International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8, Issue-1, May 2019

A Suite of Load Balancing Algorithms and Service Broker Policies for Cloud: A Quantitative Analysis with Different User Grouping Factor in Cloud

Kumar PJ, Suganya P, Komal Malhotra, Payal Yadav

Abstract: Several load balancing algorithms available to balance load in cloud data centers along with various service broker policies to select data centers based on the origin of user request. The user response time and data center request servicing time are important metrics to analyze the performance of cloud from the user perspective. We investigate the performance of Round Robin, and throttle based load balancing algorithms along with service broker policies such as closest data center and optimize response time with different user grouping factor in cloud environment. i.e., when the user grouping factor is set as 100 and 500. We present the results of simulation with various configurations as mentioned above, performed in CloudAnalyst, an open source tool. The results obtained would help the researchers, academicians, industrialist and cloud Vendors to effectively design their applications that might be deployed in cloud.

Index Terms: Round Robin, Data Centers, Virtual Machines, Load Balancing, Cloud Computing

I. INTRODUCTION

A set of load balancing algorithms exist in the literature to balance load in cloud data centers [15]. Round Robin load balancing, Throttled load balancing and eaually spread current execution are some of the popular algorithms. A set of service brokering policies are also available to select a particular data center based on the location of user request. The performance of these algorithms and policies differ in situations where the parameters such as bandwidth and device characteristics differ largely and moreover when the user request grouping factor is varied. We consider two different user grouping factor such as 500 and 1000 to observe the performance of the existing load balancing and service broker policies. The basic constituting components of computing such as platform, infrastructure, hardware and software are provided as service by the cloud vendors [15]. We evaluate the performance of well-known load balancing algorithms such as Round Robin, ESCE and Throttled (TLB) along with various service broker policies such as closest data center, optimize response time and dynamically reconfigure with homogeneous cloud setup. The rest of the paper is organized below. Section II presents the literature review on the existing methodologies and survey performed by various researchers in the literature. Section III presents the configuration scenario. Section IV presents the simulation results and interpretations. Section V concludes the paper with future works.

Revised Manuscript Received on 30 May 2019.

* Correspondence Author
Dr Kumar PJ , Associate Professor, School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu,India
P suganya , Assistant Professor, School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu,India
Komal Malhotra, Student of M.C.A, School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu,India
Payal Yadav,, Student of M.C.A, School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu,India

II. LITERATURE SURVEY

Considering the computing capacity of the servers the workload are assigned to servers in a model named CLB(Cloud Load Balancing) [1]. The proposed work ensures evenly distribution of tasks to all servers. A comprehensive analysis has been made by [2] on various load balancing techniques for Cloud. A suit of load balancing approaches has been proposed by [3] which ensures the improved metrics on system performance. An increase around 14% in energy consumption is obtained by the approach proposed by [4]. A QoS based approach has been proposed by [5]. Load balancing for systems which handles massive data has been proposed by [6]. The authors in [7] proposed artificial bee based algorithm for load balancing in cloud computing. The authors in [8][11][12] proposed algorithms to minimize the response time and access time in cloud based on the QoS parameters of the user. The source code of the simulator used in this paper can be obtained from the online sources as mentioned in [9] [10]. The authors in [13][14] discusses about various simulation environment available for cloud computing and also discusses the execution of load balancing algorithm in it.

III CLOUD COMPUTING ARCHITECUTRE AND LOAD BALANCING

Software as a Service [SaaS]	Multimedia Applications and Web services	Google App and Facebook

Platform as a Service [PaaS]	Software Tools	Microsoft Azure
Infrastructure as a Service [IaaS]	Infrastructure and Hardware (CPU, Memory and Bandwidth)	Amazon EC2, Data centers

Table I: Cloud Computing Architecture [15]

Based on the access to the cloud it can be categorized into public and private. Private cloud mostly behaves as a private network owned by an organization. It offers more security but restricted to the users of the organization only. In contrast, public cloud is open to users who can have access to it. But lack of security exist in public cloud compared to the private cloud. Cloud vendors offers various services such as Platform as a service, infrastructure as a service and software as a service. The cloud users can avail the services through internet. Selection of data centers are performed by service broker policies that selects a data center based on a policy such as closest one. Load balancing algorithms are used to select virtual machines based on the policy to allocate task on it. [15]

IV ANALYZING THE PERFORMANCE OF RR AND TLB LOAD BALANCING ALGORITHMS WITH CDC AND ORT SERVICE BROKER POLICIES WITH DIFFERENT USER GROUPING FACTOR.

A. SIMULATION CONFIGURATION

Table II shows the settings of bandwidth between different regions in Cloud. The bandwidth is varied between 500 Mbps to 2500 Mbps. Table III shows the configuration of physical hosts in data center. The virtual machine configuration is set as shown in Table IV .Various attributes of physical hosts such as memory , storage, bandwidth , processor count and processor speed are set as shown in the Table III.

Region/ Region	0	1	2	3	4	5
0	2000	1000	1000	1000	1000	1000
1	1000	800	1000	1000	1000	1000
2	1000	1000	2500	1000	1000	1000
3	1000	1000	1000	1500	1000	1000
4	1000	1000	1000	1000	500	1000
5	1000	1000	1000	1000	1000	2000

Table II: Bandwidth Configuration between various Regions

ID	Memory(Mb)	Storage (Mb)	Available BW	Number of Processors	Processor Speed	VM Policy
0	2048	1000	1000	3	1000	TIME_SHARED
1	2048	1000	1000	3	1000	TIME_SHARED
2	2048	100	1000	3	100	TIME_

	Region	Processor Architecture	OS	VMM	Cost Per VM $/Hour	Memory cost $/second	Storage Cost $/s	Data Transfer Cost $/Gb
DC 1	3	X86	Linux	Xen	1	0.05	1	1
DC 2	2	X86	Linux	Xen	1	0.05	1	1
DC 3	0	X86	Linux	Xen	1	0.05	1	1
DC 4	1	X86	Linux	Xen	1	0.05	1	1
DC 5	4	X86	Linux	Xen	1	0.05	1	1

		0			0	SHARED

Table III: Host Configuration for Virtual Machines in Data Center

Table IV: Data Center Configuration

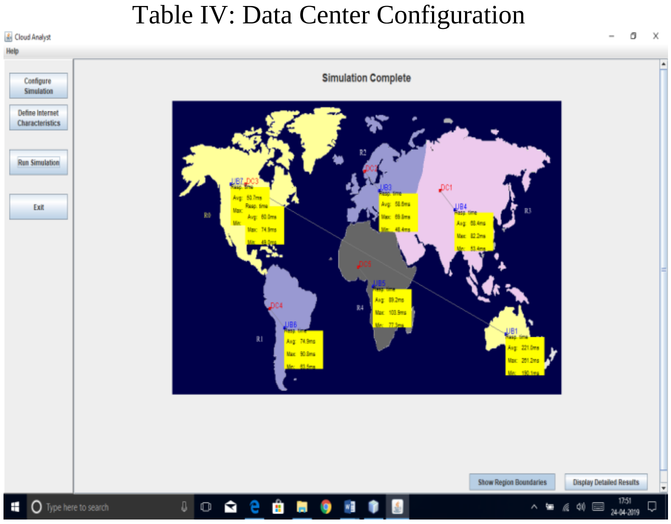

Fig.1: Placement of Data Centers and User Base in Cloud Analyst Simulation Tool

B. RESULTS AND DISCUSSION

Table V shows the overall summary of simulation done using Round Robin load balancing algorithm with closest data center as service broker policy. The user grouping factor is set to 100. The overall user response time is obtained as 211.29 ms. The data center processing time is 11.09ms.

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	211.29	44.37	659.44
Data Center processing time:	11.09	0.10	76.68

Table V : Simulation Result of RR with CDC (Grouping Factor=100)

The individual user base (UB1, UB2..UB6) response time is shown in the Table VI with average, minimum and maximum values.

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	340.00	168.46	528.26
UB2	122.60	47.44	207.52
UB3	109.90	44.37	185.70
UB4	144.67	49.62	254.66
UB5	326.23	61.15	659.44
UB6	224.17	54.14	421.14

Table VI : Average Response Time in RR with CDC (Grouping Factor=100)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the
Table VII with average, minimum and maximum values.

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	10.70	0.11	30.91
DC2	13.68	0.11	51.85
DC3	11.30	0.10	76.68
DC4	9.80	0.10	21.00
DC5	9.73	0.10	20.99

Table VII : Data Center Request Processing Time in RR with CDC (Grouping Factor=100)

Table VIII shows the overall summary of simulation done using Round Robin load balancing algorithm with closest data center as service broker policy. The user grouping factor is set to 500. The overall user response time is obtained as 207.85 ms. The data center processing time is 9.14ms.

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	207.85	46.84	640.66
Data Center processing time:	9.14	0.10	52.55

Table VIII : Simulation Result of RR with CDC (Grouping Factor=500)

The individual user base (UB1, UB2..UB6) response time is shown in the Table IX with average, minimum and maximum values

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	336.86	170.38	506.80
UB2	122.34	50.21	219.93
UB3	109.58	46.84	187.83
UB4	143.44	51.07	272.08
UB5	315.90	74.26	640.66
UB6	218.83	59.54	426.59

Table IX : Average Response Time in RR with CDC (Grouping Factor=500)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the Table X with average, minimum and maximum values

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	9.09	0.10	30.97
DC2	9.92	0.11	33.69
DC3	9.31	0.10	52.55
DC4	8.74	0.10	30.97
DC5	8.48	0.10	30.93

Table X : Data Center Request Processing Time in RR with CDC (Grouping Factor=500)

Table XI shows the overall summary of simulation done using Round Robin (RR) load balancing algorithm with optimize response time (ORT) as service broker policy. The user grouping factor is set to 100. The overall user response time is obtained as 211.29 ms. The data center processing time is 11.10ms.

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	211.29	43.90	4087.93
Data Center processing time:	11.10	0.10	3970.53

Table XI : Simulation Result of RR with ORT (Grouping Factor=100)

The individual user base (UB1, UB2..UB6) response time is shown in the Table XII with average, minimum and maximum values

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	339.98	178.43	519.21
UB2	122.60	45.51	214.42
UB3	110.03	43.90	4087.93
UB4	144.65	46.91	257.25
UB5	326.16	58.96	642.53
UB6	224.17	53.68	439.57

Table XII : Average Response Time in RR with ORT (Grouping Factor=100)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the Table X with average, minimum and maximum values

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	10.70	0.10	30.97
DC2	13.77	0.10	3970.53
DC3	11.31	0.10	58.88
DC4	9.79	0.13	21.00
DC5	9.73	0.10	20.99

Table XIII : Data Center Request Processing Time in RR with ORT (Grouping Factor=100)

Table XIV shows the overall summary of simulation done using Round Robin (RR) load balancing algorithm with optimize response time (ORT) as service broker policy. The user grouping factor is set to 500. The overall user response time is obtained as 207.84 ms. The data center processing time is 9.16ms

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	207.84	48.12	641.97
Data Center processing time:	9.16	0.10	55.98

Table XIV : Simulation Result of RR with ORT (Grouping Factor=500)

The individual user base (UB1, UB2..UB6) response time is shown in the Table XV with average, minimum and maximum values

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	336.88	164.35	522.60
UB2	122.24	48.12	212.17
UB3	109.53	48.17	184.58
UB4	143.61	51.46	255.95
UB5	315.79	73.95	641.97
UB6	218.87	62.37	442.39

Table XV : Average Response Time in RR with ORT (Grouping Factor=500)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the Table X with average, minimum and maximum values

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	9.13	0.10	30.93
DC2	9.93	0.10	34.07
DC3	9.31	0.10	55.98
DC4	8.75	0.10	30.97
DC5	8.52	0.10	30.98

Table XVI : Data Center Request Processing Time in RR with ORT (Grouping Factor=500)

Table XVII shows the overall summary of simulation done using Throttle (TLB) load balancing algorithm with closest data center (CDC) as service broker policy. The user grouping factor is set to 100. The overall user response time is obtained as 210.69 ms. The data center processing time is 9.88ms

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	210.69	44.37	659.44
Data Center processing time:	9.88	0.10	26.43

Table XVII : Simulation Result of TLB with CDC (Grouping Factor=100)

The individual user base (UB1, UB2..UB6) response time is shown in the Table XVIII with average, minimum and maximum values

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	339.76	168.46	528.26
UB2	121.59	47.44	207.52
UB3	108.09	44.37	185.24
UB4	144.25	49.38	253.74
UB5	326.18	61.15	659.44
UB6	224.14	54.14	427.37

Table XVIII : Average Response Time in TLB with CDC (Grouping Factor=100)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the Table XIX with average, minimum and maximum values

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	9.78	0.11	19.26
DC2	10.27	0.11	21.62
DC3	9.89	0.10	26.43
DC4	9.73	0.10	18.76
DC5	9.72	0.10	17.90

Table XIX : Data Center Request Processing Time in TLB with CDC (Grouping Factor=100)

Table XX shows the overall summary of simulation done using Throttle (TLB) load balancing algorithm with closest data center (CDC) as service broker policy. The user grouping factor is set to 500. The overall user response time is obtained as 208.23 ms. The data center processing time is 8.86ms

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	208.23	46.84	640.66
Data Center processing time:	8.86	0.10	45.61

Table XX : Simulation Result of TLB with CDC (Grouping Factor=500)

The individual user base (UB1, UB2..UB6) response time is shown in the Table XXI with average, minimum and maximum values

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	337.24	170.38	504.86
UB2	122.53	50.21	204.81
UB3	109.65	46.84	186.95
UB4	143.81	51.07	274.08
UB5	316.61	74.26	640.66
UB6	219.41	59.54	426.59

Table XXI : Average Response Time in TLB with CDC (Grouping Factor=500)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the Table XXII with average, minimum and maximum values

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	8.82	0.10	20.89
DC2	9.34	0.11	30.11
DC3	8.93	0.10	45.61
DC4	8.61	0.10	20.73
DC5	8.51	0.10	20.43

Table XXII: Data Center Request Processing Time in TLB with CDC (Grouping Factor=500)

Table XXIII shows the overall summary of simulation done using Throttle (TLB) load balancing algorithm with Optimize Response Time (OTR) as service broker policy. The user grouping factor is set to 100. The overall user response time is obtained as 210.68 ms. The data center processing time is 9.88 ms

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	210.68	43.90	642.53
Data Center processing time:	9.88	0.10	25.46

Table XXIII: Simulation Result of TLB with ORT (Grouping Factor=100)

The individual user base (UB1, UB2..UB6) response time is shown in the Table XXIV with average, minimum and maximum values

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	339.75	178.43	519.21
UB2	121.57	45.51	208.31
UB3	108.12	43.90	183.81
UB4	144.23	46.91	257.25
UB5	326.13	58.96	642.53
UB6	224.12	53.68	421.35

Table XXIV : Average Response Time in TLB with ORT (Grouping Factor=100)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the Table XXV with average, minimum and maximum values

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	9.77	0.10	19.04
DC2	10.27	0.10	20.24
DC3	9.89	0.10	25.46
DC4	9.73	0.13	17.12
DC5	9.72	0.10	17.33

Table XXV : Data Center Request Processing Time in TLB with ORT (Grouping Factor=100)

Table XXVI shows the overall summary of simulation done using Throttle (TLB) load balancing algorithm with Optimize Response Time (OTR) as service broker policy. The user grouping factor is set to 500. The overall user response time is obtained as 208.23 ms. The data center processing time is 8.86 ms

	Avg (ms)	Min (ms)	Max (ms)
Overall response time:	208.23	48.12	645.97
Data Center processing time:	8.86	0.10	38.70

Table XXVI : Simulation Result of TLB with ORT (Grouping Factor=500)

The individual user base (UB1, UB2..UB6) response time is shown in the Table XXVII with average, minimum and maximum values

Userbase	Avg (ms)	Min (ms)	Max (ms)
UB1	337.23	164.35	522.60
UB2	122.42	48.12	212.17
UB3	109.63	48.17	183.26
UB4	143.97	51.46	252.12
UB5	316.53	73.95	645.97
UB6	219.46	62.37	442.39

Table XXVII : Average Response Time in TLB with ORT (Grouping Factor=500)

The individual data center request servicing time (DC1, DC2…DC5) response time is shown in the Table XXVIII with average, minimum and maximum values

Data Center	Avg (ms)	Min (ms)	Max (ms)
DC1	8.84	0.10	29.52
DC2	9.35	0.10	29.48
DC3	8.92	0.10	38.70
DC4	8.61	0.10	20.65
DC5	8.53	0.10	20.61

Table XXVIII : Data Center Request Processing Time in TLB with ORT (Grouping Factor=500)

Based on the results obtained through simulation when the user base grouping factor is set to 500 , Round Robin load balancing algorithm with optimize response time as a service broker policy produces less user base average response time 207.84 ms compared to TLB with other service broker policies. when the user base grouping factor is set to 100 , Throttle load balancing algorithm with optimize response time as a service broker policy produces less user base average response time 210.68 ms compared to RR with other service broker policies. when the user base grouping factor is set to 500 , Throttle load balancing algorithm with optimize response time as a service broker policy produces higher user base average response time 208.23 ms. when the user base grouping factor is set to 100 , Round Robin load balancing algorithm with closest data center as a service broker policy produces higher user base average response time 211.29 ms. TLB with OTR produces less data center processing time compared to all other algorithms with a values of 8.86 ms when the user grouping factor is set to 500. RR with CDC produces less data center processing time with a value of 9.14ms when the user grouping factor is set to 100.

V. CONCLUSION

We have investigated the performance of various load balancing algorithm for cloud such as Round Robin and Throttle based with various service broker policies such as closest data center and optimize response time using a simulator Cloud Analyst, an open source simulation tool available for cloud. The performance of load balancing algorithms is tested under various user group factoring values such as 500 and 1000. The performance metrics such as average user response time and data center request serving time obtained through simulation with various combination of load balancing algorithms and service broker policies are presented in this paper which would help researchers, cloud vendors and academicians to design and configure data centers and user bases effectively

REFERENCES

1. Shang-Liang ChenYun-Yao Chen Suang-HongKuo, “CLB: A novel load balancing architecture and algorithm for cloud services, Computers & Electrical Engineering, Volume 58, Pages 154-160, February 2017.
2. Minxian Xu Wenhong Tian Rajkumar Buyya, “A survey on load balancing algorithms for virtual machines placement in cloud computing”, Concurrency and computation Practice and Experience, Wiley Online Library, Volume 29, Issue 12, march 2017.
3. Einollah Jafarnejad Ghomia Amir Masoud RahmaniaNooruldeenNasih Qaderb, “Load-balancing algorithms in cloud computing: A survey”, Journal of Network and Computer Applications, Volume 88, , Pages 50-71, June 2017
4. Seyed Ebrahim Dashti & Amir Masoud Rahmani , “Dynamic VMs placement for energy efficiency by PSO in cloud computing”, Journal of Experimental & Theoretical Artificial Intelligence, Volume 28, Issue 1-2: Advances and Applications of Swarm Intelligence, 2016
5. M. H. Ghahramani, MengChu Zhou,Chi Tin Hon, “Toward cloud computing QoS architecture: analysis of cloud systems and cloud services”, EEE/CAA Journal of Automatica Sinica, Volume 4, Issue 1, PP:6-18, Jan 2017
6. Jianhua Peng, Ming Tang, Ming Li & Zhiqin Zha,, “A Load Balancing Method For Massive Data Processing Under Cloud Computing Environment”, Intelligent Automation & Soft Computing, Issue 4, PP: 547-553, 2017.
7. Yao, J.H., Ju-hou, Load Balancing Strategy of Cloud Computing Based On Artificial Bee Algorithm in Computing Technology and Information Management (ICCM), IEEE: Seoul. p. 185 - 189. 2012
8. Kumar P.J., Ilango P., “MQRC: QoS aware multimedia data replication in cloud”, International Journal of Biomedical Engineering and Technology, Vol.No:25, Issue.No:2/3/4, PP:250-266, 2017
9. cloudsim.cloudbus;Available from: http://www.cloudbus.org/cloudsim/.
10. https://sourceforge.net/projects/cloudanalystnetbeans/
11. Kumar P.J., Ilango P., “BMAQR: Balanced multi attribute QoS aware replication in HDFS”, International Journal of Internet Technology and Secured Transactions, Vol.No:8, Issue.No:2, PP:195-208,2018.
12. P.J.Kumar, P.Ilango, “An Optimized Replica Allocation Algorithm Amidst of Selfish Nodes in MANET”, Wireless Personal Communications,Vol.No:94, Issue.No:4, PP:2719-2738,2017.
13. Pakize, S.R., S.M. Khademi, and A. Gandomi, Comparison Of CloudSim, CloudAnalyst And CloudReports Simulator in Cloud Computing. International Journal of Computer Science And Network Solutions, 2: p. 19-27, 2014
14. Ray, S. and A. De Sarkar, Execution Analysis of Load Balancing Algorithms In Cloud Computing Environment. International Journal on Cloud Computing: Services and Architecture (IJCCSA), 2(5): p. 1-13,2012..
15. Hafiz Jabr Younis, Alaa Al Halees, Mohammed Radi, “Hybrid Load Balancing Algorithm in Heterogeneous Cloud Environment”, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-5 Issue-3, July 2015 PP: 61-65

AUTHORS PROFILE

Dr Kumar PJ is currently working as Associate Professor in the School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu. He pursued his under graduation in Engineering with Electrical and Electronics as specialization, Master in Engineering with Computer Science as specialization and Ph.D in Mobile Adhoc Networks. His area of interest in research includes Operating systems, Distributed systems, Mobile Adhoc Networks and Internet of Things. He has published numerous papers in reputed journals with impact factor and filed an Indian patent recently for his innovative work in Routing protocols for MANET.

Suganya P is currently working as Assistant Professor in the School of Information Technology and Engineering, VIT, Vellore, Tamil Nadu. She pursued her under graduation in Engineering with Electrical and Electronics as specialization, Master in Information Technology with Networking as specialization . Her area of interest in teaching includes Python Programming, Microprocessor&Micro Controllers and Internet of Things. She has published numerous papers in reputed journals .

Komal Malhotra is currently a student of MCA in VIT, Vellore, India. She pursued her under graduation in Computer applications from S.R.P.A.B College, Pathankot, Punjab. She worked in industry as Web developer for a couple of years. She has done various projects in cloud computing. Her programming interest includes languages like C, C++, Java and C#

Payal Yadav is currently a student of MCA in VIT, Vellore, India. She pursued her under graduation in Computer applications. She has done various projects in cloud computing and Computer Networks as part of her academic requirements using simulators like QualNET and Cloud Analyst. Her programming interest includes languages like C, C++ and Java