Efficient Resources Allocation for Different Jobs in Cloud

K. Dinesh
PG Scholar
Department of CSE
Anna University of Technology
Coimbatore

G. Poornima
Assistant Professor
Department of IT
Sri Ramakrishna Engineering
College
Coimbatore

K. Kiruthika
UG Scholar
Department of CSE
Bannari Amman Institute of
Technology
Sathyamangalam

ABSTRACT

Cloud Computing is an emerging technique in recent years that provides computing as a services. In order to maximize resources utilization, many scheduling algorithms were analyzed and implemented. Job scheduling using Berger model is one of the algorithm for scheduling jobs. The combination of Berger model and Neural Network would overcome the disadvantage of Berger Model i.e., incompletion of task when tasks-resources match is not achieved. In this work, the submitted jobs are classified based on different parameters like bandwidth, memory, Completion time and Resources Utilization. The classified user tasks are passed to the neural network. Neural network consists of input layer, hidden layer and output layer. With the help of hidden layer, the jobs are matched with the resources by adjusting weight. The performance of the system has been improved by means of efficient use of bandwidth, reducing a completion time which in turn improves resources utilization. CloudSim, a simulation tool has been used to simulate and the results shows reduced completion time and increased performance of the system.

Keywords

Cloud Computing, Job Scheduling, Neural Network, Berger model, CloudSim.

1. INTRODUCTION

Cloud Computing is a pay as per usage model (e.g., gas, electricity etc,) and provide resources as per needs. The characteristics of cloud computing are rapid elasticity, measured service, on demand self-service, ubiquitous network access, location-independent resources polling. Cloud computing is a business paradigm. It is also enhancing highly use of data, resources and promote for data sharing and storing through internet. It can also focuses on scheduling job and resources, providing security for storing data.

Cloud Computing is service oriented model not user oriented one. Cloud computing can mainly provide four different service like Infrastructure as a service (IaaS), Software as a Service (SaaS), Platform as a Service (PaaS) and Communication as a service (CaaS) [1]. At the time of providing different services, there are several problems addressed. Among those addressed problems, job scheduling is one of the major problems for improving job utilization. Generally scheduling can be followed for giving equal preference to jobs and promote the efficiency use of resources. There are many open sources tools available for developing Cloud Computing like Eucalyptus systems, open Nebula etc.,

Job scheduling in cloud computing is mainly focused for improving the resources utilization such as bandwidth, memory and reduce completion time. In the below Fig 1, described that nodes or servers are establish communication with each other through two different links namely variable link and fixed link. Nodes are connected through wireless medium (i.e., Variable link). Nodes are physically connected using fixed link (wired medium). The nodes are connected through Internet.
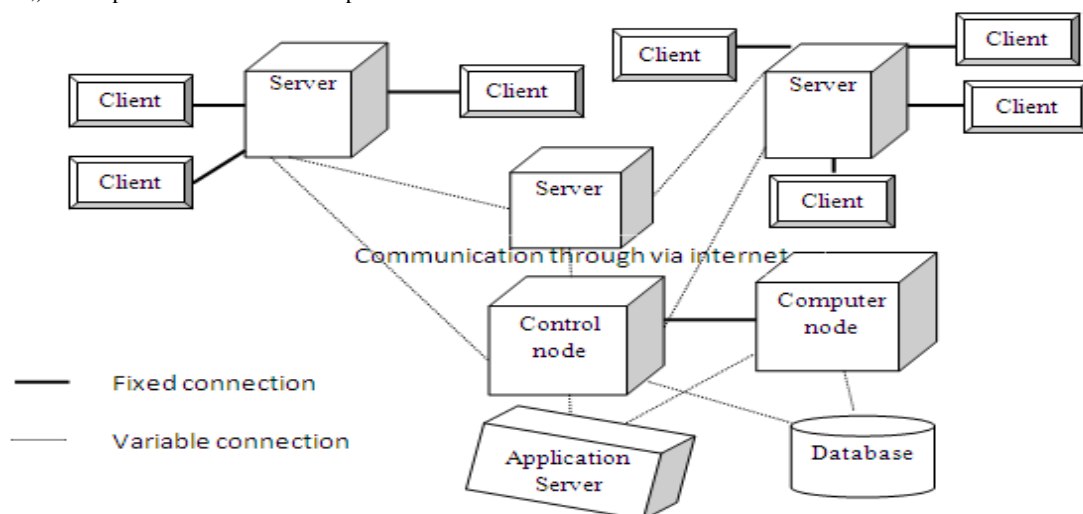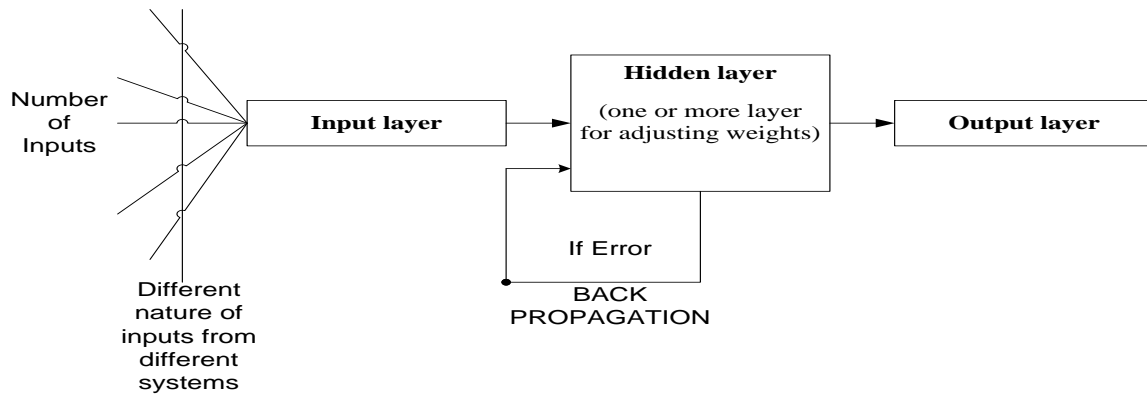

Fig 1: Cloud Computing

Fig 2: Neural Network

In Fig 2 illustrates the Neural Network structure. It is an information processing inspired by biological nervous system. It consists of large number of interconnected systems (nodes) and that is working together. Neural network is also learning from experience (i.e.: by example). There are two types of learning in neural network like supervised learning and unsupervised learning. It consists of three layers namely input layer, hidden layer and output layer. Many nodes or systems submit their input through input layer. Hidden layer is used to processing those input values. Back propagation algorithm can be used for solving error occur during processing.

The jobs are submitted from different users and also different environment like different operating system, different bandwidth etc. The heterogeneity is avoided by using virtualization. Virtualization is defined as the ability of running many operating systems in a single physical server and also sharing the hardware of the corresponding physical system[2]. Virtualization acts as an intermediate between operating system and hardware. Several virtual machines can also run in a single physical server. Virtualization is mainly used for resources consolidation and management. There are two types of virtualization such as Para virtualization and full virtualization. Full virtualization uses the entire system resources, which are abstracted by virtualization. Para virtualization uses only a little portion of system resources.

2. RELATED WORK

BAOMIN XU et al. [3] described the job scheduling using Berger model. In that scheduling the different classified user tasks are mapped with system resources. The user tasks classified based on the QoS parameters such as size, bandwidth and expectation time. The tasks are mapping with the system resources and if matched and also the task is completed within the expectation time. This reflects the tasks to be executed successfully. Otherwise, adjust the general expectation and map with the system resources. After adjusting, the tasks can complete within expectation time. Suppose after adjusting the expectation, the tasks cannot map with system resources and the tasks remain incomplete.
FERGUSON D et al. [4] proposed two principles in Micro-economic model. In that model have been involved in Micro-economic model; it is auctioning market and commodity market. In commodity approach to complete a single task, bundled resources and individual resources are provided. After the equilibrium only the price of a particular resource is fixed. Based on the CPU cycles, disk space, network bandwidth and memory space the resources are classified into different classes. It has maintained among demand and supply. The result will lead to increase in price, suppose both demand increase and supply decrease or vice versa. The first come first serve

principle is followed and also further price cannot be changed after achieving equilibrium condition. In grid environment, auctioning is also another algorithm for allocating resources. KUMARAN SUBRAMONIAM et al. [5] described various resources available in distributed system. Reassembling of resources in distributed resources compared with traditional resources and reflects in increase the reliability and performance of the system. The performances of the systems are depending on two different factors such as jobs and processors. Jobs is performing three functions such as compute budget set, preference relation and generates most preferred bid. Similarly processors also perform three actions like action resources, update price and advertising. The execution of this algorithm is similar to producer and customer problem and relations between them. For minimizing the complexity of large complex problems is split into set of smaller problems based on dependence in the economic model.
GOMOLUCH J et al. [6] explained an idea of computing resources allocation. The author described four different approaches in this work. They are state based (current snapshot of resources), Pre- emptive (resources migration in different systems), non Pre-emptive (execute in congregation itself) and Model based (predicts system state). There are three algorithm also discussed like round robin (RR), Proportional Share Protocol (PSP) and Continuous Double Actions (CDA). Message delay, processing delay, task creation, and number of server and speed of the servers are considered as parameters in distributed systems. The communication overhead of auctions is minimized in both State Based and Non pre-emptive based system model. System based and pre-emptive is dynamic and flexible in nature (e.g. divide and conquer).
REGEV O et al. [7] proposed a new Popcorn model. Popcorn is intermediate between single virtual parallel computing and the developers (programmers). It also provides connection to internet. It is also a global distributed system. Market based product is the principal behind Popcorn algorithm. In economic model, selling indicates jobs and buying indicates parallel program (like CPU time). The term market is used to connect the jobs and virtual parallel computer. There are no direct connection between seller and buyer (e.g. online trade). Similarly there is no connection exist among programmer and virtual parallel computer. There are two methods in Popcorn namely repeat vickrey auction and double auction mechanism.
CARL G et al. [8] analyzed simple neural network. It consists of input layer, hidden layer and output layer. In input layer N numbers of inputs are given and exemplar label is used in hidden layer. N inputs are passing to M nodes in hidden layers. Hidden layer also contain classes label to map the M nodes with output layers. In fuzzy neural network is a combination of AND and OR for finding maxima and minima for mapping input and

output nodes. Its rule based and back propagation algorithm is used for re-classify and mapping input with output labels.

KOTHALIL GOPALAKRISHNAN ANILKUMAR et al. [9] analyzed that job scheduling in a neural network. It is based on job priority assigner system. Scheduling is achieved for allocation of resources over time to perform collection of tasks. If the tasks are well trained then the ratio of missing rate is reduced. The relative priority is used to determine higher priority and change priority at any time. Back propagation algorithm is commonly used in well supervised learning. Mean squared error [MSE] is small when error propagation is high.

UDO SEIFFERT [10] explained the processing speed of the processor. It depends on the clock speeds, processor layouts and so on. Neural network is a technical model similar to biological model. It's evident of dealing with large scale networks and increasing training times. Artificial neural networks focus on highly accessing parallel computer like PC, workstation cluster and multiple processor machines. The hardware platforms checks for availability, suitability (no of nodes, communication loads, numerical complexity). The survey of the network contain message passing interface in homogeneous computer cluster. In neural network consists of input layer, hidden layer and output layer. The hidden layer is used for mapping the input layer with output layer through adjusting weight matrix. The back propagation algorithm is not best suitable if more than two hidden layers.

ASHA GOWDA KAREGOWDA et al. [11] combined the idea of BPN and GA using neural network. Neural networks are one of the data mining analytical tools for predicting the future from the past knowledge. Model selection for neural networks must choose various factors such as selection of number of hidden layers and relevant corresponding input, connection weight and output. Neural network is learning from experience. There are several learning techniques available. They are genetic algorithm (GAs), particle swarm optimization (PSO), ant colony optimization to find the network weights. Back propagation networks (BPN) are novel structure of information processing systems and supervised learning. Genetic algorithm is an optimized technique for natural selection and natural genetics. The individuals in the genetic space are called chromosome. There are three operation in genetic algorithm i.e., reproduction, crossover and mutation. Further, encoding can be classified as binary encoding, value encoding, permutation encoding and tree encoding. Different types of crossover by and large used are one point crossover and so on.

A.MAITHILI et al [12] discussed the idea of artificial neural network for diagnosis of cancer. An artificial neural network is a collection of interconnecting artificial neurons. Artificial neurons are used for solving artificial intelligence problems. There are two approaches in artificial neural networks like feed-forward networks and feedback networks. Feed-forward networks are a straight forward technique and can't use loops or feedback technique. Feedback networks are dynamic in nature. It's changing their 'static' state continuously to reach equilibrium point. Infrastructure as a service (IaaS) in cloud computing can use virtualization rather than hardware for any operation.

M.R.RAWTANI [13] proposed an new idea of knowledge management in cloud computing. Knowledge management contains large amount of different information. Intelligence concepts consist of four fundamental principles like data, information, knowledge and wisdom. Knowledge is a combination of experience (past knowledge), context etc, It contains a high value for making decision. Knowledge in cloud computing achieves efficient storage, multi-tenant environment, elasticity, scalable services, reliable access, high performance computing, fault tolerance and security.

SATHISH NARAYANA SRIRAMA et al [14] discussed map-reduce algorithm. Cloud computing contain virtual machine to solve resources scientific problems. The scientific computer application in cloud is running by using map reduce technique. The algorithm can be sub-divided into several classes. Map-reduce technique can be combined with iterative algorithm for solving scientific problems. Hodoop framework is inspired by Map-reduce technique and Google File system. Map-reduce technique is used for creating public cloud in university so that the resources can be utilized efficiently.

3. ARCHITECTURE

Large numbers of jobs are submitted from different users. The architecture diagram describes that how and when the tasks are executed. In Fig 3, describes that large number of tasks from different users. The user tasks are classified based on different parameters such as bandwidth, memory, size and completion time. The priority assigner is used to assign priority of different tasks along with deadline. Job queue is used to execute the remaining tasks based on FIFO. If user tasks is mapped with system resources and execute within expectation time then tasks remains complete. Otherwise, the tasks is re-classified and mapped with system resources until completion of tasks. .
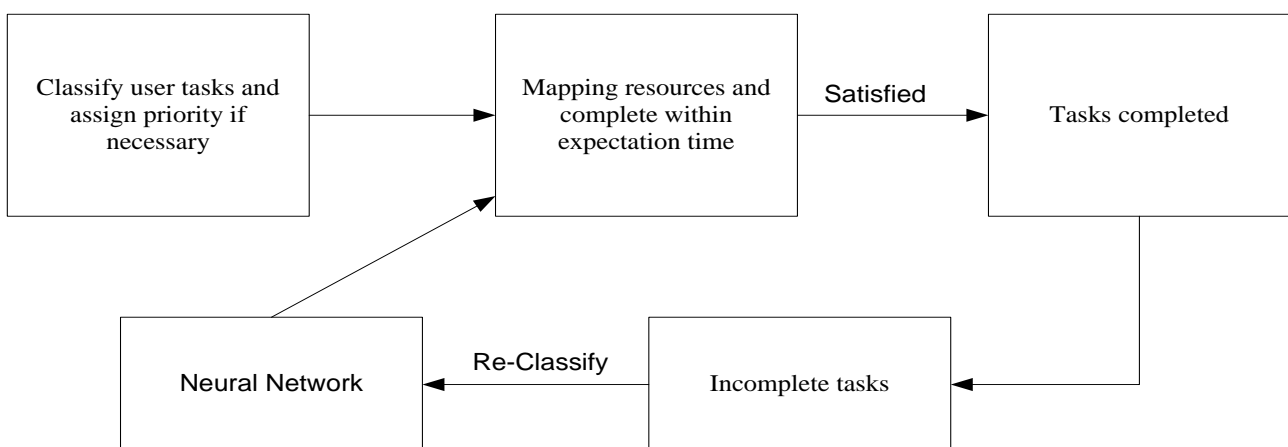

Fig 3: System Architecture

3.1 Characteristic of Jobs

3.1.1 Independent Jobs

The various Jobs submitted from different users are independent of each others. Each job can be sub divided into large number of tasks. DAG (Directed Acyclic Graph) is used to represent the possible dependency exists among this tasks. Large numbers of users submit their tasks at any time and various requirements.

3.1.2 De-Centralized Server

In the different jobs submitted from different users are run in different or same virtual machine of physical resources.

3.2 Scheduling Model

Based on the user inputs, the tasks are classified into two types like jobs with deadline and with deadline. Job queue is used to maintain the various jobs submitted from different users. Dispatching queue and also periodically update the job queue.

3.3 Task Mapping

The tasks completed, when the ratio between the actual time taken and expected time is less than one. Initially check whether the requesting resources less than available resources, the task may be mapped with system resources otherwise the tasks can be mapped with some other resources.

X=Requesting Resources/Available Resources   (1)

If X<1, then tasks may mapped with system resources. Otherwise, the tasks cannot be mapped with that resource.

The bandwidth can be calculated by measuring the distance between two points by Euclidean distance. Euclidean distance (ED) is calculated by square root of difference between two points[15]

$$ED= \sqrt{(X_1 - Y_1) + \cdots . (X_n - Y_{n)}}$$ (2)

Calculate the expectation time using expectation function and compare with actual completion time. Jobs are divided into number of independent tasks based on different parameters. The classified tasks are grouped by cluster.

Fig 4: denotes that the data flow diagram. There are different jobs submitted from various users. The jobs are classified into different independent tasks. The classified tasks are grouped by clustering.
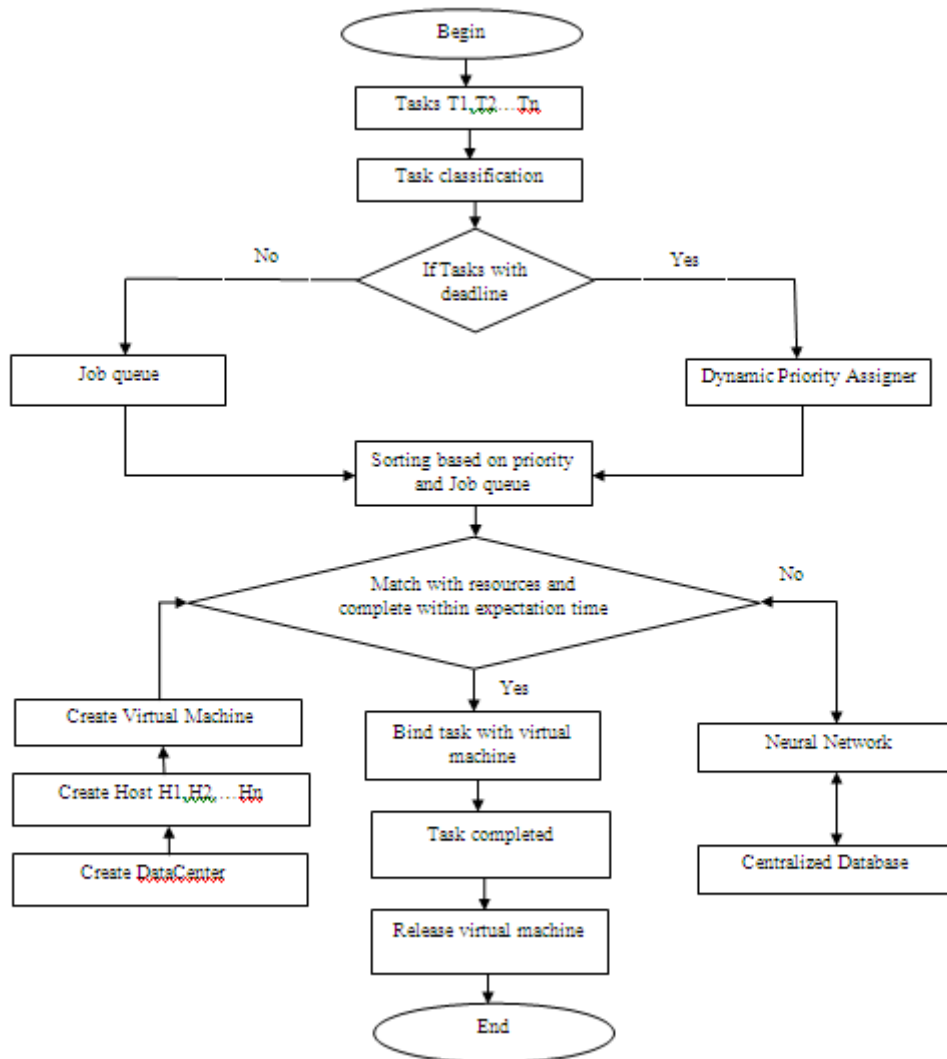

Fig 4: Data Flow Diagram

Check whether the user submits the different tasks along with deadline or not. If the user submit the tasks with deadline, then assign dynamic priority for those tasks.

The priority may be modified dynamically. Otherwise, if the tasks submitted by the users without deadline and then tasks are stored in a job queue. Finally, combine and sort the tasks in a priority assigner and also job queue. For mapping of resources with user tasks, virtual machines are created. If the user tasks are mapped with system resources and also tasks are completed within expectation time and then bind the user tasks with corresponding virtual machine. Otherwise, neural network can be used for mapping the user tasks with system resources. Before adjusting the weights of hidden layer for mapping the user tasks with system resources, just check in centralized database if already exists. In neural network, the weights of the hidden layers are adjusted for mapping the input layer with output layer. Back propagation algorithm is used for reassigning the weights of the hidden layer if the tasks cannot map with system resources. Feedback mechanism is used for reclassify the user tasks in a neural network. If the tasks are mapped with system resources and completed within expectation time and then store the results in centralized database for future reference. If the condition is satisfied and the user tasks are mapped with system resources, afterwards just bind the virtual machine with the user tasks for execution. After the tasks completed, the corresponding virtual machines can be released.

3.4 Pseudo Code for Resources Mapping

Step 1: The jobs are submitted from different users and classified based on dependence.

Step 2: Calculate the bandwidth and completion time of resources.

for tasks t=1 to n

{

$$ED= \sqrt{(X_1 - Y_1) + \cdots . (X_n - Y_{n)}}$$

}

Creating the number of virtual machines

{

Create and select virtual machines (Vm) according to the user's tasks requirements (Tn)

}

Step 3: Check whether the tasks are mapped with VM and complete within expectation time.

Step 4: If the tasks cannot complete, then the neural network is used for adjusting hidden layer for efficiently mapping with resources and repeat step 3.

Step 5: If step 3, executes successfully means then tasks completed and release corresponding virtual machine and update centralized database for future reference.

The centralized database is used for reducing the complexity of the resources mapping in future. This data are considered as the learning from experience in a neural network.

4. SIMULATED RESULTS

This section shows simulation and their result obtained from this work. In this simulation Neural Network is a learn from experience or examples. The simulation can run with the help of JAVA because it's also collection of package.

4.1 Simulation tool

In cloud computing, CloudSim is used as a simulation tool to implement this process. In CloudSim, the jobs or tasks are submitted by users. The submitted jobs are classified into different tasks. The user tasks are classified based on the parameters such as bandwidth, memory, completion time and resources utilization etc,. In CloudSim, the tasks are added to scheduling algorithm[16]. After scheduling, CloudSim used to bind the single user tasks with single virtual machine through Datacenter Broker and run the tasks with the help of virtual machine. CloudSim is implemented with the help of java. It consists of several datacenter for storing data. In a network, simulation contain server virtual machine can run in a single physical system. For binding Cloudlet (user tasks) with virtual machine CloudSim is used and allow the tasks for execution. It is used for supporting the search of reasonable virtual machine for a particular user tasks.

4.2 Performance Comparsion

The performanceof existing Berger model is compared with this work. The parameters consider are bandwidth utilization, CPU utilization and Completion time. Deponding on these parameters, the cost of the system is calculated.

Table 1: Specifies the values of bandwidth, CPU Utilization and Completion time

No of jobs	Bandwidth		CPU		Time	
	BM	NN	BM	NN	BM	NN
10	571	317	474	325	1688	1699
15	2102	300	2250	381	10976	7329
20	2108	478	1857	546	8397	5016
25	1030	508	1081	621	8459	5378

Hence in the below table clearly showed that completion time, bandwidth and CPU utilization are efficiently used compare with existing work. So the cost of this work is compactiably low when compare with existing work.

From the table, the chart shows that this work use only optimial bandwidth than existing work. Here the number of jobs (i.e., cloudlets in CloudSim) increases the bandwidth is improving linearly. But the allocated virtual machine's bandwidth of this work is less than the exisiting model. Based on these comparison and the experimental results shows better bandwidth utilization than existing work.
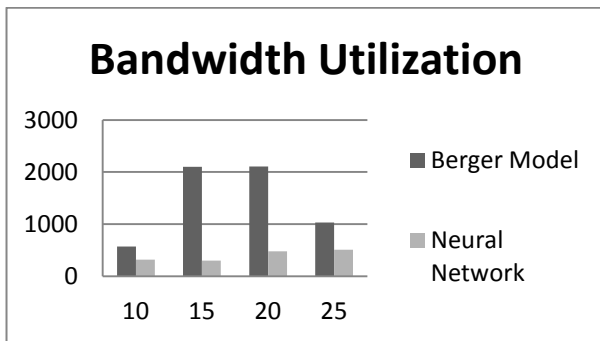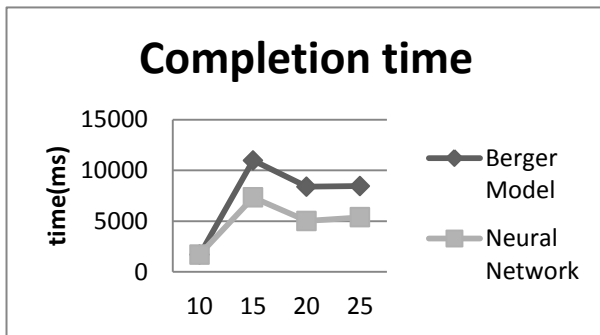
Fig 5: Bandwidth Utilizaion

Fig 6: Completion Time

From the above line chart shows that this work reduces the completion time than existing work. The line chart is drawn between no of cloudlets and completion time in millisecond (ms).the cloudlets are mapped with the resources hurriedly with their past experience. The above results obtained from CloudSim simulation clearly shows that completion is reduced in this work than existing work.

5. CONCLUSION

Various algorithms for mapping the resources with jobs had analyzed and compared. From this work, it is clearly showed that completion time, bandwidth utilization and CPU utilization are effectively utilized compare with existing work. The cost of the system also decreased due to efficient use of resources. Neural network is used either for reclassification of unmapped jobs or fast and efficient mapping from past knowledge. Further enhancement can be made by clustering all the jobs submitted from various users.

6. REFERENCES

[1] .Peeyush Mathur, " Cloud Computing: New Challenge to the entire Computer Industry", international conference on Parallel, Distributed and Grid computing(PDGC), 2010.

[2] Vmware, "Virtualization Overview", online @ "vmware.com/pdf/virtualization.pdf".

[3] Baomin Xu, Chunyan Zhao , Enzhao Hu, Bin Hu, "Job Scheduling algorithm using Berger model in Cloud Environment" ,Elsevier in Advances in Engineering Software, Vol 42 , No. 7, Pp. 419-425, 2011.

[4] Ferguson D, Yemini Y, Nikolaou, "C. Microeconomic algorithms for load balancing in distributed computer systems", Proceedings of the eighth international conference on distributed systems, San Jose: IEEE Press, Vol. 2, No. 30, Pp. 491–9,1988.

[5] Kumaran Subramoniam, Muthucumaru Maheswaran and Michel Toulouse, "Towards a Micro-Economic Model for Resource Allocation in Grid Computing Systems" Electrical and Computer Engineering, IEEE CCECE, Vol. 2,Pp. 782- 785, 2002.

[6] Gomoluch J, Schroeder M., "Market-Based resource allocation for Grid Computing: a model and simulation", In: Endler M, Schmidt D, editors. Proceedings of the first international workshop on middleware for Grid Computing (MGC 2003). Rio de Janeiro: Springer-Verlag, Vol. 6, No. 5, Pp. 211-218, 2009.

[7] Regev O, Nisan N, "The popcorn market – An online market for computational resources", Proceedings of the first international conference on information and computation economies Charleston: ACM Press, Vol.28,No1–2, Pp. 177-189,2000.

[8] Carl G.Looney and Sergiu Dascalu, "A Simple Fuzzy Neural Network", University of Nevada Reno,Vol 9, issue 2, Pp 89557, 2009.

[9] Kothalil Gopalakrishnan Anilkumar and Thitipong Tanprasert, "Neural Network Based Priority Assignment for Job Scheduler", Faculty of science and Technology, Assumption University Bangkok, Thailland Vol 9, issue 3,pp-181-186.

[10] Udo Seiffert,"Artificial Neural Networks on Massively Parallel Computer Hardware", European Symposium on Artificial Neural Networks, pp. 319-330 april 2002.

[11] Asha Gowda Karegowda, A.S.Manjunath, M.A.Jayaram, "Application of Genetic Algorithm Optimized Neural Network Connection Weights for Medical Diagnosis of Pima Indiana Diabetes", International journal on soft computing(IJSC), Vol.2, No.2, May 2011.

[12] Maithili A, Vasantha kumara R, Rajamanickam S, "Neural Networks Cum Cloud Computing Approach in Diagnosis of Cancer", International Journal of Engineering Research and Application(IJERA), Vol. 2, Issue 2, pp.428-435, Mar-Apr 2012.

[13] Rawtani R M, "Achieving Knowledge Management Through Cloud Computing", 8th Convention PLANNER 2012.

[14] Satish Narayana Srirama, Pelle Jakovits, Eero Vainikko, "Adapting Scientific Computing Problems to Clouds using MapReduce", Elsevier, may 2011.

[15] Venkatesa .Kumar V and Dinesh K,"Job Scheduling Using Fuzzy Neural Network in Cloud Environment",bonfring international journal of man machine interface, Vol 2, no 1, march 2012.

[16] Rajkumar Buyya, Rajiv Ranjan and Rodrigo N. Calheiros, "Modeling and simulation of scalable Cloud Computing environments and the CloudSim Toolkit: challenges and opportunities", Proceedings of the seventh high performance Computing and simulation conference(HPCS 2009, ISBN: 978-1- 4244-49071), Leipzig, Germany. New York, USA: IEEE Press, June 21–24, 2009.