EBU7501: Cloud Computing
Week 1, Day 2: SLA and Virtualisation

Dr. Gokop Goteng

Lecture Aim and Outcome

- Aim
  - The aim of this lecture is to emphasise the importance of Service Legal Agreement (SLA) between service providers and service subscribers and to introduce virtualisation as a key component of the cloud to students.
- Outcome
  - At the end of this class students should learn the following:
    - Know the answers to service providers and service subscribers' responsibilities
    - Create virtual machines using hypervisor
    - Know the different commercial and open source virtualisation technologies
    - Know the different types of virtualisations

Lecture Outline

- **SLA: Service Level Agreement**
- **Types of SLAs**
- **Quality of Service (QoS) in SLAs**
- **Virtualization**
- **Types of Virtualizations**
- **Structure of Virtualization**
- **Architecture of Microsoft Virtualization Technology**
- **The Hypervisor or Virtual Machine Manager (VMM)**
- **Architecture of Xen Virtualization Technology**
- **Virtualization Technologies**
- **The Big Players in Virtualization**
- **Advantages of Virtualization**
- **Challenges of Virtualization**

SLA: Service Level Agreement

- Service Level Agreement (SLA) in cloud computing is a service-based contractual legal agreement between a cloud service provider (e.g. Amazon Web Services, Google, Microsoft Azure Cloud, Salesforce.com) and service requester/subscriber (customer) who will be the user of these resources or services (e.g. compute, storage, software, backup, databases, network, security, support, etc) that are provided as services on a demand basis.
- SLAs are the first document that is required before a cloud resource will be used
  - The provider's responsibilities are defined
  - The subscriber's responsibilities and entitlements are also defined in the contractual document
  - Subscribers can be individual persons/organisations or a group of persons/organisations
  - There is a billing and accounting system embedded in the contract
    - Normally the subscriber's credit/debit card information is registered
  - Failure of either the provider or subscriber to meet their obligations may result in violation of the contract and may carry some penalties which can be financial or withdrawal of the services.

Types of SLAs

- Service-based SLA
  - This is based offering shared resources that are used by many customers.
  - This is the kind of SLA offered to cloud users
- Customer-based SLA
  - This is an agreement between an individual customer or individual group to provide non-shared specific services
- Multi-level SLA
  - Corporate-Level SLA
    - This covers all users of a particular organisation
    - This is common to generic services
  - Customer-Level SLA
    - This covers particular user group
  - Service-Level SLA
    - This covers specific type of services for users

Boundaries in the Cloud Community

- Organisational boundary
  - This is the physical perimeter that surrounds IT resources that are owned and governed by a cloud organisation or provider
- Trust boundary
  - This is when the cloud consumer extends trust beyond its physical organisation to include those resources that are consumed from sources provided by the cloud service provider

Characteristics of the Cloud

- On-demand usage
  - Cloud consumers can use cloud resources only when they need it and be charged only on the amount of usage
- Multi-tenancy
  - This describes the characteristic of cloud resources running as “instances” that can be used/shared by many cloud consumers at the same time
  - It is also known as Resource Pooling
- Elasticity
  - This is the ability of cloud resources to scale up or down automatically based on the settings and contractual agreements of the service
- Ubiquitous access
  - This is the ability of cloud services to be accessed anywhere any time from many platforms, technologies and devices
- Measured usage
  - This is the ability of cloud platforms to keet track and record of the amount of IT resources used (consumed) by cloud service consumers
  - The time stamp, consumer identity, location and cost metrics logs should ke kept for accounting and auditing purpose
- Resilience
  - This is the ability to distribute/replicate IT resources to physical locations as redundant and failover mechanisms in case of system failures
  - This is a backup strategy for cloud platforms

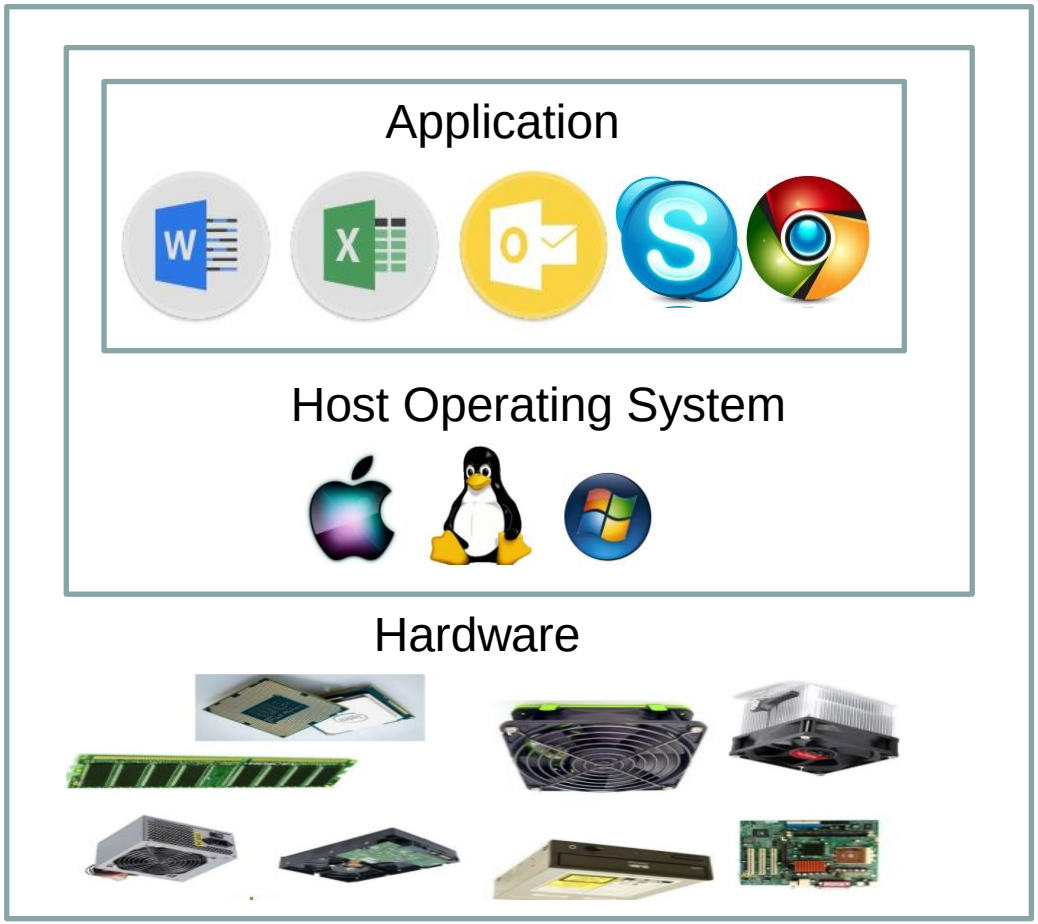
Quality of Service (QoS) in SLAs

- SLAs are based on quality of service (QoS)
- QoS
  - This is a combination of defined metrics relating to customer's requirements and satisfaction of services offered by a service provider
  - They can be quantitative or qualitative in nature
- QoS Cloud-Based Parameters
  - Reliability and availability
  - Security and trust
  - Scalability
  - Data integrity and data protection legal issues
  - Resilience
  - Data recovery time
  - Support

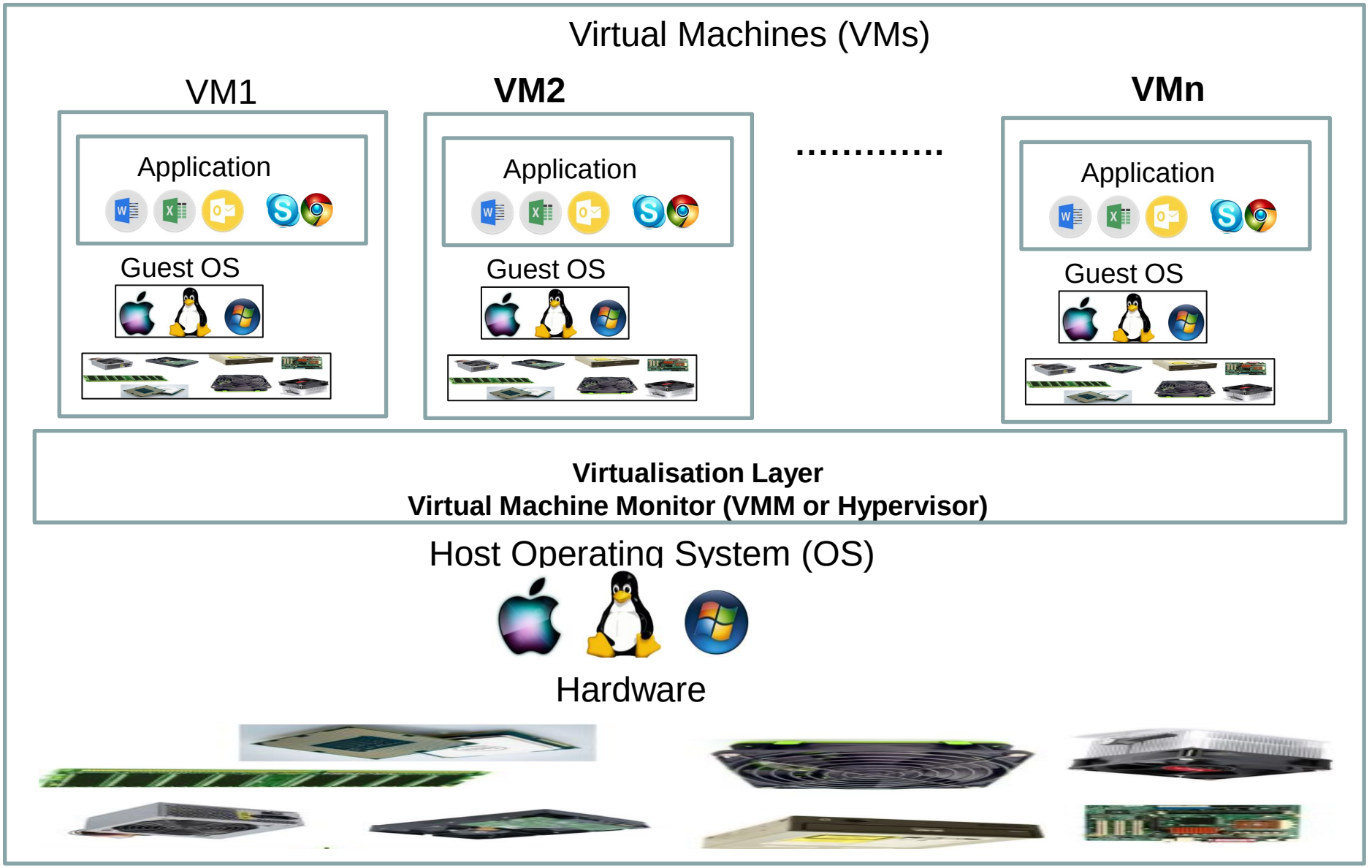
Virtualisation

- In computer science, “Virtualisation” is the process of creating the “virtual” or “software” version of hardware (computer, processor, network, disk storage, memory, keyboard, etc), database, operating system and application from their “actual” or “real” copies using virtual machine monitor (VMM) to emulate the capabilities and functionalities of the physical copy.
  - It is a computer architecture technology in which multiple virtual machines (VMs) are multiplexed (incorporated into a single system) in the same hardware and its resource to serve as an abstraction of the physical server where the guest OS is not bound or attached to the physical hardware (Hwang, Fox and Dongarra, 2012)
  - The virtual version can be complete computer system with operating system (OS), hard disk, memory, CPU and network interface
  - The virtual computer consists of “guest” OS that runs on top of the “host” OS of the physical computer
  - The guest and host OSs can be different (one can be Windows and the other Linux)

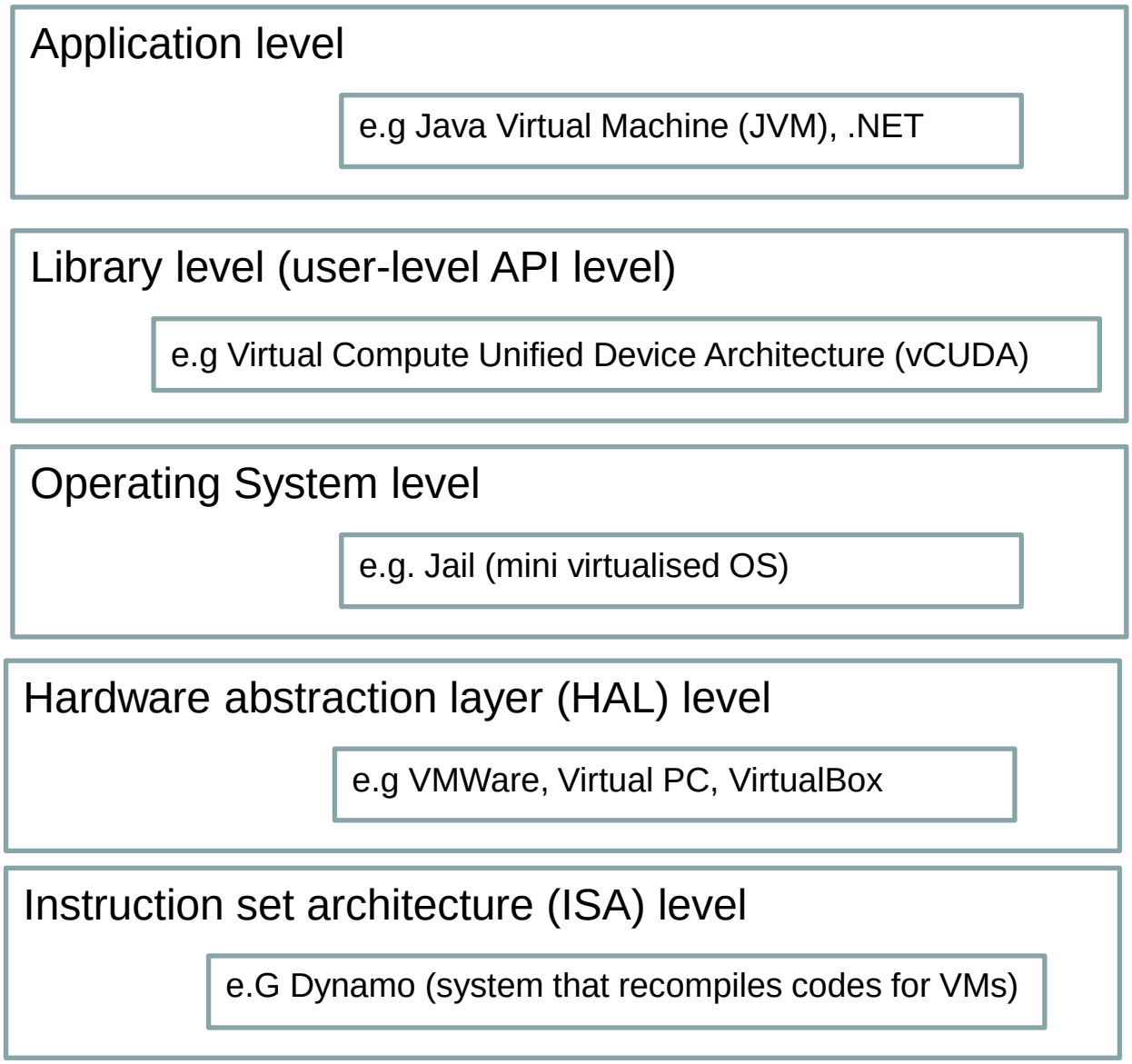
Computer for Transition to Virtualisation



Computer after Virtualisation



Implementation Levels of Virtualisation



Implementation Levels of Virtualisation

- Instruction Set Architecture (ISA) level
  - ISA level is where instructions using computer codes are interpreted to emulate and extend the functions of physical components to the virtual components
- Hardware abstraction level
  - This is where virtual copies of the actual physical hardware are created and the virtual copies eh VMs are located directly on top of the actual physical hardware
- Operating system level
  - This is the layer between actual traditional OS and user applications.
- Library support level
  - These are APIs (application programming interfaces) exposed by user level libraries rather than complex calls by OS
- User-application level
  - This is the virtal applications that run on VMs for users to use them

Types of Virtualisations

- Desktop Virtualization
  - This is the creation of one or multiple virtual or “logical” desktop computers from a physical desktop machine
  - This is the most popular type of virtualization
- Hardware Virtualization
  - This is the creation of virtual computer with OS that represents and act as a real physical computer
  - The virtual machine is created from the physical machine
  - Virtual desktop infrastructure (VDI), which is a form of desktop virtualization is also a hardware virtualization
  - The virtual machine is separated from the physical machine by a hypervisor or virtual machine manager (VMM)
  - The virtual machine runs a guest operating system which can be different than the host OS
    - For example the guest OS can be Linux and the host OS can be Windows

Types of Virtualisations

- Nested Virtualisation
  - This is the type of virtualization in which virtual machine and its virtual computing resources are run on another virtual machine.
  - This means running one or more hypervisors in another hypervisor.
- CPU virtualization
  - A VM is a duplicate (copy) of the entire computer system, but majority of instructions on the VM are run on the host through the hypervisor.
  - However, to reduce much of the heavy lifting on the host, virtual CPU can be created, this is known as CPU virtualization
- Memory virtualization
  - This is a process where by virtual memory is created and run on the VM and controlled by the virtual CPU of the VM but cannot contact the memory of the physical machine. The VMM coordinates the actions on addresses of the physical memory.
- I/O virtualization
  - I/O virtualization is the process of managing the routing of I/O requests between virtual devices and the shared physical hardware

Virtual Clusters

- A physical cluster is a collection of servers (physical machines) which are tightly or loosely interconnected by a physical network such as the local area network (LAN) to work on solving common problems
- Virtual cluster on the other hand, is the collection of VMs interconnected by a virtual network across several physical networks
- Physical and virtual cluster comparisons:
  - Virtual cluster nodes can be either physical or virtual machines
  - Virtual clusters can have multiple VMs running different guest OSs
  - The purpose of using VMs is to integrate and consolidate multiple functionalities on the same server there by enhancing server optimum utilisation and application flexibility
  - In virtual cluster, VMs can be replicated in multiple servers to promote distributed/parallel computation, fault tolerance and disaster recovery
  - The number of nodes in virtual cluster can grow or shrink dynamically such as in a physical P2P system
  - The failure of VMs will not cause a shutdown to the host machine, but the failure of any physical nodes may disable some VMs installed on the failing physical node.

Migration of VM and Virtual Resources

- The use of virtual clusters enables VMs to be moved from one cluster to another or its functionalities moved to another in the case of any failure
- If the physical host fails, all its residing VM guest machines will fail. This can be mitigated using VM life migration. The migration copies the VM state file (snapshot) from the storage area to the host machine
- There are 4 ways to manage a virtual cluster
  - Using Guest-based Manager, in which the cluster manager resides on the quest system with the virtual cluster consisting of multiple VMs.
  - Using Host-based Manager by building a cluster manager on the host system which supervises the guest system and can restart the guest system on another host machine
  - Using independent cluster manager on both the host and guest systems
  - Using an integrated cluster manger in which VMs can be live-migrated from one physical machine to another and a VM can replace another VM

VM Migration Steps

- There are six steps in VM migration:
  - **Steps 0 and 1: Start migration**
    - This is where preparations for migrating VM are made. In this steps you determine the migrating VM and its destination host. This can be done manually or by automated strategies known as “load balancing” and “server consolidation”
  - **Step 2: Transfer memory**
    - The entire states of the VM are stored in the memory of the VM, so transferring its memory to the destination host ensures continuity of the states of the VM
  - Step 3: Suspend the VM and copy the last portion of the data
    - The migrating VM’s execution is suspended when the last round’s memory data is transferred. Non memory components such as the network and CPU should also be sent to the destination host system.
  - Steps 4 and 5: Commit and activate the new host
    - After all the required data is copied to the destination host, the VM reloads and starts execution of programs. The network connection is redirected to the new VM with its dependencies and the process finishes by removing the original VM from the source host.
- You can migrate memory, file system and network known as memory migration, file system migration and network migration respectively

Cloud infrastructure Mechanism

- These are the main building blocks of any cloud infrastructure
- They are:
  - Logical network perimeter
    - This is a virtual network boundary that can encompass and isolate a group of related cloud-based resources that may be physically distributed
    - For example isolate cloud resources from non authorised users
  - Virtual server
    - This is a virtualisation software that emulates and behave exactly like a physical server
    - Multiple virtual servers can run on one physical server that are shared by multiple cloud users
  - Cloud usage monitor
    - This is an autonomous software used for collecting and processing cloud resource usage data
  - Resource replication
    - This is the creation of multiple instances of the same cloud resource to various locations used by multiple cloud users that are physically distributed around the world
  - Ready-made environment
    - This is the environment provided by PaaS

Cloud Mechanisms

- Automated scaling listener
  - This is a cloud service agent that monitors and tracks communications between service consumers and cloud services for dynamic scaling up or down resources automatically
- Load balancer
  - This a cloud runtime mechanism that distributes workload (data processing, computation, network traffic, etc) across two or more cloud resources for better performance. This is the automated horizontal scaling in cloud platforms
- SLA (Service Level Agreement) monitor
  - This is used to observe the runtime performance of cloud services to ensure that they are fulfilling the contractual Quality of Service (QoS) requirements between service providers and consumers
- Pay-per-use monitor
  - This ensures that cloud-based resource usage is in accordance with pre-defined pricing parameters and agreements
- Audit monitor
  - This system collects audit tracking data for networks and cloud resources in support of regulatory and contractual obligations
- Failover system
  - This mechanism is used to increase the reliability and availability of cloud resources using clustering technologies to provide redundant or standby resource instances in case of any failure

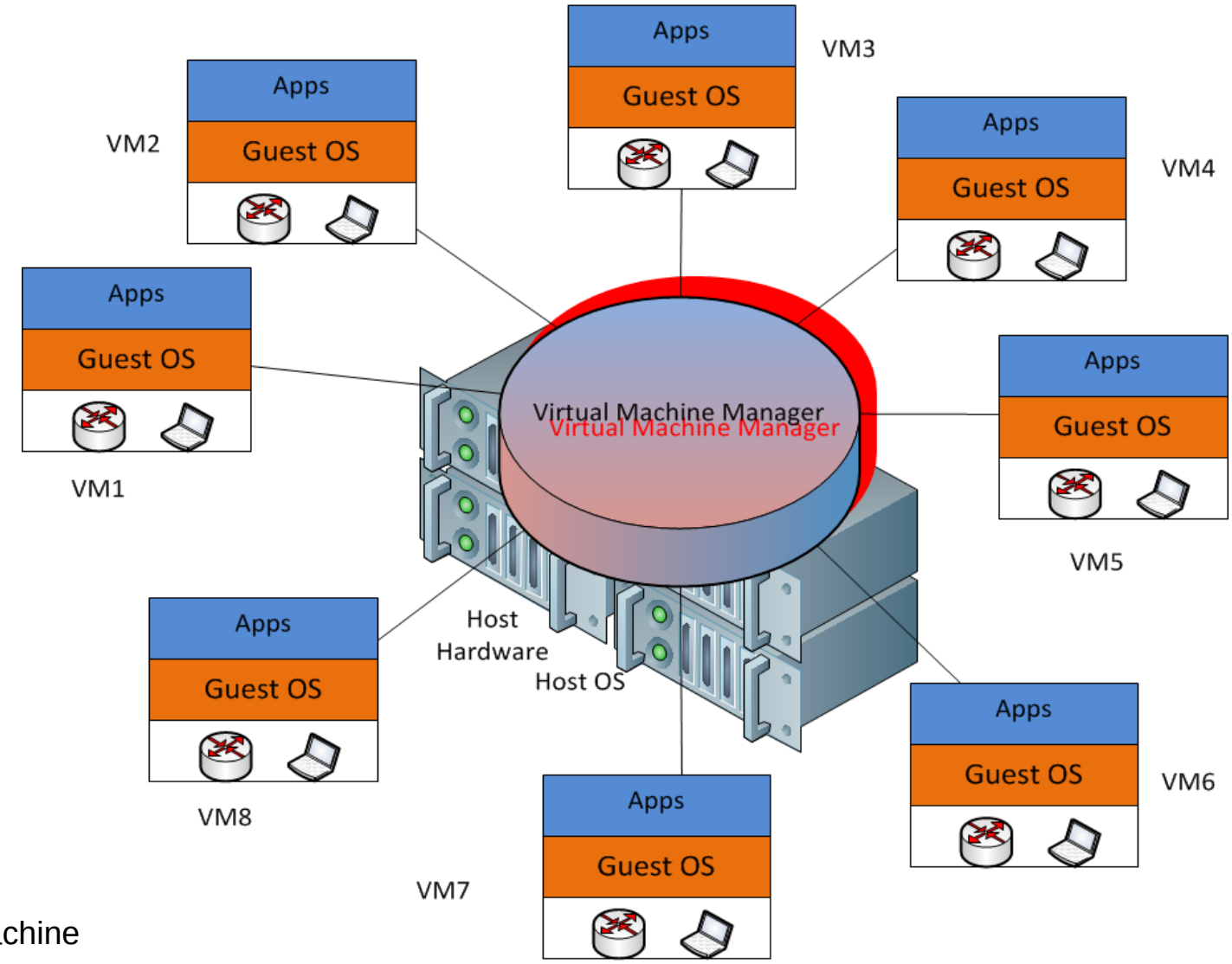
Cloud Mechanisms Continued…

- Hypervisor
  - This mechanism is used to generate virtual server instances from physical servers
- Resource cluster
  - The cloud resources are geographically distributed and this mechanism groups the resources instances together for high performance
- Multi-device broker
  - This mechanism converts and exchange information to multiple cloud service consumers running different devices/platforms but are accessing the same cloud service provider's cloud resources
- State management database
  - This is a storage database device that is used to temporarily maintain (persist) state data for software programs to improve performance as is done in caching for memory performance in computer systems

More Examples of Types of Virtualization

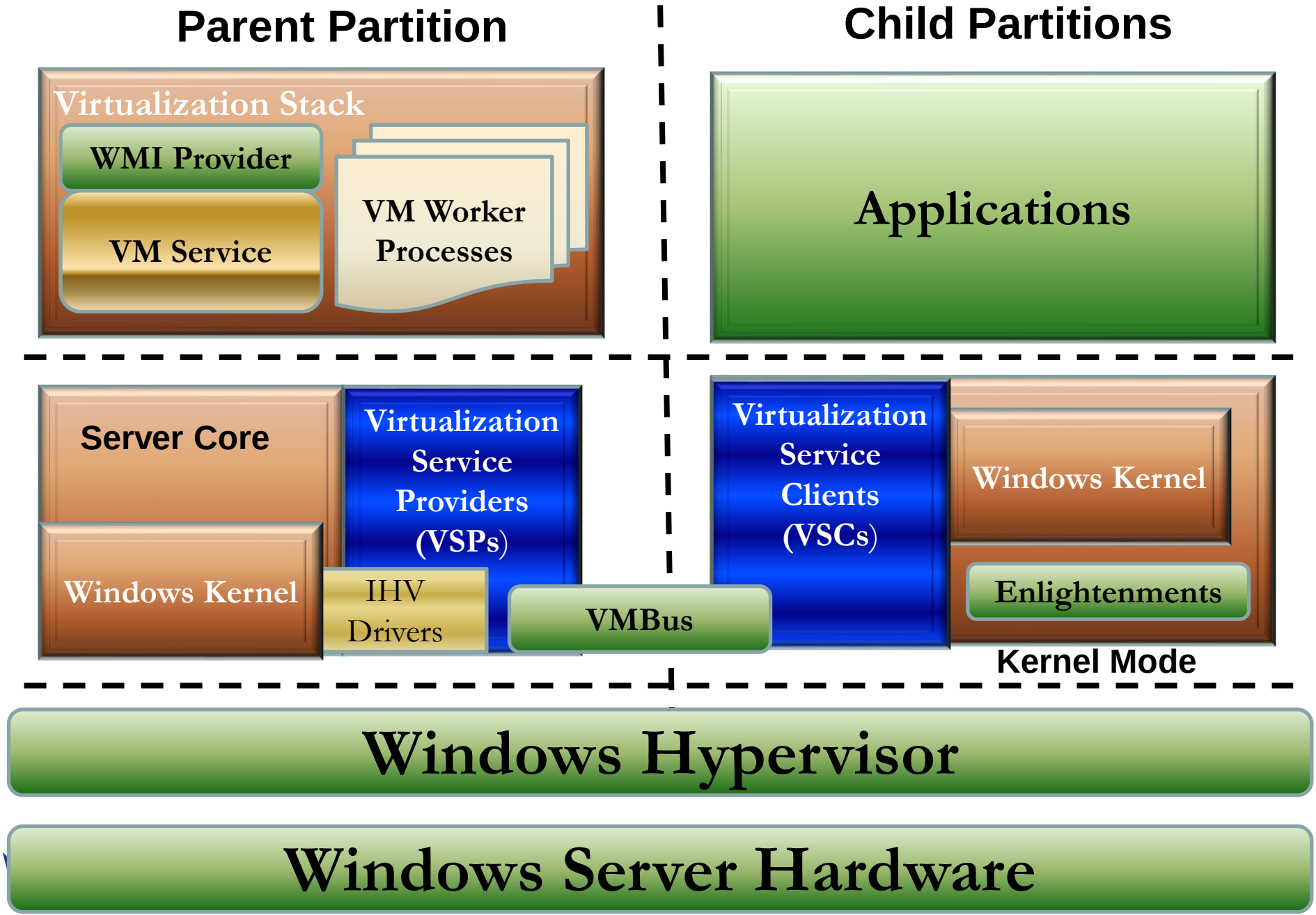
- Other forms of virtualization
  - Software virtualization
  - Data virtualization
  - Database virtualization
  - Network virtualization
  - Virtual private network
  - Storage virtualization
  - Memory virtualization
  - Server virtualization

Structure of Virtualization



Key:
VM-Virtual Machine

Architecture of Microsoft Virtualization Technology



Some Terminologies used in MS Virtualization

- Parent Partition
  - Logical hard drive division responsible for running the hypervisor layer and creating/managing child partitions
- Child Partitions
  - Logical hard drive division responsible for running isolated OSs
- Virtualization Stack
  - Collections of resources including the hypervisor which makes up the Microsoft Hyper-V environment
- Virtualization Service Providers (VSPs)
  - A parent partition side server component that handles I/O requests
- Virtualization Service Clients (VSCs)
  - A child partition component which sends request to VSP

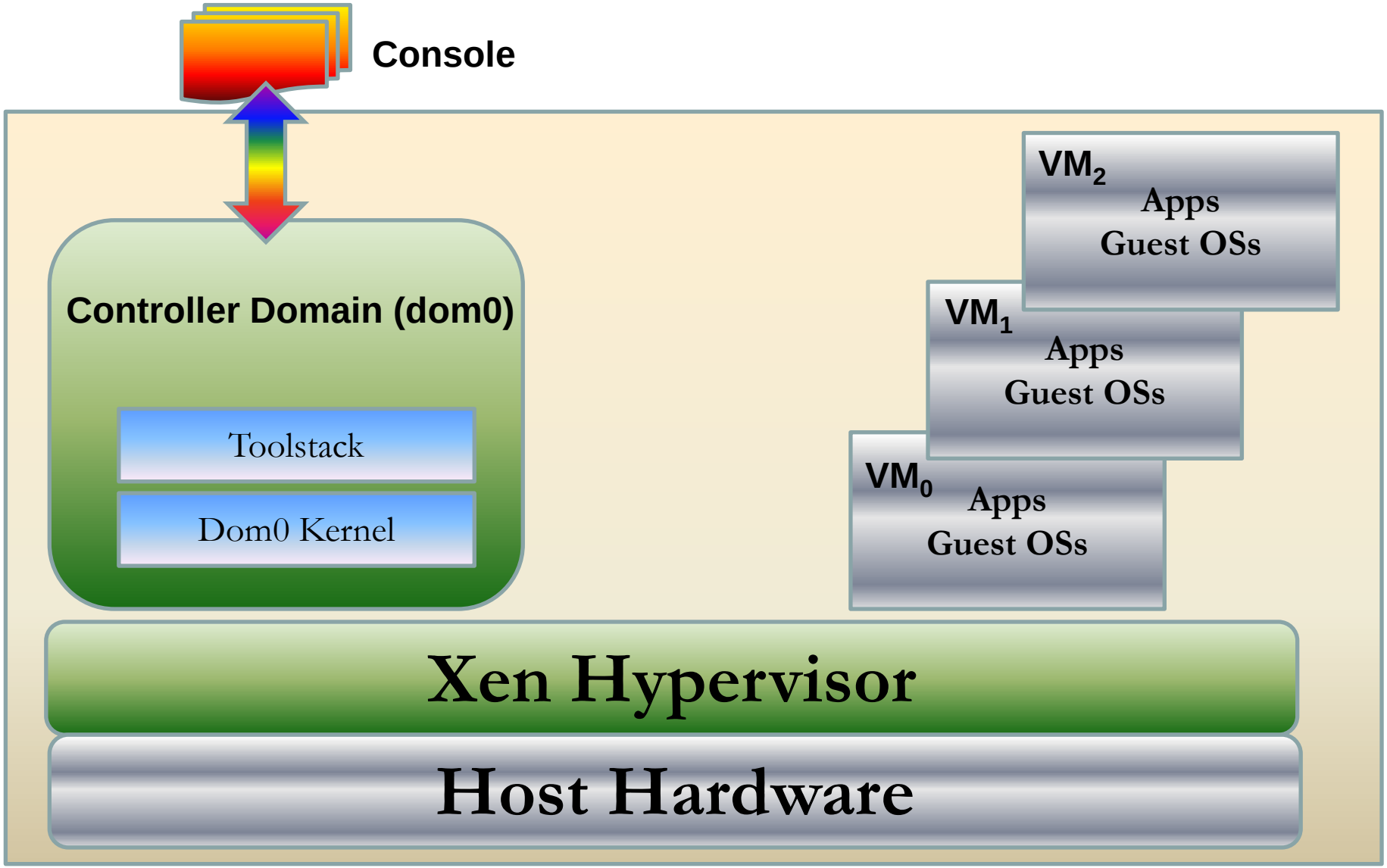
Some Terminologies used in MS Virtualization

- VMBus
  - A communication system for sending requests and data between virtual machines
- WMI (Windows Management Instrumentation) Provider
  - Manages virtual devices
- Windows Kernel
  - Manages I/O requests from parent and child components
- IHV (Independent Hardware Vendor) drivers
  - These are replaced by Microsoft drivers
- VM Worker Processes
  - Runs and configure one instance of a VM at a time

The Hypervisor or Virtual Machine Manager (VMM)

- Manages a minimum set of hardware components
  - Processors, physical address space
- It schedules and isolate virtual components to enhance loose coupling
- Provides security to the virtual resources
- Enhance performance of hardware
- Simplifies management of virtual infrastructure
- It supports virtualization processes and actions

Architecture of Xen Virtualization Technology



Xen Virtualization Technology

- Host Hardware
  - Consists of the real hardware such as processor, I/O devices, memory, hard disk, network interface
- Xen Hypervisor
  - It is a very lean software layer with about 150, 000 lines of codes that runs directly on the host hardware and manages the CPUs, memory, I/O devices
- Guest Domains/Virtual Machines
  - These are virtual machines running individual OSs and applications
- Control Domain or Domain 0 (dom0)
  - Is a specialised virtual machine with special privileges to access the hardware directly, handles all access to I/O functions and interacts with other VMs
  - It exposes its control interface to the outside world which is used to control it
  - It is the first VM to be started
  - Without dom0, the Xen Hypervisor will not function
- Toolstack and Console
  - Domain 0 contains a control stack known as Toolstack whick enables users to create, destroy and configure virtual machines
  - Toolstack exposes an interface that uses that can be activated via a command line, graphical user interface (GU) or cloud orchestration tools such as OpenStack and CloudStack

Virtualization Technologies

- Host Components
  - These are the “real” and tangible hardware, operating systems (OS), networks and databases that formed part of the “Host” machine or datacenter
- Guest Components
  - These are the systems created from the host machine
  - This includes the OS, network, databases and servers
  - They are the “virtual” components such as virtual machines (VMs)
- Hypervisor or Virtual Machine Manager
  - The system that manages virtual machines
    - To create, edit, start, stop, control and view performance and running statistics of virtual machines
- Networking
  - It uses networking technologies for distributed systems
- Web Services
  - Web services help to provide virtual machines over the internet and cloud
- Front end systems
  - The interface between users and the applications running on virtual machines

Advantages of Virtualization

- Less cost compared to having a datacenter
  - Reduction in cost of hardware, energy, cooling, expertise and administration
- Reduction in time of deployment and implementation
- Increased efficiency of using hardware
- Increased scalability
- Increase reliability and availability of resources
- Quick response to data disaster recovery
- Flexibility of systems specifications on the fly
- Ease of creating servers that run on different operating systems
- Ease of backups of complete system as snapshots
- Ease of migrating running systems

Challenges of Virtualization

- Security vulnerability as it is based on using shared resources
- Support for the system may require certain level of expertise
- Requires reliable backup systems to recover data during disaster
- Failure of host systems may affect the whole guest systems
- Datacenters generate huge energy that affect the environment
- Bandwidth, network and latency performances may not be as good as using physical servers

The Future of Virtualization

- It is and it will still be the dominant component of cloud computing in the future
- Increased decoupling of applications and encapsulating them on virtual machines
- Better energy consumption OS and VMM to manage energy used by datacenters
- It will play a dominate economic factor in the IT market
- More scalable and robust disaster recovery and backup systems will evolve
- Better security models will also evolve for users to have more confidence in virtualization

Class Task

- Sketch and label a simple diagram of virtualisation
- Explain the different components of virtualisation
- What type of virtualisation is most popular and why?
- Differentiate between cloud computing and virtualisation.
- List 3 advantages of virtualisation
- What is SLA?
- What is the importance of SLA?
- What is QoS? List 4 QoS parameters.