PPT ON
CLOUD APPLICATION DEVELOPMENT
Course Code : ACS011
VII SEM (IARE-R16)

UNIT-I

INTRODUCTION AND CLOUD APPLICATION DEVELOPMENT

Definition of cloud computing

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released.

Characteristics:

- Great Availability of Resources
- On-demand Self-service
- Easy Maintenance
- Large Network Access
- Availability
- Automatic System
- Economical
- Security
- Pay as you go

Cont..

Benefits

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service
- Lower costs
- Ease of utilization
- Quality of Service
- Reliability
- Outsourced IT management
- Simplified maintenance and upgrade
- Low Barrier to Entry

Challenges of Cloud Computing

- Although the computers of today are highly intelligent and sophisticated they have their own limitations. The computer cannot think on its own, since it does not have its own brain.

- It can only do what is has been programmed to do. It can execute only those jobs that can be expressed as a finite set of instructions to achieve a specific goal.

- The computers do not learn from previous experience nor can they arrive at a conclusion without going through all the intermediate steps.

Cloud models

Defining Infrastructure as a Service (IaaS):

- The Application layer forms the basis for Software as a Service (SaaS),while the Platform layer forms the basis for Platform as a Service (PaaS) models.

- Infrastructure as a Service (IaaS) creates what may be determined to be a utility computing model

- Infrastructure as a Service (IaaS) is a cloud computing service model in which hardware is virtualized in the cloud.

- In this particular model, the service vendor owns the equipment: servers, storage, network infrastructure, and so forth.

Cont..

IaaS workloads:

- The fundamental unit of virtualized client in an IaaS deployment is called a workload.
- A workload simulates the ability of a certain type of real or physical server to do an amount of work.
- In cloud computing, a provisioned server called an instance is reserved by a customer.
- A client would reserve a machine equivalent required to run each of these workloads

Cont..

Defining Platform as a Service (PaaS):

- The Platform as a Service model describes a software environment in which a developer can create customized solutions within the context of the development tools that the platform provides.

- Platforms can be based on specific types of development languages, application frameworks, or other constructs.

- In a PaaS model, customers may interact with the software to enter and retrieve data, perform actions, get results, and to the degree that the vendor allows it, customize the platform involved.

Cont..

Defining Software as a Service (SaaS):

- The most complete cloud computing service model is one in which the computing hardware and software, as well as the solution itself, are provided by a vendor as a complete service offering.

- Software as a Service (SaaS) may be succinctly described as software that is deployed on a hosted service and can be accessed globally over the Internet, most often in a browser.

- SaaS applications come in all shapes and sizes, and include custom software such as billing and invoicing systems, Customer Relationship Management (CRM) applications.

Cont..

SaaS characteristics

- The software is available over the Internet globally through a browser on demand.
- The typical license is subscription-based or usage-based and is billed on a recurring basis.
- The software and the service are monitored and maintained by the vendor, regardless of where all the different software components are running.
- Reduced distribution and maintenance costs and minimal end-user system costs generally make SaaS applications cheaper to use than their shrink-wrapped versions.

Cont..

- Such applications feature automated upgrades, updates, and patch management and much faster rollout of changes.

- SaaS applications often have a much lower barrier to entry than their locally installed competitors, a known recurring cost, and they scale on demand.

- All users have the same version of the software so each user's software is compatible with another's.

- SaaS supports multiple users and provides a shared data model through a single-instance, multi-tenancy model

Cloud Types

Most people separate cloud computing into two distinct sets of models:

- **Deployment models:** This refers to the location and management of the cloud's infrastructure.
- **Service models:** This consists of the particular types of services that you can access on a cloud computing platform.

This is a very useful demarcation that is now widely accepted.

- **The NIST model**

  The United States government is a major consumer of computer services and, therefore, one of the major users of cloud computing networks.

Cont..

The Cloud Cube Model

- The Open Group maintains an association called the Jericho Forum whose main focus is how to protect cloud networks.

The four dimensions of the Cloud Cube Model are:

- Physical location of the data
- Ownership
- Security boundary
- Sourcing

Cont..

Deployment models:

A deployment model defines the purpose of the cloud and the nature of how the cloud is located.

- **Public cloud:** The public cloud infrastructure is available for public use alternatively for a large industry group and is owned by an organization selling cloud services.
- **Private cloud:** The private cloud infrastructure is operated for the exclusive use of an organization.
- **Hybrid cloud:** A hybrid cloud combines multiple clouds where those clouds retain their unique identities, but are bound together as a unit.
- **Community cloud:** A community cloud is one where the cloud has been organized to serve a common function or purpose.

Cont..

Service models:

In the deployment model, different cloud types are an expression of the manner in which infra- structure is deployed.

Three service types have been universally accepted:

- **Infrastructure as a Service:** IaaS provides virtual machines, virtual storage, virtual infra-structure, and other hardware assets as resources that clients can provision.
- **Platform as a Service**: PaaS provides virtual machines, operating systems, applications, services, development frameworks, transactions, and control structures.
- **Software as a Service**: SaaS is a complete operating environment with applications, management, and the user interface.

Grid Computing

- Grid computing is a combination of resources from multiple administrative domains to reach a common.

- The grid is multiple owned, it could be owned by several companies. Interconnection network is mostly internet with high latency and low bandwidth.

- The security in the grid is public and private based on authentication and mapping user to an account.

- The grid computing is usually used in predictive modeling and simulations, engineering design and automation, energy resources exploration.

Cont..

Advantages:

- One of the advantages of grid computing that you don’t need to buy large servers for applications that can be split up and farmed out to smaller commodity type servers, secondly it’s more efficient in use of resources.
- Also the grid environments are much more modular and don't have much points of failure.
- The policies in the grid it can be managed by the grid software, beside that upgrading can be done without scheduling downtime, and jobs can be executed in parallel speeding performance

Cont..

Disadvantages:

- It needs fast interconnect between computers resources, and some applications may need to be pushed to take full advantage of the new model.

- licensing across many servers may make it forbidden for some applications.

- the grid environments include many smaller servers across various administrative domains also political challenges associated with sharing resources especially across different admin domains.

Utility Computing

- Utility Computing refers to a type of computing technologies and business models which provide services and computing resources to the customers.

- This model has the advantage of a low cost with no initial price to reach the computer resources.

- This repackaging of computing services is the foundation of the shift to on demand computing, software as a service and cloud computing models.

-

- Utility computing is kind of virtualization, that means the whole web storage space.

Cont..

Advantages of utility computing:

- Here are some benefits of utility computing such as that the client doesn't have to buy all the hardware, software and licenses needed to do business.
- Instead, the client relies on another party to provide these services.
- It also gives companies the option to subscribe to a single service and use the same suite of software throughout the entire client organization.
- And it offers compatibility of all the computers in large companies

Cont..

Disadvantages of utility computing:

- There's some issues are considered as disadvantages such as reliability which means the service could be stopped from the utility computing

- company for any reason such as a financial trouble or equipment problems.

- Also utility computing systems can also be attractive targets for hackers, and much of the responsibility of keeping the system safe falls to the provider

Cloud computing at Amazon

- Amazon introduced a computing platform that has changed the face of computing in the last decade.
- First, it installed a powerful computing infrastructure to sustain its core business, e-commerce.
- Then Amazon discovered that this infrastructure could be further extended to provide affordable and easy-to-use resources for enterprise computing as well as computing for the masses.
- In mid-2000 Amazon introduced Amazon Web Services (AWS), based on the IaaS delivery model. In this model the cloud service provider offers an infrastructure consisting of compute and storage servers.
- An application developer is responsible for installing applications on a platform of his or her choice and managing the resources provided by Amazon

Cont..

- AMI images create an exact copy of the original image but without configuration-dependent information such as the hostname or the MAC address.
- A user can:

(i) Launch an instance from an existing AMI and terminate an instance;

(ii) start and stop an instance;

(iii) create a new image;

(iv) add tags to identify an image; and

(v) reboot an instance.

EC2 is based on the Xen virtualization strategy

The Google perspective

- Google's effort is concentrated in the area of Software-as-a-Service (SaaS). It is estimated that the number of servers used by Google was close to 1.8 million.
- Services such as Gmail, Google Drive, Google Calendar, Picasa, and Google Groups are free of charge for individual users and available for a fee for organizations.
- These services are running on a cloud and can be invoked from a broad spectrum of devices, including mobile ones such as iPhones, iPads, Black-Berrys, and laptops and tablets.
- The data for these services is stored in data centers on the cloud.

Cont..

- Google Groups allows users to host discussion forums to create messages online or via email.

- Google is also a leader in the Platform-as-a-Service (PaaS) space. AppEngine is a developer platform hosted on the cloud.

- The database for code development can be accessed with Google Query Language (GQL) with a SQL-like syntax.

- Google Drive is an online service for data storage that has been available since April 2012.

Microsoft Windows Azure and online services

- Azure and Online Services are, respectively, PaaS and SaaS cloud platforms from Microsoft.

- Windows Azure is an operating system, SQL Azure is a cloud-based version of the SQL Server, andAzure AppFabric (formerly .NET Services) is a collection of services for cloud applications.

- Windows Azure has three core components Compute, which provides a computation environment; Storage for scalable storage;

- it interconnects nodes consisting of servers, high-speed connections, and switches

Cloud based services and applications

Health care:

- There are different definition available of cloud computing, one of them is as “Cloud computing is a new way of delivering computing resources and services.
- There are many researcher and experts believe that it can improve health care services, benefit of health care research, and change the face of health information technology.”
- However, as with new innovation type, cloud computing should be more rigorously evaluated before its widespread adoption.
- Cloud computing is making its way in many field and healthcare.

Cont

- This is to use benefits of clouds services of minimum cost, effective use of resources and maximized availability of services.

- However, like in other fields healthcare is hesitant to embrace the cloud computing environment because of concerns of data privacy

- The accelerating migration to cloud computing represents a change for the way the healthcare industry sources its information technology.

Different architectures are available for Cloud Computing, broadly divide the cloud architecture into two parts: 1. Front End 2. Back End

Cont..

Energy Systems:

- Electrical power has become an indispensable part of modern day life. Hebra and NIST styled today’s electric power system as a multifaceted system of power generation.

- With the global economy more reliant on sustainable development of energy, a series of problems, such as energy shortage, electricity blackout and global warming are gaining attention.

- A smart grid is an electricity network that uses digital and other advanced technologies to monitor and manage the transport of electricity from all generation sources to meet the varying electricity demands of end-users.

Cont..

- Energy demand from the users changes dynamically in different time-periods

  A clear concept of cloud computing applications in smart grid.

- A highlight on cost effective cloud based power dispatching in smart grid

.• The rest of the paper is organized as follows.

Cont..

Transportation:

- IBM introduces urban traffic management system in the year of 1956.
- Today, transportation research and development is no longer a field dominated by civil, mechanical, operations research, and other traditional engineering and management disciplines.
- Rather, computer sciences, control, communication, the Internet, and methods developed in artificial intelligence (AI).
- Cloud computing control the traffic allocation process provides optimal solution with five stages specification.

Cont..

- In the first phase, computers were huge and costly, so mainframes were usually shared by many terminals. In the 1960s

- the whole traffic management system shared the resources of one computer in a centralized model.

- Introduction of large-scale integrated circuits and the miniaturization of computer technology.

- the IT industry welcomed the second transformation in computing paradigm.

Cont..

Manufacturing

- The Business Benefits of Cloud in Manufacturing Another important question addressed in the survey was how companies were justifying their investment .

- By far the biggest factor for manufacturers, which is consistent with overall results, is to save on hardware spend.

- Although hardware cost reduction is the dominant factor, the data shows the percentages for what was top ranked.

- For example, the second-most-frequent choice as the top business benefit was reducing software license costs

Cont..

- The Business Benefits of Cloud in Manufacturing Another important question addressed in the survey was how companies were justifying their investment

- By far the biggest factor for manufacturers, which is consistent with overall results, is to save on hardware spend.

- Although hardware cost reduction is the dominant factor, the data shows the percentages for what was top ranked

- For example, the second-most-frequent choice as the top business benefit was reducing software license costs, which switches positions with IT staff productivity in the overall sample

Cont..

Government:

- over the past 10 years Internet and Web-based services have grown rapidly and has been used by many companies.

- However, the cost of data storage and the power consumption by the hardware is increased.

- In these studies, they found a new solution to answer their challenges, to use and to get maximum benefit from the resources and it was nothing but cloud computing.

- Yes, this new technology is what which can answer thousands of their hardware and software needs.

Cont..

Here are the five key features of cloud computing

- Service demand on self. Using this feature when needed the customer can easily and automatically access to computing facilities like server, net, storage and soon from any provider.
- Ubiquitous network access. It implies that the facilities are accessible on the net and they can be used following standard methods.
- Location-independent resource pooling. This features pools different customers needed resources in the same place dynamically by the providers.

Cont..

- Cloud computing for mobile world or, rather, Mobile Cloud Computing (MCC) is a well accepted concept that aims at using cloud computing techniques for storage and processing of data on mobile devices, thereby reducing their limitations.

- While it must be noted that there were only 42.8 million Mobile Cloud Computing subscribers in 2008.

- This underlines the importance of cloud computing for mobile.

- Company users can share resources and applications without a high level of capital expenditure on hardware and software resources.

Cont..

Application development:

- Cloud computing is rapidly emerging as a new paradigm for delivering IT services as utility-oriented services on subscription basis.

- The rapid development of applications and their deployment in Cloud computing environments in efficient manner is a complex task.

- Examine the offerings in this category, and provide the basis for helping readers to understand basic application platform opportunities in Cloud by technology's.

UNIT -II
CLOUD ARCHITECTURE, PROGRAMMING MODEL

Cloud computing architecture

- Cloud computing architecture refers to the components and subcomponents required for cloud computing.

- These components typically consist of a front end platform, back end platforms, a cloud based delivery, and a network.

- Combined, these components make up cloud computing architecture

- The cloud technology architecture also consists of front-end platforms called the cloud client which comprises servers, thin & fat client, tablets &.

The Conceptual Reference Model

- Which identifies the major actors, their activities and functions in cloud computing.

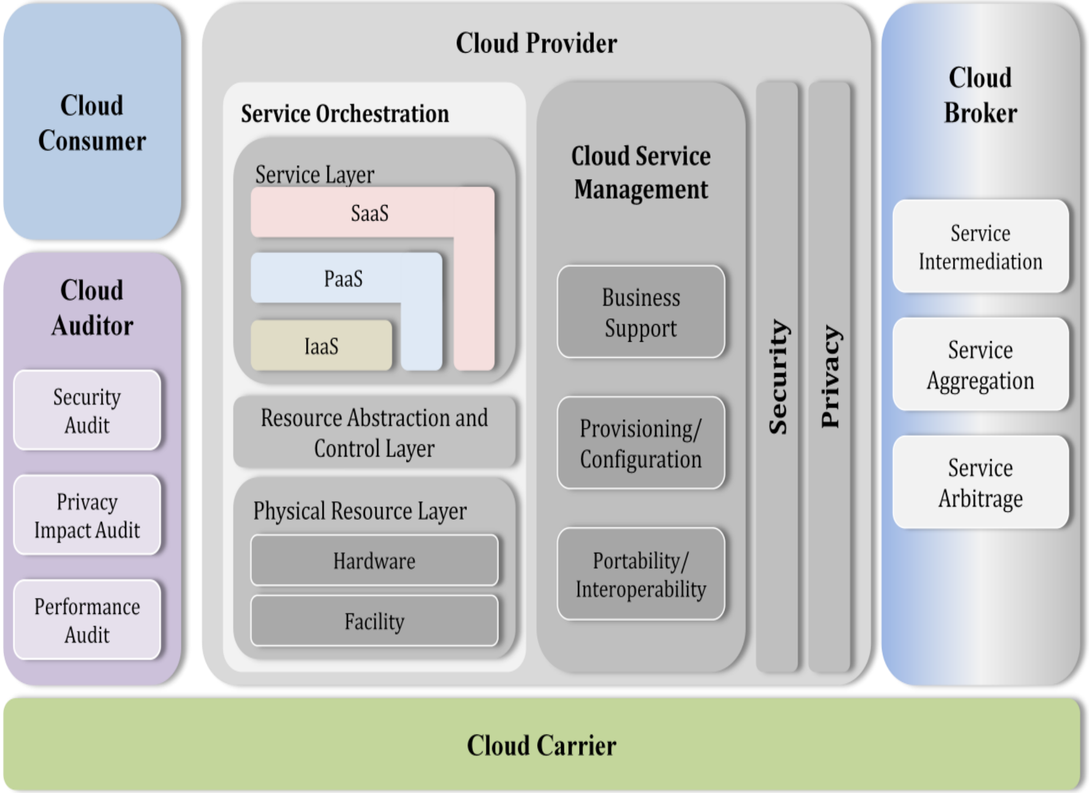
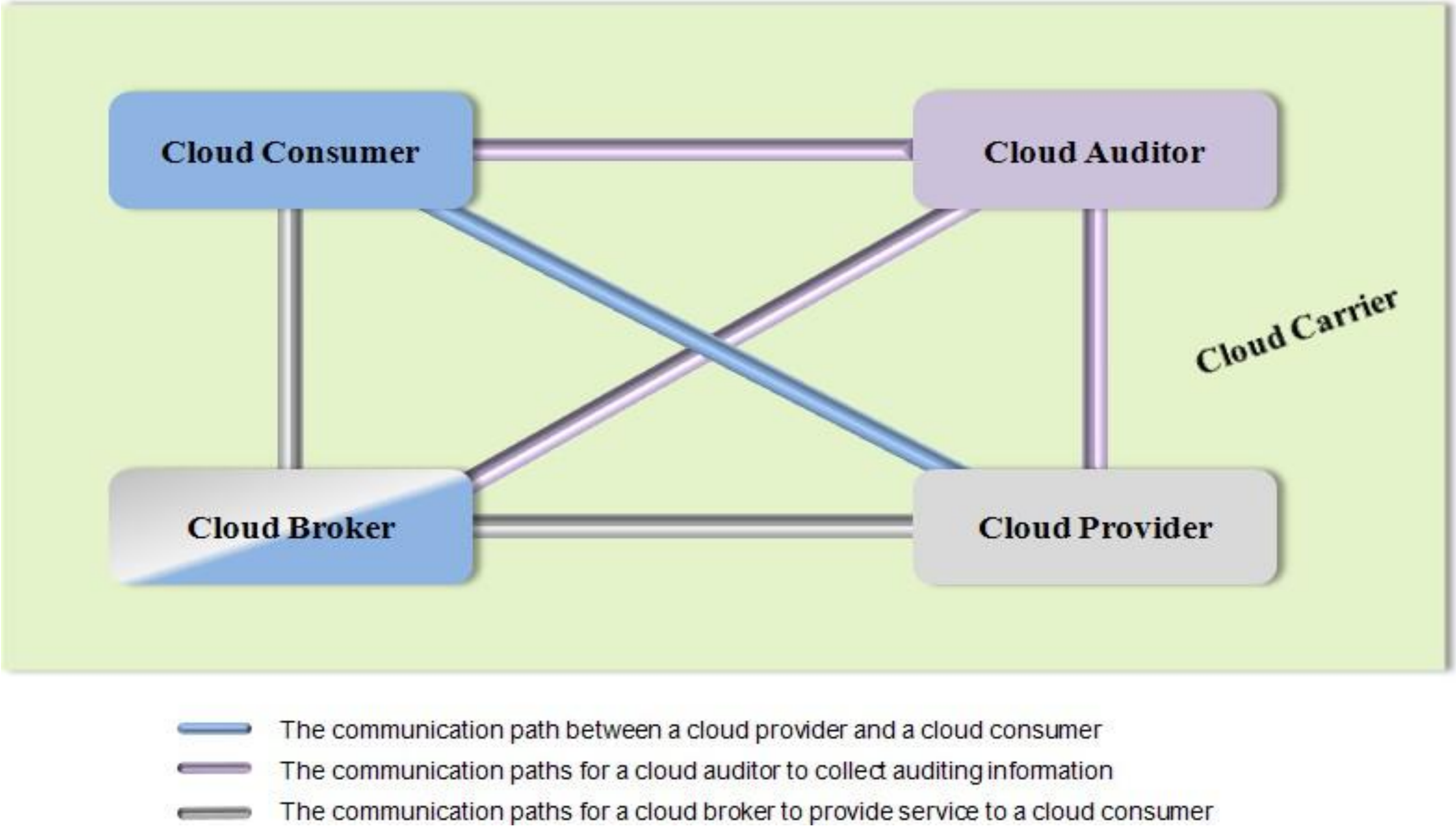


Figure 1: The Conceptual Reference Model

Cont..

- **Cloud Consumer:** A person or organization that maintains a business relationship with, and uses service from, Cloud Providers.
- **Cloud Provider:** A person, organization, or entity responsible for making a service available to interested parties.
- **Cloud Auditor:** A party that can conduct independent assessment of cloud services, information system operations, performance and security of the cloud implementation.
- **Cloud Broker:** An entity that manages the use, performance and delivery of cloud services, and negotiates relationships between Cloud Providers and Cloud Consumers.
- **Cloud Carrier:** An intermediary that provides connectivity and transport of cloud services from Cloud Providers to Cloud Consumers.

Cont..



- The communication path between a cloud provider and a cloud consumer
- The communication paths for a cloud auditor to collect auditing information
- The communication paths for a cloud broker to provide service to a cloud consumer

Figure 2: Interactions between the Actors in Cloud Computing

Cont..

- **Example Usage Scenario 1**: A cloud consumer may request service from a cloud broker instead of contacting a cloud provider directly.

- The cloud broker may create a new service by combining multiple services or by enhancing an existing service.

- In this example, the actual cloud providers are invisible to the cloud consumer and the cloud consumer interacts directly with the cloud broker

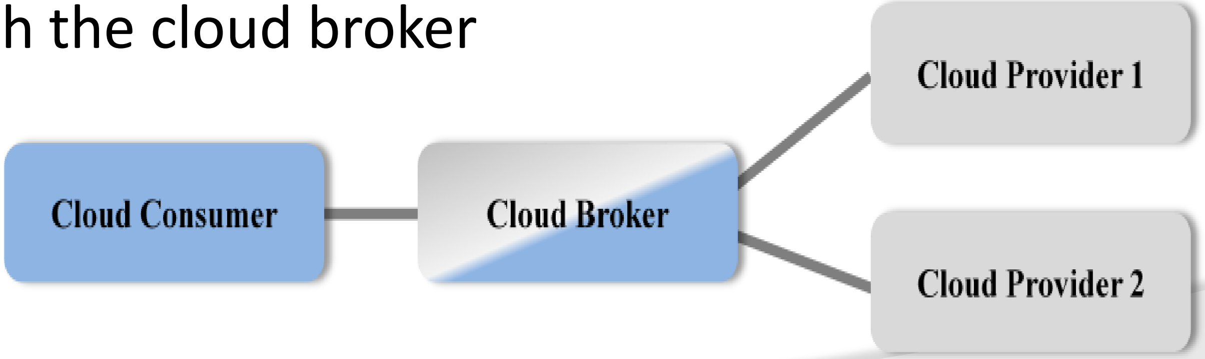


Figure 3: Usage Scenario for Cloud Brokers

Cont..

Example Usage Scenario 2:

- Cloud carriers provide the connectivity and transport of cloud services from cloud providers to cloud consumers.

- A cloud provider participates in and arranges for two unique service level agreements (SLAs), one with a cloud carrier (e.g. SLA2) and one with a cloud consumer (e.g. SLA1).

- A cloud provider arranges service level agreements (SLAs) with a cloud carrier and may request dedicated and encrypted connections to ensure the cloud services are consumed at a consistent level according to the contractual obligations with the cloud consumers.

Cont..

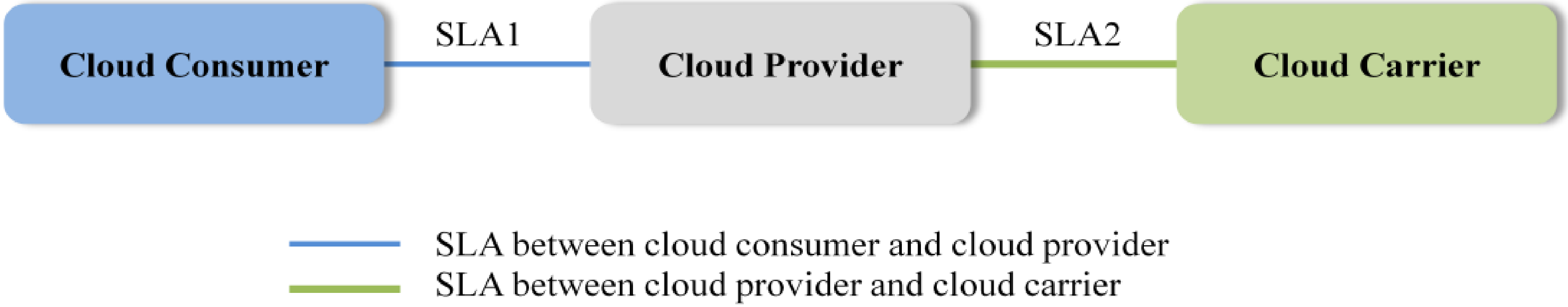

Cloud Consumer
SLA1
Cloud Provider
SLA2
Cloud Carrier
SLA between cloud consumer and cloud provider
SLA between cloud provider and cloud carrier

Figure : Usage Scenario for Cloud Carriers

Example Usage Scenario 3:

For a cloud service, a cloud auditor conducts independent assessments of the operation and security of the cloud service implementation. The audit may involve interactions with both the Cloud Consumer and the Cloud Provider.

Cont..

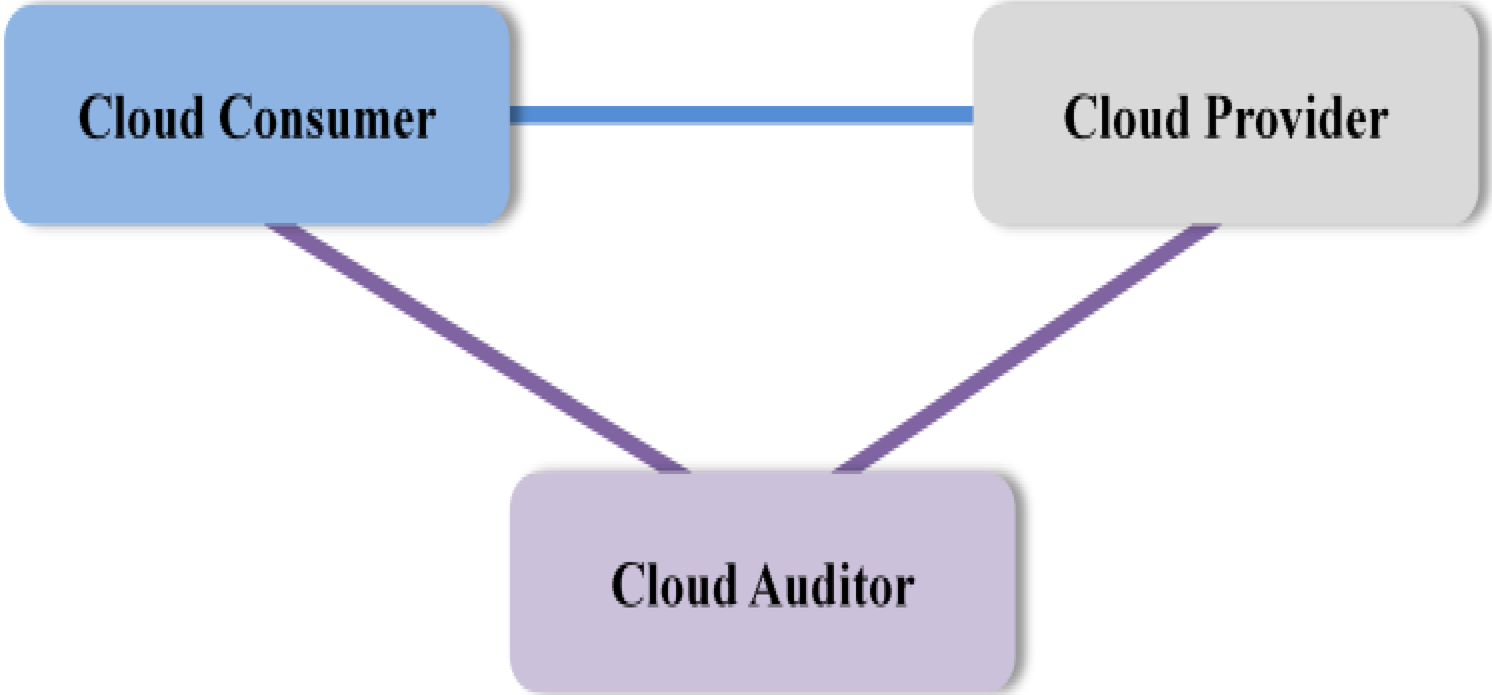


Figure 5: Usage Scenario for Cloud Auditors

The Three-Tier Model of Cloud Computing

- Cloud services are usually divided into three basic levels, or tiers, that are differentiated by the level of abstraction they present to consumers.

- The lowest tier is infrastructure-as-a-service (IaaS). With IaaS, users run software on machines owned and configured by a third party.

- IaaS makes up the hardware and software that supports the whole system, from the server to storage and networks.

- The next tier is platform-as-a-service (PaaS). PaaS is basically a platform that supports the complete life-cycle of building and delivering applications and services over the Web.

Cont..

- PaaS provides services to deploy, test, host, and maintain applications within the same integrated development environment.
- With PaaS, each platform component – whether middleware, messaging, integration, or communication – is provided as a service.
- The most visible layer to end users is software-as-a-service (SaaS), which rests on top of the PaaS layer.
- With SaaS, users can access pre-baked services simply by navigating to them via a Web browser, without bothering with the hardware and software details.

Cont..

Architectural styles of cloud applications

- Cloud transitions can be difficult to begin. Transitions can be difficult to design and plan, as much of the diligence now falls on the customer side.

- This change is a double-edged sword; it cuts both ways. It enables the customer to have significantly more control over designs, technical choices, economics, and risk.

- The baseline compute component takes into account a web layer, application layer, and database layer, each having some level of storage. Storage attributes will change based on design requirements.

Cont..

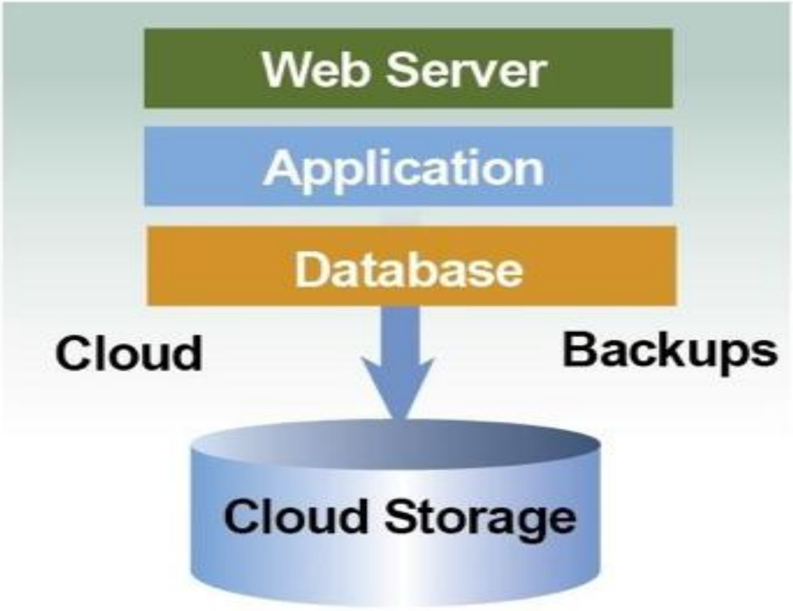


Fig 6.three-tier-model1.png

Cont..

- This type of layering is called tiering. Most designs will have three or four tiers.
- Tiers are typically the number of individual isolated layers between the environment entry point and the destination data.
- As an example, a three-tier architecture has a web layer, app layer, and database layer. A single-server architecture will have all three layers residing on the same virtual or physical server.

In this article, we will cover the following topics:

- Baseline architecture types
- OSI model and layer description
- Complex architecture types
- Architecting for hybrid clouds

Cont..

Baseline architecture types

The various types of baseline architectures are as follows.

- **Single server**

Single server templates represent the use of one server, virtual or physical, that contains a web server, an application, and a database.

Single server architectures are not very common, as they have inherent security risks as one compromise can compromise all.

These architectures are commonly deployed for development work, allowing developers to quickly build functionality without having to deal.

Cont..

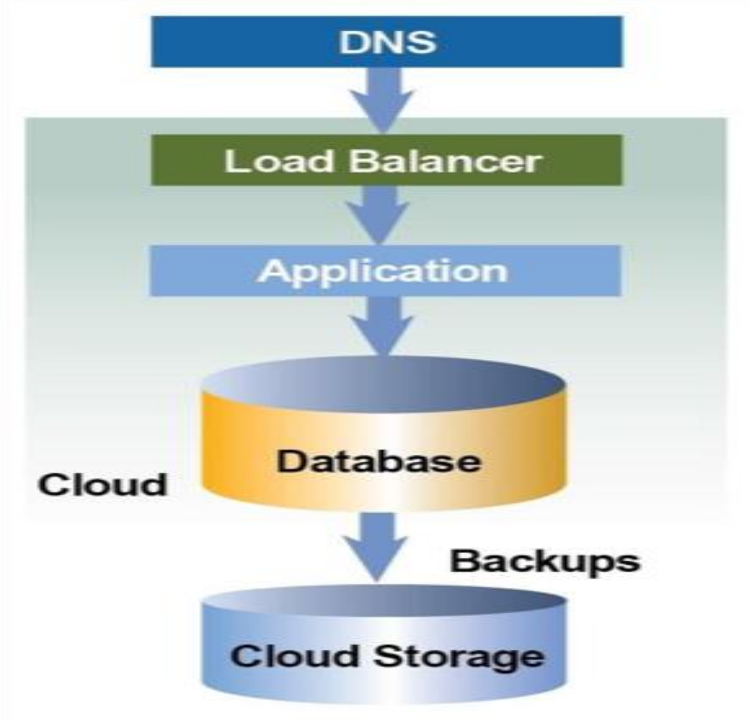


Fig.7 Single-site non-redundant-cloud architecture

Cont..

- Single-site architectures take the single server architecture and split all of the layers into their own compute instances, creating the three-tier architecture mentioned.

- There are two versions of single-site architectures:
  - ➢ non- redundant and
  - ➢ Redundant

❖ **Non-redundant three-tier architectures**

Non-redundant three-tier architectures (at right) are used to save on costs and resources but must accept a higher risk.

A single failure in any component, a single point of failure, can stop traffic flowing correctly into or out of the environment.

Cont..

- **Redundant three-tier architectures**

  Redundant three-tier architectures add another set of the same components for redundancy.

  Designing redundant infrastructures requires a well thought out plan for the components within each layer (horizontal scaling), as well as a plan for how the traffic will flow from one layer to another (vertical scaling).

- **Single points of failure**

- In redundant architectures, duplicate components eliminate the single point of failure present when only one device or component is present in the layer.

Cont..

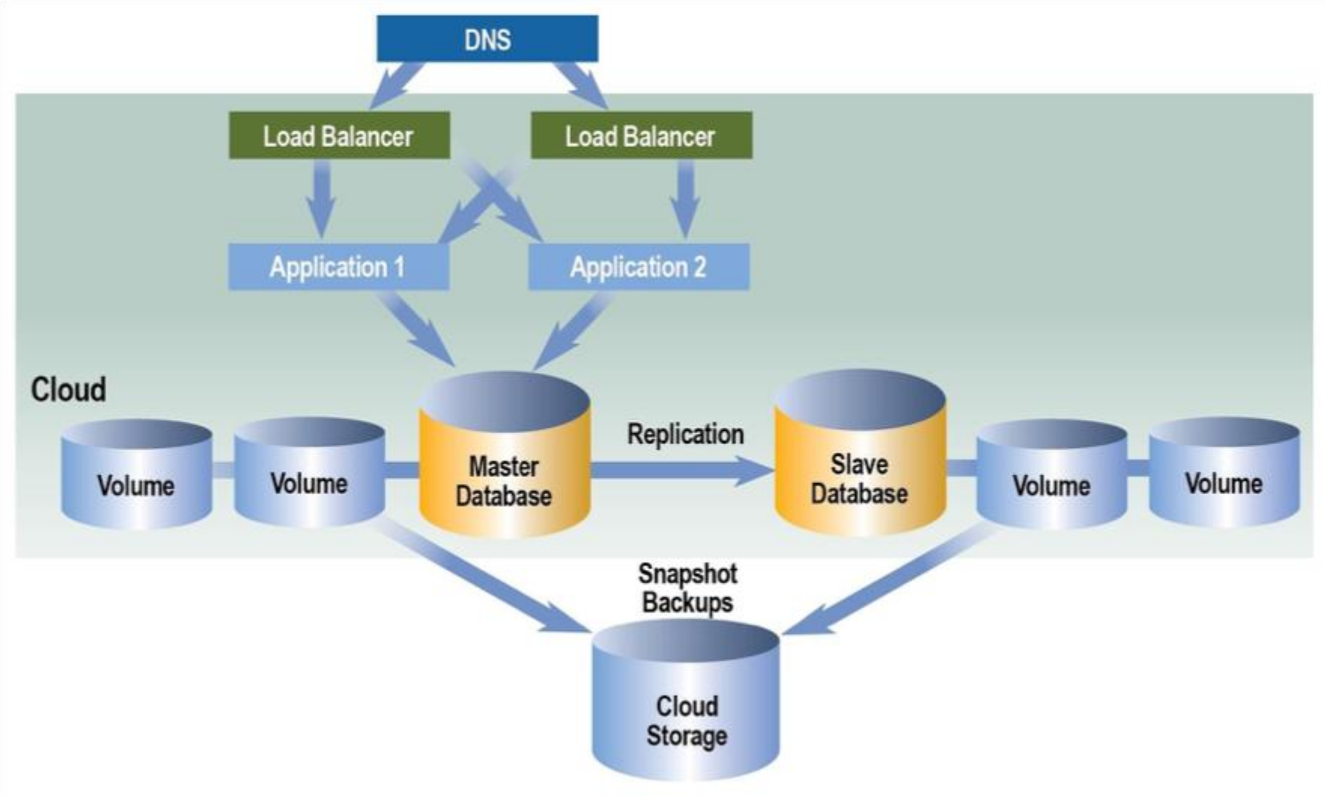


Fig .8redundant-cloud-architecture3.

Cont..

Redundancy versus resiliency

- Redundancy and resiliency are often confused. They are related, but not interchangeable.
- Resiliency, from the word resolve, relates to how to find solutions after a problem has occurred.
- Redundancy is before the issue. Resiliency is after.
- For example, redundant databases with replication can be utilized. Multiple components and copies of data create a redundant design.

Cont..

Data-intensive computing

- Data-intensive computing is a class of parallel computing applications which use a data parallel approach to process large volumes of data typically terabytes or petabytes in size and typically referred to as big data.

- Computing applications which devote most of their execution time to computational requirements are deemed compute-intensive

- The principle of collection of the data and programs or algorithms is used to perform the computation.

- The programming model utilized.

Cont..

Compute-intensive

- Compute-intensive is a term that applies to any computer application that demands a lot of computation, such as meteorology programs and other scientific applications.

- **Compute-intensive applications**

  There are a number of characteristics that can make these applications unsuitable for traditional Java EE programming models

- The need for asynchronous submission and start of work

- The need for work to run for extended periods of time

Cont..

Compute-intensive programming model is centered around two basic concepts:

- The use of jobs to submit and manage work asynchronously
- A minor extension to the asynchronous beans programming model to support work that runs for an extended period

UNIT - III

CLOUD RESOURCE VIRTUALIZATION

Cloud resource virtualization

Cloud resource virtualization:

- Resource Virtualization. resource virtualization is to create a layer of abstraction between actual physical hardware providing resources and the logical or semantic activities which consume those resources

Basics of virtualization

- In computing, virtualization means to create a virtual version of a device or resource
- Three Types Of Server Virtualization. There are three ways to create virtual servers: full virtualization, para-virtualization and OS-level virtualization. They have little in common. Physical server is called host.

Cont..

Types Of Virtualization Techniques

- Virtualization in Cloud Computing is making a virtual platform of server operating system and storage devices.

- Cloud Computing can also be known as services and application delivered to help the virtualized environment. This environment can be either public or private.

- Virtualization is technology that separates functions from hardware, while clouds rely on that split.

- Assuming intranet access, internet access, or both is already established, virtualization is what creates clouds.

Cont..

Software called a hypervisor sits on top of physical hardware and abstracts the machine's resources.

- Hardware/Server Virtualization.
- Network Virtualization.
- Storage Virtualization.
- Memory Virtualization.
- Software Virtualization.
- Data Virtualization.
- Desktop virtualization.

Full virtualization uses a special kind of software called a hypervisor. The hypervisor interacts directly with the physical server's CPU and disk space.

Cont..

Hypervisor

- Examples of this type of hypervisor include VMware Fusion, Oracle Virtual Box.

- In contrast, a Type 1 hypervisor (also called a bare metal hypervisor) is installed directly on physical host server hardware just like an operating system.

- Microsoft Hyper-V, codenamed Viridian and formerly known as Windows Server Virtualization, is a native hypervisor; it can create virtual machines on x86-64 systems running Windows.

- A server computer running Hyper-V can be configured to expose individual virtual machines to one or more networks.

Cont..

Disadvantages:

- Virtual machines are less efficient than real machines because they access the hardware indirectly.

- Running software on top of the host operating system means that it will have to request access to the hardware from the host.

- Virtualization can increase IT agility, flexibility and scalability while creating significant cost savings.

- Vitalizing your computing environments and networks brings many benefits – including lower costs, better backup and disaster recovery, a better testing environment, and better migration to the cloud.

A Virtual Machine Monitor (VMM)

- A Virtual Machine Monitor (VMM) is a software program that enables the creation, management and governance of virtual machines (VM)

- manages the operation of a virtualized environment on top of a physical host machine.

- Examples of this type of hypervisor include VMware Fusion, Oracle Virtual Box, Oracle VM for x86, Solaris Zones, Parallels and VMware Workstation.

- Three Types Of Server Virtualization. There are three ways to create virtual servers: full virtualization, para-virtualization and OS-level virtualization.

Cont..

- Virtualization limits costs by reducing the need for physical hardware systems.

- Virtual machines more efficiently use hardware, which lowers the quantities of hardware and associated maintenance costs, and reduces power and cooling demand.

Disadvantages:

- Virtual machines are less efficient than real machines because they access the hardware indirectly.

- Running software on top of the host operating system means that it will have to request access to the hardware from the host. That will slow the usability

Taxonomy of Virtual Machine

- Contexts in source publication. Virtual machine (VM) is an abstraction of computing resources presented to services to allow them to operate simultaneously on the same physical hardware infrastructure.

- VMs can be classified into two main categories: process and system virtual machines.

- A VM may be defined as a software implementation of a computing and operational environment on which system software like an operating system (OS)

- The principal focus of virtualization is to maximization the hardware utilization and to minimize hardware.

Virtualization and Virtual Machine

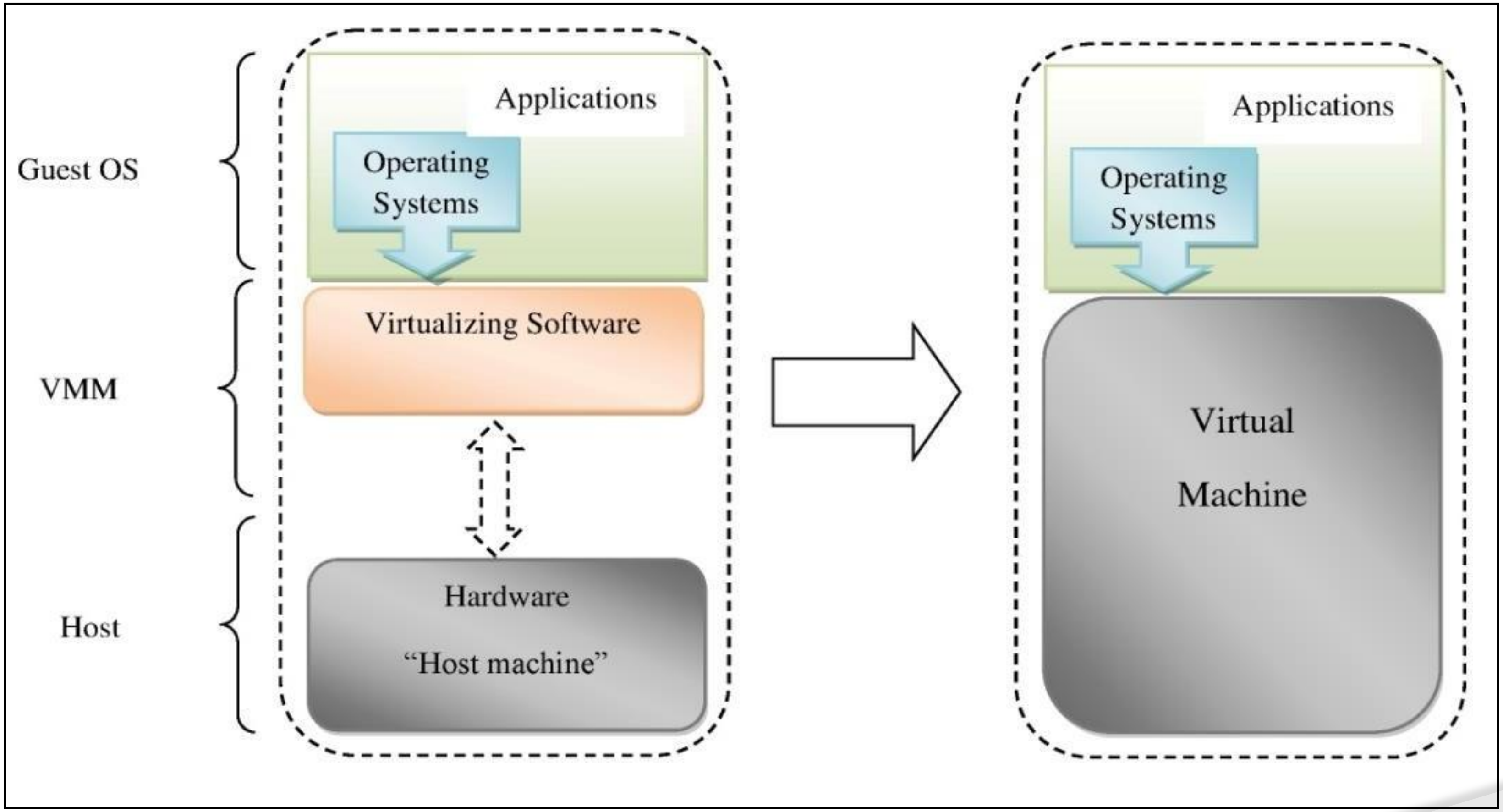
- A virtual machine is basically defined by “an efficient, isolated duplicate of a real machine.

- Virtual machines have no direct interaction with any real hardware. A VM may be defined as a software implementation of a computing

- VMs are created within the Virtualization layer; the operating system, running on the servers or data centers, can be referred as Host Operating System.

- The operating systems that run on the virtualization platform can be referred as Virtual Machine Monitor (VMM).

System Virtual Machine



Process Virtual Machine:

- A process virtual machine sometimes referred as an application virtual machine, which runs as a normal application process inside operating systems and also supports a single process.

- Process Virtual machine is created when that particular process is started and also destroyed when it exits.

- A process VM can able to provide a high-level abstraction that of a high-level programming language.

- A schematic diagram of process virtual machine presented. Above the bare metal hardware of host machine.

Cont..

We categorized the guest operating system virtualization techniques into three ways such as:

- Full virtualization
- Paravirtualization
- Hardware-assisted virtualization

❖ **Full virtualization:**

- Full virtualization requires a virtualizable architecture; the hardware is fully exposed to the guest OS which runs unchanged and this ensure that this direct execution mode is efficient.

- In Full Virtualized architecture, Guest Operating systems are provided with all the services and that is given be physical computing systems that include virtualized memory, virtual devices and also virtual BIOS.

Cont..

- In full virtualization operating system doesn’t know that it is being virtualized.

Paravirtualization

- Paravirtualization is a process that is used to virtualized a guest OS.

- It is mainly helpful because it provides better performance than hardware-assisted or full virtualization.

- The reasons for adopting paravirtualization are as follows: firstly, some features of the hardware can’t be virtualized. Secondly, it is used to present to the users a simpler interface.

- Paravirtualized Guest operating system is capable of communicating directly with the hypervisor and for that reason.

Hardware-assisted virtualization

- The shortcomings of paravirtualization which has been raised due to modification of hardware that allows the guest operating system to make communication with the hypervisor.

- Hardware-assisted guest operating system (level-1) is capable of communicating directly with the hypervisor (level-0) and for that reason.

Characteristics of Virtualization

- Consolidation
- Migration and cloning
- Stability and security
- Paravirtualization

UNIT-IV
CLOUD RESOURCE MANAGEMENT AND SCHEDULING

Cloud Resource Management and Scheduling

- Resource management is a core function of any man-made system.
- It affects the three basic criteria for the evaluation of a system: performance, functionality, and cost.
- An inefficient resource management has a direct negative effect on performance and cost and an indirect effect on the functionality of a system.
- Indeed, some functions provided by the system may become too expensive or may be avoided due to poor performance.

Cont..

- Cloud resource management is extremely challenging because of the complexity of the system.

- The strategies for resource management associated with the three cloud delivery models, IaaS, PaaS, and SaaS, differ from one another.

- In all cases the cloud service providers are faced with large, fluctuating loads that challenge the claim of cloud elasticity.

- Auto Scaling can be used for unplanned spike loads, provided that (a) there is a pool of resources that can be released or allocated on demand and (b) there is a monitoring system that allows a control loop to decide in real time to reallocate resources

Policies and mechanisms for resource management

- A policy typically refers to the principal guiding decisions, whereas mechanisms represent the means to implement policies.

- Separation of policies from mechanisms is a guiding principle in computer science.

- Cloud resource management policies can be loosely grouped into five classes:

  1. Admission control.

  2.Capacity allocation.

  3. Load balancing.

  4. Energy optimization.

  5. Quality-of-service (QoS) guarantees.

Cont..

- A system may not accept an additional workload that would prevent it from completing work already in progress or contracted.

- Limiting the workload requires some knowledge of the global state of the system. In a dynamic system such knowledge, when available, is at best obsolete.

- Capacity allocation means to allocate resources for individual instances; an instance is an activation of a service.

- Load balancing and energy optimization can be done locally, but global load balancing and energy optimization policies encounter the same difficulties as the one we have already discussed.

Resource bundling

Combinatorial auctions for cloud resources:

- Resources in a cloud are allocated in bundles, allowing users get maximum benefit from a specific combination of resources.

- Indeed, along with CPU cycles, an application needs specific amounts of main memory, disk space, network bandwidth, and so on.

- Resource bundling complicates traditional resource allocation models and has generated interest in economic models.

- **Combinatorial Auctions:** Auctions in which participants can bid on combinations of items, or packages, are called combinatorial auctions.

Cont..

- **Pricing and Allocation Algorithms**. A pricing and allocation algorithm partitions the set of users into two disjoint sets, winners and losers.

1.Be computationally tractable. Traditional combinatorial auction algorithms such as Vickey-ClarkeGroves (VLG) fail this criteria, because they are not computationally tractable.

2. Scale well. Given the scale of the system and the number of requests for service, scalability is a necessary condition.

3. Be objective. Partitioning in winners and losers should only be based on the price πu of a user's bid.

4. Be fair. Make sure that the prices are uniform. All winners within a given resource pool pay the same price.

Cont..

- The function to be maximized is max x,p f (x, p).
- The constraints in correspond to our intuition:

(a) the first one states that a user either gets one of the bundles it has opted for or nothing; no partial allocation is acceptable.

(b) The second constraint expresses the fact that the system awards only available resources; only offered resources can be allocated.

(c) The third constraint is that the bid of the winners exceeds the final price.

(d) The fourth constraint states that the winners get the least expensive bundles in their indifference set.

(e) The fifth constraint states that losers bid below the final price.

(f) The last constraint states that all prices are positive numbers

Fair queuing

- Computing and communication on a cloud are intimately related.
- Interconnection networks allow cloud servers to communicate with one another and with users.
- When the load exceeds its capacity, a switch starts dropping packets because it has limited input buffers for the switching fabric and for the outgoing links, as well as limited CPU cycles.
- Thus, a scheduling algorithm has to manage several quantities at the same time: the bandwidth, the amount of data each flow is allowed to transport;

Cont..

- A fair allocation of the bandwidth does not have an effect on the timing of the transmission.

- A possible strategy is to allow less delay for the flows using less than their fair share of the bandwidth.

- This proposes the introduction of a quantity called the bid, Ba i , and scheduling the packet transmission based on its value.

- The bid is defined as Ba I = Pa I + maxFa i−1, R t a i − δ , (6.30) with δ a nonnegative parameter.

- The properties of the FQ algorithm, as well as the implementation of a nonpreemptive version of the algorithms, are analyzed.

Start-time fair queuing

- A hierarchical CPU scheduler for multimedia operating systems was proposed.

- The basic idea of the start-time fair queuing (SFQ) algorithm is to organize the consumers of the CPU bandwidth in a tree structure;

- The root node is the processor and the leaves of this tree are the threads of each application.

- A scheduler acts at each level of the hierarchy. The fraction of the processor bandwidth, B, allocated to the intermediate node i is Bi B = wi n j=1 wj (6.31) with wj, 1 j n, the weight of the n children of node I;

Borrowed virtual time

- The objective of the borrowed virtual time (BVT) algorithm is to support low-latency dispatching of real-time applications as well as a weighted sharing of the CPU among several classes of applications.

- Like SFQ, the BVT algorithm supports scheduling of a mix of applications, some with hard, some with soft real-time constraints, and applications demanding only a best effort.

- Thread i has an effective virtual time, Ei , an actual virtual time, Ai , and a virtual time warp, Wi.

- The scheduler thread maintains its own scheduler virtual time (SVT), defined as the minimum actual virtual time Aj of any thread.

Cont..

- In this case a latency-sensitive thread gains dispatching preference as

  $E_i \leftarrow A_i$ if warpBack = OFF, $A_i - W_i$ if warpBack = ON

- The algorithm measures the time in minimum charging units (mcu) and uses a time quantum called context switch allowance (C), which measures the real time a thread is allowed to run when competing with other threads, measured in multiples of mcu.
- Typical values for the two quantities are mcu = 100 μsec and C = 100 msec.
- A thread is charged an integer number of mcu. Context switches are triggered by traditional events.
- Context switching also occurs when a thread becomes runnable after sleeping.When the thread τ_i becomes runnable after sleeping, its actual virtual time is updated as follows: $A_i \leftarrow \max[A_i, SVT]$.

Cloud scheduling subject to deadlines

- Often, an SLA specifies the time when the results of computations done on the cloud should be available.

Task Characterization and Deadlines.

- Real-time applications involve periodic or aperiodic tasks with deadlines. A task is characterized by a tuple (Ai,σi, Di), where Ai is the arrival time, σi > 0 is the data size of the task, and Di is the relative deadline.

- Instances of a periodic task, q i , with period q are identical, q i ≡ q , and arrive at times A0, A1,... Ai,..., with Ai+1 − Ai = q. The deadlines satisfy the constraint Di Ai+1 and generally the data size is the same, σi = σ. The individual instances of aperiodic tasks, i , are different.

Cont..

- Their arrival times Ai are generally uncorrelated, and the amount of data σi is different for different instances. The absolute deadline for the aperiodic task i is (Ai + Di). We distinguish hard deadlinesfrom soft deadlines

- In the first case, if the task is not completed by the deadline, other tasks that depend on it may be affected and there are penalties; a hard deadline is strict and expressed precisely as milliseconds or possibly seconds.

- Soft deadlines play more of a guideline role and, in general, there are no penalties. Soft deadlines can be missed by fractions of the units used to express them.

Cont..

- **System Model:** we consider only aperiodic tasks with arbitrarily divisible workloads. The application runs on a partition of a cloud, a virtual cloud with a head node called S0 and n worker nodes S1, S2,..., Sn.

- The system is homogeneous, all workers are identical, and the communication time from the head node to any worker node is the same.

- The head node distributes the workload to worker nodes, and this distribution is done sequentially. In this context there are two important problems:

  1. The order of execution of the tasks i .

  2. The workload partitioning and the task mapping to worker nodes

Cont..

Scheduling Policies: The most common scheduling policies used to determine the order of execution of the tasks are:

- First in, first out (FIFO). The tasks are scheduled for execution in the order of their arrival. Earliest deadline first (EDF). The task with the earliest deadline is scheduled first.

- Maximum workload derivative first (MWF). The workload derivative DCi(nmin) of a task i when nmin nodes are assigned to the application, is defined as DCi(nmin) = Wi nmin i + 1 − Wi nmin i , (6.69) with Wi(n) the workload allocated to task i when n nodes of the cloud are available; if E(σi, n) is the execution time of the task, then Wi(n) = n × E(σi, n).

Scheduling *MapReduce* applications subject to deadlines

MapReduce applications on the cloud subject to deadlines. Several options for scheduling Apache *Hadoop*, an open-source implementation of the *MapReduce* algorithm, are:

- The default FIFO schedule.
- The Fair Scheduler.
- The Capacity Scheduler.
- The Dynamic Proportional Scheduler.
- A recent paper applies the deadline scheduling framework analyzed to *Hadoop* tasks.

- We make two assumptions for our initial derivation:
- *map* and the *reduce* task, respectively, are the same for all servers.
- Load equipartition.

Resource management and dynamic application scaling

- The demand for computing resources, such as CPU cycles, primary and secondary storage, and network bandwidth, depends heavily on the volume of data processed by an application.

- The demand for resources can be a function of the time of day, can monotonically increase or decrease in time, or can experience predictable or unpredictable peaks.

- For example, a new Web service will experience a low request rate when the service is first introduced and the load will exponentially increase if the service is successful.

Cont..

- The elasticity of a public cloud, the fact that it can supply to an application precisely the amount of resources it needs and that users pay only for the resources they consume are serious incentives to migrate to a public cloud

- A very important first step in application processing is to identify the type of application and map it accordingly.

- To scale up and down a compute-intensive application, a good strategy is to increase or decrease^{b} the number of V^{b}Ms or instances.

- Because the load is relatively stable, the overhead of starting up or terminating an instance does not increase significantly the computing time or the cost.

UNIT – V
CLOUD SECURITY

Cloud Security

- Security has been a concern since the early days of computing, when a computer was isolated in a room and a threat could be posed only by malicious insiders.

- The Pandora’s box1 of threats opened wide once computers were able to communicate with one another.

- The security of computing and communication systems takes on a new urgency as society becomes increasingly dependent on the information infrastructure.

- Nowadays, even the critical infrastructure of a nation can be attacked by exploiting flaws in computer security

Cont..

- One of the consequences of the breathtaking pace of development of information science and technology is that standards, regulations.
- As a result, many issues related to privacy, security, and trust in cloud computing are far from settled.
- The pool of resources of a cloud service provider can be dispersed over several countries or even several continents.
- Since information can freely cross national borders there is a need for international regulations to be adopted by the countries where data centers of cloud computing providers are located

Cloud security risks

- Some believe that it is very easy, possibly too easy, to start using cloud services without a proper understanding of the security risks and without the commitment to follow the ethics rules for cloud computing.
- A first question is: What are the security risks faced by cloud users? There is also the possibility that a cloud could be used to launch large-scale attacks against other components of the cyber infrastructure.
- The next question is: How can the nefarious use of cloud resources be prevented? There are multiple ways to look at the security risks for cloud computing.
- A recent paper identifies three broad classes of risk: traditional security threats, threats related to system availability, and threats related to third-party data control.

Cont..

- Traditional threats are those experienced for some time by any system connected to the Internet, but with some cloud-specific twists

  The favorite means of attack are distributed denial-of-service (DDoS) attacks, which prevent legitimate users accessing cloud services;

- 1 phishing;
- 2 SQL injection;
- 3 or cross-site scripting.
- Cloud servers host multiple VMs, and multiple applications may run under each VM. Multiternency in conjunction with VMM vulnerabilities could open new attack channels for malicious users.

Privacy and privacy impact assessment

- The term privacy refers to the right of an individual, a group of individuals, or an organization to keep information of a personal or proprietary nature from being disclosed to others.
- Many nations view privacy as a basic human right. The Universal Declaration of Human Rights, Article 12, states:
- “No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation.
- The U.S. Constitution contains no express right to privacy

Cont..

- Targets primarily public clouds where privacy has an entirely new dimension because the data, often in an unencrypted form, resides on servers owned by a CSP.

- Services based on individual preferences, the location of individuals, membership in social networks, or other personal information present a special risk.

- The owner of the data cannot rely exclusively on the CSP to guarantee the privacy of the data.

Operating system security

- An operating system (OS) allows multiple applications to share the hardware resources of a physical system, subject to a set of policies.

- A critical function of an OS is to protect applications against a wide range of malicious attacks such as unauthorized access to privileged information, tempering with executable code, and spoofing.

- Such attacks can now target even single-user systems such as personal computers, tablets, or smart phones. Data brought into the system may contain malicious code.

Cont..

- A trusted-path mechanism is required to prevent malicious software invoked by an authorized application to tamper with the attributes of the object and/or with the policy rules.

- A trusted path is also required to prevent an impostor from impersonating the decider agent.

- Another question is how an OS can protect itself and the applications running under it from malicious mobile code attempting to gain access to the data and the other resources.

- Java Security Manager uses the type-safety attributes of Java to prevent unauthorized actions of an application running in a "sandbox."

Virtual machine security

- The hybrid and the hosted VM models in respectively, expose the entire system to the vulnerability of the host operating system; thus, we will not analyze these models.

- Virtual security services are typically provided by the VMM. Another alternative is to have a dedicated security

- Indeed, a VMM controls the execution of privileged operations and can thus enforce memory isolation as well as disk and network access.

- The VMMs are considerably less complex and better structured than traditional operating systems; thus, they are in a better position to respond to security attacks

Cont..

- A major challenge is that a VMM sees only raw data regarding the state of a guest operating system

- whereas security services typically operate at a higher logical level, e.g., at the level of a file rather than a disk block.

- A guest OS runs on simulated hardware, and the VMM has access to the state of all virtual machines operating on the same hardware.

- The state of a guest virtual machine can be saved, restored, cloned, and encrypted by the VMM.

Cont..

The security group involved with the NIST project has identified the following VMM- and VM-based threats:

- **MM-based threats:**

1. Starvation of resources and denial of service for some VMs. Probable causes:

(a) badly configured resource limits for some VMs;

(b) a rogue VM with the capability to bypass resource limits set in the VMM.

2. VM side-channel attacks. Malicious attacks on one or more VMs by a rogue VM under the same VMM. Probable causes:

(a) lack of proper isolation of inter-VM traffic due to misconfiguration of the virtual network residing in the VMM;

(b) limitation of packet inspection devices to handle high-speed traffic, e.g., video traffic;

Cont..

(c) presence of VM instances built from insecure VM images, e.g., a VM image having a guest OS without the latest patches.

3. Buffer overflow attacks.

- **VM-based threats:**

1. Deployment of rogue or insecure VM. Unauthorized users may create insecure instances from images or may perform unauthorized administrative actions on existing VMs.

2. Presence of insecure and tampered VM images in the VM image repository. Probable causes:

(a) lack of access control to the VM image repository;

(b) lack of mechanisms to verify the integrity of the images, e.g., digitally signed image

Security of virtualization

- The relationship between virtualization and security is a complex one and has two distinct aspects: virtualization for security and the security of virtualization.
- Two of the problems associated with virtual environments:

(a) the negative effect on performance due to the additional overhead; and

(b) the need for more powerful systems to run multiple virtual machines. In this section we take a closer look at the security of virtualization.

One of the most important virtues of virtualization is that the complete state of an operating system running under a virtual machine is captured by the VM.

Cont..

There are several useful implications regarding this fact

1. Ability to support the IaaS delivery model. In this model a user selects an image matching the local environment used by the application

2. Increased reliability. An operating system with all the applications running under it can be replicated and switched to a hot standby in case of a system failure.

3. Straightforward mechanisms to implement resource management policies:

Cont..

4. Improved intrusion prevention and detection. In a virtual environment a clone can look for known patterns in system activity and detect intrusion.

5. Secure logging and intrusion protection. Intrusion detection can be disabled and logging can be modified by an intruder when implemented at the OS level.

- A first type of undesirable effects of virtualization leads to the diminished ability of an organization to manage its systems and track their status:

New Virtualization System-Specific Attacks

- **VM jumping/guest hopping**

- Attackers take advantage of hypervisor escape vulnerabilities to "jump" from one VM to another
- **VM attacks**

- Attacks during deployment and duplication
- Deletion of virtual images
- Attacks on control of virtual machines
- Code/file injection into virtualization file structure
- **VM migration**

- VM migration is transfer of guest OS from one physical server

Cont..

VM migration attack

- If migration protocol is unencrypted, susceptible to man-in-the-middle attack
- Allows arbitrary state in VM to be modified
- In default configuration, Xen Motion is susceptible (no encryption)
- VMware’s VMotion system supports encryption
- Proof-of-concept developed by John Oberheide at the Univ. of Michigan

- **Management server attacks**
- Exploit management console vulnerabilities that divulge password information

Cont..

- **Administrative VM attacks – exploit vulnerabilities to:**
- Cause a denial of service by halting the system
- Cause a denial of service by crashing the administrative VM
- Obtain passwords that are stored in clear text
- Exploit buffer overflows in exposed services to execute arbitrary code
- Exploit vulnerable services to gain elevated privileges
- Bypass authentication

- **Guest VM attacks – exploit vulnerabilities to:**
- Gain elevated privileges
- Crash the virtual machine

Cont..

Hypervisor attacks – exploit vulnerabilities to:

Cause the hypervisor to crash

Escape from one guest VM to another

- **Hyperjacking**

Consists of installing a rogue hypervisor

- One method for doing this is overwriting page files on disk that contain paged-out kernel code
- Force kernel to be paged out by allocating large amounts of memory
- Find unused driver in page file and replace its dispatch function with shell code
- Take action to cause driver to be executed
- Shell code downloads the rest of the malware

Cont..

- Host OS is migrated to run in a virtual machine
- Has been demonstrated for taking control of Host OS
- Hyper jacking of hypervisors may be possible, but not yet demonstrated
- Hypervisors will come under intense scrutiny because they are such attractive targets