KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed University Established Under Section 3 of UGC Act 1956)
Coimbatore - 641021.
(For the candidates admitted from 2016 onwards)
Department of Computer Science, Applications & Information Technology

SUBJECT : CLOUD COMPUTING
SEMESTER : I **L T P C**
SUBJECT CODE: 18CSP104 CLASS : I M.Sc .CS **4 0 0 4**

COURSE OBJECTIVE

- To learn about the basic things involved in cloud computing and its architecture.
- To know about the services such as IaaS, PaaS, SaaS, IDaaS and CaaS.
- To understand the Virtualization Technologies.
- To understand the Information Security, Privacy and Compliance Risks.
- To learn commercial Google Web services – Open Nebula.

COURSE OUTCOME

On successful completion of the course the student should be able to:

- Understand cloud architecture and model.
- Explore cloud infrastructure.
- Learn Threat issues and Database Integrity Issues.
- Learn Open Source and Commercial Clouds such as Microsoft Azure, Amazon EC2.

UNIT-I

Introduction to Cloud Computing -Characteristics of Cloud Computing -Paradigm shift - Benefits of cloud computing - Disadvantages of cloud computing- Role of Open Standards- Cloud Computing Architecture: Cloud computing stack-Public cloud -Private cloud -Hybrid cloud -Community cloud

UNIT –II

Infrastructure as a Service (IaaS) -Platform as a Service (PaaS) -Software as a Service (SaaS) - Identity as a Service (IDaaS) -Compliance as a Service (CaaS)- Cloud storage.

UNIT -III

Virtualization Technologies -Load Balancing and Virtualization -Advanced load balancing -The Google cloud - Hypervisors -Virtual machine types -VMware vSphere - Machine Imaging - Porting Applications -The Simple Cloud API - AppZero Virtual Application Appliance

UNIT-IV (3rd book)

Cloud Information Security Objectives -Confidentiality Integrity and Availability -Cloud Security Services - Relevant Cloud Security Design Principles -Cloud Computing Risk Issues - The CIA Triad

Privacy and Compliance Risks -Threats to Infrastructure Data and Access Control -Cloud Access Control Issues -Database Integrity Issues -Cloud Service Provider Risks- Architectural Considerations
General Issues- Trusted Cloud Computing -Identity Management and Access Control

UNIT -V

Case Study on Open Source and Commercial Clouds: Microsoft Azure- Amazon EC2-Google Web services – Open Nebula.

SUGGESTED READINGS

1. Barrie Sosinsky .(2010). Cloud Computing Bible .New Delhi: Wiley- India
2. Rajkumar Buyya, James Broberg, & Andrzej, M. Goscinski. (2011). New Delhi: Tata Mc-Graw Hill.
3. Ronald, L. Krutz, Russell Dean Vines. (2010). Cloud Security: A Comprehensive Guide to Secure Cloud Computing. New Delhi: Wiley –India
4. Dr Kumar Saurabh.(2012). Cloud Computing (2nd ed.). New Delhi: Wiley India.
5. Anthony T.Velte Toby J.Velte Robert Elsenpeter. (2010). Cloud Computing Practical Approach (1st ed.). New Delhi:Tata McGraw Hill.
6. Nikos Antonopoulos, Lee Gillam. (2012). Cloud Computing: Principles Systems and Applications . Springer.
7. Giovanni Toraldo. (2012). Open Nebula 3 Cloud Computing.

WEB SITES

W1: en.wikipedia.org/wiki/Cloud_Computing

W2: www.ibm.com/cloud-computing/in/en/

W3: www.oracle.com/CloudComputing

W4: www.microsoft.com/en-us/cloud/default.aspx

W5: https://azure.microsoft.com/en-in/case-studies

W6: https://aws.amazon.com/solutions/case-studies

W7: https://cloud.google.com/

W8: https://opennebula.org/about/technology

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

Established Under Section 3 of UGC Act, 1956)

Coimbatore – 641021, INDIA

Department of Computer Science, Applications & Information Technology

Lecture Plan

Subject Name: Cloud Computing **Subject Code: 18CSP104**

Semester: I **Class: I Msc. CS**

Staff: Dr.S.Manju Priya

S.No	Topics	No. of Periods Required	Reference Materials
Unit-I : Introduction to Cloud Computing			
1	Introduction to Cloud Computing	1hr	SR1: pg.1-4
2	Characteristics of Cloud Computing, Paradigm shift	1hr	SR1: pg.13-15
3	Benefits of cloud computing, Disadvantages of cloud computing	1hr	SR1: pg.16-19
4	Role of Open Standards	1hr	SR1: pg.19-22
5	Cloud Computing Architecture: Cloud computing stack	1hr	SR1: pg.45-58
6	Public cloud , Private cloud	1hr	SR3: pg.44-48
7	Hybrid cloud , Community cloud	1hr	SR1: pg.49-50
8	Recapitulation and Discussion of Important Questions	1hr	-
	Total Hours	**8 hrs**	
Unit-II : Cloud Platforms			
1	Infrastructure as a Service (IaaS)	1hr	SR1: pg.66-70
2	Platform as a Service (PaaS)	1hr	SR1: pg.70-71
3	Software as a Service (SaaS)	1hr	SR1: pg.71-76

4	Identity as a Service (IDaaS)	1hr	SR1: pg.76-84
5	Compliance as a Service (CaaS)	1hr	SR1: pg.87-88
6	Cloud storage	1hr	SR3: pg.58
7	Virtual Storage Containers	1hr	SR3: pg.59
8	Recapitulation and Discussion of Important Questions	1hr	-
	Total Hours	**8 hrs**	
Unit - III : Cloud Virutalization			
1	Virtualization Technologies	1hr	SR1: pg.93-94
2	Load Balancing and Virtualization	1hr	SR1: pg.95
3	Advanced load balancing	1hr	SR1: pg.96
4	The Google cloud	1hr	SR1: pg.97-100
5	Hypervisors	1hr	SR1: pg.100
6	Virtual machine types , VMware vSphere	1hr	SR1: pg.100-106
7	Machine Imaging	1hr	SR1: pg.107-108
8	Porting Applications	1hr	SR1: pg.108
9	The Simple Cloud API	1hr	SR1: pg.109
10	AppZero Virtual Application Appliance	1hr	SR1: pg.109-111
11	Recapitulation and Discussion of Important Questions	1hr	-
	Total Hours	**11 hrs**	
UNIT-IV : Cloud Security			
1	Cloud Information Security Objectives	1hr	SR3: pg.62
2	Confidentiality, Integrity and Availability	1hr	SR3: pg.63

3	Cloud Security Services, Relevant Cloud Security Design Principles	1hr	SR3: pg.66-68
4	Cloud Computing Risk Issues, The CIA Triad	1hr	SR3: pg.125-127
5	Privacy and Compliance Risks, Threats to Infrastructure Data and Access Control	1hr	SR3: pg.127-140
6	Cloud Access Control Issues, Database Integrity Issues	1hr	SR3: pg.145-147
7	Cloud Service Provider Risks, Architectural Considerations, General Issues	1hr	SR3: pg.147, 177-178
8	Trusted Cloud Computing, Identity Management and Access Control	1hr	SR3: pg.188, 204
9	Recapitulation and Discussion of Important Questions	1hr	
	Total Hours	**9 hrs**	
	UNIT-V : Case Study		
1	Case Study on Open Source and Commercial Clouds	1hr	W5
2	Microsoft Azure	1hr	W5
3	Example	1hr	W5
4	Amazon EC2	1hr	W6
5	Example	1hr	W6
6	Google Web services	1hr	W7
7	Google applications	1hr	W7
8	Open Nebula	1hr	W8
9	Recapitulation and Discussion of Important Questions	1hr	
10	Discussion of Previous ESE Question Papers	1hr	
11	Discussion of Previous ESE Question Papers	1hr	
12	Discussion of Previous ESE Question Papers	1hr	
	Total Hours	**12 hrs**	
	Total Number of periods (8hr+9hr+11hr+7hr+13hr)	**48 hrs**	

Suggested Readings

SR1: Barrie Sosinsky .(2010). Cloud Computing Bible .New Delhi: Wiley- India

SR2: Rajkumar Buyya, James Broberg, & Andrzej, M. Goscinski. (2011). New Delhi: Tata Mc-Graw Hill.

SR3: Ronald, L. Krutz, Russell Dean Vines. (2010). Cloud Security: A Comprehensive Guide to Secure Cloud Computing. New Delhi: Wiley –India

SR4: Dr Kumar Saurabh.(2012). Cloud Computing (2nd ed.). New Delhi: Wiley India.

SR5: Anthony T.Velte Toby J.Velte Robert Elsenpeter. (2010). Cloud Computing Practical Approach (1st ed.). New Delhi:Tata McGraw Hill.

SR6: Nikos Antonopoulos, Lee Gillam. (2012). Cloud Computing: Principles Systems and Applications. Springer.

SR7: Giovanni Toraldo. (2012). Open Nebula 3 Cloud Computing.

Web References

W1: en.wikipedia.org/wiki/Cloud_Computing

W2: www.ibm.com/cloud-computing/in/en/

W3: www.oracle.com/CloudComputing

W4: www.microsoft.com/en-us/cloud/default.aspx

W5: https://azure.microsoft.com/en-in/case-studies

W6: https://aws.amazon.com/solutions/case-studies

W7: https://cloud.google.com/

W8: https://opennebula.org/about/technology

UNIT-I

Introduction to Cloud Computing -Characteristics of Cloud Computing -Paradigm shift - Benefits of cloud computing - Disadvantages of cloud computing- Role of Open Standards- Cloud Computing Architecture: Cloud computing stack-Public cloud -Private cloud -Hybrid cloud -Community cloud

Introduction to Cloud Computing

Cloud Computing provides us means by which we can access the applications as utilities over the internet. It allows us to create, configure, and customize the business applications online.

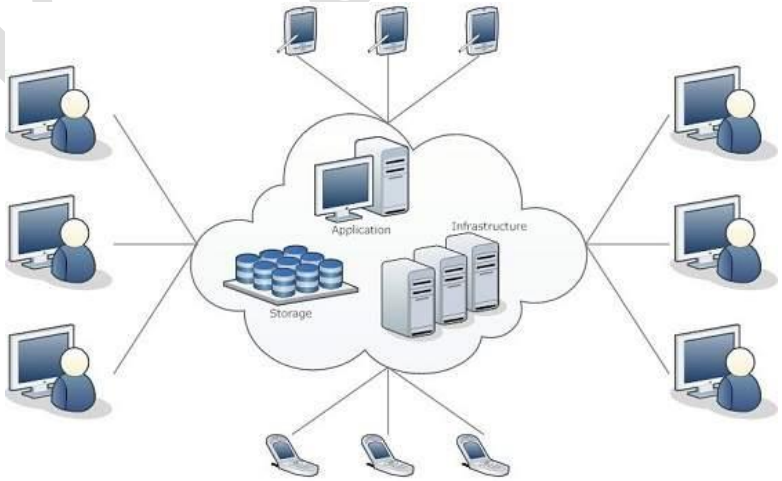
What is Cloud?

The term **Cloud** refers to a **Network** or **Internet.** In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over public and private networks, i.e., WAN, LAN or VPN.

Applications such as e-mail, web conferencing, customer relationship management (CRM) execute on cloud.

What is Cloud Computing?

Cloud Computing refers to **manipulating, configuring,** and **accessing** the hardware and software resources remotely. It offers online data storage, infrastructure, and application.



Cloud computing offers **platform independency,** as the software is not required to be installed locally on the PC. Hence, the Cloud Computing is making our business applications **mobile** and **collaborative.**

Characteristics of Cloud Computing

An IT environment requires a specific set of characteristics to enable the remote provisioning of scalable and measured IT resources in an effective manner. These characteristics need to exist to a meaningful extent for the IT environment to be considered an effective cloud.

The following six specific characteristics are common to the majority of cloud environments:

- On-Demand Usage
- Ubiquitous Access
- Multi-tenancy (Resourcing Pooling)
- Elasticity (and Scalability)
- Measured Usage
- Resiliency

Cloud providers and cloud consumers can assess these characteristics individually and collectively to measure the value offering of a given cloud platform. Although cloud-based services and IT resources will inherit and exhibit individual characteristics to varying extents, usually the greater the degree to which they are supported and utilized, the greater the resulting value proposition.

On-Demand Usage

A cloud consumer can unilaterally access cloud-based IT resources giving the cloud consumer the freedom to self-provision these IT resources. Once configured, usage of the self-provisioned IT resources can be automated, requiring no further human involvement by the cloud consumer or cloud provider. This results in an *on-demand usage* environment. Also known as "on-demand self-service usage," this characteristic enables the service-based and usage-driven features found in mainstream clouds.

Ubiquitous Access

Ubiquitous Access represents the ability for a cloud service to be widely accessible. Establishing ubiquitous access for a cloud service can require support for a range of devices, transport protocols, interfaces, and security technologies. To enable this level of access generally requires that the cloud service architecture be tailored to the particular needs of different cloud service consumers.

Multitenancy (and Resource Pooling)

The characteristic of a software program that enables an instance of the program to serve different consumers (tenants) whereby each is isolated from the other, is referred to as *multitenancy*. A cloud provider pools its IT resources to serve multiple cloud service consumers by using multitenancy models that frequently rely on the use of virtualization technologies. Through the use of multitenancy technology, IT resources can be dynamically assigned and reassigned, according to cloud service consumer demands.

Resource pooling allows cloud providers to pool large-scale IT resources to serve multiple cloud consumers. Different physical and virtual IT resources are dynamically assigned and reassigned according to cloud consumer demand, typically followed by execution through statistical multiplexing. Resource pooling is commonly achieved through multitenancy technology, and therefore encompassed by this multitenancy characteristic.

Figures 1 and 2 illustrate the difference between single-tenant and multitenant environments.

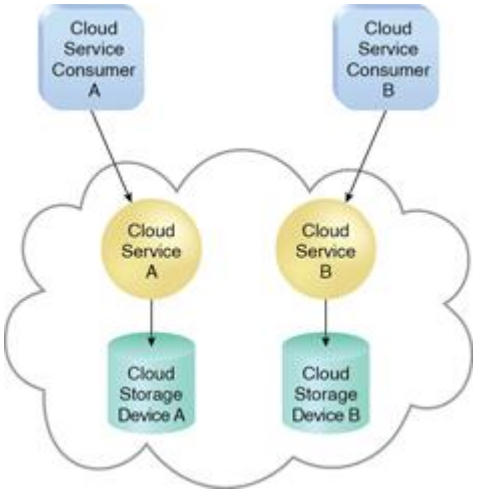


Figure 1 - In a single-tenant environment, each cloud consumer has a separate IT resource instance.

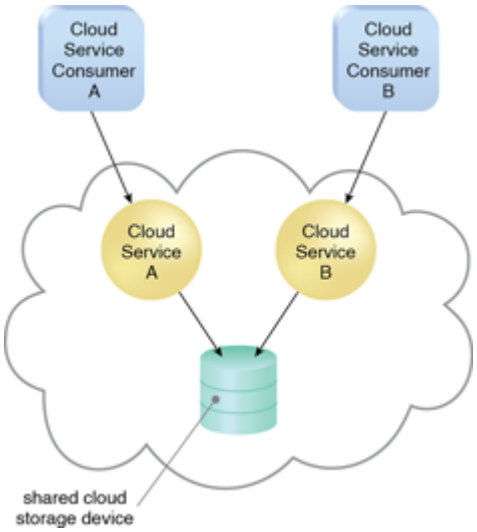


Figure 2 - In a multitenant environment, a single instance of an IT resource, such as a cloud storage device, serves multiple consumers.

As illustrated in Figure 1, multitenancy allows several cloud consumers to use the same IT resource or its instance while each remains unaware that it may be used by others.

Elasticity

Elasticity is the automated ability of a cloud to transparently scale IT resources, as required in response to runtime conditions or as pre-determined by the cloud consumer or cloud provider. Elasticity is often considered a core justification for the adoption of cloud computing, primarily due to the fact that it is closely associated with the Reduced Investment and Proportional Costs benefit. Cloud providers with vast IT resources can offer the greatest range of elasticity.

Measured Usage

The *measured usage* characteristic represents the ability of a cloud platform to keep track of the usage of its IT resources, primarily by cloud consumers. Based on what is measured, the cloud provider can charge a cloud consumer only for the IT resources actually used and/or for the timeframe during which access to the IT resources was granted. In this context, measured usage is closely related to the on-demand characteristic.

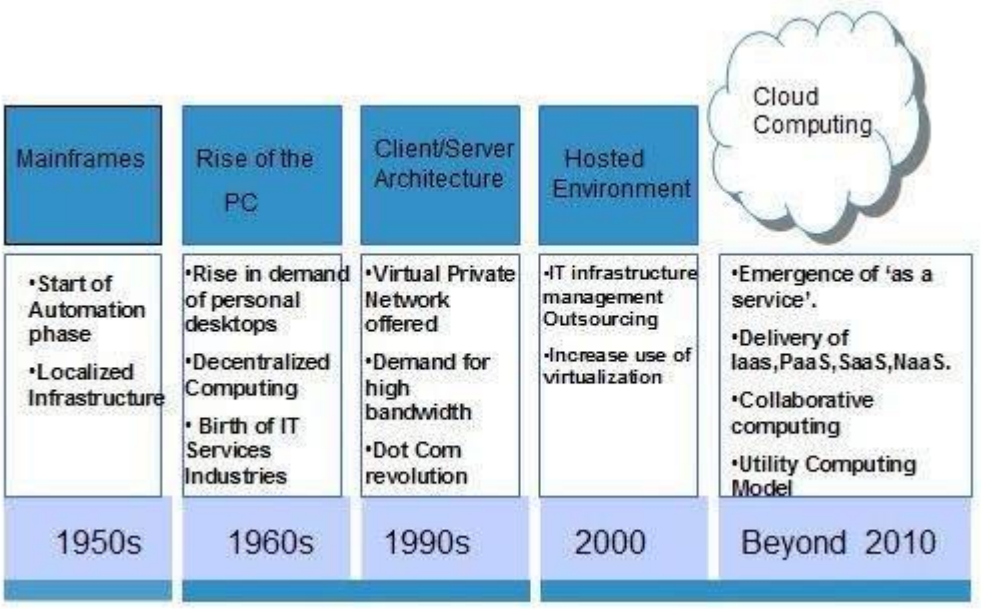
Measured usage is not limited to tracking statistics for billing purposes. It also encompasses the general monitoring of IT resources and related usage reporting (for both cloud provider and cloud consumers).

Resiliency

Resilient computing is a form of failover that distributes redundant implementations of IT resources across physical locations. IT resources can be pre-configured so that if one becomes deficient, processing is automatically handed over to another redundant implementation. Within cloud computing, the characteristic of resiliency can refer to redundant IT resources within the same cloud (but in different physical locations) or across multiple clouds. Cloud consumers can increase both the reliability and availability of their applications by leveraging the resiliency of cloud-based IT resources

Paradigm shift

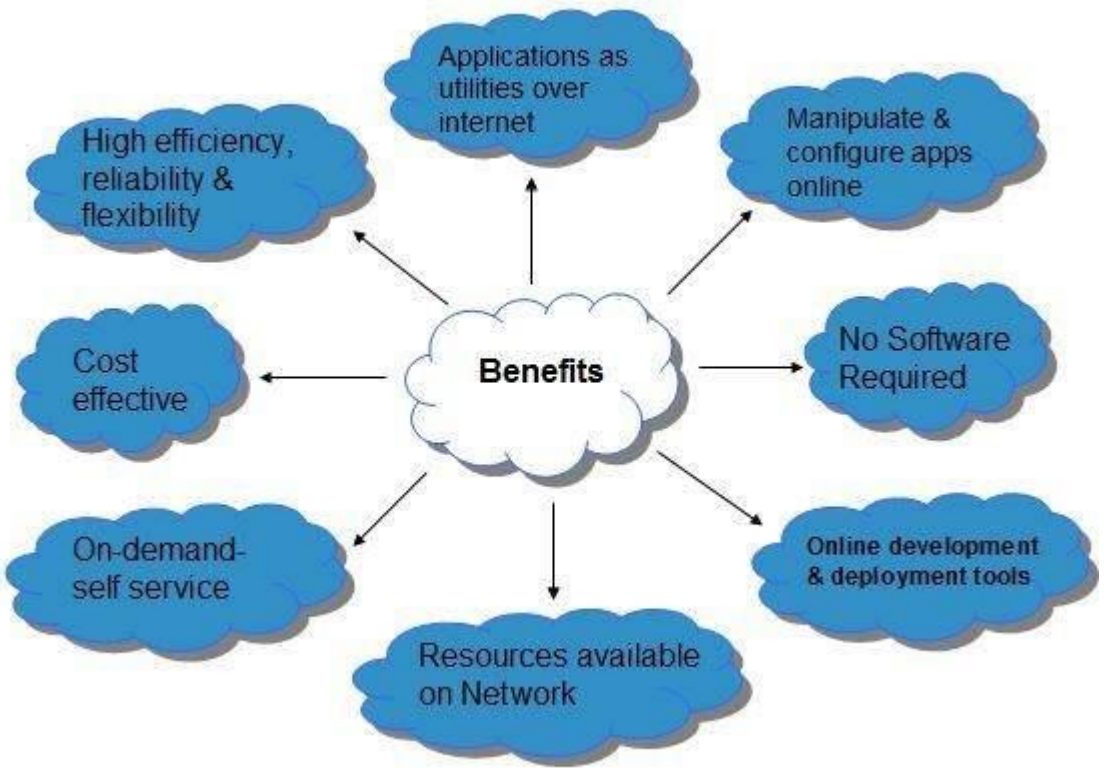
The concept of **Cloud Computing** came into existence in the year 1950 with implementation of mainframe computers, accessible via **thin/static clients.**Since then, cloud computing has been evolved from static clients to dynamic ones and from software to services. The following diagram explains the evolution of cloud computing:



Benefits of cloud computing

Cloud Computing has numerous advantages. Some of them are listed below -

- One can access applications as utilities, over the Internet.
- One can manipulate and configure the applications online at any time.
- It does not require to install a software to access or manipulate cloud application.
- Cloud Computing offers online development and deployment tools, programming runtime environment through **PaaS model.**
- Cloud resources are available over the network in a manner that provide platform independent access to any type of clients.
- Cloud Computing offers **on-demand self-service.** The resources can be used without interaction with cloud service provider.
- Cloud Computing is highly cost effective because it operates at high efficiency with optimum utilization. It just requires an Internet connection
- Cloud Computing offers load balancing that makes it more reliable.



Disadvantages of cloud computing

Although cloud Computing is a promising innovation with various benefits in the world of computing, it comes with risks. Some of them are discussed below:

Security and Privacy

It is the biggest concern about cloud computing. Since data management and infrastructure management in cloud is provided by third-party, it is always a risk to handover the sensitive information to cloud service providers.

Although the cloud computing vendors ensure highly secured password protected accounts, any sign of security breach may result in loss of customers and businesses.

Lock In

It is very difficult for the customers to switch from one **Cloud Service Provider (CSP)** to another. It results in dependency on a particular CSP for service.

Isolation Failure

This risk involves the failure of isolation mechanism that separates storage, memory, and routing between the different tenants.

Management Interface Compromise

In case of public cloud provider, the customer management interfaces are accessible through the Internet.

Insecure or Incomplete Data Deletion

It is possible that the data requested for deletion may not get deleted. It happens because either of the following reasons

- Extra copies of data are stored but are not available at the time of deletion
- Disk that stores data of multiple tenants is destroyed.

Role of Open Standards

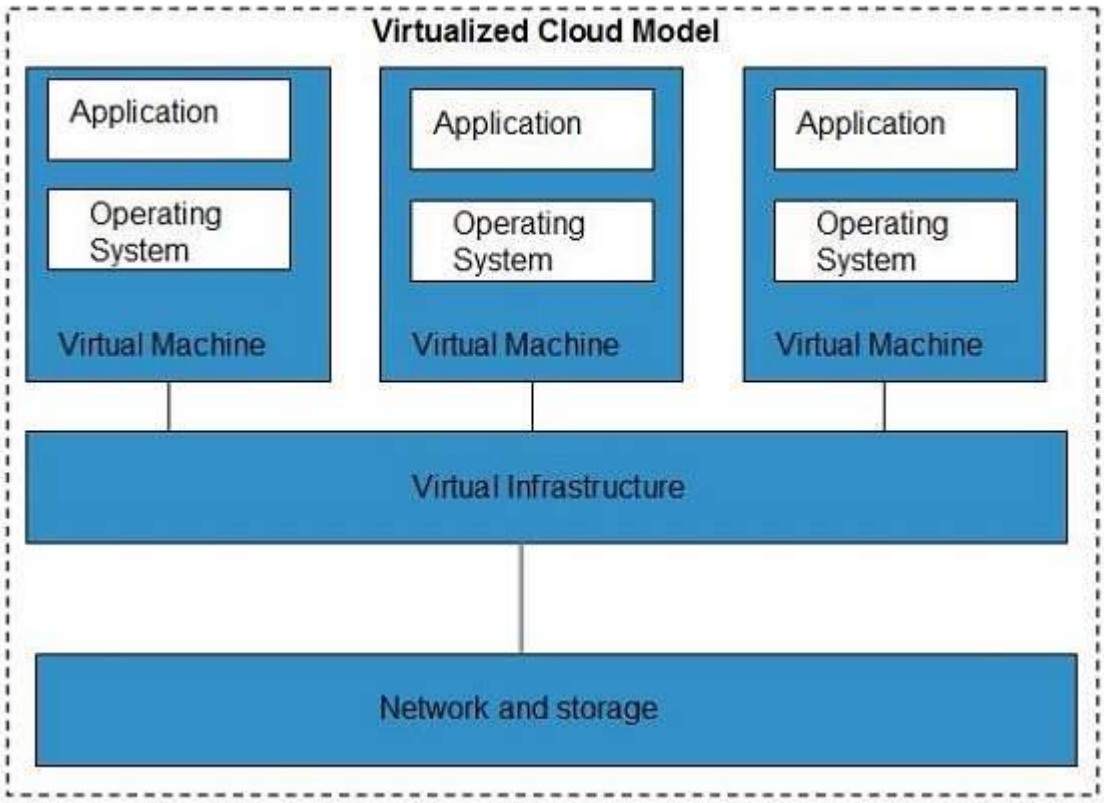
There are certain technologies working behind the cloud computing platforms making cloud computing flexible, reliable, and usable. These technologies are listed below:

- Virtualization
- Service-Oriented Architecture (SOA)

- Grid Computing
- Utility Computing

Virtualization

Virtualization is a technique, which allows to share single physical instance of an application or resource among multiple organizations or tenants (customers). It does this by assigning a logical name to a physical resource and providing a pointer to that physical resource when demanded.

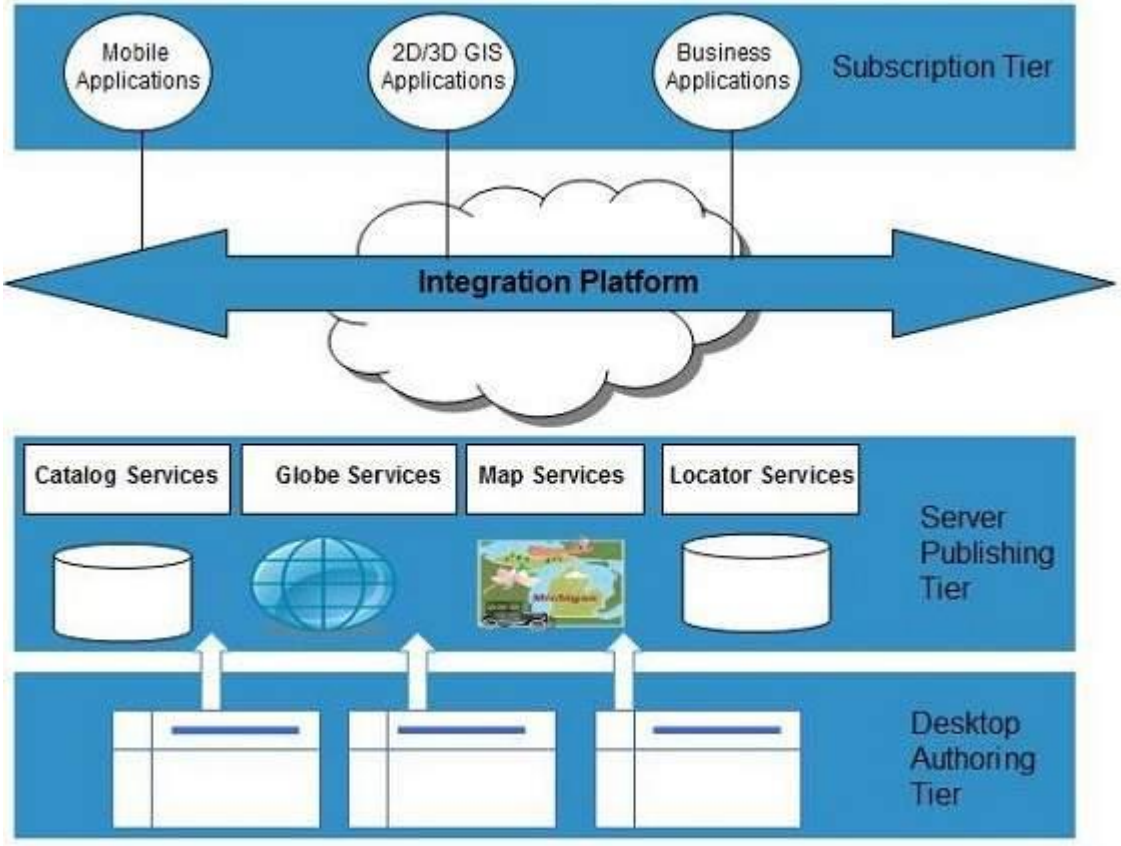


The **Multitenant** architecture offers **virtual isolation** among the multiple tenants. Hence, the organizations can use and customize their application as though they each have their instances running.

Service-Oriented Architecture (SOA)

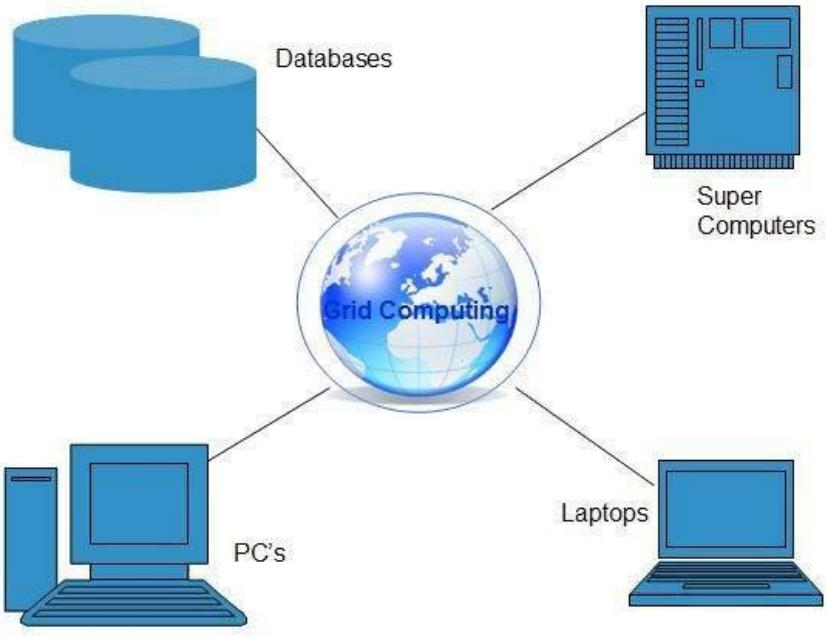
Service-Oriented Architecture helps to use applications as a service for other applications regardless the type of vendor, product or technology. Therefore, it is possible to exchange the data between applications of different vendors without additional programming or making changes to services.

The cloud computing service oriented architecture is shown in the diagram below.



Grid Computing

Grid Computing refers to distributed computing, in which a group of computers from multiple locations are connected with each other to achieve a common objective. These computer resources are heterogeneous and geographically dispersed. Grid Computing breaks complex task into smaller pieces, which are distributed to CPUs that reside within the grid.



Utility Computing

Utility computing is based on **Pay-per-Use model.** It offers computational resources on demand as a metered service. Cloud computing, grid computing, and managed IT services are based on the concept of utility computing.

Cloud Computing Architecture: Cloud computing stack

The cloud creates a system where resources can be pooled and partitioned as needed. Cloud architecture can couple software running on virtualized hardware in multiple locations to provide an on-demand service to user-facing hardware and software. It is this unique combination of abstraction and metered service that separates the architectural requirements of cloud computing systems from the general description given for an n-tiered Internet application. Many descriptions of cloud computing describe it in terms of two architectural layers: A client as a front end The "cloud" as abackend.

This is a very simplistic description because each of these two components is composed of several component layers, complementary functionalities, and a mixture of standard and proprietary protocols. Cloud computing may be differentiated from older models by describing

an encapsulated information technology service that is often controlled through an Application Programming Interface (API), thus modifying the services that are delivered over thenetwork.

A cloud can be created within an organization's own infrastructure or outsourced to another datacenter. While resources in a cloud can be real physical resources, more often they are virtualized resources because virtualized resources are easier to modify and optimize. A compute cloud requires virtualized storage to support the staging and storage of data. From a user's perspective, it is important that the resources appear to be infinitely scalable, that the service be measurable, and that the pricing be metered.

Composability

Applications built in the cloud often have the property of being built from a collection of components, a feature referred to as composability. A composable system uses components to assemble services that can be tailored for a specific purpose using standard parts. A composable component must be:

• **Modular:** It is a self-contained and independent unit that is cooperative, reusable, and replaceable.

• **Stateless:** A transaction is executed without regard to other transactions orrequests.

It isn't an absolute requirement that transactions be stateless, some cloud computing applications provide managed states through brokers, transaction monitors, and service buses. In rarer cases, full transactional systems are deployed in the clouds, but these systems are harder to architect in a distributedarchitecture.

Although cloud computing doesn't require that hardware and software be composable, it is a highly desirable characteristic from a developer or user's standpoint, because it makes system design easier to implement and solutions more portable andinteroperable.

There is a tendency for cloud computing systems to become less composable for users as the services incorporate more of the cloud computing stack. From the standpoint of an IaaS (Infrastructure as a Service) vendor such as Amazon Web Services, GoGrid, or Rackspace, it

makes no sense to offer non-standard machine instances to customers, because those customers are almost certainly deploying applications built on standard operating systems such as Linux, Windows, Solaris, or some other well-known operating system.

In the next step up the cloud computing stack, PaaS (Platform as a Service) vendors such as Windows Azure or Google AppEngine may narrow the definition of standard parts to standard parts that work with their own platforms, but at least from the standpoint of the individual platform service provider, the intent is to be modular for their own developers.

When you move to the highest degree of integration in cloud computing, which is SaaS (Software as a Service), the notion of composability for users may completely disappear. AnSaaS vendor such as Quicken.com or Salesforce.com is delivering an application as a service to a customer, and there's no particular benefit from the standpoint of the service provider that the customer be able to compose its own custom applications. A service provider reselling anSaaS may have the option to offer one module or another, to customize the information contained in the module for a client, to sell the service under their own brand, or to perform some other limited kind of customization, but modifications are generally severelylimited.

This idea that composability diminishes going up the cloud computing stack is from the user's point of view. If you are a PaaS or SaaS service provider and your task is to create the platform or service presented to the developer, reseller, or user, the notion of working with a composable system is still a very powerful one. A PaaS or SaaS service provider gets the same benefits from a composable system that a user does—these things, among others:

- Easier to assemblesystems
- Cheaper systemdevelopment
- More reliableoperation
- A larger pool of qualifieddevelopers
- A logical designmethodology

Infrastructure

Most large Infrastructure as a Service (IaaS) providers rely on virtual machine technology to deliver servers that can run applications. Virtual servers described in terms of a machine image or instance have characteristics that often can be described in terms of realservers delivering a certain number of microprocessor (CPU) cycles, memory access, and network bandwidth to customers. Virtual machines are containers that are assigned specific resources. The software that runs in the virtual machines is what defines the utility of the cloud computing system.

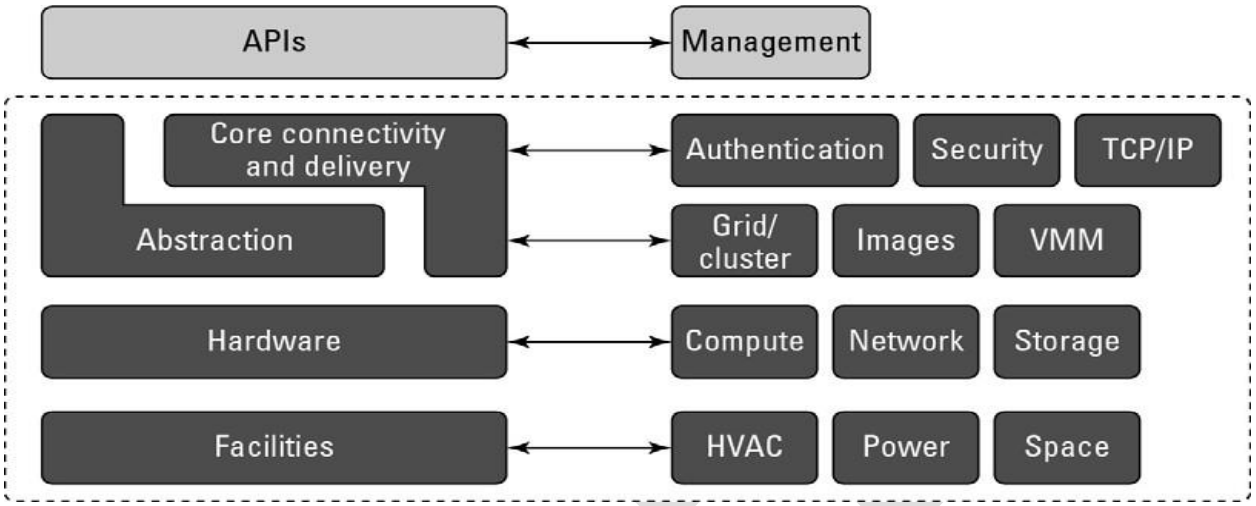
The below figure shows the portion of the cloud computing stack that is defined as the "server." In the diagram, the API is shown shaded in gray because it is an optional component that isn't always delivered with the server. The VMM component is the Virtual Machine Monitor, also called a hypervisor. This is the low-level software that allows different operating systems to run in their own memory space and manages I/O for the virtual machines.

The notion of a virtual server presents to an application developer a new way of thinking about and programming applications. For example, when a programmer is creating software that requires several different tasks to be performed in parallel, he might write an application that creates additional threads of execution that must be managed by the application. When a developer creates an application that uses a cloud service, the developer can attach to the appropriate service(s) and allow the application itself to scale the program execution. Thus, an application such as a three-dimensional rendering that might take a long time for a single server to accomplish can be scaled in the cloud to many servers at once for a short period of time, accomplishing the task at a similar or lower price but at a much faster rate.

In future applications, developers will need to balance the architectural needs of their programs so their applications create new threads when it is appropriate or create new virtual machines. Applications will also need to be mindful of how they use cloud resources, when it is appropriate to scale execution to the cloud, how to monitor the instances they are running, and when not to expand their application's usage of the cloud. This will require a new way of thinking about

application development, and the ability to scale correctly is something that will have to be architected into applications from the groundup.

This architectural diagram illustrates the portion of the cloud computing stack that is designated as theserver.



Platforms

A platform in the cloud is a software layer that is used to create higher levels of service. Many different Platform as a Service (PaaS) providers offer services meant to provide developers with different capabilities. • Salesforce.com's Force.comPlatform

• Windows AzurePlatform

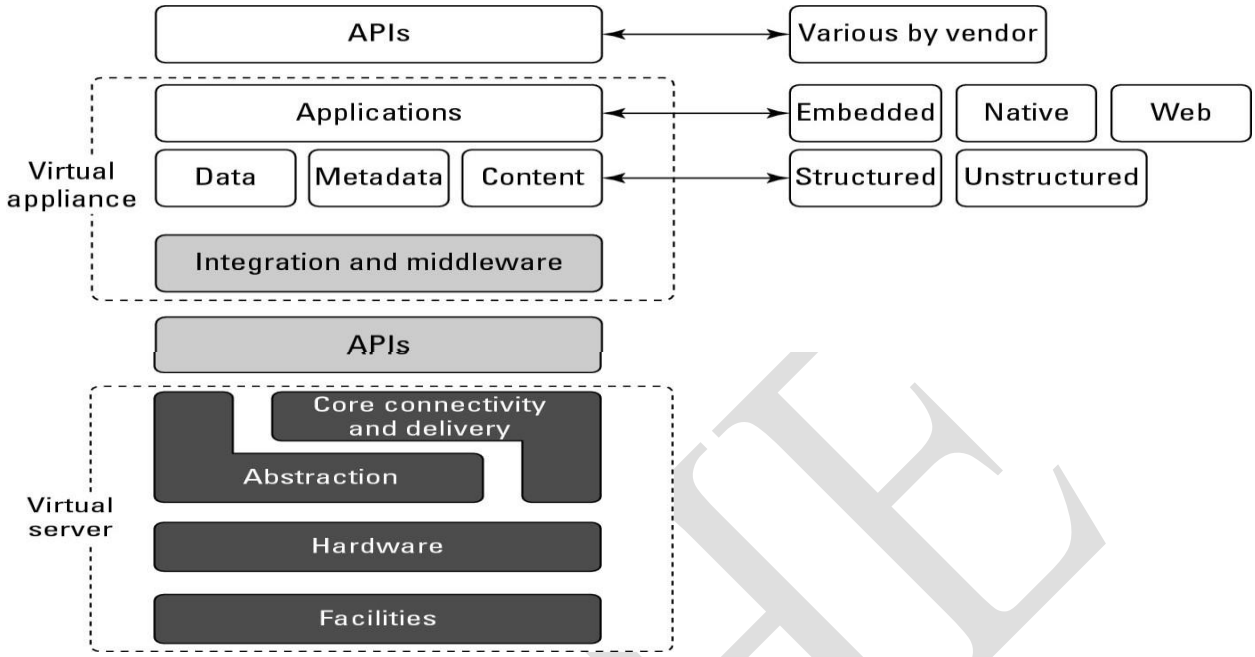
• Google Apps and the GoogleAppEngine

These three services offer all the hosted hardware and software needed to build and deploy Web applications or services that are custom built by the developer within the context and range of capabilities that the platform allows. Platforms represent nearly the full cloud software stack, missing only the presentation layer that represents the user interface. This is the same portion of the cloud computing stack that is a virtual appliance and is shown in Figure. What separates a platform from a virtual appliance is that the software that is installed is constructed from components and services and controlled through the API that the platform provider publishes.

It makes sense for operating system vendors to move their development environments into the cloud with the same technologies that have been successfully used to create Web applications. Thus, you might find a platform based on a Sun xVM hypervisor virtual machine that includes a NetBeans Integrated Development Environment (IDE) and that supports the Sun GlassFish Web stack programmable using Perl or Ruby. For Windows, Microsoft would be similarly interested in providing a platform that allowed Windows developers to run on a Hyper- V VM, use the ASP.NET application framework, support one of its enterprise applications such as SQL Server, and be programmable within Visual Studio—which is essentially what the Azure Platform does. This approach allows someone to develop a program in the cloud that can be used by others.

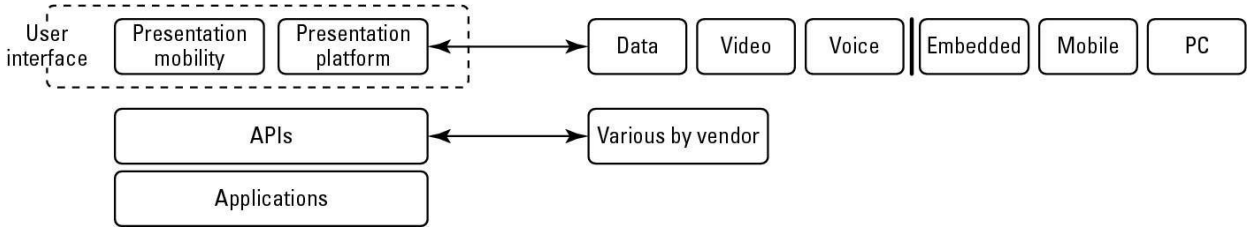
Platforms often come replete with tools and utilities to aid in application design and deployment. Depending upon the vendor, you may find developer tools for team collaboration, testing tools, instrumentation for measuring program performance and attributes, versioning, database and Web service integration, and storage tools. Most platforms begin by establishing a developer community to support the work done in the environment.

A virtual appliance is software that installs as middleware onto a virtual machine.



Just as a virtual appliance may expose itself to users through an API, so too an application built in the cloud using a platform service would encapsulate the service through its own API. Users would then interact with the platform, consuming services through that API, leaving the platform to manage and scale the service appropriately. Many platforms offer user interface development tools based on HTML, JavaScript, or some other technology. As the Web becomes more media-oriented, many developers have chosen to work with rich Internet environments such as Adobe Flash, Flex, or Air, or alternatives such as Windows Silverlight. A user interface abstracts away the platform API, making those services managed through the UI. Figure shows the top portion of the cloud computing stack, which includes the API and the presentationfunctionality.

The top of the cloud computing interface includes the user interface and the API for the application layer.



The Application Programming Interface is one of the key differentiators separating cloud computing from the older models of Internet applications, because it is the means for instantiating resources needed to support applications. An API can control data flow, communications, and other important aspects of the cloud application. Unfortunately, each cloud vendor has their own cloud API, none of them are standard, and the best you can hope for is that eventually the major cloud vendor's APIs will interoperate and exchange data. For now, the use of proprietary APIs results in vendor lock-in, which is why you are advised to choose systems that implement APIs based on open standards.

Virtual Appliances

Applications such as a Web server or database server that can run on a virtual machine image are referred to as virtual appliances. The name *virtual appliance* is a little misleading because it conjures up the image of a machine that serves a narrow purpose. Virtual appliances are software installed on virtual servers—application modules that are meant to run a particular machine instance or image type. A virtual appliance is a platform instance. Therefore, virtual appliances occupy the middle of the cloud computing stack.

A virtual appliance is a common deployment object in the cloud, and it is one area where there is considerable activity and innovation. One of the major advantages of a virtual appliance is that you can use the appliances as the basis for assembling more complex services, the appliance being one of your standardized components. Virtual appliances remove the need for application configuration and maintenance from your list of system management chores.

You run across virtual appliances in IaaS systems such as Amazon's Elastic Compute Cloud

(EC2) . Amazon Machine Images are virtual appliancesthat have been packaged to run on the gird of Xen nodes that comprise the Amazon Web Service's EC2 system. Shown in Figure 3.4, the AMI library

http://developer.amazonwebservices.com/connect/kbcategory.jspa?categoryID=171) includes a variety of operating systems both proprietary and open source, a set of enterprise applications such as Oracle BPM, SQL Server, and even complete application stacks such as LAMP (Linux, Apache, MySQL, and PHP). Amazon has negotiated licenses from these vendors that are part of your per-use pricing when you run these applications on their servers.

Virtual appliances are far easier to install and run than an application that you must set up yourself. However, virtual appliances are also much larger than the application themselves would be because they are usually bundled with the operating system on which they are meant to run. An application that is 50 or 100MB might require a virtual appliance that is 500MB to 1GB in size. Usually, when a virtual appliance is created, the operating system is stripped of all excess functionality that isn't required by the appliance, because the appliance is meant to be used as is.

Amazon Machine Images are a collection of virtual appliances that you can install on their Xen hypervisor servers.

Virtual appliances have begun to affect the PC industry in much the same way that application stores have affected the cell phone industry. You can find various Web sites that either sell or distribute ready-to-use virtual appliances in various forms. Perhaps the best developed of these marketplaces is VMware's Virtual Appliances site (http://www.vmware.com/appliances/) shown in Figure 3.5. These appliances are certified by VMware to be ready to use in the enterprise.

Among the other places you can find virtual appliances are at the Web sites of the various operating system vendors, such as Ubuntu, Xen (http://www xen.org/), and others, including these:

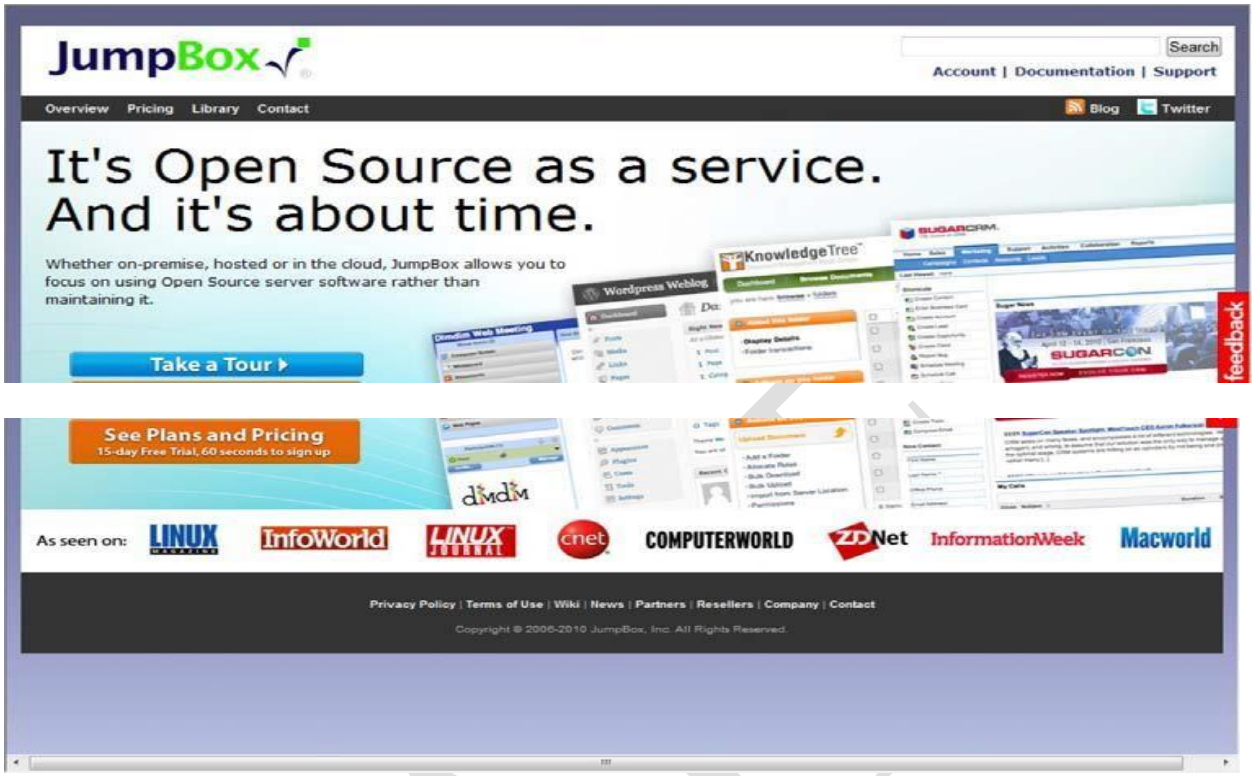
• **Bagvapp**(http://bagside.com/bagvapp/) offers virtual appliances, including ones based on Windows, all of which run on VMwarePlayer.

- **HelpdeskLive**(http://helpdesklive.info/download/VirtualBox%20VDI%20free%20images html) offers various Linux distributions upon which you can build a virtualmachine.
- **Jcinacio**(http://www.jcinacio.com/) has Ubuntuappliances.

VMware's Virtual Appliance marketplace (http://www.vmware.com/appliances/) sells virtual appliances that run on VMware's hypervisor in cloud computing applications.

- **Jumpbox**(http://www.jumpbox.com) offers open source virtual appliances installed by them as a managed service. Jumpbox offers virtual appliances for many applications including Bugzilla, DokuWiki, Drupal, Joomla!,Nagios, OpenVPN, PostgreSQL, Redmine, WordPress, and many others. Figure 3.6 shows the Jumpbox homepage.
- **QEMU** (http://www.qemu.org/) is a CPU emulator and virtual machinemonitor.
- **Parallels** (http://ptn.parallels.com/ptn) hosts a variety of appliances that includes Linux distros, server software, and otherproducts.
- **ThoughtPolice**(http://www.thoughtpolice.co.uk/vmware/) offers appliances based on a variety of Linuxdistributions.
- **VirtualBox**(http://www.virtualbox.org/) is a virtual machine technology now owned by Oracle that can run various operating systems and serves as a host for a variety of virtualappliances.
- **Vmachines**(http://www.vmachines net/) is a site with desktop, server, and security-related operating systems that run onVMware.

Jumpbox (http://www.jumpbox.com/) is an open-source virtual appliance installation and management service.



Converting a virtual appliance from one platform to another isn't an easy proposition. Efforts are underway to create file format standards for these types of objects that make this task easier. The best known of these file formats is the Open Virtualization Format (OVF), the work of the Distributed Management Task Force (DMTF) group. Nearly all major virtualization platform vendors support OVF, notably VMware, Microsoft, Oracle, and Citrix.

Communication Protocols

Cloud computing arises from services available over the Internet communicating using the standard Internet protocol suite underpinned by the HTTP and HTTPS transfer protocols. The other protocols and standards that expose compute and data resources in the cloud either format data or communications in packets that are sent over these two transportprotocols.

In order to engage in interprocess communication (IPC) processes, many client/server protocols have been applied to distributed networking over the years. Various forms of RPC (Remote Procedure Call) implementations (including DCOM, Java RMI, and CORBA) attempt to solve

the problem of engaging services and managing transactions over what is essentially a stateless network. The first of the truly Web-centric RPC technologies was XML-RPC, which uses platform-independent XML data to encode program calls that are transported over HTTP, the networking transport to which nearly everyone is connected.

As Internet computing became more firmly entrenched over the last decade, several efforts began to better define methods for describing and discovering services and resources. The most widely used message-passing standard at the moment is the Simple Object Access Protocol (SOAP), which essentially replaces XML-RPC. SOAP uses XML for its messages and uses RPC and HTTP for message passing. SOAP forms the basis for most of the Web services stacks in use today. If you examine the XML file used in a SOAP transaction, you find that it contains a message and the instructions on how to use the message. The message has a set of rules that are translated into application instances and datatypes, and it defines the methods that must be used to initiate procedure calls and then return a response.

Several standards have emerged to allow the discovery and description of Web-based resources. The most commonly used model for discovery and description used with SOAP messaging is the Web Services Description Language (WSDL), a World Wide Web Consortium (http://www.w3.org/2002/ws/desc/) Internet standard. WSDL lets a Web service advertise itself in terms of a collection of endpoints or ports associated with a specific network address (URL) that can be addressed using XML messages to provide a service. In WSDL, a service is a container that performs a set of functions that are exposed to Web protocols. Taken together, the protocol and port are a binding to which messages are passed and operations are performed. A bound service is one that responds to any valid HTTP request sent to it. The important thing to remember about WSDL is that it defines a Web service's public interface.

Using WSDL and SOAP, a number of extensions were created that allow various Web services to describe additional sets of properties and methods that they could provide. These extensions fall under the name WS-*, or the “WS-star” specifications. A number of WS-* extensions are in

common use, with the following being the most widely used:

- WS-Addressing
- WS-Discovery
- WS-Eventing
- WS-Federation
- WS-MakeConnection
- WS-Messaging
- WS-MetadataExchange
- WS-Notification
- WS-Policy
- WS-ResourceFramework
- WS-Security
- WS-Transfer
- WS-Trust

These different specifications provide a standard means of adding metadata to a SOAP message by modifying the message header while maintaining the message body structure. In this way, a standard method for metadata exchange is piggybacked onto the WSDL XML message. Each of these different WS-* specifications is in a different state of development.

You use these various WS-* services in your daily work. For example, the Web Services Dynamic Discovery specification (WS-Discovery) is a specification for multicast discovery on a LAN (Local Area Network) that is extended to Web services, most often as SOAP over UDP (User Datagram Protocol). When you open the Network Neighborhood in Windows and use the People Near Me feature, WS-Discovery goes into action and shows you discoverable resources. These WS-* services carried over XML messages using the SOAP protocol access remote server applications in ways that are becoming increasingly complex. Whereas earlier methods for client/server provided a means through a gateway like CGI to access media content on servers, the current data communications burden servers with accepting and processing very complex

requests or engaging their clients in sophisticated negotiations that seek to minimize the amount of processing that must be done and the information that must be exchanged as the response. None of this type of rich media servicing was ever envisaged in the construction of the Internet, and all of it is essentially a kludge.

Over the years, a variety of platform-specific RPC specifications, such as DCOM (Distributed Common Object Model) and CORBA (Common Object Response Broker Architecture), were developed to allow software components that ran on different computers to interoperate with one another. As SOAP and WS-* were developing, those protocols began to build into their specifications server application features from these other technologies in a more platform-independent protocol. What was really needed was a method for standardizing resources on the Web, which is where the idea of REST comesin.

REST stands for Representational State Transfer, and it owes its original description to the work of Roy Fielding, who was also a co-developer of the HTTP protocol. REST assigns a global identifier to a resource so there is a uniform method for accessing information sources. That identifier is a URI expressed in HTTP form. Given a resource then at a known address, various network clients in the form of what are called *user agents* can then communicate with that resource using HTTP commands (requests) to exchange information in the form of documents or files. Typical data transfers might use XML, text, an image file, a JSON document, or some other standard or agreed upon format to perform the data exchange. A transaction following the rules of REST is therefore considered to be RESTful, and this is the basis for cloud computing transactions to be initiated, processed, and completed in most modernimplementations.

While REST is heavily used, it is not the only data interchange standard that is used by cloud services. Another example of a data exchange standard is the Atom or Atom Publishing Protocol (APP). Atom is a syndication format that allows for HTTP protocols to create and update

information. Microsoft's ADO.NET Data Services Framework is another system for transferring data using a RESTful transaction and standard HTTP commands.

Cloud services span the gamut of computer applications: audio and video streaming, instant messaging, and so forth. Each of these areas uses protocols developed for network use and adapted for use by Web services. The impact of cloud computing on network communication is to encourage the use of open-network protocols in place of proprietary protocols. For example, in the area of instant messaging, which is a major cloud service, the SIMPLE protocol (which stands for the Session Initiated Protocol for Instant Messaging and Presence Leveraging Extensions) is in widespread use today. SIMPLE is based on the IETF Session Initiation Protocol (SIP) standard and is an open standard. Although most early IM (Instant Messaging) services used proprietary standards, many IM services now support SIMPLE because without this support, it would be hard for the different services to interoperate. Similarly, in the area of VoIP, you find that the open XMPP or the Extensible Messaging and Presence Protocol isused.

Applications

Although the cloud computing stack encompasses many details that describe how clouds are constructed, it is not a perfect vehicle for expressing all the considerations that one must account for in any deployment. An important omission arises from the nature of distributed Web applications and the design of Internet protocols as a stateless service. The Internet was designed to treat each request made to a server as an independent transaction. Therefore, the standard HTTP commands are all atomic in nature: GET to read data, PUT to writedata, and so on. While stateless servers are easier to architect and stateless transactions are more resilient and can survive outages, much of the useful work that computer systems need to accomplish are stateful. Here's the classic example. When you go to a reservation system to purchase something, you query inventory, reserve the item, and then pay for it. In a multiuser system, if you don't have a stateful system, you cannot know whether the item you reserved has already been taken by another user before you can enter your payment for the item. Should you decide you don't want the item at some later time, it is much easier to restore the item to inventory and return payments or

make other adjustments if you can roll back all the steps as a transactional unit.

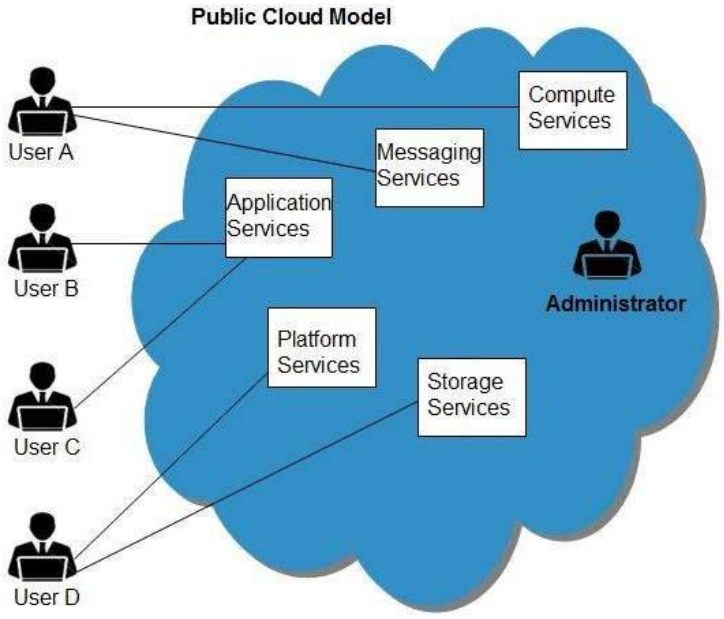
Much of the really hard development efforts that have gone into making the Web useful in commerce have centered around creating mechanisms to change a set of stateless transactions into stateful ones. The development of transaction servers, message queuing servers, and other middleware is meant to bridge this problem. Cloud computing is no exception to this problem, and to an extent it amplifies the problem by not only making transactions stateless but also virtualizing resources so transactions are always occurring in physically different locations. In cloud computing, a variety of constructs are brought to bear to solve these issues, but these are the two most important concepts:

• The notion of orchestration—that process flow can be choreographed as aservice

• The use of what is referred to as a service bus that controls cloudcomponents

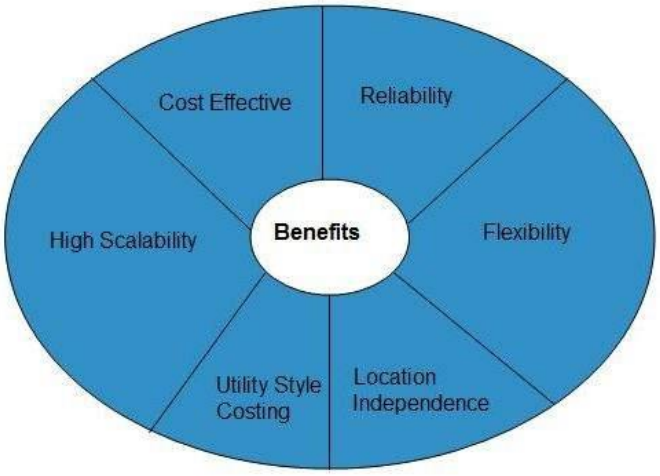
Public cloud

Public Cloud allows systems and services to be easily accessible to general public. The IT giants such as **Google, Amazon** and **Microsoft** offer cloud services via Internet. The Public Cloud Model is shown in the diagram below.



Benefits

There are many benefits of deploying cloud as public cloud model. The following diagram shows some of those benefits:



Cost Effective

Since **public cloud** shares same resources with large number of customers it turns out inexpensive.

Reliability

The **public cloud** employs large number of resources from different locations. If any of the resources fails, public cloud can employ another one.

Flexibility

The public cloud can smoothly integrate with private cloud, which gives customers a flexible approach.

Location Independence

Public cloud services are delivered through Internet, ensuring location independence.

Utility Style Costing

Public cloud is also based on **pay-per-use** model and resources are accessible whenever customer needs them.

High Scalability

Cloud resources are made available on demand from a pool of resources, i.e., they can be scaled up or down according the requirement.

Disadvantages

Here are some disadvantages of public cloud model:

Low Security

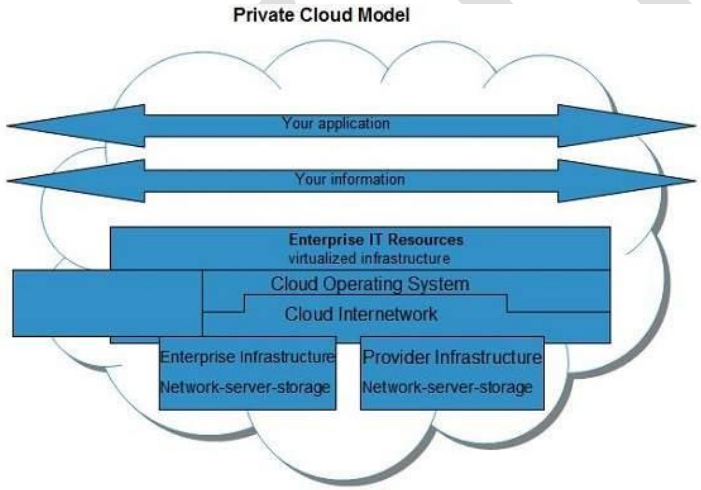
In **public cloud model,** data is hosted off-site and resources are shared publicly, therefore does not ensure higher level of security.

Less Customizable

It is comparatively less customizable than private cloud.

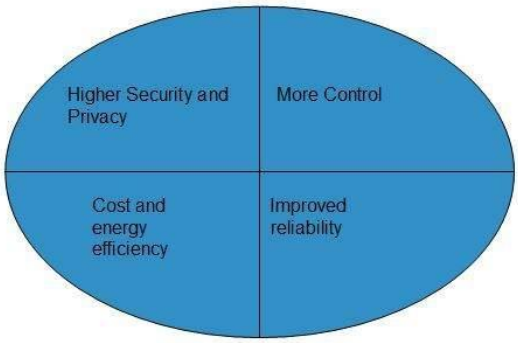
Private Cloud

Private Cloud allows systems and services to be accessible within an organization. The Private Cloud is operated only within a single organization. However, it may be managed internally by the organization itself or by third-party. The private cloud model is shown in the diagram below.



Benefits

There are many benefits of deploying cloud as private cloud model. The following diagram shows some of those benefits:



High Security and Privacy

Private cloud operations are not available to general public and resources are shared from distinct pool of resources. Therefore, it ensures high **security** and **privacy.**

More Control

The **private cloud** has more control on its resources and hardware than public cloud because it is accessed only within an organization.

Cost and Energy Efficiency

The **private cloud** resources are not as cost effective as resources in public clouds but they offer more efficiency than public cloud resources.

Disadvantages

Here are the disadvantages of using private cloud model:

Restricted Area of Operation

The private cloud is only accessible locally and is very difficult to deploy globally.

High Priced

Purchasing new hardware in order to fulfill the demand is a costly transaction.

Limited Scalability

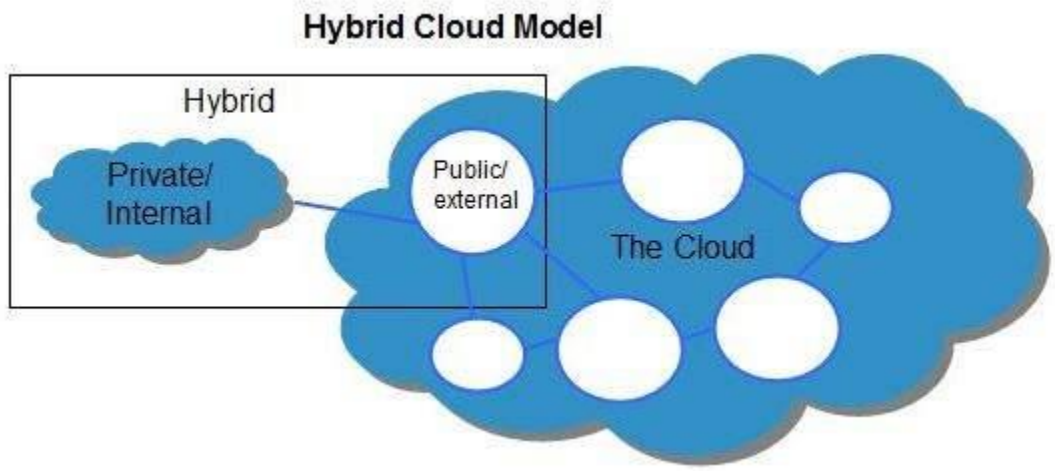
The private cloud can be scaled only within capacity of internal hosted resources.

Additional Skills

In order to maintain cloud deployment, organization requires skilled expertise.

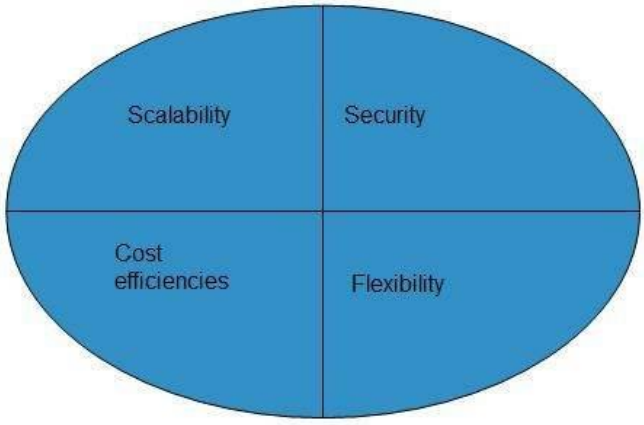
Hybrid cloud

Hybrid Cloud is a mixture of **public** and **private** cloud. Non-critical activities are performed using public cloud while the critical activities are performed using private cloud. The Hybrid Cloud Model is shown in the diagram below.



Benefits

There are many benefits of deploying cloud as hybrid cloud model. The following diagram shows some of those benefits:



Scalability

It offers features of both, the public cloud scalability and the private cloud scalability.

Flexibility

It offers secure resources and scalable public resources.

Cost Efficiency

Public clouds are more cost effective than private ones. Therefore, hybrid clouds can be cost saving.

Security

The private cloud in hybrid cloud ensures higher degree of security.

Disadvantages

Networking Issues

Networking becomes complex due to presence of private and public cloud.

Security Compliance

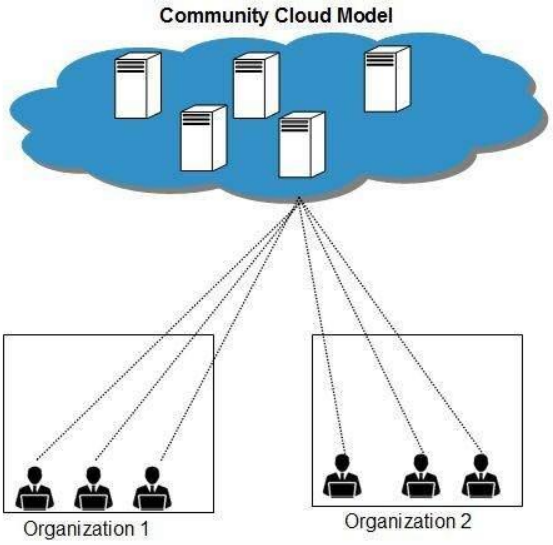
It is necessary to ensure that cloud services are compliant with security policies of the organization.

Infrastructure Dependency

The **hybrid cloud model** is dependent on internal IT infrastructure, therefore it is necessary to ensure redundancy across data centers.

Community Cloud

Community Cloud allows system and services to be accessible by group of organizations. It shares the infrastructure between several organizations from a specific community. It may be managed internally by organizations or by the third-party. The Community Cloud Model is shown in the diagram below.



Benefits

There are many benefits of deploying cloud as **community cloud model.**

Cost Effective

Community cloud offers same advantages as that of private cloud at low cost.

Sharing Among Organizations

Community cloud provides an infrastructure to share cloud resources and capabilities among several organizations.

Security

The community cloud is comparatively more secure than the public cloud but less secured than the private cloud.

Issues

- Since all data is located at one place, one must be careful in storing data in community cloud because it might be accessible to others.
- It is also challenging to allocate responsibilities of governance, security and cost among organizations.

POSSIBLE QUESTIONS

6 Marks

1. Explain the characteristics of cloud computing.
2. List the advantages and disadvantages of cloud computing
3. Discuss the role of open standards in cloud.
4. Compare Private and Public cloud
5. Elaborate on Cloud computing architecture
6. Compare community and hybrid cloud.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

ONE MARK QUESTIONS

DEPARTMENT OF CS, CA & IT

STAFF NAME: Dr.S.MANJU PRIYA **SUBJECT NAME: CLOUD COMPUTING**

SUB.CODE: 18CSP104 **UNIT I** **SEMESTER: V**

S.NO	Question	Choice1	Choice2	Choice3	Choice4	Ans
1	_____refers to applications and services that run on a distributed network using virtualized	Cloud Computing	Virtual Computing	Cloud Storage	Cloud Networking	Cloud Computing
2	The _____infrastructure is operated for the exclusive use of an organization.	Public Cloud	Private Cloud	Community Cloud	Hybrid Cloud	Private Cloud
3	_____ is a complete operating environment with applications, management, and the user interface	CaaS	PaaS	IaaS	SaaS	SaaS
4	The _____is something that you can obtain under contract from your vendor.	QoS	QpS	QtS	QaS	QoS

5	_____refers to the components and subcomponents required for Cloud Computing	Cloud Computing	Cloud Computing	Cloud Computing	Cloud Based Delivery	Cloud Computing
6	_____refers to the location and management of the cloud's infrastructure.	Service models	Deployment models	Development models	Business models	Deployment models
7	_____ is one of the cloud applications in use.	Cloud backup	Cloud storage	Cloud service	Cloud Networking	Cloud backup
8	A_____ cloud is one where the cloud has been organized to serve a common function or purpose.	Hybrid cloud	Community cloud	Private cloud	Public cloud	Community cloud
9	A _____ cloud combines multiple clouds are bound together as a unit.	Community cloud	Public cloud	Private cloud	Hybrid cloud	Hybrid cloud
10	Google App Engine is an example of _____ services.	CaaS	PaaS	IaaS	SaaS	PaaS
11	SQL Azure is an example of _____ services.	CaaS	PaaS	IaaS	SaaS	SaaS
12	The Backend platforms are called as_____	Intercloud	Mobile device	Storage	Fat client	Storage
13	_____is taking the physical hardware and going completely virtual	IaaS	PaaS	Daas	SaaS	IaaS
14	Cloud networking is a_____ network	Non Agile	Agile	Latency	Low Latency	Agile
15	_____constitute the first expression of cloud computing	Community Cloud	Private Cloud	Public Cloud	Hybrid Cloud	Public cloud
16	A fundamental characteristic of public clouds is _____	Security	High bandwidth	QoS	multi tenancy	multi tenancy
17	_____ is most commonly implemented in PaaS solutions that support hybrid clouds.	Dynamic provisioning	Provisioning	Distributed Mapping	Mapping	Dynamic provisioning
18	_____ are distributed systems created by integrating the services of different clouds to	Community cloud	Private cloud	Public cloud	Hybrid cloud	Community cloud

19	From an architectural point of view, a _____ is most likely implemented over multiple	Community cloud	Private cloud	Public cloud	Hybrid cloud	Community cloud
20	_____abstracts the details of system implementation from users and developers.	Cloud Computing	Virtual Computing	Cloud Storage	Cloud Networking	Cloud Computing
21	_____consists of the particular types of services that you can access on a cloud computing	Development models	Deployment models	Service models	Business models	Service models
22	_____is an example of IaaS service providers	Oracle on Demand	GoogleApps	Force.com	Eucalyptus	Eucalyptus
23	A cloud computing deployment lets someone else manage your computing infrastructure while you	Outsourced IT Staffing	Outsourced IT	QoS	Outsourced IT deployment	Outsourced IT management
24	All cloud computing applications suffer from the inherent latency that is intrinsic in their	MAN	WAN	LAN	LAN & MAN	WAN
25	Cloud computing is a _____system	stateful	stateup	stateless	statedown	stateless
26	A single area of concern in cloud computing is_____	privacy and network	security and storage	storage and network	privacy and security	privacy and security
27	The use of the word “cloud” makes reference to the____and___essential concepts.	Abstraction& Virtualization	Services & applications	Virtualization & Services	Abstraction& applications	Abstraction & Virtualization
28	_____provides virtual machines, operating systems, applications, services, development	IaaS	PaaS	Daas	SaaS	PaaS
29	Expand EC2	Elastic Cloud Compute	Extended Compute	Elastic Compute	Extended Cloud	Elastic Compute Cloud
30	_____can be rapidly and elastically provisioned.	Data	Network	Information	Resources	Resources
31	_____ is one of the services that are heavily deployed on cloud computing systems.	VoIP	IPoV	TCP	UDP	VoIP
32	_____creates a single point of failure.	Fat Clients	The Zero Clients	Thick Clients	Cloud Clients	The Zero Clients

33	_____provides the equivalent of installed applications in the traditional delivery of	IaaS	Daas	SaaS	PaaS	SaaS
34	_____ are open systems in which fair competition between different solutions can	Community cloud	Public cloud	Private cloud	Hybrid cloud	Community cloud
35	Science clouds are an interesting example of _____	Public cloud	Community cloud	Private cloud	Hybrid cloud	Community cloud
36	_____are appealing and provide a viable option to cut IT costs and reduce capital expenses.	Private cloud	Community cloud	Public cloud	Hybrid cloud	Public cloud
37	Customer information protection is an aspect of _____	Private cloud	Community cloud	Public cloud	Hybrid cloud	Private cloud
38	In most cases the _____ option prevails because of the existing IT infrastructure.	Public cloud	Community cloud	Private cloud	Hybrid cloud	Private cloud
39	From an architectural point of view, a_____ is most likely implemented over multiple	Community cloud	Public cloud	Hybrid cloud	Private cloud	Community cloud
40	_____ cloud is used for healthcare industry.	Private cloud	Public cloud	Hybrid cloud	Community cloud	Community cloud
41	_____has the least levels of integrated functionality.	IaaS	PaaS	Daas	SaaS	IaaS
42	_____has the most levels of integrated functionality.	IaaS	SaaS	Daas	PaaS	SaaS
43	Expand SLA	Strorage Level	Service Level	Service Level	Storage Level Applications	Service Level Agreement
44	_____ is not a benefit of cloud computing.	Resource pooling	Rapid elasticity	Infinite data	Measured service	Infinite data
45	If your application needs large amounts of data transfer, _____ may not be the best model for you.	Distributed computing	Load balancing	Virtualization	Cloud computing	Cloud computing

46	Institutions such as government and military agencies will not consider _____ as an option for processing or storing their sensitive data.	Public cloud	Hybrid cloud	Private cloud	Community cloud	Public cloud
47	_____ is not an opertaion of Quality of Service.	Data replication	Queries	System monitoring	Disaster recovery	Queries
48	_____is the inability to scale on demand and to efficiently address peak loads	Public cloud	Hybrid cloud	Private cloud	Community cloud	Private cloud
49	_____address scalability issues by leveraging external resources for exceeding capacity demand	Public cloud	Hybrid cloud	Private cloud	Community cloud	Hybrid cloud
50	One of the fundamental components of PaaS middleware is the mapping of _____ onto the cloud infrastructure	Dynamic applications	Standalone applications	Standard applications	Distributed applications	Distributed applications
51	Cloud computing represents a _____ in the way in which systems are deployed.	Real time applications	Real Paradigm	Infinitely Scalable	Measurable Service	Real Paradigm shift
52	The scale of cloud computing networks and their ability to provide_____ makes them highly reliable.	Lower costs	Ease of utilization	Load balancing and failover	Simplified maintenance and upgrade	Load balancing and failover
53	Cloud computing industry continues to address _____ concerns, if you have an application that works with sensitive data.	Security	Privacy	Storage	Bigdata	Security
54	_____is not an architectural standards in Cloud computing.	Grid computin	Distributed computing	Autonomic systems	Standardized Web services	Distributed computing
55	_____ is one of the large IaaS cloud service providers	Rackspace.com	Salesforce.com	GoGrid.com	Openstack.com	Rackspace.com
56	_____share common concerns such as their mission, policies, security, regulatory compliance needs, and so on.	Public cloud	Hybrid cloud	Community cloud	Private cloud	Community cloud

57	IDaaS Stands for_____	Infrastructure as a Service	Independent as a Service	Interdependent as a	Identity as a Service	Identity as a Service
58	_____ has a number of operating systems and some enterprise applications that they offer on a	Eucalyptus	Amazon	MS Azure	GoGrid	Amazon
59	_____ are open systems in which fair competition between different solutions can	Public cloud	Hybrid cloud	Community cloud	Private cloud	Community cloud
60	_____represents the ability for a cloud service to be widely accessible.	Multitenancy	On-Demand	Ubiquitous Access	Resiency	Ubiquitous Access
61	_____is a CPU emulator and virtual machine monit	Parallels	QEMU	Jumpbox	Vmachines	QEMU

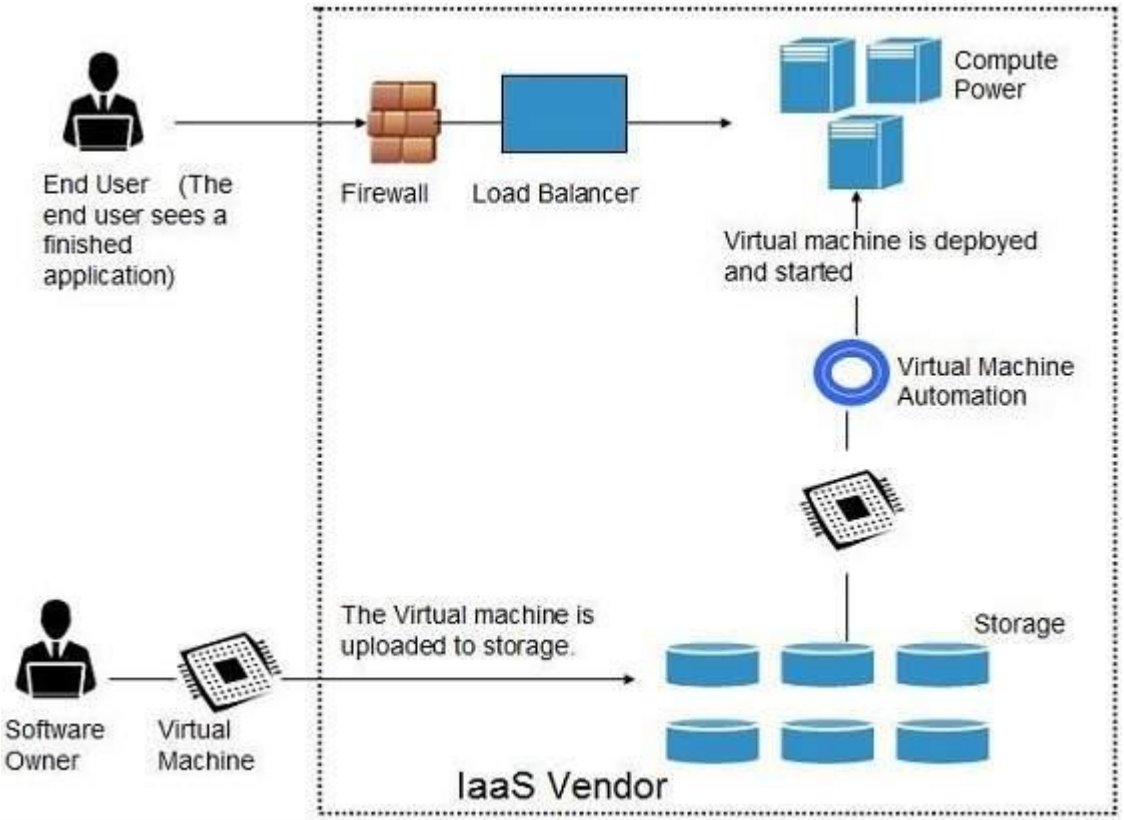
UNIT –II
Infrastructure as a Service (IaaS) -Platform as a Service (PaaS) -Software as a Service (SaaS) -Identity as a Service (IDaaS) -Compliance as a Service (CaaS)- Cloud storage

Infrastructure as a Service (IaaS)

Infrastructure-as-a-Service provides access to fundamental resources such as physical machines, virtual machines, virtual storage, etc. Apart from these resources, the IaaS also offers:

- Virtual machine disk storage
- Virtual local area network (VLANs)
- Load balancers
- IP addresses
- Software bundles

All of the above resources are made available to end user via **server virtualization.** Moreover, these resources are accessed by the customers as if they own them.



Benefits

IaaS allows the cloud provider to freely locate the infrastructure over the Internet in a cost-effective manner. Some of the key benefits of IaaS are listed below:

- Full control of the computing resources through administrative access to VMs.
- Flexible and efficient renting of computer hardware.
- Portability, interoperability with legacy applications.

Full control over computing resources through administrative access to VMs

IaaS allows the customer to access computing resources through administrative access to virtual machines in the following manner:

- Customer issues administrative command to cloud provider to run the virtual machine or to save data on cloud server.
- Customer issues administrative command to virtual machines they owned to start web server or to install new applications.

Flexible and efficient renting of computer hardware

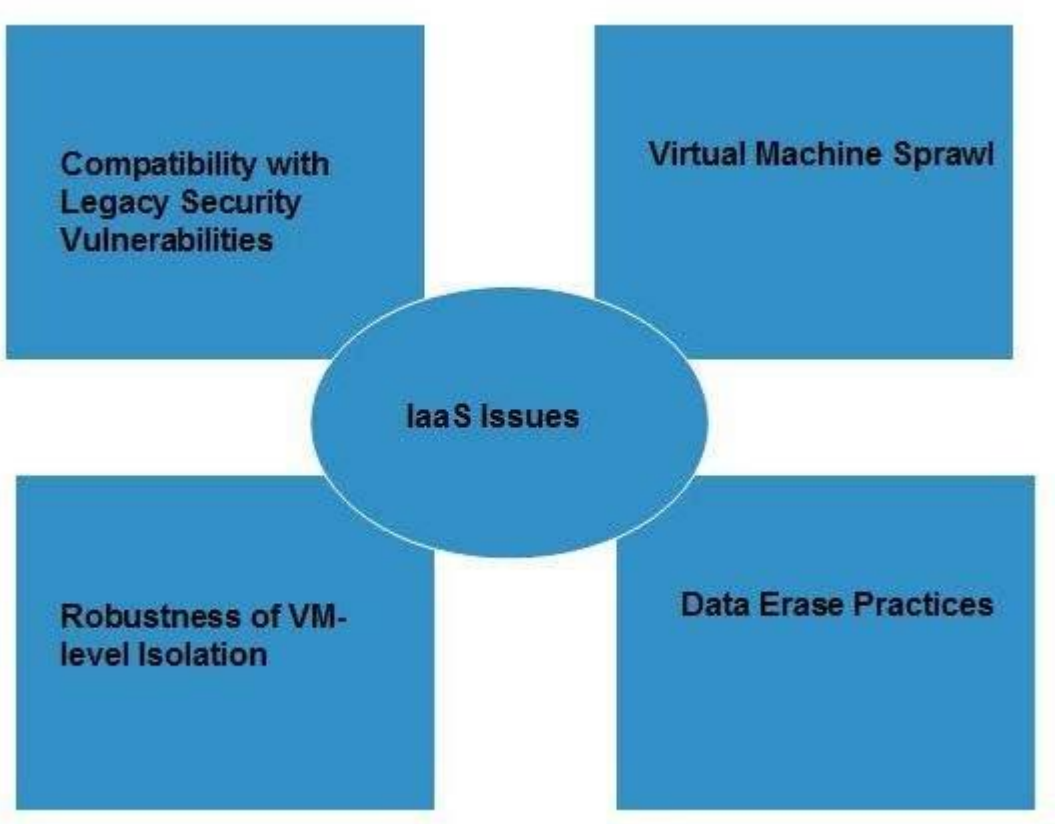
IaaS resources such as virtual machines, storage devices, bandwidth, IP addresses, monitoring services, firewalls, etc. are made available to the customers on rent. The payment is based upon the amount of time the customer retains a resource. Also with administrative access to virtual machines, the customer can run any software, even a custom operating system.

Portability, interoperability with legacy applications

It is possible to maintain legacy between applications and workloads between IaaS clouds. For example, network applications such as web server or e-mail server that normally runs on customer-owned server hardware can also run from VMs in IaaS cloud.

Issues

IaaS shares issues with PaaS and SaaS, such as Network dependence and browser based risks. It also has some specific issues, which are mentioned in the following diagram:



Compatibility with legacy security vulnerabilities

Because IaaS offers the customer to run legacy software in provider's infrastructure, it exposes customers to all of the security vulnerabilities of such legacy software.

Virtual Machine sprawl

The VM can become out-of-date with respect to security updates because IaaS allows the customer to operate the virtual machines in running, suspended and off state. However, the provider can automatically update such VMs, but this mechanism is hard and complex.

Robustness of VM-level isolation

IaaS offers an isolated environment to individual customers through hypervisor. Hypervisor is a software layer that includes hardware support for virtualization to split a physical computer into multiple virtual machines.

Data erase practices

The customer uses virtual machines that in turn use the common disk resources provided by the cloud provider. When the customer releases the resource, the cloud provider must ensure that next customer to rent the resource does not observe data residue from previous customer.

Characteristics

Here are the characteristics of IaaS service model:

- Virtual machines with pre-installed software.
- Virtual machines with pre-installed operating systems such as Windows, Linux, and Solaris.
- On-demand availability of resources.
- Allows to store copies of particular data at different locations.
- The computing resources can be easily scaled up and down.

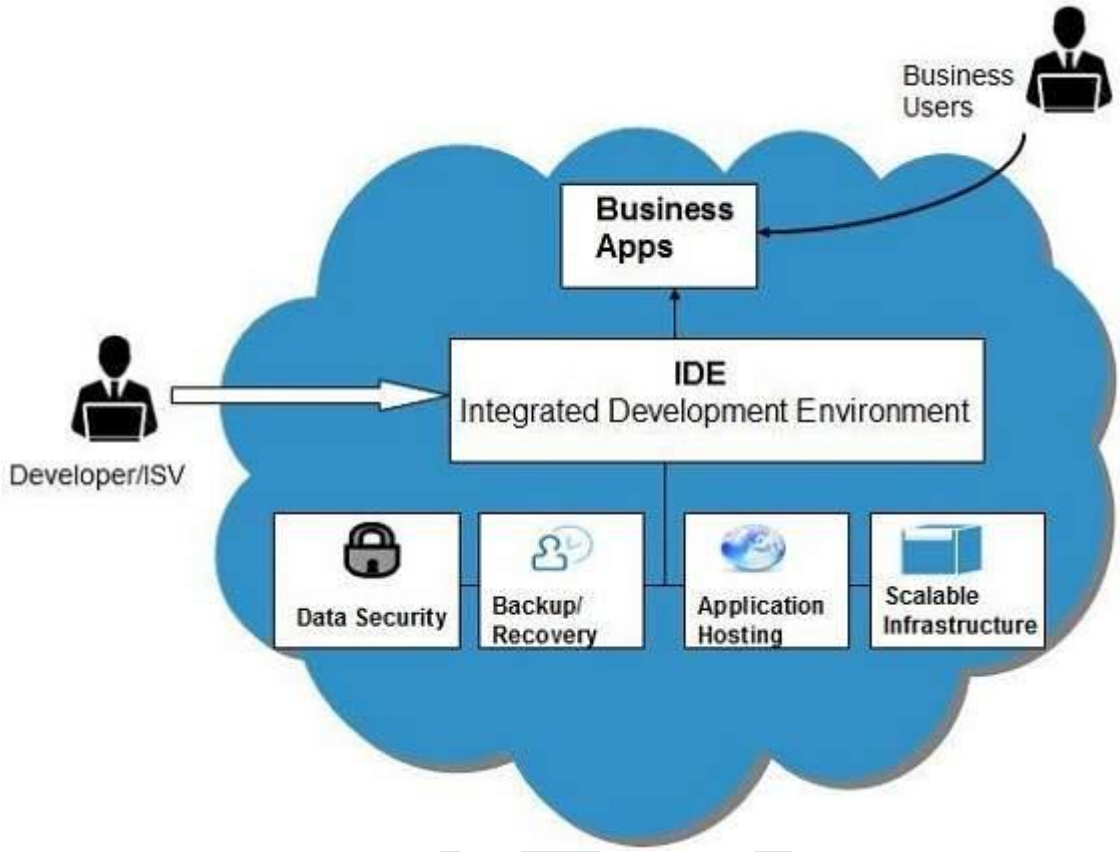
Platform as a Service (PaaS)

Platform-as-a-Service offers the runtime environment for applications. It also offers development and deployment tools required to develop applications. PaaS has a feature of **point-and-click** tools that enables non-developers to create web applications.

App Engine of Google and **Force.com** are examples of PaaS offering vendors. Developer may log on to these websites and use the **built-in API** to create web-based applications.

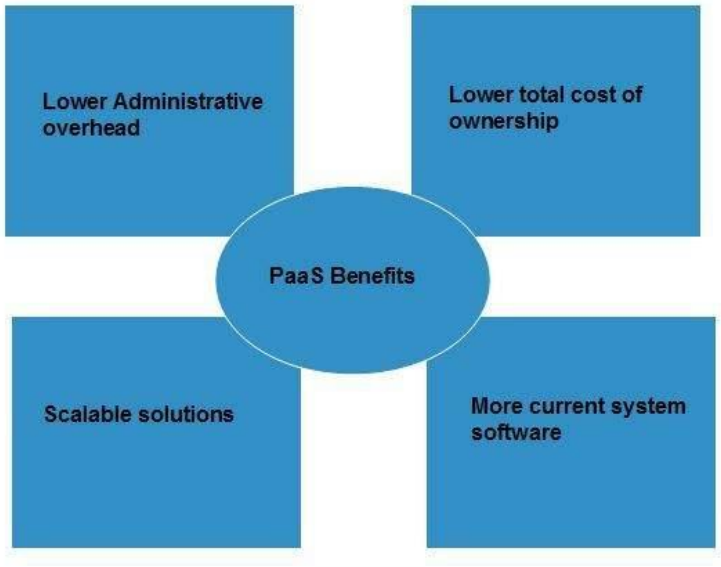
But the disadvantage of using PaaS is that, the developer **locks-in** with a particular vendor. For example, an application written in Python against API of Google, and using App Engine of Google is likely to work only in that environment.

The following diagram shows how PaaS offers an API and development tools to the developers and how it helps the end user to access business applications.



Benefits

Following are the benefits of PaaS model:



Lower administrative overhead

Customer need not bother about the administration because it is the responsibility of cloud provider.

Lower total cost of ownership

Customer need not purchase expensive hardware, servers, power, and data storage.

Scalable solutions

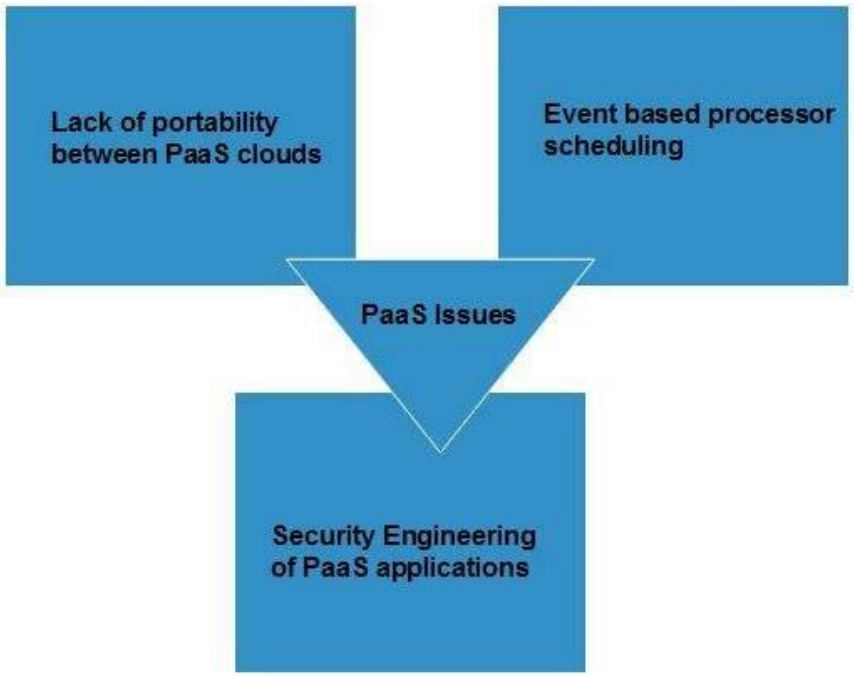
It is very easy to scale the resources up or down automatically, based on their demand.

More current system software

It is the responsibility of the cloud provider to maintain software versions and patch installations.

Issues

Like **SaaS, PaaS** also places significant burdens on customer's browsers to maintain reliable and secure connections to the provider's systems. Therefore, PaaS shares many of the issues of SaaS. However, there are some specific issues associated with PaaS as shown in the following diagram:



Lack of portability between PaaS clouds

Although standard languages are used, yet the implementations of platform services may vary. For example, file, queue, or hash table interfaces of one platform may differ from another, making it difficult to transfer the workloads from one platform to another.

Event based processor scheduling

The PaaS applications are event-oriented which poses resource constraints on applications, i.e., they have to answer a request in a given interval of time.

Security engineering of PaaS applications

Since PaaS applications are dependent on network, they must explicitly use cryptography and manage security exposures.

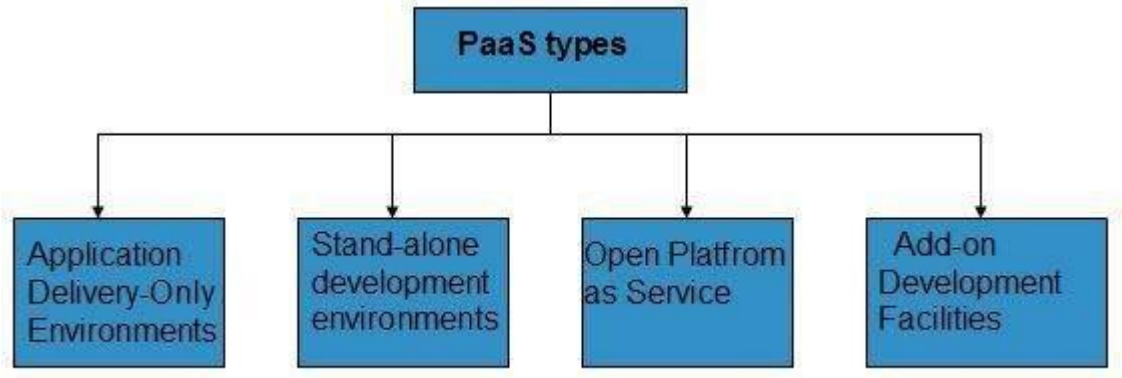
Characteristics

Here are the characteristics of PaaS service model:

- PaaS offers **browser based development environment.** It allows the developer to create database and edit the application code either via Application Programming Interface or point-and-click tools.
- PaaS provides **built-in security, scalability,** and **web service interfaces.**
- PaaS provides built-in tools for defining **workflow, approval processes,** and business rules.
- It is easy to integrate PaaS with other applications on the same platform.
- PaaS also provides web services interfaces that allow us to connect the applications outside the platform.

PaaS Types

Based on the functions, PaaS can be classified into four types as shown in the following diagram:



Stand-alone development environments

The **stand-alone PaaS** works as an independent entity for a specific function. It does not include licensing or technical dependencies on specific SaaS applications.

Application delivery-only environments

The **application delivery PaaS** includes **on-demand scaling** and **application security.**

Open platform as a service

Open PaaS offers an **open source software** that helps a PaaS provider to run applications.

Add-on development facilities

The **add-on PaaS** allows to customize the existing SaaS platform.

Software as a Service (SaaS)

Software-as–a-Service (SaaS) model allows to provide software application as a service to the end users. It refers to a software that is deployed on a host service and is accessible via Internet. There are several SaaS applications listed below:

- Billing and invoicing system
- Customer Relationship Management (CRM) applications
- Help desk applications
- Human Resource (HR) solutions

Some of the SaaS applications are not customizable such as **Microsoft Office Suite.** But SaaS provides us **Application Programming Interface (API),**which allows the developer to develop a customized application.

Characteristics

Here are the characteristics of SaaS service model:

- SaaS makes the software available over the Internet.
- The software applications are maintained by the vendor.
- The license to the software may be subscription based or usage based. And it is billed on recurring basis.
- SaaS applications are cost-effective since they do not require any maintenance at end user side.
- They are available on demand.
- They can be scaled up or down on demand.
- They are automatically upgraded and updated.
- SaaS offers shared data model. Therefore, multiple users can share single instance of infrastructure. It is not required to hard code the functionality for individual users.
- All users run the same version of the software.

Benefits

Using SaaS has proved to be beneficial in terms of scalability, efficiency and performance. Some of the benefits are listed below:

- Modest software tools
- Efficient use of software licenses
- Centralized management and data
- Platform responsibilities managed by provider
- Multitenant solutions

Modest software tools

The SaaS application deployment requires a little or no client side software installation, which results in the following benefits:

- No requirement for complex software packages at client side
- Little or no risk of configuration at client side
- Low distribution cost

Efficient use of software licenses

The customer can have single license for multiple computers running at different locations which reduces the licensing cost. Also, there is no requirement for license servers because the software runs in the provider's infrastructure.

Centralized management and data

The cloud provider stores data centrally. However, the cloud providers may store data in a decentralized manner for the sake of redundancy and reliability.

Platform responsibilities managed by providers

All platform responsibilities such as backups, system maintenance, security, hardware refresh, power management, etc. are performed by the cloud provider. The customer does not need to bother about them.

Multitenant solutions

Multitenant solutions allow multiple users to share single instance of different resources in virtual isolation. Customers can customize their application without affecting the core functionality.

Issues

There are several issues associated with SaaS, some of them are listed below:

- Browser based risks
- Network dependence
- Lack of portability between SaaS clouds

Browser based risks

If the customer visits malicious website and browser becomes infected, the subsequent access to SaaS application might compromise the customer's data.

To avoid such risks, the customer can use multiple browsers and dedicate a specific browser to access SaaS applications or can use virtual desktop while accessing the SaaS applications.

Network dependence

The SaaS application can be delivered only when network is continuously available. Also network should be reliable but the network reliability cannot be guaranteed either by cloud provider or by the customer.

Lack of portability between SaaS clouds

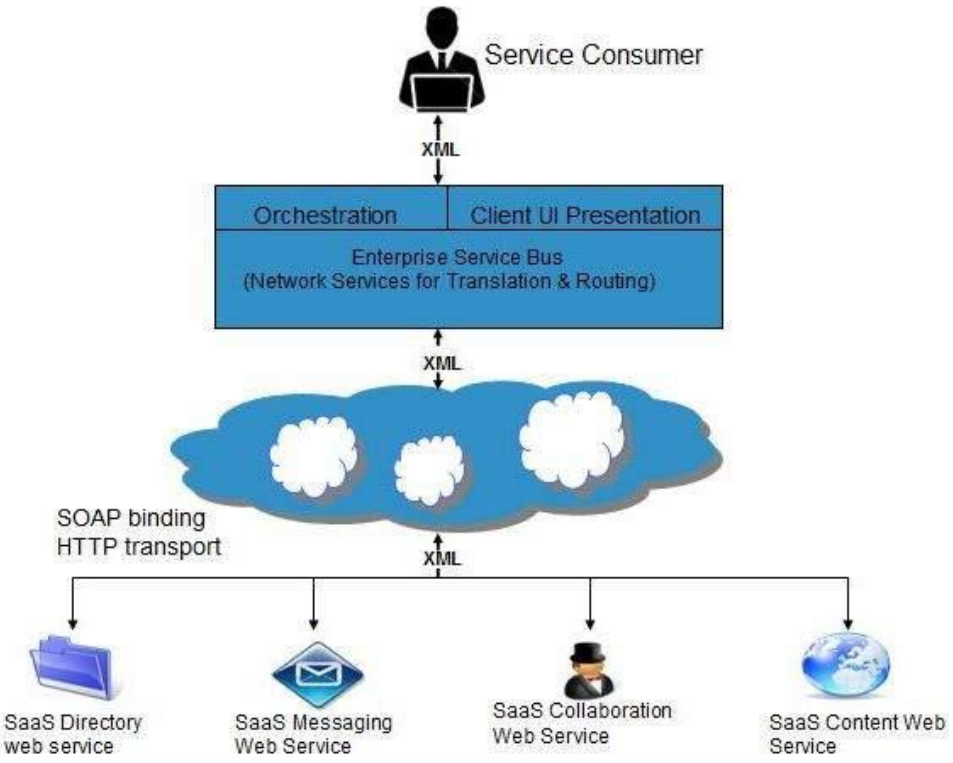
Transferring workloads from one SaaS cloud to another is not so easy because work flow, business logics, user interfaces, support scripts can be provider specific.

Open SaaS and SOA

Open SaaS uses those SaaS applications, which are developed using open source programming language. These SaaS applications can run on any open source operating system and database. Open SaaS has several benefits listed below:

- No License Required
- Low Deployment Cost
- Less Vendor Lock-in
- More portable applications
- More Robust Solution

The following diagram shows the SaaS implementation based on SOA:



Identity as a Service (IDaaS)

Employees in a company require to login to system to perform various tasks. These systems may be based on local server or cloud based. Following are the problems that an employee might face:

- Remembering different username and password combinations for accessing multiple servers.
- If an employee leaves the company, it is required to ensure that each account of that user is disabled. This increases workload on IT staff.

To solve above problems, a new technique emerged which is known as **Identity-as–a-Service (IDaaS).**

IDaaS offers management of identity information as a digital entity. This identity can be used during electronic transactions.

Identity

Identity refers to set of attributes associated with something to make it recognizable. All objects may have same attributes, but their identities cannot be the same. A unique identity is assigned through unique identification attribute.

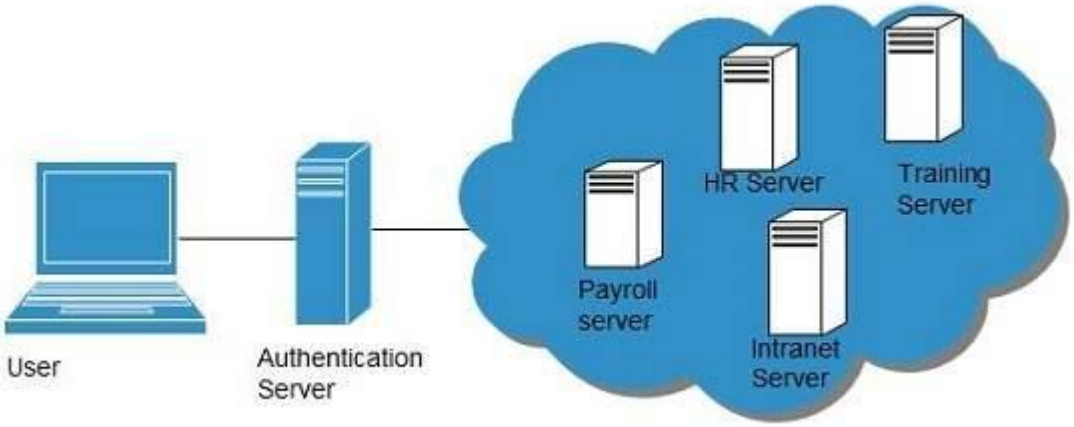
There are several **identity services** that are deployed to validate services such as validating web sites, transactions, transaction participants, client, etc. Identity-as-a-Service may include the following:

- Directory services
- Federated services
- Registration
- Authentication services
- Risk and event monitoring
- Single sign-on services
- Identity and profile management

Single Sign-On (SSO)

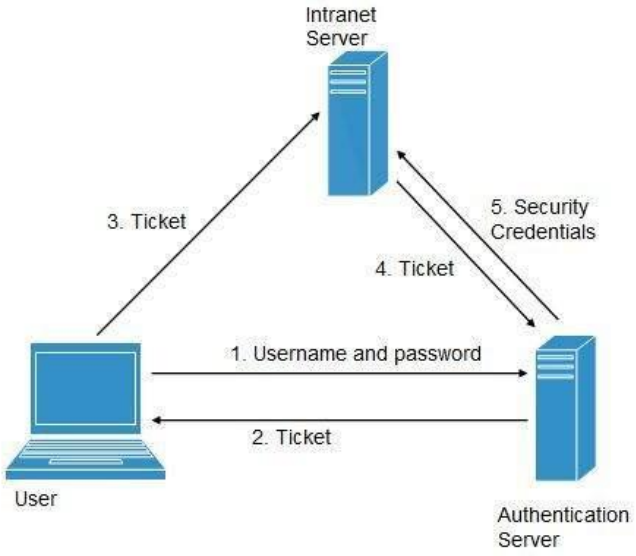
To solve the problem of using different username and password combinations for different servers, companies now employ Single Sign-On software, which allows the user to login only one time and manage the access to other systems.

SSO has single authentication server, managing multiple accesses to other systems, as shown in the following diagram:



SSO Working

There are several implementations of SSO. Here, we discuss the common ones:



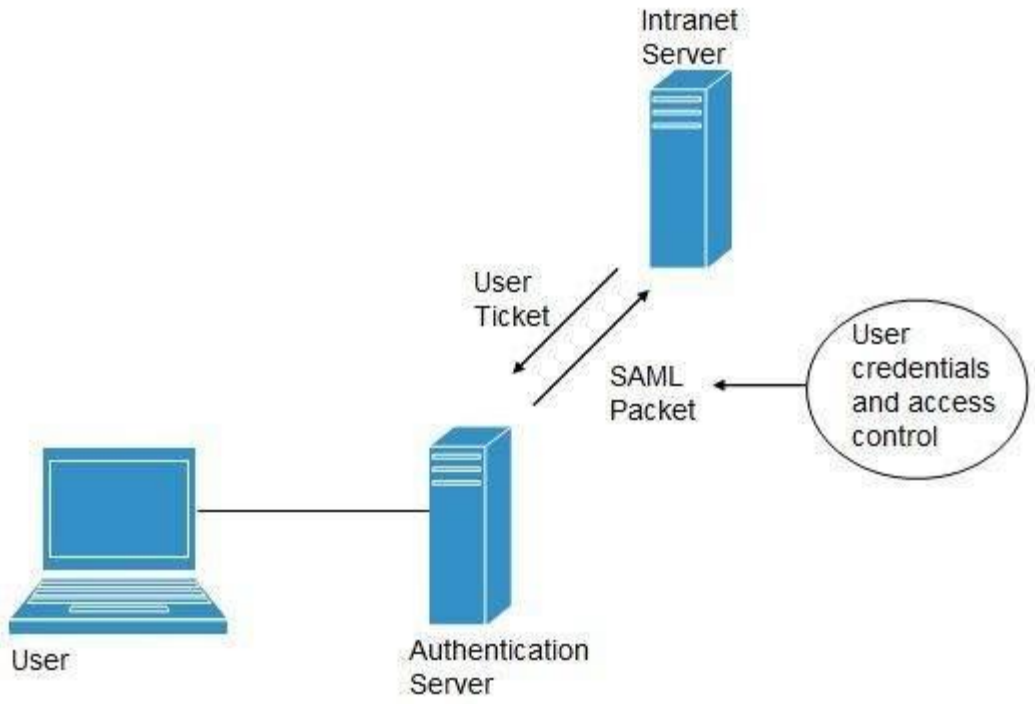
Following steps explain the working of Single Sign-On software:

- User logs into the authentication server using a username and password.
- The authentication server returns the user's ticket.
- User sends the ticket to intranet server.
- Intranet server sends the ticket to the authentication server.
- Authentication server sends the user's security credentials for that server back to the intranet server.

If an employee leaves the company, then disabling the user account at the authentication server prohibits the user's access to all the systems.

Federated Identity Management (FIDM)

FIDM describes the technologies and protocols that enable a user to package security credentials across security domains. It uses **Security Markup Language (SAML)** to package a user's security credentials as shown in the following diagram:



OpenID

It offers users to login into multiple websites with single account. Google, Yahoo!, Flickr, MySpace, WordPress.com are some of the companies that support OpenID.

Benefits

- Increased site conversation rates
- Access to greater user profile content
- Fewer problems with lost passwords
- Ease of content integration into social networking sites

Compliance as a Service (CaaS

Compliance as a service is offered for specific standards, such as PCI-DSS or HIPAA. Typically, what this means is that, in addition to the servers, you are purchasing a variety of **compliance services** such as data encryption, disaster recovery, reporting, vulnerability scanning, etc. There are many more cloud providers who are recognizing the need to provide their users with regulation-compliant services and more and more businesses are also going to their cloud providers to be compliant to certain standards.

Simplifies the Process

This includes the needed encryption levels and the types of data that need to be hidden and/or given extra protection. It also streamlines the compliance process, as most cloud providers do not just provide service, but also offer education and resources to help businesses simplify administration based on their obligations to certain regulations.

Automatic Updates

Cloud providers that provide compliance-as-a-service offerings need to keep up with the ever-changing sets of regulations and standards their service aims to comply with. They adjust their service based on these changes to remain compliant. As a subscriber to the service, you would not have to worry about updating your system according to these changes because the cloud provider will be rolling out the updates automatically to all their users.

These benefits make it easier for you to choose between trusting your cloud provider to provide compliance-as-a-service solutions or trying to meet compliance requirements on

your own. In order to implement CaaS, some companies are organizing what might be referred to as "vertical clouds," clouds that specialize in a vertical market. Examples of vertical clouds that advertise CaaS capabilities include the following:

• **athenahealth** (http://www.athenahealth.com/) for the medical industry

• **bankserv** (http://www.bankserv.com/) for the banking industry

• **ClearPoint PCI** Compliance-as-a-Service for merchant transactions under the Payment Card Industry Data Security Standard

• **FedCloud** (http://www fedcloud.com/) for government

• **Rackserve PCI** Compliant Cloud (http://www rackspace.com/; another PCI CaaS service)

It's much easier to envisage a CaaS system built inside a private cloud where the data is under the control of a single entity, thus ensuring that the data is under that entity's secure control and that transactions can be audited. Indeed, most of the cloud computing compliance systems to date have been built using private clouds.

It is easy to see how CaaS could be an incredibly valuable service. A well-implemented CaaS service could measure the risks involved in servicing compliance and ensure or indemnify customers against that risk. CaaS could be brought to bear as a mechanism to guarantee that an e- mail conformed to certain standards, something that could be a new electronic service of a network of national postal systems—and something that could help bring an end to the scourge of spam.

Cloud storage

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

Storage Devices

Storage devices can be broadly classified into two categories:

- Block Storage Devices
- File Storage Devices

Block Storage Devices

The **block storage devices** offer raw storage to the clients. These raw storage are partitioned to create volumes.

File Storage Devices

The **file Storage Devices** offer storage to clients in the form of files, maintaining its own file system. This storage is in the form of Network Attached Storage (NAS).

Cloud Storage Classes

Cloud storage can be broadly classified into two categories:

- Unmanaged Cloud Storage
- Managed Cloud Storage

Unmanaged Cloud Storage

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

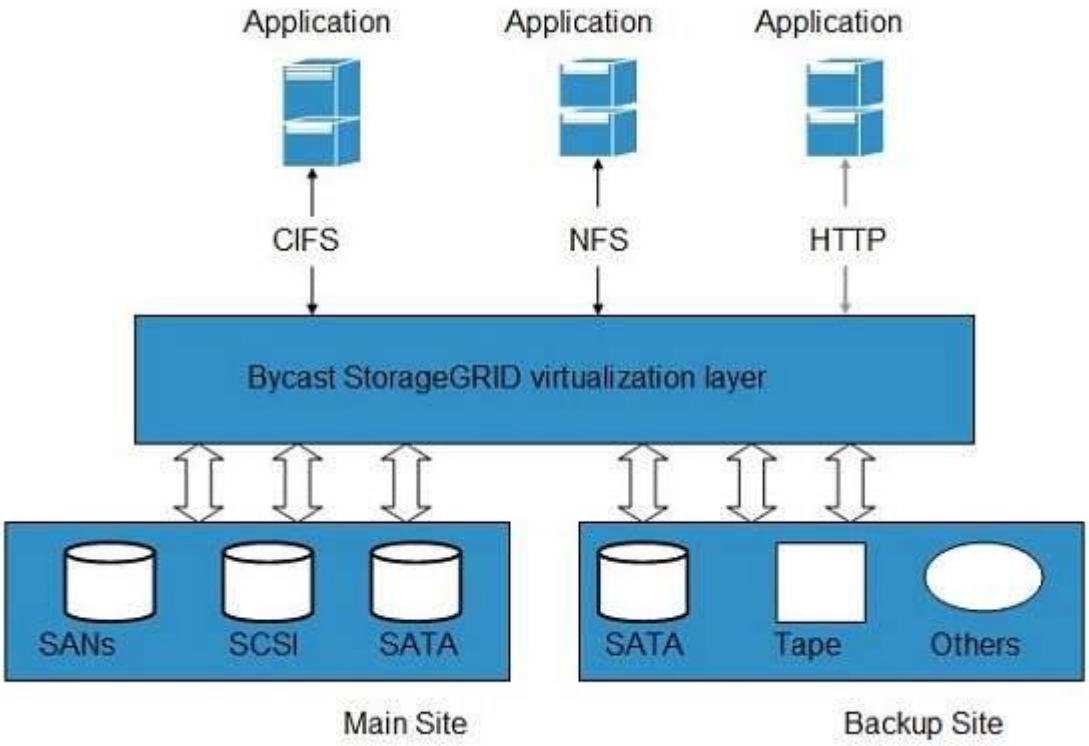
Managed Cloud Storage

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

Creating Cloud Storage System

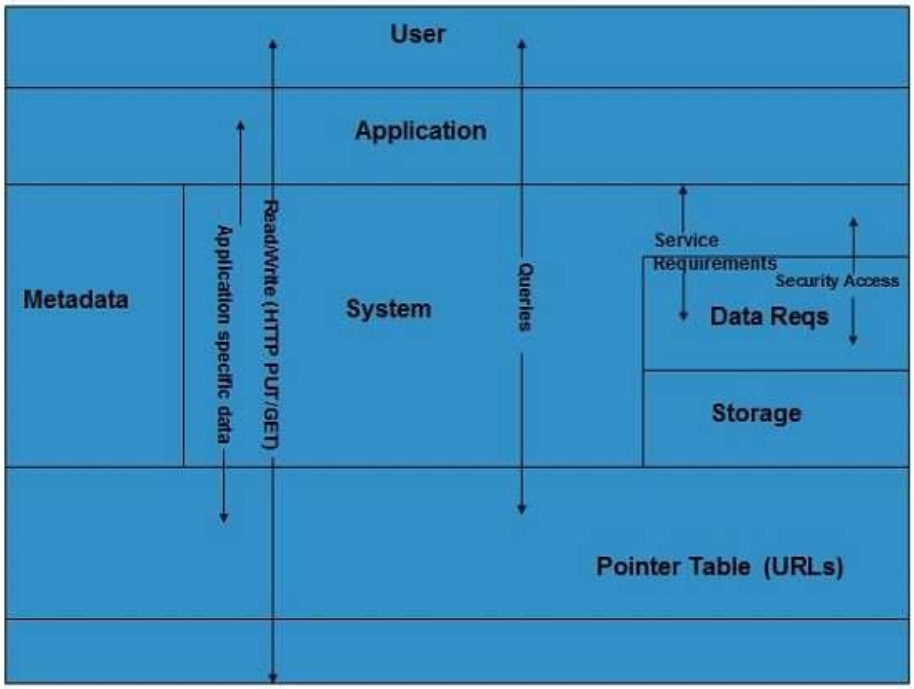
The cloud storage system stores multiple copies of data on multiple servers, at multiple locations. If one system fails, then it is required only to change the pointer to the location, where the object is stored.

To aggregate the storage assets into cloud storage systems, the cloud provider can use storage virtualization software known as **StorageGRID.** It creates a virtualization layer that fetches storage from different storage devices into a single management system. It can also manage data from **CIFS** and **NFS** file systems over the Internet. The following diagram shows how StorageGRID virtualizes the storage into storage clouds:



Virtual Storage Containers

The **virtual storage containers** offer high performance cloud storage systems. **Logical Unit Number (LUN)** of device, files and other objects are created in virtual storage containers. Following diagram shows a virtual storage container, defining a cloud storage domain:



Challenges

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

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

POSSIBLE QUESTIONS

6 MARKS

1. Compare the characteristics of PaaS and SaaS.
2. Discuss the role of CaaS.
3. Discuss the role of Cloud Storage.
4. Compare the characteristics of IaaS and SaaS.
5. Explain about PaaS.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

ONE MARK QUESTIONS

DEPARTMENT OF CS, CA & IT

STAFF NAME: Dr.S.MANJU PRIYA **SUBJECT NAME: CLOUD COMPUTING**

SUB.CODE: 18CSP104 **UNIT II** **SEMESTER: I**

S.NO	Question	Choice1	Choice2	Choice3	Choice4	Ans
1	We can broadly partition cloud computing into____ layers that form a cloud computing	3	4	5	6	4
2	The _____ forms the basis for Software as a Service.	Application layer	Datalink layer	Platform layer	Network layer	Application layer
3	_____ is a cloud computing service model in which hardware is virtualized in the cloud.	Development as a Service	Platform as a Service	Infrastructure as a Service	Software as a Service	Infrastructure as a Service
4	The fundamental unit of virtualized client in an IaaS deployment is called a _____	Workload	Scheduling	Infrastructure	Operating System	Workload

5	A_____ would reserve a machine equivalent required to run each of these workloads.	Client	Server	Network	Host	Client
6	A group of users within a particular instance is called a _____	Aggregation	Silos	Pod	Network	Pod
7	____ are the cloud computing equivalent of compute islands.	Aggregation	Silos	Pod	Network	Silos
8	The one example that is most quoted as a PaaS offering is _____	Force.com	Amazon Web	Google's App Engine	Quickbase	Google's App Engine platform
9	An _____ is a set of characteristics or traits that make something recognizable or known.	Infrastructure	Integrity	Identity	Infinite data	Identity
10	_____ is one of the more expensive and complex areas of network computing.	Software protection	Information protection	Identity protection	Privacy	Identity protection
11	XRI stands for _____	eXtensible Resource	eXtensible Resource	eXtensible Raw	eXtensible Raw Identifier	eXtensible Resource
12	In order to implement CaaS, some companies are organizing what might be referred to as _____	Horizontal clouds	Vertical clouds	Fed clouds	Rack serve clouds	Vertical clouds
13	_____ is mainly meant for developers and to support applications built using Web services.	Unmanaged cloud storage	Managed cloud	Vertical cloud storage	Horizontal cloud storage	Managed cloud storage
14	_____ can manage data from CIFS and NFS file systems over HTTP networks.	GoGRID	SystemGRID	StorageGRID	DataGRID	StorageGRID
15	Snapshots are also known as_____	Differential backups	Image backups	Point-in-time backups	Incremental backups	Point-in-time backups
16	A _____ creates a full backup first and then periodically synchronizes the full copy with the live version.	Incremental backup	Differential backup	Point-in-time backup	Reverse delta backup	Reverse delta backup
17	CDMI stands for_____	Cloud Data Module	Cloud Data Management	Cloud Data Management	Cloud Data Module	Cloud Data Management

18	VIM stands for _____	Vendor Interface	Vendor Interface	Vendor Identity	Vendor Identity	Vendor Interface
19	_____can access objects stored in the cloud by using standard HTTP command and the REST protocol to manipulate those objects.	VIM	CDMI	CMID	VMI	CDMI
20	ROA stands for_____	Reverse Oriented	Return Oriented	Resource Oriented	Rapid Oriented	Resource Oriented
21	The ability to provide storage on demand from a storage pool is referred to as _____	Static provisioning	Dynamic provisioning	Thick provisioning	Thin provisioning	Thin provisioning
22	Open file backup systems are _____	Less Expensive	Expensive	Exchangable	Non Exchangable	Expensive
23	Continuous Data Protection (CDP) also known as_____	Porting	Imaging	Mirroring	Manipulating	Mirroring
24	An _____ allows a system to do what is referred to as a bare metal restore.	Image backup	Incremental backup	Differential backup	Point-in-time backup	Image backup
25	_____is an example of software that supplies Image backup.	Carbonite	Ghost	Apple's Time Machine	SQL Server	Ghost
26	_____is an example of software that supplies Point-in-time backup.	Carbonite	Ghost	Apple's Time Machine	SQL Server	Carbonite
27	_____ virtualizes storage into storage clouds.	SystemGRID	StorageGRID	DataGRID	GoGRID	SystemGRID
28	_____ is a direct competitor to Amazon's S3 service.	Nirvanix	Iron Mountain	Rackspace	EMC Atmos	Rackspace
29	Most of the user-based applications that work with cloud storage are of_____ type.	Unmanaged storage	Managed storage	Web storage	Rack storage	Unmanaged storage
30	_____ offer faster data transfers, but impose additional overhead on clients.	File storage devices	Block storage	Network Attached	Web storage devices	Block storage devices

31	Most of the cloud computing compliance systems to date have been built using_____	Public clouds	Private clouds	Community clouds	Hybrid clouds	Private clouds
32	FedCloud is used for_____	Banking industry	Medical industry	Government	Merchant transactions	Government
33	In IaaS, the virtualized resources are mapped to _____	Real systems	Virtual systems	Sophisticated systems	Dynamic systems	Real systems
34	The work done in IaaS can be measured by the number of _____	Data Per Minute	Process Per Minute	Transactions Per Minute	Clients Per Minute	Transactions Per Minute
35	In cloud computing, a provisioned server called an ____ is reserved by a customer.	Input	Instance	Application	Output	Instance
36	From an architectural standpoint, _____ in an IaaS infrastructure is assigned its own private network.	Server	Client	New User	Host	Client
37	_____ limits broadcast and multicast traffic because Data Link Layer in networking is not	Amazon Web Service's	Google's App Engine	Quickbase routing	Rackspace routing	Amazon Web Service's
38	Consider a transactional eCommerce system, for which a typical stack contains____components.	4	3	5	6	5
39	Pods are managed by a _____	Google App Engine	Cloud Storage	Amazon Web Services	Cloud Control System	Cloud Control System
40	____ are processing domains that are sealed off from the outside.	Silos	Aggregation	Pod	Network	Silos
41	_____ can be based on specific types of development languages, application frameworks,	Infrastructure	Softwares	Platforms	Compliance	Platforms
42	The ____ is responsible for all the operational aspects of the service, for maintenance, and for managing the product(s) lifecycle.	Hardware	Vendor	Software	Infrastructure	Vendor
43	A developer might write an application in a programming language like ____ using the Google	C++	C	Python	Pascal	Python

44	Gmail is an offering of _____	DaaS	SaaS	PaaS	CaaS	PaaS
45	The_____ is available over the Internet globally through a browser on demand.	Software	Hardware	Platform	Compliance	Software
46	_____applications feature automated upgrades, updates, and patch management and much faster rollout of changes.	SaaS	DaaS	PaaS	CaaS	SaaS
47	_____ supports multiple users and provides a shared data model through a single-instance, multi-tenancy model.	DaaS	PaaS	SaaS	IaaS	SaaS
48	_____is at the heart of the Internet as a service that provides identity authorization and lookup.	Distributed Name Service	Distributed Name	Domain Name Service	Domain Name	Domain Name Service
49	ORM stands for_____	Object Role Management	Object Role Model	Object Right Model	Object Right Management	Object Role Model
50	Cloud computing_____applications must rely on a set of developing industry standards to provide interoperability.	DaaS	IDaaS	SaaS	IaaS	IDaaS
51	XACML stands for_____	eXtensible Access	eXtensible Access	eXtensible Assertion	eXtensible Assertion	eXtensible Access Control
52	SAML stands for_____	Security Authenticatio n Markup Language	Security Attributes Markup Language	Security Acess Markup Language	Security Assertion Markup Language	Security Assertion Markup Language
53	SPML stands for_____	Security Provisioning Markup	Services Provisioning Markup	Services Processing Markup	Security Processing Markup	Services Provisioning Markup

54	XACML standards is used for_____	Provisioning	Audit	Authentication	Authorization	Authorization
55	SPML standards is used for_____	Provisioning	Audit	Authentication	Authorization	Provisioning
56	InforCards and OpenID standards is used for_____	Provisioning	Audit	Authentication	Authorization	Authentication
57	The _____ standard applies to the unique identity of the URL.	OpenID	SPML	SAML	XACML	OpenID
58	_____exposes its storage to clients in the form of files, maintaining its own file system.	Web Stroage devices	Network Attached	File storage devices	Block storage devices	Network Attached
59	_____ service currently in beta allows developers to store their data in Google's cloud storage infrastructure.	Atmos	Platypus	Eucalyptus	Nirvanix	Platypus
60	The _____was designed to be a fault-tolerant network that could survive a nuclear attack.	Internet	Intranet	Intercloud	Security	Internet

UNIT -III

Virtualization Technologies -Load Balancing and Virtualization -Advanced load balancing -The Google cloud - Hypervisors -Virtual machine types -VMware vSphere - Machine Imaging - Porting Applications -The Simple Cloud API - AppZero Virtual Application Appliance

Virtualization Technologies

The dictionary includes many definitions for the word "cloud." A cloud can be a mass of water droplets, gloom, an obscure area, or a mass of similar particles such as dust or smoke. When it comes to cloud computing, the definition that best fits the context is "a collection of objects that are grouped together."

It is that act of grouping or creating a resource pool that is what succinctly differentiates cloud computing from all other types of networked systems. The benefits of pooling resources to allocate them on demand are so compelling as to make the adoption of these technologies a priority.

Without resource pooling, it is impossible to attain efficient utilization, provide reasonable costs to users, and proactively react to demand. In this chapter, you learn about the technologies that abstract physical resources such as processors, memory, disk, and network capacity into virtual resources.

When you use cloud computing, you are accessing pooled resources using a technique called virtualization. Virtualization assigns a logical name for a physical resource and then provides a pointer to that physical resource when a request is made. Virtualization provides a means to manage resources efficiently because the mapping of virtual resources to physical resources can be both dynamic and facile. Virtualization is dynamic in that the mapping can be assigned based on rapidly changing conditions, and it is facile because changes to a mapping assignment can be nearly instantaneous.

These are among the different types of virtualization that are characteristic of cloud computing:

- **Access:** A client can request access to a cloud service from any location.
- **Application:** A cloud has multiple application instances and directs requests to an instance based on conditions.
- **CPU:** Computers can be partitioned into a set of virtual machines with each machine being assigned a workload. Alternatively, systems can be virtualized through load-balancing technologies.
- **Storage:** Data is stored across storage devices and often replicated for redundancy.

To enable these characteristics, resources must be highly configurable and flexible. You can define the features in software and hardware that enable this flexibility as conforming to one or more of the following mobility patterns:

- **P2V:** Physical to Virtual
- **V2V:** Virtual to Virtual
- **V2P:** Virtual to Physical
- **P2P:** Physical to Physical
- **D2C:** Datacenter to Cloud
- **C2C:** Cloud to Cloud
- **C2D:** Cloud to Datacenter
- **D2D:** Datacenter to Datacenter

Virtualization is a key enabler of the first four of five key attributes of cloud computing:

- **Service-based:** A service-based architecture is where clients are abstracted from service providers through service interfaces.
- **Scalable and elastic:** Services can be altered to affect capacity and performance on demand.
- **Shared services:** Resources are pooled in order to create greater efficiencies.

- **Metered usage:** Services are billed on a usage basis.
- **Internet delivery:** The services provided by cloud computing are based on Internet protocols and formats.

Load Balancing And Virtualization

The technology used to distribute service requests to resources is referred to as *load balancing.* Load balancing can be implemented in hardware, as is the case with F5's BigIP servers, or in software, such as the Apache mod_proxy_balancerextension, the Pound load balancer and reverse proxy software, and the Squid proxy and cache daemon.

Load balancing is an optimization technique; it can be used to increase utilization and throughput, lower latency, reduce response time, and avoid system overload.

The following network resources can be load balanced:

- Network interfaces and services such as DNS, FTP, and HTTP
- Connections through intelligent switches
- Processing through computer system assignment
- Storage resources
- Access to application instances

Without load balancing, cloud computing would very difficult to manage. Load balancing provides the necessary redundancy to make an intrinsically unreliable system reliable through managed redirection.

It also provides fault tolerance when coupled with a failover mechanism. Load balancing is nearly always a feature of server farms and computer clusters and for high availability applications.

A load-balancing system can use different mechanisms to assign service direction. In the simplest load- balancing mechanisms, the load balancer listens to a network port for service requests. When a request from a client or service requester arrives, the load

balancer uses a scheduling algorithm to assign where the request is sent. Typical scheduling algorithms in use today are round robin and weighted round robin, fastest response time, least connections and weighted least connections, and custom assignments based on other factors.

A session ticket is created by the load balancer so that subsequent related traffic from the client that is part of that session can be properly routed to the same resource. Without this session record or persistence, a load balancer would not be able to correctly failover a request from one resource to another. Persistence can be enforced using session data stored in a database and replicated across multiple load balancers. Other methods can use the client's browser to store a client-side cookie or through the use of a rewrite engine that modifies the URL. Of all these methods, a session cookie stored on the client has the least amount of overhead for a load balancer because it allows the load balancer an independent selection of resources.

The algorithm can be based on a simple round robin system where the next system in a list of systems gets the request. Round robin DNS is a common application, where IP addresses are assigned out of a pool of available IP addresses. Google uses round robin DNS, as described in the next section.

Load Balancing and Virtualization

IT industry is growing each day and so is the need for computing and storage resources. Large quantities of data are generated and exchanged over the network which further necessitates the need of more and more computing resources. Organizations, to better capitalize their investment, are opening their infrastructure to new found virtualization technologies like Cloud computing.

Cloud has helped enterprises leverage the benefits of computing resources which are shared over a virtualized environment. A lot of enterprises are already using cloud-based services in one or the other form. This brings us to the concept of load balancing in cloud.

What is load balancing in Cloud computing?

A website or a web-application can be accessed by a plenty of users at any point of time. It becomes difficult for a web application to manage all these user requests at one time. It may even result in system breakdowns. For a website owner, whose entire work is dependent on his portal, the sinking feeling of website being down or not accessible also brings lost potential customers.

Here, the load balancer plays an important role.

Cloud Load balancing is the process of distributing workloads and computing resources across one or more servers. This kind of distribution ensures maximum throughput in minimum response time. The workload is segregated among two or more servers, hard drives, network interfaces or other computing resources, enabling better resource utilization and system response time. Thus, for a high traffic website, effective use of cloud load balancing can ensure business continuity. The common objectives of using load balancers are:

- To maintain system firmness.
- To improve system performance.
- To protect against system failures.

Cloud providers like **Amazon Web Services (AWS)**, **Microsoft Azure** and **Google o**ffer cloud load balancing to facilitate easy distribution of workloads. For ex: **AWS offers Elastic Load balancing (ELB) technology** to distribute traffic among EC2 instances. Most of the AWS powered applications have ELBs installed as key architectural component.

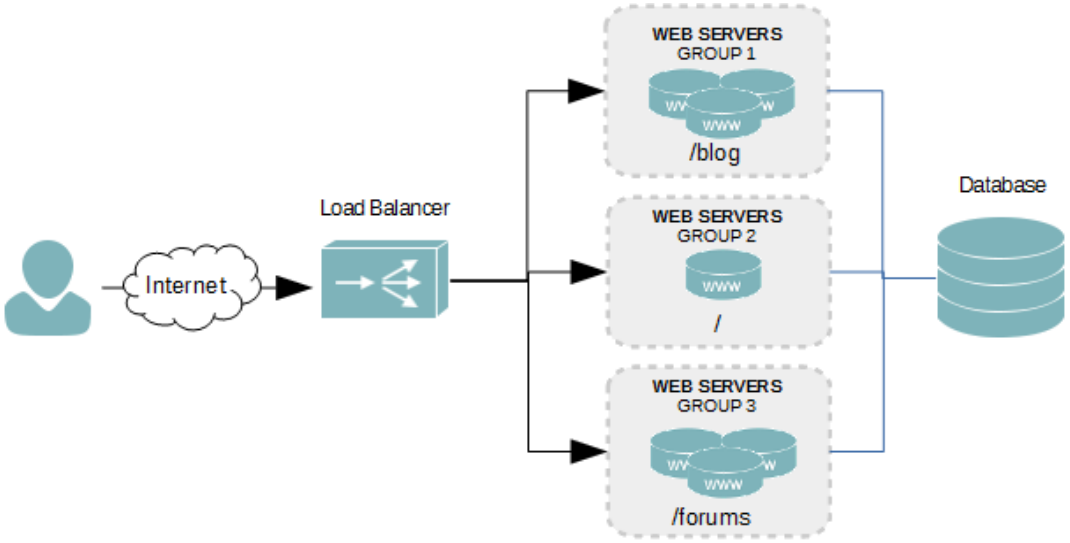
Similarly, **Azure's Traffic Manager** allocates its cloud servers' traffic across multiple datacenters.

How does load balancing work?

Here, load refers to not only the website traffic but also includes CPU load, network load and memory capacity of each server. A load balancing technique makes sure that each system in the network has same amount of work at any instant of time. This means neither any of them is excessively over-loaded, nor under-utilized.

The load balancer distributes data depending upon how busy each server or node is. In the absence of a load balancer, the client must wait while his process gets processed, which might be too tiring and demotivating for him.

Various information like jobs waiting in queue, CPU processing rate, job arrival rate etc. are exchanged between the processors during the load balancing process. Failure in the right application of load balancers can lead to serious consequences, data getting lost being one of them.



Different companies may use different load balancers and multiple load balancing algorithms like static and dynamic load balancing. One of the most commonly used methods is **Round-robin load balancing.**

It forwards client request to each connected server in turn. On reaching the end, the load balancer loops back and repeats the list again. The major benefit is its ease of implementation. The load balancers check the system heartbeats during set time intervals to verify whether each node is performing well or not.

What are the advantages of Cloud Load Balancing?

a) High Performing applications

Cloud load balancing techniques, unlike their traditional on-premise counterparts, are less expensive and simple to implement. Enterprises can make their client applications work faster and deliver better performances, that too at potentially lower costs.

b) Increased scalability

Cloud balancing takes help of cloud's scalability and agility to maintain website traffic. By using efficient load balancers, you can easily match up the increased user traffic and distribute it among various servers or network devices. It is especially important for ecommerce websites, who deals with thousands of website visitors every second. During sale or other promotional offers they need such effective load balancers to distribute workloads.

c) Ability to handle sudden traffic spikes

A normally running University site can completely go down during any result declaration. This is because too many requests can arrive at the same time. If they are using cloud load balancers, they do not need to worry about such traffic surges. No matter how large the request is, it can be wisely distributed among different servers for generating maximum results in less response time.

d) Business continuity with complete flexibility

The basic objective of using a load balancer is to save or protect a website from sudden outages. When the workload is distributed among various servers or network units, even if one node fails the burden can be shifted to another active node.

Thus, with increased redundancy, scalability and other features load balancing easily handles website or application traffic.

ZNetLive offers Cloud VPS and **Dedicated Servers** on latest generation of hardware resources for full scalability and fault-tolerance for ensuring high-availability of websites. Our technical experts ensure that your website is always up and performing high with 99.99% uptime guarantees.

Advanced load balancing

The more sophisticated load balancers are workload managers. They determine the current utilization of the resources in their pool, the response time, the work queue length, connection latency and capacity, and other factors in order to assign tasks to each resource. Among the features you find in load balancers are polling resources for their health, the ability to bring standby servers online (priority activation), workload weighting based on a resource's capacity (asymmetric loading), HTTP traffic compression, TCP offload and buffering, security and authentication, and packet shaping using content filtering and priority queuing.

An Application Delivery Controller (ADC) is a combination load balancer and application server that is a server placed between a firewall or router and a server farm providing Web services. An Application Delivery Controller is assigned a virtual IP address (VIP) that it maps to a pool of servers based on application specific criteria. An ADC is a combination network and application layer device. You also may come across ADCs referred to as a content switch, multilayer switch, or Web switch.

These vendors, among others, sell ADC systems:

- A10• Networks (http://www.a10networks.com/)
- Barracuda Networks (http://www.barracudanetworks.com/)
- Brocade Communication Systems (http://www.brocade.com/)
- Cisco Systems (http://www.cisco.com/)
- Citrix Systems (http://www.citrix.com/)

- F5 Networks (http://www f5.com/)
- Nortel Networks (http://www nortel.com/)
- Coyote Point Systems (http://www.coyotepoint.com/)
- Radware (http://www radware.com/)

An ADC is considered to be an advanced version of a load balancer as it not only can provide the features described in the previous paragraph, but it conditions content in order to lower the workload of the Web servers. Services provided by an ADC include data compression, content caching, server health monitoring, security, SSL offload and advanced routing based on current conditions. An ADC is considered to be an application accelerator, and the current products in this area are usually focused on two areas of technology: network optimization, and an application or framework optimization. For example, you may find ADC's that are tuned to accelerate ASP.NET or AJAX applications.

An architectural layer containing ADCs is described as an Application Delivery Network (ADN), and is considered to provide WAN optimization services. Often an ADN is comprised of a pair of redundant ADCs. The purpose of an ADN is to distribute content to resources based on application specific criteria. ADN provide a caching mechanism to reduce traffic, traffic prioritization and optimization, and other techniques. ADN began to be deployed on Content Delivery Networks (CDN) in the late 1990s, where it added the ability to optimize applications (application fluency) to those networks. Most of the ADC vendors offer commercial ADN solutions.

In addition to the ADC vendors in the list above, these are additional ADN vendors, among others:

- Akamai Technologies (http://www.akamai.com/)
- Blue Coat Systems (http://www.bluecoat.com/)
- CDNetworks (http://www.cdnetworks.com/)

- Crescendo Networks (http://www.crescendonetworks.com/)
- Expand Networks (http://www.expand.com/)
- Juniper Networks (http://www.juniper net/)

Google's cloud is a good example of the use of load balancing, so in the next section let's consider how Google handles the many requests that they get on a daily basis.

The Google cloud

Google Cloud Platform is a suite of public cloud computing services offered by Google. The platform includes a range of hosted services for compute, storage and application development that run on Google hardware. Google Cloud Platform services can be accessed by software developers, cloud administrators and other enterprise IT professionals over the public internet or through a dedicated network connection.

Overview of Google Cloud Platform offerings

Google Cloud Platform offers services for compute, storage, networking, big data, machine learning and the internet of things (IoT), as well as cloud management, security and developer tools. The core cloud computing products in Google Cloud Platform include:

- Google Compute Engine, which is an infrastructure-as-a-service (IaaS) offering that provides users with virtual machine instances for workload hosting.
- Google App Engine, which is a platform-as-a-service (PaaS) offering that gives software developers access to Google's scalable hosting. Developers can also use a software developer kit (SDK) to develop software products that run on App Engine.
- Google Cloud Storage, which is a cloud storage platform designed to store large, unstructured data sets. Google also offers database storage options, including Cloud Datastore for NoSQL nonrelational storage, Cloud SQL for MySQL fully relational storage and Google's native Cloud Bigtable database.
- Google Container Engine, which is a management and orchestration system for Dockercontainers that runs within Google's public cloud. Google Container Engine is based on the Google Kubernetes container orchestration engine.

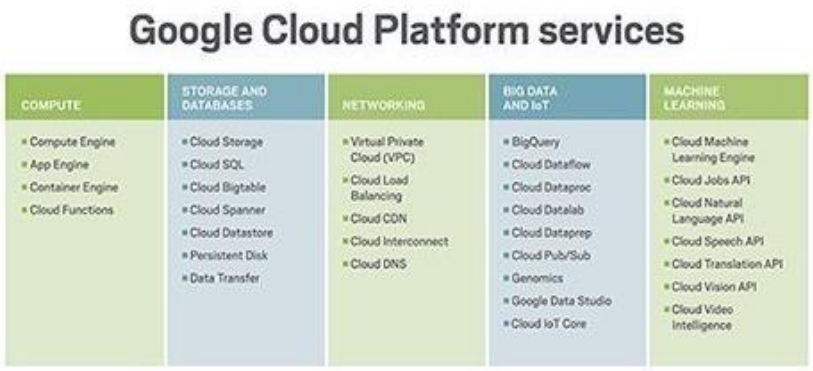
Google Cloud Platform offers application development and integration services. For example, Google Cloud Pub/Sub is a managed and real-time messaging service that allows messages to be exchanged between applications. In addition, Google Cloud Endpoints allows developers to create services based on RESTful APIs, and then make those services accessible to Apple iOS, Android and JavaScript clients. Other offerings include Anycast DNS servers, direct network interconnections, load balancing, monitoring and logging services.

Higher-level services

Google continues to add higher-level services, such as those related to big data and machine learning, to its cloud platform. Google big data services include those for data processing and analytics, such as Google BigQuery for SQL-like queries made against multi-terabyte data sets. In addition, Google Cloud Dataflow is a data processing service intended for analytics; extract, transform and load (ETL); and real-time computational projects. The platform also includes Google Cloud Dataproc, which offers Apache Spark and Hadoop services for big data processing.

For artificial intelligence (AI), Google offers its Cloud Machine Learning Engine, a managed service that enables users to build and train machine learning models. Various APIs are also available for the translation and analysis of speech, text, images and videos.Google also provides services for IoT, such as Google Cloud IoT Core, which is a series of managed services that enables users to consume and manage data from IoT devices.

The Google Cloud Platform suite of services is always evolving, and Google periodically introduces, changes or discontinues services based on user demand or competitive pressures. Google's main competitors in the public cloud computing market include Amazon Web Services (AWS) and Microsoft Azure.



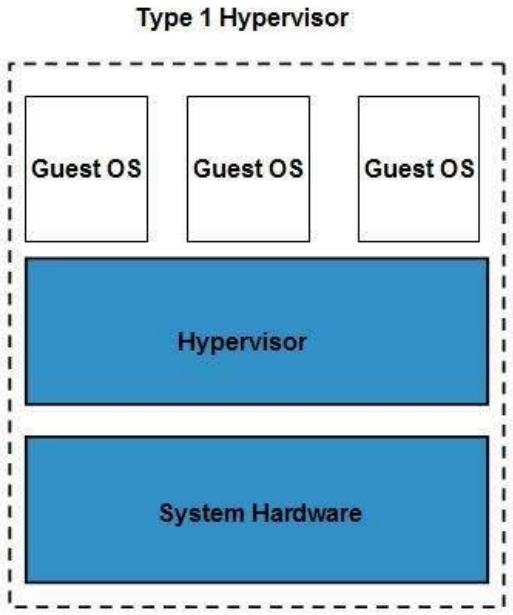
Google Cloud Platform pricing options

Like other public cloud offerings, most Google Cloud Platform services follow a pay-as-you-go model in which there are no upfront payments, and users only pay for the cloud resources they consume. Specific terms and rates, however, vary from service to service.

Hypervisor

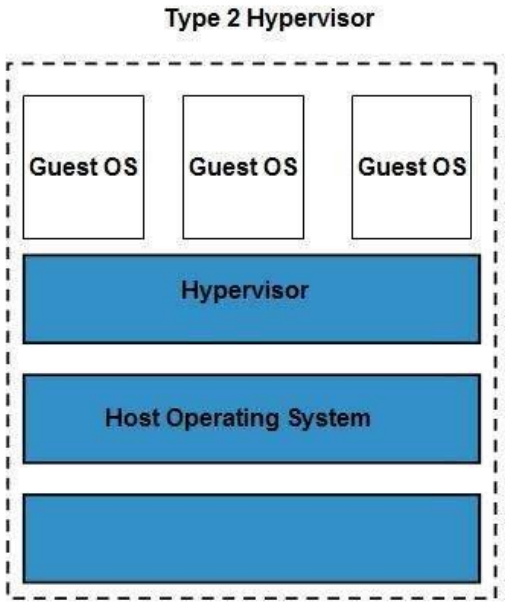
The **hypervisor** is a firmware or low-level program that acts as a Virtual Machine Manager. There are two types of hypervisor:

Type 1 hypervisor executes on bare system. LynxSecure, RTS Hypervisor, Oracle VM, Sun xVM Server, VirtualLogic VLX are examples of Type 1 hypervisor. The following diagram shows the Type 1 hypervisor.



The **type1 hypervisor** does not have any host operating system because they are installed on a bare system.

Type 2 hypervisor is a software interface that emulates the devices with which a system normally interacts. Containers, KVM, Microsoft Hyper V, VMWare Fusion, Virtual Server 2005 R2, Windows Virtual PC and **VMWare workstation 6.0** are examples of Type 2 hypervisor. The following diagram shows the Type 2 hypervisor.



Virtual machine types

Virtual machines are becoming more common with the evolution of virtualization technology. Virtual machines are often created to perform certain tasks that are different than tasks performed in a host environment.

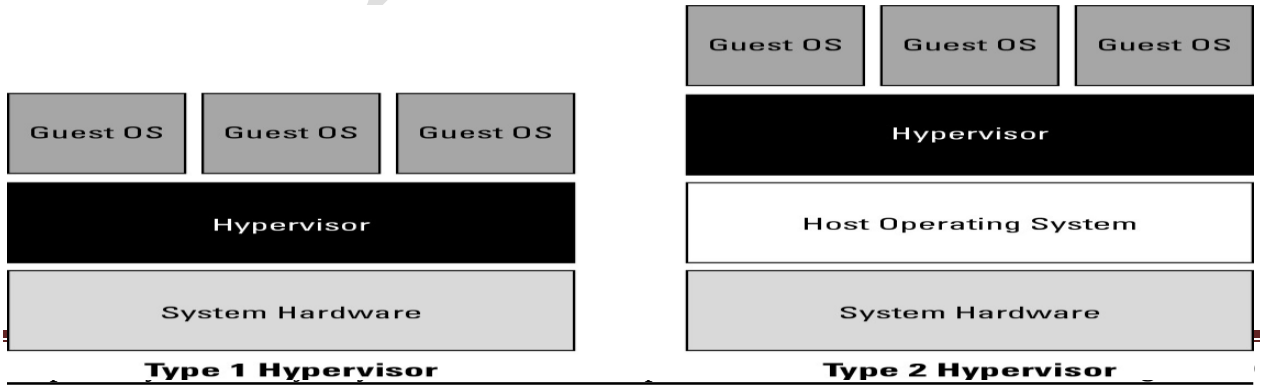
Virtual machines are implemented by software emulation methods or hardware virtualization techniques. A low-level program is required to provide system resource access to virtual machines, and this program is referred to as the hypervisor or Virtual Machine Monitor (VMM). A hypervisor running on bare metal is a Type 1 VM or native VM. Examples of Type 1 Virtual Machine Monitors are LynxSecure, RTS Hypervisor, Oracle VM, Sun xVM Server, VirtualLogix VLX, VMware ESX and ESXi, and Wind River VxWorks, among others. The operating system loaded into a virtual machine is referred to as the guest operating system, and there is no constraint on running the same guest on multiple VMs on a physical system. Type 1 VMs have no host operating system because they are installed on a bare system.

An operating system running on a Type 1 VM is a full virtualization because it is a complete simulation of the hardware that it is running on.

Some hypervisors are installed over an operating system and are referred to as Type 2 or hosted VM. Examples of Type 2 Virtual Machine Monitors are Containers, KVM, Microsoft Hyper V, Parallels Desktop for Mac, Wind River Simics, VMWare Fusion, Virtual Server 2005 R2, Xen, Windows Virtual PC, and VMware Workstation 6.0 and Server, among others. This is a very rich product category. Type 2 virtual machines are installed over a host operating system; for Microsoft Hyper-V, that operating system would be Windows Server. In the section that follows, the Xen hypervisor (which runs on top of a Linux host OS) is more fully described. Xen is used by Amazon Web Services to provide Amazon Machine Instances (AMIs).

The below figure shows the diagram of Type 1 and Type 2 hypervisors. On a Type 2 VM, a software interface is created that emulates the devices with which a system would normally interact. This abstraction is meant to place many I/O operations outside the virtual environment, which makes it both programmatically easier and more efficient to execute device I/O than it would be inside a virtual environment. This type of virtualization is sometimes referred to as *paravirtualization,* and it is found in hypervisors such as Microsoft's Hyper-V and Xen. It is the host operating system that is performing the I/O through a para-API.

VMware's vSphere cloud computing infrastructure model

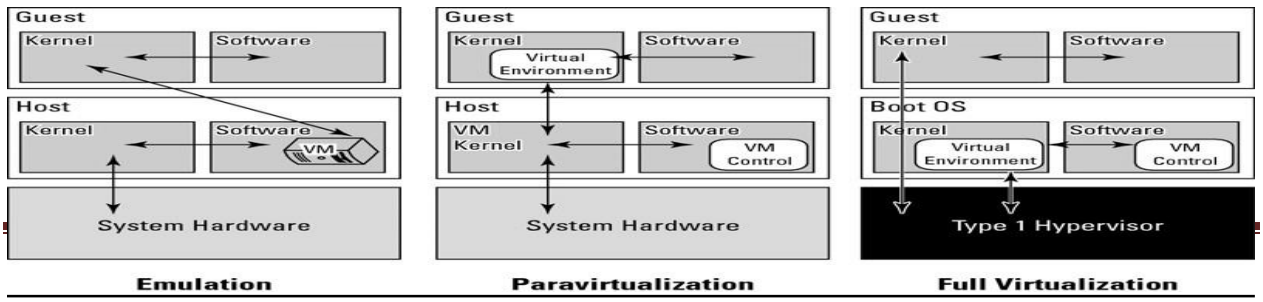


The below figure shows the difference between emulation, paravirtualization, and full virtualization. In emulation, the virtual machine simulates hardware, so it can be independent of the underlying system hardware. A guest operating system using emulation does not need to be modified in any way. Paravirtualization requires that the host operating system provide a virtual machine interface for the guest operating system and that the guest access hardware through that host VM. An operating system running as a guest on a paravirtualization system must be ported to work with the host interface. Finally, in a full virtualization scheme, the VM is installed as a Type 1 Hypervisor directly onto the hardware. All operating systems in full virtualization communicate directly with the VM hypervisor, so guest operating systems do not require any modification.

Guest operating systems in full virtualization systems are generally faster than other virtualization schemes. The Virtual Machine Interface (VMI) open standard (http://vmi ncsa.uiuc.edu/) that VMware has proposed is an example of a paravirtualization API. The latest version of VMI is 2.1, and it ships as a default installation with many versions of the Linux operating system.

Most folks run the Java Virtual Machine or Microsoft's .NET Framework VM (called the Common Language Runtime or CLR) on their computers. A process virtual machine instantiates when a command begins a process, the VM is created by an interpreter, the VM then executes the process, and finally the VM exits the system and is destroyed. During the time the VM exists, it runs as a high-level abstraction.

Emulation, paravirtualization, and full virtualization types



Applications running inside an application virtual machine are generally slow, but these programs are very popular because they provide portability, offer rich programming languages, come with many advanced features, and allow platform independence for their programs. Although many cloud computing applications provide process virtual machine applications, this type of abstraction isn't really suitable for building a large or high-performing cloud network, with one exception.

The exception is the process VMs that enable a class of parallel cluster computing applications. These applications are high-performance systems where the virtual machine is operating one process per cluster node, and the system maintains the necessary intra-application communications over the network interconnect. Examples of this type of system are the Parallel Virtual Machine (PVM; see http://www.csm.ornl.gov/pvm/pvm home.html) and the Message Passing Interface (MPI; see http://www mpi-forum.org/). Some people do not consider these application VMs to be true virtual machines, noting that these applications can still access the host operating system services on the specific system on which they are running. The emphasis on using these process VMs is in creating a high-performance networked supercomputer often out of heterogeneous systems, rather than on creating a ubiquitous utility resource that characterizes a cloud network.

Some operating systems such as Sun Solaris and IBM AIX 6.1 support a feature known as *operating system virtualization.* This type of virtualization creates virtual servers at the operating system or kernel level. Each virtual server is running in its own virtual environment (VE) as a virtual private server (VPS). Different operating systems use different names to describe these machine instances, each of which can support its own guest OS. However, unlike true virtual machines, VPS must all be running the same OS and the same version of that OS. Sun Solaris 10 uses VPS to create what is called Solaris Zones. With IBM AIX, the VPS is called a System Workload Partition (WPAR). This

type of virtualization allows for a dense collection of virtual machines with relatively low overhead. Operating system virtualization provides many of the benefits of virtualization previously noted in this section.

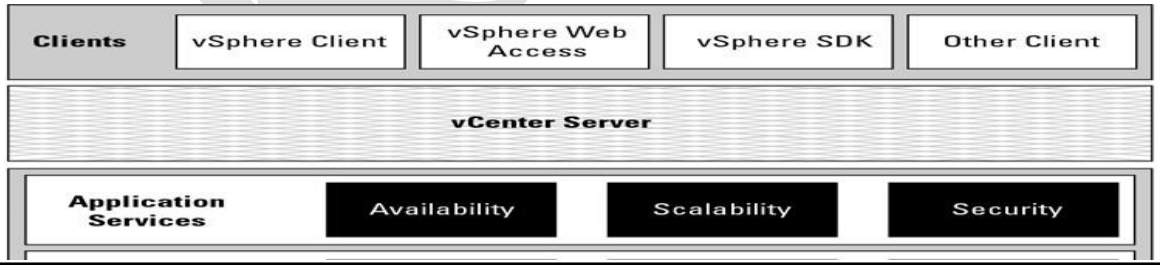
VMware vSphere

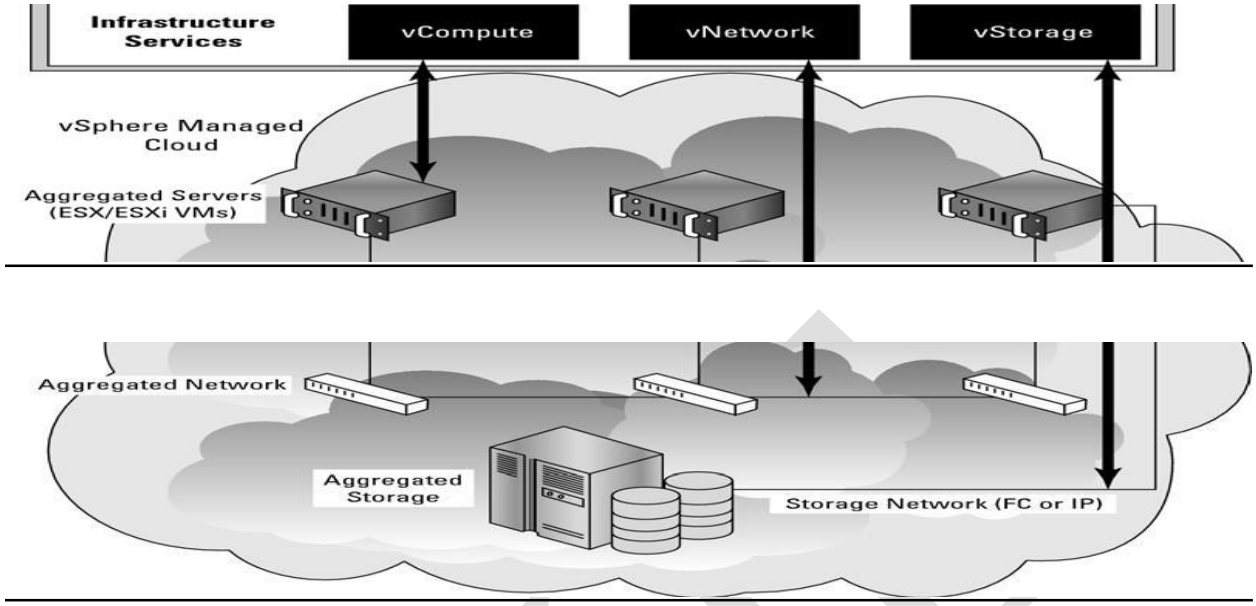
VMware vSphere is a management infrastructure framework that virtualizes system, storage, and networking hardware to create cloud computing infrastructures. vSphere is the branding for a set of management tools and a set of products previously labeled VMware Infrastructure. vSphere provides a set of services that applications can use to access cloud resources, including these:

- **VMware vCompute:** A service that aggregates servers into an assignable pool
- **VMware vStorage:** A service that aggregates storage resources into an assignable pool
- **VMware vNetwork:** A service that creates and manages virtual network interfaces
- **Application services:** Such as HA (High Availability) and Fault Tolerance
- **vCenter Server:** A provisioning, management, and monitoring console for VMware cloud infrastructures.

The figure below shows an architectural diagram of a vSphere cloud infrastructure.

VMware's vSphere cloud computing infrastructure model





A vSphere cloud is a pure infrastructure play. The virtualization layer that abstracts processing, memory, and storage uses the VMware ESX or ESXi virtualization server. ESX is a Type 1 hypervisor; it installs over bare metal (a clean system) using a Linux kernel to boot and installs the vmkernel hypervisor (virtualization kernel and support files). When the system is rebooted, the vmkernel loads first, and then the Linux kernel becomes the first guest operating system to run as a virtual machine on the system and contains the service console.

VMware is a very highly developed infrastructure and the current leader in this industry. A number of important add-on products are available for cloud computing applications. These are among the more notable products:

- **Virtual Machine File System (VMFS):** A high-performance cluster file system for an ESX/ESXi cluster.
- **VMotion:** A service that allows for the migration of a virtual machine from one physical server to another physical server while the virtual server runs continuously and without any interruption of ongoing transactions. The ability to live migrate virtual machines is considered to be a technological tour de force and a differentiator from other

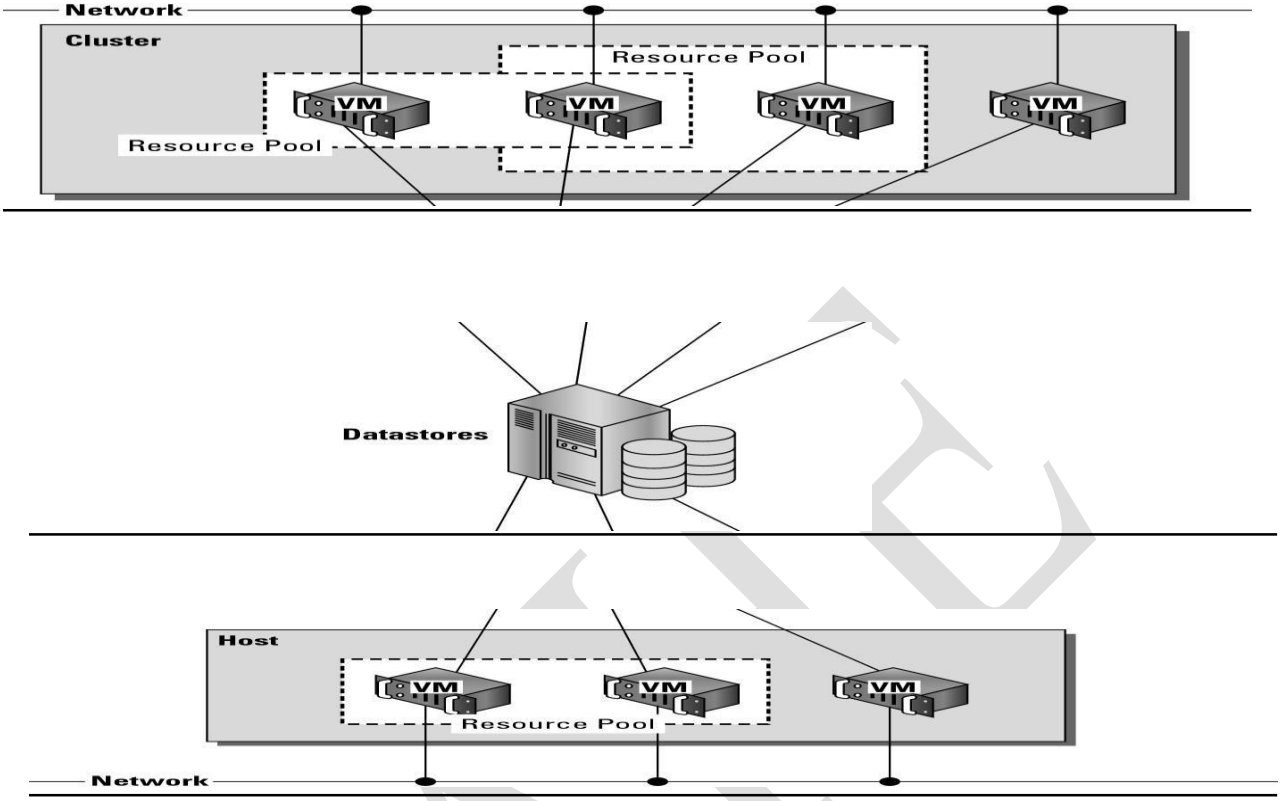
virtual machine system vendors.

- **Storage VMotion:** A product that can migrate files from one datastore to another datastore while the virtual machine that uses the datastore continues to run.
- **Virtual SMP:** A feature that allows a virtual machine to run on two or more physical processors at the same time.
- **Distributed Resource Scheduler (DRS):** A system for provisioning virtual machines and load balancing processing resources dynamically across the different physical systems that are in use. A part of the DRS called the distributed power management (DPM) module can manage the power consumption of systems.
- **vNetwork Distributed Switch (DVS):** A capability to maintain a network runtime state for virtual machines as they are migrated from one physical system to another. DVS also monitors network connections, provides firewall services, and enables the use of third-party switches such as the Cisco Nexus 1000V to manage virtual networks.

You can get a better sense of how the different resources are allocated by vSphere into a virtual set of components by examining the figure. Physical computers can be standalone hosts or a set of clustered systems. In either case, a set of virtual machines can be created that is part of a single physical system or spans two or more physical systems.

You can define a group of VMs as a Resource Pool (RP) and, by doing so, manage those virtual machines as a single object with a single policy. Resource Pools can be placed into a hierarchy or nested and can inherit properties of their parent RP. As more hosts or cluster nodes are added or removed, vSphere can dynamically adjust the provisioning of VMs to accommodate the policy in place. This fine tuning of pooled resources is required to accommodate the needs of cloud computing networks.

Virtual infrastructure elements



The datastore shown at the center of Figure 5.4 is a shared storage resource. These storage resources can be either Direct Attached Storage (DAS) of a server using SCSI, SAS, or SATA connections, Fibre Channel disk arrays/SANs, iSCSI disk arrays/SANs, or Network Attached Storage (NAS) disk arrays. Although the lines drawn between the datastore and different VMs indicate a direct connection, with the exception of DAS, the other storage types are shared storage solutions.

Storage virtualization is most commonly achieved through a mapping mechanism where a logical storage address is translated into a physical storage address. Block-based storage such as those used in SANs use a feature called a Logical Unit Identifier (LUN**)** with specific addresses stored in the form of an offset called the Logical Block Address (LBA). The address space mapping then maps the address of the logical or virtual disk (vdisk) to the logical unit on a storage controller. Storage virtualization may be done in software or in hardware, and it

allows requests for virtualized storage to be redirected as needed.

Similarly, network virtualization abstracts networking hardware and software into a virtual network that can be managed. A virtual network can create virtual network interfaces (VNICs) or virtual LANs (VLANS) and can be managed by a hypervisor, operating system, or external management console. In a virtualized infrastructure such as the one presented in this section, internal network virtualization is occurring and the hypervisor interacts with networking hardware to create a pseudo-network interface. External network virtualization can be done using network switches and VLAN software. The key feature that makes virtual infrastructure so appealing for organizations implementing a cloud computing solution is flexibility. Instantiating a virtual machine is a very fast process, typically only a few seconds in length. You can make machine images of systems in the configuration that you want to deploy or take snapshots of working virtual machines. These images can be brought on-line as needed.

Machine Imaging

A system image makes a copy or a clone of the entire computer system inside a single container such as a file. The system imaging program is used to make this image and can be used later to restore a system image. Some imaging programs can take snapshots of systems, and most allow you to view the files contained in the image and do partial restores.

A prominent example of a system image and how it can be used in cloud computing architectures is the Amazon Machine Image (AMI) used by Amazon Web Services to store copies of a virtual machine. An AMI is a file system image that contains an operating system, all appropriate device drivers, and any applications and state information that the working virtual machine would have.

When you subscribe to AWS, you can choose to use one of its hundreds of canned AMIs or to create a custom system and capture that system's image to an AMI. An AMI can be

for public use under a free distribution license, for pay-per-use with operating systems such as Windows, or shared by an EC2 user with other users who are given the privilege of access.

The AMI file system is not a standard bit-for-bit image of a system that is common to many disk imaging programs. AMI omits the kernel image and stores a pointer to a particular kernel that is part of the AWS kernel library. Among the choices are Red Hat Linux, Ubuntu, Microsoft Windows, Solaris, and others. Files in AMI are compressed and encrypted, and an XML file is written that describes the AMI archive. AMIs are typically stored in your Amazon S3 (Simple Storage System) buckets as a set of 10MB chunks.

Machine images are sometimes referred to as "virtual appliances"—systems that are meant to run on virtualization platforms. AWS EC2 runs on the Xen hypervisor, for example. The term *virtual appliance* is meant to differentiate the software image from an operating virtual machine. The system image contains the operating system and applications that create an environment. Most virtual appliances are used to run a single application and are configurable from a Web page.

Virtual appliances are a relatively new paradigm for application deployment, and cloud computing is the major reason for the interest in them and for their adoption. This area of WAN application portability and deployment, and of WAN optimization of an application based on demand, is one with many new participants. Certeon (http://www.certeon.com/), Expand Networks (http://www.expand.com/), and Replify (http://www replify.com/) are three vendors offering optimization appliances for VMware's infrastructure.

Porting Applications

Cloud computing applications have the ability to run on virtual systems and for these systems to be moved as needed to respond to demand. Systems (VMs running applications), storage, and network assets can all be virtualized and have sufficient flexibility to give acceptable distributed WAN application performance. Developers who

write software to run in the cloud will undoubtedly want the ability to port their applications from one cloud vendor to another, but that is a much more difficult proposition. Cloud computing is a relatively new area of technology, and the major vendors have technologies that don't interoperate with one another.

The Simple Cloud API

If you build an application on a platform such as Microsoft Azure, porting that application to Amazon Web Services or GoogleApps may be difficult, if not impossible. In an effort to create an interoperability standard, Zend Technologies has started an open source initiative to create a common application program interface that will allow applications to be portable. The initiative is called the Simple API for Cloud Application Services (http://www.simplecloud.org/), and the effort has drawn interest from several major cloud computing companies. Among the founding supporters are IBM, Microsoft, Nivanix, Rackspace, and GoGrid.

Simple Cloud API has as its goal a set of common interfaces for:

- **File Storage Services:** Currently Amazon S3, Windows Azure Blob Storage, Nirvanix, and Local storage is supported by the Storage API. There are plans to extend this API to Rackspace Cloud Files and GoGrid Cloud Storage.
- **Document Storage Services:** Amazon SimpleDB and Windows Azure Table Storage are currently supported. Local document storage is planned.
- **Simple Queue Services:** Amazon SQS, Windows Azure Queue Storage, and Local queue services are supported.

AppZero Virtual Application Appliance

Applications that run in datacenters are captive to the operating systems and hardware platforms that they run on. Many datacenters are a veritable Noah's Ark of computing. So moving an application from one platform to another isn't nearly as simple as moving a machine image from one system to another. The situation is further complicated by the fact that applications are tightly coupled with the operating systems on which they run. An

application running on Windows, for example, isn't isolated from other applications. When the application loads, it often loads or uses different Dynamic Link Libraries (DLL), and it is through the sharing or modification of DLLs that Windows applications get themselves in trouble. Further modifications include modifying the registry during installation. These factors make it difficult to port applications from one platform to another without lots of careful work. If you are a Platform as a Service (PaaS) application developer, you are packaging a complete software stack that includes not only your application, but the operating system and application logic and rules as well. Vendor lock-in for you application is assured.

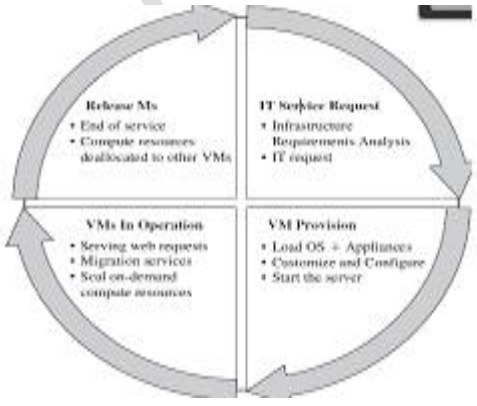
The ability to run an application from whatever platform you want is not one of the characteristics of cloud computing, but you can imagine that it is a very attractive proposition. While the Simple Cloud API is useful for applications written in PHP, other methods may be needed to make applications easily portable. One company working on this problem is AppZero (http://www.appzero.com/), and its solution is called the Virtual Application Appliance (VAA).

The AppZero solution creates a virtual application appliance as an architectural layer between the Windows or the UNIX operating system and applications. The virtualization layer serves as the mediator for file I/O, memory I/O, and application calls and response to DLLs, which has the effect of sandboxing the application. The running application in AppZero changes none of the registry entries or any of the files on the Windows Server.

VAA creates a container that encapsulates the application and all the application's dependencies within a set of files; it is essentially an Application Image for a specific OS. Dependencies include DLL, service settings, necessary configuration files, registry entries, and machine and network settings. This container forms an installable server- side application stack that can be run after installation, but has no impact on the underlying operating system. VAAs are created using the AppZero Creator wizard, managed with the

AppZero Admin tool, and may be installed using the AppZero Director, which creates a VAA runtime application. If desired, an application called AppZero Dissolve removes the VAA virtualization layer from the encapsulated application and installs that application directly into the operating system.

Installations can be done over the network after the AppZero application appliance is installed. Therefore, with this system, you could run applications on the same Windows Server and eliminate one application from interfering with another; applications would be much more easily ported from one Windows system to another. AppZero's approach provides the necessary abstraction layer that frees an application from its platform dependence. As shown in Figure 5.3, the cycle starts by a request delivered to the IT department, stating the requirement for creating a new server for a particular service. This request is being processed by the IT administration to start seeing the servers' resource pool, matching these resources with the requirements, and starting the provision of the needed virtual machine. Once it is provisioned and started, it is ready to provide the required service according to an SLA, or a time period after which the virtual is being released; and free resources, in this case, won't be needed.



VM Provisioning Process Provisioning a virtual machine or server can be explained and illustrated as in Figure.

Steps to Provision VM.

Here, we describe the common and normal steps of provisioning a virtual server:

- Firstly, you need to select a server from a pool of available servers (physical servers with enough capacity) along with the appropriate OS template you need to provision the virtual machine.
- Secondly, you need to load the appropriate software (operating system you selected in the previous step, device drivers, middleware, and the needed applications for the service required).
- Thirdly, you need to customize and configure the machine (e.g., IP address, Gateway) to configure an associated network and storage resources.
- Finally, the virtual server is ready to start with its newly loaded software.

Typically, these are the tasks required or being performed by an IT or a data center's specialist to provision a particular virtual machine. To summarize, server provisioning is defining server's configuration based on the organization requirements, a hardware, and software component (processor, RAM, storage, networking, operating system, applications, etc.). Normally, virtual machines can be provisioned by manually installing an operating system, by using a preconfigured VM template, by cloning an existing VM, or by importing a physical server or a virtual server from another hosting platform.

Physical servers can also be virtualized and provisioned using P2V (physical to virtual) tools and techniques (e.g., virt-p2v). After creating a virtual machine by virtualizing a physical server, or by building a new virtual server in the virtual environment, a template can be created out of it. Most virtualization management vendors (VMware, XenServer, etc.) provide the data center's administration with the ability to do such tasks in an easy way. Provisioning from a template is an invaluable feature, because it reduces the time

required to create a new virtual machine. Administrators can create different templates for different purposes. For example, you can create a Windows 2003 Server template for the finance department, or a Red Hat Linux template for the engineering department. This enables the administrator to quickly provision a correctly configured virtual server on demand. This ease and flexibility bring with them the problem of virtual machine's sprawl, where virtual machines are provisioned so rapidly that documenting and managing the virtual machine's life cycle become a challenge .

POSSIBLE QUESTIONS

6 MARKS

1. Explain about virtualization technologies in cloud.
2. Describe the role of advanced load balancing in cloud computing.
3. What are load balancing? Elaborate the advantages of Cloud Load Balancing.
4. What are hypervisors? Explain the different types of hypervisors.
5. Describe about Virtual machine types.
6. Give a clear explanation on AppZero Virtual Application Appliance.
7. What are Google clouds? Explain its role in cloud services.
8. Explain the role of VMware vSphere in cloud.
9. Explain the role of machine images in cloud.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

ONE MARK QUESTIONS

DEPARTMENT OF CS, CA & IT

STAFF NAME: Dr.S.MANJU PRIYA **SUBJECT NAME: CLOUD COMPUTING**

SUB.CODE: 18CSP104 **UNIT III** **SEMESTER: I**

S.NO	Question	Choice1	Choice2	Choice3	Choice4	Ans
1	When you use cloud computing, you are accessing pooled resources using a technique called	Load Balancing	Virtualization	Hypervisors	Infrastructure	Virtualization
2	____are the services provided by cloud computing are based on Internet protocols and formats.	Internet delivery	Shared servic	service-based architecture	Internet storage	Internet delivery
3	The technology used to distribute service requests to resources is referred to as _____	load balancing	Virtualization	Storage resources	Application delivery controller	load balancing
4	ADC stands for_____	Application Direct	Application Delivery	Application Delivery	Application Direct Cloud	Application Delivery

5	_____ is the single most heavily visited site on the Internet	MSN	Rediff	Yahoo	Google	Google
6	Google supports over_____country specific versions of the Google index	40	10	20	30	30
7	When you initiate a Google search, your query is sent to a _____ server	Google App	DNS	Cloud	Amazon	DNS
8	_____crawls the Web and collects document information	Google autorun	Google Trojan	Google virus	Google Bot	Google Bot
9	Google doesn't use _____ virtualization	Middleware	Software	Hardware	Malware	Hardware
10	_____virtualizes systems and resources by mapping a logical address to a physical address	Infrastructure	Load balancing	Hypervisors	Virtualization	Load balancing
11	VMI stands for _____	Virtual Machine	Virtual Machine	Virtual Model	Virtual Machine	Virtual Machine
12	_____ provides a set of services that applications can use to access cloud resources	vSphere	VMware	vStorage	vCenter	vSphere
13	ADN stands for_____	Application Development	Amazon Delivery	Application Delivery	Amazon Development	Application Delivery
14	_____provide a caching mechanism to reduce traffic, traffic prioritization and optimization, and other techniques	Amazon Delivery Network	Application Delivery Network	Amazon Development Controller	Application Delivery Controller	Application Delivery Network
15	A_____ can contain thousands of servers.	Google servers	Google cluster	Google app engine	Google clients	Google cluster
16	Google doesn't use _____virtualization	Software	Physical	Hardware	Logical	Hardware

17	Resources are pooled in order to create greater efficiencies is called_____	Shared services	Resource sharing	Resource pooling	Scalable services	Shared services
18	_____can be enforced using session data stored in a database and replicated across multiple load balancers	Consistent	Dialog control	Persistence	Synchronization	Persistence
19	_____ algorithm examines the words and the relationships of one word to another.	Load Balancing	Google	Amazon	Round Robin	Google
20	Word relationships are mapped against the main index to create a list of documents called_____	Query Index	Index	Inverted index	Hash Table	Inverted index
21	_____provide the capability of running multiple machine instances, each with their own operating system	Machines	Shared machines	Local machines	Virtual machines	Virtual machines
22	The downside of _____technologies is that having resources indirectly addressed means there is some	Shared machine	Local machine	Virtual machine	Machine	Virtual machine
23	VMM stands for	Virtual Machine	Virtual Model	Virtual Monitor	Virtual Monitor model	Virtual Machine
24	Some operating systems such as Sun Solaris and IBM AIX 6.1 support a feature known as	OS Virtualization	Cloud OS	Simple OS	Linux OS	OS Virtualization
25	Amazon SimpleDB and Windows Azure Table Storage are _____	File Storage Services	Document Storage	Simple Queue	Data Storage Services	Document Storage
26	DLL stands for_____	Dynamic Link Libraries	Dynamic Library Link	Data Link Layer	Data Link Libraries	Dynamic Link Libraries
27	VAA stands for_____	Virtual Amazon	Virtual Appliance	Virtual Application	Virtual Amazon	Virtual Application
28	VAAs are created using the _____ Creator wizard	Simple cloud App	AppZero	vsphere app	Hypervisor app	AppZero

29	The latest version of VMI is _____	2.1	2.2	1.9	2	2.1
30	_____ is a management infrastructure framework that virtualizes system, storage, and networking hardware to create cloud computing infrastructures	VMware	VMware vSphere	vSphere	Vsphere vNetwork	VMware vSphere
31	C2D stands for_____	Cloud to Datacenter	Cloud to Data	Cloud to Development	Cloud to Distributor	Cloud to Datacenter
32	An _____ is considered to be an advanced version of a load balancer	ADN	ADC	ACD	AND	ADC
33	_____is a good example of the use of load balancing	Web Server	Virtualization	Hypervisors	Google's cloud	Google's cloud
34	_____is dynamic in that the mapping can be assigned based on rapidly changing conditions	Load Balancing	Hypervisors	Virtualization	Appzero	Virtualization
35	The _____ takes the result of a query and composes the Web page from that result.	URL	Web server	IP address	Web client	Web server
36	A system virtual machine is also known as_____	A hardware virtual	a software virtual	a cloud machine	a multipurpose	A hardware virtual machine
37	_____ is found in hypervisors such as Microsoft's Hyper-V and Xen	Full virtualization	Emulation	Vmware	Para virtualization	Para virtualization
38	vNetwork is a_____	Application service	Infrastructure service	Clients	Network service	Infrastructure service
39	Security is a_____	Clients	Infrastructure service	Application service	Network service	Application service
40	_____is a very highly developed infrastructure and the current leader in this industry	vSphere	VMware	vStorage	vCenter	VMware

41	A service that allows for the migration of a virtual machine from one physical server to another physical server while the virtual server runs continuously and without any interruption of ongoing transactions is called_____	Virtual Machine File System	VMotion	Virtual SMP	Vmare	VMotion
42	A feature that allows a virtual machine to run on two or more physical processors at the same time is _____	Virtual SMP	Distributed Resource Scheduler	Virtual Machine File System	Virtual Network	Virtual SMP
43	Block-based storage such as those used in sANs use a feature called a_____	Logical Unit Identifier	Logical Block	Logical Unit Block	Logical Block Identifier	Logical Unit Identifier
44	An_____ is a file system image that contains an operating system, all appropriate device drivers, and any applications and state information that the working virtual machine would have.	ADN	ADC	AWS	AMI	AMI
45	Without _____, cloud computing would very difficult to manage	Virtualization	Hypervisors	Load balancing	VMware	Load balancing
46	A service that allows for the migration of a virtual machine from one physical server to another physical server is called_____	Virtual Machine File System	VMotion	VMware	vSphere	VMotion
47	A _____ makes a copy or a clone of the entire computer system inside a single container such as a file	System image	Virtual image	Canned image	Stored image	System image
48	Machine images are sometimes referred to as _____	Virtual images	Virtual appliances	Canned images	Stored images	Virtual appliances

49	Developers who write software to run in the cloud will undoubtedly want the ability to _____ their applications from one cloud vendor to another	Port	Compress	Encrypt	Translate	Port
50	_____has started an open source initiative to create a common application program interface that will allow applications to be portable.	Microsoft	GoGrid	Zend Technologies	Rackspace	Zend Technologies
51	There are _____ key attributes of cloud computing	4	5	6	7	5
52	Services are billed on a usage basis are called_____	Metered usage	Billed usage	Daily usage	Rented usage	Metered usage
53	Which of the following is not an Network interfaces and services _____	FTP	DNS	Cloud computing	HTTP	Cloud computing
54	_____ is not an OS.	Sun Solaris	IBM AIX 6.1	Linux	vNetwork	vNetwork
55	A _____ can create virtual network interfaces (VNICs) or virtual LANs (VLANS)	Wide Area Network	Distributed Network	Virtual Network	Local Area Network	Virtual Network

UNIT-IV

Cloud Information Security Objectives -Confidentiality Integrity and Availability -Cloud Security Services - Relevant Cloud Security Design Principles -Cloud Computing Risk Issues -The CIA Triad - Privacy and Compliance Risks -Threats to Infrastructure Data and Access Control -Cloud

Cloud Information Security Objectives

Developing secure software is based on applying the secure software design principles that form the fundamental basis for software assurance. Software assurance has been given many definitions, and it is important to understand the concept. The Software Security Assurance Report defines software assurance as "the basis for gaining justifiable confidence that software will consistently exhibit all properties required to ensure that the software, in operation, will continue to operate dependably despite the presence of sponsored (intentional) faults. In practical terms, such software must be able to resist most attacks, tolerate as many as possible of those attacks it cannot resist, and contain the damage and recover to a normal level of operation as soon as possible after any attacks it is unable to resist or tolerate."

The U.S. Department of Defense (DoD) Software Assurance Initiative defines software assurance as "the level of confidence that software functions as intended and is free of vulnerabilities, either intentionally or unintentionally designed or inserted as part of the software." The Data and Analysis Center for Software (DACS) requires that software must exhibit the following three properties to be considered secure:

Dependability — Software that executes predictably and operates correctly under a variety of conditions, including when under attack or running on a malicious host

Trustworthiness — Software that contains a minimum number ofv vulnerabilities or no vulnerabilities or weaknesses that could sabotage the software's dependability. It must also be resistant to malicious logic.

Survivability (Resilience) — Software that is resistant to or tolerant ofv attacks and has the ability to recover as quickly as possible with as little harm as possible

Seven complementary principles that support information assurance are confidentiality, integrity, availability, authentication, authorization, auditing, and accountability. These concepts are summarized in the following sections.

Confidentiality Integrity and Availability

Confidentiality, integrity, and availability are sometimes known as the CIA triad of information system security, and are important pillars of cloud software assurance.

Confidentiality

Confidentiality refers to the prevention of intentional or unintentional unauthorized disclosure of information. Confidentiality in cloud systems is related to the areas of intellectual property rights, covert channels, traffic analysis, encryption, and inference.

Intellectual property rights — Intellectual property (IP) includes inventions, designs, and artistic, musical, and literary works. Rights to intellectual property are covered by copyright laws, which protect creations of the mind, and patents, which are granted for new inventions.

Covert channels — A covert channel is an unauthorized and unintendedv communication path that enables the exchange of information. Covert channels can be accomplished through timing of messages or inappropriate use of storage mechanisms.

Traffic analysis — Traffic analysis is a form of confidentiality breach that canv be accomplished by analyzing the volume, rate, source, and destination of message traffic, even if it is encrypted. Increased message activity and high bursts of traffic can indicate a major event is occurring. Countermeasures to traffic analysis include maintaining a near-constant rate of message traffic and disguising the source and destination locations of the traffic.

Encryption — Encryption involves scrambling messages so that they cannotv be read by an unauthorized entity, even if they are intercepted. The amount of effort (work factor) required to decrypt the message is a function of the strength of the encryption key and the robustness and quality of the encryption algorithm.

Inference — Inference is usually associated with database security. Inferencev is the ability of an entity to use and correlate information protected

at one level of security to uncover information that is protected at a higher security level.

Integrity

The concept of cloud information integrity requires that the following three principles are met:

- Modifications are not made to data by unauthorized personnel or processes.
- Unauthorized modifications are not made to data by authorized personnelv or processes.
- The data is internally and externally consistent — in other words, the internal information is consistent both among all sub-entities and with the real-world, external situation.

Availability

Availability ensures the reliable and timely access to cloud data or cloud computing resources by the appropriate personnel. Availability guarantees that the systems are functioning properly when needed. In addition, this concept guarantees that the security services of the cloud system are in working order. A denial-of-service attack is an example of a threat against availability. The reverse of confidentiality, integrity, and availability is disclosure, alteration, and destruction (DAD).

Cloud Security Services

Additional factors that directly affect cloud software assurance include authentication, authorization, auditing, and accountability, as summarized in the following sections.

Authentication

Authentication is the testing or reconciliation of evidence of a user's identity. It establishes the user's identity and ensures that users are who they claim to be. For example, a user presents an identity (user ID) to a computer login screen and then has to provide a password. The computer system authenticates the user by verifying that the password corresponds to the individual presenting the ID.

Authorization

Authorization refers to rights and privileges granted to an individual or process that enable access to computer resources and information assets. Once a user's identity and authentication are established, authorization levels determine the extent of system rights a user can hold.

Auditing

To maintain operational assurance, organizations use two basic methods: system audits and monitoring. These methods can be employed by the cloud customer, the cloud provider, or both, depending on asset architecture and deployment.

- A system audit is a one-time or periodic event to evaluate security.

- Monitoring refers to an ongoing activity that examines either the systemv or the users, such as intrusion detection.

Information technology (IT) auditors are often divided into two types: internal and external. Internal auditors typically work for a given organization, whereas external auditors do not. External auditors are often certified public accountants (CPAs) or other audit professionals who are hired to perform an independent audit of an organization's financial statements. Internal auditors usually have a much broader mandate than external auditors, such as checking for compliance and standards of due care, auditing operational cost efficiencies, and recommending the appropriate controls.

IT auditors typically audit the following functions:

- System and transaction controls
- Systems development standards
- Backup controls
- Data library procedures
- Data center security
- Contingency plans

In addition, IT auditors might recommend improvements to controls, and they often participate in a system's development process to help an organization avoid costly reengineering after the system's implementation.

An audit trail or log is a set of records that collectively provide documentary evidence of processing, used to aid in tracing from original transactions

forward to related records and reports, and/or backward from records and reports to their component source transactions. Audit trails may be limited to specific events or they may encompass all of the activities on a system.

Audit logs should record the following:

- The transaction's date and time
- At which terminal the transaction was processed
- Various security events relating to the transaction

In addition, an auditor should examine the audit logs for the following:

- Amendments to production jobs
- Production job returns
- Computer operator practices
- All commands directly initiated by the user
- All identification and authentication attempts
- Files and resources accessed

Accountability

Accountability is the ability to determine the actions and behaviors of a single individual within a cloud system and to identify that particular individual. Audit trails and logs support accountability and can be used to conduct postmortem studies in order to analyze historical events and the individuals or processes associated with those events. Accountability is related to the concept of nonrepudiation, wherein an individual cannot successfully deny the performance of an action.

Relevant Cloud Security Design Principles

Historically, computer software was not written with security in mind; but because of the increasing frequency and sophistication of malicious attacks against information systems, modern software design methodologies include security as a primary objective. With cloud computing systems seeking to meet multiple objectives, such as cost, performance, reliability, maintainability, and security, trade-offs have to be made. A completely secure system will exhibit poor performance characteristics or might not function at all.

Technically competent hackers can usually find a way to break into a computer system, given enough time and resources. The goal is to have a system that is secure enough for everyday use while exhibiting reasonable performance and reliability characteristics.

In 1974 paper that is still relevant today, Saltzer and Schroeder of the University of Virginia addressed the protection of information stored in a computer system by focusing on hardware and software issues that are necessary to support information protection. The paper presented the following 11 security design principles:

- Least privilege
- Separation of duties
- Defense in depth
- Fail safe
- Economy of mechanism

- Complete mediation
- Open design
- Least common mechanism
- Psychological acceptability
- Weakest link
- Leveraging existing components

The fundamental characteristics of these principles are summarized in the following sections.

Least Privilege

The principle of *least privilege* maintains that an individual, process, or other type of entity should be given the minimum privileges and resources for the minimum period of time required to complete a task. This approach reduces the opportunity for unauthorized access to sensitive information.

Separation of Duties

Separation of duties requires that completion of a specified sensitive activity or access to sensitive objects is dependent on the satisfaction of a plurality of condi- tions. For example, an authorization would require signatures of more than one individual, or the arming of a weapons system would require two individuals with different keys. Thus, separation of duties forces collusion among entities in order to compromise the system.

Defense in Depth

Defense in depth is the application of multiple layers of protection wherein a subsequent layer will provide protection if a previous layer is breached.

The Information Assurance Technical Framework Forum (IATFF), an organization sponsored by the National Security Agency (NSA), has produced a document titled the "Information Assurance Technical Framework" (IATF) that provides excellent guidance on the concepts of defense in depth.

The IATFF encourages and supports technical interchanges on the topic of information assurance among U.S. industry, U.S. academic institutions, and U.S. government agencies.

The IATF document 3.1 stresses the importance of the *people* involved, the *operations* required, and the *technology* needed to provide information assurance and to meet the organization's mission.

The defense-in-depth strategy as defined in IATF document 3.1 promotes application of the following information assurance principles:

- **Defense in multiple places** — Information protection mechanisms placed in a number of locations to protect against internal and external threats
- **Layered defenses** — A plurality of information protection and detection mechanisms employed so that an adversary or threat must negotiate a series of barriers to gain access to critical information

- **Security robustness** — An estimate of the robustness of information assurance elements based on the value of the information system component to be protected and the anticipated threats

- **Deploy KMI/PKI** — Use of robust key management infrastructures (KMI) and public key infrastructures (PKI)

- **Deploy intrusion detection systems** — Application of intrusion detection mechanisms to detect intrusions, evaluate information, examine results, and, if necessary, take action

Fail Safe

Fail safe means that if a cloud system fails it should fail to a state in which the security of the system and its data are not compromised. One implementation of this philosophy would be to make a system default to a state in which a user or process is denied access to the system. A complementary rule would be to ensure that when the system recovers, it should recover to a secure state and not permit unauthorized access to sensitive information. This approach is based on using permissions instead of exclusions.

In the situation where system recovery is not done automatically, the failed system should permit access only by the system administrator and not by other users, until security controls are reestablished.

Economy of Mechanism

Economy of mechanism promotes simple and comprehensible design and imple- mentation of protection mechanisms, so that unintended access paths

do not exist or can be readily identified and eliminated.

Complete Mediation

In complete meditation, every request by a subject to access an object in a computer system must undergo a valid and effective authorization procedure. This mediation must not be suspended or become capable of being bypassed, even when the information system is being initialized, undergoing shut-down, being restarted, or is in maintenance mode. Complete mediation entails the following:

1. Identification of the entity making the access request
2. Verification that the request has not changed since its initiation
3. Application of the appropriate authorization procedures
4. Reexamination of previously authorized requests by the same entity

Open Design

There has always been an ongoing discussion about the merits and strengths of security designs that are kept secret versus designs that are open to scrutiny and evaluation by the community at large. A good example is an encryption system. Some feel that keeping the encryption algorithm secret makes it more difficult to break. The opposing philosophy believes that exposing the algo-rithm to review and study by experts at large while keeping the encryption key secret leads to a stronger algorithm because the experts have a higher probability of discovering weaknesses in it. In general, the latter approach has proven more effective, except in the case of organizations such as the National

Security Agency (NSA), which employs some of the world's best cryptographers and mathematicians.

For most purposes, an open-access cloud system design that has been evaluated and tested by a myriad of experts provides a more secure authentication method than one that has not been widely assessed. Security of such mechanisms depends on protecting passwords or keys.

Least Common Mechanism

This principle states that a minimum number of protection mechanisms should be common to multiple users, as shared access paths can be sources of unauthorized information exchange. Shared access paths that pro- vide unintentional data transfers are known as *covert channels*. Thus, the *least common mechanism* promotes the least possible sharing of common security mechanisms.

Psychological Acceptability

Psychological acceptability refers to the ease of use and intuitiveness of the user interface that controls and interacts with the cloud access control mechanisms. Users must be able to understand the user interface and use it without having to interpret complex instructions.

Cloud Computing Risk Issues

In addition to the risks and threats inherent in traditional IT computing, cloud computing presents an organization with its own set of security issues.

The CIA Triad

The three fundamental tenets of information security — confidentiality, integrity, and availability (CIA) — define an organization's security posture. All of the information security controls and safeguards, and all of the threats, vulnerabilities, and security processes are subject to the CIA yardstick.

Confidentiality

Confidentiality is the prevention of the intentional or unintentional unauthorized disclosure of contents. Loss of confidentiality can occur in many ways. For example, loss of confidentiality can occur through the intentional release of private company information or through a misapplication of network rights.

Some of the elements of telecommunications used to ensure confidentiality are as follows:

- Network security protocols
- Network authentication services
- Data encryption services

Integrity

Integrity is the guarantee that the message sent is the message received and that the message is not intentionally or unintentionally altered. Loss of integrity can occur through an intentional attack to change information (for example, a website defacement) or, more commonly, unintentionally (data is accidentally altered by an operator). Integrity also contains the concept of nonrepudiation

of a message source, which we will describe later.

Some of the elements used to ensure integrity include the following:

- Firewall services
- Communications security management
- Intrusion detection services

Availability

This concept refers to the elements that create reliability and stability in networks and systems. It ensures that connectivity is accessible when needed, allowing authorized users to access the network or systems.

Also included in that assurance is the guarantee that security services for the security practitioner are usable when they are needed. The concept of availability also tends to include areas in an information system (IS) that are traditionally not thought of as pure security (such as guarantee of service, performance, and up time), yet are obviously affected by breaches such as a denial-of-service (DoS) attack.

Some of the elements that are used to ensure availability are as follows:

- Fault tolerance for data availability, such as backups and redundant disk systems
- Acceptable logins and operating process performance
- Reliable and interoperable security processes and network security mechanisms

Privacy and Compliance Risks

One area that is greatly affected by cloud computing is privacy. It's important to remember that although the control of cloud computing privacy has many threats and vulnerabilities in common with non cloud processes and infrastructure, it also has unique security issues.

For example, a successful identity theft exploit can result in a privacy loss that has a huge impact on an enterprise. The organization can suffer short-term losses due to remediation, investigation, and restitution costs. It can also incur longer term problems for the organization due to loss of credibility, confidence, and negative publicity.

Another mistake organizations often make is in assigning responsibility for privacy controls to the IT dept, rather than a business unit that owns the data. Information systems security frameworks have defined, standardized processes that apply to cloud computing — and its potential privacy breaches. This section examines the legal and standard processes that affect privacy control in the cloud.

An individual's right to privacy is embodied in the fundamental principles of privacy:

- **Notice** — Regarding the collection, use, and disclosure of personally identifiable information (PII)
- **Choice** — To opt out or opt in regarding disclosure of PII to third parties
- **Access** — By consumers to their PII to permit review and correction of information

- **Security** — To protect PII from unauthorized disclosure
- **Enforcement** — Of applicable privacy policies and obligations

Proving compliance depends on the number of transactions that they process. Merchants and service providers with higher levels of transactions have to pass an on-site audit every year. Those with lower levels of transactions must submit documentation stating that they meet the requirements, called a self-assessment questionnaire.

The PCI DSS contains the following set of 12 high-level requirements that are supported by a series of more detailed requirements:

- Install and maintain a firewall configuration to protect cardholder data.
- Do not use vendor-supplied defaults for system passwords and other security parameters.
- Protect stored cardholder data.
- Encrypt transmission of cardholder data across open, public networks.
- Use and regularly update antivirus software.
- Develop and maintain secure systems and applications.
- Restrict access to cardholder data based on the business's need to know.
- Assign a unique ID to each person with computer access.
- Restrict physical access to cardholder data.
- Track and monitor all access to network resources and

cardholder data.

- Regularly test security systems and processes.
- Maintain a policy that addresses information security.

Information Privacy and Privacy Laws

There are many types of legal systems in the world, and they differ in how they treat evidence, the rights of the accused, and the role of the judiciary. These laws have a significant privacy impact on cloud computing environments, yet vary widely.

Examples of these different legal systems are common law, Islamic and other religious law, and civil law. The common law system is employed in the United States, United Kingdom, Australia, and Canada. Civil law systems are used in France, Germany, and Quebec, to name a few.

Organizations develop and publish privacy policies that describe their approach to handling PII. The websites of organizations usually have their privacy policies available to read online, and these policies usually cover the following areas:

- Statement of the organization's commitment to privacy
- The type of information collected, such as names, addresses, credit card numbers, phone numbers, and so on
- Retaining and using e-mail correspondence
- Information gathered through cookies and Web server logs and how that information is used

- How information is shared with affiliates and strategic partners
- Mechanisms to secure information transmissions, such as encryption and digital signatures
- Mechanisms to protect PII stored by the organization
- Procedures for review of the organization's compliance with the privacy policy
- Evaluation of information protection practices
- Means for the user to access and correct PII held by the organization
- Rules for disclosing PII to outside parties
- Providing PII that is legally required

Privacy Legislation

The following list summarizes some important legislation and recommended guidelines for privacy:

- *The Cable Communications Policy Act* provides for discretionary use of PII by cable operators internally but imposes restrictions on disclosures to third parties.
- *The Children's Online Privacy Protection Act* (COPPA) is aimed at providing protection to children under the age of 13.
- *Customer Proprietary Network Information Rules* apply to telephone com- panies and restrict their use of customer information both internally and to third parties.

- *The Financial Services Modernization Act* (*Gramm-Leach-Bliley*) requires financial institutions to provide customers with clear descriptions of the institution's policies and procedures for protecting the PII of customers.
- *The Telephone Consumer Protection Act* restricts communications between companies and consumers, such as telemarketing.
- *The 1973 U.S. Code of Fair Information Practices* states that:

1. There must not be personal data record-keeping systems whose very existence is secret.
2. There must be a way for a person to find out what information about them is in a record and how it is used.
3. There must be a way for a person to prevent information about them, which was obtained for one purpose, from being used or made avail- able for another purpose without their consent.

Health Insurance Portability and Accountability Act (HIPAA)

An excellent example of the requirements and application of individual privacy principles is in the area of health care. The protection from disclosure and misuse of a private individual's medical information is a prime example of a privacy law. Some of the common health care security issues are as follows:

- Access controls of most health care information systems do not provide sufficient granularity to implement the principle of least privilege among users.
- Most off-the-shelf applications do not incorporate adequate

information security controls.

- Systems must be accessible to outside partners, members, and some vendors.
- Providing users with the necessary access to the Internet creates the poten- tial for enabling violations of the privacy and integrity of information.
- Criminal and civil penalties can be imposed for the improper disclosure of medical information.
- A large organization's misuse of medical information can cause the public to change its perception of the organization.
- Health care organizations should adhere to the following information privacy principles (based on European Union principles):
  - An individual should have the means to monitor the database of stored information about themselves and should have the ability to change or correct that information.
  - Information obtained for one purpose should not be used for another purpose.
  - Organizations collecting information about individuals should ensure that the information is provided only for its intended use and should provide safeguards against the misuse of this information.

The existence of databases containing personal information should not be kept secret.

Threats to Infrastructure Data and Access Control

To properly understand the threats that cloud computing presents to the computing infrastructure, it's important to understand communications security techniques to prevent, detect, and correct errors so that integrity, availability, and the confidentiality of transactions over networks may be maintained.
This includes the following:

- Communications and network security as it relates to voice, data, multi- media, and facsimile transmissions in terms of local area, wide area, and remote access networks
- Internet/intranet/extranet in terms of firewalls, routers, gateways, and various protocols

Common Threats and Vulnerabilities

A threat is simply any event that, if realized, can cause damage to a system and create a loss of confidentiality, availability, or integrity. Threats can be malicious, such as the intentional modification of sensitive information, or they can be accidental — such as an error in a transaction calculation or the accidental deletion of a file.

A *vulnerability* is a weakness in a system that can be exploited by a threat. Reducing the vulnerable aspects of a system can reduce the risk and impact of threats on the system. For example, a password-generation tool, which helps users choose robust passwords, reduces the chance that users will select poor

passwords (the vulnerability) and makes the password more difficult to crack (the threat of external attack).

Common threats to both cloud and traditional infrastructure include the following:

- **Eavesdropping** — Data scavenging, traffic or trend analysis, social engineer- ing, economic or political espionage, sniffing, dumpster diving, keystroke monitoring, and shoulder surfing are all types of eavesdropping to gain information or to create a foundation for a later attack. Eavesdropping is a primary cause of the failure of confidentiality.
- **Fraud** — Examples of fraud include collusion, falsified transactions, data manipulation, and other altering of data integrity for gain.
- **Theft** — Examples of theft include the theft of information or trade secrets for profit or unauthorized disclosure, and physical theft of hardware or software.
- **Sabotage** — Sabotage includes denial-of-service (DoS) attacks, production delays, and data integrity sabotage.
- **External attack** — Examples of external attacks include malicious crack- ing, scanning, and probing to gain infrastructure information, demon dialing to locate an unsecured modem line, and the insertion of a malicious code or virus.

This section explores the most common types of attacks. Although these attacks are constantly evolving, most networked systems attacks can be grouped

into several general areas.

Logon Abuse

Logon abuse can refer to legitimate users accessing services of a higher security level that would normally be restricted to them. Unlike network intrusion, this type of abuse focuses primarily on those users who might be legitimate users of a different system or users who have a lower security classification.

Masquerading is the term used when one user pretends to be another user, such as an attacker socially engineering passwords from an Internet Service Provider (ISP).

Inappropriate System Use

This style of network abuse refers to the nonbusiness or personal use of a network by otherwise authorized users, such as Internet surfing to inappropriate content sites (travel, pornography, sports, and so forth). As per the International Information Systems Security Certification Consortium (ISC) Code of Ethics and the Internet Advisory Board (IAB) recommendations, the use of networked services for other than business purposes can be considered abuse of the system. While most employers do not enforce extremely strict Web surfing rules, occasional harassment litigation may result from employees accessing pornography sites and employees operating private Web businesses using the company's infrastructure.

Eavesdropping

This type of network attack consists of the unauthorized interception of network

traffic. Certain network transmission methods, such as satellite, wireless, mobile, PDA, and so on, are vulnerable to eavesdropping attacks. *Tapping* refers to the physical interception of a transmission medium (like the splicing of a cable or the creation of an induction loop to pick up electromagnetic emanations from copper). Eavesdropping can take one of two forms:

- **Passive eavesdropping** — Covertly monitoring or listening to transmissions that are unauthorized by either the sender or receiver
- **Active eavesdropping** — Tampering with a transmission to create a covert signaling channel, or actively probing the network for infrastructure information

Eavesdropping and probing are often the preliminary steps to session hijack- ing and other network intrusions. Covert channel eavesdropping refers to using a hidden, unauthorized network connection to communicate unauthorized information. A covert channel is a connection intentionally created to transmit unauthorized information from inside a trusted network to a partner at an outside, untrusted node.

War walking (or war driving) refers to scanning for 802.11-based wireless net- work information by either driving or walking with a laptop, a wireless adapter in promiscuous mode, some type of scanning software such as NetStumbler or AiroPeek, and a Global Positioning System (GPS).

Network Intrusion

This type of attack refers to the use of unauthorized access to break into a network primarily from an external source. Unlike a logon abuse attack, the intruders

are not considered to be known to the company. Most common hacks belong to this category. Also known as a *penetration attack*, it exploits known security vulnerabilities in the security perimeter.

Back doors are very hard to trace, as an intruder will often create several avenues into a network to be exploited later. The only real way to ensure that these avenues are closed after an attack is to restore the operating system from the original media, apply the patches, and restore all data and applications.

Piggy-backing, in the network domain, refers to an attacker gaining unauthorized access to a system by using a legitimate user's connection. A user leaves a session open or incorrectly logs off, enabling an unauthorized user to resume the session.

Denial-of-Service (DoS) Attacks

The DoS attack might use some of the following techniques to overwhelm a target's resources:

- Filling up a target's hard drive storage space by using huge e-mail attach- ments or file transfers
- Sending a message that resets a target host's subnet mask, causing a dis- ruption of the target's subnet routing
- Using up all of a target's resources to accept network connections, result- ing in additional network connections being denied

Session Hijacking Attacks

Unauthorized access to a system can be achieved by session hijacking. In this

type of attack, an attacker hijacks a session between a trusted client and network server. The attacking computer substitutes its IP address for that of the trusted client and the server continues the dialog, believing it is communicating with the trusted client. High jacking attacks include IP spoofing attacks, TCP sequence number attacks, and DNS poisoning.

Fragmentation Attacks

IP fragmentation attacks use varied IP datagram fragmentation to disguise their TCP packets from a target's IP filtering devices. The following are two examples of these types of attacks:

- A *tiny fragment attack* occurs when the intruder sends a very small frag- ment that forces some of the TCP header field into a second fragment. If the target's filtering device does not enforce minimum fragment size, this illegal packet can then be passed on through the target's network.
- An *overlapping fragment attack* is another variation on a datagram's zero-offset modification. Subsequent packets overwrite the initial packet's destination address information, and then the second packet is passed by the target's filtering device. This can happen if the target's filtering device does not enforce a minimum fragment offset for fragments with non-zero offsets.

Cloud Access Control Issues

The cost of access control in the cloud must be commensurate with the value of the information being protected. The value of this information is determined through qualitative and quantitative methods. These methods incorporate fac-

tors such as the cost to develop or acquire the information, the importance of the information to an organization and its competitors, and the effect on the organization's reputation if the information is compromised.

Proper access controls enable full availability. Availability ensures that a system's authorized users have timely and uninterrupted access to the information in the system. The additional access control objectives are reliability and utility. Access control must offer protection from an unauthorized, unanticipated, or unintentional modification of information. This protection should preserve the data's internal and external consistency. The confidentiality of the information must also be similarly maintained, and the information should be available on a timely basis. These factors cover the integrity, confidentiality, and availability components of information system security.

Accountability is another facet of access control. Individuals on a system are responsible for their actions. This accountability property enables system activities to be traced to the proper individuals. Accountability is supported by audit trails that record events on both the system and the network. Audit trails can be used for intrusion detection and for the reconstruction of past events. Monitoring individual activities, such as keystroke monitoring, should be accomplished in accordance with the company policy and appropriate laws. Banners at logon time should notify the user of any monitoring being conducted.

The following measures compensate for both internal and external access violations:

- Backups

- RAID (Redundant Array of Independent Disks) technology
- Fault tolerance
- Business continuity planning
- Insurance

Database Integrity Issues

Database integrity requires the following three goals:

- Prevention of the modification of information by unauthorized users
- Prevention of the unauthorized or unintentional modification of informa- tion by authorized users
- Preservation of both internal and external consistency:
  - *Internal consistency* — Ensures that internal data is consistent. For example, assume that an internal database holds the number of units of a particular item in each department of an organization. The sum of the number of units in each department should equal the total num- ber of units that the database has recorded internally for the whole organization.
  - *External consistency* — Ensures that the data stored in the database is consistent with the real world. Using the preceding example, external consistency means that the number of items recorded in the database for each department is equal to the number of items that physically exist in that department.

Cloud Service Provider Risk

Using virtualized systems introduces many new risks, while maintaining many if not most of the risks inherent in using traditional systems. The publication by the Burton Group, "Attacking and Defending Virtual Environments," groups these risk as follows:

- All existing attacks still work.
- As a separate system that must be protected, the hypervisor is risk additive.
- Aggregating separate systems into VMs increases risk.
- An untrusted hypervisor with a trusted VM has a higher risk than a trusted hypervisor with an untrusted VM.

Based on these parameters, we can identify several areas of risk to virtualized systems, including the following:

- **Complexity of configuration** — Virtual systems add more layers of com- plexity to networks and systems, greatly increasing the possibility of improper configuration or the induction of heretofore unseen vulnerabilities.
- **Privilege escalation** — A hacker may be able to escalate his or her privi- leges on a system by leveraging a virtual machine using a lower level of access rights, then attack a VM with a higher level of security controls through the hypervisor.
- **Inactive virtual machines** — Virtual machines that are not active (i.e., are dormant), could store data that is sensitive. Monitoring access to that data in a dormant VM is virtually

impossible, but provides a security risk through the loss of or access to the VM. Also, monitoring tools for VM systems are not as mature as traditional tools, but are expected to improve quickly.

- **Segregation of duties** — A virtualized system poses risk to organiza- tions through the improper definition of user access roles. Because the VM provides access to many type of components from many directions, proper segregation of duties may be difficult to maintain.
- **Poor access controls** — The virtual machine's hypervisor facilitates hard- ware virtualization and mediates all hardware access for the running virtual machines. This creates a new attack vector into the VM, due to its single point of access. Therefore, the hypervisor can expose the trusted network through poorly designed access control systems, deficient patch- ing, and lack of monitoring. This vulnerability also applies to virtualized databases.

It is important for the information security professional to understand and identify other types of attacks. These attacks are summarized in the following sections.

Back-Door

A back-door attack takes place using dial-up modems or asynchronous external connections. The strategy is to gain access to a network through bypassing of control mechanisms, getting in through a "back door" such as a modem.

Spoofing

Intruders use IP spoofing to convince a system that it is communicating with a known, trusted entity in order to provide the intruder with access to the system. IP spoofing involves alteration of a packet at the TCP level, which is used to attack Internet-connected systems that provide various TCP/IP services. The attacker sends a packet with an IP source address of a known, trusted host instead of its own IP source address to a target host. The target host may accept the packet and act upon it.

Man-in-the-Middle

The man-in-the-middle attack involves an attacker, A, substituting his or her public key for that of another person, P. Then, anyone desiring to send an encrypted message to P using P's public key is unknowingly using A's public key. Therefore, A can read the message intended for P. A can then send the message on to P, encrypted in P's real public key, and P will never be the wiser. Obviously, A could modify the message before resending it to P.

Replay

The replay attack occurs when an attacker intercepts and saves old messages and then tries to send them later, impersonating one of the participants. One method of making this attack more difficult to accomplish is through the use of a random number or string called a *nonce*. For example, if Bob wants to com- municate with Alice, he sends a nonce along with the first message to Alice. When Alice replies, she sends the nonce back to Bob, who verifies that

it is the one he sent with the first message. Anyone trying to use these same messages later will not be using the newer nonce. Another approach to countering the replay attack is for Bob to add a timestamp to his message.

This timestamp indicates the time that the message was sent. Thus, if the message is used later, the timestamp will show that an old message is being used.

TCP Hijacking

In this type of attack, an attacker steals, or hijacks, a session between a trusted client and network server. The attacking computer substitutes its IP address for that of the trusted client, and the server continues the dialog believing it is communicating with the trusted client.

Social Engineering

This attack uses social skills to obtain information such as passwords or PIN numbers to be used against information systems. For example, an attacker may impersonate someone in an organization and make phone calls to employees of that organization requesting passwords for use in maintenance operations.

The following are additional examples of social engineering attacks:

- E-mails to employees from a cracker requesting their passwords to validate the organizational database after a network intrusion has occurred
- E-mails to employees from a cracker requesting their passwords because work has to be done over the weekend on the system
- E-mails or phone calls from a cracker impersonating an official

who is conducting an investigation for the organization and requires passwords for the investigation

- Improper release of medical information to individuals posing as doctors and requesting data from patients' records
- A computer repair technician convinces a user that the hard disk on his or her PC is damaged and irreparable and installs a new hard disk. The technician then takes the hard disk, extracts the information, and sells the information to a competitor or foreign government.

Dumpster Diving

Dumpster diving involves the acquisition of information that is discarded by an individual or organization. In many cases, information found in trash can be very valuable to a cracker. Discarded information may include technical manuals, password lists, telephone numbers, credit card numbers, and organization charts. Note that in order for information to be treated as a trade secret, it must be adequately protected and not revealed to any unauthorized individuals. If a document containing an organization's trade secret information is inadvertently discarded and found in the trash by another person, the other person can use that information, as it was not adequately protected by the organization.

Password Guessing

Because passwords are the most commonly used mechanism to authenticate users to an information system, obtaining passwords is a common and effective attack approach. Gaining access to a person's password can be obtained by physically looking around their desk for notes with the password, "sniffing"

the connection to the network to acquire unencrypted passwords, social engineering, gaining access to a password database, or outright guessing. The last approach can be done in a random or systematic manner.

An effective means to prevent password guessing is to place a limit on the number of user attempts to enter a password. For example, a limit could be set such that a user is "locked out" of a system for a period of time after three unsuccessful tries at entering the password. This approach must be used carefully, however. For example, consider the consequences of employing this type of control in a critical application such as a Supervisory Control and Data Acquisition (SCADA) System. SCADA systems are used to run real-time processes such as oil refineries, nuclear power stations, and chemical plants.

Consider the consequences of a panicked operator trying to respond to an emergency in the plant, improperly typing in his or her password a number of times, and then being locked out of the system. Clearly, the lock-out approach should be carefully evaluated before being applied to systems requiring rapid operator responses.

Trojan Horses and Malware

Trojan horses hide malicious code inside a host program that seems to do something useful. Once these programs are executed, the virus, worm, or other type of malicious code hidden in the Trojan horse program is released to attack the workstation, server, or network, or to allow unauthorized access to those devices. Trojans are common tools used to create back doors into the network for later exploitation by crackers. Trojan horses can be carried via Internet traffic

such as FTP downloads or downloadable applets from websites, or distributed through e-mail.

Some Trojans are programmed to open specific ports to allow access for exploitation. If a Trojan is installed on a system it often opens a high-numbered port. Then the open Trojan port could be scanned and located, enabling an attacker to compromise the system.

A *logic bomb* is an instantiation of a Trojan horse that is activated upon the occurrence of a particular event. For example, the malicious code might be set to run when a specific piece of code is executed or at a certain time and date. Similarly, *a time bomb* is set to activate after a designated period of time has elapsed.

Architectural Considerations

A variety of factors affect the implementation and performance of cloud security architecture. There are general issues involving regulatory requirements, adherence to standards, security management, information classification, and security awareness. Then there are more specific architecturally related areas, including trusted hardware and software, providing for a secure execution environment, establishing secure communications, and hardware augmentation through micro architectures. These important concepts are addressed in this section.

General Issues

A variety of topics influence and directly affect the cloud security architecture.

They include such factors as compliance, security management, administrative issues, controls, and security awareness.

Compliance with legal regulations should be supported by the cloud security architecture. As a corollary, the cloud security policy should address classification of information, what entities can potentially access information, under what conditions the access has to be provided, the geographical jurisdiction of the stored data, and whether or not the access is appropriate. Proper controls should be determined and verified with assurance methods, and appropriate personnel awareness education should be put in place.

Compliance

In a public cloud environment, the provider does not normally inform the clients of the storage location of their data. In fact, the distribution of processing and data storage is one of the cloud's fundamental characteristics. However, the cloud provider should cooperate to consider the client's data location requirements. In addition, the cloud vendor should provide transparency to the client by supplying information about storage used, processing characteristics, and other relevant account information. Another compliance issue is the accessibility of a client's data by the provider's system engineers and certain other employees. This factor is a necessary part of providing and maintaining cloud services, but the act of acquiring sensitive information should be monitored, controlled, and protected by safeguards such as separation of duties. In situations where information is stored in a foreign jurisdiction, the ability of local law enforcement agencies to access a client's sensitive data is a concern. For example, this scenario might occur when a government entity conducts a computer forensics investigation of a cloud provider under suspicion of illegal activity.

Security Management

Security architecture involves effective security management to realize the benefits of cloud computation. Proper cloud security management and administration should identify management issues in critical areas such as access control, vulnerability analysis, change control, incident response, fault tolerance, and disaster recovery and business continuity planning. These areas are enhanced and supported by the proper application and verification of cloud security controls.

Controls

The objective of cloud security controls is to reduce vulnerabilities to a tolerable level and minimize the effects of an attack. To achieve this, an organization must determine what impact an attack might have, and the likelihood of loss. Examples of loss are compromise of sensitive information, financial embezzlement, loss of reputation, and physical destruction of resources. The process of analyzing various threat scenarios and producing a representative value for the estimated potential loss is known as a *risk analysis (RA)*. Controls function as countermeasures for vulnerabilities. There are many kinds of controls, but they are generally categorized into one of the following four types:

- **Deterrent controls** — Reduce the likelihood of a deliberate attack.
- **Preventative controls** — Protect vulnerabilities and make an attack unsuccessful or reduce its impact. Preventative controls inhibit attempts to violate security policy.
- **Corrective controls** — Reduce the effect of an attack.

- **Detective controls** — Discover attacks and trigger preventative or corrective controls. Detective controls warn of violations or attempted violations of security policy and include such controls as intrusion detection systems, organizational policies, video cameras, and motion detectors.

Complementary Actions

Additional activities involved in cloud security management include the following:

- Management and monitoring of service levels and service-level agreements
- Acquisition of adequate data to identify and analyze problem situations through instrumentation and dashboards

Reduction of the loss of critical information caused by lack of controls.

- Proper management of data on an organization's distributed computing resources. Data centralized on the cloud reduces the potential for data loss in organizations with large numbers of laptop computers and other personal computing devices.
- Monitoring of centrally stored cloud information, as opposed to having to examine data distributed throughout an organization on a variety of computing and storage devices.
- Provisioning for rapid recovery from problem situations.

Cloud security management should also foster improved capabilities to conduct forensic analysis on cloud-based information using a network forensic model. This model will provide for more rapid acquisition and verification of evidence, such as taking advantage of automatic hashing that is applied when

storing data on a cloud.

Cloud security management can also be enhanced by the selective use of automation and by the application of emerging cloud management standards to areas such as interoperable security mechanisms, quality of service, accounting, provisioning, and API specifications. APIs provide for control of cloud resources through program interfaces, and remote APIs should be managed to ensure that they are documented and consistent.

Cloud security management should address applications with the goal of enterprise cost containment through scalability, pay as you go models, on-demand implementation and provisioning, and reallocation of information management operational activities to the cloud.

Information Classification

Another major area that relates to compliance and can affect the cloud security architecture is information classification. The information classification process also supports disaster recovery planning and business continuity planning.

Information Classification Objectives

There are several good reasons to classify information. Not all data has the same value to an organization. For example, some data is more valuable to upper management, because it aids them in making strategic long-range or short-range business direction decisions. Some data, such as trade secrets, formulas, and new product information, is so valuable that its loss could create a significant problem for the enterprise in the marketplace — either by creating

public embarrassment or by causing a lack of credibility Information classification has the longest history in the government sector. Its value has long been established, and it is a required component when securing trusted systems. In this sector, information classification is used primarily to prevent the unauthorized disclosure of information and the resultant failure of confidentiality.

Information classification supports privacy requirements and enables regulatory compliance. A company might wish to employ classification to maintain a competitive edge in a tough marketplace. There might also be sound legal reasons for an organization to employ information classification on the cloud, such as to minimize liability or to protect valuable business information.

Information Classification Benefits

In addition to the aforementioned reasons, employing information classification has several clear benefits to an organization engaged in cloud computing. Some of these benefits are as follows:

- It demonstrates an organization's commitment to security protections.
- It helps identify which information is the most sensitive or vital to an organization.
- It supports the tenets of confidentiality, integrity, and availability as it pertains to data.
- It helps identify which protections apply to which information.

- It might be required for regulatory, compliance, or legal reasons.

Information Classification Concepts

The information that an organization processes must be classified according to the organization's sensitivity to its loss or disclosure. The information system owner is responsible for defining the sensitivity level of the data. Classification according to a defined classification scheme enables security controls to be properly implemented.

The following classification terms are typical of those used in the private sector and are applicable to cloud data**:**

- **Public data** — Information that is similar to unclassified information; all of a company's information that does not fit into any of the next cat- egories can be considered public. While its unauthorized disclosure may be against policy, it is not expected to impact seriously or adversely the organization, its employees, and/or its customers.
- **Sensitive data** — Information that requires a higher level of classification than normal data. This information is protected from a loss of confidentiality as well as from a loss of integrity due to an unauthorized alteration. This classification applies to information that requires special precautions to ensure its integrity by protecting it from unauthorized modification or deletion. It is information that requires a higher-than-normal assurance of accuracy and

completeness.

- **Private data** — This classification applies to personal information that is intended for use within the organization. Its unauthorized disclosure could seriously and adversely impact the organization and/or its employees. For example, salary levels and medical information are considered private.

- **Confidential data** — This classification applies to the most sensitive business information that is intended strictly for use within the organization. Its unauthorized disclosure could seriously and adversely impact the organization, its stockholders, its business partners, and/or its customers. This information is exempt from disclosure under the provisions of the Freedom of Information Act or other applicable federal laws or regulations. For example, information about new product development, trade secrets, and merger negotiations is considered confidential.

Classification Criteria

Several criteria may be used to determine the classification of an information object:

- **Value** — Value is the number one commonly used criteria for classifying data in the private sector. If the information is valuable to an organization or its competitors, then it needs to be classified.

- **Age** — The classification of information might be lowered if the informa- tion's value decreases over time. In the U.S. Department of Defense, some classified documents are automatically

declassified after a predetermined time period has passed.

- **Useful life** — If the information has been made obsolete due to new information, substantial changes in the company, or other reasons, the information can often be declassified.
- **Personal association** — If information is personally associated with specific individuals or is addressed by a privacy law, it might need to be classified. For example, investigative information that reveals informant names might need to remain classified.

Information Classification Procedures

There are several steps in establishing a classification system. These are the steps in priority order:

1. Identify the appropriate administrator and data custodian. The data custodian is responsible for protecting the information, running backups, and performing data restoration.
2. Specify the criteria for classifying and labeling the information.
3. Classify the data by its owner, who is subject to review by a supervisor.
4. Specify and document any exceptions to the classification policy.
5. Specify the controls that will be applied to each classification level.
6. Specify the termination procedures for declassifying the information or for transferring custody of the information to

another entity.

7. Create an enterprise awareness program about the classification controls.

Trusted Cloud Computing

Trusted cloud computing can be viewed as a computer security architecture that is designed to protect cloud systems from malicious intrusions and attacks, and ensure that computing resources will act in a specific, predictable manner as intended. A trusted cloud computing system will protect data in use by hypervisors and applications, protect against unauthorized access to information, provide for strong authentication, apply encryption to protect sensitive data that resides on stolen or lost devices, and support compliance through hardware and software mechanisms.

Trusted Computing Characteristics

In a cloud computational system, multiple processes might be running concurrently. Each process has the capability to access certain memory locations and to execute a subset of the computer's instruction set. The execution and memory space assigned to each process is called a *protection domain*. This domain can be extended to virtual memory, which increases the apparent size of real memory by using disk storage. The purpose of establishing a protection domain is to protect programs from all unauthorized modification or executional interference. A *trusted computing base (TCB)* is the total combination of protection mechanisms within a computer system, which includes the hardware, software, and firmware that are trusted to enforce a security policy. Because the TCB components are responsible for enforcing the

security policy of a computing system, these components must be protected from malicious and untrusted processes. The TCB must also provide for memory protection and ensure that the processes from one domain do not access memory locations of another domain. The *security perimeter* is the boundary that separates the TCB from the remainder of the system. A *trusted path* must also exist so that users can access the TCB without being compromised by other processes or users. Therefore, a *trusted computer system* is one that employs the necessary hardware and software assurance measures to enable its use in processing multiple levels of classified or sensitive information. This system meets the specified requirements for reliability and security.

Another element associated with trusted computing is the *trusted platform module (TPM)*. The TPM stores cryptographic keys that can be used to attest to the operating state of a computing platform and to verify that the hardware and software configuration has not been modified. However, the standard TPM cannot be used in cloud computing because it does not operate in the virtualized cloud environment. To permit a TPM version to perform in the cloud, specifications have been generated for a virtual TPM (VTM)[4] that provides software instances of TPMs for each virtual machine operating on a trusted server.

Trusted computing also provides the capability to ensure that software that processes information complies with specified usage policies and is running unmodified and isolated from other software on the system. In addition, a trusted computing system must be capable of enforcing mandatory access control (MAC) rules. MAC rules are discussed in more detail later in this chapter.

Numerous trust-related issues should be raised with, and satisfied by, a cloud provider. They range from concerns about security, performance, cost, control, availability, resiliency, and vendor lock in.

Additional factors that inspire trust include the following:

- Use of industry-accepted standards.
- Provision for interoperability and transparency.
- Robust authentication and authorization mechanisms in access control.
- Management of changing personnel and relationships in both the cloud client and provider organizations.
- Establishment of accountability with respect to security and privacy require- ments in a multi-party, flexible service delivery setting.
- Use of information system security assurance techniques and metrics to establish the effectiveness of hardware and software protection mechanisms.
- Establishment of effective policies and procedures to address multiple legal jurisdictions associated with cloud international services and com- pliance requirements.
- Application of Information Rights Management (IRM) cryptographic techniques to protect sensitive cloud-based documents and provide an audit trail of accesses and policy changes. IRM prevents protected docu- ments from screen capture, being printed, faxed, or forwarded, and can prohibit

messages and attachments from being accessed after a specified period of time.

Also, because of the high volume of data that is being moved around in various locations, authorization privileges and rights management constraints must be attached to the data itself to restrict access only to authorized users.

Because of legal and forensic requirements, a trusted cloud provider should also have a Security Information and Event Management (SIEM) capability that can manage records and logs in a manner that meets legal constraints. An SEIM is a software mechanism that provides for centralized acquisition, storage, and analysis of recorded events and logs generated by other tools on an enterprise network.

Information stored in a SEIM can be used for data mining to discover significant trends and occurrences, and to provide for reliable and legally acceptable storage of information. It can also be used by report generators, and provide for backup of log data that might be lost at the source of the data.

Secure Execution Environments and Communications

In a cloud environment, applications are run on different servers in a distributed mode. These applications interact with the outside world and other applications and may contain sensitive information whose inappropriate access would be harmful to a client. In addition, cloud computing is increasingly being used to manage and store huge amounts of data in database applications that are also co-located with other users' information. Thus, it is extremely important for the cloud supplier to provide a secure execution environment and secure

communications for client applications and storage.

Secure Execution Environment

Configuring computing platforms for secure execution is a complex task; and in many instances it is not performed properly because of the large number of parameters that are involved. This provides opportunities for malware to exploit vulnerabilities, such as downloading code embedded in data and having the code executed at a high privilege level.

In cloud computing, the major burden of establishing a secure execution environment is transferred from the client to the cloud provider. However, protected data transfers must be established through strong authentication mechanisms, and the client must have practices in place to address the privacy and confidentiality of information that is exchanged with the cloud. In fact, the client's port to the cloud might provide an attack path if not properly provisioned with security measures. Therefore, the client needs assurance that computations and data exchanges are conducted in a secure environment. This assurance is affected by trust enabled by cryptographic methods. Also, research into areas such as compiler-based virtual machines promises a more secure execution environment for operating systems.

Another major concern in secure execution of code is the widespread use of "unsafe" programming languages such as C and C++ instead of more secure languages such as object-oriented Java and structured, object-oriented C#.

Secure Communications

As opposed to having managed, secure communications among the computing resources internal to an organization, movement of applications to the cloud requires a reevaluation of communications security. These communications apply to both data in motion and data at rest.

Secure cloud communications involves the structures, transmission methods, transport formats, and security measures that provide confidentiality, integ- rity, availability, and authentication for transmissions over private and public communications networks. Secure cloud computing communications should ensure the following:

- **Confidentiality** — Ensures that only those who are supposed to access data can retrieve it. Loss of confidentiality can occur through the inten- tional release of private company information or through a misapplication of network rights. Some of the elements of telecommunications used to ensure confidentiality are as follows:
  - Network security protocols
  - Network authentication services
  - Data encryption services
- **Integrity** — Ensures that data has not been changed due to an accident or malice. Integrity is the guarantee that the message sent is the message received and that the message is not intentionally or unintentionally altered. Integrity also contains the concept of nonrepudiation of a message source. Some of the constituents of integrity are as follows:

- Firewall services
- Communications Security Management
- Intrusion detection services

- **Availability** — Ensures that data is accessible when and where it is needed, and that connectivity is accessible when needed, allowing authorized users to access the network or systems. Also included in that assurance is the guarantee that security services for the security practitioner are usable when they are needed. Some of the elements that are used to ensure avail- ability are as follows:

  - Fault tolerance for data availability, such as backups and redundant disk systems
  - Acceptable logins and operating process performances
  - Reliable and interoperable security processes and network security mechanisms

APIs

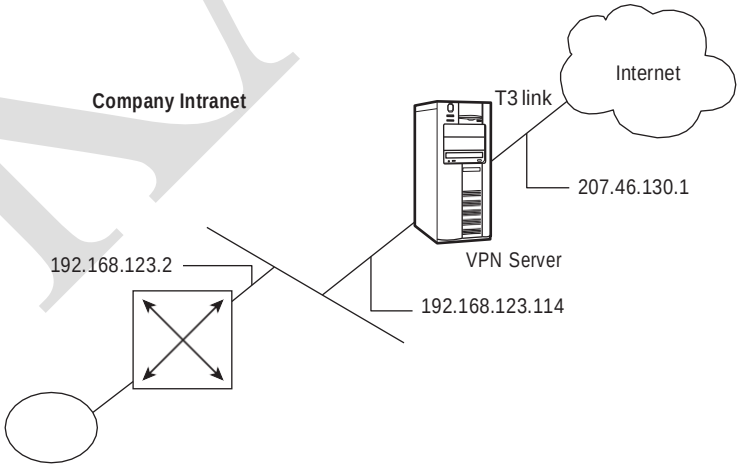
Common vulnerabilities such as weak antivirus software, unattended computing platforms, poor passwords, weak authentication mechanisms, and inadequate intrusion detection that can impact communications must be more stringently analyzed, and proper APIs must be used.

Virtual Private Networks

Another important method to secure cloud communications is through a virtual private network (VPN). A VPN is created by building a secure communications link between two nodes by emulating the properties of a

point-to-point private link. A VPN can be used to facilitate secure remote access into the cloud, securely connect two networks together, or create a secure data tunnel within a network.

The portion of the link in which the private data is encapsulated is known as the *tunnel*. It may be referred to as a secure, encrypted tunnel, although it's more accurately defined as an encapsulated tunnel, as encryption may or may not be used. To emulate a point-to-point link, data is encapsulated, or wrapped, with a header that provides routing information. Most often the data is encrypted for confidentiality. This encrypted part of the link is considered the actual virtual private network connection. Figure below shows a common VPN configuration with example IP addresses for remote access into an organization's intranet through the Internet. Address 192.168.123.2 designates the organization's router.



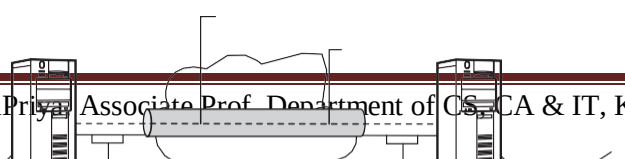
The two general types of VPNs relevant to cloud computing are remote access and network-to-network. These VPN types are described in the fol- lowing sections.

Remote Access VPNs

A VPN can be configured to provide remote access to corporate resources over the public Internet to maintain confidentiality and integrity. This con- figuration enables the remote user to utilize whatever local ISP is available to access the Internet without forcing the user to make a long-distance or 800 call to a third-party access provider. Using the connection to the local ISP, the VPN software creates a virtual private network between the dial-up user and the corporate VPN server across the Internet.

Network-to-Network VPNs

A VPN is commonly used to connect two networks, perhaps the main corporate LAN and a remote branch office LAN, through the Internet. This connection can use either dedicated lines to the Internet or dial-up connections to the Internet. However, the corporate hub router that acts as a VPN server must be connected to a local ISP with a dedicated line if the VPN server needs to be available 24/7. The VPN software uses the connection to the local ISP to create a VPN tunnel between the branch office router and the corporate hub router across the Internet. Figure below shows a remote branch office connected to the corporate main office using a VPN tunnel through the Internet.



VPN Tunneling

Tunneling is a method of transferring data from one network to another network by encapsulating the packets in an additional header. The additional header provides routing information so that the encapsulated payload can traverse the intermediate networks.

For a tunnel to be established, both the tunnel client and the tunnel server must be using the same tunneling protocol. Tunneling technology can be based on either a Layer 2 or a Layer 3 tunneling protocol. These layers correspond to the Open Systems Interconnection (OSI) Reference Model.

Tunneling, and the use of a VPN, is not intended as a substitute for encryption/ decryption. In cases where a high level of security is necessary, the strongest possible encryption should be used within the VPN itself, and tunneling should serve only as a convenience.

A popular tunneling protocol for network-to-network connectivity is IPSec, which encapsulates IP packets in an additional IP header. IPSec operates at the Network Layer of the OSI Reference Model and allows multiple simultaneous tunnels. IPSec contains the functionality to encrypt and authenticate IP data. It is built into the new IPv6 standard and is used as an add-on to the current

IPv4. IPSec tunnel mode allows IP packets to be encrypted and then encapsulated in an IP header to be sent across a corporate IP Intranetwork or a public IP Internetwork, such as the Internet.

IPSec uses an authentication header (AH) to provide source authentication and integrity without encryption, and it uses the Encapsulating Security Payload (ESP) to provide authentication and integrity along with encryption. With IPSec, only the sender and recipient know the key. If the authentication data is valid, then the recipient knows that the communication came from the sender and was not changed in transit.

Public Key Infrastructure and Encryption Key Management

To secure communications, data that is being exchanged with a cloud should be encrypted, calls to remote servers should be examined for imbedded malware, and digital certificates should be employed and managed. A certification process can be used to bind individuals to their public keys as used in public key cryptography. A *certificate authority (CA)* acts as notary by verifying a person's identity and issuing a certificate that vouches for a public key of the named individual. This certification agent signs the certificate with its own private key. Therefore, the individual is verified as the sender if that person's public key opens the data.

The certificate contains the subject's name, the subject's public key, the name of the certificate authority, and the period in which the certificate is valid. To verify the CA's signature, its public key must be cross-certified with another CA. (The X.509 standard defines the format for public key certificates.) This

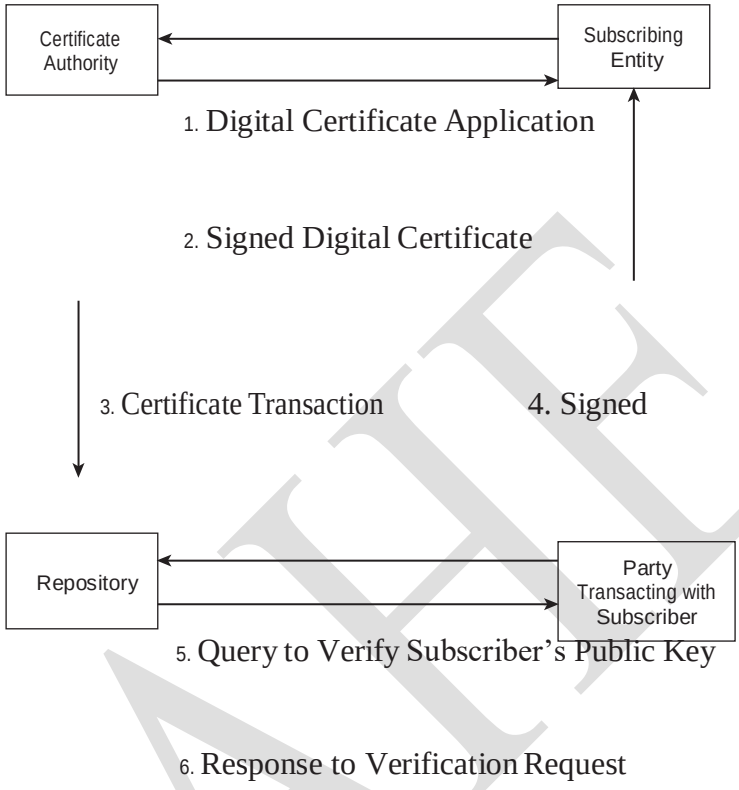
certificate is then sent to a repository, which holds the certificates and *certificate revocation lists (CRLs)* that denote the revoked certificates.

The integration of digital signatures and certificates and the other services required for e-commerce is called the *public key infrastructure (PKI)*. These services provide integrity, access control, confidentiality, authentication, and nonrepudiation for electronic transactions. The PKI includes the following elements:

- Digital certificates
- Certificate authority (CA)
- Registration authorities

Policies and procedures

- Certificate revocation
- Nonrepudiation support
- Timestamping
- Lightweight Directory Access Protocol (LDAP)
- Security-enabled applications



Digital Certificates

The digital certificate and management of the certificate are major components of PKI. Remember: The purpose of a digital certificate is to verify to all that an individual's public key — posted on a public "key ring" — is actually his or hers. A trusted, third-party CA can verify that the public key is that of the named individual and then issue a certificate attesting to that fact. The CA accomplishes the certification by digitally signing the individual's public key and associated information.

Certificates and CRLs can be held in a repository, with responsibilities

defined between the repository and the CA. The repository access proto- col determines how these responsibilities are assigned. In one protocol, the repository interacts with other repositories, CAs, and users. The CA deposits its certificates and CRLs into the repository. The users can then access the repository for this information.

Directories and X.500

In PKI, a repository is usually referred to as a *directory*. The directory contains entries associated with an object class. An object class can refer to individuals or other computer-related entities. The class defines the attributes of the object. Attributes for PKI are defined in RFC 2587, "Internet X.509 Public Key Infrastructure LDAP v2 Schema," by Boeyen, Howes, and Richard, published in April 1999. Additional information on attributes can be found in RFC 2079, "Definition of an X.500 Attribute Type and an Object Class to Hold Uniform Resource Identifiers (URLs)," by M. Smith, published in January 1997.

The X.509 certificate standard defines the authentication bases for the X.500 directory. The X.500 directory stores information about individuals and objects in a distributed database residing on network servers. Some of the principal definitions associated with X.500 include the following:

- Directory user agents (DUAs) — Clients
- Directory server agents (DSAs) — Servers
- Directory Service Protocol (DSP) — Enables information exchanges between DSAs
- Directory Access Protocol (DAP) — Enables information

exchanges from a DUA to a DSA

- Directory Information Shadowing Protocol (DISP) — Used by a DSA to duplicate or "shadow" some or all of its contents

DSAs accept requests from anonymous sources as well as authenticated requests. They share information through a *chaining* mechanism.

The Lightweight Directory Access Protocol

The Lightweight Directory Access Protocol (LDAP) was developed as a more efficient version of the DAP and has evolved into a second version. LDAP servers communicate through referrals — that is, a directory receiving a request for information it does not have will query the tables of remote directories. If it finds a directory with the required entry, it sends a referral to the requesting directory.

LDAP provides a standard format to access the certificate directories. These directories are stored on network LDAP servers and provide public keys and corresponding X.509 certificates for the enterprise. A directory contains information such as individuals' names, addresses, phone numbers, and public key certificates. The standards under X.500 define the protocols and information models for computer directory services that are independent of the platforms and other related entities. LDAP servers are subject to attacks that affect availability and integrity.

For example, denial-of-service attacks on an LDAP server could prevent access to the CRLs and thus permit the use of a revoked certificate. The DAP protocol in

X.500 was unwieldy and led to most client implementa- tions using LDAP. LDAP version 3 provides extensions that offer shadowing and chaining capabilities.

X.509 Certificates

The original X.509 certificate (CCITT, *The Directory-Authentication Framework*, Recommendation X.509, 1988) was developed to provide the authentication foundation for the X.500 directory. Since then, a version 2 and a version 3 have been developed. Version 2 of the X.509 certificate addresses the reuse of names, and version 3 provides for certificate extensions to the core certificate fields. These extensions can be used as needed by different users and different applications.

The Consultation Committee, International Telephone and Telegraph, International Telecommunications Union (CCITT-ITU)/International Organization for Standardization (ISO) has defined the basic format of an X.509 certificate. This structure is outlined in Figure .

Version
Serial Number
Algorithm Identifer • Algorithm • Parameters
Issuer

Period of Validity
Subject
Subject's Public Key Public Key Algorithm Parameters
Signature

The CCITT-ITU/ ISO X.509 certificate format

If version 3 certificates are used, the optional extensions field can be used. It comes before the signature field components in the certificate. Some typi- cal extensions are the entity's name and supporting identity information, the attributes of the key, certificate policy information, and the type of the subject. The digital signature serves as a tamper-evident envelope.

Some of the different types of certificates that are issued include the following:

- **CA certificates** — Issued to CAs, these certificates contain the public keys used to verify digital signatures on CRLs and certificates.
- **End entity certificates** — Issued to entities that are not CAs, these certifi- cates contain the public keys that are needed by the certificate's user in order to perform key management or verify a digital signature.
- **Self-issued certificates** — These certificates are issued by an entity

to itself to establish points of trust and to distribute a new signing public key.

- **Rollover certificates** — These certificates are issued by a CA to transition from an old public key to a new one.

Certificate Revocation Lists

Users check the certificate revocation list (CRL) to determine whether a digital certificate has been revoked. They check for the serial number of the signature. The CA signs the CRL for integrity and authentication purposes. A CRL is shown in Figure for an X.509 version 2 certificate.

Version
Signature
Issuer
Thisupdate (Issue Date)
Nextupdate (Date by which the next CRL will be issued)
Revoked Certificates (List of Revoked Certificates)
CRLExtensions
SignatureAlgorithm
SignatureValue

CRL format (version 2)

The CA usually generates the CRLs for its population. If the CA generates the CRLs for its entire population, the CRL is called a *full CRL*.

Key Management

Obviously, when dealing with encryption keys, the same precautions must be used as with physical keys to secure the areas or the combinations to the safes. The following sections describe the components of key management.

Key Distribution

Because distributing secret keys in symmetric key encryption poses a problem, secret keys can be distributed using asymmetric key cryptosystems. Other means of distributing secret keys include face-to-face meetings to exchange keys, sending the keys by secure messenger, or some other secure alternate channel. Another method is to encrypt the secret key with another key, called a *key encryption key*, and send the encrypted secret key to the intended receiver. These key encryption keys can be distributed manually, but they need not be distributed often. The X9.17 Standard (ANSI X9.17 [Revised], "American National Standard for Financial Institution Key Management [Wholesale]," American Bankers Association, 1985) specifies key encryption keys as well as data keys for encrypting the plain-text messages.

Key distribution can also be accomplished by splitting the keys into different parts and sending each part by a different medium.

In large networks, key distribution can become a serious problem because in

an *N*-person network, the total number of key exchanges is *N*(*N*–1)/2. Using public key cryptography or the creation and exchange of session keys that are valid only for a particular session and length of time are useful mechanisms for managing the key distribution problem.

Keys can be *updated* by generating a new key from an old key. If, for example, Alice and Bob share a secret key, they can apply the same transformation function (a hash algorithm) to their common secret key and obtain a new secret key.

Key Revocation

A digital certificate contains a timestamp or period for which the certificate is valid. Also, if a key is compromised or must be made invalid because of business- or personnel-related issues, it must be revoked. The CA maintains a CRL of all invalid certificates. Users should regularly examine this list.

Key Recovery

A system must be put in place to decrypt critical data if the encryption key is lost or forgotten. One method is *key escrow*. In this system, the key is subdivided into different parts, each of which is encrypted and then sent to a different trusted individual in an organization. Keys can also be escrowed onto smart cards.

Key Renewal

Obviously, the longer a secret key is used without changing it, the more it is subject to compromise. The frequency with which you change the key is a direct function of the value of the data being encrypted and transmitted. Also, if the same secret key is used to encrypt valuable data over a relatively long period of time, you risk compromising a larger volume of data when the key

is broken. Another important concern if the key is not changed frequently is that an attacker can intercept and change messages and then send different messages to the receiver.

Key encryption keys, because they are not used as often as encryption keys, provide some protection against attacks. Typically, private keys used for digital signatures are not frequently changed and may be kept for years.

Key Destruction

Keys that have been in use for long periods of time and are replaced by others should be destroyed. If the keys are compromised, older messages sent with those keys can be read.

Keys that are stored on disks, EEPROMS, or flash memory should be overwritten numerous times. One can also destroy the disks by shredding and burning them. However, in some cases, it is possible to recover data from disks that were put into a fire. Any hardware device storing the key, such as an EPROM, should also be physically destroyed.

Older keys stored by the operating system in various locations in memory must also be searched out and destroyed.

Multiple Keys

Usually, an individual has more than one public/private key pair. The keys may be of different sizes for different levels of security. A larger key size may be used for digitally signing documents, whereas a smaller key size may be used for encryption. A person may also have multiple roles or responsibilities

wherein they want to sign messages with a different signature. One key pair may be used for business matters, another for personal use, and another for some other activity, such as being a school board member.

Distributed versus Centralized Key Management

A CA is a form of centralized key management. It is a central location that issues certificates and maintains CRLs. An alternative is *distributed key management*, in which a "chain of trust" or "web of trust" is set up among users who know each other. Because they know each other, they can trust that each one's public key is valid. Some of these users may know other users and can thus verify their public key. The chain spreads outward from the original group. This arrangement results in an informal verification procedure that is based on people knowing and trusting each other.

Identity Management and Access Control

Identity management and access control are fundamental functions required for secure cloud computing. The simplest form of identity management is logging on to a computer system with a user ID and password. However, true identity management, such as is required for cloud computing, requires more robust authentication, authorization, and access control. It should determine what resources are authorized to be accessed by a user or process by using technology such as biometrics or smart cards, and determine when a resource has been accessed by unauthorized entities.

Identity Management

Identification and authentication are the keystones of most access control systems. Identification is the act of a user professing an identity to a system, usually in the form of a username or user logon ID to the system. Identification establishes user accountability for the actions on the system. User IDs should be unique and not shared among different individuals. In many large organizations, user IDs follow set standards, such as first initial followed by last name, and so on. In order to enhance security and reduce the amount of information available to an attacker, an ID should not reflect the user's job title or function.

Authentication is verification that the user's claimed identity is valid, and it is usually implemented through a user password at logon. Authentication is based on the following three factor types:

- **Type 1** — Something you know, such as a personal identification number (PIN) or password
- **Type 2** — Something you have, such as an ATM card or smart card
- **Type 3** — Something you are (physically), such as a fingerprint or retina scan

Sometimes a fourth factor, something you do, is added to this list. Something you do might be typing your name or other phrases on a keyboard. Conversely, something you do can be considered something you are.

Two-factor authentication requires two of the three factors to be used in the authentication process. For example, withdrawing funds from an ATM machine

requires two-factor authentication in the form of the ATM card (something you have) and a PIN number (something you know).

Passwords

Because passwords can be compromised, they must be protected. In the ideal case, a password should be used only once. This "one-time password," or OTP, provides maximum security because a new password is required for each new logon. A password that is the same for each logon is called a *static password*. A password that changes with each logon is termed a *dynamic password*. The chang- ing of passwords can also fall between these two extremes. Passwords can be required to change monthly, quarterly, or at other intervals, depending on the criticality of the information needing protection and the password's frequency of use. Obviously, the more times a password is used, the more chance there is of it being compromised. A *passphrase* is a sequence of characters that is usually longer than the allotted number for a password. The passphrase is converted into a virtual password by the system.

In all these schemes, a front-end authentication device or a back-end authentication server, which services multiple workstations or the host, can perform the authentication. Passwords can be provided by a number of devices, including tokens, memory cards, and smart cards.

Tokens

Tokens, in the form of small, hand-held devices, are used to provide passwords. The following are the four basic types of tokens:

- Static password tokens

1. Owners authenticate themselves to the token by typing in a secret password.

2. If the password is correct, the token authenticates the owner to an information system.

- Synchronous dynamic password tokens, clock-based

1. The token generates a new, unique password value at fixed time intervals that is synchronized with the same password on the authentication server (this password is the time of day encrypted with a secret key).

2. The unique password is entered into a system or workstation along with an owner's PIN.

3. The authentication entity in a system or workstation knows an owner's secret key and PIN, and the entity verifies that the entered password is valid and that it was entered during the valid time window.

- Synchronous dynamic password tokens, counter-based

1. The token increments a counter value that is synchronized with a counter in the authentication server.

2. The counter value is encrypted with the user's secret key inside the token and this value is the unique password that is entered into the system authentication server.

3. The authentication entity in the system or workstation knows the user's secret key and the entity verifies that the entered password is valid by performing the same encryption on its identical counter value.

- Asynchronous tokens, challenge-response

1. A workstation or system generates a random challenge string, and the owner enters the string into the token along with the proper PIN.

2. The token performs a calculation on the string using the PIN and generates a response value that is then entered into the workstation or system.

3. The authentication mechanism in the workstation or system performs the same calculation as the token using the owner's PIN and challenge string and compares the result with the value entered by the owner. If the results match, the owner is authenticated.

Memory Cards

Memory cards provide nonvolatile storage of information, but they do not have any processing capability. A memory card stores encrypted passwords and other related identifying information. A telephone calling card and an ATM card are examples of memory cards.

Smart Cards

Smart cards provide even more capability than memory cards by incorporat- ing additional processing power on the cards. These credit-card-size devices comprise microprocessor and memory and are used to store digital signatures, private keys, passwords, and other personal information.

Biometrics

An alternative to using passwords for authentication in logical or technical access control is *biometrics*. Biometrics is based on the Type 3 authentication mechanism — something you are. Biometrics is defined as an automated means of identifying or authenticating the identity of a living person based on physiological or behavioral

characteristics. In biometrics, identification is a one-to-many search of an individual's characteristics from a database of stored images. Authentication is a one-to-one search to verify a claim to an identity made by a person. Biometrics is used for identification in physical controls and for authentication in logical controls.

There are three main performance measures in biometrics:

- **False rejection rate (FRR) or Type I Error** — The percentage of valid subjects that are falsely rejected.
- **False acceptance rate (FAR) or Type II Error** — The percentage of invalid subjects that are falsely accepted.
- **Crossover error rate (CER)** — The percentage at which the FRR equals the FAR. The smaller the CER, the better the device is performing.

In addition to the accuracy of the biometric systems, other factors must be considered, including enrollment time, throughput rate, and acceptability. *Enrollment time* is the time that it takes to initially register with a system by providing samples of the biometric characteristic to be evaluated. An acceptable enrollment time is around two minutes. For example, in fingerprint systems the actual fingerprint is stored and requires approximately 250KB per finger for a high-quality image. This level of information is required for one-to-many searches in forensics applications on very large databases.

In finger-scan technology, a full fingerprint is not stored; rather, the features extracted from this fingerprint are stored by using a small template that requires approximately 500 to 1,000 bytes of storage. The original fingerprint cannot be reconstructed from this template. Finger-scan technology is used for one-to-one verification by using smaller databases. Updates of the enrollment information might be required because some biometric characteristics, such as voice and signature, might change over time.

The *throughput rate* is the rate at which the system processes and identifies or authenticates individuals. Acceptable throughput rates are in the range of 10 subjects per minute. *Acceptability* refers to considerations of privacy, invasiveness, and

psychological and physical comfort when using the system. For example, a concern with retina scanning systems might be the exchange of body fluids on the eyepiece. Another concern would be disclosing the retinal pattern, which could reveal changes in a person's health, such as diabetes or high blood pressure.

Collected biometric images are stored in an area referred to as a *corpus*. The corpus is stored in a database of images. Potential sources of error include the corruption of images during collection, and mislabeling or other transcription problems associated with the database. Therefore, the image collection process and storage must be performed carefully with constant checking. These images are collected during the enrollment process and thus are critical to the correct operation of the biometric device.

The following are typical biometric characteristics that are used to uniquely authenticate an individual's identity:

- **Fingerprints** — Fingerprint characteristics are captured and stored. Typical CERs are 4–5%.
- **Retina scans** — The eye is placed approximately two inches from a camera and an invisible light source scans the retina for blood vessel patterns. CERs are approximately 1.4%.
- **Iris scans** — A video camera remotely captures iris patterns and characteristics. CER values are around 0.5%.
- **Hand geometry** — Cameras capture three-dimensional hand characteristics. CERs are approximately 2%.
- **Voice** — Sensors capture voice characteristics, including throat vibra- tions and air pressure, when the subject speaks a phrase. CERs are in the range of 10%.
- **Handwritten signature dynamics** — The signing characteristics of an individual making a signature are captured and recorded. Typical char- acteristics including writing pressure and pen direction. CERs are not published at this time.

Other types of biometric characteristics include facial and palm scans.

Implementing Identity Management

Realizing effective identity management requires a high-level corporate com- mitment and dedication of sufficient resources to accomplish the task. Typical undertakings in putting identity management in place include the following:

- Establishing a database of identities and credentials
- Managing users' access rights
- Enforcing security policy
- Developing the capability to create and modify accounts
- Setting up monitoring of resource accesses
- Installing a procedure for removing access rights
- Providing training in proper procedures

An identity management effort can be supported by software that automates many of the required tasks.

The Open Group and the World Wide Web Consortium (W3C) are working toward a standard for a global identity management system that would be interoperable, provide for privacy, implement accountability, and be portable. Identity management is also addressed by the XML-based eXtensible Name Service (XNS) open protocol for universal addressing. XNS provides the fol- lowing capabilities:

- A permanent identification address for a container of an individual's personal data and contact information
- Means to verify whether an individual's contact information is valid
- A platform for negotiating the exchange of information among different entities

Access Control

Access control is intrinsically tied to identity management and is necessary to preserve the confidentiality, integrity, and availability of cloud data.

These and other related objectives flow from the organizational security policy. This policy is a high-level statement of management intent regarding the control of access to information and the personnel who are authorized to receive that information.

Three things that must be considered for the planning and implementation of access control mechanisms are threats to the system, the system's vulnerability to these threats, and the risk that the threats might materialize. These concepts are defined as follows:

- **Threat** — An event or activity that has the potential to cause harm to the information systems or networks
- **Vulnerability** — A weakness or lack of a safeguard that can be exploited by a threat, causing harm to the information systems or networks
- **Risk** — The potential for harm or loss to an information system or network; the probability that a threat will materialize

Controls

Controls are implemented to mitigate risk and reduce the potential for loss. Two important control concepts are *separation of duties* and the principle of *least privilege*. Separation of duties requires an activity or process to be performed by two or more entities for successful completion. Thus, the only way that a security policy can be violated is if there is collusion among the entities. For example, in a financial environment, the person requesting that a check be issued for payment should not also be the person who has authority to sign the check. Least privilege means that the entity that has a task to perform should be provided with the minimum resources and privileges required to complete the task for the minimum necessary period of time.

Control measures can be administrative, logical (also called technical), and physical in their implementation.

- Administrative controls include policies and procedures, security awareness training, background checks, work habit checks, a review of vacation history, and increased supervision.

- Logical or technical controls involve the restriction of access to systems and the protection of information. Examples of these types of controls are encryption, smart cards, access control lists, and transmission protocols.
- Physical controls incorporate guards and building security in general, such as the locking of doors, the securing of server rooms or laptops, the protection of cables, the separation of duties, and the backing up of files.

Controls provide accountability for individuals who are accessing sensi- tive information in a cloud environment. This accountability is accomplished through access control mechanisms that require identification and authentica- tion, and through the audit function. These controls must be in accordance with and accurately represent the organization's security policy. Assurance procedures ensure that the control mechanisms correctly implement the security policy for the entire life cycle of a cloud information system.

In general, a group of processes that share access to the same resources is called a *protection domain,* and the memory space of these processes is isolated from other running processes.

Models for Controlling Access

Controlling access by a subject (an active entity such as an individual or process) to an object (a passive entity such as a file) involves setting up access rules. These rules can be classified into three categories or models.

Mandatory Access Control

The authorization of a subject's access to an object depends upon labels, which indicate the subject's *clearance*, and the *classification or sensitivity* of the object. For example, the military classifies documents as unclassified, confidential, secret, and top secret. Similarly, an individual can receive a clearance of confidential, secret, or top secret and can have access to documents classified at or below his or her specified clearance level. Thus, an individual with a clearance of "secret" can have access to secret and confidential documents with a restriction. This restriction is that the individual must have a *need to know* relative to the clas- sified documents involved. Therefore, the

documents must be necessary for that individual to complete an assigned task. Even if the individual is cleared for a classification level of information, the individual should not access the information unless there is a need to know. *Rule-based access control* is a type of mandatory access control because rules determine this access (such as the cor- respondence of clearance labels to classification labels), rather than the identity of the subjects and objects alone.

Discretionary Access Control

With discretionary access control, the subject has authority, within certain limi- tations, to specify what objects are accessible. For example, access control lists (ACLs) can be used. An access control list is a list denoting which users have what privileges to a particular resource. For example, a *tabular listing* would show the subjects or users who have access to the object, e.g., file X, and what privileges they have with respect to that file.

An *access control triple* consists of the user, program, and file, with the cor- responding access privileges noted for each user. This type of access control is used in local, dynamic situations in which the subjects must have the discretion to specify what resources certain users are permitted to access. When a user within certain limitations has the right to alter the access control to certain objects, this is termed a *user-directed discretionary access control*. An identity-based access control is a type of discretionary access control based on an individual's identity. In some instances, a hybrid approach is used, which combines the features of user-based and identity-based discretionary access control.

Nondiscretionary Access Control

A central authority determines which subjects can have access to certain objects based on the organizational security policy. The access controls might be based on the individual's role in the organization (role-based) or the subject's responsibili- ties and duties (task-based). In an organization with frequent personnel changes, nondiscretionary access control is useful because the access controls are based on the individual's role or title

within the organization. Therefore, these access controls don't need to be changed whenever a new person assumes that role.

Access control can also be characterized as *context-dependent* or *content-dependent*. Context-dependent access control is a function of factors such as location, time of day, and previous access history. It is concerned with the environment or context of the data. In content-dependent access control, access is determined by the information contained in the item being accessed.

Single Sign-On (SSO)

Single sign-on (SSO) addresses the cumbersome situation of logging on mul- tiple times to access different resources. When users must remember numerous passwords and IDs, they might take shortcuts in creating them that could leave them open to exploitation. In SSO, a user provides one ID and password per work session and is automatically logged on to all the required applications. For SSO security, the passwords should not be stored or transmitted in the clear. SSO applications can run either on a user's workstation or on authentication servers. The advantages of SSO include having the ability to use stronger passwords, easier administration of changing or deleting the passwords, and less time to access resources. The major disadvantage of many SSO implementations is that once users obtain access to the system through the initial logon, they can freely roam the network resources without any restrictions.

Authentication mechanisms include items such as smart cards and magnetic badges. Strict controls must be enforced to prevent a user from changing con- figurations that another authority sets.

SSO can be implemented by using scripts that replay the users' multiple log- ins or by using authentication servers to verify a user's identity, and encrypted authentication tickets to permit access to system services.

Enterprise access management (EAM) provides access control management services to

Web-based enterprise systems that include SSO. SSO can be provided in a number of ways. For example, SSO can be implemented on Web applications residing on different servers in the same domain by using nonpersistent, encrypted cookies on the client interface. This task is accomplished by providing a cookie to each application that the user wishes to access. Another solution is to build a secure credential for each user on a reverse proxy that is situated in front of the Web server. The credential is then presented each time a user attempts to access protected Web applications.

POSSIBLE QUESTIONS

6 Marks

1. Explain the security objectives of a cloud computing.
2. Explain about CIA Triad
3. Analyze on Cloud Service Provider Risks with real-time environment
4. Explain in detail about Identity Management and Access Control in cloud computing
5. Elucidate the Cloud Information Security Objectives and its features
6. Describe about Threats to Infrastructure Data and Access Control in cloud
7. Enlighten the cloud security services
8. Discuss on Cloud Access Control Issues with relate to the database.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

ONE MARK QUESTIONS

DEPARTMENT OF CS, CA & IT

STAFF NAME: Dr.S.MANJU PRIYA **SUBJECT NAME: CLOUD COMPUTING**

SUB.CODE: 18CSP104 **UNIT IV** **SEMESTER: I**

S.NO	Question	Choice1	Choice2	Choice3	Choice4	Ans
1	DACS stands for_____	Data and Analysis	Data and Analysis	Data and Analysis	Data and Analysis	Data and Analysis Center
2	CIA stands for_____	Confidential, Independent and Analysis	Confidentiality, Integrity,	Cloud Integrity Assurance	Cloud Independent Analysis	Confidentiality, Integrity, and Availability
3	_____ ensures the reliable and timely access to cloud data or cloud computing resources by the appropriate personnel.	Integrity	Availability	Confidentiality	Accessibility	Availability
4	_____refers to the prevention of intentional or unintentional unauthorized disclosure of	Integrity	Security	Confidentiality	Privacy	Confidentiality

5	There are ____ types of cloud security services.	3	2	4	5	3
6	_____ is the testing or reconciliation of evidence of a user's identity.	Authentication	Authorization	Auditing	Accessibility	Authentication
7	_____refers to rights and privileges granted to an individual.	Authentication	Auditing	Authorization	Accessibility	Authorization
8	DoS stands for_____	Denial-of-System	Denial-of-Security	Denial-of-Software	Denial-of-Service	Denial-of-Service
9	_____can be regarded as the collection, use, and disclosure of personally identifiable information.	Notice	Access	Security	Enforcement	Notice
10	The Telephone Consumer Protection Act restricts communications between companies and consumers, such as_____	Telesales	Telegraph	Telemarketing	Teleconsulting	Telemarketing
11	A____is simply any event that, if realized, can cause damage to a system and create a loss of confidentiality, availability, or integrity	Threat	Theft	Vulnerability	Fraud	Threat
12	____ is a weakness in a system that can be exploited by a threat.	Eavesdropping	Virus	Fraud	Vulnerability	Vulnerability
13	_____includes denial-of-service attacks, production delays, and data integrity.	Sabotage	External attack	Eavesdropping	Vulnerability	Sabotage
14	_____ is a primary cause of the failure of confidentiality.	Vulnerability	Threat	Sabotage	Eavesdropping	Eavesdropping
15	_____can refer to legitimate users accessing services of a higher security level that would	Sabotage	Eavesdropping	Inappropriate System Use	Logon abuse	Logon abuse

16	_____refers to the physical interception of a transmission medium.	Hub	Cable	Fiber optics	Tapping	Tapping
17	_____can be used for intrusion detection and for the reconstruction of past events.	Audit trails	Accountability	Fault Tolerance	Backups	Audit trails
18	RAID stands for_____	Redundant Array of Important	Redundant Array of Independent	Random Access of Independent	Random Access of Important	Redundant Array of Independent
19	The objective of cloud security controls is to reduce _____to a tolerable level and minimize the effects of an attack.	Vulnerabilities	Threat	Fraud	Theft	Vulnerabilities
20	_____can reduce the unauthorized actions attempted by personnel.	Trusted Cloud	Security Awareness	Security Provision	Security Access	Security Awareness
21	A_____ is the total combination of protection mechanisms within a computer system.	Trusted platform module	Trusted computing base	Trusted computing network	Trusted software module	Trusted computing base
22	The_____stores cryptographic keys that can be used to attest to the operating state of a computing platform.	Trusted computing network	Trusted computing base	Trusted platform module	Trusted software module	Trusted platform module
23	VPN stands for_____	Virtual Public	Virtual Private	Virtual Pin Network	Virtual Protected	Virtual Private Network
24	The simplest form of _____ is logging on to a computer system with a user ID and password.	Access Management	Identity Control	Access Control	Identity Management	Identity Management
25	_____ is verification that the user's claimed identity is valid	Authentication	Authorization	Auditing	Accessibility	Authentication
26	_____provides maximum security because a new password is required for each new logon.	OTP	OPT	POT	TOP	OTP

27	A password that is the same for each logon is called a _____	One Time Password	Static Password	Special Password	Dynamic Password	Static Password
28	A password that changes with each logon is termed a _____	One Time Password	Static Password	Special Password	Dynamic Password	Dynamic Password
29	_____provide nonvolatile storage of information, but they do not have any processing capability.	Memory cards	Biometrics	Smart cards	Passwords	Memory cards
30	An alternative to using passwords for authentication in logical or technical access control is _____	Biometrics	Passwords	Memory cards	Private key	Biometrics
31	_____is a one-to-one search to verify a claim to an identity made by a person.	Accessibility	Authentication	Authorization	Auditing	Authentication
32	There are ____ main performance measures in biometrics.	3	2	4	5	3
33	False rejection rate is also known as_____	Type II Error	Type I Error	Type III Error	Type IV Error	Type I Error
34	Typical Crossover Error Rates of Hand geometry are _____	approximately 2%.	approximately 1%.	approximately 3%.	approximately 4%.	approximately 2%.
35	Typical Crossover Error Rates of Voice are _____	12%	10%	11%	15%	10%
36	_____addresses the cumbersome situation of logging on multiple times to access different resources.	Super sign-on	Situation sign-on	Sign single-on	Single sign-on	Single sign-on
37	_____provides access control management services to Web-based enterprise systems that include SSO.	Framework access management	Environment access management	Enterprise access management	Enterprise control access management	Enterprise access management

38	An _____ is a list denoting which users have what privileges to a particular resource.	File control list	Access control list	Program control list	Data control list	Access control list
39	_____ controls involve the restriction of access to systems and the protection of information.	Administrative	Logical	Physical	Risk	Logical
40	_____controls incorporate guards and building security in general.	Risk	Physical	Administrative	Logical	Physical
41	_____are implemented to mitigate risk and reduce the potential for loss.	Security	Privacy	Controls	Priviledge	Controls
42	A _____ can be violated is if there is collusion among the entities.	Security policy	Privacy policy	Network policy	Access policy	Security policy
43	The potential for harm or loss to an information system or network is called_____	Threat	Risk	Vulnerability	Fraud	Risk
44	XNS stands for_____	eXtensible Name Service	eXtensible Number Service	eXtensible Name Security	eXtensible Number Security	eXtensible Name Service
45	_____means to verify whether an individual's contact information is valid.	eXtensible Number Service	eXtensible Name Security	eXtensible Name Service	eXtensible Number Security	eXtensible Name Service
46	_____ for senior managers, functional managers, and business unit managers.	Technical security training	Awareness training	Advanced training	Security training	Security training

47	_____for IT support personnel and system administrators.	Technical security training	Awareness training	Advanced training	Security training	Technical security training
48	_____for specific departments or personnel groups with security-sensitive positions.	Technical security	Awareness training	Advanced training	Security training	Awareness training+H190
49	Confidentiality, integrity, and availability are important pillars of _____	Cloud security services	Cloud Information Security	Cloud software assurance	Cloud Independent Analysis	Cloud software assurance
50	_____concept guarantees that the security services of the cloud system are in working order.	Availability	Confidentiality	Accessibility	Integrity	Availability
51	_____is the guarantee that the message sent is the message received and that the message is not intentionally or unintentionally altered.	Integrity	Security	Privacy	Confidentiality	Integrity
52	_____ include management and monitoring of service levels and service-level agreements.	Cloud access service	Cloud Information Security	Cloud security management	Cloud access security	Cloud security management
53	In general, a computer security awareness and training program should encompass ____steps.	6	7	5	9	7
54	There are___ types of training are related to cloud security.	7	4	6	5	5
55	_____ can be viewed as a computer security architecture.	Cloud Information Security	Trusted cloud computing	Cloud security management	Cloud access security	Trusted cloud computing

UNIT -V

Case Study on Open Source and Commercial Clouds: Microsoft Azure- Amazon EC2-Google Web services – Open Nebula.

Case Study on Open Source and Commercial Clouds:

Microsoft Azure

Microsoft Azure is an ever-expanding set of cloud services to help your organisation meet your business challenges. It is the freedom to build, manage and deploy applications on a massive, global network using your favorite tools and frameworks.

The Essentials

Amazon's AWS has a range of offerings that fall under IaaS, and each of these is categorized into four classes:

- content delivery and storage,
- compute,
- networking, and
- database.

No matter which IaaS offering you get, you will be using Amazon's identity and security services such as AWS CloudHSM's key storage service and Amazon's own Active Directory. Not only that, but AWS offerings also have a range of management tools that users can use, including AWS Config, AWS Cloudtrail, and Cloudwatch.

Azure, on the other hand, also has four classes of offerings:

- Data management and databases,
- compute,
- networking, and
- performance.

Security and management tools include Active Directory Federation Services, Azure Active Directory, Multi-Factor Auth, among others, as well as a range of integrations for Azure monitoring and performance tweaks.

Azure has multiple app deployment options for developers. Including App Services, Cloud Services, Service Fabric, Container Service, Functions, Batch, WebJobs and more. No matter what type of application you are developing, Microsoft has great tools in place to help deploy and scale it.

AWS offers similar solutions with Container Service, Elastic Beanstalk, Lambda, and Batch. AWS does not have as many options or features on the app hosting side. Microsoft has flexed their knowledge of developer tools to have a little bit of an advantage for hosting cloud apps.

Containers seem to be the preferred mechanism to deploy apps in the future, especially for open source applications. Look for more and more advancements in hosting containerized apps in the cloud. **Hybrid clouds are easier with Azure**, partly because Microsoft has foreseen the need for hybrid clouds early on. Azure offers substantial support for hybrid clouds, where you can use your onsite servers to run your applications on the Azure Stack. You can even set your computer resources to tap cloud-based resources when necessary.

Amazon EC2

Amazon Elastic Compute Cloud (Amazon EC2) provides scalable computing capacity in the Amazon Web Services (AWS) cloud. Using Amazon EC2 eliminates your need to invest in hardware up front, so you can develop and deploy applications faster. You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure security and networking, and manage storage. Amazon EC2 enables you to scale up or down to handle changes in requirements or spikes in popularity, reducing your need to forecast traffic.

Features of Amazon EC2

Amazon EC2 provides the following features:

- Virtual computing environments, known as *instances*
- Preconfigured templates for your instances, known as *Amazon Machine Images (AMIs)*, that package the bits you need for your server (including the operating system and additional software)

- Various configurations of CPU, memory, storage, and networking capacity for your instances, known as *instance types*
- Secure login information for your instances using *key pairs* (AWS stores the public key, and you store the private key in a secure place)
- Storage volumes for temporary data that's deleted when you stop or terminate your instance, known as *instance store volumes*
- Persistent storage volumes for your data using Amazon Elastic Block Store (Amazon EBS), known as *Amazon EBS volumes*
- Multiple physical locations for your resources, such as instances and Amazon EBS volumes, known as *regions* and*Availability Zones*
- A firewall that enables you to specify the protocols, ports, and source IP ranges that can reach your instances using*security groups*
- Static IPv4 addresses for dynamic cloud computing, known as *Elastic IP addresses*
- Metadata, known as *tags*, that you can create and assign to your Amazon EC2 resources
- Virtual networks you can create that are logically isolated from the rest of the AWS cloud, and that you can optionally connect to your own network, known as *virtual private clouds* (VPCs)

Using Google Web Services

Google is the prototypical cloud computing services company, and it supports some of the largest Web sites and services in the world. Google uses automated technology to index the Web. It makes its search service available to users as a standard search engine and to developers as a collection of special search tools limited to various areas of content. The application of Google's searches to content aggregation has led to enormous societal changes and to a growing trend of disintermediation.

The most important commercial part of Google's activities is its targeting advertising business: AdWords and AdSense. Google has developed a range of services including Google

Analytics that supports its targeted advertising business.

Google applications are cloud-based applications. The range of application types offered by Google spans a variety of types: productivity applications, mobile applications, media delivery, social interactions, and many more.

Exploring Google Applications

The bulk of Google's income comes from the sales of target advertising based on information that Google gathers from your activities associated with your Google account or through cookies placed on your system using its AdWords system. The company is highly profitable, and that has allowed Google to create a huge infrastructure as well as launch many free cloud-based applications and services that this chapter details. These applications are offered mostly on a free usage model that represents Google's Software as a Service portfolio. A business model that offers cloud-based services for free that are ―good enough‖ is very compelling. While Google is slowly growing a subscription business selling these applications to enterprises, its revenue represents only a small but growing part of Google's current income.

Google's cloud computing services falls under two umbrellas. The first and best-known offerings are an extensive set of very popular applications that Google offers to the general public. These applications include Google Docs, Google Health, Picasa, Google Mail, Google Earth, and many more.

Google's cloud-based applications have put many other vendors' products—such as office suites, mapping applications, image-management programs, and many other categories of traditional shrink-wrapped software—under considerable pressure.

The second of Google's cloud offerings is its Platform as a Service developer tools. In April 2008, Google introduced a development platform for hosted Web applications using Google's

infrastructure called the Google App Engine (GAE). The goal of GAE is to allow developers to create and deploy Web applications without worrying about managing the infrastructure necessary to have their applications run. GAE applications may be written using many high-level programming languages (most prominently Java and Python) and the Google App Engine Framework, which lowers the amount of development effort required to get an application up and running. Goggle also allows a certain free level of service so that the application must exceed a certain level of processor load, storage usage, and network bandwidth (Input/Output) before charges are assessed.

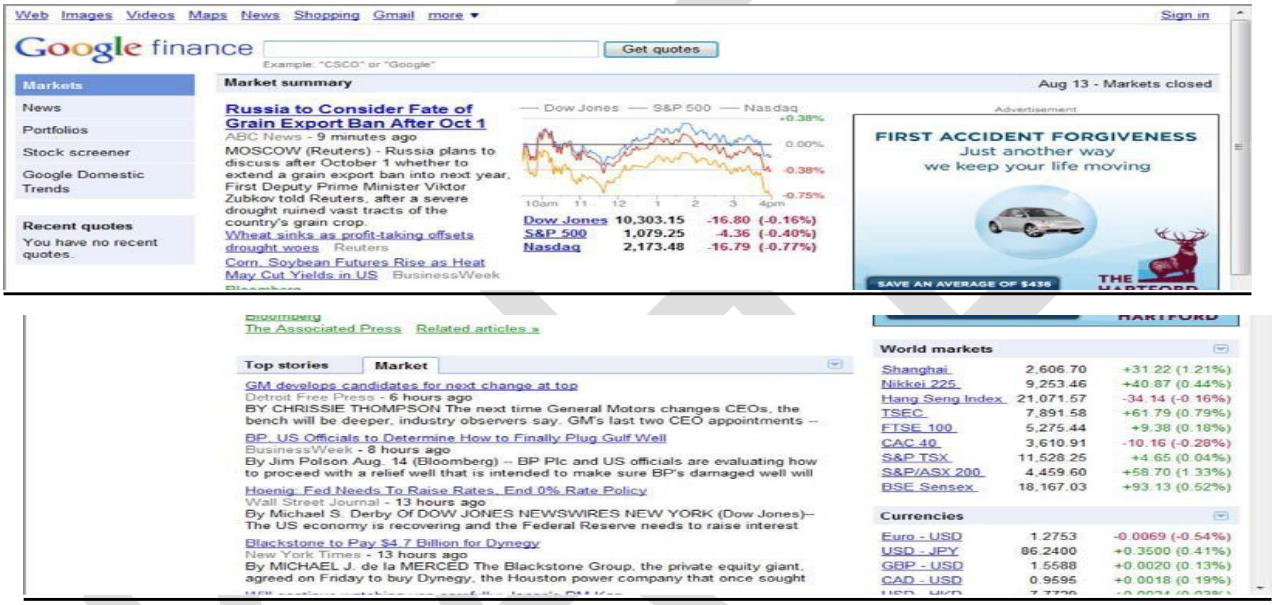
More Google Products equals fewer commercial products.

Surveying the Google Application Portfolio

It is fair to say that nearly all the products in Google's application and service portfolio are cloud computing services in that they all rely on systems staged worldwide on Google's one million plus servers in nearly 30 datacenters. Roughly 17 of the 48 services listed leverage Google's search engine in some specific way. Some of these search-related sites search through selected content such as Books, Images, Scholar, Trends, and more. Other sites such

as Blog Search, Finance, News, and some others take the search results and format them into an Aggregation page. The below figure shows one of these aggregation pages: Google Finance.

Google's Finance page at http://www.google.com/finance/ is an example of an aggregation page provided by results from Google's search engine.



Enterprise offerings

As Google has built out its portfolio, it has released special versions of its products for the enterprise. The following are among Google's products aimed at the enterprise market:

• **Google Commerce Search** (http://www.google.com/commercesearch/): This is a search service for online retailers that markets their products in their site searches with a number of navigation, filtering, promotion, and analytical functions.

• **Google Site Search** (http://www.google.com/sitesearch/): Google sells its search engine customized for enterprises under the Google Site Search service banner. The user enters a search string in the site's search, and Google returns the results from that site.

• **Google Search Appliance** (http://www.google.com/enterprise/gsa): This server can be deployed within an organization to speed up both local (Intranet) and Internet searching. The three versions of the Google Search Appliance can store an index of up to 300,000 (GB-1001), 10 million (GB-5005), or 30 million (GB-8008) documents. Beyond indexing, these appliances have document management features, perform custom searches, cache content, and give local support to Google Analytics and Google Sitemaps.

• **Google Mini** (http://www.google.com/enterprise/mini/): The Mini is the smaller version of the GSA that stores 300,000 indexed documents.

Google Apps for Business is the commercial versions of the company's productivity suites.

Many of Google's productivity applications are quite capable, but none is a state-of-the-art client you might expect to find in a locally installed office suite. When compared one-on-one to Microsoft Office applications, Google's online offerings give users the essential features for a fraction of the Microsoft Office price.

AdWords

AdWords (http://www.google.com/AdWords) is a targeted ad service based on matching

advertisers and their keywords to users and their search profiles. This service transformed Google from a competent search engine into an industry giant and is responsible for the majority of Google's revenue stream. AdWords' two largest competitors are Microsoft adcenter (http://adcenter microsoft.com/) and Yahoo! Search Marketing (http://searchmarketing.yahoo.com/).

Ads are displayed as text, banners, or media and can be tailored based on geographical location, frequency, IP addresses, and other factors. AdWords ads can appear not only on Google.com, but on AOL search, Ask.com, and Netscape, along with other partners. Other partners belonging to the Google Display Network can also display AdSense ads. In all these cases, the AdWords system determines which ads to match to the user searches.

Using OpenNebula

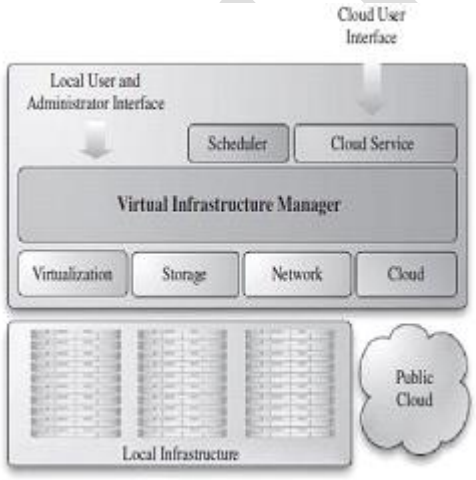
OpenNebula is an open and flexible tool that fits into existing data center's environments to build any type of cloud deployment. OpenNebula can be primarily used as a virtualization tool to manage your virtual infrastructure, which is usually referred to as private cloud. OpenNebula supports a hybrid cloud to combine local infrastructure with public cloud-based infrastructure, enabling highly scalable hosting environments. OpenNebula also supports public clouds by providing cloud's interfaces to expose its functionality for virtual machine, storage, and network management. OpenNebula is one of the technologies being enhanced in the Reservoir Project , European research initiatives in virtualized infrastructures, and cloud computing.

OpenNebula architecture is shown in Figure which illustrates the existence of public and private clouds and also the resources being managed by its virtual manager.

OpenNebula is an open-source alternative to these commercial tools for the dynamic management of VMs on distributed resources. This tool is supporting several research lines in advance reservation of capacity, probabilistic admission control, placement

optimization, resource models for the efficient management of groups of virtual machines, elasticity support, and so on. These research lines address the requirements from both types of clouds namely, private and public.

OpenNebula and Haizea. Haizea is an open-source virtual machine-based lease management architecture developed by Sotomayor et al. ; it can be used as a scheduling backend for OpenNebula. Haizea uses leases as a fundamental resource provisioning abstraction and implements those leases as virtual machines, taking into account the overhead of using virtual machines when scheduling leases.



Haizea also provides advanced functionality such as:

- Advance reservation of capacity.
- Best-effort scheduling with backfilling.
- Resource preemption (using VM suspend/resume/migrate).
- Policy engine, allowing developers to write pluggable scheduling policies in Python.

Aneka

Manjrasoft Aneka is a .NET-based platform and framework designed for building and deploying distributed applications on clouds. It provides a set of APIs for transparently exploiting distributed resources and expressing the business logic of applications by using the preferred programming abstractions. Aneka is also a market-oriented cloud

platform since it allows users to build and schedule applications, provision resources, and monitor results using pricing, accounting, and QoS/SLA services in private and/or public cloud environments.

It allows end users to build an enterprise/private cloud setup by exploiting the power of computing resources in the enterprise data centers, public clouds such as Amazon EC2 , and hybrid clouds by combining enterprise private clouds managed by Aneka with resources from Amazon EC2 or other enterprise clouds built and managed using technologies such as XenServer.

Aneka also provides support for deploying and managing clouds. By using its Management Studio and a set ofWeb interfaces, it is possible to set up either public or private clouds, monitor their status, update their configuration, and perform the basic management operations.

Aneka Architecture.

Aneka platform, consists of a collection of physical and virtualized resources connected through a network. Each of these resources hosts an instance of the Aneka container representing the runtime environment where the distributed applications are executed. The container provides the basic management features of the single node and leverages all the other operations on the services that it is hosting. The services are broken up into fabric, foundation, and execution services. Fabric services directly interact with the node through the platform abstraction layer (PAL) and perform hardware profiling and dynamic resource provisioning. Foundation services identify the core system of the Aneka middleware, providing a set of basic features to enable Aneka containers to perform specialized and specific sets of tasks. Execution services directly deal with the scheduling and execution of applications in the cloud.

POSSIBLE QUESTIONS

6 MARKS

1. Discuss about Google Web services with real-time environment.
2. Describe the role of Open Nebula in the cloud environment.
3. Discuss the features of Microsoft Azure in the cloud computing.
4. Elaborate the role of Amazon EC2 in the cloud.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established Under Section 3 of UGC Act 1956)

Coimbatore – 641 021.

ONE MARK QUESTIONS

DEPARTMENT OF CS, CA & IT

STAFF NAME: Dr.S.MANJU PRIYA **SUBJECT NAME: CLOUD COMPUTING**

SUB.CODE: 18CSP104 **UNIT V** **SEMESTER: I**

S.NO	Question	Choice1	Choice2	Choice3	Choice4	Ans
1	_____ is a cloud operating system built on top of Microsoft datacenters infrastructure	Microsoft Windows	Oracle Azure	VB Azure	Java Azure	Microsoft Windows Azure
2	_____ are the core components of Microsoft Windows Azure,	Storage services	Compute services	Product services	Infrastructure services	Compute services
3	Expand BLOBs	Binary Large Objects	Blocking Large Objects	Between Large Objects	Blocking Last Objects	Binary Large Objects

4	_____ is a comprehensive middleware for developing, deploying, and managing applica tions on the cloud	Microsoft Windows Azure	AppFabric	SQL Azure	Oracle Azure	AppFabric
5	_____is one of the most important and heavily trafficked Web sites in the world	Yahoo.com	Google.com	Amazon.com	MSN.com	Amazon.com
6	_____ is a Platform as a Service (PaaS) cloud-based Web hosting service on Google's infrastructure.	SQL Azure	Web Server Gateway	Amazon Web Services	Google App Engine	Google App Engine
7	_____is a service that allows developers to quickly access data persisted on Windows Azure storage.	Web Server Gateway	Azure Cache	Amazon Web Services	SQL Azure	Azure Cache
8	_____ supports databases with a maximum size of 1 GB or 5 GB.	Web Edition	Business Edition	Standard Edition	Special Edition	Web Edition
9	_____supports databases with a maximum size from 10 GB to 50 GB	Web Edition	Standard Edition	Business Edition	Special Edition	Business Edition
10	In Which Year did Amazon.com made its Web service platform available to developers on a usage-basis model.	2006	2005	2007	2008	2006
11	_____ is the world's largest online retailer with net sales in $24.51 billion, according to their 2009 annual report.	Google.com	Amazon.com	Force.com	Rackspace.com	Amazon.com
12	_____ is the hourly rate with no long-term commitment.	On-Demand Instance	Reserved Instances	Spot Instance	Timing Instance	On-Demand Instance
13	There are currently _____different EC2 service zones or regions	Three	Four	Five	Six	Four
14	_____is the prototypical cloud computing Services Company.	Google	Yahoo	Amazon	IBM	Google

15	_____supports some of the largest Web sites and services in the world	Yahoo	Amazon	Google	IBM	Google
16	SEO stands for_____	Search Engine Optimization	Small Engine Optimizatio	Secret Engine Optimization	Sophisticated Engine Optimization	Search Engine Optimization
17	Online content that isn't indexed by search engines belongs to what has come to be called the_____	Crawl Web	Deep Web	Dark Web	Open Web	Deep Web
18	World's number two Web site, is called_____	Facebook	Google	Yahoo	MSN	Facebook
19	_____ is a prominent example of a site that isn't indexed in search engines	Facebook	Google	Yahoo	MSN	Facebook
20	The success of the ad is measured by what is called the_____	Money through Click	Click-through rate	Minimum Click through	Maximum through Click	Click-through rate
21	_____ supports Dynamic Web services based on common standards.	SQL Azure	Azure Cloud	Google App Engine	Amazon EC2	Google App Engine
22	_____ can be used to run and scale PHP Web applications on Azure.	Worker Roles	Web Roles	Virtual machine Roles	Storage Roles	Web Roles
23	_____ can be used to host Tomcat and serve JSP-based applications	Web Roles	Virtual machine	Worker roles	Storage Roles	Worker roles
24	A single block blob can reach ___ in dimension	100 GB	200 GB	150 GB	160 GB	200 GB
25	_____type of blob is optimized for random access and can be used to host data different from streaming	Page blobs	Block blobs	Storage blobs	Cluster blobs	Page blobs

26	Access to SQL Azure is based on the _____ protocol	User Datagram	Transmission Control	Tabular Data Stream	Stream Oriented	Tabular Data Stream
27	_____ is foundation layer + set of developer services	Azure Platform	Google Platform	Amazon Platform	Azure Infrastructure	Azure Platform
28	_____provides access to e-mail and display names within your app	Amazon web services	App Engine	SQL Azure	Azure Cloud	App Engine
29	_____eliminates the need for an application to develop its own authentication system	Amazon web services	SQL Azure	App Engine	Azure Cloud	App Engine
30	_____has a distributed datastore system that supports queries and transactions.	Google App Engine	Azure Cloud	SQL Azure	Amazon EC2	Google App Engine
31	The Datastore in Google App Engine is _____	Non-relational	Relational	Schema based	Standard	Non-relational
32	The _____ uses an optimistic concurrency control and maintains strong consistency.	Entity Group	User API	Datastore	Attribute Group	Datastore
33	_____manage entities as a single group, and entity groups are stored together in the system so operations can be performed faster.	Queries	Datastore	Entity Group	Transactions	Transactions
34	Applications can use the_____ to determine whether a user belongs to a specific group	Admin API	User API	Third party API	API	User API
35	Applications running in GAE are isolated from the underlying operating system, which Google describes as running in a _____	Sandbox	Web pages	Well known ports	Protocols	Sandbox
36	The pricing scheme of Google in the Outgoing bandwidth measured in GB is _____ per GB.	$0.20	$0.22	$0.10	$0.12	$0.12

37	The pricing scheme of Google in the Incoming bandwidth measured in GB is ____ per GB	$0.20	$0.22	$0.10	$0.12	$0.10
38	The pricing scheme of Google in the Stored data measured in GB per month is ____ per GB/month	$0.20	$0.10	$0.12	$0.15	$0.15
39	The pricing scheme of Google in the CPU time measured in CPU hours is _____ per hour.	$0.10	$0.22	$0.20	$0.12	$0.10
40	WSGI stands for_____	Window Server	Web Server Gateway	Web Server Gateway	Window Server	Web Server Gateway
41	_____ supports the feature Task queues and task scheduling	Google App Engine	Azure Cloud	SQL Azure	Amazon EC2	Google App Engine
42	_____ can appear not only on Google.com, but on AOL search, Ask.com, and Netscape, along with other partners	AdSense ads	AdDeep web ads	AdWords ads	AdServer ads	AdWords ads
43	The ____system determines which ads to match to the user searches	AdWords	AdSense	AdDeep web	AdServer	AdWords
44	Advertisers bid on _____that are used to match a user to their product or service	Exact words	Keywords	Retrieved words	Keys	Keywords
45	_____can be deployed within an organization to speed up both local (Intranet) and Internet searching	Google Site Search	Google Commerce Search	Google Search Appliance	Google Mini	Google Search Appliance

46	_____is the smaller version of the Google Search Appliance .	Google Mini	Google Commerce	Google Site	Google App	Google Mini
47	_____ includes Pages without links.	Crawl Web	Deep Web	Dark Web	Open Web	Deep Web
48	_____ includes Private or limited access Web pages and sites	Crawl Web	Dark Web	Deep Web	Open Web	Deep Web
49	_____ includes Information contained in sources available through executable code such as JavaScript	Deep Web	Crawl Web	Dark Web	Open Web	Deep Web
50	_____ can be useful in allowing content that isn't browsable to be crawled	Deep Web	Crawl Web	Dark Web	Sitemaps	Sitemaps
51	Web crawlers are also called as _____	Spiders or Robots	Dark Robots	Dark Spiders	Shadowers	Spiders or Robots
52	Content on pages is scanned up to a certain number of words and placed into an _____	Header	Index	Table	Database	Index
53	_____is an example of an aggregation page provided by results from Google's search engine	Google's app page	Google's Store page	Google's Finance page	Google's Box page	Google's Finance page
54	_____which lowers the amount of development effort required to get an application up and running.	Azure Cloud Framework	Google App Engine Framework	SQL Azure Framework	Amazon EC2 Framework	Google App Engine Framework

55	Google's cloud computing services falls under_____umbrellas	2	3	4	5	2
56	The bulk of Google's income comes from the sales of_____	Social Networking	Applications	Target advertising	Blogger	Target advertising
57	Google is always tweaking the algorithm to prevent _____strategies from gaming the system	Search Engine Results Page	Search Engine Data	Search Engine	Search Engine Optimization	Search Engine Optimization
58	_____represent the units of deployment of Web applications within the Azure infrastructure	Web Roles	Virtual machine	Worker roles	Storage Roles	Web Roles
59	_____service is optimal to store large text or binary files.	Blocks	Blobs	Storage	Cluster	Blobs
60	A connection is the Service Bus element that is priced by Azure on a _____basis	Yearly	Weekly	Pay-as-you-go	Monthly	Pay-as-you-go
61	is a .NET-based platform and framework designed for building and deploying distributed applications on clouds	Haizea	Manjrasoft Aneka	AdWords	google Analytics	Manjrasoft Aneka
62	____clouds are easier with Azure	Private	Public	Community	Hybrid	Hybrid
63	Static IPv4 addresses for dynamic cloud computing, is known as ____ IP addresses	dynamic	elastic	portable	static	elastic

Reg.No _____

[18CSP104]

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established Under Section 3 of UGC Act 1956)

For the candidates admitted in 2018 onwards

First Semester

M.Sc COMPUTER SCIENCE

FIRST INTERNAL EXAMINATION – SEPTEMBER 2018

Cloud Computing

Class : I M.Sc CS **Time: 2 hrs**

Date & Session: 12.09.18 & AN **Marks: 50**

PART – A (20 * 1 = 20 Marks)

1. _____ is a complete operating environment with applications, management, and the user interface.

a. CaaS b. PaaS c. IaaS d. **SaaS**

2. ____model refers to the location and management of the cloud's infrastructure.

a. Service b. Deployment c. Development d. Business

3. _____ cloud is used for healthcare industry.

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

4. _____ virtualizes storage into storage clouds.

a. SystemGRID b. StorageGRID c. DataGRID d. GoGRID

5. The work done in IaaS can be measured by the number of _____per minute.

a. Data b. Process **c. Transactions** d. Clients

6. _____provide the capability of running multiple machine instances, each with their own operating system.

a. MultiMachines b. Shared machines c. Local machines d. **Virtual machines**

7. The latest version of VMI is _____.

a. 2.1 b. 2.2 c. 1.9 d. 2

8. An _____ is considered to be an advanced version of a load balancer.

a. ADN **b. ADC** c. ACD d. AND

9. Gmail is an offering of _____.

a. DaaS b. SaaS **c. PaaS** d.CaaS

10. SPML standards is used for_____.

a. Provisioning b. Audit c. Authentication d. Authorization

11. The _____ standard applies to the unique identity of the URL.

a. OpenID b. SPML c. SAML d. XACML

12. _____ is one of the large IaaS cloud service providers.

a. Rackspace.com b. Salesforce.com c. GoGrid.com d. Openstack.com

13. Which of the following is not an Network interfaces and services?

a. FTP b. DNS **c. Cloud computing** d. HTTP

14. A fundamental characteristic of public clouds is _____.

a. Security b. High bandwidth c. QoS **d. multi tenancy**

15. Services are billed on a usage basis are called_____.

a. **Metered usage** b. Billed usage c. Daily usage d. Rented usage

16. _____ is not an OS.

a. Sun Solaris b. IBM AIX 6.1 c. Linux **d. vNetwork**

17. Customer information protection is an aspect of _____.

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

18. C2D stands for_____.

a**. Cloud to Datacenter** b. Cloud to Data

c. Cloud to Development d. Cloud to Distributor

19. FedCloud is used for_____.

a. Banking industry b. Medical industry **c. Government** d. Merchant transactions

20. Without _____, cloud computing would be very difficult to manage.

a. Virtualization b. Hypervisors c**. Load balancing** d. Vmware

PART – B (3* 2 = 6 Marks)

21. What is cloud computing?

Cloud Computing refers to **manipulating, configuring,** and **accessing** the hardware and software resources remotely. It offers online data storage, infrastructure, and application.

22. List the benefits of IaaS.

IaaS allows the cloud provider to freely locate the infrastructure over the Internet in a cost-effective manner. Some of the key benefits of IaaS are listed below:

- Full control of the computing resources through administrative access to VMs.
- Flexible and efficient renting of computer hardware.
- Portability, interoperability with legacy applications.

23. Mention the applications of Google cloud.

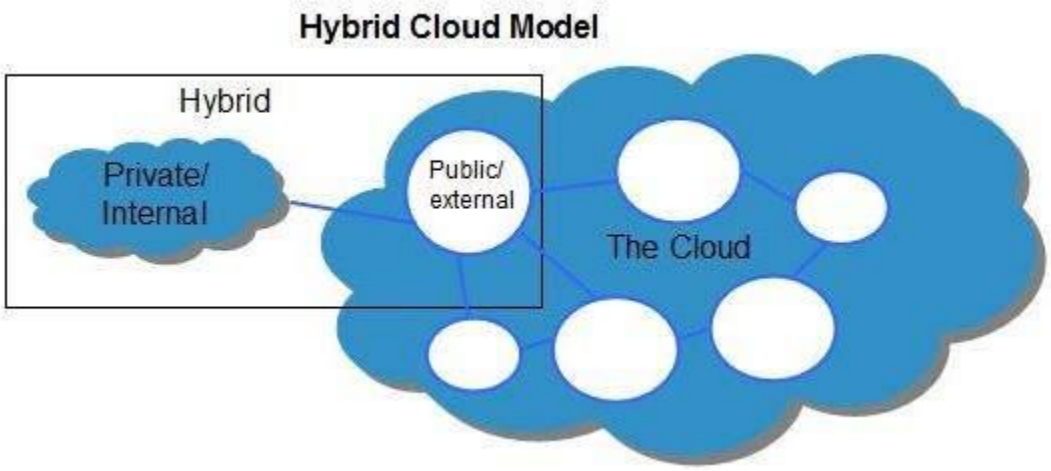
Google Cloud Platform is a suite of public cloud computing services offered by Google. The platform includes a range of hosted services for compute, storage and application development that run on Google hardware. Google Cloud Platform offers application development and integration services. For example, Google Cloud Pub/Sub is a managed and real-time messaging service that allows messages to be exchanged between applications.

PART – C (3 * 8 = 24 Marks)

21. a. Explain (i) Hybrid cloud (ii) Community cloud (Or)

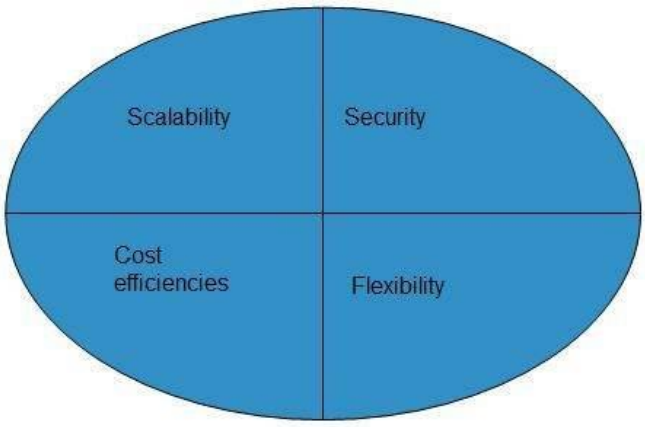
Hybrid cloud

Hybrid Cloud is a mixture of **public** and **private** cloud. Non-critical activities are performed using public cloud while the critical activities are performed using private cloud. The Hybrid Cloud Model is shown in the diagram below.



Benefits

There are many benefits of deploying cloud as hybrid cloud model. The following diagram shows some of those benefits:



Scalability

It offers features of both, the public cloud scalability and the private cloud scalability.

Flexibility

It offers secure resources and scalable public resources.

Cost Efficiency

Public clouds are more cost effective than private ones. Therefore, hybrid clouds can be cost saving.

Security

The private cloud in hybrid cloud ensures higher degree of security.

Disadvantages

Networking Issues

Networking becomes complex due to presence of private and public cloud.

Security Compliance

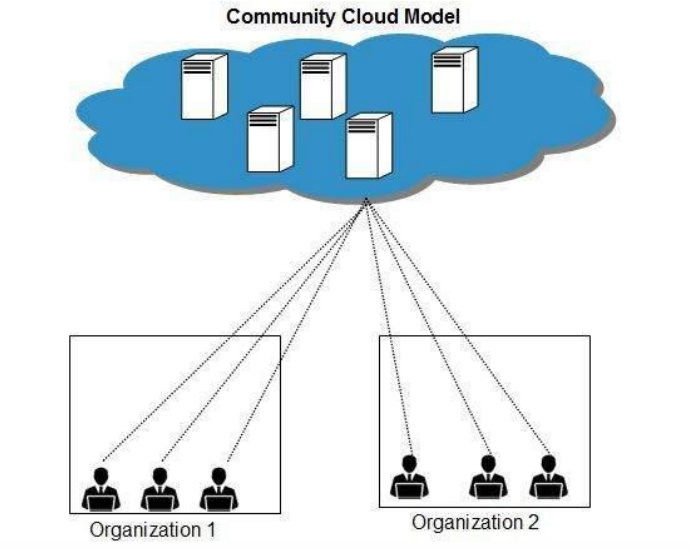
It is necessary to ensure that cloud services are compliant with security policies of the organization.

Infrastructure Dependency

The **hybrid cloud model** is dependent on internal IT infrastructure, therefore it is necessary to ensure redundancy across data centers.

Community Cloud

Community Cloud allows system and services to be accessible by group of organizations. It shares the infrastructure between several organizations from a specific community. It may be managed internally by organizations or by the third-party. The Community Cloud Model is shown in the diagram below.



Benefits

There are many benefits of deploying cloud as **community cloud model.**

Cost Effective

Community cloud offers same advantages as that of private cloud at low cost.

Sharing Among Organizations

Community cloud provides an infrastructure to share cloud resources and capabilities among several organizations.

Security

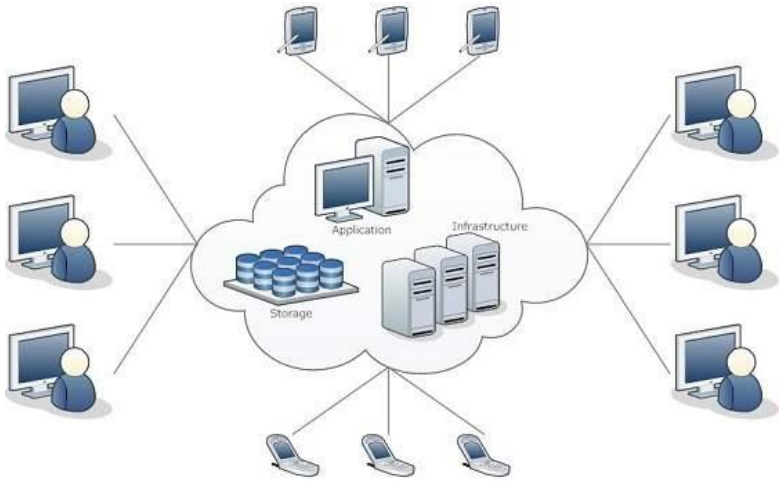
The community cloud is comparatively more secure than the public cloud but less secured than the private cloud.

Issues

- Since all data is located at one place, one must be careful in storing data in community cloud because it might be accessible to others.
- It is also challenging to allocate responsibilities of governance, security and cost among organizations.

b. Describe the characteristics and advantages of cloud computing

Cloud Computing refers to **manipulating, configuring,** and **accessing** the hardware and software resources remotely. It offers online data storage, infrastructure, and application.



Cloud computing offers **platform independency,** as the software is not required to be installed locally on the PC. Hence, the Cloud Computing is making our business applications **mobile** and **collaborative.**

Characteristics of Cloud Computing

An IT environment requires a specific set of characteristics to enable the remote provisioning of scalable and measured IT resources in an effective manner. These

characteristics need to exist to a meaningful extent for the IT environment to be considered an effective cloud.

The following six specific characteristics are common to the majority of cloud environments:

- On-Demand Usage
- Ubiquitous Access
- Multi-tenancy (Resourcing Pooling)
- Elasticity (and Scalability)
- Measured Usage
- Resiliency

Cloud providers and cloud consumers can assess these characteristics individually and collectively to measure the value offering of a given cloud platform. Although cloud-based services and IT resources will inherit and exhibit individual characteristics to varying extents, usually the greater the degree to which they are supported and utilized, the greater the resulting value proposition.

On-Demand Usage

A cloud consumer can unilaterally access cloud-based IT resources giving the cloud consumer the freedom to self-provision these IT resources. Once configured, usage of the self-provisioned IT resources can be automated, requiring no further human involvement by the cloud consumer or cloud provider. This results in an *on-demand usage* environment. Also known as "on-demand self-service usage," this characteristic enables the service-based and usage-driven features found in mainstream clouds.

Ubiquitous Access

Ubiquitous Access represents the ability for a cloud service to be widely accessible. Establishing ubiquitous access for a cloud service can require support for a range of devices, transport protocols, interfaces, and security technologies. To enable this level of access generally requires that the cloud service architecture be tailored to the particular needs of different cloud service consumers.

Multitenancy (and Resource Pooling)

The characteristic of a software program that enables an instance of the program to serve different consumers (tenants) whereby each is isolated from the other, is referred to as *multitenancy*. A cloud provider pools its IT resources to serve multiple cloud service consumers by using multitenancy models that frequently rely on the use of virtualization technologies. Through the use of multitenancy technology, IT resources can be dynamically assigned and reassigned, according to cloud service consumer demands.

Resource pooling allows cloud providers to pool large-scale IT resources to serve multiple cloud consumers. Different physical and virtual IT resources are dynamically assigned and reassigned according to cloud consumer demand, typically followed by execution through statistical multiplexing. Resource pooling is commonly achieved through multitenancy technology, and therefore encompassed by this multitenancy characteristic.

Figures 1 and 2 illustrate the difference between single-tenant and multitenant environments.

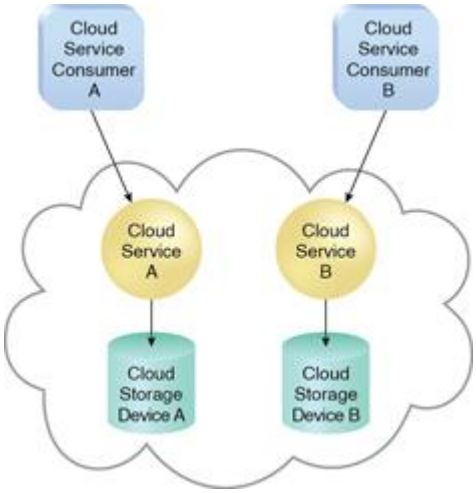


Figure 1 - In a single-tenant environment, each cloud consumer has a separate IT resource instance.

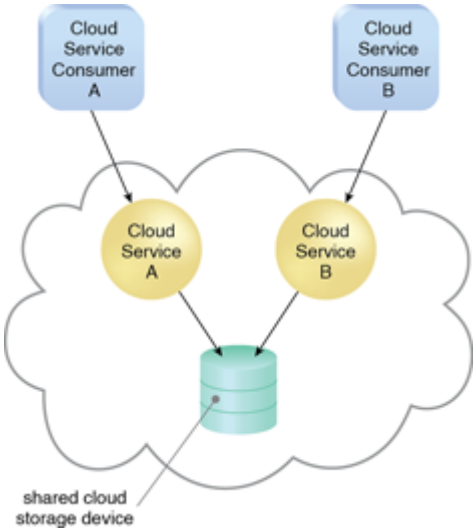


Figure 2 - In a multitenant environment, a single instance of an IT resource, such as a cloud storage device, serves multiple consumers.

As illustrated in Figure 1, multitenancy allows several cloud consumers to use the same IT resource or its instance while each remains unaware that it may be used by others.

Elasticity

Elasticity is the automated ability of a cloud to transparently scale IT resources, as required in response to runtime conditions or as pre-determined by the cloud consumer or cloud provider. Elasticity is often considered a core justification for the adoption of cloud computing, primarily due to the fact that it is closely associated with the Reduced Investment and Proportional Costs benefit. Cloud providers with vast IT resources can offer the greatest range of elasticity.

Measured Usage

The *measured usage* characteristic represents the ability of a cloud platform to keep track of the usage of its IT resources, primarily by cloud consumers. Based on what is measured, the cloud provider can charge a cloud consumer only for the IT resources actually used and/or for the timeframe during which access to the IT resources was granted. In this context, measured usage is closely related to the on-demand characteristic.

Measured usage is not limited to tracking statistics for billing purposes. It also encompasses the general monitoring of IT resources and related usage reporting (for both cloud provider and cloud consumers).

Resiliency

Resilient computing is a form of failover that distributes redundant implementations of IT resources across physical locations. IT resources can be pre-configured so that if one becomes deficient, processing is automatically handed over to another redundant implementation. Within cloud computing, the characteristic of resiliency can refer to redundant IT resources within the same cloud (but in different physical locations) or across multiple clouds. Cloud consumers can increase both the reliability and availability of their applications by leveraging the resiliency of cloud-based IT resources

Benefits of cloud computing

Cloud Computing has numerous advantages. Some of them are listed below -

- One can access applications as utilities, over the Internet.
- One can manipulate and configure the applications online at any time.
- It does not require to install a software to access or manipulate cloud application.
- Cloud Computing offers online development and deployment tools, programming runtime environment through **PaaS model.**
- Cloud resources are available over the network in a manner that provide platform independent access to any type of clients.
- Cloud Computing offers **on-demand self-service.** The resources can be used without interaction with cloud service provider.
- Cloud Computing is highly cost effective because it operates at high efficiency with optimum utilization. It just requires an Internet connection
- Cloud Computing offers load balancing that makes it more reliable.

22. a. Compare the characteristics of PaaS and SaaS **(Or)**

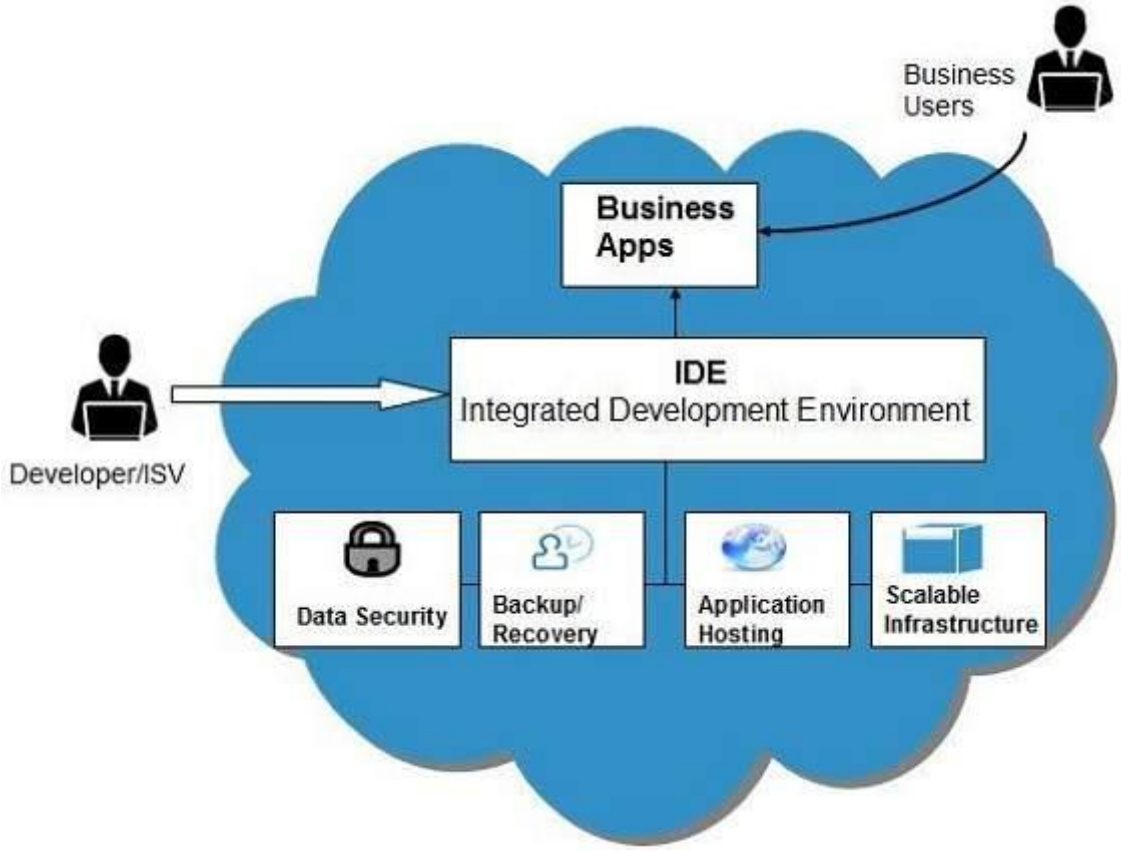
Platform as a Service (PaaS)

Platform-as-a-Service offers the runtime environment for applications. It also offers development and deployment tools required to develop applications. PaaS has a feature of **point-and-click** tools that enables non-developers to create web applications.

App Engine of Google and **Force.com** are examples of PaaS offering vendors. Developer may log on to these websites and use the **built-in API** to create web-based applications.

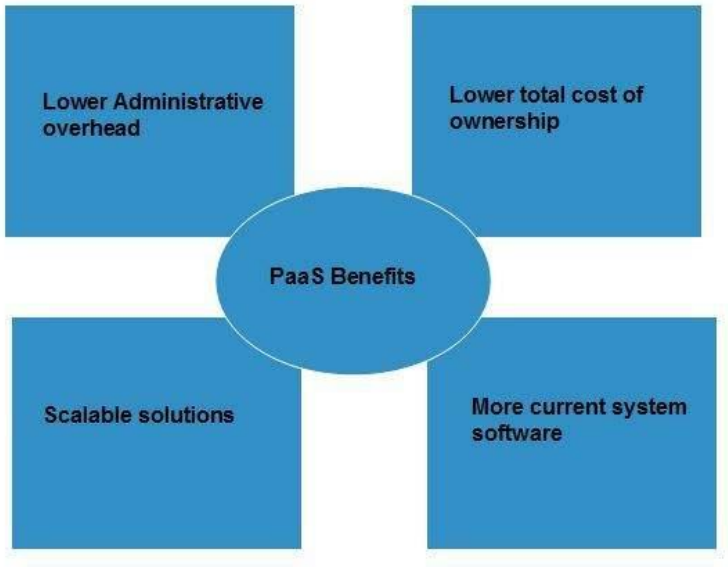
But the disadvantage of using PaaS is that, the developer **locks-in** with a particular vendor. For example, an application written in Python against API of Google, and using App Engine of Google is likely to work only in that environment.

The following diagram shows how PaaS offers an API and development tools to the developers and how it helps the end user to access business applications.



Benefits

Following are the benefits of PaaS model:



Lower administrative overhead

Customer need not bother about the administration because it is the responsibility of cloud provider.

Lower total cost of ownership

Customer need not purchase expensive hardware, servers, power, and data storage.

Scalable solutions

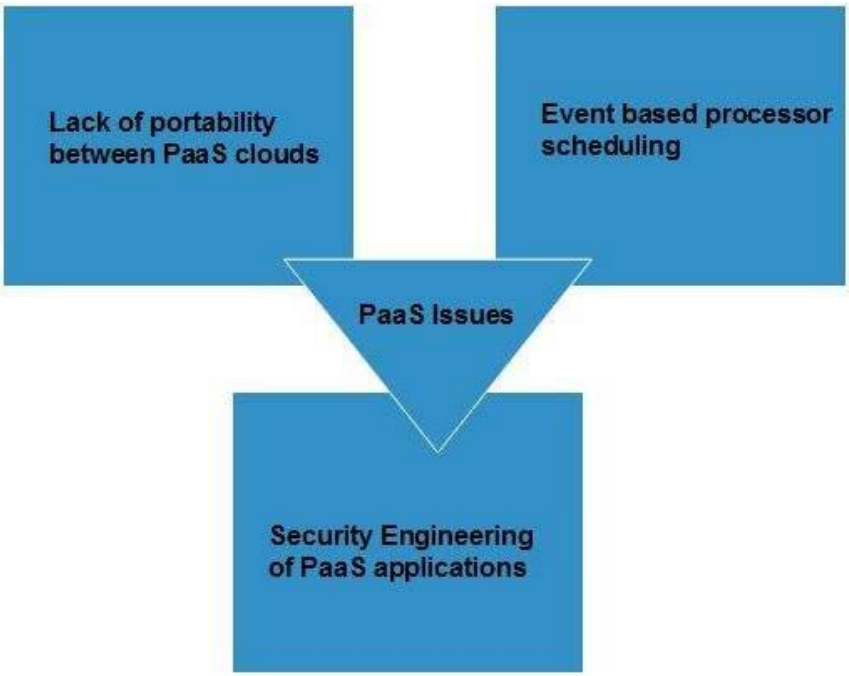
It is very easy to scale the resources up or down automatically, based on their demand.

More current system software

It is the responsibility of the cloud provider to maintain software versions and patch installations.

Issues

Like **SaaS, PaaS** also places significant burdens on customer's browsers to maintain reliable and secure connections to the provider's systems. Therefore, PaaS shares many of the issues of SaaS. However, there are some specific issues associated with PaaS as shown in the following diagram:



Lack of portability between PaaS clouds

Although standard languages are used, yet the implementations of platform services may vary. For example, file, queue, or hash table interfaces of one platform may differ from another, making it difficult to transfer the workloads from one platform to another.

Event based processor scheduling

The PaaS applications are event-oriented which poses resource constraints on applications, i.e., they have to answer a request in a given interval of time.

Security engineering of PaaS applications

Since PaaS applications are dependent on network, they must explicitly use cryptography and manage security exposures.

Characteristics

Here are the characteristics of PaaS service model:

- PaaS offers **browser based development environment.** It allows the developer to create database and edit the application code either via Application Programming Interface or point-and-click tools.
- PaaS provides **built-in security, scalability,** and **web service interfaces.**
- PaaS provides built-in tools for defining **workflow, approval processes,** and business rules.

- It is easy to integrate PaaS with other applications on the same platform.
- PaaS also provides web services interfaces that allow us to connect the applications outside the platform.

Software as a Service (SaaS)

Software-as–a-Service (SaaS) model allows to provide software application as a service to the end users. It refers to a software that is deployed on a host service and is accessible via Internet. There are several SaaS applications listed below:

- Billing and invoicing system
- Customer Relationship Management (CRM) applications
- Help desk applications
- Human Resource (HR) solutions

Some of the SaaS applications are not customizable such as **Microsoft Office Suite.** But SaaS provides us **Application Programming Interface (API),**which allows the developer to develop a customized application.

Characteristics

Here are the characteristics of SaaS service model:

- SaaS makes the software available over the Internet.
- The software applications are maintained by the vendor.
- The license to the software may be subscription based or usage based. And it is billed on recurring basis.
- SaaS applications are cost-effective since they do not require any maintenance at end user side.
- They are available on demand.
- They can be scaled up or down on demand.
- They are automatically upgraded and updated.
- SaaS offers shared data model. Therefore, multiple users can share single instance of infrastructure. It is not required to hard code the functionality for individual users.
- All users run the same version of the software.

Benefits

Using SaaS has proved to be beneficial in terms of scalability, efficiency and performance. Some of the benefits are listed below:

- Modest software tools
- Efficient use of software licenses
- Centralized management and data
- Platform responsibilities managed by provider
- Multitenant solutions

Modest software tools

The SaaS application deployment requires a little or no client side software installation, which results in the following benefits:

- No requirement for complex software packages at client side
- Little or no risk of configuration at client side
- Low distribution cost

Efficient use of software licenses

The customer can have single license for multiple computers running at different locations which reduces the licensing cost. Also, there is no requirement for license servers because the software runs in the provider's infrastructure.

Centralized management and data

The cloud provider stores data centrally. However, the cloud providers may store data in a decentralized manner for the sake of redundancy and reliability.

Platform responsibilities managed by providers

All platform responsibilities such as backups, system maintenance, security, hardware refresh, power management, etc. are performed by the cloud provider. The customer does not need to bother about them.

Multitenant solutions

Multitenant solutions allow multiple users to share single instance of different resources in virtual isolation. Customers can customize their application without affecting the core functionality.

Issues

There are several issues associated with SaaS, some of them are listed below:

- Browser based risks
- Network dependence
- Lack of portability between SaaS clouds

Browser based risks

If the customer visits malicious website and browser becomes infected, the subsequent access to SaaS application might compromise the customer's data.

To avoid such risks, the customer can use multiple browsers and dedicate a specific browser to access SaaS applications or can use virtual desktop while accessing the SaaS applications.

Network dependence

The SaaS application can be delivered only when network is continuously available. Also network should be reliable but the network reliability cannot be guaranteed either by cloud provider or by the customer.

b. Discuss the role of Cloud Storage.

Cloud storage

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

Storage Devices

Storage devices can be broadly classified into two categories:

- Block Storage Devices
- File Storage Devices

Block Storage Devices

The **block storage devices** offer raw storage to the clients. These raw storage are partitioned to create volumes.

File Storage Devices

The **file Storage Devices** offer storage to clients in the form of files, maintaining its own file system. This storage is in the form of Network Attached Storage (NAS).

Cloud Storage Classes

Cloud storage can be broadly classified into two categories:

- Unmanaged Cloud Storage
- Managed Cloud Storage

Unmanaged Cloud Storage

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

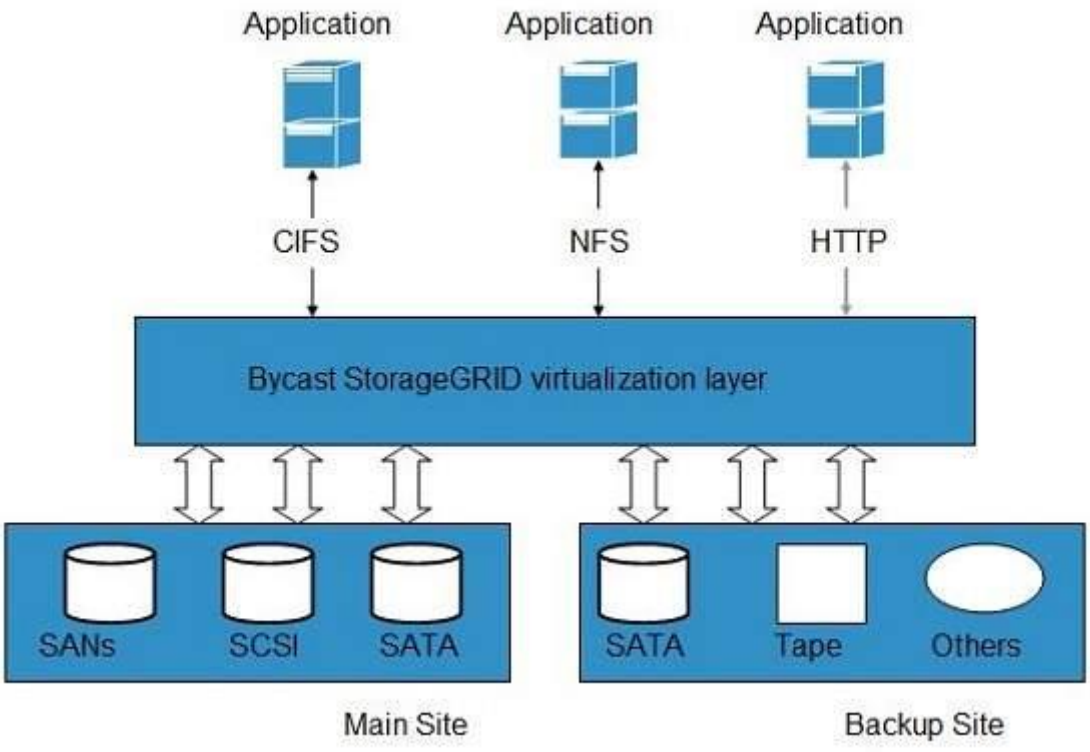
Managed Cloud Storage

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

Creating Cloud Storage System

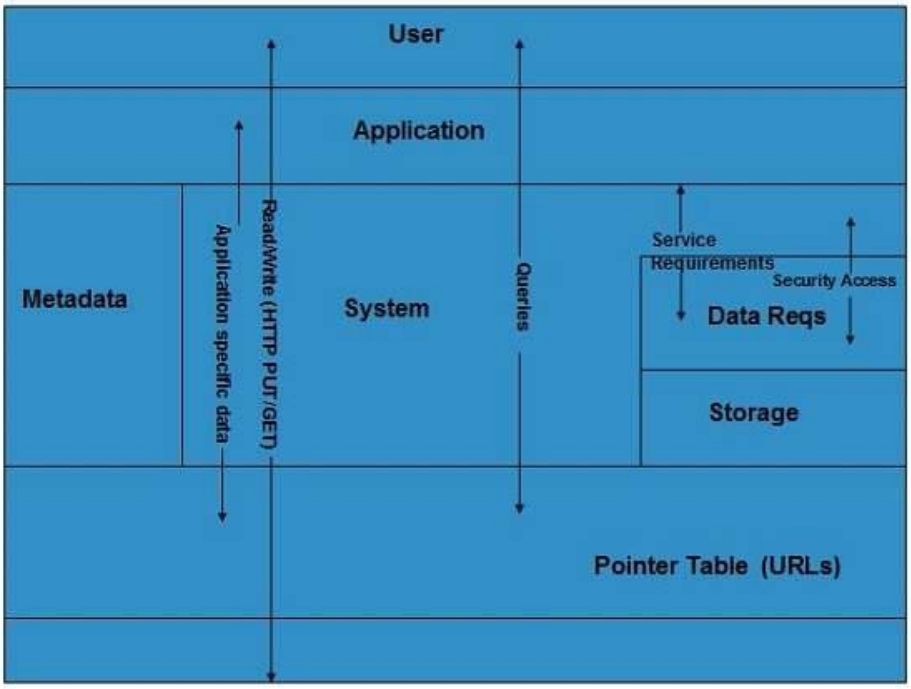
The cloud storage system stores multiple copies of data on multiple servers, at multiple locations. If one system fails, then it is required only to change the pointer to the location, where the object is stored.

To aggregate the storage assets into cloud storage systems, the cloud provider can use storage virtualization software known as **StorageGRID.** It creates a virtualization layer that fetches storage from different storage devices into a single management system. It can also manage data from **CIFS** and **NFS** file systems over the Internet. The following diagram shows how StorageGRID virtualizes the storage into storage clouds:



Virtual Storage Containers

The **virtual storage containers** offer high performance cloud storage systems. **Logical Unit Number (LUN)** of device, files and other objects are created in virtual storage containers. Following diagram shows a virtual storage container, defining a cloud storage domain:



Challenges

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

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

23. a. Explain about virtualization technologies in cloud. (Or)

Virtualization Technologies

The dictionary includes many definitions for the word "cloud." A cloud can be a mass of water droplets, gloom, an obscure area, or a mass of similar particles such as dust or smoke. When it comes to cloud computing, the definition that best fits the context is "a collection of objects that are grouped together."

It is that act of grouping or creating a resource pool that is what succinctly

differentiates cloud computing from all other types of networked systems. The benefits of pooling resources to allocate them on demand are so compelling as to make the adoption of these technologies a priority.

Without resource pooling, it is impossible to attain efficient utilization, provide reasonable costs to users, and proactively react to demand. In this chapter, you learn about the technologies that abstract physical resources such as processors, memory, disk, and network capacity into virtual resources.

When you use cloud computing, you are accessing pooled resources using a technique called virtualization. Virtualization assigns a logical name for a physical resource and then provides a pointer to that physical resource when a request is made. Virtualization provides a means to manage resources efficiently because the mapping of virtual resources to physical resources can be both dynamic and facile. Virtualization is dynamic in that the mapping can be assigned based on rapidly changing conditions, and it is facile because changes to a mapping assignment can be nearly instantaneous.

These are among the different types of virtualization that are characteristic of cloud computing:

- **Access:** A client can request access to a cloud service from any location.
- **Application:** A cloud has multiple application instances and directs requests to an instance based on conditions.
- **CPU:** Computers can be partitioned into a set of virtual machines with each machine being assigned a workload. Alternatively, systems can be virtualized through load-balancing technologies.
- **Storage:** Data is stored across storage devices and often replicated for redundancy.

To enable these characteristics, resources must be highly configurable and

flexible. You can define the features in software and hardware that enable this flexibility as conforming to one or more of the following mobility patterns:

- **P2V:** Physical to Virtual
- **V2V:** Virtual to Virtual
- **V2P:** Virtual to Physical
- **P2P:** Physical to Physical
- **D2C:** Datacenter to Cloud
- **C2C:** Cloud to Cloud
- **C2D:** Cloud to Datacenter
- **D2D:** Datacenter to Datacenter

Virtualization is a key enabler of the first four of five key attributes of cloud computing:

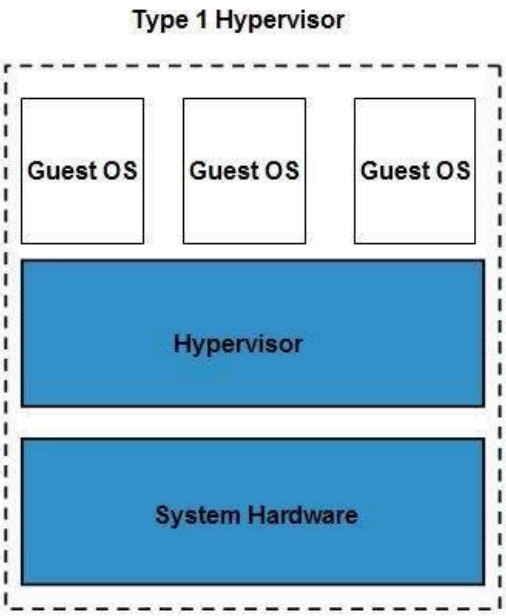
- **Service-based:** A service-based architecture is where clients are abstracted from service providers through service interfaces.
- **Scalable and elastic:** Services can be altered to affect capacity and performance on demand.
- **Shared services:** Resources are pooled in order to create greater efficiencies.
- **Metered usage:** Services are billed on a usage basis.
- **Internet delivery:** The services provided by cloud computing are based on Internet protocols and formats.

b. What are hypervisors? Explain the different types of hypervisors.

Hypervisor

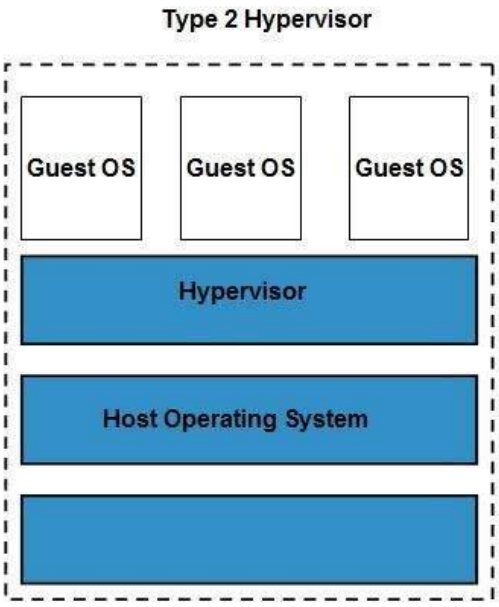
The **hypervisor** is a firmware or low-level program that acts as a Virtual Machine Manager. There are two types of hypervisor:

Type 1 hypervisor executes on bare system. LynxSecure, RTS Hypervisor, Oracle VM, Sun xVM Server, VirtualLogic VLX are examples of Type 1 hypervisor. The following diagram shows the Type 1 hypervisor.



The **type1 hypervisor** does not have any host operating system because they are installed on a bare system.

Type 2 hypervisor is a software interface that emulates the devices with which a system normally interacts. Containers, KVM, Microsoft Hyper V, VMWare Fusion, Virtual Server 2005 R2, Windows Virtual PC and **VMWare workstation 6.0** are examples of Type 2 hypervisor. The following diagram shows the Type 2 hypervisor.



Reg.No _____

[18CSP104]

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established Under Section 3 of UGC Act 1956)

For the candidates admitted in 2018 onwards

First Semester

M.Sc COMPUTER SCIENCE

SECOND INTERNAL EXAMINATION – OCTOBER 2018

Cloud Computing

Class : I M.Sc CS **Time: 2 hrs**

Date & Session: 25.10.18 & AN **Marks: 50**

PART – A (20 * 1 = 20 Marks)

1. _____ is the single most heavily visited site on the Internet.

a. MSN b. Rediff c. Yahoo d. **Google**

2. Google doesn't use _____ virtualization.

a. Middleware b. Software **c. Hardware** d. Malware

3. A_____ can contain thousands of servers.

a. Google servers **b. Google cluster** c. Google app engine d. Google clients

4. _____ is a cloud operating system built on top of Microsoft datacenters infrastructure.

a. Microsoft Windows Azure b. Oracle Azure c. VB Azure d. Java Azure

5. _____is one of the most important and heavily trafficked Web sites in the world.

a. Yahoo.com b. Google.com **c. Amazon.com** d. MSN.com

6. Web crawlers are also called as _____.

a. Spiders or Robots b. Dark Robots c. Dark Spiders d. Shadowers

7._____ is the total combination of protection mechanisms within a computer system.

a. Trusted platform module **b. Trusted computing base**

c. Trusted computing network d. Trusted software module

8. There are ____ main performance measures in biometrics.

a. 3 b. 2 c. 4 d. 5

9. _____controls incorporate guards and building security in general.

a. Risk **b. Physical** c. Administrative d. Logical

10. There are currently _____different EC2 service zones or regions.

a. Three **b. Four** c. Five d. Six

11. _____ is a prominent example of a site that isn't indexed in search engines.

a. Facebook b. Google c. Yahoo d. MSN

12. _____are implemented to mitigate risk and reduce the potential for loss.

a. Security b. Privacy **c. Controls** d. Privilege

13.The Data store in Google App Engine is _____.

a. Non-relational b. Relational c. Schema based d. Standard

14. _____refers to the prevention of intentional or unintentional unauthorized disclosure of information.

a. Integrity b. Security **c. Confidentiality** d. Privacy

15._____ clouds are easier with Azure.

a. hybrid b. private c. public d. shared

16. The ____system determines which ads to match to the user searches.

a. AdWords b. AdSense c. AdDeep web d. AdServer

17._____ is the testing or reconciliation of evidence of a user's identity.

a. Authentication b. Authorization c. Auditing d. Accessibility

18. ____ is a weakness in a system that can be exploited by a threat.

a. Eavesdropping b. Virus c. Fraud d. **Vulnerability**

19. A connection is the Service Bus element that is priced by Azure on a _____basis.

a. Yearly b. Weekly **c. Pay-as-you-go** d. Monthly

20. Machine images are sometimes referred to as _____.

a. Virtual images **b. Virtual appliances** c. Canned images d. Stored images

PART – B (3* 2 = 6 Marks)

21. What is CIA Triad?

The three fundamental tenets of information security — confidentiality, integrity, and availability (CIA) — define an organization's security posture. Confidentiality is the prevention of the intentional or unintentional unauthorized disclosure of contents. *Integrity* is the guarantee that the message sent is the message received and that the message is not intentionally or unintentionally altered. Availability refers to the elements that create reliability and stability in networks and systems.

22. List the 11 security design principles in cloud.

the following 11 security design principles:

- Least privilege
- Separation of duties
- Defense in depth
- Fail safe
- Economy of mechanism
- Complete mediation
- Open design
- Least common mechanism
- Psychological acceptability
- Weakest link
- Leveraging existing components

23. Mention the risk associated with Cloud Service Providers.

Complexity of configuration, Privilege escalation, Inactive virtual machines, Segregation of duties AND Poor access controls are the risk associated with Cloud Service Providers.

PART – C (3 * 8 = 24 Marks)

24. a. Give a clear explanation on AppZero Virtual Application Appliance. (Or)

Applications that run in datacenters are captive to the operating systems and hardware platforms that they run on. Many datacenters are a veritable Noah's Ark of computing. So moving an application from one platform to another isn't nearly as simple as moving a machine image from one system to another. The situation is further complicated by the fact that applications are tightly coupled with the operating systems on which they run. An application running on Windows, for example, isn't isolated from other applications. When the application loads, it often loads or uses different Dynamic Link Libraries (DLL), and it is through the sharing or modification of DLLs that Windows applications get themselves in trouble. Further modifications include modifying the registry during installation. These factors make it difficult to port applications from one platform to another without lots of careful work. If you are a Platform as a Service (PaaS) application developer, you are packaging a complete software stack that includes not only your application, but the operating system and application logic and rules as well.

Vendor lock-in for you application is assured.

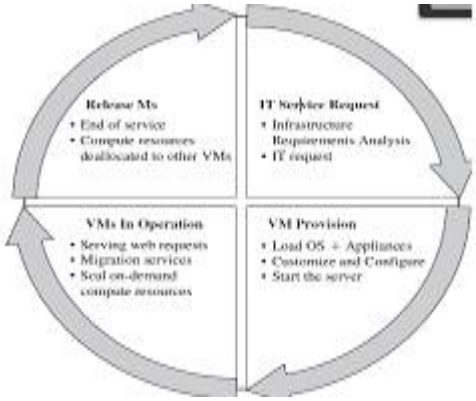
The ability to run an application from whatever platform you want is not one of the characteristics of cloud computing, but you can imagine that it is a very attractive proposition. While the Simple Cloud API is useful for applications written in PHP, other methods may be needed to make applications easily portable. One company working on this problem is AppZero (http://www.appzero.com/), and its solution is called the Virtual Application Appliance (VAA).

The AppZero solution creates a virtual application appliance as an architectural layer between the Windows or the UNIX operating system and applications. The virtualization layer serves as the mediator for file I/O, memory I/O, and application calls and response to DLLs, which has the effect of sandboxing the application. The running application in AppZero changes none of the registry entries or any of the files on the Windows Server.

VAA creates a container that encapsulates the application and all the application's dependencies within a set of files; it is essentially an Application Image for a specific OS. Dependencies include DLL, service settings, necessary configuration files, registry entries, and machine and network settings. This container forms an installable server- side application stack that can be run after installation, but has no impact on the underlying operating system. VAAs are created using the AppZero Creator wizard, managed with the AppZero Admin tool, and may be installed using the AppZero Director, which creates a VAA runtime application. If desired, an application called AppZero Dissolve removes the VAA virtualization layer from the encapsulated application and installs that application directly into the operating system.

Installations can be done over the network after the AppZero application appliance is installed. Therefore, with this system, you could run applications on the same

Windows Server and eliminate one application from interfering with another; applications would be much more easily ported from one Windows system to another. AppZero's approach provides the necessary abstraction layer that frees an application from its platform dependence. As shown in Figure 5.3, the cycle starts by a request delivered to the IT department, stating the requirement for creating a new server for a particular service. This request is being processed by the IT administration to start seeing the servers' resource pool, matching these resources with the requirements, and starting the provision of the needed virtual machine. Once it is provisioned and started, it is ready to provide the required service according to an SLA, or a time period after which the virtual is being released; and free resources, in this case, won't be needed.



VM Provisioning Process Provisioning a virtual machine or server can be explained and illustrated as in Figure.

Steps to Provision VM.

Here, we describe the common and normal steps of provisioning a virtual server:

- Firstly, you need to select a server from a pool of available servers (physical servers with enough capacity) along with the appropriate OS template you need to provision the virtual machine.
- Secondly, you need to load the appropriate software (operating system you selected in the previous step, device drivers, middleware, and the needed applications for the service required).
- Thirdly, you need to customize and configure the machine (e.g., IP

address, Gateway) to configure an associated network and storage resources.

- Finally, the virtual server is ready to start with its newly loaded software.

Typically, these are the tasks required or being performed by an IT or a data center's specialist to provision a particular virtual machine. To summarize, server provisioning is defining server's configuration based on the organization requirements, a hardware, and software component (processor, RAM, storage, networking, operating system, applications, etc.). Normally, virtual machines can be provisioned by manually installing an operating system, by using a preconfigured VM template, by cloning an existing VM, or by importing a

physical server or a virtual server from another hosting platform. Physical servers can also be virtualized and provisioned using P2V (physical to virtual) tools and techniques (e.g., virt-p2v). After creating a virtual machine by virtualizing a physical server, or by building a new virtual server in the virtual environment, a template can be created out of it. Most virtualization management vendors (VMware, XenServer, etc.) provide the data center's administration with the ability to do such tasks in an easy way. Provisioning from a template is an invaluable feature, because it reduces the time required to create a new virtual machine. Administrators can create different templates for different purposes. For example, you can create a Windows 2003 Server template for the finance department, or a Red Hat Linux template for the engineering department. This enables the administrator to quickly provision a correctly configured virtual server on demand. This ease and flexibility bring with them the problem of virtual machine's sprawl, where virtual machines are provisioned so rapidly that documenting and managing the virtual machine's life cycle become a challenge .

b. Describe machine imaging and porting applications.

Machine Imaging

A system image makes a copy or a clone of the entire computer system inside a single container such as a file. The system imaging program is used to make this image and can be used later to restore a system image. Some imaging programs can take snapshots of systems, and most allow you to view the files contained in the image and do partial restores.

A prominent example of a system image and how it can be used in cloud computing architectures is the Amazon Machine Image (AMI) used by Amazon Web Services to store copies of a virtual machine. An AMI is a file system image that contains an operating system, all appropriate device drivers, and any applications and state information that the working virtual machine would have.

When you subscribe to AWS, you can choose to use one of its hundreds of canned AMIs or to create a custom system and capture that system's image to an AMI. An AMI can be for public use under a free distribution license, for pay-per-use with operating systems such as Windows, or shared by an EC2 user with other users who are given the privilege of access.

The AMI file system is not a standard bit-for-bit image of a system that is common to many disk imaging programs. AMI omits the kernel image and stores a pointer to a particular kernel that is part of the AWS kernel library. Among the choices are Red Hat Linux, Ubuntu, Microsoft Windows, Solaris, and others. Files in AMI are compressed and encrypted, and an XML file is written that describes the AMI archive. AMIs are typically stored in your Amazon S3 (Simple Storage System) buckets as a set of 10MB chunks.

Machine images are sometimes referred to as “virtual appliances”—systems that are meant to run on virtualization platforms. AWS EC2 runs on the Xen hypervisor, for example. The term *virtual appliance* is meant to differentiate the software image from an operating virtual machine. The system image contains the

operating system and applications that create an environment. Most virtual appliances are used to run a single application and are configurable from a Web page.

Virtual appliances are a relatively new paradigm for application deployment, and cloud computing is the major reason for the interest in them and for their adoption. This area of WAN application portability and deployment, and of WAN optimization of an application based on demand, is one with many new participants. Certeon (http://www.certeon.com/), Expand Networks (http://www.expand.com/), and Replify (http://www replify.com/) are three vendors offering optimization appliances for VMware's infrastructure.

Porting Applications

Cloud computing applications have the ability to run on virtual systems and for these systems to be moved as needed to respond to demand. Systems (VMs running applications), storage, and network assets can all be virtualized and have sufficient flexibility to give acceptable distributed WAN application performance. Developers who write software to run in the cloud will undoubtedly want the ability to port their applications from one cloud vendor to another, but that is a much more difficult proposition. Cloud computing is a relatively new area of technology, and the major vendors have technologies that don't interoperate with one another.

25. a. Elaborate on Privacy and Compliance Risks in Cloud Computing. (Or)

One area that is greatly affected by cloud computing is privacy. It’s important to remember that although the control of cloud computing privacy has many threats and vulnerabilities in common with non cloud processes and infrastructure, it also has unique security issues.

For example, a successful identity theft exploit can result in a privacy loss that has a huge

impact on an enterprise. The organization can suffer short-term losses due to remediation, investigation, and restitution costs. It can also incur longer term problems for the organization due to loss of credibility, confidence, and negative publicity.

Another mistake organizations often make is in assigning responsibility for privacy controls to the IT dept, rather than a business unit that owns the data. Information systems security frameworks have defined, standardized processes that apply to cloud computing — and its potential privacy breaches. This section examines the legal and standard processes that affect privacy control in the cloud.

An individual's right to privacy is embodied in the fundamental principles of privacy:

- **Notice** — Regarding the collection, use, and disclosure of personally identifiable information (PII)
- **Choice** — To opt out or opt in regarding disclosure of PII to third parties
- **Access** — By consumers to their PII to permit review and correction of information
- **Security** — To protect PII from unauthorized disclosure
- **Enforcement** — Of applicable privacy policies and obligations

Proving compliance depends on the number of transactions that they process. Merchants and service providers with higher levels of transactions have to pass an on-site audit every year. Those with lower levels of transactions must submit documentation stating that they meet the requirements, called a self-assessment questionnaire.

The PCI DSS contains the following set of 12 high-level requirements that are supported by a series of more detailed requirements:

- Install and maintain a firewall configuration to protect cardholder data.
- Do not use vendor-supplied defaults for system passwords and other security parameters.
- Protect stored cardholder data.

- Encrypt transmission of cardholder data across open, public networks.
- Use and regularly update antivirus software.
- Develop and maintain secure systems and applications.
- Restrict access to cardholder data based on the business's need to know.
- Assign a unique ID to each person with computer access.
- Restrict physical access to cardholder data.
- Track and monitor all access to network resources and cardholder data.
- Regularly test security systems and processes.
- Maintain a policy that addresses information security.

Information Privacy and Privacy Laws

There are many types of legal systems in the world, and they differ in how they treat evidence, the rights of the accused, and the role of the judiciary. These laws have a significant privacy impact on cloud computing environments, yet vary widely.

Examples of these different legal systems are common law, Islamic and other religious law, and civil law. The common law system is employed in the United States, United Kingdom, Australia, and Canada. Civil law systems are used in France, Germany, and Quebec, to name a few.

Organizations develop and publish privacy policies that describe their approach to handling PII. The websites of organizations usually have their privacy policies available to read online, and these policies usually cover the following areas:

- Statement of the organization's commitment to privacy
- The type of information collected, such as names, addresses, credit card numbers, phone numbers, and so on
- Retaining and using e-mail correspondence
- Information gathered through cookies and Web server logs and how

that information is used

- How information is shared with affiliates and strategic partners
- Mechanisms to secure information transmissions, such as encryption and digital signatures
- Mechanisms to protect PII stored by the organization
- Procedures for review of the organization's compliance with the privacy policy
- Evaluation of information protection practices
- Means for the user to access and correct PII held by the organization
- Rules for disclosing PII to outside parties
- Providing PII that is legally required

Privacy Legislation

The following list summarizes some important legislation and recommended guidelines for privacy:

- *The Cable Communications Policy Act* provides for discretionary use of PII by cable operators internally but imposes restrictions on disclosures to third parties.
- *The Children's Online Privacy Protection Act* (COPPA) is aimed at providing protection to children under the age of 13.
- *Customer Proprietary Network Information Rules* apply to telephone companies and restrict their use of customer information both internally and to third parties.
- *The Financial Services Modernization Act* (*Gramm-Leach-Bliley*) requires financial institutions to provide customers with clear descriptions of the institution's policies and procedures for protecting the PII of customers.
- *The Telephone Consumer Protection Act* restricts communications between companies and consumers, such as telemarketing.

- *The 1973 U.S. Code of Fair Information Practices* states that:

1. There must not be personal data record-keeping systems whose very existence is secret.
2. There must be a way for a person to find out what information about them is in a record and how it is used.
3. There must be a way for a person to prevent information about them, which was obtained for one purpose, from being used or made avail- able for another purpose without their consent.

b. Analyze on Trusted Cloud Computing with real-time environment.

Trusted cloud computing can be viewed as a computer security architecture that is designed to protect cloud systems from malicious intrusions and attacks, and ensure that computing resources will act in a specific, predictable man- ner as intended. A trusted cloud computing system will protect data in use by hypervisors and applications, protect against unauthorized access to informa- tion, provide for strong authentication, apply encryption to protect sensitive data that resides on stolen or lost devices, and support compliance through hardware and software mechanisms.

Trusted Computing Characteristics

In a cloud computational system, multiple processes might be running concur- rently. Each process has the capability to access certain memory locations and to

execute a subset of the computer's instruction set. The execution and memory space assigned to each process is called a *protection domain*. This domain can be extended to virtual memory, which increases the apparent size of real memory by using disk storage. The purpose of establishing a protection domain is to protect programs from all unauthorized modification or executional interference. A *trusted computing base (TCB)* is the total combination of protection mechanisms within a computer system, which includes the hardware, software, and firmware that are trusted to enforce a security policy. Because the TCB components are responsible for enforcing the security policy of a computing system, these components must be protected from malicious and untrusted

processes. The TCB must also provide for memory protection and ensure that the processes from one domain do not access memory locations of another domain. The *security perimeter* is the boundary that separates the TCB from the remainder of the system. A *trusted path* must also exist so that users can access the TCB without being compromised by other processes or users. Therefore, a *trusted computer system* is one that employs the necessary hardware and software assurance measures to enable its use in processing multiple levels of classified or sensitive information. This system meets the specified requirements for reliability and security.

Another element associated with trusted computing is the *trusted platform module (TPM)*. The TPM stores cryptographic keys that can be used to attest to the operating state of a computing platform and to verify that the hardware and software configuration has not been modified. However, the standard TPM cannot be used in cloud computing because it does not operate in the virtualized cloud environment. To permit a TPM version to perform in the cloud, specifications have been generated for a virtual TPM (VTM)[4] that provides software instances of TPMs for each virtual machine operating on a trusted server.

Trusted computing also provides the capability to ensure that software that processes information complies with specified usage policies and is running unmodified and isolated from other software on the system. In addition, a trusted computing system must be capable of enforcing mandatory access con- trol (MAC) rules. MAC rules are discussed in more detail later in this chapter. Numerous trust-related issues should be raised with, and satisfied by, a cloud provider. They range from concerns about security, performance, cost, control, availability, resiliency, and vendor lock in.

Additional factors that inspire trust include the following:

- Use of industry-accepted standards.
- Provision for interoperability and transparency.
- Robust authentication and authorization mechanisms in access control.
- Management of changing personnel and relationships in both the cloud client and provider organizations.

- Establishment of accountability with respect to security and privacy requirements in a multi-party, flexible service delivery setting.
- Use of information system security assurance techniques and metrics to establish the effectiveness of hardware and software protection mechanisms.
- Establishment of effective policies and procedures to address multiple legal jurisdictions associated with cloud international services and com- pliance requirements.
- Application of Information Rights Management (IRM) cryptographic techniques to protect sensitive cloud-based documents and provide an audit trail of accesses and policy changes. IRM prevents protected docu- ments from screen capture, being printed, faxed, or forwarded, and can prohibit messages and attachments from being accessed after a specified period of time.

Also, because of the high volume of data that is being moved around in vari- ous locations, authorization privileges and rights management constraints must be attached to the data itself to restrict access only to authorized users.

Because of legal and forensic requirements, a trusted cloud provider should also have a Security Information and Event Management (SIEM) capability that can manage records and logs in a manner that meets legal constraints. An SEIM is a software mechanism that provides for centralized acquisition, storage, and analysis of recorded events and logs generated by other tools on an enterprise network.

Information stored in a SEIM can be used for data mining to discover signifi- cant trends and occurrences, and to provide for reliable and legally acceptable storage of information. It can also be used by report generators, and provide for backup of log data that might be lost at the source of the data.

Secure Execution Environment

Configuring computing platforms for secure execution is a complex task; and in many instances it is not performed properly because of the large number of parameters that are

involved. This provides opportunities for malware to exploit vulnerabilities, such as downloading code embedded in data and having the code executed at a high privilege level.

In cloud computing, the major burden of establishing a secure execution environment is transferred from the client to the cloud provider. However, protected data transfers must be established through strong authentication mechanisms, and the client must have practices in place to address the privacy and confidentiality of information that is exchanged with the cloud. In fact, the client's port to the cloud might provide an attack path if not properly provisioned with security measures. Therefore, the client needs assurance that computations and data exchanges are conducted in a secure environment. This assurance is affected by trust enabled by cryptographic methods. Also, research into areas such as compiler-based virtual machines promises a more secure execution environment for operating systems.

Another major concern in secure execution of code is the widespread use of "unsafe" programming languages such as C and C++ instead of more secure languages such as object-oriented Java and structured, object-oriented C#.

Secure Communications

As opposed to having managed, secure communications among the computing resources internal to an organization, movement of applications to the cloud requires a reevaluation of communications security. These communications apply to both data in motion and data at rest.

Secure cloud communications involves the structures, transmission methods, transport formats, and security measures that provide confidentiality, integ- rity, availability, and authentication for transmissions over private and public communications networks. Secure cloud computing communications should ensure the following:

- **Confidentiality** — Ensures that only those who are supposed to access data can retrieve it. Loss of confidentiality can occur through the inten- tional release of private

company information or through a misapplication of network rights. Some of the elements of telecommunications used to ensure confidentiality are as follows:

- Network security protocols
- Network authentication services
- Data encryption services

- **Integrity** — Ensures that data has not been changed due to an accident or malice. Integrity is the guarantee that the message sent is the message received and that the message is not intentionally or unintentionally altered. Integrity also contains the concept of nonrepudiation of a message source. Some of the constituents of integrity are as follows:

- Firewall services
- Communications Security Management
- Intrusion detection services

- **Availability** — Ensures that data is accessible when and where it is needed, and that connectivity is accessible when needed, allowing authorized users to access the network or systems. Also included in that assurance is the guarantee that security services for the security practitioner are usable when they are needed. Some of the elements that are used to ensure avail- ability are as follows:

- Fault tolerance for data availability, such as backups and redundant disk systems
- Acceptable logins and operating process performances
- Reliable and interoperable security processes and network security mechanisms

26. a. Explain about Open Nebulla and its components. **(Or)**

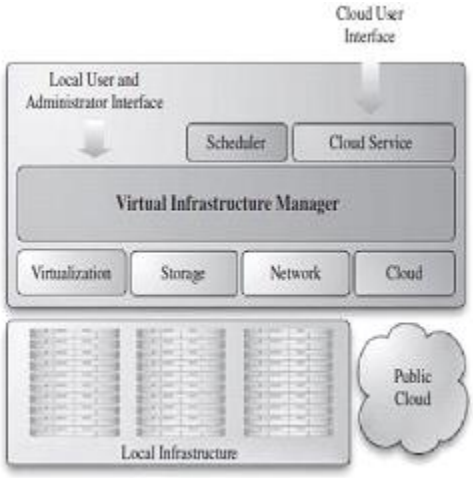
OpenNebula is an open and flexible tool that fits into existing data center's environments to build any type of cloud deployment. OpenNebula can be primarily used as a virtualization tool to manage your virtual infrastructure, which is usually referred to as private cloud. OpenNebula supports a hybrid cloud to combine local

infrastructure with public cloud-based infrastructure, enabling highly scalable hosting environments. OpenNebula also supports public clouds by providing cloud's interfaces to expose its functionality for virtual machine, storage, and network management. OpenNebula is one of the technologies being enhanced in the Reservoir Project , European research initiatives in virtualized infrastructures, and cloud computing.

OpenNebula architecture is shown in Figure which illustrates the existence of public and private clouds and also the resources being managed by its virtual manager.

OpenNebula is an open-source alternative to these commercial tools for the dynamic management of VMs on distributed resources. This tool is supporting several research lines in advance reservation of capacity, probabilistic admission control, placement optimization, resource models for the efficient management of groups of virtual machines, elasticity support, and so on. These research lines address the requirements from both types of clouds namely, private and public.

OpenNebula and Haizea. Haizea is an open-source virtual machine-based lease management architecture developed by Sotomayor et al. ; it can be used as a scheduling backend for OpenNebula. Haizea uses leases as a fundamental resource provisioning abstraction and implements those leases as virtual machines, taking into account the overhead of using virtual machines when scheduling leases.



Haizea also provides advanced functionality such as:

- Advance reservation of capacity.
- Best-effort scheduling with backfilling.
- Resource preemption (using VM suspend/resume/migrate).
- Policy engine, allowing developers to write pluggable scheduling policies in Python.

Aneka

Manjrasoft Aneka is a .NET-based platform and framework designed for building and deploying distributed applications on clouds. It provides a set of APIs for transparently exploiting distributed resources and expressing the business logic of applications by using the preferred programming abstractions. Aneka is also a market-oriented cloud platform since it allows users to build and schedule applications, provision resources, and monitor results using pricing, accounting, and QoS/SLA services in private and/or public cloud environments.

It allows end users to build an enterprise/private cloud setup by exploiting the power of computing resources in the enterprise data centers, public clouds such as Amazon EC2 , and hybrid clouds by combining enterprise private clouds managed by Aneka with resources from Amazon EC2 or other enterprise clouds built and managed using technologies such as XenServer.

Aneka also provides support for deploying and managing clouds. By using its Management Studio and a set ofWeb interfaces, it is possible to set up either public or private clouds, monitor their status, update their configuration, and perform the basic management operations.

Aneka Architecture.

Aneka platform, consists of a collection of physical and virtualized resources connected through a network. Each of these resources hosts an instance of the Aneka container representing the runtime environment where the distributed applications

are executed. The container provides the basic management features of the single node and leverages all the other operations on the services that it is hosting. The services are broken up into fabric, foundation, and execution services. Fabric services directly interact with the node through the platform abstraction layer (PAL) and perform hardware profiling and dynamic resource provisioning. Foundation services identify the core system of the Aneka middleware, providing a set of basic features to enable Aneka containers to perform specialized and specific sets of tasks. Execution services directly deal with the scheduling and execution of applications in the cloud.

b. Discuss about Amazon EC2 with cloud software environments.

Amazon Elastic Compute Cloud (Amazon EC2) provides scalable computing capacity in the Amazon Web Services (AWS) cloud. Using Amazon EC2 eliminates your need to invest in hardware up front, so you can develop and deploy applications faster. You can use Amazon EC2 to launch as many or as few virtual servers as you need, configure security and networking, and manage storage. Amazon EC2 enables you to scale up or down to handle changes in requirements or spikes in popularity, reducing your need to forecast traffic.

Features of Amazon EC2

Amazon EC2 provides the following features:

- Virtual computing environments, known as *instances*
- Preconfigured templates for your instances, known as *Amazon Machine Images (AMIs)*, that package the bits you need for your server (including the operating system and additional software)
- Various configurations of CPU, memory, storage, and networking capacity for your instances, known as *instance types*
- Secure login information for your instances using *key pairs* (AWS stores the public key, and you store the private key in a secure place)
- Storage volumes for temporary data that's deleted when you stop or terminate your instance, known as *instance store volumes*

- Persistent storage volumes for your data using Amazon Elastic Block Store (Amazon EBS), known as *Amazon EBS volumes*
- Multiple physical locations for your resources, such as instances and Amazon EBS volumes, known as *regions* and*Availability Zones*
- A firewall that enables you to specify the protocols, ports, and source IP ranges that can reach your instances using*security groups*
- Static IPv4 addresses for dynamic cloud computing, known as *Elastic IP addresses*
- Metadata, known as *tags*, that you can create and assign to your Amazon EC2 resources
- Virtual networks you can create that are logically isolated from the rest of the AWS cloud, and that you can optionally connect to your own network, known as *virtual private clouds* (VPCs)