

# **vSphere Networking and Converged IO with Blade Servers**

Julian Wood

UK VMware User Group – 15<sup>th</sup> November 2012

#UKVMUG

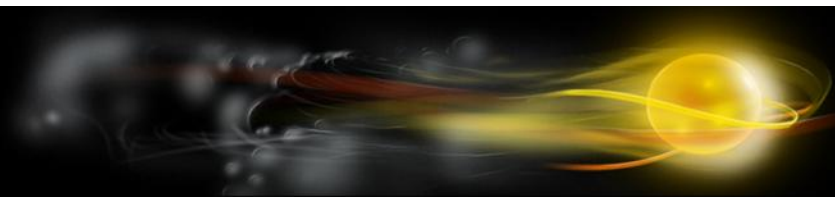


**Julian Wood**

IT Infrastructure Architect & Blogger

[www.WoodITWork.com](http://www.WoodITWork.com)

@julian\_wood



# What is converged IO?



LAN and SAN/NAS traffic over same cables and switches

10Gbps Ethernet (10GbE)

Fibre Channel-over-Ethernet (FCoE)

Lossless Ethernet

Converged Enhanced Ethernet (CEE) or Data Center Bridging (DCB)

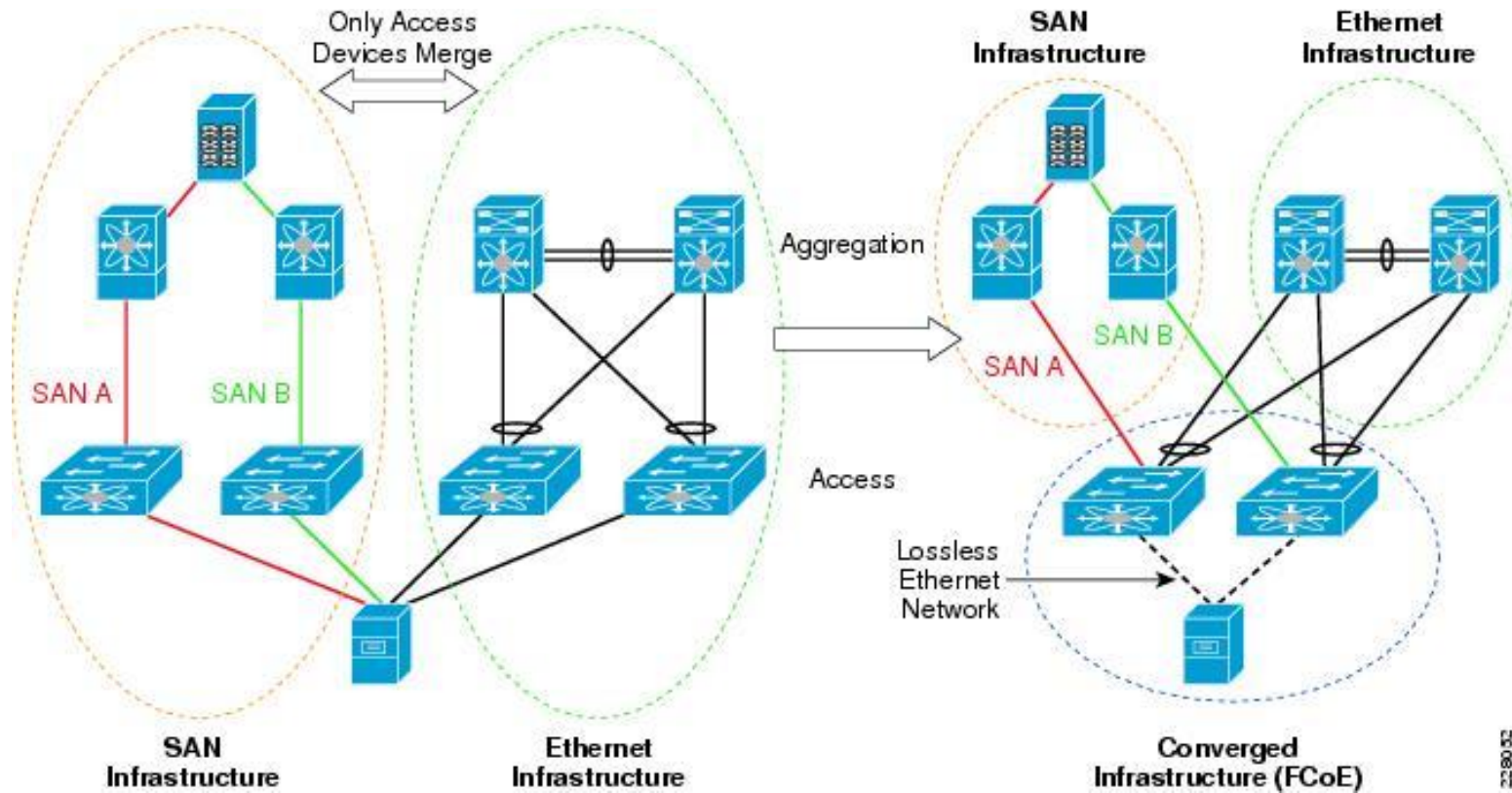
FCoE tunnels storage traffic through Ethernet & requires CEE/DCB

Infiniband RDMA ( Remote Direct Memory Access) over CEE = RoCEE

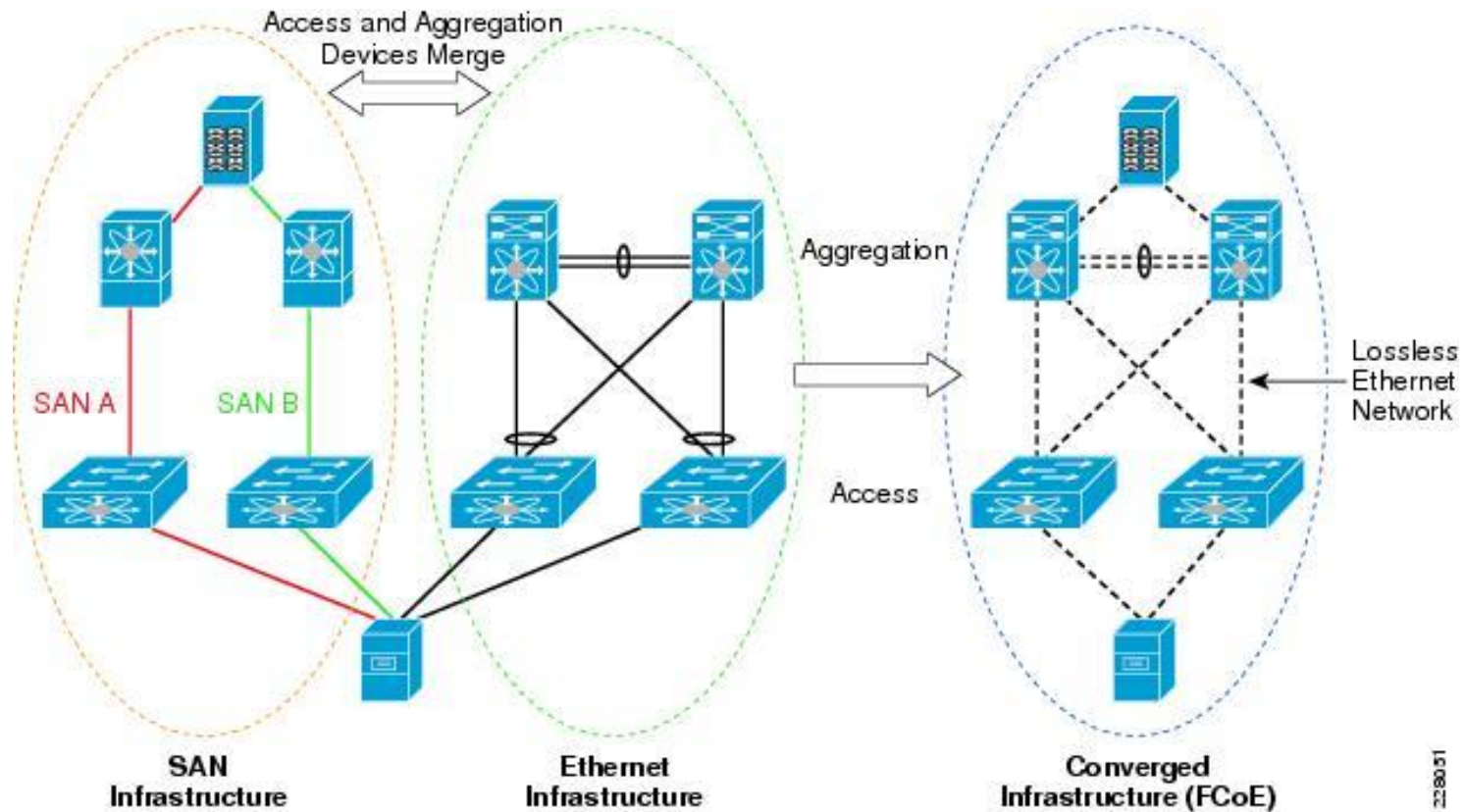
Converged IO Nic Adapter (CNA)



# What is converged IO?



# What is converged IO?



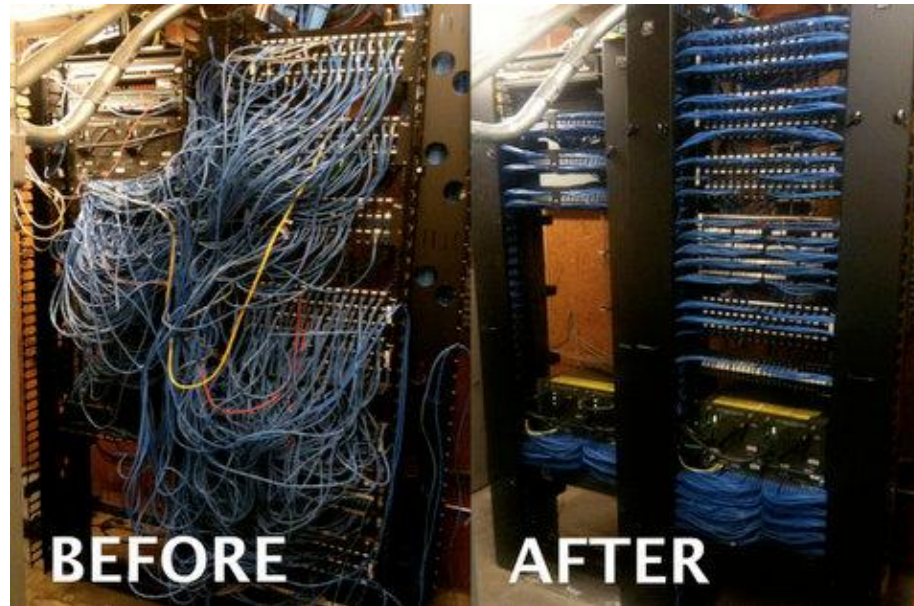
Who does it?





# Why would you use it?

Reduce stuff and costs



Savings:

30% - 50% in capital expenditures

70% - 80% in cabling

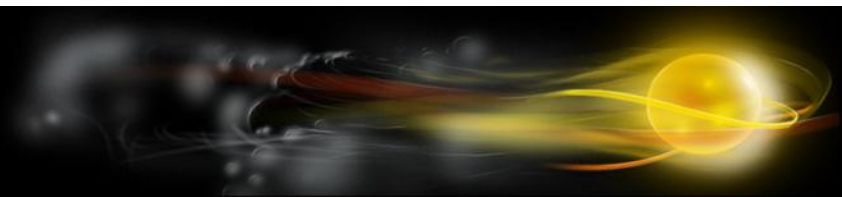
# Why would you use it?



utilise cheap bandwidth more efficiently

get rid of fiber channel,  
without getting rid of fiber channel

blades pre-provisioning  
power + management + network  
blades force consolidation

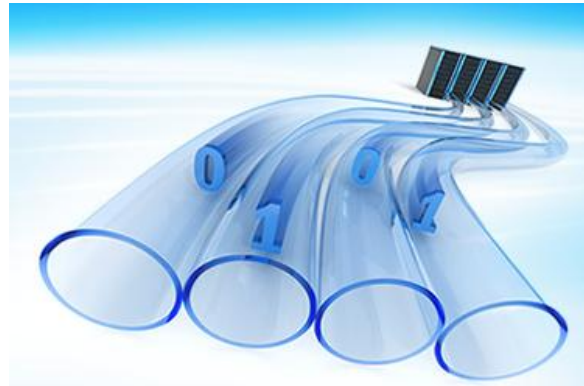




# How?

## Physical Convergence

consolidate multiple 1Gb Nics to single 10Gb Nic  
consolidate Ethernet and Fiber Channel to FCoE

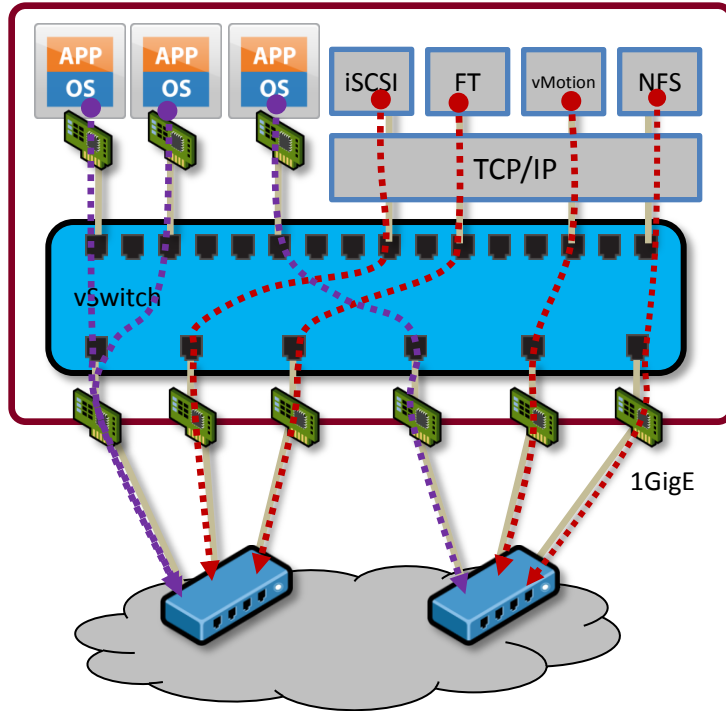


## Logical Divergence

carve up 10Gb Nic to multiple logical Nics

# Ethernet Multi 1Gb -> 10GbE

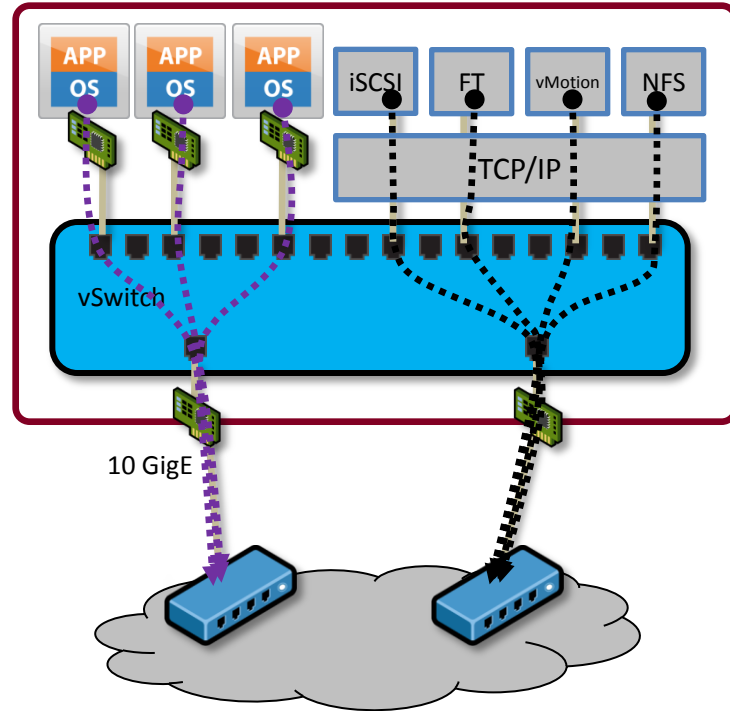
1GigE pNICs



Dedicated NICs for traffic types

Bandwidth assured by dedicated physical NICs

10 GigE pNICs

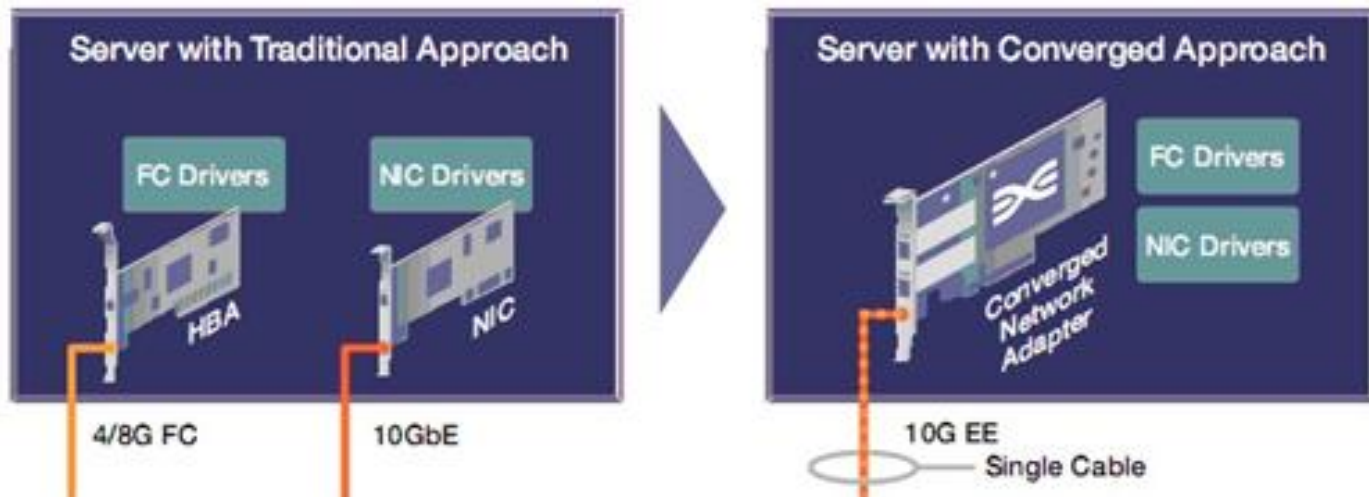


Traffic converged to two 10 GigE NICs

Traffic types could dominate others through oversubscription

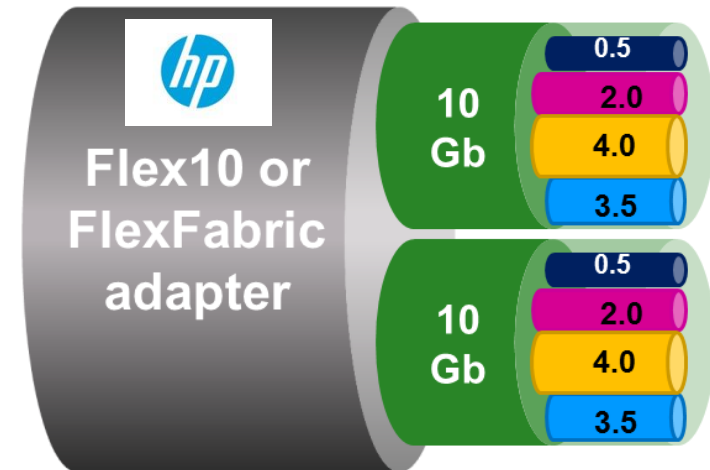
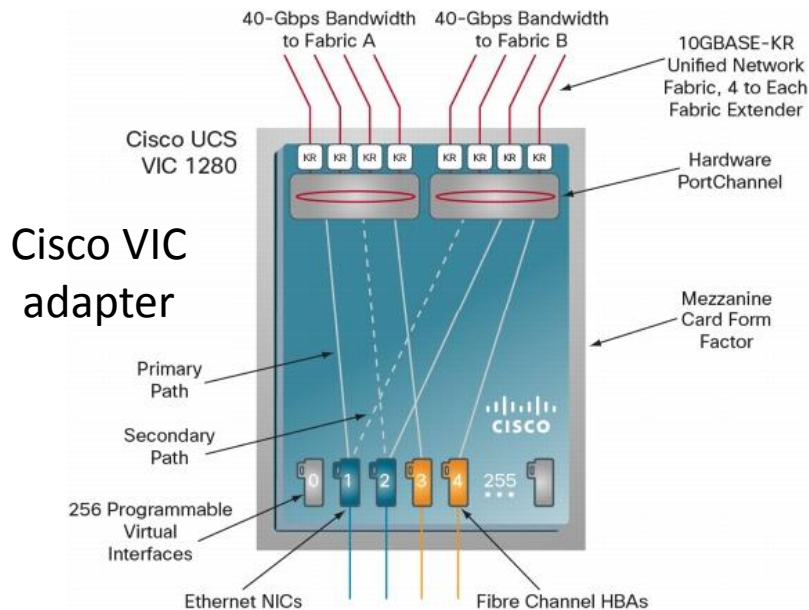
# Ethernet + Fiber Channel -> FCoE

logically carve up CNA



host sees 2 x 10GbE + 2 x HBA

# Carve 10GbE -> logical Nics



host sees ? x ?GbE + (? x HBA)

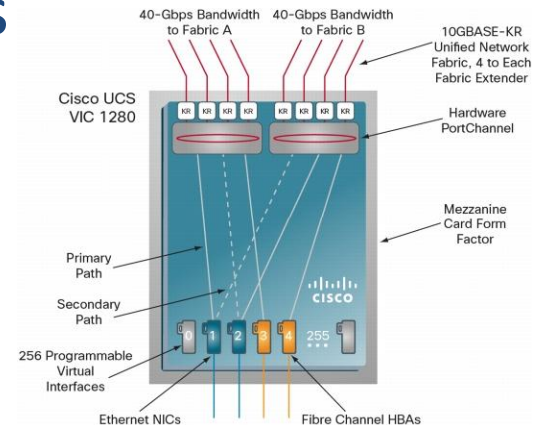
# Cisco UCS (Unified Computing System)

## VIC Cards

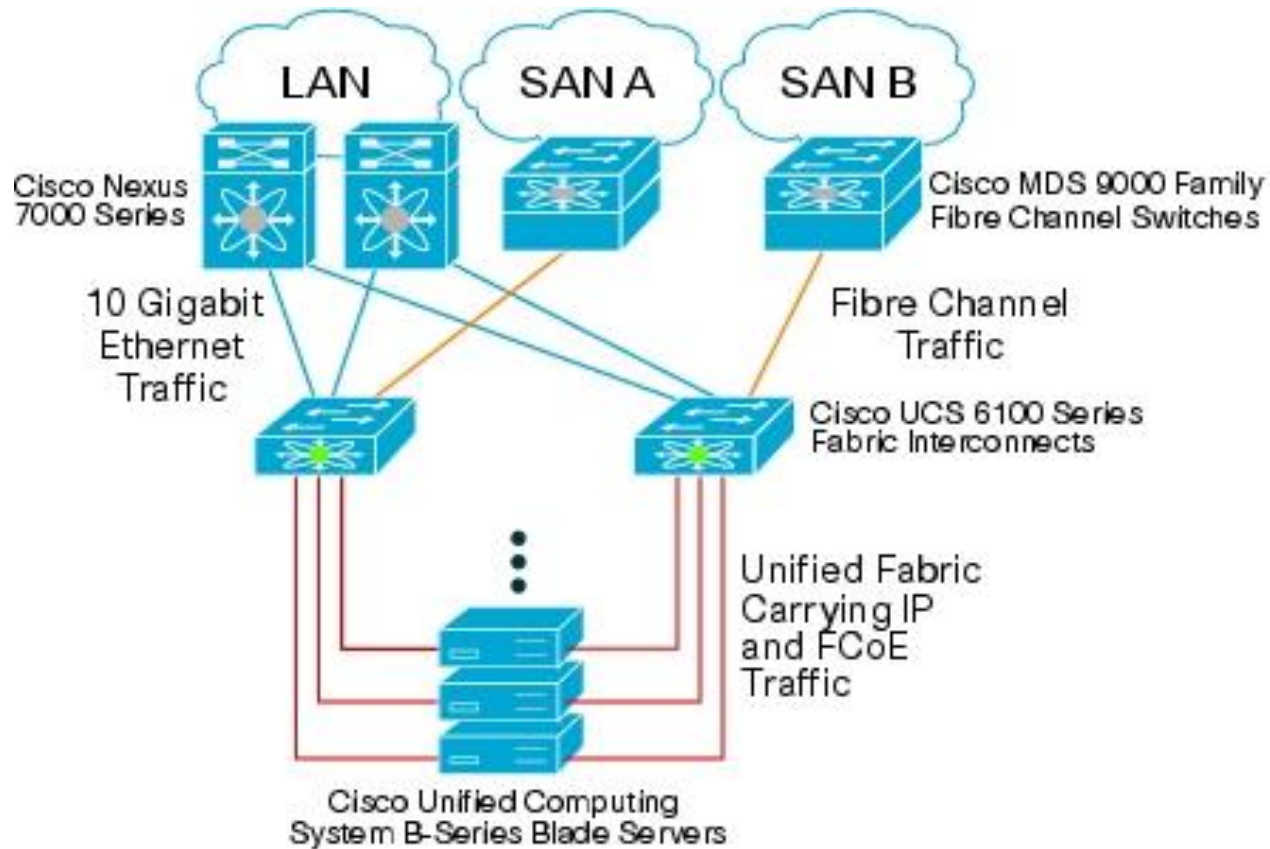
256 vNICs and vHBA and present them to the host.  
uses QoS Lanes to segregate traffic

switching external to chassis  
remote line cards  
fabric extenders (FEX)

Unified Computing System  
CISCO

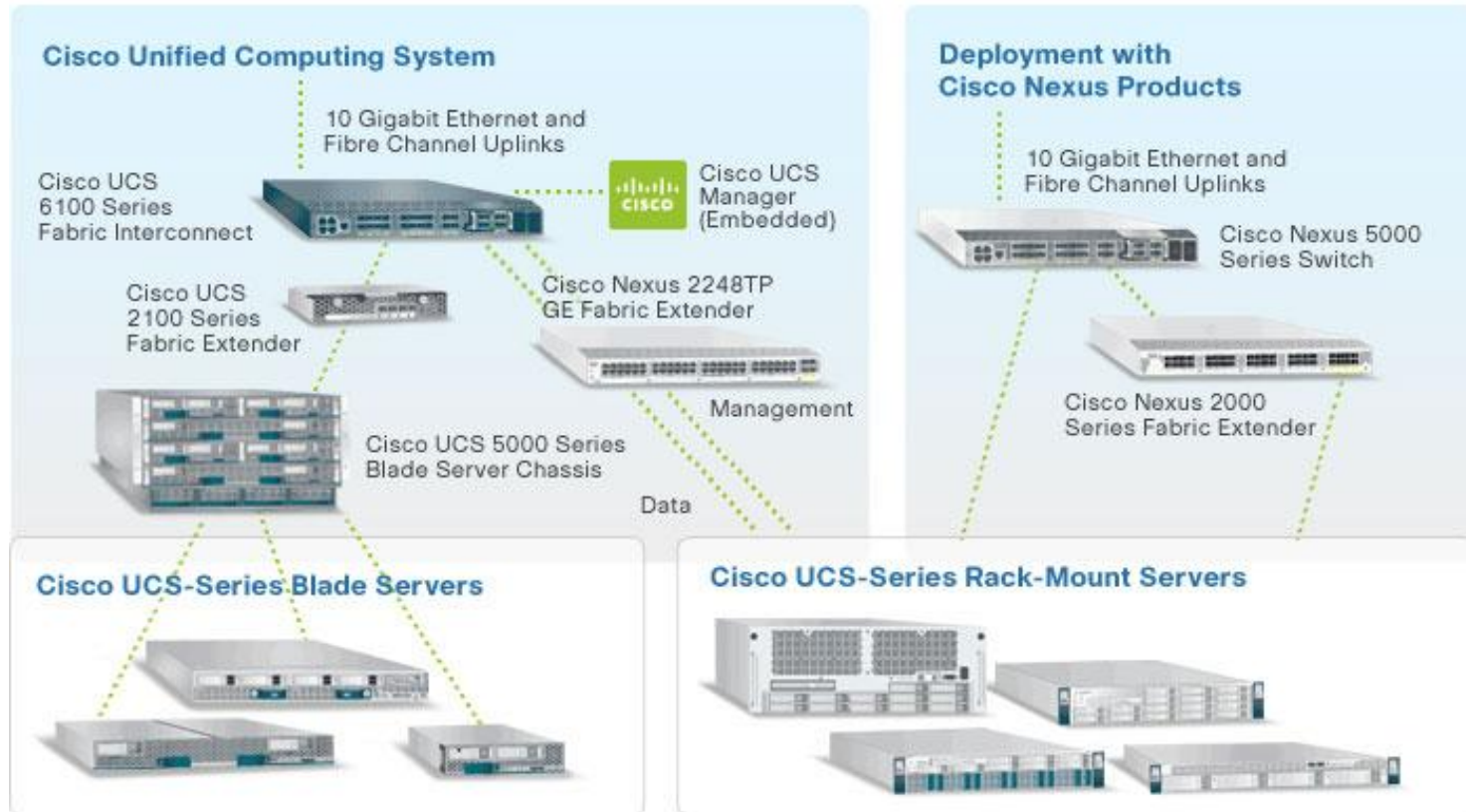


# Cisco UCS (Unified Computing System)





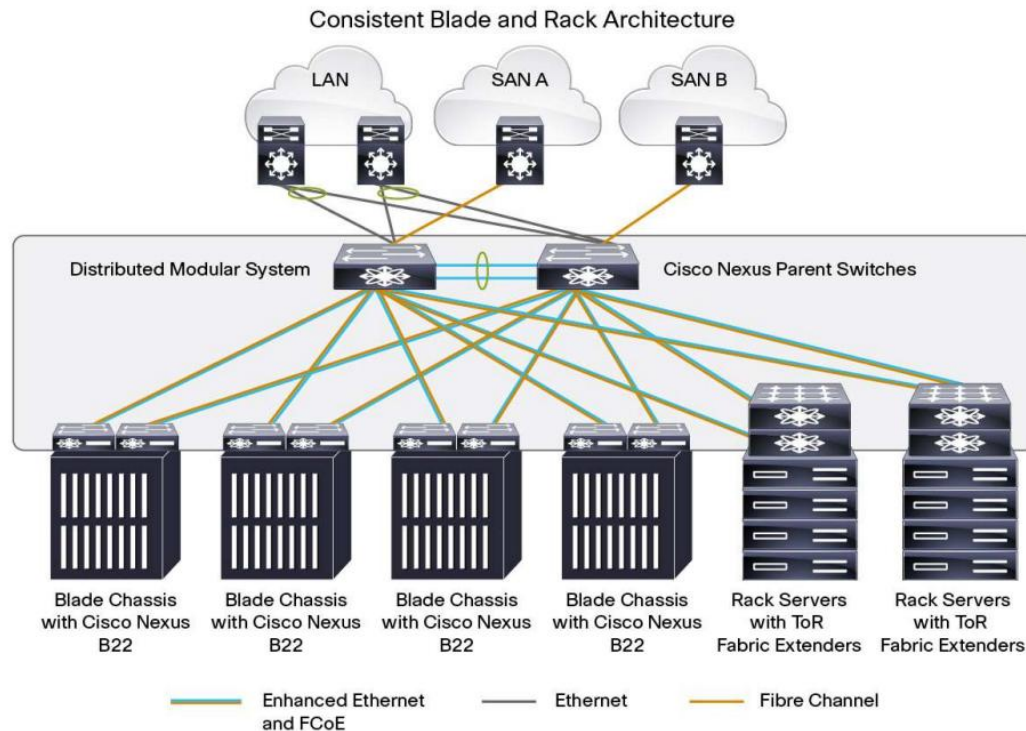
# Cisco UCS (Unified Computing System)



# HP BladeSystem + Cisco

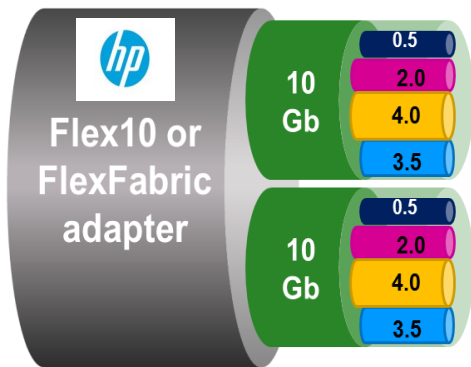
## Cisco Fabric Extender for HP BladeSystem (B22HP)

### 2 x 10GbE + Remote Line Cards (FEX)



# HP BladeSystem Virtual Connect

8 x FlexNics per blade



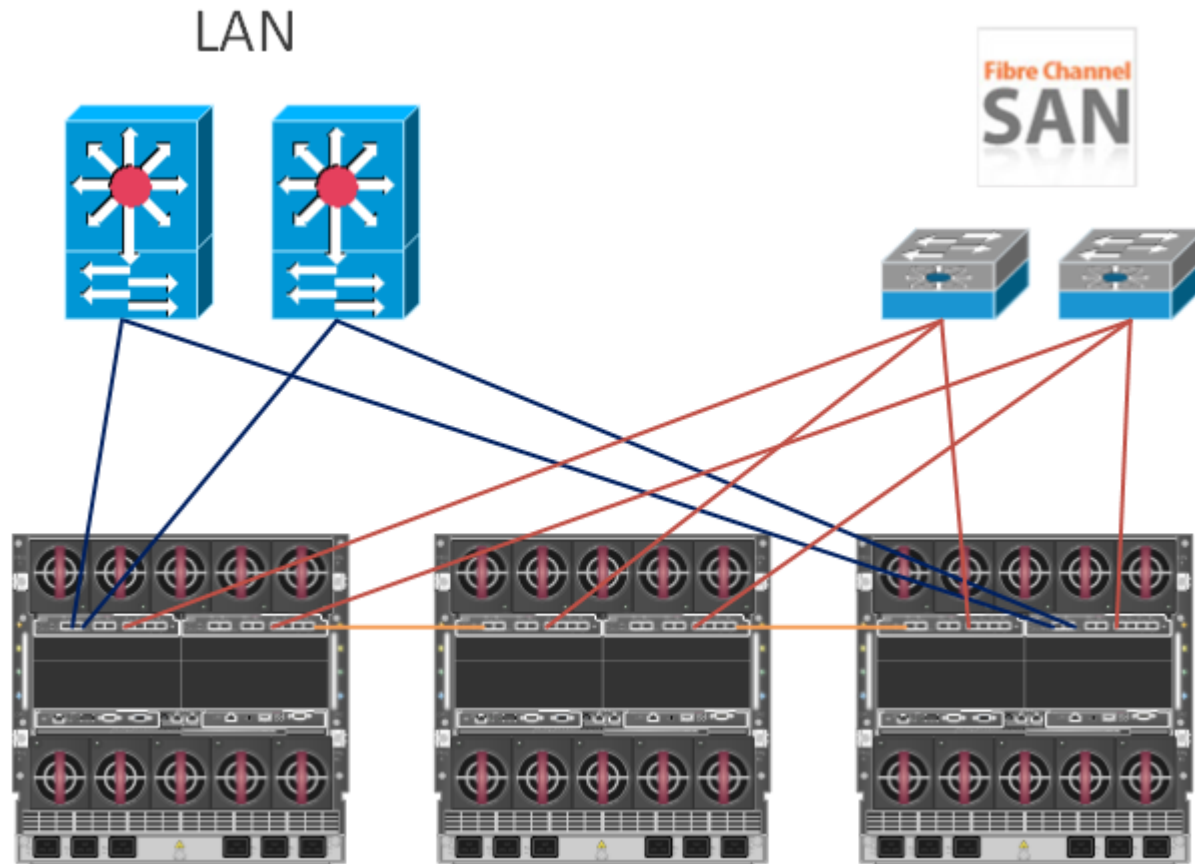
FlexFabric for FCoE  
6 x vNIC + 2 x vHBA

Flex-10 for Ethernet  
8 x vNIC

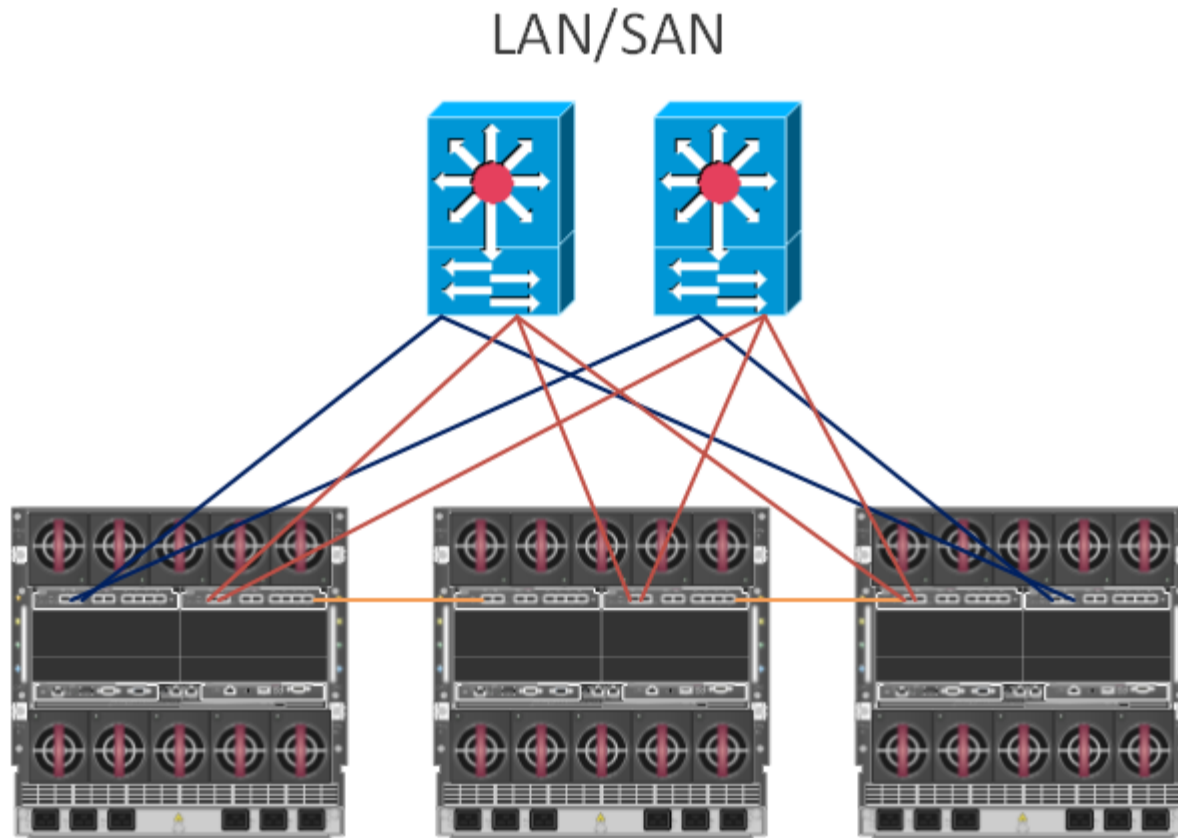


Virtual Connect switching internal to chassis  
no LACP across switches  
rate limiting to segregate traffic

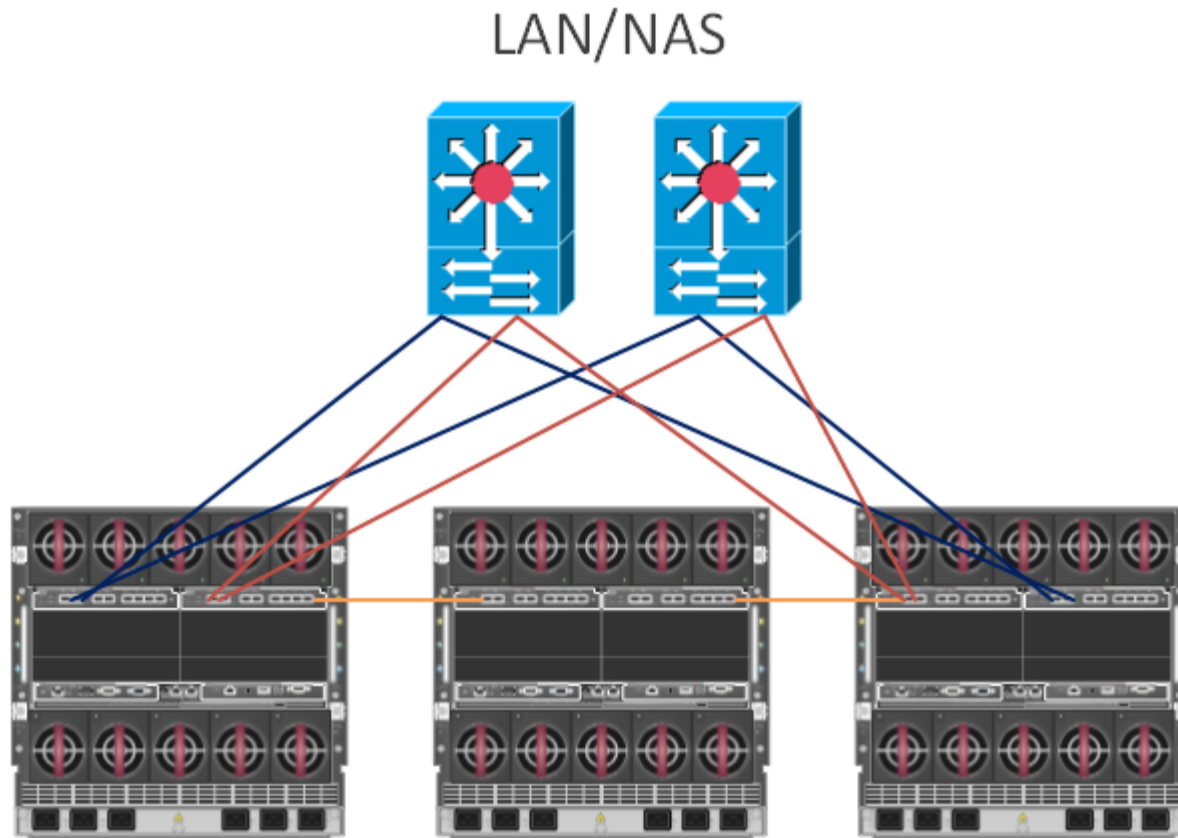
# HP BladeSystem Virtual Connect



# HP BladeSystem Virtual Connect

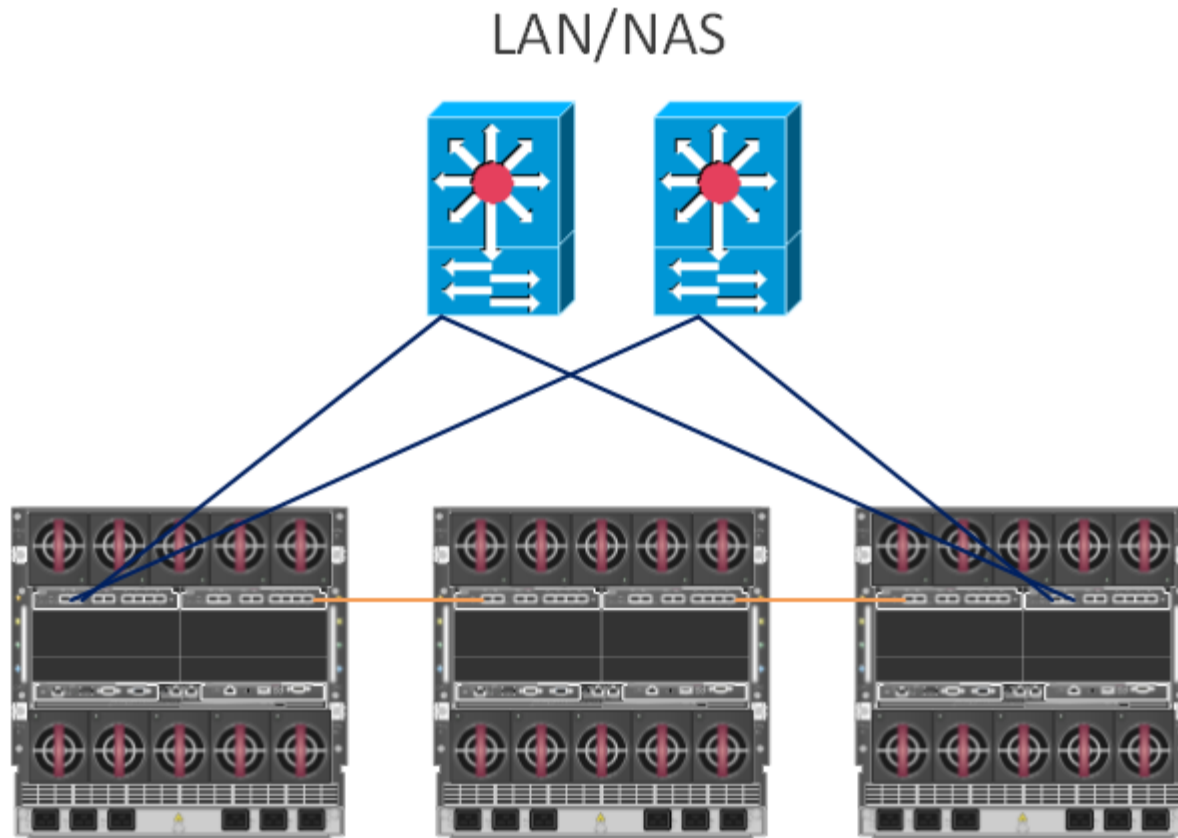


# HP BladeSystem Virtual Connect





# HP BladeSystem Virtual Connect



# Traffic Protection & Security

management

VM...

DMZ

backup

vMotion

ethernet storage

FCoE Storage

FT

Options:

physical hardware

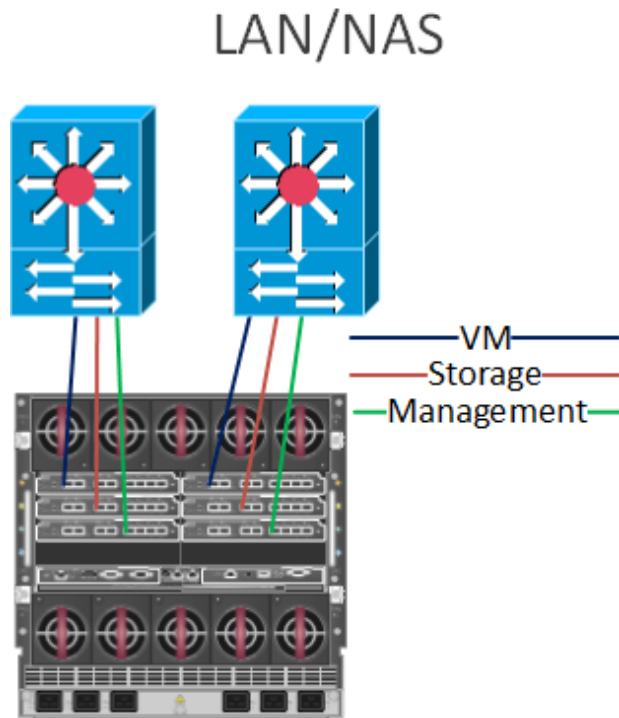
logical hardware

software

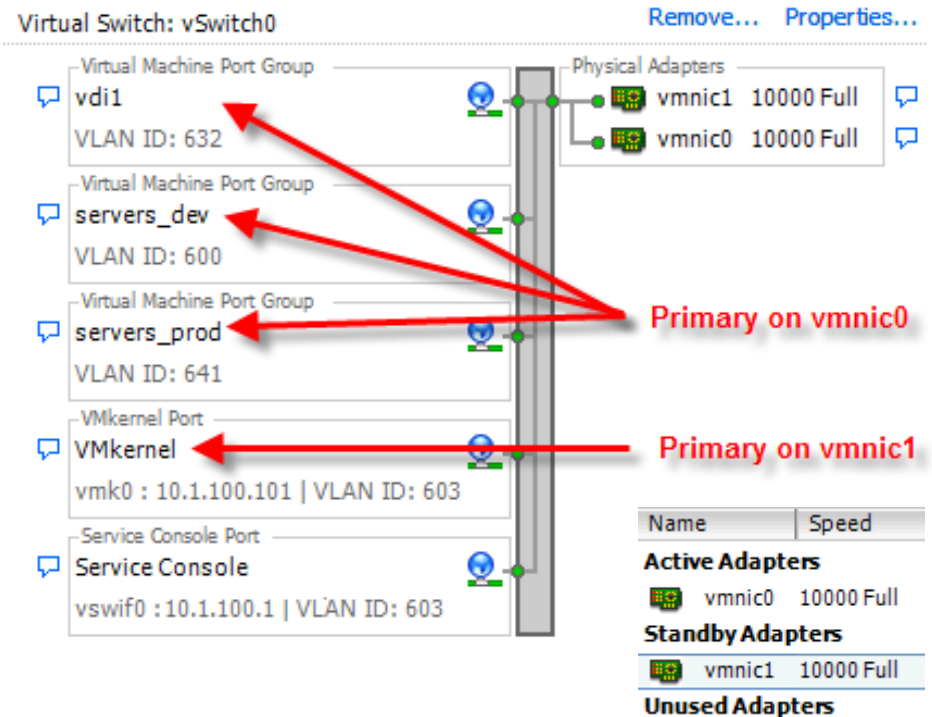


# Physical Hardware Separation

physical separation  
by pNics & uplinks



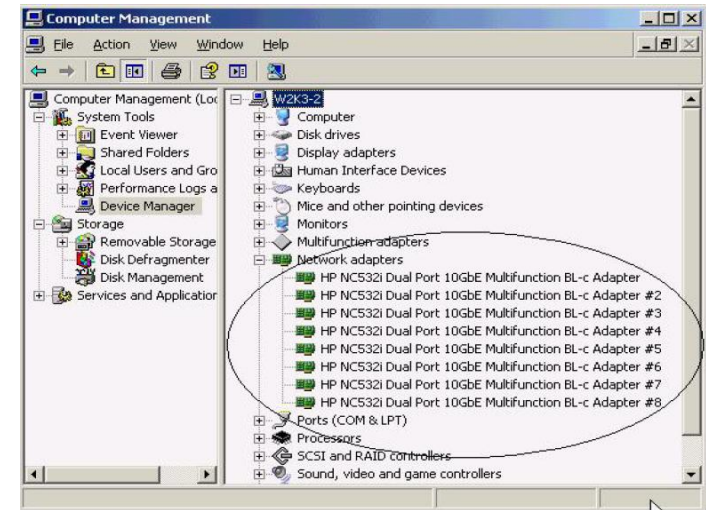
port group traffic direction



# Logical Hardware Separation

Cisco VIC/HP FlexNics  
multiple PCI Nics to host

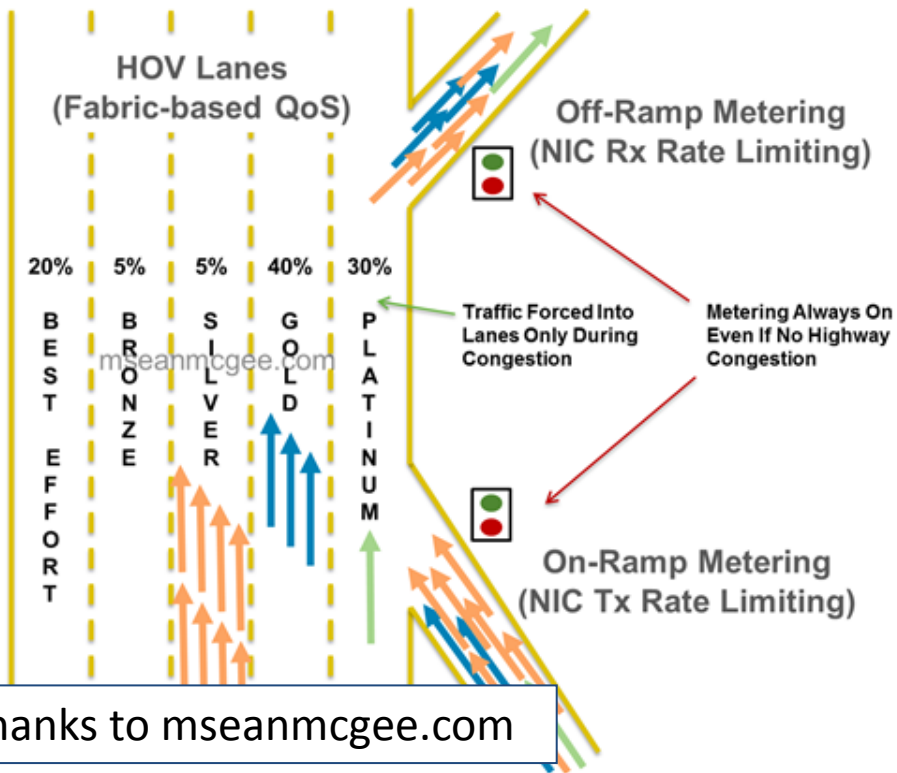
Device	Speed	Configured	Switch	MAC Address
<b>Broadcom Corporation NC532i Dual Port 10GbE Multifunction BL-C Adapter</b>				
vmnic0	2500 Full	Negotiate	vSwitch0	00:17:a4:77:24:f0
vmnic1	2500 Full	Negotiate	vSwitch0	00:17:a4:77:24:f2
vmnic2	2500 Full	Negotiate	None	00:17:a4:77:24:f4
vmnic3	2500 Full	Negotiate	None	00:17:a4:77:24:f6
vmnic4	2500 Full	Negotiate	None	00:17:a4:77:24:f8
vmnic5	2500 Full	Negotiate	None	00:17:a4:77:24:fa
vmnic6	2500 Full	Negotiate	None	00:17:a4:77:24:fc
vmnic7	2500 Full	Negotiate	None	00:17:a4:77:24:fe



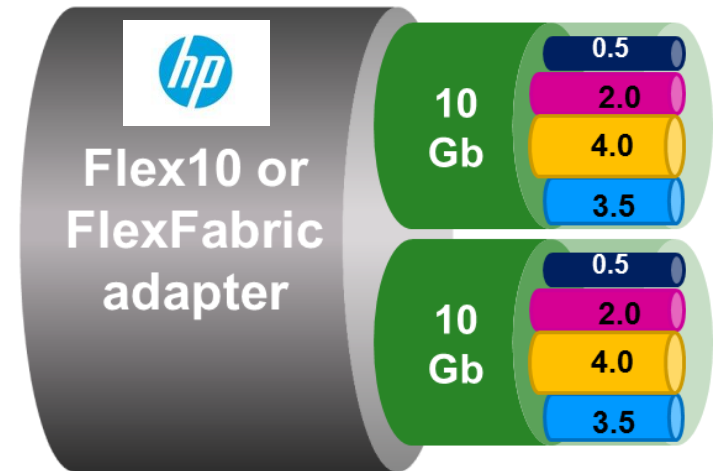
multiple traffic lanes  
convergence at the switch

# Logical Hardware Separation

Cisco VICs = 8 x QoS lanes  
(queuing/shares)

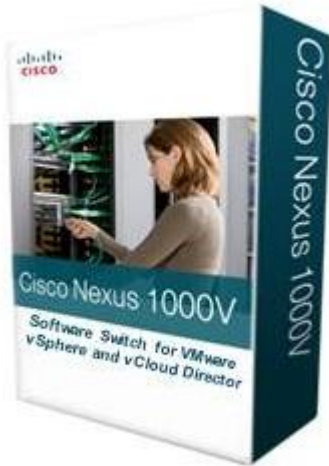


HP FlexNics  
rate limiting on send  
(policing)



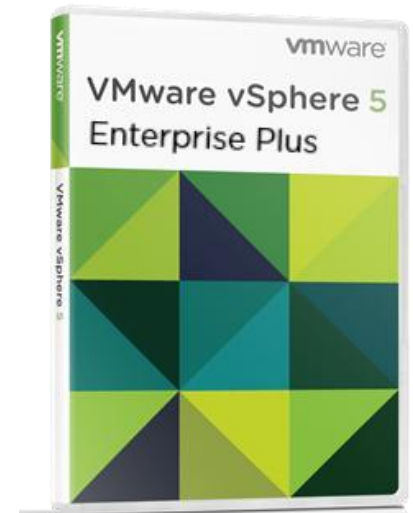
thanks to mSeanmcgee.com

# Software Segregation



Cisco Nexus 1000V Virtual Switch  
QoS (802.1p)

VMware vSphere Distributed Switch (VDS)  
Network I/O Control (NIOC)





# Cisco Nexus 1000V

virtual switch installed on ESXi  
replaces VMware switch  
acts as line card of modular switch  
each VM on a port of 1000V  
managed like physical servers  
port profiles = network config  
port profiles -> vCenter Port Groups  
QoS

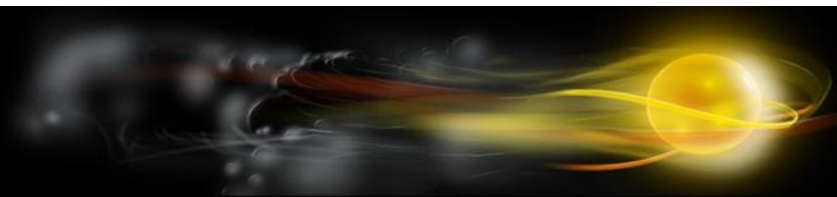


# Cisco Nexus 1000V QoS

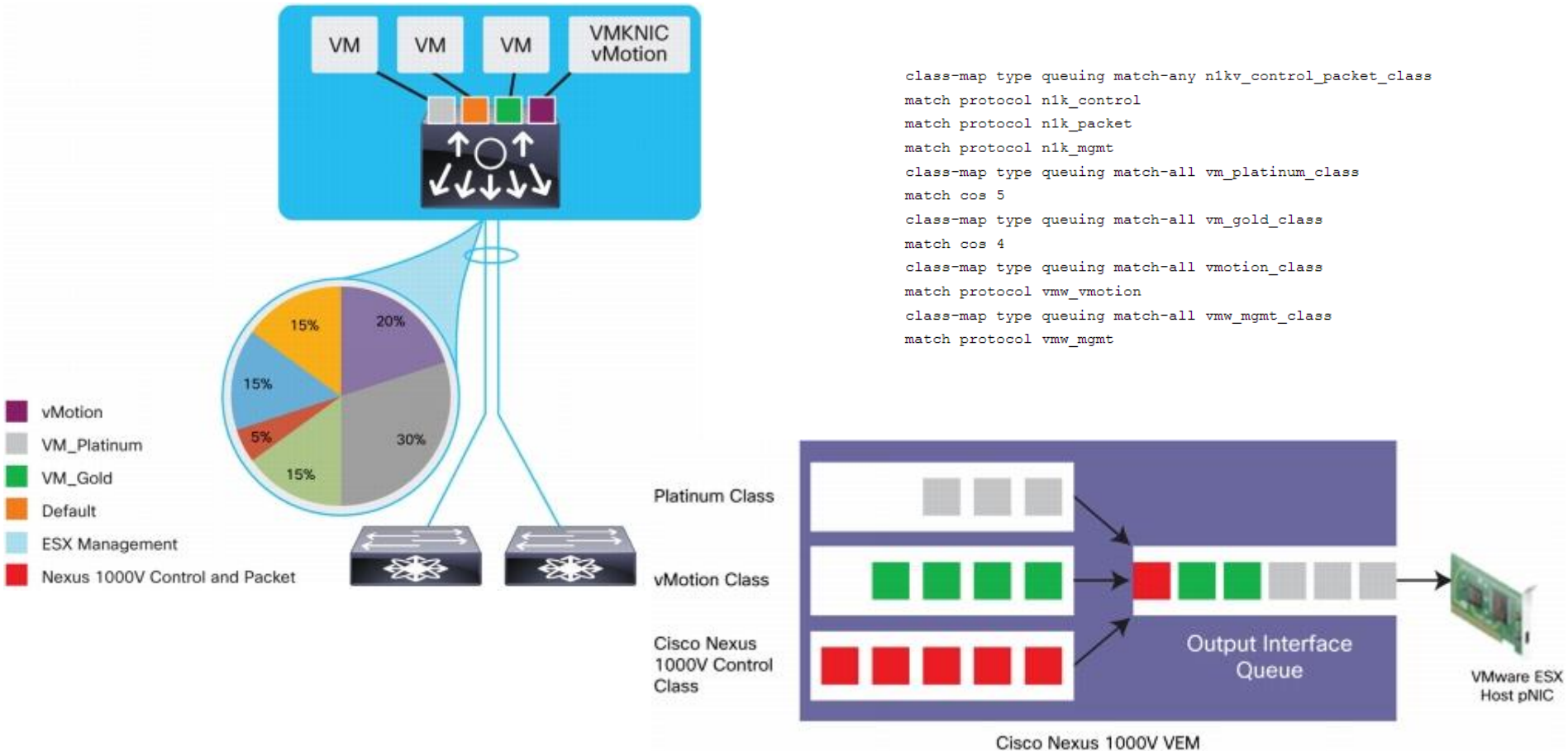


traffic marking  
before hitting adapter  
similar to UCS VIC QoS  
not FCoE/UDP

class-based weighted fair queuing (CBWFQ)  
guaranteed minimum bandwidth = shares  
all the way up to the switch  
802.1p world boundary (no HP VC)



# Cisco Nexus 1000V QoS



# VMware Network I/O Control

requires vSphere Distributed Switch (Ent+)

outbound only, not marked

before hitting adapter

not FCoE/UDP

Load-based Teaming

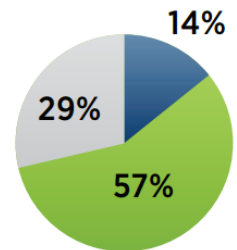
isolation, shares, limits

resource pools

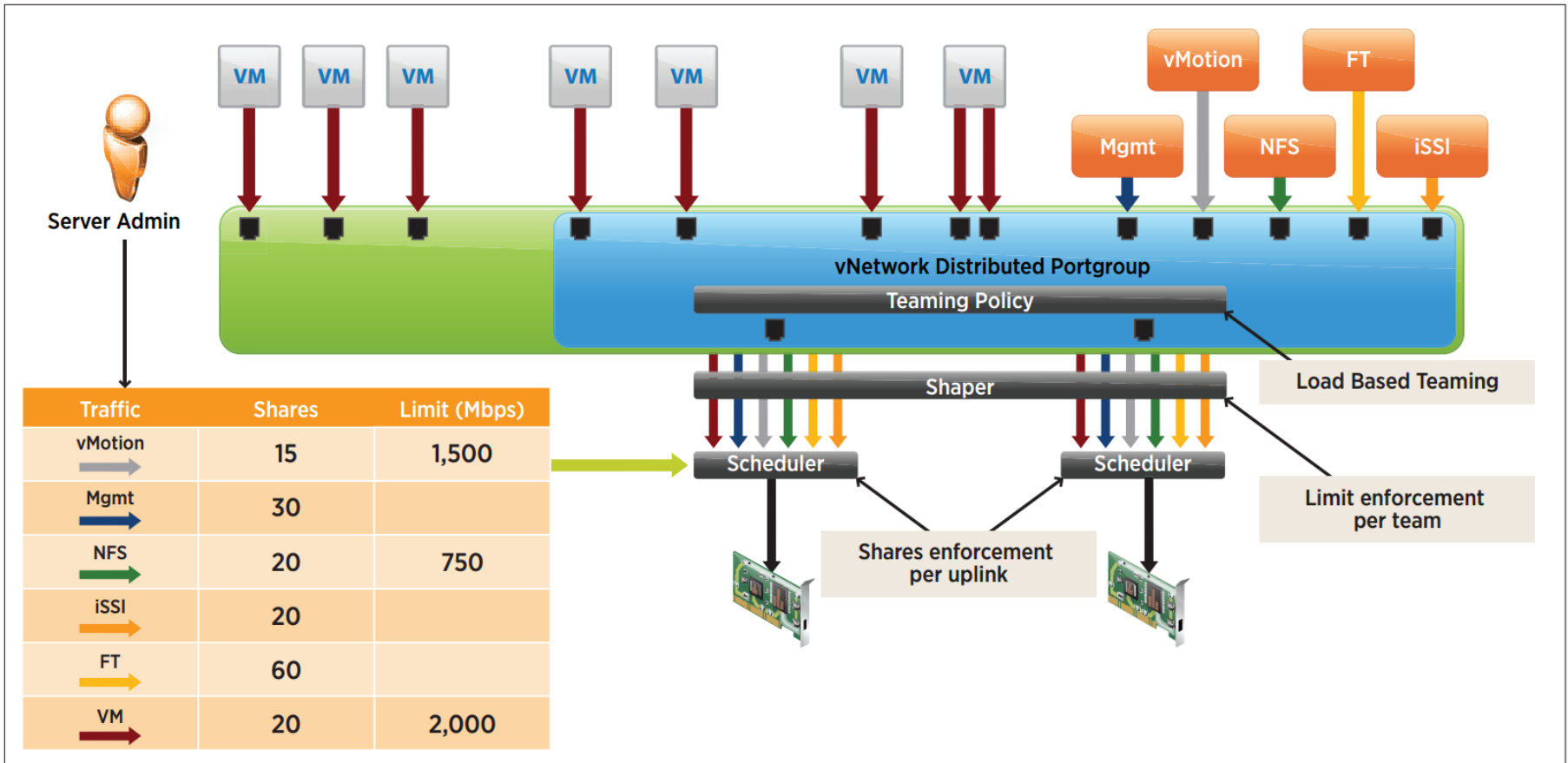
VM  
NFS  
iSCSI

vMotion  
FT  
Management

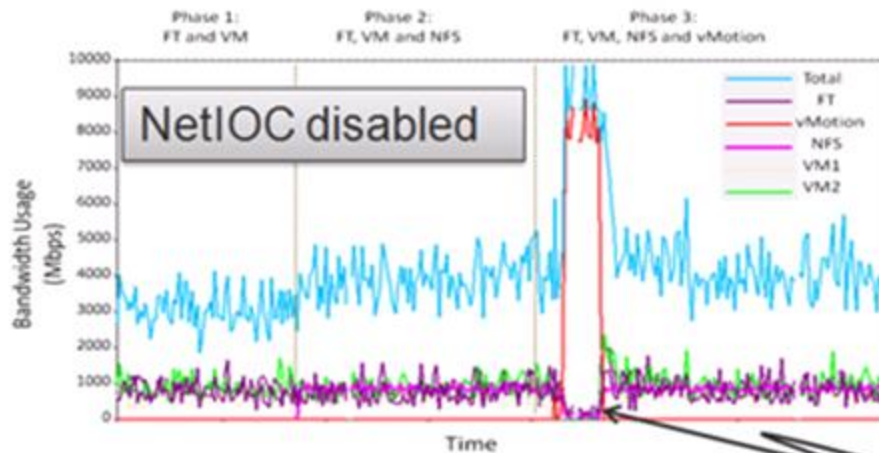
vSphere Replication  
User Defined



# VMware Network I/O Control



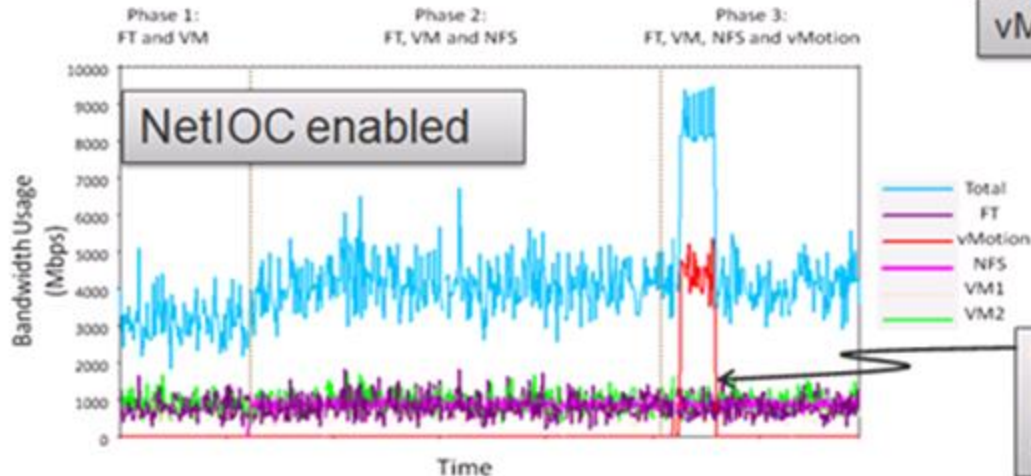
# VMware Network I/O Control



Shares & Limits as Configured

<b>Summary</b>			
Total number of physical adapters: <b>2</b>			
Total network bandwidth capacity: <b>20000 Mbit/s</b>			
Network I/O control: <b>Enabled</b>			
Network resource pool	Host limit - Mbit/s	Physical adapter shares	Shares value
FT Traffic	Unlimited	High	100
iSCSI Traffic	Unlimited	Custom	75
vMotion Traffic	Unlimited	Normal	50
Management Traffic	Unlimited	Normal	50
NFS Traffic	Unlimited	Custom	75
Virtual Machine Traffic	Unlimited	Custom	75

NFS, VM, and FT traffic take a dip during vMotion



NFS, VM, and FT traffic not affected by concurrent vMotion



# Future

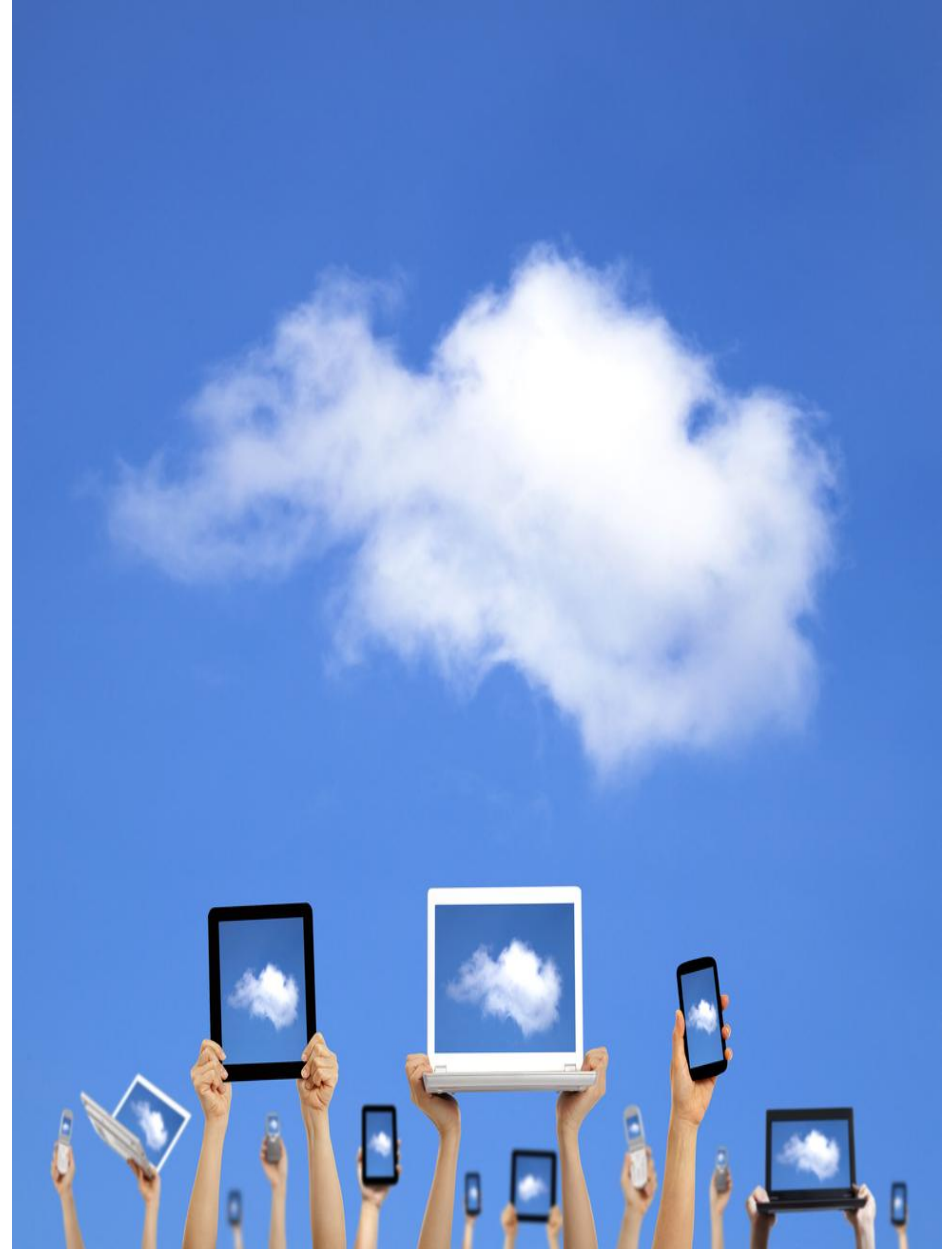
more bandwidth

standard switch is on notice

VXLAN

local storage (VSA/vSAN)

VMware vVolumes



# After all that...

10 GbE bandwidth is a lot!

logical separation or 2 x 10GbE with QoS/NIOC

10GbE minus FCoE

where is the switch?

availability/load balancing at the host

all uplinks active, don't waste bandwidth

pod approach

testing

VBlock/FlexPod/HP BladeSystem Matrix



# Resources

[wooditwork.com](http://wooditwork.com)

[bradhedlund.com](http://bradhedlund.com)

[mseanmcgee.com](http://mseanmcgee.com)

[rayheffer.com](http://rayheffer.com)

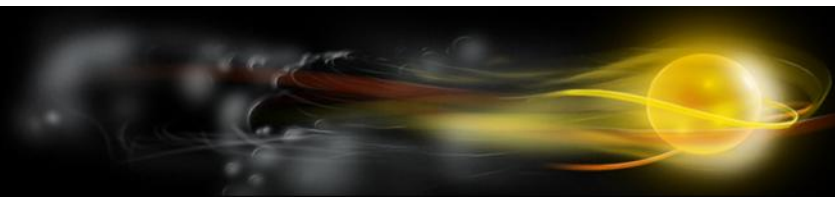
[www.vmware.com/files/pdf/techpaper/VMW\\_Netioc\\_BestPractices.pdf](http://www.vmware.com/files/pdf/techpaper/VMW_Netioc_BestPractices.pdf)

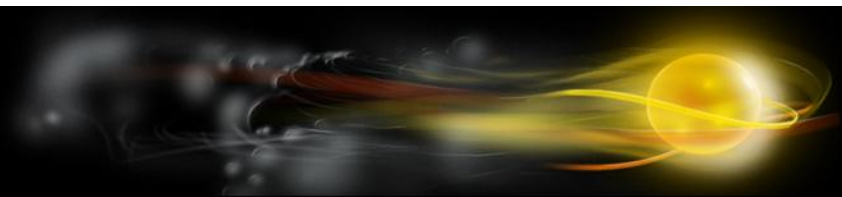
[www.cisco.com/en/US/products/ps9902/index.html](http://www.cisco.com/en/US/products/ps9902/index.html)

[www.cisco.com/en/US/docs/solutions/Enterprise/Data\\_Center/UF\\_FCoE\\_final.html](http://www.cisco.com/en/US/docs/solutions/Enterprise/Data_Center/UF_FCoE_final.html)

[h18004.www1.hp.com/products/blades/virtualconnect/index.html](http://h18004.www1.hp.com/products/blades/virtualconnect/index.html)

[www.viktorious.nl/2012/11/12/featured-vmug-presentation-designing-a-vsphere-deployment-for-cisco-ucs-blades](http://www.viktorious.nl/2012/11/12/featured-vmug-presentation-designing-a-vsphere-deployment-for-cisco-ucs-blades)





# **vSphere Networking and Converged IO with Blade Servers**

Julian Wood

UK VMware User Group – 15<sup>th</sup> November 2012

#UKVMUG

