International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6, August, 2019

Efficient Fuzzy Based Real Time Scheduling of Workflows on Cloud

R. A. Kulkarni, S.H.Patil N.Balaji

Abstract: ***The technological advancements in service oriented architectures have given rise to Cloud computing as a new platform in the area of distributed computing. Virtualization is a prominent technology adopted by cloud while providing pay as per usage model. Cloud provides the computing resources in terms of different type of virtual machines. VM migration is a fundamental phenomena in cloud systems while providing uninterrupted service and load balancing to its users according to service level agreement. The virtualization and migration poses many challenges and complexities because of many uncertainties. Uncertainties are the inaccurate data in the form of task parameters, variation in the VM performance etc. As the Fuzzy logic can be employed effectively in decision making with inaccurate data we propose fuzzy logic based real time scheduling algorithm to deal with the uncertainties. The proposed algorithm in this paper performs in a better way in meeting of deadlines of real time workflows under uncertain data and VM migrations.***

Index Terms**: *Cloud, Fuzzy Logic.. Real time Scheduling, Virtualization, VM migration***

I. INTRODUCTION

Cloud computing models a category of sophisticated on demand services to its users. It provides a way for service providers and individuals to access resources located anywhere on demand. Cloud provides various pay-as-use access to computing, storage and network resources. While providing these services there is service provider and end user who takes these services. In [22] the author has defined cloud computing as a computing system consisting of large number of inter connected and virtualized parallel and distributed resources which can be dynamically provided to the customers as a unified resource based on some agreement between the provider and end used. National Institute of Computer Technology (NIST) has defined the cloud computing as Model for providing on demand network based access to pool of resources of various types such as network, storage, applications, servers etc. These resources can be configured, provisioned and released dynamically based on consumer needs with very less interaction from the provider [1]. Cloud provides different type of services to its end users. These services are Infrastructure as a service (IaaS), Platform as a service (PaaS) and Software as a Service (SaaS). While providing Infrastructure as service, cloud mainly uses the virtualization technology. Virtualization allows to treat single physical machine as multiple virtual machines hosting multiple operating systems. Virtualization technology comes across many challenges in supporting real-time applications as these applications have to meet the timing constraints while running on dynamic, shared and virtualized computing environments. Resource provisioning and scheduling in cloud systems has been studied a lot in cloud computing however real time scheduling, scheduling under uncertainties and managing deadline constrained workflows has drawn less attention. A workflow on the cloud consists of group of tasks to be executed on the provided resources. The workflows have to meet deadlines in satisfying the timing requirements of real time scheduling. While meeting the timing constraints of workflow and resources within the cloud environment, these services are prone to considerable uncertainties. Uncertainty [3] may be because of inaccurate estimation of runtime, sharing of common resources with other Virtual machines and lack of exact knowledge about the system and because of dependencies among the tasks. For the effective performance, the sources of uncertainty like virtualization, jobs arrival, migration, resource availability have to be handled in a better way.

Fuzzy logic being an effective approach to handle uncertain, imprecise data helps in providing solution to decision-making problems. In our proposed solution we are using fuzzy logic in deciding the priority of tasks of an workflow in meeting the deadlines and migrating the VMs to other hosts in case of overloading and the failure of VMs. We are proposing an algorithm to analyze the effect of VM migration and Fuzzy mapping of VMs on meeting deadline of real time workflows on cloud during a fault processing or during periods of computational overload. By dynamically adjusting the fuzzy mapping functions it is possible to get better performance in meeting deadline of workflows.

Rest of the paper is arranged as follows. Section II provides the summary of the related work in the field of VM management, uncertainty, fuzzy logic etc. Section III contains the architectural details and methodology of the proposed system, Section IV discusses about the experimental setups and results. We present conclusion and the future work in section V.

Revised Manuscript Received on October 30, 2019.
* Correspondence Author
Mrs. R.A.Kulkarni*, Dept of Comp.Engg, PICT Pune, India
Dr.S.H.Patil, Prof in CSE Dept, BVCOE, BV university, PUNE, India
Dr.N.Balaji, Prof in ECE Dept, JNTU Kakinada University,COE, VIZIANAGARAM, INDIA

II. RELATED WORK

Cloud computing which is a form of utility computing offers resources required for the workflow applications on pay-per-use model. While providing these services adhering to the quality of service is of almost importance.

As there is service level agreement between service provider and end user in providing good quality of service the cloud service provides has to face many challenges as discussed in [2]. These challenges are because of the various uncertainties which are seen [20]. In cloud systems there are many factors which cause uncertainties like virtualization, migration, fault tolerance etc. Understanding the sources of these uncertainties and proposing the solution can be done with the help of fuzzy logic. These uncertainties further complicate the performance of real time scheduling of deadline sensitive workflows on cloud systems [3]. They have discussed about two approaches for real-time virtual machine scheduling which have been made as a part of hypervisor design.

In [10] Shamsi has addressed the challenges and probable solution for meeting the resources requirements and timing constraints of data intensive applications on Hadoop platform. Brendam Jennings [11] have discussed about the various aspects of resource management challenges in cloud environment. Fei Teng [12] have proposed the Paused Rate Monotonic (PRM) algorithm. In their algorithm they have analyzed performance of hard real time tasks on Map reduce cloud environment in which number of tasks executed by dividing the computations into smaller tasks and combining the results of various tasks. Performance of the algorithm has been analyzed in theory. Chitra Devi [6] and Rhymend have done the work related to an load balancing and scheduling of various jobs in cloud settings considering the facts related to size of VMs computing capacity, no. of tasks involved in the job and how the jobs are dependent on each other. Chen-Wei [7] has presented a hyper-heuristic scheduling algorithm (HHSA) based on the usage of heuristics in order to find improved scheduling solutions for cloud computing systems. Fahmy [8] has used fuzzy logic based algorithm for scheduling on single processor the a periodic jobs with deadline of soft real-times. The single processor is assumed to be with heavy loads as it is in use by multiple users. Tom Springer [14] has presented a new scheduling method for real-time jobs in an embedded system. In their method they have used hierarchical scheduling to make provision for a resource based allocation scheme while using a fuzzy logic based feedback scheduler to catch the environmental changes within the application. Jyoti sahni [16] has proposed "a dynamic cost-effective deadline-constrained heuristic algorithm" which works on public cloud to schedule the scientific workflows. The proposed algorithm makes use of advantages provided by Cloud systems. It also considers the performance variability of virtual machines and instance acquisition delay to apply a just-in-time schedule of workflows related to scientific applications which are deadline constrained at lesser costs. The authors in [4] have worked on virtual machine consolidation method to have uniform energy consumption and quality of service requirements using the fuzzy logic and heuristic based approach. In their research work, they have used VM selection based on fuzzy logic. It selects VM to be migrated from an overloaded host in order to save the energy usage. They have incorporated the migration policy using least migration time or with some threshold which will improve the performance of selected policy. A new overload detection algorithm has also been tried based on utilization of VMs like mean, median and standard deviation.

M.A Rodriguez and Rajkumar Buyya [18] have presented an algorithm using an optimization technique of heuristics like meta heuristics and particle swarm optimization (PSO). In their algorithm the focus is on minimization of overall workflow execution cost while meeting the QOS requirements of deadline. The authors in [19] have studied the performance variance among resources while scheduling the soft deadline constrained applications by utilizing the idle time of provided resources and with the help of creating multiple copies of the same task and the budgetary requirements for the same. In [21] authors have carried out a detailed study on classification of various VM migration schemes. They have presented a qualitative investigation of currently used VM migration schemes. Anton Beloglazov and Rajkumar Buyya [22] have proposed a new approach for host overload detection. The algorithm used by them works on Markov chain model to maximize the intermigration time. It considers stationary workload and the state configuration. They have also proposed to apply a heuristic algorithm to handle dynamic and non-stationary workloads using the Multi size Sliding Window workload estimation method.

III. PROPOSED METHODOLOGY

Cloud Computing has been evolved on many different prominent research areas such as architecture designs for providing functionality as service, grid computing, utility computing and virtualization technology. While providing infrastructure as a service a single physical host will be divided into numbers of virtual machines. Each virtual machine will be responsible for performing assigned task to it as that of machine. As there will be number of virtual machines concurrently running on the same physical host the performance of the VMs may degrade. Virtual machine migration is also a quite frequent phenomena in cloud systems which may also effect the performance the time constrained workflows. As a part of our proposed work the fuzzy based logic has been studied and implemented in handling the virtual machine migration and mapping issues.

Real time scheduling on cloud systems is characterized by timeliness and predictability. Timeliness tells us how closely the scheduler can meet the timing requirements and predictability provides the information related to how much deviation is there in delivered timeliness. The various scheduling algorithms in cloud such as Optimized-Resource Scheduling Algorithm [14], Improved Cost-Based Algorithm for Task Scheduling: [24] etc. have implemented the scheduling algorithms for improving the resource utilization. The other algorithms like Particle Swarm Optimization-based Heuristic for Scheduling [25] etc. have also utilized the various scheduling algorithms for better resource utilization or load balancing.

These approaches have not considered the effect of overloading, failure and migration on meeting of deadline of real time workflows on cloud. In our proposed approach we are considering the effect of all these factors on meeting the deadline of real time workflows. Fuzzy logic plays important role in working with uncertain data. In our approach to work with uncertainties we are proposing fuzzy based mapping and migration of VMs on which the tasks are actually executed. Fig. 1 shows the architectural design of our fuzzy based system.

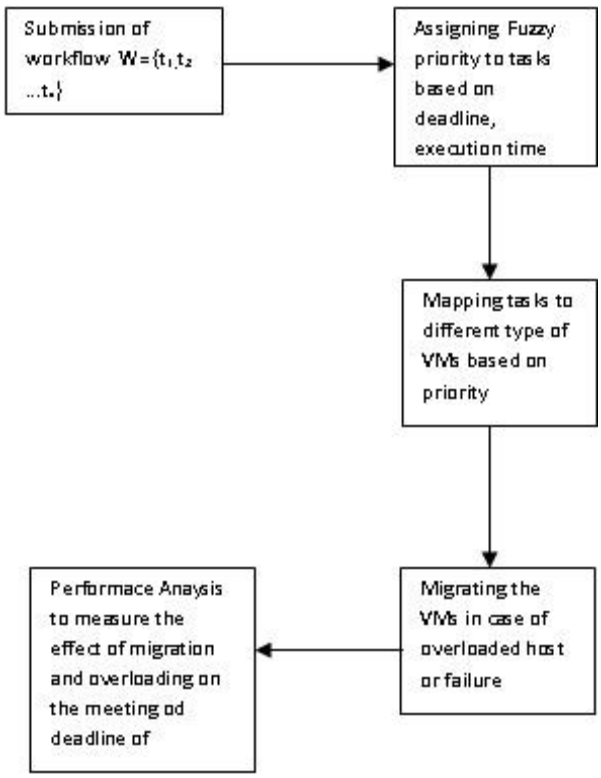


Fig 1. Architecture of Fuzzy based Scheduling System

Our methodology works by giving priorities to various tasks of workflow based on fuzzy mappings. Based on the deadlines of tasks and priority, the tasks will be mapped to different type of VMs. After the allocation of tasks to VMs because of sharing of resources a host may be overloaded. In case of overloading of host, few of the VMs have to be moved from the overloaded host to other hosts which may be under loaded or idle at that time. In our study we are exploring the effect of VM relocation on meeting of deadline of real time tasks. The algorithmic steps of the proposed method has been given in Algorithm 1.

Algorithm 1.

Algorithm for our proposed system:

1. Input the workflow having a set of tasks along with various attributes values like the expected execution time (EET), Assigned priority (AP) and task deadline(TDL).
2. Fuzzify the input parameters the expected execution time (EET), Assigned priority (AS) and task deadline(TDL).
3. Based on the inference rules calculate the overall priority.
4. Defuzzification the output i.e. overall priority.
5. Based on the priority, map the tasks to different type of VMs.
6. Check for a situation of overloaded hosts based on threshold and with the help of fuzzy logic with minimum migration time to identify the host to Live migrate the VMs.
7. Evaluate performance of new method for deadline misses of the workflow.

In our algorithm we are considering the least migration time as the policy for migrating the VM from an overloaded Host. In identifying the overload, we are using the threshold such that CPU usage of all the VMs of a host is more than the CPU threshold as given in the equation 1. The CPU $_{threshold}$ can be decided on some previous history on the usage of cloud resources.

$$CPU_u(host) \geq CPU_{threshold} \qquad (1)$$

Least Moving Time (LMT) policy selects the VM which can be migrated within Least time limit[4]. The transferring or moving time of VM can be calculated with the parameters of memory utilization of VM and the available bandwidth. At a given time t_m, the above mentioned Migration Control policy finds VM for migrating as selected by the equation 2. RAM(a) is the Radom Access Memory (RAM) usage of VM $_a$ and RAM(b) is the RAM usage of VM_b. NET($_{host}$) means the available bandwidth for migration and V_h is the set of all VMs of host h. So this methodology checks all the VMs of a overloaded host for required transfer time and migrates a VM whose amount of memory to be moved is least among all the VMs in host h.

$$VM(a) \in V_h \mid \forall\, VM(\mathrm{b}) \in V_h, \frac{RAM(a)}{NET_h} \leq \frac{RAM(b)}{NET_h} \qquad (2)$$

Even VM has to be migrated from the host if it has been failed in completing the work which has been allocated to it. I both the cases VM will be transferred from one host the identified host. In our experimentation study we have considered the effect of moving the VM under overloading and failure on meeting the deadline of real time tasks.

IV. RESULT ANALYSIS:

Experimental setup we have used CloudSim as simulation tool [9]. To integrate the fuzzy logic with basic CloudSim classes we have used jFuzzylogic language extension tool [10]. We have studies and analyzed our experiment for workflows consisting of set of tasks. While assigning tasks of a workflow it is very natural that the one task may be dependent on the output of other tasks. Such dependencies have been identified with the help of task dependency graphs. Task dependency graphs have been realized with directed acyclic graphs (DAG) .Computational complexity of our algorithm with DAG :

Assuming that the workflow WF (TA, ED) to be scheduled is having TA tasks and ED edges.

Cloud service providers provide different type of VMs such as Small, Medium and Large based on the capacity of RAM and processor speed.

Let m indicate the maximum number of VM types . Since WF is a DAG, the maximum number of edges in WF with n as number of Tasks is

$$(n-1)(n-2)2 \cong (n^2) \quad (3)$$

Since a task may have a maximum of (n-1)children tasks, The time required to consider all the dependencies among the tasks is an order of three $O(n^3)$. After the analysis of tasks belonging to workflow WF these tasks will be mapped to various available VMs. Following table shows assignment of fuzzy priorities to various tasks of a workflow.

Table 1 : Fuzzy priority assignment

Task	Expected Execution Time EET	Assigned Priority AP	Task Deadline TDL	Fuzzy Priority
T1	7	8	10	~10
T2	4	3	15	~12
T3	6	4	19	~14
T4	3	3	25	~17
T5	2	9	12	~9
T6	15	5	50	~20
T7	10	15	45	~18
T8	20	7	24	~25
T9	8	8	23	~24
T10	9	9	18	~23

The mapping functions used for fuzzy logic are the triangular functions and provided inference rules are AND inference rules. As fuzzy logic maps the region of values instead of binary logic of one or two we expect performance improvement with our proposed method.

Table (3) shows the comparison of performance of our algorithm with weighted round robin (WRR). It has been observed that using Fuzzy inference rules for assigning priorities and for VM migrations gives better results in terms of less number of Deadline misses.

The migration of VMs involves identification of the host machine based on the criteria of least migration time. Here we have used pre-copy migration technique for actual migration and the host to be migrated is the one which requires less amount of time. Pre-copy migration stops the VM at the source for very short period. As our overall goal is the meet the deadline of real time workflows we have selected with pre-copy migration methodology.

Table 2: Mapping of Tasks to different VMs

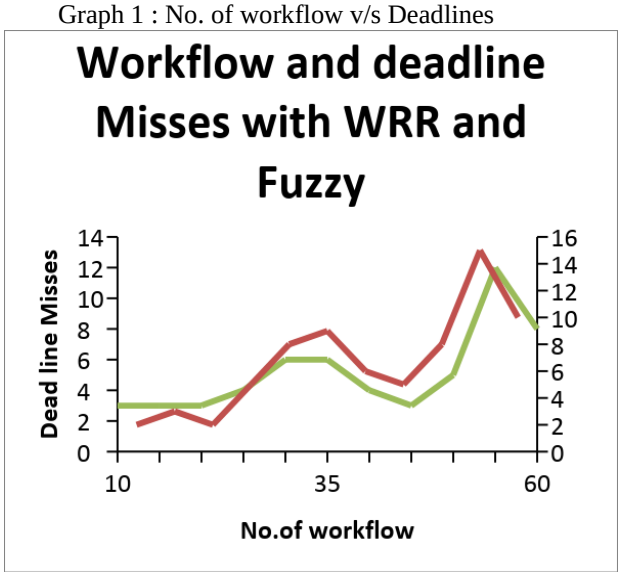
Task	**Fuzzy Priority**	**Mapped VM**
T1	~10	**Vm Small**
T2	~12	**Vm Medium**
T3	~14	**Vm Medium**
T4	~17	**Vm Medium**
T5	~9	**Vm Small**
T6	~20	**Vm Large**
T7	~18	**Vm Medium**
T8	~25	**Vm Large**
T9	~24	**Vm Large**
T10	~23	**Vm Large**

Table 3: Comparison of WRR and Fuzzy Algorithm

No. of Workflows	No. of Missed Deadlines in WRR	No. Missed deadlines in Fuzzy Algorithm
10	2	3
15	3	3
20	2	3
25	5	4
30	8	6
35	9	6
40	6	4
45	5	3
50	8	5
55	15	12
60	10	8

The comparison table 3 shows that our algorithm works better when there are more number of tasks. Following graph 1 shows the results of comparison.

Graph 1 : No. of workflow v/s Deadlines



V. CONCLUSION AND FUTURE WORK :

Cloud computing provides services to its end users based on demand and users are charged as per the usage and as per the agreed terms of Service to satisfy QOS. Here as a part of our work we have tried to analyze the effect of using fuzzy logic in managing the VMs which are essential part of any cloud service. We have employed Least migration time as Migration policy. Using of fuzzy logic has been shown that it helps in meeting the time limit of real time when we have less number of VMs. This approach can be improvised with other rule based techniques. In future we can apply VM transfers or movement to reduce the amount of energy requirement. The effect of varying network bandwidth can also be considered which has not been emphasized in this research work.

REFERENCES

1. P. Mell and T. Grance, "The NIST Definition of Cloud Computing," US Nat'l Inst. of Science and Technology, 2011; http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.
2. Ehab NabielAlkhanak, Sai Peck Lee, Saif Ur Rehman Khan,"Cost-aware challenges for workflow scheduling approaches in cloud computing environments: Taxonomy and opportunities",Future Generation Computer Systems,ElseVier 50 (2015) 3–21
3. Marisol García-Valls, Tommaso Cucinotta, Chenyang Lu"Challenges in real-time virtualization and predictable cloud computing".
4. Mohammad A H, Monil and Rashedur M. R,"VM consolidation approach based on heuristics, fuzzy logic, and migration control" Journal of Cloud Computing: Advances, Systems and Applications (2016) 5:8DOI 10.1186/s13677-016-0059-7
5. Ehab NabielAlkhanak, Sai Peck Lee, Saif Ur Rehman Khan, "Cost-aware challenges for workflow scheduling approaches in cloud computing environments: Taxonomy and opportunities",Future Generation Computer Systems,ElseVier 50 (2015) 3–21
6. D. Chitra Devi and V. RhymendUthariaraj "Load Balancing in Cloud Computing Environment Using Improved Weighted Round Robin Algorithm for Nonpreemptive Dependent Tasks",The Scientific World Journal Volume 2016, Article ID 3896065,
7. Chun-Wei Tsai,Wei-chang Huang, M-S Chieng,Ming-chao Chiang and Chu-Sing Yang," A Hyper-heuristic Scheduling Algorithm for Cloud" ,IEEE transactions on cloud computing Vol-2, No2 April -June 2014.
8. M.M.M. Fahmy, "A fuzzy algorithm for scheduling non-periodic job on soft real-time single processor system" , Ain Shams Engineering Journal (2010) 1, 31–38
9. Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, César A. F. De Rose, and R Buyy," CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning", Algorithms Software: Practice and Experience (SPE), Volume 41, Number 1, Pages: 23-50, ISSN: 0038-0644, Wiley Press, New York, USA, January, 2011.
10. JawwadShamsi,· Muhammad Ali Khojaye · Mohammad Ali Qasmi, "Data-Intensive Cloud Computing: Requirements, Expectations, Challenges, and Solutions ", J Grid Computing (2013) 11:281–310 DOI 10.1007/s10723-013-9255-6
11. Brendan Jennings Rolf Stadler " Resource Management in Clouds: Survey and Research Challenges " J NetwSyst Manage DOI10.1007/s10922-014-9307-7
12. Fei Teng, Frédéric Magoulès · Lei Yu · Tianrui Li " A novel real-time scheduling algorithm and performance analysis of a MapReduce-based cloud ",J Supercomputing (2014) 69:739–765 DOI 10.1007/s11227-014-1115-z
13. Avtar Singh and Kamlesh Dutta " A novel real-time scheduling algorithm. and performance analysis of a MapReduce-based cloud" IEEK Transactions on Smart Processing and Computing, vol. 2, no. 6,December 2013.
14. Tom Springer, Steffen Peter Tony Givargis "Fuzzy Logic Based
15. Adaptive Hierarchical Scheduling for Periodic Real-Time Tasks", EWiLi'15, October 8th, 2015, Amsterdam, The Netherlands.
16. Kai Hwang , Xiaoying Bai , Yue Shi , Muyang Li, Wen-Guang Chen ,Yongwei Wu "Cloud Performance Modeling with Benchmark Evaluation of Elastic Scaling Strategies", IEEE transactions on parallel and distributed systems, vol. 27, no. 1, January 2016.
17. Jyothi sahni,Deo Prakash Vidyarthi," A cost effective deadline-
18. constrained dynamic scheduling algorithm for scientific workflow in a cloud environment", IEEE transactions on cloud computing, vol. 6, no. 1, January-march 2018.
19. Cingolani, Pablo, and Jesus Alcala-Fdez. "jFuzzyLogic: a robust and flexible Fuzzy-Logic inference system language implementation." Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on. IEEE, 2012.
20. M.A Rodriguez and RajKumar Buyya "Deadline Based Resource
21. Provisioning and Scheduling Algorithm for Scientific Workflows on Clouds", IEEE transactions on cloud computing, vol. 2, no. 2, april-june 2014
22. Rodrigo N. Calheiros, Rajkumar Buyya "Meeting deadlines of scientific workflows in public clouds with tasks replication" IEEE trasaction on parallel and distributed systems vol 25,No 7,July 2014
23. Andrei Tchernykh, Uwe Schwiegelsohn, Vassil Alexandrov, El-ghazali Talbi, "Towards Understanding Uncertainty in Cloud Computing Resource Provisioning", Procedia Computer Science Volume 51, 2015, Pages 1772–1781
24. RajaWasim Ahmad · Abdullah Gani "Virtual machine migration in cloud data centers: a review, taxonomy, and open research issues", J Supercomputing (2015) 71:2473–2515
25. DOI 10.1007/s11227-015-1400-5
26. Anton Beloglazov and Rajkumar Buyya," Managing Overloaded Hosts for Dynamic Consolidation of Virtual Machines in Cloud Data Centers under Quality of Service Constraints", IEEE transactions on parallel and distributed systems, vol. 24, no. 7, July 2013 .
27. Hai Zhong1, 2, Kun Tao1, Xuejie Zhang1, 2," An Approach to Optimized Resource Scheduling Algorithm for Open-source Cloud Systems ", in Fifth Annual China Grid Conference,2010.
28. Y. Yang, K. Liu, J. Chen, X. Liu, D. Yuan and H. Jin, An Algorithm in SwinDeW-C for Scheduling Transaction-Intensive Cost-Constrained Cloud Workflows, Proc. of 4th IEEE International Conference on e-Science, 374-375, Indianapolis, USA, December 2008.
29. Cui Lin, Shiyong Lu," Scheduling Scientific Workflows Elastically for Cloud Computing" in IEEE 4th International Conference on Cloud Computing, 2011.

AUTHORS PROFILE

Mrs R.A.Kulkarni is working as Asst.Prof in PICT,Pune. She is pursuing her PhD at JNTUH in the domain of Cloud Computing. Her main areas of

Dr.Suhas H Patil is currently working as professor in BVCOE, He is working as Research Paper Evaluator for World Scientific Engineering and Academic Society WSEAS, SCIEI and IASER. His main areas of interest are operating systems, Web technologies, Information systems etc. He has more than 224 research papers to his credit in reputed national and international journals and conferences. Member of International Society of Computer Associations (ISCA). Senior Member of Science and Engineering Institute (SSEI), Member of Editorial Review Board of International Journal of Intelligence Science (IJIS), Life Member, Numismatic Society of Kolkata, India (NCK). Member of Teacher Training and Education (MTTE), Education Management Professionals (EMP) and IEEE computer Society Members on LinkedIn (IEEE),member of ASDF(Association of Scientists not-for-profit organization focusing on research and development

Dr.N Balaji is currently working as Professor of ECE & Vice-Principal at JNTU Kakinada University, COE, Narasaraopet. His areas of interest are VLSI architecture and signal processing.