# Data Center Virtualization

Dr. Peter J. Welcher,
Chesapeake Netcraftsmen

Cisco Mid-Atlantic User's Group
Columbia MD – 4/27/10
Washington DC – 4/29/10

Slides Copyright 2009, Cisco, used with permission (and thanks).
Slides added by Netcraftsmen are identified.



## About the Speaker

- **Dr. Pete Welcher**
  - Cisco CCIE #1773, CCSI #94014, CCIP, CCDE written
  - Specialties: Network Design, Data Center, QoS, MPLS, Wireless, Large-Scale Routing & Switching, High Availability
  - Customers include large enterprises, federal agencies, hospitals, universities, cell phone provider
  - Taught many of the Cisco router / switch courses, developed some, including revisions to DESGN and ARCH courses
  - Reviewer for many Cisco Press books and book proposals
  - Presented lab session on MPLS VPN Configuration at Networkers 2005, 2006, 2007, BGP in 2008 and 2009, CCIP: Data Center Design in 2009
- **Over 27 blogs, 140 articles, prior seminars, posted**
  - http://www.netcraftsmen.net/welcher/

## Objectives

- In this presentation I hope to:
  - Look at why virtualization is needed and useful
  - Look at various types of virtualization with a Data Center focus
  - Discuss some design examples to share ideas on how virtualization might help in your network
  - Understand the benefits of vmware ESX and Cisco 1000v (and their network impact)
- Consequently:
  - The topic coverage will be broad not too deep
  - WAY too many slides, too little time – will present some slides quickly

Chesapeake NETCRAFTSMEN 3 CNC content Copyright 2010

## Agenda

- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion

Chesapeake NETCRAFTSMEN 4 Copyright 2010

## Virtualization: One Definition

**Virtualization**
is the pooling and abstraction of resources and services in a way that masks the physical nature and boundaries of those resources and services from their users

http://www.gartner.com/DisplayDocument?id=399577




## What's Virtualization?

- One as many
  - Single physical device acting as multiple virtual devices
    - E.g. contexts (ASA contexts, Nexus 7K vDC's, …)
  - VMWare and servers as VM's
  - VLANs, VRFs segmenting single links and/or routers
- Many as one
  - Clustering / stacking, whereby multiple physical boxes become logically one virtual box
  - Example: 6500 VSS, Nexus vPC
- Emulation
  - Example: pseudo-wires (EoMPLS, etc.)

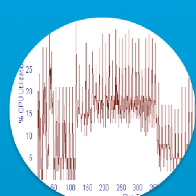

CNC content

## Why Virtualize?

- Servers
  - One app, one box
  - (Seriously underused hardware) x (many boxes)
  - One app per blade continues that trend
- "Death by (small) boxes" (servers, network)
  - Device count drives up Operations costs
- Underused boxes cost:
  - Procurement system costs, purchase price, vendor support, admin, space, power, cabling, operations support, ....
- However, separate boxes are sometimes used to reduce complexity
  - Everything in one (two) chassis means you have to be careful with those chassis
- Compromise: LOGICALLY separate boxes, or virtualization

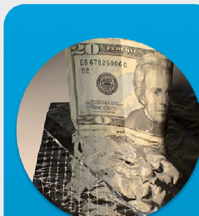


## Four Drivers Behind Virtualization



**Hardware Resources Underutilized**
- CPU utilizations ~ 10% - 25%
- One server – One Application
- Multi-core even more under-utilized

**Data Centers are running out of space**
- Last 10+ years of major server sprawl
- Exponential data growth
- Server consolidation projects just a start



**Rising Energy Costs**
- As much as 50% of the IT budget
- In the realm of the CFO and Facilities Mgr. now!

**Administration Costs are Increasing**
- Number of operators going up
- Number of Management Applications going up

**Operational Flexibility**

## But It's Not Just Servers!

- Clutter of many project-specific Server Load Balancers
  – MS or Linux Load Balancing, various vendor appliances, now virtualized SLB appliances, Cisco CSM's, …
- Firewalls proliferating
  – Firewall contexts
- Replication of environments
  – Dev, Test, Prod: similar, sometimes hand-me-down hardware
  – Can use separate contexts instead

9 CNC content Copyright 2010

## Other Significant Benefits

- Virtualization addresses several key aspects:
  – Ability to quickly spawn test and development environments
  – Provides failover capabilities to applications that can't do it natively
  – Maximizes utilization of resources (compute & I/O capacity)
  – Server portability (migrate a server from one host to the other)
- Virtualization is not limited to servers and OS
  – Network
  – Storage
  – Application
  – Desktop

Isolation
Roll-Back
Abstraction
Portability
Deployment

10 Copyright 2010

## Data Center Building Blocks

**Applications**

**Application Networking Services**
Application Delivery and Application Optimization

**Virtualization**
Network, Server, Storage and Management

| Transport Infrastructure<br>Eth, FC, DCE, WAN, MAN | Compute Infrastructure<br>OS, Hardware, Firmware | Storage Infrastructure<br>SAN, NAS, DAS |
|---|---|---|

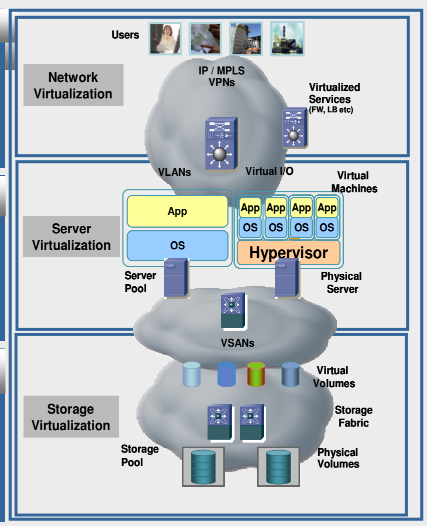


## Virtualization Is Not Limited to OS/Server

**Network Virtualization**
- Segmentation and security
- Higher resource flexibility
- Improved capacity utilization

**Server Virtualization**
- Consolidation of physical servers
- Virtual Machine mobility
- Rapid application deployment with VMs

**Storage Virtualization**
- Segmentation and security
- Improved data mgmt. & compliance
- Non-disruptive provisioning & migration



## Impact of VMWare

- Right now, server virtualization is driving a lot of change
  - VMWare gives the ability to change server sizing and storage sizing issues without disruption
  - vMotion gives the ability to take physical chassis out of service without service disruption
  - Not to mention load-shifting, high availability for VM's, etc.
  - Some costs are:
    - Data center infrastructure designs changing rapidly
    - Need to manage VM proliferation, use of shared resources (CPU, RAM, SAN)
- Coming next:
  - Data center virtualization (clouds)
    - Modulo addressing cloud security considerations
  - Per-application infrastructure virtualization

Chesapeake NETCRAFTSMEN 13 CNC content Copyright 2010

## Some Observations

- Hidden lesson (to me): automation requires NOT hand-crafting solutions
- Needed: system + network + SAN architecture
  - (or a small set of architectures)
  - Think: application or service required components description (along with how they fit together)
- Stop doing one-offs
  - Do a small number of variations of hardware environments supporting software environments
- Racking, cabling costs (and labor time) are getting too expensive
  - Avoid them via virtualization
  - Use less cabling (10+ G links, FCoE)

Chesapeake NETCRAFTSMEN 14 CNC content Copyright 2010

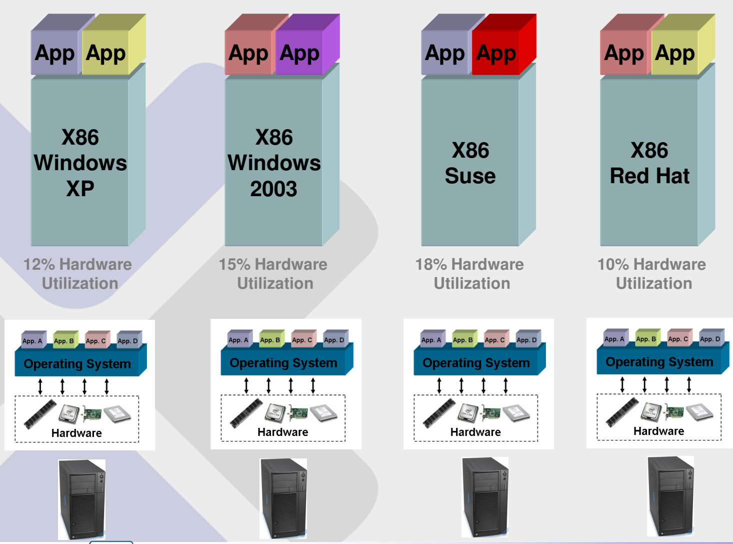
## Agenda

- Virtualization: Getting Motivated!
- **Compute Resource Virtualization**
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion

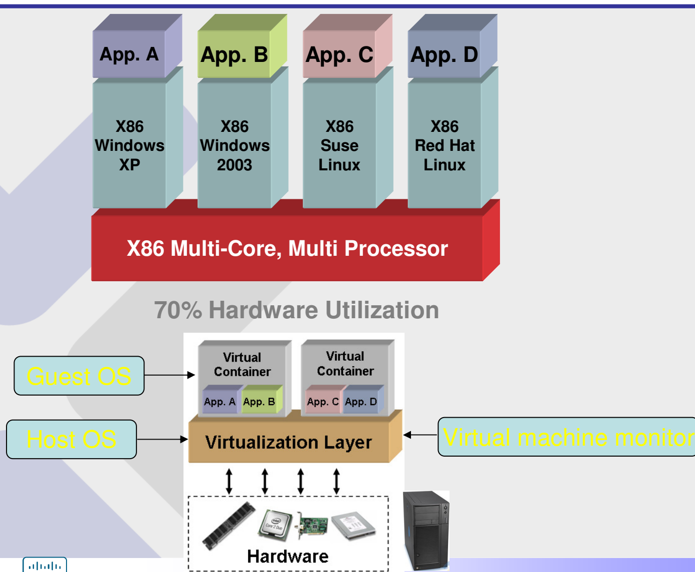


## Going from Here…

Evolution of Virtualization



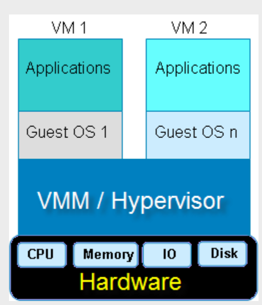
## … to There





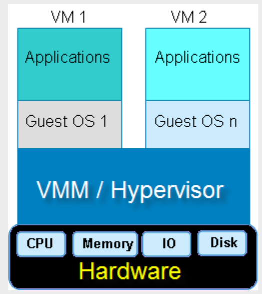

## Native/Full Virtualization (Type-1)

- VMM runs on 'bare metal'
- VMM virtualizes (emulates) hardware
  - Virtualizes x86 ISA, for example
- Guest OS unmodified
- VMs: Guest OS+Applications run under the control of VMM
- Examples
  - VMware ESX Server, Microsoft Hyper-V
  - IBM z/VM
  - Linux KVM (Kernel VM)



## What to Virtualize

- Ideally all components
- CPU
  - –Privileged Instructions
  - –Sensitive Instructions
- Memory
- I/O
  - –Network
  - –Block/Disk
- Interrupt





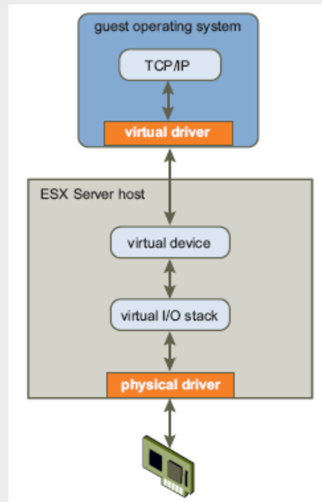


## A Closer Look at VMware's ESX™

- Full virtualization
  - –Runs on bare metal
  - –Referred to as 'Type-1 Hypervisor'
- ESX is the OS (and of course the VMM)
  - –ESX has Linux scripting / shell capabilities
  - –ESXi does not – smaller, less « attack surface »
- ESX handles privileged executions from Guest kernels
  - –Emulates hardware when appropriate
- Uses 'Trap and Emulate' and 'Binary Translation'
- Guest OS run as if it were business as usual
  - –Except they really run in user mode (including their kernels)

## What About Networking?

- Users naturally expect VMs to have access to network
- VMs don’t directly control networking hardware
  - –Physical NIC is usually shared between multiple VMs
- When a VM communicates with the outside world, it:
  - –... passes the packet to its local device driver ...
  - –... which in turns hands it to the virtual I/O stack
  - –... which in turns passes it to the physical NIC
- ESX gives VMs several device driver options:
  - –Strict emulation of Intel’s e1000
  - –Strict emulation of AMD’s PCnet 32 Lance
  - –VMware vmxnet: paravirtualized!
- VMs have MAC addresses that appear on the wire

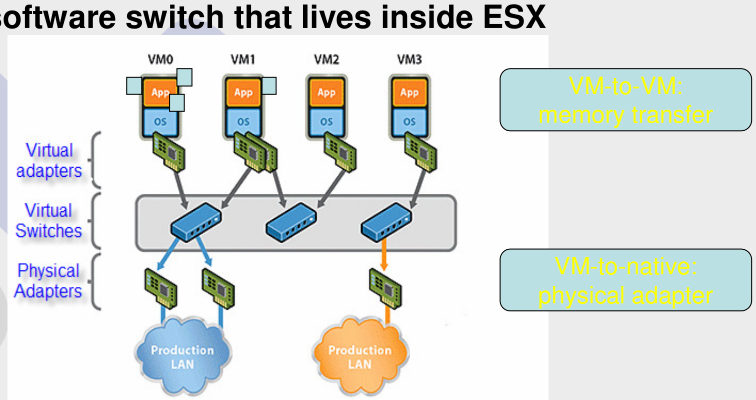







## Virtual Adapters and Virtual Switches

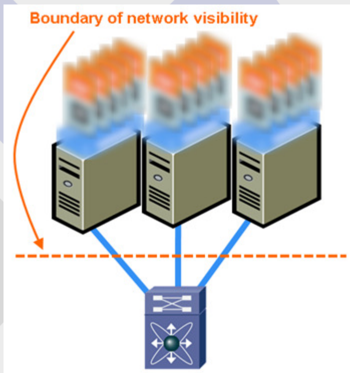
- VM-to-VM and VM to native host traffic handled via software switch that lives inside ESX




- Note: speed and duplex are irrelevant with virtual adapters http://www.vmware.com/files/pdf/virtual_networking_concepts.pdf

## What Does This Mean for the LAN Admin?

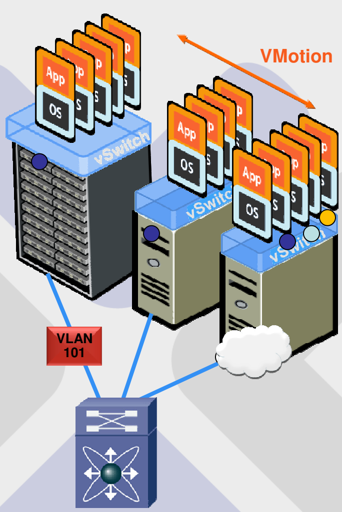
- **To the LAN administrator, the picture is blurry**




- LAN role typically limited to provisioning a trunk to ESX
- No visibility into VM-to-VM traffic
- Troubleshooting performance or connectivity issues challenging



## Solution: Cisco's Virtual Switch (Nexus 1000-V)




Problems:

- **VMotion may move VMs across physical ports—policy must follow**
- **Impossible to view or apply policy to locally switched traffic**
- **Cannot correlate traffic on physical links—from multiple VMs**

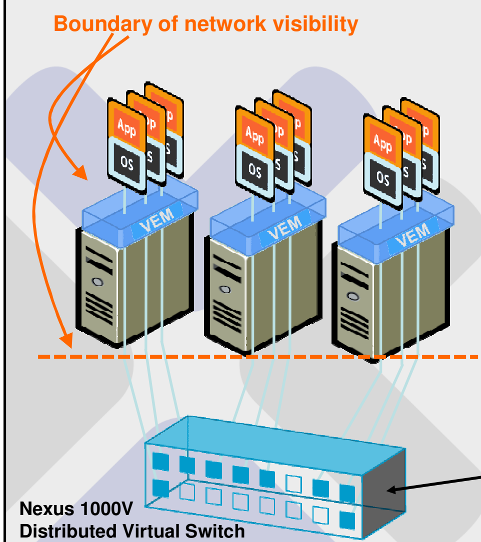
Nexus 1000-V:

- **Extends network to the VM**
- **Consistent services**
- **Coordinated, coherent management**



## Virtual Networking with Cisco's Nexus 1000-V

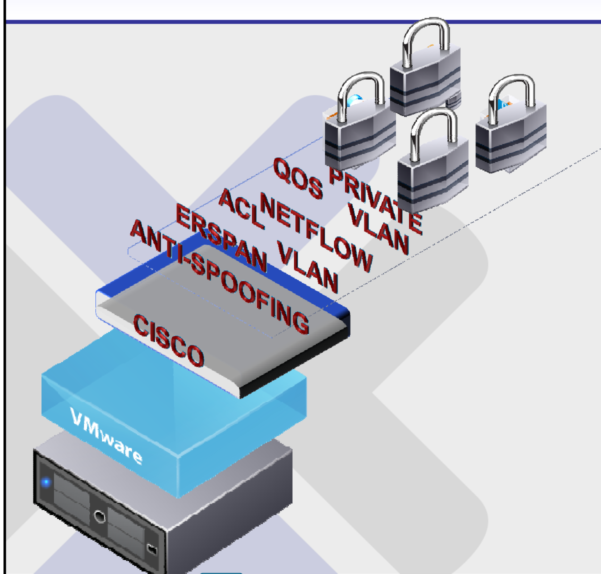



- Nexus 1000V provide visibility down to the individual VMs
- Policy can be configured per-VM
- Policy can move around within the ESX cluster

Cisco NX-OS
Command Line Interface!



## Nexus 1000V Key Features

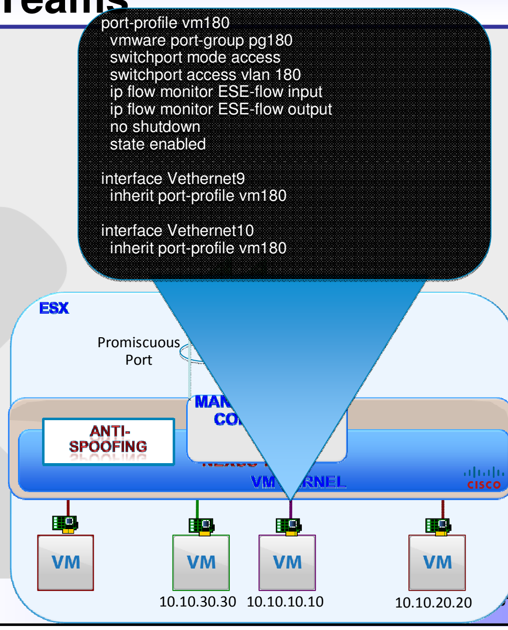



- Includes Key Cisco Network & Security features
- Addressing Issues for:
  - VM Isolation
  - Separation of Duties
  - VM Visibility

## Separation of Duties: Network and Server Teams

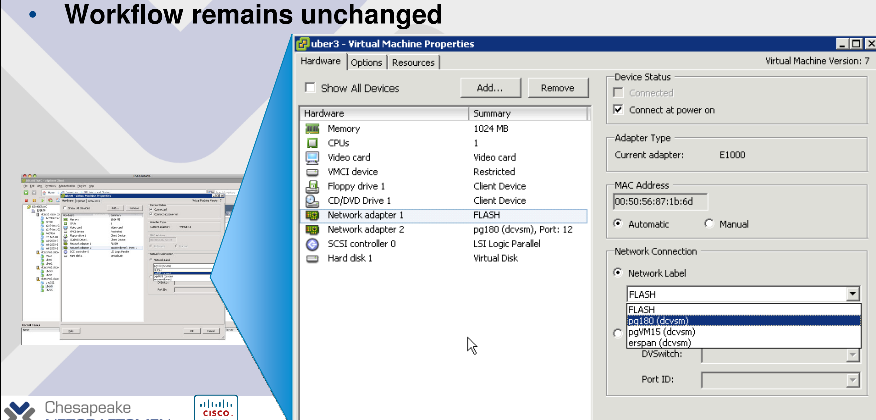
### Port Profiles

- A network feature macro
- Example: Features are configured under a port profile once and can be inherited by access ports
- Familiar IOS look and feel for network teams to configure virtual infrastructure

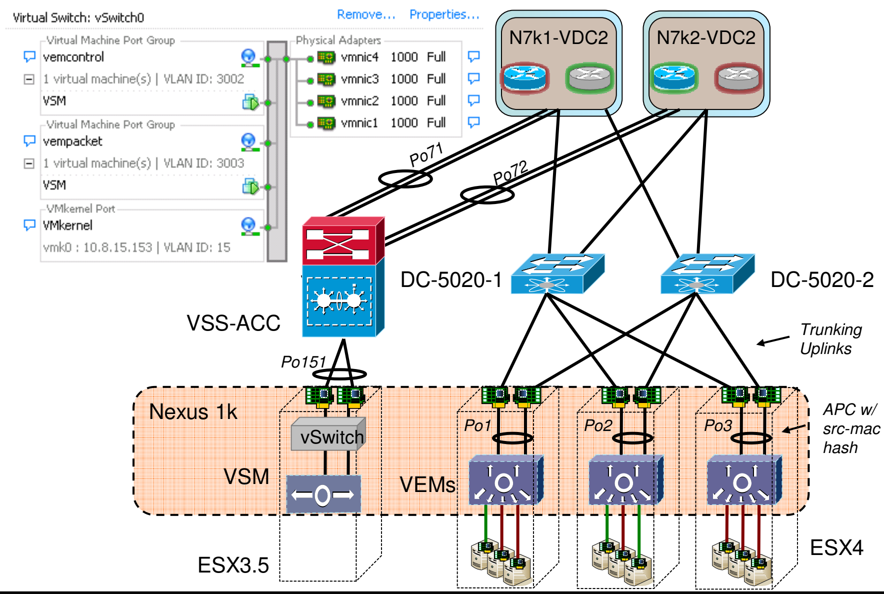



## Separation of Duties: Network & Server Teams

- Nexus 1000V automatically enables port groups in Virtual Center via API
- Server Admin uses Virtual Center to assign vnic policy from available port groups
- Nexus 1000V automatically enables VM connectivity at VM power-on
- Workflow remains unchanged



## Virtual Access Layer Nexus 1000v




## Lessons Learned

- Data Center moves
- If you don't have 1000v or comparable information, replacing & troubleshooting get "interesting"
  - You need to know EXACTLY which blade NIC is cabled to which switch port
  - The switch config is, in effect, your documentation
    - Impossible to technically verify active / passive vNICs
    - Alternative: extensive server admin discussions and followup
  - To upgrade such a switch, map old port to new port, replicate features, cross your fingers…
  - Your visibility and control is really per blade server, not per-VM

## Some Other Thoughts

- VMotion requires SAN
  - In some form (iSCSI, NFS, FC, etc.)
  - Claimed that well-designed iSCSI and NFS can give performance comparable to FC
  - Except perhaps for high-end servers with high IO rates
- Tiered SAN expected
  - Less costly approaches where suitable
  - FC / high performance arrays, etc. where needed
- VMotion for Storage requires SAN
  - Provides flexible re-allocation of disk resources
  - Non-disruptive if done properly



## Agenda

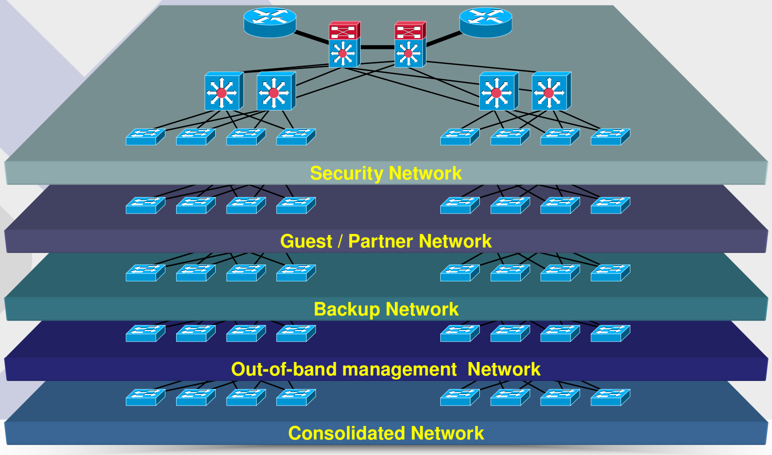
- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion



# What Is Network Virtualization?

- Overlay of physical topologies (N:1)
- N physical networks maps to 1 physical network

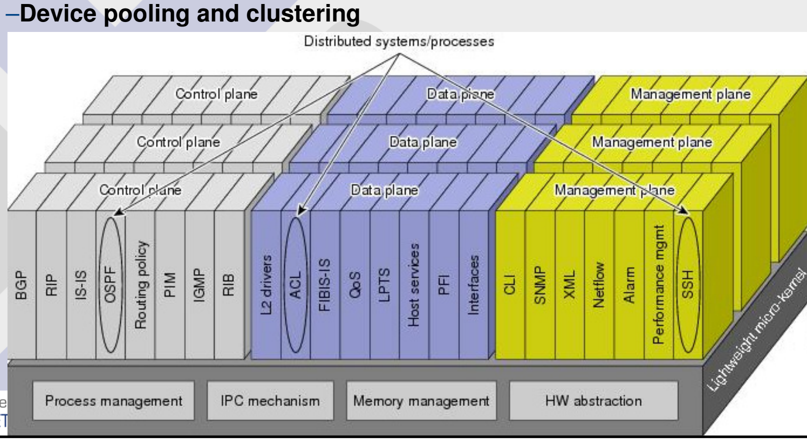






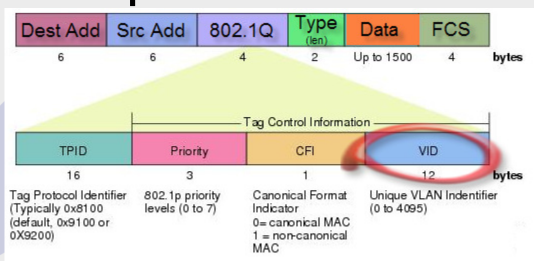
# Network Virtualization Classification

- Generally speaking, four areas in network virtualization
  - –Control-plane virtualization
  - –Data-plane virtualization
  - –Management plane virtualization
  - –Device pooling and clustering

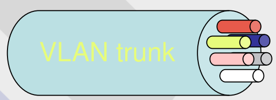


## Data Plane Virtualization

- Simple example: Virtual LANs




- 802.1Q: 12 bits → up to 4096 VLANs on same physical cable

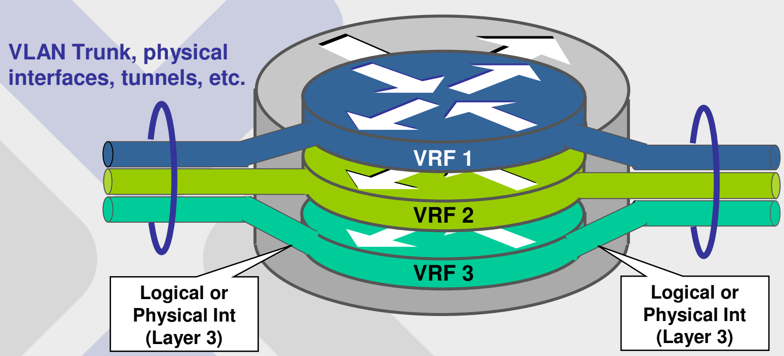






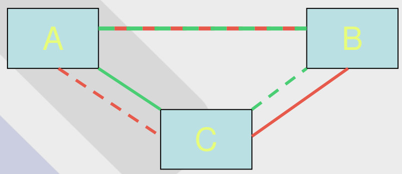
## Another Data Plane Virtualization Example

- The VRF: Virtual Routing and Forwarding instance



## Control-Plane Virtualization 'for VLANs'

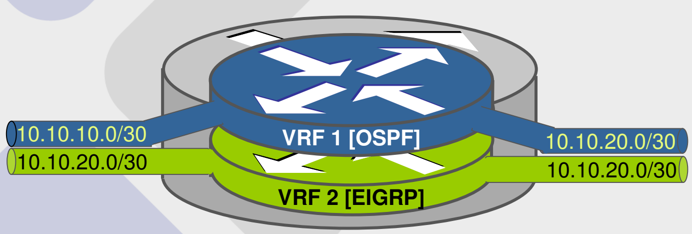
- Example: Spanning-tree protocol
  - Loop-breaker in Ethernet topologies
- How is it virtualized?
  - Per-VLAN spanning-tree
- What's in it for me?
  - Allows multiple logical topologies to exist on top of one physical topology using the Good Old 'odd/even' VLAN balancing scheme






## Control-Plane Virtualization 'for VRFs'

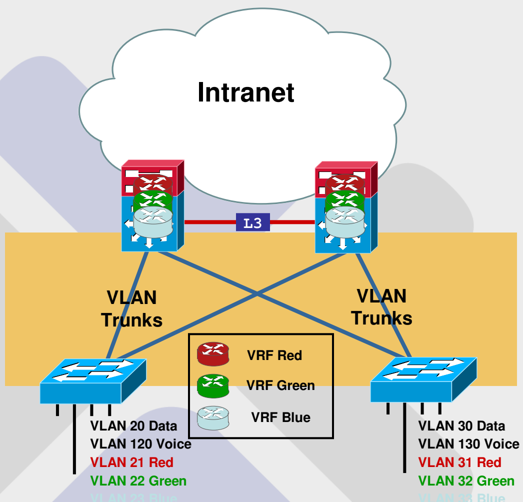
- Example: per VRF routing protocol
  - One VRF could run OSPF while another runs EIGRP
- Goal
  - Isolation of routing and forwarding tables
  - Allows overlapping IP addresses between VRFs





# Intersection of VLANs and VRFs

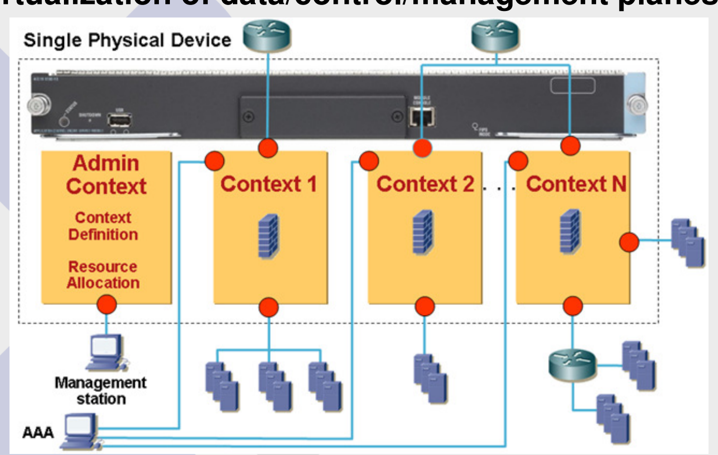



- It is easy to map VLANs to VRFs at the distribution layer
- Provides safe and easy way to isolate logical networks
- No uncontrolled leaking from one to the other
- Maximizes use of physical infrastructure



# ASA / FWSM: Device Partitioning

- Example: Firewall Services Module virtual contexts
- Virtualization of data/control/management planes



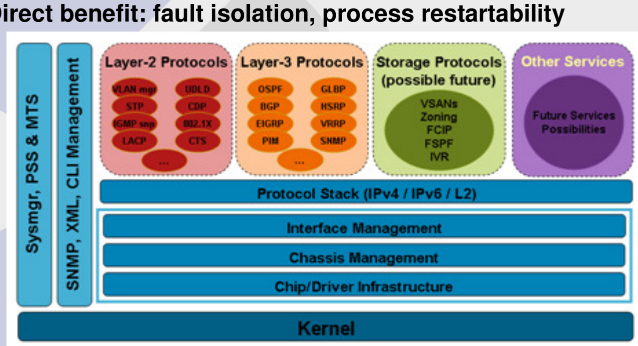
# FWSM Example: Device Partitioning

- Mix of control, data and management plane virtualization techniques
  - –The ‘changeto’ command: switch from one context to the other, similar to running multiple terminal session on a Linux system
- Not much in common with OS/Server virtualization
  - –No isolation between contexts, no VMM, single OS image
  - –Not even a one-to-one mapping between a process and a context
- Virtualization here is essentially a classification problem
  - – Inbound interface, destination MAC address
  - →These two values allow data plane to assign traffic to right context
  - →Concept of virtual interface throughout the packet processing chain
- CNC Homework: how does the ASA differ?



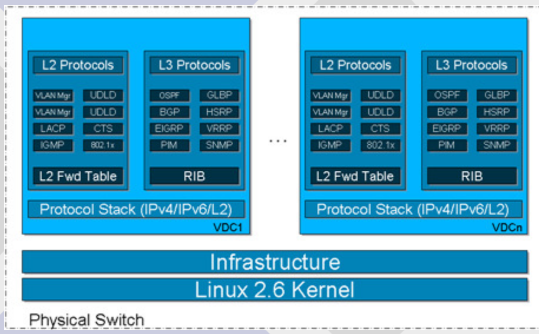
# Another Example: Nexus 7000

- Nexus 7000 runs Cisco’s NXOS
  - –Very different internal architecture compared to classic IOS
- NXOS is a true multiprogramming OS
  - –Features a Linux kernel and user-space processes
  - –Most features (BGP, HSRP, EIGRP, etc.) are individual processes
  - –Direct benefit: fault isolation, process restartability



# Nexus 7000's Virtual Device Contexts

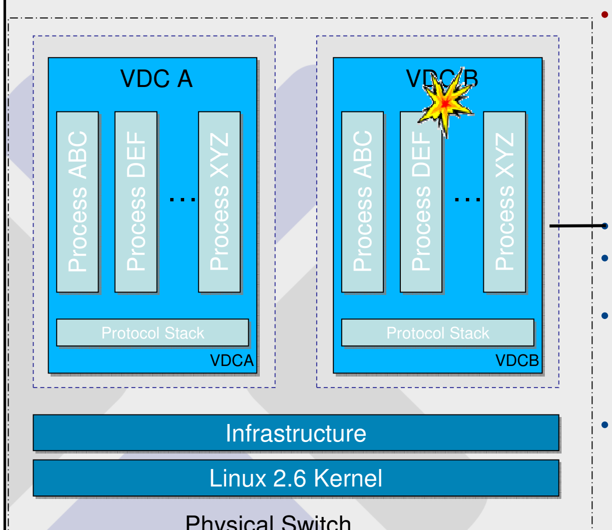
- OS and hardware architecture allow robust virtualization implementation
- VDC concept: up to 4 individual partitions
  - Concept of switchto/switchback and per-vDC access/isolation
- Somewhat like host-based virtualization






# NX-OS Virtual Device Contexts

VDC Fault Domain

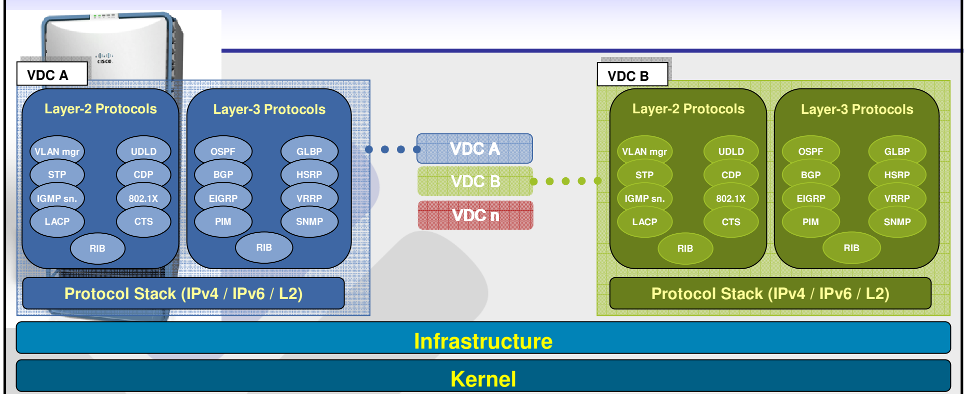



- A VDC builds a Fault Domain around all running processes within that VDC — faults within a running process or entire VDC are isolated from other device contexts
- Fault Domain
- Process "DEF" in VDC B crashes
- Processes in VDC A are not affected and will continue to run unimpeded
- This is a function of the process modularity of the OS and a VDC specific IPC context



## Virtual Device Contexts (VDCs)

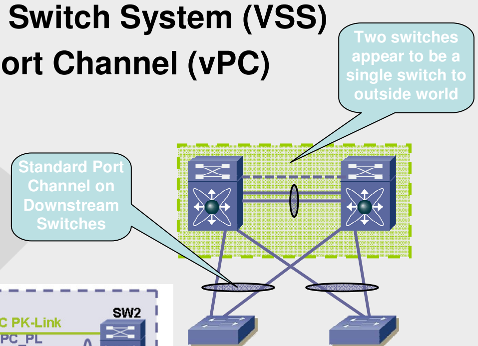


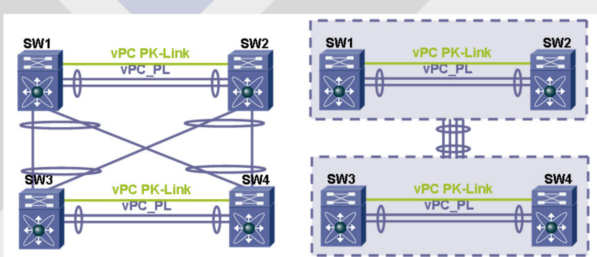

- VDC—Virtual Device Context
  - **Flexible separation/distribution of** Software Components
  - **Flexible separation/distribution of** Hardware Resources
  - **Securely delineated** Administrative Contexts
- VDCs are not…
  - **The ability to run different OS levels on the same box at the same time**
  - **Similar to host-based OS virtualization: single 'hypervisor' handles all h/w resources**



## Device Pooling and/or Clustering

- **Catalyst 6500's Virtual Switch System (VSS)**
- **Nexus 7000's Virtual Port Channel (vPC)**
- **It's really clustering**
- **Clever packet classification**



# Agenda

- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion



# Current Network Challenges

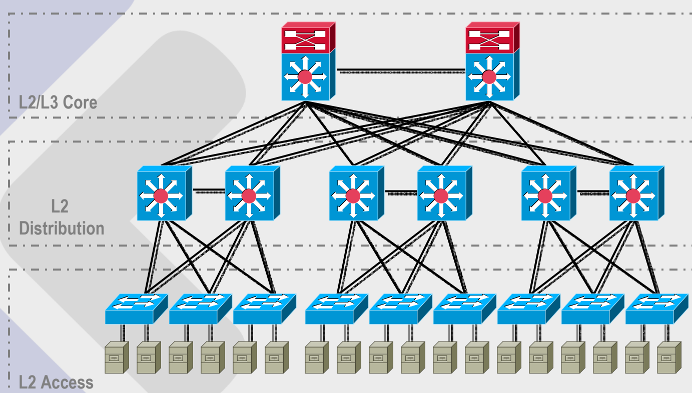
## Data Center

**Traditional Data Center designs are requiring ever increasing Layer 2 adjacencies between server nodes due to prevalence of virtualization technology. However, they are pushing the limits of Layer 2 networks, placing more burden on loop-detection protocols such as Spanning Tree…**

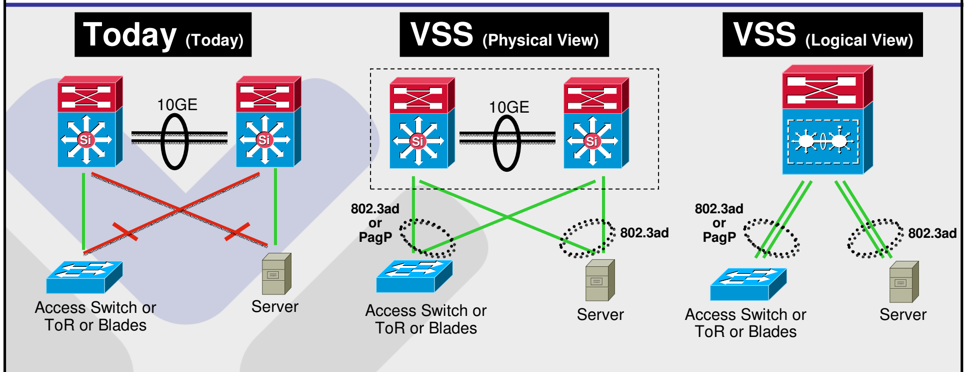
FHRP, HSRP, VRRP
Spanning Tree
Policy Management

Single active uplink per VLAN (PVST), L2 reconvergence, excessive BPDUs

Dual-Homed Servers to single switch, Single active uplink per VLAN (PVST), L2 reconvergence



## Catalyst 6500 Virtual Switching System 1440
Overview

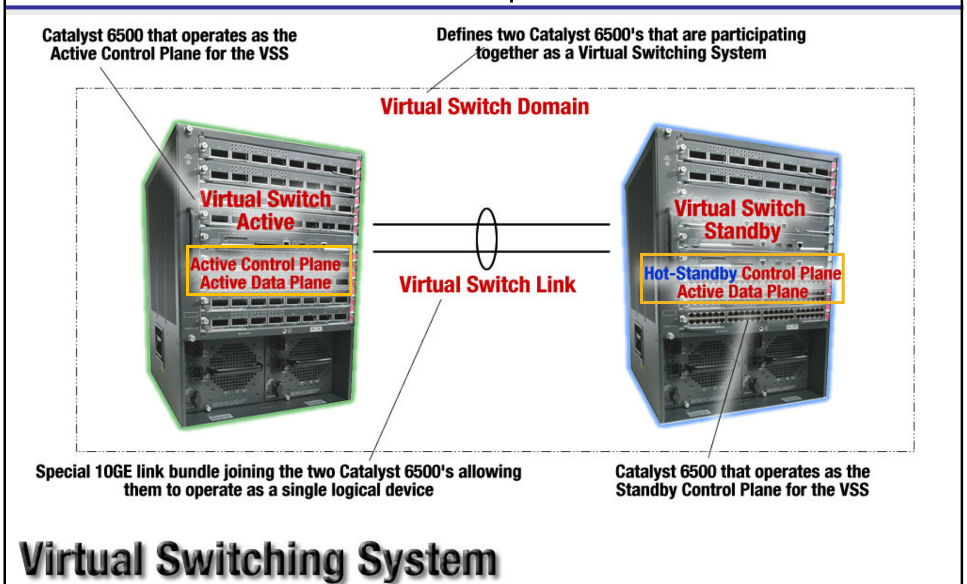



**Simplifies** operational Manageability via Single point of Management, Elimination of STP, FHRP etc

**Doubles** bandwidth utilization with Active-Active Multi-Chassis Etherchannel (802.3ad/PagP) Reduce Latency

**Minimizes** traffic disruption from switch or uplink failure with Deterministic subsecond Stateful and Graceful Recovery (SSO/NSF)

## Introduction to Virtual Switching System
Concepts



# Virtual Switching System

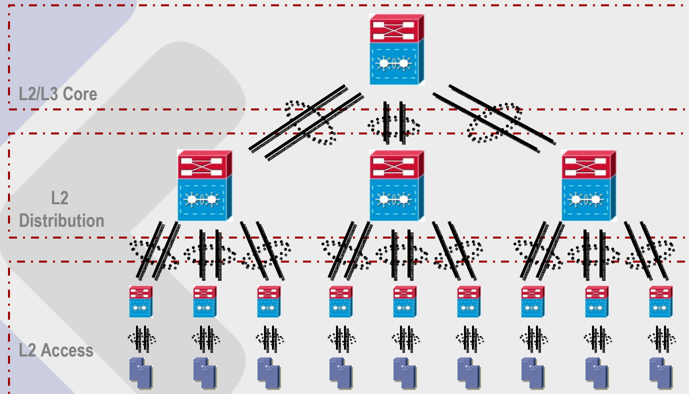
## Data Center

**A Virtual Switching System-enabled Data Center allows for maximum scalability so bandwidth can be added when required, but still providing a larger Layer 2 hierarchical architecture free of reliance on Spanning Tree…**

Single router node, Fast L2 convergence, Scalable architecture

Dual Active Uplinks, Fast L2 convergence, minimized L2 Control Plane, Scalable

Dual-Homed Servers, Single active uplink per VLAN (PVST), Fast L2 convergence

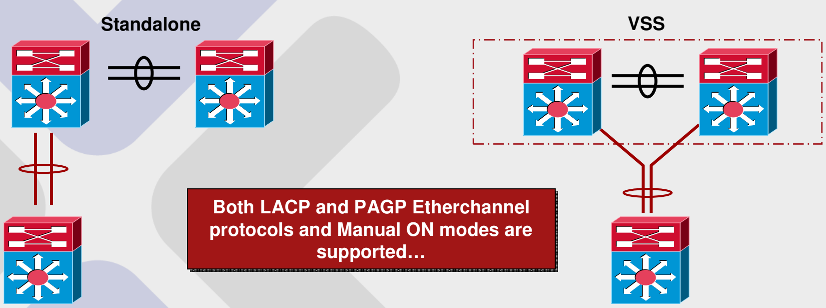





# Virtual Switching System Architecture

## Multichassis EtherChannel (MEC)

**Prior to Virtual Switching System, Etherchannels were restricted to reside within the same physical switch. In a Virtual Switching environment, the 2 physical switches form a single logical network entity - therefore Etherchannels can now also be extended across the 2 physical chassis…**




**Both LACP and PAGP Etherchannel protocols and Manual ON modes are supported…**

**Regular Etherchannel on single chassis**

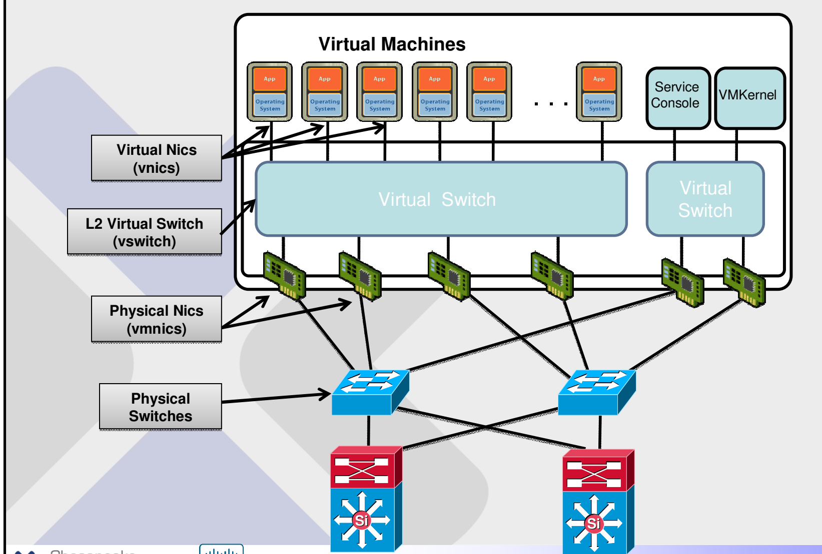
**Multichassis EtherChannel across 2 VSS-enabled chassis**



# Overview

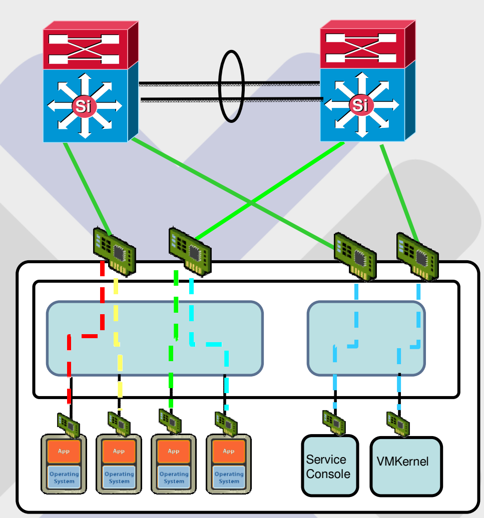
## VMWare ESX Virtual Networking






# VM Based NIC Teaming

## Virtual Port-ID Based or Virtual MAC-Based



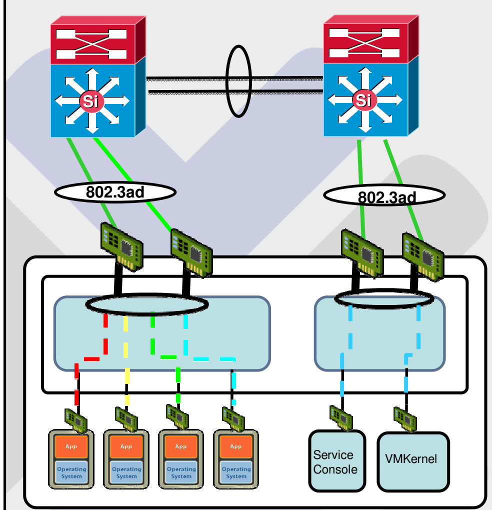

### Advantages

- Switch Redundancy

### Disadvantages

- Unequal traffic distribution possible
- VM bandwidth limited to mapped physical NIC capacity
- VMotion/IP Storage limited to 1 Physical NIC bandwidth capacity

## VM Based NIC Teaming
## IP Hash NIC Teaming




**Advantages**

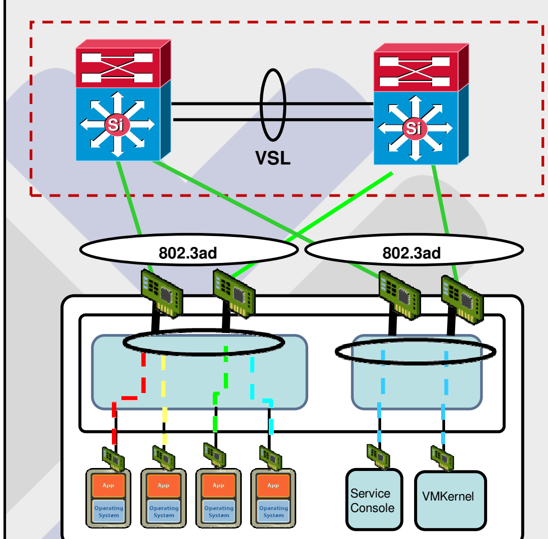
- Better Bandwidth availability for VM and Service/VMotion/IP Storage Traffic

**Disadvantages**

- No Switch Redundancy

Chesapeake NETCRAFTSMEN 55 Copyright 2010

## VM Based NIC Teaming
## NIC Teaming Across VSS Catalyst Switches




- Maximum Bandwidth for VM and Service/VMotion/IP Storage Traffic with Granular Load Balancing
- Increased Availability with Link Aggregation Across Two Separate Physical Catalyst 6500
- Simpler configuration on Catalyst Switch
- Maintains separation between VM traffic and Service/VMotion/IP Storage traffic
- Allows scaling VM traffic and Service/VMotion/IP Storage to all available NICs

Chesapeake NETCRAFTSMEN 56 Copyright 2010

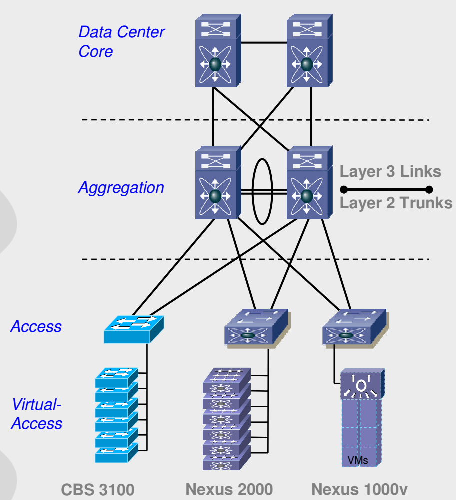
## Agenda

- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion



## Building the Access Layer using Virtualized Switching

- Virtual Access Layer
  - Still a single logical tier of layer-2 switching
  - Common control plane with virtual hardware and software based I/O modules
- Cisco Nexus 2000
  - Switching fabric extender module
  - Acts as a virtual I/O module supervised by Nexus 5000
- Nexus 1000v
  - Software-based Virtual Distributed Switch for server virtualization environments.

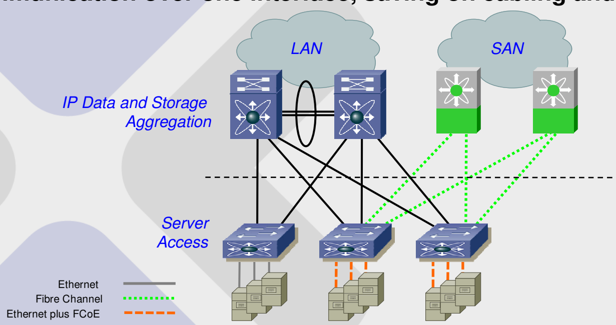




# Migration to a Unified Fabric at the Access Supporting Data and Storage

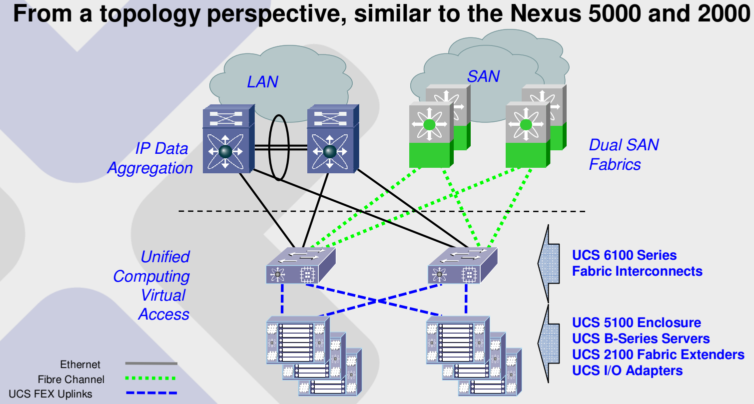
- Nexus 5000 Series switches support integration of both IP data and Fibre Channel over Ethernet at the network edge
- FCoE traffic may be broken out on native Fibre Channel interfaces from the Nexus 5000 to connect to the Storage Area Network (SAN)
- Servers require Converged Network Adapters (CNAs) to consolidate this communication over one interface, saving on cabling and power






# Cisco Unified Computing System (UCS)

- A cohesive system including a virtualized layer-2 access layer supporting unified fabric with central management and provisioning
- Optimized for greater flexibility and ease of rapid server deployment in a server virtualization environment
- From a topology perspective, similar to the Nexus 5000 and 2000 series

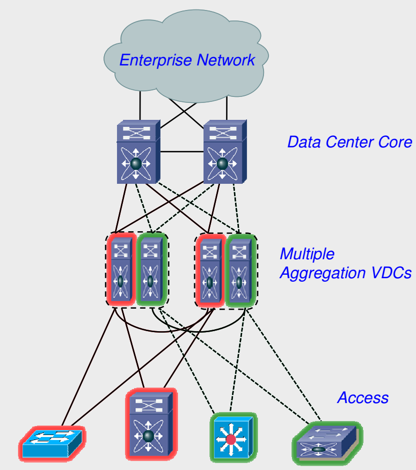




# Virtual Device Context Example: Multiple Aggregation Blocks

- Single physical pair of aggregation switches used with multiple VDCs
  - Access switches dual-homed into one of the aggregation VDC pairs
  - Aggregation blocks only communicate through the core layer
- Design considerations:
  - Ensure control plane requirements of multiple VDCs do not overload Supervisor or I/O Modules
  - Where possible consider dedicating complete I/O Modules to one VDC (CoPP in hardware per-module)
  - Ports or port-groups may be moved between aggregation blocks (DC pods) without requiring re-cabling






# Agenda

- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion

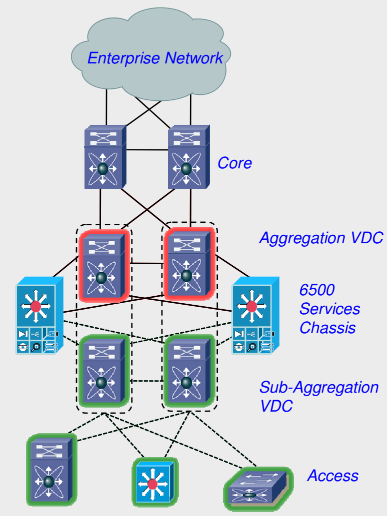


# Virtual Device Context Example:

Services VDC Sandwich

- **Multiple VDCs used to "sandwich" services between switching layers**
  - **Allows services to remain transparent (layer-2) with routing provided by VDCs**
  - **May be leveraged to support both services chassis and appliances**
- **Design considerations:**
  - **Access switches requiring services are connected to sub-aggregation VDC**
  - **Access switches not requiring services may be connected to aggregation VDC**
  - **Allows firewall implementations not to share interfaces for ingress and egress**
  - **Facilitates virtualized services by using multiple VRF instances in the sub-aggregation VDC**

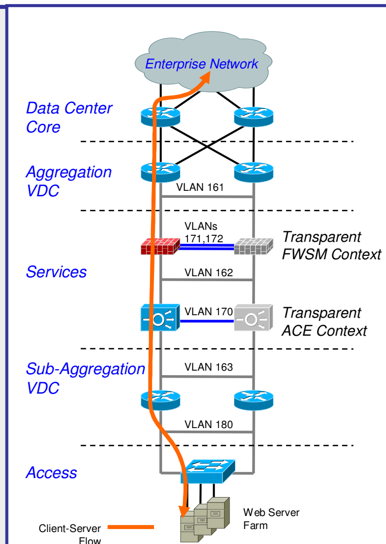






# Using Virtualization and Service Insertion to Build Logical Topologies

- **Logical topology example using services VDC sandwich physical model**
  - **Layer-2 only services chassis with transparent service contexts**
  - **VLANs above, below, and between service modules are a single IP subnet**
  - **Sub-aggregation VDC is a layer-3 hop running HSRP providing default gateway to server farm subnets**
  - **Multiple server farm VLANS can be served by a single set of VLANs through the services modules**
  - **Traffic between server VLANs does not need to transit services device, but may be directed through services using virtualization**

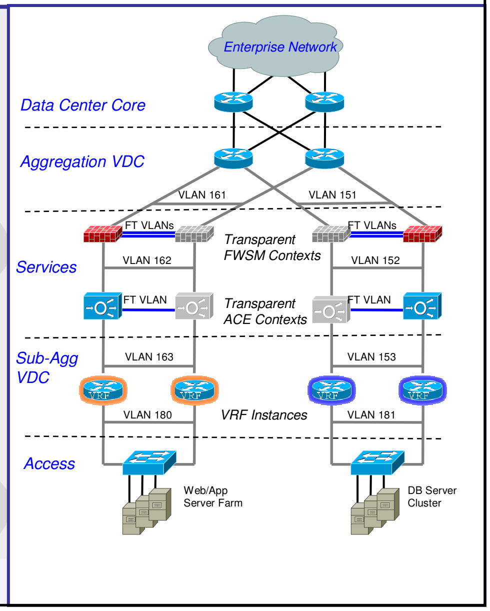




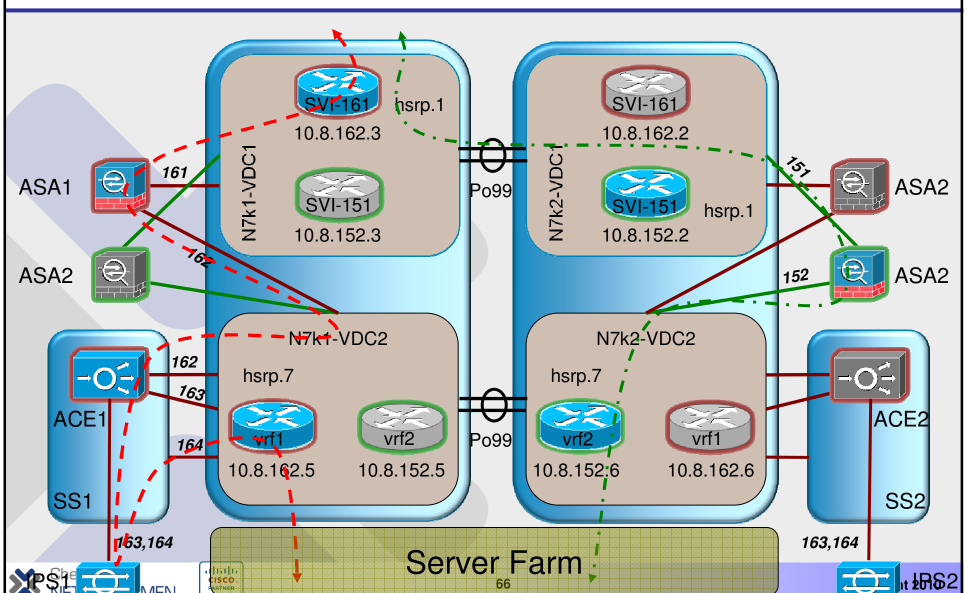

 

# Using Virtualization and Service Insertion to Build Logical Topologies

- Logical Topology to support multi-tier application traffic flow
  – Same physical VDC services chassis sandwich model
  – Addition of multiple virtual contexts to the transparent services modules
  – Addition of VRF routing instances within the sub-aggregation VDC
  – Service module contexts and VRFs are linked together by VLANs to form logical traffic paths
  – Example Web/App server farm and Database server cluster homed to separate VRFs to direct traffic through the services




# Service Pattern Active-Active: Client-to-Server



## Agenda

- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services

- Storage Virtualization
- Data Center Interconnect
- Conclusion

Chesapeake NETCRAFTSMEN 67 Copyright 2010

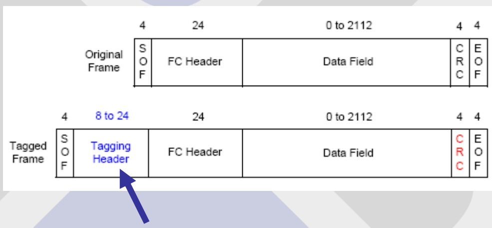
## Storage Virtualization: Terminology?

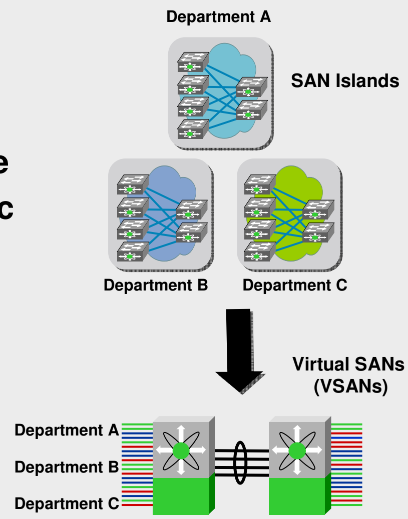
- Storage virtualization englobes various concepts
- Definitions may vary based on your interlocutor
  - –For some, storage virtualization starts at virtual volumes
  - –For others, it starts with Virtual SANs
- Example: unified I/O
  - –Storage virtualization, network virtualization, both?
- First things first: the basics
  - –VSANs, FlexAttach, NPIV, NPV, Unified I/O, Virtual Volumes

Chesapeake NETCRAFTSMEN 68 Copyright 2010

# Just Like There Are VLANs, There Are VSANs

- SAN islands
  - –Duplication of hardware resources
  - –Just-in-time provisioning
- VSAN: consolidation of SANs on one physical infrastructure
- Much like VLANs, VSAN traffic carries a tag

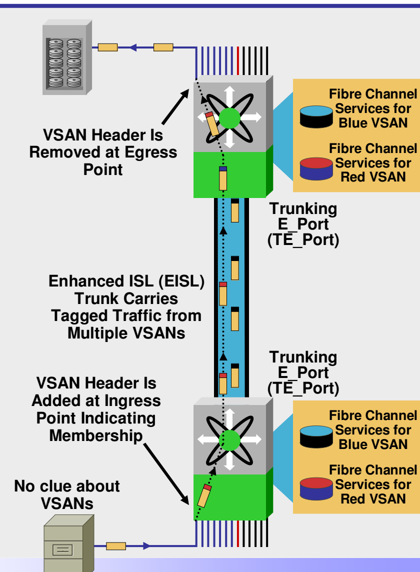





# VSAN Tagging Two Primary Functions

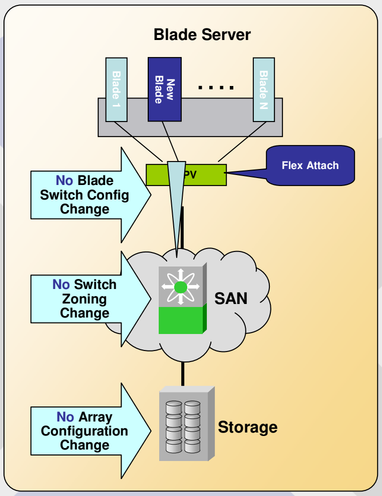
- Hardware-based isolation of tagged traffic belonging to different VSANs
  - –No special drivers or configuration required for end nodes (hosts, disks, etc.)
  - –Traffic tagged at Fx_Port ingress and carried across EISL (enhanced ISL) links between switches
- Create independent instance of Fibre Channel services for each newly created VSAN—services include:
  - –Zone server, name server, management server, principle switch election, etc.
  - –Each service runs independently and is managed/configured independently





## WWN Virtualization: FlexAttach

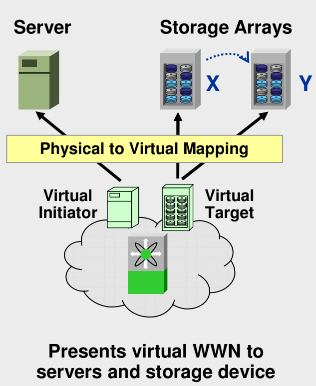



- HBAs have World-Wide-Names
  - –They're burnt-in like MAC addresses
- FlexAttach assigns a WWN to a port
  - –Each F-Port is assigned a virtual WWN
  - –Burnt-in WWN is NAT'd to virtual WWN
- Benefits
  - –Same WWN on a given port
  - –Control over WWN assignment
  - –Replacing failed HBA or host simple



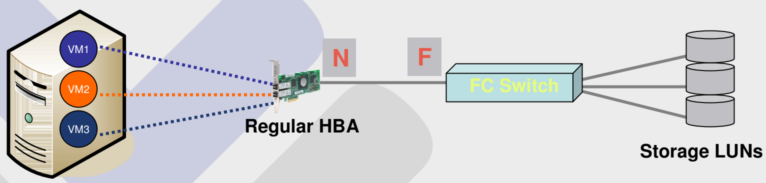
## SAN Device Virtualization

- Allows provisioning with virtualized servers and storage devices
- Significantly reduces time to replace HBAs and Storage devices
  - –No reconfiguration of zoning, VSANs, etc. required on MDS
  - –No need to reconfigure storage array LUN masking after replacing HBAs
  - –Eliminates re-building driver files on AIX and HP-UX after replacing storage



## VM-Unaware Storage

- Traditional scenario: 3 VMs on ESX, one physical HBA

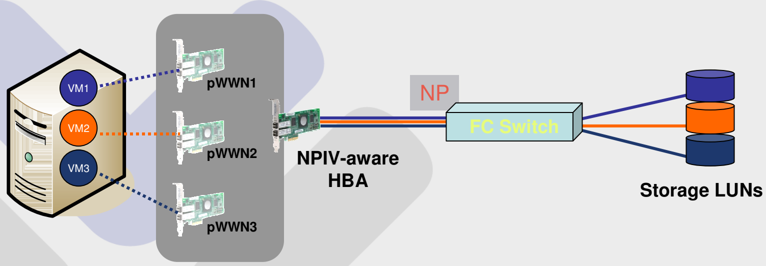



- VMs don't have WWNs. Only physical HBA does.
  - –No VM-awareness inside SAN fabric
  - –No VM-based LUN masking for instance



## VM-Aware Storage: NPIV

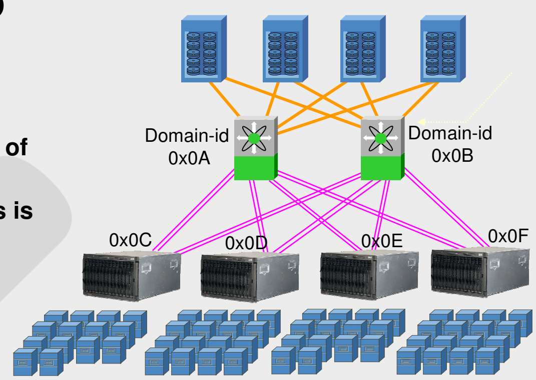
- NPIV stands for N_Port ID Virtualizer




- Now each VM has its own port WWN
- Fabric sees those WWNs
  - –VM-aware zones or LUN masking

# Domain ID Explosion

- Blade servers: domain ID explosion!
  - –Each FC switch inside blade servers use single domain ID
  - –Theoretical maximum number of Domain IDs is 239 per VSAN
  - –Supported number of domains is quite smaller:
    - EMC: 40 domains
    - Cisco Tested: 75
    - HP: 40 domains
- Manageability
  - –Lots of switches to manage
  - –Possible domain-ID overlap
  - –Possible FSPF reconfiguration



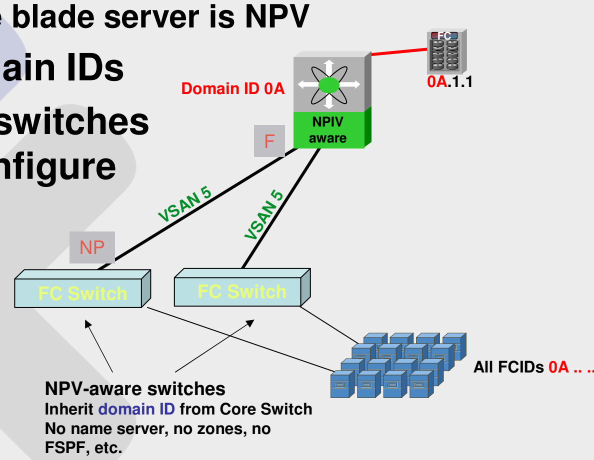



# Solution: N-Port Virtualizer (NPV)

- What is NPV?
  - –NPV enables the switch to act as a proxy for connected hosts
  - –Switch in NPV mode is no longer a switch
  - –NPV switch does not use a Domain ID
  - – Inherits Domain ID from upstream fabric switch
    - No longer limited to Domain ID boundaries
- Manageability
  - –Far less switches to manage – NPV very much plug and play
  - –NPV-enabled switch is now managed like a NPIV enabled host
  - –Eliminates the need for server administrators to manage the SAN

# N-Port Virtualization (NPV): An Overview

- NPV topology
  - –Switch inside blade server is NPV
- Reduces domain IDs
- Blade server switches simpler to configure



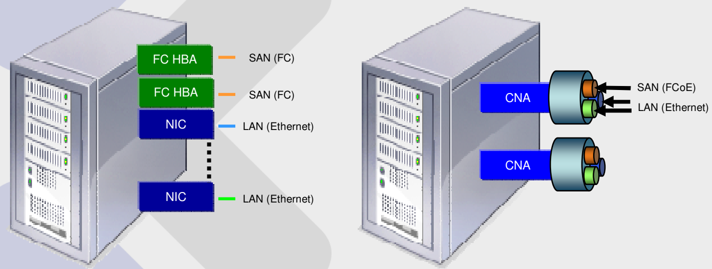



# Differences Between NPIV and NPV

- NPIV (N-Port ID Virtualization)
  - –Functionality geared towards server's host bus adapters (HBA)
  - –NPIV provides a means to assign multiple Server Logins to a single physical interface
  - –The use of different virtual pWWN allows access control (zoning) and port security to be implemented at the application level
  - –Usage applies to applications such as VMWare, MS Virtual Server and Linux Xen
- NPV (N-Port Virtualizer)
  - –Functionality geared towards MDS fabric switches (MDS 9124, MDS 9134, Nexus 5000 and blade switches)
  - –NPV provides the FC switch's connections (uplink) to act as server connections – instead of acting like a standard ISL
  - –Utilizes NPIV type functionality to allow multiple server logins from other switch ports to use NP-port uplink

## Unified I/O?

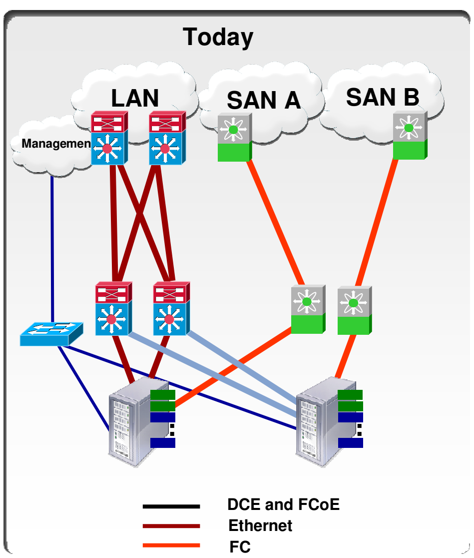
- Consolidation of FC and Ethernet traffic on same infrastructure
- New protocols (FCoE, ‘Data Center Ethernet’) for guaranteed QoS levels per traffic class







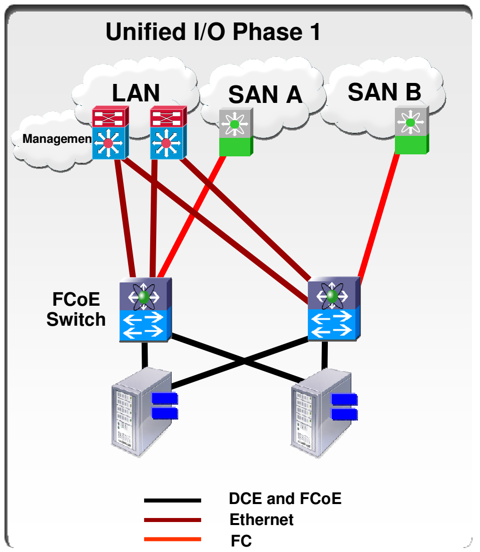
## Un-Unified I/O Today




Today

- Parallel LAN/SAN Infrastructure
- Inefficient use of Network Infrastructure
- 5+ connections per server – higher adapter and cabling costs
  - Adds downstream port costs; cap-ex and op-ex
  - Each connection adds additional points of failure in the fabric
- Longer lead time for server provisioning
- Multiple fault domains – complex diagnostics
- Management complexity – firmware, driver-patching, versioning

# Unified I/O Today

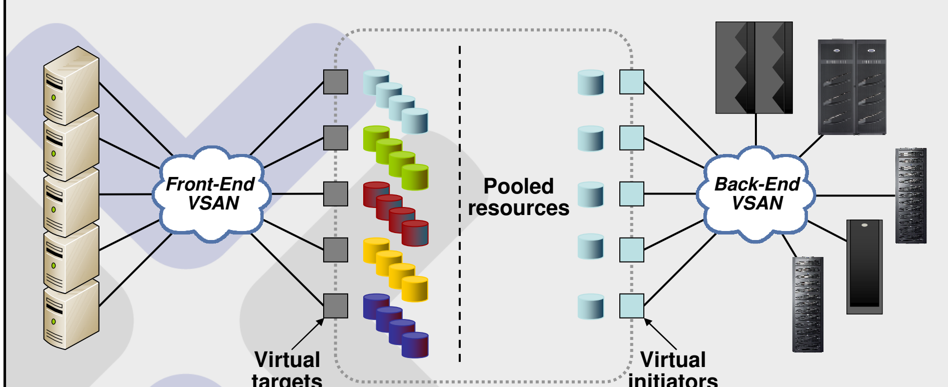



Unified I/O Phase 1

- Reduction of server adapters
- Simplification of access layer and cabling
- Gateway free implementation—fits in installed base of existing LAN and SAN
- L2 Multipathing Access—Distribution
- Lower TCO
- Fewer Cables
- Investment Protection (LANs and SANs)
- Consistent Operational Model

# Storage Virtualization Logical Topology

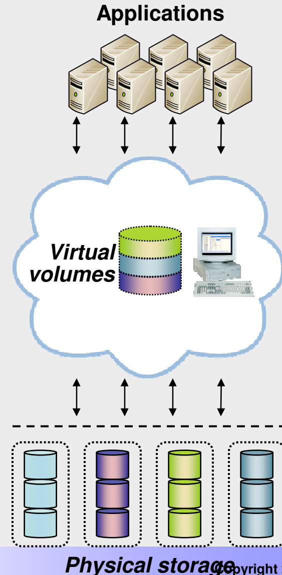


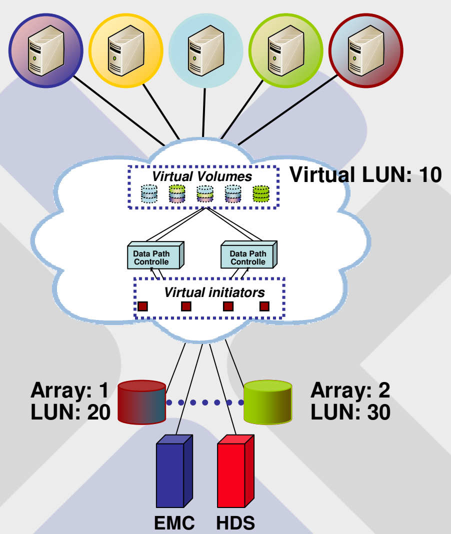
# Network-Based Volume Management

- Simplify volume presentation and management
  - –Create, delete, change storage volumes
  - –Provides front-end LUN Masking and mapping of storage volume to hosts
- Centralize management and control
  - –Single Invista console to manage virtual volumes, clones, and mobility jobs
- Reduce management complexity of a heterogeneous storage
  - –Single management interface to allocate and reallocate storage resources





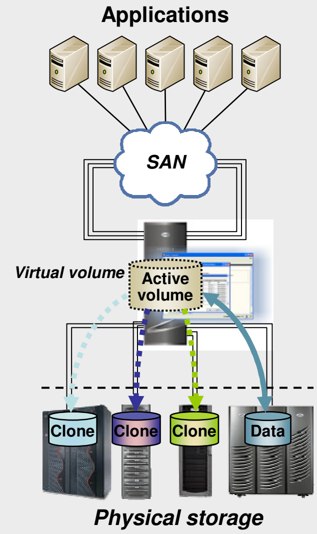

# Dynamic Volume Mobility Explained




- Virtualization
  - –Hosts see Storage Virtualization as an array
  - –Presents virtual volumes to hosts
  - –Maps virtual volumes to physical volumes
- To move a volume:
  - –Select source and target volumes
  - –Network synchronizes the volumes, then changes the virtual-physical mapping
  - –No I/O disruption to host

## Heterogeneous Point-in-Time Copies

- Create point-in-time copies
  - –Source and clone can be on different, heterogeneous storage arrays
- Enable replication across heterogeneous storage
  - –Leverage existing storage investments
  - –Reduce replication storage capacity and management costs
- Maximize replication benefits to support service levels
  - –Backup and recovery
  - –Testing, development, and training
  - –Parallel processing, reporting, and queries



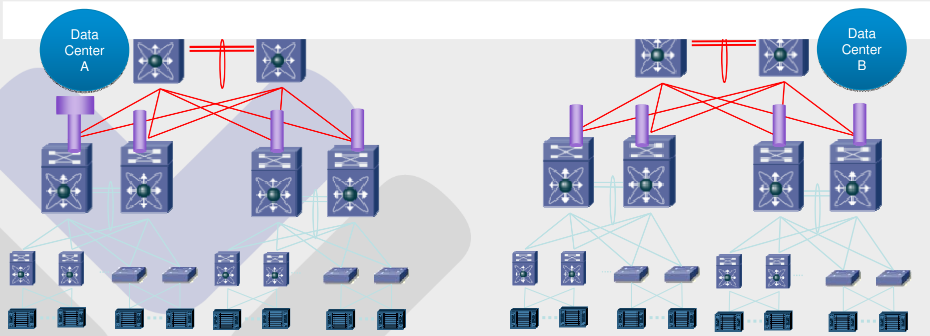

Chesapeake NETCRAFTSMEN 85 Copyright 2010

## Agenda

- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion

Chesapeake NETCRAFTSMEN 86 Copyright 2010

# Problem Statement – LAN extensions

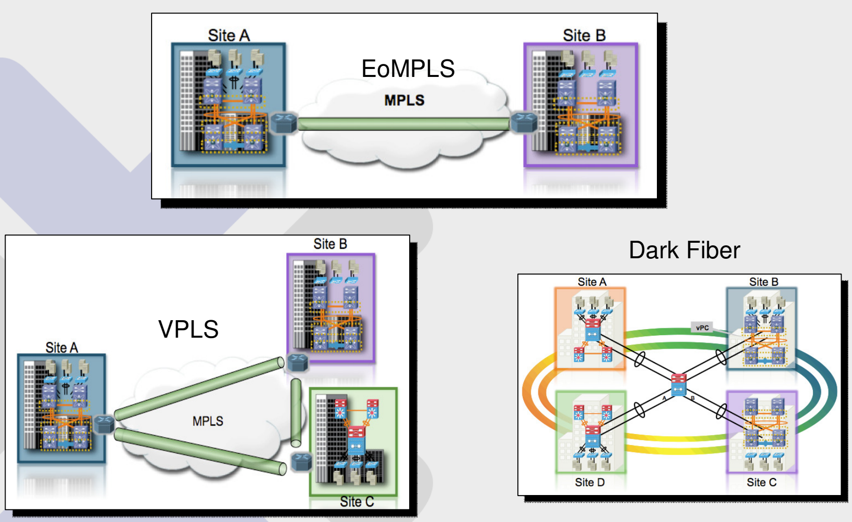



- Certain Applications require L2 connectivity among peers
  - Clusters (Veritas, MSFT)
  - vMotion
  - Home-brewed apps
- Within and between Data Centers

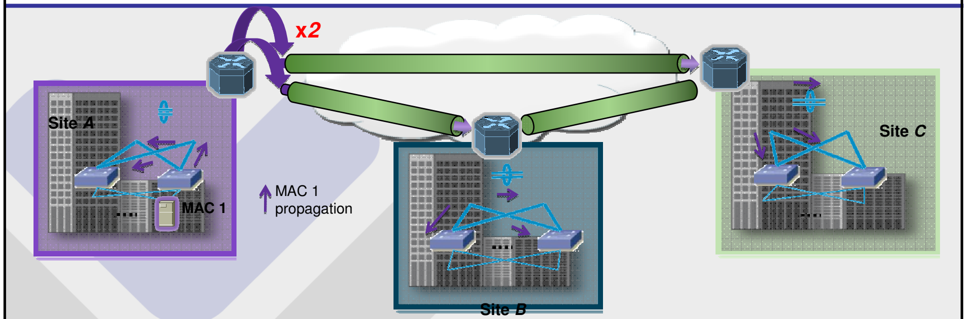
- Uses:
  - Server migrations
  - Disaster recovery and resiliency
  - High rate encryption may require an L2 transport between sites
  - Distributed Active-Active DCs



# Traditional Layer 2 Data Center Interconnect



# Traditional Layer 2 DCI: *Data Plane MAC Learning*




- Layer 2 VPN technologies use a Data Plane driven learning mechanism.
- This mechanism is the same as the one used by classical Ethernet bridges.
- When the frame is received with an unknown source MAC address, the source MAC address is programmed in the bridge table.

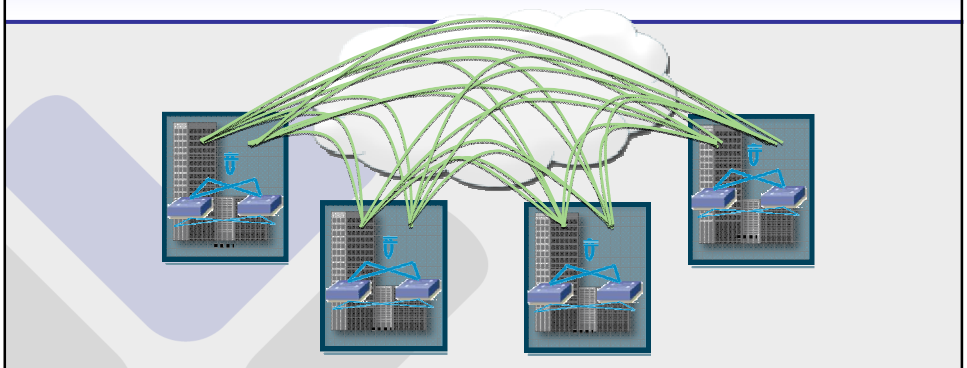
- When a bridge receives a frame and its destination MAC is not in the MAC table, the frame is flooded on the bridge domain.
- This is referred to as ***unknown unicast flooding***.
- As the flood travels throughout the entire bridge domain, it triggers learning of its source MAC address over multiple hops.

This flooding behavior causes failures to propagate to every site in the L2-VPN

89

Copyright 2010

# Traditional Layer 2 DCI: *Circuit Switching*



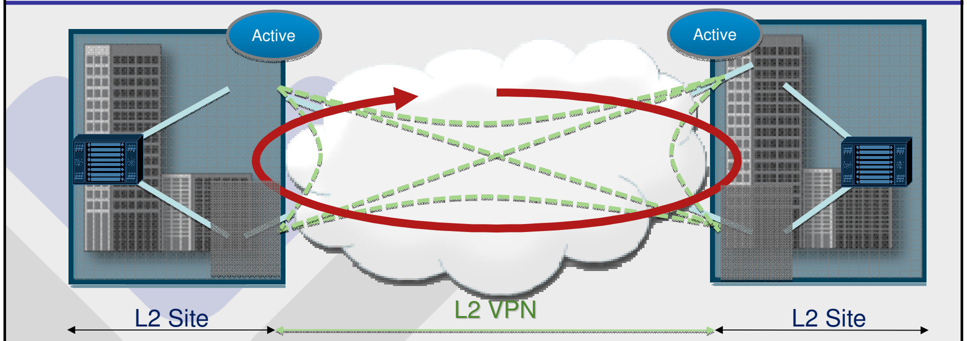
- Before any learning can happen a ***full mesh*** of circuits must be available.
- Circuits are usually statically predefined.
- For *N* sites, there will be ***N*(N-1)/2*** circuits. Operational challenge!

- Scalability is impacted as the number of sites increases.
- Head-end replication for multicast and broadcast
- Complex addition and removal of sites.

90

Copyright 2010

## Traditional Layer 2 DCI: *Loop Prevention*



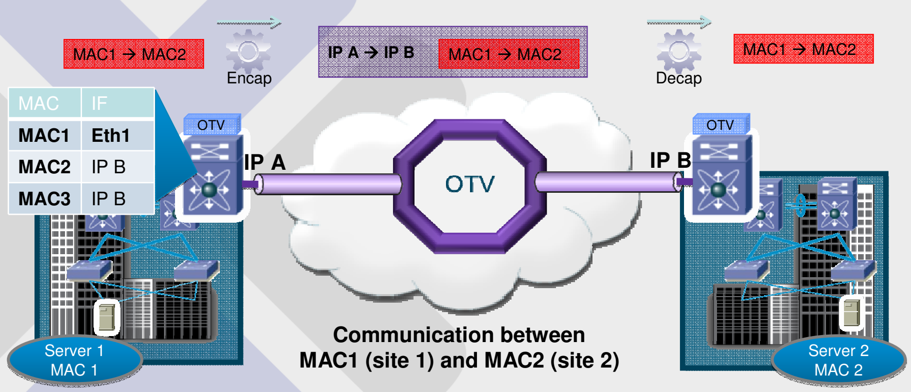

- Coordination between edge devices on the same site is needed.
- One of the edge devices becomes the designated active device.
- The designed active device can be at the device level or per VLAN.

- STP is often extended across the sites of the Layer 2 VPN.
- Very difficult to manage as the number of sites grows.
- Malfunctions on one site will likely impact all sites on the VPN.



## Overlay Transport Virtualization at a Glance

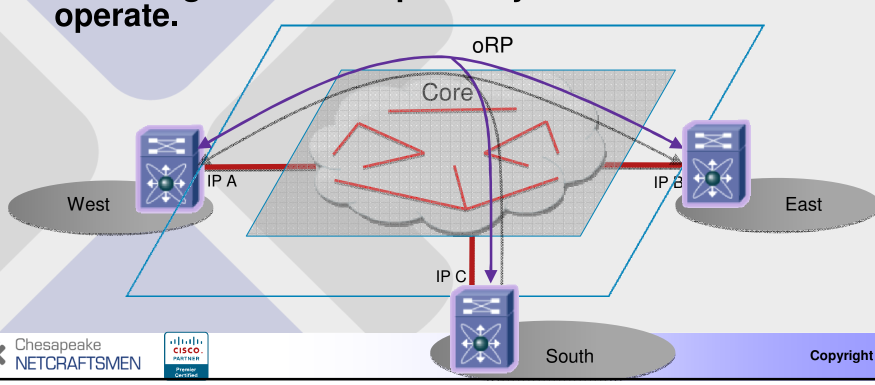
- Ethernet traffic between sites is encapsulated in IP: "MAC in IP"
- Dynamic encapsulation based on MAC routing table
- No Pseudo-Wire or Tunnel state maintained




Communication between MAC1 (site 1) and MAC2 (site 2)

# OTV: MAC Tables

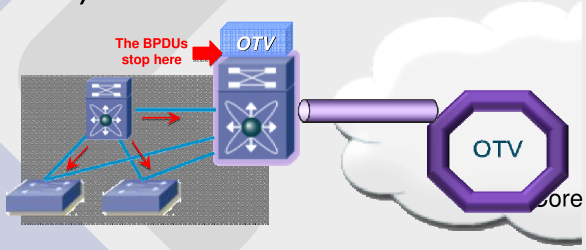
- OTV uses a protocol to proactively advertise MAC reachability (control-plane learning). We will refer to this protocol as the "overlay Routing Protocol" (oRP).
- oRP runs in the background once OTV has been configured.
- No configuration is required by the user for oRP to operate.






# Overlay Transport Virtualization Benefits

- STP BPDU's – not forwarded on overlay network, OTV device participates in STP on campus side
- Unknown unicasts not forwarded on overlay – assumption that no hosts are silent or uni-directional (workarounds if not)
- Proxy ARP keeps ARP traffic local, reduces overlay broadcast traffic
- OTV prevents loops from forming via control of device forwarding for a site (VLAN for site OTV edge devices to communicate on)



# Improving Traditional Layer 2 VPNs

- Data Plane Learning → Control Plane Learning
  - Moving to a Control Plane protocol that proactively advertises MAC addresses and their reachability instead of the current flooding mechanism.
- Circuit Switching → Packet Switching
  - No static tunnel or pseudo-wire configuration required
  - A Packet Switched approach would allow for the replication of traffic closer to the destination, which translates into much more efficient bandwidth utilization in the core.
- Loop Prevention → Automatic Multi-homing
  - Ideally a multi-homed solution should allow load balancing of flows within a single VLAN across the active devices in the same site, while preserving the independence of the sites.

    STP confined within the site (each site with its own STP Root bridge)

Chesapeake NETCRAFTSMEN 95 

# Overlay Transport Virtualization: Tech Pillars

OTV is a "MAC in IP" technique for supporting Layer 2 VPNs over any transport.

**Packet Switching**
- No Pseudo-Wire State Maintenance
- Optimal Multicast Replication
- Multi-point Connectivity
- Point-to-Cloud Model

OTV

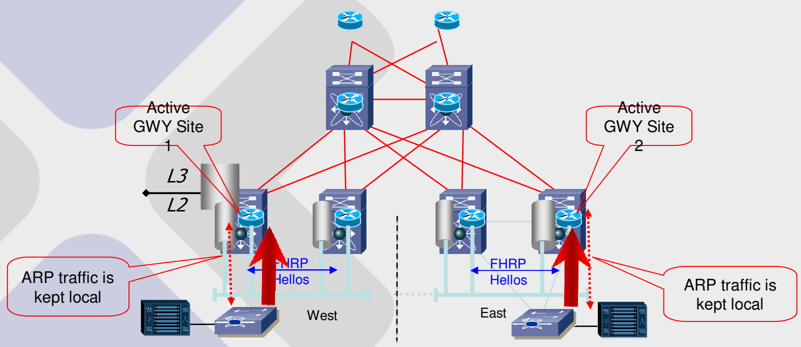
**Protocol Learning**
- Built-in Loop Prevention
- Preserve Failure Boundary
- Seamless Site Addition/Removal
- Automated Multi-homing

Chesapeake NETCRAFTSMEN 96 

# OTV: Egress Routing Localization

- HSRP hellos can be filtered at the OTV site edge.
- A single FHRP group will now have an active GWY on each site.
- No special configuration of the FHRP is required.
- ARP requests for A are intercepted at the OTV edge to ensure the replies are from the local active GWY.
- Optimal Egress Router choice.






# OTV Configuration

The CLI is subject to change prior to FCS.

```
interface Overlay0
description otv-demo
otvexternal-interface Ethernet1/1
otvgroup-address 239.1.1.1 data-group-range 232.192.1.2/32
otvadvertise-vlan 100-150
otvsite-vlan100
```

Connect to the core. Used to join the core mcast groups. Their IP addresses are used as source IP for the OTV encap

ASM/Bidir group in the core used for oRP.

SSM group range used to carry the site mcast traffic data.

Site VLANs being extended by OTV

VLAN used **within** the Site for communication between the site's Edge Devices

## Agenda

- Virtualization: Getting Motivated!
- Compute Resource Virtualization
- Network Virtualization
- Virtualization with VSS
- Virtualization with Nexus
- Adding Services
- Storage Virtualization
- Data Center Interconnect
- Conclusion



## The Future?

- Bigger faster ESXi servers
  - IBM has announced Power7 chips (8-way cores) and servers that are claimed to support up to 640 VM's (32 processors x 20 VM's each)
  - Intel Nehalem-EX may be roughly comparable
  - Intel has put out experimental 48-way cores (lower clock rate)?
  - Faster CPU + more cores reduces CPU limitation on # VM's
  - Cisco (and now others) technology reduces memory issues capping # of VM's
  - Do the math: 128 processors x perhaps 30 VM's each? (3840 VM's in a rack?). 50 VM's per? (6400 VM's in a rack??)
- Fewer, faster network connections
  - Do you use N x 10 G or 40 G or 100 G to such a box?
  - Especially with FCoE thrown in
  - Greatly reduce cable tangle to 6 – 10 NIC + HBA adapters
  - Further shrink size of chassis

## The Future – 2

- More SAN
  - VMotion and other desirable techniques require SAN
  - Your business depends on it – speed and reliability are key
  - Consistent SAN management practices and SAN virtualization enhance flexibility and reliability
  - SAN de-deplication, SAN-based backup, etc. are the icing on the cake
- Cloud computing
  - Some mix, low risk servers may well end up in cloud
  - Crucial servers, big DB's, high risk servers remain internal?



## Virtualization – What is in for me?

- Virtualization is an overloaded term
- A collection of technologies to allow a more flexible usage of hardware resources
- Assembled in an end to end architecture these technologies provide the agility to respond to business requirements

## Summary

- Virtualization of the network infrastructure improves utilization and offers new deployment
models in the data center
- Flexible service models readily account for application requirements
- Security is a process not a product; virtualization allows for efficient application of security policies
- The application is the beneficiary of all these developments



## Any Questions?

- For a copy of the presentation, email me at pjw@netcraftsmen.net
  - I'll post a link in a blog article
- About Chesapeake Netcraftsmen:
  - Cisco Premier Partner
  - Cisco Customer Satisfaction Excellence rating ☆
  - We wrote the original version of the Express Foundations courses required for VAR Premier Partner status (and took and passed the tests), and the recent major CCDA/CCDP refresh
  - Cisco Advanced Specializations:
    - Advanced Unified Communications (and IP Telephony)
    - Advanced Wireless
    - Advanced Security
    - Advanced Routing & Switching
    - Advanced Data Center Networking Infrastructure
  - Deep expertise in Routing and Switching (several R&S and four double CCIE's)
  - We do network / security / net mgmt / unified communications / data center Design and Assessment

CISCO PARTNER Premier Certified

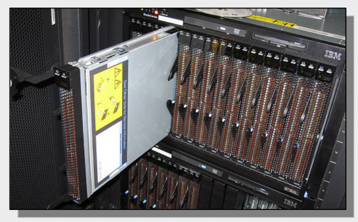


Extra Slides:
Virtualization

## Servers: 2000+

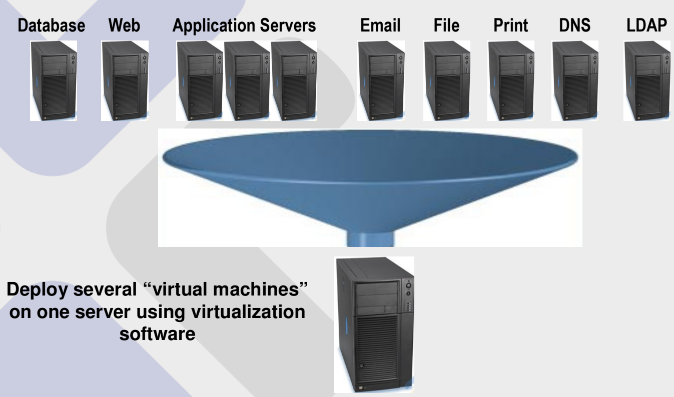
- One app, one server
- Focus on reducing footprint
  - "Rack" form factors (6-20 servers per cabinet)
  - "Blade" form factors (30-60 servers per cabinet)
  - Helped alleviate some of the footprint issues
  - Power and heat still a problem
- The more powerful the CPU
  - The lower server utilization!
  - Average server utilization ranges between 4–10%
  - Still one application per server





## Servers: Virtualization Is the Key

- Apply Mainframe Virtualization Concepts to x86 Servers:
- Use virtualization software to partition an Intel / AMD server to work with several operating system and application "instances"



# Virtualization Landscape

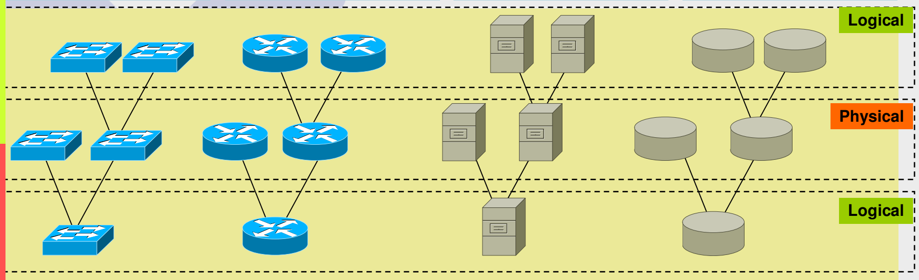
Consolidation

**Services**
- Unified Communication
- Content Network
- WAN Acceleration

**Network**
- VLAN - MPLS
- VPN
- Virtual Switch - VRF

**Compute**
- Server Virtualization
- Virtual Appliances
- Virtual Context

**Storage**
- VSAN
- vHBA
- NPIV




Scaling

**Services**
- Call control
- Web Server
- Video Server
- File Server

**Network**
- HSRP/GLBP/VRRP
- VSS
- WCCP

**Compute**
- SLB VIP
- Unified Computing
- Cloud Computing

**Storage**
- Service Profiles
- Logical Volumes



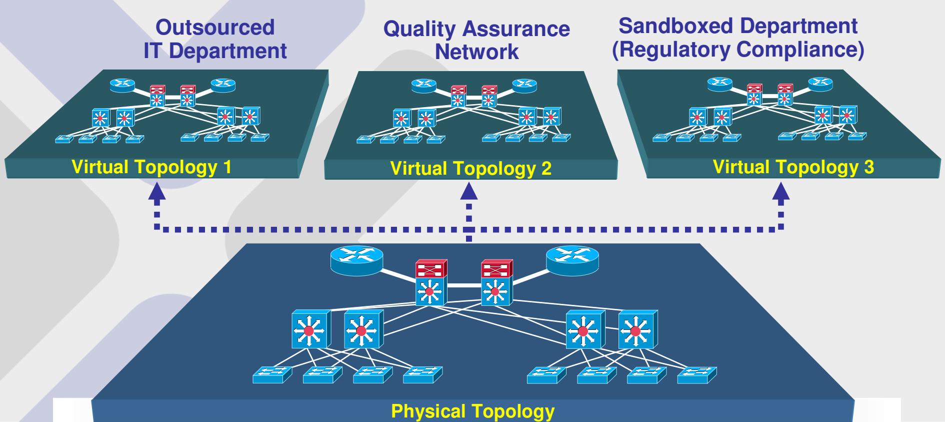
# Several Ways to Virtualize

- Container-based
  - Linux VServer
- Paravirtualization
  - Xen, VMware ESX (device drivers), Microsoft Hyper-V
- Host-based
  - Microsoft Virtual Server, VMware Server and Workstation
- Native ('Full') virtualization
  - VMware ESX, Linux KVM, Microsoft Hyper-V, Xen

Extra Slides:
Network Virtualization

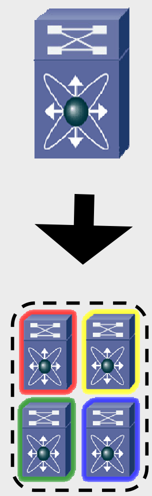
# What Is Network Virtualization?

- Overlay of logical topologies (1:N)
- One physical network supports N virtual networks



# Nexus 7000 Series Virtual Device Contexts (VDCs)

- Virtualization of the Nexus 7000 Series Chassis
  - Up to 4 separate virtual switches from a single physical chassis with common supervisor module(s)
  - Separate control plane instances and management/CLI for each virtual switch
  - Interfaces only belong to one of the active VDCs in the chassis, external connectivity required to pass traffic between VDCs of the same switch
- Designing with VDCs
  - VDCs serve a "role" in the topology similar to a physical switch; core, aggregation, or access
  - Multiple VDC example topologies have been validated within Cisco by ESE and other teams
  - Two VDCs from the same physical switch should not be used to build a redundant network layer – physical redundancy is more robust



# Virtualization Inside a VDC

**Nexus**

**VDC**

| | |
|---|---|
| VLAN VLAN VLAN | VRF VRF VRF |
| VLAN VLAN VLAN | VRF VRF VRF |
| VLAN VLAN VLAN | VRF VRF VRF |

**VDC**

| | |
|---|---|
| VLAN VLAN VLAN | VRF VRF VRF |
| VLAN VLAN VLAN | VRF VRF VRF |
| VLAN VLAN VLAN | VRF VRF VRF |

**VDC**

| | |
|---|---|
| VLAN VLAN VLAN | VRF VRF VRF |
| VLAN VLAN VLAN | VRF VRF VRF |
| VLAN VLAN VLAN | VRF VRF VRF |

**Scalability:**

- 4K VLANs/VDC
- 256 VRFs/VDC
- 4 VDCs

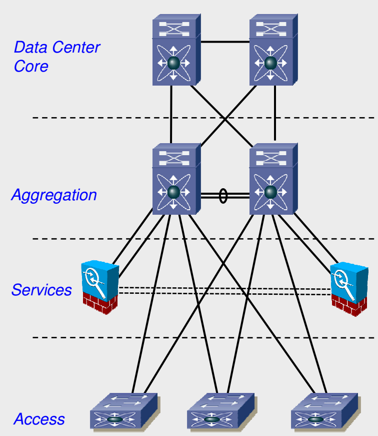
 

# Extra Slides:
# Data Center Service Insertion

## Data Center Service Insertion:

Direct Services Appliances

- Appliances directly connected to the aggregation switches
  - Service device type and Routed or Transparent mode can affect physical cabling and traffic flows.
- Transparent mode ASA example:
  - Each ASA dependant on one aggregation switch
  - Separate links for fault tolerance and state traffic either run through aggregation or directly
  - Dual-homed with interface redundancy feature is an option
  - Currently no EtherChannel supported on ASA



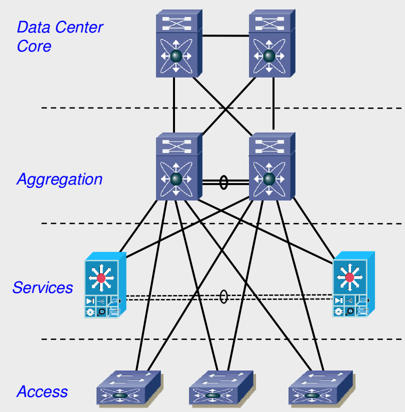

 

# Data Center Service Insertion:

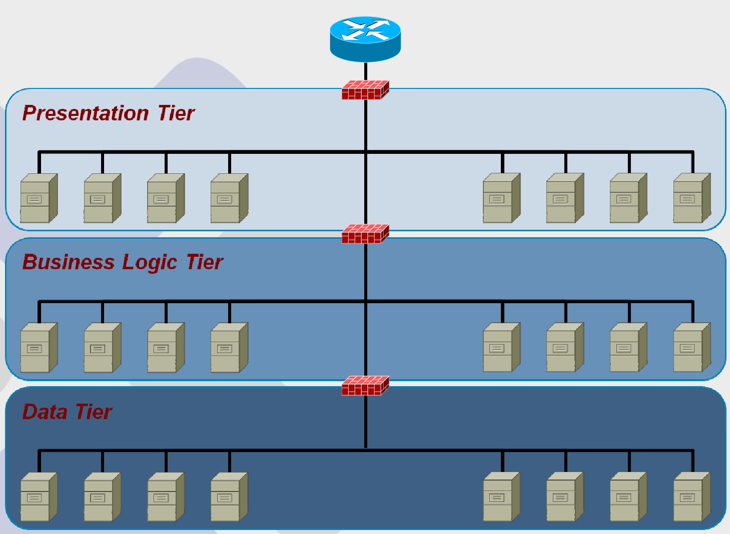
External Services Chassis

- **Dual-homed Catalyst 6500**
  - **Services do not depend on a single aggregation switch**
  - **Direct link between chassis for fault-tolerance traffic, may alternatively trunk these VLANs through Aggregation**
- **Dedicated integration point for multiple data center service devices**
  - **Provides slot real estate for 6500 services modules**
  - **Firewall Services Module (FWSM)**
  - **Application Control Engine (ACE) Module**
  - **Other services modules, also beneficial for appliances**

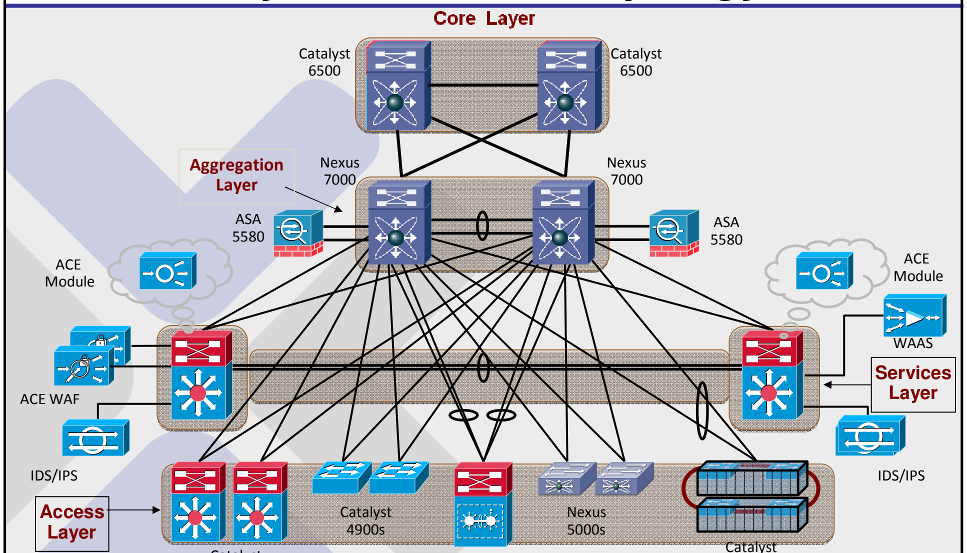





# ISV Network View



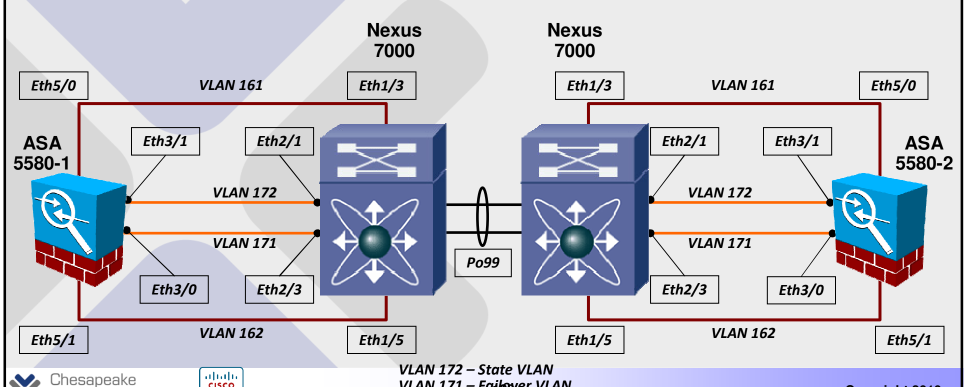
# Physical Solution Topology





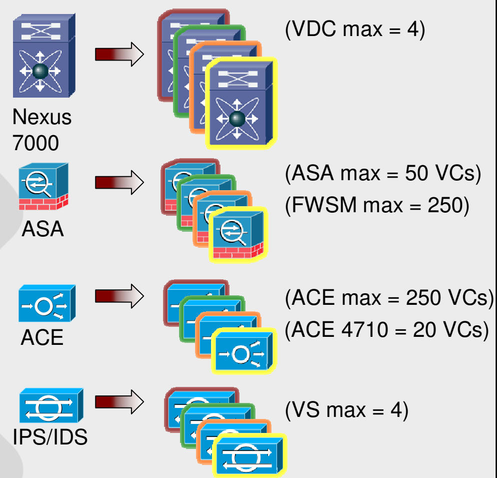

# Data Center Design Aggregation Layer DMZ

- Redundant physical chassis provide virtual platform
- Physical interfaces allocated to independent VDCs and ASA virtual contexts
- Fault tolerance and state VLANs leverage VDC2 Po99



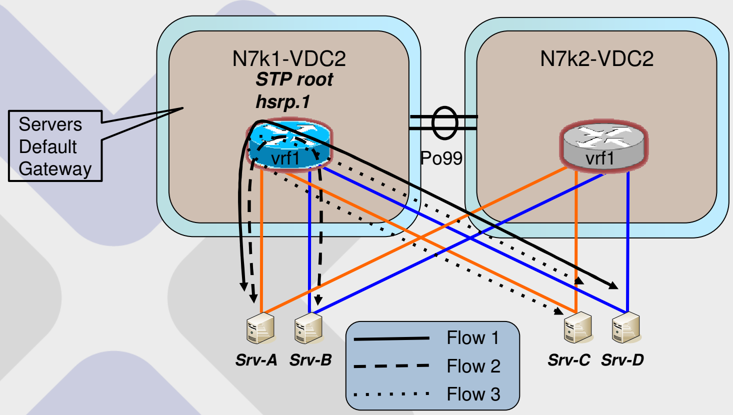
## Active-Active Solution Virtual Components

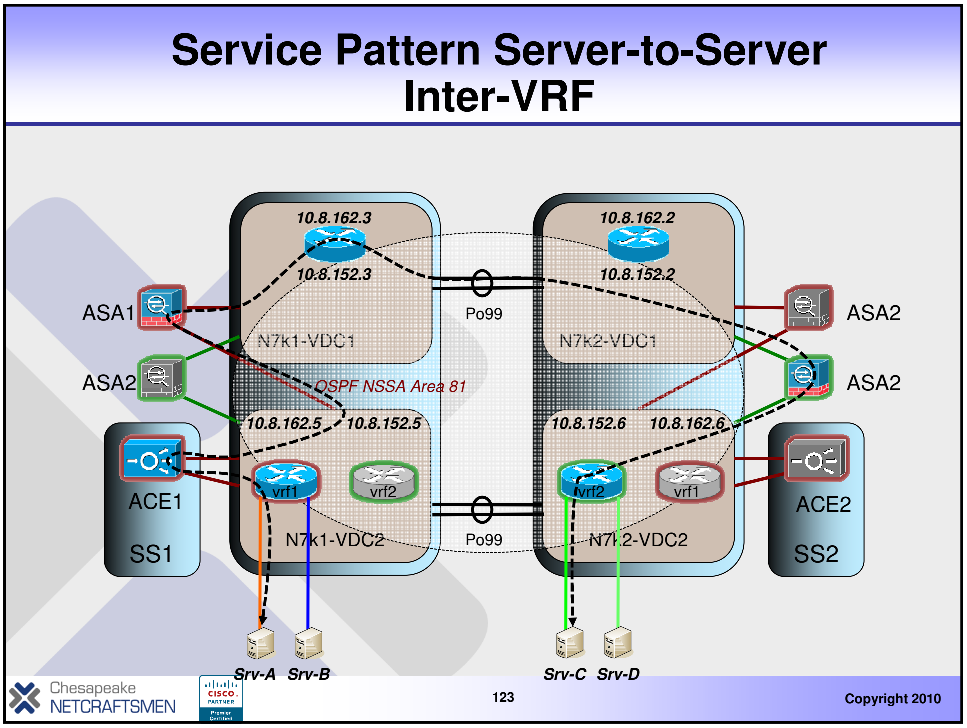
- **Nexus 7000**
  - **VDCs, VRFs, SVIs**
- **ASA 5580**
  - **Virtual Contexts**
- **ACE Service Module**
  - **Virtual Contexts, Virtual IPs (VIPs)**
- **IPS 4270**
  - **Virtual Sensors**
- **Virtual Access Layer**
  - **Virtual Switching System**
  - **Nexus 1000v**
  - **Virtual Blade Switching**

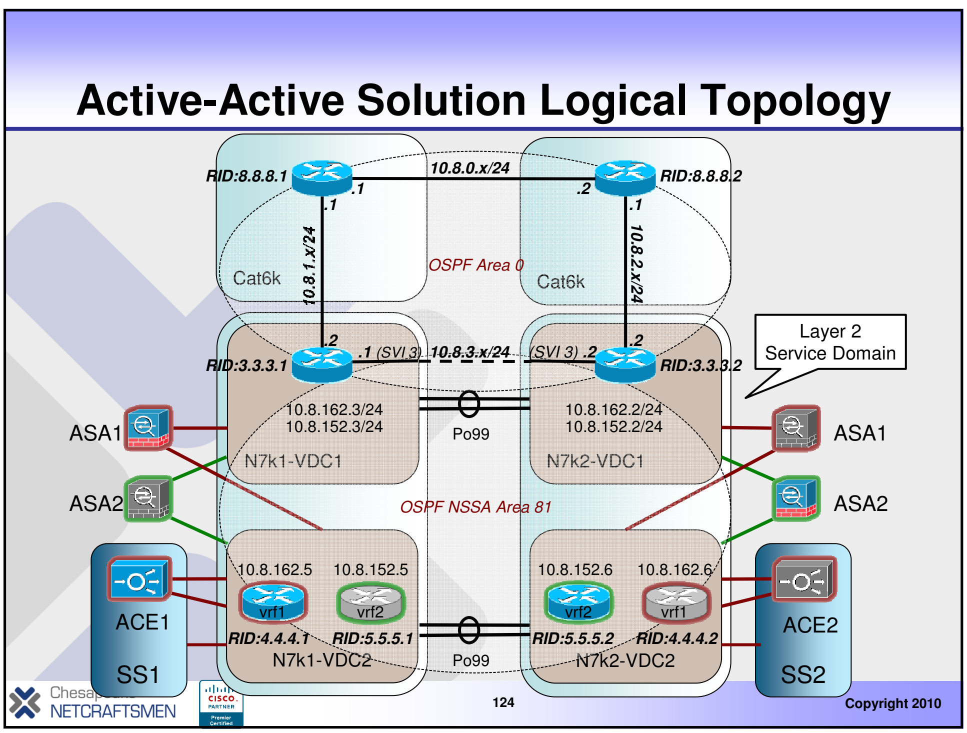


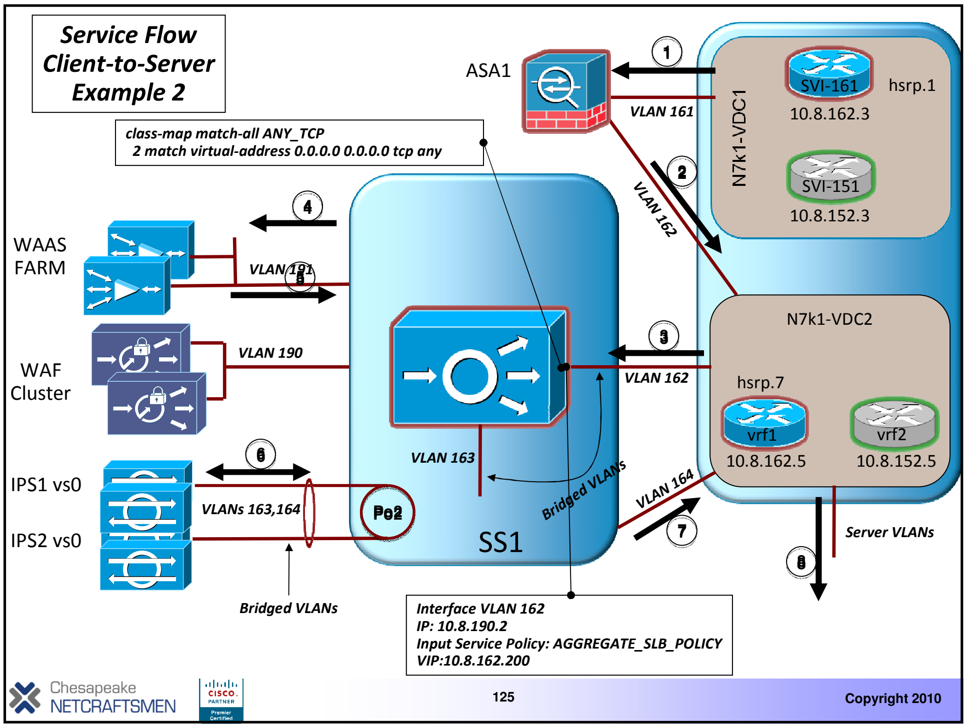



## Service Pattern Server-to-Server Intra-VRF

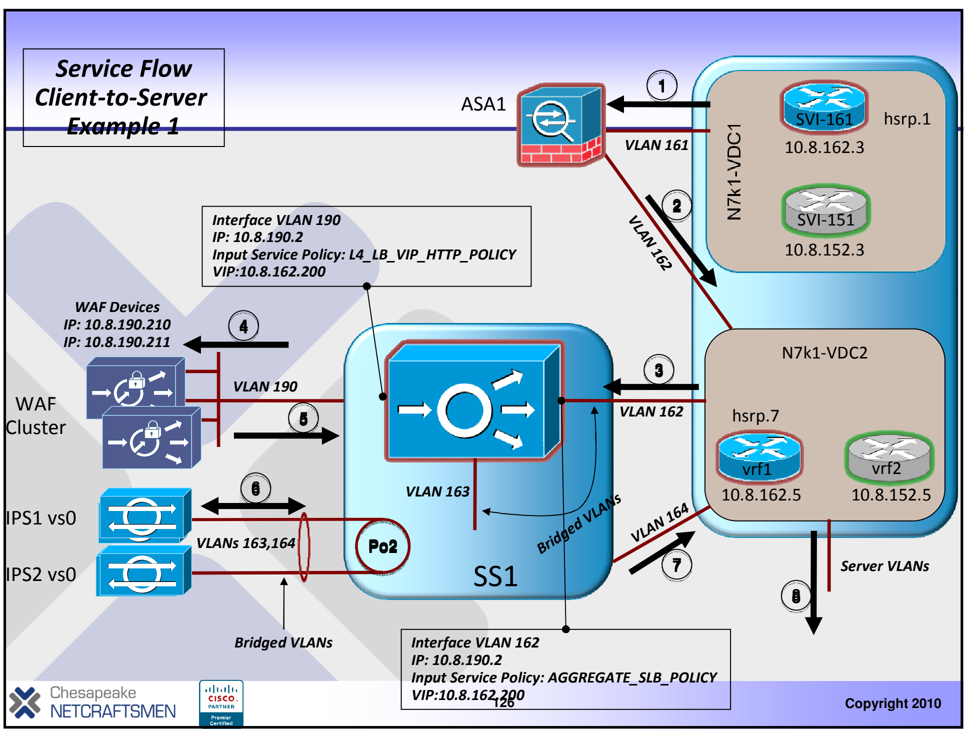



Service Pattern Server-to-Server Inter-VRF
10.8.162.3
10.8.152.3
10.8.162.2
10.8.152.2
ASA1
ASA2
ASA2
ASA2
Po99
N7k1-VDC1
N7k2-VDC1
OSPF NSSA Area 81
10.8.162.5
10.8.152.5
10.8.152.6
10.8.162.6
ACE1
SS1
vrf1
vrf2
vrf2
vrf1
ACE2
SS2
N7k1-VDC2
Po99
N7k2-VDC2
Srv-A
Srv-B
Srv-C
Srv-D
Chesapeake NETCRAFTSMEN
123
Copyright 2010


Active-Active Solution Logical Topology
RID:8.8.8.1
.1
10.8.0.x/24
.2
RID:8.8.8.2
.1
.1
10.8.1.x/24
OSPF Area 0
10.8.2.x/24
Cat6k
Cat6k
Layer 2 Service Domain
.2
.1 (SVI 3)
10.8.3.x/24
(SVI 3) .2
.2
RID:3.3.3.1
RID:3.3.3.2
10.8.162.3/24
10.8.152.3/24
10.8.162.2/24
10.8.152.2/24
ASA1
ASA1
Po99
N7k1-VDC1
N7k2-VDC1
ASA2
ASA2
OSPF NSSA Area 81
10.8.162.5
10.8.152.5
10.8.152.6
10.8.162.6
ACE1
vrf1
vrf2
vrf2
vrf1
ACE2
RID:4.4.4.1
RID:5.5.5.1
RID:5.5.5.2
RID:4.4.4.2
N7k1-VDC2
Po99
N7k2-VDC2
SS1
SS2
NETCRAFTSMEN
124
Copyright 2010

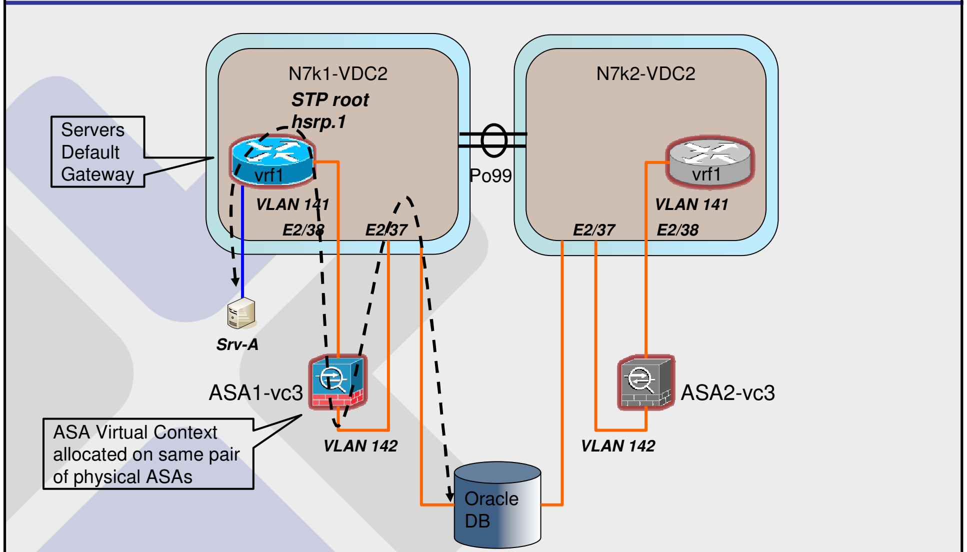


Service Flow Client-to-Server Example 2
ASA1
VLAN 161
class-map match-all ANY_TCP
2 match virtual-address 0.0.0.0 0.0.0.0 tcp any
N7k1-VDC1
SVI-161
hsrp.1
10.8.162.3
SVI-151
10.8.152.3
VLAN 162
WAAS FARM
VLAN 191
WAF Cluster
VLAN 190
N7k1-VDC2
hsrp.7
vrf1
10.8.162.5
vrf2
10.8.152.5
VLAN 163
IPS1 vs0
IPS2 vs0
VLANs 163,164
Po2
SS1
Bridged VLANs
VLAN 164
Server VLANs
Interface VLAN 162
IP: 10.8.190.2
Input Service Policy: AGGREGATE_SLB_POLICY
VIP:10.8.162.200
Chesapeake NETCRAFTSMEN
125
Copyright 2010

Service Flow Client-to-Server Example 1
ASA1
VLAN 161
N7k1-VDC1
SVI-161
hsrp.1
10.8.162.3
SVI-151
10.8.152.3
VLAN 162
Interface VLAN 190
IP: 10.8.190.2
Input Service Policy: L4_LB_VIP_HTTP_POLICY
VIP:10.8.162.200
WAF Devices
IP: 10.8.190.210
IP: 10.8.190.211
WAF Cluster
VLAN 190
N7k1-VDC2
hsrp.7
vrf1
10.8.162.5
vrf2
10.8.152.5
VLAN 163
IPS1 vs0
IPS2 vs0
VLANs 163,164
Po2
SS1
Bridged VLANs
VLAN 164
Server VLANs
Interface VLAN 162
IP: 10.8.190.2
Input Service Policy: AGGREGATE_SLB_POLICY
VIP:10.8.162.200
126
Chesapeake NETCRAFTSMEN
Copyright 2010

## Service Pattern Intra-VRF with Services




Chesapeake NETCRAFTSMEN

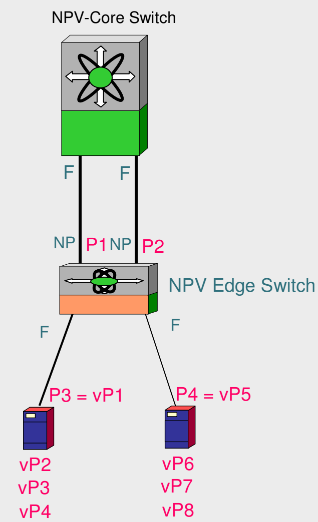
Bond142: 10.8.141.151



Extra Slides:
SAN Virtualization

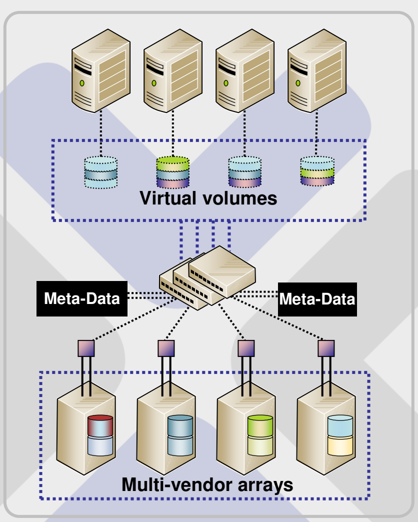
# Nested NPIV

- When NP port comes up on a NPV edge switch, it first FLOGI and PLOGI into the core to register into the FC Name Server
- End Devices connected on NPV edge switch does FLOGI but NPV switch converts FLOGI to FDISC command, creating a virtual PWWN for the end device and allowing to login using the physical NP port.
- NPIV capable devices connected on NPV switch will continue FDISC login process for all virtual PWWN which will go through same NP port as physical end device






# SAN-Based Storage Virtualization

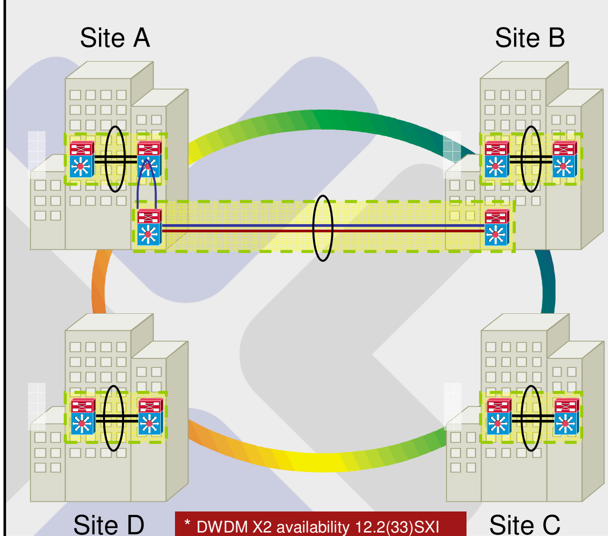



- Performance architecture
  - –Leverages next-generation “intelligent” SAN switches
- Scalable architecture
  - –Split-path architecture for high performance
  - –A “stateless” virtualization architecture does not store any information written by the application.
  - –High speed, high throughput data mapping
    - •Purpose-built ASICs (DPP) that handle and redirect I/O at line speed, with almost no additional latency
  - –Based on instructions provided by the Meta-Data Appliances
- Provides advanced functionality
- Supports heterogeneous environments

Extra Slides:
Data Center Interconnect and OTV

# DC Interconnect LAN Extension

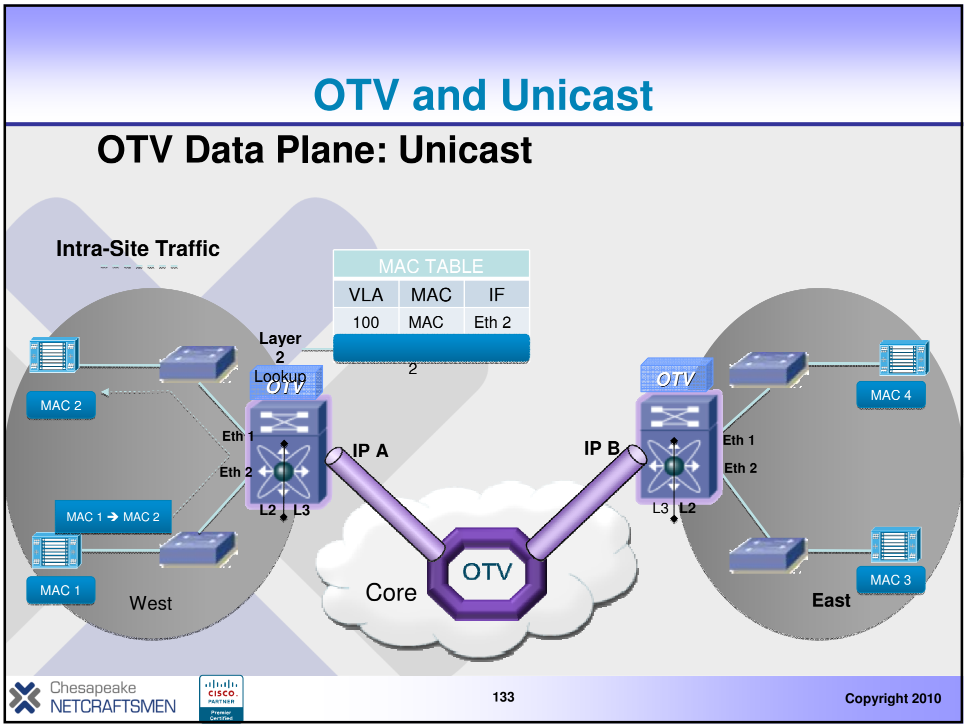
VSS Over Dark Fiber – *Multiple DC's*

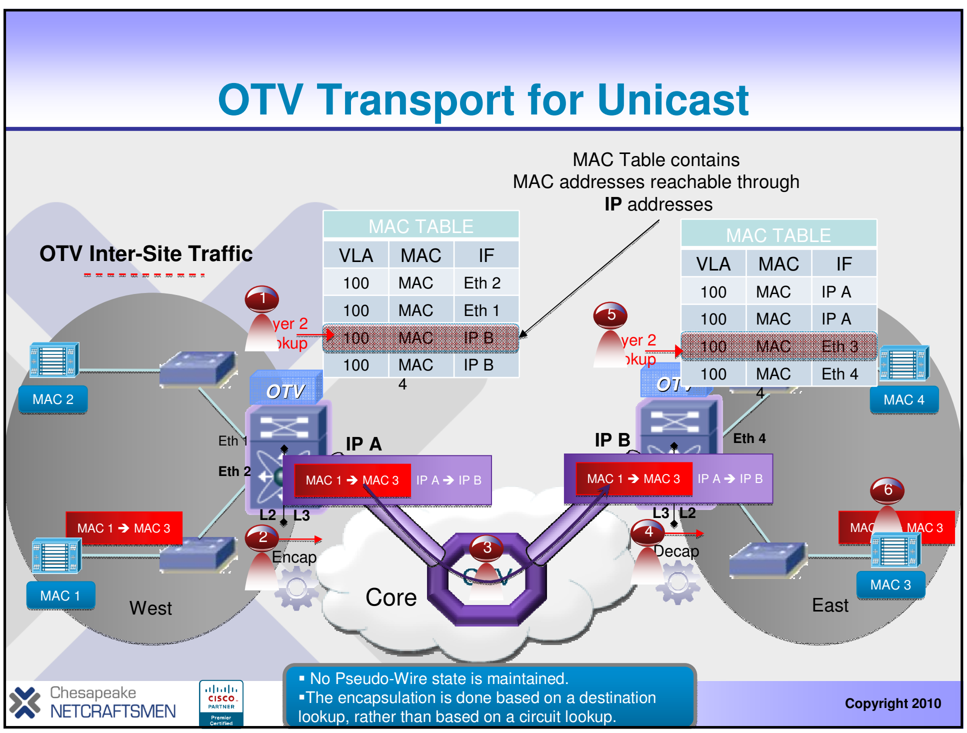



- Assumes dark fiber between sites
- Distance limitations are given by DWDM
- Number of sites can be
  2 or more
- **Add 2 switches in main data centers**
- Switches use separate lambda to interconnect
- These switches will form a VSS
- VSL is 10Gbps

* DWDM X2 availability 12.2(33)SXI

132 Copyright 2010


OTV and Unicast
OTV Data Plane: Unicast
Intra-Site Traffic
MAC TABLE
VLA MAC IF
100 MAC Eth 2
2
Layer 2 Lookup
OTV
MAC 2
MAC 1 → MAC 2
MAC 1
MAC 4
MAC 3
Eth 1
Eth 2
L2 L3
IP A
IP B
OTV
Core
West
East
Chesapeake NETCRAFTSMEN
133
Copyright 2010


OTV Transport for Unicast
MAC Table contains
MAC addresses reachable through
IP addresses
OTV Inter-Site Traffic
MAC TABLE
VLA MAC IF
100 MAC Eth 2
100 MAC Eth 1
100 MAC IP B
100 MAC IP B
4
MAC TABLE
VLA MAC IF
100 MAC IP A
100 MAC IP A
100 MAC Eth 3
100 MAC Eth 4
4
Encap
Decap
MAC 1 → MAC 3
IP A → IP B
IP A
IP B
Eth 1
Eth 2
Eth 4
L2 L3
MAC 2
MAC 1
MAC 4
MAC 3
Core
West
East
Chesapeake NETCRAFTSMEN
▪ No Pseudo-Wire state is maintained.
▪The encapsulation is done based on a destination lookup, rather than based on a circuit lookup.
Copyright 2010

## OTV Scalability Targets for the (First) Release

| | Multi-Dimensional | Uni-dimensional |
|---|---|---|
| Overlays | 3 | 64 |
| Number of sites | 3 | 10 |
| VLANs per Overlay | 128 | 128 |
| MACs across all sites | 25K | 32K |
| MACs on each site | 8K | 8K |
| Multicast Data Groups | 50 | 1000 |

DISCLAIMER
These are targets and are still subject to additional testing