Analytic Review of Mathematical model for non-linear programming problem formulation: A novel Queuing Theory approach using stochastic cloudlet request evaluation in cloud computing environment

Srinivasa Rao Gundu[1], *Dr. T.Anuradha*[2]

Ph.D. Research scholar, Department of Computer Science, Dravidian University, AP, India [1]

Professor & Registrar, Department of Computer Science, Dravidian University, AP, India [2]

Abstract: ***Cloud computing is a new service made available to the end user with pay-and -use type, which are provided using the data centers, and with internet for providing different cloud based applications to the end users using subscription based, that are provided using sharable distributed computing resources concept. Generally these data centers are located at different locations and different end-users needs different requirements at the same time may get overloaded which is a challenging task. Quality of cloud service gets decreased due to the reason at the level of resources it is not possible to make decisive decision to provide better services. It is a study to investigate the application of Queuing Theory to the cloud computing and aimed evaluates the Stochastic Cloudlet requests which arrive at the Data centers. Analytic review is made and a Mathematical model for non-linear programming problem is formulated.***

Keywords: ***Cloud computing, pay-and -use, Data center, Sharable Distributed Computing Resources Concept (SDCRC), End-users, Cloud resources, Heuristic algorithm***

I. Introduction

In 1940's IBM started Mainframes, which are large computational units leveraging multiple processing. Mainframes are used for the purpose of processing huge amount of data, like ERP applications. These mainframe systems are always 'on' without any interruption in services. Modifications in mainframe applications are not a simple task in view of upcoming technologies like mobile computing. In the year 1969, 'Leonard Weinrock' as part of ARPANET, started internet. The Internet is a network of networks. Mainframes are followed by cluster computing in the late 1970s. Cluster computing is based on Amdahl's law. Clusters are low-cost alternatives to mainframes. These are commodity based. Cluster computing needs higher bandwidth networks. The clusters are controlled by specially designed tools. Clusters are the reason behind the evolution of cloud computing. Cluster computing also has some disadvantages such as programmability issues, problems in finding the faults. Clusters are difficult to handle by the non-professionals. Cluster computing is followed by Grid computing in the late 1990s. Grids belong to different organizations. They share computational power [1]. It is a dynamic aggregation of heterogeneous nodes. Grid computing also has some disadvantages such as; it needs fast internet connection between the computing resources. It has non-interactive job submission. Cloud computing technology has succeeded from the grid concept. Cloud computing systems are generally established in bigger data centres. These data centres would be hosted by a company.

The hosting company provides the services to the customers or the end users. Before the cloud computing technology came in to existence, the traditional business applications were existed. They are very complicated. They needed different types of Hardware, and Softwares to be installed, configure, test, run, secure and update periodically by a group of experts. In the recent days the cloud computing has shown a predominant growth in the area of internet based network technology. It is possible due to the reason of growth of communication technology. Cloud technology uses 'Sharable Distributed Computing Resources' (SDCR) which is also called as virtualization. SDCR Concept can be used for the massive usage of large scale internet based applications. Generally cloud computing is a distributed computing facility [2]. Distributed computing facility is provided to the third party. John McCarthy is the father of this model of computing. Today cloud computing is a common and major computing resource. Cloud computing tremendously changed the technology aspects worldwide. These applications are available to the end users because of sharable resources such as servers, storage of data, services. Amazon, Apple, Google, HP, IBM, Microsoft, Oracle are a few examples of cloud service providers.

II. Working Models of Cloud Computing

Cloud computing has two types of models. (A) Deployment model (B) Service model, they are

A. Deployment model: It refers to the way cloud computing is made available to the end user. Generally Cloud computing Deployment model is a control based infrastructure mode provisioning the services to the end user. There are four types of deployment models. They are

(1). Public deployment model: General public could access in this model. It can be made available at a lower cost. Ex: Google, Amazon, Microsoft. These services are available with the help of internet. Advantages of this model are, it will have less cost, reliable, flexible, location independent and highly scalable. It has disadvantages also such as it has less security, less privacy, and less control.

The real time examples of public models are i).Face book : which is a popular free social networking website which allows the registered users to create their profiles, later provides a facility to upload the photos, videos, exchange the messages. It provides a shared environment of their data with trusted members of other face book users like family, friends and colleagues. This technology came in to the existence on 4th February 2004 by its founder Mark Zuken berg [3].

(2). Private deployment model: This deployment model is available for fewer people such as private use. It takes more cost and good security can be provided. This deployment model is intended within an organization. Advantages of this model are high level privacy, more control. It has disadvantages also such area restricted, inflexible, and limited scalability. The real time examples of private models are bank cloud network which does not join the public cloud or a website which is hosted with any third party hosting provider are also the real time examples of private models.

(3).Hybrid deployment model: Amalgamation public and also private. Cost for this establishment is medium and the security it provides is moderate. Advantages of this model are scalable, flexible, cost efficient and secured. It has disadvantages also such as network issues, security compliance, and infrastructural dependency. The real time examples of hybrid models are i).Cloud stack: it is an open source resource for implementing the cloud services [4]. These are designed and deploy and manage large networks of virtual machines, highly scalable IaaS cloud computing platform.

(4).Community deployment model: The services of this community model are available to a community only. Establishment cost is medium. This is a secured model. An advantage of this model is less cost effective and more secured. It has disadvantages also such as data is accessible to others. Its management is very hard. The real time examples of community models are government services such as healthcare schemes.

B. Service models: refers the way the cloud computing provides its service to the end user. Generally Cloud computing Service model is a facility made available to the end user on demand via internet from a cloud provider's server. There are three types of service models. They are

(1).SaaS (Software as a Service): This would offer the software for the end user as a service. SaaS makes the software available over the internet. It is controlled by the cloud services provider or the one who sells it. Licenses can be subscribed to software, based on the requirements. Its applications are cost effective and available on demand. SaaS applications are automatically upgraded.

Ex: Salesforce.com, Billing systems, Invoicing systems, Customer relationship applications, Human resource solutions [5]. The advantage of this model is modest software tools are available which does not need to deploy. Efficient usage of software licenses, platform responsibilities are managed by the provider. Multi tenant solutions, centralized management, and data are also available along with this. It has the disadvantages also such as browsing level problems, network system level dependencies and portability among connected SaaS cloud environment. The real time examples of SaaS models are i).Salesforce.com: is a cloud computing and social enterprise SaaS provider founded in 1999.

(2). PaaS (Platform as a Service): models a platform as a service. It offers browser based development environment, provides built in security web services interfaces, it is very simple in order to interact with other application programs within the identical cloud computing technology platform, it would also provide web technological services to interface to connect other applications which are on other platforms[6].

Ex: Google App Engine. Generally, there will be less administration control and these are not portable. The benefits of PaaS architecture are less administration look up, lesser total cost ownership, solutions with scalability, latest system software. The PaaS disadvantages are less portability, the scheduling mechanism is event based, and less secure. The real time examples of PaaS models are i).Apprenda: is a PaaS provider founded in 2007 ii). Pivotal CF: is a PaaS provider. iii). Red hat open shift: PaaS provider from Red hat Company.

(3). IaaS (Infrastructure as a Service): provides access to basic computing resources for example computers, virtual machines, load balancers, IP addresses, and software bundles.

Ex: Amazon web services. The advantages of this service model is, full control over the resources, flexible, efficient renting of hardware, portable, and interoperable with legacy applications[7]. The disadvantage of this IaaS model is compatibility machines with preinstalled software, chances of virtual machine collapse, data erase practices are possible. The real time examples of IaaS models are i). Cisco Meta pod: is an IaaS provider from CISCO.

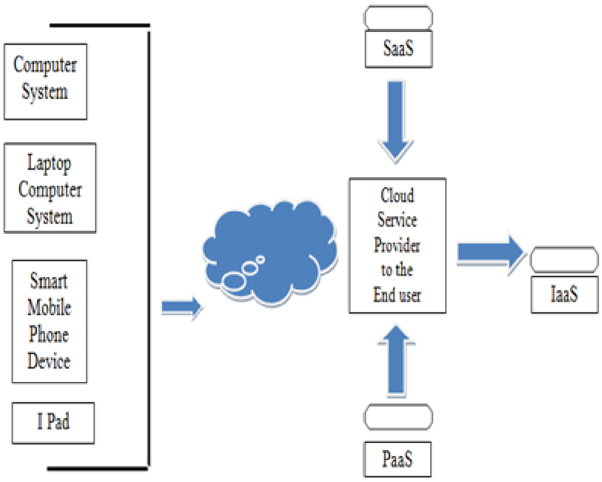

Fig. 1 Cloud computing working mechanism from user point of view

III. Cloud computing at Data centre level – current status

The data centre and cloud architectures are keeping on advancing, to address the requirements of expansive scale multi server farms in clouds.

i).In cloud computing the clients would have the capability to obtain to their applications, data from anywhere and whenever they need. It needs a distributed computing framework. Data will not be limited to a hard disk on the user's personal computers or enterprise server system.

ii).Cloud computing would reduce the equipment costs to a lesser level. Cloud computing architecture will require lesser devices on the user side. It does not require any fastest computer with more memory.

iii).Distributed computing frameworks would provide vast accessibility to personal computers applications. The companies need not buy software licenses of every time. The software can be used on a subscription base, provided by the cloud vendor. Generally the softwares are three types such as Free Softwares which are not needed to pay any amount for them to use, but cannot be modified, Ex: Google Apps. The other is Open softwares which are possible to modify and use them, Ex: Apache Tomcat Server. Third is proprietary softwares which are need to purchase and not possible to change, Ex: Ms Office.

iv).Servers and advanced hardware devices would consume up to room size. A few companies would lease a space to store the servers and their databases since the company does not need to have its accessibility on the location[8].

v).Software companies can reduce its huge investments on software and hardware by using the services of cloud data centres [9].

IV. Existing Queuing Theory Approach and the problems

Queuing theory and scheduling have both developed techniques to deal with dynamic resource allocation and sequencing. In the scheduling literature, dynamic problems are often handled by periodically solving static, deterministic scheduling problems. Such approaches apply existing scheduling methods to the combinatorics but are only able to reason about a short time horizon. Because the scheduling approaches typically solve an NP-complete problem during each scheduling period, it is assumed that the "time pressure" of the problem is low. That is, one time unit in the scheduling model represents a reasonable amount of "real time" and therefore it has the time to use optimization techniques to improve the schedule. With high time pressure, it is likely that online decisions must be made by a polynomial dispatching rule or policy.

The dynamic scheduling issues that they are involved in are characterized through a stream of jobs arriving stochastically over time. Each and every job requires a combo of assets, sequentially and/or in parallel, for different processing times. The existence of any designated job and its corresponding traits will not be known except its arrival. Nonetheless, it can be make the assumption, as in queuing conception however unlike in common scheduling settings, that it has the stochastic information of the distribution of job arrivals and traits. Jobs could require an intricate routing by means of to be had resources, which may be heterogeneous however with recognized and deterministic capacities. Such a scheduling trouble could have both short- and lengthy-time period targets.

Fixing dynamic scheduling issues is difficult each because of the combinatorics of the interaction of jobs, resources, and time, and due to the stochastic: to make a determination, one could use most effective the understanding that is recognized with simple task at a decision factor and the stochastic properties of situations that will arise at some point. Once the highest quality value of a controllable parameter is decided, it turns into a constant attribute of the queue.

Queuing manage issues, on the contrary, are dynamic: the intention in such problems is traditionally to investigate a choicest action to take when the method is in an exact state. In each queuing design and control problems, the goal is to search out premiere values for the controllable parameters of the queue. These parameters incorporate the number of machines to be had for processing arriving jobs, the buffer capability, the advent price of jobs to the queue(s), the service charges of the computing device(s), as well as any blend of these.

Furthermore, queuing thought ordinarily reviews systems with easy combinatorics, as such systems are more amenable to rigorous analysis of their stochastic methods. The study deals with formulation of optimization programming problems with the objective of exploring the decision Parameters which regulates the Virtual Machine. Computation of different stages continuous time dependent high capacity required user request incidence rates can be done with stochastic models.

These mathematical models have the limitations that they are formulated with the assumptions of handling multiple problems independently and separately. However high capacity required user request and less capacity required user request all are happened simultaneously. Keeping this research gap in mind some stochastic non-linear programming problems were formulated with the ratio of average normal average capacity required request from user.

Let A and B be two parameters that influence the Virtual Machines like throughput, makespan etc.

Let T- Transformation ratio of normal capacity required User request in to high capacity required user Request.

I-Initial number of user requests the needs less activity.

K-Initial number of user requests that needs high Capacity.

Then the objective function (M) which predicts the maximizing cases of high capacity required request. Let A and B be two parameters that influence the Virtual Machines like throughput, makespan etc.

$$M=I /(T-I/A-B)+(K-T I/A-B)$$

User request handling capacity, server capacity and the geographic location plays and important role. Mathematical Modelling studies are the meaning full alternatives to handle the Virtual Machines. The model construction should be stochastic models will provide the most relevant picture about the prevailing conditions. Stability is an important idea in queuing conception and types the predominant a part of the theoretical evaluation. Informally, a method is stable if its queues stay bounded over time.

The soundness of a system is dependent on the scheduling policy it employs: for a given set of hindrance parameters, one coverage could stabilize the approach whilst an extra might now not. Abilities of whether a process is steady for a given job arrival expense, processing price and scheduling policy is viewed a precursor to more distinctive evaluation and is essential for useful functions.

V. Simulation tool for the Investigation- Cloud sim

When a particular algorithm is designed and developed, it needs to be tested in a simulation environment. Cloud sim is such an environment which gives the user to test his/her web oriented applications to test. To make such a test, it needs to change many number of parameters and check again and again.

Cloud sim can be used to model data centers, host, service brokers, scheduling and allocation policies of a large scaled cloud platform. Cloud helps in VM provisioning at two levels: At the host level and at the VM level. At the host level, it is possible to specify how much of the overall processing power of each core will be assigned to each VM. At the VM level, the VM assigns a fixed processing power to the individual cloudlets. This is termed as VM Scheduling.

Every Datacenter should register with Cloud Information Service registry (CIS). The user requests are mapped to suitable cloud providers. The available list of cloud providers are issued by the CIS. Among the list, the cloud providers which are capable of providing the needed services are identified and a perfect matching provider is submitted with the application.

VI. Step wise Description of CloudSim Data Flow

Step 1 : Cloud user sends the job request to the User Interface

Step 2 : User Interface analyzes the request according to the Service Level Agreement (SLA).

Step 3 : User Interface assigns the tasks to cloud broker. Cloud broker is also termed as Data Center broker.

Step 4 : Cloud broker (Data Center broker) divides tasks into same sized cloudlets. Cloudlet is also termed as task units.

Step 5 : Cloud broker sends the cloudlets to Virtual Machine Manager (VMM).

Step 6 : Each data center entity registers with the Cloud Information Service (CIS) registry.

Step 7 : Cloud broker sends the resource request to Cloud Information service (CIS). Then Cloud broker consults the CIS to get hold of the record of resources which can recommend hardware services that match user's infrastructure requirements.

Step 8 : The cloud broker gets information about the availability of the data center and resources from the CIS.

Step 9 : Virtual machine manager (VMM) creates the virtual machine.

Step 10 : Data Center entity invokes UpdateVMProcessing for every host that manages it because concerned VMs manage each task unit. That's why the processes should be updated.

Step 11 : At the view point of host, invocation of UpdateVMProcessing triggers an UpdateCloudletProcessing process that orders every Virtual Machine (VM) to update its task unit status (executing, finish, suspend) with the data center entity.

Step 12 : VM analyze the approximate execution time and sends to host machine.

Step 13 : Host machine analyzes the smallest time to the next event.

Step 14 : Data center provides the information about the execution time and different resources such as Operating System, VMM used, RAM size, MIPS, a number of cloudlets, the number of CPU, storage capacity etc.

Step 15 : Request for execution of the cloudlet is sent to the virtual machine by VMM.

Step 16 : Cloudlet is being executed in the VM.

Step 17 : VM sends the executed cloudlets to the VMM.

Step 18 : After completing the execution, VM releases the resources for further use.

Step 19 : CIS updates the registry according to the information sent by the data center.

Step 20 : VMM further passes the executed cloudlets to cloud broker.

Step 21 : Cloud broker combines all the executed cloudlets together to form the task.

Step 22 : Cloud broker sends the completed task to the User Interface.

Step 23 : After completion of the task, User Interface can either expire the session or make another renewal request.

Step 24 : If the session is expired, then User Interface sends the executed task to the cloud user.

VII. Simulation Environment

There are two scheduling strategies defined, They are Time-shared scheduling policy and Space-shared scheduling policy. These scheduling strategies can be applied to both VMs and task units. CloudSim allows two level VM provisioning: firstly, at the host level, and secondly, at the VM level. At the host level, it determines the amount of processing power of each processing core required for each VM. And at the second level, each task unit is assigned a fixed amount of processing power by the VM. At each level, space-shared and time-shared policies are implemented in CloudSim. These two policies differ with the performance of the task units or application services. In this scenario, a host (having two processing cores) receives request for hosting two VMs in such a way that one VM needs two cores and executes four task units. VM1 is assigned to the task units namely T1, T2, T3and T4, and the second one is to host T5, T6, T7 and T8.

Space-shared policy:
Step 1: Accepted task units are arranged in the ready queue.
Step 2: Find out if there are any free processing cores available or not.
Step 3: If available, then task units are put in the execution queue.
Step 4: First task unit is assigned to the hosting VM.
Step 5: After completion of first task, the next task is assigned.
Step 6: If the queue is empty, it checks for the new task.
Step 7: Then it repeats from the step 1.
Step 8: End

Note 1: As each VM requires two cores, only one VM can run at a given instance of time. Only after completion of VM1, VM2 can be assigned the cores.
Note 2: Each task unit requires only one core.

Time-shared policy:
Step 1: Accepted task units are arranged in the ready queue.
Step 2: All the task units are simultaneously assigned to VM. These task
units are dynamically context switched during their life cycle.
Step 3: When the queue is empty, it checks for the new task.
Step 4: Then it repeats from the step 1.
Step 5: End

Note: Each VM gets a time slice on every individual processing core and the time slices are distributed among the task units.

In Time-shared Policy , VMs are allocated on the basis of space-shared policy and within a VM, task units are allocated to the processing cores on time-shared basis. As mentioned before, since each VM requires two cores, only one VM can run at a given instance of time. Only after completion of VM1, VM2 can be assigned the cores (Space-shared). Therefore, during a VM lifespan, all the assigned task units are dynamically context switched during their life cycle (Time-shared).

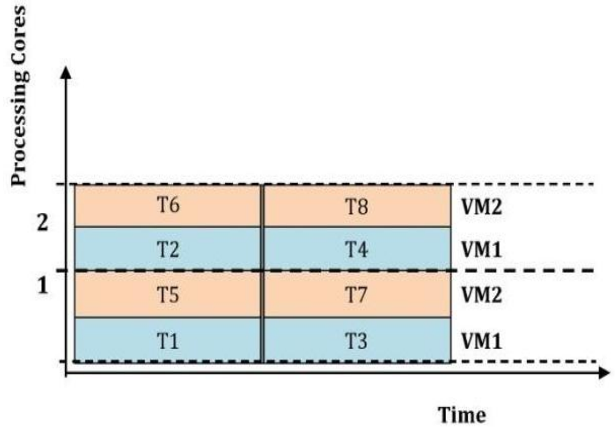

Fig. 2 VM in time-shared and Cloudlet in space-shared

In Space -Shared Policy, VMs are allocated in the basis of time-shared policy and within a VM, task units are allocated to the processing cores in space-shared basis. Each VM gets a time slice on every individual processing core and the time slices are distributed among the task units (Time-shared). As the processing cores are shared, VMs are getting fewer amounts of processing power. But, at any given instance of time, only one task can be allocated to each processing core (Space-shared).

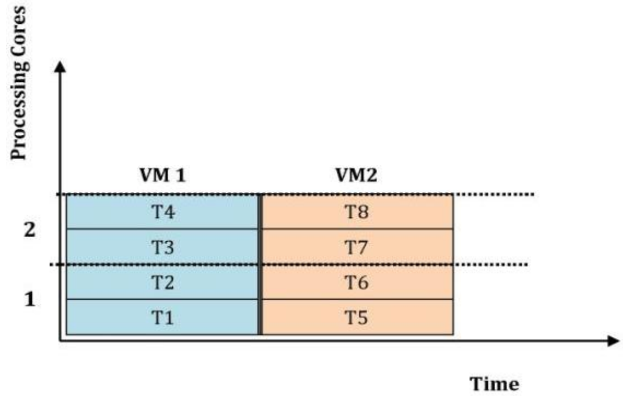

Fig. 3. Shared and Cloudlet in time-shared

It is needed to compare the average completion time of the task units.

It is configured the simulation environment as stated above. In first case, VMs are allocated in space-shared policy and task units are in time-shared policy. Secondly, VMs are in time-shared policy and task units are in space-shared policy. After the creation of VMs, the task units are submitted, and VMs and task units are arranged accordingly. The tests were conducted on a 32-bit Intel Core i5 machine having 2.60 GHz and 3 GB RAM running windows 7 Professional and Java Development Kit 8 (32 bit).

The main goal of the experiment is to compare the average completion time of the said scheduling scenario and to find out the minimum average completion time. In this experimental set up, simulation creates a data center Broker dynamically and also subject to other constraints this simulation is done. The test simulation environment for measuring the average task completion time includes data center, VMs, and task units (cloudlets). During simulation, a data center is created, which has the characteristics like x86 of architecture, Linux as operating system, Xen as VMM. The simulation environment consists of two hosts; each host has been modeled to have 1000 MIPS, 2 GB of RAM memory, and 1 TB of storage. VM runs inside a host, sharing the host list (the number of hosts present in the simulation) with other VMs. It processes the cloudlets. This processing happens according to a policy (time-shared or space shared), defined by the cloudlet scheduler. The host can submit cloudlets to the VM to be executed. According to the requirement, the following inputs have been taken to create the VMs and cloudlets.

IV. Conclusion

The field of queuing thought that offers with prescriptive items is regularly referred to as the design and control of queuing techniques. The stochastically arriving user requests and allocation of tasks are needed to specify how time assignments should be made at any time when a new job arrives or completes processing. Nevertheless, the agenda as a consequence of this type of policy, while being of fine first-class in expectation, may be far from most suitable for the distinctive recognition of uncertainty that occurs. A formulation is made to fix the issue of Queuing design issues using a stochastic model.

ACKNOWLEDGMENT

I sincerely thank and express deep sense of gratitude to my Research supervisor Prof. T. Anuradha (Registrar Dravidian University)for her Inspiration and Mentorship for this Paper.

References

[1] Dr. P. Neelakantan,"A Study on E-Learning and Cloud Computing" , International Journal of Scientific Research in Computer Science, Engineering and Information Technology(IJSRCSEIT), ISSN : 2456-3307, Volume 3 , Issue 1, PP 1534-1539

[2] Cloud Computing Protected: Security Assessment Handbook,John Rhoton,January 2013, International Journal of Computer Sciences and Engineering Vol.6, October 2018, E-ISSN: 2347-2693, 2018, IJCSE All Rights Reserved 260

[3] Cloud Computing Building the Infrastructure for Cloud Security,Raghuram Yeluri,March 2014

[4] Cloud Computing Scheduling and Isolation in Virtualization A study on scheduling and isolation for data centers and cloud computing,Gaurav Somani,September 29 2010

[5] Cloud Computing Management Strategies For The Cloud Revolution How Cloud Computing is Transforming Business and Why You Can't Afford to Be Left Behind,Charles Babcock May 22 2010

[6] Enterprise Cloud Computing A Strategy Guide for Business and Technology Leaders,Peter Fingar , Andy Mulholland & Jon Pyke,April 23 2010

[7] Cloud Computing Above the Clouds Managing Risk in the World of Cloud Computing,Kevin T McDonald,February 23 2010

[8] Mithilesh Mittal, Pradeep Sharma and Pankaj Kumar Gehlot, ―A Comparative Study of Security Issues& Challenges of Cloud Computing, International Journal of Scientific Research in Computer Science and Engineering, Vol.1, Issue.5, pp.9-15, 2013.

[9] S. Renu, OHH Parveen, ―Biometric Based Approach for Data Sharing in Public Cloud, International Journal of Advanced Research in Computer and Communication Engineering Vol.4, Issue.2 pp.1-9, 2015

BIOGRAPHY

First Author (Corresponding Author) Mr.G.Srinivasa Rao pursed Bachelor of Science from Osmania University of Hyderabad, India in 2006 and Master of Computer Applications from Osmania University in year 2013. He is currently pursuing Ph.D. in the Department of Computer Sciences, Dravidian University, Kuppam since 2014. Thesis Submitted – Awaiting for VIVA VOCE Examination. Presently working as Senior Software Trainer as Re-joined and Continued after submission of Ph.D. Thesis from Feb- 2019 to till date. He is a member of Internet Society Global Member, also the Member of International Association of Engineers. Presently Associate Reviewer & Member of the International Board of Reviewers International Journal of Community Development and Management Studies (IJCDMS), INFORMING SCIENCE INSTITUTE. His main research work focuses on Load balancing, Cloud computing. His Research Interests are Cloud Computing, Mobile Cloud Computing, Data Mining, Big Data, Artificial Intelligence and Internet of things. His publication record as follows, published paper in relevance to the current research in Web of Science -1 Paper, and published papers in UGC Approved Journals -3, published papers in IEEE Conferences in relevance to other topic-2, Published papers in relevance to other topic- 1, and attended 14 Conferences.

Second Author (Research Supervisor)

Prof. T. Anuradha has pursued Master of Computer Applications from Sri Padmavathi Mahila Viswa Vidyalayam, Tirupati. She has pursued her Ph.D. from Sri Padmavathi Mahila Viswa Vidyalayam, Tirupati. She is working as the Professor in the Department of Computer Science, and Presently she is the Registrar of Dravidian University. Her Areas of Research Interests are Data Mining & data warehousing, Neural Networks, Cloud computing, Wireless Sensor Networks. Under her Guidance 2 M.Phil Degrees are awarded and 1 pursuing, 2 Ph.D s are awarded ,5 pursuing and 1 submitted,. She has the Publication record of 17 International Journals, 2 International Conference Proceedings , 4 International Conferences , 8 National Conference Proceedings , 3 National Conferences. She Organized 2 Conferences.