CHAPTER 1

INTRODUCTION TO CLOUD COMPUTING

CHAPTER OUTLINE

After studying this chapter, students will be able to understand the:

- Overview of Clouding Computing
- Evolution of Cloud Computing
- Characteristic of Cloud Computing
- Type of Cloud and its cloud services
- Benefits and challenges of cloud computing
- Applications cloud computing
- Cloud storage
- Cloud services requirements
- Cloud and dynamic requirements,
- Cloud adoption

ANALOGY OF CLOUD COMPUTING

Once upon a time, every family, town, farm, or hamlet had its own water well. Today, shared public utilities provide them the clean water by just turning on the faucet; cloud computing operates similarly. Cloud computing services, like water from your kitchen faucet, may be turned on and off as required. As with the water business, a staff of devoted specialists ensures that the service supplied is safe, secure, and available 24 hours a day, seven days a week. When you do not turn on the water, you are not only conserving water, but you are also not paying for resources you do not need right now.

OVERVIEW OF CLOUD COMPUTING

Cloud computing has impacted practically every aspect of IT in the last decade. All of these programs, including sales, marketing, finance, and support, are being redesigned to take advantage of the cloud's quick access, no download, and pay-as-you-go features. The cloud is altering the way applications are built, tested, and deployed, according to Gartner, resulting in a substantial shift in application development priorities. Cost is a big factor, but so are agility, adaptability, and the speed with which new applications may be deployed.

> **"You don't generate your own electricity. Why generate your own computing?"**
>
> ***Jeff Bezos, Amazon.***

Instead of residing on your desktop or inside your employer's network, all of the computer hardware and software you use is offered as a service by another firm and accessible through the Internet, generally in a fully seamless manner. To you, the user, it does not matter where the hardware and software are situated or how they work—simply it is someplace on the Internet. Cloud computing allows users to access computer system resources on-demand, including storage and computational power, without having to manage them directly. The word refers to data centers that are accessible over the Internet to a large number of people. Functions from central servers are commonly dispersed across numerous locations in today's large clouds. Clouds can be private (for a single company), public (for multiple organizations), or a mixture of the two (hybrid.) Amazon Web Services (AWS) is the largest public cloud.

> "Cloud computing is a model for enabling ubiquitous, 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. This cloud model is composed of five essential characteristics, three service models, and four deployment models."
>
> ***National Institute of Standards & Technology (NIST)***

Different computer services are supplied remotely to clients through the Internet in cloud computing. Clients often pay providers a monthly or annual service charge to acquire access to systems that provide software as a service, platforms as a service and infrastructure as a service to subscribers. Clients that use cloud computing services can benefit from a range of features, based on their specific company needs at the moment. For any company that decides to use the cloud computing model to acquire IT services, the days of big capital investments in software and IT infrastructure are now a thing of the past. The ability to access robust IT resources on an as-needed basis is leveling the playing field for small and medium-sized businesses. It gives them the needed tools and technologies to compete in the global marketplace without having to invest in on-premises IT resources. Clients that subscribe to computing services offered via the "cloud" can significantly lower their IT service costs while also gaining access to more agile and adaptable business computing capabilities.

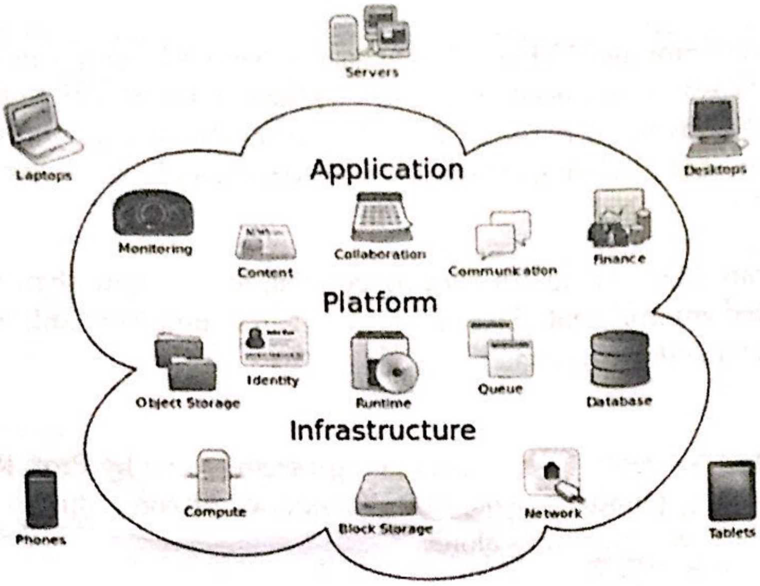


Figure 1.1 Cloud Computing

Companies that own, house, and run enormous buildings full of servers provide cloud computing. Data centers are the facilities that power the applications, services, and virtual machines that are offered to clients. Customers use them to store data, install applications, and construct virtual networks.

> Cloud computing is the realization of utility computing for the masses, where traditional IT services are now virtualized and provided via modular reference architectures that are created by the providers and vendors rather than end-users.
>
> *Chris Poelkar*

«Everything is stored remotely and distributed via web-based connections, which is why we call it the cloud. All of this data is not kept in one place; instead, it is accessed by anybody who has access to the Internet. Cloud computing services are used by businesses because they are less costly than purchasing pricey computing gear. Users just rent the computing resources of a provider's data center to virtualize the tools they require. Cloud services are available on a pay-per-use basis, allowing businesses to spend money on what they need right now while knowing that they can grow up later. The number of cloud services available is growing all the time, but they all follow the same delivery strategy.

Computing Resources should be always available when you need them, and you have control over overturning them on or off to ensure there is no lack of resources or wastage happens. You should be able to scale (increase or decrease the resource) when necessary. The cloud providers should have sufficient capacity to meet customer's needs. Related processes including billing, resource provisioning, and deployment should be self-service and automated, involving much less manual processing. If a machine where the cloud service is hosted fails, the cloud provider should be able to start failover service immediately. Cloud providers should be able to provide customer reliability service, committing to uptimes of their service. For the used cloud services, customers pay the cloud provider as a utility-based subscription, just like paying your electricity bill.

EVOLUTION OF CLOUD COMPUTING

The Early 1960s

The computer scientist John McCarthy, come up with the concept of timesharing and enabling the organization to simultaneously use an expensive mainframe. This computing is described as a significant contribution to the development of the Internet, and a pioneer of Cloud computing.

1969

The idea of an "Intergalactic Computer Network" or "Galactic Network" (a computer networking concept similar to today's Internet) was introduced by J.C.R. Licklider, who was responsible for enabling the development of ARPANET (Advanced Research Projects Agency Network). His vision was for everyone on the globe to be interconnected and being able to access programs and data at any site, from anywhere.

1970

Using virtualization software like VMware; it became possible to run more than one Operating System simultaneously in an isolated environment. It was possible to run a completely different Computer (virtual machine) inside a different Operating System.

1997

The first known definition of the term "Cloud Computing" seems to be by Prof. Ramnath Chellappa in Dallas in 1997 – "A computing paradigm where the boundaries of computing will be determined by economic rationale rather than technical limits alone."

1999

The arrival of Salesforce.com in 1999 pioneered the concept of delivering enterprise applications via a simple website. The services firm covered the way for both specialist and mainstream software firms to deliver applications over the Internet. Salesforce started to offer users the ability to run applications over the internet, marking the beginning of Software as a Service.

2002

Amazon introduced its web-based retail services. The Cloud Computing Infrastructure Model gave them the flexibility to use their computer's capacity much more efficiently.

2003

The first public release of Xen, which creates a Virtual Machine Monitor (VMM) also known as a hypervisor, a software system that allows the execution of multiple virtual guest operating systems simultaneously on a single machine.

2006

Amazon launched Amazon Web Services, which offers online services to other websites, or clients. One of Amazon Web Services' sites, called Amazon Mechanical Turk, provides a variety of Cloud-based services including storage, computation, and "human intelligence." Amazon expanded its cloud services. First was its Elastic Compute Cloud (EC2), which allowed people to access computers and run their applications on them, all on the cloud. Then they brought out the Simple Storage Service (S3). This introduced the pay-as-you-go model to both users and the industry as a whole, and it has become standard practice now.

Google launched the Google Docs services. Google Docs was originally based on two separate products, Google Spreadsheets, and Writely. Google purchased Writely, which offers renters the ability to save documents, edit documents, and transfer them into blogging systems which documents are compatible with Microsoft Word. Google Spreadsheets, which was acquired from 2Web Technologies, in 2005 is an Internet-based program allowing users to develop, update, and edit spreadsheets, and to share the data online.

2007

IBM, Google, and several universities joined forces to develop a server farm for research projects needing both fast processors and huge data sets. The University of Washington was the first to sign up and use resources provided by IBM and Google. Carnegie Mellon University, MIT, Stanford University, the University of Maryland, and the University of California at Berkeley, quickly followed suit. The universities immediately realized computer experiments can be done faster and for less money if IBM and Google were supporting their research. Since much of the research was focused on problems IBM and Google had interests in, they also benefited from the arrangement.

Netflix launched its video streaming service, using the Cloud, and provided support for the practice of "binge-watching."

2008

Google announced a preview release of App Engine, a developer tool that allowed users to run their web applications on Google infrastructure.

Microsoft announced the Windows Azure Platform.

Eucalyptus offered the first AWS API compatible platform, which was used for distributing private Clouds. In the same year, NASA's OpenNebula provided the first open-source software for deploying Private and Hybrid Clouds. Many of its most innovative features focused on the needs of major businesses.

2010

Cloud Storage launched. Google App Engine for Business announced, offering added management and support features tailored specifically for the enterprise. BigQuery and Prediction API announced in limited preview.

Windows Azure Platform commercially available. Rackspace Hosting and NASA jointly launched an open-source cloud-software initiative called OpenStack.

2011

Google App Engine officially leaves preview and becomes a "fully supported Google product". IBM introduced the IBM SmartCloud framework, in support of Smarter Planet. Then, Apple launched iCloud, which focuses on storing more personal information such as photos, music, videos, etc. Microsoft began advertising the Cloud on television, making the general public aware of its ability to store photos, or video, with easy access.

2012

Google Compute Engine was released as an addition to the Google Cloud Platform. The offering stands as the IaaS component of the platform enabling users to launch virtual machines on-demand, which can be launched from the standard images or custom images created by users.

Oracle introduced the Oracle Cloud, offering the three basics for business, IaaS (Infrastructure-as-a-Service), PaaS (Platform-as-a-Service), and SAAS (Software-as-a-Service).

2013

Google Cloud Storage begins automatically encrypting each Storage object's data and metadata under the 128-bit Advanced Encryption Standard (AES-128). PHP support for Google App Engine goes into preview, making PHP on App Engine available for everyone.

IBM acquires SoftLayer, which marks IBM's entry into cloud computing.

2014

Google Cloud SQL graduates to General Availability (GA), supporting automatic encryption, 99.95% uptime SLA, and support for databases up to 500GB. Google acquired Firebase and Stackdriver making them a part of GCP.

Windows Azure was renamed, Microsoft Azure. IBM introduced Bluemix, a Platform-as-a-Service (PaaS).

2015

Financial Conduct Authority FCA declares cloud technologies as "acceptable" for use by financial services firms.

Microsoft Azure Cloud Switch was introduced as a cross-platform Linux distribution. Currently known as SONiC.

Google Cloud DNS goes into GA along with the expansion of load balancing solutions to 12 additional points of presence. The Beta version of Bigtable is launched.

2016

Salesforce.com, a cloud computing company that makes money primarily through its customer relationship management product suite, selects Amazon Web Services as its preferred public cloud infrastructure provider. However, Salesforce.com does not plan to move entirely to Amazon but rather uses Amazon only to meet infrastructure expansion needs in new geographical areas and for specific use cases.

IBM rolled the SoftLayer brand under its Bluemix brand of PaaS offerings.

2017

AWS launches i3 instances, a new generation of instances with large SSDs intended to be used for high throughput datastores. The instances are more than 50% cheaper than the corresponding previous generation i2 instances and have larger memory.

MIT math professor Andrew V. Sutherland breaks the record for the largest ever Google Compute Engine cluster with 220,000 cores on Preemptible Virtual Machines.

IBM unified its entire cloud portfolio as IBM Cloud. Google, IBM, and Lyft announced the Alpha release of Istio: a new open-source project that provides a uniform way to help connect, secure, manage and monitor microservices.

2018

Amazon acquired the aws.com domain from Earth Networks, formerly known as Automated Weather Source.

According to Deloitte, spending on IT-as-a-Service for data centers, software and services reached $547B by the end of 2018. As per the salesforce study, cloud computing spending is growing at 4.5x the rate of IT spending since 2009 and is expected to grow at better than 6x the rate of IT spending through 2020.

2019

Amazon announced AWS Outposts, which is a fully managed service that extends AWS infrastructure, AWS services, APIs, and tools to virtually any customer datacenter, co-location space, or on-premises facility for a truly consistent hybrid experience.

2020

Microsoft said that there was a 775% increase in Microsoft Teams usage in Italy due to the COVID-19 pandemic. The company estimates there are now 44 million daily active users of Teams worldwide.

CLOUD COMPUTING AT A GLANCE

5 ESSENTIAL CHARACTERISTICS • 3 SERVICE MODELS • 4 DEPLOYMENT MODELS

5 Essential Characteristics:

- Broad Network Access
- Rapid Elasticity
- Measured Service
- On-Demand Self-Service
- Resource Pooling

3 Service Models:

- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)

4 Deployment Models

- Public
- Private
- Hybrid
- Community

Advantages

- **Flexibility:** Cloud services give you the flexibility to access it anywhere.
- **Availability:** At the same time, it is available 24×7.
- **Scalability:** You can provide additional instances and terminate it as and when it is required.
- **Cost of Ownership:** There is no Cost of Ownership. There is no need for investment as you do not own any servers.
- **Metered Services:** It gives you a metered reading as to how much you need to pay and how much you are utilizing.

Disadvantages

- **Downtime:** It Impacts the entire service if there is downtime.
- **Vulnerability to attack:** Anything that is accessible through the Internet is vulnerable to attack. This can be rectified by patching and ensuring that it is accessible to only those who require it.

CLOUD COMPUTING ENVIRONMENT

The cloud computing environment is a novel and effective method of delivering IT-related services. In reality, this method is the dream of many business owners who want to acquire all of their IT services in one location in minimal effort. This type of service is becoming increasingly popular in the information age since it assists entrepreneurs in swiftly and efficiently resolving all IT difficulties inside one organization. The cloud computing environment is all about IT and what IT requires: various types of software and hardware, pay-per-use, or subscription-based services delivered both via the Internet and in real-time. Performance management, backup, and recovery, configuration management, helpdesk, data center automation, etc. are popular services in the cloud environment. On-demand services are gaining traction among developers because of their flexibility.

Cloud computing intends to power the next generation of data centers by allowing application service providers to lease data center capabilities to deploy applications based on customer QoS (Quality of Service) needs. The formulation, setup, and deployment requirements for cloud applications are all distinct.

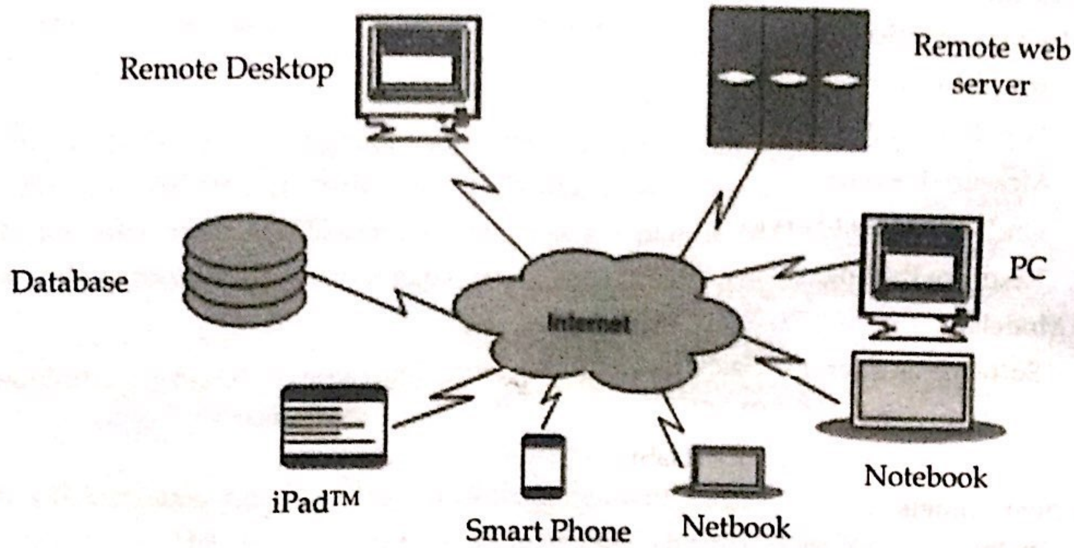


Figure 1.2 Components of Cloud Environment

An extendable simulation framework that allows modeling and simulation of cloud computing infrastructures are employed to make the cloud communication process easier. The simulation toolkit such as CloudSim allows for the modeling and development of one or more virtual machines (VMs) on a simulated data center node, as well as the mapping of workloads to appropriate VMs. It also enables the modeling of many data centers, allowing for research on the federation and related rules for VM movement for dependability and autonomous application scalability.

Despite the rapid development in multiple sectors, the field of integrated development environments (IDEs) has remained relatively unaffected. More than 15 million engineers, teams, and organizations across the world continue to utilize desktop IDEs as their primary workstations.

Why hasn't the development environment, like so many other applications, been transferred to the cloud swiftly? This is a big question the cloud market leader needs to answer.

WHAT IS WRONG WITH THE DESKTOP ENVIRONMENT?

Desktop development environments are becoming outdated, failing more often, and causing productivity issues for developers.

1. **Complicated configuration management**. The substantial configuration management process for a developer's workspace turns developers into part-time system administrators, responsible for their mini-data center running entirely on the desktop. This is time-consuming, error-prone, and challenging to automate. Many developers have multiple computers and are forced to repeat these tasks on each machine. There is no way to synchronize the configurations of components across different machines, and each machine requires similar hardware and operating systems to operate the components identically.
2. **Decreased productivity**. Many IDEs are memory and disk hogs, with significant boot times. They are so resource-hungry they can starve other applications, such as the Web browser. The net effect is a less productive developer due to a slower machine.
3. **Limited accessibility**. Desktop developer workspaces are not accessible via mobile devices.
4. **Poor collaboration**. These days most developers work as part of a team, so communication and collaboration are critical. But desktop IDEs must outsource collaboration to communication systems outside the developer's workflow, forcing developers to continuously switch between developing within the IDE and communicating with their team via other means.

THE SOLUTION IS A CLOUD ENVIRONMENT

1. **Development workspace into the cloud**. To solve the issues of desktop development, the complete development workspace must be moved to the cloud. The developer's environment consists of an IDE, a local build system, a local runtime to test and debug locally edited code, and the connections between these components and their dependencies using tools like Continuous Integration or central services like Web Services, specialized data stores, legacy applications, or partner-provided services.
2. **Centralized.** The centralized cloud-based workplace makes it simple to share. Developers can invite others into their workplace to collaborate on projects, such as co-editing, co-building, or co-debugging. Developers may speak with one another in the workplace, which has changed the way pair programming, code reviews, and classroom instruction are done.
3. **System efficiency**. The cloud can increase system efficiency and density by allocating a customizable slice of available memory and computation resources to each workspace. Of course, there is always more work to be done, and we are still far from fully exploiting the limitless possibilities that cloud computing provides developers. However, the advantages are already apparent.

MAJOR SERVICES MODELS OF CLOUD COMPUTING

Because cloud computing is the dynamic delivery of information technology resources and capabilities as a service through the Internet, we all utilize it daily, whether we realize it or not. When you sit at your computer and input a query into Google, the computer on your desk plays a minor role in locating the information you require. The words you input are quickly sent over the Internet to one of Google's hundreds of thousands of clustered PCs, which dig up your results and return them to you as quickly as possible. When you conduct a Google search, the actual job of discovering your answers may be performed by a computer in California, Dublin, Tokyo, or Beijing; you have no idea—and probably do not care! The same goes for web-based email services like Gmail and Outlook, as well as social networking sites like Facebook and Twitter. We use Google Images to store our photos, and we use Google Drive, Dropbox, or OneDrive to keep our files.

Cloud computing is a modern form of document preparation through the Internet. Simply go on to a web-based service like Google Docs and use Web-based software to generate a document, spreadsheet, presentation, or whatever you want. Instead of putting your thoughts into a program on your computers like Microsoft Word or Open Office, you are using equivalent software on a PC in one of Google's global data centers. You may access the document you create remotely, on a Web server, from any Internet-connected computer, anywhere in the globe, at any time. Do you have any idea where it is kept? No way! Do you mind where it is kept? No, once again! When you use a Web-based service like this, you are "contracting out" or "outsourcing" part of your computing needs to a firm like Google: they pay for the cost of building and maintaining the software, and they recoup the cost through advertising and other paid-for services.

SAAS (SOFTWARE AS A SERVICE)

SaaS allows clients to access software programs from anywhere using a browser. Software as a service (SaaS) is sometimes known as "on-demand software." Because service providers store programs and their related data at their location, clients may access SaaS apps from anywhere over the web. The fundamental advantage of SaaS is that it has a reduced cost of usage, as subscription costs are substantially lower than what is generally required under the traditional software delivery model. Subscribing to the SaaS style of software delivery eliminates licensing fees, installation charges, maintenance fees, and support expenses that are commonly connected with the conventional style of software delivery. Google Apps and internet-based email applications such as Gmail, Outlook, and others are examples of SaaS.

> **Cloud computing is the on-demand delivery of computing power, database storage, applications, and other IT resources through a cloud services platform via the internet with pay-as-you-go pricing.**
>
> aws.amazon.com

PAAS (PLATFORM AS A SERVICE)

Clients can create and distribute customized apps in a hosted environment through the web using PaaS (Platform as a Service). It represents a novel approach to software development that is fast gaining traction. Salesforce.com is an example of PaaS. The fundamental benefit of PaaS, like SaaS, is a cheaper cost of usage, as subscriber costs are substantially lower than what is generally required when deploying traditional software development, testing, and deployment technologies. PaaS companies manage platform maintenance and system updates, making business software development more efficient and cost-effective.

> "With the cloud, individuals and small businesses can snap their fingers and instantly set up enterprise-class services."
>
> **Roy Stephan**

IAAS (INFRASTRUCTURE AS A SERVICE)

IaaS (Infrastructure as a Service) enables clients to access IT gear and resources on a pay-as-you-go basis from a remote location. **HaaS** is another name for it (hardware as a service). IBM, Google, and Amazon.com are among the major IaaS providers. Virtualization is used in IaaS to create and manage infrastructure resources in the "cloud." Small startup businesses benefit greatly from IaaS since it enables them to progressively build their IT infrastructure without having to make huge financial expenditures in hardware and peripheral systems.

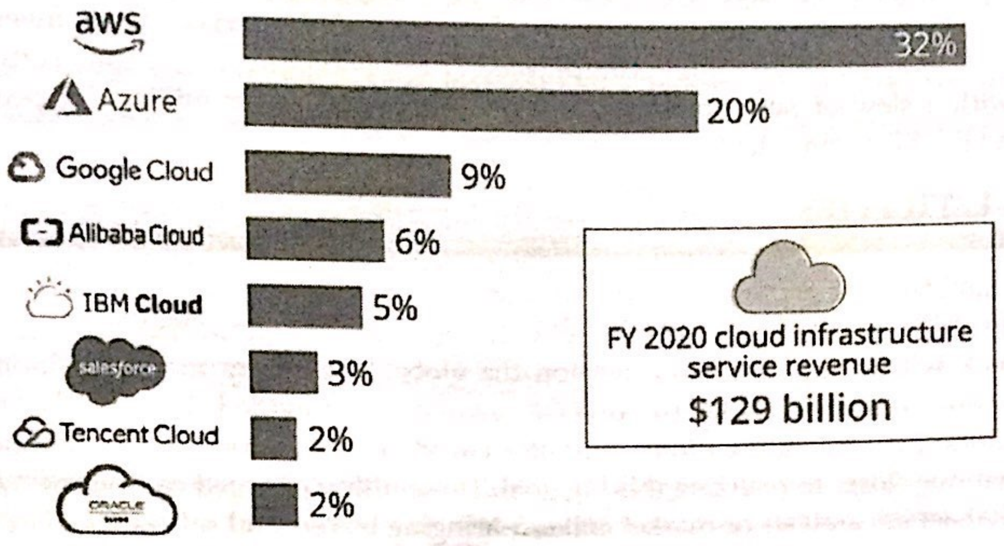
Worldwide IaaS Public Cloud Services Market Share, 2018-2019 (Millions of U.S. Dollars)

Company	2019 Revenue	2019 Market Share (%)	2018 Revenue	2018 Market Share (%)	2018-2019 Growth (%)
Amazon	19,990.4	45.0	15,495.0	47.9	29.0
Microsoft	7,949.6	17.9	5,037.8	15.6	57.8
Alibaba	4,060.0	9.1	2,499.3	7.7	62.4
Google	2,365.5	5.3	1,313.8	4.1	80.1
Tencent	1,232.9	2.8	611.8	1.9	101.5
Others	8,858	19.9	7,425	22.9	19.3
Total	**44,456.6**	**100.0**	**32,382.2**	**100.0**	**37.3**

Source: Gartner (August 2020)

Key players in Cloud Computing

Category	Characteristics
SaaS	SalesForce.com (CRM), Clarizen.com (Project Management), Google Documents, Google Mail (Automation)
Paas	Google AppEngine, Microsoft Azure, Manjrasoft Aneka
IaaS	Amazon EC2 and S3, GoGrid, Nirvanix



Source: Synergy Research Group

Figure 1.3: Worldwide market share of leading service providers in 2020.

Visions Of Cloud Computing

In the short period since its inception, cloud computing has advanced a lot. Let us take a look at what Cloud computing technology may become in the future.

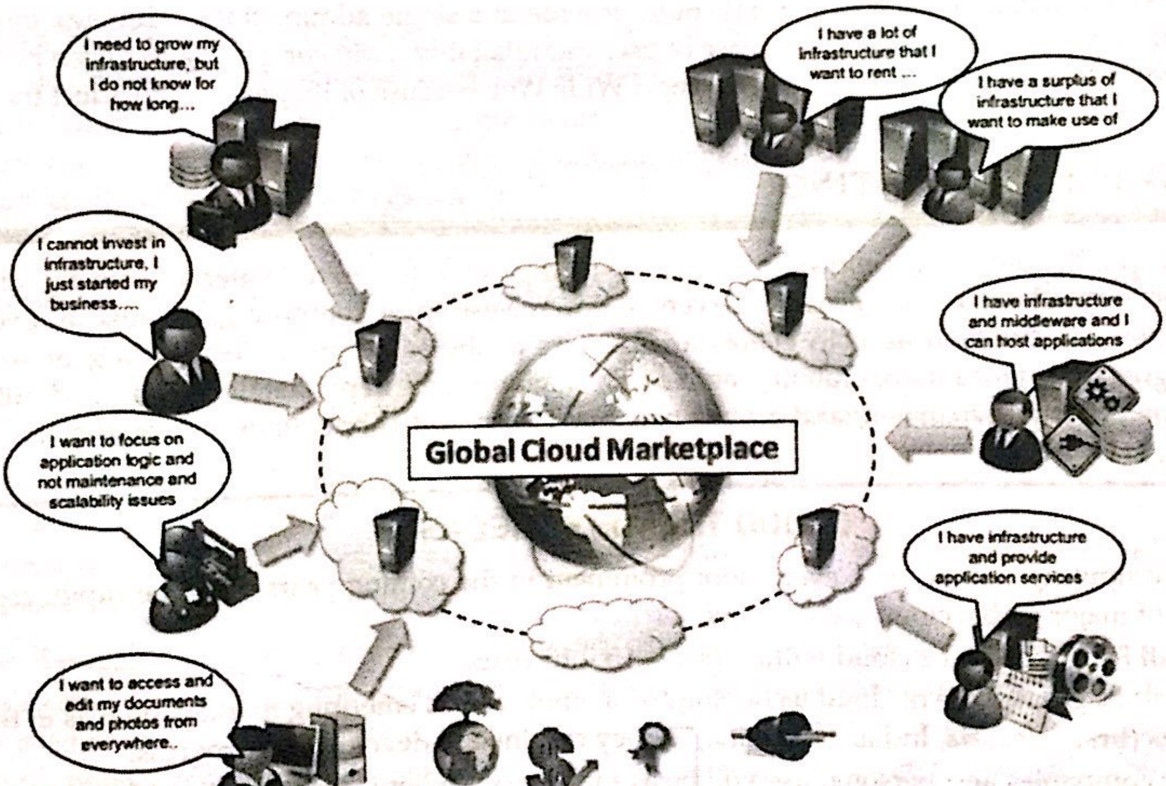


Figure 1.4 Global Cloud Marketplace

Service Provisioning Model

The computing utility based on the service provisioning model foresees a huge revolution of the whole computing sector in the twenty-first century, with computer services readily available on-demand, similar to other utility services already accessible in society.

Similarly, users of computer services (consumers) are only required to pay providers when they utilize computer services. Furthermore, customers are no longer needed to make large investments or face challenges in constructing and maintaining complicated IT infrastructure. As a result, software developers are confronted with a slew of new obstacles in developing software for millions of people to use as a service rather than on their computers.

COMPUTER UTILITIES

The Internet's establishment represented the first step in realizing this great 21st-century vision of computer utilities by establishing a global network of computer networks that allows individual computers to speak with computers anywhere on the globe. The ability to use an infinite quantity of dispersed computing resources held by diverse owners is provided by this internetworking of independent machines. As a result, various computer paradigms have been suggested and implemented in recent years to move closer to realizing this big goal. These utility-oriented computing systems are used by applications that act as catalysts or market makers, bringing buyers and sellers together. This generates many trillion dollars in utility/pervasive computing revenue.

GRID COMPUTING

Grid computing allows for the sharing, selection, and aggregation of a wide range of physically distributed resources, such as supercomputers, storage systems, data sources, and specialized devices, owned by various organizations, to solve large-scale resource-intensive problems in science, engineering, and commerce. The motivation for Grid computing was initially driven by large-scale, resource (computational and data)-intensive scientific applications that required more resources than a single computer (PC, workstation, supercomputer) could provide in a single administrative domain, inspired by the electrical power Grid's pervasiveness, ease of use, and reliability. Grid computing has been heralded as the next revolution after the Internet and the World Wide Web because of its potential to affect the twenty-first century.

PEER-TO-PEER COMPUTING

Peer-to-peer (P2P) computing allows peer nodes (computers) to share material directly and in a decentralized way. There are no clients or servers in pure peer-to-peer computing. Because all peer nodes are equal, they can function as both clients and servers at the same time. Cost-sharing or reduction, resource aggregation and interoperability, enhanced scalability and dependability, enhanced autonomy, anonymity or privacy, dynamism, and ad-hoc communication and cooperation are some of the aims of P2P computing.

CLOUD TRENDS FORECASTS

- Cloud computing will become even more prominent in the coming years with the rapid, continued growth of major global cloud data centers.
- 50% of all IT will be in the cloud within the next 5 – 10 years.
- There will be greater use of cloud technology as a whole across emerging markets such as in the BRIC countries (Brazil, Russia, India, and China) as they continue to develop and progress.
- Data for companies and personal use will be available everywhere in standardized formats, allowing us to easily consume and interact with one another at an even greater level.
- The security and reliability of cloud computing will continue to evolve, ensuring that data will be even more secure with numerous techniques employed.
- We will not even consider 'cloud' as the key technology; instead, we will focus on the services and applications that it enables.
- Combining cloud technology with the Internet of Things (IoT), Wearable, and 'Bring Your Own Device' (BYOD) will become the norm in personal and working lives, so much so that the presence of cloud technology as an enabler will be overlooked.

- Worldwide end-user spending on public cloud services is forecast to grow 23.1% in 2021 to total $332.3 billion, up from $270 billion in 2020, according to the latest forecast from Gartner, Inc.
- Gartner forecasts Software as a service (SaaS) to remain the largest market segment and is predicted to reach $122.6 billion in 2021. Infrastructure-as-a-service (IaaS) and desktop-as-a-service (DaaS) will see the highest growth in 2021, 38.5% and 67.7% respectively.

SERVICE COMPUTING

The relationship between business processes and IT services is the emphasis of service computing, which allows business operations to be easily automated using IT services. Service-Oriented Architecture (SOA) and Web Services are two examples of services computing technology. The SOA allows distributed systems to interact and exchange data with one another, allowing service consumers and providers to discover and deliver services in a consistent manner. Web Services allows for the operation of self-contained business processes through the Internet.

MARKET ORIENTED COMPUTING

It looks at computer resources from an economic standpoint, implying that resource consumers will have to pay resource suppliers to access them. As a result, it can provide benefits such as incentivizing resource providers to contribute their resources for others to use and profit from, regulating the supply and demand of computing resources at the market, incentivizing resource users to back off when necessary, and eliminating the need for a central coordinator.

VIRTUALIZED COMPUTE AND STORAGE TECHNOLOGIES

Cloud computing is the most recent paradigm to emerge, promising dependable services offered through next-generation data centers based on virtualized computation and storage technologies. Details of virtualization will be discussed later in unit IV.

CHARACTERISTICS OF CLOUD COMPUTING

National Institute of Standards and Technology (NIST) is an agency under the scope of the US Department of Commerce. NIST is responsible for defining standards in Science and Technology. The Computer Security Division of NIST has provided a formal Definition and Characteristics of Cloud computing.

According to the official NIST definition, "*cloud computing is a model for enabling ubiquitous, 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.*"

The NIST definition lists five essential characteristics of cloud computing: on-demand self-service, broad network access, resource pooling, rapid elasticity or expansion, and measured service. It also lists three "service models" (software, platform, and infrastructure), and four "deployment models" (private, community, public, and hybrid) that together categorize ways to deliver cloud services. The definition is intended to serve as a means for broad comparisons of cloud services and deployment strategies, and to provide a baseline for discussion from what is cloud computing to how to best use cloud computing.

NIST'S FIVE ESSENTIAL CHARACTERISTICS OF CLOUD COMPUTING

1. On-demand self-service
2. Broad network access
3. Resource pooling
4. Rapid Elasticity
5. Measured service

On-Demand Self service

Email, Application Network, and Server services may all be delivered without requiring contact with each service provider. Self-service refers to when a customer completes all of the steps required to obtain a service without having to go through an IT department. For example - The consumer's request is then automatically processed by the cloud infrastructure, without human intervention on the provider's side.

Broad Network Access

Cloud services are accessible over the Internet and may be accessed using a standard mechanism that allows for usage by a variety of clients, including mobile phones and laptops.

Resource Pooling

Different physical and virtual resources are dynamically assigned and reassigned according to the customers' needs, with the provider's computer capabilities pooled together to serve numerous clients. The client has little control or awareness of the specific location of the delivered resources but may be able to define location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth.

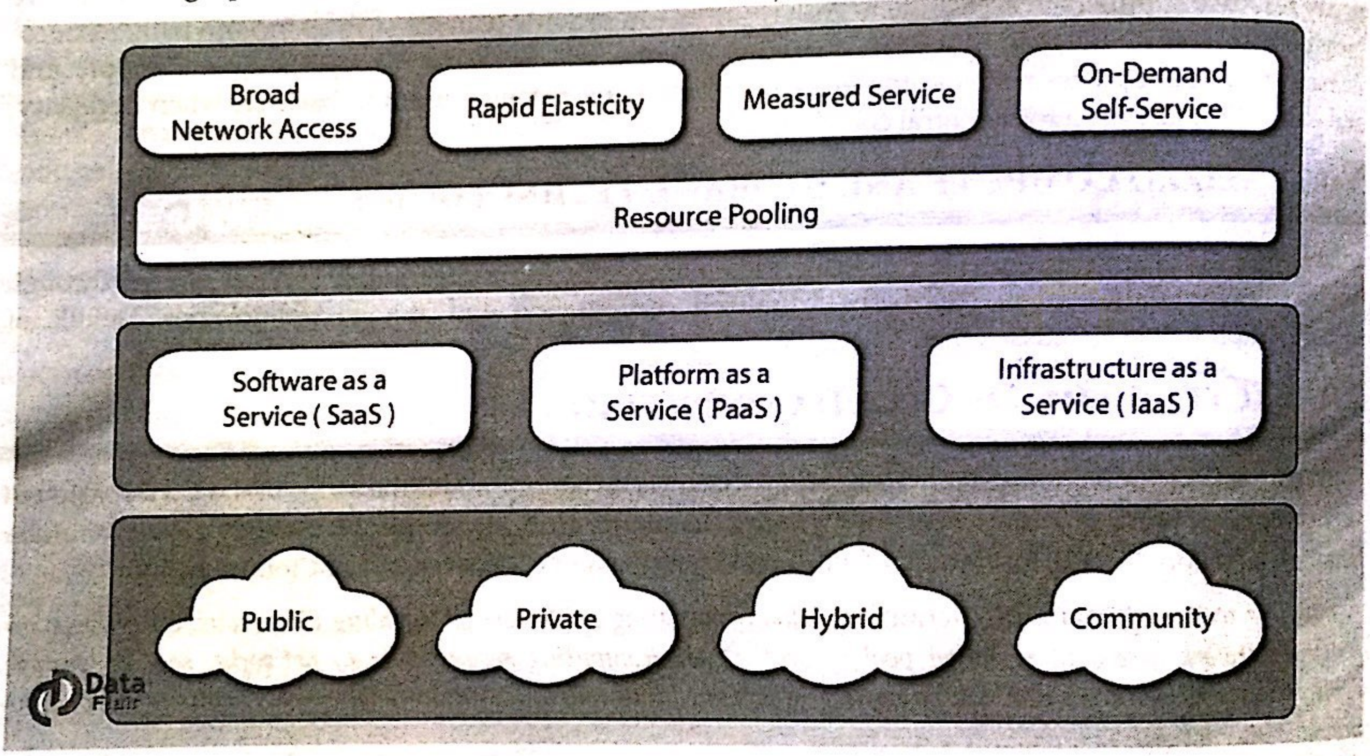


Figure 1.5: Depiction of the cloud definition by NIST

Rapid Elasticity

Capabilities of cloud can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

Measured Service

Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (*e.g., storage, processing, bandwidth, and active use account*). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

Advantages and Disadvantages of Cloud Computing

Advantages

Cloud computing offers numerous advantages both to end-users and businesses of all sizes. The obvious huge advantage is that we no more have to support the infrastructure or have the knowledge necessary to develop and maintain the infrastructure, development environment, or application, as were things up until recently. The burden has been lifted and someone else is taking care of all that. Businesses are now able to focus on their core business by outsourcing all the hassle of IT infrastructure. Some of the most important advantages of cloud computing are discussed here. Those will include both a company's and an end user's perspectives.

1. **Cost efficiency**. This is the biggest advantage of cloud computing, achieved by the elimination of the investment in stand-alone software or servers. By leveraging the cloud's capabilities, companies can save on licensing fees and at the same time eliminate overhead charges such as the cost of data storage, software updates, management, etc. The cloud is in general available at much cheaper rates than traditional approaches and can significantly lower the overall IT expenses. At the same time, convenient and scalable charging models have emerged (such as one-time payment and pay-as-you-go), making the cloud even more attractive.

2. **Convenience and continuous availability**. Public clouds offer services that are available wherever the end-user might be located. This approach enables easy access to information and accommodates the needs of users in different time zones and geographic locations. As a side benefit, collaboration booms since it is now easier than ever to access, view, and modify shared documents and files.

3. **Increased storage capacity**. The cloud can accommodate and store much more data compared to a personal computer and in a way offers almost unlimited storage capacity. It eliminates worries about running out of storage space and at the same time it spares businesses the need to upgrade their computer hardware, further reducing the overall IT cost.

4. **Backup and recovery**. The process of backing up and recovering data is simplified since those now reside on the cloud and not on a physical device. The various cloud providers offer reliable and flexible backup/recovery solutions. In some cases, the cloud itself is used solely as a backup repository of the data located in local computers.

5. **Cloud is environmentally friendly**. The cloud is, in general, more efficient than the typical IT infrastructure and it takes fewer resources to compute, thus saving energy. For example, when servers are not used, the infrastructure normally scales down, freeing up resources and consuming less power. At any moment, only the resources that are truly needed are consumed by the system.

6. **Resiliency and redundancy**. Cloud deployment is usually built on a robust architecture thus providing resiliency and redundancy to its users. The cloud offers automatic failover between hardware platforms out of the box, while disaster recovery services are also often included.

7. **Scalability and performance**. The cloud systems are generally quick to deploy and have ease of integration. Such systems can be up and running in a very short period, making quick deployment a key benefit. On the same aspect, the introduction of a new user in the system happens instantaneously, eliminating waiting periods.

8. **Device diversity and location independence**. Cloud computing services can be accessed via a plethora of electronic devices that can have access to the Internet. These devices include not only traditional PCs but also smartphones, tablets, etc. With the cloud, the "Bring Your Own Device" (BYOD) policy can be easily adopted, permitting employees to bring personally owned mobile devices to their workplace.

PROS & CONS OF CLOUD AT A GLANCE

PROS

- Lower upfront costs and reduced infrastructure costs.
- High computing capabilities.
- Reliability due to multiple redundant sites.
- Scalable and Flexible; Growth of applications becomes easy.
- Scale up or down at short notice.
- Only pay for what you use.
- No worries on maintenance as that's the provider's concern.
- Accessible through various devices (PCs, Mobiles, Tablets, etc.)
- Overall environmental benefit (lower carbon emissions) of many users efficiently sharing large systems.

CONS

- Higher ongoing operating costs if any firm needs a large number of services.
- Greater dependency on service providers so the risk of being locked into proprietary or vendor-recommended systems.
- Potential privacy and security risks of putting valuable data on someone else's system in an unknown location.
- Dependency on an Internet connection.

QUESTIONS TO THINK

- What happens if your supplier suddenly decides to stop supporting a product or system you have come to depend on? What is the chance of such a thing happening?
- If lots of people migrate to the cloud, what does that imply for the future development of the Internet?

Disadvantages

As made cloud computing is a tool that offers enormous benefits to its adopters. However, being a tool, it also comes with its set of problems and inefficiencies.

1. **Security issues**. How safe is your data? Cloud computing means Internet computing. So, we should not be using cloud computing applications that involve using or storing data that we are not comfortable having on the Internet. That being said, established, reliable cloud computing vendors will have the latest, most sophisticated data security systems possible as they want your business and realize that data security is a big concern.
2. **Dependency**. One of the major disadvantages of cloud computing is the implicit dependency on the provider. This is what the industry calls vendor lock-in since it is difficult, and sometimes impossible, to migrate from a provider once we have rolled with him. If a user wishes to switch to some other provider, then it can be really painful and cumbersome to transfer huge data from the old provider to the new one. This is another reason why we should carefully and thoroughly contemplate all options when picking a vendor.
3. **Technical difficulties and downtime**. Certainly, the smaller business will enjoy not having to deal with the daily technical issues and will prefer handing those to an established IT company, however, we should keep in mind that all systems might face dysfunctions from time to time. Outage and downtime are possible even to the best cloud service providers, as the past has shown.
4. **Limited control**. Since the applications and services run on remote, third-party virtual environments, companies and users have limited control over the function and execution of the hardware and software. Moreover, since remote software is being used, it usually lacks the features of an application running locally.

5. **Inflexibility**. Be careful when choosing a cloud computing vendor so that the vendor is not locking your business into using their proprietary applications or formats. Also, make sure that the cloud computing users can be added or removed as necessary with the growth or contraction of the business.

6. **Increased weakness**. Related to the security and privacy issues mentioned before, note that cloud-based solutions are exposed on the Internet and are thus a more vulnerable target for malicious users and hackers. Nothing on the Internet is completely secured and even the biggest players suffer from serious attacks and security breaches. Due to the interdependency of the system, if there is a compromise in one of the machines where data is stored, there might be a leakage of personal information to the world.

CapEx and OpEx

Traditionally, organizations depended on in-house data center models, which required a significant financial commitment in the form of purchasing space, equipment, software, and a crew to manage and maintain everything. These businesses were guaranteed robust data monitoring and protection, but at a hefty expense. Today's businesses want the same level of protection and supervision as in the past, but also greater flexibility and cost-effective solutions. The Cloud is the solution.

Capital Expenditures, often known as **CapEx**, are the funds spent by your company on fixed assets. Whole systems and servers, printers and scanners, and air conditioners and generators are examples of IT equipment that belong under this category. You purchase these things once and they will benefit your firm for many years. Maintenance of such goods is also termed CapEx because it increases their usefulness and lifespan. These are large investments, generally a one-time buy, that are meant to help your company over a longer period. The procurement cost appears on your balance sheet and is subsequently depreciated over time. Small firms sometimes struggle with CapEx expenditure since it typically comprises the high cost of a one-time purchase. The more money spent on capital expenditures, the less free cash flow there is for the remainder of the firm.

We also have Operating Expenses, abbreviated as OpEx, which are the cash that an organization requires to operate its day-to-day operations. Subscription fees for cloud services, website hosting, and web domain registrations are examples of such expenditures. OpEx is expenses for pay-as-you-go things that are taken from your business's income as they occur. For a variety of reasons, most firms now prefer OpEx over CapEx. The amount paid to a vendor for leasing is OpEx; it is incurred as part of the day-to-day operations of the firm.

The cloud service provider invests in CapEx, while their customer invests in a pay-as-you-go approach or OpEx. This implies that the data will be kept in the cloud service provider's data, alongside the data of many other organizations, in what is known as a public cloud. Your provider will handle the maintenance and updates. This is a suitable strategy for firms that do not have the resources of an in-house IT team.

If a company requires a private cloud, it must develop its own IT infrastructure using the CapEx model. Following an OpEx, they can then provide the resources to other users or divisions inside the organization. A company can also use resources from both private and public clouds, which is known as a Hybrid Cloud solution. A hybrid Cloud may be controlled by a single solution and can provide your company with greater cost-control flexibility.

Difference between CapEx and OpEx

CapEx	OpEx
Investment of big amounts on types of equipment (e.g., Computer system, software) and significant staffs required for integration.	Buy on an as-needed basis, so you are never investing in something that can't be used profitably immediately.
Equipment failures need to be managed by the company. Inhouse maintenance staff are required.	Maintenance is the responsibility of the cloud service provider and in most of the cases it's automatic. Customers don't need to worry.

Overprovisioning and wasted capacity. Getting stuck with the capacity you don't need.	Customers can buy what they need and can resize on the fly depending upon the need.
HVAC (Heating, ventilation, and air conditioning), electricity, facility rental, insurances are the responsibility of the company.	None of those are to be considered by the customer.
Configuration errors can be costly and difficult to remediate.	Software-defined services can be easily configured.
Taking a very long time, usually through a difficult process, to adopt a new capacity.	Stand up assets as you need them, delete what you no longer need.
Significant investment is required on monitoring tools with limited automation.	Monitoring and alerting are built into the service with significant automated remediation opportunities.
The total cost of ownership (TCO) contains many hidden charges.	TCO is completely transparent.

COMPONENTS OF CLOUD COMPUTING

The right implementation of various components is required for successful cloud computing deployment. Cloud computing will not be viable without any of these components. These elements are difficult to accomplish by one person alone. Cloud computing will necessitate people with a variety of skills, experiences, and backgrounds. It is no surprise that cloud computing is a very expensive business because it will demand more workers in the sector. However, even with the additional costs that a corporation may incur, the benefits given by cloud computing far outweigh the initial costs. Because of a lack of resources, some might turn to a cloud computing vendor, but others have the means to develop their cloud computing apps, platforms, and hardware. However, components must be constructed with the intention of excellent performance in any case.

The End User – The Client

The client will use the hardware components, application, and everything else created for cloud computing. Nothing will be possible without the client. The client might be in one of two forms: a hardware component or a software and hardware combination. Although it is a frequent misconception that cloud computing relies exclusively on the Internet, certain systems require pre-installed programs to enable a smooth transition. The hardware, on the other hand, will be the launch platform for everything. Optimization is based on two fronts: local hardware capacity and software security. The program will run smoothly thanks to improved hardware and security.

The Service – the Functions in Cloud Computing

Cloud computing is always used for a specific reason. Businesses have adopted cloud computing as a simpler method to perform business procedures, which is one of the key reasons for its popularity. Cloud computing is all about processes, and the services that are launched. As a result, cloud computing must always deal with processes that have a predictable outcome. The optimization of services is dependent on two factors: the application's correct development and the end-user. The user may utilize the service in situations when their device has a significant impact on their experience.

The Application – Backbone of Service

The application is frequently confused with the service. Although it is somewhat right because it supplies the functionalities, the application is fundamentally different because the service is implemented through the application. This is where software developers must concentrate their efforts to ensure that the program performs as planned. The application's optimization is based on the creators' real coding. The program may perform as predicted after comprehensive load handling, security, and functionality testing

The Platform – "Soft Infrastructure" for the Application

The application is directly connected to the server in conventional web pages or apps that do not deal with cloud computing. In cloud computing, the application is launched to another application called the platform. Ajax (Asynchronous JavaScript and XML) or Ruby on Rails are common examples of programming platforms. Those who choose to use cloud computing services must now adhere to a set of programming languages that can be used on the platform. Although most programming languages may be run on a variety of platforms, cloud computing requires a sophisticated application with real-time updating capabilities.

The Storage – The Warehouse of Cloud Computing

Storage allows the program to store all its data as well as the functionalities that the service may supply. The data and information are stored in the storage facility. Storage optimization is dependent on how well the storage facility is protected from various attacks and the availability of backups. Could computing is always about service consistency and availability, which means that storage must be available at all times.

The infrastructure – The Backbone of Cloud Computing

Every function, service, and the ability of storage to provide the needed data is only possible through optimized infrastructure. This could be considered as the platform behind the storage as the infrastructure helps the storage deal with load issues. The infrastructure is a platform wherein it weights the ability of the storage against the number of requests. The infrastructure can make some changes with techniques such as load balancing.

XaaS (ANYTHING AS A SERVICE)

XaaS is an abbreviation for Anything-as-a-Service (or sometimes also called Everything-as-a-Service). XaaS highlights how enterprises throughout the world are using the as-a-Service model to supply, well, everything. It defines a broad range of cloud computing and remote access services It acknowledges the large variety of goods, tools, and technology that are now supplied as a service to users over the internet. Any IT function, in essence, may be converted into service for corporate consumption. Instead of an upfront purchase or license, the service is paid for through a flexible consumption approach.

The list of XaaS examples is limitless. Many types of IT resources and services are being supplied in this manner. Cloud computing models are broadly classified into three types as Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). As XaaS is a broad word referring to the provision of anything as a service, Disaster recovery as a service (DRaaS), Communications as a service (CaaS), Network as a service (NaaS), Database as a service (DBaaS), Storage as a service (STaaS), Desktop as a service (DaaS), Monitoring as a service (MaaS), Marketing as a service, Healthcare as a service, Malware as a service, etc. can also be considered as further examples.

In the book *Cloud Computing and SOA Convergence in your Enterprise, David S. Linthicum* mentions 11 major categories or patterns of cloud computing technology.

1. Storage-as-a-service
2. Database-as-a-service
3. Information-as-a-service
4. Process-as-a-service
5. Application-as-a-service
6. Platform-as-a-service
7. Integration-as-a-service
8. Security-as-a-service
9. Management/governance-as-a-service
10. Testing-as-a-service
11. Infrastructure-as-a-service

1. **Storage-as-a-service** (*also known as disk space on demand*) allows any application that requires storage space to use a storage area that is physically located at a remote site but is conceptually a local storage resource. This is the most basic component of cloud computing, and it is a component or pattern that most other cloud computing components rely on.

2. **Database-as-a-service** allows you to take use of the services of a remotely hosted database, share it with other users, and have it performed as if it were local. Different suppliers offer different options, but the power is in the database technology, which costs thousands of dollars in hardware and software licensing.

3. **Information-as-a-service** is known as the capacity to consume any sort of information, hosted remotely, using a well-defined interface such as an API. Stock price information, address confirmation, and credit reporting are a few examples.

4. **Process-as-a-service** is a remote resource that may link numerous other resources, such as services and data, together to form business processes, either within the same cloud computing resource or remotely. A business process may be thought of as a meta-application that excels systems and combines critical services and information into a sequence to form a process. Because these processes are often easier to adapt than apps, those that use these on-demand process engines benefit from increased agility.

5. **Application-as-a-service**, sometimes known as software-as-a-service **(SaaS)**, refers to any application that is supplied to an end-user over the Web platform, often via a browser. While many people connect application-as-a-service with corporate software like Salesforce SFA, office automation programs like Google Docs, Gmail, and Google Calendar are also apps-as-a-service.

6. **Platform-as-a-Service (PaaS)** is a full platform that includes application development, interface development, database development, storage, testing, and other services that are supplied to subscribers via a remotely hosted platform. Modern platform-as-a-service providers, based on the classic time-sharing model, offer the opportunity to construct enterprise-class applications for local or on-demand usage for a minimal subscription fee or for free.

7. **Integration-as-a-service** refers to the capacity to supply an entire integration stack from the cloud, including application interface, semantic mediation, flow control, and integration design, among other things. In essence, integration-as-a-service combines the majority of the capabilities and functionalities of traditional Enterprise Application Integration (EAI) technology with the convenience of a cloud-based service.

8. **Security-as-a-service** refers to the capacity to offer essential security services remotely through the Internet. While the majority of security services are basic, increasingly advanced services such as identity management are being offered through the cloud platform.

9. **Management/Governance-as-a-service** is any on-demand service that provides the ability to manage one or more cloud services. Simple things like topology, resource usage, virtualization, and uptime management are classic examples. Governance solutions are also becoming accessible, allowing users to enforce established regulations on data and services, for example.

10. **Testing-as-a-service** refers to the ability to test local or cloud-delivered systems utilizing remotely hosted testing tools and services. While a cloud service necessitates testing in and of itself, testing-as-a-service solutions may test other cloud apps, Web sites, and internal corporate systems without requiring a hardware or software footprint within the company.

11. **Infrastructure-as-a-service** (IaaS) can also be referred to as a 'datacenter-as-a-service' because of its ability to remotely access computing resources. In essence, you lease a physical server that is yours to do with as you like and serves as your data center, or at least a portion of one, for all practical reasons. The difference between this method and more widespread cloud computing is that instead of utilizing an interface and a metered service, you have complete access to the system and its software.

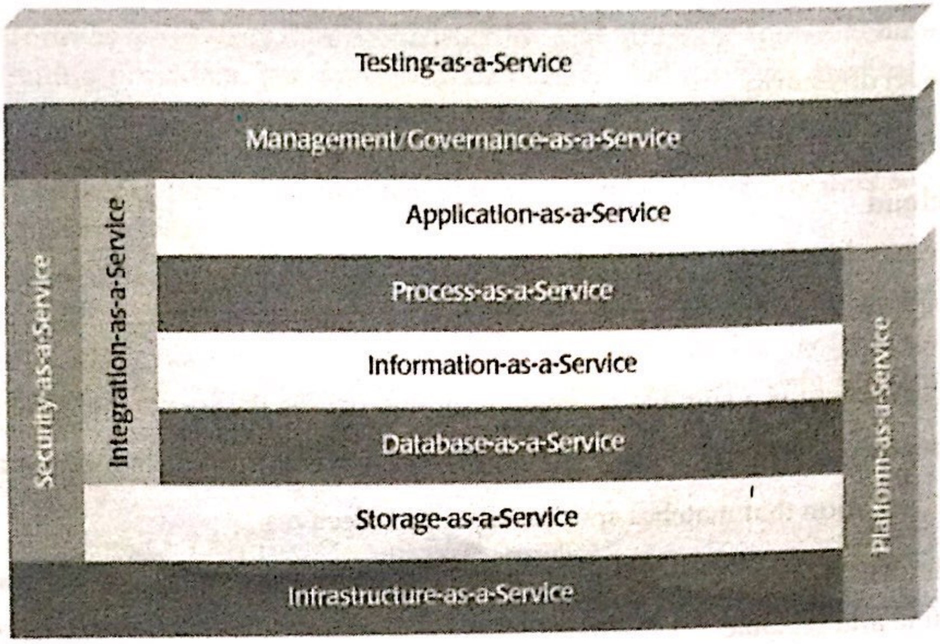


Figure 1.6 Patterns of cloud computing technology

TYPES OF CLOUD AND ITS CLOUD SERVICES

Cloud computing is a wide word that refers to a set of services that provide organizations with a low-cost way to expand their IT capacity and usefulness. Businesses may select where, when, and how they employ cloud computing to provide an effective and dependable IT solution based on their individual needs. We will look at the many forms of cloud computing, including the three primary deployment methods and the cloud services that may be hosted in these settings.

TYPES OF CLOUD

There are three primary types of cloud environments, which are also known as cloud deployment models. Depending on their individual needs, businesses can operate apps on public, private, or hybrid clouds.

Public Cloud

A public cloud environment is held by an outsourced cloud provider and is available to numerous enterprises on a pay-per-use basis over the Internet. This deployment strategy offers services and infrastructure to enterprises looking to reduce IT operations expenses, but the cloud provider is in charge of creating and maintaining the resources. Public clouds are perfect for small and medium-sized organizations that have a limited budget and need a rapid and easy platform to deploy IT resources.

Pros of a public cloud

- Easy scalability
- No geographical restrictions
- Cost-effective
- Highly reliable
- Easy to manage

Cons of a public cloud

- Not considered the safest option for sensitive data

Private Cloud

This cloud deployment architecture is a one-of-a-kind infrastructure controlled by a single company. It provides a more regulated environment with more centralized access to IT resources inside the organization. This model can be hosted outside or controlled in-house. Although private cloud hosting might be costly, it may provide a better level of security and greater autonomy in customizing the storage, networking, and computing components to meet the needs of bigger enterprises.

Pros of a private cloud

- Improved level of security
- Greater control over the server
- Customizable

Cons of a private cloud

- Harder to access data from remote locations
- Requires IT expertise

Hybrid Cloud

A hybrid cloud environment is a suitable alternative for enterprises that want the benefits of both private and public cloud deployment options. A hybrid cloud model, which combines the two models, delivers a more customized IT solution that matches specific business objectives.

Pros of a hybrid cloud

- Highly flexible and scalable
- Cost-effective
- Enhanced security

Cons of a hybrid cloud

- Communication at the network level may be conflicted as it's used in both private and public clouds.

Community Cloud

Community clouds, though not as widely used as the other three types, are a collaborative, multi-tenant platform utilized by numerous diverse businesses to share the same applications. Users are often from the same industry or field and have similar concerns about security, compliance, and performance. A community cloud is a private cloud that operates similarly to a public cloud. The platform is privately maintained, either in a data center or on-premises. Within that environment, authorized users are then separated. Government agencies, healthcare organizations, financial services corporations, and other professional communities frequently employ these installations.

Pros of a community cloud

- Scalability
- Compliant with any industry regulations
- Highly flexible

Cons of a community cloud

- Shared storage and bandwidth
- Data security can be a concern due to shared space

CLOUD SERVICES

Cloud computing service models are divided into three categories: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) (SaaS). There are evident distinctions between the three in terms of what they can provide a business in terms of storage and resource sharing, but they may also interact with one another to build a single comprehensive cloud computing platform.

1. **IaaS (Infrastructure as Service):** This is the most prevalent cloud computing service architecture since it provides the underlying infrastructure of virtual servers, networks, operating systems, and data storage devices. It provides the flexibility, stability, and scalability that many organizations desire from the cloud while eliminating the need for hardware in the office. This makes it excellent for small and medium-sized businesses seeking a cost-effective IT solution to support business growth. IaaS is a completely outsourced pay-for-use service that is accessible as a public, private, or hybrid infrastructure.

2. **PaaS (Platform-as-a-Service):** The infrastructure and software foundation is deployed here by cloud computing providers, but enterprises can design and run their applications. PaaS allows for the rapid and easy creation of web apps, and the service is versatile and resilient enough to sustain them. PaaS solutions are scalable and suited for corporate scenarios with several developers working on a single project. It is also useful in cases when an existing data source (such as a CRM program) must be used.

3. **SaaS (Software as a Service):** This cloud computing approach entails the distribution of software through the internet to numerous enterprises that pay via subscription or pay-per-use. It is a useful tool for CRM and apps that require a lot of online or mobile access, such as mobile sales management software. SaaS is handled from a single place, removing the need for enterprises to worry about maintaining it themselves, and thus is suitable for short-term initiatives.

Types of Cloud and cloud services will be discussed in more detail in unit 2.

Cloud Computing Data Storage

Cloud Storage is a service that allows saving data on an offsite storage system managed by a third party and is made accessible by a **web services API.**

STORAGE DEVICES

Storage devices can be broadly classified into two categories:

- Block Storage Devices
- File Storage Devices

Block Storage Devices: The block storage devices offer raw storage to the clients. These raw storages are partitioned to create volumes.

File Storage Devices: The file storage devices offer storage to clients in the form of files, maintaining their file system. This storage is in the form of Network Attached Storage (NAS).

CLOUD STORAGE CLASSES

Cloud storage can be broadly classified into two categories:

- Unmanaged Cloud Storage
- Managed Cloud Storage

Unmanaged Cloud Storage

Unmanaged cloud storage means the storage is preconfigured for the customer. The customer can neither format nor install his file system or change drive properties.

Managed Cloud Storage

Managed cloud storage offers online storage space on-demand. The managed cloud storage system appears to the user to be a raw disk that the user can partition and format.

Creating a Cloud Storage System

Various copies of data are stored on numerous servers in multiple locations by the cloud storage system. If one of the systems fails, all that is necessary is to update the reference to the object's storage location. The cloud provider aggregates storage assets into a cloud storage system using virtualization technologies. StorageGRID is the name of the technology, which produces a virtualization layer that consolidates storage from several storage devices into a single management system. It can also use the Internet to handle data from CIFS and NFS file systems.

The following diagram shows how StorageGRID virtualizes the storage into storage clouds:

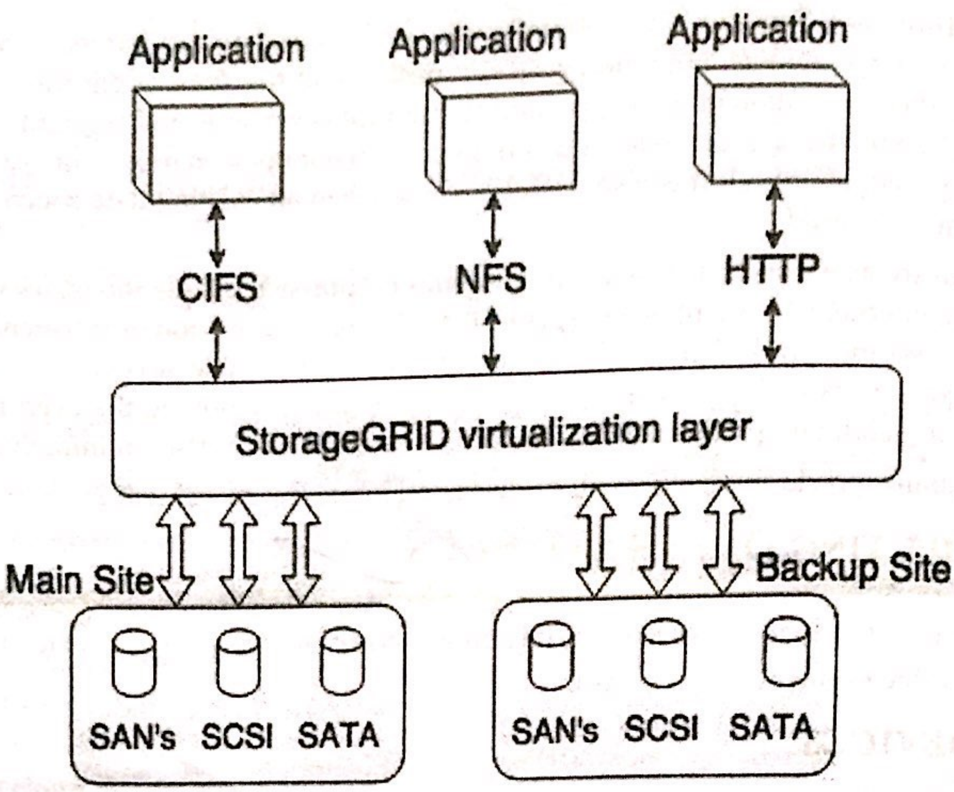


Figure 1.7: storage clouds virtualization with StorageGRID

CHALLENGES

Storing the data in the cloud is not that simple task. Apart from its flexibility and convenience, it also has several challenges faced by the customers. The customers must be able to:

- Get provision for additional storage on-demand.
- Know and restrict the physical location of the stored data.
- Verify how data was erased.
- Have access to a documented process for disposing of data storage hardware.
- Have administrator access control over data.

CLOUD SERVICE REQUIREMENTS

Cloud computing consists of hardware and software resources made available on the Internet as managed third-party services. Cloud services requirements typically provide access to advanced software applications and high-end networks of server computers. Service providers create cloud computing systems to serve common business or research needs.

Examples of cloud computing services include:

1. **Virtual IT.** It configures and utilizes remote third-party servers as extensions to a company's local IT network.
2. **Software.** They utilize commercial software applications or develop and remotely host custom-built applications.
3. **Network storage.** The backup or archive data across the Internet to a provider without needing to know the physical location of storage.

Cloud computing systems in general are designed for scalability to support large numbers of customers and surges in demand. Service providers are responsible for installing and maintaining core technology within the cloud. Some customers prefer this model because it limits their manageability burden. However, customers cannot directly control system stability in this model and are highly dependent on the provider instead.

Cloud computing systems are normally designed to closely track all system resources, which enables providers to charge customers according to the resources each consumes. Some customers will prefer this so-called metered billing approach to save money, while others will prefer a flat-rate subscription to ensure predictable monthly or yearly costs. Using a cloud computing environment generally requires you to send data over the Internet and store it on a third-party system. The privacy and security risks associated with this model must be weighed against alternatives.

CLOUD COMPUTING PLANNING

It is vital to analyze your business requirements before deploying apps to the cloud. The following are some of the issues to think about:

- Data Security and Privacy Requirement
- Budget Requirements
- Type of cloud - public, private, or hybrid
- Data backup requirements
- Training requirements
- Dashboard and reporting requirements
- Client access requirements
- Data export requirements

It is crucial to have well-compiled planning to achieve all of these needs. Various planning processes that an organization must go through before transferring their entire operation to the cloud are addressed here. Each of these planning phases is described with the help of the following diagram:

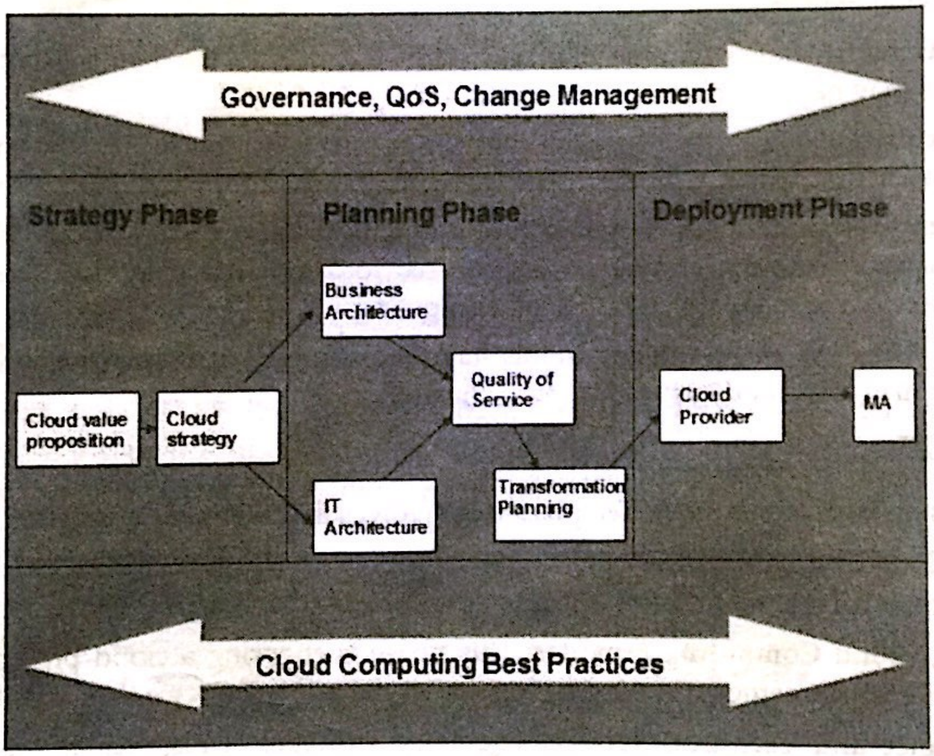


Figure 1.8: Cloud Computing Planning

Strategy Phase

In this phase, we analyze the strategy problems that customers might face. There are two steps to perform this analysis:

- Cloud Computing Value Proposition
- Cloud Computing Strategy Planning

Cloud Computing Value Proposition

In this book, we look at the elements that influence clients when using cloud computing and the prima challenges they may tackle. These are the most important factors:

- IT management simplification
- Operation and maintenance cost reduction
- Business mode innovation
- Low-cost outsourcing hosting
- High service quality outsourcing hosting.

All of the above analysis helps in decision-making for future development.

CLOUD COMPUTING STRATEGY PLANNING

The plan is developed based on the findings of the previous step's study. In this stage, a strategy paper created based on the circumstances that a client could encounter when using cloud computing.

Planning Phase

This stage analyzes faults and dangers in the cloud application to guarantee that cloud computing satisfying the customers' business objectives. The following planning steps are included in this phase:

- Business Architecture Development
- IT Architecture Development
- Requirements on Quality-of-Service Development
- Transformation Plan Development

1. **Business Architecture Development:** In this stage, we identify the dangers that cloud computing applications may have from a business standpoint.
2. **IT Architecture Development:** We determine the applications that support business operations, as well as the technology necessary to support enterprise applications and data systems, in this stage.
3. **Requirements on Quality-of-Service Development:** Non-functional needs like dependability security, disaster recovery, and so on are referred to as quality of service. These non-functional variables have a role in the success of cloud computing.
4. **Transformation Plan Development:** In this stage, we create all of the plans we will need to convert our present firm to a cloud computing environment.

Deployment Phase

This phase focuses on both of the above two phases. It involves the following two steps:

- Selecting Cloud Computing Provider
- Maintenance and Technical Service

1. **Selecting Cloud Computing Provider:** This phase is choosing a cloud provider based on their Service Degree Agreement (SLA), which specifies the level of service they will deliver.
2. **Maintenance and Technical Service:** The cloud provider is responsible for maintenance and technical support. They must guarantee that the services they provide are of high quality.

CLOUD COMPUTING TECHNOLOGIES

Certain technologies are at work behind the scenes of cloud computing platforms, making them more adaptable, dependable, and useful. The following technologies are listed:

- Virtualization,
- Service-Oriented Architecture (SOA)
- Grid Computing
- Utility Computing

Virtualization

Virtualization is a method that allows several organizations to share a single physical instance of an application or resource. It accomplishes this by giving a physical resource a logical name and delivering a reference to that physical resource when needed. Details of virtualization are discussed in Unit IV.

Service-Oriented Architecture

Regardless of the manufacturer, product, or technology, Service-Oriented Architecture enables the usage of applications as a service for other applications. As a result, data may be exchanged across apps from various manufacturers without the need for further programming or service modifications. The graphic below depicts the cloud computing service-oriented architecture.

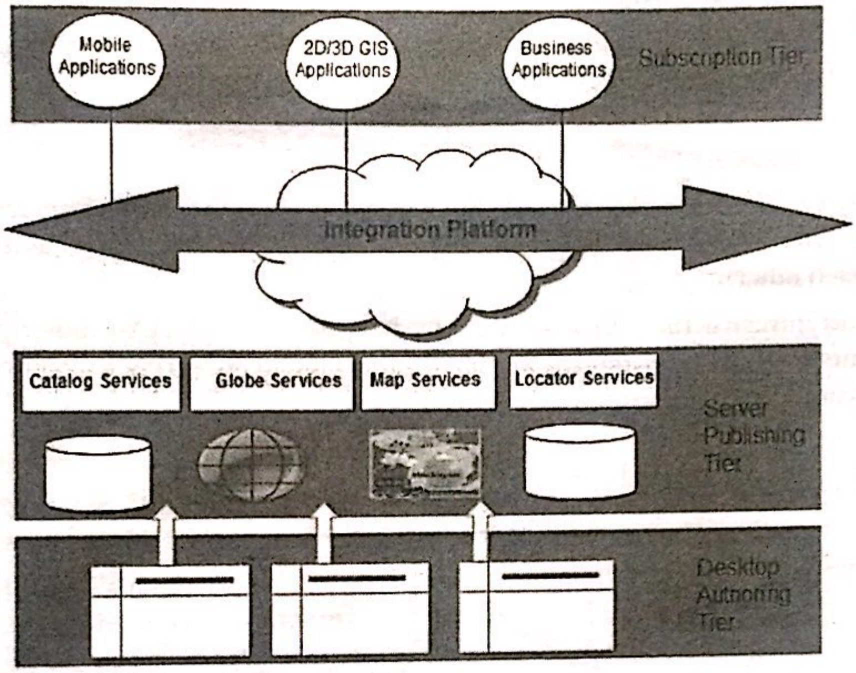


Figure 1.9: Service-oriented architecture

Grid computing

Grid computing is a type of distributed computing in which several computers from various places collaborate to achieve a shared goal. These computing resources are diverse and distributed around the globe. Grid computing divides large tasks into smaller chunks that are dispersed across the grid's CPUs.

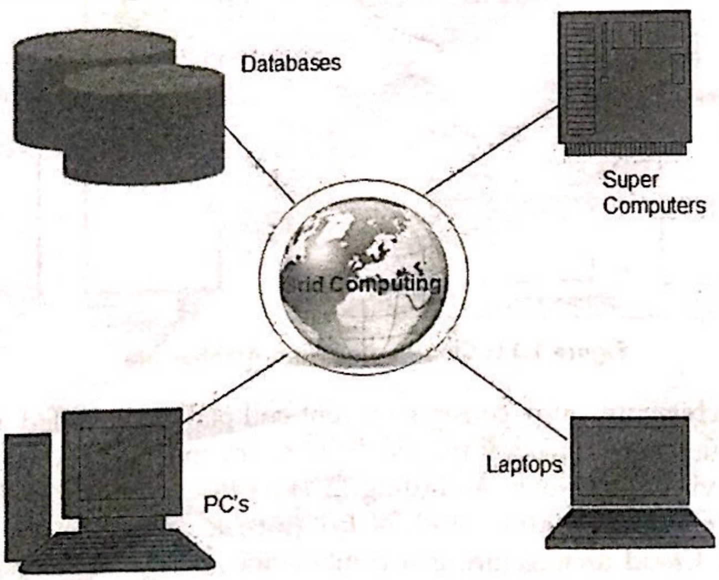


Figure 1.10: Basics of Grid Computing

Utility Computing

Utility computing is a service delivery paradigm in which a service provider makes computing resources and infrastructure management available to customers on an as-needed basis and charges them based on consumption rather than a fixed cost. The utility model, like other kinds of on-demand computing, attempts to optimize resource efficiency and/or decrease related costs. The Pay-per-Use concept is used in utility computing. As a metered service, it provides computing resources on demand. Utility computing is the foundation of cloud computing, grid computing, and managed IT services.

CLOUD COMPUTING ARCHITECTURE

The cloud infrastructure is closely related to its architecture & comprises many cloud components which are loosely connected.

The broad divisions of cloud architecture are:

- Front-end
- Back-end

It is the responsibility of the back end to provide the security of data for cloud users along with the traffic control mechanism. The server also provides the middleware which helps to connect devices & communicate with each other.

Businesses used cloud infrastructures to work with these applications. Unlike subscription-based models of pricing, the payment structure of the cloud enables the user to subscribe to vendor services & cloud infrastructures are paid on a 'pay-per-use basis.

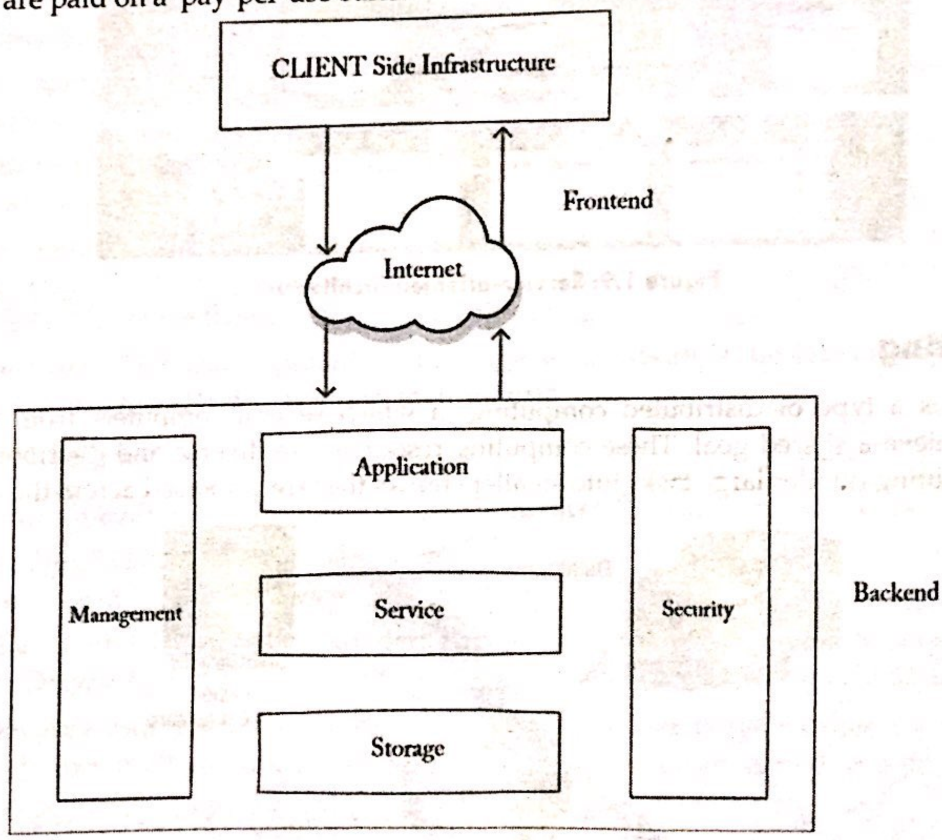


Figure 1.11: Cloud Computing Architecture

The cloud technology architecture also consists of front-end platforms called the cloud client which comprises servers, thin & fat client, tablets & mobile devices. The interaction is done through middleware or via a web browser or virtual sessions. According to *Jason Bloomberg of ZapThink*, the cloud-oriented architecture can essentially be the building block of IoT (Internet of Things) in which anything can be connected to the Internet. Cloud architecture is a combination of both services-oriented architecture & event-driven architecture. So, cloud architecture encompasses all elements of the cloud environment.

Cloud Computing Infrastructure

Servers, storage devices, networks, cloud management software, deployment software, and platform virtualization make up cloud infrastructure.

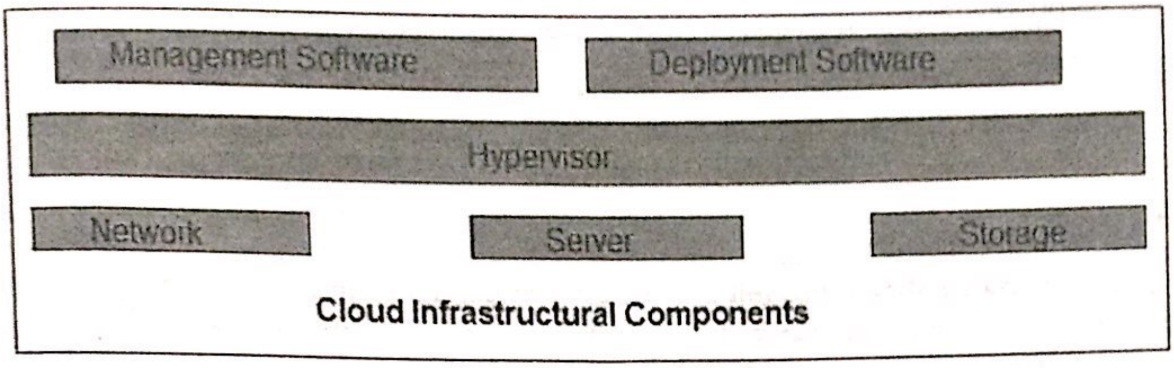


Figure 1.12: Cloud Infrastructural Components

1. **Management Software:** It helps to maintain and configure the infrastructure.
2. **Deployment Software:** It helps to deploy and integrate the application on the cloud.
3. **The hypervisor** is a firmware or low-level program that acts as a Virtual Machine Manager. It allows sharing the single physical instance of cloud resources between several tenants.
4. **Network:** It is a crucial part of cloud infrastructure. It enables cloud services to be connected through the Internet. It is also feasible to supply networks as a service through the Internet, allowing customers to define network routes and protocols.
5. **Server:** The server facilitates resource sharing computations and provides other services such as resource allocation and de-allocation, resource monitoring, and security.
6. **Storage:** Multiple storage replicates are kept in the cloud. Cloud computing is more dependable because if one of the storage resources fails, data may be pulled from another.

Infrastructural Constraints

Fundamental constraints that cloud infrastructure should implement are shown in the following diagram:

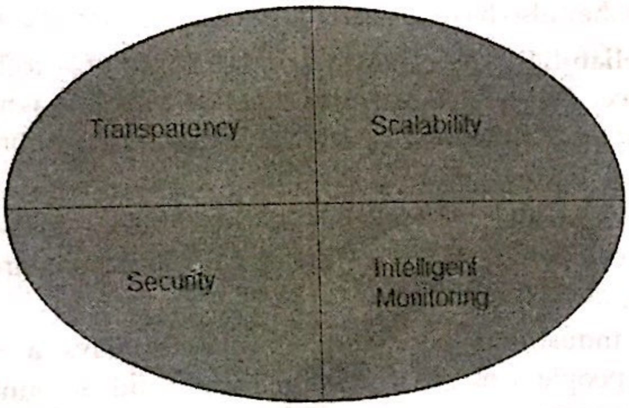


Figure 1.13: Fundamental Constraints

Transparency

In a cloud setting, virtualization is essential for sharing resources. However, the demand cannot be met with a single resource or server. As a result, transparency in resources, load balancing, and application is required so that we can grow them as needed.

Scalability

Because it includes setup complexity or even re-architecting the network, scaling up an application delivery solution is not as simple as scaling up an application. As a result, application delivery solutions must be scalable, which necessitates the use of virtual infrastructure that allows resources to be quickly supplied and de-provisioned.

Intelligent Monitoring

Application solution delivery will need to be capable of intelligent monitoring to ensure transparency and scalability.

Security

The cloud's big data center should be built with security in mind. In a mega data center, the control node, which serves as an entrance point, must also be secure.

CHALLENGES FOR THE CLOUD COMPUTING

Cloud computing is used for enabling global access to mutual pools of resources such as services, apps, data, servers, and computer networks. It is done on either a third-party server located in a data center or a privately-owned cloud. This makes data-accessing contrivances more reliable and efficient, with nominal administration effort. Because cloud technology depends on the allocation of resources to attain consistency and economy of scale, similar to a utility, it is also fairly cost-effective, making it the choice for many small businesses and firms.

But there are also many challenges involved in cloud computing, and if you are not prepared to deal with them, you will not realize the benefits. Here are six common challenges you must consider before implementing cloud computing technology.

1. **Security and Data privacy:** This is the major issue that is related to storing and securing data. Security risks of cloud computing have become the top concern. Sensitive and personal information that is kept in the cloud should be defined as being for internal use only, not to be shared with third parties.
2. **Cost:** Cloud computing itself is affordable but tuning the platform according to the company's needs can be expensive. Companies can save some money on system maintenance, management, and acquisitions. But they also have to invest in additional bandwidth.
3. **Service Provider Reliability:** The capacity and capability of a technical service provider are as important as the price. The service provider must be available when you need them. The main concern should be the service provider's sustainability and reputation.
4. **Downtime:** Downtime is a significant shortcoming of cloud technology. No seller can promise a platform that is free of possible downtime. Cloud technology makes small companies reliant on their connectivity, so companies with an untrustworthy internet connection probably want to think twice before adopting cloud computing.
5. **Password Security:** Industrious password supervision plays a vital role in cloud security. However, the more people you are accessing your cloud account with, the less secure it is. Anybody aware of your passwords will be able to access the information you store there.
6. **Vendor lock-in:** Entering a cloud computing agreement is easier than leaving it. "Vendor lock-in" happens when altering providers is either excessively expensive or just not possible. It could be that the service is nonstandard or that there is no viable vendor substitute.
7. **Compliance Concerns:** There are many compliance issues affecting cloud computing based on the type of data and application for which the cloud is being used.
8. **Lack of standards:** Clouds have documented interfaces; however, no standards are associated with these and thus it is unlikely that most clouds will be interoperable. The open grid forum is developing an open cloud computing Interface to resolve this issue and the Open Cloud

Consortium is working on cloud computing standards and practices. The findings of these groups will need to mature, but it is not known whether they will address the needs of the people deploying the services and the specific interfaces; these services need. However, keeping up to date on the latest standards as they evolve will allow them to be leveraged if applicable.

9. **Continuously evolving:** User requirements are continuously evolving, as are the requirements for interfaces, networking, and storage. This means that a "cloud", especially a public one, does not remain static and is also continuously evolving.

CLOUD AND DYNAMIC INFRASTRUCTURE

Cloud computing is an extension of existing corporate computing systems dating back to the first instance of networked computing, rather than a new revolution. The difference nowadays is that virtualization has advanced in practically every facet of the data center. A dynamic infrastructure is a collection of data center resources such as computation, networking, and storage that can provide and adapt itself automatically as workload demands vary. IT administrators can also manually manage these resources.

To identify, virtualize, categorize, and track data center resources, dynamic infrastructure relies largely on software. Regardless of whether they are physically located in one or more data centers, these resources are combined into pools. IT teams may build and manage various service tiers by categorizing data center resources, ensuring that more demanding applications receive more computing and storage resources. Dynamic infrastructure software automatically distributes resources from suitable pools to satisfy workload demands. When workload demands grow, the software adds resources to the pool, and when workload needs drop, the program returns resources to the pool – a process known as workload balancing.

This cloud dynamic paradigm must be capable of intercepting application and data traffic, interpreting the present context, and instructing the cloud infrastructure on how to serve the request in the most efficient manner possible. Scalability, flexibility, extensibility, and management are among these needs.

The question now is, what does this new dynamic computing architecture look like, and what is necessary to qualify as a "cloud" above and beyond the ordinary tools we have today? Integration of all parts of the infrastructure to build the cloud is a critical strategic decision. This includes everything from the bare metal to the users to all of the elements in between. In addition, depending upon the job, there are several methods to see the interplay of various functions inside the architecture. Several functional component blocks make up the cloud computing architecture. Consider computational resources or deployment environments, which are grouped into certain tiers of a pyramid as seen in the diagram. The depth of technical skills required to create or deploy each layer is represented by the layer's breadth.

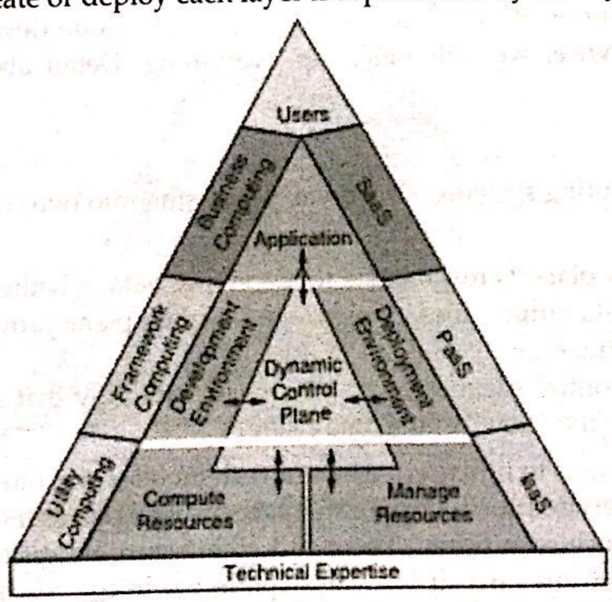


Figure 1.14: Computational resources or deployment environments grouped into layers

Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) are the tiers of the pyramid. Users access applications at the top of the pyramid, while a dynamic control plane in the middle, traverses all others and offers real-time communication, information coordination, and flow control across the levels. Each component must exist in some state or another to optimize the benefit of cloud architecture. For cloud environments to be operationally efficient and on-demand, dynamic control plane elements, for example, are required at every tier of the cloud architecture.

DESIGNING THE CLOUD

All computer resources obtained through cloud computing are delivered as services, such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), Software as a Service (SaaS), and so on. As indicated in the diagram, the cloud design may also be described as breaking down application-related components in such a manner that they may be supplied as a service. Before designing an application, it is important to understand the phrase "Service Oriented Architecture" (SOA). Designing for the cloud entails efficiently leveraging computer resources over the Internet.

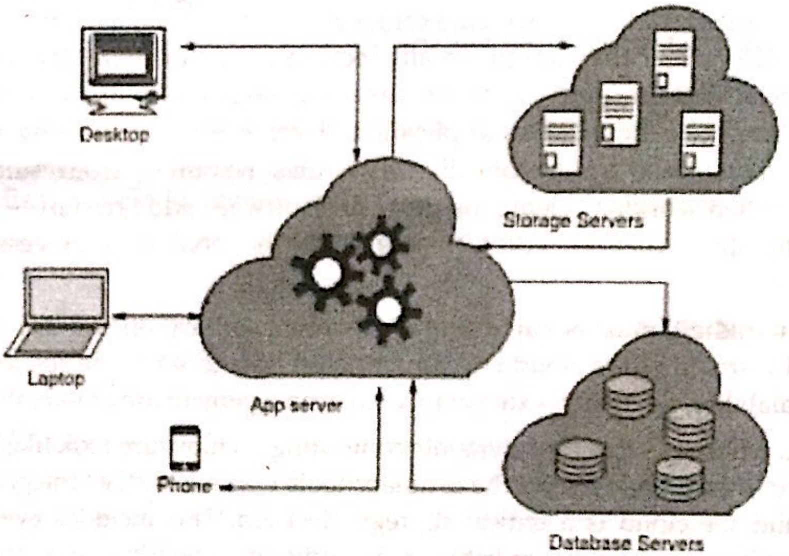


Figure 1.15: Basic components of the cloud

If we do not know how to utilize SOA efficiently, we will waste more money on computer resources that are not being used properly, and we will have trouble scaling up and down. If we build our application architecture to be more reliant on SOA, it will prove to be a smart cloud design, but we must also consider proper coding skills, otherwise, we will muck up everything. Detail about SOA is discussed in the upcoming unit.

Dynamic Control Plane

Traditional traffic and computing systems often break processing into two components: **the data** plane and **the control** plane.

1. **Data plane:** The data plane is responsible for receiving data, whether it is from a system or users, and returning it as data output, files, or answers. The data plane provides the foundation for traffic flow to and from destinations.
2. **Control plane:** The control plane is more concerned with how that data is managed with context and policy. It changes the "how" of the data plane.

Cloud architecture's main goal is to link users who are with mobiles and moving between LANs, WLANs, Internet connections, and services to the applications they use. The users can also travel across cloud centers depending on the business's needs. The cloud architecture necessitates a dynamic control plane that monitors the data and ensures that it is always connected in the best possible manner as hardware resources and servers are decommissioned, as applications are moved from development to production, and as entire applications are moved from the internal data center to a cloud provider. The dynamic control plane must be able to intercept traffic as it travels across the cloud, interpret the information, and direct the cloud architecture on how to connect the user to the right application instance as quickly as possible.

1. **Intercept traffic.** The dynamic control plane must be able to see all traffic between the user and the application, as well as traffic throughout the whole cloud platform. The dynamic control plane cannot accomplish any of the other things it needs to accomplish without being able to intercept traffic and data requests. It must be able to intercept both the metadata and the context of the communication, in addition to seeing the actual flow. The data plane, as well as any components that function inside it, must be visible to the dynamic control plane.

2. **Interpret the data.** The data and application flow information are insufficient. The dynamic control plane must be able to comprehend contextual variables in connection to individual requests, corporate policies, and other application and cloud traffic. To make sensible judgments at any given movement, the dynamic control plane must continually analyze the situation and policy.

3. **Instruct the cloud architecture.** The dynamic control plane must direct the architecture on how to link the two endpoints after it has all of the available information and has assessed the context. To adapt the present delivery model to satisfy the stated demands, the dynamic control plane must interface with the infrastructure and data plane. This might entail directing requests to a new instance of the program or a new data center, adjusting compression and encryption settings, or even telling other architectural components to construct or remove resources required to provide the program or data. The dynamic control plane may also be required to simply restrict access depending on policies and circumstances at any particular movement.

CLOUD ARCHITECTURE — IAAS

The basic layer of the cloud architecture, which is made up of basic components such as computation and management resources, needs the greatest technical expertise to create and install. The components most typically given by companies who supply IaaS solutions to their clients make up the basic basis upon which a cloud is created, as depicted in figure 1.14.

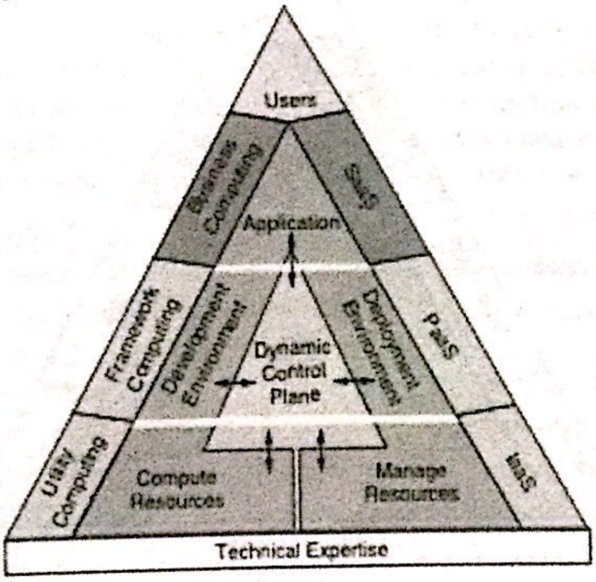


Figure 1.14*: Computational resources or deployment environments grouped into layers

Infrastructure as a Service (IaaS) is a cloud computing concept in which the whole infrastructure is provided on demand. This nearly invariably takes the shape of a virtualized infrastructure and infrastructure services that allow customers to deploy virtual machines as components that can be managed through a console.

The cloud provider manages the actual resources such as servers, storage, and networks, while the user manages the infrastructure built on top of those components. Each component in this layer of the architecture is responsible for supplying actionable data to the other components and completing particular activities to complete an auto-provisioning or decommissioning scenario effectively. IaaS is sometimes referred to as utility computing since it treats compute resources similarly to how utilities like energy are handled. When demand for capacity grows, the supplier allocates more computer resources. As demand for capacity falls, so does the number of computing resources available. This supports cloud architecture's on-demand as well as "pay-per-use" features.

There the two components of IaaS:

1. **Compute resources.** Compute resources are one of the most fundamental components of the cloud, consisting of bare-metal resources such as CPU, memory, and storage, which eventually power cloud-based apps. This might be a hosting service provider with hundreds, or thousands of server systems built and ready for members to utilize.
2. **Manage resources**. The components required to transform bare metal into useable server platforms with the requisite CPU, memory, and disk resources to run the programs that will be created on them are known as manage resources. Manage resources is also in charge of continuing to monitor resource requirements and ensuring that the program obtains all compute resources it requires, as well as transferring the program or locating new resources. This component is sometimes confused with virtual machine management or software provisioning systems, which can take bare metal and install operating systems, patches, and application logic, as well as apply higher-level network connection like IP addressing and many more.

CLOUD ARCHITECTURE — PAAS

Many apps are built on top of infrastructure services and operate on software platforms. These platforms, which might include environments like Oracle or ASP.NET, allow organizations to construct bespoke apps without having to worry about the complexities that lie behind the platforms. Many platforms such as Java EE, Google AppEngine can be used to build architecture frameworks and enterprise applications.

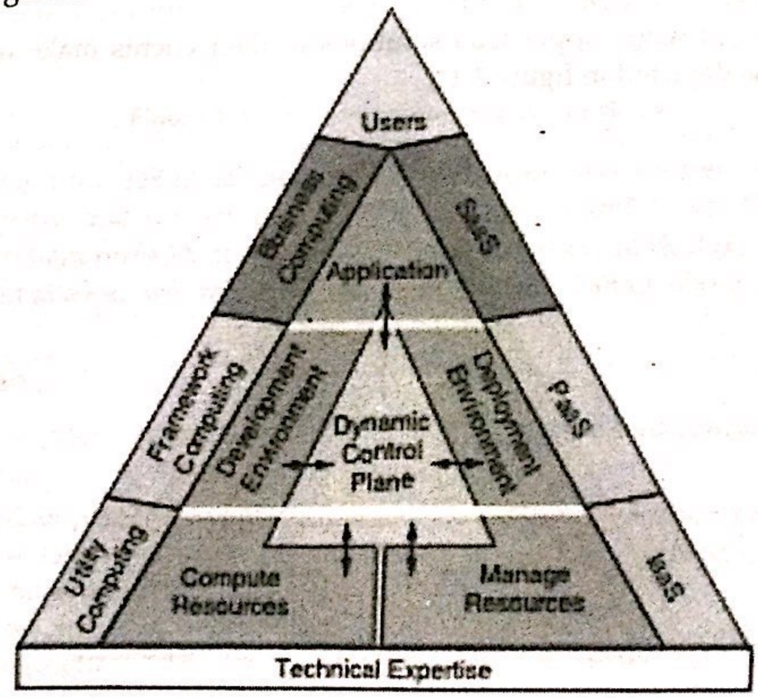


Figure 1.14*: Computational resources or deployment environments grouped into layers

Platform as a Service (PaaS) is a cloud computing architecture in which a specific development and deployment platform is used as the foundation for deployments, such as Java EE, IBM WebSphere, Oracle, Google Apps,.NET, and BizTalk. Only applications designed for the respective platform may be installed in the cloud, making these clouds exclusive.

PaaS is a type of framework computing in which the platform offered serves as the foundation upon which applications are built. These programs are platform-specific and incorporate platform-specific extensions or services, such as Amazon's SimpleDB, that cannot be migrated to other settings. Framework computing is derived from architectures that provide a layer of capabilities and services that shield developers from the underlying specifics. As seen in Figure 1.14, this strategy allows for faster application development and deployment.

1. **Development environment**: All platforms necessitate a development environment where applications are created, created, tested, and validated away from the production environment. Traditional integrated development environments (IDEs) can be set to deploy to resources inside a PaaS environment, or they may be integrated directly into the PaaS offering. Microsoft Visual Studio-derived environments may connect to both internal and external instances of Microsoft-specific platforms, allowing for offline development of PaaS-deployed apps. Less standard and more proprietary options, such as Salesforce.com's Force.com, are increasingly providing a PaaS-hosted development environment through which developers can build, test, and deploy their solutions.
2. **Deployment environment**: Once the program is ready to be consumed by end-users, the next step is to deploy it into production. The run-time environment in which the applications are deployed is essentially this capability. The expectation of on-demand scalability associated with PaaS that does not present in earlier incarnations is the distinction between the PaaS run-time environment and that of hosted or even classic enterprise installed systems. This on-demand scalability may be achieved by putting the environment on a general IaaS or by constructing and linking the development, deployment, and dynamic control planes explicitly. The latter creates a platform-specific IaaS, with the needed components structured specially to provide and decommission resources based on the deployment environment's particular demands.

Cloud Architecture — SaaS

Software as a Service (SaaS) is a cloud computing paradigm in which prebuilt programs such as Customer Relationship Management (CRM), Sales Force Automation (SFA), word processing, spreadsheets, and Human Resource Management (HRM) are made available to clients via a web browser or other local interface such as a mobile device app. These apps are typically customizable, but the user does not need to be bothered with the underlying infrastructure, development platform, or real execution.

The sole component, in this case, is application. Users engage with the program, whether it is the outcome of constructing utility computing, followed by a platform, or it is just an application installed on a server. Users are unconcerned about how it was developed, where it is stored, or the computational resources necessary to provide it. They just want it to be available when they need it, responsive and performant enough to be helpful, and secure regardless of where, when, or how they use it.

Cloud Adoption and Rudiments

Cloud-computing adoption has been increasing rapidly, with cloud-specific spending expected to grow at more than six times the rate of general IT spending through 2020. While large organizations have successfully implemented specific software-as-a-service (SaaS) solutions or adopted a cloud-first strategy for new systems, many are struggling to get the full value of moving the bulk of their enterprise systems to the cloud.

Cloud Adoption is a strategic move by organizations of reducing cost, mitigating risk, and achieving scalability of database capabilities. Cloud adoption may be up to various degrees in an organization, depending on the depth of adoption. The depth of adoption yields insight into the maturity of best practices, enterprise-ready cloud services availability.

Organizations that go ahead with the strategic decision of adopting cloud-based technologies have to identify potential security thefts and controls, required to keep the data and applications in the cloud secure. Hence there is a need for compliance assessment during cloud adoption. The following measures are taken for compliance assessment to ensure security and accountability of data and applications in the cloud services:

- Matching the security requirements of the organization with the security capabilities of the cloud service provider
- Analyzing the security policies of the cloud service provider along with the history of transparency and security-related practices
- Proper understanding of the technical aspects of data and traffic flow
- Proper understanding and documentation of the roles and responsibilities of the cloud service provider
- Understanding of the certifications and compliances that can be leveraged from the cloud service provider

A variety of industries benefit from cloud adoption, including healthcare, marketing and advertising, retail, finance, and education.

Benefits include

1. **Healthcare**: Fueled by digital and social consumer behaviors and the need for secure and accessible electronic health records, hospitals, clinics, and other medical organizations are using cloud computing for document storage, marketing, and human resources.
2. **Marketing and Advertising**: In an industry dependent on social media, as well as the quick creation and publishing of customer-relevant content, agencies are using hybrid cloud adoption strategies to deliver critical client messages to their local and worldwide audiences.
3. **Retail**: A successful e-commerce strategy requires a sound Internet strategy. With the help of cloud adoption, Internet-based retail can effectively market to customers and save their product data for less money.
4. **Finance**: Efficient expense management, human resources, and customer communications are three of the most important business needs of today's finance organizations. For these reasons, financial services institutions are now placing their email platforms and marketing tools in the cloud.
5. **Education**: Internet-based education opportunities are now more popular than ever. The cloud allows universities, private institutions, and public schools to provide learning, homework, and grading systems online. Cloud applications will increase the efficiency of the educational system nationwide or statewide through the latest technologies which are also made available to the less privileged education centers not merely the top institutions. The geographical location would no longer be a barrier to class attendance. This not only enhances knowledge sharing but also fortifies team collaboration.
6. **Manufacturing**. Manufacturers take cloud applications as the most important tool for IT advancement in the future. It allows the IT manufacturers to avoid the technical as well as business issues that would otherwise have taken place in their own data centers. This also allows them to pay for the utilized services, hence becoming a cost-effective solution. In cloud computing, the huge remote internet servers are accessed by the manufacturers just the same way as their own.
7. **Business and professional services**. The manufacturers are now able to be virtually present at every location in the world through employing three major components namely, Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). This enables the manufacturers to explore business opportunities globally and link with the IT infrastructure.
8. **Government**. Scalable, on-demand cloud computing services can help government organizations focus on mission-critical objectives while helping to reduce IT costs. The benefits of employing cloud computing and cloud applications are uncountable. There is no way the industries would not employ it to maximize their benefits and profits. However, organizations and institutions need to select the right solutions to get benefited from cloud computing.
9. **Large Companies and Corporations**: Corporate environments require the largest IT investment. Enterprise cloud adoption results in significant bottom-line savings by improving efficiency, eliminating the need for a large security and maintenance staff, and lowering the cost of server space.
10. **Small and Mid-Size Companies**: An increase in staff, clientele, and projects often require small and mid-size organizations to quickly grow their IT infrastructure. Engaging in cloud computing allows for efficient and cost-effective scalability that takes minutes instead of days.

11. **Entrepreneurs and Startups**: Choosing the cloud instead of an expensive IT infrastructure reduces startup costs and up-front software investments. Software-as-a-service (SaaS) vendors now commonly offer a monthly fee subscription model.

Additional industries benefiting from cloud adoption include construction, real estate, and not-for-profit. No matter the business focus, companies of varying sizes can benefit from this technology revolution. Cloud computing can provide travel and hospitality companies with many convenient solutions in this competitive world. Without any paper expense, it allows people to book the desired flights for traveling, reserve places for accommodation, and not only that but also make online payments through the company's accounts system. The customers will not have to connect to the airlines and travel service individually as all of that can be dealt with through a single transaction system. Apart from the other services, cloud computing can potentially perk up the huge and portioned health industry. This industry comprises patients, insurance companies, hospitals, laboratories, and physicians. The cloud benefits the health industry in innumerable ways. Better quality patient data updates, easier updates, better communication with the patients, higher consumer lifestyle choices, and more appropriate choice of following the specified courses of patient's treatments are some of the benefits gained through cloud computing.

INTRODUCTION TO DISTRIBUTED COMPUTING

A computer processing approach in which various elements of a program run concurrently on two or more machines that communicate with each other over a network is known as distributed computing. Distributed computing is a subset of segmented or parallel computing; however, the latter term is most usually used to describe processing in which separate sections of a program execute concurrently on two or more processors in the same machine. While both forms of processing need that a program is segmented—divided into portions that may run concurrently—distributed computing requires that the segmentation of the program take into consideration the many settings in which the various pieces of the program will operate. Two computers, for example, are likely to have distinct file systems and physical components. Distributed computing is a technique in which software system components are shared among numerous computers to increase efficiency and performance. It is confined to applications that use components shared by computers in a certain geographical region.

Distributed computing just means that something is shared among multiple systems which may also be in different locations. The Distributed Computing Environment is a widely used industry standard that supports distributed computing.

In the workplace, distributed computing has frequently meant locating multiple phases in business processes in the most effective locations on a computer network. For example, in a typical 3-tier distribution, user interface processing is performed on the user's PC, business processing is handled on a distant computer, and database access and processing is performed on another computer that gives centralized access to various business activities. The client/server communications model is commonly used in this type of distributed computing.

A distributed computing architecture is made up of several client workstations that are outfitted with extremely lightweight software agents and one or more specialized distributed computing management servers. When the agents operating on the client computers detect that the computer is idle, they notify the management server that the machine is not in use and is ready for a processing job. After then, the agents request an application bundle. When the client machine receives this application package to process from the management server, it executes the application software whenever it has free CPU cycles and delivers the results back to the management server. When the user returns and needs the resources again, the management server provides the resources that were used to execute various activities while the user was away.

Distributed systems offer many benefits over centralized systems such as:

1. **Reliability, high fault tolerance:** A system crash on one server does not affect other servers.
2. **Scalability:** The system can easily be expanded by adding more machines as needed.
3. **Flexibility:** It makes it easy to install, implement and debug new services.

4. **Redundancy:** Several machines can provide the same services, so if one is unavailable, work does not stop. Additionally, because many smaller machines can be used, this redundancy does not need to be prohibitively expensive.
5. Fast calculation speed: A distributed computer system can have the computing power of multiple computers, making it faster than other systems.
6. **Openness:** Since it is an open system, it can be accessed both locally and remotely.
7. **High performance:** Compared to centralized computer network clusters, it can provide higher performance and better cost performance.
8. **Latency:** Low latency is achieved through distributed systems. If a node is closer to the user, the distributed system ensures that traffic from that node is received by the system. As a result, the user may realize that it takes considerably less time to serve them.

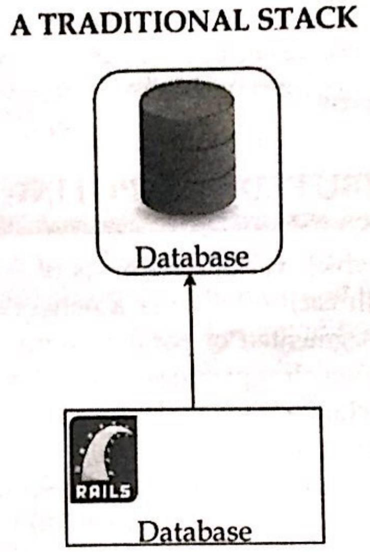


Figure 1.16: Traditional Computing

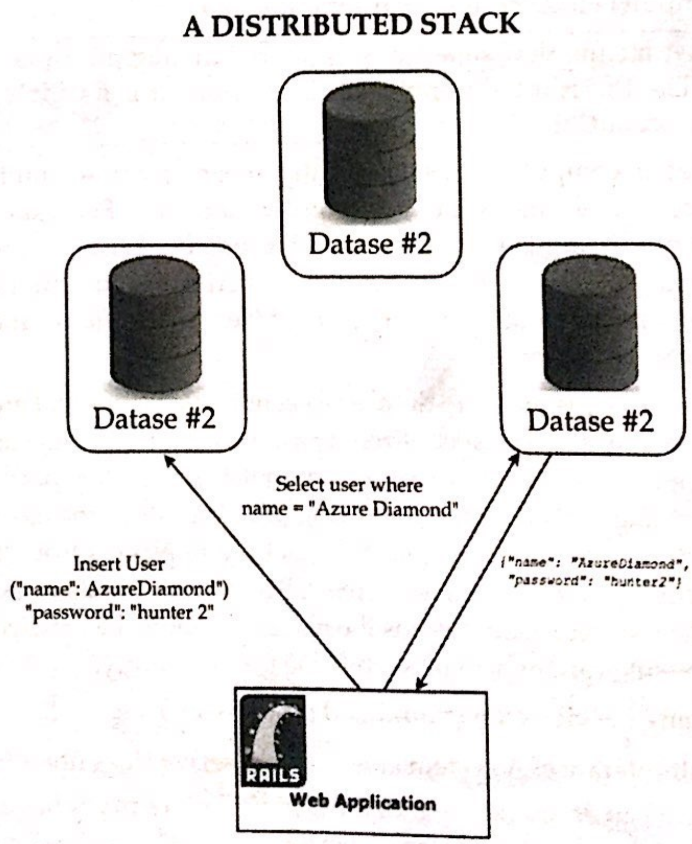


Figure 1.17: An Architecture that can be considered distributed

SYSTEM MODELS FOR DISTRIBUTED AND CLOUD COMPUTING

Distributed and cloud computing systems are made up of many independent computer nodes. These node machines are linked hierarchically by SANs, LANs, or WANs. A few LAN switches can easily join hundreds of PCs into a functional cluster with today's networking technologies. A WAN can link several local clusters together to build a very large cluster of clusters. In this manner, a vast system with millions of computers connected to edge networks may be built.

Clusters, peer-to-peer (P2P) networks, computer grids, and Internet clouds are the four types of distributed systems. These four system types may involve hundreds, thousands, or even millions of computers as participating nodes in terms of the node number. At varying levels, these machines work jointly, cooperatively, or collaboratively. In diverse technical and application areas, the table items characterize these four system types.

A **computing cluster** is made up of interconnected stand-alone computers that work together to form a single integrated computing resource. The network of computing nodes is normally homogenous with dispersed control running Unix/Linux and is connected by LAN/SAN. They are well-suited to high-performance computing.

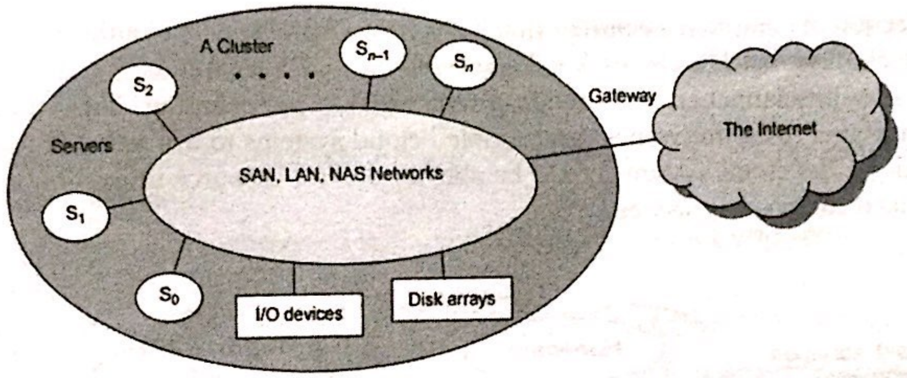


Figure 1.18: A cluster of servers interconnected by high bandwidth SAN or LAN with shared IO and disk

Every peer in a **P2P network** serves as both a client and a server. Peers join and exit the network on their own. There is no requirement for centralized coordination or a centralized database. All client machines function autonomously, allowing them to freely join or exit the system. This means that there is no master-slave relationship among the peers. There is no requirement for centralized coordination or a centralized database. The system self-organizes and has dispersed control. A P2P network, unlike a cluster or grid, does not employ a specialized interconnection network.

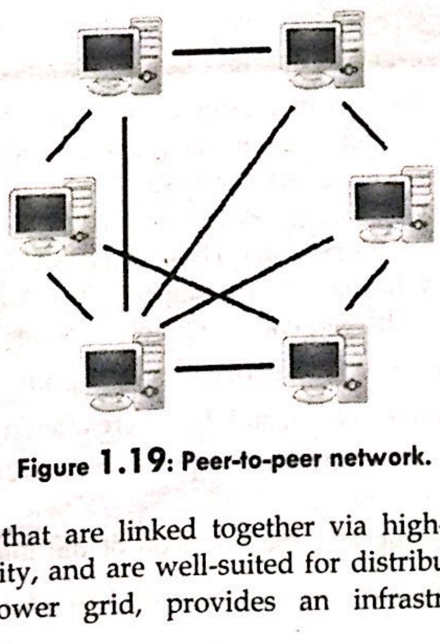


Figure 1.19: Peer-to-peer network.

Grids are heterogeneous groups that are linked together via high-speed networks. They are server-oriented, offer authenticated security, and are well-suited for distributed supercomputing. A computing grid, like an electric utility power grid, provides an infrastructure that connects computers, software/middleware, humans, and sensors. The grid is built on a regional, national, or worldwide scale using LANs, WANs, or Internet backbones. A grid's computers include servers, clusters, and supercomputers. A grid system may be accessed through PCs, laptops, and mobile devices.

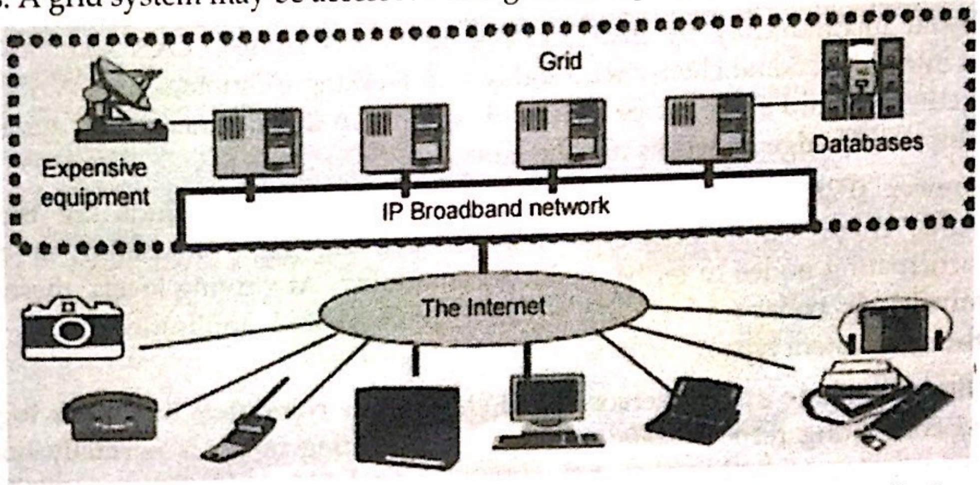


Figure 1.20: Grid providing computing utility, data, and information services through resource sharing and cooperation.

A **cloud** is a collection of computer resources that have been virtualized. A cloud can handle a wide range of workloads, including batch-style backend processes as well as interactive and user-facing apps. Workloads can be swiftly launched and scaled up with fast VM provisioning. The virtualization of server resources has improved cost-effectiveness and enabled cloud systems to use reduced costs to benefit both users and providers. The cloud system should be able to monitor resource utilization in real-time so that allocations may be rebalanced as needed.

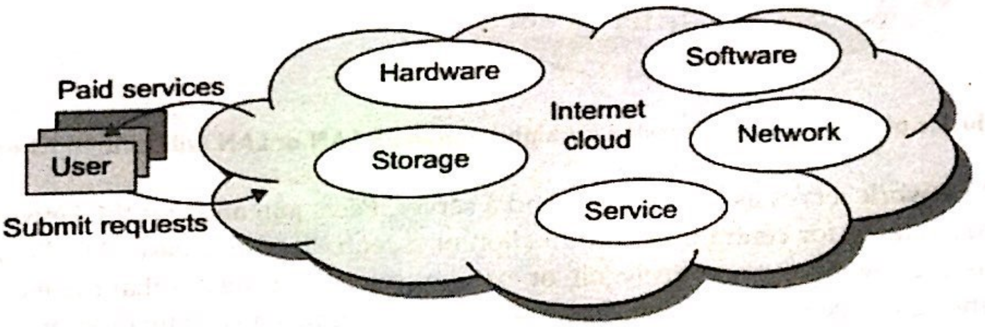


Figure 1.21: Internet Cloud and its components

GRID COMPUTING CONCEPT

Grid computing emerged as an extension of cluster computing in the early 1990s. Grid computing, analogous to the electricity grid, suggested a new technique to accessing big processing capacity, massive storage facilities, and a range of services. Users can "use" resources in the same manner that they use conventional utilities like electricity, gas, and water. Grids began as aggregations of geographically distributed clusters linked via Internet connections. These clusters belonged to several organizations, and agreements were negotiated for them to share processing capacity. A computer grid, as opposed to a "big cluster," was a dynamic aggregation of heterogeneous computer nodes on a national or even global scale.

Several developments made possible the diffusion of computing grids:

- clusters became quite common resources, and they were often underutilized.
- new problems were requiring computational power that went beyond the capability of single clusters; and
- the improvements in networking and the diffusion of the Internet made possible long-distance, high-bandwidth connectivity.

All of these factors contributed to the creation of grids, which currently serve a large number of people worldwide. Cloud computing is frequently seen as grid computing's heir. Cloud computing is deployed in huge data centers hosted by a single enterprise that delivers services to others. Clouds are distinguished by their almost endless capacity, tolerability to faults, and constant availability, like in the case of mainframes. In many situations, the computer nodes that comprise the infrastructure of computing clouds, such as clusters, are commodity computers. A cloud vendor's services are utilized on a pay-per-use basis, and clouds fully fulfill the utility concept provided by grid computing.

Grid Computing Infrastructures

Grid infrastructure is the bedrock upon which successful grid applications are built. This infrastructure is a complicated amalgamation of a variety of capabilities and resources recognized for the specific challenge and environment at hand.

Developers/service providers must evaluate the following considerations in the early phases of implementing any Grid Computing application architecture to define the basic infrastructure support necessary for that environment:

- What problem(s) are we trying to solve for the user? How do we address grid enablement simpler, while addressing the user's application simpler? How does the developer (programmatically) help the user to be able to quickly gain access and utilize the application to best fit their problem resolution needs?
- How difficult is it to use the grid tool? Are grid developers providing a flexible environment for the intended user community?
- Is there anything not yet considered that would make it easier for grid service providers to create tools for the grid, suitable for the problem domain?
- What are the open standards, environments, and regulations grid service providers must address?

In several phases of deployment, a Grid Computing infrastructure component must handle multiple potentially complex areas.

These areas are:

- Security
- Resource management
- Information services
- Data management

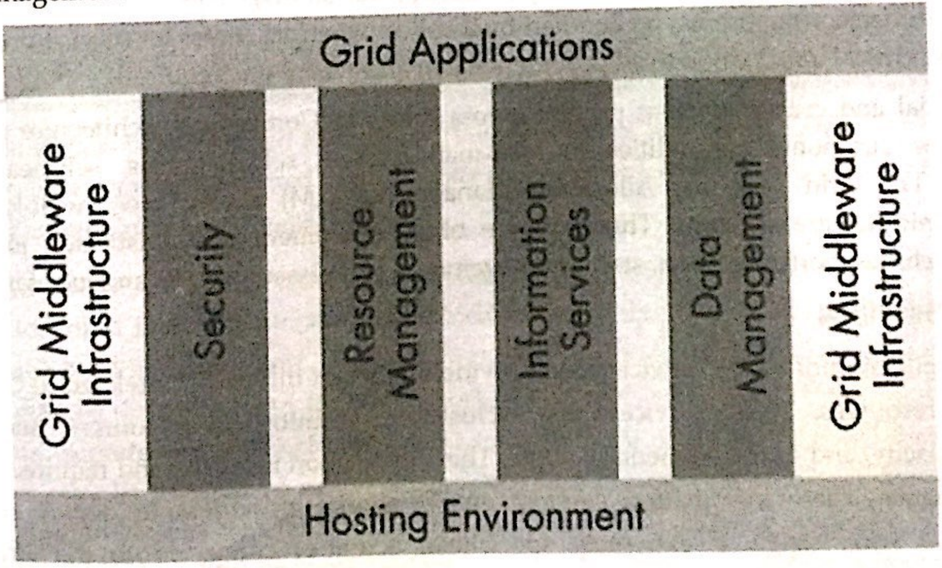


Figure 1.22: Grid Computing infrastructure component

Security

In the security schemes of a Grid Computing Environment, the diverse nature of resources and their different security regulations are intricate and complex. These computer resources are housed across a variety of security domains and platforms. Local security integration, safe identity mapping, safe access/authentication, safe federation, and trust management must all be addressed by middleware systems.

Other security needs are frequently focused on data integrity, confidentiality, and information privacy. Grid Computing data sharing must be secured using secure communication channels such as SSL/TLS which are frequently used in conjunction with secure message exchange protocols such as *WS-Security*. The Grid Security Infrastructure is the most well-known security infrastructure for grid security (GSI). Most GSI solutions support single sign-on, heterogeneous platform integration, and secure resource access/authentication.

Resource Management

Because of the enormous quantity and varied potential of Grid Computing resources, resource management is an important effort area in Grid Computing settings. These resource management scenarios frequently comprise resource detection, inventory, fault isolation, provisioning, monitoring, a range of autonomic capabilities, and service-level management tasks. The most intriguing component of the resource management area is selecting the appropriate resource from the grid resource pool based on the service-level requirements, and then efficiently provisioning them to meet user needs.

A work management system, for example, where the resource management feature detects the work, assigns the appropriate resources for task execution, partitions the task if required, and offers task status feedback to the user. This procedure comprises transporting the data required for various calculations to the appropriate Grid Computing resources, as well as procedures for dispatching the task results.

It is critical to recognize that many service providers can host Grid Computing resources in a variety of areas, including security, administration, networking, and application functions. Additionally, operational and application resources may be housed on many hardware and software platforms. In addition to this complexity, Grid Computing middleware must offer effective resource monitoring to collect the necessary usage, availability, and other information matrices.

The security and capacity of the grid service provider to reach out and probe into other service provider domains to gather and reason about crucial operational information is one of the consequences of this reality (i.e., to reach across a service provider environment to ascertain firewall and router volume-related specifics, or networking switch status, or application server-status). This frequently gets problematic across several dimensions and must be handled by a collective effort across all service providers, such as communicating critical information to all providers when and where it is required.

Another essential and crucial element present across the Grid Computing architecture is provisioning, which provides autonomic capabilities for self-management, self-diagnosis, self-healing, and self-configuration. The Grid Resource Allocation Manager (GRAM) is the most well-known resource management middleware solution. This resource offers customers a sophisticated job management solution that includes work allocation, status management, data dissemination, and job start/stop.

Information Services

The primary focus of information services is to provide important information related to Grid Computing infrastructure resources. These services rely exclusively on information sources, such as resource availability, capacity, and usage, to mention a few. This information is crucial and required feedback from resource managers. These information services enable service providers to deploy resources more efficiently for a wide range of extremely specialized activities connected to the Grid Computing infrastructure solution.

Furthermore, developers and providers may build grid systems that mimic portals and use meta-schedulers and meta-resource managers. These indicators, when combined with resource policies, aid in service-level management (SLA). This information is resource-specific and is supplied following the schema for that resource. To turn this resource-specific data into relevant information sources for the end-user, we may require higher-level indexing services, data aggregators, and transformers.

For example, one resource may provide information on operating systems, while another may provide information on hardware configuration, and we can then group this resource information, reason with it, and then suggest a "best" price combination on selecting the operating system on other specific hardware. This combinatorial technique to reasoning is relatively simple in a Grid Computing environment since all critical resources are shared, as is the information associated with the resources.

Data Management

The single most crucial asset in a Grid Computing system is data. This data may be sent into the resource, and the resource's output may be used to execute a specified job. Data migration in a geographically distributed system can soon produce scalability issues if the infrastructure is not correctly constructed. It is commonly established that data must be kept close to the calculation where it will be used. In every Grid Computing system, data mobility requires secure data transfers, both to and from the corresponding resources. Current data management advancements are mostly focused on virtualized data storage technologies such as storage area networks (SAN), network file systems, dedicated storage servers, and virtual databases. These virtualization mechanisms in data storage solutions and common access mechanisms (e.g., relational SQLs, Web services, and so on) assist developers and providers in designing data management concepts into the Grid Computing infrastructure with far greater flexibility than traditional approaches.

Some of the factors that developers and suppliers must consider when making decisions involve selecting the best data management strategy for Grid Computing infrastructures. This comprises the size of the data repositories, the geographical distribution of resources, the security needs, replication and caching strategies, and the underlying technologies used for storage and data access.

Today, the most significant endeavor in this field is the Open Grid Service Architecture (OGSA) and its associated standard projects. The OGSA provides a uniform interface solution to grid services, with all information neatly encoded using XML as the standard. This standardizes information services and resource management for Grid Computing systems.

GRID ARCHITECTURE AND STANDARDS

The Grid's design is sometimes defined in terms of "layers," each of which performs a distinct purpose. In general, upper levels are more concerned with the user (user-centric), and lower levels are more concerned with computers and networks (hardware-centric).

The network layer, which ensures connection for the Grid's resources, lies at the absolute bottom of everything. On top of it is the resource layer, which is made up of the real Grid resources, such as computers, storage systems, electronic data catalogs, and even sensors such as telescopes or other equipment that may be directly linked to the network. The middleware layer offers the tools that allow different elements (servers, storage, networks, etc.) to participate in a unified Grid environment. The middleware layer may be thought of as the intelligence that connects the different pieces - the Grid's "brain"! The application layer is the structure's topmost layer, and it contains all of the user applications (science, engineering, business, and finance), portals, and development toolkits that support the apps. This is the layer that grid users will "see."

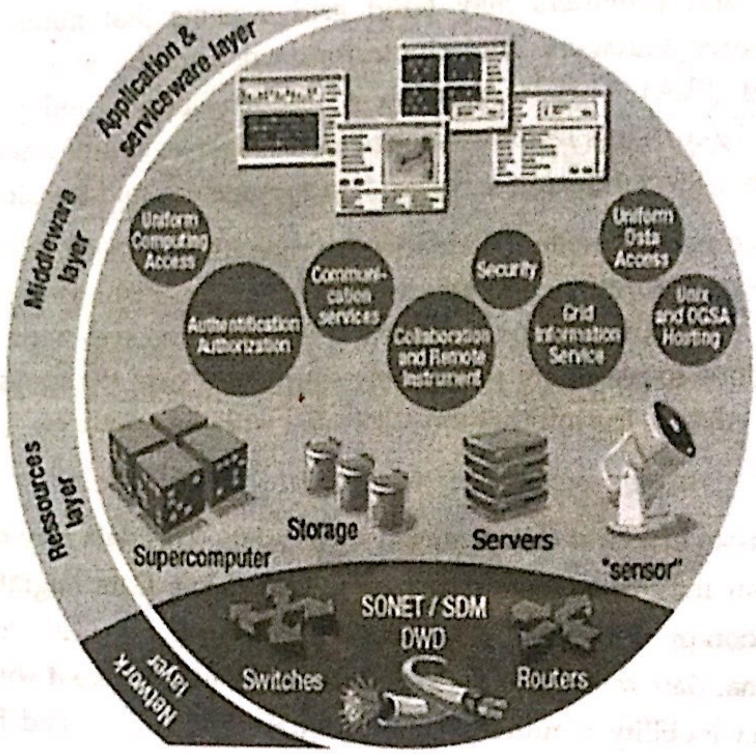


Figure 1.17 Grid Environment

CHARACTERISTICS OF GRIDS

1. **Large scale**: A grid must be able to deal with several resources ranging from just a few to millions.
2. **Geographical distribution**: Grid resources may be spread geographically.
3. **Heterogeneity**: A grid hosts both software and hardware resources that can be ranging from data, files, software components or programs to sensors, scientific instruments, display devices, personal digital organizers, computers, supercomputers, and networks.
4. **Resource sharing and coordination**: Resources in a grid belong to different organizations that allow other organizations (*i.e., users*) to access them. The resources must be coordinated to provide aggregated computing capabilities.
5. **Multiple administrations**: Each organization may establish different security and administrative policies under which resources can be accessed and used.
6. **Accessibility attributes**: Transparency, dependability, consistency, and pervasiveness are attributed typically to grid resource access. A grid should be seen as a single virtual computing environment and must assure the delivery of services under established Quality of Service requirements. A grid must be built with standard services, protocols, and interfaces thus hiding the heterogeneity of the resources while allowing its scalability. A grid must grant access to available resources by adapting to dynamic environments where resource failure is common.

ELEMENTS OF GRID

Grid computing combines elements such as distributed computing, high-performance computing, and disposable computing depending on the application of the technology and the scale of operation. Grids can create a virtual supercomputer out of the existing servers, workstations, and personal computers. Present-day grids encompass the following types:

- **Computational grids,** in which machines will set aside resources to "number crunch" data or provide coverage for other intensive workloads
- **Scavenging grids,** commonly used to find and harvest machine cycles from idle servers and desktop computers for use in resource-intensive tasks (scavenging is usually implemented in a way that is unobtrusive to the owner/user of the processor)

- **Data grids**, which provide a unified interface for all data repositories in an organization, and through which data can be queried, managed and secured.
- **Market-oriented grids**, which deal with price-setting and negotiation, grid economy management, and utility-driven scheduling and resource allocation.

THE KEY COMPONENTS OF GRID COMPUTING

- **Resource management**: a grid must be aware of what resources are available for different tasks
- **Security management**: the grid needs to take care that only authorized users can access and use the available resources
- **Data management**: data must be transported, cleansed, parceled, and processed
- **Services management**: users and applications must be able to query the grid in an effective and efficient manner

More formally, a grid computing environment may be thought of as a computer system made up of several logical hierarchical layers. The figure depicts these layers. Grid fabric resources, grid security infrastructure, core grid middleware, user-level middleware and resource aggregators, grid programming environment and tools, and grid applications are all part of it.

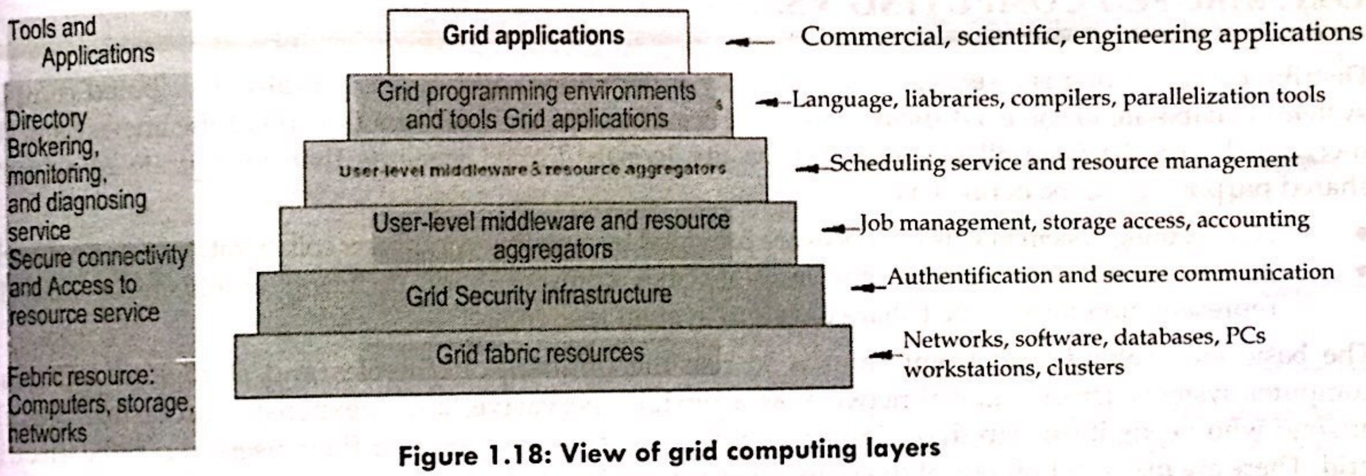


Figure 1.18: View of grid computing layers

SETI@home Project, LHC Computing Grid, NFCR Centre for Computational Drug Discovery are examples of Grid Computing Projects.

> **Open Grid Forum**
>
> The Open Grid Foundation (OGF) is an open community dedicated to accelerating the growth and acceptance of applied distributed computing. OGF fulfills its mission through open forums that foster community, investigate trends, exchange best practices, and codify these best practices into standards.

The major components of a grid computing system can be identified into various categories from different perspectives as follows:

- Functional view
- Physical view
- Service view

Basic elements of a grid from a functional view are decided depending on the grid design and its expected use. Some of the functional constituents of a grid are:

- Security (Grid security infrastructure),
- Resource Broker,
- Scheduler,
- Data Management,
- Job and resource management,
- Resources.

Based on the physical viewpoint, a grid is a collection of:

- Networks,
- Computation (Processors),
- Storage,
- Scientific Instruments,
- Software and Licenses.

Service view of grid computing adheres to the concepts of:

- Service-oriented architectures (SOA)
- Simple Object Access Protocol (SOAP)
- Web services standards
- Web Services Description Language (WSDL)
- Web Services Inspection Language (WSIL)
- Universal Description, Discovery, and Integration (UDDI)
- .NET
- Web Services Resource Framework (WSRF)

DISTRIBUTED COMPUTING VS. GRID COMPUTING

Distributed computing is a setting in which many independent and geographically distributed computer systems collaborate to solve a difficult issue, with each computer solving a portion of the answer and then merging the results from all machines. These are loosely linked systems that collaborate to achieve a shared purpose. It can be defined as:

- A computing system in which services are provided by a pool of computers collaborating over a network.
- A computing environment that may involve computers of differing architectures and data representation formats that share data and system resources.

The basic idea behind Grid Computing is to use the optimal CPU cycles and storage of millions of computer systems across a global network as a flexible, pervasive, and inexpensive accessible pool that anyone who needs it can tap into, similar to how power companies and their users share the electrical grid. There are many definitions of the term 'Grid computing:

- A service for sharing computer power and data storage capacity over the Internet
- An ambitious and exciting global effort to develop an environment in which individual users can access computers, databases, and experimental facilities simply and transparently, without having to consider where those facilities are located.
- Grid computing is a model for allowing companies to use a large number of computing resources on-demand, no matter where they are located.

Since 1980, two technological improvements, computer CPU power, and network bandwidth have made distributed computing a more viable concept. As a result of these technologies, it is not only conceivable but also simple to assemble a huge number of computer systems to solve complicated computing power or storage needs. However, the number of truly distributable apps remains rather restricted, and the constraints remain substantial (standardization, interoperability, etc.).

Traditional distributed computing can be characterized as a subset of grid computing. Some of the differences between these two are:

- Distributed computing is the management or pooling of hundreds or thousands of computer systems, each of which has a limited amount of memory and processing capacity. Grid computing, on the other hand, has certain unique qualities. It is concerned with the effective exploitation of a pool of heterogeneous systems with optimal workload management, leveraging an enterprise's full computational resources (servers, networks, storage, and information) functioning in concert to produce one or more big pools of computing resources. Grid computing has no restrictions on users, departments, or origins.

Grid computing is distinguished from traditional distributed computing by its ability to facilitate processing across different administrative domains. Grids provide a method for optimizing the use of information technology resources inside an organization through the virtualization of computer resources. Its support for numerous administration policies, as well as secure authentication and authorization systems, allowing it to be spread across a local, metropolitan, or wide-area network.

Comparing Cloud Computing with Grid Computing

Basis for comparison	Cloud Computing	Grid Computing
Definition and basic difference	Cloud computing is defined as the use of distant computers, often housed on the internet, to store or manage data. This information might come from any computer or server. The cloud enables a user to ensure on-demand access to data stored in the cloud at any time. Cloud computing is a term used to describe a new type of computing that is based on network technologies. It has hardware and software that is integrated and networked.	Grid computing is a distributed architecture in which numerous computers are linked to address a certain task. Grid computing involves connecting all servers or personal computers to a shared network through WAN and assigning autonomous duties to each machine. All of these systems can communicate with one another either directly or through scheduling systems. It is similar to a cluster, except that each node on the grid has its manager. A single system manages all of these resources. Grid computing connects systems in various areas over WAN.
Types	After its evolution, cloud computing deployments has been segregated into: ● Public Clouds ● Private Clouds ● Community Clouds ● Hybrid Clouds	Grid computing has also following types: ● Distributed Computing systems ● Distributed Information systems ● Distributed Pervasive Systems
Goals	Cloud computing is primarily concerned with lowering costs and increasing profits. It also aims to improve scalability, as well as availability and dependability.	Grid computing is network-centric and so has a large-scale purpose. It focuses on resource sharing, ubiquitous, consistent, and dependable access to data, storage capacity, and computing power. It also emphasizes the delivery of a computer as a utility.
Pros	Cloud computing has many advantages such as: ● Clouds can store large amounts of data properly and safely. Data stored in the cloud is highly secure and can be accessed whenever needed. ● Cloud is easily accessible from any part of the world. You just need to have internet connectivity ● Cloud runs on the latest network and all data centers being secure ensure that it provides the best performance. ● It is cost-efficient and has fast backup and data restoration it also has automatic software updates.	Grid computing also has its advantages as below: ● Grid computing is useful in dealing with idle energy in computers. It is more efficient to put it into more sensible use. ● It helps to save money when huge projects are involved. Grid computing helps in distributing and splitting up the work into multiple computers. ● Whenever a failure occurs it would not stop the work as another computer will pick up the work, making this system more reliable. ● Space is saved and access to additional resources is made possible.
User Management	A centralized system managing the entire cloud can work with this setup or it can be delegated to any third party.	The management is decentralized, and it also has virtual organization-based management.

INTRODUCTION TO MOBILE COMPUTING

Mobile computing is a technology that allows data, voice, and video to be transmitted wirelessly from a computer or other wirelessly equipped device without the need for a fixed physical link.

The main concept involves:

- Mobile communication
- Mobile hardware
- Mobile software

Mobile communication

Mobile communication refers to the infrastructure put in place to ensure continuous and dependable communication through mobile devices. These would include things like protocols, services, bandwidth, and portals that are required to facilitate and support the aforementioned services. At this point, the data format is also determined. This prevents collisions with other current systems that provide the same service. Because the media is unguided, the underlying infrastructure is mostly radio wave oriented. That is, the signals are transmitted over the air to intended equipment capable of receiving and transmitting comparable types of signals.

Mobile Hardware

Mobile hardware consists of mobile devices or device components that receive or access mobility services. They would include portable laptops, smartphones, tablets, and Personal Digital Assistants, etc. These gadgets will have a receptor medium that can send and receive signals. These devices are set up to work in full-duplex mode, which means they can send and receive signals at the same time. They do not have to wait for one device to finish communicating before the other may begin conversing.

Mobile Software

The application that runs on the mobile device is referred to as mobile software. It is concerned with the features and needs of mobile apps. This is the mobile device's engine. In other words, it is the appliance's operating system. It is the crucial component that allows the mobile device to function. Because mobility is the primary consideration, this sort of computing assures that users are not tethered or anchored to a specific physical place but may work from anywhere. It covers every element of wireless communication. *For example,* iOS, Android, Windows Mobile, etc.

MOBILE CLOUD COMPUTING

In the era of growing technology, Mobile Cloud Computing (MCC) has emerged as a prominent conversation topic for modernizing the internet computing infrastructure. MCC provides a large platform for the IT industry, allowing mobile users to study their data anywhere, at any time, by utilizing Cloud services. Cloud computing, internet technology, and mobile computing are all combined in Mobile Cloud Computing. This makes it more user-friendly for them to keep their data on the cloud with simplicity. Despite significant progress, the MCC continues to fall short of user expectations owing to privacy and security concerns with regulated data.

Mobile Cloud Computing (MCC) is the marriage of cloud computing, mobile computing, and wireless networks to give rich computational resources to mobile users, network operators, and cloud computing providers. MCC's ultimate objective is to enable the execution of sophisticated mobile apps on a mobile device while providing a rich user experience. MCC offers an opportunity for both mobile network operators and cloud providers. MCC can be defined more broadly as "a rich mobile computing technology that leverages unified elastic resources of diverse clouds and network technologies toward unrestricted functionality, storage, and mobility to serve a multitude of mobile devices anywhere, anytime via the channel of Ethernet or Internet regardless of heterogeneous environments and platforms based on the pay-as-you-go model."

MCC Architecture

Computations are conducted remotely rather than on the device using MCC, allowing resource-constrained mobile devices to use computational resources from a variety of cloud-based services. There are four categories of cloud-based resources in MCC: distant immobile clouds, nearby immobile computing entities, proximal mobile computing entities, and hybrid resources (a combination of the other three models). Cloudlets and surrogates are member applications of proximal immobile computing entities, whereas giant clouds like Amazon EC2 are in the far immobile groupings. Smartphones, tablets, portable devices, and wearable computing devices are examples of proximal mobile computing entities, the third type of cloud-based resource.

Issues with MCC

MCC is the combination of mobile computing, cloud computing, and communication networks that creates several complex challenges such as Mobile Computation Offloading, Seamless Connectivity, Long WAN Latency, Mobility Management, Context-Processing, Energy Constraint, Vendor/data Lock-in, Security and Privacy, Elasticity that hinder MCC success and adoption.

Along with these here are some issues of MCC:

- **Architectural issues**: A reference design for diverse MCC environments is essential for unlocking the power of mobile computing towards unconstrained ubiquitous computing.
- **Energy-efficient transmission**: MCC imposes frequent transfers between the cloud platform and mobile devices; yet because wireless networks are stochastic, the transmission protocol must be properly constructed.

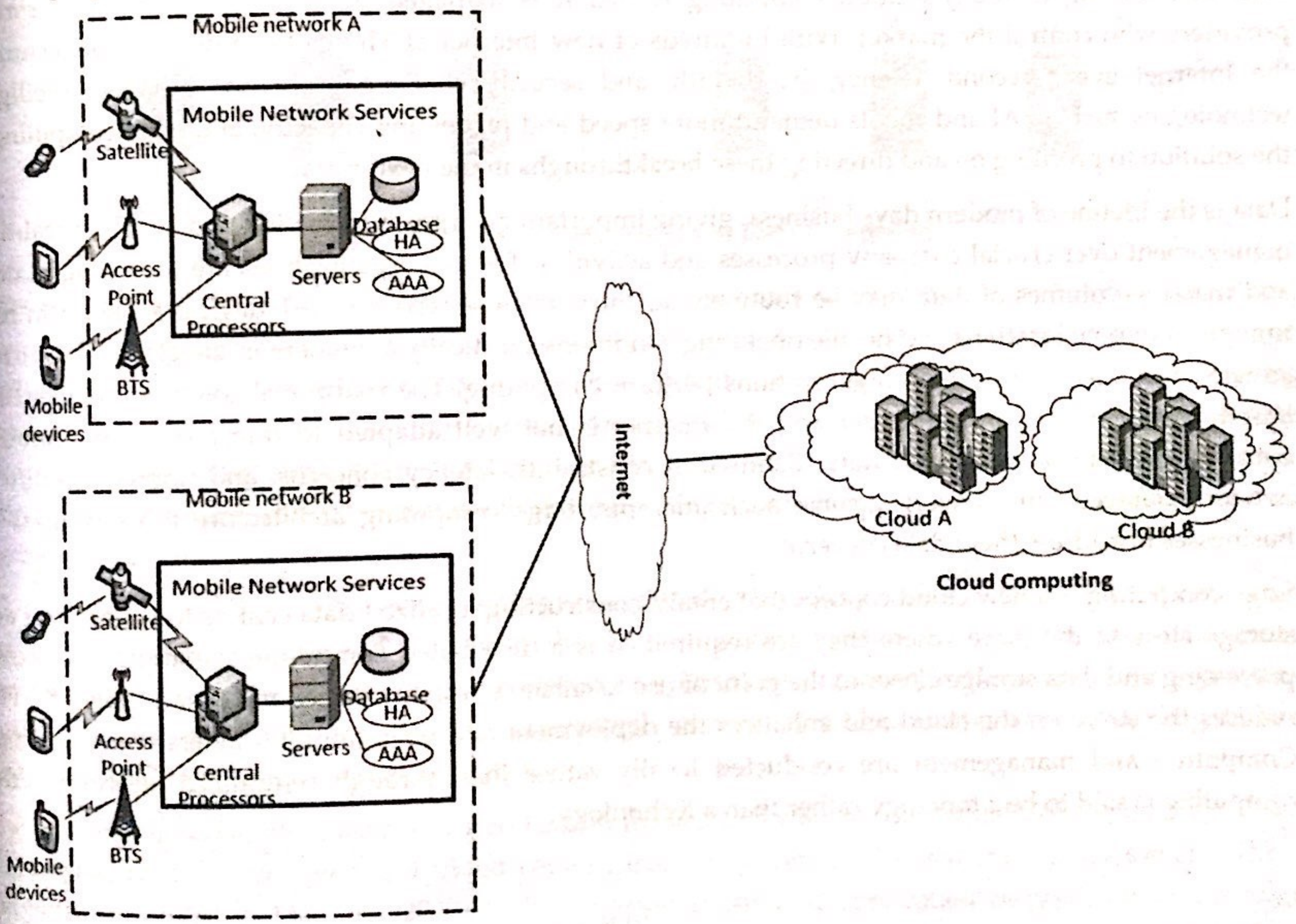


Figure 1.19: Architecture of MCC

- **Context-awareness issues:** Context-aware and socially aware computing are inseparably linked characteristics of modern mobile devices. Designing resource-efficient environment-aware apps are critical to realizing the vision of mobile computing amid heterogeneous converging networks and computing devices.
- **Live VM migration issues:** Exercising resource-intensive mobile applications through Virtual Machine (VM) migration-based application offloading entails encapsulating the application in a VM instance and migrating it to the cloud, which is a difficult task due to the additional overhead of deploying and managing VM on mobile devices.
- **Mobile communication congestion issues:** Mobile data traffic is skyrocketing due to ever-increasing mobile user demands for cloud services, which has an impact on mobile network operators and necessitates future efforts to ensure efficient communication between mobile and cloud endpoints.
- **Trust, security, and privacy issues:** Trust is critical to the success of the emerging MCC paradigm. This is because the data, together with the code/component/application/complete VM, is offloaded to the cloud for execution. Furthermore, similarly to software and mobile application piracy, MCC application development approaches are vulnerable to piracy. *Pirax* is well recognized as the first dedicated framework for combating application piracy in MCC standards.

INTRODUCTION TO EDGE COMPUTING

One shortcoming of today's cloud computing is that it is managed by a limited number of service providers who control the market. With hundreds of new Internet of Things (IoT) devices connecting to the Internet every second, latency, bandwidth, and security challenges are unavoidable. Intelligent technologies such as AI and robots demand more speed and processing capacity, and edge computing is the solution to profiting on and directing these breakthroughs in the next years.

Data is the lifeline of modern days business, giving important business information and enabling real-time management over crucial company processes and activities. Today's organizations are drowning in data, and massive volumes of data may be routinely acquired from sensors and IoT devices working in real-time from remote locations and hostile operating conditions practically anywhere in the globe. This virtual stream of data is altering how organizations perform computing. The traditional computer architecture, based on a centralized data center and the internet is not well adapted to transferring continuously expanding rivers of real-world data. Bandwidth constraints, latency concerns, and unpredictability in network outages can all work against such attempts. Edge computing architecture is being used by businesses to address these data concerns.

Edge computing is a new cloud concept that entails constructing localized data centers for computing and storage close to the place where they are required. It is a distributed computing paradigm that moves processing and data storage closer to the point of use to enhance response times and save bandwidth. This reduces the stress on the cloud and enhances the deployment and operation of a diverse range of apps. Computing and management are conducted locally rather than through centralized networks. Edge computing is said to be a topology rather than a technology.

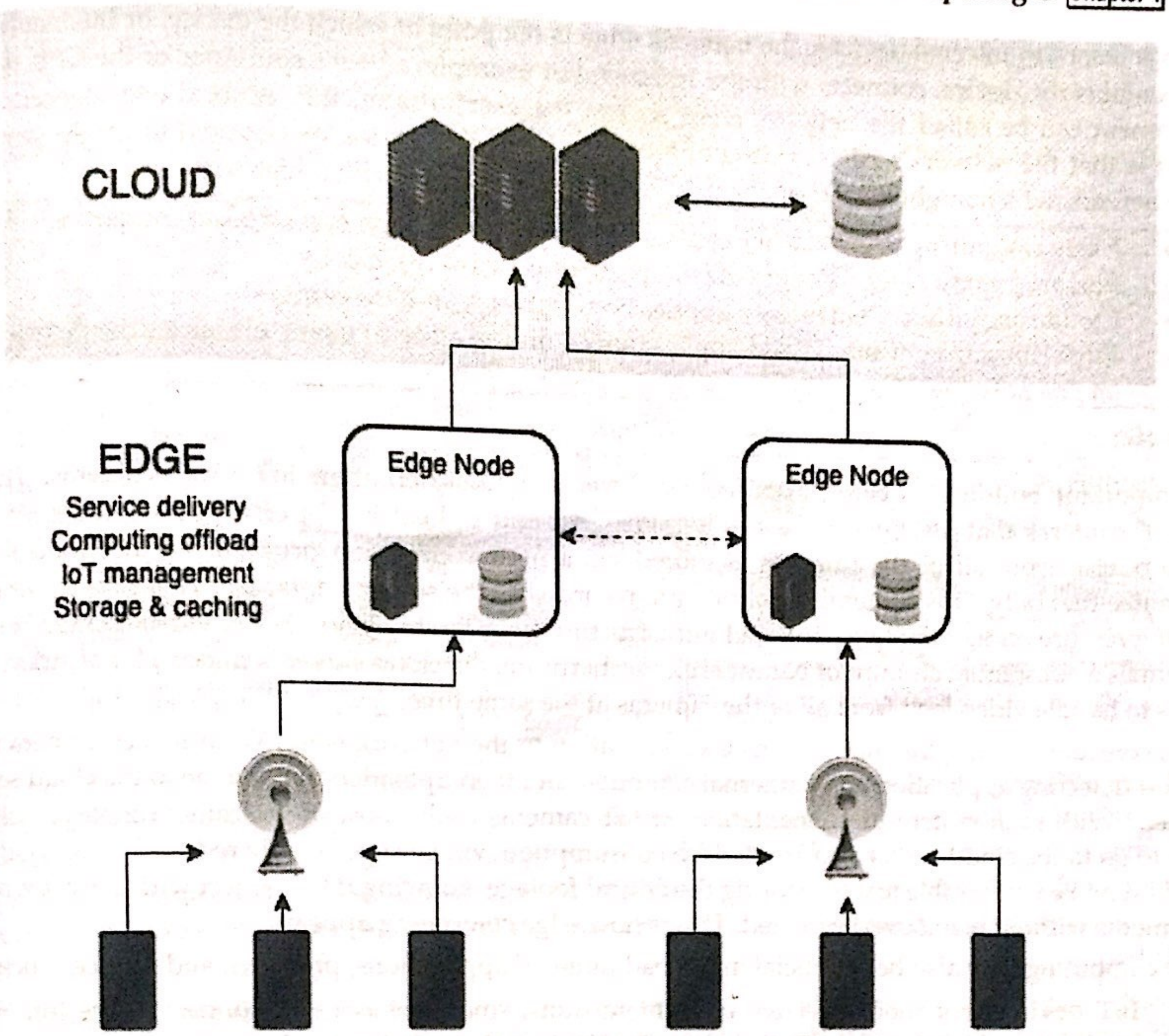


Figure 1.20: Basic Infrastructure of Edge Computing

Edge computing shifts certain storage and computation resources away from the central data center and closer to the source of the data. Rather than sending raw data to a central data center for processing and analysis, that work is done where the data is created - whether that's in a retail store, on a manufacturing floor, or throughout a smart city. Only the outcomes of such edge computing activities, such as real-time business insights, equipment repair projections, or other actionable solutions, are routed back to the main data center for review and other human interactions. Edge computing, as a result, is redefining IT and business computing.

Edge computing eliminates the latency concerns associated with talking with a central server. It also saves bandwidth by storing locally and connects to the cloud only when needed. A locally managed data center provides a greater level of information protection for enterprises dealing with privacy and regulatory problems. Edge computing will become increasingly important in harnessing and controlling linked devices and IoT innovations as they progress.

The roots of edge computing may be traced back to the late 1990s when content delivery networks (CDNs) were established to offer web and video content from edge servers located near to customers. These networks expanded in the early 2000s to house applications and application components at the edge servers, culminating in the first commercial edge computing services, which housed applications such as dealer locators, shopping carts, real-time data aggregators, and ad insertion engines.

Gartner estimates that by 2025, 75% of data will be processed outside the traditional data center or cloud.

When it comes to Internet devices, the network edge is the point at which the device, or the local network that contains the device, connects with the Internet. For example, a user's computer or the CPU within an IoT camera can be called the network edge, so can the user's router, ISP, or local edge server. The key lesson is that the network's edge is geographically close to the device, as opposed to origin servers and cloud servers, which might be extremely far away from the devices with which they connect.

- Early computing: Centralized applications only running on one isolated computer.
- Personal computing: Decentralized applications running locally.
- Cloud computing: Centralized applications running in data centers
- Edge computing: Centralized applications running close to users, either on the device itself or on the network edge

Example:

An important building is being guarded by dozens of high-definition IoT video cameras. These are "dumb" cameras that just produce a raw video signal and upload it to a cloud server indefinitely. The video output from all of the cameras is routed via a motion-detection program on the cloud server to guarantee that only clips containing activity are recorded to the server's database. This puts continual and severe pressure on the building's Internet infrastructure since the huge number of video footage being sent consumes a substantial amount of bandwidth. Furthermore, the cloud server is under a lot of strain since it needs to handle video feed from all of the cameras at the same time.

Now, consider moving the motion sensor computation to the network edge. What if each camera ran the motion-detecting application on its internal computer and then uploaded the footage to the cloud server as needed? With such system implementation on IoT cameras itself, most of the camera footage will never have to go to the cloud server and bandwidth consumption will be significantly reduced. The cloud server would now be responsible just for storing the critical footage, allowing it to connect with a greater number of cameras without being overburdened. This is how edge computing appears.

Edge computing may also be beneficial in a broad range of applications, products, and services, including:

- **IoT devices**: For more effective user interactions, smart devices that connect to the Internet can benefit from running code on the device itself rather than on the cloud.
- **Self-driving cars**: As autonomous vehicles must react in real-time, rather than waiting for commands from a server so edge computing comes in handy for best performance.
- **Efficient caching:** An application can change how the content is cached by executing code on a CDN edge network, allowing it to provide information to consumers more effectively.
- **Medical monitoring devices:** It is critical for medical equipment to act in real-time rather than waiting for instructions from a cloud server.
- **Video conferencing:** As interactive live video consumes a significant amount of bandwidth, relocating backend functions closer to the source of the video can reduce lag and latency.

FOG COMPUTING

The phrase fog computing, coined by Cisco, refers to a cloud computing alternative. This strategy capitalizes on the twin challenges of the creation of computing devices and the opportunity offered by the data generated by those devices by placing specific resources and transactions at the network's edge.

Fog computing is a form of decentralized computing infrastructure that distributes data, processing, storage, and applications between the data source and the cloud. Fog computing, like edge computing, brings the cloud's benefits and capabilities closer to the point where data is generated and acted upon. Many people use the terms fog computing and edge computing interchangeably since both involve bringing intelligence and processing closer to where the data is generated. This is often done to enhance efficiency, but it can also be done for security and compliance reasons. Processing latency is eliminated or significantly reduced by locating storage and computing systems as close to the applications, components and devices that require them as feasible. This is especially essential for devices connected to the Internet of Things, which create huge volumes of data. Because they are closer to the data source, these devices have much lower latency in fog computing.

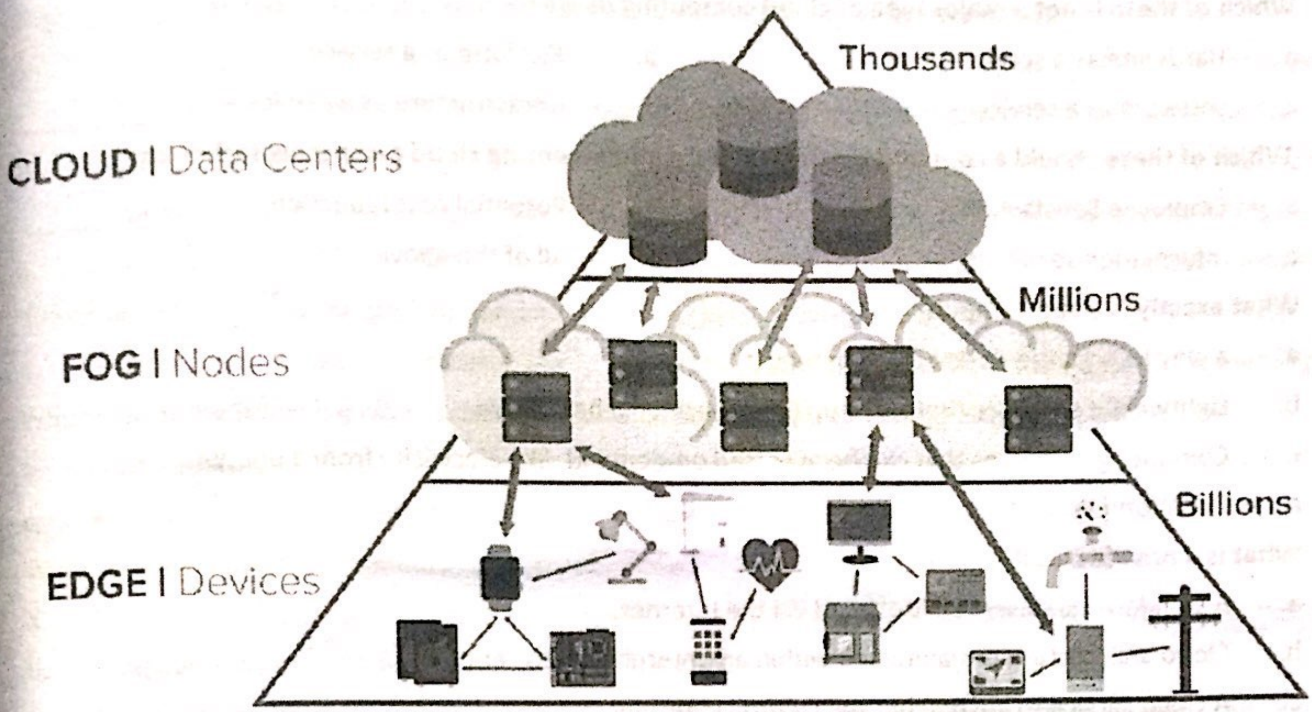


Figure 1.21: Basics of Fog Computing

Within the edge computing architecture, fog computing is the standard that allows repeatable, structured, and scalable performance. Edge computing is a subset of fog computing in which data is created, processed, and stored nearby. Edge processing, as well as the infrastructure and network connections required for data transmission, are all part of fog computing. This is because while both fog and mobile edge computing strive to reduce latency and enhance efficiency, they handle data in somewhat different places.

Edge computing generally occurs where sensors are linked to equipment to collect data—there is a physical connection between the data source and the processing point. Although fog computing reduces the distance between the processing site and the data source, it does so by performing edge computing operations within an IoT gateway or fog node with LAN-connected processors or within the LAN hardware itself. As a result, there is a greater physical distance between the processor and the sensors, yet there is no increased delay.

OBJECTIVE QUESTIONS

1) **Which of the following is an essential concept related to Cloud?**

a. Reliability b. Productivity c. Abstraction d. All of the mentioned

2) **________ has many of the characteristics of what is now being called cloud computing.**

a. Internet b. Software c. Web Service d. All of the mentioned

3) **Which of the following cloud concepts is related to pooling and sharing of resources?**

a. Polymorphism b. Abstraction c. Virtualization d. None of the mentioned

4) **Which of the following is Cloud Platform by Amazon?**

a. Azure b. AWS c. Cloudera d. All of the mentioned

5) Which of these is not a major type of cloud computing usage?

a. Hardware as a service
b. Platform as a service
c. software as a service
d. Infrastructure as a service

6) Which of these should a company consider before implementing cloud computing technology?

a. Employee Satisfaction
b. Potential cost reduction
c. Information sensitivity
d. All of the above

7) What exactly is cloud computing?

a. a way to organize desktop computers
b. Lightweight software that takes up little space on a hard drive
c. Computing resources that can be accessed on-demand, like electricity from a utility.\
d. The World Wide Web

8) What is a private cloud?

a. A standard cloud service is offered via the Internet.
b. Cloud architecture is maintained within an enterprise data center.
c. A cloud service to anyone but the cultural elite.
d. All of the above

9) Amazon Web Services is which type of cloud computing distribution model?

a. Software as a Service
b. Platform as a Service
c. Infrastructure as a Service
d. All of the above

10) This is a software distribution model in which applications are hosted by a vendor of or service provider and made available to customers over a network, typically the Internet.

a. PaaS
b. IaaS
c. SaaS
d. All of the above

11) What is the name of the organization helping to foster security standards for cloud computing?

a. Cloud Security Standards Working.
b. Cloud Security Alliance.
c. Cloud Security WatchDog.
d. Security in the Cloud Alliance

12) What facet of cloud computing helps to guard against downtime and determines cost?

a. Service-level agreements
b. Application programming interfaces
c. Virtual private networks
d. Bandwidth fees

13) Virtual Machine Ware (VMware) is an example of

a. Infrastructure Service
b. Platform Service
c. Software Service
d. All of the above

14) .... feature allows you to optimize your system and capture all possible transactions.

a. Scalability
b. Reliability
c. Elasticity
d. None of the above

15) Which of the following is owned by an organization selling cloud services?

a. Public
b. Private
c. Community
d. Hybrid

16) Which of the following architectural standards is working with the cloud computing industry?

a. Service-oriented architecture
b. Standardized Web Services
c. Web-application frameworks
d. All of the Above

17) ________ refers to the location and management of the cloud's infrastructure.

a. Service
b. Deployment
c. Application
d. None of the mentioned

18) The ________ cloud infrastructure is operated for the exclusive use of an organization.
   a. Public b. Private c. Community d. All of the mentioned

19) ________ cloud is one where the cloud has been organized to serve a common function or purpose.
   a. Public b. Private c. Community d. All of the mentioned

20) ________ provides virtual machines, virtual storage, virtual infrastructure, and other hardware assets
   a. IaaS b. SaaS c. PaaS d. All of the mentioned

21) Which one of the following options can be considered as the Cloud?
   a. Hadoop b. Intranet c. Web Applications d. All of the mentioned

22) Which one of the following cloud concepts is related to sharing and pooling the resources?
   a. Polymorphism b. Virtualization
   c. Abstraction d. None of the mentioned

23) Which of the following statements is not true?
   a. Through cloud computing, one can begin with very small and become big in a rapid manner.
   b. All applications benefit from deployment in the Cloud.
   c. Cloud computing is revolutionary, even though the technology it is built on is evolutionary.
   d. None of the mentioned

24) Which one of the following refers to the non-functional requirements like disaster recovery, security, reliability, etc.
   a. Service Development b. Quality of service
   c. Plan Development d. Technical Service

25) Cloud computing architecture is a combination of?
   a. service-oriented architecture and grid computing
   b. Utility computing and event-driven architecture.
   c. Service-oriented architecture and event-driven architecture.
   d. Virtualization and event-driven architecture

26) Which one of the following can be considered as the example of the Front-end?
   a. Web Browser b. Google Compute Engine
   c. Cisco Metapod d. Amazon Web Services

27) Which of the following provides the Graphic User Interface (GUI) for interaction with the cloud?
   a. Client b. Client Infrastructure
   c. Application d. Server

28) In Virtualization, which architecture provides the virtual isolation between the several tenants?
   a. IT Architecture b. Multitenant
   c. Deployment d. Business Architecture

29) On which one of the following utility computing is based?
   a. Grid Computing Model b. SOA Model
   c. virtual isolation Model d. Pay-per-Use model

30) Which one of the following statements is true about Service-Oriented Architecture?
   a. It is possible to exchange data between applications from different vendors without using additional programming.
   b. It provides computational resources on-demand as a metered service.
   c. Service-Oriented Architecture allows using the application as a service for other applications.
   d. Both A and C

31) In Grid Computing, which types of computer resources are there?

a. heterogeneous dispersed.
b. geographically dispersed.
c. Both A and B
d. None of the above

32) Which one of the following given programs provides isolation (abstraction) and partitioning?

a. System hypervisor
b. Software hypervisor
c. Hardware hypervisor
d. Virtualization hypervisor

33) Which one of the following PaaS types involves on-demand scaling and application security?

a. Stand-alone development environments
b. Open Platform as a service
c. Application delivery-only environments
d. Add-on development facilities

34) The resources like IP addresses and VLANs are provided to the end-users by which of the following?

a. Server virtualization
b. Client virtualization
b. End-user virtualization
d. IaaS

35) Which one of the following is a kind of open standard protocol?

a. SOAP b. WSDL c. DHML d. SIMPLE

36) In the virtual appliance, the content can be _____.

a. structured b. unstructured c. Both A and B d. None of the above

37) Which of the following is the most commonly used model for description and discovery and is also used with SOAP messaging?

a. DHML b. VMC c. WSDL d. SOA

38) The infrastructure as a service (IaaS) provides a type of isolated environment to each customer individually by using_____.

a. renting
b. virtual machine sprawl
c. security vulnerabilities
d. hypervisor

39) To provide more secure authentication, which of the following is required at least?

a. three-factor authentication.
b. two-factor authentication.
c. four-factor authentication.
d. None of the above

40) Which one of the following statements is correct about the FIDM?

a. SAML has access control
b. It typically uses the Security Markup Language (SAML) for packaging the user's security credentials.
c. It describes the protocols and technologies.
d. All of the above

41) Which of the following can be referred to as the location and management of the cloud's infrastructure?

a. Service
b. Deployment
c. Application
d. None of the mentioned

42) Which of the following is the deployment model?

a. Public b. Private c. Hybrid d. All of the mentioned

43) Which one of the following is the wrong statement?

a. Public Cloud may be managed by the constituent organization(s) or by a third party
b. A community cloud may be managed by the constituent organization(s) or by a third party
c. Private clouds may be either on- or off-premises
d. None of the mentioned

44) **Which one of the following is a type of Cloud that is organized in such a way to serve the common purpose or the functions?**

a. Public b. Private c. Community d. All of the mentioned

45) **Which one of the following is considered a type of cloud computing model involving the three different service models together?**

a. CPI b. SIP c. SPI d. All of the mentioned

46) **The term "CaaS" stands for_____ as a service?**

a. Compliance as a service b. Computer as a service

c. Community as a service d. Communication as a service

47) **Which one of the following is a type of infrastructure as a service?**

a. EC2 b. EC1 c. EC10 d. Hybrid

48) **All cloud computing applications suffer from the inherent _______ that is intrinsic in their WAN connectivity.**

a. Propagation b. Latency c. Noise d. All of the mentioned

QUESTION

1. Explain Cloud Computing with examples. Describe its properties.
2. What are the characteristics of cloud computing as per NIST? Explain.
3. Explain the working principle of Cloud Computing.
4. Describe Cloud services with examples.
5. What is the typical communication mechanism used in cloud computing? Explain.
6. Explain the vision of Cloud Computing.
7. What is cloud interoperability? Describe.
8. Explain scalability and fault tolerance.
9. What is Grid Computing? Give the major differences between Grid and Cloud computing.
10. What are the various challenges for the cloud?
11. Explain the innovative characteristics of cloud computing.
12. Describe the technologies on which cloud computing relies.
13. Define cloud computing and identify its core features.
14. What are the major distributed computing technologies that led to cloud computing?
15. Explain utility computing.
16. Describe the vision introduced by cloud computing.
17. What is the major advantage of cloud computing? Describe.
18. Briefly summarize the challenges still open in cloud computing.
19. How is cloud development different from traditional software development? Explain.
20. Explain cloud adoption and rudiments.
21. What is a cloud ecosystem? Explain cloud analytics.
22. What is Virtual Desktop Infrastructure? Explain its importance.
23. How is cloud development different from traditional software development? Describe.
24. Give a brief about the evolution of cloud computing.

25. Define the terms IaaS, SaaS, and PaaS with examples.
26. What is the difference between parallel and distributed computing?
27. Identify the reasons that parallel processing constitutes an interesting option for computing.
28. What is a distributed system? What are the components that characterize it? Explain.
29. Define cloud computing. Explain its features, services, and deployment models.
30. Clarify in brief, how the cloud helps reduce capital expenditure? Explain the cloud services provided Google from a perspective of a user.
31. Will cloud services replace the existing desktop environment in the future? Analyze and give your view.
32. What is mobile cloud computing? Explain its architecture and issues.
33. What is grid computing? What kinds of applications are best suited to grid computing? Describe the workin mechanism of grid computing.
34. What is Grid Computing, describe along with basic architecture? Explain the lexicons of Grid Computing?
35. Explain cloud computing. Describe its features, services, and deployment models.
36. What is Grid Computing? Explain the standards and architecture of grid computing.
37. List some of the challenges in cloud computing.
38. What is Grid Computing? Explain the standards and architecture of grid computing.
39. What is mobile cloud computing? Explain its architecture and issues.
40. Define cloud computing and identify its core features.
41. What are the major distributed computing technologies that led to cloud computing?
42. Describe the vision introduced by cloud computing.
43. Explain the major advantage of cloud computing.
44. Briefly summarize the challenges still open in cloud computing.

❑❑❑