**Atvdd\_AOC\_01**

**Atividade de Pesquisa – Cloud Computing**

**Colocar Nome e R.A.**

Efetuar uma pesquisa - **Cloud Computing** ou Computação em Nuvens.

**Apresentar:**

* Surgimento;
* Características;
* Segurança;
* Arquitetura;
* Vantagens e Desvantagens.

**Apresentar as Fontes de Consulta.**

# Introdução:

Este trabalho tem como objetivo apresentar uma breve atividade de pesquisa sobre *Cloud Computing* (Computação em Nuvem) tendo em conta os tópicos: surgimento do conceito de *Cloud Computing*, características, segurança, arquitetura além de suas vantagens e desvantagens.

*Cloud Computing* é considerado um conjunto de serviços de TI (Tecnologia da Informação) fornecidos ao cliente através de uma rede juntamente com a habilidade de se reduzir os requisitos de serviço desejado. A Computação em Nuvem tem o potencial de eliminar os requisitos de infraestrutura com alto custo para serviços e soluções baseadas em TI usados pela indústria, prometendo oferecer uma arquitetura de TI flexível e acessível através da *Internet* (Rede Mundial de Computadores) e dispositivos portáteis.

A Computação em Nuvem é fornecida por acesso sob demanda, através da Internet, para utilização de recursos computacionais, de aplicações, servidores (servidores físicos e virtuais), armazenamento de dados, ferramentas de desenvolvimento, capacidade de rede com hospedagem em Data Centers remotos gerenciados por um provedor de serviços (CSP – Cloud Services Provider). Os CSPs tornam estes recursos disponíveis mediante uma taxa de assinatura mensal ou cobrança de acordo com o uso.

# Breve histórico do surgimento da Computação em Nuvem

A ideia atual de Computação em Nuvem data dos anos 1950 quando *mainframes* de larga escala eram disponibilizados para universidades e corporações. A infraestrutura colossal de *hardware* de um *mainframe* era instalada no que era chamada de Sala do Servidor (já que esta sala era capaz de comportar apenas um *mainframe*). Múltiplos usuários eram capazes de acessar os recursos do *mainframe* através dos “*dumb-terminals*” (terminais-burros) que eram estações com a única funcionalidade de facilitar o acesso ao *mainframe*.

Devido ao alto custo de se comprar e manter um *mainframe*, as organizações não eram capazes de oferecer um dispositivo para cada usuário. Se tornou prática permitir que múltiplos usuários compartilhassem acesso a mesma camada de armazenamento de dados e recursos de CPU, ou seja, poder de processamento, de qualquer estação. Com a disponibilização de um *mainframe* de acesso compartilhado, uma oprganização poderia ter um melhor retorno de investimento feito em seu equipamento.

O cientista da computação norte-americano, John McCarthy , inventor do termo “Inteligência Artificial”, defendeu a proposta de “time-sharing” ou computação por tempo compartilhado. McCarthy, disse que a computação poderia permitir que um computador fosse utilizado simultaneamente por dois ou mais usuários. Desta forma, as pessoas poderiam realizar tarefas, aproveitando o período de tempo disponível dos recursos. Na visão do cientista, o compartilhamento, além de reduzir gastos, permitiria pagamento somente pelo período utilizado.

Esse modelo foi apresentado por McCarthy durante um discurso no Massachusetts Institute of Technology (MIT), nos EUA, em 1961. Ele sugeriu a criação da “Utility Computing” ou computação como serviço de utilidade pública, no mesmo sentido do fornecimento de água, luz ou telefone, que chega às residências ou empresas, sem que as pessoas saibam de onde vem.

Em 1963, DARPA (Defense Advanced Research Projects Agency – Agência de Projetos de Pesquisa Avançada de Defesa), presenteou o MIT (Massachussets Institute of Technology) com 2 milhões de dólares para o Projeto MAC (MIT Project on Mathematics and Computation). O fundo incluía um requerimento para que o MIT desenvolvesse uma tecnologia que permitisse o uso de um computador por duas ou mais pessoas simultâneamente. O projeto definiu-se inicialmnente como uma forma primitiva de Computação em Nuvem, com apenas duas ou três possoas ralizando acesso simultâneo.

O Projeto MAC inicialmente foi dirigido pelo Cientista da Computação Robert M.Fano, tendo o Cientista da Computação Fernando José Corbató como membro fundador. O termo Projeto foi usado no lugar de Laboratório para isnpirar o pessoal do MIT a juntar-se à iniciativa sem desafiliar-se dos seus laboratórios à época. Uma das primeiras contribuições do projeto foi expandir e fornecer o hardware para o software CTSS (*Corbató’s 1961 Compatible Time-Sharing System*) o que permitiu com que multiplos usuários em terminais espalhados pelos departamentos do MIT rodassem programas centralizados em uma única máquina *mainframe*.

Em 1969, . C. R. Licklider (Joeph Carl Robnett Licklider), um Psicólogo e Cientista da Computação estado-unidense, ajudou a desenvolver a ARPANET (Rede de Agências para Projetos de Pesquisa Avançadas - *Advanced Research Projects Agency Network*) o embrião da rede mundial de computadores atualmente conhecida como Internet, e promoveu a ideia de “*Intergalactic Computer Network*” (Rede Intergalática de Computadores) todas as pessoas do planeta estariam iterconectadas por computadores capazes de acessar informação de qualquer lugar. J. C. R. Licklider acreditava que o CTSS facilitaria grandemente a eficiência, redução de custos, a ganharia tempo ao permitir que vários usuários compartilhassem dos mesmos recursos computacionais de um enorme mainframe em vez de empregar pequenas máquinas individuais.

Dentro de seis meses da criaçlão do Projeto MAC, 200 usuários estavam aptos a acessar o sistem de 10 departamentos diferentes no MIT. Em 1969, *Project MAC*, *Bell Laboratories*, e *General Electric* juntamente desenvolveram o Multics (*Multiplexed Information and Computing Service*) que evolouiu de computador de tempo compartilhado para um sistema de computador online e incorporou funcionalidades como compartilhamento de arquivos, gerenciamento e sistema de segurança em seu *design*. O complexo sistema agora poderia suportar 300 conexões simultâneas em 1000 terminais do MIT e levou a Bell Labs a empregar uma forma mais simples do sistema operacional UNIX.

A palavra “virtualização” começou a ser usada para descrever esta situação, embora seu significado tenha mudado posteriormente para algo mais amplo.

O termo “*Cloud*” (Nuvem) e sua primeira referência data dos anos 1990 e origina-se no mundo das telecomunicações, onde historicamente estas companhias ofereciam apenas uma conexão dedicada de dados ponto a ponto. Estas companhias começaram a oferecer o serviço de Rede Virtual Privada (VPN – *Virtual Private Network*). Em vez de construir uma infraestrutura física para permitir que mais usuários tivessem suas próprias conexões, as empresas de telecomunicações forneceram aos usuários acesso compartilhado à mesma infraestrutura física. Essa mudança permitiu que as empresas de telecomunicações redirecionassem o tráfego de rede em tempo real conforme necessário, levando a um melhor equilíbrio da rede e maior controle sobre o uso da largura de banda por um custo mais baixo. O termo “*Cloud*” em *Cloud Computing* refere-se portanto, a *Internet* e sua infraestrutura.

O significado de virtualização começou a ser difundido nos anos 1970, e hoje é definida como a criação de uma Máquina Virtual, que atua como um computador real com um sistema operacional totalmente funcional. O conceito de virtualização desenvolve-se juntamente com a Internet, a partir do momento que as VPNs se tornaram um negócio rentável levando ao desenvolvimento da Infraestrutura de Computação em Nuvem Moderna.

# Características essenciais da Computação em Nuvem

Cloud computing platforms provide easy access, scalability, reliability, reconfigurability, and high performance from its resources over the Internet without complex infrastructure management by customers.

Cloud computing is nothing but the integration of Distributed Computing, Parallel Computing, Utility Computing along with Network Storage, Virtualization, Load Balance,

High Available, and various other related technologies. Cloud computing is defined in several ways, but National Institute of Standards and Technology (NIST) defines that “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”

Grid computing is a distributed computing system where a group of computers coupled to form a virtual machine to perform large tasks. Utility computing is a method of packaging several services such as facilities, storage, and so on for billable IT resources. Autonomic computing is a system that has selfmanagement capabilities. Cloud computing capabilities depend on the clusters deployed (grid computing) with various functionalities of Utility and Autonomic computing.

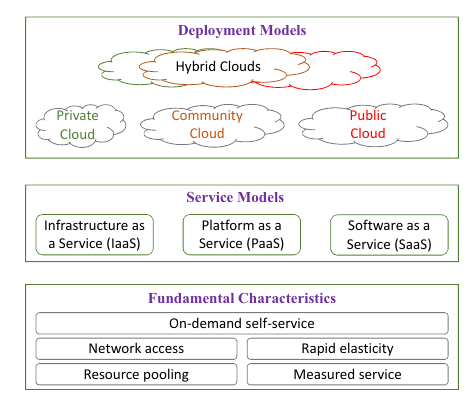
## fundamental characteristics, delivery models, and deployment

## models.

• Utility computing : Around 1960, the processing or computing power prices were high for any purpose, so they came up with the idea of sharing computing resources. Its goal is to integrate servers, storage systems, and applications distributed around the world to share with multiple users. Sharing would allow the users to use and share the computer resources and customers using the resources can pay for the services used for the period they have used services only [4].

• Grid computing : is the process of solving the massive computing problems into smaller problems parts and solve these problems on simple or low-performance machines or computers to get the final result for large problems by distributing the tasks among various machines on the grid [3].

• Cloud computing : The concept of cloud computing is very similar to that of utility computing and grid computing. With the evolution of technologies over a couple of decades, cloud computing is possible to reach its goal in the last couple of years. In this decade cloud computing has matured a lot in terms of technology to meet all its demands.

2) Hybrid Cloud Computing

Nowadays some of the organizations started using both private and public cloud services, for various reasons such as cross-platform evaluation, to check applicability in real time scenarios across multiple platforms, and so on. Usage of public cloud service providers services even when they have private cloud services has to lead to the development of hybrid cloud computing, to have compatibility or connectivity between different cloud computing services [16].

3) Mobile Cloud Services

As the usage of mobile devices has increased in the last decade, data generated with individuals has also increased tremendously. To gain more customers, many of the cloud service providers have started supporting mobile devices for using or accessing cloud services. Using these mobile applications or interfaces customers can store their data or access the services provided by cloud service providers [17].

4) Cloud Security

With the evolution of technology and cloud services many of the users are using the cloud services, but still, there is an essential problem in the cloud which needs to be addressed, such as, data security in the cloud. There are many encryption techniques and security protocols to protect the data, but with the rapid growth in technology and processing power available to attackers it becomes easy to break some of the existing encryption techniques. So, still, there is scope for new encryption techniques and security protocols for safe and secure future operations of cloud computing [18].

5) Cloud Design

Development of cloud services has lead to rapid adoption of the services even in traditional markets. Cloud-centric markets have advantages over conventional markets because it is convinent for younger generation to use the services. Cloud based markets attracting more new customers without physically being present in many locations with various stores in each city [19].

**Characteristics**

Cloud computing is a distribution of a massive computational power accessible over the Internet, rather than on the local machines. Organizations with their private data centers also work on a similar principle. Cloud computing allows organizations to move their resources where they need more processing power for their applications instead of wasting the resources that are not utilized at their full potential. They are moving their traditional processing powers to centralized processing of data in their data centers. The shift allows computing or processing power as a commodity which can be traded over the Internet. Common characteristics of cloud computing are large scale, virtualized, low cost, geographically distributed, service oriented, resilient computing, and advanced security for services. Apart from above characteristics, cloud should also have following fundamental characteristics [2] as shown in Fig. 2:

1) On-demand self-service : A end-users with the need to use the computing resources at a particular time (e.g., CPU time, network storage, software, and so on) automatically and conveniently, start and stop using them without any human interference.

2) Network access : The computing resources delivered over the Internet and can be used by various applications from different types of devices such as laptops, desktops, and mobile devices as per the end-users requirement and availability.

3) Resource pooling : Cloud service providers pools (groups) all of their computing resources together to serve multiple end-users using multi-tenancy model “with different physical and virtual resources dynamically assigned and reassigned according to consumer demand” [2]. End-user should be able to use resources irrespective of their location to support location independent resource pooling.

4) Physical transparency or Rapid elasticity : End-users can change their resource capabilities automatically to scale up whenever they want to use more resources and release them once they are done using the services to scale down. For end-users, the resources are available for configuring with simple steps to scale up their operations and vice versa. These resources are not limited to end-users; they increase the usage of services to meet their peak requirements at any time.

5) Pay peruse or Measured Service : Even though all the resources are pooled and shared among multiple tenants, the end-users needs to be charged only for the cloud services they have used. This has to be taken care with a proper mechanism to measure the services used by each customer.

III. CHARACTERISTICS OF C LOUD COMPUTING

TABLE I. CLOUD COMPUTING VS. GRID COMPUTING CHARACTERISTICS

|  |  |  |
| --- | --- | --- |
| Characteristic | Cloud Computing | Grid Computing |
| Service oriented | Yes | Yes |
| Loose coupling | Yes | Half |
| Strong fault tolerant | Yes | Half |
| Business model | Yes | No |
| Ease use | Yes | Half |
| TCP/IP based | Yes | Half |
| High security | Half | Half |
| Virtualization | Yes | Half |

The cloud computing, grid computing, High performance computing (HPC) or supercomputing and data center computing all belong to parallel computing. HPC focuses on scientific computing which is computing intensive and delay sensitive. So high processing performance and low delay are the most important criterias in HPC. Grid computing is based on HPC center. Many connected HPC centers form a large grid which owns a powerful underlying concept – service-oriented architectures (SOA). Some other creative and impressive concepts like utility computing and autonomic computing do not come into reality. The cloud computing which is based on data center is much more widely accepted than grid computing. Data center which doesn’t only pursue powerful processing performance and low delay is more balanced than HPC center. The comparable characteristics of cloud computing and grid computing are listed in Table I. The “yes” and “no” stand for cloud computing or grid computing has the special characteristic or not. The “half” means not owning the whole characteristic to a certain extent. This paper doesn’t pay much attention on the similarities and difference between them and focuses on the essential characteristics of cloud computing. The more detail comparisons can see [6].

A. Conceptional characteristic – service oriented

The service oriented concept is similar to but more practical than the concept of SOA in grid computing. Abstraction and accessibility are two keys to achieve the service oriented conception. Through virtualization and other technologies, the underlying architecture is abstracted without exposing much to user. So it is opacity to cloud user. Abstraction reduces both the need for cloud user to learn the detail of cloud architecture and the threshold of application development. At the same time, the key elements of underlying architecture can be simply accessed by cloud user. Cloud user can consume all the capacity easily by exploring system parameters such as processing performance and storage capacity. In general, according to the type of provided capability, the services of cloud computing are broadly divided into three categories: Infrastructures-as-Service (IaaS), Platform-as-a-Service (PaaS), and Software-as-a-Service (SaaS) [5, 6].

Infrastructure-as-a-Service is the delivery of huge computing resources such as the capacity of processing, storage and network. Taking storage as an example, when a user use the storage service of cloud computing, he just pay the consuming part without buying any disks or even knowing nothing about the location of the data he deals with. Sometimes the IaaS is also called Hardware-as-a-Service (HaaS).

Platform-as-a-Service generally abstracts the infrastructures and supports a set of application program interface to cloud applications. It is the middle bridge between hardware and application. Because of the importance of platform, many big companies want to grasp the chance of predominating the platform of cloud computing as Microsoft does in personal computer time. The well known examples are Google App Engine [3] and Microsoft’s Azure Services Platform [15]. Software-as-a-Service aims at replacing the applications running on PC. There is no need to install and run the special software on your computer if you use the SaaS. Instead of buying the software at a relative higher price, you just follow the pay-per-use pattern which can reduce you total cost. The concept of SaaS is attractive and some software runs well as cloud computing, but the delay of network is fatal to real time or half real time applications such as 3D online game.

There are also many cloud resources can not rank into infrastructure, platform or software. Apple’s App store [20] is a creative and famous cloud computing in wireless area. Software services are sold in pay-per-use style. But running on terminals such as 3G phones instead of huge data centers is different from SaaS. In online game area, powerful game servers supply the interactions of millions game players. Game players use the capability of cloud computing without much waking up to this technology. The electric books resources in Amazon [21] are also services in cloud computing. These services hardly have substitution and building another EC2 is much more easer than owning so many electric resources. All these services are as important as IaaS, PaaS and SaaS and should be fully studied.

B. Technical characteristic 1 – loose coupling

The loose coupling is the technical fundament of cloud computing and goes beyond the loose coupling method of application interaction. Through virtualization or other technologies, the infrastructures are separated in logic or physic. The behavior of one part hardly affects other parts. For example, the platform is an abstract layer which can isolate different applications running on it. Most important of all, whole cloud computing runs in a client-server model. The clients or cloud users connect loosely with servers or cloud providers. All the users have almost no data or control dependence. But the data dependence plays a key roll in HPC. We can just have formalizations:

Users make up of user sets Uset 1 , Uset 2 , ..., Usetm

( m ≥ 1 ). Providers make up of provider sets Pset 1 , Pset 2 , ..., Pset n ( n ≥ 1 ). User set Uset i loose coupling with provider set Pset j is described as Set ( Useti , Pset j ) .

Three properties are showed as following:

User sets are independent: Uset i ∩ Uset j = φ ( 0 ≤ i, j ≤ m , i ≠ j ).

Provider sets are independent: Pseti ∩ Pset j = φ ( 0 ≤ i, j ≤ m , i ≠ j ).

The loose coupling (cloud user connects to cloud provider) sets are independent:

Set ( Useti1 , Pset j1 ) ∩ Set ( Uset i 2 , Pset j 2 ) = φ .

Taking search as a simple example, the providers are Google, Yahoo! and Bing. The search service users can not use three search engines at the same time (the absolute time) and can divided into independent user sets UsetGoogle ∩ UsetYahoo! ∩ Uset Bing = φ . The data centers

behind simple search interfaces are independent between the three giants Pset Google ∩ PsetYahoo ! ∩ Pset Bing = φ . The loose coupling sets are independent Set ( Uset Google , Pset Google ) ∩ Set ( UsetYahoo ! , PsetYahoo! ) ∩ Set ( Uset Bing , Pset Bing ) = φ .

The independent cloud users can induce many other features such as stateless, scalability and strong fault tolerant. An opposite example is the tight coupling of HPC systems which focus on solving scientific problems. Usually, there are too many data dependences or global synchronizations in different iterations to bear the high delay among computing nodes. This kind of systems use high speed network e.g. InfiniBand instead of industrial standard Ethernet which is much cheaper and widely supported. It is hard to imagine spending tens of minutes to achieve a global synchronization in HPC. But in cloud computing one time of global data synchronization may cost several hours or even several days.

C. Technical characteristic 2 – strong fault tolerant

There are many fault tolerant methods in parallel computing. At low-level, there always exist some fault correction mechanisms with specific hardware. At high-level, many specific applications are studied with methods aiming at algorithms. Checking point is one of the most effective methods at middle-level. In large scale parallel computer systems, the interval of two failures may be shorter than application execution time. For example, some scientific computing applications run for weeks or even longer but there may be several trivial or fatal errors during the whole runtime. The fault tolerant technology becomes critical in this condition. Otherwise it has only less chance to complete the time consuming computing tasks. Because a minimum error is unacceptable and redoing costs too much time in many scientific applications, so the whole computation states which are saved periodically on stable storage will roll back to a special checking point if an error occurs. It is unnecessary to keep the whole states of cloud computing systems. There is almost no dependence between two transactions. The failure of one transaction does not affect the other one and partly failure of system will not cause chain reaction.

There are mainly four places where faults maybe occur in cloud computing: provider-inner, provider-across, provider- user and user-across. If a fault occurs in provider, the backup or redundancy of provider will substitute for the failed part. Stop services and restart are another common method if the services are not on time or urgent. The loose coupling nature of provider e.g. data center makes this kind of faults not hard be deal with.

If a fault occurs among providers, the provider-across transaction will be canceled and return with an error hint. Redirecting to other providers is a universal method which involves load balance of whole cloud system. Fortunately, there are only fewer transactions, which are caused by background management in the main, involving more than one provider. It needs only to run background management one time per day or even per week. There are too many reasons such as network congestion, browser collapse, request time out, provider busy and hacker attack can cause faults between provider and user. If not involving some key elements, these faults are omitted and user can try next time. The Byzantine fault tolerant algorithms are very important on the aspect of technology because malicious provider-user errors are increasingly common and can cause faulty nodes to exhibit arbitrary behavior which is hard to deal with [22]. If involving key elements which cause real lost to user such as money in personal account, additional operations are needed to ensure the security of transaction. At the same time, the system log and credit of provider can deal with these accidents. And corresponding laws are supposed to solve all these at last. User does not only connect with provider but also other users. Many users attend activities and share several critical resources. On this condition, unsafe accessing critical resources can cause chaos in cloud computing systems. There exist hardware level, operating system level and software level methods to protect critical resources. The provider will sit on the fence to arbitrate dispute among users. If all these are not enough, the law will stand in the end. So it is not terrible when a fault occur.

Both the mature of technology and society guarantee the strong fault tolerant characteristic of cloud computing.

D. Economic characteristic – business model

The business model is the key characteristic to distinguish grid computing and cloud computing. The grid computing is mainly supported by government and academe. On the one hand, this determines the nature of grid computing: the impulse of profit is not strong enough. On the other hand, the grid computing is a research for future development of information technology. But the cloud computing is mainly supported by gigantic IT companies. They plan that all investments on cloud computing should get return on investment (ROI) in the near future or beat market competitors in the long run.

There are many business models especially how-to-pay models in cloud computing. Pay-per-use may be the favorite one in many cases. This is almost the same as the concept of utility computing. The capacity of processing, storage and network in cloud computing is utility service as water, electricity and gas in society [11]. These utility services can be available whenever the user requires them at any time in modern human society. Users pay service providers based on their usage of these utility services. There are two categories of cloud users: end user and median user.

Cloud services are ends in themselves for end user. End user consumes cloud services for self use. Median user consumes cloud services and cost efficiently supplies professional services to others. End user sometimes doesn’t pay for cloud services directly. For example, online game players pay for special game according to how long they stay online. And part charge is defrayed to maintain the running of cloud system. This process is opacity to end user. Median user usually pays for consumed cloud services directly. They save money on jumping to the market quickly. For median user, it is no need to manage complex hardware and software, learn how to use tools and gain experience with cloud computing technology. But the business of cloud computing is far beyond these. There are many free accesses to cloud computing. It is free to search with Google, send email with hotmail and find new friends with Facebook. These conditions rarely appear in using other social utility services. Maybe air is an exception but clean air may be not free anymore in the future. We still don’t know how long we can breathe freely due to air pollution. The strategy of sustaining these free cloud services is out of the scope of this paper. Attention economy may be the key idea.

E. User experience characteristic – ease use

User experience which belongs to the subject of human computer interaction is an important criterion when evaluating whether an application is successful or not. In cloud computing, user experience improves a lot than its ancestors like grid computing. The cloud service is a means toward the end of providing a good experience for cloud user. The valuable services should be easily accessed by cloud user. The core of user experience is achieving ease use. Ease use is not only simple but also elegant.

There are three reasons why cloud computing should be ease use:

First, most cloud providers offer Internet-based interfaces which are simpler than other application program interfaces (API). These interfaces are simple and elegant enough to hide the business processing behind. The interfaces can stay the same ignoring whether the business processing has changed or not.

Second, user experience of web applications is full studied. So the user interfaces are independent of content. The development of web application has a full suit of flow which can be divided into three stages including user need analysis, function design and program implementation. In top- down method, the user experience design is the fundamental of whole function design. The facets of the user experience are useful, usable valuable, desirable, findable, credible and accessible [23].

Third, the web 2.0 increases the interactions between web users and providers. The web was originally designed to transport hypertext. As the rapid and rich developments of increasingly sophisticated contents are appearing, web is usually used as a remote software interface. The web 2.0 is supposed to be the continuum of user experience and blurs the line between software and the Internet. The emerging AJAX technology makes web applications and services are becoming more software-like [24].

All these reduce information technology overhead for the end user. Search in the Internet is very simple: a web page with an input text and a confirm button.

F. Other characteristics

There are other important characteristics such as TCP/IP based, virtualization and high security. TCP/IP gives reliable delivery, a connection-oriented service between remote applications. TCP/IP is widely used in cloud computing. Although the network protocols may be private in the back end of data center, most cloud users connect to providers through TCP/IP. The HTTP protocol over TCP/IP or Internet inspires the user experience characteristics. Cloud resources are often virtualized as a service over the Internet. Up to the present, many cloud computing infrastructures consist of data centers. Data center uses virtualization technologies which abstract the commonness of infrastructure in different levels.

High security of cloud computing is achieved mainly through three ways. First, the loose coupling makes cloud computing system run well when part of it is destroyed. Second, the abstraction, virtualization and privation of cloud provider avoid exposing the details of corresponding implementations. Third, technology cooperating with law is the guard of cloud computing.

**Segurança:**

CLOUD SECURITY

Cloud security has evolved with the adoption of Cloud Computing. The concept of cloud security is more more critical with the adoption of cloud services by more users. With many of the users around the world, if the cloud services are not protected adequately, it leaves vast amounts of customers data vulnerable to attackers from all over the worlds.

Cloud security can be achieved in several forms, protection against the Network attacks, Software attacks, Intrusion Detection, Access control, Analysis of abnormal behavior, Analysis of Virus, Analysis of Malware, Analysis of Trojans and so on.

Security measures, which ensures cloud security are presented below:

1) Password : To secure the cloud services from simple attacks against the access controls, users are encouraged to use a unique password for accessing the cloud-based services. Customers should not use simple passwords or reuse the password which has been used on some other services over the Internet. Cloud service providers should make sure that there is no direct relation to user names and passwords stored in their database. In case there is a breach on the cloud service provider, it makes it hard for attackers to match the user names to passwords [29].

2) Access Recovery : Customers should use confidential details or questions, for recovering their access control to the cloud in case they forgot their password. This information can be used to recover access to the Cloud. Users should not use the information which can be gained by using social engineering or just checking some information on their social networking profiles. As most of the personal details are posted on the networking websites. Using such information, attackers can easily gain access to the Cloud without knowing the person.

3) Encryption : Using a good encryption technique by cloud service providers always protects the customer’s data, such as Homomorphic Encryption. Usage of a homomorphic encryption technique is still not completely feasible in real time scenarios [30].

4) Password Management : As discussed in the first point, users should not reuse their passwords, and cloud service providers should encourage them to use strong passwords with special characters, symbol, alphabets, and numbers. It is tough for users to remember all of their user names and passwords. They need to have a proper management tool for storing their user names and 5 passwords to protect them from anyone getting access to them [31].

5) Multi-factor authentication : Multi-factor authentication adds an extra layer of security to the traditional approach to access the Cloud services instead of user name and password. To access the cloud services using multi-factor authentication, customers need to have two or more factors to access the cloud services to authenticate them as a genuine user of the Cloud. These factors can be based on anything such as knowledge (something known to the user, such as an other password), something user has (Biometric features), and something user possesses (RSA key or USB based keys or random text sent to their mobile) [32].

Cloud services providers should support multi-factor authentication methods and encourage the customers to use the Multi-factor authentication instead of using simple authentication using user name and password. In this way, it will be easy to defend against unauthorized access to customers data even if someone has customers credentials; they wont be having access to other factors.

6) Login Monitor : Cloud service providers and customers need to monitor recent devices used to access cloud services. Based on that information users can identify if someone has logged in with their credentials and change their passwords in case of a suspicious login from unknown devices or locations. Cloud service providers need to improve the login statistics with proper details for all the devices connected to access the Cloud Services for all the customers [33].

7) Personal Devices : Customers should be careful where they are logging in to cloud to access the services. They should avoid using someone else device, as they might have key loggers (a program which saves all the keys pressed on a device, while the program is running). In those devices, if they have such applications, attackers will gain user credential for the Cloud compromising security for customers [34].

8) Virus, Malware, and Trojans : Customers should have good anti-virus and anti-spyware applications on their devices. If they dont have proper protection of their devices, which they use to use the cloud services might have some viruses or malware which store the user credentials and gain access to the cloud services leaving their personal and confidential details into the unauthorized persons or attackers. It would be a good habit for users to have good anti-virus and anti-spyware applications to protect their personal devices [35].

Cloud computing at present has matured a lot and solves many of the simple security aspects. Still, there are many open challenges which need to be addressed for more growth in Cloud Computing Industry.

SECURITY ISSUES IN CLOUD COMPUTING

1. Software-as-a-service (SaaS)

Through this model, the service provider of cloud provides database and application software access. SaaS is a software with high demands. The problems which are faced with this application is with its security which are naturally centered revolving around the access and the stored information as almost all the models which are responsible for the data sharing security issues leave these 2 issues over the costumers of the SaaS. It is very much important and also the responsibility of every user to know the type of information they share with the cloud and who else is authorized to use that information. The users must know the level of protection they are provided with by the service provider.

Considering the provider of the SaaS’s role is very significant in relat ion to the access of the information and the processes of the organization. Advancement like rising of golden eye ransom ware and Xcode ghost highlights that the attackers knows the cloud and software provider’s value and consider them as a vector through which they can fire attack over the larger assets. This is resulting in increase of the focus of these attackers over this type of potential vulnerability. To protect the information, the user must be Technologies [ICICT 2021] alert and scrutinize the security programs of the provider of cloud computation.

Top 10 security issues of applications of cloud which are faced with SaaS are mentioned below:

i. The applications of cloud do not provide a clear visible picture about what data is within it.

ii. Data theft from a cloud application through malicious actor

iii. The control in respect the accessibility of the sensitive data is incomplete

iv. Inability in reference of monitoring the data in transferring from/to cloud applications.

v. Cloud applications being provisioned outside of IT visibility (e.g., shadow IT)

vi. For managing the issues and development of security of the applications of cloud, the available staffs are not sufficient or skilled.

vii. Inability in reference of preventing malicious inside misuse of data or data theft.

viii. High tech fire attacks and threats against providers.

ix. Inability in reference of assessing the operation’s security of the cloud application

x. Inability in reference of maintaining regulatory compliance.

2. Infrastructure-as-a-service(IaaS)

IaaS is a way of providing the user with virtual or physical machines like Hyper-V or virtual bo x which operate virtual machine. Protection is data is not an easy task in IaaS. As the responsibilities of the user increases to OS, network traffic as well as applications, more and more threats sums up. Organizations should not delay in considering the evolutions in attacks that has extended beyond the data which is the center of the risk associated with the IaaS. Lately, many malicious actors has conducted computing resources’ hostile takeover for mining crypto currency. These resources are then further used as an virtual weapon to attack vector against other elements of the infrastructure of the enterprise and also against the third party.

When an infrastructure is built in the cloud, assessing your abilities is important in order for preventing the data theft and accessing of control. Hardening and securing orchestration tool, tracking the modification of the resources for identifying abnormal behaviors, addition of network analysing of both east – west and north to south traffic as a potential signal and to determine who is permitted to enter data into it are the ways which enhancing as standard measures to protect the infrastructure of cloud deployments at scale.

“Below are the Top 10 cloud security issues experienced with infrastructure-as-a-service (IaaS)

i. Cloud workloads and accounts being created outside of IT visibility (e.g., shadow IT)

ii. Data Incomplete control over who can access sensitive

iii. Data theft hosted in cloud infrastructure by malicious actor

iv. Lack of staff with the skills to secure cloud infrastructure

v. Lack of visibility into what data is in the cloud

vi. Inability to prevent malicious insider theft or misuse of information

vii. Lack of consistent security controls over multi-cloud and on-premises environments

viii. Advanced threats and attacks against cloud infrastructure

ix. Inability to monitor cloud workload systems and applications for vulnerabilities

x. Lateral spread of an attack from one cloud workload to another”

3. Platform-as-a-service (PaaS)

The provider of this model avails the user with features for accessing the OS, web servers and execution environment of programming language. This model acts as bridge between IaaS and SaaS.

As per the NIST, the model of the cloud comprises of 4 core deployment models such as hybrid cloud, private cloud, community cloud and public cloud.

Fine-tuned control which is available with the environment of the private cloud is considered as important factor for the process of decision making for allocating resources to private vs public cloud. Additional level available for controlling and supplemental protection in private clouds can compensate for other foundation and it might make contributions to a practical transition. With all these factors, the organizations should keep in consideration that the maintenance of fine-tuned control creates difficulty. Presently, much of the efforts are taken by the service providers upon themselves. Simplification of the management of security can be made by the users which can decrease the difficulty by abstracting the controls. This amalgamates private and public cloud platforms across and above hybrid, virtual as well as physical environment.

Below are the Top 5 cloud security issues experienced with private cloud:

i. Consistent spanning of control in relation to the securit is lacking in the virtualized and traditional server private cloud infrastructure.

ii. Hike in the infrastructure’s complexity results in more effort/time of maintenance and implementation.

iii. Skilled staff is available as per the requirement for managing the software defined data centre’s security.

iv. Visibility is not complete over the software defined data centre’s security.

v. Newly developed advance level attacks and threats.

III. SECURITY THREATS IN CLOUD COMPUTING

A force which act from outside through which the nodes which existed in one state gets transferred to another is Technologies [ICICT 2021] termed as a threat. The data is stored in the node and this node provides the user with a platform for using the application in services form. Significant numbers of intrusions or attacks are available occurring within the applications of the cloud.

The 3 service models of the cloud provides various services to the user and also discloses data‘s issue of security as well as risks which are availab le within the systems of cloud.

1. SQL Injection Attack.

This is a virtual attack made to a computer and it mostly damages the SaaS. This attack damages SaaS the most because of the poor design of application. It also completes the execution of the commands of SQL (unauthorized) through taking benefits of insecure interface. These types are attacks are programmed for accessing unauthorized data which is under protection and not allowed to access publically.

2. Abuse And Nefarious Use of Cloud Computing

The hackers gain advantage of shortcomings in the process of authentic registrations of cloud. Further, they are provided with services of SaaS, PaaS, IaaS. It is possible for hackers to make their move with suspectible activities like Phishing and/or spamming. These threats are available in all the 3 layers.

3. Net Sniffers

It is also threat associated with SaaS. Through this type of threat, the hacker gains the access via applications. This enables them in capturing packets which flows within a network and also the data if they are transited through the captured packets unencrypted. If this happens, the data become available to everyone.

4. Session Hijacking

Over a protected network, it is an attack on the security of a user session. When a website is logged in by a user, a new session starts in that server. The new session comprises of all the data and the information of the user which the server uses so that password won’t be needed every time the user enters a new page. With all the needed knowledge, the hackers can enter a running sess ion and succeeds in gaining access of that session identifier via HTTP. Session identifier is used by the server in order to identify the user for that particular session. This session hijacking is used by the hacker for gaining the control over the session identifier which further enables them in gaining unauthorized control over the user’s information. Cross site scripting, session fixat ion, session side-jacking and session prediction are the most commonly known session hijacking attacks.

5. Man In The Middle Attack

MITM attack is another kind of session hijacking in which a sniffer is used by the hackers to hack the communication among the devices through which data collection is done and hacker further transmits the data. An independent connection is es tablished by the hackers with the user’s device and the user is convinced that the connection is direct and private. But in reality, the hackers control the session completely. It is a big threat to the SaaS model.

6. Denial of Services

resources unavailable for the user virtually. This interruption may be both temporary and permanent.

7. Flooding Attacks

This is a type of “denial of service attack” which is used for increasing the conjunction of the network through flooding the network with various types of traffic in a large amount. This type of attack happens when the hackers overweight the services or the network with packets containing data. The server is attacked by it with connections which will remain incomplete and as a result it end with filling the buffer memory of the host with redundant and unused data. At the end when buffer is left with no space, the server won’t be able to make any type of connections. This will result as the “denial of service”. This attack occurs in IaaS and PaaS layers of the model of cloud.

8. Privacy Breach Organization as well as users stores their data and information in the cloud.

Therefore, any type of breach in the cloud will hack the informat ion available of those users who are authorized. This will enable the unauthorized users in accessing the private information of the users which further might lead to unauthorized and unethical activities with the stored information. This will mostly affect the users of the SaaS model.

IV. COUNTER MEASURES

The infrastructure of the cloud computing comprises of a provider of the services which is responsible for providing resources for computing for the end user. For assuring the best possible services, it is important for the service providers to ensure the users regarding the security safety of the cloud. Through applying methods of advanced security as well as defining stringent security policies, this may be done.

1. DevSecOps processes — DevSecOps and DevOps are continuously been observed in order to decrease the options of vulnerability and explo itations, enhance the quality of the codes, deployment of features and hiking the application’s speed. Including security procedures, advancement and QA in the units of the business/applications team rather than depending upon a single security verification team is important for the operations as per the demands of the today’s businesses.

2. Automated application deployment and management tools —Hike in the speed and amount of security threats in combination with the insufficient skills in relations to the security leads to the fact that even the professional with highest experience of security cannot keep up. With the help of automation, ordinary tasks can be removed and it also supplement the human work benefits with that of machines which a basic element of advanced operations of IT.

3. Unified security with centralized management across all services and providers — It is not possible for a single vendor, service or product to deliver all the things but this can be delivered through multip le management tool which reduces the difficulty so that something can slip by. In combination with an open integration fabric, the system of Technologies [ICICT 2021]

This is an attack in the layer of SaaS and through this attack; the hackers make the services and the network unified management decreases the difficulty to great extent by combining the parts and restructuring the flow of the work.

**Arquitetura:**

CLOUD COMPUTING ARCHITECTURE

A. Describing the NIST Model According to NIST, five major actors have been identified in cloud computing [7]. The NIST Conceptual Reference Model diagram in [7] shows the actors which are discussed below.

a. Cloud Consumer: A person or organization that starts and keeps a business association with and requires services from suppliers of cloud services [22].

b. Cloud Provider: A person, organization engaged in supplying cloud computing services to interested persons or organizations.

c. Cloud Auditor: An organisation in charge of conducting independent evaluation of cloud computing, and determining the systems effectiveness and security.

d. Cloud Broker: A third-party organisation or individual that serves as an intermediary between cloud consumers and cloud providers . He/she is useful for negotiating terms and conditions of the contract for the purchase of cloud services.

e. Cloud Carrier: An intermediary person, organisation or entity that provides connectivity and transport of cloud services from cloud provider to cloud consumers. The diagram showing interactions between the actors in cloud computing in [7] describes the relationship amongst actors in cloud computing. Highlighting the relevance of each party in actualising the delivery, use and maintenance of cloud services. In addition, NIST also provides five characteristics of the cloud [11]

a. On-demand self–service:

Computing capabilities are allocated to users based on their predefined requirements. Capabilities such as server processing time and storage spaces are allocated automatically without human interference.

b. Broad Network Access: Thick or thin clients (ranging from tablets, laptops, workstations to mobile phones) enjoy access to the cloud by using standard mechanisms.

c. Resource pooling: Multi-tenant models pool together resources in order to provide services to multiple consumers. These services can be customised to satisfy the peculiarity of each consumers’ requirements.

d. Rapid Elasticity: This describes the cloud providers’ ability to rapidly deploy scalable resources at the request of consumers. The cloud architecture must be able to seamlessly scale up and down amongst each requests of users, ensuring that the capabilities appear unlimited to the consumer.

e. Measured Service: The cloud provider monitors and controls certain aspects of the cloud service to ensure effective resource usage and overall predictive planning. This iachieved by the use of a metering capability embedded within the system.

B. Cloud Computing Services

A cloud architecture can be divided into the back end and front end . The front end is made visible to the user through connections to the Internet, allowing user interactions with the system [11]. The back end comprises the various cloud services models.

1) Software-as-a-Service (SaaS)

The user is offered a hosted set of software running on a platform and infrastructure owned by the cloud provider [2]. Applications are designed and developed to be simultaneously accessed by various cloud consumers over the Internet [11] . The hosted application is managed by the CSP, who maintains and ensures up-to-date running of the system. The hosted application supports multitenancy, it is available on demand and can be scaled up on down [11].

Some SaaS providers run on other cloud provider’s PaaS or IaaS offerings [2]. Examples of SaaS [21]:

a. Email and Office Productivity: Email applications, word editors and processors, spreadsheets applications, presentations applications are typical examples in this category.

b. Billing: There are applications designed to monitor and manage customer billing. This is determined by users’ system usage and subscriptions to products and services.

c. Customer Relationship Management (CRM): CRM are typical call-centre applications.

d. Financials: These are applications useful for tracking and reporting financial activities including processing of expenditure, generating invoices, payroll, and managing taxes.

2) Platform–as-a-Service (PaaS) PaaS is a development service offered to the user through the Internet [2] . The user does not require any software installation or hardware requirements, thereby saving cost. It is a middleware upon which applications are built [11] . PaaS has built–in tools, built–in-security and web service interfaces for the deployed applications [11]. The deployed application can be integrated with other applications on the same platform and interfaced with other applications outside the platform [11]. PaaS has software comprising a database, middleware and development tools [2]. Examples of PaaS [21]:

a. Business Intelligence.

b. Database.

c. Development and Testing.

d. Integration.

e. Application Deployment.

3) Infrastructure–as-a–Service.

This is delivery of servers, storage, network and operating system, as a service [2]. IaaS provides an abstract machine with operating system already installed and configured [11]. IaaS enables data to be stored in different geographical locations . IaaS providers control activities in the cloud data centres while allowing users the flexibility to deploy and manage software services themselves [2]. The user has access to a virtual computer, storage, network infrastructure, computing resources for deploying and running software [6]. The cloud provider only manages the software and hardware, such as servers, storage devices, host OS and hypervisor for virtualization [5]. A typical cloud architecture diagram services available to cloud users is shown in [5]. Examples of IaaS [21]:

a. Content Delivery Networks (CDNs): CDNs record user content and files to improve the system performance such as speed and the cost associated with the delivery content for web-based systems. This is useful for handling diverse kinds of content for delivery to any website or mobile app.

b. Backup and Recovery: This provides ability for seamless backup and restoration of files.

c. Compute: This involves server requirements for maintaining cloud systems that can be configured and provisioned dynamically.

d. Storage: Highly scalable storage ability useful for recording activities of applications, file backups and recovery and storing files are also available.

C. Cloud Computing Deployment Types

a. Private Cloud. A private cloud is exclusively setup and run for a particular enterprise, but third party organizations are given access to manage them on the behalf of the cloud owner [11]. The private cloud can be operated on-premise or off-premise. Private cloud has privacy, security and control. The cost and energy efficiency is also good [11]. Private clouds have limited scalability and are restricted to an area. Public Cloud. They are operated by a CSP, who owns the infrastructure and data centres. The infrastructure is on premise and enterprises can access services on-demand and pay-as-you-go basis [11]. Services are made available to an organizations and users over a public network through a browser [3] [6].

b. Public clouds are location independent, reliable and highly scalable, but less secure and not customizable [11].

c. Community Cloud. Community cloud is hosted by several organizations or institutions sharing common interest. Typical examples are universities using it for learning and research. Organisations may decide to manage the cloud system themselves on site or off site and may also decide to outsource the daily running of the system to a third party organization [11].

d. Hybrid Cloud. Hybrid Cloud represents a combination of either a selection or all cloud deployment types i.e. private, public or community cloud. Core activities are hosted on a private cloud, while less essential services are outsourced to a public cloud. Each of the cloud remains a unique entity, but linked together by standardized technology [11]. Hybrid clouds are subjected to network and security issues.

D. Inter Cloud Inter cloud can be simply referred to as an interconnection of clouds, like the network of networks [11]. This allows connection of multiple cloud infrastructure to make more computing resources available. There are two types of inter-cloud: the federated cloud and Multi cloud.

a. Federation Cloud: A federation cloud is an inter-cloud setup where a set of cloud providers wilfully combine their cloud infrastructure, this enables easy sharing of resources amongst the participating organizations. Every cloud provider voluntarily collaborates to exchange resources.

b. Multi-Cloud: Multi-cloud involves the use multiple independent cloud architecture acting as a single architecture. For example, organisations or enterprises could run a particular cloud activity that requires enormous resources on a private cloud and run other cloud activities requiring lower level of network resources or storage capability on a public cloud.

**Service Model**

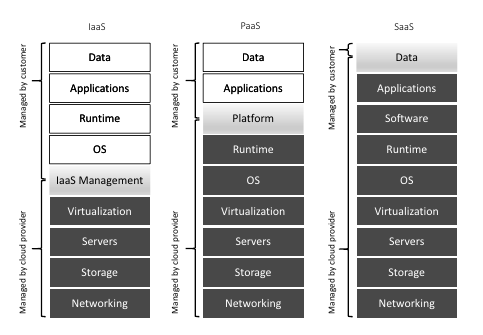
Cloud computing can be segregated into the following service levels: Infrastructure as a Service, Platform as a Service, and Software as a Service as shown in Fig. 2. Fig. 3 presents he separation between service models with control of cloud service provider and customer of different underlying concepts in each model.

1) Infrastructure as a Service (IaaS) : Customers will get the services for a complete computing infrastructure over the Internet. Example: Amazon EC2 [10] and S3 [20].

2) Platform as a Service (PaaS) : In PaaS, customers will get the platform for the development of software applications. Example: Microsoft Azure [8] and Google AppEngine [9].

3) Software as a Service (SaaS) : Customers will be provided with the Software over the Internet. In this model, users will not get the software; instead, they get the web-based software from the service providers to the intended work. Example: Dropbox [21] and Office365[22].

4) Others-as-a-Service: Backup-as-a-service, storage-as-a-service, logistics-as-a-service, operating systems as a service(OSaaS), framework as a service(FaaS), database as a service(DaaS), network as a service(NaaS) and other models apart from the 3 primary ones were discussed in [19], [9], [23], [29] and [33]. However, only 13% of the papers examined discussed Others-as-a- Service.

 **Deployment Model**

Cloud computing services are provided over the following deployment models [2] as shown in Fig. 2:

1) Public cloud : provides the cloud services for end-users by allowing them to access the services from the Internet. So, these cloud services are publicly accessible. The Cloud service provider provides the required infrastructure for end-users.

A cloud is called a "public cloud" when the services are rendered over a network that is open for public use. Public cloud services may be free.[103] Technically there may be little or no difference between public and private cloud architecture, however, security consideration may be substantially different for services (applications, storage, and other resources) that are made available by a service provider for a public audience and when communication is effected over a non-trusted network. Generally, public cloud service providers like Amazon Web Services (AWS), Microsoft and Google own and operate the infrastructure at their data center and access is generally via the Internet. AWS and Microsoft also offer direct connect services called "AWS Direct Connect" and "Azure ExpressRoute" respectively, such connections require customers to purchase or lease a private connection to a peering point offered by the cloud provider.

2) Private cloud : model is used within the organizations to meet the cloud requirements across various levels in the organizations. They will maintain their infrastructure to set up cloud services. This model will reduce the cost of using the cloud services for the organization in the long run with extra security as these private cloud models are deployed in their private networks behind their firewalls. Private cloud model can be accessed inside their network or by authenticating the user at their firewall.

Private cloud is cloud infrastructure operated solely for a single organization, whether managed internally or by a third-party, and hosted either internally or externally. Undertaking a private cloud project requires significant engagement to virtualize the business environment, and requires the organization to reevaluate decisions about existing resources. It can improve business, but every step in the project raises security issues that must be addressed to prevent serious vulnerabilities. Self-run data centers are generally capital intensive. They have a significant physical footprint, requiring allocations of space, hardware, and environmental controls. These assets have to be refreshed periodically, resulting in additional capital expenditures. They have attracted criticism because users "still have to buy, build, and manage them" and thus do not benefit from less hands-on management, essentially "[lacking] the economic model that makes cloud computing such an intriguing concept".

3) Community cloud : model is used to deploy the cloud infrastructure that can be shared between the several organizations with similar goals. It is comparable to a private cloud but shared among some organizations.

4) Hybrid cloud deployment model : model is used where the customers use more than one model to meet the goals of their organizations or end-users them self.

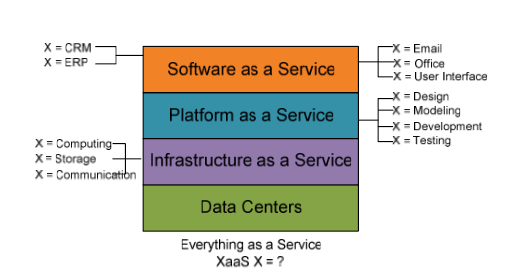
Hybrid cloud is a composition of two or more clouds (private, community or public) that remain distinct entities but are bound together, offering the benefits of multiple deployment models. Hybrid cloud can also mean the ability to connect collocation, managed and/or dedicated services with cloud resources. Gartner, Inc. defines a hybrid cloud service as a cloud computing service that is composed of some combination of private, public and community cloud services, from different service providers. A hybrid cloud service crosses isolation and provider boundaries so that it can't be simply put in one category of private, public, or community cloud service. It allows one to extend either the capacity or the capability of a cloud service, by aggregation, integration or customization with another cloud service.

Varied use cases for hybrid cloud composition exist. For example, an organization may store sensitive client data in house on a private cloud application, but interconnect that application to a business intelligence application provided on a public cloud as a software service. This example of hybrid cloud extends the capabilities of the enterprise to deliver a specific business service through the addition of externally available public cloud services. Hybrid cloud adoption depends on a number of factors such as data security and compliance requirements, level of control needed over data, and the applications an organization uses.

Another example of hybrid cloud is one where IT organizations use public cloud computing resources to meet temporary capacity needs that cannot be met by the private cloud. This capability enables hybrid clouds to employ cloud bursting for scaling across clouds. Cloud bursting is an application deployment model in which an application runs in a private cloud or data center and "bursts" to a public cloud when the demand for computing capacity increases. A primary advantage of cloud bursting and a hybrid cloud model is that an organization pays for extra compute resources only when they are needed. Cloud bursting enables data centers to create an in-house IT infrastructure that supports average workloads, and use cloud resources from public or private clouds, during spikes in processing demands. The specialized model of hybrid cloud, which is built atop heterogeneous hardware, is called "Cross-platform Hybrid Cloud". A cross-platform hybrid cloud is usually powered by different CPU architectures, for example, x86-64 and ARM, underneath. Users can transparently deploy and scale applications without knowledge of the cloud's hardware diversity. This kind of cloud emerges from the raise of ARM-based system-on-chip for server-class computing.

# Deployment models of Cloud Computing

2.1. A Hierarchical View of Cloud Computing Most of the current clouds are built on top of modern data centers. It incorporates Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS), and provides these services like utilities, so the end users are billed by how much they used. Figure 1 shows a hierarchical view for cloud computing.

Figure 1: Hierarchical View of Cloud Computing

Data Centers: This is the foundation of cloud computing which provides the hardware the clouds run on. Data centers are usually built in less populated areas with cheaper energy rate and lower probability of natural disasters. Modern data centers usually consist of thousands of inter-connected servers.

Infrastructure as a Service: Built on top of data centers layer, IaaS layer virtualizes computing power, storage and network connectivity of the data centers, and offers it as provisioned services to consumers. Users can scale up and down these computing resources on demand dynamically. Typically, multiple tenants coexist on the same infrastructure resources [1]. Examples of this layer include Amazon EC2, Microsoft Azure

Platform.

Platform as a Service: PaaS, often referred as cloudware,

provides a development platform with a set of services to assist

application design, development, testing, deployment,

monitoring, hosting on the cloud. It usually requires no

software download or installation, and supports geographically

distributed teams to work on projects collaboratively. Google

App Engine, Microsoft Azure, Amazon Map Reduce/Simple

Storage Service are among examples of this layer.

Software as a Service: In SaaS, Software is presented to the

end users as services on demand, usually in a browser. It saves

the users from the troubles of software deployment and

maintenance. The software is often shared by multiple tenants,

automatically updated from the clouds, and no additional license

needs to be purchased. Features can be requested on demand,

and are rolled out more frequently. Because of its service

characteristics, SaaS can often be easily integrated with other

mashup applications. An example of SaaS is Google Maps, and

its mashups across from the internet. Other examples include

Salesforce.com and Zoho productivity and collaboration suite.

The dividing lines for the four layers are not distinctive.

Components and features of one layer can also be considered to

be in another layer. For example, data storage service can be

considered to be either in as IaaS or PaaS. Figure 1 suggests a

hierarchical relationship among the different layers; however, it

does not mean the upper layer has to be built on top its

immediate lower layer. For example, a SaaS application can be

built directly over IaaS, instead of PaaS.

In the cloud computing environment, everything can be

implemented and treated as a service. Figure 1 shows a few

examples of what can be treated as a service in different layers.

3. Existing Cloud Computing Architectures

Both academia and industry have been active on cloud

computing research, and several cloud computing architectures

have been proposed. In [5], IBM considers current single-

providers cloud as limited resource, and the lack of

interoperability among cloud providers prevents deployment

across different clouds. A cloud computing architecture named

Reservoir was proposed to create a federation from multiple

cloud providers which acts as a global fabric of resources that

can guarantee the required SLA. In Reservoir architecture, the

computational resources within a site are partitioned by a

virtualization layer into virtual execution environments (VEEs).

A service application is decomposed into a set of software

components/services running on VEEs on the same or different

VEEs within a site or across from different sites. However,

Reservoir architecture does not allow a component/service to

run on its duplicates on different VEEs; Moreover, computing

resources are abstracted as hosting service which might not be

necessarily true for all clouds. In [6], a software platform for

.NET based cloud computing named Aneka was introduced.

Aneka is a customizable and extensible service oriented runtime

environment that enables developers to build .NET applications

with the supports of APIs and multiple programming models.

Aneka is a service-oriented, pure PaaS cloud solution. In [7],

Rajkumar and his colleagues explained a market-oriented cloud

architecture in detail used by Aneka, which regulates the supply

and demand of cloud resources to achieve market equilibrium,

adds economic incentives for both cloud consumers and

providers, and promotes QoS-based resource allocation

mechanisms that differentiates service request based on their

utility. The key component of this architecture is SLA (Service

Level Agreement) Resource Allocator which is consisted of

Service Request Examiner and Access Control, VM (Virtual

Machines) monitor, Service Request Monitor, and Request

Dispatcher. Based on the feedback from VM and Service

Request monitors, the dispatcher routes the requests from

users/brokers to the cloud resources that can fulfill their QoS

requirements. In [8], Huang and her colleagues from IBM

described a service oriented cloud computing platform that

enables web-delivery of application-based services with a set of

common business and operational services. The platform

supports multi-tenancy feature by utilizing single application

instance model. The isolation among tenants is taken care by the

underline design. Other services include subscription

management, federated ID management, application firewall,

etc.

3.1. Issues with Current Clouds

Current cloud computing has following characteristics:

Users are often tied with one cloud provider: Even though

up-front cost for a cloud computing deployment is reduced and

long term lease is eliminated, much effort and money is spent on

developing the application for a specific cloud platform which

makes it difficult to migrate the same application onto a

different cloud. Often, migration simply may mean

redevelopment. For example, applications deployed on Amazon

EC2 cannot be migrated easily due its particular storage

framework [9].

Computing components are tightly coupled: This can be

clearly explained using an analogy. Suppose one wants a new

computer, this person has the choices of either buying a ready-

to-use computer from a manufacturer (buying) or purchasing the

components separately and building the computer in a DIY style

(building). The advantages of building over buying include

wider selection of components, flexibility to customize, and

cheaper cost [10]. However, as the computing resources over

the internet, current cloud implementations do not allow this

kind of flexibility. If a customer opts to use Amazon S3 storage

service, he is then stuck with other cloud computing services

Amazon provides, such as EC2, Elastic Map Reduce.

Lack of SLA supports: Currently, SLA is an obstacle that

prevents wide adoption for cloud computing. Cloud computing

infrastructure services such as EC2 are not yet able to sign the

SLA needed by companies that want to use cloud computing for

serious business deployment [11]. Moreover, business is

dynamic. Static SLA is not able to adapt to the changes in

business needs as cloud computing promises to.

Lack of Multi-tenancy supports: Multi-tenancy can support

multiple client tenants simultaneously to achieve the goal of

cost effectiveness. Currently, one has three types of multi-

tenancy enablement approaches: virtualization, mediation and

sharing [12]. To achieve the full potential of multi-tenancy,

three issues remain to be solved [12]:

1. Resource sharing: To reduce the hardware, software and

management cost of each tenant.

2. Security isolation: To prevent the potential invalid access,

conflict and interference among tenants.

3. Customization: To support tenant-specific UI, access control,

process, data, etc.

Lack of Flexibility for User Interface: UI is an important part

of the application, and user experience can be a major

evaluation factor for a business application. However,

cloud/SaaS users are limited with UI choices because UI

composition frameworks, such as the one proposed in [13], have

not been integrated with cloud computing.

Advantages of using Cloud Computing

Cloud computing simplied software development, busi-

ness process, and accessing the services over the In-

ternet. The traditional way of accessing services has

changed with cloud computing. Adoption of the cloud

has reduced costs, made an effective business model, and

offers a great scale of exibility for using the services.

Many organizations have adopted cloud services and

beneted by moving their services to the cloud. With

the adoption of cloud, organizations are improving cross-

platform collaboration between the developers, allowing

them to do more innovations on their IT capabilities,

which in turn helps the organizations to grow their

business and get more revenue [1].

Referências

Swati I. Bairagi e Ankur O. Bang, “Cloud Computing: History, Architecture, Security Issues”, Special Issue National Conference “CONVERGENCE 2015”, 28 March 2015, International Journal of Advent Research in Computer and Electronics (IJARCE) - (E-ISSN: 2348-5523)

J. Surbiryala e C. Rong, "Cloud Computing: History and Overview," 2019 IEEE Cloud Summit, Washington, DC, USA, 2019, pp. 1-7, doi: 10.1109/CloudSummit47114.2019.00007.

W. -T. Tsai, X. Sun and J. Balasooriya, "Service-Oriented Cloud Computing Architecture," 2010 Seventh International Conference on Information Technology: New Generations, Las Vegas, NV, USA, 2010, pp. 684-689, doi: 10.1109/ITNG.2010.214.

C. Gong, J. Liu, Q. Zhang, H. Chen and Z. Gong, "The Characteristics of Cloud Computing," 2010 39th International Conference on Parallel Processing Workshops, San Diego, CA, USA, 2010, pp. 275-279, doi: 10.1109/ICPPW.2010.45.

M. K. Sasubilli and V. R, "Cloud Computing Security Challenges, Threats and Vulnerabilities," 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 2021, pp. 476-480, doi: 10.1109/ICICT50816.2021.9358709.

I. Odun-Ayo, M. Ananya, F. Agono and R. Goddy-Worlu, "Cloud Computing Architecture: A Critical Analysis," 2018 18th International Conference on Computational Science and Applications (ICCSA), Melbourne, VIC, Australia, 2018, pp. 1-7, doi: 10.1109/ICCSA.2018.8439638.

Deployment models of Cloud Computing [<https://forum.huawei.com/enterprise/en/deployment-models-of-cloud-computing/thread/1010535-893>]

A Brief History of Cloud Computing - IBM [<https://www.ibm.com/cloud/blog/cloud-computing-history>]

# Britannica - Project MAC [<https://www.britannica.com/topic/Project-Mac>]

Dataversity - A Brief History of Cloud Computing [<https://www.dataversity.net/brief-history-cloud-computing/>]