P, Lamb T, Palhares A, Sadok D, Kelner J, Melander B, Mångs J, 2009. Resource Allocation in Clouds: Concepts, Tools and Research Challenges. Brazilian Symposium on Computer Networks and Distributed Systems.
[6] Grossman R, 2009. The Case for Cloud Computing. IEEE Computer Society, IT Pro.
[7] Manvi S, Shyam G K, 2014. Resource Management for Infrastructure as a Service (IaaS) in Cloud computing: A survey. ELSEVIER - Journal of Network and Computer Application, http://dx.doi.org/10.1016/ j.jnca.2013.10.004.
[8] Mell P, Grance T, 2011. The NIST Definition of Cloud Computing. Computer Security Division, Information Technology Laboratory.
[9] S. Mousavi, A. Mosavi, and A. R. Varkonyi-Koczy. 2018. A Load Balancing Algorithm for Resource Allocation in Cloud Computing, pages 289-296. Springer International Publishing, Cham.
[10] Pallis G, 2010. Cloud Computing - The New Frontier of Internet Computing. IEEE Computer Society, IEEE INTERNET COMPUTING.
[11] Parikh S M, 2013. A Survey on Cloud Computing Resource Allocation Techniques. 4th Nirma University International Conference on Engineering, 978-1-4673-1719-1/12.
[12] V Krishna Reddy, B Thirumala Rao, Dr L S S Reddy, P Sai Kiran, 2011. Research Issues in Cloud Computing. Global Journal of Computer Science and Technology Volume 11 Issue 11 Version 1.0 Online ISSN: 0975-4172 and Print ISSN: 0975-4350.
[13] Sadashiv N, Dilipkumar S, 2011. Cluster, Grid and Cloud Computing. IEEE 6th International Conference on Computer Science and Education (ICCSE), 978-1-4244-9718-8/11.
[14] Xu L, Li J, 2016. Building Efficient Resource Management Systems in the Cloud: Opportunities and Challenges. International Journal of Grid and Distributed Computing, Vol 9, No 3. pp. 157-172.
[15] Yuan Y, Liu W, 2011. Efficient resource management for cloud computing. IEEE International Conference on System Science, Engineering Design and Manufacturing Informatization.
[16] Zhang L, Zhang J, Fiaidhi J, J. Morris Chang, 2012. Hot Topics in Cloud Computing. IEEE Computer Society, ITPro, pg 1520-9202.
[17] Zhang S, Yan H, Chen K, 2012. Research on Key Technologies of Cloud Computing. ELSEVIER - International Conference on Medical Physics and Biomedical Engineering, Physics Procedia 33, pg 1791 – 1797.
[18] Zissis D, Lekkas D, 2008. Addressing cloud computing security issues. ELSEVIER - Future Generation Computer Systems 28, pg 583–592.