Chapter: 01

INTRODUCTION TO THE CLOUD COMPUTING

Mohd Naved Ul Haq*
Faculty, Glocal School of Science and Technology,
Glocal University, Saharanpur, U.P.
**Correspondence to: naved@theglocaluniversity.in*

Mr. Anuj Kumar
Faculty, Glocal School of Science and Technology,
Glocal University, Saharanpur, U.P.

DOI: https://doi.org/10.52458/9788196869434.2023.eb.grf.ch-01
Ch.Id:-GU/GRF/EB/ETCSIA/2023/Ch-01

ABSTRACT

The way that businesses manage and provide IT services has been completely transformed by the disruptive technology known as cloud computing. With previously unthinkable scalability, cost-efficiency, and flexibility, this paradigm shift is characterized by remote storage, processing, and access to data and applications through the Internet. This abstract offers a general overview of cloud computing, delving into its essential elements and demonstrating its tremendous influence on both organizations and people.

To accommodate a range of demands and preferences, cloud deployment types include public, private, hybrid, and community clouds. Reduced capital costs, enhanced scalability, and resource accessibility from anywhere are all benefits of cloud computing that encourage remote work and collaboration. It also prompts questions about security, data privacy, and possible vendor lock-in, though. To foster innovation and competitiveness in a quickly changing digital landscape, it is crucial to solve these issues as cloud usage keeps increasing while maximizing the advantages that cloud computing provides.

Keywords: *Cloud Computing, Cloud Service Models, IaaS, PaaS, SaaS*

INTRODUCTION

Despite the excitement surrounding numerous cloud services and activities, many people continue to have doubts about "the cloud" as a concept. Cloud computing has been officially defined by the National Institute of Standards and Technology (NIST) as of October 2009[1]. The definition's writers stated that cloud computing is a paradigm that is always growing, with the public and commercial sectors constantly debating and improving its definitions, use cases, underlying technologies, problems, risks, and advantages. They made a point of stating that these descriptions, qualities, and traits will inevitably change and develop through time. A casual examination of cloud computing books supports this idea. While there may appear to be several definitions of what the cloud means on the surface, the majority of them center around the commonly acknowledged

Although there may initially appear to be different definitions of what a cloud is, most of them center on the accepted concepts outlined in the NIST definition. The goal of this definition is to cover all of these foundational ideas, and it is usually considered as providing a balanced and all-encompassing framework. In-depth discussions of the five essential qualities, three service models, and four deployment types specified in the NIST definition of cloud computing will be provided in this chapter. It will also examine

instances and cloud-based offerings that perfectly illustrate each of the key ideas in the definition.[1]

CHARACTERISTICS OF THE CLOUD

It's critical to have a firm grasp of the characteristics that distinguish cloud-based services and what makes a service "cloud-like." The National Institute of Standards and Technology (NIST) has listed five defining characteristics of cloud computing:

i. On-demand self-service

ii. Access to a large network

iii. Meter-based services

iv. Elasticity

v. Resource pooling

i. On-Demand Self-Service: On-Demand Self-Service, cloud users can independently provision and manage computing resources like virtual machines, storage, and apps without needing help from the cloud service provider's staff. Through a user-friendly interface, this feature offers customers a great degree of flexibility and control by allowing them to scale resources up or down, distribute storage, and use different cloud services as needed. Cloud computing is a flexible and effective technological paradigm because it enables businesses and individuals to react quickly to shifting market demands and to optimize their resource allocation with no delays or complications.

ii. Access to a large network: A fundamental aspect of cloud computing is broad network access, which guarantees that services and resources are easily accessible online from a variety of devices. Users can access the cloud remotely and with flexibility by connecting their computers, smartphones, or tablets to it. This feature encourages remote work, mobility, and collaboration by enabling users to engage with cloud services via web browsers or application interfaces from almost anywhere with an internet connection. Cloud computing is a flexible option for companies and individuals in the contemporary digital landscape since it eliminates physical barriers and improves accessibility.

iii. Meter based services: Metered Services, a crucial aspect of cloud computing, include tracking and charging customers based on their real resource consumption. Instead of charging a set price to users, cloud service providers

utilize a pay-as-you-go approach where users are only paid for the specific resources and services they use. Users can adjust their usage up or down as needed with this pricing strategy, allowing charges to be transparent and flexible with actual demand. It is a cost-effective feature that makes cloud computing a viable and cost-effective alternative for enterprises of all sizes by eliminating up-front capital expenses and ensuring that customers only pay for what they use.

iv. **Elasticity:** Elasticity, a key characteristic of cloud computing, enables resources to be swiftly and dynamically scaled up or down to match shifting workloads. This skill makes sure that businesses can quickly and effectively, without requiring manual involvement, respond to fluctuating needs. When there is a surge in traffic, cloud services can automatically provision more resources and release them as the need decreases. Elasticity encourages cost-effectiveness by removing the need for companies to retain excess capacity to manage sporadic surges in activity. It is an important factor in flexibility, resilience, and resource optimization, enabling cloud customers to exactly match their resource allocation to their present needs while minimizing expenses.

v. **Resource pooling:** The effective distribution and sharing of computing resources across numerous users is known as resource pooling and is a fundamental aspect of cloud computing. A common pool of resources from cloud service providers' physical and virtual resources is combined, and according to demand, users are dynamically assigned from this pool. This multi-tenant strategy maximizes resource use while enhancing cost-effectiveness and scalability. The shared infrastructure reduces the need for specialized gear for users, allowing service providers to offer their services more affordable. Resource pooling is essential for maximizing the effective use of computer resources, improving overall performance, and increasing the cost-effectiveness of cloud services.[2][3]

EXPLORING CLOUD SERVICE MODELS:

There are three ways to provide customers with cloud services, according to the NIST [2] definition of the cloud:

i. Software as a service (SaaS)

ii. Platform as a service (PaaS)

iii. Infrastructure as a service (IasS)

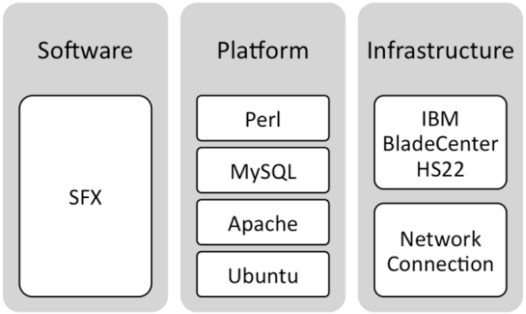

Figure 1: Cloud Service Models

i. Infrastructure as a Service (IaaS):

IaaS, a cloud computing service model, provides users with critical computing resources in a virtualized setting. Virtual machines, storage, and network infrastructure are frequently included in these resources. The unique feature of IaaS is that it offers a highly adaptable and scalable infrastructure while letting customers maintain control over the operating system, applications, and data.

In an IaaS model, the cloud service provider is responsible for managing the physical infrastructure, including servers, storage arrays, and networking hardware. They are in charge of the upkeep, safety, and scalability of these supporting elements. On the other side, users are free to deploy their software, set up their virtual machines, and control the apps they use.

IaaS is a popular option for businesses that wish to keep control of their software stack while minimizing the administrative burden of managing physical hardware. As users can quickly scale resources up or down to satisfy shifting demands, it is a great solution for situations where workloads fluctuate. For example, a business might use IaaS to power its development environments, databases, and web applications. Through this arrangement, they can quickly provide resources as needed, ensuring peak performance without the accompanying capital costs of running an internal data center.

ii. Platform as a Service (PaaS):

Software as a Platform (PaaS) represents a step beyond IaaS, delivering a full development and deployment platform in addition to infrastructure resources. For the development, testing, and deployment of applications, PaaS offers a full range of tools, services, and environments. In this paradigm, users are in charge of the apps and data

while the platform looks after the supporting infrastructure, such as the operating system.

PaaS is especially enticing to developers since it simplifies the software development process, lowers operational complexity, and promotes team communication. The platform maintains the hardware, networking, and other infrastructure components, allowing developers to completely concentrate on developing code. This allows for a quicker time to market by dramatically accelerating the creation of applications.

Platforms as a service (PaaS) provide a selection of databases, middleware, and application services. They frequently offer support for a variety of frameworks and programming languages, allowing them to handle many different kinds of applications. Popular PaaS services include Heroku, Microsoft Azure App Service, and Google App Engine.

To develop and deliver its applications, for instance, a mobile app development company might employ a PaaS paradigm. Developers may focus on creating feature-rich apps while the platform takes care of the challenging aspects of server management and scaling.

iii. Software as a Service (SaaS):

The cloud service paradigm with the most customer focus is called Software as a Service (SaaS). SaaS provides complete software applications via the internet, which customer's access through web browsers. In contrast to IaaS and PaaS, where users retain control over elements like the operating system and application code, a SaaS model delegates administration of the complete software stack, including infrastructure management, maintenance, and security, to the service provider.

SaaS applications span a wide range of software categories, such as email services (like Gmail), customer relationship management (CRM) tools (like Salesforce), productivity tools (like Microsoft Office 365), and collaboration software (like Slack). SaaS is ideal for people and companies looking for affordable, pre-configured solutions that need little to no setup or configuration.

SaaS's simplicity and usability are two of its main benefits. The provider manages upgrades, backups, and security, and users may access the software from almost any device with an internet connection. As a result, enterprises can concentrate on using the software rather than keeping it up-to-date, which lessens the workload on IT staff.

In conclusion, IaaS, PaaS, and SaaS are considered as 3 main service models covered by cloud computing. Each model provides a different degree of control and accountability, allowing users to choose the one that most closely matches their unique requirements and goals. Cloud computing offers adaptable solutions for both enterprises and individuals, whether it be for infrastructure management, application development, or ready-to-use software.

UNDERSTANDING CLOUD DEPLOYMENT MODELS

Within the domain of cloud computing, disagreements are common, especially in regard to the precise definition of what a cloud is. This disagreement is never more evident than when considering the various cloud computing deployment models. To further complicate this ongoing discussion, the National Institute of Standards and Technology (NIST) has developed a framework that divides cloud deployment into four different models.

i. Public Cloud

ii. Community Cloud

iii. Private Cloud

iv. Hybrid Cloud

i. Public Cloud:

In terms of deployment models for cloud computing, the public cloud may be the most well-known and popular. It is defined by cloud service providers giving the general public access to computing resources and services via the internet, which enables a wide range of consumers. It is well-suited for a broad range of use cases, from hosting websites and applications to data storage and processing. It offers rapid deployment, cost-efficiency, and easy access to a broad set of services and resources. While it is highly accessible and flexible, users should carefully consider security, data privacy, and regulatory compliance when using public cloud services.

The public cloud deployment model's salient characteristics and specifics include:

- **Accessibility:** Anyone with an internet connection can use the resources in the public cloud. With the help of this approach, businesses and individuals can use subscription-based or pay-per-use computing services.

- **Efficiency in Terms of Costs:** Utility-based pricing is the norm for public clouds. Since there are no upfront capital expenses, users are only charged based on their actual usage, which can result in significant cost savings.
- **Scalability:** Users can simply scale resources up or down to handle changing workloads thanks to public clouds' excellent scalability. This adaptability is especially beneficial for new organizations and those with fluctuating computing needs.
- **Managed Infrastructure:** The servers, storage, networking, and data centers that make up the underlying infrastructure are managed and maintained by public cloud service providers. Users are relieved of operational duties including hardware upkeep, security, and upgrades thanks to this.

  Examples: Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform (GCP), IBM Cloud, and Oracle Cloud are notable public cloud service providers.

ii. Community Cloud:

The community cloud is a less common but highly specialized deployment model, shared among a specific group of organizations that have common interests or compliance requirements. It allows members of the community to pool their resources while maintaining a degree of separation and control.

Key features and details of the community cloud deployment model include:

- **Shared Infrastructure:** Multiple organizations within the same community share the same cloud infrastructure and services, allowing them to distribute costs and pool resources efficiently.
- **Customization:** Organizations within the community can customize the cloud environment to meet their specific needs while adhering to shared community standards and requirements. This customization ensures that the cloud environment aligns with the community's unique goals.
- **Security and Compliance:** Community clouds are designed to meet the specific security and compliance requirements of the community members. This makes them particularly well-suited for highly regulated industries such as healthcare, education, and government.
- **Cost-Sharing:** Organizations within the community share the costs of infrastructure, maintenance, and management. This results in significant cost

savings, making it an attractive option for organizations with common interests.

Examples: Various community cloud solutions are tailored to specific sectors and industries. One example is the Federal Risk and Authorization Management Program (FedRAMP), which is designed to meet the cloud security requirements of U.S. government agencies.

Community clouds are an excellent choice for organizations with shared interests or regulatory requirements. By pooling resources, organizations within the community can achieve cost savings and compliance while maintaining a high level of control over the cloud environment. These deployments cater to sectors where data privacy, security, and specialized compliance are paramount.

iii. Private Cloud:

Private clouds are dedicated to a single organization, providing a higher degree of control, security, and customization. In this model, the organization owns and manages its cloud infrastructure, which can be hosted on-premises or by a third-party provider.

Key features and details of the private cloud deployment model include:

- **Dedicated Resources:** Private clouds offer exclusive, dedicated computing, storage, and networking resources for a single organization. This ensures greater control and predictability over resource allocation.
- **Customization:** Organizations have the freedom to customize their private cloud environments, tailoring them to their specific needs. This includes configuring network settings, security policies, and resource allocation to meet specific requirements.
- **Security and compliance:** Private clouds are recognized for their strong security and data privacy protections. They are ideal for sectors with strict regulatory compliance requirements, such healthcare, finance, and government, where sensitive data and compliance with regulations are crucial.
- **Control:** Businesses maintain total control over all aspects of their private cloud infrastructure, including the hardware, software, and security

regulations. For companies with stringent governance, security, and performance requirements, this level of control is ideal.

Examples: Private cloud solutions encompass various offerings, including VMware Cloud, OpenStack, Microsoft Azure Stack, and Oracle Cloud at Customer.

Private clouds are particularly suitable for organizations that require complete control over their infrastructure, data, and applications. They are ideal for highly regulated industries and companies that need to maintain sensitive data on-premises or in a private hosted environment. Private clouds offer customization, security, and reliability, although they may involve higher upfront costs and ongoing management responsibilities.

iv. Hybrid Cloud:

The hybrid cloud is a deployment model that combines elements of both public and private clouds. It allows data and applications to be shared and moved between these environments, providing a balance between cost-efficiency and control.

Key features and details of the hybrid cloud deployment model include:

- **Integration:** Hybrid clouds involve the seamless integration of public and private cloud environments. This integration is often facilitated through cloud orchestration tools and APIs, enabling data and workloads to be moved between the two environments.
- **Flexibility:** Businesses can utilize public cloud resources for scalability during peak demand, while maintaining critical data and applications in a private cloud for security, compliance, and control.
- **Cost Optimization:** The hybrid model enables organizations to optimize costs by utilizing the public cloud for non-sensitive workloads, and the private cloud for mission-critical or regulated data. It ensures that businesses only pay for the computing resources they require.
- **Resource Redundancy:** Hybrid clouds offer redundancy and disaster recovery capabilities by allowing data and applications to be mirrored between public and private cloud environments. This redundancy enhances business continuity and mitigates risks.

Examples: Numerous cloud providers offer hybrid cloud solutions, including AWS, Azure, Google Cloud, and IBM Cloud. Additionally, organizations can create custom hybrid setups by integrating their private and public cloud resources.

The hybrid cloud is becoming a best choice for businesses seeking to leverage the strengths of both private and public clouds. It is particularly valuable for enterprises with fluctuating workloads, diverse application requirements, and the need for disaster recovery and redundancy. By combining public and private cloud resources, organizations can optimize costs, enhance scalability, and ensure data resilience.

In summary, the four primary deployment models in cloud computing—public, private, hybrid, and community—offer organizations a range of options to align with their unique requirements, objectives, and preferences. The choice of a deployment model depends on factors such as the organization's industry, data sensitivity, regulatory compliance, scalability needs, and budget constraints. Cloud computing provides the flexibility to select the right deployment model for specific use cases, ensuring that businesses can optimize their resources and technology investments.

CONCLUSION

NIST's (National Institute of Standards and Technology) explanation of cloud computing. The framework is laid by this founding document for understanding cloud technology's fundamental ideas. People can identify the distinctive qualities that cloud services have by looking at the characteristics that make up the cloud. Furthermore, comprehending the four NIST deployment models clarifies the many contexts in which cloud systems might be used.

The three service models—Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS)—offer a defined framework for how cloud providers deliver their services to clients. These models show the various ways that users can access and use cloud resources, providing flexibility and customization based on particular requirements. The field of cloud computing is constantly evolving, with new developments and changes taking place all the time. Undoubtedly, the meanings and conceptions associated with the cloud will change as additional suppliers enter the market. New products and creative concepts will surface, adding to the technology's dynamism. Despite this, one thing is for certain: cloud computing is here to stay and will keep revolutionizing how IT departments operate by providing greater efficiency, scalability, and agility.

REFERENCES

1. *Mell, P., & Grance, T. (2010). October 7, 2009, "The NIST Definition of Cloud Computing,". National Institute of Standards and Technology.*
2. *Ul Haq, M. N., & Kumar, N. (2021). A novel data classification-based scheme for cloud data security using various cryptographic algorithms. International Review of Applied Sciences and Engineering.*
3. *Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.*