International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 6 (2015) pp. 14421-14432

http://www.ripublication.com

An Overview of Approaches and Application of Cloud Computing

Prasanta Kumar Bal[1],Manoranjan Pradhan[2], Sateesh Kumar Pradhan[3]
[1,2]Department of Comp. Sc. & Engg., GITA, Bhubaneswar, India
[1]prasanta_bal@yahoo.co.in
[2]manoranjanpradhan72@gmail.com
[3]Department of Computer Science, Utkal University, India
sateeshind@yahoo.com

Abstract

Now a days a great deal of attention is being paid to cloud computing both in research and among end users. Since cloud computing is cost effective, highly available and easily scalable, most of enterprises are in their way to reduce their computing cost through the means of virtualization. Cloud Computing offers better computing through improved utilization and reduced administration and infrastructure costs. With the advent of cloud computing, resources are used as an aggregated virtual computer. This paper provides an overview of some basic concepts of cloud computing and the most popular cloud computing approaches and its applications with strength and limitation.

Keywords: Cloud Computing, SaaS, IaaS, PaaS, Virtualization.

Introduction

From the past few years, there has been a rapid progress in Cloud Computing. Cloud Computing delivers a wide range of resources like computational power, computational platforms, storage and applications to users via internet. The major Cloud providers in the current market segment are Amazon, Google, IBM, Microsoft, Salesforce, etc.. . Cloud Computing has successfully managed to advertise itself as one of the fastest growing service models on the Internet. Many large scale IT providers, such as Amazon [1] and IBM [2], share their data centers, through virtualization concepts, for the public consumption of their computational resources. As a result, cloud consumers can minimize many startup financial overheads plus receive an increase in the availability and scalability for their cloud-hosted applications. Moreover, cloud consumers can enjoy on-demand service with the ease of Pay-As-You-Go subscription. The National Institute of Standard and Technology (NIST) defines Cloud Computing as the model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., Networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction [3]. This paper describes an overview of approaches and applications of cloud computing. Here cloud computing is providing the technological underpinnings for new ways to collect, process, and store massive amounts of information. Section 2 describes the overview of cloud computing. Section 3 describes the cloud computing approaches. Section 4 describes the strength and limitation of cloud computing approach. Section 5 provides the concluding remarks and section 6 presents future research work.

Overview of Cloud Computing

Cloud is a computing model that refers to both the applications derived as services over the Internet, the hardware and system software in the datacenters that provide those services. Cloud Computing is treated as the high potential paradigm used for deployment of applications on Internet. This concept also explains the applications that are broaden to be accessible through the Internet. Cloud applications use large data centers and effective servers that host web applications and services. Each of the current definitions of cloud systems addresses cloud systems from a distinct perspective. According to NIST (National Institute of standards and Technology) [4], "Cloud computing ('cloud') is an evolving term that describes the development of several existing technologies and approaches to computing into something different. Cloud separates application and information resources from the underlying infrastructure, and the mechanisms used to deliver them". According to Ian Foster et al. [5], "Cloud computing is a large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the Internet". There are a few key points in this definition. First, cloud computing is a specialized distributed computing paradigm; it differs from traditional ones in that

- It is massively scalable.
- Can be encapsulated as an abstract entity that delivers different levels of services to customers outside the cloud.
- It is driven by economies of scale; the services can be dynamically configured via virtualization or other approaches and delivered on demand.

NIST defines cloud computing in terms of five essential characteristics, three service models, and four deployment models. They are summarized in visual form in Figure 1 and explained below as in [4, 5].

Essential Characteristics of Cloud Computing

Cloud services exhibit five essential characteristics that demonstrate their relation to, and differences from, traditional computing approaches [4]:

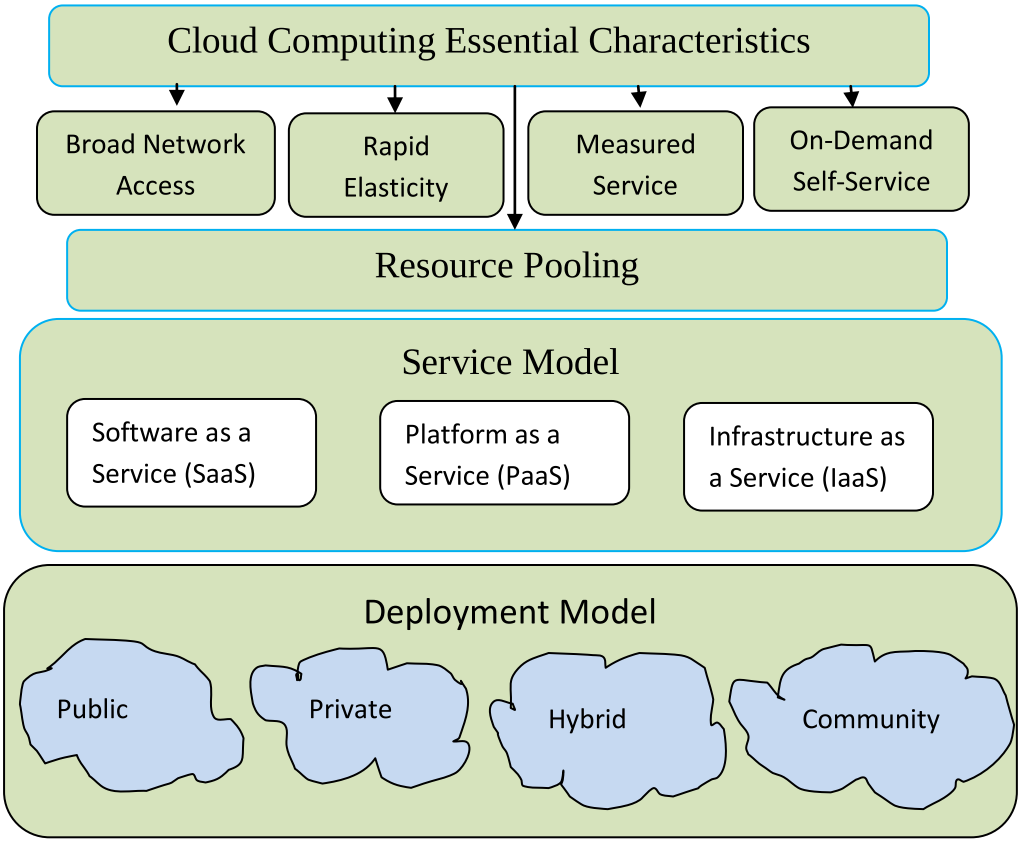

Figure 1: NIST Visual Model of Cloud Computing Definition

On-Demand Self-Service
A consumer can unilaterally provision computing capabilities as needed and automatically, without human interaction with a service provider.

Broad Network Access
Computing capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms (e.g. mobile phones, laptops, and PDAs) as well as other traditional or cloud based software services.

Resource Pooling
A provider pools computing resources to serve several consumers using a multi-tenant model, which dynamically assigns and reassigns physical and virtual resources according to consumer demand. There is a degree of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources.

Rapid Elasticity
Capabilities can be rapidly and elastically provisioned, in most cases automatically and rapidly released to quickly scale out and scale in. For a consumer, the capabilities appear to be unlimited and can be purchased in any quantity at any time.

Measured Service
Cloud systems automatically control and optimize resource usage by leveraging a metering capability according to the type of service. Usage can be monitored, controlled, and reported, providing transparency for both the provider and the consumer.

Cloud Service Models
In general, clouds offer services at three different levels [5]: IaaS, PaaS, and SaaS. However, some providers can expose services at multiple levels. According to NIST, the cloud model is composed of three service models:

Software as a Service (SaaS): The capability provided to the consumer is to use the provider’s applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser (e.g., web-based email), or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings. SaaS delivers software that is remotely accessible by consumers through the Internet with a usage-based pricing model. E.g., Live Mesh from Microsoft allows files and folders to be shared and synchronized across multiple devices.

Platform as a Service (PaaS): The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment. PaaS offers a high-level integrated environment to build, test, and deploy custom applications as in Google’s App Engine [6]. Inside this layer resides the middleware system, a portable component for both grid and cloud systems. Examples include WSO2 Stratos [7], Windows Azure [8], and our middleware HIMAN [9, 10, and 11].

Infrastructure as a Service (IaaS): The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components (e.g., host firewalls) [12].

Infrastructure as a Service (IaaS) provisions hardware, software, and equipments to deliver software application environments with a resource usage-based pricing model. Infrastructure can scale up and down dynamically based on application resource needs. Typical examples are Amazon EC2 (Elastic Cloud Computing) Service [13], Eucalyptus [14], Microsoft Private Cloud [15].

Cloud Deployment Models

There are four deployment models for cloud services, with derivative variations that address specific requirements: According to NIST, the cloud model is composed of four deployment models which is stated in figure 2.

Private cloud: The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises [12].

Community cloud: The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises [12].

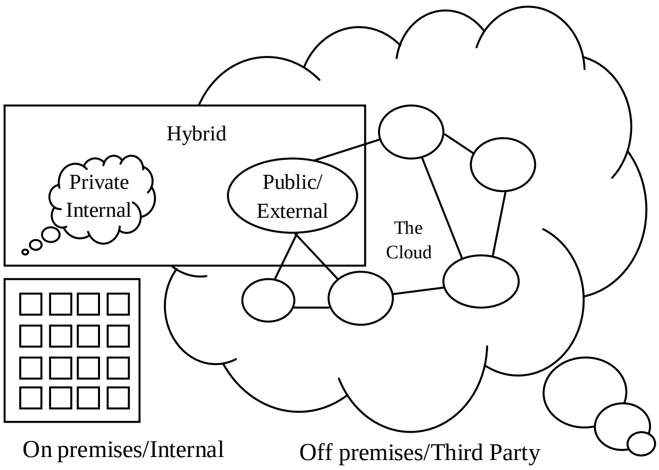

Figure 2: Cloud Deployment Models

Public cloud: The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider [12].

Hybrid cloud: The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds) [12].

Cloud Computing Approaches

The Google Approach to Cloud Computing

All services and applications of Google is based on its internally developed cloud-based computing infrastructure. Following is the description of how Google has built and how they use it. Rather than building a system from a moderate number of very high-performance computers, Google builds their cloud hardware as a cluster containing a much larger number of commodity computers[16]. They can optimize their costs, power consumption, and space needs by making appropriate tradeoffs. Another key aspect of Google's design philosophy is optimization of their system software for the specific applications they plan to build on it. Because Google is building both the system and application software, they know what their applications require and accordingly design their system software on meeting exactly those requirements.

Google File System

The core of Google's design philosophy is the architecture of the Google File System (GFS)[17]. GFS serves as the foundation of Google's cloud software stack and resembles a Unix-like file system to its users. GFS is optimized for storing very large files of size more than 1 GB because Google's applications typically manipulate files of this size. One way that Google implements this optimization is by changing the smallest unit of allocated storage in the file system from the 8 KB block size to 64 MB. Using a 64 MB block size results in much higher performance for large files at the expense of very inefficient storage utilization for files that are substantially smaller than 64 MB. Another important design choice Google makes in GFS is to optimize the performance of the types of I/O access patterns they expect to see most frequently. A typical file system is designed to support both sequential and random reads and writes reasonably efficiently. Because Google's applications typically write a file sequentially once and then only read it, GFS is optimized to support append operations.

Map Reduce

Layered with on top of GFS, Google's MapReduce framework is the heart of the computational model for their approach to cloud computing [18,19]. The basic idea behind Google's computational model is that a software developer writes a program containing two simple functions—*map* and *reduce*—to process a collection of data. Google's underlying runtime system then divides the program into many small tasks, which are then run on hundreds or thousands of hosts in the cloud. The runtime system also ensures that the correct subset of data is delivered to each task. The detail architecture of MapReduce is beyond the scope of this paper.

Big Table

The last major component of Google's approach to cloud computing is their Bigtable data storage system [20]. In many respects, Bigtable superficially resembles a relational database management system (RDBMS). Both store data in tabular form with labeled rows and columns, and they allow data to be searched using the row name (and possibly the column name) as keys. Both also allow new data to be added, data in existing rows to be updated, and data to be deleted.

Google claims that Bigtable can store petabytes of data across thousands of servers [20]. In this same paper, Google reported that their Google Maps imagery data consumed 70 TB of space in Bigtable. They also reported that their Crawl Project (not described in the paper) data consumed 800 TB of space in a Bigtable table.

Hadoop

After publishing a series of scientific papers describing their approach to cloud computing and their successful experiences using it [16, 19, 20], the approach generated a great deal of interest and enthusiasm outside of Google. Taking cue from these papers, the open source software development community created an implementation of Google's cloud approach called Hadoop [21]. Hadoop is part of the Apache open source project and contains an implementation of MapReduce and a GFS-like distributed file system called Hadoop Distributed File System (HDFS). HBase, a related Apache project, is an open source implementation of Bigtable [22]. As with GFS, HDFS stores large files across large clusters in sequences of blocks. Replication is also included within HDFS. Hadoop MapReduce and HDFS, as well as HBase, employ master/slave architectures very similar to the approach Google took in designing its corresponding systems. Unlike Google's systems, which are written in C++, Hadoop and HBase are written in Java.

Amazon Approach To Cloud Computing

Though best known for selling books online, Amazon is also actively provides services that allow developers to take advantage of their computing technology. Amazon Web Services provides developers use of open APIs to access Amazon's vast infrastructure . By using these APIs, developers can create interfaces and access the computing infrastructure provided by Amazon on a fee-based schedule. Software developers, start-up companies, and established companies in need of reliable computing power are members of a large and growing crowd using Amazon services. Amazon Elastic Compute Cloud or EC2 [23] is one such service. The EC2 provides virtualization for developers to load Amazon-managed machines with their preferred software environments. The developers first create an Amazon Machine Instance (AMI) with the operating system, custom configuration settings, libraries, and all needed applications. After AMI is created , it is loaded into the Amazon Simple Storage Service (AS3) and receives a unique identifier. The unique identifier can then be used to run as many instances of the AMI as needed using Amazon's APIs. Amazon also provides a set of prebuilt AMIs that can be used by developers. AMIs can be sized to the requirements of individual applications. AMIs fall into categories ranging from a small instance to an extra-large instance. A small instance has less memory, virtual cores, storage, and I/O performance than a large one. Similar to a timesharing system, Amazon bills users by the instance-hour. As the size of memory, number of cores, or other features increases, the instance-hour fee increases.

Microsoft Approach To Cloud Computing

Microsoft's cloud environment Azure Services Platform was launched in October 2008 [24, 25]. It provides computing and storage resources for consumers to develop and host applications, Microsoft is also offering cloud applications that are already developed and ready for consumption. The Azure Service Platform is built on the Windows Azure cloud operating system, which provides a development, hosting, and management environment for cloud applications. Services including Live Services, SQL Services and .NET Services are available on top of the Azure operating system. Potential users can also download an Azure SDK and Azure tools for Microsoft Visual Studio to simulate the Azure framework during the preview period. Once Azure is launched for commercial use it will be priced using a consumption-based model. Consumption will be measured in compute time, bandwidth, and storage and transactions (put and gets). Microsoft is using a combination of Microsoft .NET framework and the Microsoft Visual Studio development tools to provide a base for developers to easily launch new solutions in the cloud. It is noted that both applications running in the cloud and outside of the cloud can use the Azure cloud platform.

Other Cloud Computing Approaches and Applications

Besides Amazon, Google, and Microsoft other organizations including Dell, IBM, Oracle, and some universities. IBM is providing a variety of cloud-based services by using existing functionality and capabilities of the IBM Tivoli portfolio [26]. Tivoli is a collection of products and software services that can be used as building blocks to support IBM Service Management software. IBM's cloud-based services, which target independent software vendors (ISVs), offer design of cloud infrastructures, use of worldwide cloud computing centers, and integration of cloud services.

Strength and Limitation of Cloud Computing Approach

Since cloud computing technique gains attention from enough people ,it is seen as solution to every problem. But the fact is that some situation warrants the existing technology better than what cloud provides. As long as transaction model is concerned , the RDBMS is a better one than Google's Bigtable. Similarly the Amazon approach to cloud computing is ideal for small organizations or organizations with unpredictable computing usage requirements. For large organizations or organizations that process particularly sensitive data, this approach may not make sense. With the Amazon approach, a user is effectively renting computing resources. Renting computing resources may not be the most cost effective use of funds for a large corporation. As an organization grows in size and importance, the *value* of its data also increases dramatically. Cloud computing approaches use parallelism to improve the computational performance of applications. The Google MapReduce framework is particularly good at this so long as the problem fits the framework. Other approaches to high performance computing have similar constraints. It's very important for developers to understand the underlying algorithms in their software and then match the algorithms to the right framework. Single-threaded software will not run faster on a cloud, or even on a single computer with multiple processing cores, unless the software is modified to take advantage of the additional processing power. Along Many problems cannot also be easily broken up into pieces that can run independently on many machines.

Conclusion

The cloud computing approaches described in this paper have been applied successfully to a wide range of problems and also we have provided some strength and limitation of each approach. As it gains momentum, range of application area of cloud computing will increase. Like other approaches to high performance computing, cloud computing is providing the technological underpinnings for new ways to collect, process, and store massive amounts of information. Based on what has been demonstrated thus far, ongoing research efforts, and the continuing advancements of computing and networking technology, we believe that cloud computing is poised to have a major impact on our society's data centric commercial and scientific endeavors.

Future Research Work

Undoubtedly cloud provides enough computing facilities and much progress has already been made in this field, still there are a number of research areas that still need to be explored. One of the most important issues of security and vulnerabilities should be addressed by using Intrusion Detection System.

References

[1]. Amazon Web Servies, \Amazon Elastic Compute Cloud, (Amazon EC2)." http://aws.amazon.com/ec2/. Accessed: Jan 2015.
[2]. IBM, \IBM Smart Cloud." http://www.ibm.com/cloud-computing/us/en/index.html. Accessed: Jan 2015.
[3]. NIST, \Cloud Computing Synopsis and Recommendations." http://csrc.nist.gov/publications/nistpubs/800-146/sp800-146.pdf, Jan 2015. Accessed: Jan 2015.
[4]. "Security Guidance for Critical Areas of Focus in Cloud Computing", Cloud Security Alliance, http://www.cloudsecurityalliance.org/ guidance/csaguide.pdf
[5]. Foster, I.; Yong Zhao; Raicu, I.; Lu, S., "Cloud Computing and Grid Computing 360-Degree Compared", Grid Computing Environments Workshop, 2008. GCE '08, vol., no., pp.1-10, 12-16 Nov. 2008.
[6]. Google App Engine, http://code.google.com/appengine/, 2008.
[7]. http://wso2.com/cloud/stratos/, June 2013.
[8]. "Windows Azure System" http://www.microsoft.com/windowsazure/ windowsazure/

[9]. Ali E. El-Desoky, Hisham A., Abdulrahman A. Azab, "A Pure Peer-to-Peer Desktop Grid Framework with Efficient Fault Tolerance", ICCES 24, Nov. 2007.
[10]. Abdulrahman A. Azab, Hisham A Kholidy, "An Adaptive Decentralized Scheduling Mechanism for Peer-to-Peer Desktop Grids", (The 2008 International Conference on Computer Engineering & Systems) 25-27 Nov, 2008
[11]. Abdulrahman Azab, Hein Meling, "Broker Overlay for Decentralized Grid Management", Journal of Computer Science And Technology, January 31, 2010.
[12]. Peter Mell. (2011) 'The NIST Definition of Cloud ', Reports on Computer Systems Technology, sept., p. 7.
[13]. Amazon Elastic Compute Cloud, http://aws.amazon.com/ec2, 2008.
[14]. Eucalyptus, http://eucalyptus.cs.ucsb.edu/, 2008.
[15]. Microsoft Private cloud, http://www.microsoft.com/en-us/server-cloud/private-cloud/default.asp
[16]. Barroso LA, Dean J, Hölzle U. Web search for a planet: the Google cluster architecture. IEEE Micro. March-April; 23(2), 2003.
[17]. Ghemawat S, Gobioff H, Leung S. The Google filesystem. Proceedings ACM Symposium on Operating Systems Principles. October, 2003.
[18]. Dean J and Ghemaway S. MapReduce: Simplified data processing on large clusters. Proceedings Operating Systems Design and Implementation. December, 2004.
[19]. Dean J and Ghemaway S. MapReduce: Simplified data processing on large clusters. Communications of the ACM. January; 51(1), 2008.
[20]. Chang F, Dean J, Ghemawat S, Hsieh WC, Wallach DA, Burrows M, Chandra T, Fikes A, Gruber RE. Bigtable: A distributed storage system for structured data. ACM Trans. on Computer Systems. June; 26(2), 2008.
[21]. Hadoop Project homepage. http://hadoop.apache.org/core, 2008.
[22]. HBase Project homepage. http://hadoop.apache.org/hbase, 2008.
[23]. Amazon Web Services. Amazon Elastic Compute Cloud homepage. http://aws.amazon.com/ec2,November 2008.
[24]. Chappel D. Microsoft Azure homepage. Introducing the Azure Services Platform [Internet]. http://www.microsoft.com/azure/whatisazure.mspx, October 2008.
[25]. Microsoft Azure homepage. http://www.microsoft.com/azure, 2008.
[26]. IBM homepage. IBM launches cloud services initiative [Internet]. http://www-03.ibm.com/press/us/en/pressrelease/25341.wss, 2008.

Authors' Profile :

Prasanta Kumar Bal is a Ph. D. scholar in Computer Science under Utkal University, Bhubaneswar, India. He is presently working as an Asst. Professor in the Department of Computer Science & Engineering, Gandhi Institute for Technological Advancement (GITA), Bhubaneswar, Odisha, India. His research interests include Computer Security, Intrusion Detection, Data Mining and Cloud Computing.

Manoranjan Pradhan holds a Ph.D. Degree in Computer Science. He is presently working as a Professor and Headof the Department of Computer Science & Engineering, Gandhi Institute for Technological Advancement (GITA), Bhubaneswar, Odisha, India. He has sixteen years of teaching experience. He has published many papers in national and international journals. His research interests include Computer Security, Intrusion Detection, Soft Computing and Cloud Computing.

Sateesh Kumar Pradhan obtained his Ph.D. Degree in Computer Science from Berhampur University, India during the year 1999. He joined Berhampur University, as Lecturer in the year 1987 and promoted to Reader in 1999. He was Head, Post Graduate Department of Computer Science, Utkal University, India during 2001-2003. He was the Organizing Chair of the International Conference on Information Technology-2005. He served as a senior faculty in the Computer Engineering Department of King Khalid University, Ministry of Higher Education, Saudi Arab from September 2006 to July 2011. At present he is the Head, Post Graduate Department of Computer Science, Utkal University, Bhubaneswar, India. His research interest includes Neural Computing, Computer Architecture, Ad hoc Networks, Computer Forensic.