Data Center Architectures and Virtualization

January 25, 2011

Introduction

- Welcome to “the cloud”
  - The relationship between data and its physical infrastructure is dissolving.
  - All non-local data that is accessed from your computer, cell phone, or television is stored and managed in data centers….*somewhere*.
  - With the global build out of high performance optical networks, the location of the data center relative to the user has become irrelevant for many applications.

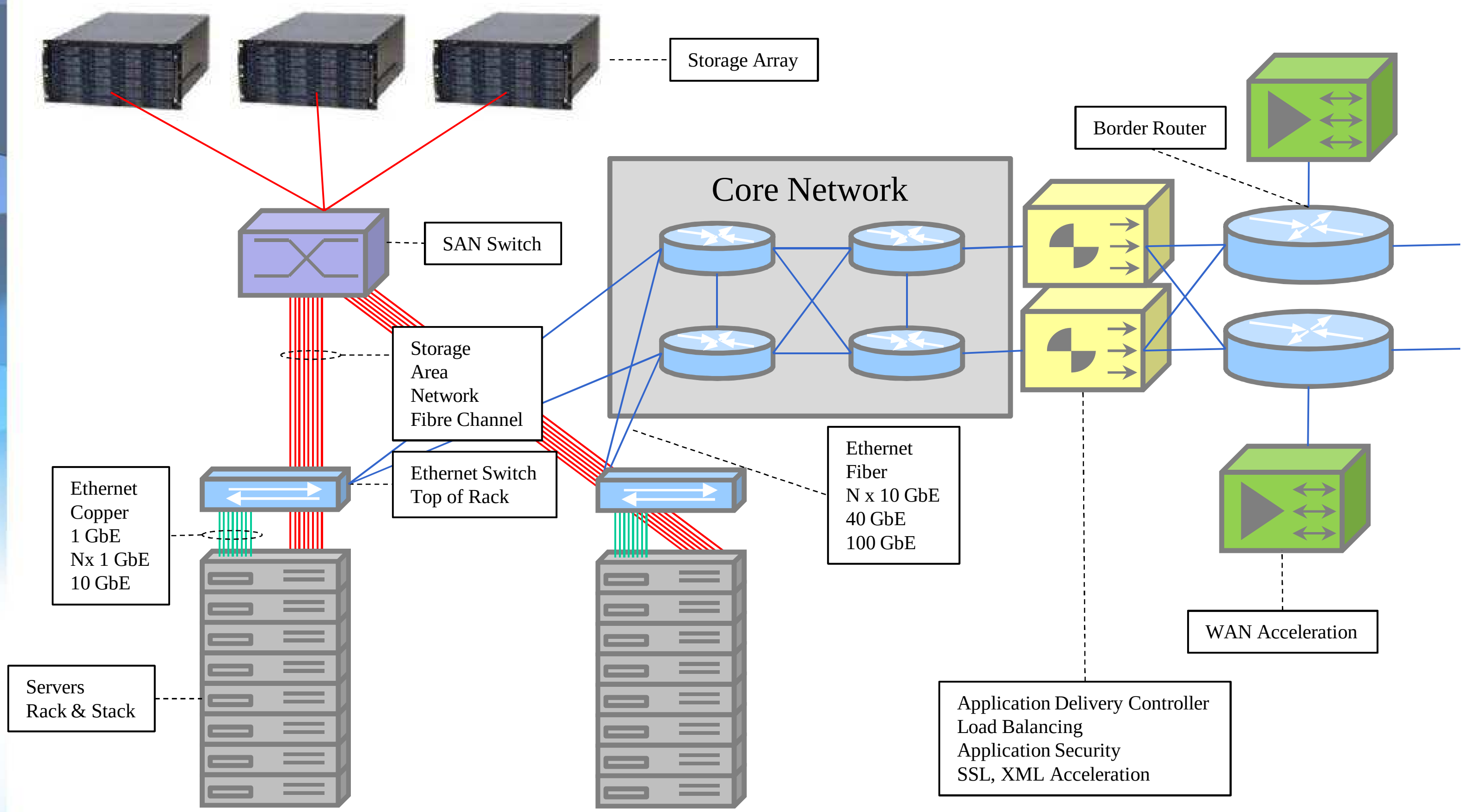
Data Center Trends and Drivers

- Data Center Trends:
  - Migration of enterprise infrastructure to third party data centers
  - Emergence of businesses with data centers as a core to the business model:
    - Content delivery
      ex: Zynga
    - Financial services
      ex: NYSE
    - Database management
      ex: salesforce.com
  - Data centers as big business:
    - Amazon
    - Google
    - Qwest
    - Verizon
- Data Center Drivers:
  - Optimization & Monetization:
    - Virtualization
    - Power & space management
    - Availability & resiliency
    - Scaling

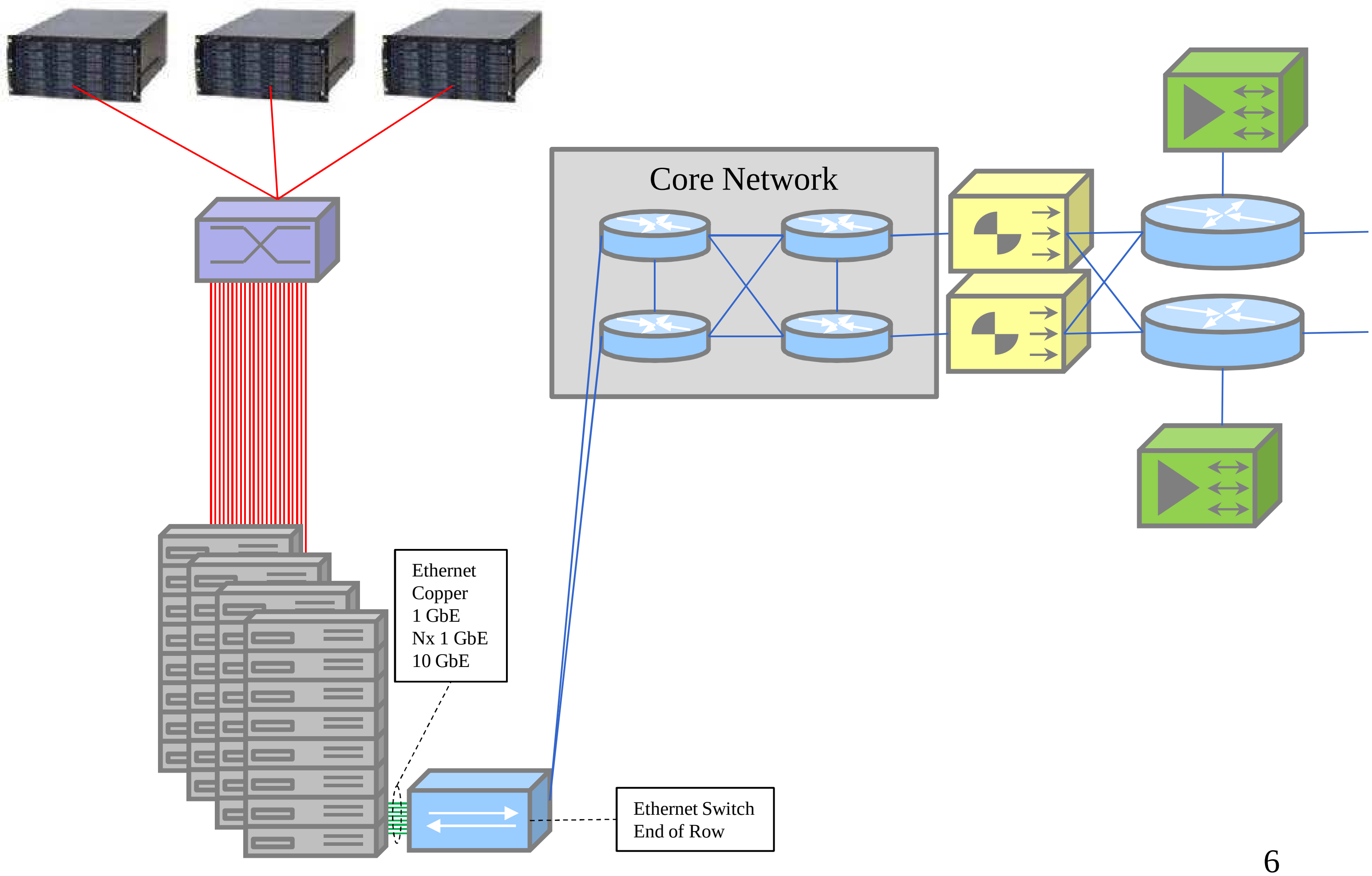
Agenda

- Data Center Prototypical Architectures
- Data Center Architectural Considerations
- Virtualization:
  - Servers
  - Networking
  - Storage
- Q & A

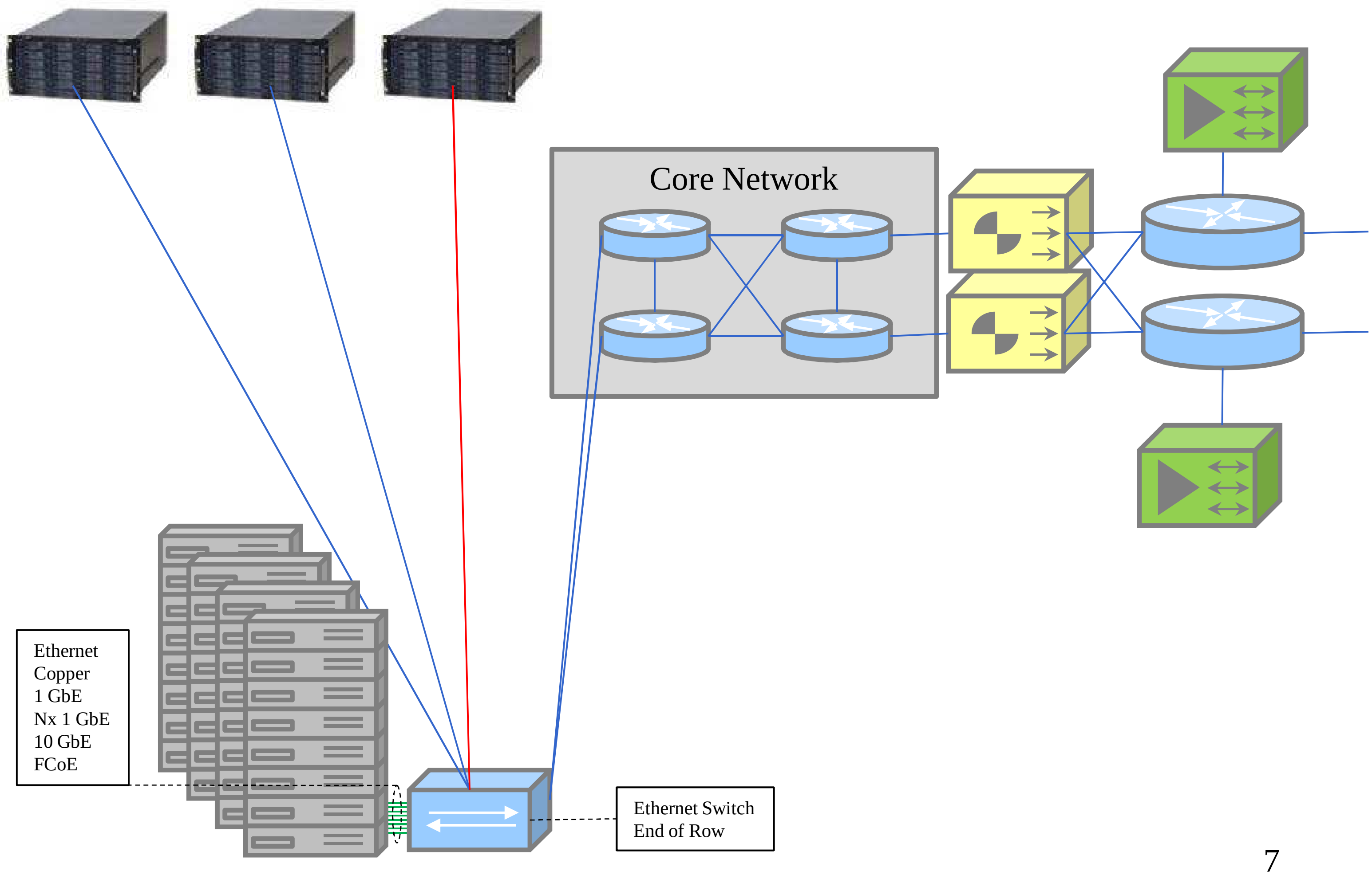
Data Center Prototypical Architectures



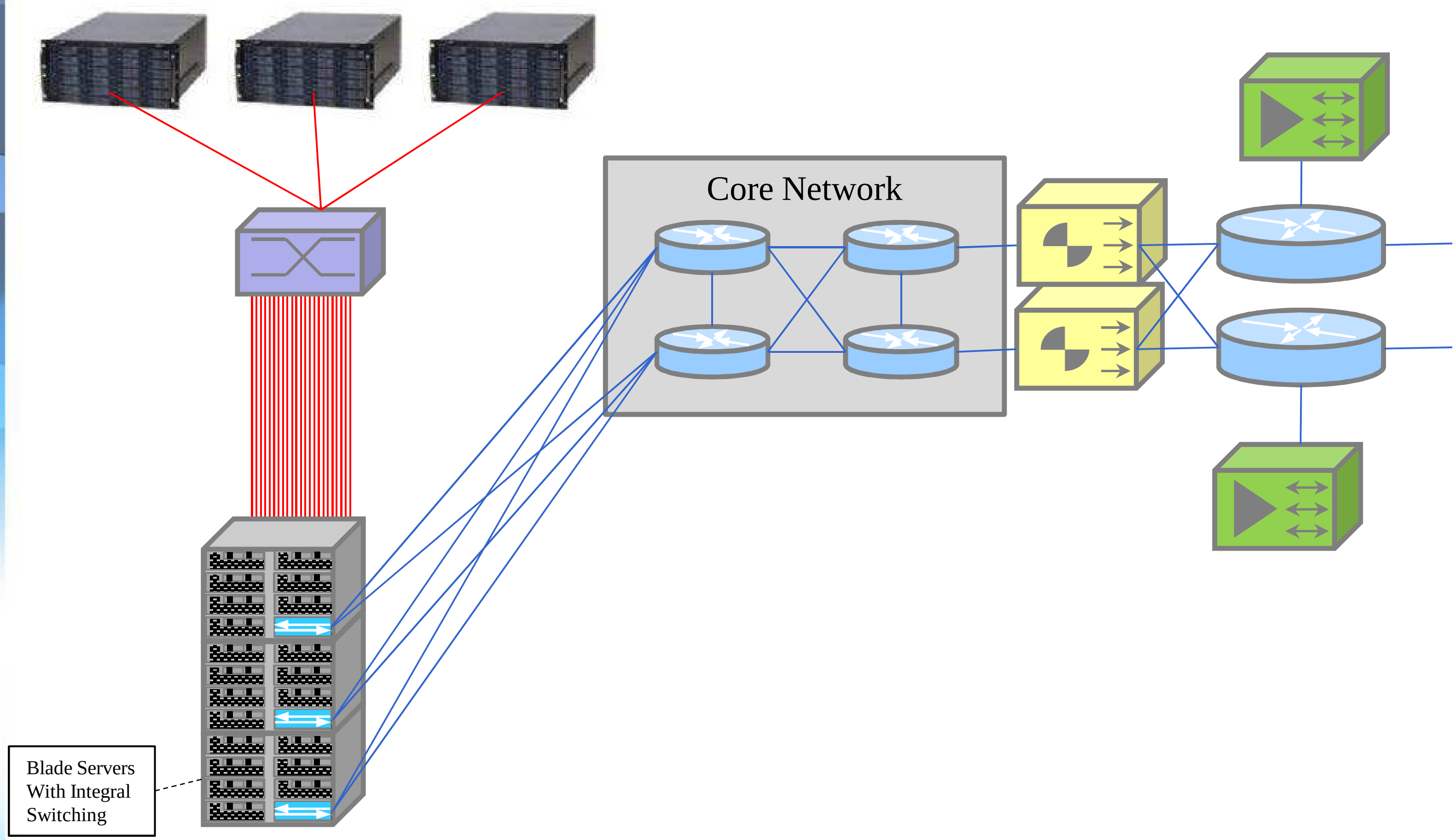
Data Center Prototypical Architectures



Data Center Prototypical Architectures



Data Center Prototypical Architectures



Data Center Architectural Considerations

- Data center operators and architects are facing a variety of challenges:
  - Power
  - Space
  - Performance
  - Availability and Resiliency
  - Security

Data Center Power Efficiency is measured as:

PUE* = Total Facility Power/
Power Delivered to IT
Equipment

The typical PUE for a data center is 2.0. This means that one watt goes to infrastructure (lighting, cooling, power delivery) for every watt delivered to IT equipment. State of the art PUE is about 1.2

* Power Utilization Efficiency

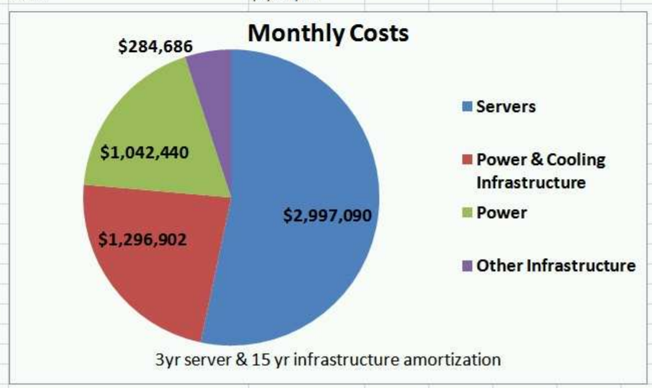
Data Center Architectural Considerations: ***Power***

Location Matters!
PUE Matters!

- In the Microsoft Analysis, the PUE is assumed to be 1.7 and the cost of power is $.07/kwh.
- With these assumptions, the cost of power is about ½ of the cost of the servers.
- If the PUE is 2.5, then the cost of power is about ½ of the cost of the servers.
- The current average retail rate of electricity is $.1045/kwh which makes the cost of power about ½ of the cost of the servers.
- In Connecticut the current average retail rate is $.1718/kwh which makes the cost of power about $2.6million/month

Microsoft Analysis

Assumptions			
Cost of power ($/kwh):	$0.07		
Cost of Facility ($):	$200,000,000.00		
Facilities Amortization:	180	(15 years)	
Number of Servers:	50,000.00		
Cost/Server ($)	$2,000.00		
Server Amortization (months)	36	(3 years)	
Size of Facility (Critical Load MW):	15,000,000.00	(15MW)	
Annual Cost of Money (%):	5%		
Average Power Usage (%):	80%	(Average % of provisioned power used)	
Power Usage Effectiveness	1.7		
Power and Cooling Infrastrucure (%)	82%	(% of infrastructure that is power & cooling)	
Network egress charges not included (workload dependent)			
Calculations			
Infrastructure	$1,581,587	[=-PMT(CostOfMoney/12,FacilityAmortization,FacilityCost,0)]	
Servers	$2,997,090	[=-PMT(CostOfMoney/12, ServerAmortization, ServerCount*ServerCost, 0)]	
Power & Cooling Infrastructure	$1,296,902	[=InfrastructureMonthly*PowerAndCoolikngInfrastructurePercentage]	
Power	$1,042,440	[=MegaWattsCriticalLoad*AveragePowerUsage/1000*PUE*PowerCost*24*365/12	
Other Infrastructure	$284,686	[=+InfrastructureMonthly-PowerAndCoolingInfrastructureMonthly]	
Full burdened Power	$2,339,342	[=+PowerAndCoolingInfrastructureMonthly+PowerMonthly]	
Total:	$5,621,117		



Data Center Architectural Considerations: ***Space***

- The space requirements of a data center are based upon the physical equipment space and the heat dissipation loads.
- A high density data center is typically comprised of racks with power of 14 KW and above and a working footprint (work cell) of about 20 square feet.
- Layout efficiency is the number of racks per 1,000 square feet of floor space. A typical range is 20 to 30.
- Watts/sq foot of work cell is the preferred metric to describe cooling load requirements.
- High density data centers can reduce the TCO significantly over low density data centers.

Intel TCO Example

	Low-Density	High-Density	Notes
Capital Cost – Building	$6,285,620	$2,393,600	$220/sq ft for CSA
Design Cost for CFD	$0	$54,440	Assumes $5/sq ft;
Capital cost taller DC	$0	$239,360	Assumes +10%
Capita cost for 30" RF.	$0	$10,880	$1 sq ft
Lighting	$126,000	$0	NPV(5 yr, i=5%)
IT Equipment (Racks)	$1,875,000	$714,000	$1.5K/ea + $1.5K/install
Oper Cost – Cooling	$1,091,000	$736,000	NPV of 5 yr with i=5%
Total Cost Delta	$9,377,620	$4,148,240	$5.2 M savings

Data Center Architectural Considerations: ***Performance***

- Data center performance is a function of specific business models. This drives purpose built architectures.
- *Financial Trading:*
  - ***Latency is Everything***

    With automated trading, 1 ms. can mean the difference between a successful trade and an unsuccessful trade. Time stamped traffic
- *Infrastructure as a Service (IaaS):*
  - ***Efficiency***

    Virtualization of servers to support more revenue per physical server. Management of capital and operational costs.
- *On Line Gaming:*
  - ***User Experience***

    Server performance…infrastructure scaling.

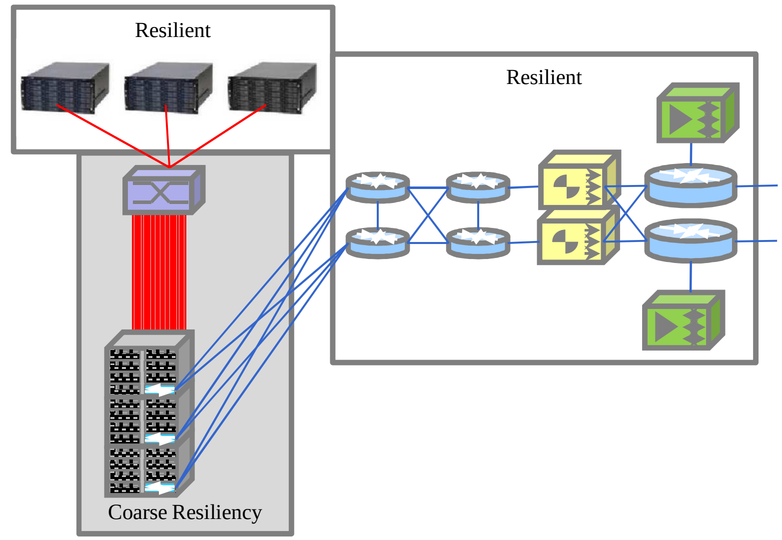
"Speaking at Oracle's OpenWorld Zynga's Cadir Lee has revealed that 215 million people have played a Zynga game at one point, about 10 percent of the world's internet population."

"…and the huge number of players actually spend quite a lot of time with the games, enough to move 1 Petabyte of data which is what Zynga sees every day."

"…and these data centers are growing at a huge rate, Zynga adds about 1,000 new servers each week."

Data Center Architectural Considerations: *Availability and Resiliency*

- As customers migrate mission critical applications to data centers, availability becomes absolutely critical.
- A set of tiered standards have been developed to encompass infrastructure resiliency in the data center:



Tier 1: composed of a single path for power and cooling distribution, without redundant components, providing 99.671% availability.

Tier II: composed of a single path for power and cooling distribution, with redundant components, providing 99.741% availability

Tier III: composed of multiple active power and cooling distribution paths, but only one path active, has redundant components, and is concurrently maintainable, providing 99.982% availability

Tier IV: composed of multiple active power and cooling distribution paths, has redundant components, and is fault tolerant, providing 99.995% availability.

94 minutes/year of downtime

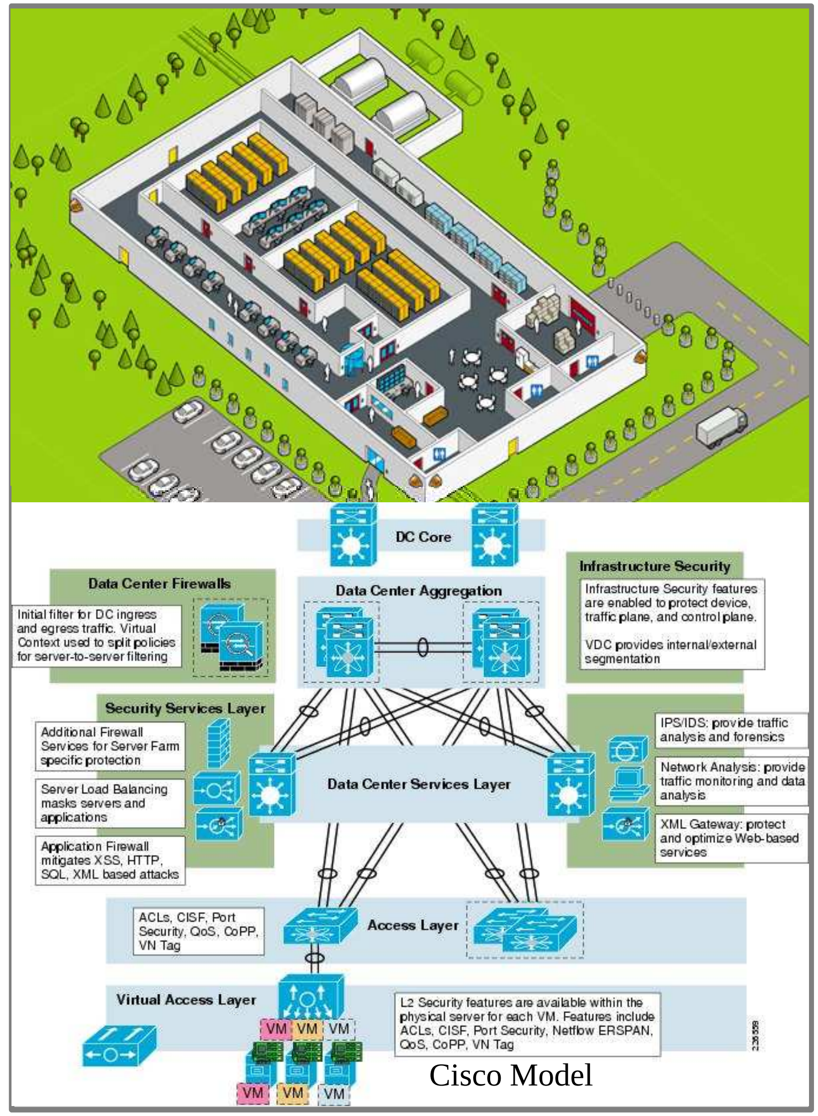
26 minutes/year of downtime

Data Center Architectural Considerations: *Security*

- Data center security is comprised of physical security and non-physical security:

- Physical security considerations:
  - ***Location***…away from earthquake zones, flood zones etc…
  - ***Redundant*** utilities and network connections.
  - ***No Windows***
  - ***Limit Access Points***
  - ***Thick Exterior Walls***
  - ***Lots of Cameras***

- Non-Physical security considerations:
  - ***Isolation***…firewalls, access lists, VLANs, physical separation of resources
  - ***Policy Enforcement***…ACLs, 802.1x, malicious code detection
  - ***Visibility***…Ensuring security policy enforcement, identifying threats or potential threats

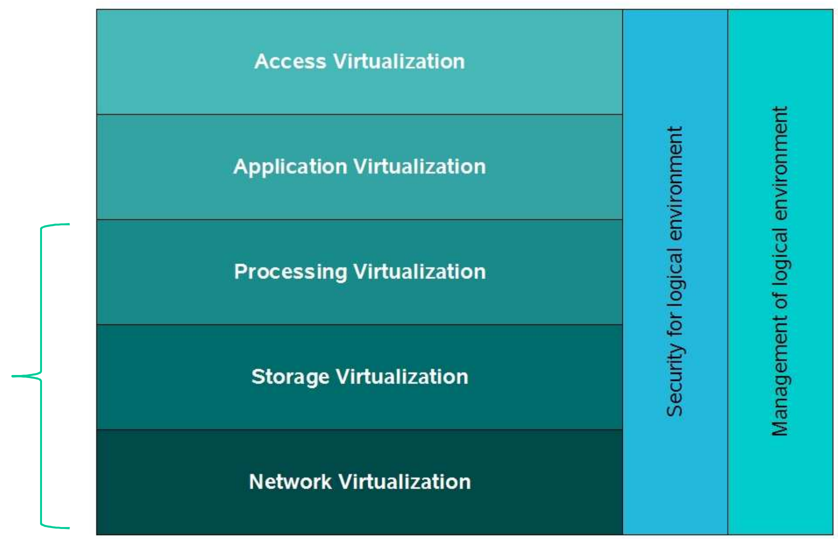


Cisco Model

Virtualization: ***Introduction***

- *Virtualization is the technique by which a physical entity can take on the appearance of multiple, fewer or different entities. This is done in the data center to improve efficiency, scaling, security, resiliency and/or administrative burden.*

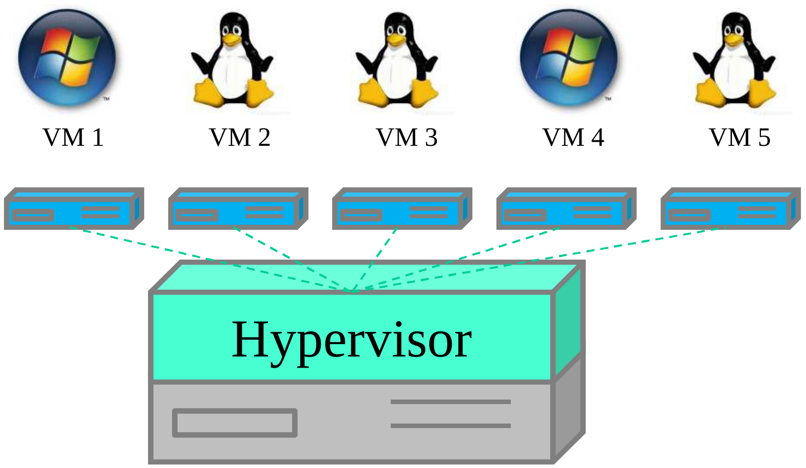
- We will examine virtualization as it applies to data center servers, networking and storage.



Kusnetzky Group © 2007

Virtualization:
Servers, Virtual Machines

- Virtual Machine (***VM***) – A virtual machine is a software implementation of a machine that executes programs like a physical machine.
- The software layer that provides server virtualization is called a ***hypervisor.***
- A hypervisor can run on bare metal (type 1) or over an existing operating system (type 2).
- Each virtual machine supports an independent instance of its own operating system.
- A virtual machine is lower performance than a physical machine in that the operating system accesses physical hardware indirectly (through the hypervisor). However, it allows servers to be more efficiently utilized for applications requiring less than the full capacity of the server.



Virtualization:
Servers: VMotion^{tm}

- VMotion^{tm} (Vmware) enables the live migration of virtual machines from one physical machine to another with zero downtime and transactional integrity.

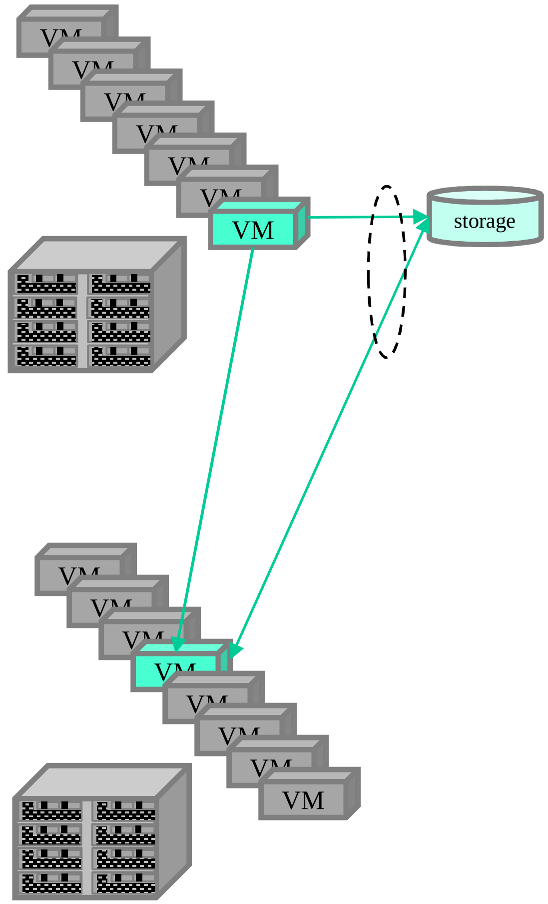
- The context of the original VM is copied as files into the data center storage (via Fibre Channel etc…).

- The local memory and execution state is copied to the target VM.

- The context is copied out of storage to the target VM.

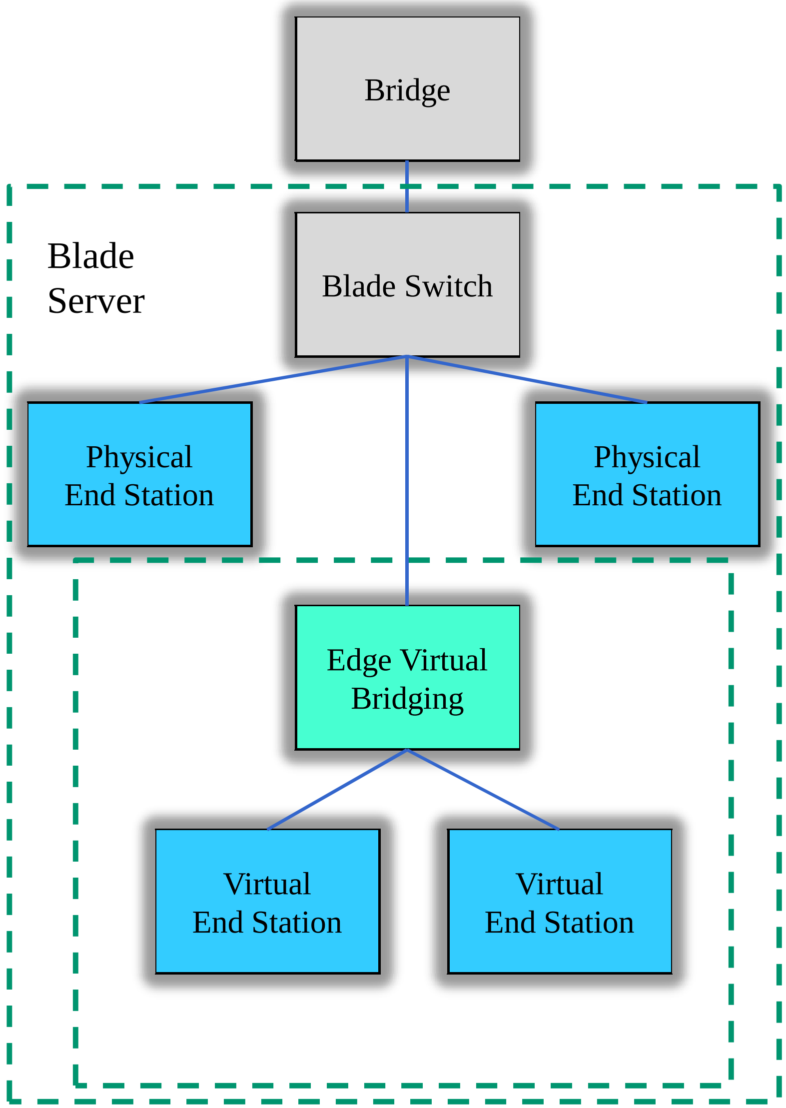
- The network association of the original VM is virtualized under the target VM and a virtual MAC address is assigned.

- The original VM is stopped and execution is initiated on the target VM.



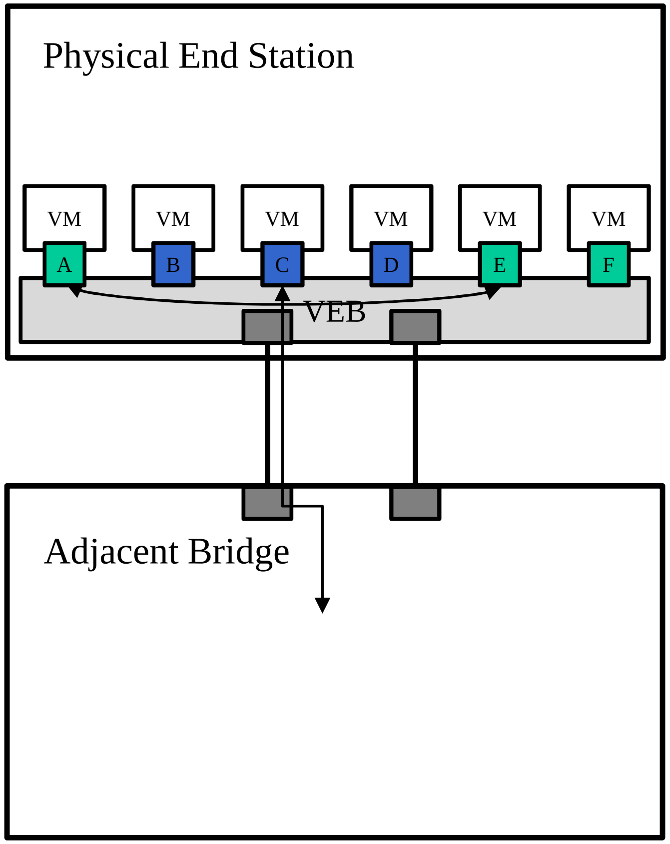
Virtualization:
Servers: Edge Virtual Bridging

- Edge Virtual Bridging (EVB) – is a virtualized switching environment where physical end stations contain multiple virtual end stations that participate in a virtual Ethernet network.



Virtualization:
Servers: Virtual Ethernet Bridge

- Virtual Ethernet Bridging (VEB) – is a virtual environment within a physical end station that supports local bridging between multiple virtual end stations.
- Forwarding is based upon MAC address and VLAN association (or port group association).
- Uplink to uplink bridging does not occur in the VEB (single active logical uplink).
- VEB does not participate in or affect spanning tree.
- VEB is not fully featured.



Virtualization:
Servers: Virtual Ethernet Port Aggregator

- Virtual Ethernet Port Aggregator (VEPA) – is a virtual environment within a physical end station that collaborates with an external adjacent bridge to provide bridging support between multiple virtual end stations and the external networks

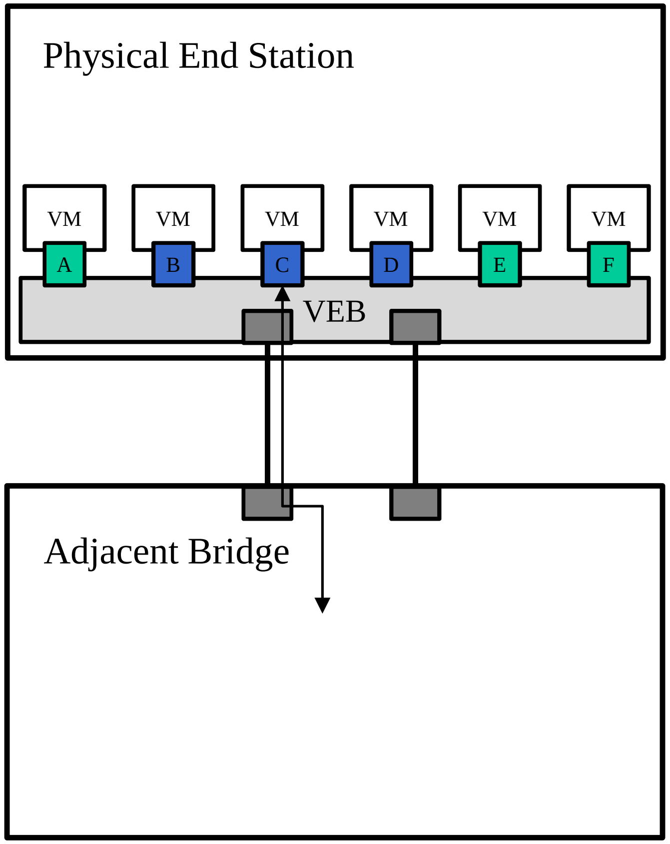
- Forwarding is based upon MAC address and VLAN association (or port group association).

- Forwarding is allowed between VM and the uplink and never between VM and VM. Does not forward between uplinks.

- VEPA does not participate in or affect spanning tree.

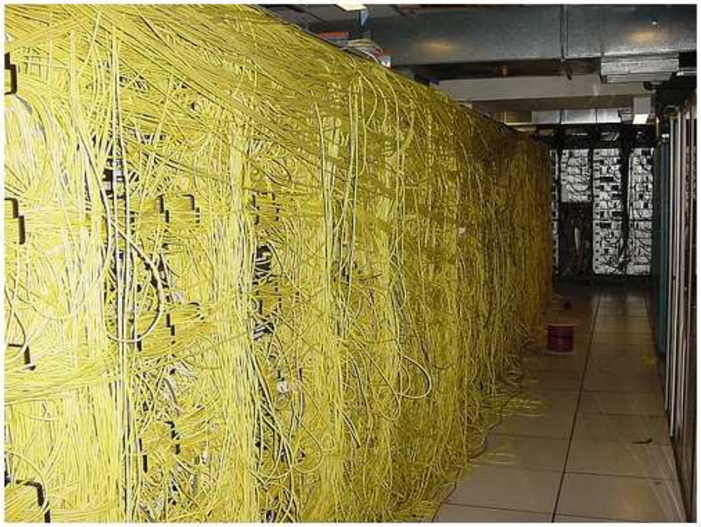
- VEPAs do not do learning.

- The advantage is that VM based forwarding can take advantage of all the features of the adjacent bridge



Virtualization: *Networking*

- Network virtualization can consolidate several physical networks into one virtual network. It can also divide a single physical network into multiple logical networks.
- Network virtualization lends itself to cost savings, security enhancement, efficiencies and flexibility.

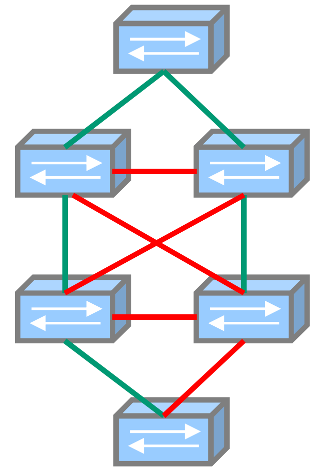


Virtualization:
Networking: Virtualized Switching System

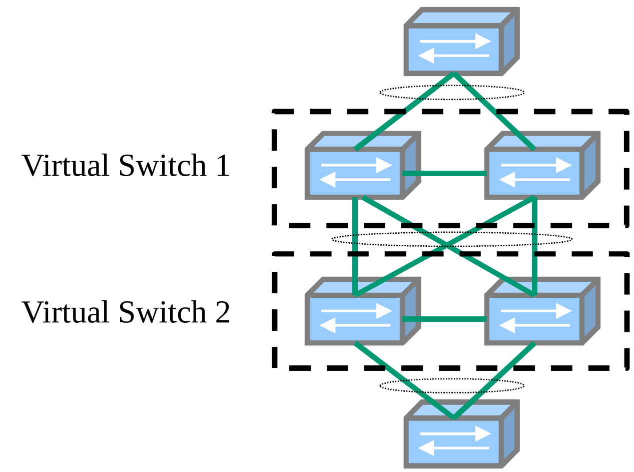
- Virtualized Switching System (VSS)* - allows two physical layer 2 switches to appear as a single virtualized layer 2 switch.
- The two member switches of the virtual switch use a virtual link to exchange control information to allow them to operate in coordination.
- The links between virtualized switches (and physical switches) can be aggregated into a multi-chassis link aggregation group.
- The advantage of virtualizing a pair of switches can be seen in the diagram:
  - In normal redundant configurations, spanning tree will prune links in order to prevent network loops.
  - The virtualized implementation locally hides interconnect that could be pruned by spanning tree, thereby significantly increasing network capacity and resiliency.

* Virtualized Switching System (VSS) is Cisco nomenclature. A number of vendors offer comparable functionality under different marketing names.

Physical Network
Spanning Tree View



Virtualized Network
Spanning Tree View



Virtualization:
Networking: Virtualized Routing and Forwarding

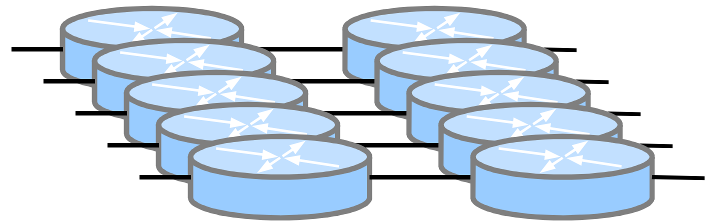
- Virtualized Routing and Forwarding (VRF) is a technology that allows multiple (independent) instances of routing tables to co-exist within the same physical router.

- The simplest implementation of VRF is VRF-lite*. In this implementation, each router in the virtual environment participates in a peer based fashion. VRF is usually associated with MPLS networks and VRF-lite is usually associated with networks that do not employ MPLS.

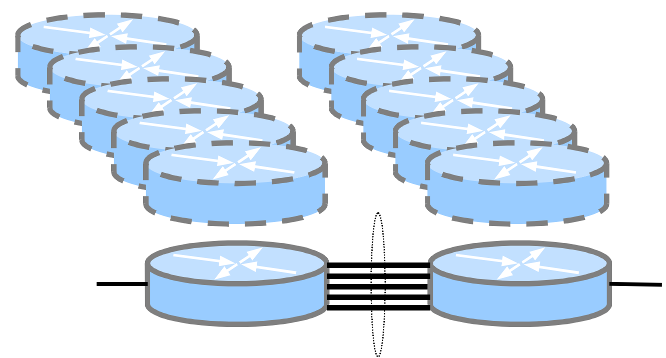
- VRF-lite creates independent instances of the route forwarding table for each VRF-lite virtualized router.

- Overlapping IP addresses are allows as each instance of the VRF-lite router is completely independent.

Physically Isolated
IP Routing



VRF-Lite Isolated
IP Routing

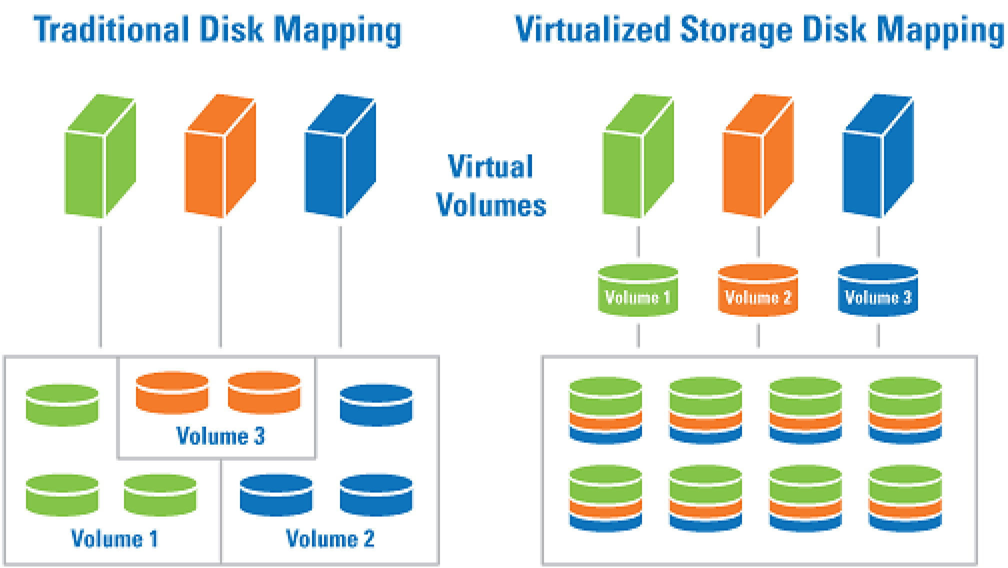


* VRF-lite is Cisco nomenclature. A number of vendors offer comparable functionality under different marketing names.

Virtualization:
Storage: Block Based Virtualization

- Block Virtualization refers to the abstraction of logical storage from physical storage so it may be accessed without regard to the physical storage.
- In block based virtualization, a single block is identified by a Logical Unit Number (LUN) and an offset known as a Logical Block Address (LBA).
- A lookup table maps the virtual LUN, LBA to the physical LUN, LBA:

- Receive a read request for vdisk LUN ID=1, LBA=33
- Perform a meta-data look up for LUN ID=1, LBA=33, and finds this maps to physical LUN ID=7, LBA1
- Sends a read request to physical LUN ID=7, LBA1
- Receives the data back from the physical LUN
- Sends the data back to the originator as if it had come from vdisk LUN ID=1, LBA33



Virtualization:
Storage: VMware Virtualization Hierarchy

- In the Vmware ESX environment, a virtual machine can attach to up to 4 virtual SCSI host bus adaptors (HBA).
- Virtual disks can be accessed by the virtual machine directly or through the virtual machine file system (VMFS). VMFS is a cluster file system. It is used to store virtual machine images to the physical disk pool.
- VMFS Features:
  - Automated file system with hierarchical directory structure
  - Optimization for virtual machines in a clustered environment
  - Lock management and distributed logical volume management
  - Dynamic data store expansion by spanning multiple storage extents
  - Clustered file system with journal logging for fast recovery
  - Encapsulation of the entire virtual machine state in a single directory

VMware Implementation