Infrastructure as a Services

Chapter outline

- IaaS

- Characteristics of IaaS

- Utilization

- Benefits and Futures

- Advantages and disadvantages
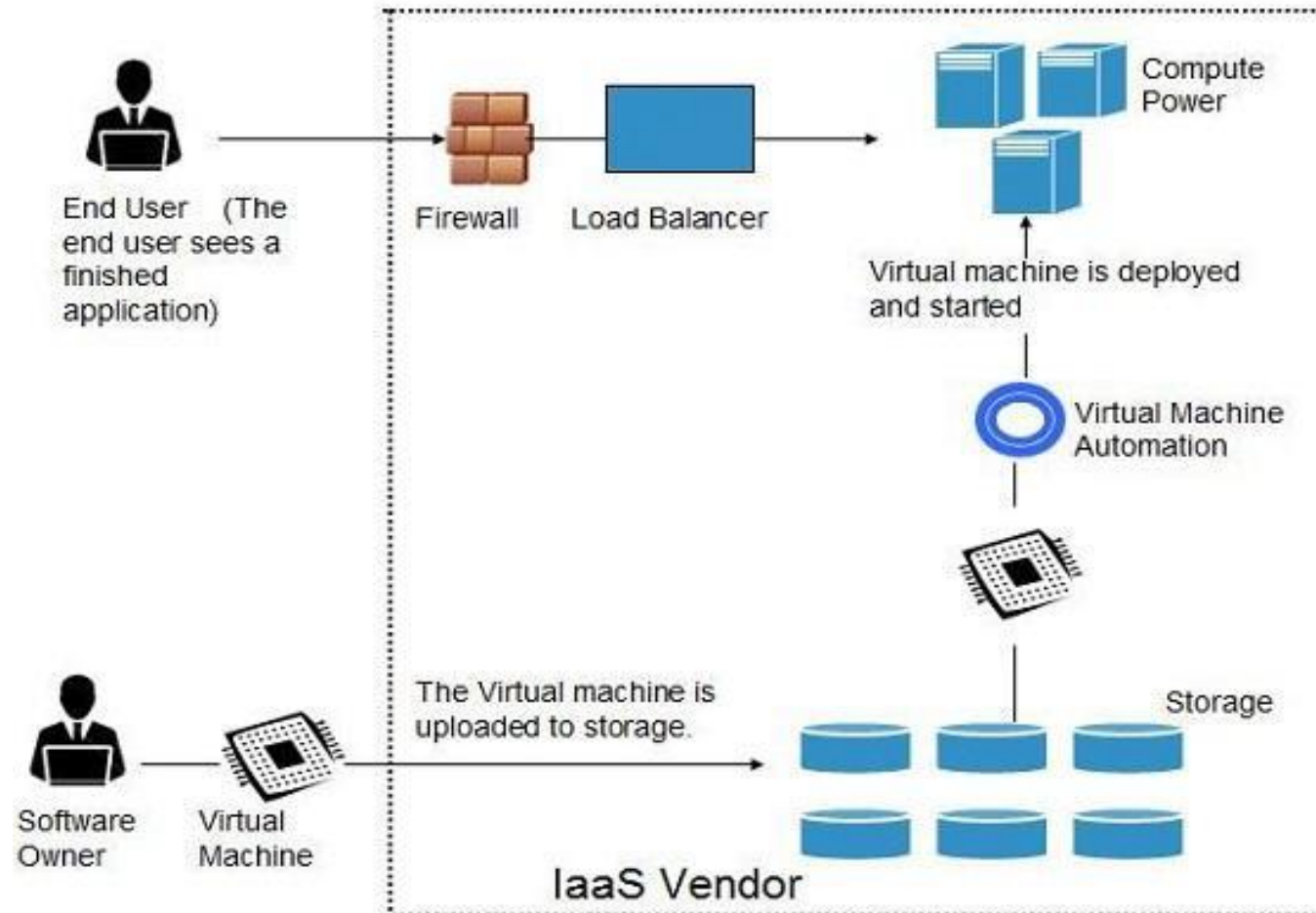
- Virtualization

Infrastructure as a Services

What is IaaS?

- All cloud computing services it provides access to computing resource in a virtualized environment, "the Cloud", across a public connection, usually the internet.

- In the case of IaaS the computing resource provided is specifically that of virtualized hardware, in other words, computing infrastructure

- IaaS provides access to resources such as physical machines, virtual machines, virtual storage.
  - Virtual machine disk storage
  - Virtual local area network (VLANs)
  - Load balancers
  - IP addresses
  - Software bundles

- Resources are made available to end user via server virtualization. Moreover, these resources are accessed by the customers as if they own them.

What is IaaS?

- Physically, the pool of hardware resource is pulled from a multitude of servers and networks usually distributed across numerous data centers, all of which the cloud provider is responsible for maintaining.

- The client, on the other hand, is given access to the virtualized components in order to build their own IT platforms.

What is IaaS?

End User (The end user sees a finished application)

Firewall

Load Balancer

Compute Power

Virtual machine is deployed and started

Virtual Machine Automation

Software Owner

Virtual Machine

The Virtual machine is uploaded to storage.

Storage

IaaS Vendor

Basic IaaS Characteristics

- Dynamic Scaling:

- Service Levels

- Rental Model

- Licensing

- Metering and Costs

- Self-Service Provisioning

Basic IaaS Characteristics

- **Dynamic Scaling**

    – Some level of uncertainty always exists when planning for IT resources.

    – One of the major benefits of IaaS for companies faced with this type of uncertainty is the fact that resources can be automatically scaled up or down based on the requirements of the application.

    – This important characteristic of IaaS is called dynamic scaling — if customers wind up needing more resources than expected, they can get them immediately (probably up to a given limit).

    – A provider or creator of IaaS typically optimizes the environment so that the hardware, the operating system, and automation can support a huge number of workloads.

Basic IaaS Characteristics

- **Service Levels**

  - Consumers acquire IaaS services in different ways. Many consumers rent capacity based on an on-demand model with no contract.
  - In other situations, the consumer signs a contract for a specific amount of storage or compute.
  - A typical IaaS contract has some level of service guarantee.
  - At the low end, a provider might simply state that the company will do its best to provide good service. If the consumers are willing to pay a premium price, they might get a mirrored service so that there are almost no change-of-service interruptions.
  - A typical service-level agreement states what the provider has agreed to deliver in terms of availability and response to demand.
  - The service level might, for example, specify that the resources will be available 99.999% of the time and that more resources will be provided dynamically if greater than 80% of any given resource is being used.

Basic IaaS Characteristics

- **The Rental Model**
  - When companies use IaaS, it's often said that the servers, storage, or other IT infrastructure components are rented for a fee based on the quantity of resources used and how long they're in use. Although this is true, there are some important differences between this rental arrangement and the traditional rental models you may be familiar with.
  - For example, when you purchase server and storage resources using IaaS services, you gain immediate virtual access to the resources you need. You aren't, however, renting the actual physical servers or other infrastructure. Don't expect a big truck to pull up to your office and deliver the servers you need to complete your project. The physical components stay put in the infrastructure service provider's data center. This concept of renting is an essential element of cloud computing, and it provides the foundation for the cost and scalability benefits of the various cloud models.
  - Within a private IaaS model, renting takes on a different focus. Although you might not charge each user to access a resource, in the charge-back model, you can allocate usage fees to an individual department based on usage over a week, month, or year. Because of the flexibility of the IaaS model, you can charge more of the budget to heavy users.

Basic IaaS Characteristics

- **Licensing**
  - The use of public IaaS has led to innovation in licensing and payment models for software you want to run in these cloud environments.
  - Note that this licensing is for the software you want to run in your cloud environment, not the license between you and the cloud provider.
  - For example, some IaaS and software providers have created a "bring your own license" (BYOL) plan so you have a way to use your software license in both traditional and cloud environments.
  - Another option is called "pay as you go" (PAYG), which generally integrates the software licenses with the on-demand infrastructure services.
  - For example, say that you're running Microsoft Windows Server and using the PAYG route. If you're paying ten cents an hour for cloud access, a few cents of that fee might be going to Microsoft.

Basic IaaS Characteristics

- **Metering and Costs**
    - Clearly, you derive a potential economic benefit by controlling the amount of resources you demand and pay for so that you have just the right match with your requirements. To ensure that users are charged for the resources they request and use, IaaS providers need a consistent and predictable way to measure usage. This process is called metering.
    - Ideally, the IaaS provider will have a transparent process for identifying charges incurred by the user. With multiple users accessing resources from the same environment, the IaaS provider needs an accurate method for measuring the physical use of resources to make sure each customer is charged the right amount.
    - IaaS providers often use the metering process to charge users based on the instance of computing consumed. An instance is defined as the CPU power and the memory and storage space consumed in an hour. When an instance is initiated, hourly charges begin to accumulate until the instance is terminated. The charge for a very small instance may be as little as two cents an hour; the hourly fee could increase to $2.60 for a large resource-intensive instance running Windows.

Basic IaaS Characteristics

- **Self-Service Provisioning**
  - You can't discuss the key characteristics of IaaS without understanding the imperative of self-service.
  - The banking ATM service is a great example of the business value of self service. Without the availability of the self-service ATM, banks would be required to use costly resources to manage activities of all their customers — even for the most repetitive tasks.
  - With an ATM, repetitive tasks can be handled easily with a self-service interface. The customer makes a direct request to perform routine transactions that conform to predefined business rules.

How IaaS can be Utilized?

The following are salient examples of how IaaS can be utilized by enterprise:

- **Enterprise infrastructure:**
  - by internal business networks, such as private clouds and virtual local area networks, which utilize pooled server and networking resources and in which a business can store their data and run the applications they need to operate day-to-day.
  - Expanding businesses can scale their infrastructure in accordance with their growth whilst private clouds (accessible only by the business itself) can protect the storage and transfer of the sensitive data that some businesses are required to handle.

- **Cloud hosting:**
  - The hosting of websites on virtual servers which are founded upon pooled resources from underlying physical servers. A website hosted in the cloud, for example, can benefit from the redundancy provided by a vast network of physical servers and on demand scalability to deal with unexpected demands placed on the website.

How IaaS can be Utilized?

- **Virtual Data Centers (VDC):** a virtualized network of interconnected virtual servers which can be used to offer enhanced cloud hosting capabilities, enterprise IT infrastructure or to integrate all of these operations within either a private or public cloud implementation.

Features and Benefits

- **Scalability:** resource is available as and when the client needs it and, therefore, there are no delays in expanding capacity or the wastage of unused capacity

- **No investment in hardware:** the underlying physical hardware that supports an IaaS service is set up and maintained by the cloud provider, saving the time and cost of doing so on the client side

- **Utility style costing:** the service can be accessed on demand and the client only pays for the resource that they actually use

- **Location independence:** the service can usually be accessed from any location as long as there is an internet connection and the security protocol of the cloud allows it

- **Physical security of data center locations:** services available through a public cloud, or private clouds hosted externally with the cloud provider, benefit from the physical security afforded to the servers which are hosted within a data canter

- **No single point of failure:** if one server or network switch, for example, were to fail, the broader service would be unaffected due to the remaining multitude of hardware resources and redundancy configurations. For many services if one entire data center were to go offline, never mind one server, the IaaS service could still run successfully.

Public Cloud and IaaS

- In Public cloud resources are dynamically provisioned via publicity accessible Web Application/ Web Services(SOAP or RESTful interfaces) from an off-side third party providers who 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.

- There are many examples for vendors who publicly provide infrastructure as
  - Amazon Elastic Compute Cloud (EC2)
  - GoGrid
  - Joyent
  - Accelerator
  - Rackspace AppNexus FlexiScale
  - Manjrasoft
  - Aneka

Public Cloud and IaaS

- Amazon EC2 reduces the time required for obtaining and booting a new server's instances to minutes, thereby allowing a quick scalable capacity and resources, up and down, as the computing requirements change.

- Amazon offers different instances' size according to (a) the resources' needs (small, large, and extra large), (b) the high CPU's needs it provides (medium and extra large high CPU instances), and (c) high-memory instances (extra large, double extra large, and quadruple extra large instance).

Private Cloud and IaaS

- A private cloud aims at providing public cloud functionality, but on private resources, while maintaining control over an organization's data and resources to meet security and governance's requirements in an organization. Private cloud exhibits a highly virtualized cloud data center located inside your organization's firewall.

- It may also be a private space dedicated for your company within a cloud vendor's data center designed to handle the organization's workloads.

- Private clouds exhibit the following characteristics:
  - 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.

Private Cloud and IaaS

- There are many examples for vendors and frameworks that provide infrastructure as a service in private setups. The best-known examples are Eucalyptus and OpenNebula.

- It is also important to highlight a third type of cloud setup named "hybrid cloud," in which a combination of private/internal and external cloud resources.

- Hybrid cloud's main function is to release resources from a public cloud and to handle sudden demand usage, which is called "*cloud bursting*."

Virtualization Technology
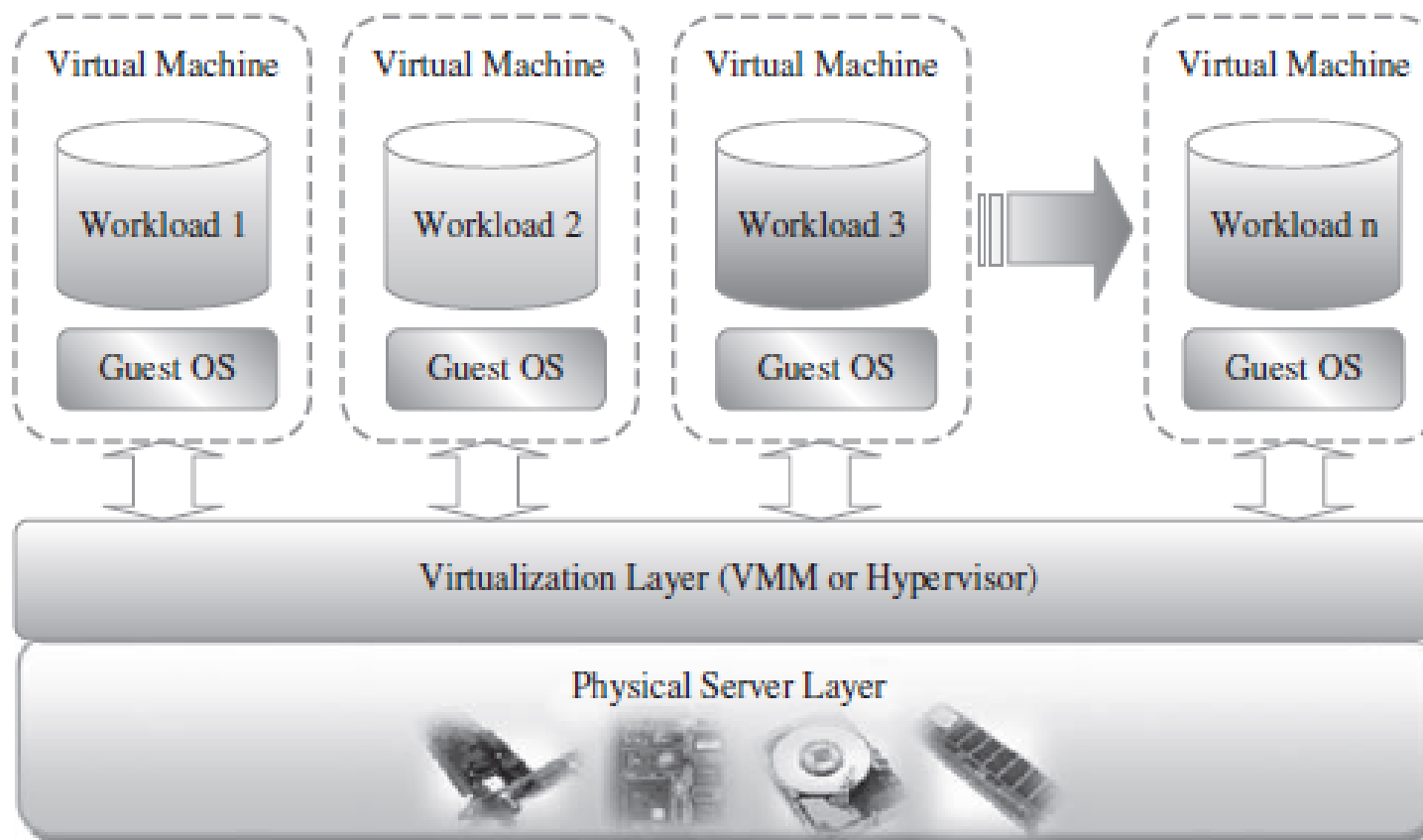
A layered virtualization technology architecture

VM Provisioning and Manageability

- VM and its major possible states of operation, which make the management and automation of VMs in virtual and cloud environments easier than in traditional computing environments.

- VM cycle starts by a request delivered to the IT department, stating the requirement for creating a new server for a particular service.

- This request is being processed by the IT administration to start seeing the servers' resource pool, matching these resources with the requirements, and starting the provision of the needed virtual machine.

- Once it is provisioned and started, it is ready to provide the required service according to an SLA, or a time period after which the virtual is being released; and free resources, in this case, won't be needed.

VM Provisioning and Manageability
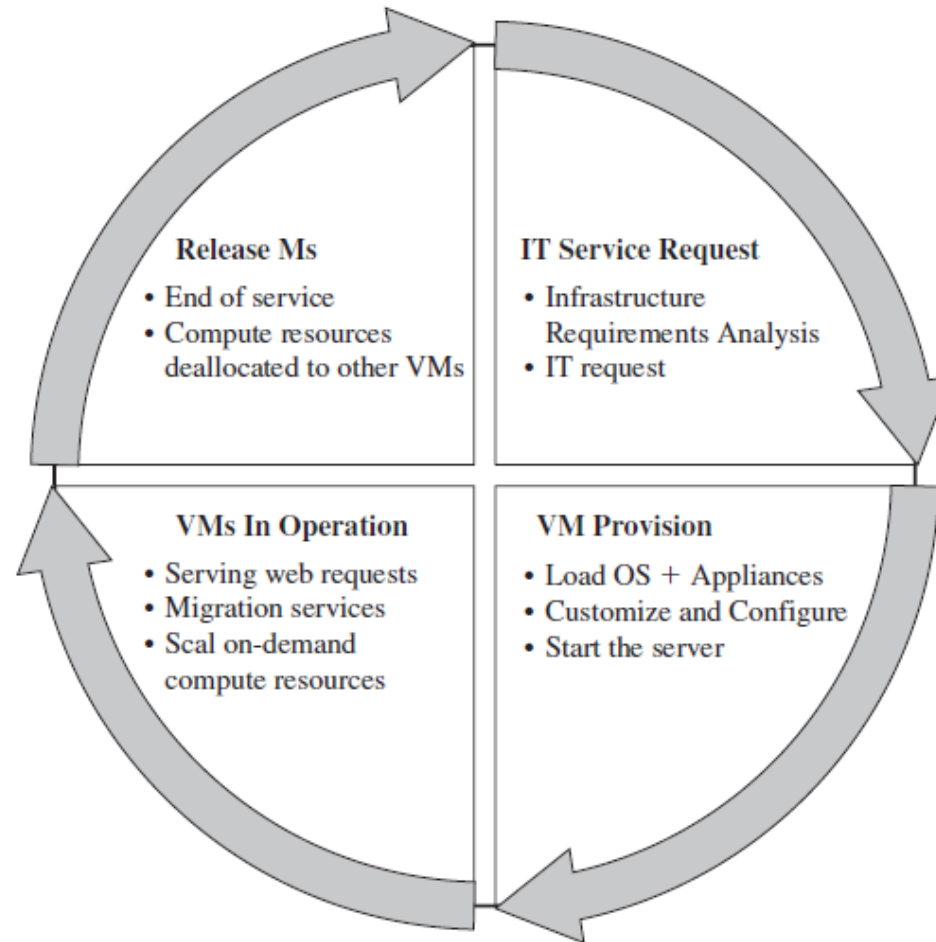
Virtual Machine Life Cycle

Virtual Machine Provisioning Process

- Firstly, you need to select a server from a pool of available servers (physical servers with enough capacity) along with the appropriate OS template you need to provision the virtual machine.

- Secondly, you need to load the appropriate software (operating system you selected in the previous step, device drivers, middleware, and the needed applications for the service required).

- Thirdly, you need to customize and configure the machine (e.g., IP address, Gateway) to configure an associated network and storage resources.

- Finally, the virtual server is ready to start with its newly loaded software.

Servers Pool

Running Provisioned VM

Load OS and Appliances

Customize and Configure

Install Patches

Start the Server

Appliances Repository

Virtual Machine Migration Service

- **VM Migration service:** in the context of virtual machines, is the process of moving a virtual machine from one host server or storage location to another.

- VM migration techniques
  - Hot/life migration
  - Cold/regular migration
  - Live storage migration of a virtual machine

- All key machines' components, such as CPU, storage disks, networking, and memory, are completely virtualized, thereby facilitating the entire state of a virtual machine to be captured by a set of easily moved data files.
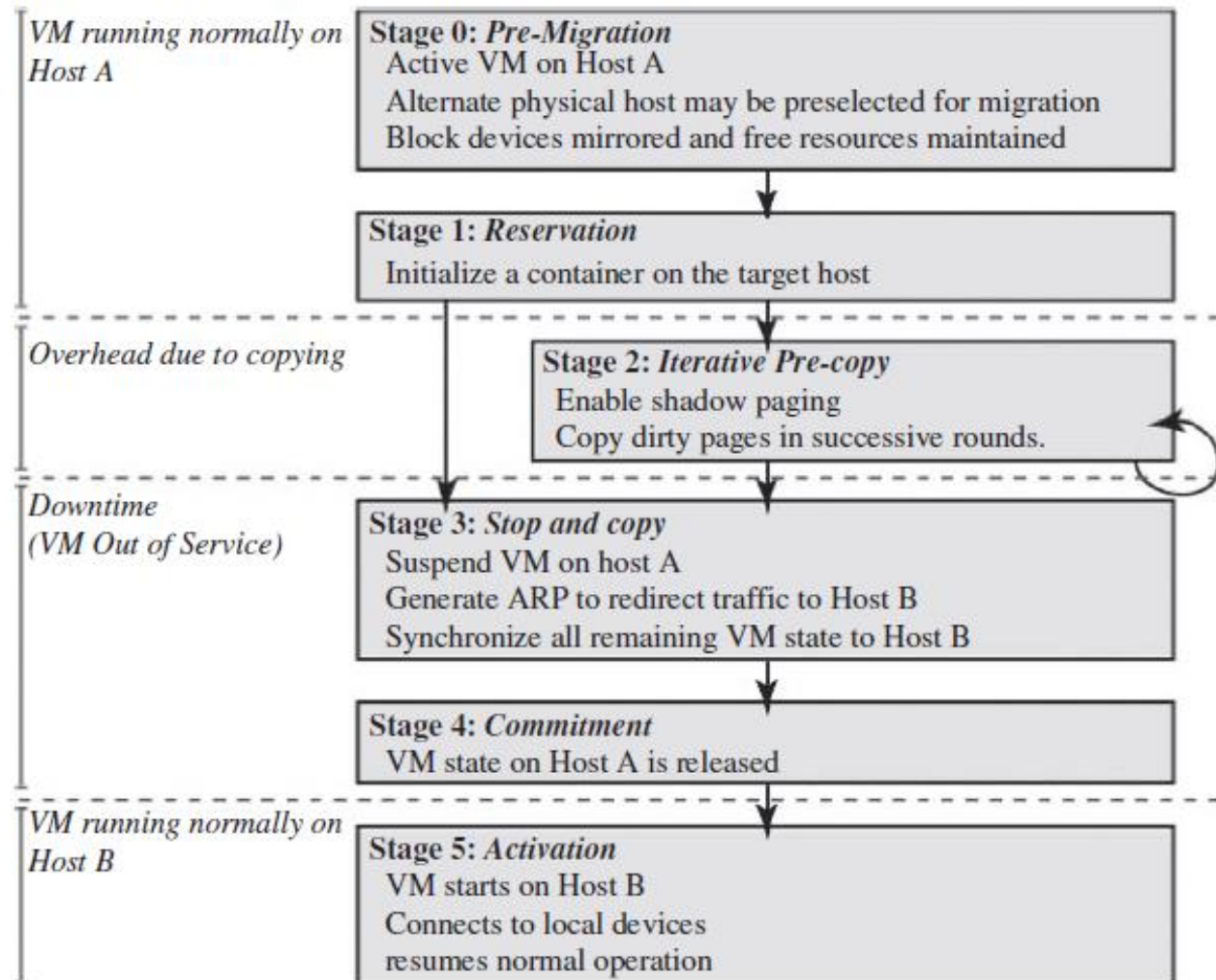
VM Migrations Techniques

Live Migration and High Availability(Hot and real-time migration):

- It can be defined as the movement of a virtual machine from one physical host to another while being powered on

- When it is properly carried out, this process takes place without any noticeable effect from the end user's point of view (a matter of milliseconds).

- One of the most significant advantages of live migration is the fact that it facilitates proactive maintenance in case of failure, because the potential problem can be resolved before the disruption of service occurs.

- Live migration can also be used for load balancing in which work is shared among computers in order to optimize the utilization of available CPU resources

- Live Migration Vendor Implementations Examples. There are lots of VM management and provisioning tools that provide the live migration of VM facility, two of which are VMware VMotion and Citrix XenServer "XenMotion."

VM Migrations Techniques

Live Migration Anatomy, Xen Hypervisor

VM running normally on Host A

Stage 0: Pre-Migration
Active VM on Host A
Alternate physical host may be preselected for migration
Block devices mirrored and free resources maintained

Stage 1: Reservation
Initialize a container on the target host

Overhead due to copying

Stage 2: Iterative Pre-copy
Enable shadow paging
Copy dirty pages in successive rounds.

Downtime (VM Out of Service)

Stage 3: Stop and copy
Suspend VM on host A
Generate ARP to redirect traffic to Host B
Synchronize all remaining VM state to Host B

Stage 4: Commitment
VM state on Host A is released

VM running normally on Host B

Stage 5: Activation
VM starts on Host B
Connects to local devices
resumes normal operation

27

VM Migrations Techniques

Regular/ Cold migration:

- Cold migration is the migration of a powered-off virtual machine. With cold migration, you have the option of moving the associated disks from one data store to another

- 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, but cold migration does not; also, in live migration for a virtual machine between two hosts, there would be certain CPU compatibility checks to be applied; while in cold migration this checks do not apply. The cold migration process is simple to implement (as the case for the VMware product).

- The configuration files, including the NVRAM file (BIOS settings), log files, as well as the disks of the virtual machine, are moved from the source host to the destination host's associated storage area.

- 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.

VM Migrations Techniques

Live Storage Migration:

- This kind of migration constitutes 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

Example

- DigitalOcean

- Linode

- Rackspace

- Amazon Web Services (AWS) EC2 and S3

- Cisco Metapod

- Microsoft Azure

- Google Compute Engine (GCE) are some popular examples of Iaas.

- Aneka

- Open Nebula

- Eucalyptus

Infrastructure as a Service (IaaS) Advantages

- Various pricing models may allow paying only for what you use. This, for example, can allow an individual or a small organization to use sophisticated development software that they could not afford if it was installed on an internal, dedicated server.

- Some IaaS Providers provide development options for multiple platforms: mobile, browser, and so on. If you or your organization want to develop software that can be accessed from multiple platforms, this might be an easy way to make that happen.

- If you have events such as high seasonal sales activity, then the <u>elasticity of the Cloud</u> with IaaS might provide an opportunity.

- The IaaS Cloud Provider may provide better security than your existing software (security—or inadequate security—can also be a disadvantage). Better security may come in part because it is critical for the IaaS Cloud Provider and is part of their main business. In-house security, on the other hand, is not usually an individual's or a organization's main business and, therefore, may not be as good as that offered by the IaaS Cloud Provider.

Infrastructure as a Service (IaaS) Advantages

- No need to manage the introduction of new releases of the development or underlying software. This is handled by the IaaS Cloud Provider.

- No need to manage the underlying data center. This is handled by the IaaS Cloud Provider.

- Usually, there is no need to manage backups. This is handled by the IaaS Cloud Provider.

- If the IaaS Cloud Provider supports failover should the software (for example, the database management software) or the data center become unavailable, that failover is a concern of the IaaS Cloud Provider and you do not need to plan for it.

Infrastructure as a Service (IaaS) Disadvantages

- When it is mandatory that the underlying hardware be of a specific type or the underlying software be modified to support the deployed application.

- There may be legal reasons that preclude the use of off-premise or out-of-counry data storage.

- Security features of the IaaS Cloud Provider may not adequate for your needs.

- If you have a need for high-speed interaction between your internal software or software in another Cloud and the IaaS Cloud Provider, relying on an Internet connection may not provide the speed that you need.