DOI:10.10543/f0299.2020. 41813

THE STATE OF CLOUD COMPUTING AND RESEARCH CHALLENGES

[#1]KOTAWAR SAVITHA, ***Assistant Professor,***

Department of Computer Science and Engineering

[#2]SHASHIKANTH ALLENKI, ***Assistant Professor,***

Department of Computer Science and Engineering,

MOTHER THERESA COLLEGE OF ENGINEERING AND TECHNOLOGY, PEDDAPALLY, TS.

ABSTRACT: Cloud computing is a new concept for internet storage and distribution. Cloud computing enables even small organizations to start small and grow as needed. Cloud computing, despite its youth, has enormous IT potential. Fundamentals of cloud computing, architecture, implementation, and future research are all discussed. We investigate cloud computing design issues and provide additional research into this emerging subject.
Keywords: Cloud computing · Data centers · Virtualization

1. INTRODUCTION

The Internet, as well as other developments in processing and storage, have made computing resources more accessible, powerful, and affordable. Because customers may borrow and return servers and data centers online, cloud computing is conceivable. Clients can control cloud platforms and assign resources as needed with service provider leases. Cloud computing was pioneered by companies such as Google, Amazon, and Microsoft. The advantages of cloud computing for enterprises are listed below.

Cloud services are billed based on consumption. Service providers can leverage the cloud without the need for servers. This cloud rents computer power.

Because it rapidly allocates and releases computing resources, cloud computing is inexpensive. As a result, service providers no longer require peak capacity. Due to low service demand, resources may be shifted to less expensive places. Infrastructure companies provide scalable data centers. The "flash crowd" demonstrates how quickly service providers may scale up to meet demand spikes. Some terms necessitate more figures.

The majority of cloud services are accessible online. They can be found in a variety of comparable gadgets. Electronic gadgets can be adjusted. Outsourcing cloud-based service infrastructure allows businesses to reduce hardware failure risks and maintenance costs. Businesses will save money on hardware upkeep and employee training.

Cloud computing benefits the IT business, but it also introduces new issues. The framework for cloud computing, use cases, and research gaps are investigated. We'll look at cloud computing design difficulties and make recommendations for further research on this fascinating area.

2. OVERVIEW OF CLOUD COMPUTING

Key cloud computing ideas are explained.

Definitions

Cloud storage is not a new concept. In the 1960s, John McCarthy saw computers as a public good. The "cloud" of the 1990s was mega-ATM networks. Google CEO Eric Schmidt coined the term "web 2.0" in 2006 to describe Google's online business model. Since then, the term "cloud computing" has been used to promote a variety of projects. The lack of a clear definition of cloud computing has created market excitement, misunderstanding, and ambiguity. We recently completed cloud computing standards. Over 20 definitions from various sources were examined in order to find common ground. We like the NIST's complete definition of "cloud computing."

Clouds provides "access to a networked pool of

ISSN: 1005-0299

configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal administration effort or service provider interaction."

"Cloud computing" may signify more than most technical terminology because it shows a new business model that makes use of existing technologies. Virtualization and utility pricing are examples of previous cloud innovations. These technologies allow cloud computing to meet modern IT efficiency and cost requirements.

Related technologies

These technologies are similar to cloud computing in the following ways: Networks are used in the computer model. Grid computing was invented because scientific computing required so much processing power. In cloud and grid computing, distributed servers carry out tasks. Through hardware and software virtualization, cloud computing enables dynamic resource provisioning and sharing.

Pay-as-you-go resources are available with utility computing. Cloud computing is a type of utility computing. Using cost-of-service pricing for efficiency. With utility-based pricing and on-demand provisioning, service providers can maximize resource utilization while lowering costs.

Virtual machines can both conceal and reveal hardware. Virtualization of mission-critical applications. Computers are used in VR servers. Cloud computing may dynamically assign device pool resources to programs thanks to virtualization.

In 2001, IBM pioneered autonomic computing by developing machines that respond to both internal and external stimuli. Automation software makes it easier to manage computing resources. The cloud, rather than simplifying infrastructure, delivers resources independently, saving money.

Cloud computing is available on demand and via the Internet. Parallel computing that is distinct from grid and autonomous computing. There are advantages and disadvantages to meeting its standards.

3. CLOUD COMPUTING ARCHITECTURE

The architectural, commercial, and operational aspects of cloud computing are examined.

A layered model of cloud computing

The diagram depicts the hardware/datacenter, infrastructure, platform, and application layers of cloud computing. 1. For your convenience, breakdowns are provided below.

Hardware manages the servers, networks, electrical infrastructure, and climatic systems. Machines are stacked in data centers. Data centers use switches, routers, and fabric to connect thousands of rack-mounted servers. Hardware factors include setup, fault tolerance, traffic, energy, and temperature control.

The virtualization layer pools storage and compute resources from physical servers using Xen, KVM, and VMware. Because virtualization enables dynamic resource allocation and other cloud computing characteristics, the infrastructure layer is critical.

The top platform layer of an infrastructure is made up of the operating system and application frameworks. Platform layer programs are accessed by VM containers. Google App Engine's platform-layer APIs can be used by web applications for storage, databases, and business logic.

The "application layer," the cloud's topmost layer, houses apps. Autonomous scalability boosts the efficiency, availability, and affordability of cloud applications. Server farms are not the same as modular cloud computing. Because of the weak interlayer connections, each layer evolves independently. This is demonstrated by the OSI network protocol suite.

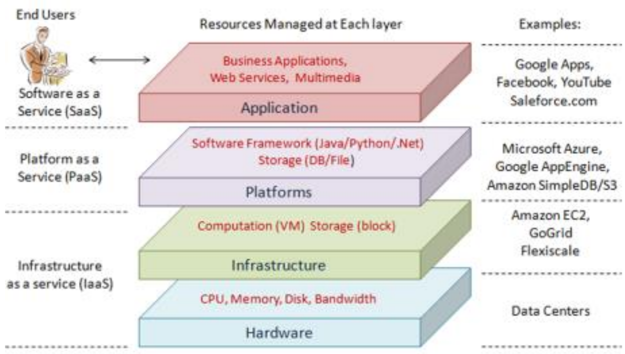

Fig. 1 Infrastructure that is hosted in the cloud..

Because of its versatility, cloud computing may serve multiple applications at a lower TCO.

Business model

Cloud computing. That is why we have the necessary equipment and software. Each new architectural layer is supported by the previous ones. Layers that intersect are eaten. SaaS, PaaS, and IaaS are examples of cloud services.

IaaS makes use of digital machines. Cloud computing companies provide infrastructure as a service. Amazon EC2, GoGrid, and Flexiscale are examples of IaaS providers.

PaaS's operating system and framework can be used by developers. PaaS is available from Microsoft Azure, Force.com, and Google App Engine.

On-demand cloud-based application service. Rackspace, SAP Business By Design, and Salesforce.com are examples of SaaS suppliers.

Fig. This diagram depicts cloud computing. 2. The levels of the cloud computing ecosystem enable PaaS providers to use IaaS clouds. Cloud computing services are widely available from Google and Salesforce. The terms "cloud" and "infrastructure" relate to PaaS and IaaS providers in this context.

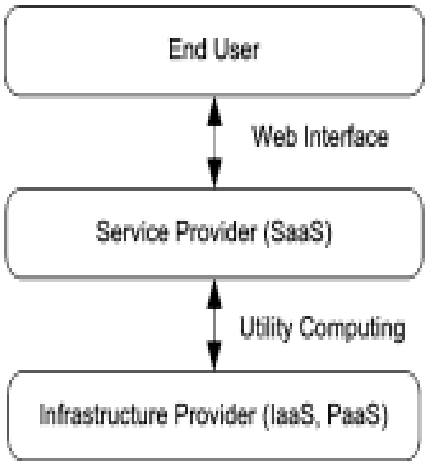

Fig. 2 A cloud-based business model

Types of clouds

It is tough to migrate an enterprise program to the cloud. Some merchants prioritize cheap over safety and dependability. Cloud computing has both advantages and disadvantages.

Providers can exchange resources over public networks. For providers, public clouds reduce risk and expense. Many corporate demands cannot be met by public cloud services due to a lack of data, network, and security management.

Each private cloud is used by a single company. Private clouds can be established by businesses or service providers. A private cloud both protects and accelerates data. Cloud computing has been chastised for its low barrier to entry and similarities to closed-source data centers.

Hybrid clouds combine private and public clouds to address their shortcomings. Hybrid clouds are made up of private and public cloud infrastructures. Public and private clouds, in contrast to hybrid clouds, are rigid. Private clouds enable the secure and scalable deployment of demand-based application services and data. Unfortunately, hybrid cloud growth necessitates the separation of public and private clouds.

VPC has several advantages over traditional cloud deployments. Public clouds are virtual private clouds. VPNs enable service providers to customize VPC topologies and firewalls. VPC is expanded by virtualizing server hardware, software, and networks. The virtualized network layer of VPC assists most organizations in migrating to cloud architecture.

Cloud computing is preferred by the majority of service providers. Low-cost computational scientific applications can be hosted in the public cloud. Clouds appear and vanish at random.

Future workplaces may make greater use of hybrid technologies. VPCs have become popular since 2009.

4. CLOUD COMPUTING CHARACTERISTICS

Cloud computing provides numerous advantages over service computing.

Customers share vendor services for cloud data centers. Companies that provide services and infrastructure frequently struggle with efficiency and administration. Cloud layer owners can focus on their separate goals by separating the task. Stakeholder engagement is reduced by multitenancy.

The infrastructure provider can dynamically assign computer resources to consumers using shared resource pooling. With dynamic source assignment, infrastructure service providers may better track resource spending. IaaS providers can benefit from VM migration by pooling servers, optimizing resources, and saving energy and cooling.

Because of the Internet, cloud services are available all over the world. Cloud services can be accessed by smartphones, PDAs, and desktops with internet connectivity. Because of localization and network efficiency, several clouds deploy worldwide data centers. The availability of multiple convenient sites may increase service provider productivity.

Cloud computing has the potential to usher in a service-oriented society. Management of services is critical. SLAs are in place for cloud computing services such as IaaS, PaaS, and SaaS. SLAs are required by all service providers.

With cloud computing, you can get assistance whenever you need it. By only purchasing what is required, dynamic resource provisioning saves money for service providers.

The ability to dynamically allocate and release resources gives service providers additional control over resource management. To cater to the "flash crowd," numerous businesses deploy automated resource management systems.

"Pay as you go" is possible with cloud services. Prices for various services vary. Hourly virtual servers are rented by SaaS enterprises from IaaS vendors. Salesforce's on-demand CRM may be charged on a per-user basis. By charging clients for their actual consumption, the utility-based pricing method reduces expenses. Spending becomes more difficult to track. VKernel allows cloud clients to monitor resource use, cost analysis, and adjustment.

5. STATE-OF-THE-ART

The introduction of innovative cloud-based software. Cloud computing has become a reality. Following that, we'll look at cloud-based data and computing services.

Cloud computing technologies

Cloud computing is discussed in this section..

a. Architectural design of data centers

Cloud computing processes and stores data in a large data center on servers, switches, and routers. Deep network design boosts distributed computing performance. Take into account durability and adaptability. In massive data centers, tiered networks have been investigated. Fig. This diagram depicts processing, storage, and networking in a data center. 3. Layers of access connect server rooms to the internet. Each building contains 20-40 servers with 1 Gbps access switches. Both aggregation switches are redundantly connected to the access switches.

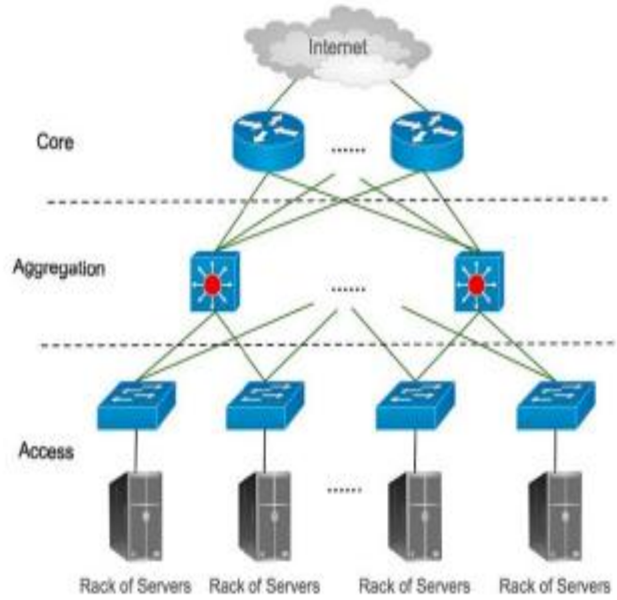

Fig. 3 Putting Together a Data Center Network

It is feasible to get 10 Gbps. The aggregation layer is responsible for balancing domain, location, and server loads. A routed, core-layer-failure-tolerant network is formed by many aggregation switches. Core routers regulate traffic both within and outside of the data center.

Network designers make use of low-cost Ethernet switches and routers. Any organization may easily adapt to a multilayered network. Data Center Network Architecture Objectives: Servers and communication should be assigned without regard for network design. Network card speed is required for data center hosts to communicate.

Moving virtual machines between networks is simplified by virtualization. Cloud hosting providers can relocate virtual computers for statistical multiplexing, modify connection patterns for densely networked hosts, and control data center temperature and electricity. Prepare your network architecture for a quick VM migration.

Failures on a large scale are unavoidable. A dependable network design should be capable of withstanding server, connection, and rack failures. Good physical connectivity is required for multicast and unicast communications. More users and servers necessitate network expansion.

Ethernet and IP network compatibility is crucial. The updated architecture should be compatible with the majority of data center Ethernet and IP equipment.

Another technological achievement is modular data centers in freight containers. MDC networks use switches to connect thousands of servers. Geodispersed MDCs near key population centers benefit from adaptable software. MDC provides redundancy because blackouts, earthquakes, and instability do not occur everywhere. Guo and colleagues preferred server-centric recursive MDCs.

b. Distributed file system over clouds

Google's File System (GFS) delivers fast, dependable data access by utilizing massive clusters of low-cost general-purpose computers. The majority of files are read in 64 MB chunks and are rarely updated. GFS in the data center improves throughput, latency, and server stability.

HDFS extends GFS to store large files across multiple computers. Data can be duplicated between servers for security purposes. Data is stored by GFS and other network nodes. The block protocol is used by HDFS nodes to transfer data blocks to users over the network. The client or browser receives all HTTP data. Communication between network nodes makes it easier to share, copy, and replicate information.

c. Distributed application framework over clouds

Clusters of servers are used in today's data centers for economic trend analysis and movie visual effects. Java EE is used in HTTP web apps.

Google developed MapReduce to handle large data cluster distributed processing. MapReduce jobs are sent to mas-ter by clients. The Master schedules jobs on task nodes strategically located across the data center. Data and Host Nodes are visible to the Master. Other rack nodes take over if the data node fails. Lowering core backbone traffic increases network throughput. Postponing labor to avoid accountability. Everything comes to a halt if the Master fails. Master makes advantage of file system logs. At launch, it looks for these data. Hadoop MapReduce was inspired by Google. Companies can use Hadoop MapReduce to process enormous amounts of data.

Commercial products

The top cloud computing service companies are contrasted.

a. Amazon EC2

Using AWS's computing, storage, and other features, businesses and people can create apps and services quickly and affordably. HTTP, REST, and SOAP are all supported by AWS.

APIs and other technologies are used by cloud clients to deploy and control Amazon EC2 server instances. Xen is used by EC2 [55]. Users can use instances to deliver new or updated code. Recent machine shots reveal significant changes. You may copy at any time. When running programs on hardware-like EC2 instances, users have nearly total software architectural control. This new capability makes Amazon's automated resource scaling more difficult.

Distribute EC2 instances. EC2 is divided into regions and ability zones. Territories are assigned availability zones. Zone isolation lowers latency between AZs within the same Region.

S3 can be used to retrieve EC2 machine images. S3 "buckets" can hold anything from one byte to five terabytes. One bucket can perform a lot of things. Determine the best location, user latency, and cost.

Using Amazon Virtual Private Cloud, AWS can be safely and easily integrated into IT infrastructures. Companies can use Amazon VPC to combine AWS compute resources such as firewalls, intrusion detection systems, and network security systems.

Amazon Cloud Watch provides near-real-time analytics to cloud clients by utilizing raw data from AWS partner services such as Amazon EC2. EC2 monitors CPU, network latency, and disk I/O.

b. Microsoft Windows Azure platform

Microsoft's Windows Azure cloud platform is divided into three parts. SQL Azure is powered by Microsoft SQL Server, while.NET Services offers a decentralized foundation for cloud and on-premises development. Windows Azure hosts both on-premises and cloud-based software.

Create Windows Azure apps. Windows supports C#, Visual Basic, C, and the.NET Framework. Windows Azure applications evolve. In lulls, ASP.NET, WCF, or both can be used to construct web apps. RESTful HTTP/HTTPS access to Windows Azure blobs, tables, and queues.

The "Huron" Data Sync database and service are hosted on SQL Azure. SQL Azure's backend database is SQL Server. It can be accessed via ADO.NET or another Windows data API. Local programs can access the cloud. Relational data synchronization using "Huron" on-premises.

Table 1 Popular things are compared

Table 1 A comparison of representative commercial products

Cloud Provider	Amazon EC2	Windows Azure	Google App Engine
Classes of Utility Computing	Infrastructure service	Platform service	Platform service
Target Applications	General-purpose applications	General-purpose Windows applications	Traditional web applications with supported framework
Computation	OS Level on a Xen Virtual Machine	Microsoft Common Language Runtime (CLR) VM; Predefined roles of app. instances	Predefined web application frameworks
Storage	Elastic Block Store; Amazon Simple Storage Service (S3); Amazon SimpleDB	Azure storage service and SQL Data Services	BigTable and MegaStore
Auto Scaling	Automatically changing the number of instances based on parameters that users specify	Automatic scaling based on application roles and a configuration file specified by users	Automatic Scaling which is transparent to users

Computation

.NET Services can be used to distribute software. Enterprises and cloud services can utilize Access Control to consolidate user authentication. Web services for Service Bus are available from both on-premises and cloud applications. Endpoints are accessed by clients via URIs.

Data center servers, virtual machines, and software are managed by fabric controller software. Transfer of software XML settings. Using this information, fabric controllers can assign servers to new apps.

c. Google App Engine

Google App Engine allows developers to host apps in the cloud. Python and Java are supported. Django, CherryPy, Pylons, web2py, and Google's own web framework are all supported by App Engine. Google is responsible for cluster setup, monitoring, failover, and application activation. The APIs handle HTTP queries, caching, and Big Table database storage and retrieval. App Engine developers have read-only access to the filesystem.

Table 1 compares three main cloud providers in terms of utility computing, application categories, processing power, storage capacity, and autonomous scaling. Cloud services abstract and manage data in a variety of ways. Multiple cloud options can meet the needs of the business.

6. RESEARCH CHALLENGES

Despite its youth, cloud computing is gaining popularity. The industrial applications are still unknown. Cloud computing research will confront problems in the future.

Automated service provisioning

The ability to acquire and release resources as needed is a key selling point for cloud computing. Service providers reallocate cloud resources to meet SLOs and reduce costs. There are numerous techniques that service providers might take. Material assets are inextricably tied to service excellence. The "flash crowd effect," or changes in demand, may hinder the performance of online services.

Mechanical assistance is not new. The concept of dynamic resource provisioning for Internet applications is well known. They often include automatic resource allocation, an application performance model that calculates the number of application instances required to achieve QoS standards at each tier, and periodic demand and resource projections. Statistical machine learning is combined with queueing and control theory to model application efficacy.

There is a distinction between proactive and reactive resource management. Asset deployment comes before asset utilization. Our reaction to unexpected demand shifts is reactive. Both resource management strategies are required in today's dynamic corporate world.

Virtual machine migration

Cloud computing spreads processing needs evenly by relocating virtual computers throughout a data center. Fast and dependable data center provisioning aids in the migration of virtual machines.

Change to a remote computer that is moving. "Live" VM migrations between Xen and VMWare recently took tens of milliseconds. According to Clark et al., live VM migration avoids process migration issues.

Server consolidation

Migration of virtual machines is difficult, but it decreases hotspots. Current approaches are incapable of immediately identifying and transferring workload hotspots. Memory delivery must be both secure and efficient, taking into account both application and physical server resources.

Cloud server consolidation lowers power costs while increasing productivity. Live VM migration solutions consolidate virtual machines (VMs) on idle servers. Some people compare data center server consolidation to vector bin packing, which is an NP-hard problem. Heuristics were used to solve the problem. Virtual machines require

Energy management

Communication has only recently been recognized as being crucial. Server consolidation should not have an impact on application performance. Different resources are required for virtual machines. If virtual machines consume too much bandwidth, memory cache, or disk I/O during server consolidation, resource congestion may occur. Server consolidation may need tracking VM footprint changes. Finally, manage your limited resources.

Another goal is the energy efficiency of cloud computing. Electricity and cooling account for more than half of a data center's operating costs. By 2020, data centers would consume 18% of US electricity, up from 1.5% in 2006. Energy efficiency is more important to infrastructure businesses. Power consumption in data centers can be lowered for environmental and legal reasons.

Data centers are becoming more environmentally friendly. There are various methods to this problem. Power-saving CPUs may cause the CPU to slow down or eliminate functions. To save energy, intelligent workload scheduling and server consolidation shut down workstations. Recent studies look at energy-efficient network topologies and protocols. To preserve electricity, all of these strategies lower application performance. There has been little research into cloud efficiency and performance.

Traffic management and analysis

Traffic analysis is critical in data centers. Many websites use traffic statistics to improve user experience. To plan and maintain networks, network operators must understand traffic trends.

When it comes to data center traffic measurement and analysis solutions, ISPs and businesses face numerous challenges. First, link density is far larger than in ISP and enterprise networks, rendering the aforementioned options inefficient. 2) In a modular data center with thousands of servers, most algorithms can only construct traffic matrices between a few hundred hosts. MapReduce workloads have an impact on both Internet and business network traffic. The network, CPU, and storage use of the application are all linked.

Few data centers keep track of network activity. Greenberg and colleagues design networks based on data center traffic characteristics such as flow volumes and concurrent flows. Benson et al. examine router SNMP traces for data center perimeter traffic.

Data security

For cloud computing, research data must be secure. Because service providers rarely have access to physical data center security, they rely on infrastructure vendors. Remote security policies can be defined but not enforced by virtual private cloud providers. To ensure application security, data access and transfer must be encrypted and auditable by the infrastructure provider. Remote attestation ensures the anonymity of cryptographic keys. The trustworthy platform module (TPM) generates a non-forgeable system summary (system state encrypted with its private key) for remote attestation. Because cloud-based VMs are transient, remote attestation is problematic. Each layer of the cloud architecture requires trust mechanisms. TPM allows for hardware-based security. VR consumers must have faith in the medium. Only trusted servers should be allowed to switch virtual machines. Recent research looks into the development of trust.

Software frameworks

Apps that require a lot of data perform well on the cloud. Hadoop and other MapReduce frameworks enable this program scale and accept faults. According to recent study, MapReduce task effectiveness and resource requirements differ depending on the application type. Sort relies heavily on I/O in Hadoop, whereas grep relies heavily on CPU. Hadoop hosts can accommodate an unlimited number of virtual machines. The number of VMs on a server limits VM bandwidth. MapReduce application efficiency and cost optimization necessitate configuration and scheduling adjustments. Application execution speeds up when bottlenecks are removed. The most difficult issues are dynamic scheduling and offline/online Hadoop task performance modeling. Energy-aware MapReduce frameworks are one option. Idle Hadoop nodes are relaxed until they get new work. Hadoop and HDFS rely heavily on energy efficiency. Performance and energy savings may be incompatible. More investigation is required to find the optimal outcome trade-off.

Storage technologies and data management

Data-intensive distributed toolkits such as MapReduce, Hadoop, and Dryad exist. GFS and HDFS internet-scale file systems are used in this approach. These file systems are distinguished from distributed file systems by their storage, access, and API. Non-POSIX interfaces cannot be used by legacy file systems or software. This subject has been well researched. Patil and colleagues created the MapReduce framework as well as scalable, concurrent data access API primitives for IBM's GPFS cluster file systems.

Novel cloud architectures

Commercial clouds are housed in centralized data centers. Although data centers are costly and energy-intensive, their design provides for management. A recent study discovered that smaller data centers require less powerful and expensive cooling systems because they are less expensive to create, more distributed, and consume less electricity. Geographic diversity benefits real-time applications such as gaming and content distribution. Valancius et al. investigated the video transmission capabilities of nano-data centers.

Another topic is cloud app hosting on user-provided infrastructure. Volunteer-managed clouds or dedicated servers save charity money while increasing efficiency. When designing, keep turnover and resource management in mind. The motivations behind these arrangements have been investigated.

7. CONCLUSION

Your servers and networks are kept up to date in the cloud. Utility computing is made possible by cloud computing.

Despite their advantages, cloud-based systems are always changing. The current investigation focuses on autonomous resource provisioning, power management, and security. We believe that this work will improve the field.

The basics, architectural blueprints, important characteristics, main technologies, and research aims of cloud computing were investigated. As cloud computing becomes more common, we anticipate that our research will shed light on design considerations and motivate further research.

REFERENCES

1. Al-Fares M et al (2008) A scalable, commodity data center net- work architecture. In: Proc SIGCOMM
2. Amazon Elastic Computing Cloud, aws.amazon.com/ec2
3. Amazon Web Services, aws.amazon.com
4. Ananthanarayanan R, Gupta K et al (2009)

Cloud analytics: do we really need to reinvent the storage stack? In: Proc of HotCloud

5. Armbrust M et al (2009) Above the clouds: a Berkeley view of cloud computing. UC Berkeley Technical Report
6. Berners-Lee T, Fielding R, Masinter L (2005) RFC 3986: uniform resource identifier (URI): generic syntax, January 2005
7. Bodik P et al (2009) Statistical machine learning makes automatic control practical for Internet datacenters. In: Proc HotCloud
8. Brooks D et al (2000) Power-aware microarchitecture: design and modeling challenges for the next-generation microprocessors, IEEE Micro
9. Chandra A et al (2009) Nebulas: using distributed voluntary re- sources to build clouds. In: Proc of HotCloud
10. Chang F, Dean J et al (2006) Bigtable: a distributed storage system for structured data. In: Proc of OSDI