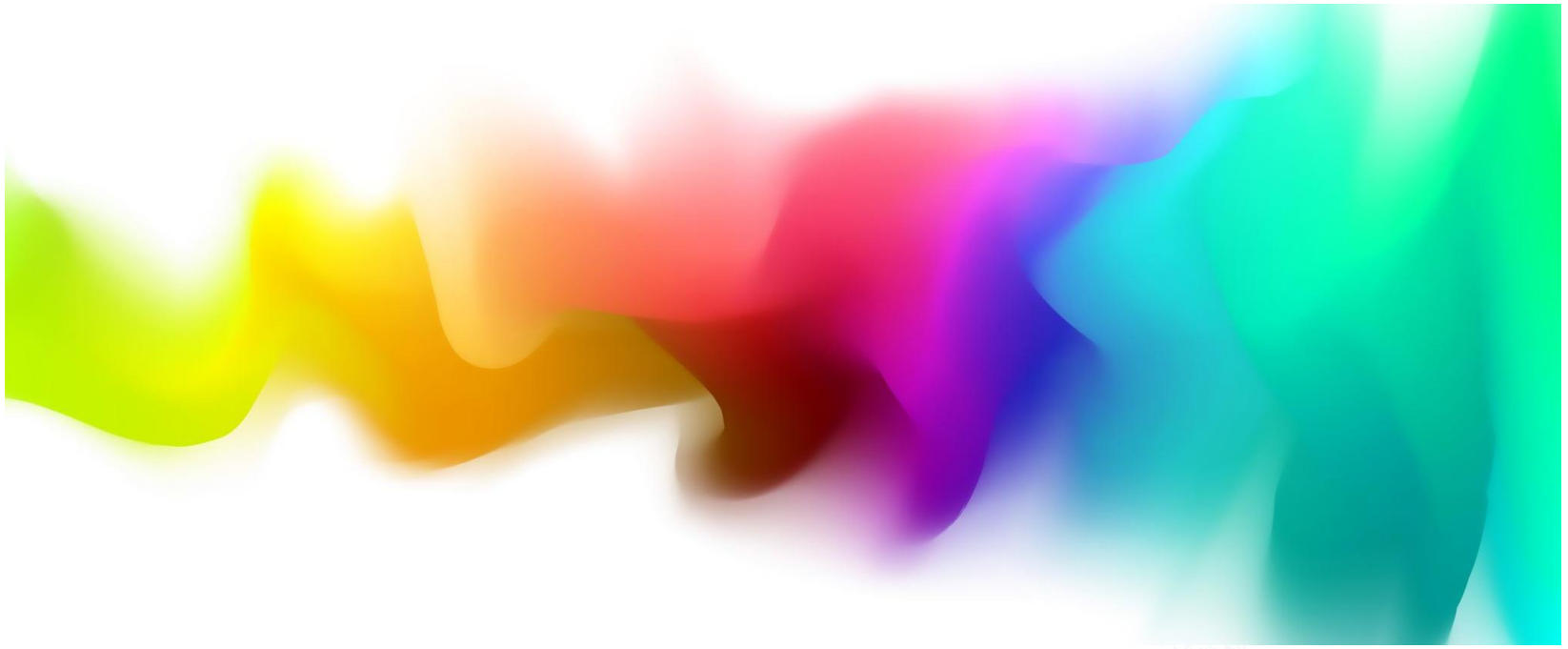
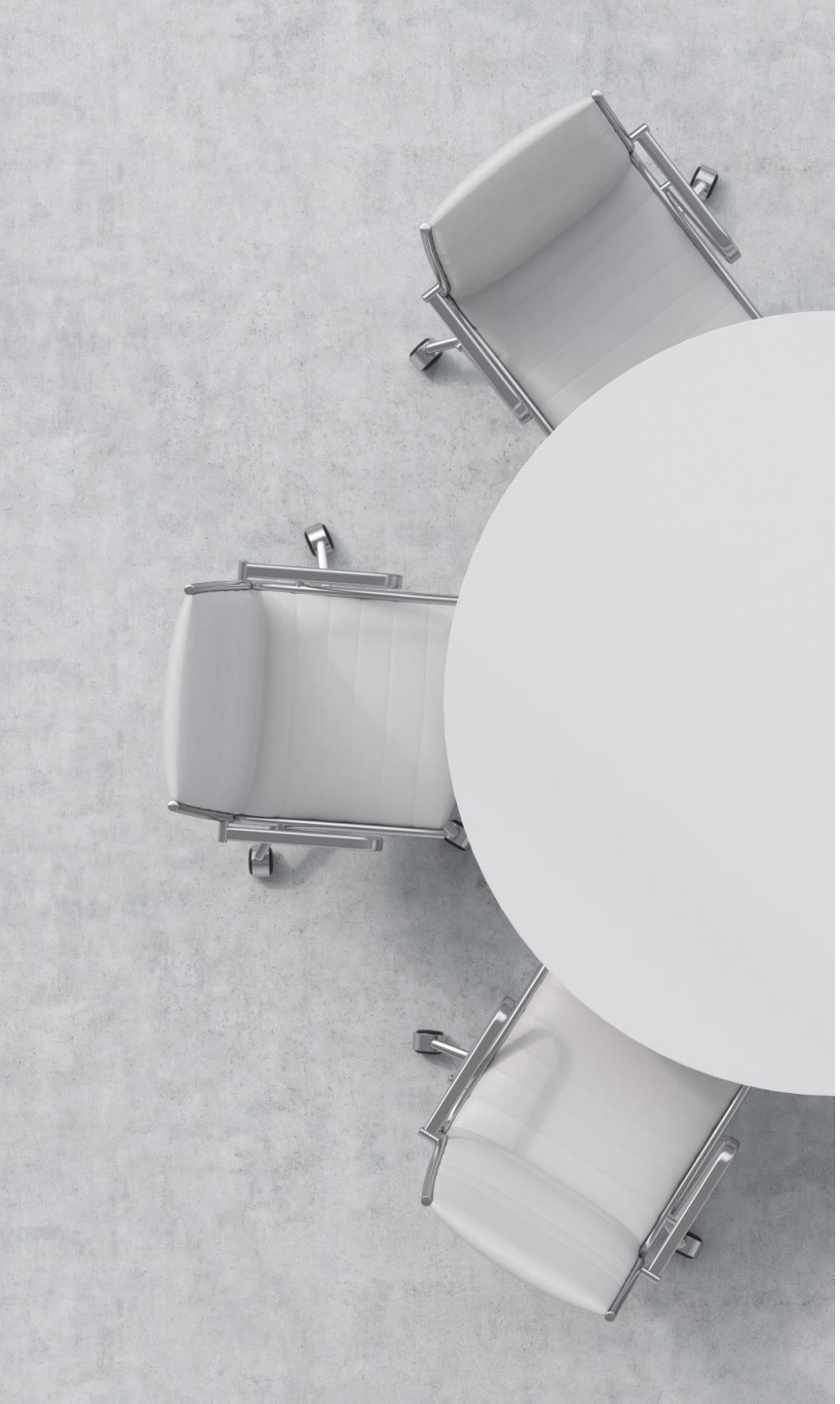


Cloud Computing

Dr. Mostafa Ammar

grading

- 7th week 30 (20+10 section & lab)
- 12th week 20 (exam/project)
- Attendance 10
- Final 40



Zoom meeting

- Join Zoom Meeting
- https://zoom.us/j/4510408387

WhatsApp group

https://chat.whatsapp.com/JasQXD5wVtzB2zM0vx26Yw

Google Classroom Link

- https://classroom.google.com/c/Mjg1NzQ1NzQ1MDM3?cjc=w2tp5q
  uClass code: w2tp5qu

Moodle 2021

- How to find your lectures ??
- Example from another course

Textbook

- Cloud Infrastructure and Services Version 2 Student Guide (EMC)
- https://drive.google.com/file/d/1TR_6pYdj_9DeuZluThUDUKuVQ86yUuU1/view?usp=sharing

Course syllabus

- Introduction to cloud computing
- Software as a service, platform as a service
- Infrastructure as a service
- Virtualization
- How data centers are built
- Physical layer in cloud
- Types Storage in cloud
- Server deployment in cloud
- Introduction to machine learning in cloud

Bonus

- Term questions and answers during lectures.

Benefits of mail students

- Access to moodle
- Access to wifi students in all AAST branches
- Office 365
- One drive 1Tera
- Microsoft teams
- Free online programs (jet brain, prezi, Amazon prime, Microsoft Azure,….)
- Create email from students portal

Attending lectures or sections

- Use your laptop not your mobile
- Follow lecture
- Install software on your laptop
- Ask if anything is not clear

contacts

- Email: mostafaammar@aast.edu
- mostafaammar79@gmail.com

Module: Introduction to Cloud Computing

Upon completion of this module, you should be able to:

- Define cloud computing
- Describe the essential cloud characteristics
- Discuss the key benefits of cloud computing
- Describe the cloud service models
- Describe cloud services brokerage
- Describe the cloud deployment models

Lesson: Cloud Computing Overview

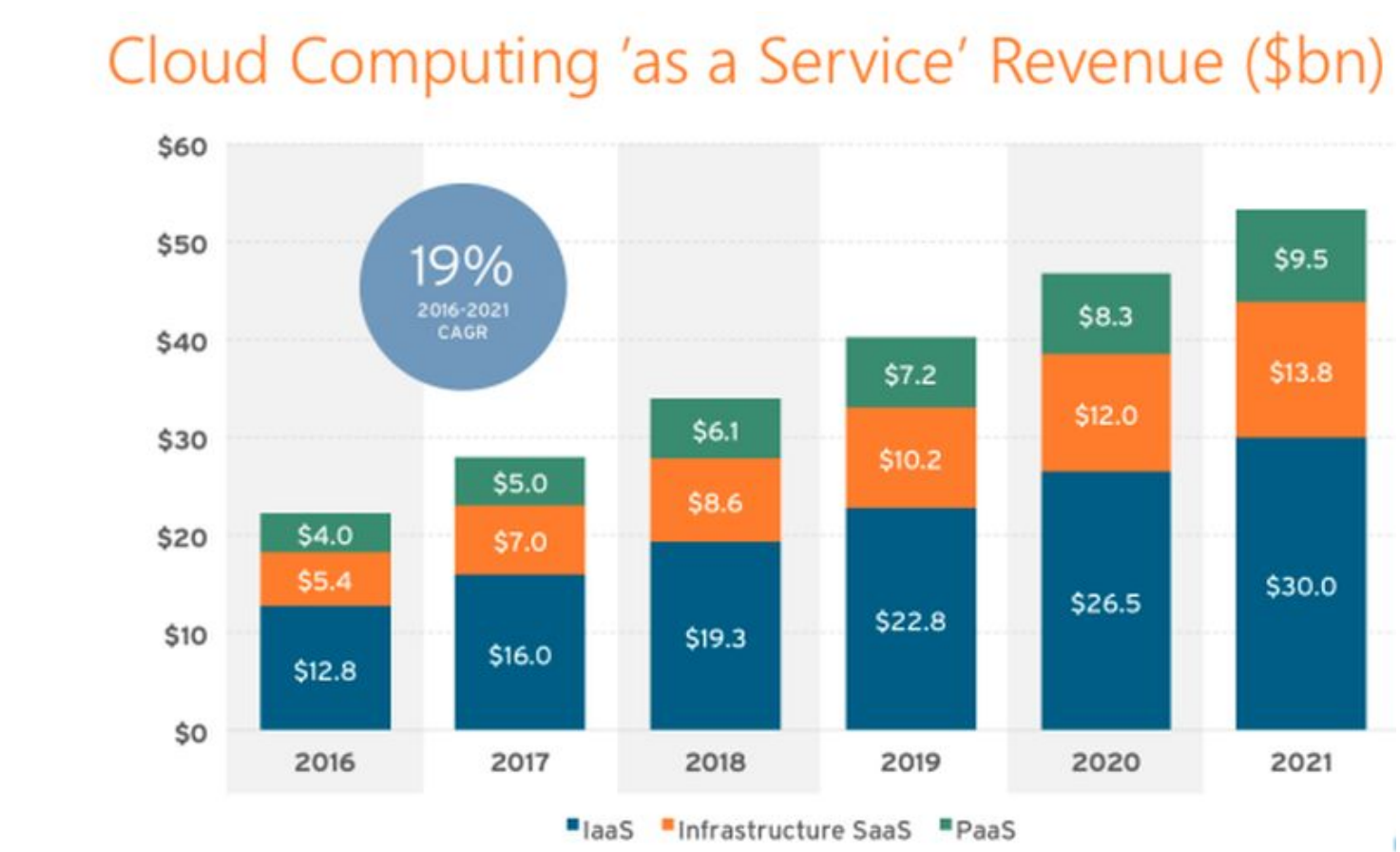
This lesson covers the following topics:

- Definition of cloud computing
- Essential characteristics of cloud computing
- Key benefits of cloud computing

The Cloud Computing Phenomenon

- Adoption of cloud computing is significantly rising in organizations
- Cloud computing is seen as a leading "disruptive" technology in the coming decade
- Cloud is driving optimization and innovation of business models in organizations
- Trends like mobility, Big Data, and social media are also influencing cloud adoption

Cloud computing adoption statistics



What is Cloud Computing?

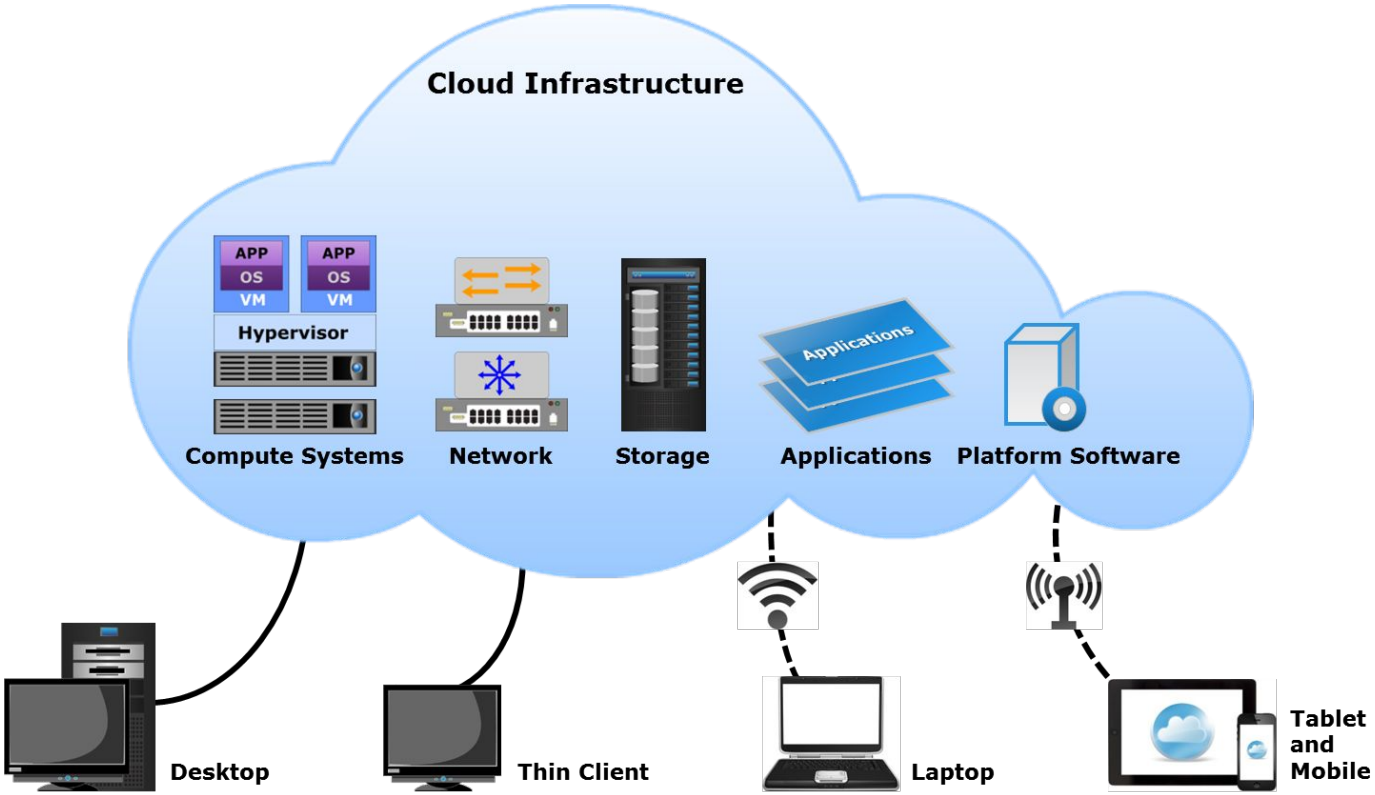
Cloud Computing

A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources, (e.g., servers, storage, networks, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.

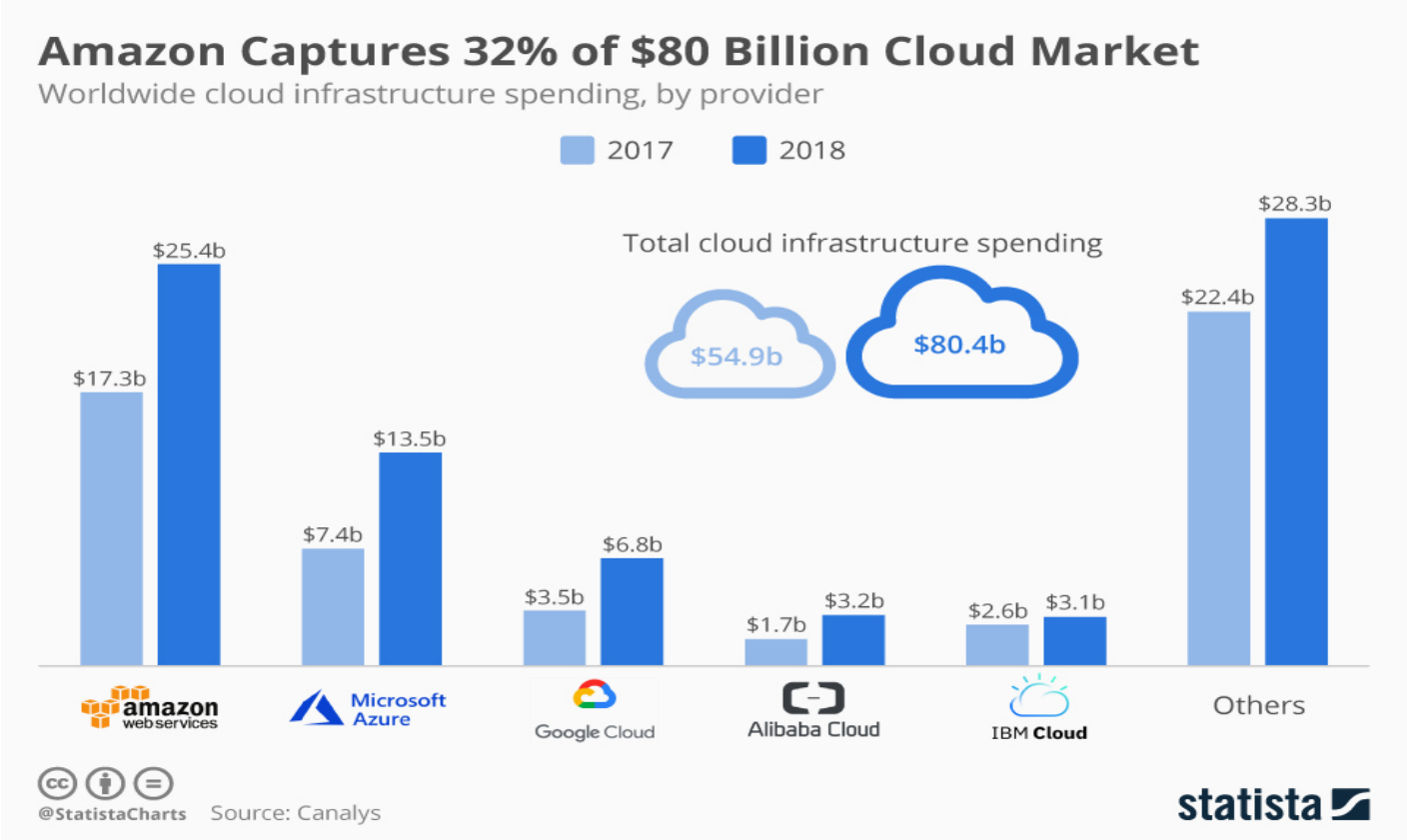
– U.S. National Institute of Standards and Technology, Special Publication 800-145

- A cloud is a collection of network-accessible IT resources
  - Consists of shared pools of hardware and software resources deployed in data centers
- Cloud model enables consumers to hire a provider's IT resources as a service

What is Cloud Computing? (Cont'd)



Cloud key players



Gartner Report

Cloud adoption statistics

- The public cloud service market is expected to reach **$623.3 billion by 2023 worldwide.**
- **83% of enterprise** workloads will be in the cloud by 2020.
- **94% of enterprises already use** a cloud service.
- **30% of all IT budgets** are allocated to cloud computing.
- **66% of enterprises** already have a central cloud team or a cloud center of excellence.
- The cloud is *already* a big deal and it's only going to keep growing for *any* foreseeable future.

AWS regions distribution worldwide

Global Network of AWS Regions

The AWS Cloud spans 76 Availability Zones within 24 geographic regions around the world, with announced plans for nine more Availability Zones and three more AWS Regions in Indonesia, Japan, and Spain.

AWS Global Infrastructure

- 24 Geographic Regions.
- 76 Availability Zones
- 3 upcoming regions
- 9 additional AZs
- 245 Countries /Territories
- 216 Points of Presence

https://aws.amazon.com/about-aws/global-infrastructure/

Cloud main concept

Essential Cloud Characteristics

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

On-demand Self-service

On-demand Self-service

A consumer can unilaterally provision computing capabilities, such as server time and network storage, as needed automatically without requiring human interaction with each service provider.

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Consumers use a web-based self-service portal to view a service catalog and request cloud services
- Enables consumers to provision cloud services in a simple and flexible manner
  - Reduces the time needed to provision new or additional IT resources

Broad Network Access

Broad Network Access

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms, (e.g., mobile phones, tablets, laptops, and workstations).

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Consumers access cloud services on any client/end-point device from anywhere over a network
- Standard mechanisms support the use of heterogeneous client platforms
  - OSI and TCP/IP protocols
  - SOAP and REST web services

Broad Network Access

Broad Network Access

Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms, (e.g., mobile phones, tablets, laptops, and workstations).

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Consumers access cloud services on any client/end-point device from anywhere over a network
- Standard mechanisms support the use of heterogeneous client platforms
  - OSI and TCP/IP protocols
  - SOAP and REST web services

Resource Pooling

Resource Pooling

The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth.

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Enables providers to improve resource utilization and to flexibly provision and reclaim resources

Datacenter (capital expense)

Cloud Computing – The IT Game Changer

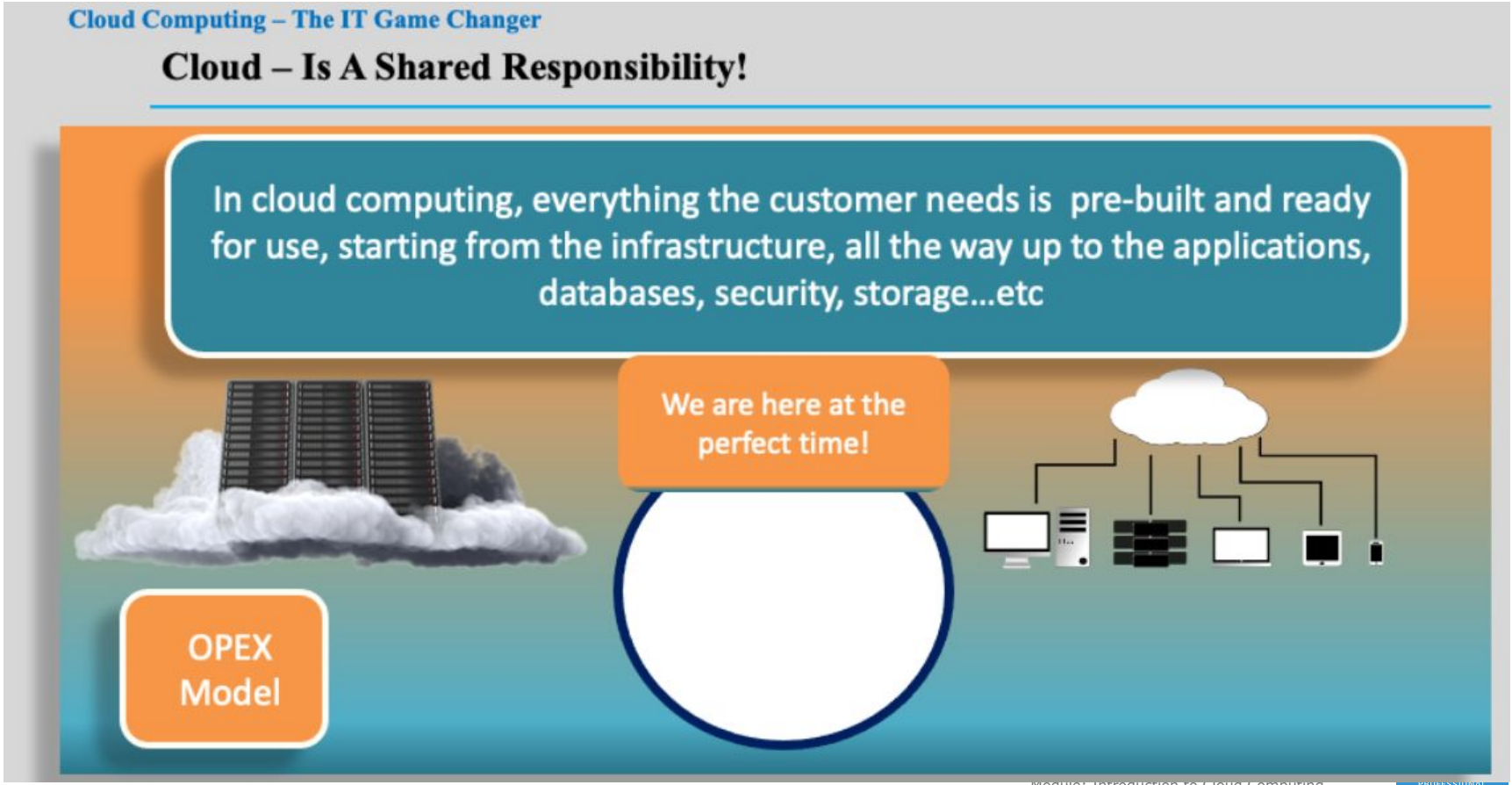
On-Premises / Data Center – All Yours to Build and Operate!

CAPEX Model

- App
- Data
- Runtime
- Middleware
- Operating System
- Virtualization
- Hardware
- Storage
- Network

Customer is 100% responsible for it : Design, Cost, Build, Operations, Optimization, Upgrades, Security…etc

Cloud (operational expense)



Resource Pooling

Resource Pooling

The provider's computing resources are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter). Examples of resources include storage, processing, memory, and network bandwidth.

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Enables providers to improve resource utilization and to flexibly provision and reclaim resources

Rapid Elasticity

Rapid Elasticity

Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Consumers can adapt to variations in workloads and maintain required performance levels
- Consumers may be able to avoid excessive costs from over-provisioning resources

Rapid Elasticity

Rapid Elasticity

Capabilities can be elastically provisioned and released, in some cases automatically, to scale rapidly outward and inward commensurate with demand. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be appropriated in any quantity at any time.

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Consumers can adapt to variations in workloads and maintain required performance levels
- Consumers may be able to avoid excessive costs from over-provisioning resources

Measured Service

Measured Service

Cloud systems automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, and active user accounts). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

– U.S. National Institute of Standards and Technology, Special Publication 800-145

- Enables billing of cloud services
- Resource monitoring helps providers with capacity and service planning
- Example: https://calculator.aws/#/createCalculator/EC2

Cloud Computing Benefits

Benefit	Description
Business agility	• Enables quick resource provisioning • Facilitates innovation • Reduces time-to-market
Reduces IT costs	• Reduces up-front capital expenditure (CAPEX) • Improves resource utilization • Reduces energy and space consumption
High availability	• Ensures resource availability based on consumer's requirements • Enables fault tolerance

Cloud Computing Benefits (Cont'd)

Benefit	Description
Business continuity	• Reduces impact of downtime • Example: Cloud-based backup
Flexible scaling	• Enables scaling of resources to meet demand • Unilateral and automatic resource scaling
Flexibility of access	• Enables access to services from anywhere • Eliminates dependency on a specific end-point device

Cloud Computing Benefits (Cont'd)

Benefit	Description
Application development and testing	• Enables application development and testing at a greater scale • Enables testing on multiple platforms
Simplified infrastructure management	• Consumers manage only those resources that are required to access cloud services
Increased collaboration	• Enables sharing and simultaneous access of resources and information
Masked complexity	• Intricacies of IT operations are hidden from end users

Cloud Computing Benefits (Cont'd)

Benefit	Description
Application development and testing	• Enables application development and testing at a greater scale • Enables testing on multiple platforms
Simplified infrastructure management	• Consumers manage only those resources that are required to access cloud services
Increased collaboration	• Enables sharing and simultaneous access of resources and information
Masked complexity	• Intricacies of IT operations are hidden from end users

Lesson Summary

During this lesson the following topics were covered:

- Definition of cloud computing
- Essential cloud characteristics
- Cloud computing benefits

Lesson: Cloud Service Models and Cloud Services Brokerage

This lesson covers the following topics:

- Infrastructure as a Service
- Platform as a Service
- Software as a Service
- Cloud services brokerage

Introduction to Cloud Service Models

- A cloud service model specifies the services and the capabilities provided to consumers
- NIST specifies three primary cloud service models:
  - Infrastructure as a Service (IaaS)
  - Platform as a Service (PaaS)
  - Software as a Service (SaaS)

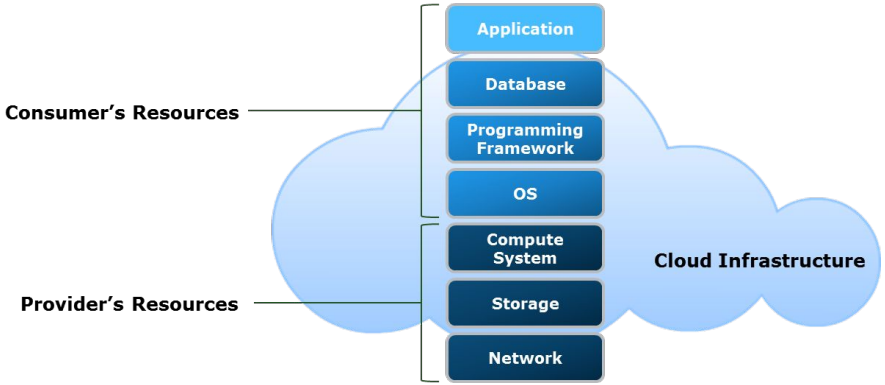
https://www.auvik.com/franklymsp/blog/aas-as-a-service-list/

Infrastructure as a Service

Infrastructure as a Service

The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, and deployed applications; and possibly limited control of select networking components , (e.g., host firewalls).

– U.S. National Institute of Standards and Technology, Special Publication 800-145

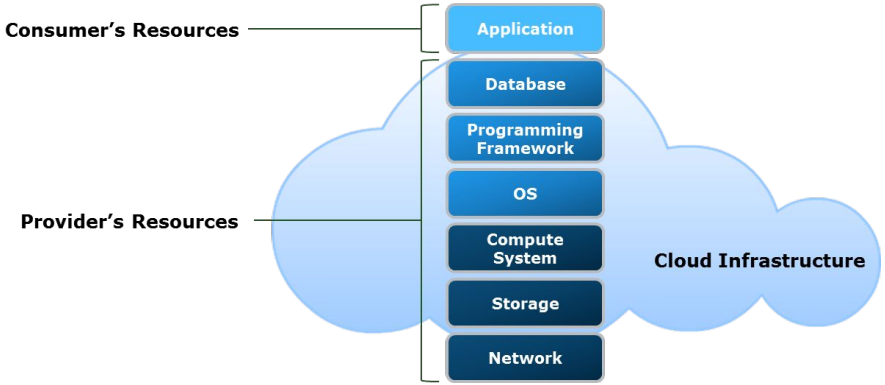


Platform as a Service

Platform as a Service

The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages, libraries, services, and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly configuration settings for the application-hosting environment.

– U.S. National Institute of Standards and Technology, Special Publication 800-145

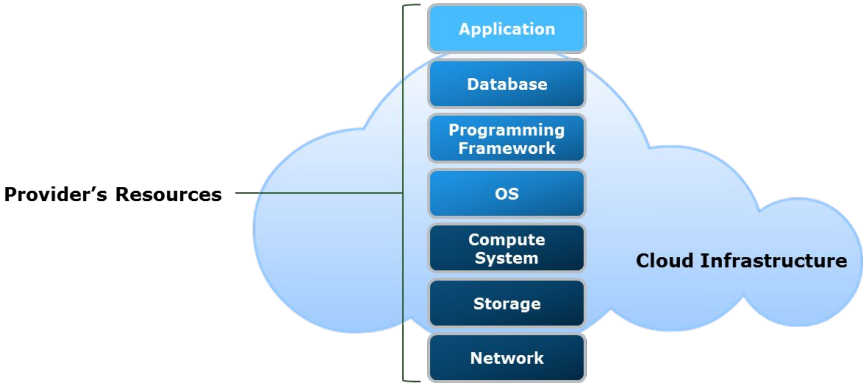


Software as a Service

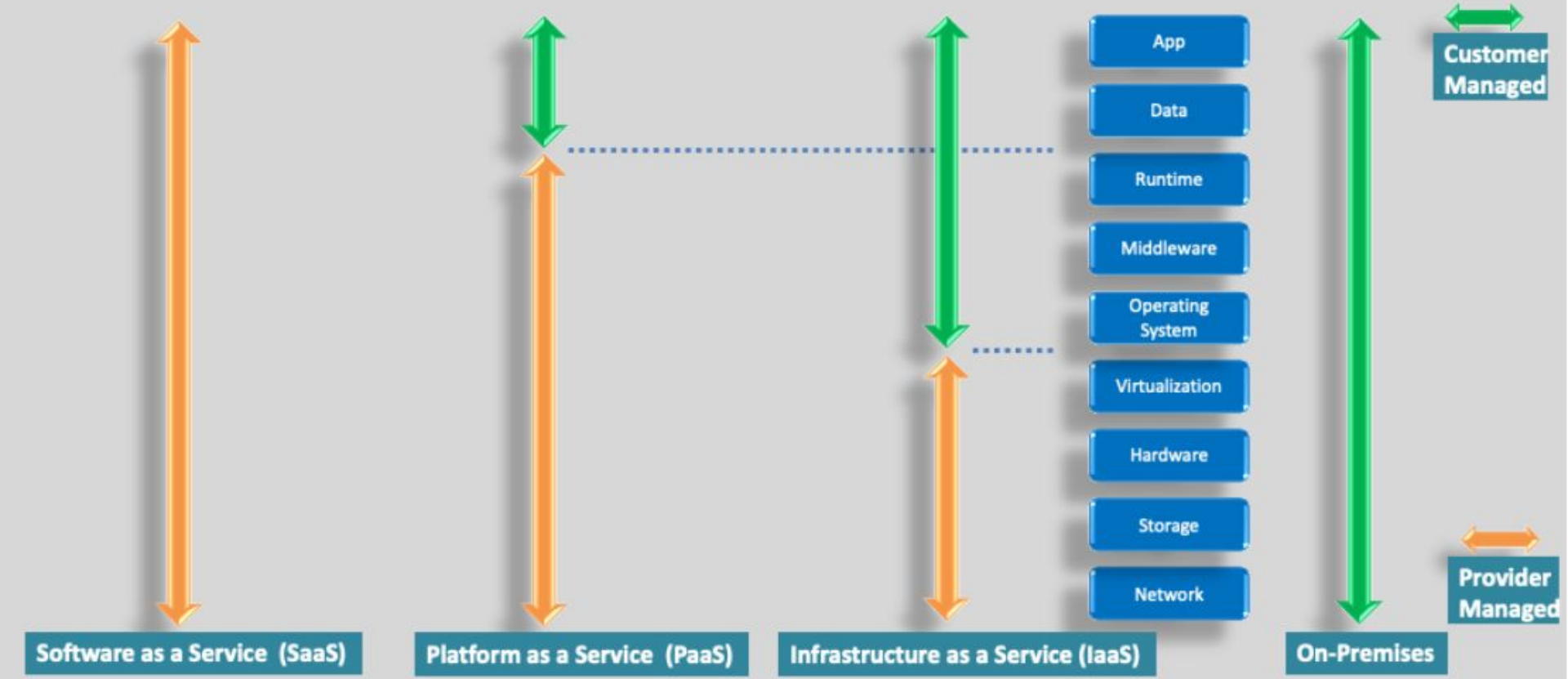
Software as a Service

The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through either a thin client interface, such as a web browser, (e.g., web-based email, or a program interface. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.

– U.S. National Institute of Standards and Technology, Special Publication 800-145



Cloud – as a Service (aaS) model



Cloud Services Brokerage (CSB)

Cloud Services Brokerage

Cloud services brokerage (CSB) is an IT role and business model in which a company or other entity adds value to one or more (public or private) cloud services on behalf of one or more consumers of that service.

– Gartner IT Glossary

- CSB is provided by a cloud broker
  - An entity that acts as an intermediary between cloud consumers and providers
- Cloud broker manages the use, performance ,and delivery of cloud services

Categories of Cloud Services Brokerage

Service intermediation

The broker enhances and adds value to a given service

Service aggregation

The broker combines and integrates multiple services into one or more new services

Service arbitrage

Similar to service aggregation except that the services being aggregated may vary

Categories of Cloud Services Brokerage

Service intermediation

The broker enhances and adds value to a given service

Service aggregation

The broker combines and integrates multiple services into one or more new services

Service arbitrage

Similar to service aggregation except that the services being aggregated may vary

Lesson Summary

During this lesson the following topics were covered:

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)
- Cloud services brokerage (CSB)

Lesson: Cloud Deployment Models

This lesson covers the following topics:

- Public cloud
- Private cloud
- Community cloud
- Hybrid cloud

Introduction to Cloud Deployment Models

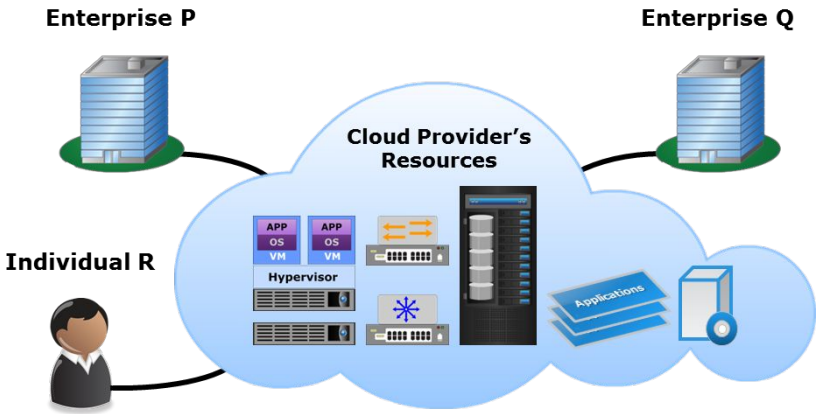
- A cloud deployment model specifies how a cloud infrastructure is built, managed, and accessed
- NIST specifies four primary cloud deployment models:
  - Public
  - Private
  - Community
  - Hybrid

Public Cloud

Public Cloud

The cloud infrastructure is provisioned for open use by the general public. It may be owned, managed, and operated by a business, academic, or government organization, or some combination of them. It exists on the premises of the cloud provider.

– U.S. National Institute of Standards and Technology, Special Publication 800-145



Private Cloud

Private Cloud

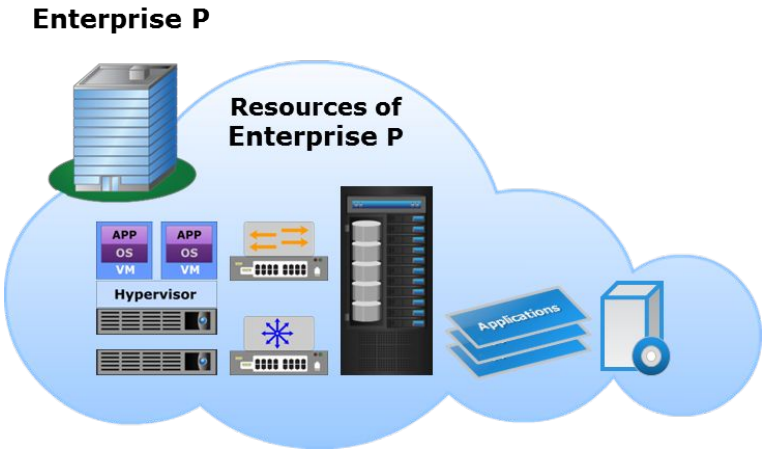
The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (for example, business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.

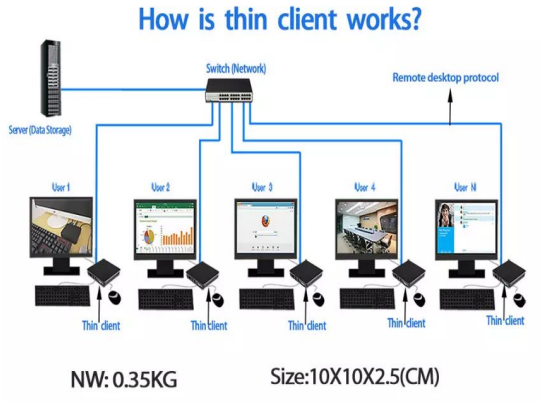
– U.S. National Institute of Standards and Technology, Special Publication 800-145

- There are two variants of private cloud:
  - On-premise
  - Externally-hosted

On-premise Private Cloud

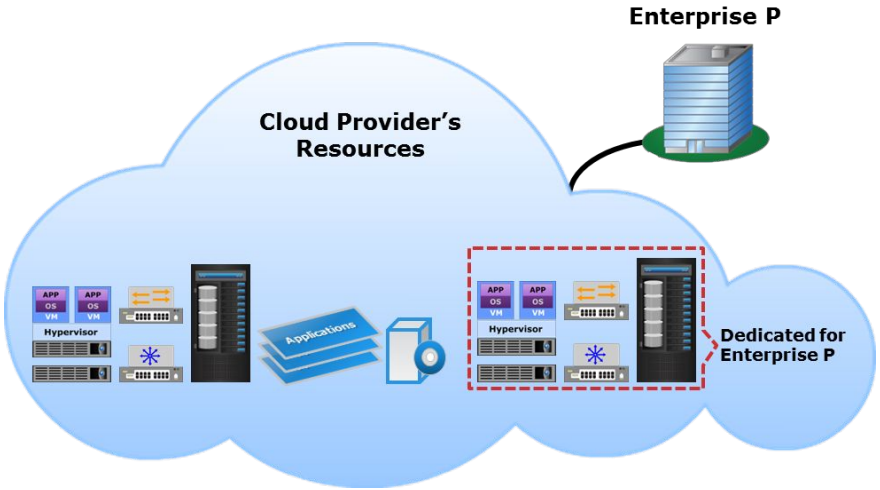
- Cloud infrastructure is deployed by an organization on its data centers within its premises
  - Provides complete control over the infrastructure and data
  - Enables standardization of IT resources, processes, and services





Externally-hosted Private Cloud

- Cloud implementation is outsourced to an external provider
- Cloud is hosted on the provider's premises and the consumers connect to it over a secure network
  - Access policies isolate the cloud resources from other tenants



Private vs. Public Cloud

- Resources are publicly shared
- Shared / Multi-Tenant infrastructure
- Internet connectivity primarily used
- Best suited for less confidential data

- Resources are for private use
- For a single customer, or for corporate business units
- Connectivity over fiber, private network and Internet
- Best suited for confidential data, client private systems

Community Cloud

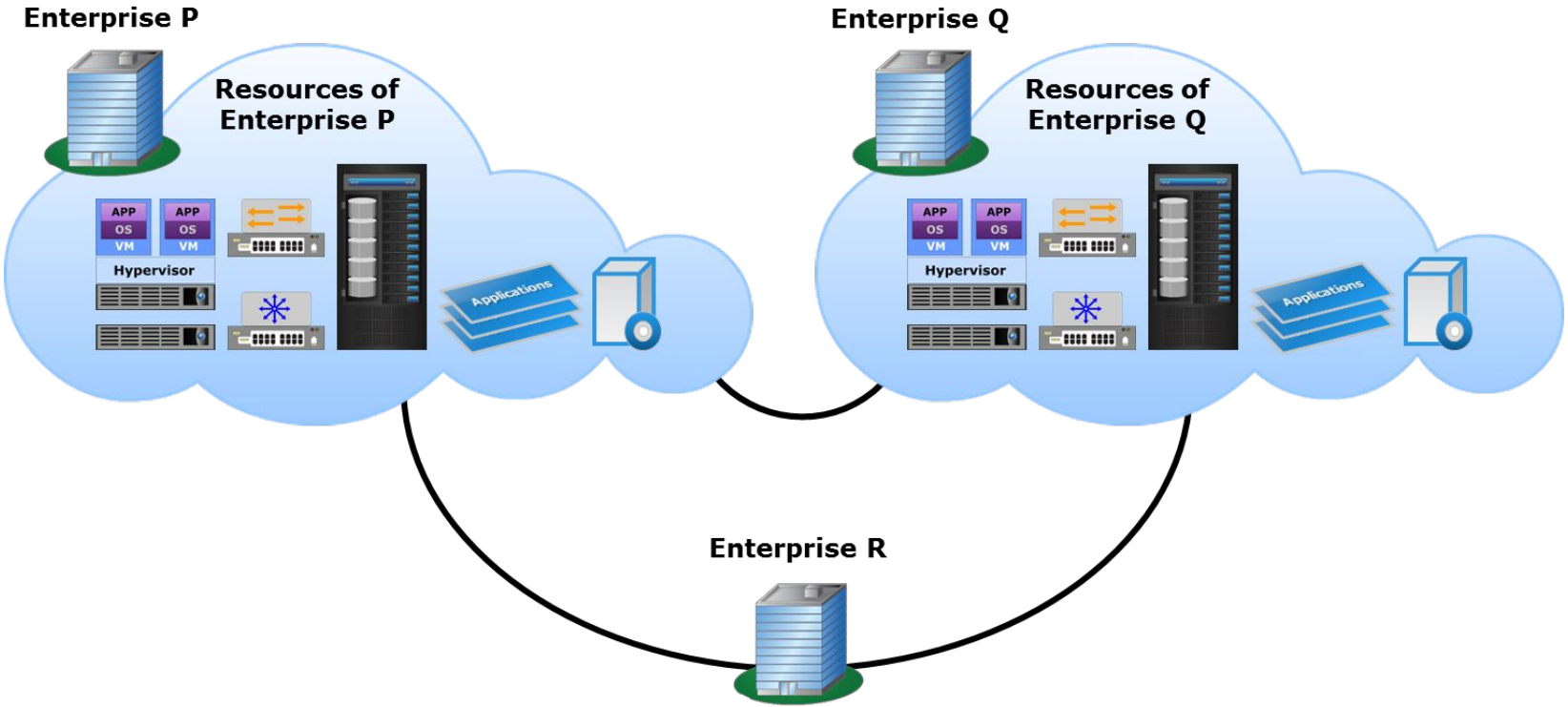
Community Cloud

The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organizations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organizations in the community, a third party, or some combination of them, and it may exist on or off premises.

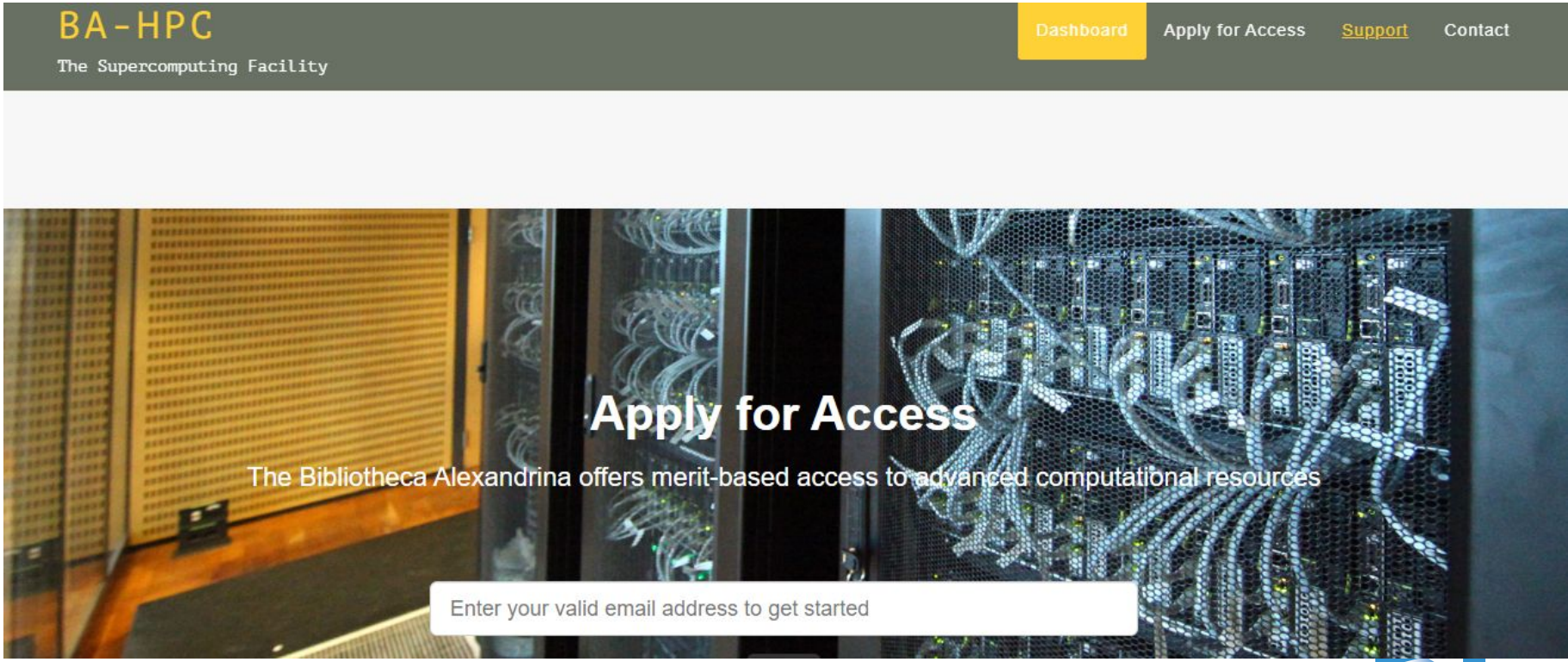
– U.S. National Institute of Standards and Technology, Special Publication 800-145

- There are two variants of community cloud:
  - On-premise
  - Externally-hosted

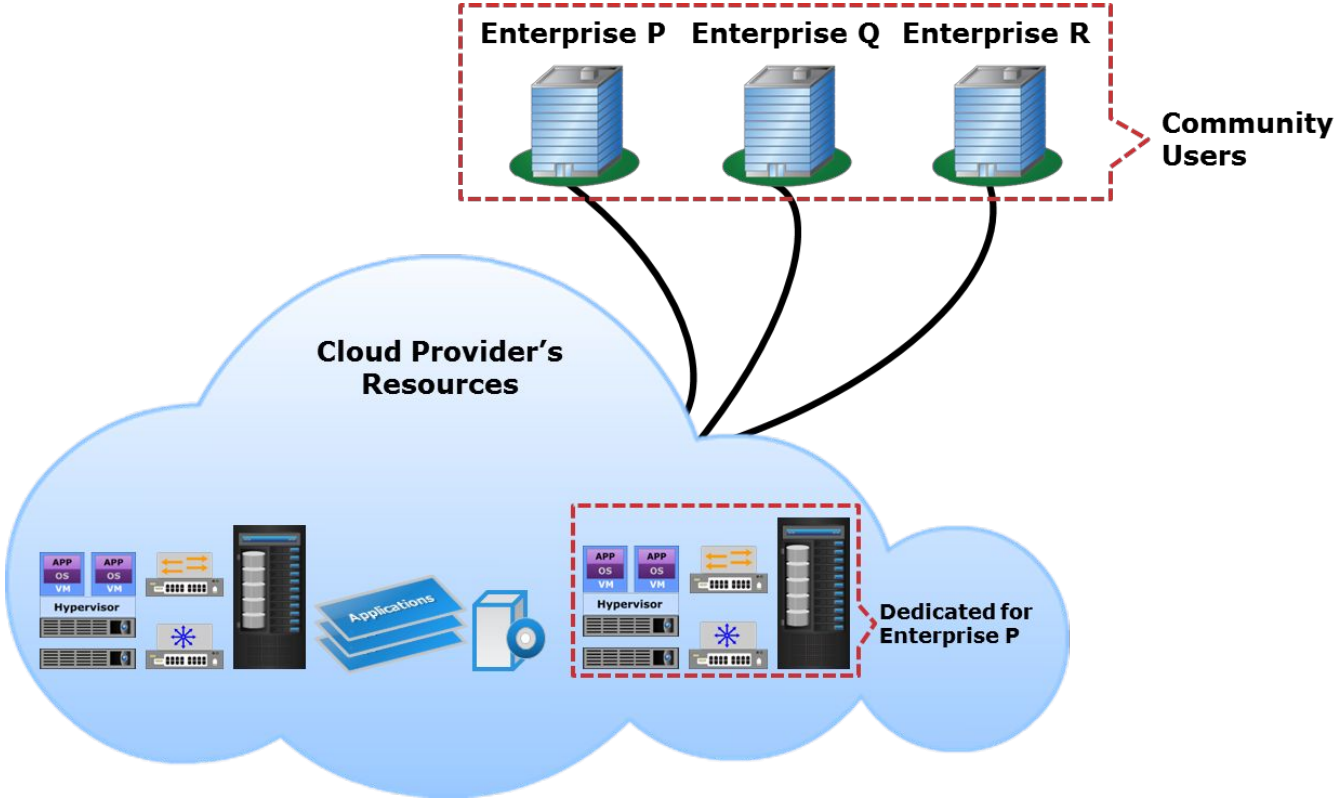
On-premise Community Cloud



Example high performance computing Alexandria university



Externally-hosted Community Cloud

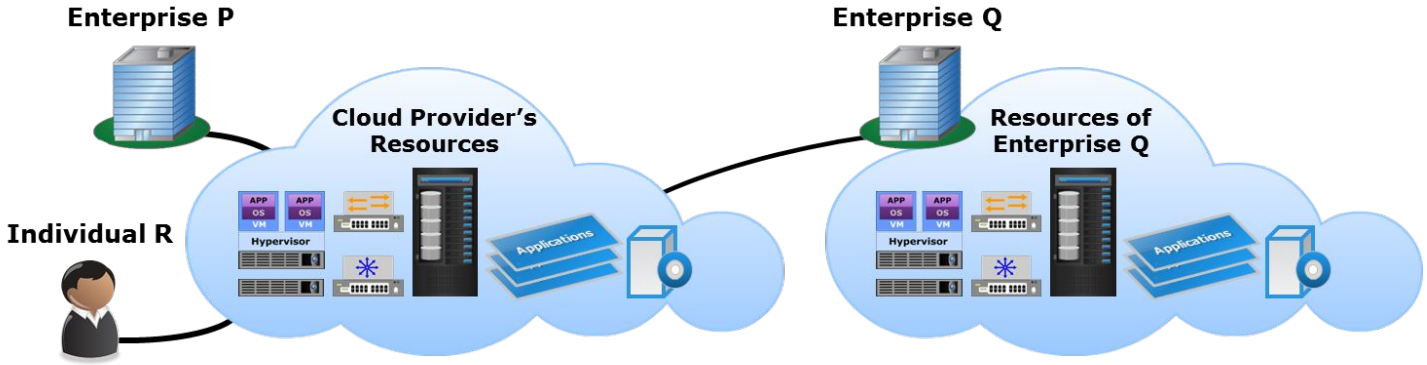


Hybrid Cloud

Hybrid Cloud

The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).

– U.S. National Institute of Standards and Technology, Special Publication 800-145



Multi Cloud

is the use of multiple cloud computing and storage services in a single heterogeneous architecture.

Organizations will continue to use multiple cloud providers for different use cases.

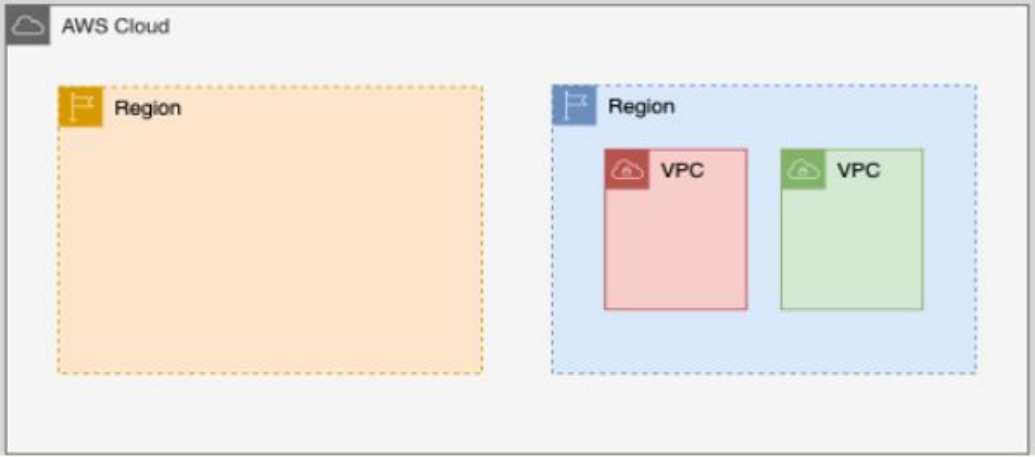
Hybrid Cloud Model Use Cases

Use case	Description
Cloud bursting	Provisioning resources for a limited time from a public cloud to handle peak workloads
Web application hosting	Hosting less critical applications such as e-commerce applications on the public cloud
Migrating packaged applications	Migrating standard packaged applications such as email to the public cloud
Application development and testing	Developing and testing applications in the public cloud before launching them

AWS Virtual Private Cloud (VPC)

Components

Virtual Private Cloud (VPC) – What is it?



- Is a virtual data center in the cloud, confined to an AWS Region, and is isolated from other VPCs by default.
- The client has full control over it.
- Can have one or more subnets
- A VPC is confined to a single AWS region.

VPC Components – CIDR Block

- The VPC's main CIDR Block can NOT be changed after it has been created.
- We can expand the VPC address pool by adding up to 4 additional secondary CIDR blocks.
- Limitations can be found here:
  https://docs.aws.amazon.com/vpc/latest/userguide/VPC_Subnets.html#add-cidr-block-restrictions

Hybrid Cloud Model Use Cases

Use case	Description
Cloud bursting	Provisioning resources for a limited time from a public cloud to handle peak workloads
Web application hosting	Hosting less critical applications such as e-commerce applications on the public cloud
Migrating packaged applications	Migrating standard packaged applications such as email to the public cloud
Application development and testing	Developing and testing applications in the public cloud before launching them

Lesson Summary

During this lesson the following topics were covered:

- Public cloud
- Private cloud: on-premise and externally-hosted
- Community cloud: on-premise and externally-hosted
- Hybrid cloud and its use cases

Concepts in Practice

- VMware vCloud Hybrid Service (vCHS)
- Pivotal Cloud Foundry
- EMC Mozy

VMware vCHS, Pivotal Cloud Foundry, and EMC Mozy

vCHS	Cloud Foundry	Mozy
• Hybrid cloud service • Provides IaaS for migrating / extending workloads to public cloud, application development, and disaster recovery	• Open-source PaaS project • Supports multiple cloud deployment models, programming languages, and database systems	• SaaS solution for secure, cloud-based online backup and recovery • Provides automatic and scheduled backups

Module Summary

Key points covered in this module:

- Definition of cloud computing
- Essential cloud characteristics
- Key benefits of cloud computing
- Cloud service models
- Cloud services brokerage
- Cloud deployment models

Module: Building the Cloud Infrastructure

Upon completion of this module, you should be able to:

- Describe the cloud computing reference model
- Describe the deployment options and solutions for building a cloud infrastructure
- Describe various factors to consider while building a cloud infrastructure

Lesson: Cloud Computing Reference Model

This lesson covers the following topics:

- Layers of cloud computing reference model
- Entities and functions of each layer
- Cross-layer functions of cloud computing reference model

Bare metal server

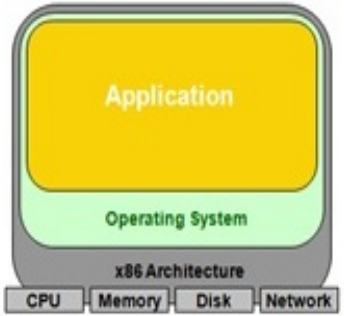
Typical datacenter

Raised floor

Air condition flow

Physical vs virtual

Traditional Physical Server

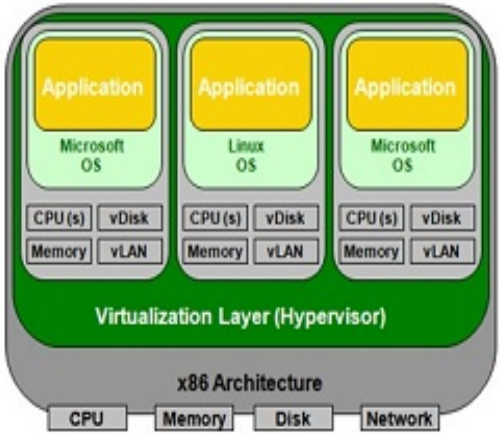


1 Physical Server, 1 Application

Traditional x86 Server Architecture

- Single operating system per machine
- Single application per machine
- Hardware components connected directly to operating system
  - CPU
  - Memory
  - Disk
  - Network Card

New Architecture: Virtual Server

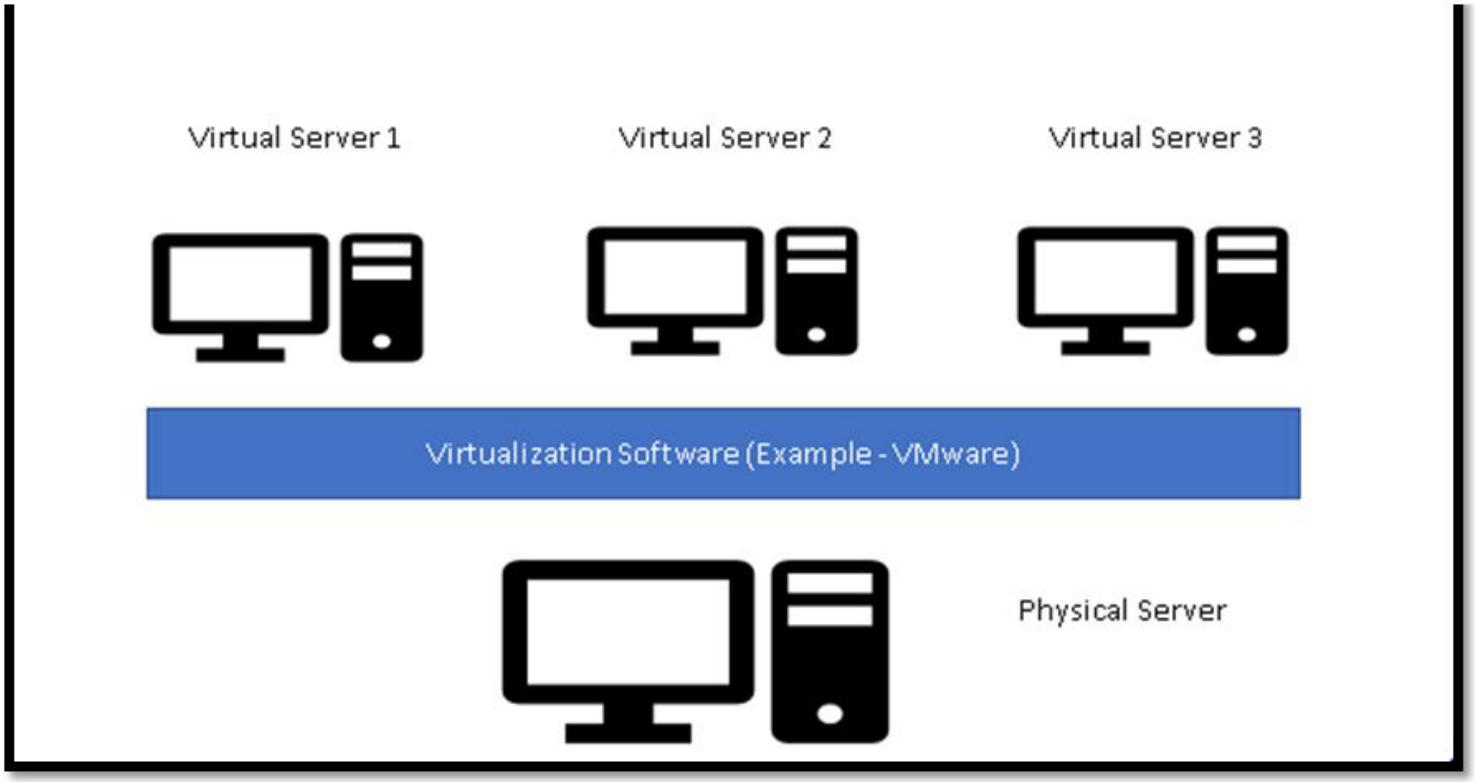


3 Virtual Servers on 1 Physical Server

Virtualization Layer

- Addition of a virtualization layer called a "hypervisor"
- Several servers can be deployed as Virtual Machines (VM) on each physical box
- Each VM has its own operating system and application
- Can run multiple, different operating systems on the same machine
- If one VM fails, other VMs are unaffected

Physical vs virtual server



Servers

Virtual vs Physical

COST
No paying for what you don't use, just pay a monthly usage charge for the virtual machine.

EFFICIENCY
Servers can be consolidated meaning that on average they are run at around 80% utilisation.

SPACE
A single physical server can host multiple virtual servers, saving space.

FLEXIBILITY
A lack of reliance on hardware allows you far greater flexibility – the servers can be moved as easily as a file.

DISASTER RECOVERY
Create a test environment for failover plans using the software included with the virtualisation platform.

concorde
IT GROUP
concordeitgroup.com

COST
There is large up-front capital expenditure, followed by smaller monthly costs for power and cooling in the data centre.

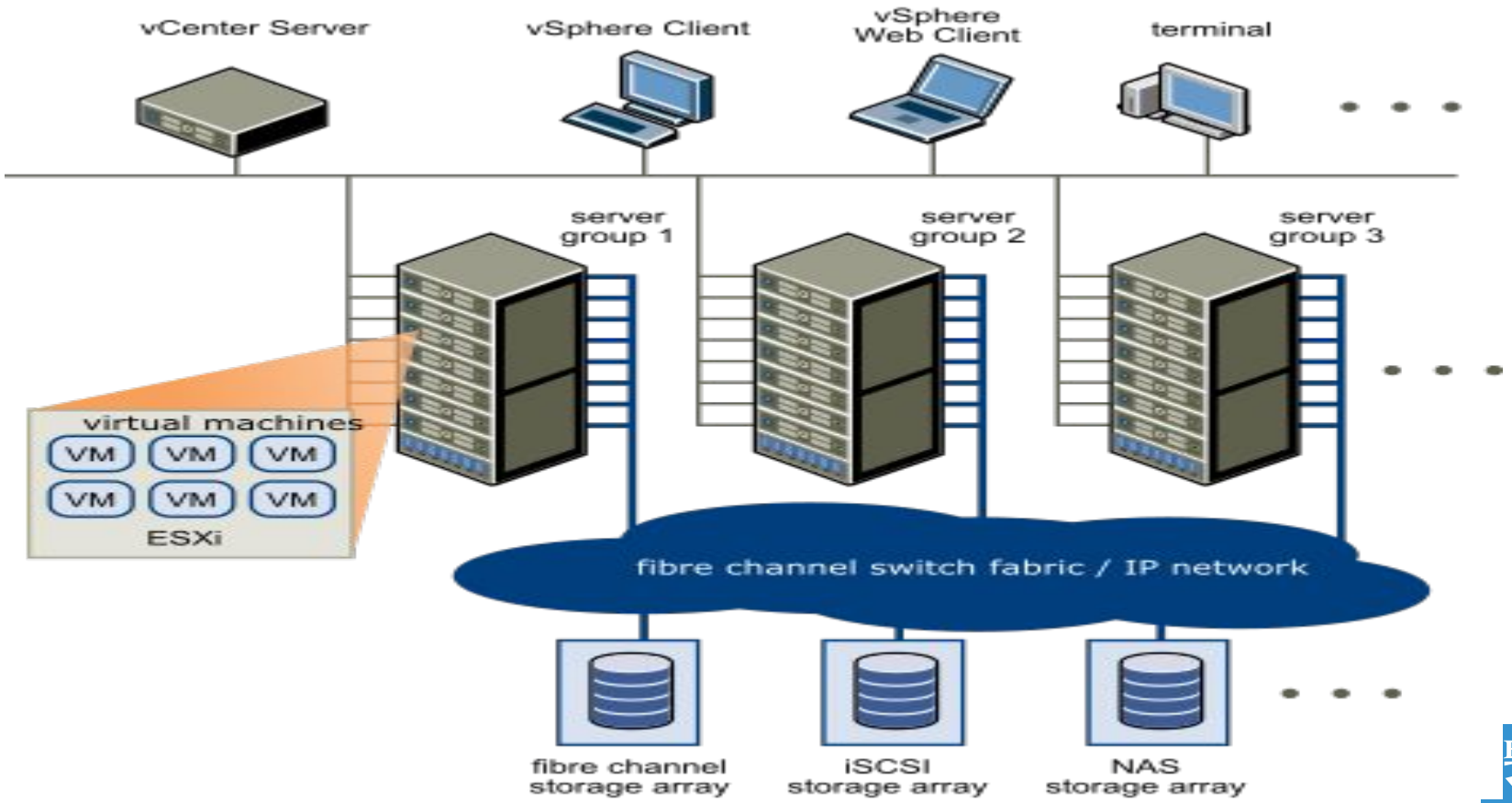
EFFICIENCY
Only 5-15% utilised on average, requiring more servers.

SPACE
Space is required for the servers, extra networking equipment and racks.

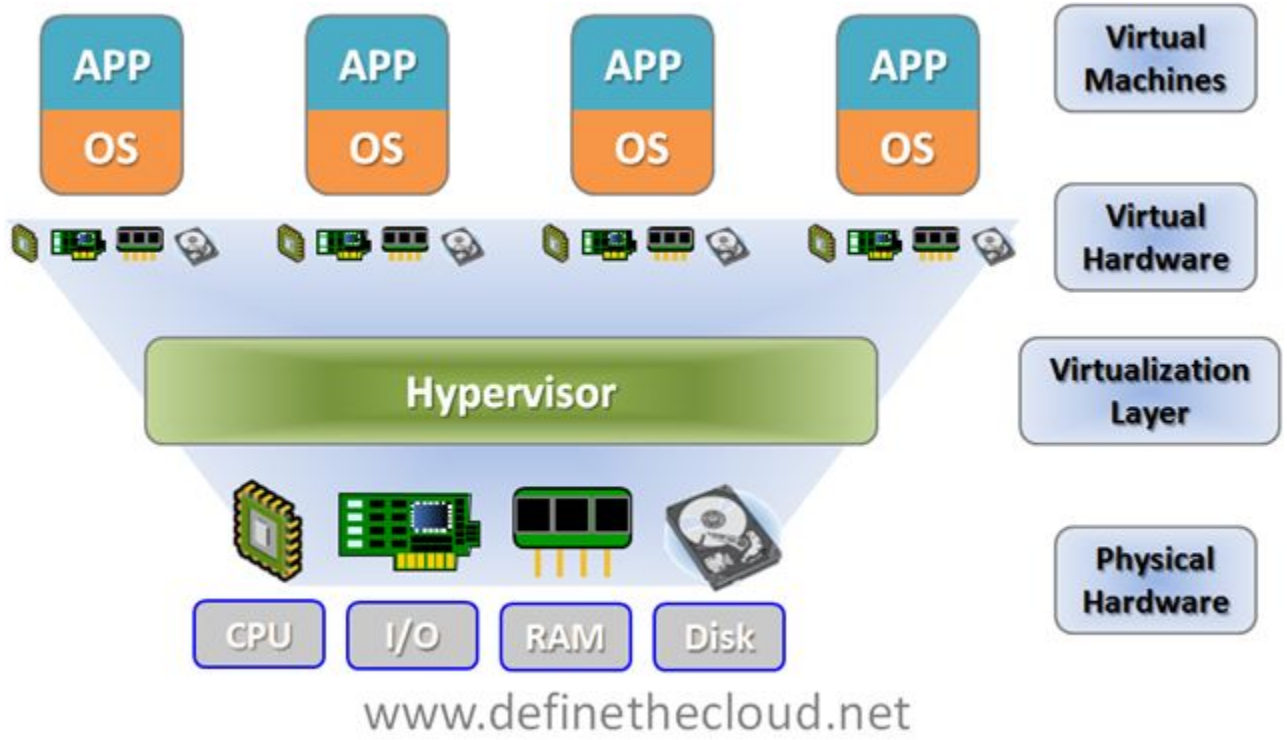
FLEXIBILITY
Any new set-up needs to mirror the existing set-up to ensure that the new server will work as required, making it far more difficult to switch.

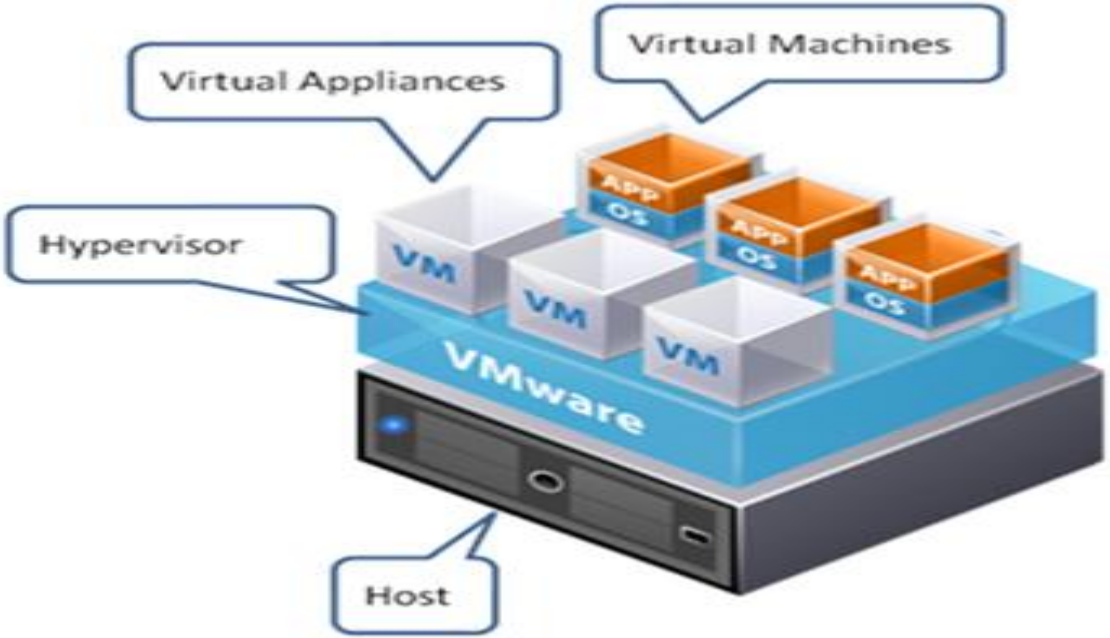
DISASTER RECOVERY
Regular back-ups must be created, and duplicate hardware is required in case of problems with the initial server hardware.

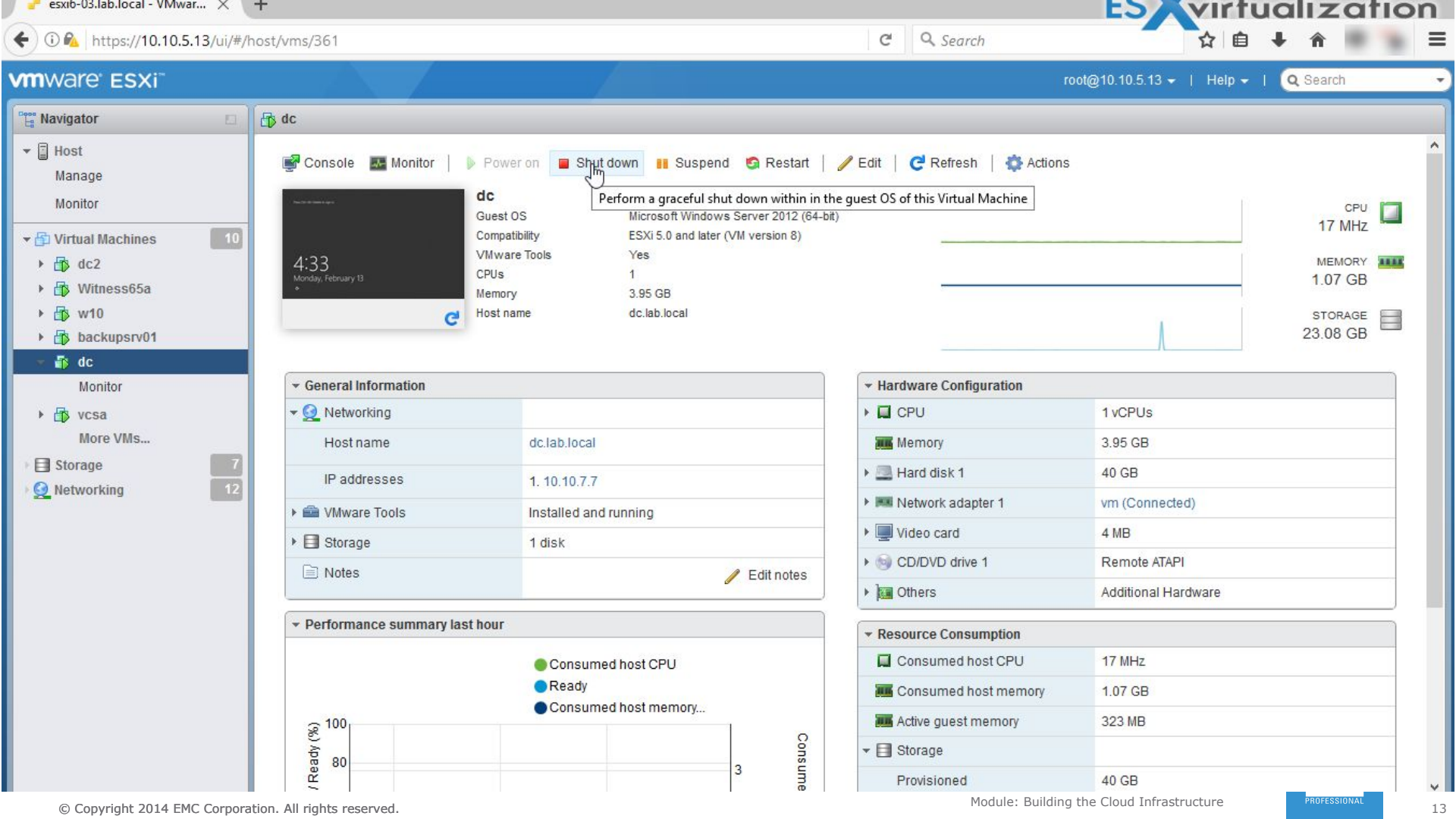
datacenter

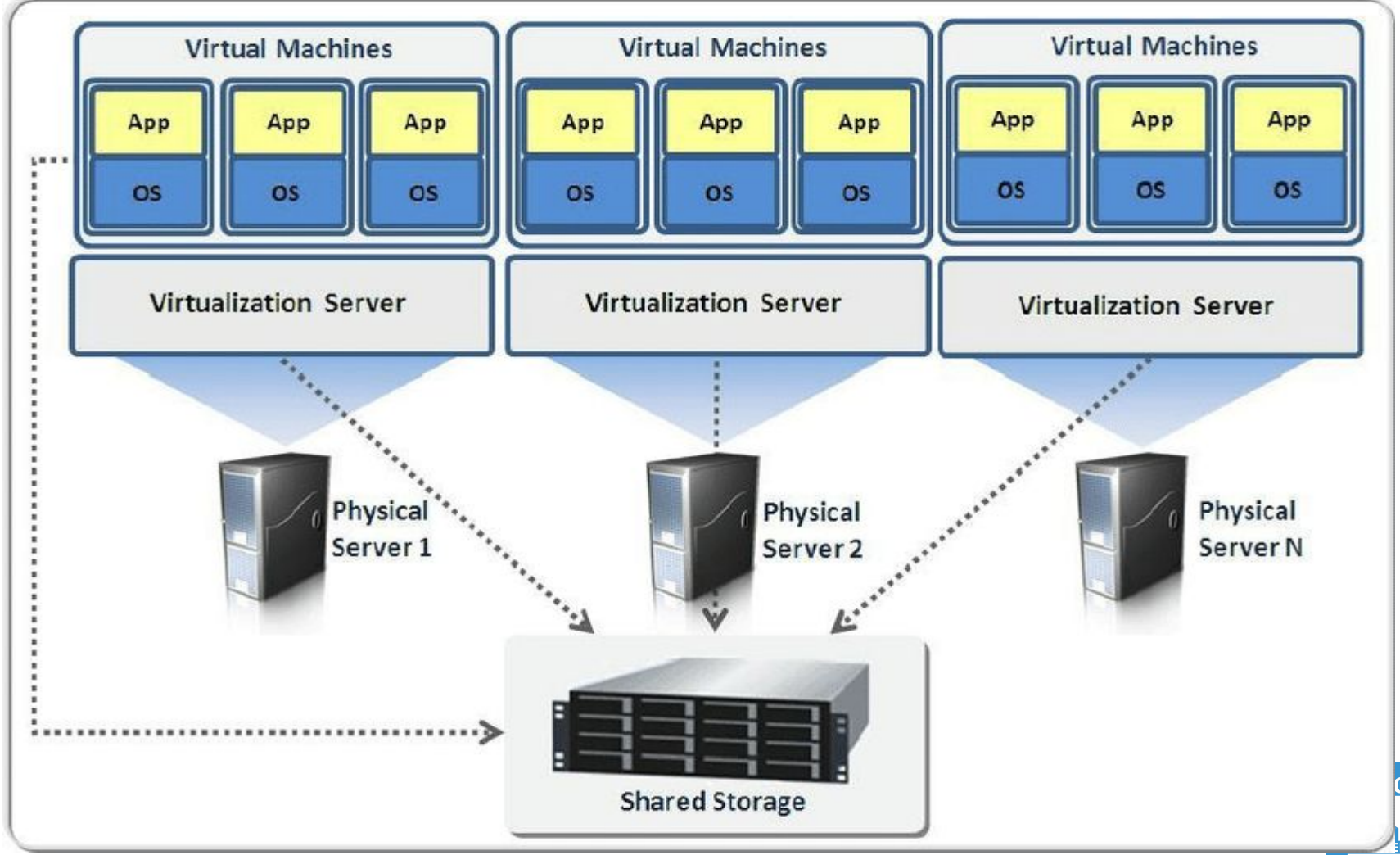


Inside a server

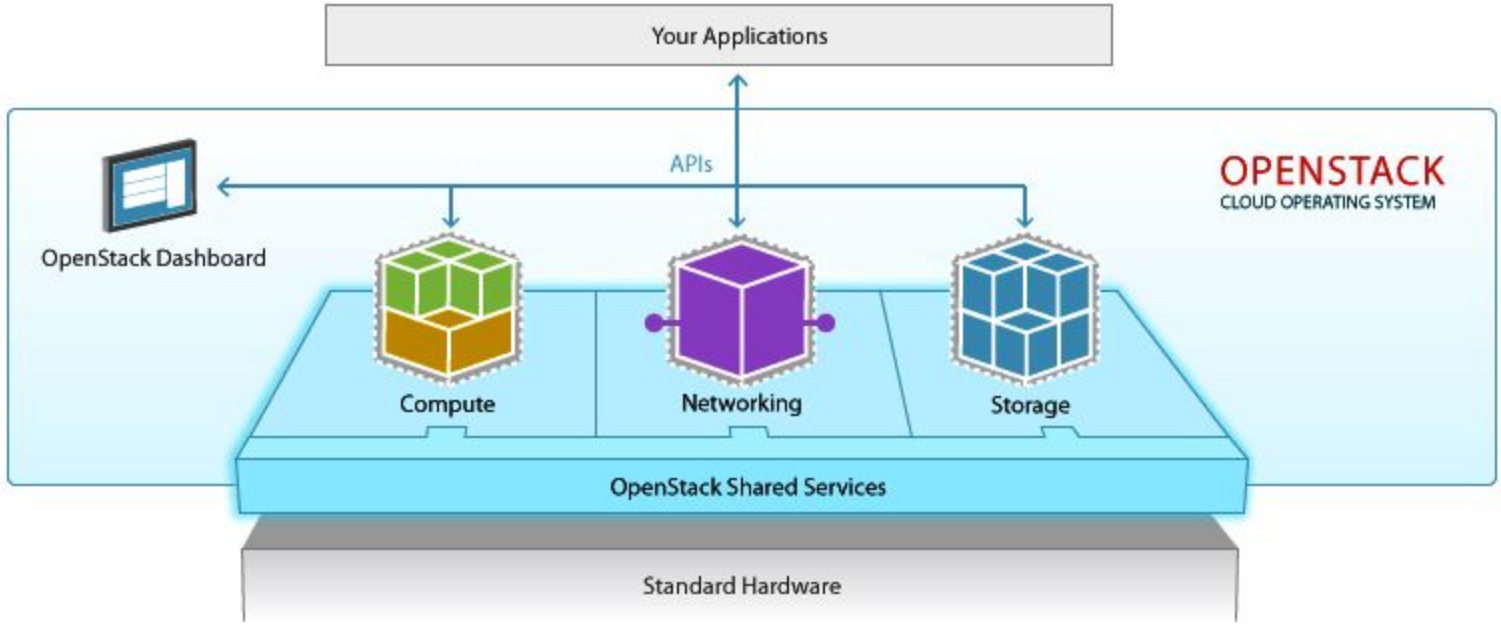








Cloud management software



What is a Reference Model?

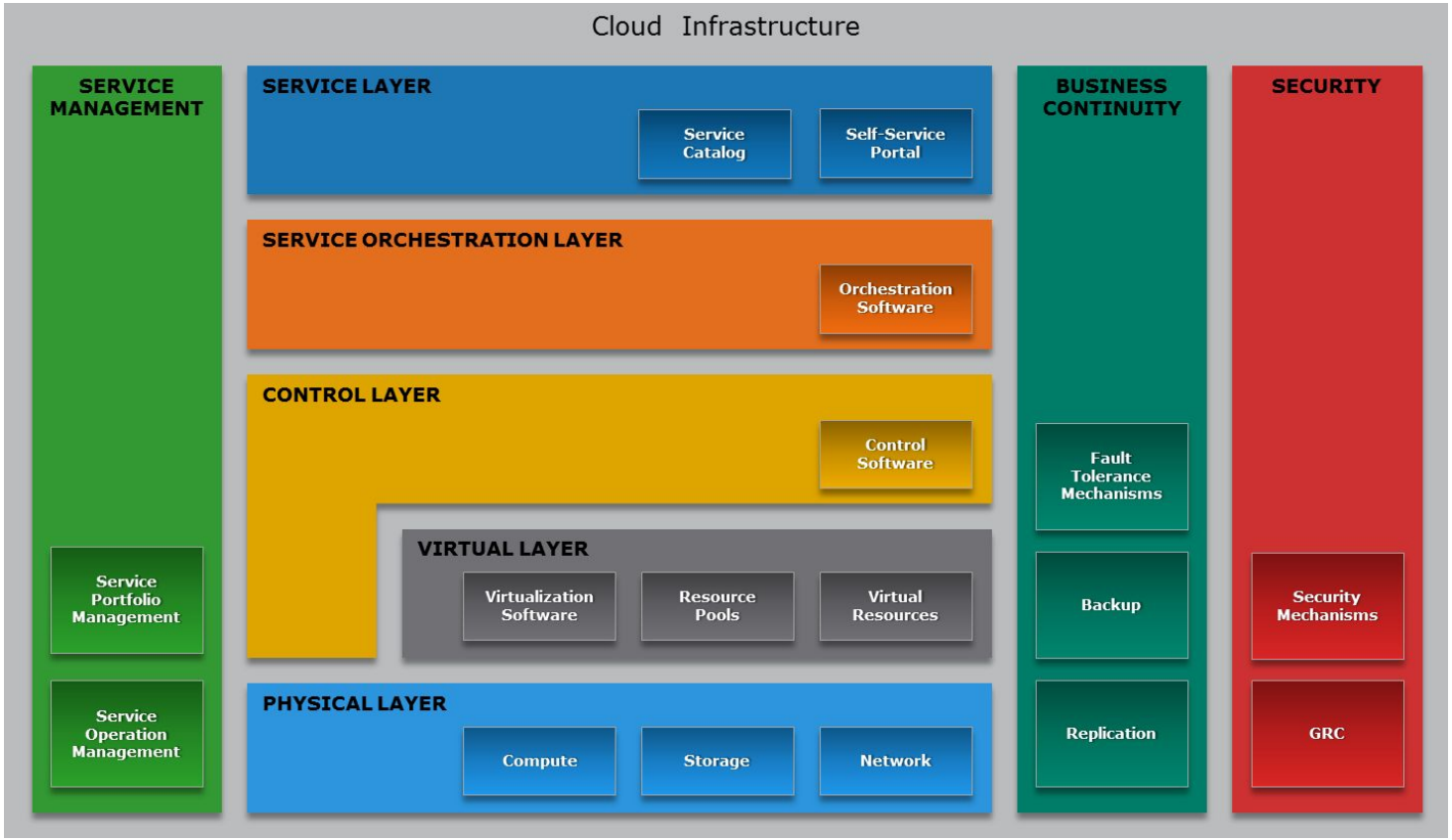
Reference Model

A reference model is an abstract framework for understanding significant relationships among the entities of some environment, and for the development of consistent standards or specifications supporting that environment. It is based on a small number of unifying concepts and may be used as a basis for education and explaining standards. It is not directly tied to any standards, technologies, or other concrete implementation details, but it does seek to provide a common semantics that can be used unambiguously across and between different implementations.

- Organization for the Advancement of Structured Information Standard (OASIS)

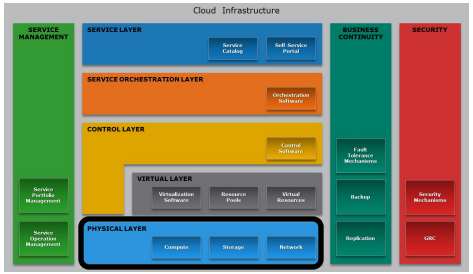
- Facilitates efficient communication of system details between stakeholders
- Provides a point of reference for system designers to extract system specifications

Cloud Computing Reference Model



Cloud Computing Layer

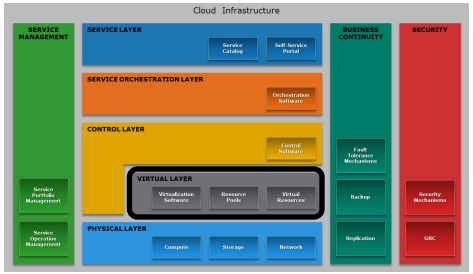
Physical Layer



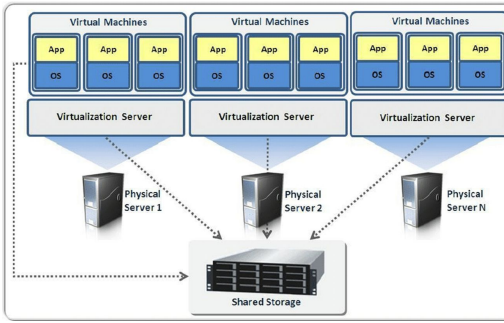
- Foundation layer of the cloud infrastructure
- Specifies entities that operate at this layer:
  - Compute systems, network devices, and storage devices
  - Operating environment, protocol, tools, and processes
- Functions of physical layer:
  - Executes requests generated by virtualization and control layer

Cloud Computing Layer

Virtual Layer

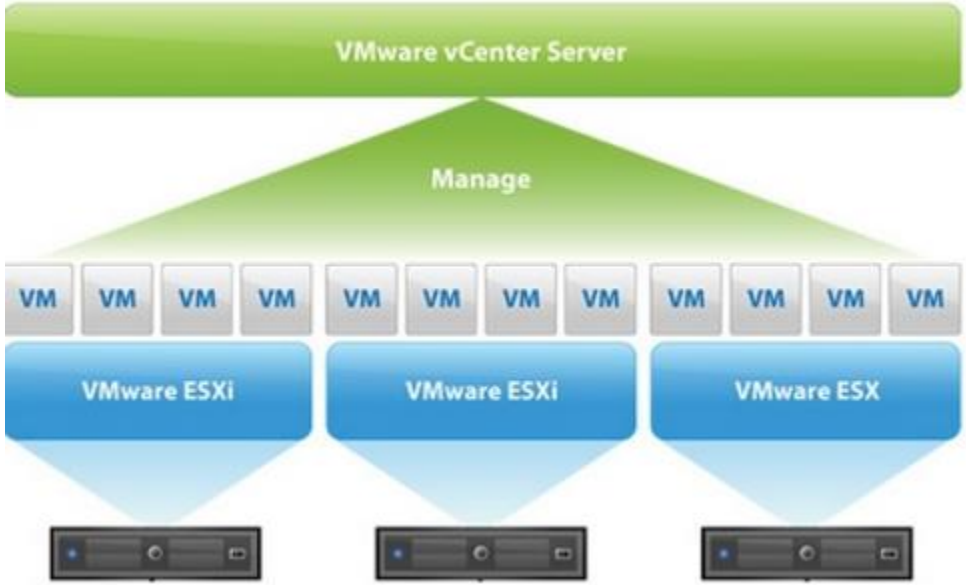


- Deployed on the physical layer
- Specifies entities that operate at this layer:
  - Virtualization software
  - Resource pools
  - Virtual resources



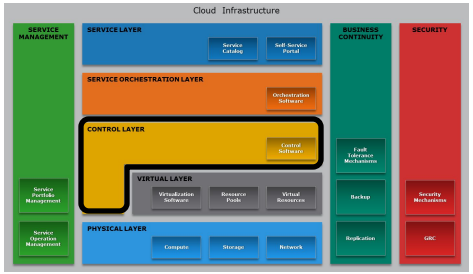
- Functions of virtual layer:
  - Abstracts physical resources and makes them appear as virtual resources
    - Enables multitenant environment, thereby improving utilization
  - Executes the requests generated by control layer

Vmware Vsphere

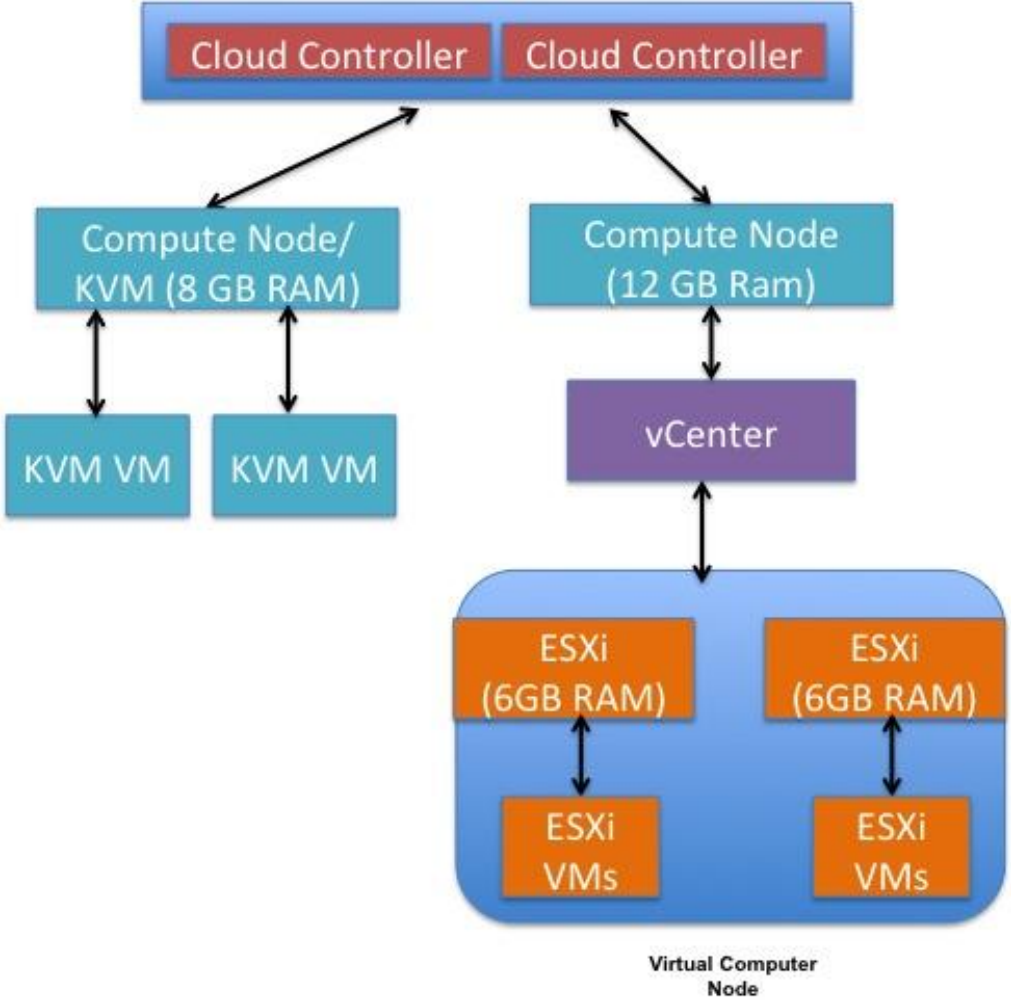


Cloud Computing Layer

Control Layer



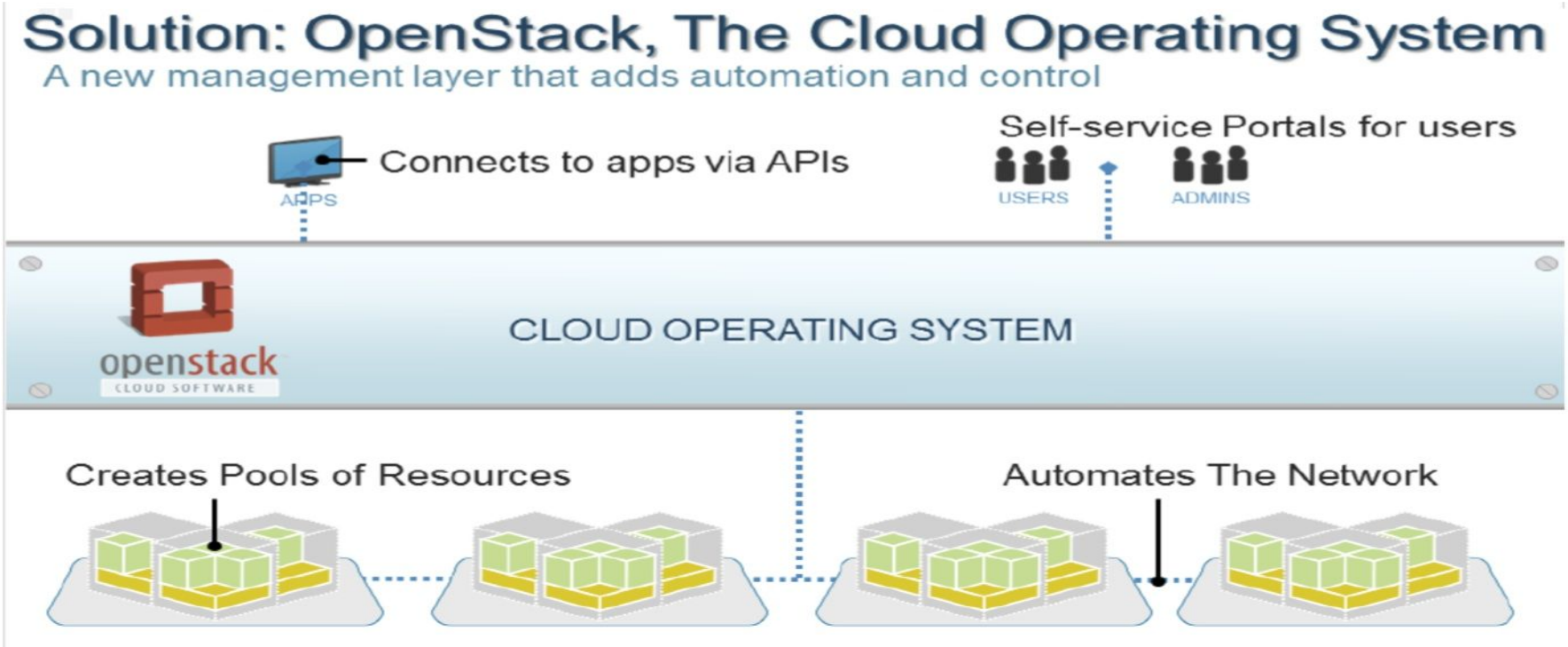
- Deployed either on virtual layer or on physical layer
- Specifies entities that operate at this layer – control software
- Functions of control layer:
  - Enables resource configuration and resource pool configuration
  - Enables resource provisioning
  - Executes requests generated by service layer
  - Exposes resources to and supports the service layer
  - Collaborates with the virtualization software and enables
    - Resource pooling and creating virtual resources
    - Dynamic allocation of resources
    - Optimizing utilization of resources



Virtual Computer Node

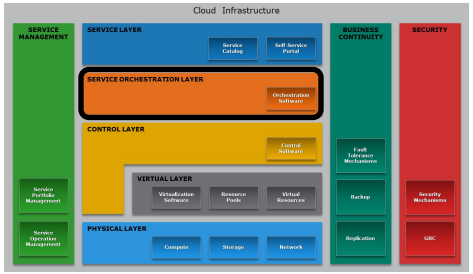
openstack

Automation and Orchestration of IT Resources



Cloud Computing Layer

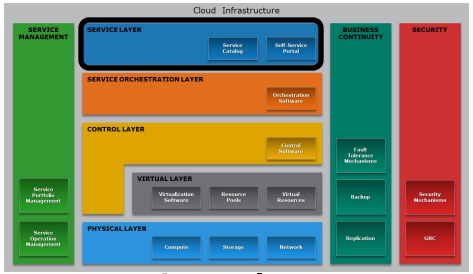
Service Orchestration Layer



- Specifies the entities that operate at this layer:
  - Orchestration software
- Functions of orchestration layer:
  - Provides workflows for executing automated tasks
  - Interacts with various entities to invoke provisioning tasks

Cloud Computing Layer

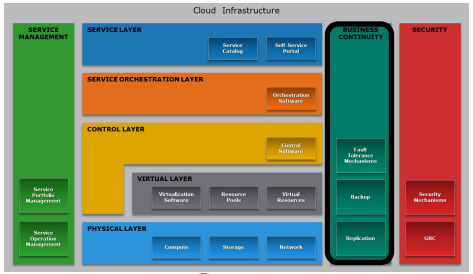
Service Layer



- Consumers interact and consume cloud resources via this layer
- Specifies the entities that operate at this layer:
  - Service catalog
  - Self-service portal
- Functions of service layer:
  - Stores information about cloud services in service catalog and presents them to the consumers
  - Enables consumers to access and manage cloud services via a self-service portal

Cross-layer Function

Business Continuity



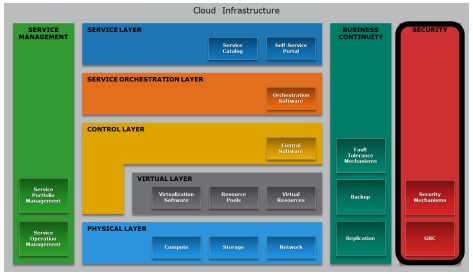
- Specifies adoption of measures to mitigate the impact of downtime:

Measures	Description
Proactive	• Business impact analysis • Risk assessment • Technology solutions deployment (backup and replication)
Reactive	• Disaster recovery • Disaster restart

- Enables ensuring the availability of services in line with SLA
- Supports all the layers to provide uninterrupted services

Cross-layer Function

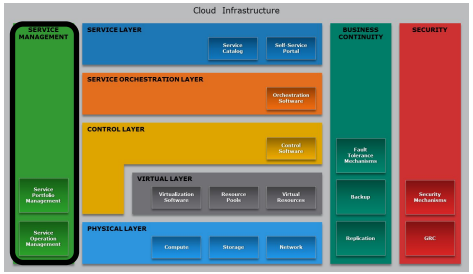
Security



- Specifies the adoption of:
  - Administrative mechanisms
    - Security and personnel policies
    - Standard procedures to direct safe execution of operations
  - Technical mechanisms
    - Firewall
    - Intrusion detection and prevention systems
    - Antivirus
- Deploys security mechanisms to meet GRC requirements
- Supports all the layers to provide secure services

Cross-layer Function

Service Management



- Specifies adoption of activities related to:

Activities	Description
Service portfolio management	• Defines service roadmap, service features, and service levels • Establishes budgeting and pricing • Deals with consumers in supporting activities *(i.e. taking orders, processing bills, and collecting payments)* • Performs market research • Collects information about competitors
Service operation management	• enables cloud administrators to manage cloud infrastructure and services • Enables infrastructure configuration and resource provisioning • Enables problem resolution • Enables capacity and availability management • Enables compliance conformance • Enables monitoring cloud services and their constituent elements

Lesson Summary

During this lesson the following topics were covered:

- Cloud computing reference model
- Entities and functions of the five layers
- Activities of the three cross-layer functions

Lesson: Options for Building a Cloud Infrastructure

This lesson covers the following topics:

- Greenfield and brownfield deployment options
- Technology solutions for building a cloud infrastructure

Deployment Options

Greenfield Deployment Option

It is typically used when an infrastructure does not exist and an organization has to build the cloud infrastructure starting from the physical layer.

Brownfield Deployment Option

It is used when some of the infrastructure entities exist, which can be transformed to cloud infrastructure by deploying the remaining entities required for the cloud infrastructure.

Solutions for Building Cloud Infrastructure

- Two solutions for building cloud infrastructure:
  - Integrating best-of-breed cloud infrastructure components
  - Cloud-ready converged infrastructure

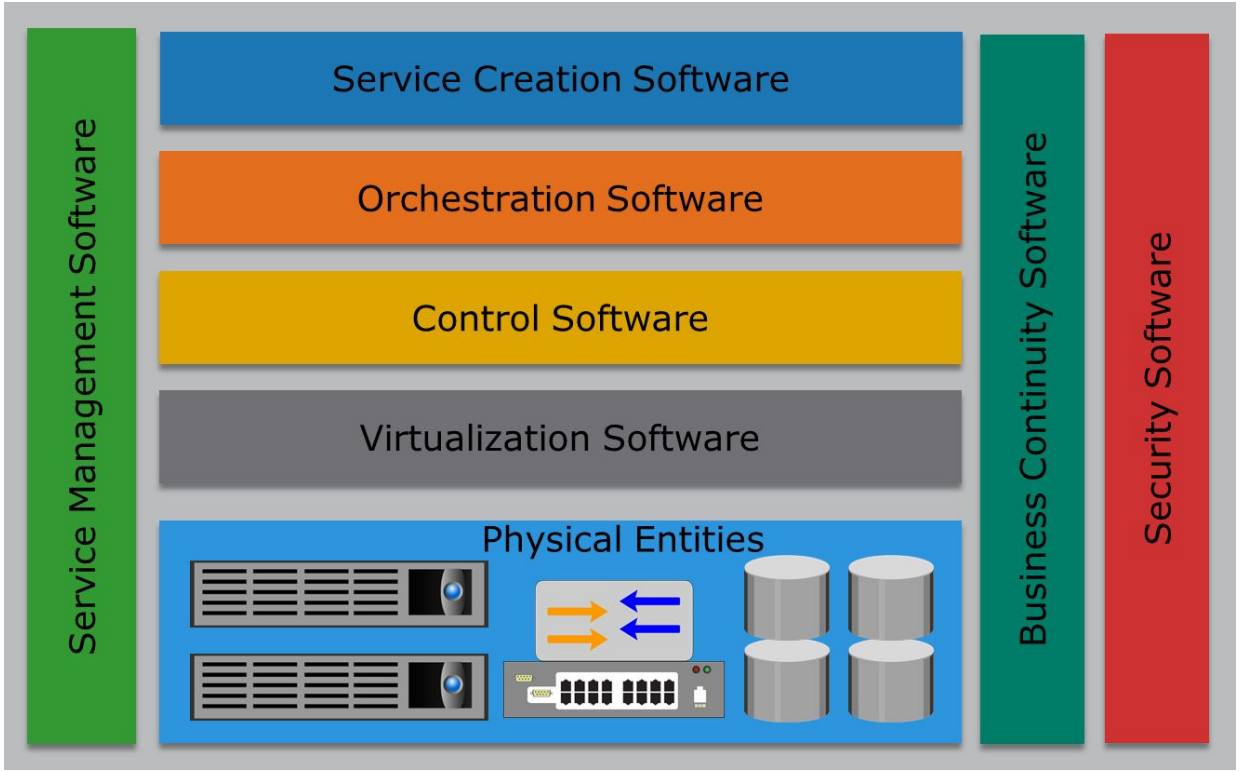
Solutions for Building Cloud Infrastructure

Integrating Best-of-breed Cloud Infrastructure Components

- Built by integrating multi-vendor infrastructure components
- Enables repurposing the existing infrastructure components
- Requires spending a significant amount of IT staff time on:
  - Evaluating individual and disparate hardware components
  - Installing and integrating infrastructure components
  - Testing hardware, middleware, and software
  - Checking compatibility of all the components
- Enables organizations to choose and switch vendors easily

Solutions for Building Cloud Infrastructure

Cloud-ready Converged Infrastructure



Lesson Summary

During this lesson the following topics were covered:

- Greenfield and brownfield deployment options
- Best-of-breed cloud infrastructure components
- Cloud-ready converged infrastructure

Lesson: Considerations for Building a Cloud Infrastructure

This lesson covers the following topics:

- Factors to consider while building a cloud infrastructure

Factors to Consider while Building a Cloud Infrastructure

- Governance
- Organization
- Finance
- Tools
- Service-level agreement and service contract

- Avoiding vendor lock-in
- Software licensing concerns
- Service model considerations
- Migration
- Testing

Governance

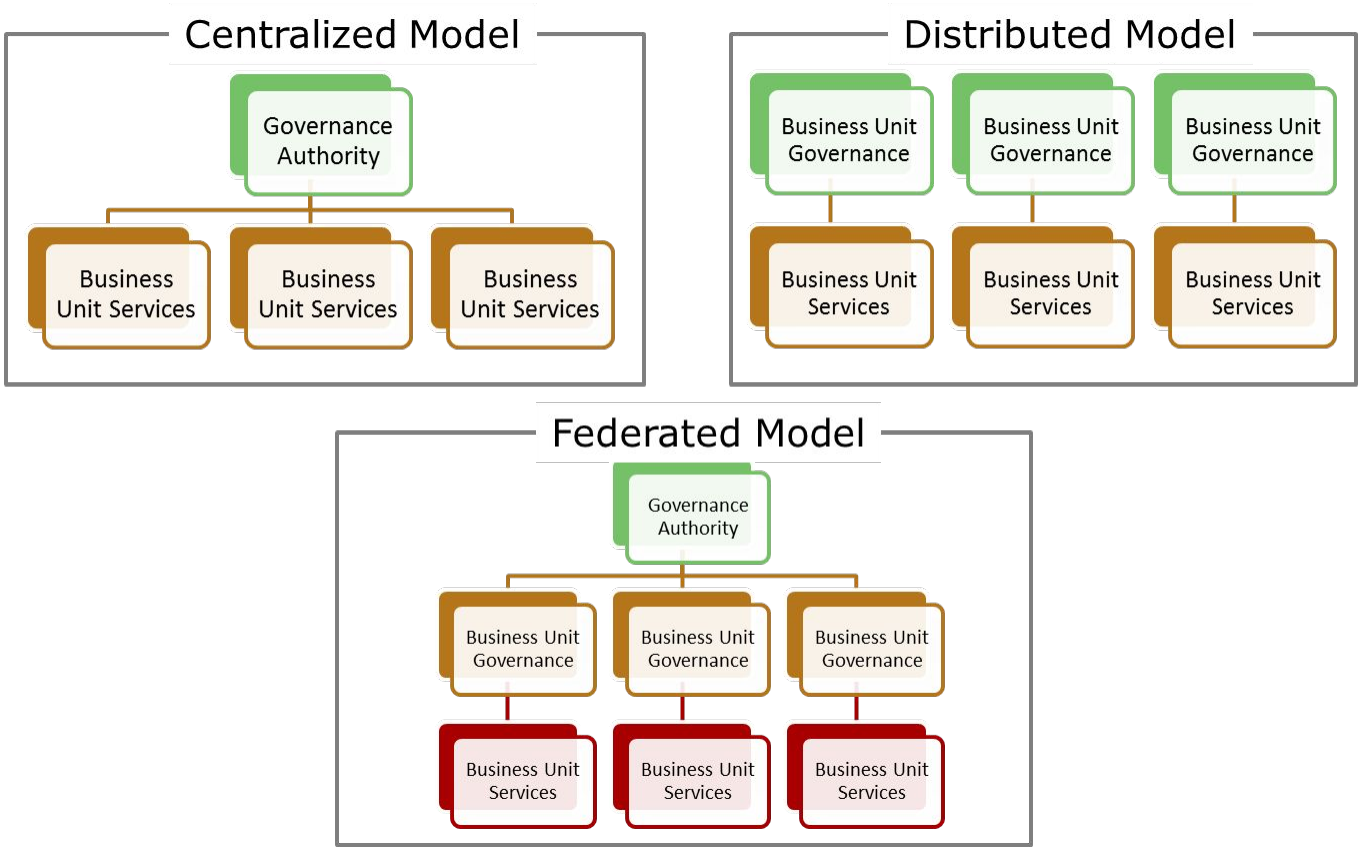
Governance

Governance is the active distribution of decision-making rights and accountability among different stakeholders in an organization. It also describes the rules and procedures for making and monitoring those decisions to determine and achieve the desired behaviors and results.

- IT governance enables the service provider to:
  - Ensure IT resources are implemented and used according to policies and procedures
  - Ensure the resources are properly controlled and maintained
  - Ensure the resources are providing value to the organization
- Instituting IT governance involves establishing a review board

  (responsible for creating rules and processes to follow to ensure that policies are being met)

Governance Models



Organization

New Roles in Cloud

Service Manager

- Key interface between clients and IT staff
- Understands consumers' needs and industry trends
- Ensures IT delivers cost-competitive services
- Manages consumers' expectations of product offerings

Account Manager

- Supports service managers in service planning, development, and deployment

Cloud Architect

- Creates detailed designs for the cloud infrastructure

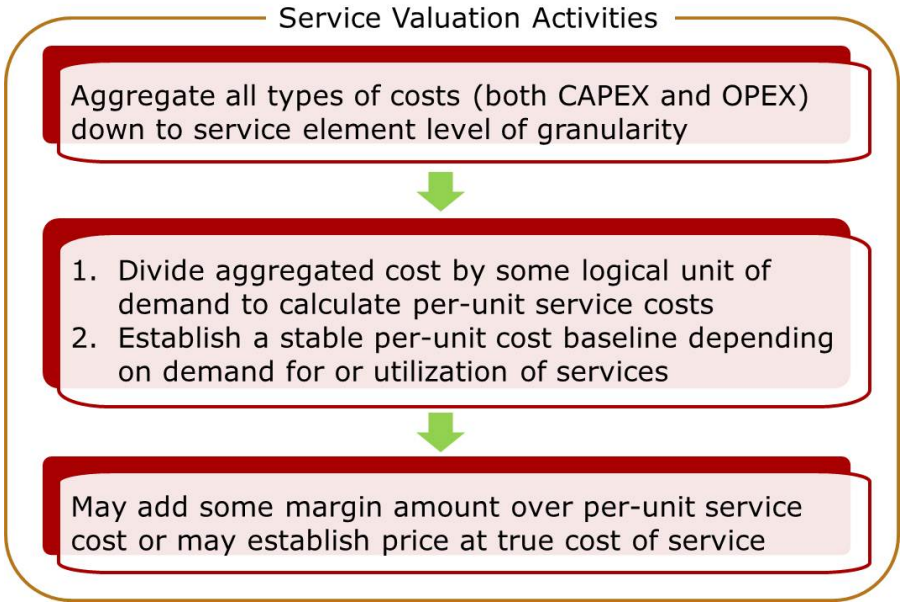
Service Operations Manager

- Streamlines service delivery and execution
- Coordinates with architecture team to define technology roadmaps and ensure SLOs are met

Finance

Service Valuation

- Determines the price (or chargeback) that a service consumer is expected to pay to meet the provider's business goal



- Recovery of cost
- Profit
- ROI goal
- Reinvestment goal

Finance

Chargeback Models

- Define how consumers need to pay for the consumed services

Model	Description
Pay-as-you-go	• Metering and pricing is based on consumption of resources • Consumers do not pay for unused resources
Subscription by time	• Cost of providing a service for a subscription period is divided among a predefined number of consumers
Subscription by peak usage	• Consumers are billed according to their peak usage of IT resources for a subscription period
Fixed cost or pre-pay	• Consumers commit needed resources upfront for committed period • Consumers pay fixed charge periodically through a billing cycle regardless of the utilization of resources
User-based	• Billing is based on the number of users logged in

Tools

- Tools play an important role in building a cloud infrastructure:
  - Virtualization and orchestration software
  - Security and business continuity software
  - Self-service portal software
- Other tools that should be considered specially when deploying hybrid cloud, community cloud, or brokerage service:
  - Cloud integration tools
  - Application Programing Interface (API)
  - Specialized connection
  - Transformation and business logic programs

Service-level Agreement and Legal Contract

Service-level Agreement

A contract negotiated between a provider and a consumer that specifies various parameters and metrics such as cost, service availability, maintenance schedules, performance levels, service desk response time, and consumer's and provider's responsibilities.

- Key points that must be included in a legal contract are:
  - Business level policies such as data privacy, data ownership, security, and jurisdiction
  - Availability and performance metrics
  - DR plan, exit plan, and penalties for not meeting SLA
  - How unexpected incidents and prolonged service outage will be handled

Avoid Vendor Lock-in

Vendor Lock-in

A situation where a consumer is unable to move readily from the current provider to another.

- Causes for vendor lock-in includes:
  - High migration cost
  - Application requires significant re-engineering for migration
  - Lack of open standards
  - Restrictions or burdensome penalties imposed by the current provider
- Vendor lock-in can be prevented by:
  - Using open standard tools, APIs, and file formats
  - Including appropriate exit clause in the agreement

Software Licensing Concerns

- Typically, relevant to IaaS and PaaS models
- Consumers can use their existing license if it is cloud enabled
- If consumer's existing license is not cloud enabled then:
  - Paying additional fees may get their license cloud enabled
  - May use software provided by the service provider
- Providers must work to understand the software license rights and its usage:
  - Prevents any non-compliance and violation of license agreements

Considerations for SaaS

- Software as a Service:
  - Ensures the software offered are thoroughly tested
  - Ensures the new features and functionalities are developed to the software to meet consumer's needs
  - Ensures applications are scalable and can handle increasingly larger consumer workloads
  - Ensures the applications are resilient and can withstand failures such as
    - Underlying component failure
    - Dependent service failure
  - Ensures the consumers are provided a secure environment

Considerations for PaaS and IaaS

- Platform as a Service:
  - Provides application development platform to the consumers
  - Supports large variety of OS, application development tools, and deployment tools
  - Ensures the consumers are provided a secure environment
  - Provides the consumer the required computing resources to operate the application
- Infrastructure as a Service:
  - Provides the consumer the required infrastructure resources to deploy their OS, application, and data
  - Ensures that the consumers are provided a secure environment

Migration

- Consumer may plan to migrate application or only data
- Two application migration strategies are:

Migration Strategy	Description
Forklift	• Entire application is migrated at once instead of in parts • Good for tightly coupled or self contained applications
Hybrid migration strategy	• Applications and its components are moved in parts • Lower-risk approach to migrate applications to the cloud • Good for application that have loosely coupled components

- For migrating data to cloud:
  - Consider copying data to cloud using replication technology
  - Consider factors such as network bandwidth, data security and integrity, and jurisdiction

Testing

DEFINE
- Define roles and responsibilities of personnel involved in testing

IDENTIFY
- Identify tools to perform test management and automation

DESIGN
- Design tests for data migration to cloud
- Design test cases to perform various testing modes
  - Stress, performance, functional, interoperability, and compatibility

TEST
- Test cloud capabilities committed by provider such as
  - Fault tolerance, disaster recovery, and security controls

Lesson Summary

During this lesson the following topics were covered:

- Governance and organization considerations
- Finance and tools considerations
- SLAs and vendor lock-in considerations
- Software and licensing considerations
- Considerations for service models
- Migration and testing considerations

Concepts in Practice

- Vblock
- EMC VSPEX

Vblock and EMC VSPEX

Vblock	VSPEX
• Integrated IT infrastructure solution for cloud deployment • Combines compute, storage, network, virtualization, security, and management software in a package • Validated solution and ready for deployment	• IT infrastructure solution for cloud deployment • Includes compute, storage, network, virtualization, and backup products • Offers choice of hypervisor, compute system, and network technology

Module Summary

Key points covered in this module:

- Cloud computing reference model
- Greenfield and brownfield deployment options
- Best-of-breed cloud infrastructure components
- Cloud-ready converged infrastructure
- Key factors to consider while building a cloud infrastructure

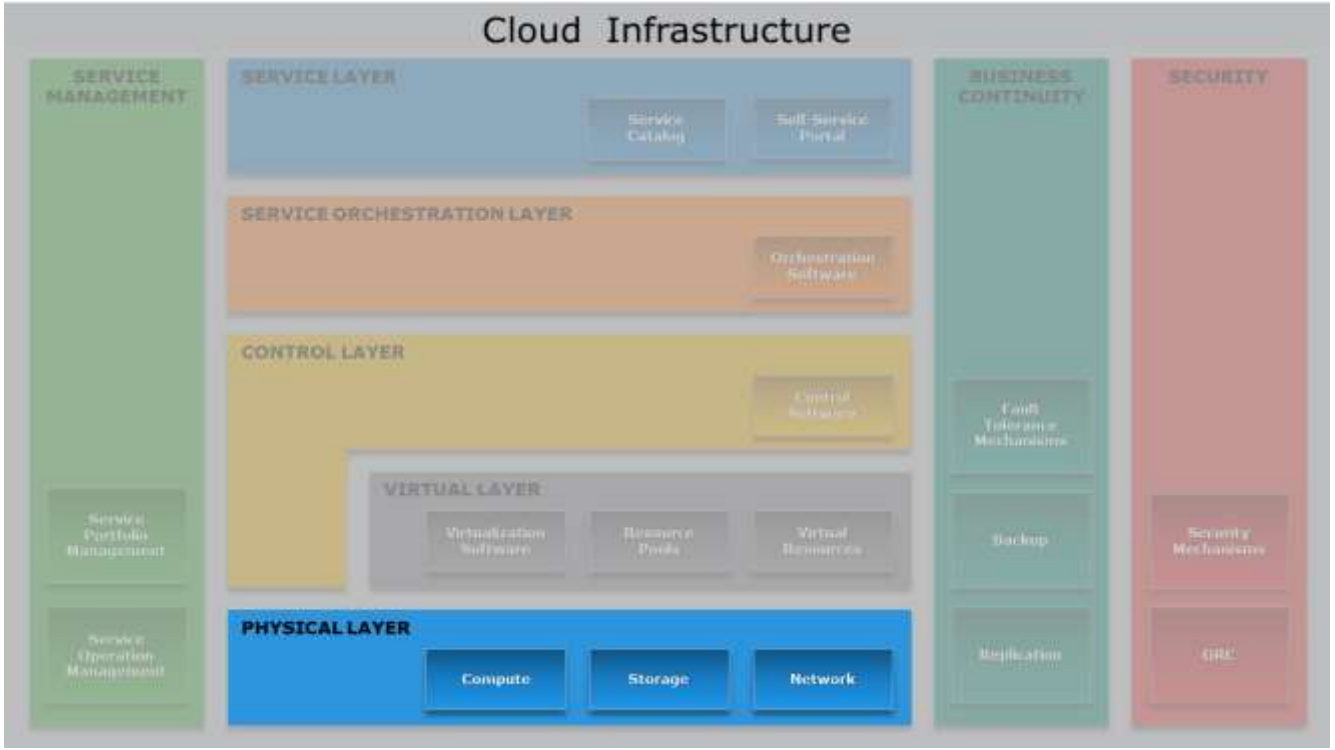
Module: Physical Layer

Upon completion of this module, you should be able to:

- Describe compute system components and types
- Describe storage system architectures
- Describe network connectivity and the types of network communication

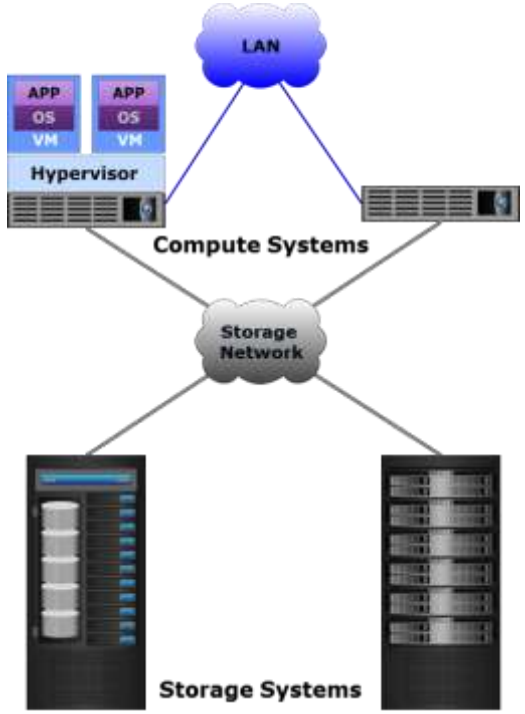
Cloud Computing Reference Model

Physical Layer



Physical Layer Overview

- The physical layer comprises physical compute, storage, and network resources
- Compute systems execute software of providers and consumers
- Storage systems store business and application data
- Networks connect compute systems with each other and with storage systems
  - Networks also connect multiple data centers or multiple clouds to one another



Lesson: Compute System

This lesson covers the following topics:

- Key components of a compute system
- Software deployed on compute systems
- Types of compute systems

Introduction to Compute System

- A computing platform (hardware, firmware, and software) that runs platform and application software
  - Executes the provider's as well as the consumers' software
  - Typically x86-based servers or hosts
- Compute systems are provided to consumers in two ways:
  - Shared hosting: Multiple consumers share compute systems
  - Dedicated hosting: Individual consumers have dedicated compute systems
- Typically providers use compute virtualization and offer compute systems in the form of virtual machines

Key Components of a Compute System

Processor	• An IC that executes software programs by performing arithmetical, logical, and input/output operations
Random-Access Memory	• A volatile data storage device containing the programs for execution and the data used by the processor
Read-Only Memory	• A semiconductor memory containing boot, power management, and other device-specific firmware
Motherboard	• A PCB that holds the processor, RAM, ROM, network and I/O ports, and other integrated components, such as GPU and NIC
Chipset	• A collection of microchips on a motherboard to manage specific functions, such as processor access to RAM and to peripheral ports

Software Deployed on Compute Systems

Self-service portal	• Enables consumers to view and request cloud services
Platform software	• Includes the software that the provider offers through PaaS
Application software	• Includes the applications that the provider offers through SaaS
Virtualization software	• Enables resource pooling and creation of virtual resources
Cloud management software	• Enables a provider to manage the cloud infrastructure and services
Consumer software	• Includes a consumer's platform software and business applications

Types of Compute Systems

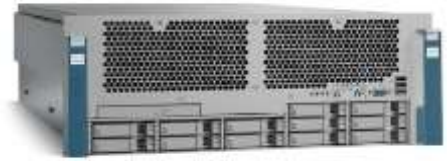
- Tower compute system
- Rack-mounted compute system
- Blade compute system

Tower Compute System

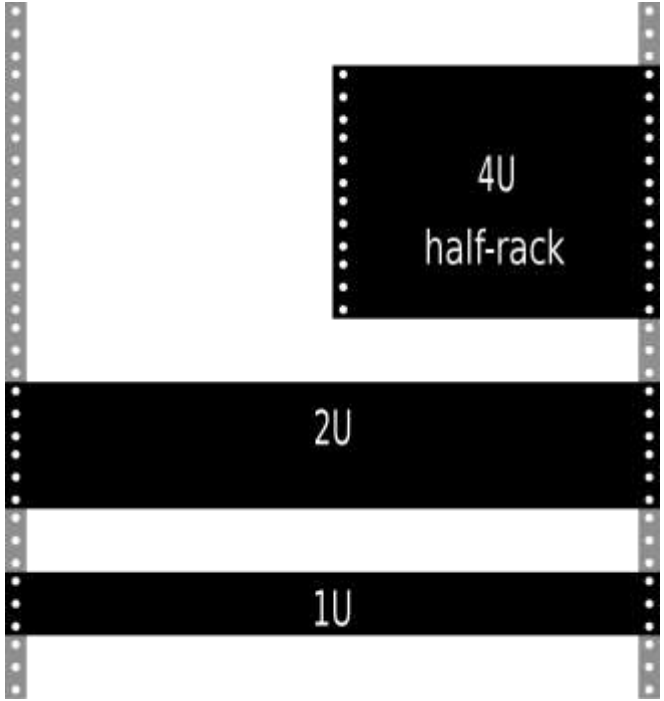
- Built in an upright enclosure called a “tower”
- Has integrated power supply and cooling
- A group of towers occupies significant floor space, requires complex cabling, and generates noise from cooling units
- Deploying in large environments may involve substantial expenditure

Rack-mounted Compute System

- Designed to be fixed on a frame called a "rack"
  - A rack is a standardized enclosure with mounting slots for vertically stacking compute systems
- Simplifies network cabling, consolidates network equipment, and reduces floor space use
- Administrators may use a console mounted on the rack to manage the compute systems



Rack unit



Blade Compute System

- Comprises an electronic circuit board with only the core processing components
- Multiple blades are housed in a blade chassis
  - The chassis provides integrated power supply, cooling, networking, and management
- Blades are interconnected via a high speed bus
- Modular design increases compute system density and scalability

Blade server and kvm conole

Lesson Summary

During this lesson the following topics were covered:

- Key components of a compute system
- Software deployed on compute systems
- Types of compute systems: tower, rack-mounted, and blade

Lesson: Storage System

This lesson covers the following topics:

- Types of storage devices
- Redundant Array of Independent Disks (RAID)
- Storage system architectures

Introduction to Storage System

- A storage system is the repository for saving and retrieving electronic data
- Providers offer storage capacity along with compute systems, or as a service
  - Storage as a Service enables data backup and long-term data retention
- Cloud storage provides massive scalability and rapid elasticity of storage resources
- Typically, a provider uses virtualization to create storage pools that are shared by multiple consumers

Types of Storage Devices

Magnetic disk drive	• Stores data on a circular disk with a ferromagnetic coating • Provides random read/write access • Most popular storage device with large storage capacity
Solid-state (flash) drive	• Stores data on a semiconductor-based memory • Very low latency per I/O, low power requirements, and very high throughput
Magnetic tape drive	• Stores data on a thin plastic film with a magnetic coating • Provides only sequential data access • Low-cost solution for long term data storage
Optical disc drive	• Stores data on a polycarbonate disc with a reflective coating • Write Once and Read Many capability: CD, DVD, BD • Low-cost solution for long-term data storage

Replace hard disk in san

Redundant Array of Independent Disks (RAID)

RAID

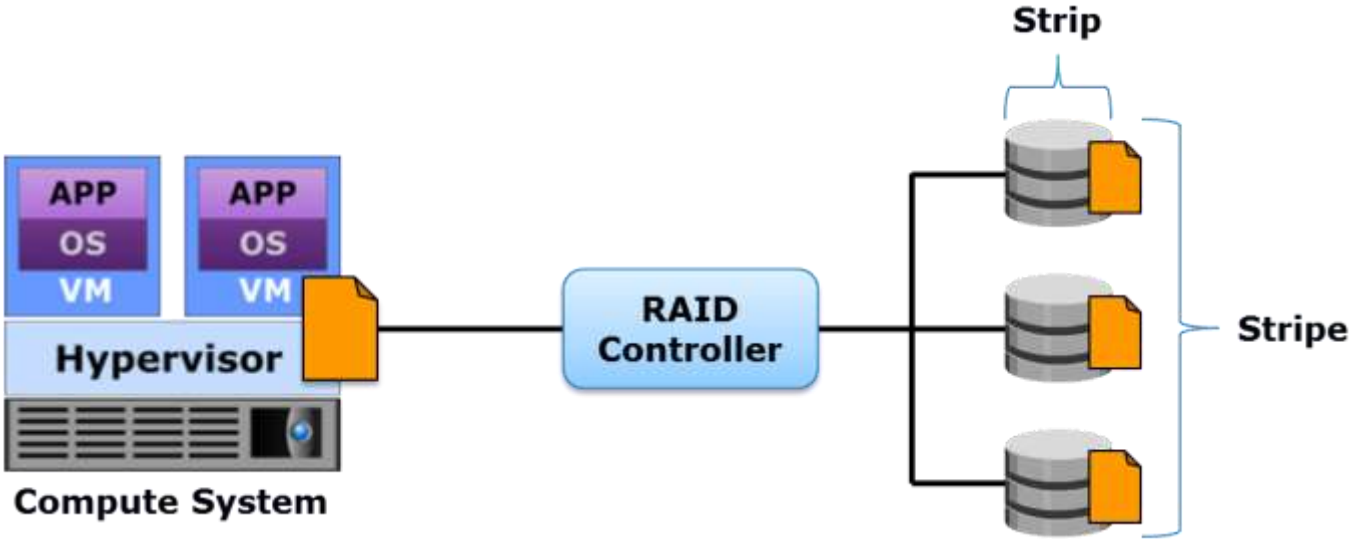
A storage technology in which data is written in blocks across multiple disk drives that are combined into a logical unit called a RAID group.

- Improves storage system performance by serving I/Os from multiple drives simultaneously
- Provides data protection against drive failures
- Three key techniques used for RAID: striping, mirroring, and parity

RAID Technique: Striping

Striping

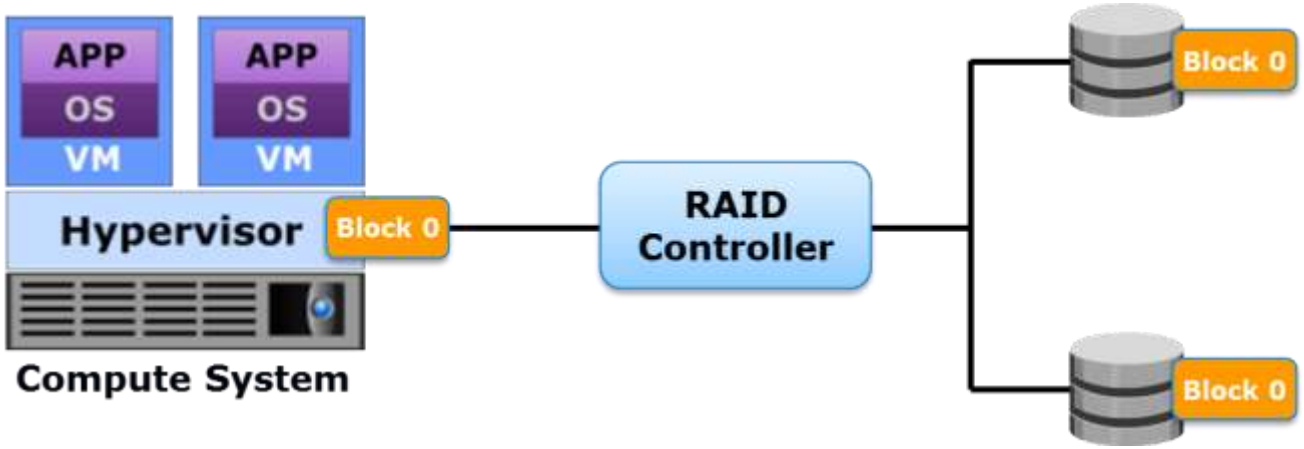
A RAID technique to spread data across multiple drives in order to use the drives in parallel.



RAID Technique: Mirroring

Mirroring

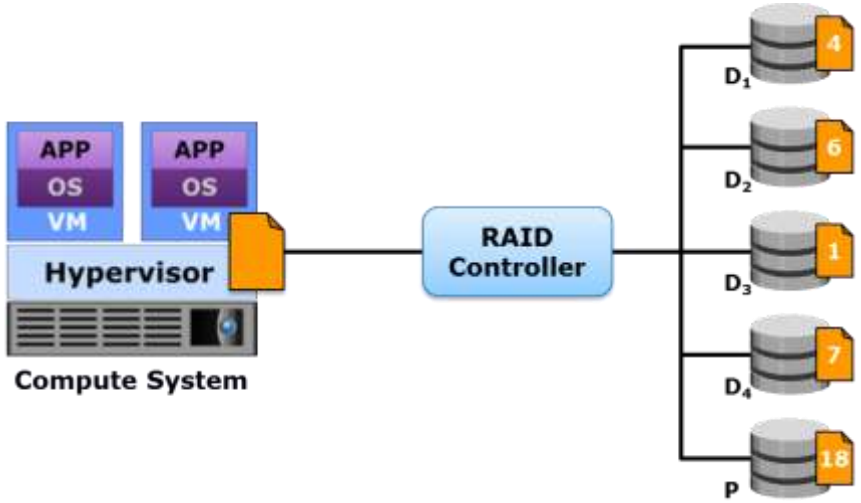
A RAID technique to store the same data simultaneously on two different drives, yielding two copies of the data.



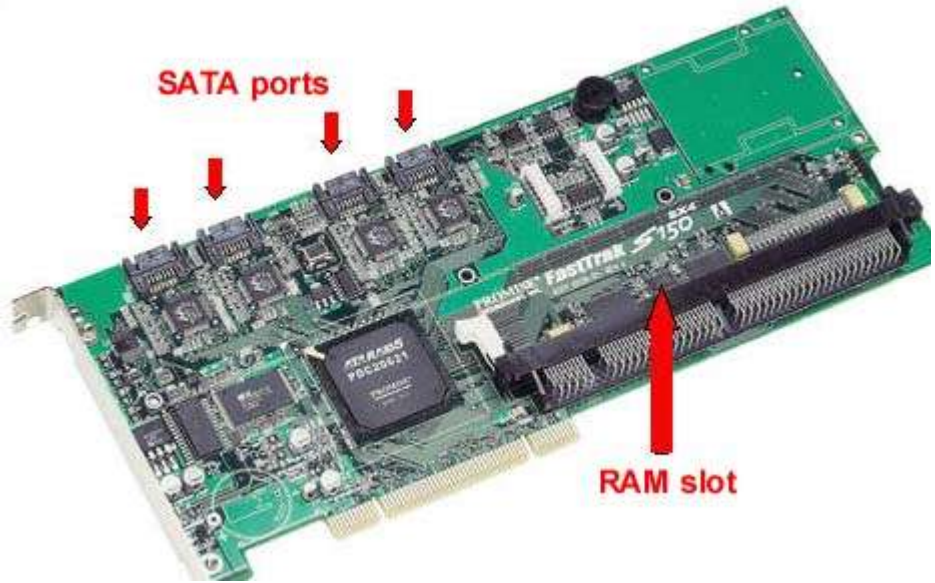
RAID Technique: Parity

Parity

A RAID technique to protect striped data from drive failure by performing a mathematical operation on individual strips and storing the result on a portion of the RAID group.



Raid controller

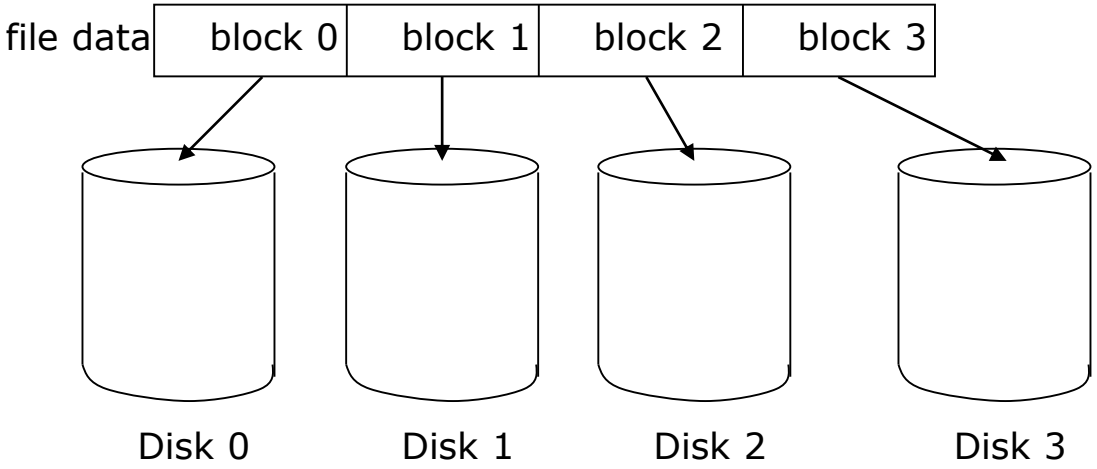


RAID

- Redundant Array of Inexpensive Disks
- Basic idea is to connect multiple disks together to provide
  - large storage capacity
  - faster access to reading data
  - redundant data
- Many different levels of RAID systems
  - differing levels of redundancy, error checking, capacity, and cost

Striping

- Take file data and map it to different disks
- Allows for reading data in parallel



Parity

Bitwise Xor Operations

Operation	Result
0 ^ 0	0
1 ^ 0	1
0 ^ 1	1
1 ^ 1	0

- Way to do error checking and correction
- Add up all the bits that are 1
  - if even number, set parity bit to 0
  - if odd number, set parity bit to 1
- To actually implement this, do an exclusive OR of all the bits being considered
- Consider the following 2 bytes

byte	parity
10110011	1
01101010	0

- If a single bit is bad, it is possible to correct it

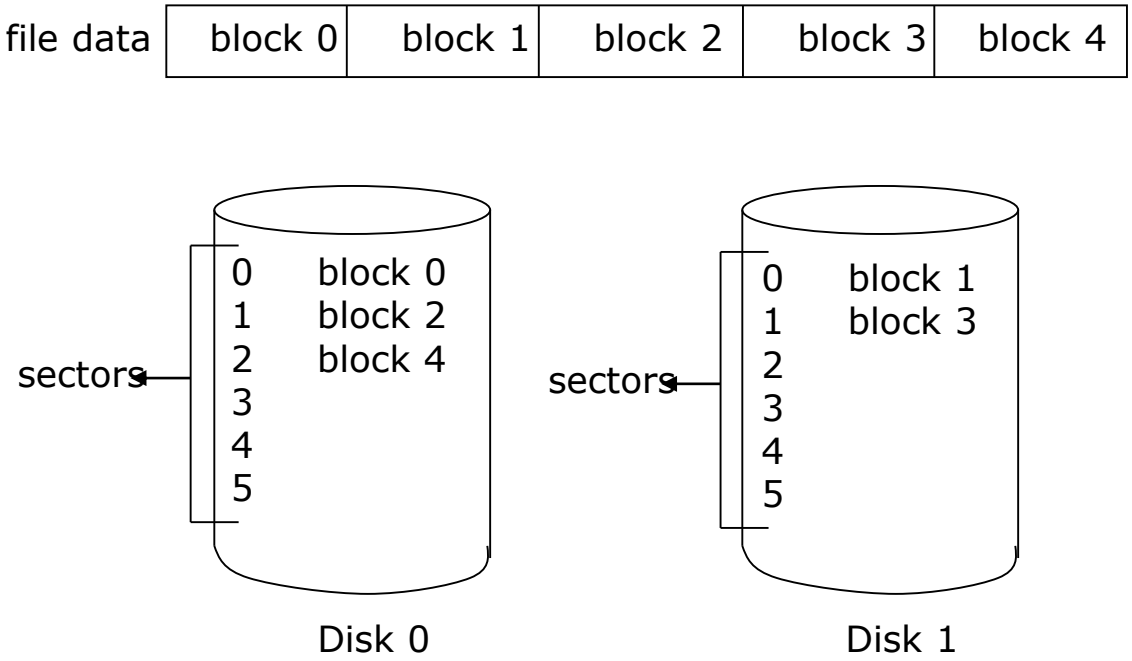
Mirroring

- Keep two copies of data on two separate disks
- Gives good error recovery
  - if some data is lost, get it from the other source
- Expensive
  - requires twice as many disks
- Write performance can be slow
  - have to write data to two different spots
- Read performance is enhanced
  - can read data from file in parallel

RAID Level-0

- Often called striping
- Break a file into blocks of data
- Stripe the blocks across disks in the system
- Simple to implement
  - disk = file block % number of disks
  - sector = file block / number of disks
- provides no redundancy or error detection
  - important to consider because lots of disks means low Mean Time To Failure (MTTF)

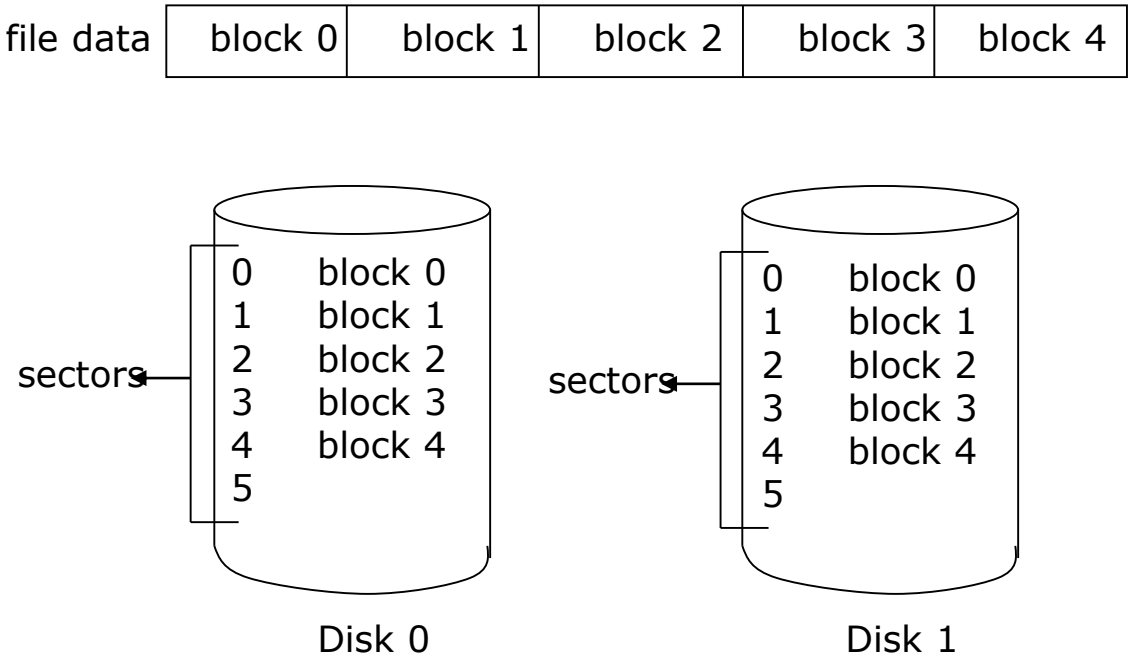
RAID Level-0



RAID Level-1

- A complete file is stored on a single disk
- A second disk contains an exact copy of the file
- Provides complete redundancy of data
- Read performance can be improved
  - file data can be read in parallel
- Write performance suffers
  - must write the data out twice
- Most expensive RAID implementation
  - requires twice as much storage space

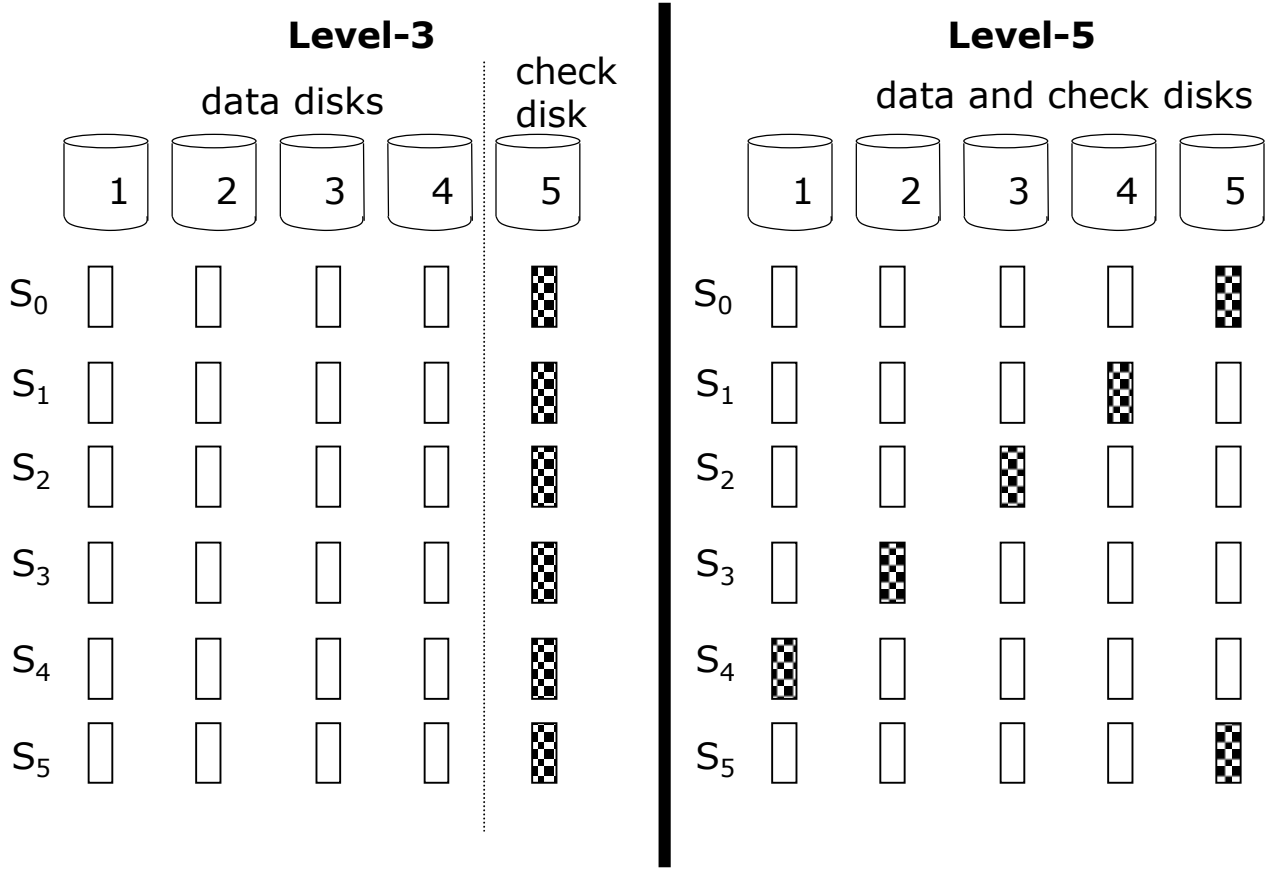
RAID Level-1



RAID Level-5

- Level-5 stripes file data and check data over all the disks
  - no longer a single check disk
  - no more write bottleneck
- Drastically improves the performance of multiple writes
  - they can now be done in parallel
- Slightly improves reads
  - one more disk to use for reading

RAID Level-5



RAID Level-10

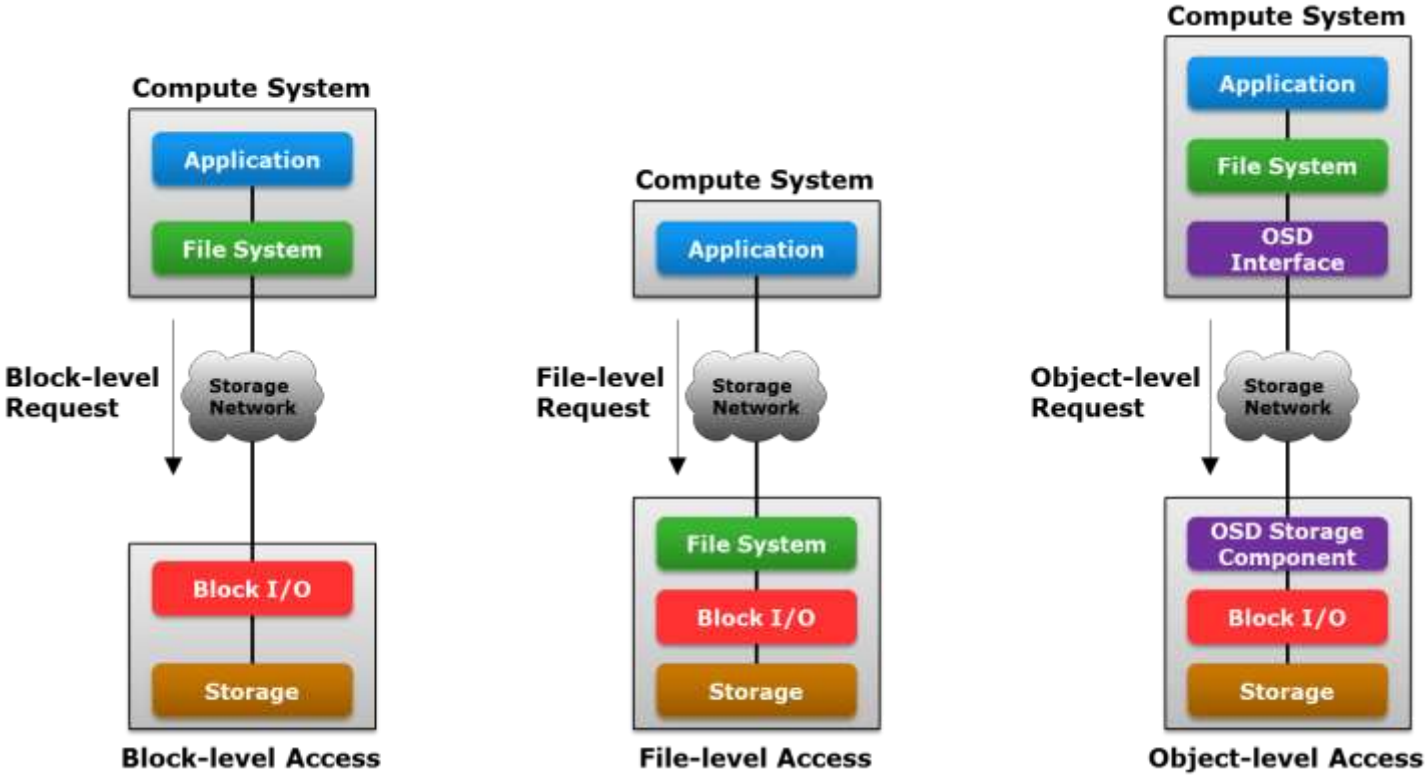
- Combine Level-0 and Level-1
- Stripe a files data across multiple disks
  - gives great read/write performance
- Mirror each strip onto a second disk
  - gives the best redundancy
- The most high performance system
- The most expensive system

Common RAID Levels

RAID 0	• Striped set with no fault tolerance
RAID 1	• Disk mirroring
RAID 1+0	• Nested RAID (striping and mirroring)
RAID 3	• Striped set with parallel access and a dedicated parity disk
RAID 5	• Striped set with independent disk access and distributed parity
RAID 6	• Striped set with independent disk access and dual distributed parity

Raid 5 and 6

Data Access Methods

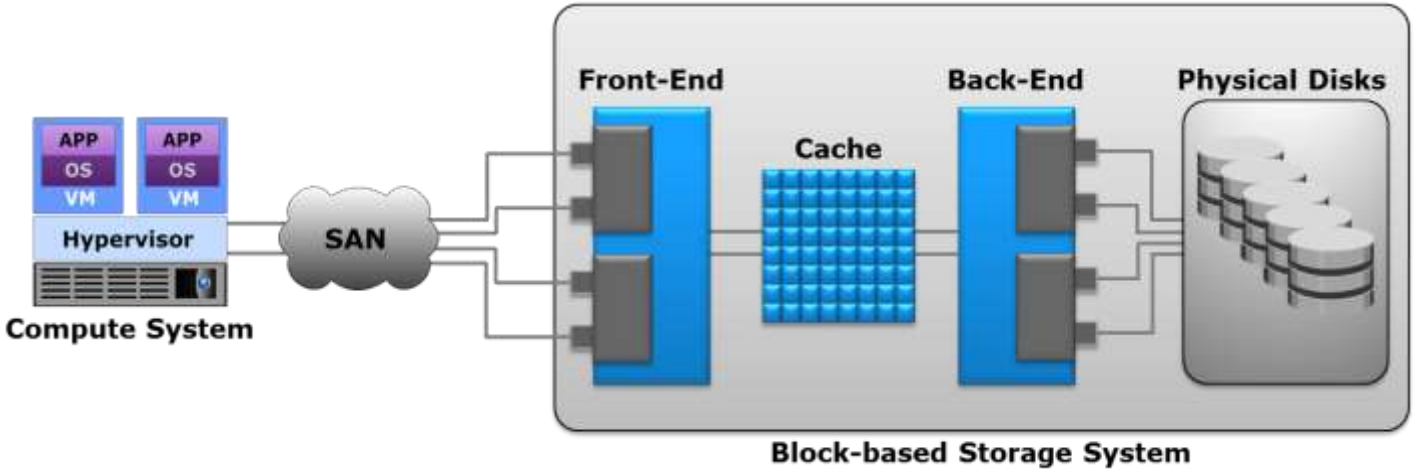


Storage System Architecture

- Storage system architectures are based on the data access methods
- Common storage system options are:
  - Block-based
  - File-based
  - Object-based
  - Unified

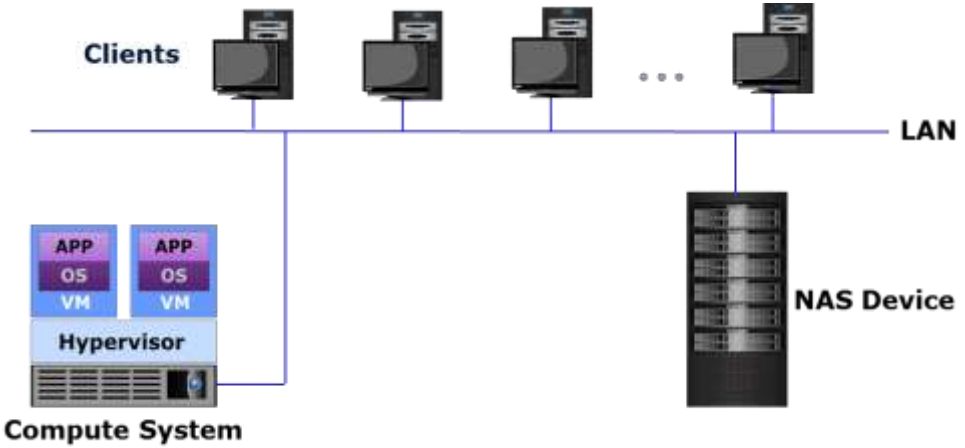
Block-based Storage System

- Enables creating and assigning storage volumes to compute systems
  - Compute system discovers the volumes as local drives
  - Required file system can be created on the volumes

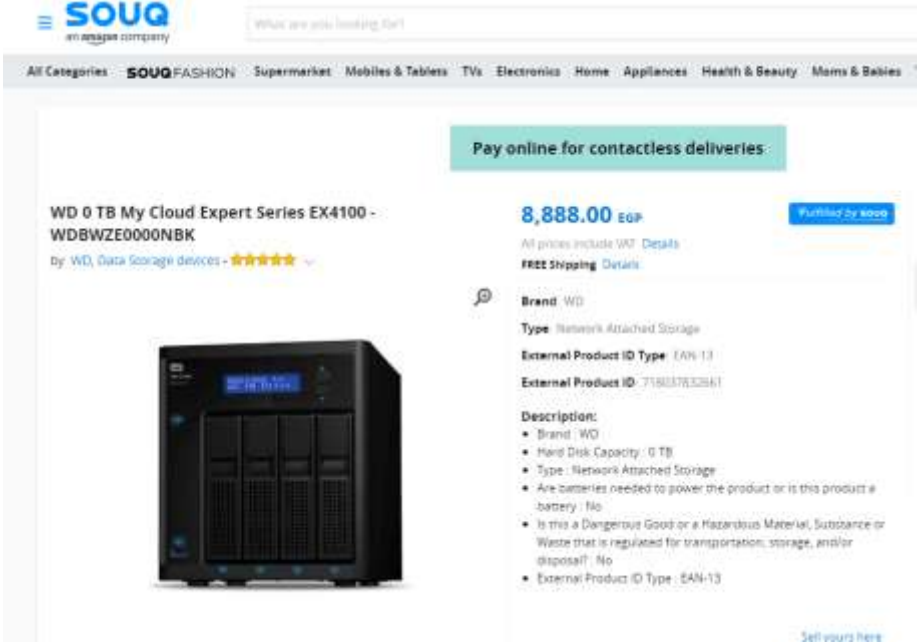


File-based Storage System

- A dedicated, high performance file server with storage (also known as Network-attached Storage)
- Enables clients to share files over an IP network
  - Supports data sharing for UNIX and Windows users
- Uses a specialized OS that is optimized for file I/O



NAS storage example

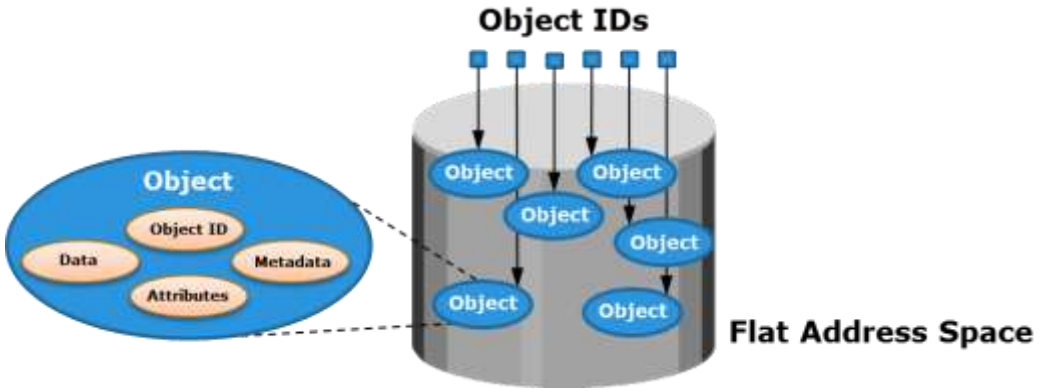


NAS Deployment Options

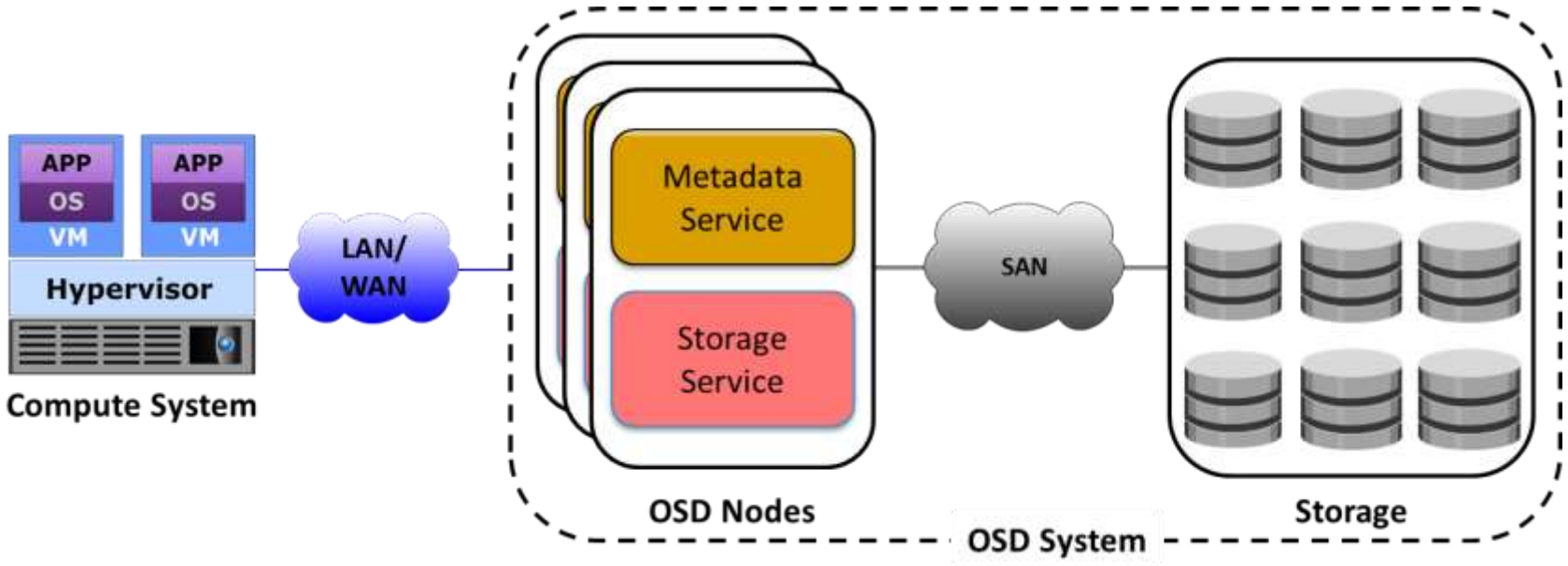
- The two common NAS deployment options are:
  - Traditional NAS (scale-up NAS)
  - Scale-out NAS
- Traditional NAS
  - Capacity and performance of a single system is scaled by upgrading or adding NAS components
- Scale-out NAS
  - Multiple processing and storage nodes are pooled in a cluster that works as a single NAS device
  - Addition of nodes scales cluster capacity and performance without disruption

Object-based Storage System

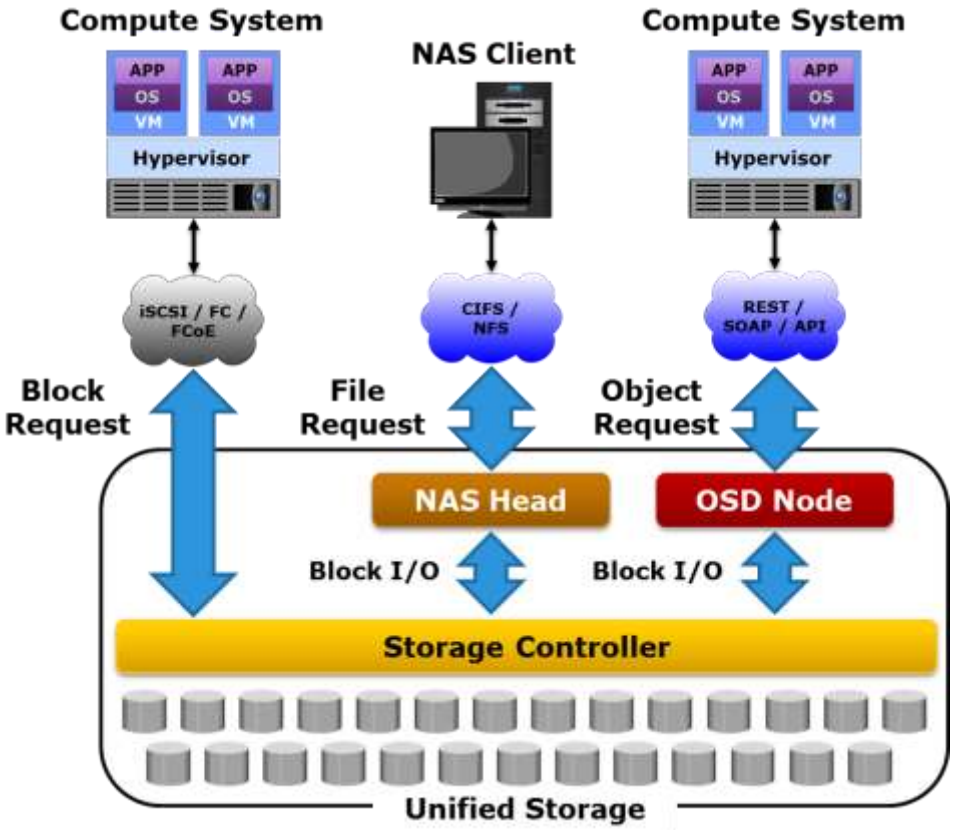
- Stores file data in the form of objects based on data contents and attributes
  - Uses a flat, non-hierarchical address space
- Object contains user data, related metadata, and user-defined attributes
  - Objects are uniquely identified using object ID



Object-based Storage System (Cont'd)



Unified Storage System



Lesson Summary

During this lesson the following topics were covered:

- Types of persistent data storage devices
- RAID and RAID techniques: striping, mirroring, and parity
- Storage system architectures: block-based, file-based, object-based, and unified

Lesson: Network

This lesson covers the following topics:

- Types of network communication
- Compute-to-compute communication
- Compute-to-storage communication
- Storage area network (SAN) classification
- Inter-cloud communication

Introduction to Networking

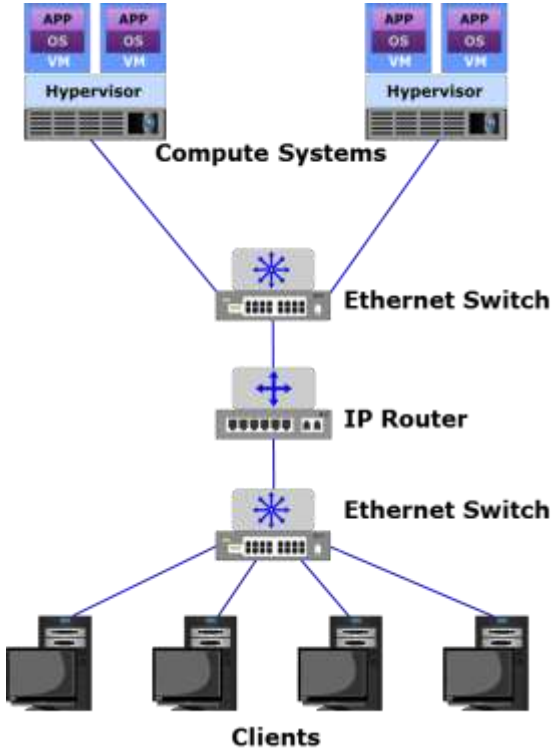
- Networking enables data transfer and sharing of IT resources between nodes across geographic regions
- Cloud consumers require a reliable and secure network to connect to a cloud and access cloud services
- Network connectivity also enables resource aggregation and service mobility across cloud data centers
- Multiple clouds may be inter-connected to enable workloads to be moved or distributed
  - For example: cloud bursting in a hybrid cloud model

Types of Network Communication

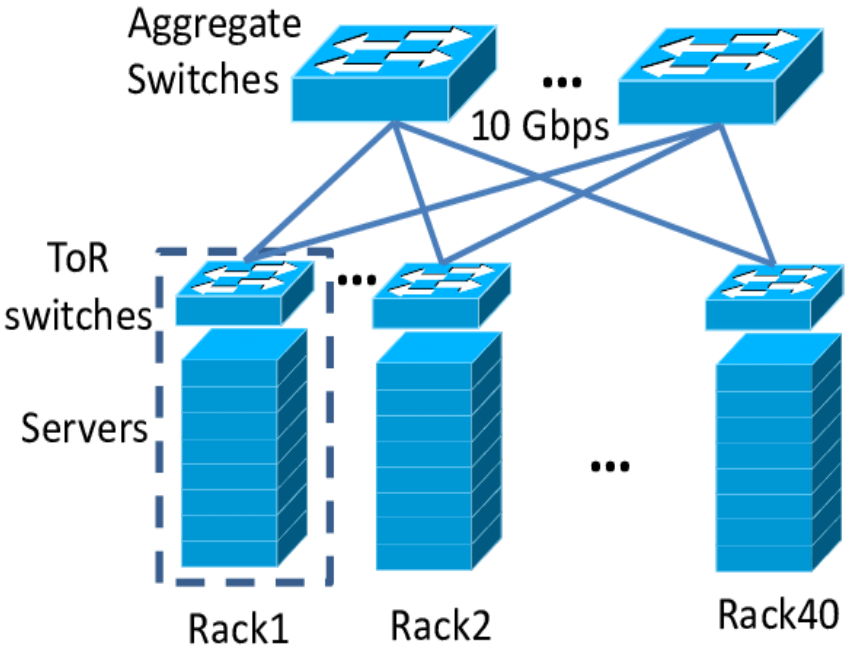
- Based on the nodes connected by a network, the network communication is broadly categorized as:
  - Compute-to-compute communication
  - Compute-to-storage communication
  - Inter-cloud communication

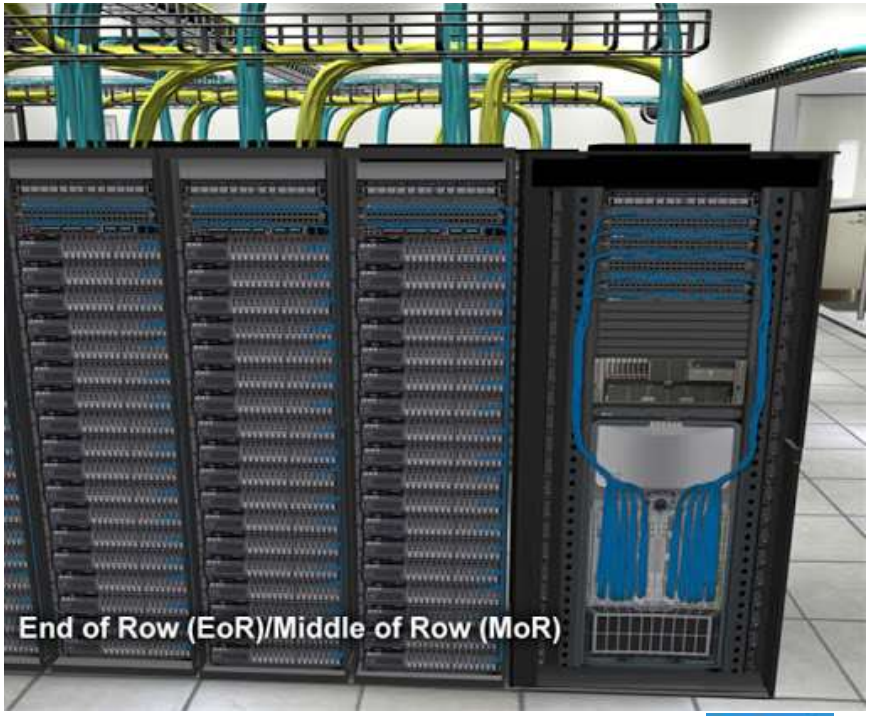
Compute-to-compute Communication

- Interconnecting physical compute systems enables compute-to-compute communication
- Compute-to-compute communication typically uses IP-based protocols
- Compute systems connect to a network through physical network card(s)
- Physical switches and routers are common interconnecting devices



Sample diagram



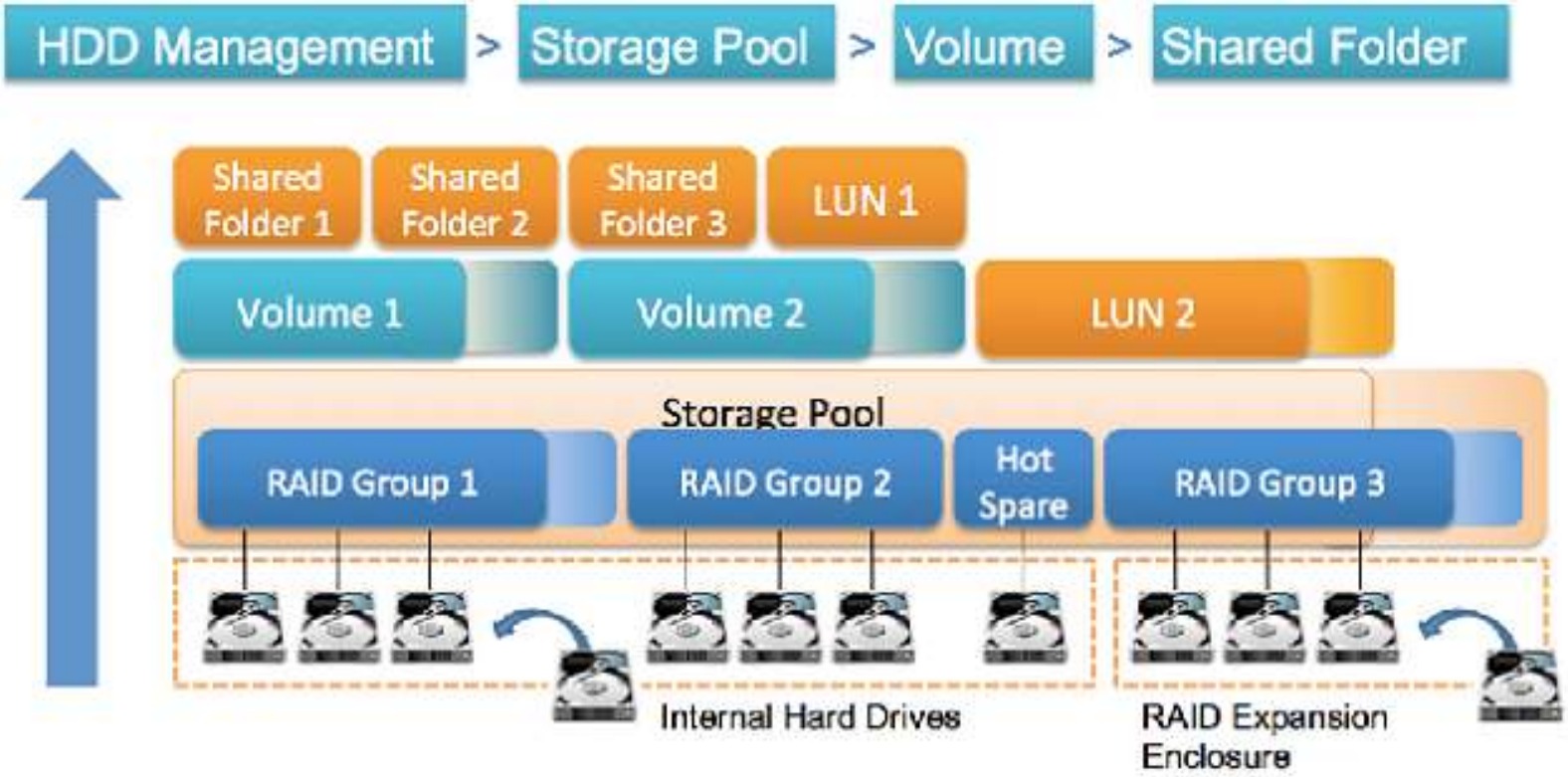


Compute-to-storage Communication

Storage Area Network (SAN)

A network that interconnects storage systems with compute systems, enabling the compute systems to access and share the storage systems.

- Based on the protocols they support, SANs can be classified as:
  - Fibre Channel SAN (FC SAN)
  - Internet Protocol SAN (IP SAN)
  - Fibre Channel over Ethernet SAN (FCoE SAN)

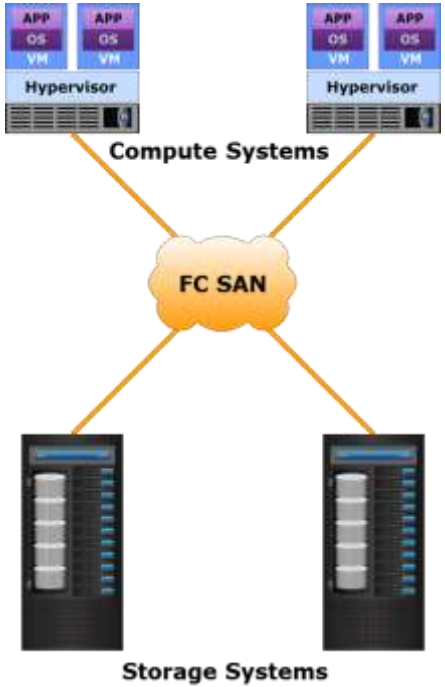


FC SAN

FC SAN

A SAN that uses Fibre Channel (FC) protocol to transport data, commands, and status information between compute and storage systems.

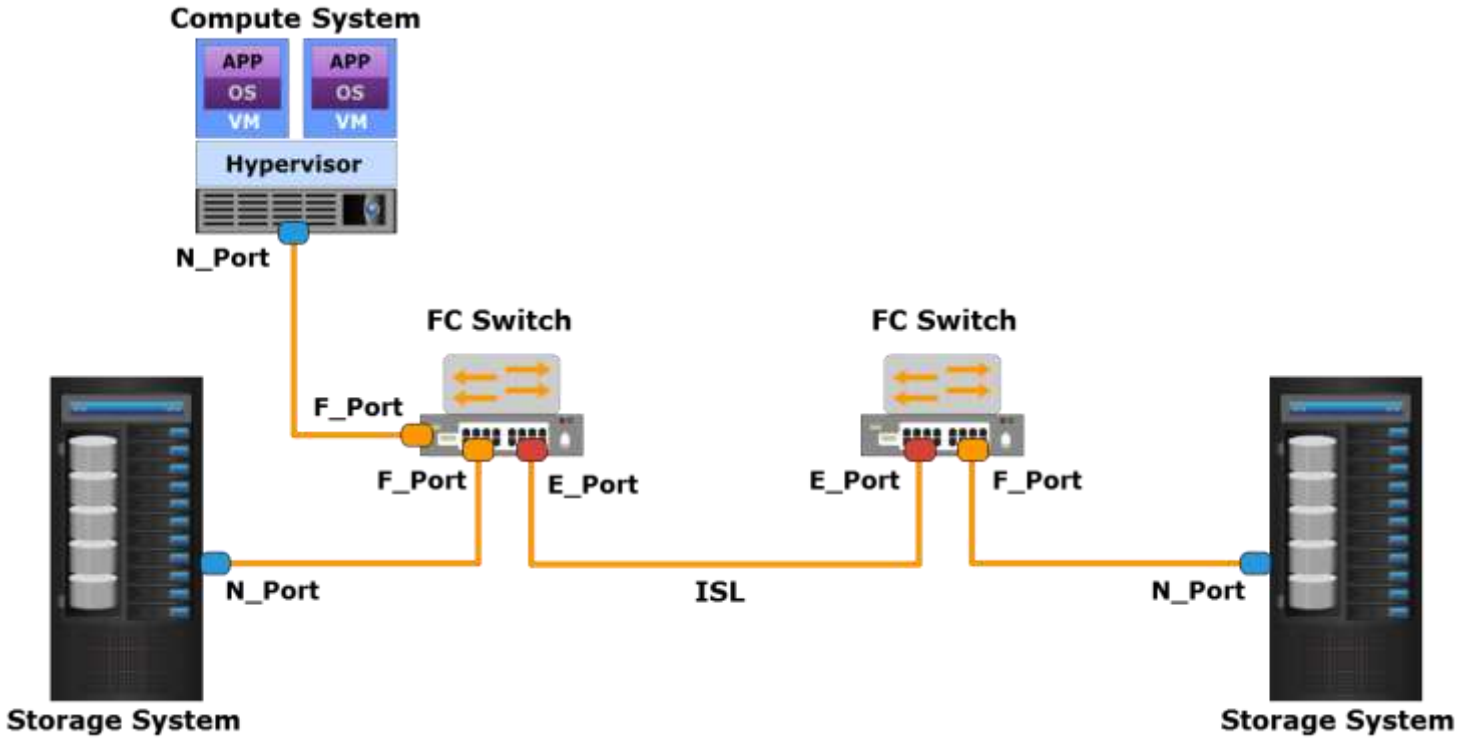
- FC provides block-level access to storage
- FC offers data transfer speeds up to 128 Gbps
- Theoretically, an FC SAN can connect approximately 15 million nodes



FC SAN Components

Network adapters	• Provide physical interface to a node for communicating with other nodes • Examples: FC HBAs and storage system front-end adapters
Cables and connectors	• Optical fiber cables are predominantly used to provide connectivity • Connectors enable cables to be swiftly connected to and disconnected from ports
Interconnecting devices	• FC switches and directors • Directors have a modular design, a higher port count, and better fault-tolerance • Switches either have a fixed port count or a modular design

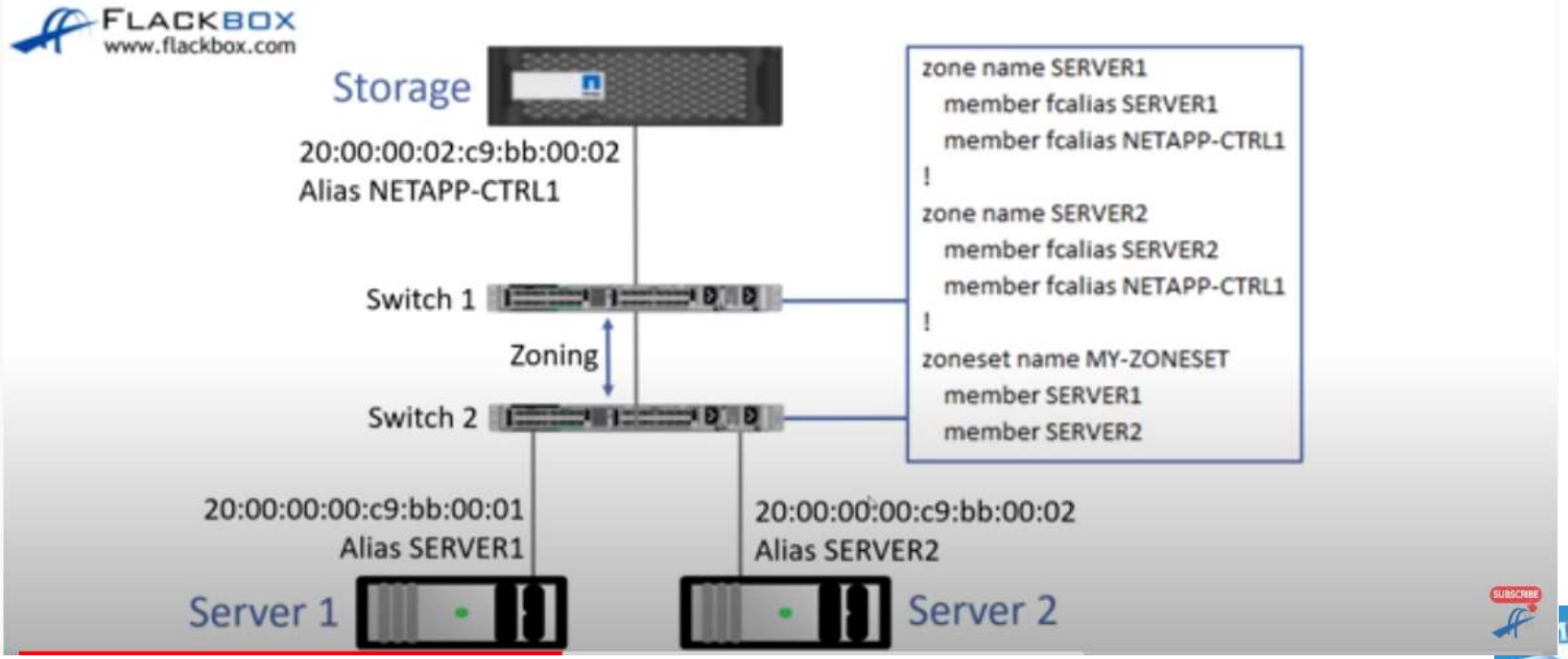
Fabric Port Types



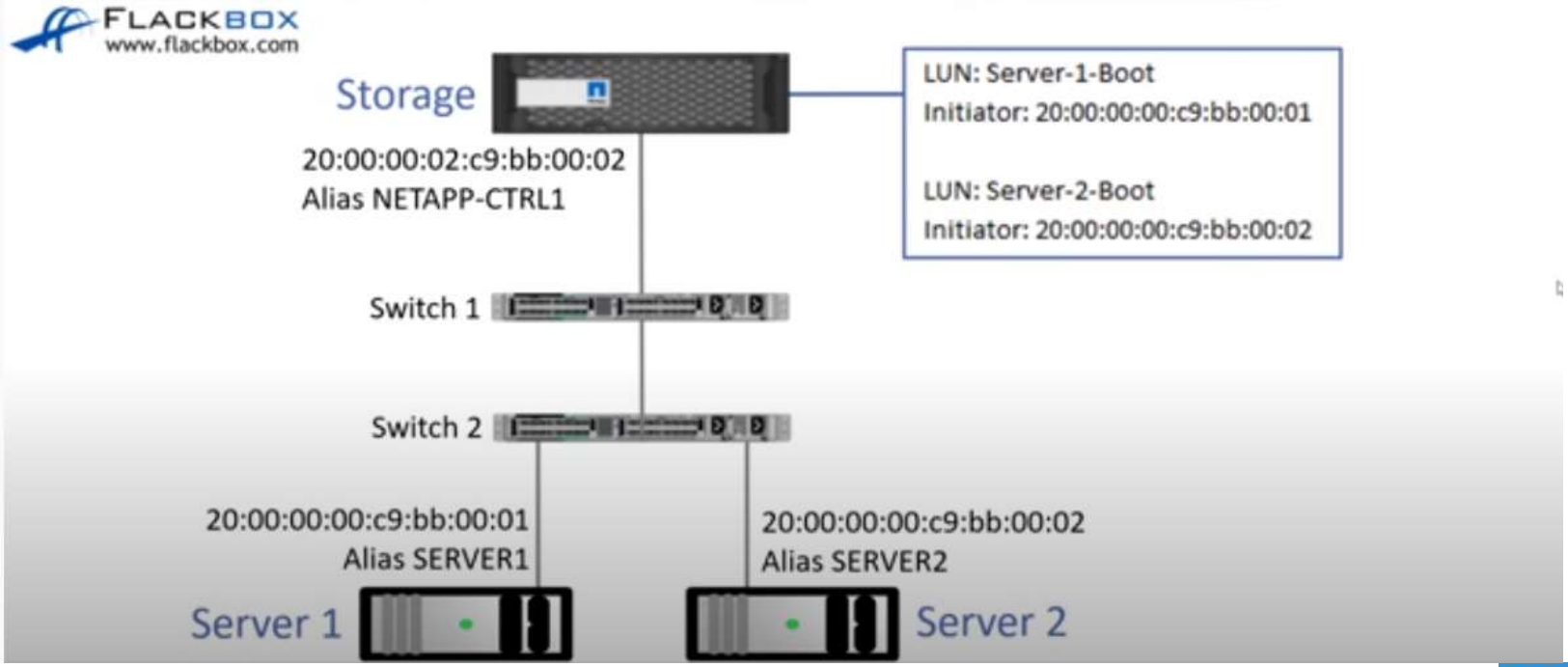
Fabric Connect and Addressing

- A fabric created with FC switches connects all nodes and enables them to communicate
- Each switch in a fabric contains a unique domain identifier (ID)
- Each network adapter is physically identified by a 64-bit World Wide Node Name (WWNN)
- Each adapter port is physically identified by a 64-bit World Wide Port Name (WWPN)
- Each adapter port in a fabric has a unique 24-bit FC address
  - Fabric assigns FC addresses to adapter ports dynamically

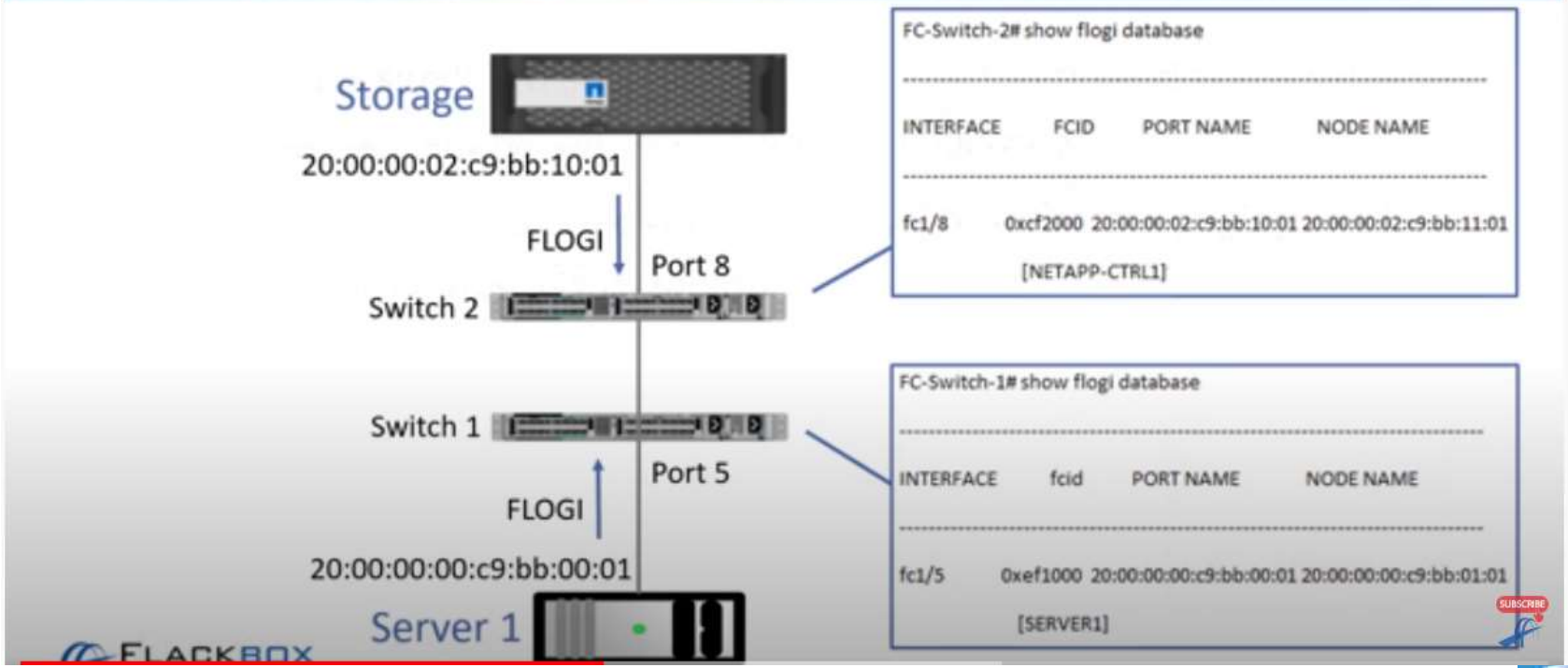
Zoning



LUN Masking



The Fabric Login



The Process Login

Zoning

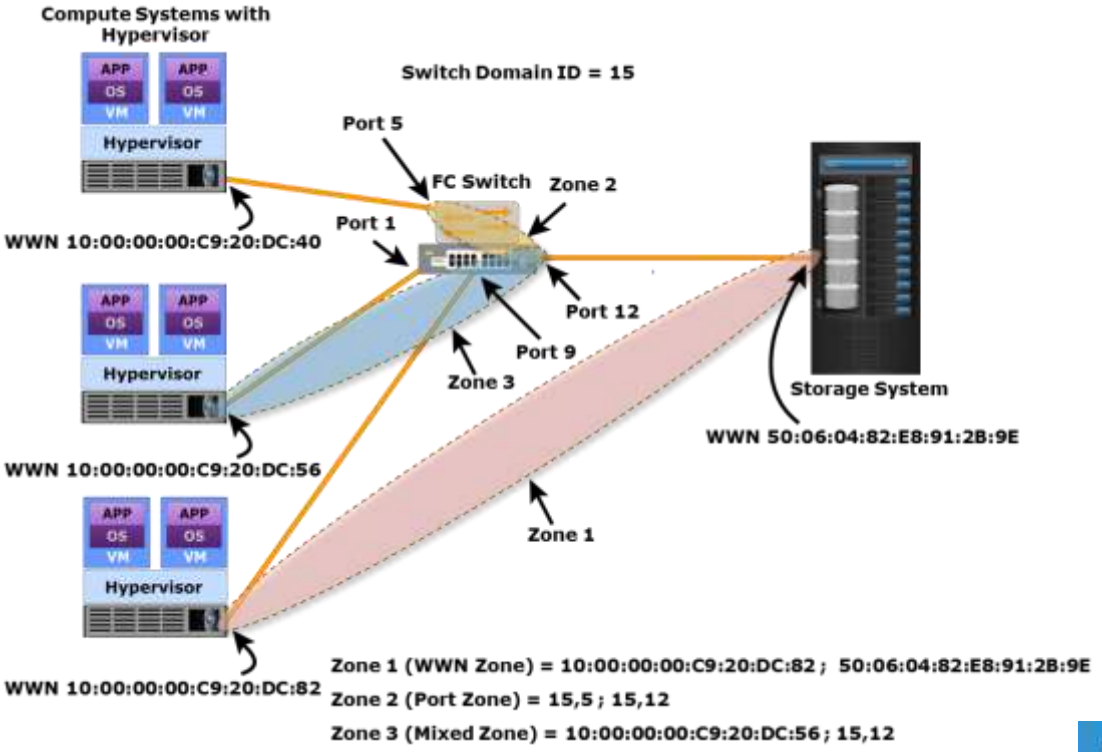
Zoning

An FC switch function that enables node ports within a fabric to be logically segmented into groups and to communicate with each other within the group.

- Both node ports and switch ports can be zone members
- Benefits:
  - Provides access control
  - Restricts RSCN traffic

Types of Zoning

- WWN zoning
- Port zoning
- Mixed zoning



IP SAN

IP SAN

A SAN that uses Internet Protocol (IP) for the transport of storage traffic. It transports block I/O over an IP-based network.

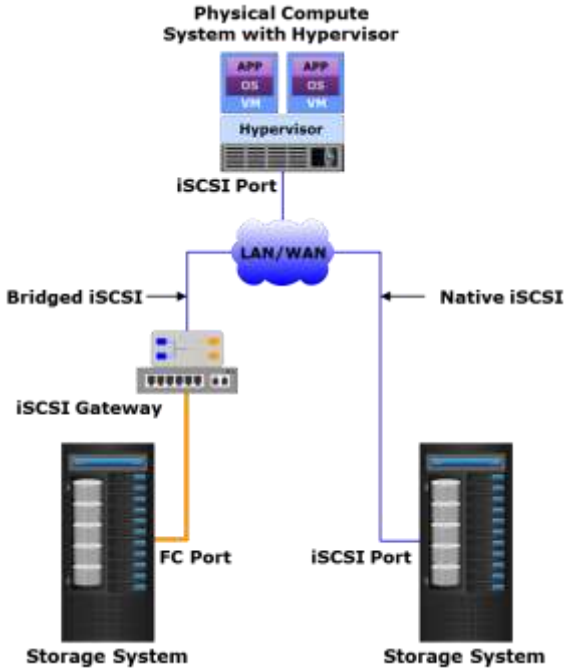
- Key drivers of IP SAN are:
  - Leveraging an existing IP-based network instead of building a new FC SAN infrastructure
  - Many robust, mature security options are available for IP networks
  - Many long-distance, disaster recovery (DR) solutions already leverage IP-based networks
- Two primary IP SAN protocols are: iSCSI and FCIP

iSCSI Networking

iSCSI

iSCSI encapsulates SCSI commands and data into IP packets that are transported over an IP-based network.

- iSCSI network components are:
  - iSCSI initiators
    - Example: iSCSI HBA
  - iSCSI targets
    - Example: storage system with iSCSI port (Native iSCSI)
    - Example: iSCSI gateway (Bridged iSCSI)
  - IP-based network



iSCSI Name

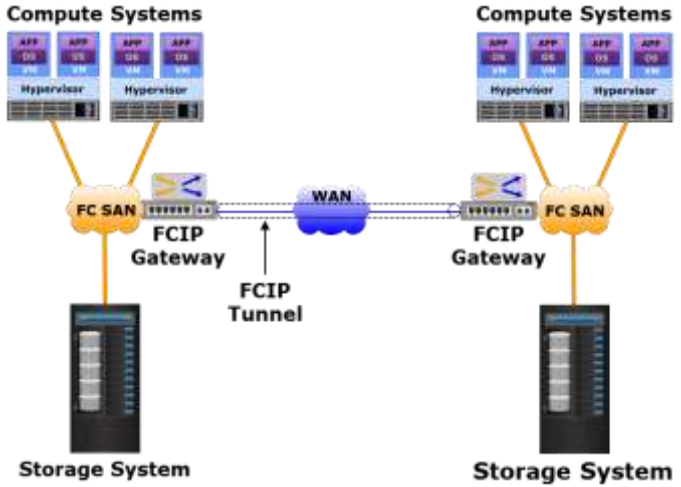
- iSCSI name is a unique iSCSI identifier that identifies initiators and targets in an iSCSI network
- The two common types of iSCSI names are:
  - iqn: iSCSI Qualified Name
    - Example: iqn.2014-02.com.example:*optional_string*
  - eui: Extended Unique Identifier
    - Example: eui.0300732A32598D26

FCIP Networking

FCIP

FCIP is an encapsulation of FC frames into IP packets that are transported between disparate FC SANs over an IP-based network through FCIP tunnel.

- An FCIP entity (e.g. FCIP gateway) exists at either end of an FCIP tunnel
  - Encapsulates FC into IP
  - Transfers IP packets to remote gateway
  - Decapsulates FC from IP
- Widely used in disaster recovery implementations



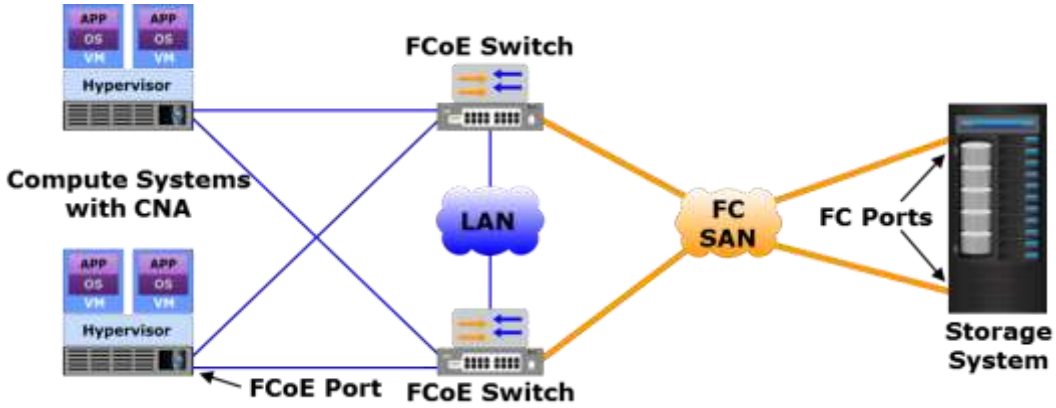
FCoE SAN

FCoE SAN

A converged enhanced Ethernet (CEE) network that uses the FCoE protocol to transport FC data along with regular Ethernet traffic over high speed Ethernet links. FCoE encapsulates FC frames into Ethernet frames.

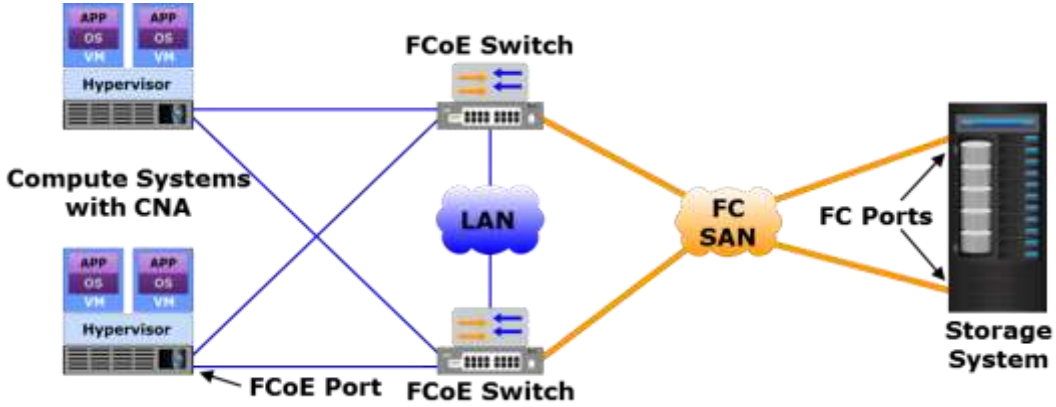
- Transfers both compute-to-compute and FC storage traffic using the same network components
  - Reduces complexity of managing multiple discrete networks
  - Reduces the number of adapters, cables, and switches, along with power and space consumption required in a data center
- Based on an enhanced Ethernet standard that ensures lossless transmission of FC traffic over Ethernet

FCoE SAN Components: CNA and S/W FCoE Adapter



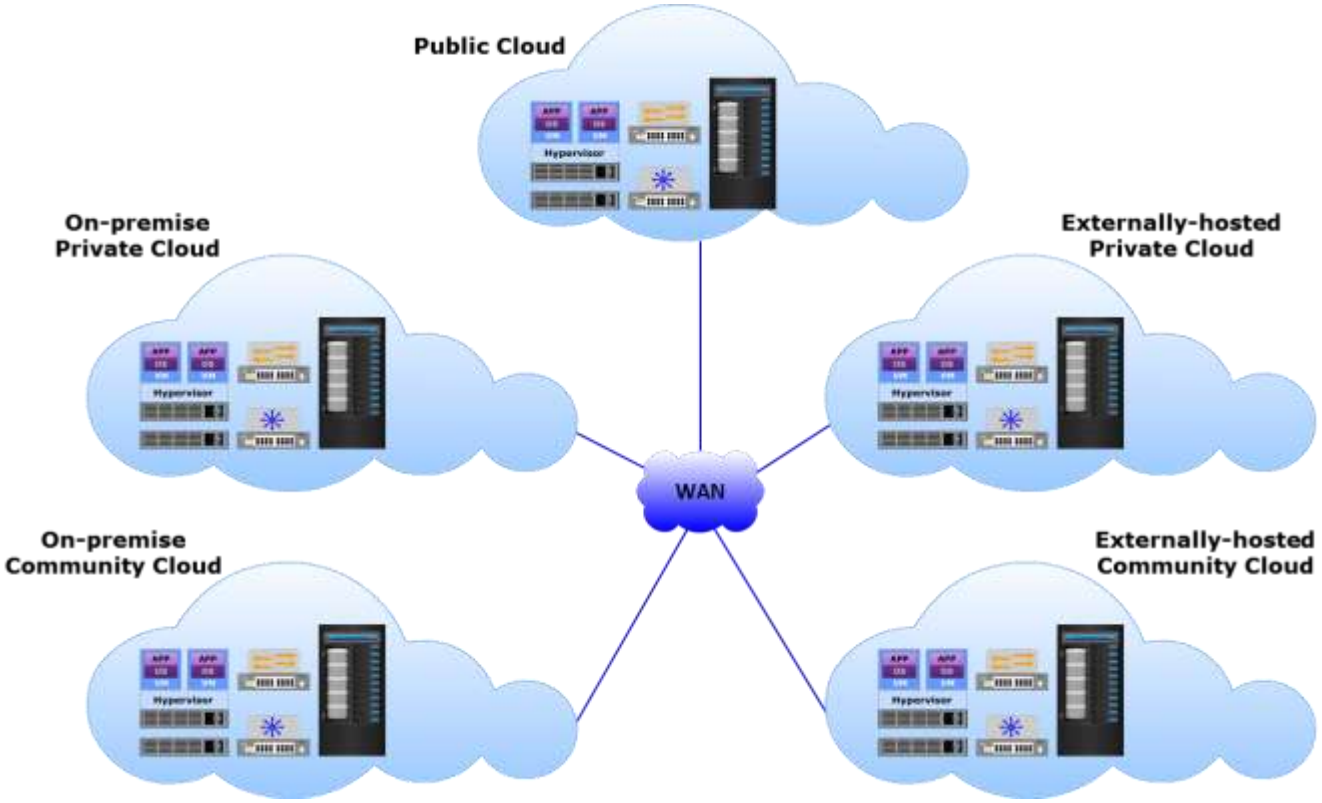
Component	Description
Converged network adapter (CNA)	• Provides functionality of both NIC and FC HBA in a single device • Encapsulates FC traffic onto Ethernet frames (FCoE traffic) • Consolidates both FC and regular Ethernet traffic over CEE links
Software FCoE adapter	• A software on the compute system performs FCoE processing • Supported NICs transfer both FCoE and regular Ethernet traffic

FCoE SAN Components: FCoE Switch and Storage Port



Component	Description
FCoE switch	• Contains Fibre Channel Forwarder (FCF), Ethernet Bridge, and a set of ports for FC, Ethernet, or FCoE connectivity • FCF encapsulates FC frames into Ethernet frames (FCoE frames) and decapsulates FCoE frames to FC frames
FCoE storage port	• Connects to FCoE switch, enabling end-to-end FCoE environment

Inter-cloud Communication



Lesson Summary

During this lesson the following topics were covered:

- Types of network communication
- Compute-to-compute communication
- Compute-to-storage communication (SAN)
- FC SAN, IP SAN, and FCoE SAN components and architectures
- Inter-cloud communication

Concepts in Practice

- EMC VMAX
- EMC VNX
- EMC ECS Appliance
- EMC Isilon
- EMC Atmos
- EMC XtremIO
- EMC Connectrix

EMC VMAX, EMC VNX, and EMC ECS Appliance

VMAX	VNX	ECS Appliance
• Family of high-end enterprise storage platforms • Block-based storage systems for mission-critical applications • High performance, reliability, availability, and scalability	• Family of unified storage platforms - Consolidates block, file, and object access • Built for SMBs and enterprises • Supports file (NFS and CIFS), FC, iSCSI, and FCoE access	• Hyper-scale storage infrastructure • Supports block, file, object, and HDFS • Provides multi-tenancy, self-service portal, and metering capabilities

EMC Isilon and EMC Atmos

Isilon	Atmos

• Scale-out NAS storage platform
• Enables pooling multiple nodes to construct a clustered NAS system
• OneFS operating environment creates single file system across the cluster
	• Scale-out object-based cloud storage platform
	  - Stores data as objects
	• Seamless scale out
	• Key cloud features include:
	  - Global namespace
	  - REST API-driven storage
	  - Multi-tenancy, metering, and self-service across tenants
	  - Metering and chargeback

EMC XtremIO and EMC Connectrix

XtremIO

- All-flash, block-based, scale-out enterprise storage array
- Uses a clustered design to grow capacity and performance as required
- A powerful OS (XIOS) manages the storage cluster
- Simplified and efficient provisioning and management

Connectrix

- Family of networked storage connectivity products including:
  - Enterprise directors
  - Departmental switches
  - Multi-purpose switches
- Multi-purpose switches support FC, iSCSI, FCIP, and FCoE protocols

Module Summary

Key points covered in this module:

- Compute system components and types
- Types of storage devices, RAID techniques, and storage system architectures
- Network connectivity and the types of network communication

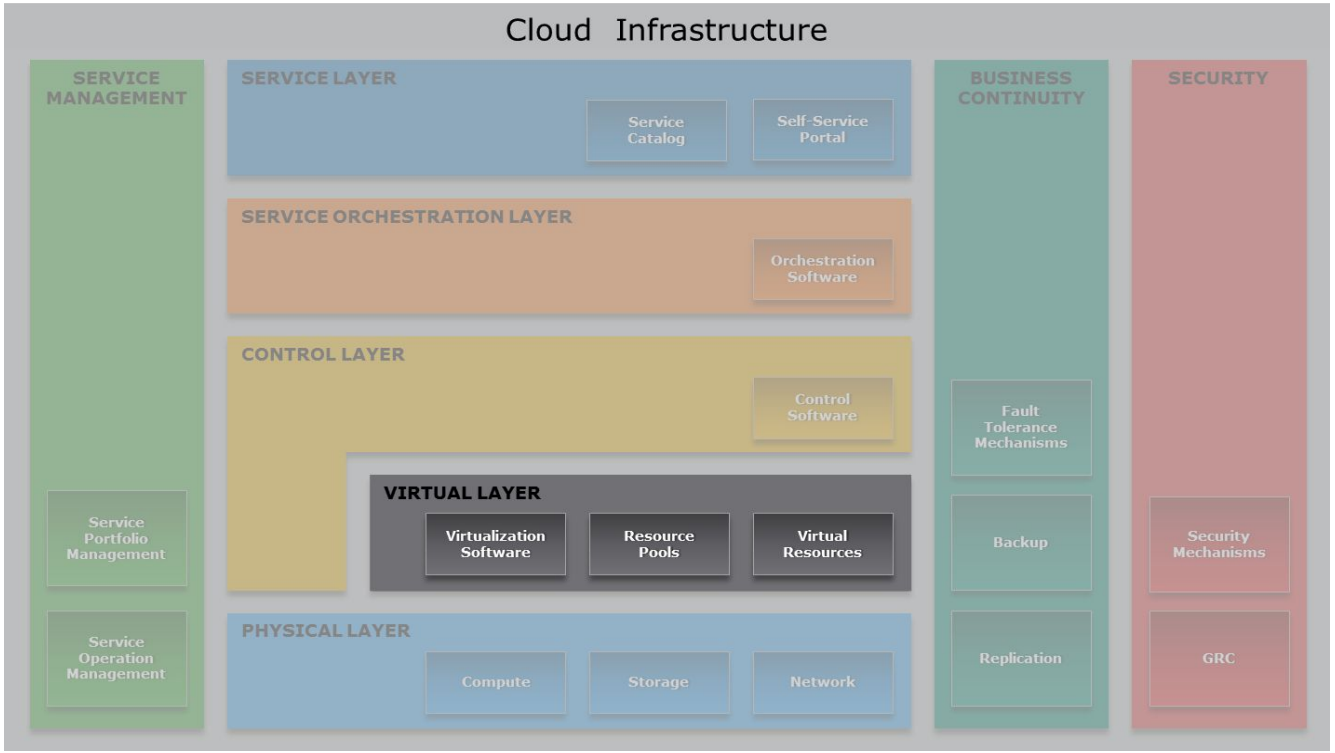
Module: Virtual Layer

Upon completion of this module, you should be able to:

- Describe the virtual layer and virtualization software
- Describe a resource pool and virtual resources

Cloud Computing Reference Model

Virtual Layer



Lesson: Virtual Layer Overview

This lesson covers the following topics:

- Virtual layer
- Virtualization software
- Resource pool
- Virtual resources

Introduction to Virtualization

Virtualization

Refers to the logical abstraction of physical resources, such as compute, network, and storage that enables a single hardware resource to support multiple concurrent instances of systems or multiple hardware resources to support single instance of system.

- Enables a resource to appear larger or smaller than it actually is
- Enables a multitenant environment improving utilization of physical resources

Benefits of Virtualization

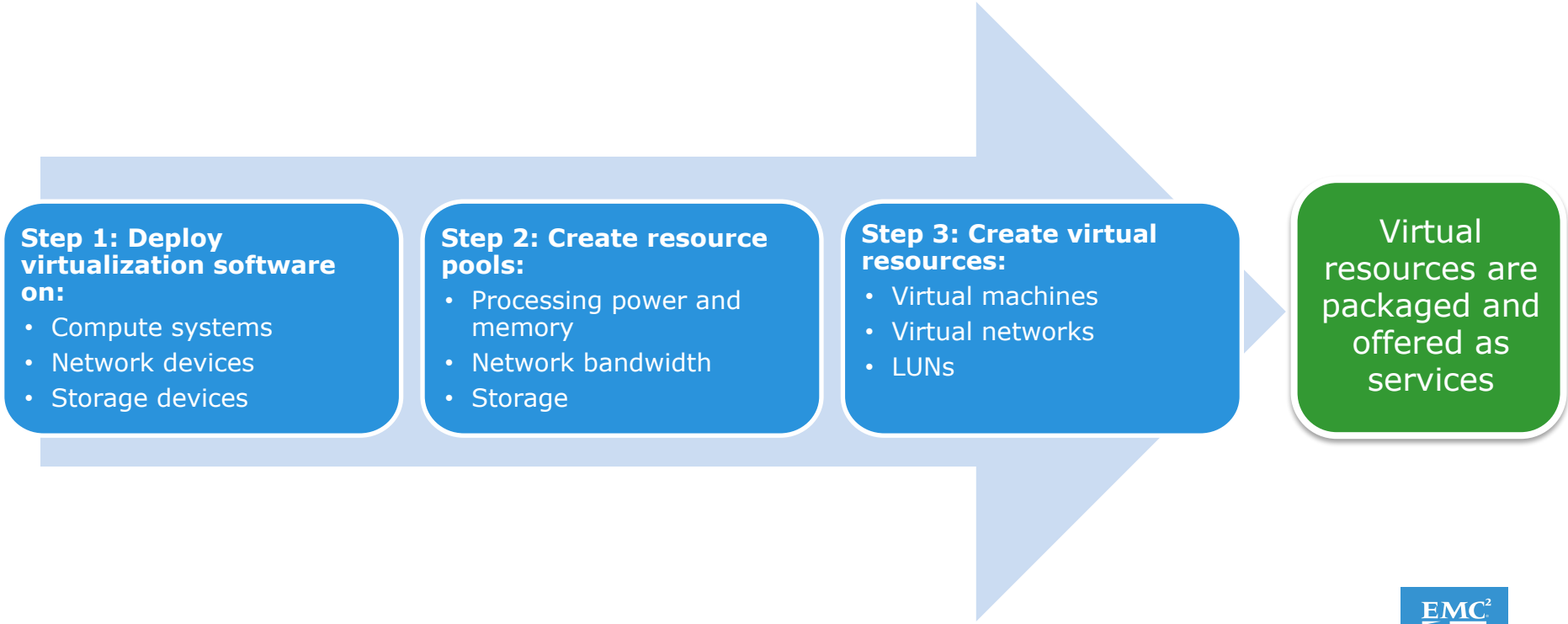
- Optimizes utilization of IT resources
- Reduces cost and management complexity
- Reduces deployment time
- Increases flexibility

Virtual Layer Overview

- Virtualized compute, network, and storage forms the virtual layer
- Enables fulfilling two characteristics of cloud infrastructure
  - Resource pooling
  - Rapid elasticity
- Specifies the entities operating at this layer
  - Virtualization software
  - Resource pools
  - Virtual resources

Virtual Layer

Virtualization Process and Operations



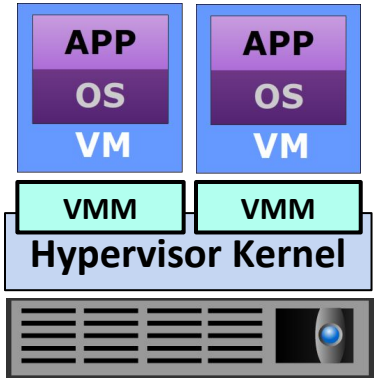
Compute Virtualization Software

Hypervisor

Hypervisor

Software that is installed on a compute system and enables multiple OSs to run concurrently on a physical compute system.

- Hypervisor kernel
  - Provides functionality similar to an OS kernel
  - Designed to run multiple VMs concurrently
- Virtual machine manager (VMM)
  - Abstracts hardware
  - Each VM is assigned a VMM
  - Each VMM gets a share of physical resources



Compute Virtualization Software (Cont'd)

Types of Hypervisor

Bare-metal Hypervisor

- It is an operating system
- Installed on a bare-metal hardware
- Requires certified hardware
- Suitable for enterprise data centers and cloud infrastructure

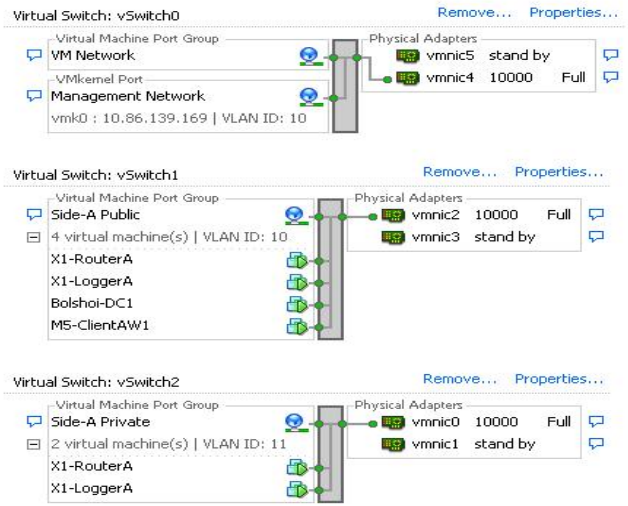
Hosted Hypervisor

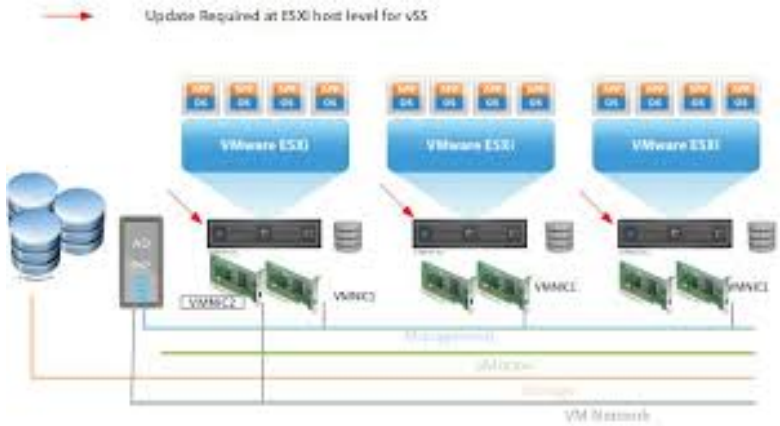
- Installed as an application on an OS
- Relies on OS, running on physical machine for device support
- Suitable for development, testing, and training purposes

Network Virtualization Software

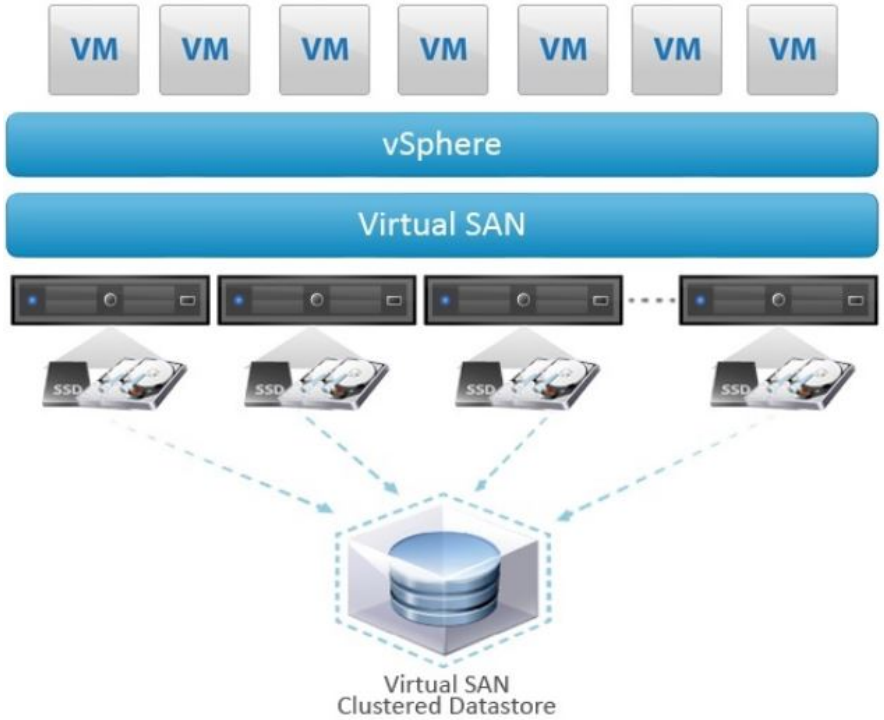
- Abstracts physical network resources to create virtual resources:
  - Virtual LAN/virtual SAN
  - Virtual Switch
- Network virtualization software can be:
  - Built into the operating environment of a network device
  - Installed on an independent compute system
    - Fundamental component for deploying software defined network
  - Hypervisor's capability

Virtual switch





vsan



Storage Virtualization Software

- Abstracts physical storage resources to create virtual resources:
  - Virtual volumes
  - Virtual disk files
  - Virtual arrays
- Storage virtualization software can be:
  - Built into the operating environment of a storage device
  - Installed on an independent compute system
    - Fundamental component for deploying software defined storage
  - Hypervisor's capability

Lesson Summary

During this lesson the following topics were covered:

- Virtual layer
- Virtualization software
- Resource pool
- Virtual resources

Lesson: Resource Pool

This lesson covers the following topics:

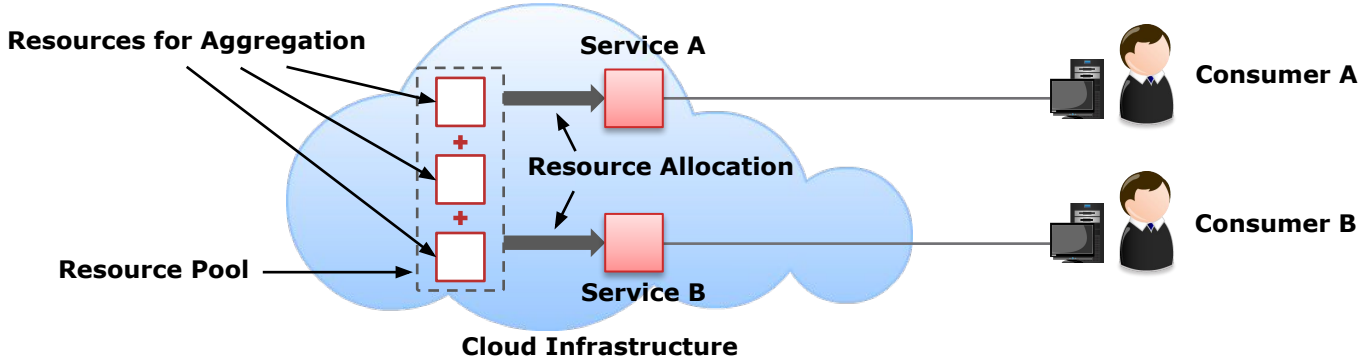
- Resource pool
- Examples of resource pooling
- Identity pool

Introduction to Resource Pool

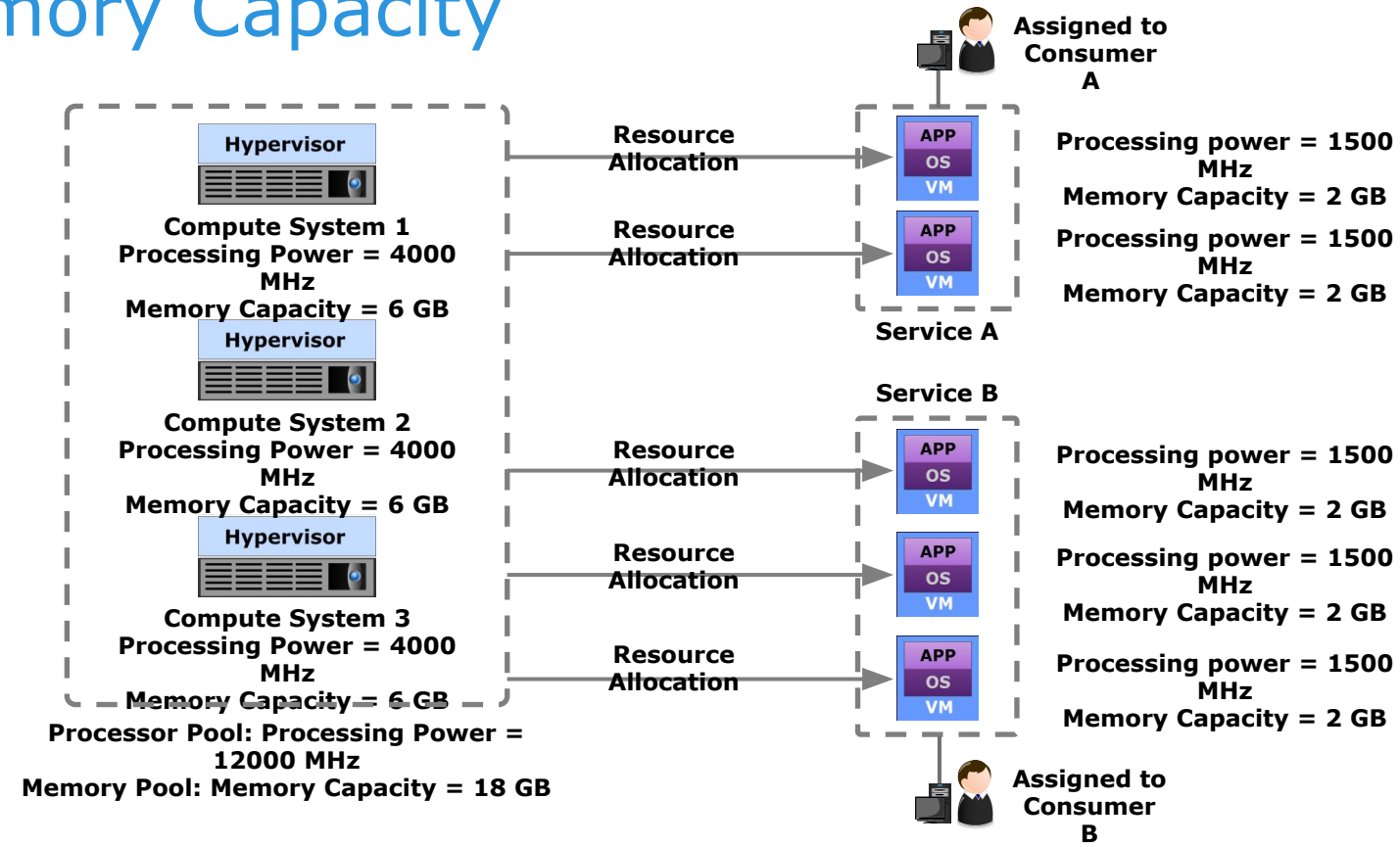
Resource Pool

A logical abstraction of the aggregated computing resources, such as processing power, memory capacity, storage, and network bandwidth that are managed collectively.

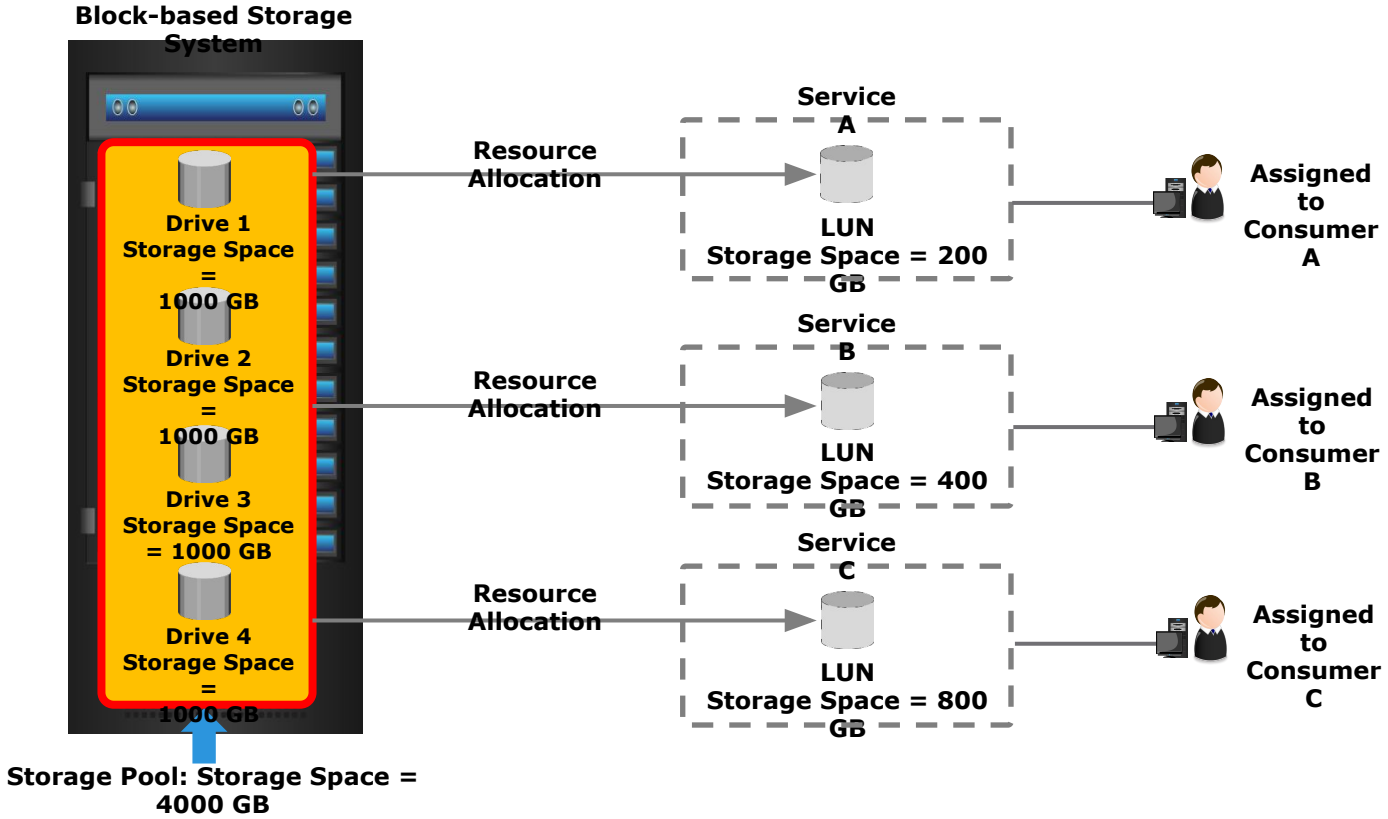
- Cloud services obtain computing resources from resource pools
  - Resources are dynamically allocated as per consumer demand
- Resource pools are sized according to service requirements



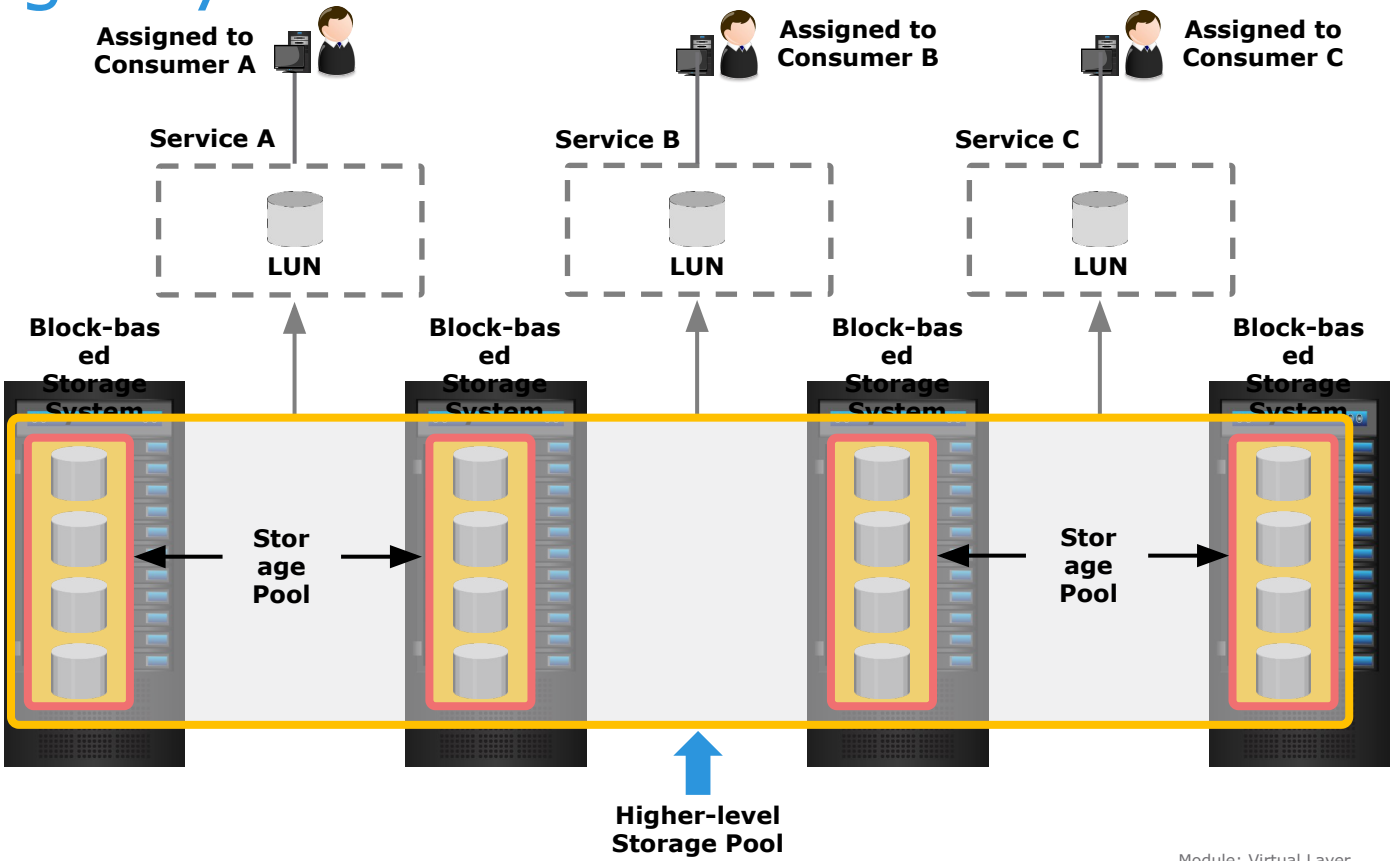
Example: Pooling Processing Power and Memory Capacity



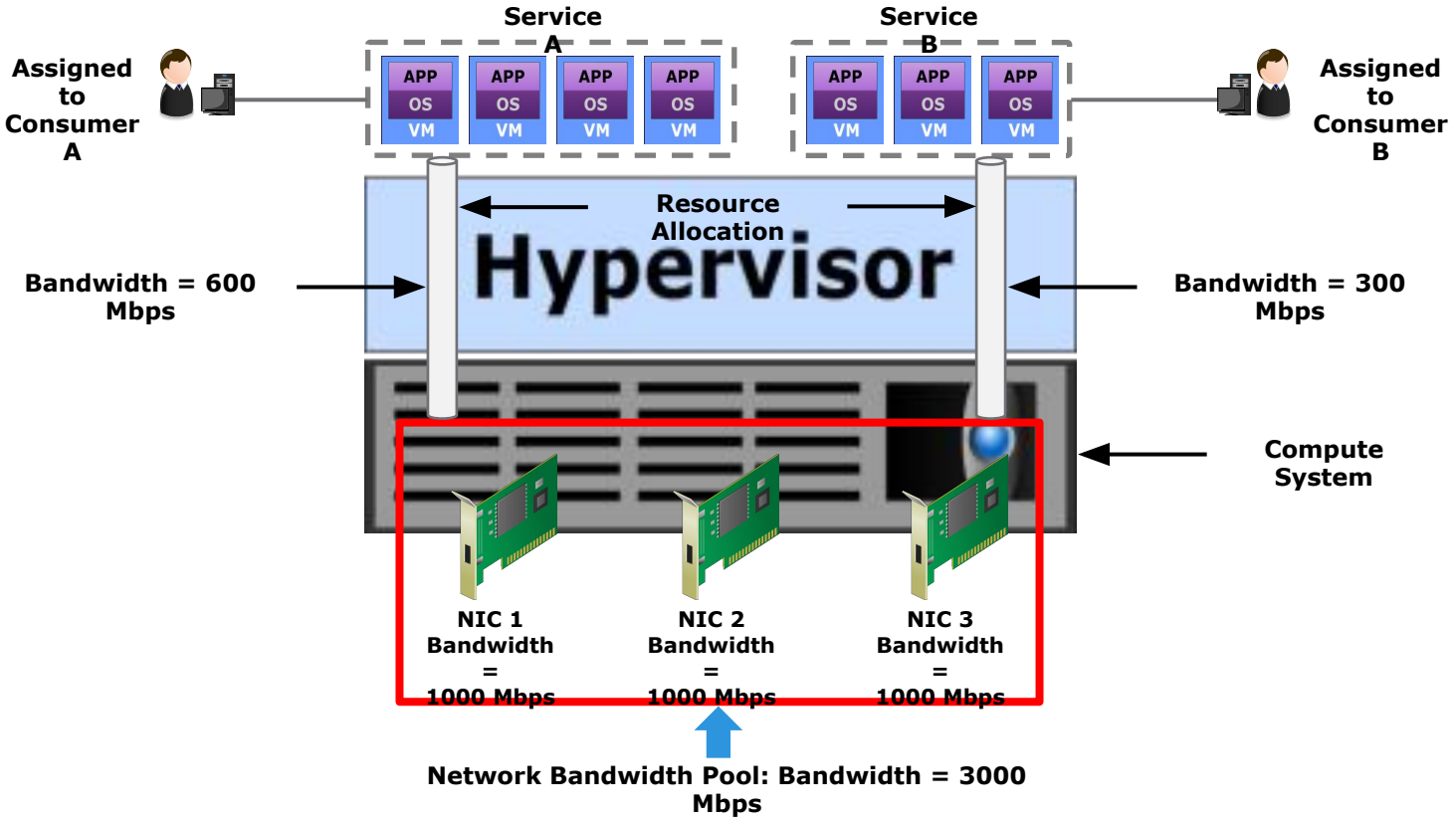
Example: Pooling Storage in a Block-based Storage System



Example: Pooling Storage Across Block-based Storage Systems

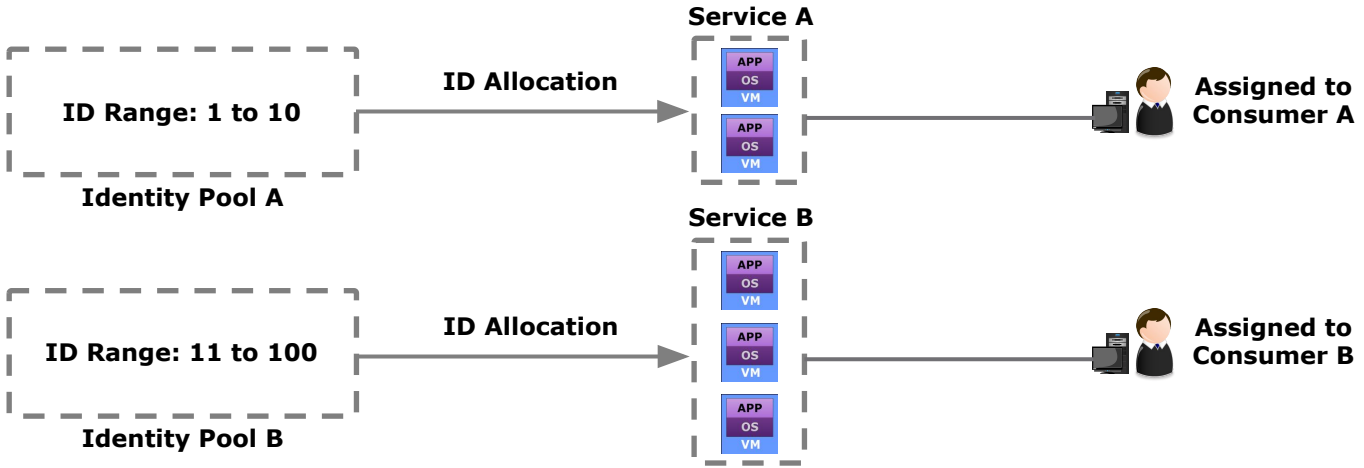


Example: Pooling Network Bandwidth of NICs



Identity Pool

- Specifies a range of network identifiers (IDs) such as virtual network IDs and MAC addresses
  - IDs are allocated from the identity pools to the elements of cloud services
- An identity pool may map to a particular service or to a group of services



Lesson Summary

During this lesson the following topics were covered:

- Resource pool
- Examples of resource pooling
- Identity pool

Lesson: Virtual Resources – I

This lesson covers the following topics:

- Virtual machine (VM) and VM hardware
- VM files and file system to manage VM files
- VM console
- VM template
- Virtual appliance
- VM network and its components

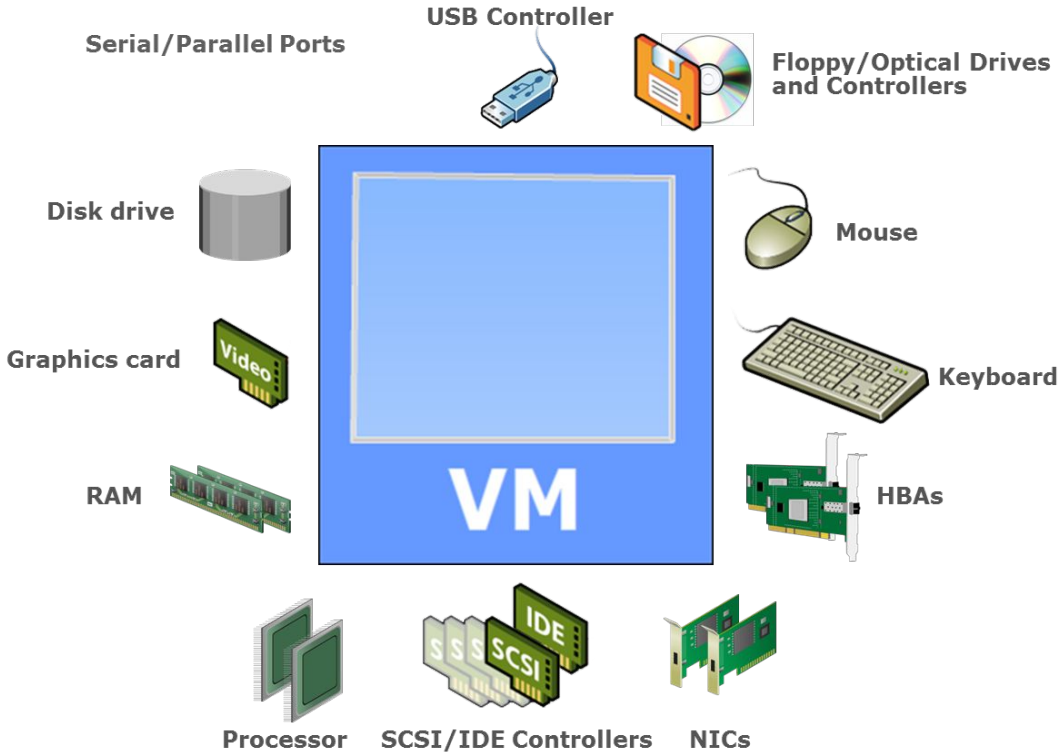
Virtual Machine (VM)

Virtual Machine

A logical compute system that, like a physical compute system, runs an OS and applications.

- Created by a hypervisor installed on a physical compute system
- Comprises virtual hardware, such as virtual processor, memory, storage, and network resources
  - Appears as a physical compute system to the guest OS
  - Hypervisor maps the virtual hardware to the physical hardware
- Provider provisions VMs to consumers for deploying applications
  - VMs on the same compute system or cluster run in isolation

VM Hardware



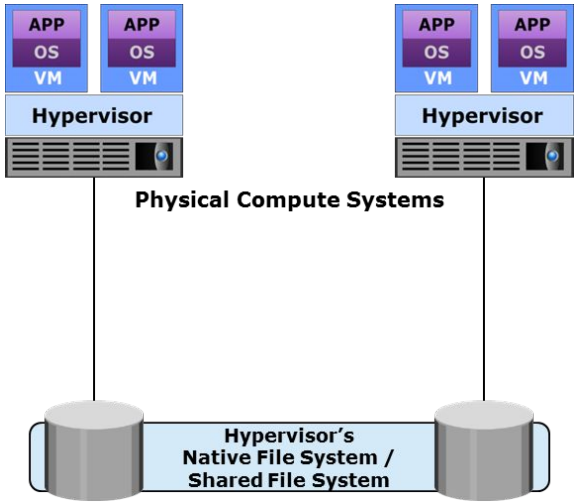
VM Files

- From a hypervisor's perspective, a VM is a discrete set of files such as:

Configuration file	• Stores information, such as VM name, BIOS information, guest OS type, memory size
Virtual disk file	• Stores the contents of the VM's disk drive
Memory state file	• Stores the memory contents of a VM in a suspended state
Snapshot file	• Stores the VM settings and virtual disk of a VM
Log file	• Keeps a log of the VM's activity and is used in troubleshooting

File System to Manage VM Files

- Hypervisor's native file system
  - Clustered file system deployed on local or external storage
  - Enables multiple hypervisors to perform concurrent reads and writes
  - Enables high availability to protect against hypervisor or compute system failure
- Shared file system
  - Enables storing VM files on remote file servers or NAS devices
  - Hypervisors have built-in NFS or CIFS clients



VM Console

- VM console is an interface to view and manage the VMs on a compute system or a cluster
- VM console may be:
  - Installed locally on a compute system
  - Web-based
  - Accessed over a remote desktop connection
- Used to perform activities such as:
  - Installing a guest OS and accessing VM BIOS
  - Powering a VM on or off
  - Configuring virtual hardware and troubleshooting

VM Template

VM Template

A master copy of a VM with standardized virtual hardware and software configuration that is used to create new VMs

- Created in two ways:
  - Converting a VM into a template
  - Cloning a VM to a template
- Steps involved in updating a VM template are:
  1. Convert the template into VM
  2. Install new software or OS/software patches
  3. Convert the VM back to a template

Virtual Appliance

Virtual Appliance

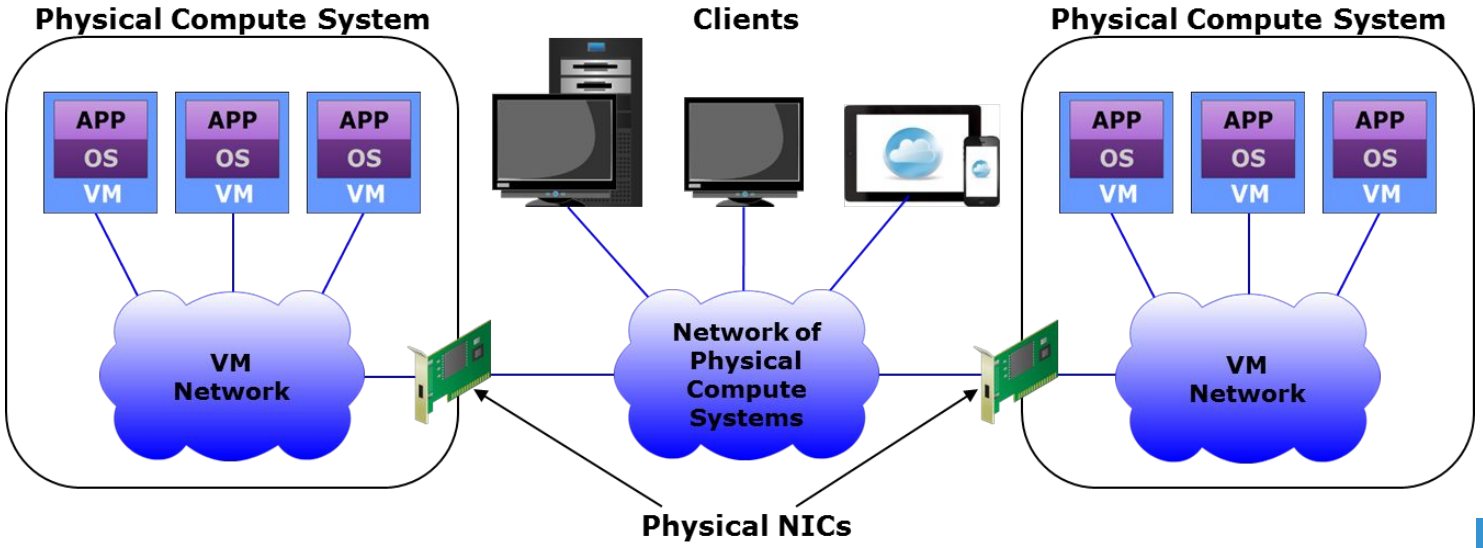
Preconfigured virtual machine(s) preinstalled with a guest OS and an application dedicated to a specific function.

- Used for functions, such as providing SaaS, routing packets, or deploying a firewall
- Simplifies the delivery and operation of an application
  - Simplifies installation and eliminates configuration issues
  - The application is protected from issues in other virtual appliances
- Typically created using Open Virtualization Format (OVF)

VM Network

VM Network

A logical network that provides Ethernet connectivity and enables communication between VMs within a compute system.



VM Network Components

Component	Description
Virtual switch	• A logical OSI Layer 2 Ethernet switch created in a compute system • Connects VMs locally and also directs VM traffic to a physical network • Forwards frames to a virtual switch port based on destination address • A distributed virtual switch can function across multiple physical compute systems
Virtual NIC	• Connects a VM to a virtual switch and functions like a physical NIC • Has unique MAC and IP addresses • Forwards the VM's network I/O in the form of Ethernet frames to the virtual switch
Uplink NIC	• A physical NIC connected to the uplink port of a virtual switch • Functions as an ISL between virtual and physical Ethernet switches • Not addressable from the network

Lesson Summary

During this lesson the following topics were covered:

- Virtual machine and VM hardware
- VM files and file system to manage VM files
- VM console
- VM template
- Virtual appliance
- VM network and its components

Lesson: Virtual Resources – II

This lesson covers the following topics:

- Logical unit number (LUN)
- Creating LUN from RAID set
- Creating LUN from storage pool

Logical Unit Number (LUN)

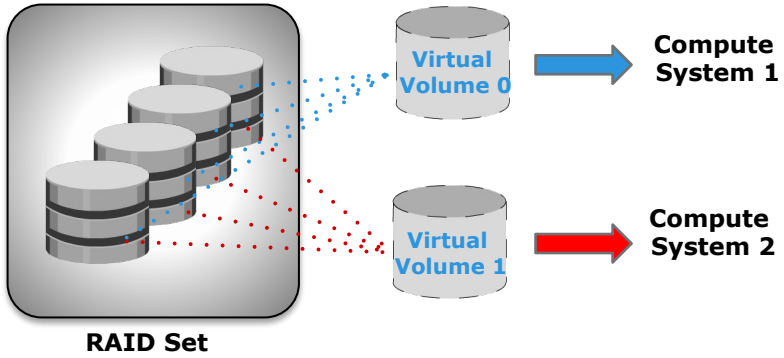
Logical Unit Number (LUN)

Abstracts the identity and internal functions of storage system(s) and appear as physical storage to the compute system.

- Mapping of virtual to physical storage is performed by the virtualization layer.
- Provider provisions LUN to consumers for storing data
  - Storage capacity of a LUN can be dynamically expanded or reduced
- LUN can be created from
  - RAID set (traditional approach)
  - Storage pool

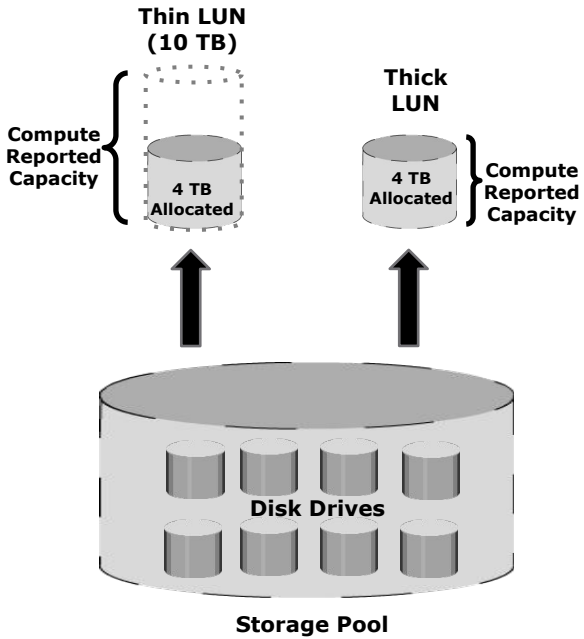
Creating LUNs from RAID Set

- LUNs are created from a RAID set by partitioning the available capacity into smaller units
  - Spread across all the physical disks that belong to a RAID set
- Suited for applications that require predictable performance



Creating LUNs from Storage Pool

- Two types of volumes are created from storage pool:
  - Thin LUN
    - Does not require physical storage to be completely allocated at the time of creation
    - Consumes storage as needed from the underlying storage pool in increments called thin LUN extents
  - Thick LUN
    - Physical storage is completely allocated at the time of creation



Use of Thin LUN

- Thin LUNs are appropriate for applications that can tolerate performance variations
  - In some cases, performance improvement is seen when using a thin volume due to striping across large number of drives in the pool
- Environments where cost, storage utilization, space, and energy efficiency is paramount
- For applications where storage space consumption is difficult to forecast
- Environment that needs optimized self provisioning

Lesson Summary

During this lesson the following topics were covered:

- LUN
- Creating LUN from RAID set
- Creating LUN from storage pool

Lesson: Virtual Resources – III

This lesson covers the following topics:

- Virtual network
- Types of virtual networks: VLAN and VSAN
- Mapping between VLANs and VSANs in an FCoE SAN

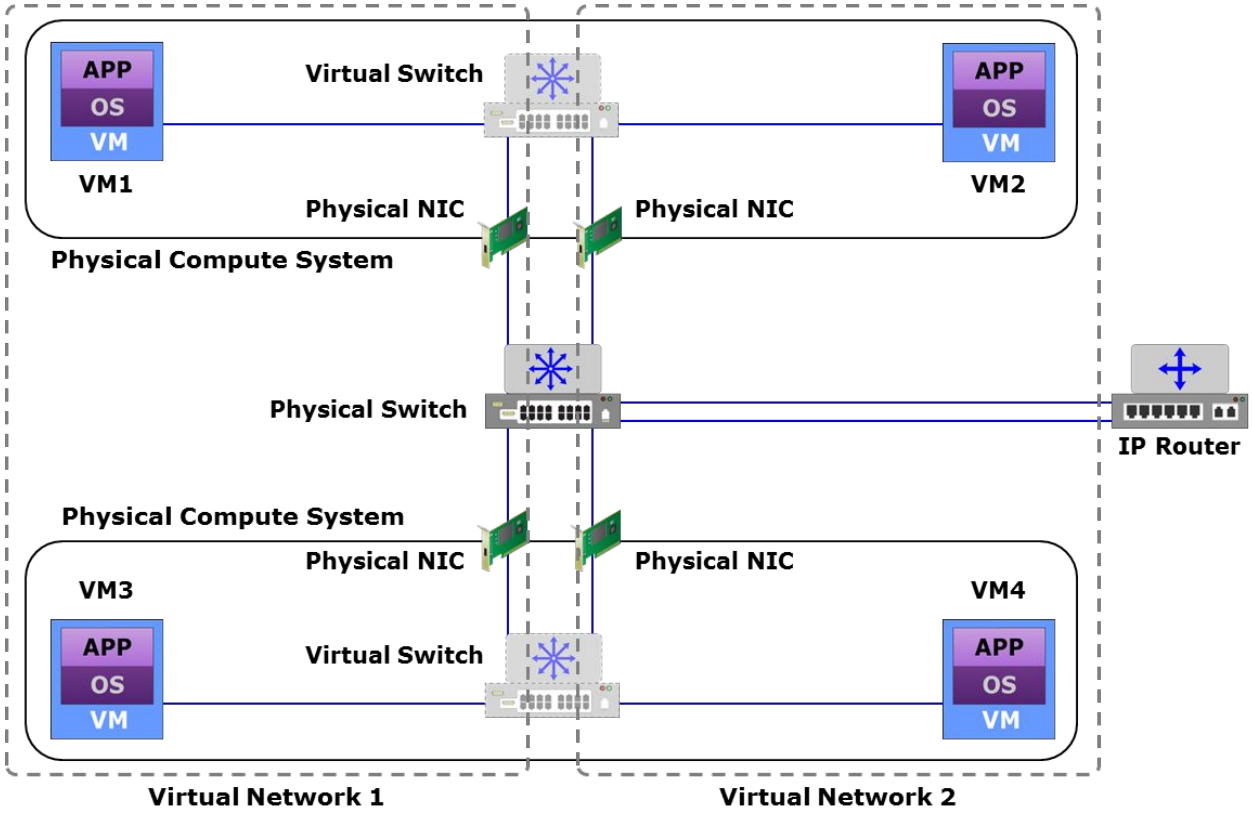
Virtual Network

Virtual Network

A software-based logical network that is either a segment of a physical network or spans across multiple physical networks.

- Appears as a physical network to the connected nodes
- Virtual networks share network components without leaking information between them
- Network traffic is routed only when two nodes in different virtual networks are communicating
- All types of networks can be virtualized, such as compute network, SAN, and VM network

Virtual Network Example



Common Types of Virtual Networks

- Virtual LAN (VLAN)
- Private VLAN (PVLAN)
- Stretched VLAN
- Virtual extensible LAN (VXLAN)
- Virtual SAN (VSAN)

Virtual LAN (VLAN)

Virtual LAN (VLAN)

A virtual network created on a LAN enabling communication between a group of nodes with a common set of functional requirements, independent of their physical location in the network.

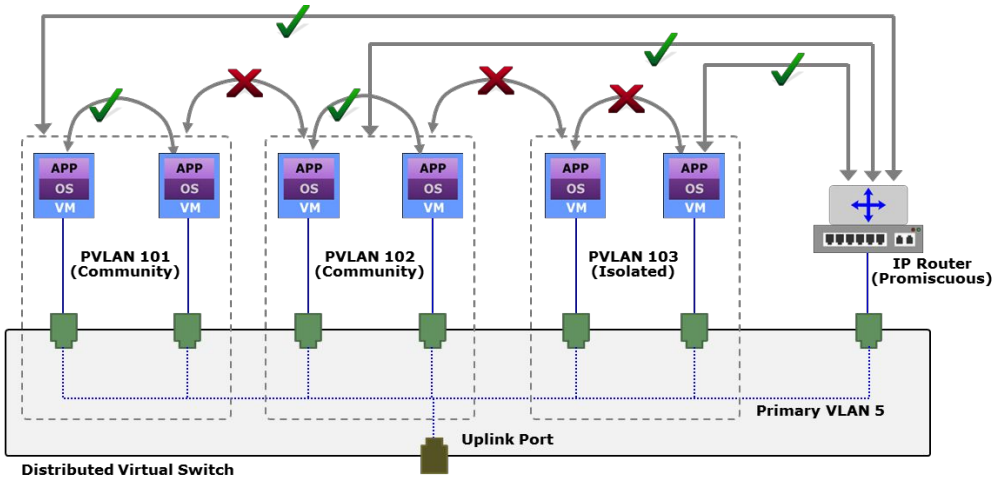
- A VLAN is identified by a unique 12-bit VLAN ID
- Configuring a VLAN:
  - Define VLAN on physical and virtual switches and assign VLAN ID
  - Configure VLAN membership based on port, MAC address, protocol, IP subnet address, or application

Private VLAN (PVLAN)

Private VLAN

A sub-VLAN that segregates the nodes within a standard VLAN, called as primary VLAN. A PVLAN can be configured as either isolated or community.

- Enables a provider to support a larger number of consumers
- Provides security between nodes on the same VLAN
- Simplifies network management

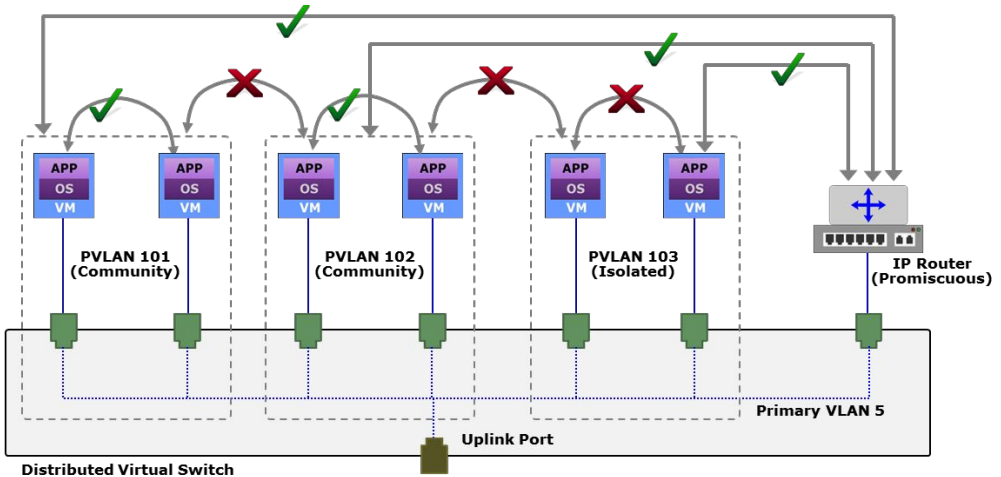


Private VLAN (PVLAN)

Private VLAN

A sub-VLAN that segregates the nodes within a standard VLAN, called as primary VLAN. A PVLAN can be configured as either isolated or community.

- Enables a provider to support a larger number of consumers
- Provides security between nodes on the same VLAN
- Simplifies network management

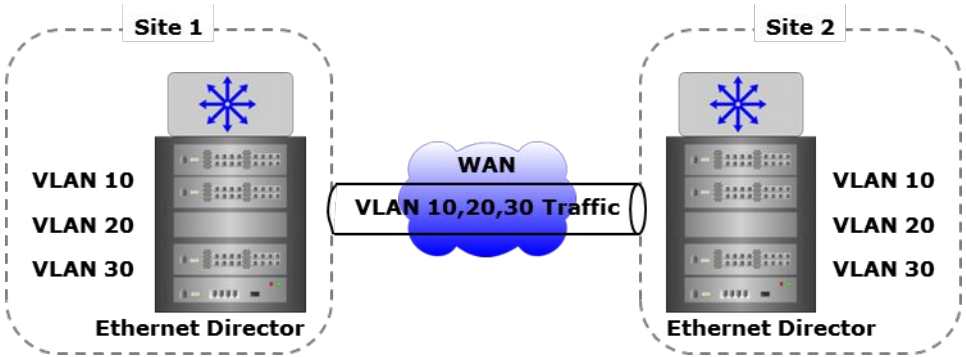


Stretched VLAN

Stretched VLAN

A VLAN that spans multiple sites and enables Layer 2 communication between a group of nodes over a Layer 3 WAN infrastructure, independent of their physical location.

- Layer 2 WAN frames are encapsulated in Layer 3 WAN packets
- Enables movement of VMs across locations without changing their network configuration



Virtual Extensible LAN (VXLAN)

Virtual Extensible LAN

A logical Layer 2 overlay network built on a Layer 3 network, which uses MAC-in-UDP encapsulation to enable communication between a group of nodes, independent of their physical location.

- VXLAN header is added to a Layer 2 frame, which is placed in a UDP-IP packet and tunneled over a Layer 3 network
  - Enables transparent Layer 2 communication between nodes over physical networks spanning Layer 3 boundaries
  - Encapsulation and decapsulation are performed by Virtual Tunnel Endpoints (VTEPs)
- 24-bit VXLAN ID provides up to 16 million VXLANs

Virtual SAN (VSAN)

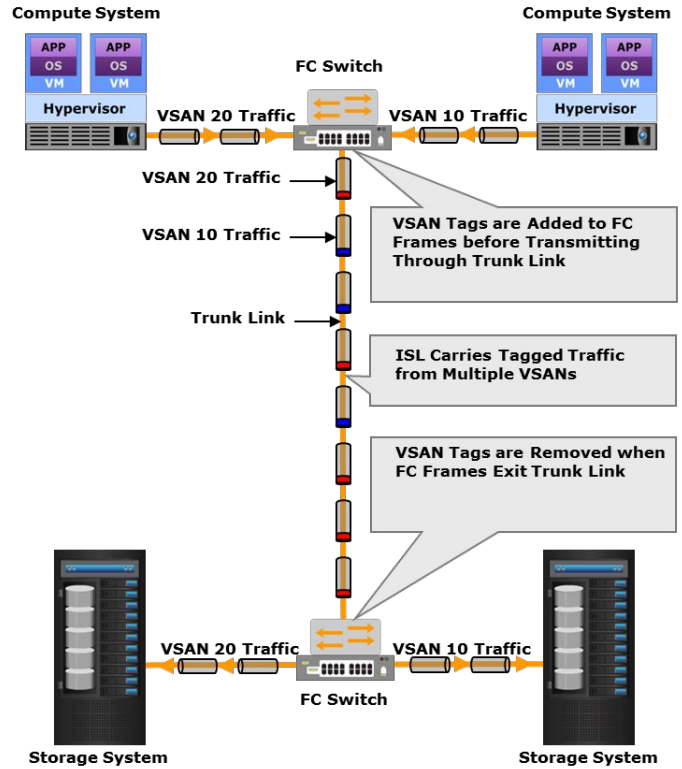
Virtual SAN

A logical fabric, created on a physical FC or FCoE SAN enabling communication between a group of nodes with a common set of requirements, independent of their physical location in the fabric.

- A VSAN has its own fabric services, configuration, and set of FC addresses
- Traffic disruptions in one VSAN do not affect other VSANs
- A VSAN may be extended across sites similar to a stretched VLAN

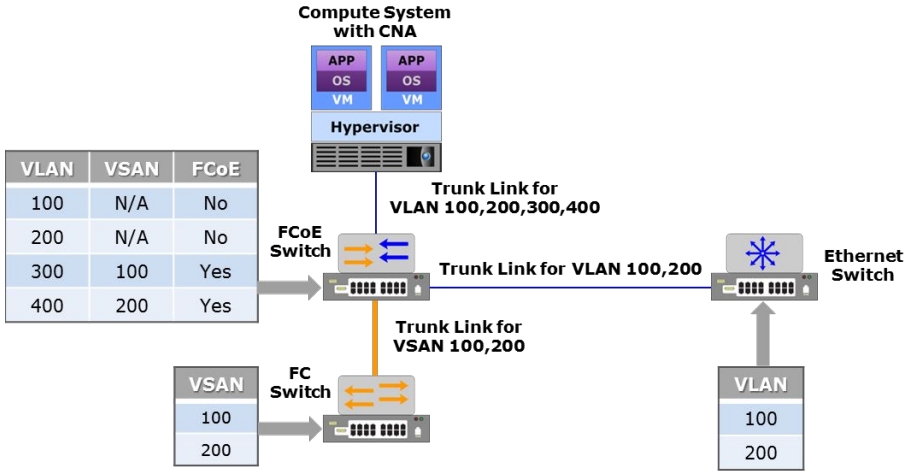
Virtual SAN (VSAN) (Cont'd)

- Configuring VSAN:
  - Define VSANs on fabric switch with specific VSAN IDs
  - Assign VSAN IDs to F_Ports to include them in the VSANs
- An N_Port connecting to an F_Port in a VSAN becomes a member of that VSAN



Mapping VLANs and VSANs in an FCoE SAN

- Mapping determines which VLAN carries a VSAN traffic
- Mapping considerations:
  - Configure a dedicated VLAN for each VSAN
  - VLANs configured for VSANs should not carry regular LAN traffic



Lesson Summary

During this lesson the following topics were covered:

- Virtual network
- Types of virtual network: VLAN, private VLAN, stretched VLAN, VXLAN, and VSAN
- Mapping between VLANs and VSANs in an FCoE SAN

Concepts in Practice

- VMware ESXi

VMware ESXi

ESXi
• Bare-metal hypervisor • Abstracts processor, memory, storage, and network resources into multiple VMs • Comprises underlying VMkernel OS that supports running multiple VMs   - VMkernel controls and manages compute resources

Module Summary

Key points covered in this module:

- Virtual layer
- Virtualization software
- Resource pool
- Virtual resources

This slide intentionally left blank.