

CLOUD COMPUTING-(18MCA43C)

UNIT – III

'Virtual Machines'

FACULTY:

Dr. R. A. Roseline, M.Sc., M.Phil., Ph.D.,
Associate Professor and Head,
Post Graduate and Research Department of Computer
Applications,
Government Arts College (Autonomous), Coimbatore – 641 018.

UNIT-III

- Virtual Machines Provisioning and Migration Services

Introduction and Inspiration

- Cloud computing builds on service-oriented architecture (SOA), grid computing, and virtualization technology
 - Offers infrastructure as a service to the end users as a public utility service
 - Based on pay-as-you-use and on-demand computing models
 - Referred to as Infrastructure as a Service (IaaS)
- The provisioning of the cloud infrastructure in data centers is a prerequisite
 - The provisioning for systems and applications on a large number of physical machines is a time-consuming process

Introduction and Inspiration (cont.)

- With low assurance on deployment's time and cost
- Two core services enable the users to get the best out of the IaaS model in public and private cloud setups
 - Virtual machine provisioning and migration services
- Much time was spent to install and provision a new server
 - When installing a new server for a certain workload to provide a service for a client
 - Check the inventory for a new machine
 - Get one, format, install OS required, install services
 - A server is needed along with lots of security patches and appliances

Introduction and Inspiration (cont.)

- With the emergence of virtualization technology and the cloud computing IaaS model
 - To provision a virtual server through a self-service interface with small steps to get what you desire with the required specifications
 - Provisioning this machine in a public cloud like Amazon Elastic Compute Cloud (EC2)
 - Or using a virtualization management software package
 - Or a private cloud management solution installed at your data center
 - Inside the organization and within the private cloud setup

Introduction and Inspiration (cont.)

- The value of virtualization and the way virtual machines are provisioned
- Whenever there was a need for performing a server's upgrade or performing maintenance tasks
 - An expensive operation to maintain or upgrade a main server that has lots of applications and users
 - With the advance of the revolutionized virtualization technology and migration services associated with hypervisors' capabilities
 - These tasks (maintenance, upgrades, patches, etc.) need no time to accomplish

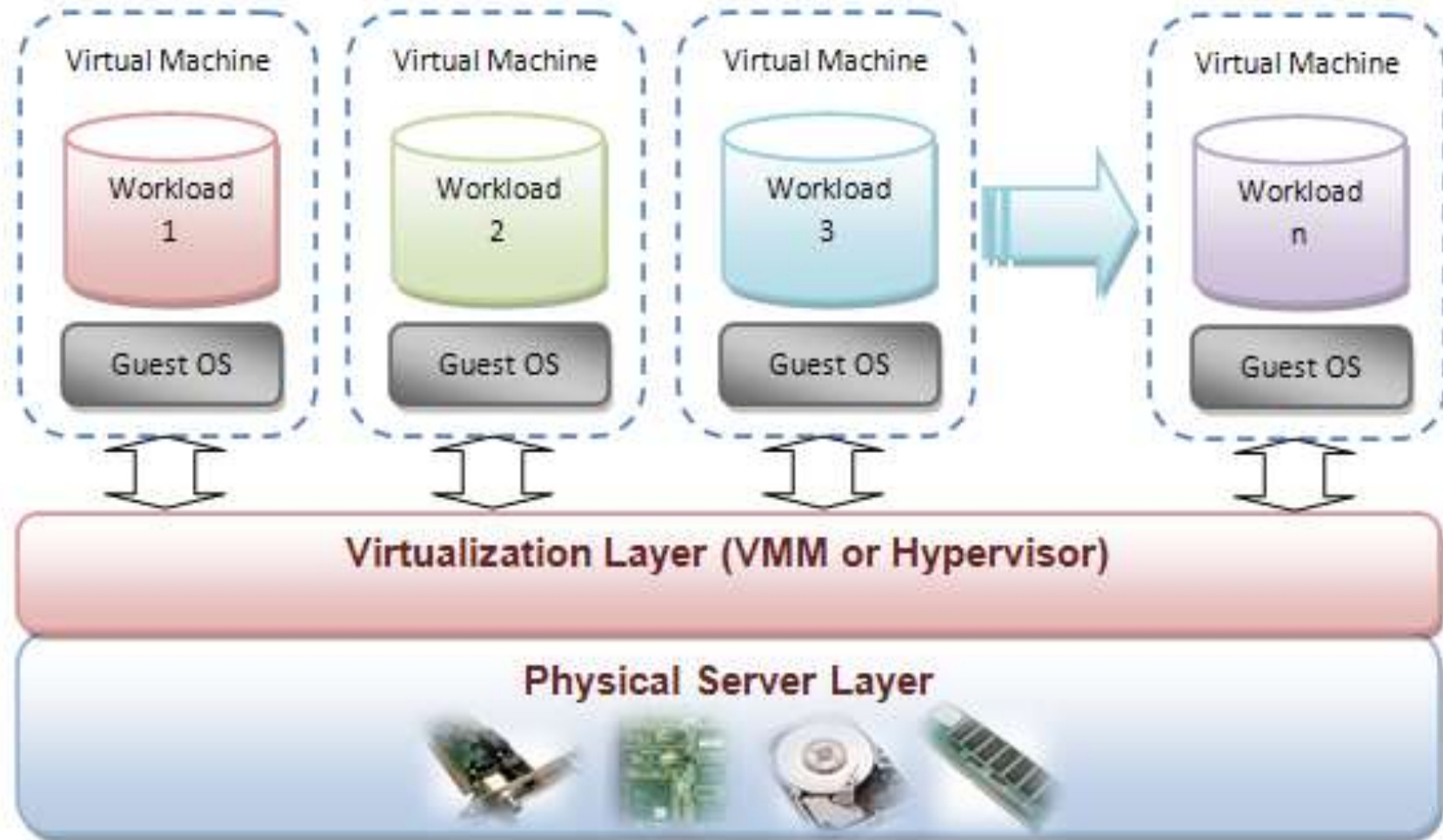
Introduction and Inspiration (cont.)

- Provisioning a new virtual machine is a matter of minutes
- Migrations of a virtual machine is a matter of milliseconds

Virtualization Technology Overview

- Virtualization facilitates the providing and management of the dynamic data center's infrastructure
 - An essential and enabling technology of cloud computing environments
 - The abstraction of the four computing resources
 - Storage, processing power, memory, and network or I/O
 - Conceptually similar to emulation
 - A system pretends to be another system
 - Virtualization is a system pretending to be two or more of the same system

Virtualization Technology Overview (cont.)



Virtualization Technology Overview (cont.)

- A layered virtualization technology architecture
 - The virtualization layer will partition the physical resource of the underlying physical server into multiple virtual machines with different workloads
 - Schedules, allocates the physical resource
 - Makes a virtual machine think it totally owns the whole underlying hardware's physical resource
 - Processor, disks, RAMs, etc.
- Virtual machine's technology manages resources in cloud computing environments

Virtualization Technology Overview (cont.)

- Improves the utilization of such resources by multiplexing many virtual machines on one physical host
 - Server consolidation
- Machines can be scaled up and down on demand with a high level of resources' abstraction
- Enables high, reliable, and agile deployment mechanisms and management of services
 - Providing on-demand cloning and live migration services which improve reliability
- Having an effective management's suite for managing virtual machines' infrastructure is critical for a cloud computing IaaS vendor

Public Cloud and Infrastructure Services

- Public cloud or external cloud
 - Resources are dynamically provisioned via publicly accessible Web applications/Web services (SOAP or RESTful interfaces) from an off-site third-party provider
 - Shares resources and bills on a fine-grained utility computing basis
 - The user pays only for the capacity of the provisioned resources at a particular time
 - Examples for vendors who publicly provide IaaS
 - Amazon Elastic Compute Cloud (EC2) GoGrid, Joyent Accelerator, Rackspace, AppNexus, FlexiScale, and Manjrasoft Aneka

Public Cloud and Infrastructure Services (cont.)

- Amazon Elastic Compute Cloud (EC2) is an IaaS service
 - Provides elastic compute capacity in the cloud
 - Leveraged via Web services (SOAP or REST), a Web-based AWS (Amazon Web Service) management console, or the EC2 command line tools
 - Provides hundreds of pre-made AMIs (Amazon Machine Images) with a variety of operating systems and pre-loaded software
 - i.e., Linux, OpenSolaris, or Windows
 - Provides complete control of computing resources

Public Cloud and Infrastructure Services (cont.)

- Run on Amazon's computing and infrastructure environment easily
- Reduces the time required for obtaining and booting a new server's instances to minutes
- Allows a quick scalable capacity and resources, up and down
 - As the computing requirements change
- Offers different instances' size according to
 - The resources' needs (small, large, and extra large)
 - The high CPU's needs it provides (medium and extra large high CPU instances)
 - High-memory instances (extra large, double extra large, and quadruple extra large instance)

Private Cloud and Infrastructure Services

- A private cloud aims at providing public cloud functionality
 - Maintaining control over an organization's data and resources to meet security and governance's requirements in an organization
 - A highly virtualized cloud data center located inside the organization's firewall
 - Also be a private space dedicated for the company within a cloud vendor's data center
 - Designed to handle the organization's workloads
- Private clouds exhibit the following characteristics:

Private Cloud and Infrastructure Services (cont.)

- Allow service provisioning and compute capability for an organization's users in a self-service manner
- Automate and provide well-managed virtualized environments
- Optimize computing resources, and servers' utilization
- Support specific workloads
- Examples for vendors and frameworks that provide IaaS in private setups
 - Eucalyptus and OpenNebula

Private Cloud and Infrastructure Services (cont.)

- A third type of cloud setup named hybrid cloud
 - A combination of private/internal and external cloud resources
 - Existing together by enabling outsourcing of noncritical services and functions in public cloud and keeping the critical ones internal
 - Releases resources from a public cloud
 - Handles sudden demand usage
 - Called cloud bursting

Distributed Management of Virtualization

- Virtualization needs a powerful management capabilities
 - Many commercial, open source products and research projects are being developed to dynamically provision virtual machines
 - e.g., OpenNebula , IBM Virtualization Manager, Joyent, and VMware DRS
 - Utilizing the physical infrastructure
 - Some commercial and scientific infrastructure cloud computing initiatives provide remote interfaces for controlling and monitoring virtual resources
 - e.g., Globus VWS, Eucalyptus and Amazon

Distributed Management of Virtualization (cont.)

- The RESERVOIR initiative
 - Grid interfaces and protocols enable the required interoperability between the clouds or infrastructure's providers
 - Needs to expand substantially on the current state-of-the-art for grid-wide accounting
 - Also to increase the flexibility of supporting different billing schemes, and accounting for services with indefinite lifetime

High Availability

- A system design protocol and an associated implementation
 - Ensures a certain absolute degree of operational continuity during a given measurement period
- Availability refers to the ability of a user's community to access the system
 - Submitting new work, updating or altering existing work, or collecting the results of the previous work
 - Unavailable: A user cannot access the system
- Services should be available all the time

High Availability (cont.)

- Along with some planned/unplanned downtime according to a certain SLA
 - SLA formalizes the service availability objectives and requirements
 - The monthly availability or downtime of a service
 - To calculate the service's credits to match the billing cycles
- Business critical services are often categorized as high availability services
 - Achieving the lowest possible amount of planned and unplanned downtime
- A virtual environment is the larger part of any organization

High Availability (cont.)

- Management of these virtual resources within this environment becomes a critical mission
- The migration services of these resources became a corner stone
 - In achieving high availability for these services hosted by VMs
- High availability allows virtual machines to automatically be restarted
 - In case of an underlying hardware failure or individual VM failure
 - If one of servers fails, the VMs will be restarted on other virtualized servers in the resource pool
 - Restoring the essential services with minimal service interruption

Cloud and Virtualization Standardization Efforts

- Standardization is important to ensure interoperability
- The prevalent standards that make cloud computing and virtualization possible
 - Distributed Management Task Force (DMTF) have produced standards for almost all the aspects of virtualization technology
 - DMTF initiated the VMAN (Virtualization Management) Initiative
 - Delivers broadly supported interoperability and portability standards for managing the virtual computing lifecycle

Cloud and Virtualization Standardization Efforts (cont.)

- VMAN's OVF (Open Virtualization Format)
 - A collaboration between industry key players
 - Dell, HP, IBM, Microsoft, XenSource, and VMware
- OVF provides a common format to package and securely distribute virtual appliances across multiple virtualization platforms
- VMAN profiles define a consistent way of managing a heterogeneous virtualized environment

OCCI and OGF

- Another standardization effort has been initiated by Open Grid Forum (OGF)
 - To deliver a standard API for cloud IaaS
 - Open Cloud Computing Interface Working Group (OCCI-WG)
 - Dedicated for delivering an API specification for the remote management of cloud computing's infrastructure
 - For allowing the development of interoperable tools for common tasks including deployment, autonomic scaling, and monitoring
 - Covering a high-level functionality required for managing the life-cycle virtual machines/workloads, running on virtualization technologies/containers and supporting service elasticity

OCCI and OGF (cont.)

- The new API for interfacing IaaS cloud computing facilities will allow
 - Consumers to interact with cloud computing infrastructure on an ad hoc basis
 - Integrators to offer advanced management services
 - Aggregators to offer a single common interface to multiple providers
 - Providers to offer a standard interface that is compatible with the available tools
 - Vendors of grids/clouds to offer standard interfaces for dynamically scalable service's delivery in their products

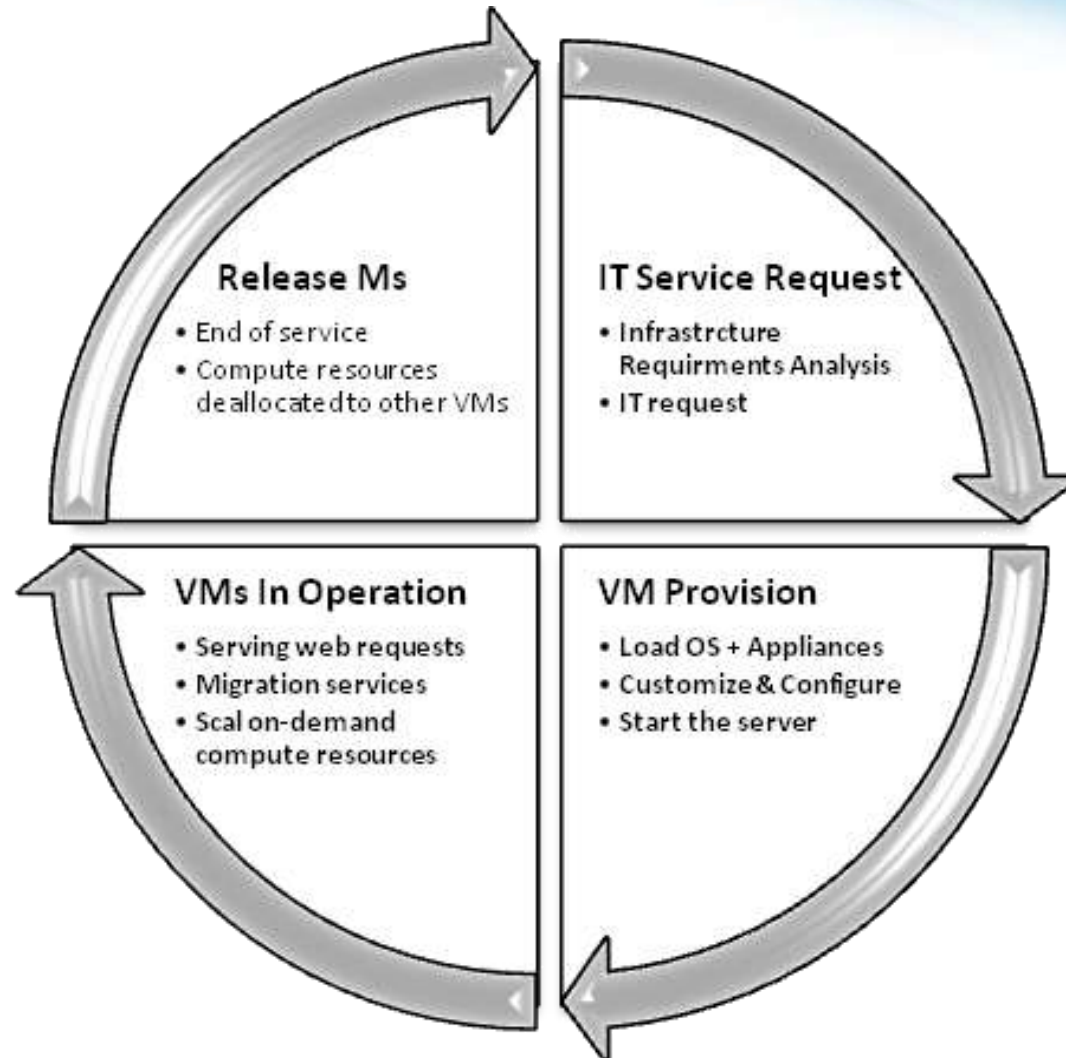
Virtual Machines Provisioning and Manageability

- An overview on the typical life cycle of VM and its major possible states of operation
 - Make the management and automation of VMs in virtual and cloud environments easier
 - Starts by a request delivered to the IT department
 - Stating the requirement for creating a new server for a particular service
 - Processed by the IT administration to start seeing the servers' resource pool
 - Matching these resources with the requirements
 - Starting the provision of the needed virtual machine

Virtual Machines Provisioning and Manageability (cont.)

- Once it is provisioned and started
 - Ready to provide the required service according to an SLA
- A time period after which the VM is being released
 - Free resources

Virtual Machines Provisioning and Manageability (cont.)



VM Provisioning Process

- The common and normal steps of provisioning a virtual server
 - Select a server from a pool of available servers along with the appropriate OS template
 - Physical servers with enough capacity
 - Load the appropriate software
 - Operating system, device drivers, middleware, and the needed applications for the service required
 - Customize and configure the machine to configure an associated network and storage resources
 - e.g., IP address, Gateway

VM Provisioning Process (cont.)

- The virtual server is ready to start with its newly loaded software
- Performed by an IT or a data center's specialist to provision a particular virtual machine
- **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.
- **Virtual machines can be provisioned**
 - By manually installing an operating system

VM Provisioning Process (cont.)



VM Provisioning Process (cont.)

- By using a preconfigured VM template
- By cloning an existing VM
- 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
- A template can be created
 - After creating a virtual machine by virtualizing a physical server or by building a new virtual server in the virtual environment

VM Provisioning Process (cont.)

- Most virtualization management vendors provide the data center's administration with the ability to do such tasks in an easy way
 - e.g., VMware, XenServer, etc.
- Provisioning from a template is an invaluable feature
 - Reduces the time required to create a new virtual machine
- Administrators can create different templates for different purposes
 - Create a Windows 2003 Server template for the finance department

VM Provisioning Process (cont.)

- Or a Red Hat Linux template for the engineering department
- Enables the administrator to quickly provision a correctly configured virtual server on demand
- The problem of virtual machine's sprawl
 - Virtual machines are provisioned so rapidly that documenting and managing the virtual machine's life cycle become a challenge

Virtual Machine Migration Services

- Migration service is the process of moving a virtual machine from one host server or storage location to another
- Different techniques of VM migration
 - Hot/live migration, cold/regular migration, and live storage migration of a virtual machine
- In this process, all key machines' components, are completely virtualized
 - e.g., CPU, storage disks, networking, memory
 - Facilitating the entire state of a virtual machine to be captured by a set of easily moved data files

Management of VM Anatomy of cloud infrastructures

- **Provisioning and Configuration Module:**

It is the lowest level of cloud and typically resides on bare hardware (as a firmware) or on the top of the hypervisor layer. Its function is to abstract the underlying hardware and provide a standard mechanism to spawn instance of virtual machine on demand. It also handles the post-configuration of the operating systems and applications residing on the VM

- **Monitoring and Optimization:**

This layer handles the monitoring of all services, storage, networking and applications components in cloud. Based on the statistics, it could perform routine functions that optimize the behavior of the infrastructure components and provide relevant data to the cloud administrator to further optimize the configuration for maximum utilization and performance,

- **Metering and Chargeback:**

This layer provides functions to measure the usage of resources in cloud. The metering module collects all the utilization data per domain per use. This module gives the cloud administrator enough data to measure ongoing utilization of resources and to create invoices based on the usage on a periodic basis.

- **Orchestration:**

Orchestration is a central to cloud operations. Orchestration converts requests from the service management layer and the monitoring, chargeback modules to appropriate action item which are then submitted to provisioning and configuration module for final closure. Orchestration updates the CMDB in the process.

- **Configuration Management Database (CMDB):**

It is a central configuration repository wherein all the meta data and configuration of different modules, resources are kept and updated in the real-time basis. The repository can then be accessed using standards protocols like SOAP by third-party software and integration components. All updates in CMDB happen in real time as requests get processed in cloud.

- **Cloud Life cycle Management Layer (CLM):**

This layer handles the coordination of all other layers in cloud. All requests internal and external are addressed to the CLM layer first. CLM may internally route requests and actions to other layers for further processing.

- **Service Catalog:**

It is central to the definition of cloud, SC defines what kind of services the cloud is capable of providing and at what cost to the end user. SC is the first thing that is drafted before a cloud is architecture. The service management layer consults SC before it processes any request for a new resource.

Scheduling techniques

- Scheduling techniques help to align the timeline, the scope and your resources. A schedule has to fit a specified timeframe and use available resources with the right skills. Given many uncertainties, variables and a possibility that resource availability or a project scope may change, it's hard to create a schedule that will last

Migrations Techniques

- Live migration
 - Also called hot or real-time migration
 - The movement of a virtual machine from one physical host to another while being powered on
 - Without any noticeable effect from the end user's point of view (a matter of milliseconds)
 - Facilitates proactive maintenance upon failure
 - The potential problem can be resolved before the disruption of service occurs
 - Used for load balancing
 - Work is shared among computers optimize the utilization of available CPU resources

Migrations Techniques (cont.)

- Live migration's mechanism
 - How memory and virtual machine states are being transferred through the network from one host A to another host B
 - e.g., the Xen hypervisor
 - The process has been viewed as a transactional interaction between the two hosts involved
 - Stage 0: Pre-Migration
 - An active virtual machine exists on the physical host A
 - Stage 1: Reservation
 - A request is issued to migrate an OS from host A to B
 - The necessary resources exist on B and on a VM container of that size

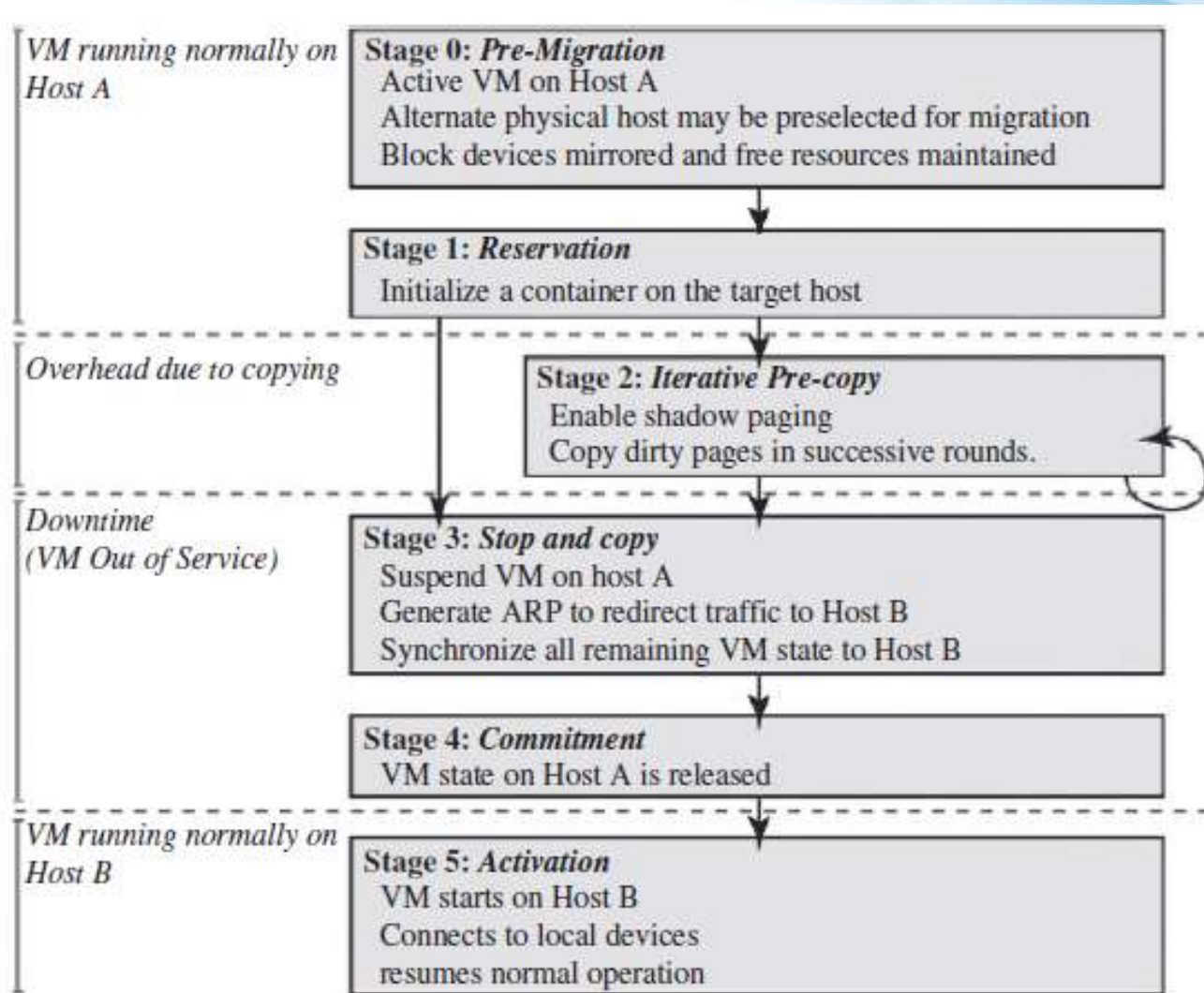
Migrations Techniques (cont.)

- **Stage 2: Iterative Pre-Copy**
 - During the first iteration, all pages are transferred from A to B
 - Iterations copy only those pages dirtied during the previous transfer phase
- **Stage 3: Stop-and-Copy**
 - Running OS instance at A is suspended
 - The network traffic is redirected to B
 - CPU state and any remaining inconsistent memory pages are then transferred
- **Stage 4: Commitment**
 - Host B indicates to A that it has successfully received a consistent OS image

Migrations Techniques (cont.)

- Host A acknowledges this message as a commitment of the migration transaction
- Host A may now discard the original VM
- Host B becomes the primary host
- **Stage 5: Activation**
 - The migrated VM on B is now activated
 - Post-migration code runs to reattach the device's drivers to the new machine and advertise moved IP addresses
- **An approach to failure management**
 - At least one host has a consistent VM image at all times during migration
 - The original host remains stable until the migration commits

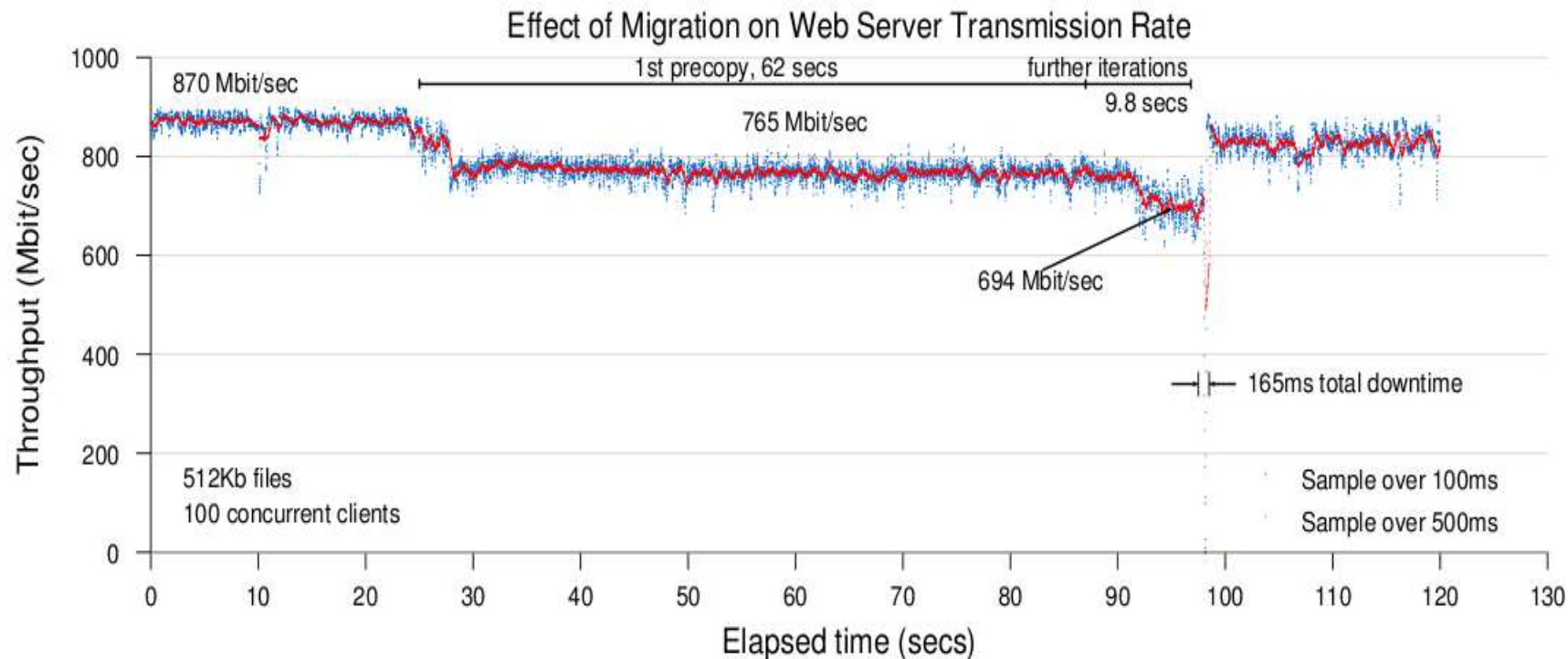
Migrations Techniques (cont.)



Migrations Techniques (cont.)

- The VM may be suspended and resumed on that host with no risk of failure
- A migration request essentially attempts to move the VM to a new host
 - On any sort of failure, execution is resumed locally
 - Aborting the migration
- Live migration effect on a running web server
 - A highly loaded server can be migrated with both controlled impact on live services and a short downtime
 - The working set of the server is rather small
 - A relatively easy case of live migration

Migrations Techniques (cont.)



Migrations Techniques (cont.)

- VM management and provisioning tools
 - Provide the live migration of VM facility
 - e.g., VMware VMotion and Citrix XenServer XenMotion
- VMware Vmotion
 - Allows users to automatically optimize and allocate an entire pool of resources
 - For maximum hardware utilization, flexibility, and availability
 - To perform hardware's maintenance without scheduled downtime
 - Along with migrating virtual machines away from failing or underperforming servers

Migrations Techniques (cont.)

- Citrix XenServer XenMotion
 - Inherited from the Xen live migrate utility
 - Provides the IT administrator with the facility to move a running VM from one XenServer to another in the same pool without interrupting the service
 - Hypothetically for zero-downtime server maintenance
 - Actually takes minutes
 - A highly available service
 - A good feature to balance the workloads on the virtualized environment
- Cold migration
 - The migration of a powered-off virtual machine

Migrations Techniques (cont.)

- The option of moving the associated disks from one data store to another
 - The virtual machines are not required to be on a shared storage
 - Live migration needs a shared storage for virtual machines in the server's pool
- In live migration for a virtual machine between two hosts, there would be certain CPU compatibility checks to be applied
 - In cold migration this checks do not apply
- The cold migration process is simple
 - The configuration files are moved from the source host to the destination host's associated storage area

Migrations Techniques (cont.)

- Including the NVRAM file (BIOS settings), log files, as well as the disks of the virtual machine
- The virtual machine is registered with the new host
- After the migration is completed, the old version of the virtual machine is deleted from the source host
- **Live Storage Migration of Virtual Machine**
 - **Moving the virtual disks or configuration file of a running virtual machine to a new data store**
 - Without any interruption in the availability of the virtual machine's service

VM Migration, SLA and On-Demand Computing

- Virtual machines' migration
 - Making it easy to adjust resource's priorities to match resource's demand conditions
 - Meeting SLAs
 - A particular VM is consuming more than its fair share of resources at the expense of other VMs on the same host
 - Eligible to either be moved to another underutilized host
 - Or to assign more resources for it in case that the host machine still has resources
 - Highly avoid the violations of the SLA
 - Fulfill the requirements of on-demand computing resources

VM Migration, SLA and On-Demand Computing (cont.)

- An integration between virtualization's management tools and SLA's management tools
 - With its migrations and performance's monitoring capabilities
 - To achieve balance in resources by migrating and monitoring the workloads
 - Meeting the SLA

Migration of Virtual Machines to Alternate Platforms

- The ability to migrate virtual machines from one platform to another
 - Numerous ways for achieving this
 - Depending on the source and target virtualization's platforms
 - And on the vendor's tools that manage this facility
 - e.g., The VMware converter handles migrations between ESX hosts, the VMware server, and the VMware workstation
 - Can also import from other virtualization platforms, such as Microsoft virtual server machines



Thank you

The Content in this Material are from the Textbooks and Reference books given in the Syllabus