Cisco Connect
Seoul, Korea
April 29-30, 2014

성공적인 VDI 구축을 위한 실전 가이드

유 승 만 과장

seyoo@cisco.com

Cisco Systems, DataCenter PSE

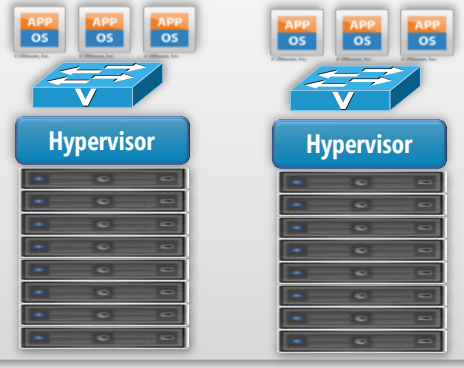
Agenda

- 성공적인 VDI 구축을 위한 도전 과제
- VDI 컴퓨팅 & 스토리지 고려사항
- 가상 네트워킹 고려사항

성공적인 VDI 구축을 위한 도전 과제

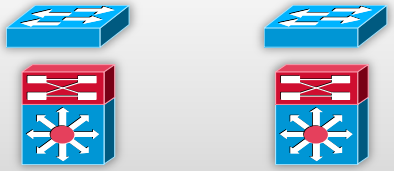
성공적인 VDI 디자인을 위한 고려사항

컴퓨팅



x86 Computing
가상 머신
가상 스위치
하이퍼바이저

네트워크



Server Switch
Core Switch

스토리지 네트워크

NAS Switch
SAN Switch

스토리지

Storage 프로세서
Disk

x86 Computing
- Cpu, Memory 선택
- IO 선택 (1/10GE , FC, iSCSI, FCoE)
- Virtualization 기술 (Intel VT, SR-IOV, VIC...)

하이퍼바이저 고려 사항
- VMWare ESX / Citrix Xen/ MS HyperV/ Redhat KVM...
- Virtual Switch에 대한 고민

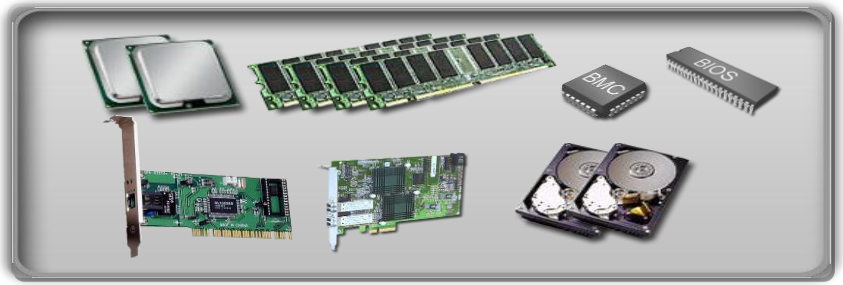
네트워크 고려 사항
- Unified Fabric (Ethernet, SAN)
- Unified IO (10GE, iSCSI, NAS, NFS, CIFS, FC, FCoE)
- Network Virtualization
- 가상화 레벨의 네트워과 물리적 네트워의 연동
- 가상화 환경의 보안
- 스토리지 프로토콜에 대한 고려 사항

Storage 고려 사항
- HDD Storage / SSD Storage 선택
- SAS, SATA, SSD disk 선택

VDI 컴퓨팅 & 스토리지 고려사항

VDI 서버 디자인

서버 타입 선택에 대한 이슈
- 랙서버? 블레이드 서버?
- 서버 집적도, 편의성



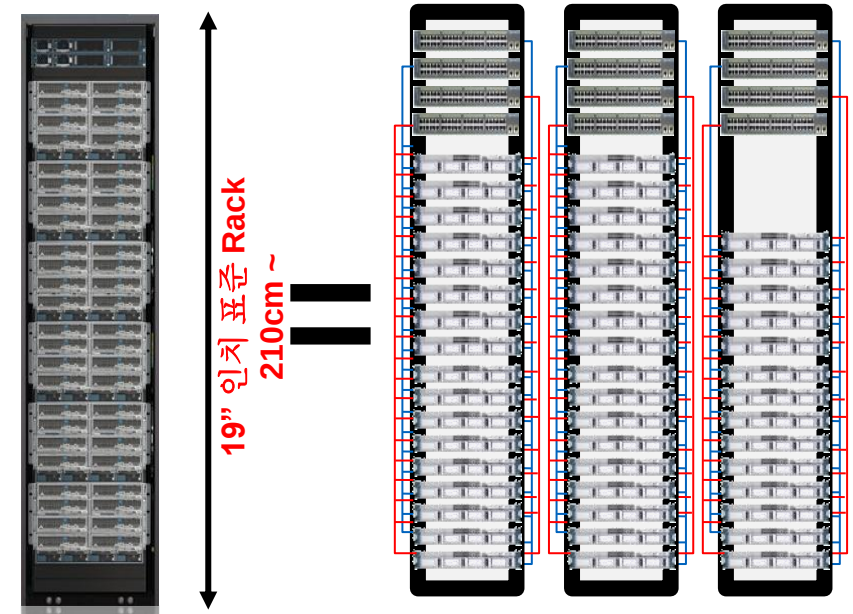
CPU 선택에 대한 이슈
- Ivy Bridge EP, EX, EN ?
- 고사양 CPU ?

Memory 선택에 대한 이슈
- 4GB , 8GB, 16GB, 32G … 저전력 메모리… Memory Speed
- 메모리 집적도

HDD 선택
- Disk Type : SATA , SAS, SSD & RPM, RAID
- IO Protocol : FC , iSCSI, NAS

IO 선택
- Bandwidth : 1GE, 10GE, 4G FC/ 8G FC
- IO : Ethernet , FCoE, FC

VDI 서버 디자인 – Rack vs Blade



Cisco UCS = 48

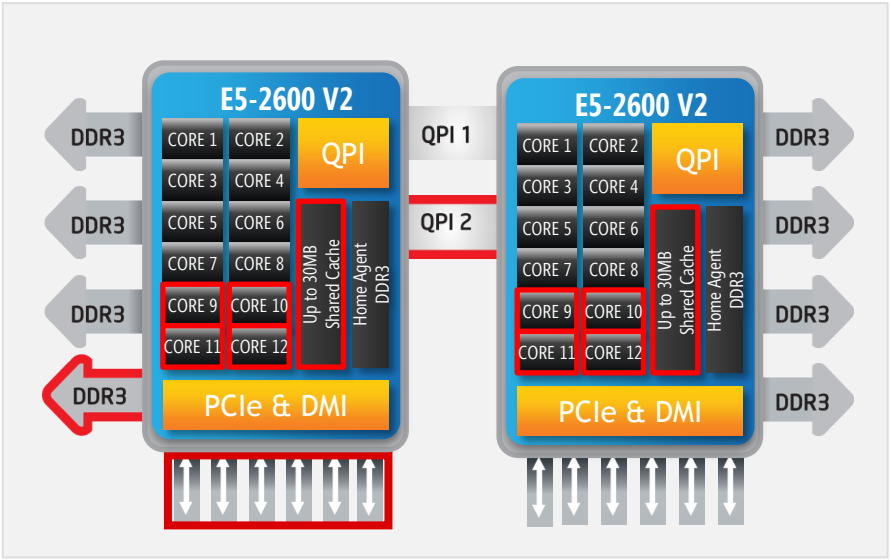
RackMount Server = 48

VDI 컴퓨팅

- 랙서버 구성시
  - 장점
    - 초기 하드웨어 투자 비용이 적게 듬
    - 하드웨어 설치 및 구성이 비교적 간단함
  - 단점
    - 상면 집적도와 확장성이 떨어짐
    - 랙단위로 네트웍 스위치, 샌스위치 컴포넌트가 구성되는 Silo 형태의 비효율적 구성
    - 1000 VDI 이상일 경우 시스템 통합관리가 어려움
- 블레이드 서버 구성시
  - 장점
    - 상면 집적도와 확장성이 우수함
    - 네트웍, 샌스위치 통합을 디자인 효율성이 우수함
    - 시스템 통합 관리가 용이함
  - 단점
    - 초기 투자 비용이 랙서버 대비 높음
    - 블레이드 하드웨어 설치,구성 및 운영에 대한 부담

VDI 서버 디자인 - CPU

기 능	E5-2600 (Sandy Bridge-EP)	E5-2600 V2 (Ivy Bridge-EP)
QPI Speed (GT/s)	6.4 / 7.2 / 8.0	
소켓당 코어	최대 8	최대 12
소켓당 쓰레드	최대 16	최대 24
Last-Level Cache(LLC)	최대 20MB	최대 30MB
Intel Turbo-Boost Technology	지원	
메모리 구성	4채널 / 채널당 3개의 RDIMMs 또는 LRDIMMs	
최대 메모리 속도	최대 1600Mhz	최대 1866Mhz
메모리 RAS 기능	ECC, Patrol Scrubbing, Demand Scrubbing, Sparing, Mirroring, Lockstep Mode, x4/x8 SDDC	
PCIe Lanes/Controllers Speed(GT/s)	40 / 10 (PCIe 3.0 @ 8 GT/s)	
최대 전력(W)	130, 115, 95, 80, 70, 60	
대기 전력(W)	15W ~ (12W for LV SKUs)	10.5W ~ (7.5W for LV SKUs)



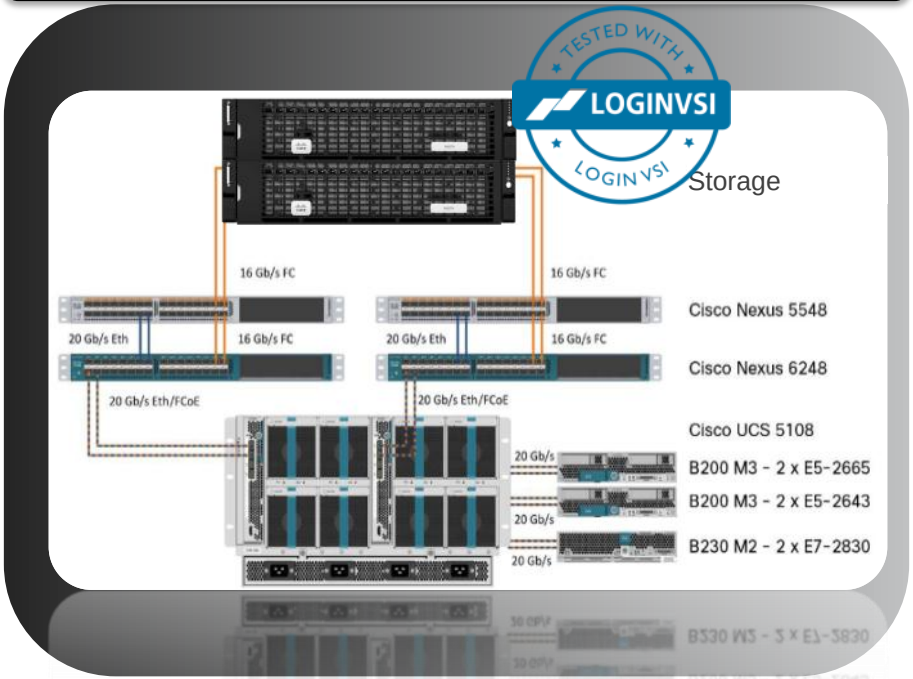
[Intel E-5 2600 v2 CPU 아키텍처]

VDI 서버 디자인 - CPU

Product ID (PID)	Intel Number	Clock Freq(GHz)	전력 소모량(W)	캐쉬 사이즈(MB)	코어	QPI	DIMM 최고 Clock(DDR3)
UCS-CPU-E5-2697B	E5-2697 v2	2.70	130	30	12	8 GT/s	1866
UCS-CPU-E5-2695B	E5-2695 v2	2.40	115	30	12	8 GT/s	1866
UCS-CPU-E5-2690B	E5-2690 v2	3.00	130	25	10	8 GT/s	1866
UCS-CPU-E5-2680B	E5-2680 v2	2.80	115	25	10	8 GT/s	1866
UCS-CPU-E5-2670B	E5-2670 v2	2.50	115	25	10	8 GT/s	1866
UCS-CPU-E5-2667B	E5-2667 v2	3.30	130	25	8	8 GT/s	1866
UCS-CPU-E5-2660B	E5-2660 v2	2.20	95	25	10	8 GT/s	1866
UCS-CPU-E5-2650B	E5-2650 v2	2.60	95	20	8	8 GT/s	1866
UCS-CPU-E5-2640B	E5-2640 v2	2.00	95	20	8	8 GT/s	1600
UCS-CPU-E5-26337B	E5-2637 v2	3.50	130	15	4	8 GT/s	1866
UCS-CPU-E5-2630B	E5-2630 v2	2.60	80	15	6	8 GT/s	1600
UCS-CPU-E5-2620B	E5-2620 v2	2.10	80	15	6	8 GT/s	1600
UCS-CPU-E5-2643B	E5-2643 v2	3.50	130	25	6	8 GT/s	1866
UCS-CPU-E5-2650LB	E5-2650L v2	1.70	70	25	10	8 GT/s	1600
UCS-CPU-E5-2630LB	E5-2630L v2	2.40	60	15	6	8 GT/s	1600
UCS-CPU-E5-2609B	E5-2609 v2	2.50	80	10	4	6.4 GT/s	1333

VDI 서버 디자인 – CPU

VDI 테스팅 토폴로지



성능 비교 CPU 스펙

- Xeon E5-2643, 4core, 3.30 Mhz
- Xeon E5-2665, 8core, 2.40 Mhz

가상화 소프트웨어

- ESX 5.0.0
- Vmware Horizon View 5.1.1

성능 측정 툴

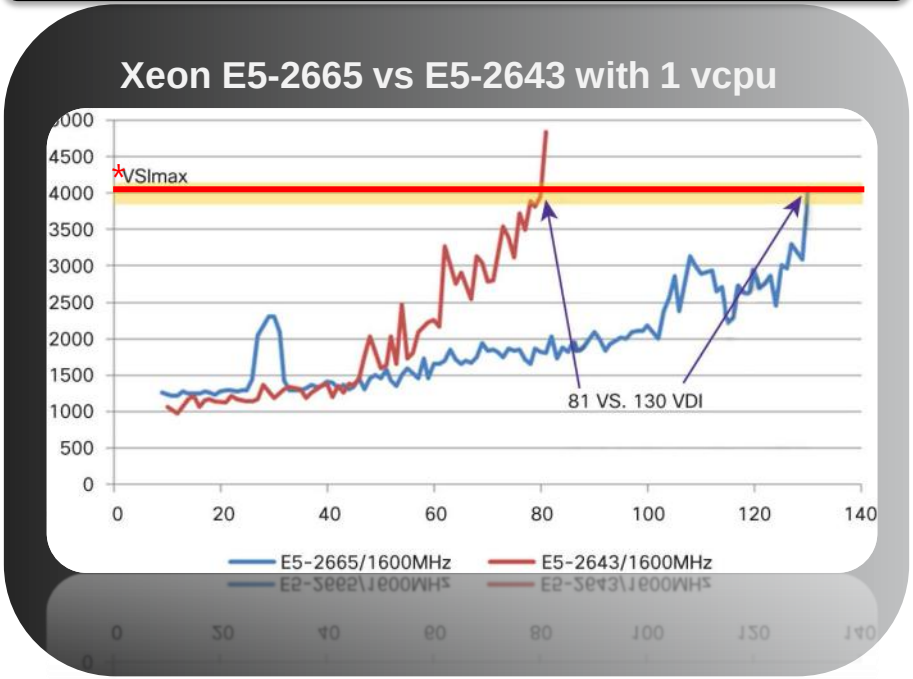
- Login VSI 3.6.1 Benchmark

VDI 운영체제

- Windows 7 32bit
- Memory 1.5G 할당

VDI 서버 디자인 – CPU

CPU (코어수)



*VSImax – 응답시간 대비 최대의 엑티브 세션/데스크탑 용량을 측정하는 기준

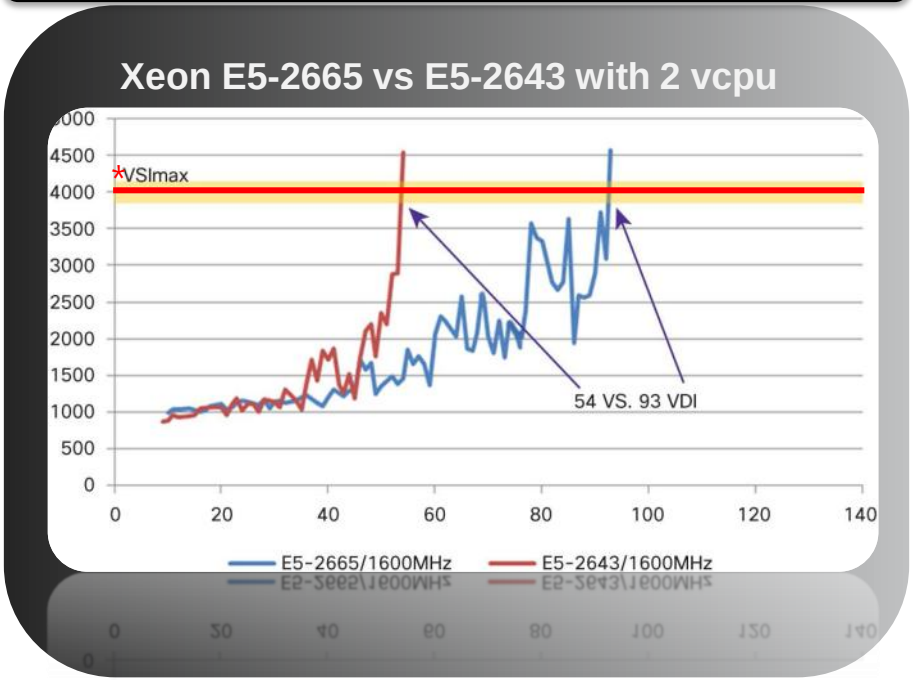
vCPU 구성 테스트 결과

- 집적도
  - E5-2643 CPU(4core, 3.3Ghz)
    - VSImax 기준 **최대 81개 VM**
  - E5-2665 CPU(8core, 2.4Ghz)
    - VSImax 기준 **최대 130개 VM**

- 결과
  - Winner
    - 코어수가 많은 E5-2665 CPU, **60%** 집적도 우수
  - 결과 분석(코어수 vs 코어 스피드)
    - **코어수가 많은 쪽**이 **CPU 오버커밋 향상 효과**가 우수 하며 집적도에 큰 영향을 미침

VDI 서버 디자인 – CPU

CPU (코어수)



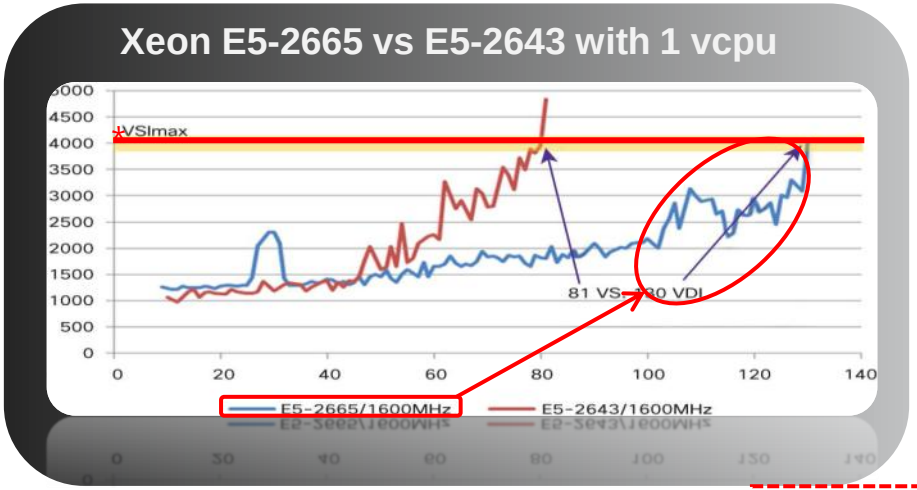
*VSImax – 응답시간 대비 최대의 엑티브 세션/데스크탑 용량을 측정하는 기준

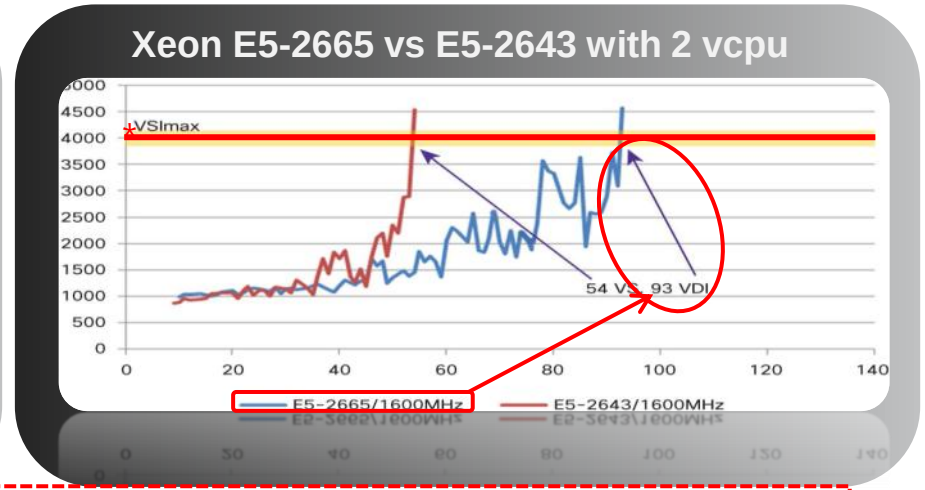
vCPU 구성 테스트 결과

- 집적도
  - E5-2643 CPU(4core, 3.3Ghz)
    - VSImax 기준 **최대 54개 VM**
  - E5-2665 CPU(8core, 2.4Ghz)
    - VSImax 기준 **최대 93개 VM**

- 결과
  - Winner
    - 코어수가 많은 E5-2665 CPU, **72%** 집적도 우수
  - 결과 분석(코어수 vs 코어 스피드)
    - **CPU 오버커밋 증가로 집적도가 조금 떨어지나 코어수가 많은 쪽이** 집적도가 더 우수함

VDI 서버 디자인 – CPU





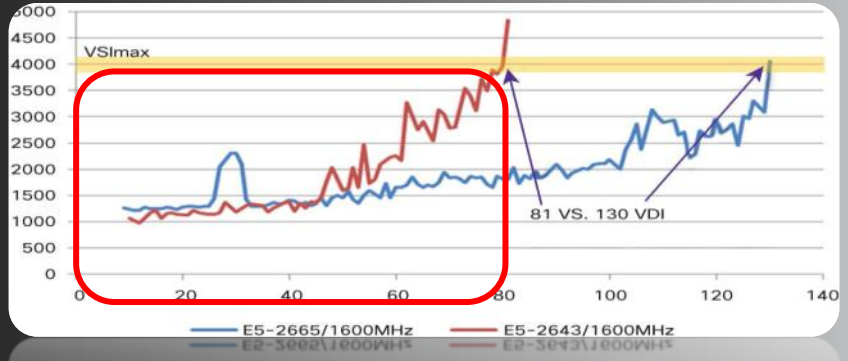
	E5-2643 가상 데스크탑	E5-2665 가상 데스크탑	E5-2665 사용시 집적도 향상 효과
1 vCPU, 1600 Mhz	81	130	**60%**
2 vCPU, 1600 Mhz	54	93	**72%**

VDI 구축에 있어 최우선 사항이 성능 보다는 고집적을 통한 가격 경쟁력 이라면,

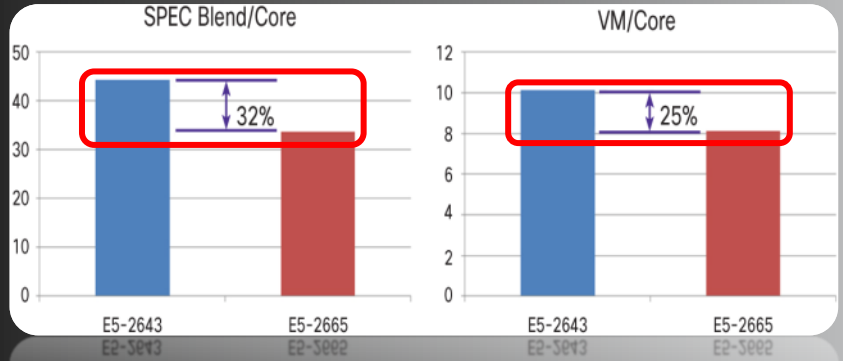
CPU 클럭 스피드 보다 Core 갯수가 높은 사양을 선택 하는 것이 좋습니다.

VDI 서버 디자인 – CPU

Xeon E5-2665 vs E5-2643 with 1 vcpu



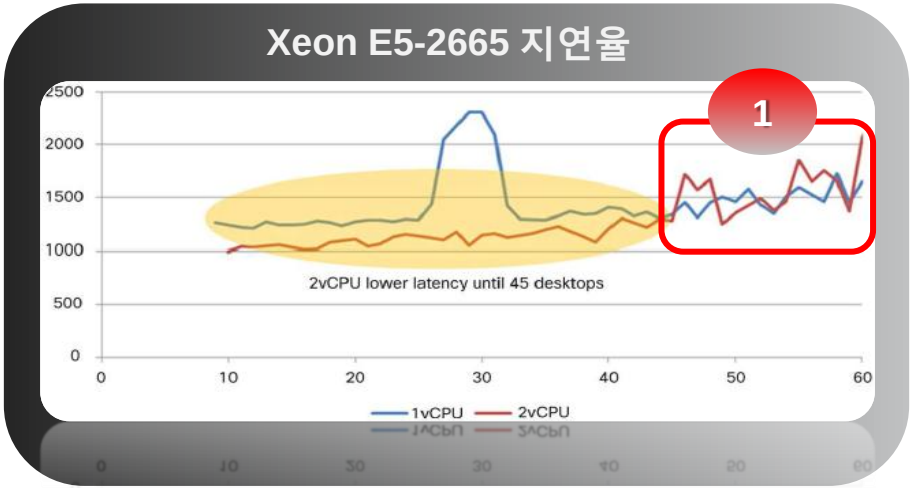
SPEC 성능 자료

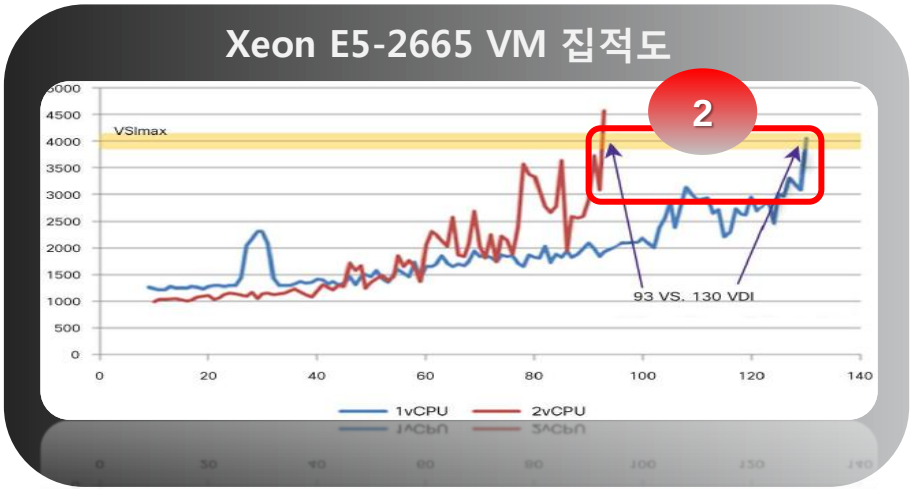


	E5-2643 VM/core	E5-2665 VM/Core	E5-2643 SPEC Blend/core	E5-2665 SPEC Blend/core	E5-2643 이점
1 vCPU, 1600 Mhz	10	8.125	44.38	33.6	성능 32% 집적도 25%

VDI 구축에 있어 최우선 사항이 성능 보다는 집적도 보다는 성능이라면 높은 클럭 스피드와 낮은 코어 집적도를 가진 CPU를 선택 하시는게 좋습니다.

VDI 서버 디자인 – vCPU





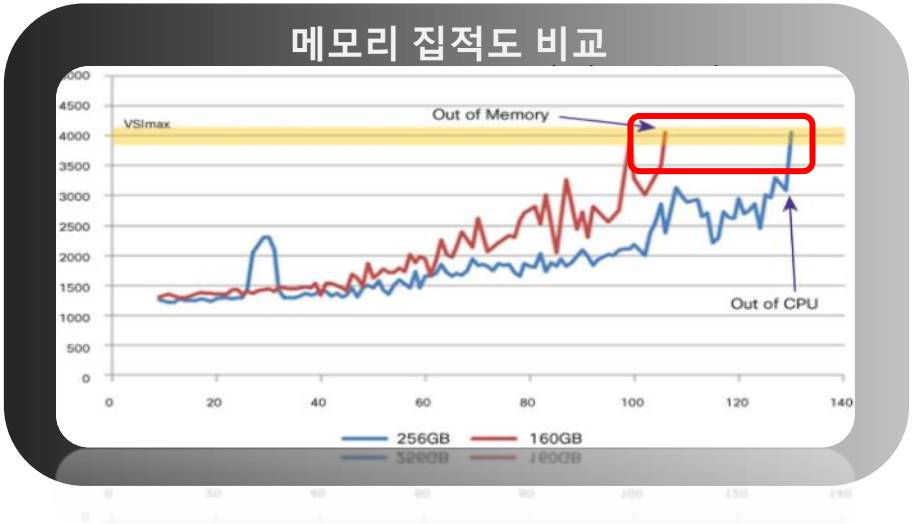
1. **1 vCPU vs 2 vCPU VM 지연율 비교**
   - 측정된 VM 집적 범위 이내에서는 2 vCPU 지연율이 소폭 우수함
   - 일정 범위 초과시 1 vCPU 환경의 지연율이 훨씬 안정적이며 우수함
2. **1 vCPU vs 2 vCPU VM 집적도 비교**
   - LSI Max 수치 기준 최대 93 VM vs 130 VM으로 2 vCPU 환경에서 약 40% 정도의 집적도 감소

VM당 vCPU를 많이 할당한다고 해서 VM의 성능이 기대치 만큼 좋아 진다거나 유저 체감율이 강화 되지는 않으며 오히려 VM의 집적도에만 큰 영향을 줄수 있습니다!!

VDI 컴퓨팅 서버 디자인 – Memory 구성

Total System Memory Size	CPU-1			CPU-2			DIMM Max Oper. Speed (MHz)	Total DIMMs in the system	Relative Perf vs. Peak Bandwidth
	Blue Slots Bank 1 (A1, B1, C1, D1)	Black Slots Bank 2 (A2,B2, C2,D2)	Black Slots Bank3 (A3,B3, C3,D3)	Blue Slots Bank 1 (E1,F1, G1,H1)	Black Slots Bank 2 (E2, F2, G2, H2)	Black Slots Bank 3 (E3, F3, G3,H3)			
32GB	4x4GB	-	-	4x4GB	-	-	N/A	8	N/A
	2x8GB	-	-	2x8GB	-	-	1866	4	0.50
64GB	4x4GB	4x4GB	-	4x4GB	4x4GB	-	N/A	16	N/A
	4x8GB	-	-	4x8GB	-	-	1866	8	1.00
96GB	4x4GB	4x4GB	4x4GB	4x4GB	4x4GB	4x4GB	N/A	24	N/A
	4x8GB	2x8GB	-	4x8GB	2x8GB	-	1866	12	Unbalanced Config*
	3x16GB	-	-	3x16GB	-	-	1866	6	0.50
	4x8GB	4x4GB	-	4x8GB	4x4GB	-	N/A	16	N/A
128GB	4x8GB	4x8GB	-	4x8GB	4x8GB	-	1866	16	0.99
	4x16GB	-	-	4x16GB	-	-	1866	8	0.99
192GB	4x8GB	4x8GB	4x8GB	4x8GB	4x8GB	4x8GB	1333	24	0.74
	4x16GB	2x16GB	-	4x16GB	2x16GB	-	1866	12	Unbalanced Config*
	4x16GB	4x8GB	-	4x16GB	4x8GB	-	1866	16	0.98
256GB	4x16GB	4x16GB	-	4x16GB	4x16GB	-	1866	16	0.98
	4x32GB	-	-	4x32GB	-	-	1866	8	0.95
384GB	4x16GB	4x16GB	4x16GB	4x16GB	4x16GB	4x16GB	1333	24	0.72
512GB	4x32GB	4x32GB	-	4x32GB	4x32GB	-	1866	16	0.90
768GB	4x32GB	4x32GB	4x32GB	4x32GB	4x32GB	4x32GB	1333	24	0.70

VDI 컴퓨팅 서버 디자인 – Memory 구성



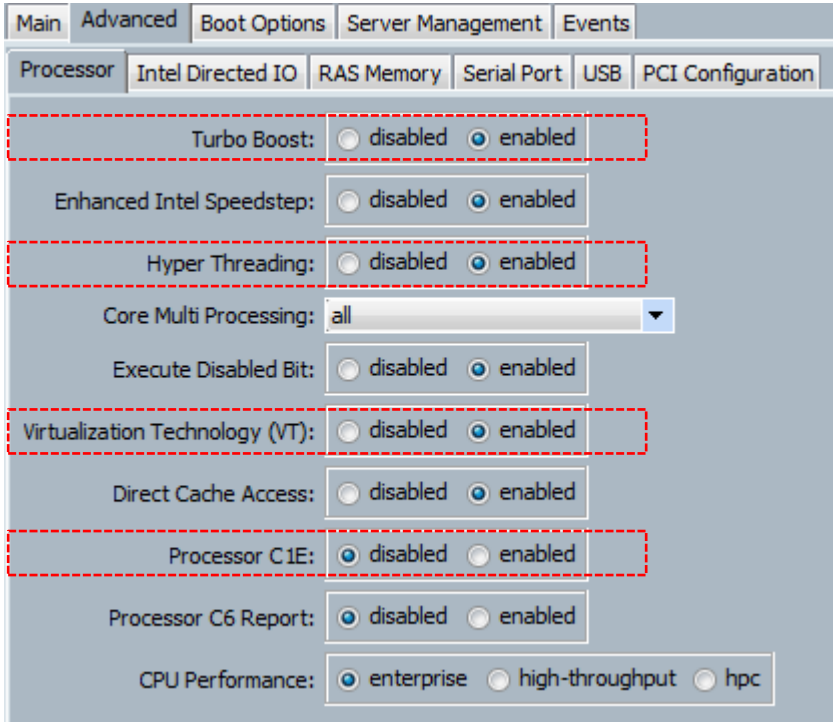
✓ **VM의 성능과 집적도는 메모리 용량에 따라 영향도가 크기 때문에 Memory 오버 커밋이 일어나지 않도록 여유있게 디자인 하시는 걸 권고 드립니다**

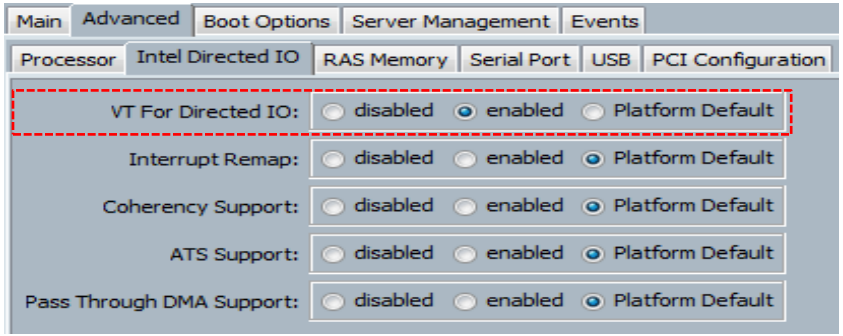
하이퍼 바이저 별 메모리 디자인시 고려사항

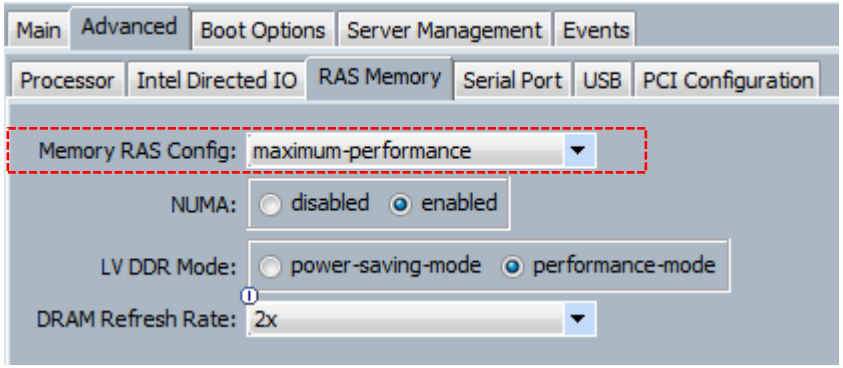
- **Vmware**
  - 메모리 오버헤드
    - 하이퍼 바이저: 약 200MB
    - **vCPU 별: 29MB (1.5GB 가상데스크탑)**
    - **공식: 200MB ESXi + VMs x (1.5G + 29MB)**
- **Microsoft**
  - 메모리 오버헤드
    - 하이퍼 바이저+호스트OS: 약 300MB+512MB
      - 베스트 프랙티스 = 2GB
    - **VM 별: 32MB (1GB 가상데스크탑)**
      - **할당량 증가시 기가비트당 8MB 환산**
    - **공식: 2GB 페어런트OS + VMs x (4GB + 56MB)**
- **Citrix**
  - 메모리 오버헤드
    - 하이퍼 바이저+control Domain: 약 128MB+752MB(Phy 32G +)
    - **VM 별: 32MB (1GB 가상데스크탑)**
      - **할당량 증가시 기가비트당 8MB 환산**
    - **공식: 880MB + VMs x (4GB + 56MB)**

VDI 컴퓨팅 서버 디자인 – BIOS 설정

BIOS 설정 – Vmware, Citrix, Microsoft

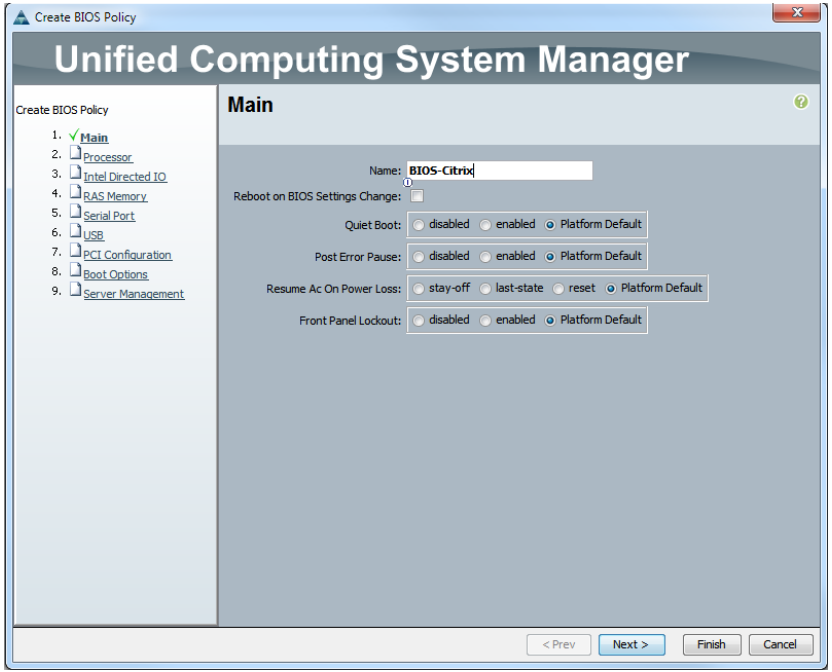




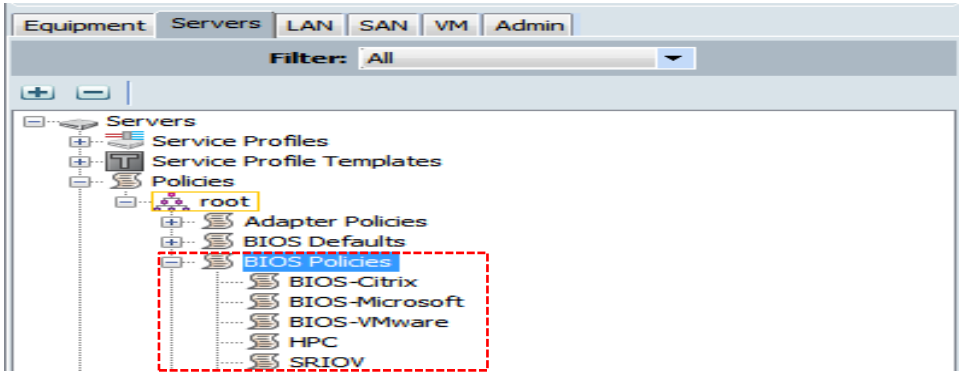


VDI 컴퓨팅 서버 디자인 – BIOS 설정

BIOS 설정

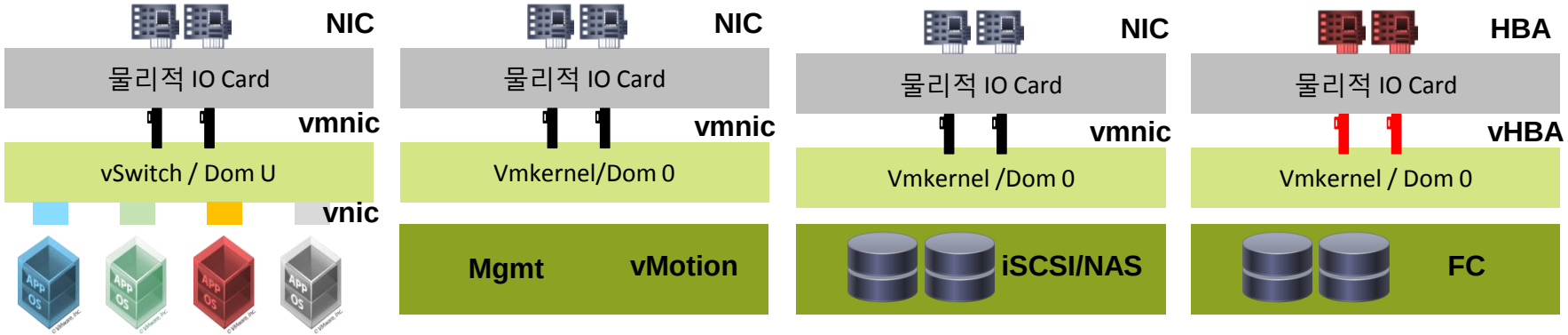


- UCSM 기반의 간편한 BIOS 설정
  - BIOS policy Tab에서 각 각의 하이퍼 바이저 특성에 맞는 BIOS 정책 생성 및 설정
  - 추 후 설치 되는 하이퍼 바이저에 따라 간편하게 정책을 적용하여 최적의 Tuning값 바로 적용 가능



VDI 컴퓨팅 서버 디자인 – IO 가상화

일반적인 x86 서버들의 가상화 구조

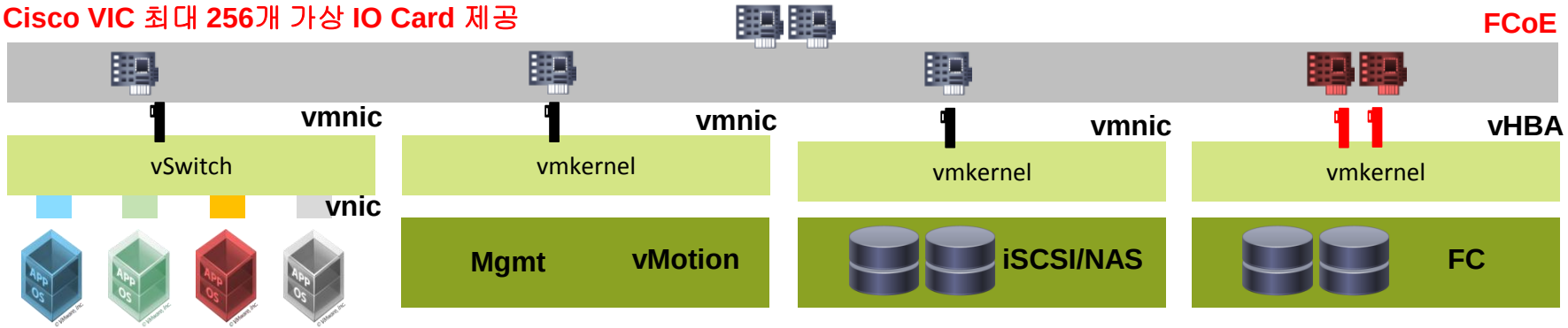


일반 x86 구성의 복잡성

- 8개 이상의 IO Port 필요 – VM Network/Dom U, VMKernel/Dom 0 Network, Storage Network
- 이중화를 위해 상단에서의 네트워크와의 Teaming 고려
- vMotion 이동성 보장을 위해 상단 네트워크와의 802.1Q 적용 및 Vlan Tagging 고려
- 일부 벤더의 가상 IO 기술을 적용하더라도, Rate Limit 기반의 비효율적인 대역폭 정책

VDI 컴퓨팅 서버 디자인 – IO 가상화

IO가사와 카드가 적용된 x86 서버들의 가상화 구조

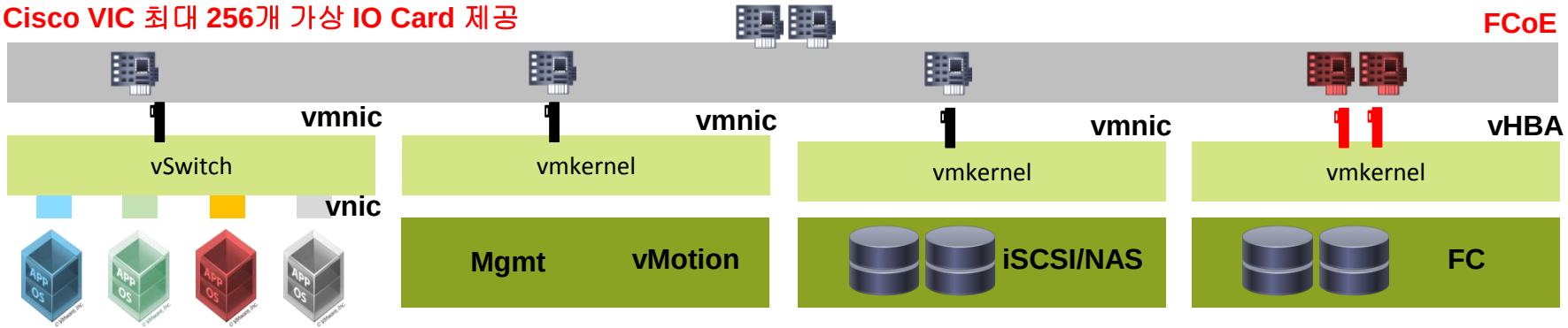


Cisco Adapter FEX 기반의 장점

- 최소 2개의 물리적 IO에서 최대 256개 가상 IO 제공 (Ethernet, iSCSI, HBA 제한 없음)
- H.W Teaming을 제공함으로 vSwitch의 이중화 구성이 필요 없음.
- 802.1Q 및 Vlan Tagging은 모두 UCS에서 적용 가능함.
- 가상화 IO 기술 적용시 Rate Limit , shapping 기술 동시 제공하여 대역폭을 효율적 사용
- CoS Tagging이 구현 가능하므로, SLA 구현에 유리

VDI 컴퓨팅 서버 디자인 – IO 가상화

IO가사와 카드가 적용된 x86 서버들의 가상화 구조

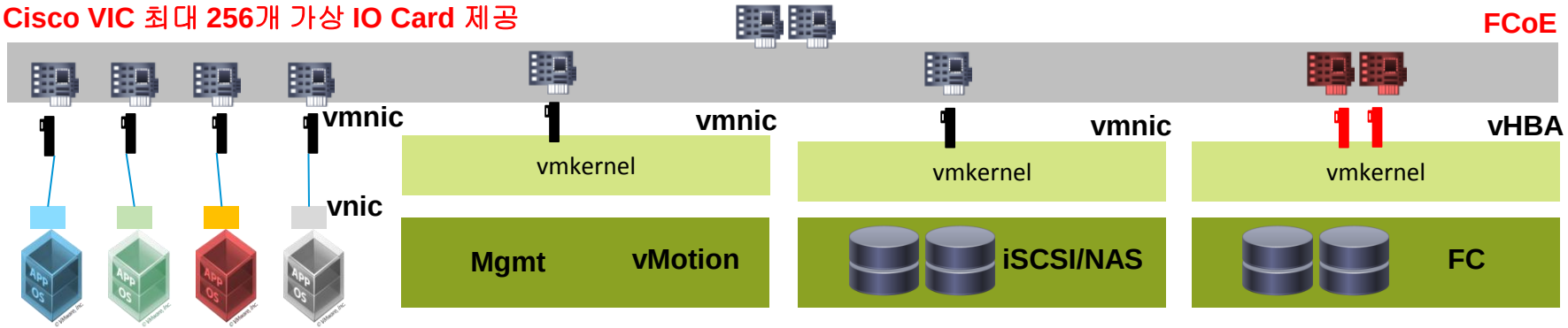


Cisco Adapter FEX 기반의 장점

- 최소 2개의 물리적 IO에서 최대 256개 가상 IO 제공 (Ethernet, iSCSI, HBA 제한 없음)
- H.W Teaming을 제공함으로 vSwitch의 이중화 구성이 필요 없음.
- 802.1Q 및 Vlan Tagging은 모두 UCS에서 적용 가능함.
- 가상화 IO 기술 적용시 Rate Limit , shapping 기술 동시 제공하여 대역폭을 효율적 사용
- CoS Tagging이 구현 가능하므로, SLA 구현에 유리

VDI 컴퓨팅 서버 디자인 – IO 가상화

IO가사와 카드가 적용된 x86 서버들의 가상화 구조

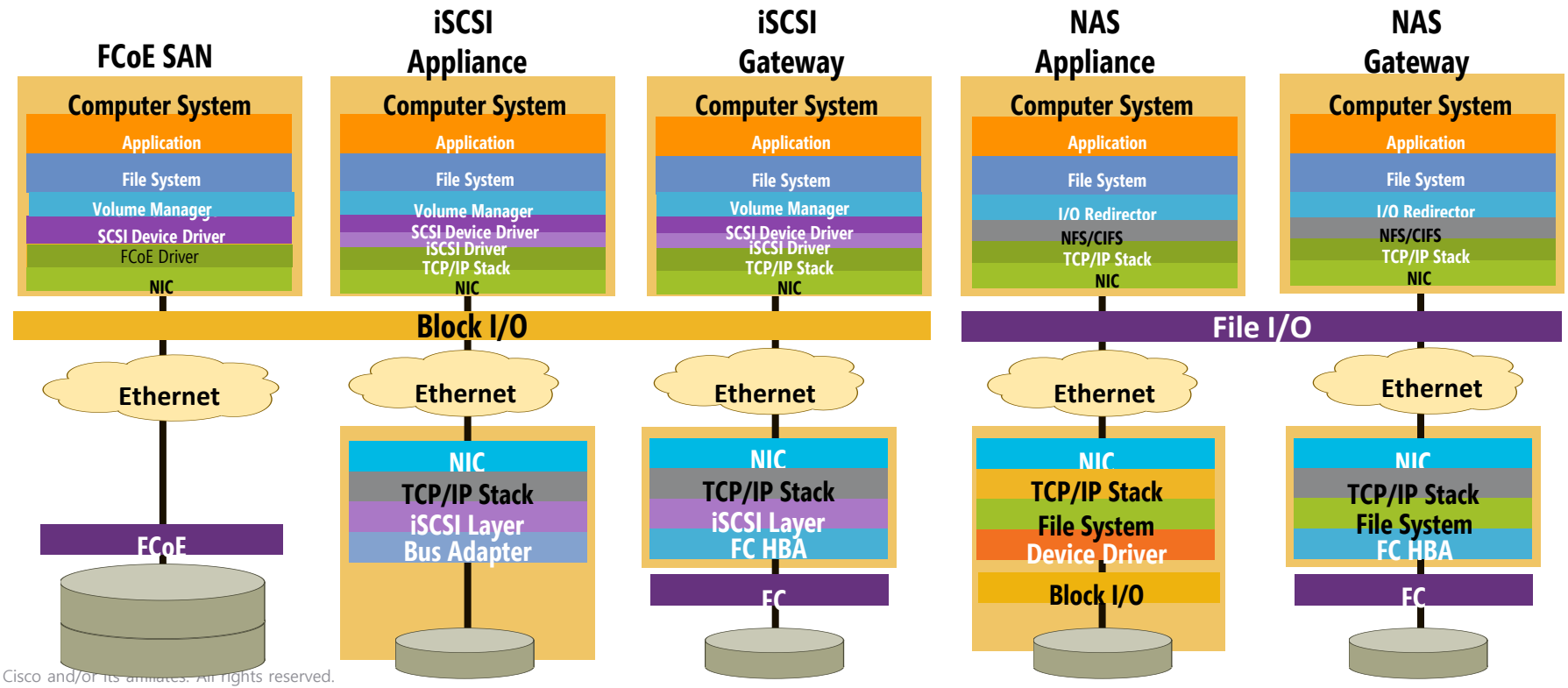


Cisco VM FEX 기반의 장점

- vSwitch를 사용하지 않는 방식으로, Hypervisor를 Bypass 함.
- 성능 30%, IO Latency 40% 이상 , CPU 20% 이상 향상 효과
- vSwitch를 사용하지 않으면서, vMotion이 가능한 UTP/VM DirectPath v2 업계 유일 구현
- VM 집적도를 타사 대비 극대화 가능함.
- 서버 가상화 또는 대규모 가상화에 매우 유리

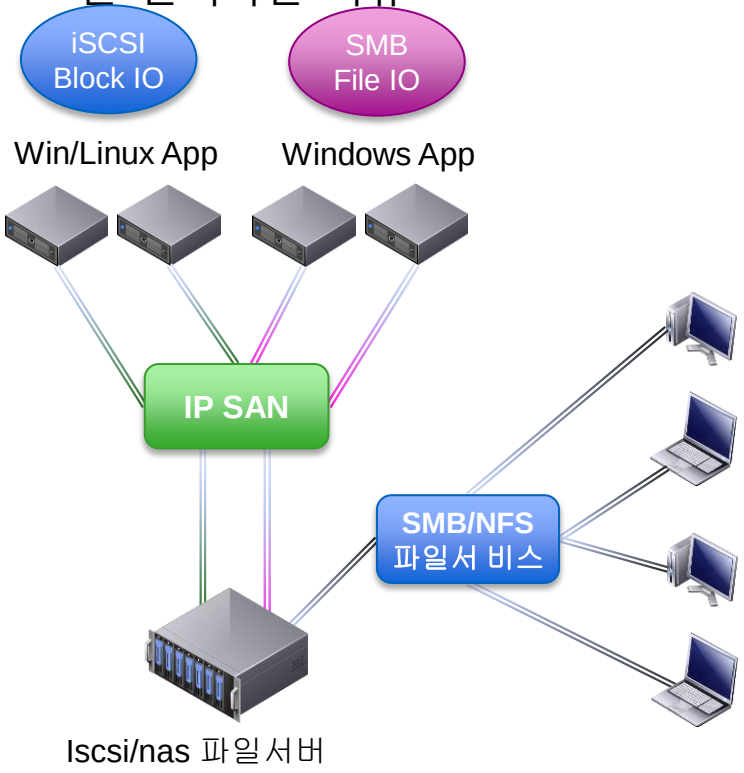
VDI 컴퓨팅 스토리지 디자인 – 스토리지 IO 선택

FC vs iSCSI vs NAS



VDI 컴퓨팅 스토리지 디자인 – 스토리지 IO 선택

iSCSI를 선택하는 이유

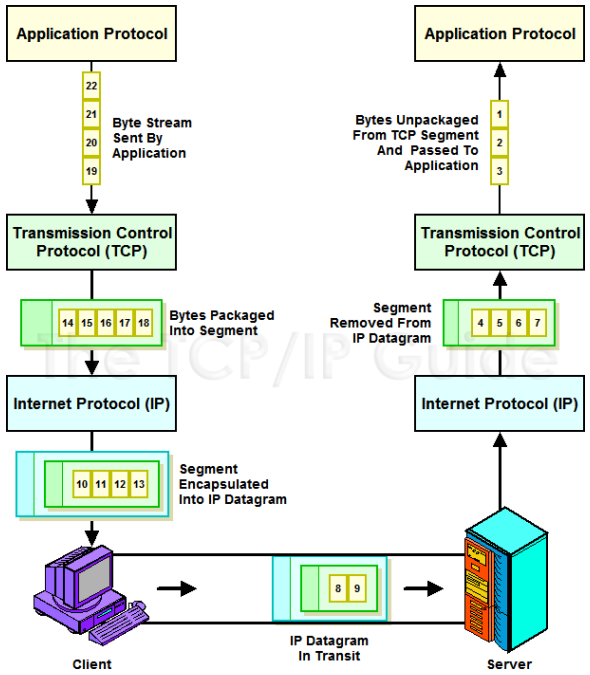


- 별도의 SAN 네트워크 불필요
- 상대적으로 빠르고 쉬운 구성
- 블록/파일 Level 스토리지의 IO 통합
- 기존 네트워크 장비활용 가능
- 이더넷 기술의 발전- 10G / 40G

1G 10G [10G Base-T] 40G

VDI 컴퓨팅 스토리지 디자인 – 스토리지 IO 선택

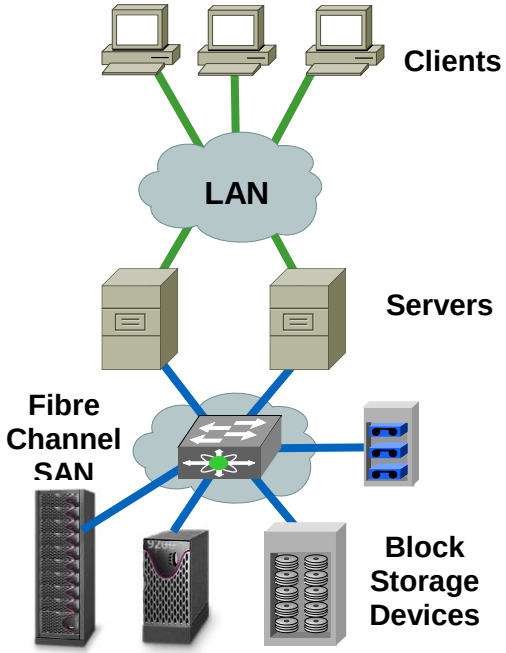
iSCSI를 선택을 주저하는 이유



- TCP/IP 의 Encap/Decap 오버헤드
- TCP/IP 는 Lossless가 아님
- iSCSI 트래픽의 보안 취약성
- FC에 대한 종교적인 믿음

VDI 컴퓨팅 스토리지 디자인 – 스토리지 IO 선택

FC를 선택 하는 이유

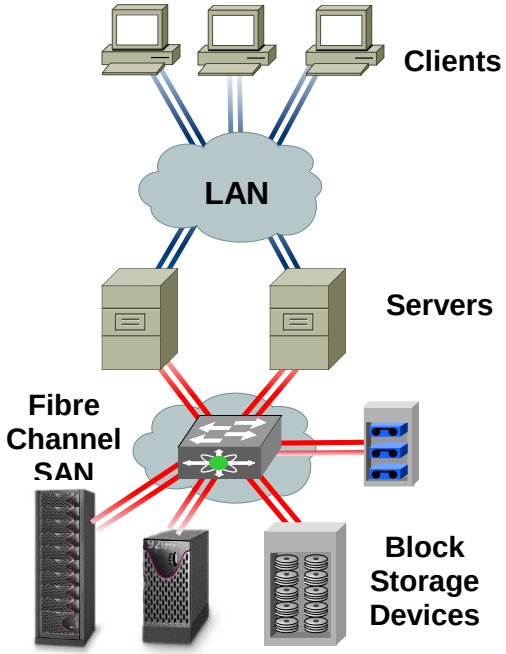


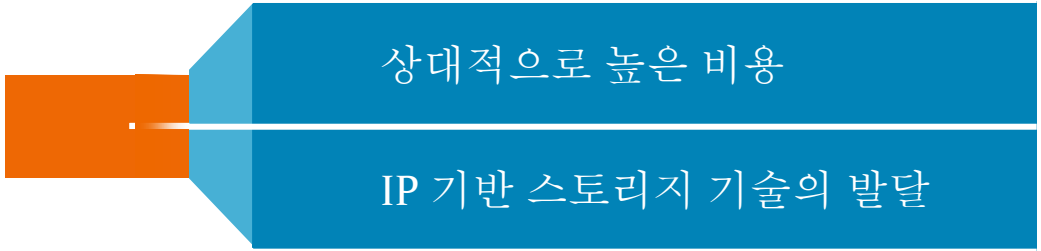
- 검증된 성능 및 안정성
- 스토리지 제조사의 권고
- 기존인프라와의 호환
- 높은 보안성

1G 2G 4G 8G 16G...

VDI 컴퓨팅 스토리지 디자인 – 스토리지 IO 선택

FC를 선택 하지 않는 이유

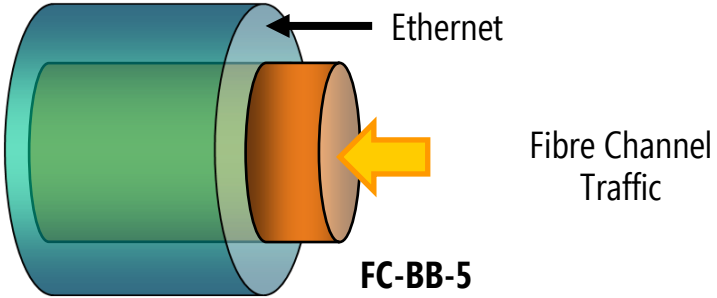


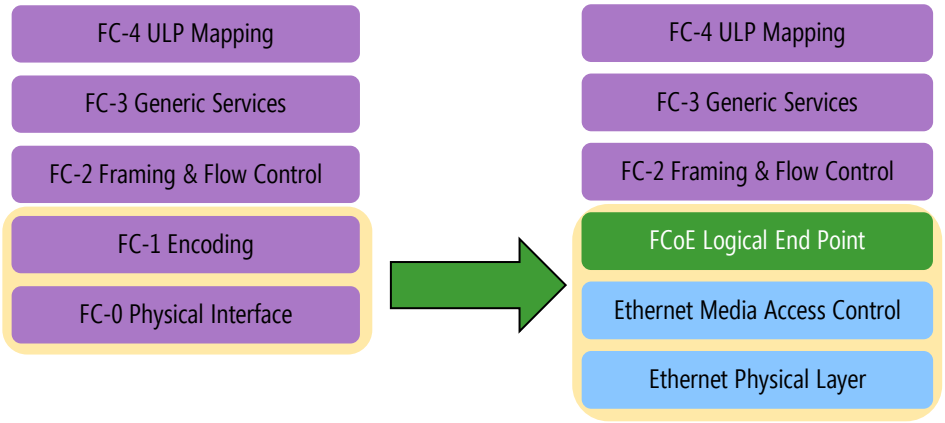


VDI 컴퓨팅 스토리지 디자인 – 스토리지 IO 선택

FCOE 환경은 어떨까요?

- FC 관점 : Ethernet 이라고 불리는 다른 형태의 케이블링을 사용하는 Fiber Channel
- Ethernet 관점 : 전송해야할 또 하나의 상위 프로토콜





VDI 컴퓨팅 스토리지 디자인 – 스토리지 IO 선택

iSCSI

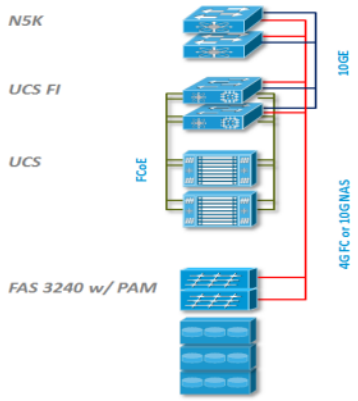
- 별도의 SAN 네트워크 불필요[저비용]
- 상대적으로 빠르고 쉬운 구성
- 블록/파일 Level 스토리지의 IO 통합
- 기존 네트워크 장비활용 가능
- 이더넷 기술의 발전- 10G / 40G

FC

- 높은 성능 및 안정성
- 스토리지 제조사의 권고
- 기존 SAN 인프라와의 호환
- 높은 보안성

- 이더넷과 FC 통합 [Unified Fabric]
- 폭넓은 FCoE 스토리지 제품 출시

VDI 컴퓨팅 스토리지 디자인 – Jumbo Frame



To FI : e 1/20
To filer : fc 2/3,4
To FI : fc 2/1,2
To N5K : e 1/20
To N5K : fc 2/1,2

Network 구성

- UCS FI to Nexus5K Switch 10GE 구성

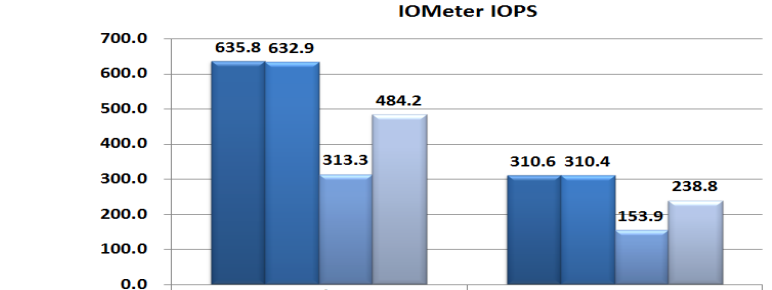
Server 구성

- UCS B200M2 & UCS B250M2 구성
- CPU 5670 2P & Memory 72GB , SAS Disk
- Xen Server 5.6 FP1 & Xen Desktop 5.0
- UCS B200M2 4set : Legacy FC or FC w/ PAM Config
- UCS B200M2 2set & UCS B250M2 : ILIO Config

Storage 구성

- Dual 4G FC per Filer Front , Dual 4G FC per Filer Backend 또는 NAS 구성

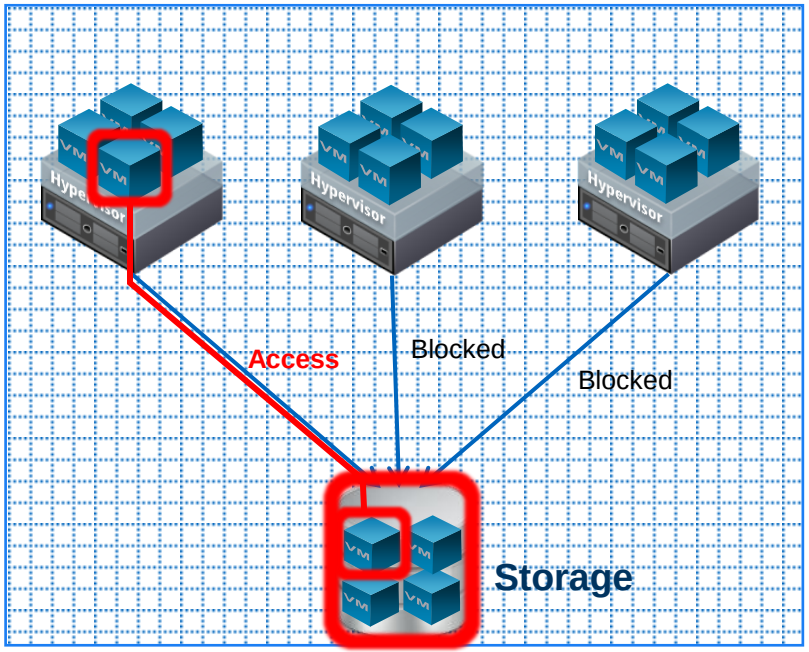
	Test Tool	테스트 방법
V3	V3 Lite 정밀 검사	V3 Lite 정밀 검사 실행 후 최종 시간 3회 평균
MS-Word	Stop Watch 실행	Open/Edit/Save as/Close 최종 시간 3회 평균
VM IO 측정	IOMeter	IOPS/전송속도/IO응답시간 최종값 3회 평균



	Read IO	Write IO
VM50	635.8	310.6
VM50 jumbo	632.9	310.4
VM100	313.3	153.9
VM100 Jumbo	484.2	238.8

- iSCSI, NAS 스토로지 운영 환경
  - Jumbo Frame 반드시 사용 권고
    - IO가 많아 질수록 월등한 성능 차이 발생
    - 대역폭이 커질 수록 성능 향상 효과 탁월
    - Jumbo Frame 설정 포인트
      - Hypervisor – ESX, Xen, Hyper-V, KVM
      - 네트웍 스위치, 스토리지

VDI 컴퓨팅 스토리지 디자인 – Sizing DataStore

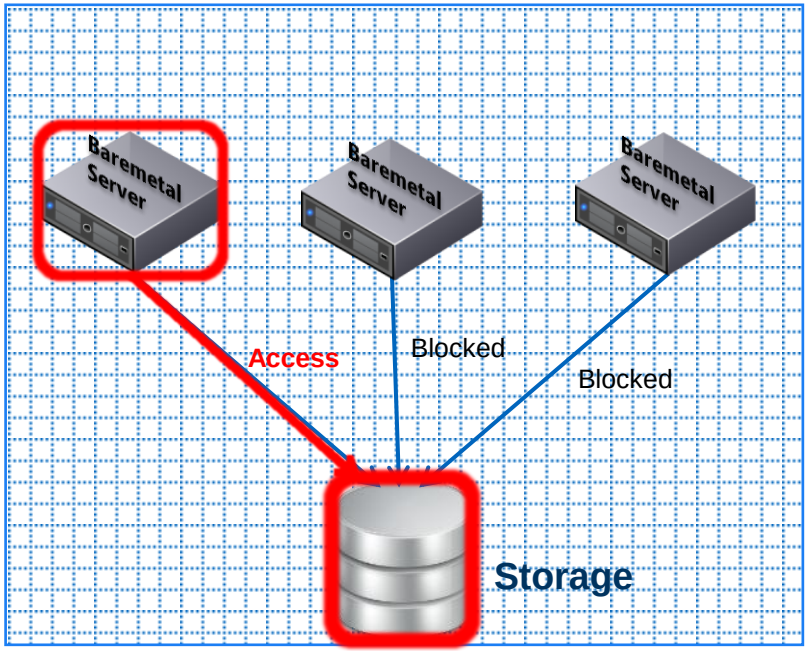


SCSI Reservation 동작 프로세스

SCSI Reservation

- **SCSI LUN Reservation**
  - iSCSI나 FC 기반의 스토리지 LUN에 대한 파일 엑세스 메카니즘
  - 해당 LUN에 write을 하기 위해 호스트 단위로 File locking이나 Meta Data Locking 을 통해서 SCSI LUN Reservation이 일어나서 File에 대한 정합성 보장
  - SCSI LUN Reservation이 임계치를 초과 할 경우 성능 저하 현상이 발생할 수 있음

VDI 컴퓨팅 스토리지 디자인 – Sizing Datastore

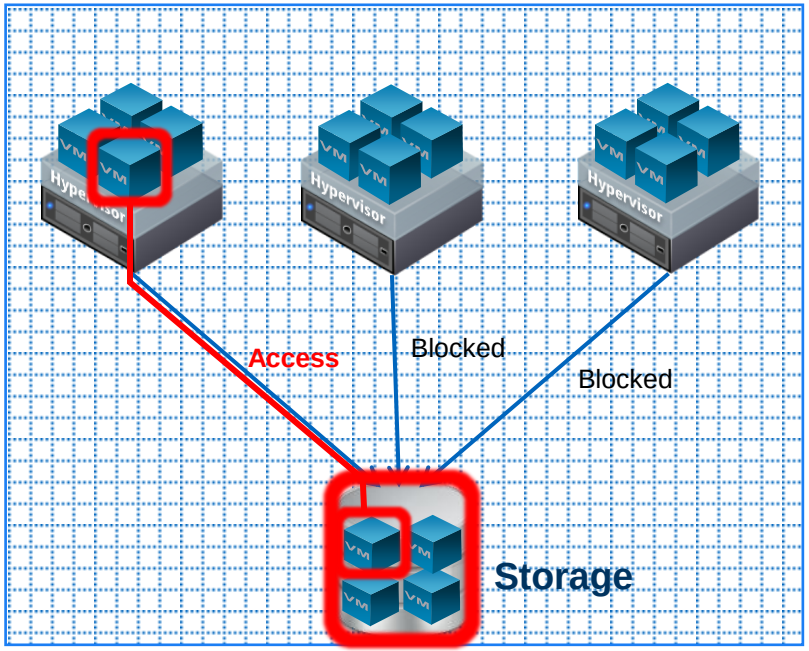


SCSI Reservation 동작 프로세스

SCSI Reservation

- **Baremetal Server SCSI LUN Reservation**
  - Baremetal Server의 경우에는 LUN에 엑세스 하는 서버 댓수가 제한 적이며 호스트당 한 개의 트랜잭션만 발생하므로 SCSI Reservation에 대한 오버헤드가 거의 발생하지 않음

VDI 컴퓨팅 스토리지 디자인 – Sizing Datastore

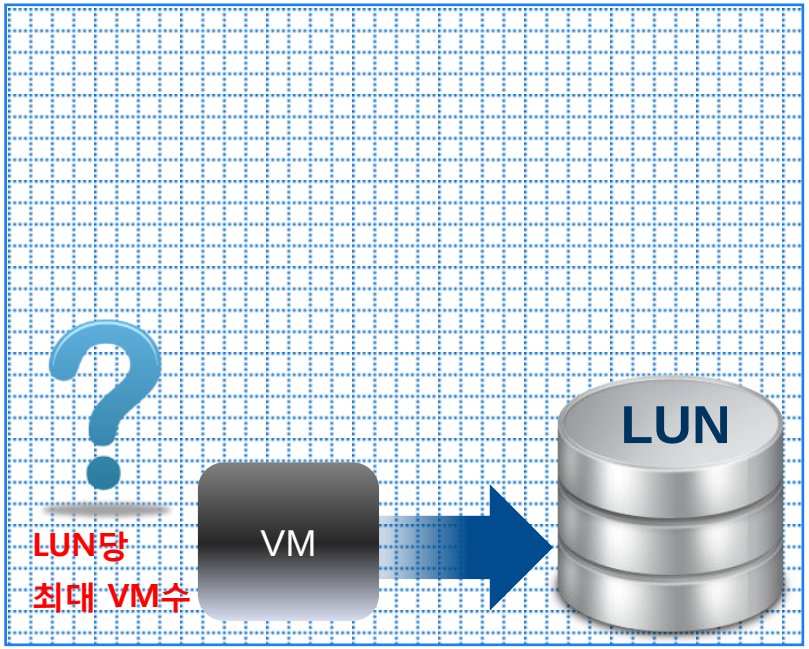


SCSI Reservation 동작 프로세스

SCSI Reservation

- **VDI SCSI LUN Reservation**
  - Hypervisor 기반의 가상화가 적용 된 경우 그 상위에서 기동 중인 각 각의 VM에 대해서 호스트 레벨에서 지속적으로 SCSI Reservation 발생하게 됨
  - 한 개의 LUN에 다수의 VM이 설치 되어 있는 디자인 환경에서 SCSI LUN reservation 오버헤드가 임계치를 최과 할 경우 심각한 성능 저하 현상이 발생 할 수 있음
  - 대상 하이퍼 바이저
    - 모든 가상화가 적용된 하이퍼 바이저
      - EX) Vsphere, Hyper-V, Xenserver, KVM etc...

VDI 컴퓨팅 스토리지 디자인 – Sizing Datastore



최적의 디스크 사이징

SCSI Reservation

- **최적의 VDI LUN 사이징**
  - ❖ FC, iSCSI 프로토콜을 사용 하는 스토리지
    - ▪ LUN당 VM 집적도는 하이퍼 바이저와 스토리지의 Offload 기능을 고려하여 디자인 필요
      - ▪ Vmware – VAAI(vStorage API for Array integration
      - ▪ Microsoft – ODX(Offloaded Data Transfer)
    - ▪ Vmware의 경우 Offload 기능이 적용 되지 않은 경우 LUN당 60개 정도의 VM 집적도를 권고하며 가상화 소프트웨어 별로 조금씩은 상이 할 수 있음

    FC, iSCSI 프로토콜을 사용 하는 스토리지
  - ❖ 디자인 권고
    - ▪ LUN의 갯수와 LUN 증가에 따른 관리 포인트의 부하를 고려 하되 LUN 당 VM의 집적도가 낮을 수록 디스크 엑세스 성능은 우수 해짐을 고려

VDI 컴퓨팅 스토리지 디자인 – HDD vs SSD

SATA, SAS, SSD 그리고 RAID

SATA 7.2K

75~100 IOPS

SATA 10K

125~150 IOPS

SATA 10K

140 IOPS

SAS 15K

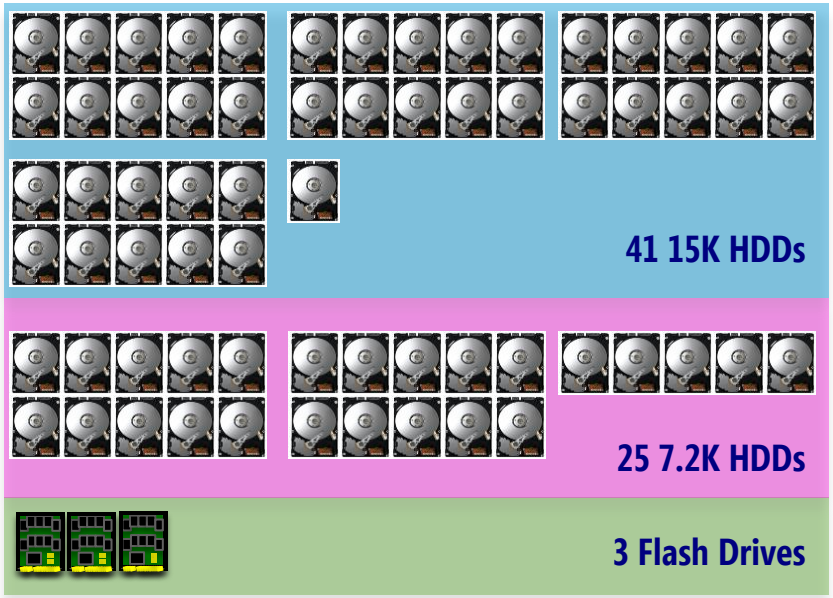
175~210 IOPS

SSD Disk

35000~ IOPS

IOPS = 1 / (평균 Latency + 평균 Seek Time)

VDI 컴퓨팅 스토리지 디자인 – HDD vs SSD



1000 유저 VDI 레퍼런스 아키텍쳐 사이징

1000 User VDI 구현 사례

- **1000 유저 VDI 레퍼런스 아키텍쳐 사이징**
  - 사용된 디스크 타입
    - SATA 15K HDD, SATA 7.2K HDD, SSD
  - 디스크 성능
    - Total IOPS : 114,950
  - 용량
    - Total Capacity : 63.2 TB

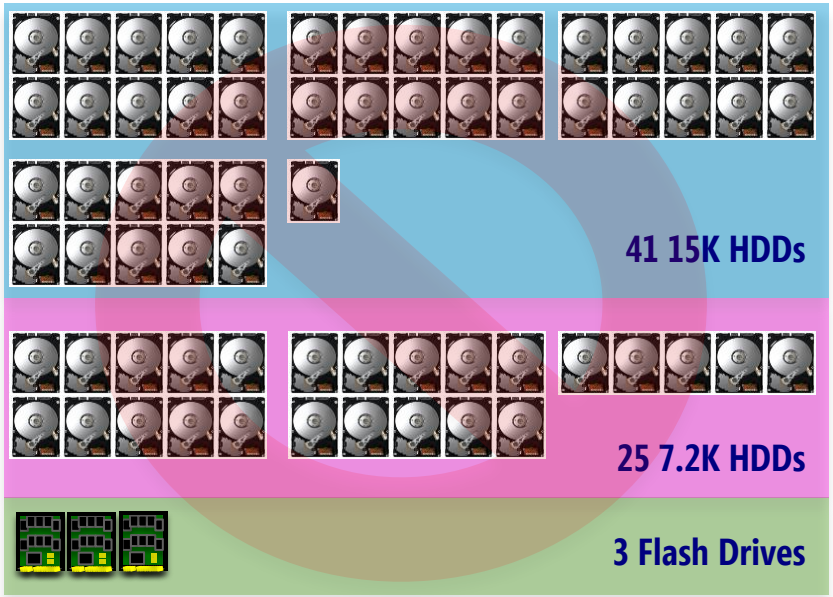
VDI 컴퓨팅 스토리지 디자인 – HDD vs SSD

1000 유저 VDI 레퍼런스 아키텍쳐 사이징

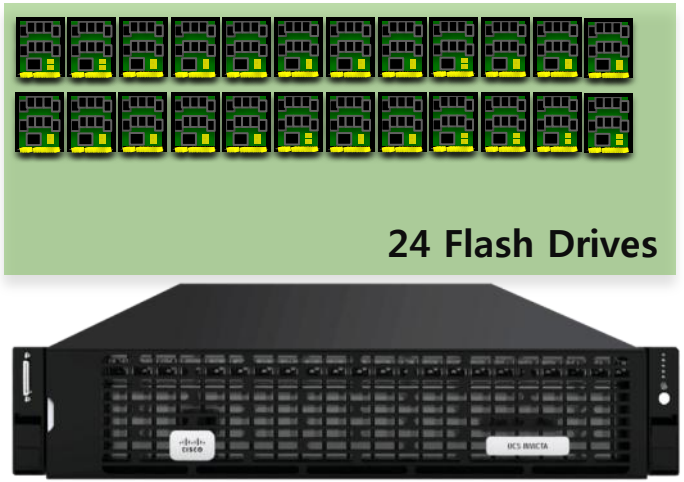
1000 User VDI 구현 사례 - SSD

- **1000 유저 VDI 레퍼런스 아키텍쳐 사이징**
  - 사용된 디스크 타입
    - 24개 SSD
  - 디스크 성능
    - Total IOPS : 180,000
  - 용량
    - Total Capacity : 64 TB (실시간 중복 제거 적용)

VDI 컴퓨팅 스토리지 디자인 – HDD vs SSD



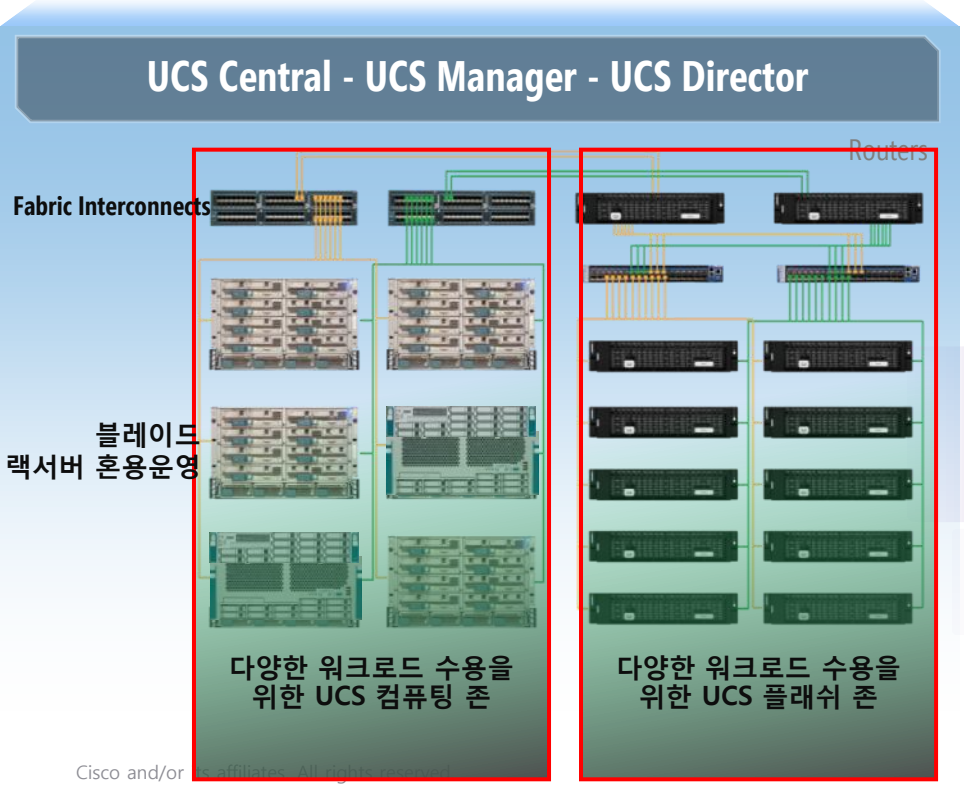
VS



성능: 114,950 IOPS
용량: 63.2 TB

성능: 200,000 IOPS
용량: 64 TB

VDI 컴퓨팅 스토리지 디자인 – HDD vs SSD



SSD Storage 응용 아키텍쳐

- **SSD 기반의 스토리지를 활용하여 두 개의 존으로 구성**
  - Performance Node
    - 고성능 I/O처리를 필요로 하는 실시간 분석, DB, 배치작업
  - Data Reduction Nodes
    - 성능 대비 최적화된 용량을 필요로 하는 이메일 서비스, 데스크톱 가상화

Performance Nodes	Batch, Data Loads OLTP Analytics
Data Reduction Nodes	Virtual Desktops Email

Cisco Connect
Seoul, Korea
April 29-30, 2014

성공적인 VDI구축을 위한 실전 가이드

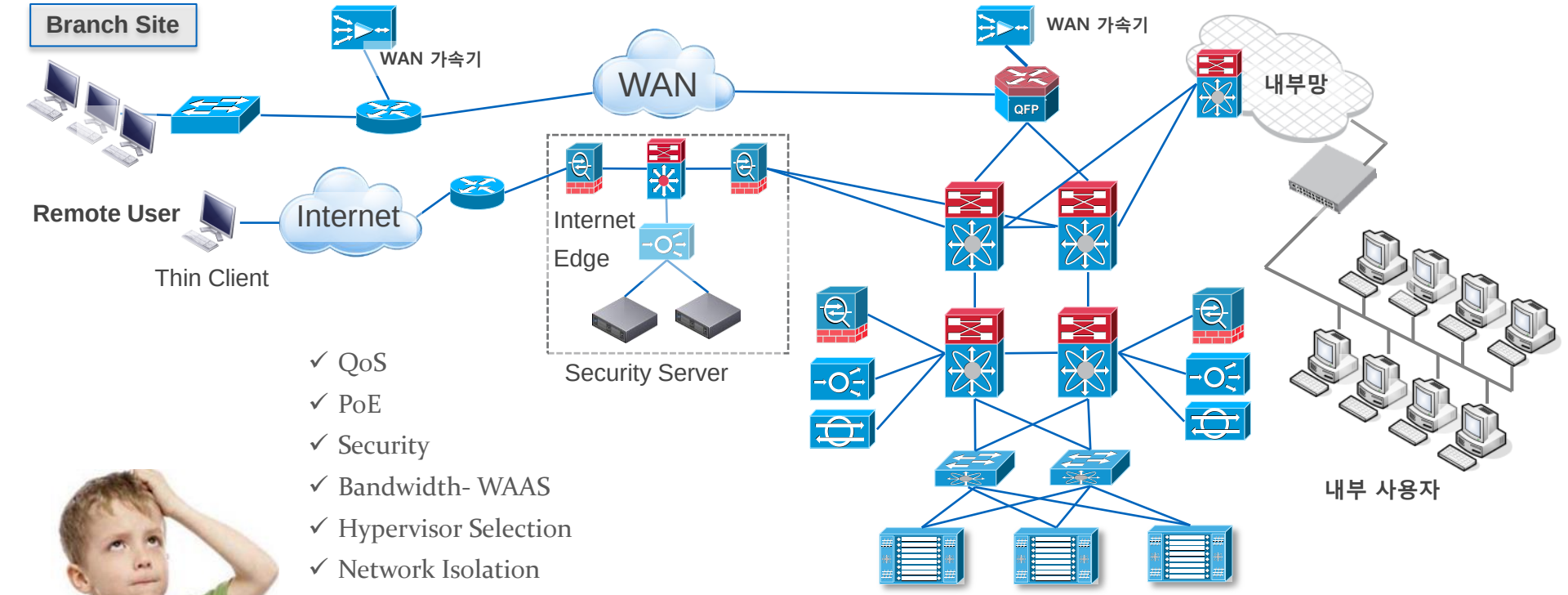
가상네트워킹 고려사항

김용우 과장

yongwkim@cisco.com

Cisco Systems, Datacenter PSE

VDI/가상 네트워킹 환경의 고려사항



- ✓ QoS
- ✓ PoE
- ✓ Security
- ✓ Bandwidth- WAAS
- ✓ Hypervisor Selection
- ✓ Network Isolation
- ✓ Storage connection
- ✓ IP Addressing scheme
- ✓ .....

Agenda

- Hypervisor 네트워킹 기초
- Hypervisor 에따른 네트워킹 구성

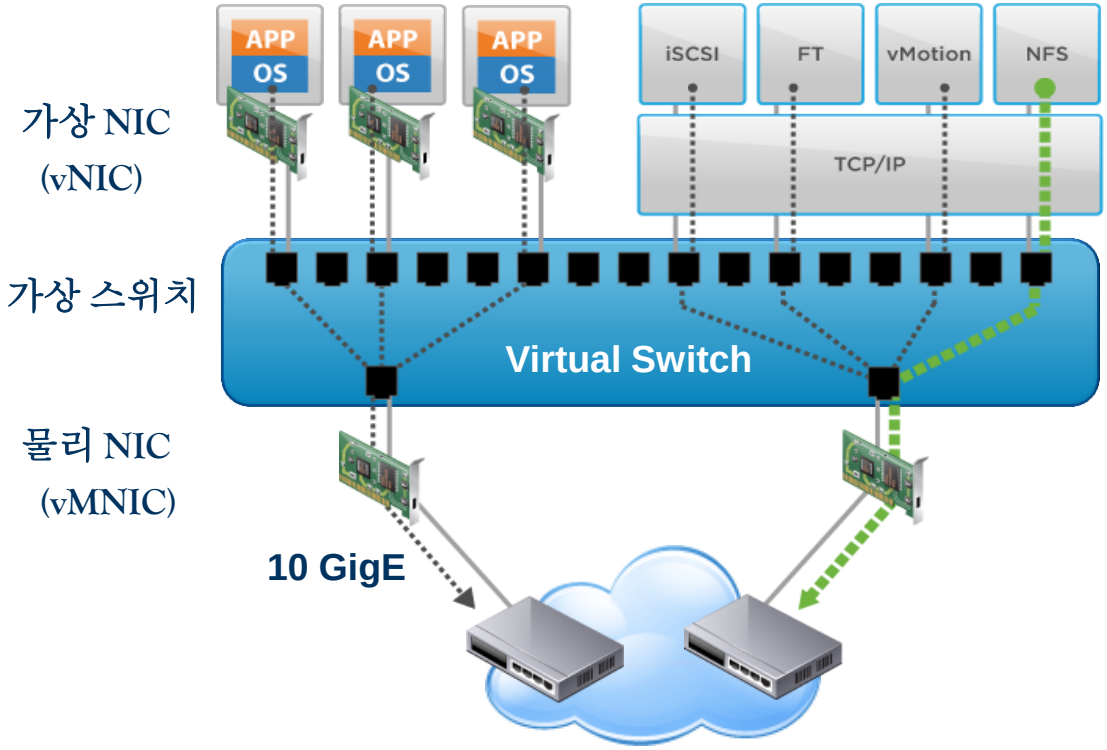
[가상 호스트를 시스코 스위치와 어떻게 연결할 것인가?]

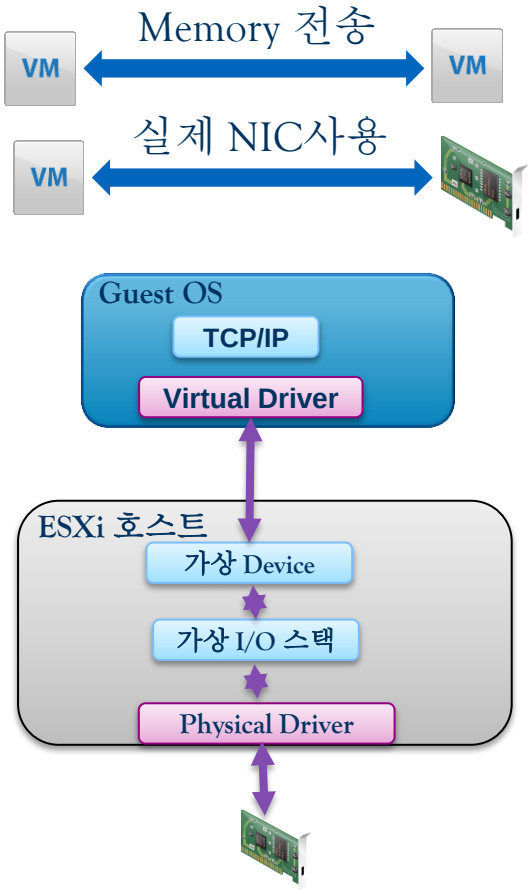
  - VMware vSphere ESXi 5.X
  - Microsoft Windows Server 2012/R2 (Hyper-V 3.0)
  - Citrix XenServer 6.x
  - Hypervisor 네트워킹 정리
- Nexus 1000v 네트워킹 구성
- UCS 네트워킹 구성 Tips
- 가상 환경에서의 보안
- 마무리

하이퍼 바이저 네트워킹 기초

Cisco Connect
Seoul, Korea
April 29-30, 2014

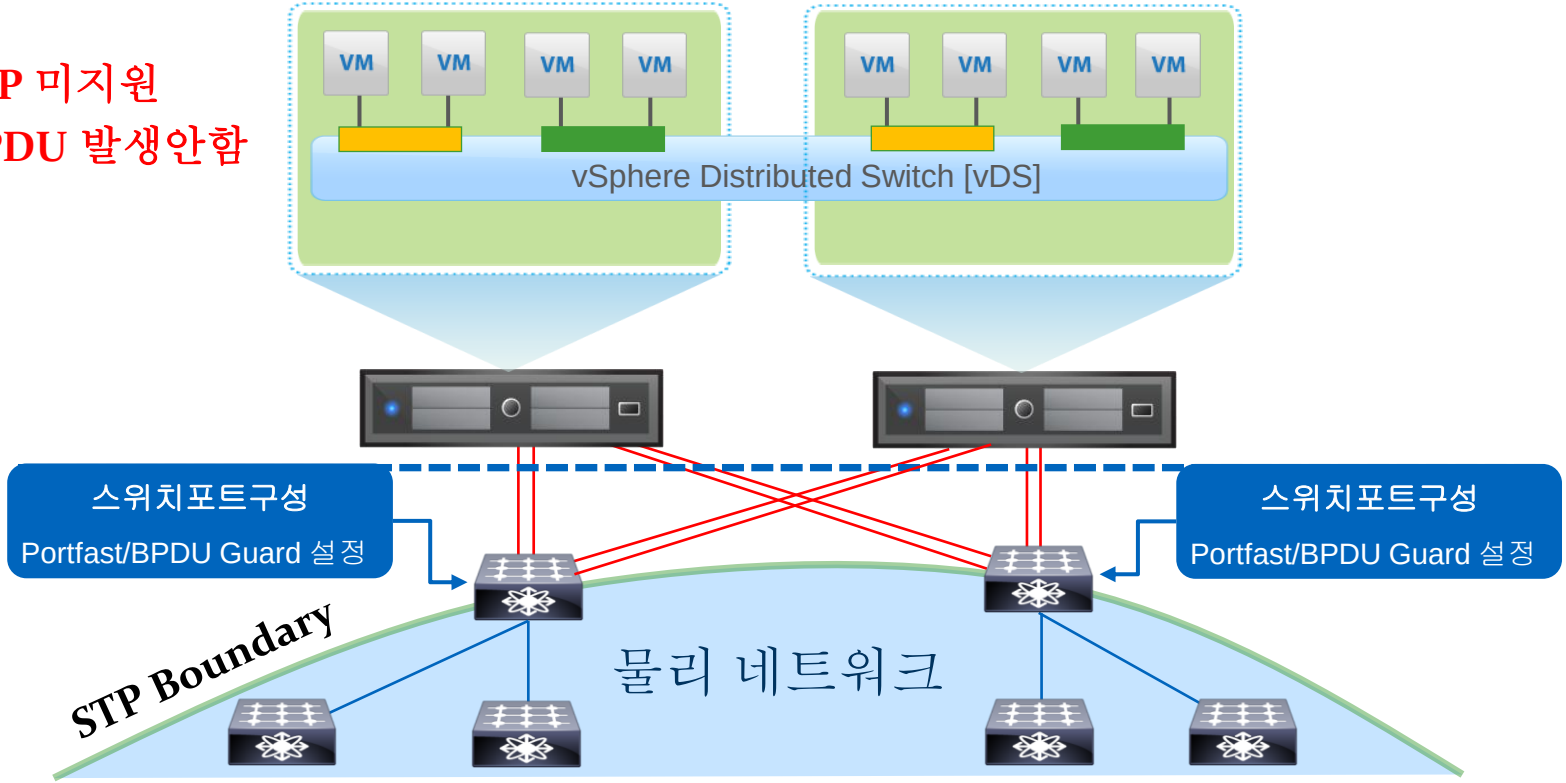
Hypervisor 네트워킹



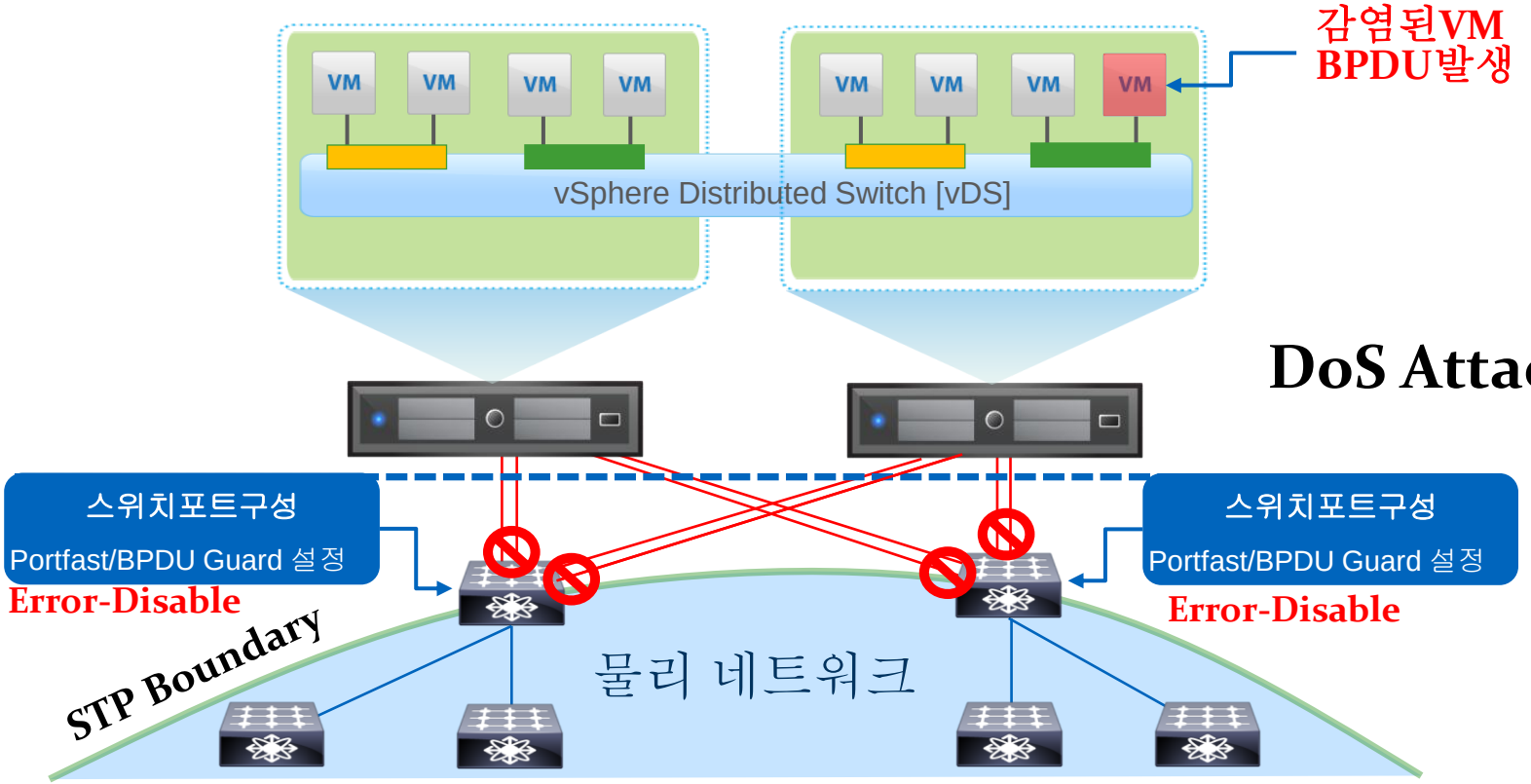


가상 스위치와 Spanning Tree

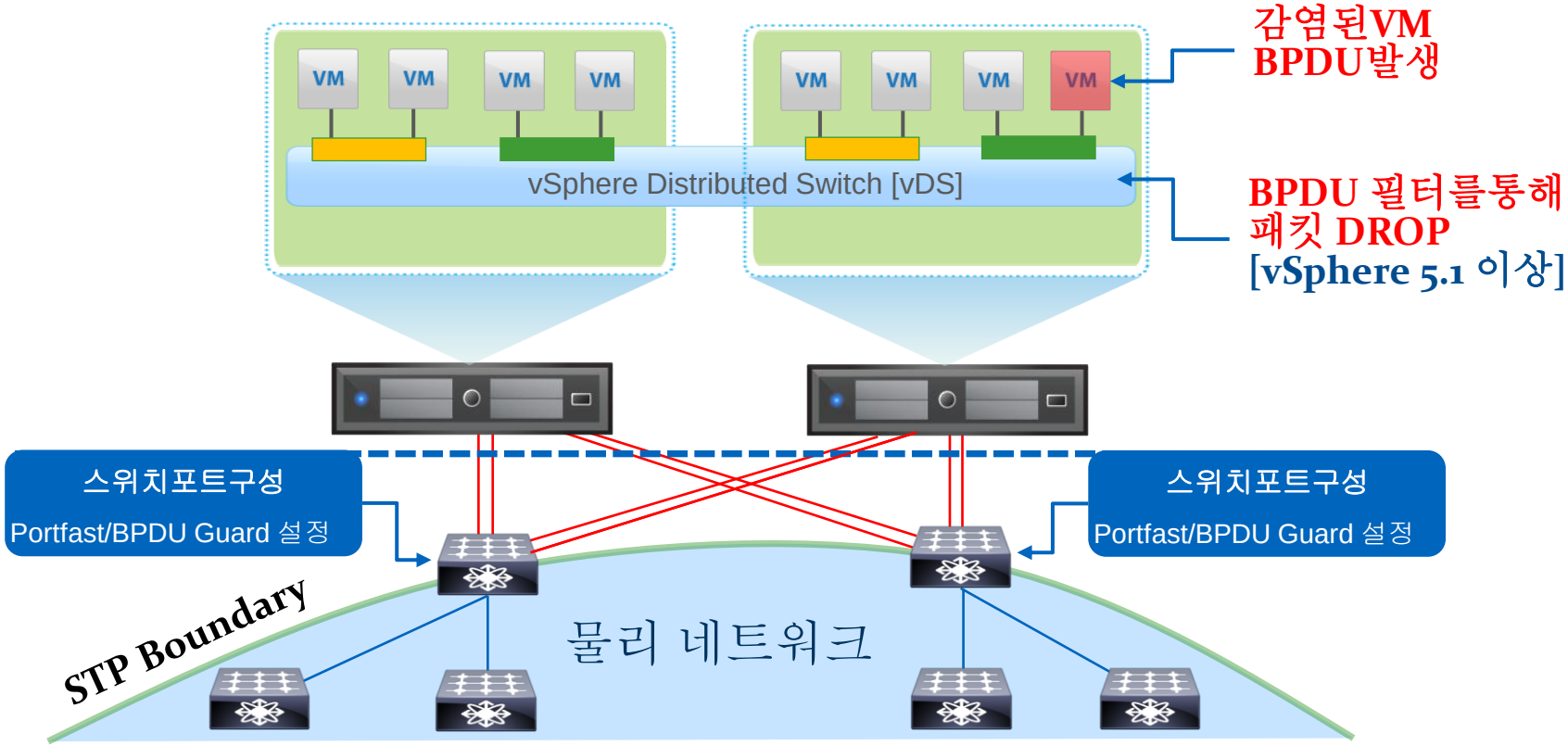
STP 미지원
BPDU 발생안함



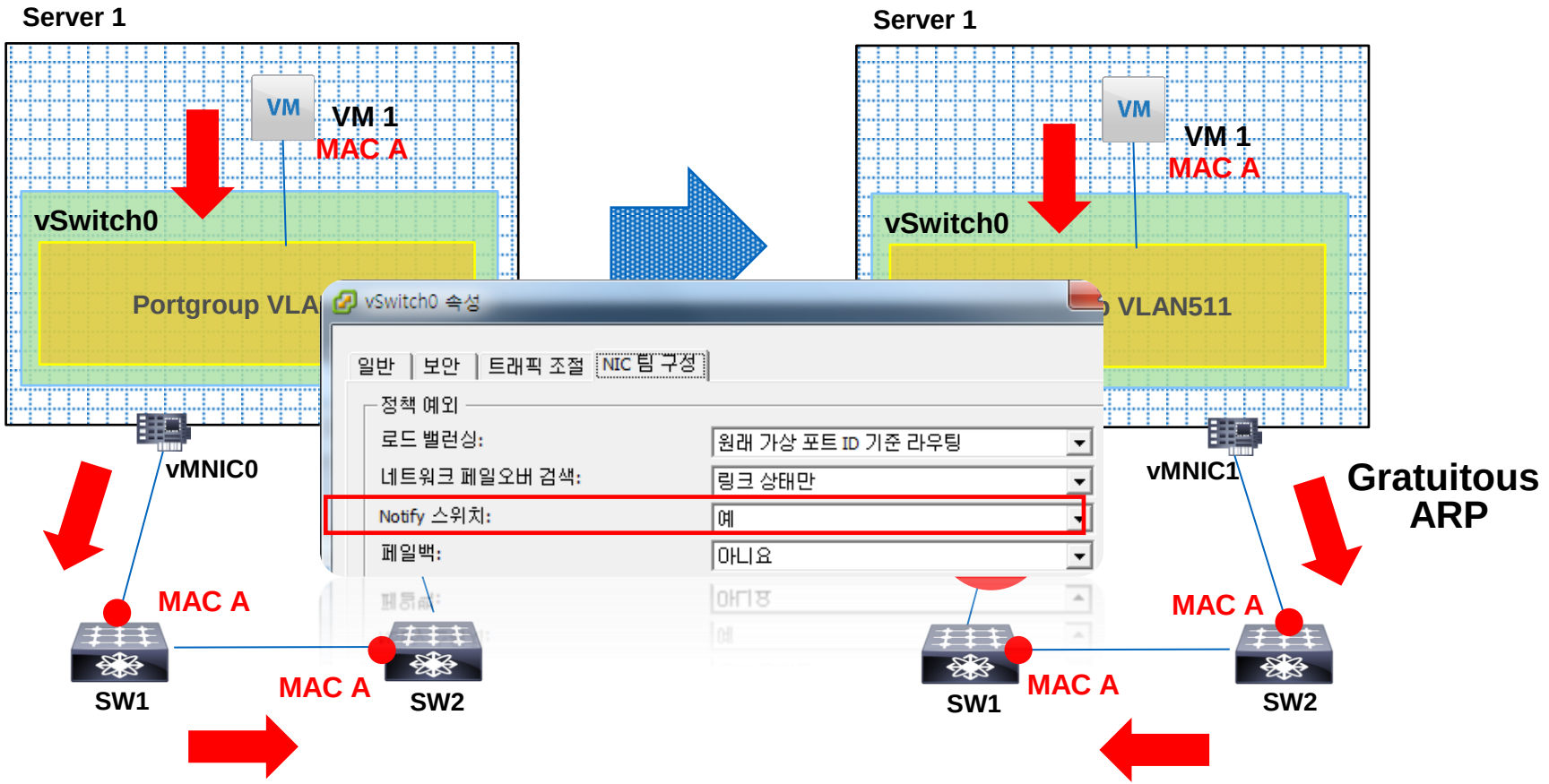
가상 스위치와 Spanning Tree



가상 스위치와 Spanning Tree – 상황별 정리

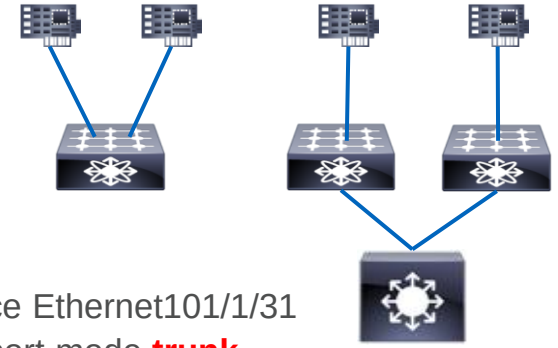


Gratuitous[그레튜이러스] ARP / Reverse ARP



Uplink 스위치 구성옵션 NX-OS

Switch Independent



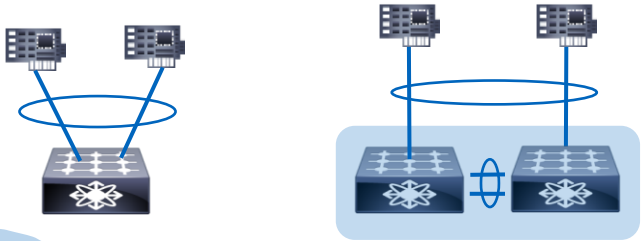
Trunk

interface Ethernet101/1/31
switchport mode **trunk**
switchport trunk allowed vlan 511
spanning-tree port type edge trunk

Access

interface Ethernet101/1/31
switchport mode **access**
switchport access vlan 511
spanning-tree port type edge

Switch dependent [Port Channel/vPC]



Static

interface Ethernet101/1/31-32
switchport mode trunk
switchport trunk allowed vlan 511
spanning-tree port type edge trunk
channel-group 300 **mode on**

LACP

interface Ethernet101/1/31-32
switchport mode trunk
switchport trunk allowed vlan 511
spanning-tree port type edge trunk
channel-group 300 **mode active**

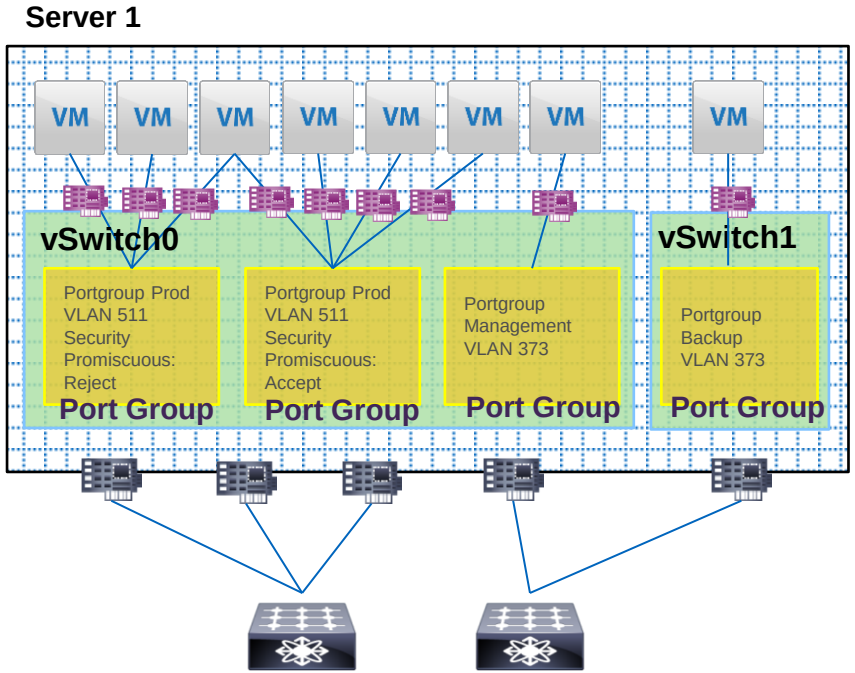
interface Port-Channel300
switchport mode trunk
switchport trunk allowed vlan 511
spanning-tree port type edge trunk

Hyper Visor - VMware vSphere ESXi

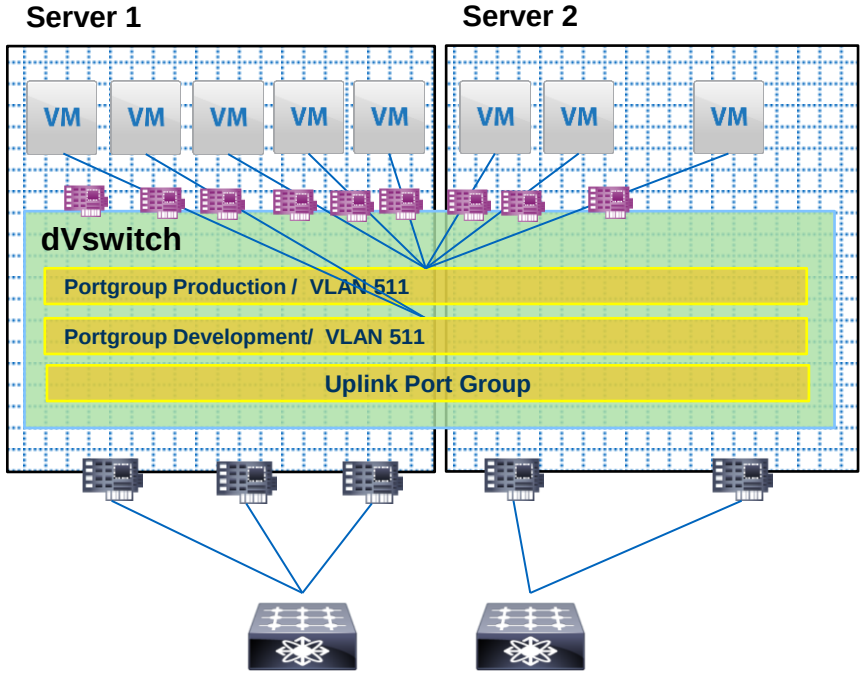
Cisco Connect
Seoul, Korea
April 29-30, 2014

vSphere Standard Switch [vSS] vs Distrubuted Switch [vDS]

vSphere Standard Switch [vSS]

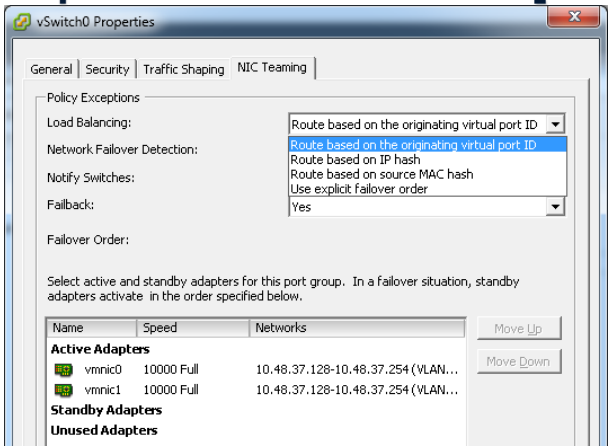


vSphere Distributed Switch [vDS]

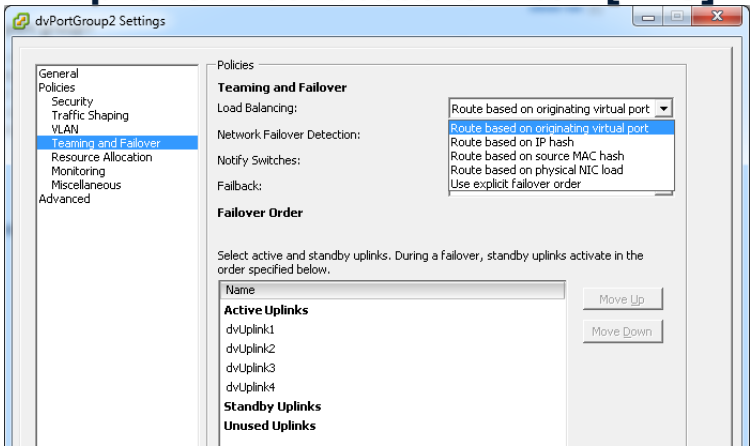


vSwitch Uplink 옵션

vSphere Standard Switch [vSS]



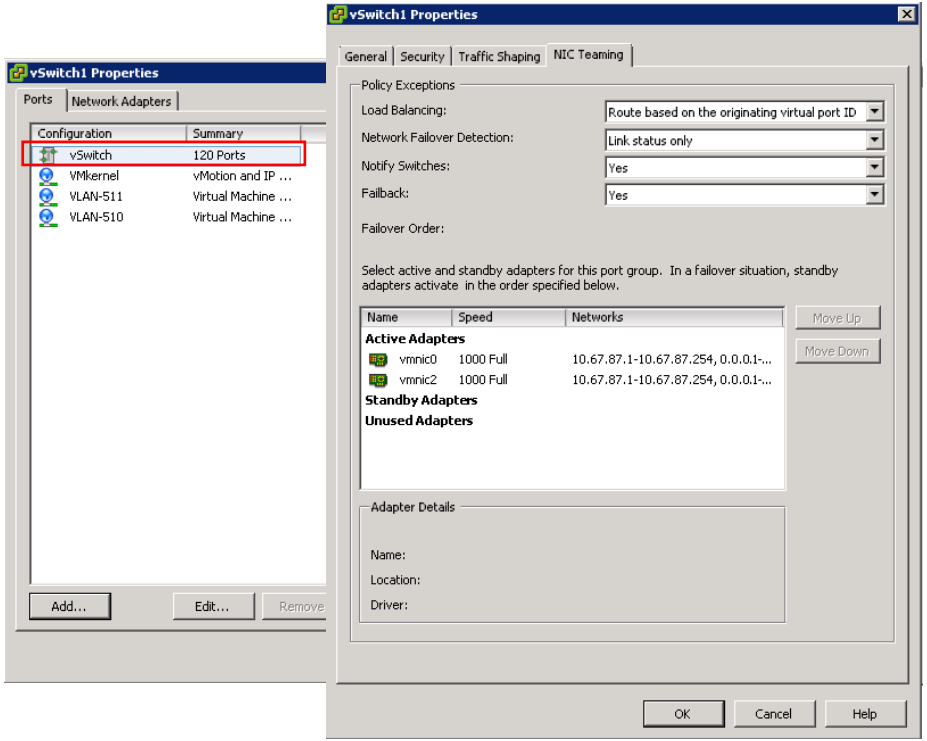
vSphere Distributed Switch [vDS]



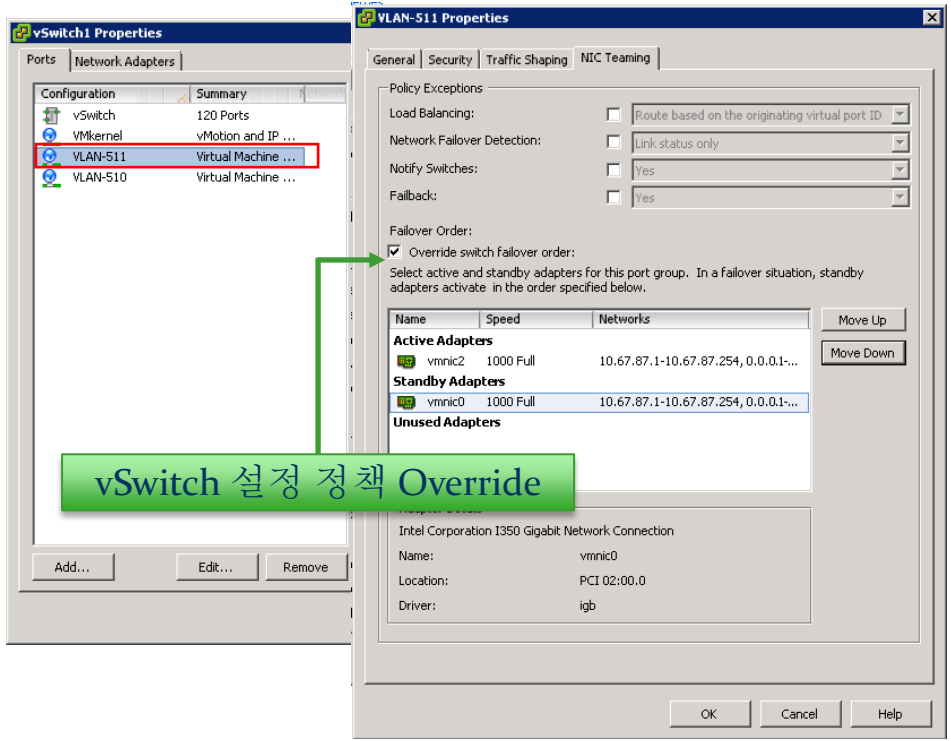
vSphere ESXi 영문버전	vSphere ESXi 한글버전	스위치 의존성
Route based on originating **virtual port ID**	원래 가상포트를 기반으로 라우팅	Switch Independent 포트채널(X)
Route based on source **MAC hash**	소스 MAC 해시 기준 라우팅	
Use **explicit failover** order	명시적 페일오버 명령사용	
Route based on **physical NIC load** (vDS)	물리적 NIC 로드기준 라우팅	
Route based on **IP hash**	IP 해시 기준 라우팅	Switch Dependent 포트채널(O)
Route based on **IP hash + LACP** (vDS)	IP 해시 기준 라우팅 + LACP	

Load Balancing: VMware Standard Switch

vSwitch : 모든 포트그룹에 적용 [Global]

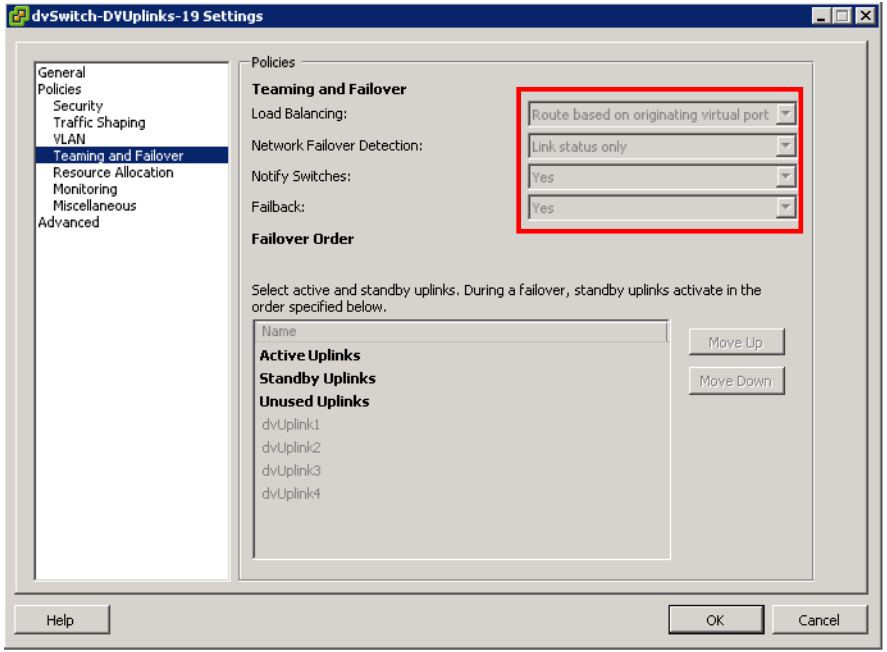


Portgroup: 해당 포트그룹에만 적용 [Local

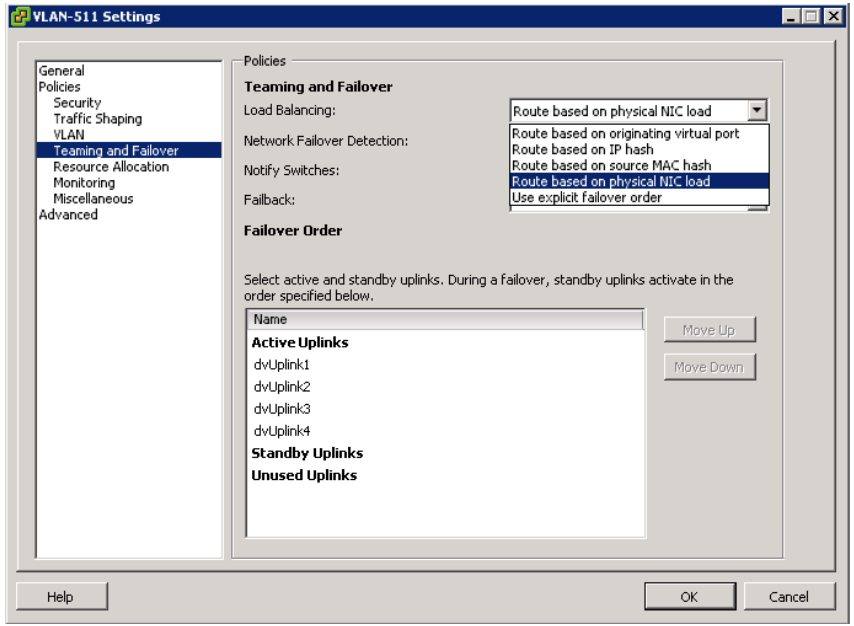


Load Balancing: VMware distributed Switch

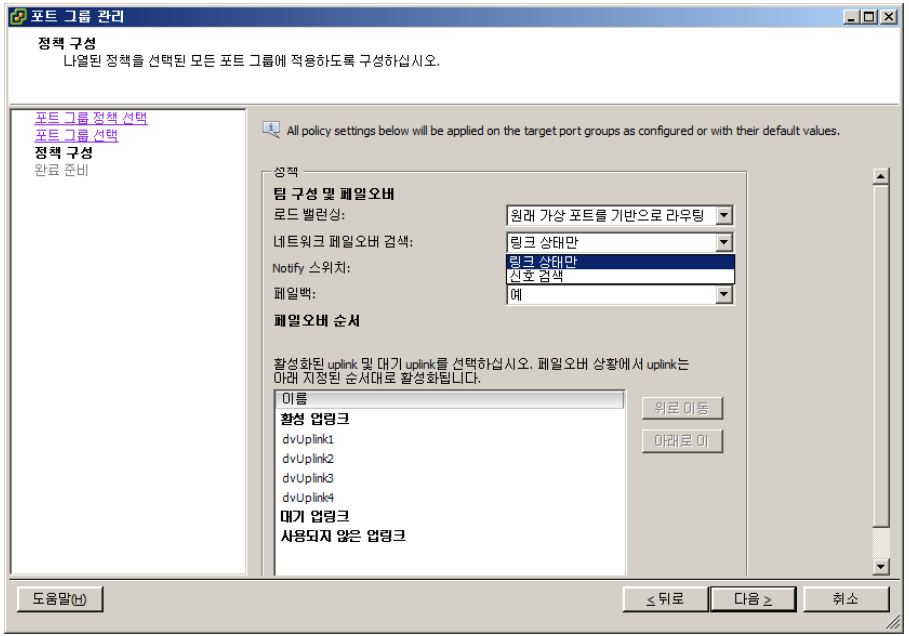
DVS 의 Uplink 포트그룹 정책 설정은 Web Client 에서만 가능 [LACP]



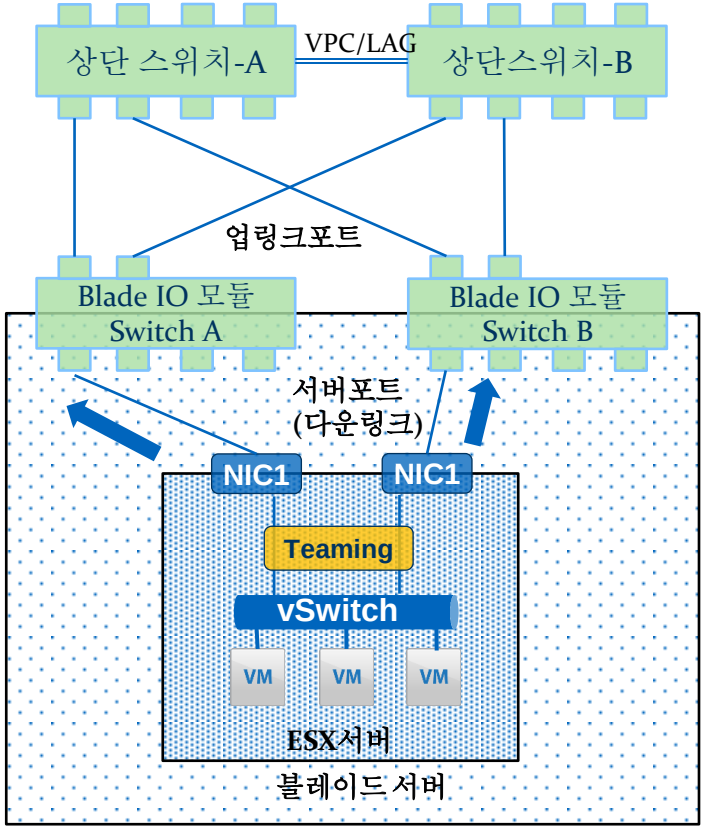
DVS 의 일반 포트그룹 정책 설정가능



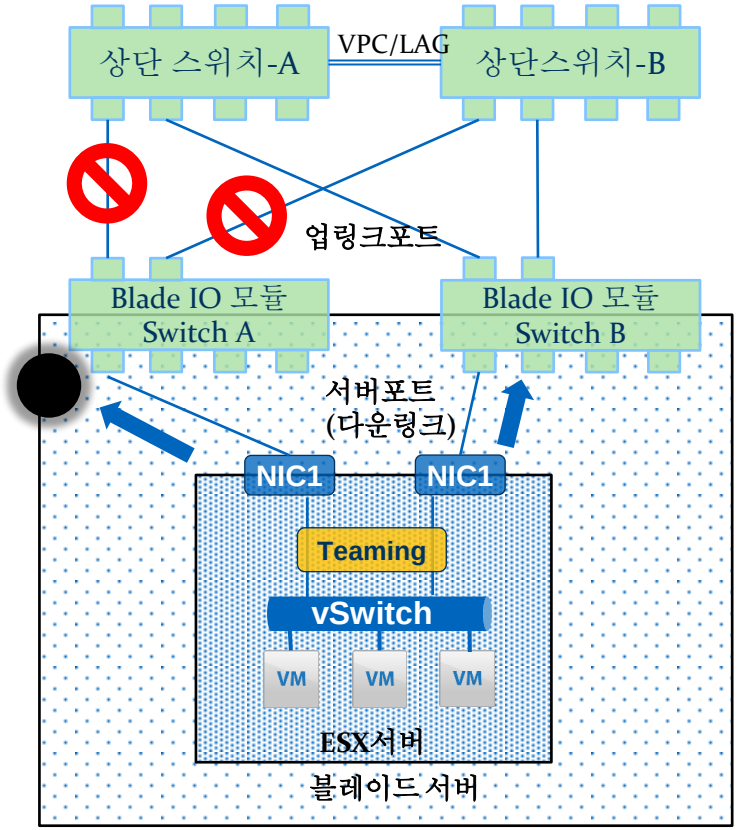
네트워크 페일오버 검색설정 – 링크상태만 [Link Status Only]

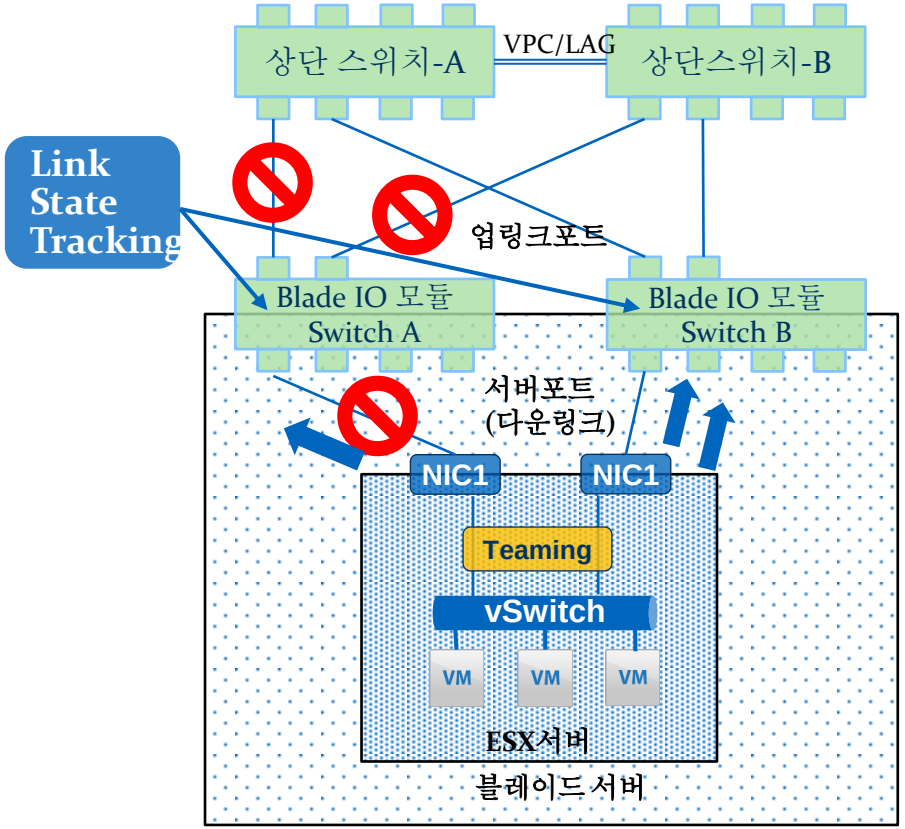


1. Link status only [Default]
2. Beacon probing



블랙홀 시나리오


상단 스위치-A
VPC/LAG
상단스위치-B
업링크포트
Blade IO 모듈
Switch A
Blade IO 모듈
Switch B
서버포트
(다운링크)
NIC1
NIC1
Teaming
vSwitch
VM
VM
VM
ESX서버
블레이드 서버


Link
State
Tracking
상단 스위치-A
VPC/LAG
상단스위치-B
업링크포트
Blade IO 모듈
Switch A
Blade IO 모듈
Switch B
서버포트
(다운링크)
NIC1
NIC1
Teaming
vSwitch
VM
VM
VM
ESX서버
블레이드 서버

Link State Tracking 설정 및 동작예시[C3012 스위치]

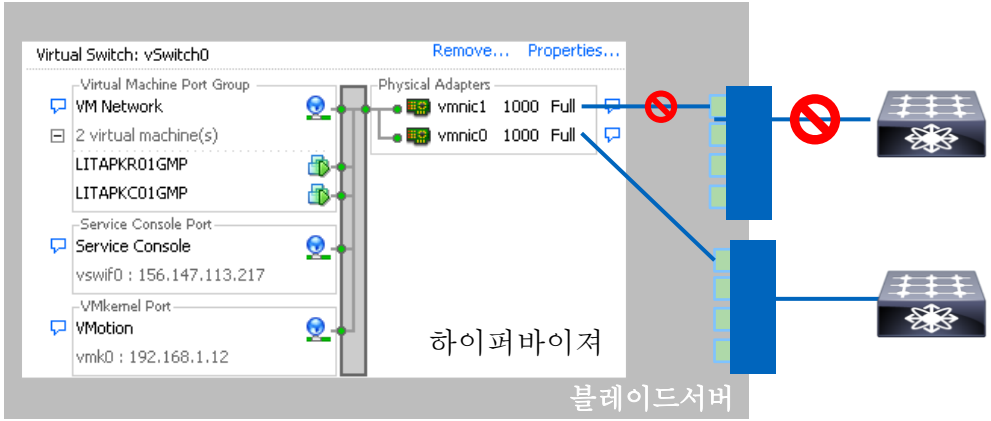
> Up-Link Port에 대한 설정 (Up-Link Port 5번)

```
!
interface GigabitEthernet0/15
 description Uplink port to Nortel L4 Switch
 link state group 1 upstream
!
```

> Down-Link Port에 대한 설정 (서버 Port 1~4번)

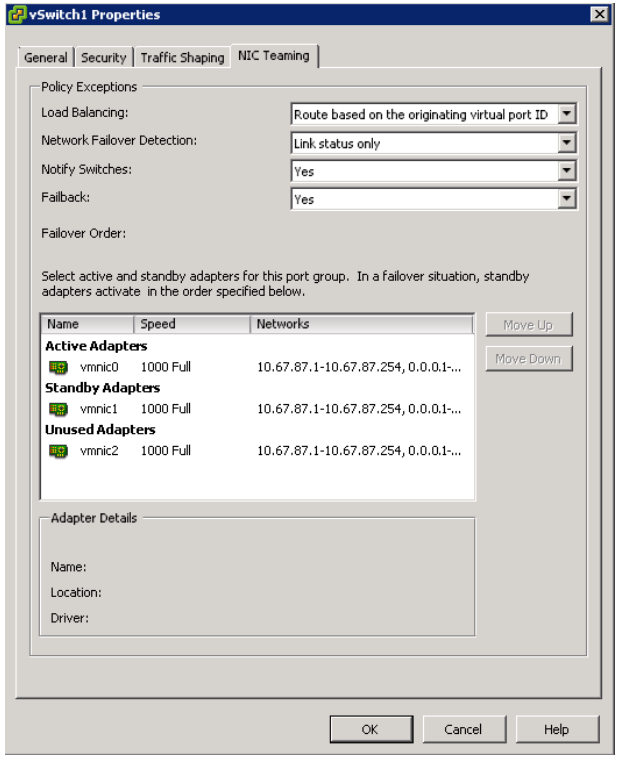
```
!
interface GigabitEthernet0/1
 description BL1-NIC0
 link state group 1 downstream
 spanning-tree portfast
!
interface GigabitEthernet0/2
 description BL2-NIC0
 link state group 1 downstream
 spanning-tree portfast
!
```

Failover 검색 : Link상태만

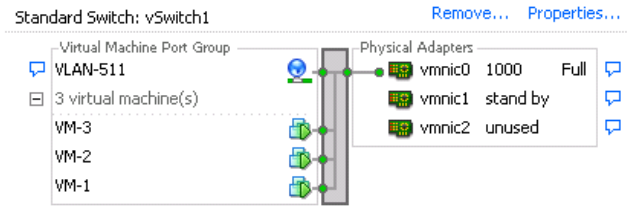



```
/var/log/vmkernel
...
Jul 14 12:41:01 bl1-esx05 vmkernel: 0:15:50:21.022 cpu4:4395)<3>bnx2: vmnic0 NIC SerDes Link is Down
```

Active / Standby / Unused 의 차이점



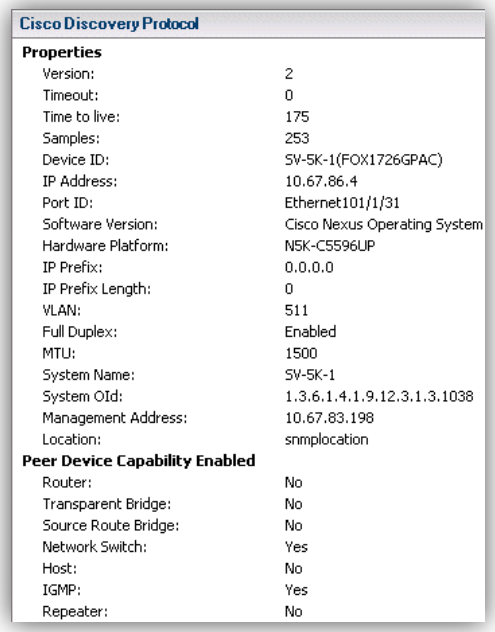
vSphere 호스트 관점



네트워크 관점

vmnic0	Ethernet101/1/31 is up
vminc1	Ethernet101/1/31 is up
vmnic2	Ethernet101/1/32 is up

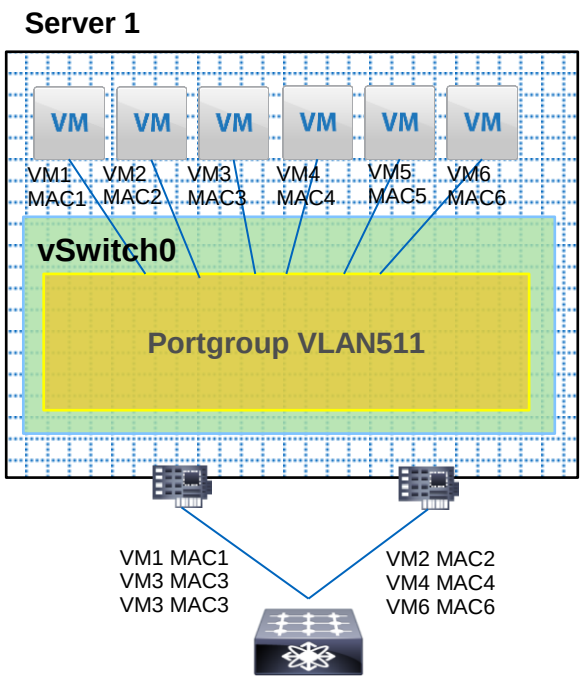
네트워크에서 해 줘야할일

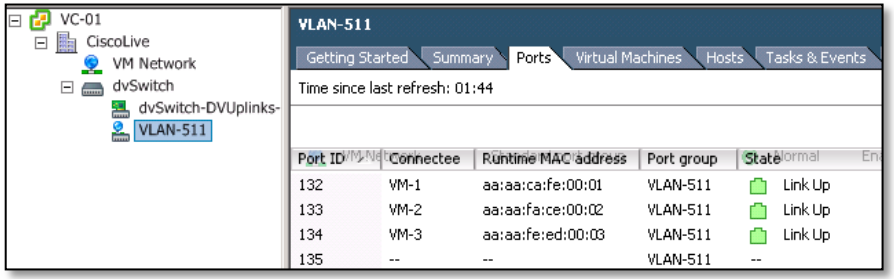


Enable cdp

Switch Independent

Route Based on Originating Virtual Port ID [포트ID 구성]



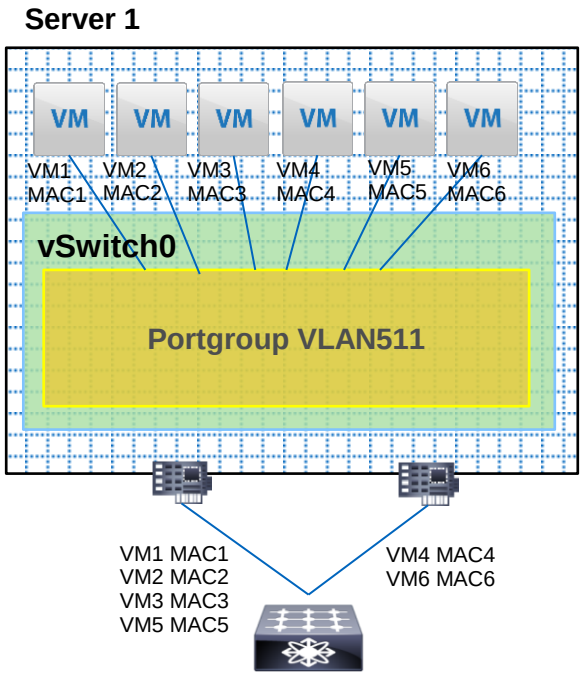



```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001     dynamic    0          F    F   Eth101/1/31
* 511       aaaa.face.0002     dynamic    0          F    F   Eth101/1/31
* 511       aaaa.feed.0003     dynamic    0          F    F   Eth101/1/32
SV-5K-1#
```

```
9:49:58am up  1:44, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,

  PORT-ID              USED-BY  TEAM-PNIC DNAME              PKTTX/s  MbTX/s
 33554433           Management        n/a vSwitch0            0.00    0.00
 33554434                vmnic3          - vSwitch0           53.10    1.52
 33554435     Shadow of vmnic3        n/a vSwitch0            0.00    0.00
 33554436                 vmk0     vmnic3 vSwitch0           53.10    1.52
 33554438 36042:VMware vCenter     vmnic3 vSwitch0            0.00    0.00
 50331649           Management        n/a DvsPortset-0        0.00    0.00
 50331650                vmnic0          - DvsPortset-0        2.16    0.00
 50331651     Shadow of vmnic0        n/a DvsPortset-0        0.00    0.00
 50331656      46658:VM-1.eth0     vmnic0 DvsPortset-0        0.98    0.00
 50331657      46777:VM-2.eth0     vmnic0 DvsPortset-0        1.18    0.00
 50331658      48692:VM-3.eth0     vmnic2 DvsPortset-0        1.18    0.00
 50331659                vmnic2          - DvsPortset-0        1.18    0.00
 50331660     Shadow of vmnic2        n/a DvsPortset-0        0.00    0.00
```

- vMware 기본설정 / 스위치 Independent
- 트래픽의 고른 분배 [가상NIC > 물리NIC]
- vNIC이 한쪽 Physical NIC 에 Pinning됨

Route Based on Source MAC Hash [소스 맥 해시]



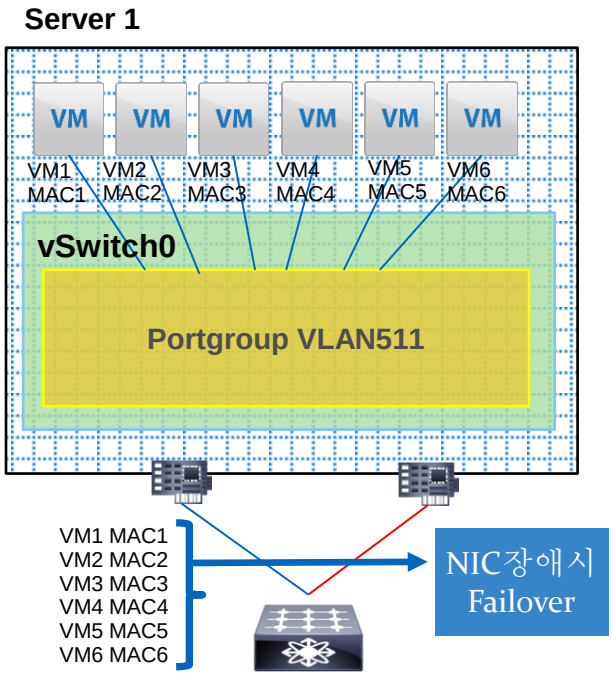

```
11:20:03am up  3:15, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01, 0
```

PORT-ID	USED-BY	TEAM-PNIC	DNAME	PKTTX/s	MbTX/s
33554433	Management	n/a	vSwitch0	0.00	0.00
33554434	vmnic3	-	vSwitch0	54.72	2.68
33554435	Shadow of vmnic3	n/a	vSwitch0	0.00	0.00
33554436	vmk0	vmnic3	vSwitch0	54.92	2.68
33554438	36042:VMware vCenter	vmnic3	vSwitch0	0.00	0.00
50331649	Management	n/a	DvsPortset-0	0.00	0.00
50331650	vmnic0	-	DvsPortset-0	1.18	0.00
50331651	Shadow of vmnic0	n/a	DvsPortset-0	0.00	0.00
50331659	vmnic2	-	DvsPortset-0	2.17	0.00
50331660	Shadow of vmnic2	n/a	DvsPortset-0	0.00	0.00
50331661	46658:VM-1.eth0	vmnic2*	DvsPortset-0	1.18	0.00
50331662	46777:VM-2.eth0	vmnic0*	DvsPortset-0	1.18	0.00
50331663	48692:VM-3.eth0	vmnic2*	DvsPortset-0	0.98	0.00

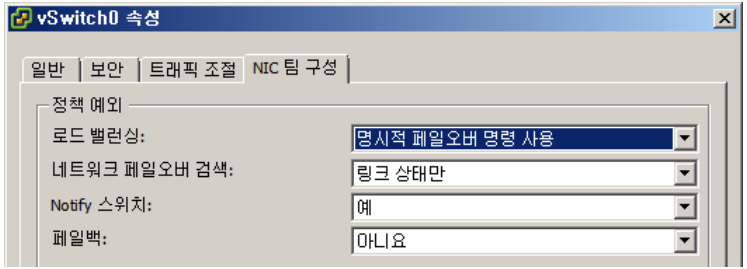
```
SV-5K-1# show mac address-table | grep aaa
* 511      aaaa.cafe.0001    dynamic   0         F    F  Eth101/1/32
* 511      aaaa.face.0002    dynamic   0         F    F  Eth101/1/31
* 511      aaaa.feed.0003    dynamic   0         F    F  Eth101/1/32
SV-5K-1#
```

- 스위치 독립적 / 소스 MAC 해시 기반
- 트래픽의 고른 분배 [Random MAC사용시]
- vNIC이 한쪽 Physical NIC 에 Pinning됨

Explicit Failover Order [명시적 페일오버]



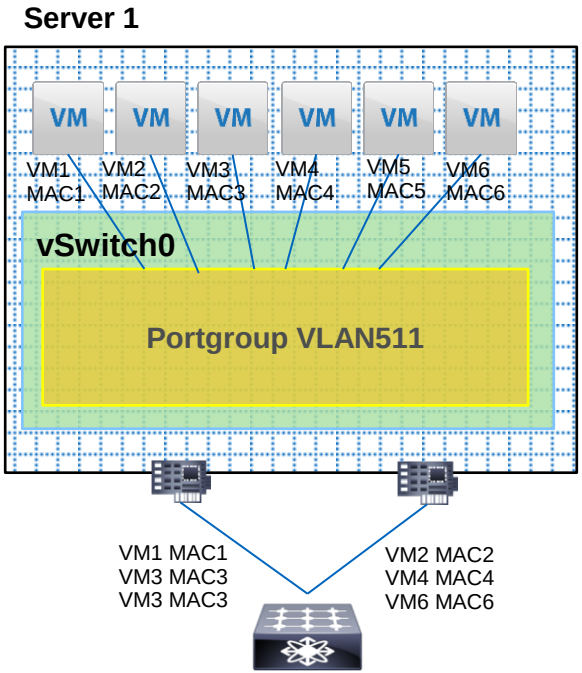
- 스위치 독립적 / 하나의 NIC만 사용
- Active/Standby 구성

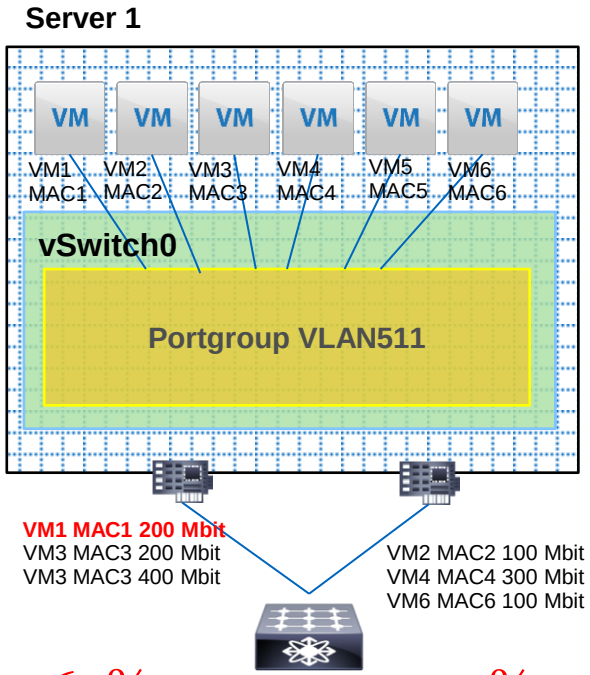



```
11:54:41am up  3:49, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,

  PORT-ID              USED-BY  TEAM-PNIC DNAME              PKTTX/s  MbTX/s
 33554433           Management        n/a vSwitch0            0.00    0.00
 33554434                vmnic3          - vSwitch0           17.13    0.34
 33554435     Shadow of vmnic3        n/a vSwitch0            0.00    0.00
 33554436                 vmk0     vmnic3 vSwitch0           17.13    0.34
 33554438 36042:VMware vCenter     vmnic3 vSwitch0            0.00    0.00
 50331649           Management        n/a DvsPortset-0        0.00    0.00
 67108865           Management        n/a vSwitch1            0.00    0.00
 67108868                vmnic0          - vSwitch1            4.33    0.00
 67108869     Shadow of vmnic0        n/a vSwitch1            0.00    0.00
 67108870                vmnic2          - vSwitch1            0.00    0.00
 67108871     Shadow of vmnic2        n/a vSwitch1            0.00    0.00
 67108872          46658:VM-1     vmnic0 vSwitch1            0.98    0.00
 67108873          46777:VM-2     vmnic0 vSwitch1            2.36    0.00
 67108874          48692:VM-3     vmnic0 vSwitch1            0.98    0.00
```

```
SV-5K-1# show int e101/1/31-32 | grep "is up"
Ethernet101/1/31 is up
Ethernet101/1/32 is up
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic    10          F    F  Eth101/1/31
* 511       aaaa.face.0002    dynamic    10          F    F  Eth101/1/31
* 511       aaaa.feed.0003    dynamic    10          F    F  Eth101/1/31
SV-5K-1# show mac address-table interface e101/1/32
SV-5K-1#
```

Route Based on Physical NIC Load [NIC 로드기반]

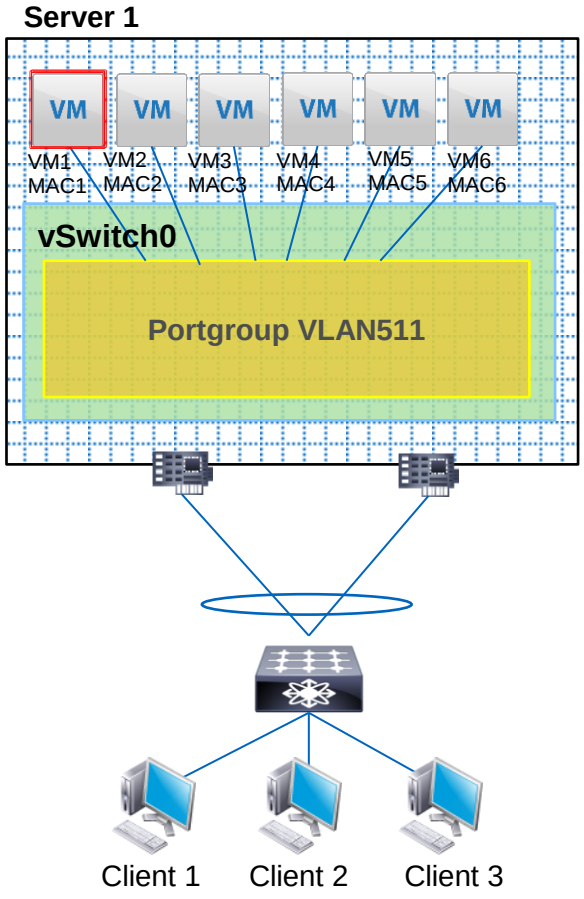




- 스위치 독립적 / 초기 포트ID 방식과 동일동작
- 한쪽 pNIC 의 RX/TX 가 75% 이상시 -> Rebalance
- VM Kernel 에서 30초간격으로 Load 검사

Switch Dependent

Route Based on IP Hash [IP 해시]



- Source IP + Destination IP 해시
- Static 802.3ad / 802.1ax 포트채널 [Mode On] 필요
- IP 구성에따라 고른 트래픽 분산 가능

```
11:01:40am up  2:56, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,
```

PORT-ID	USED-BY	TEAM-PNIC	DNAME	PKTTX/s	MbTX/s
33554433	Management	n/a	vSwitch0	0.00	0.00
33554434	vmnic3	-	vSwitch0	12.59	0.15
33554435	Shadow of vmnic3	n/a	vSwitch0	0.00	0.00
33554436	vmk0	vmnic3	vSwitch0	12.59	0.15
33554438	36042:VMware vCenter	vmnic3	vSwitch0	0.00	0.00
50331649	Management	n/a	DvsPortset-0	0.00	0.00
50331650	vmnic0	-	DvsPortset-0	127.26	0.12
50331651	Shadow of vmnic0	n/a	DvsPortset-0	0.00	0.00
50331659	vmnic2	-	DvsPortset-0	65.89	0.07
50331660	Shadow of vmnic2	n/a	DvsPortset-0	0.00	0.00
50331661	46658:VM-1.eth0	all(2)	DvsPortset-0	191.58	0.18
50331662	46777:VM-2.eth0	all(2)	DvsPortset-0	1.38	0.00
50331663	48692:VM-3.eth0	all(2)	DvsPortset-0	1.77	0.00

```
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F    Po301
SV-5K-1#
```

```
Group Port-       Type      Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
301   Po301(SU)   Eth       NONE      Eth101/1/31(P)  Eth101/1/32(P)
SV-5K-1#
```

Route Based on IP Hash Switch 설정

[Static 채널 설정]

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300

interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
```

```
SV-5K-1(config-if)# channel-group 300 mode
?
  active     Set channeling mode to ACTIVE
  on         Set channeling mode to ON
  passive    Set channeling mode to PASSIVE

SV-5K-1(config-if)# channel-group 300 mode
```

Route Based on IP Hash [스위치 포트채널 설정 없을시..]

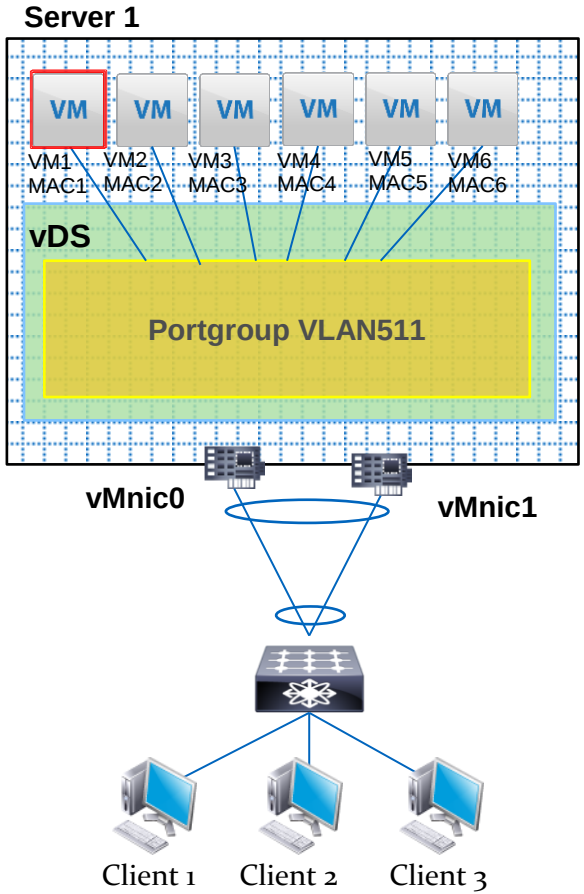
```
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/31
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/32
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/31
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   10         F    F  Eth101/1/32
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/32
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/31
```

MAC Flapping

```
SV-5K-1# show logging | grep mac
2014 Jan  3 13:38:46 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:42:07 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:45:27 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/31 and Eth101/1/32 vlan 511
2014 Jan  3 13:48:37 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:51:53 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/31 and Eth101/1/32 vlan 511
2014 Jan  3 13:54:54 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:58:01 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 14:01:03 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/31 and Eth101/1/32 vlan 511
SV-5K-1#
```

- Traffic 유실 및 서비스 장애 발생
- 조기 발견이 어렵고, 실제 서비스 개시후 트래픽 로드가 많이지면 증상이 심해짐

Route Based on IP Hash + LACP



- IP Source + Destination 해시
- IP 구성에따라 고른 트래픽 분산 가능
- Dynamic 802.3 ad (LACP Active 또는 Passive)

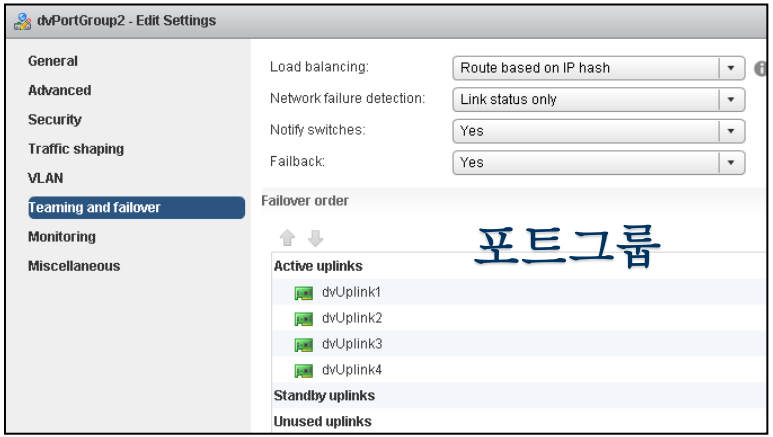
```
11:12:31am up  3:07, 565 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,

  PORT-ID              USED-BY  TEAM-PNIC DNAME              PKTTX/s  MbTX/s
  33554433                                                      0.00    0.00
  33554434                                                      0.00    0.00
  33554435     Shado                                            0.00    0.00
  33554436                                                      0.00    0.00
  33554438 36042:VMw                                            0.00    0.00
  50331649           Management        n/a DvsPortset-0         0.00    0.00
  50331650                vmnic0          - DvsPortset-0         2.36    0.00
  50331651     Shadow of vmnic0        n/a DvsPortset-0         0.00    0.00
  50331659                vmnic2          - DvsPortset-0         0.98    0.00
  50331660     Shadow of vmnic2        n/a DvsPortset-0         0.00    0.00
  50331661      46658:VM-1.eth0   all(2)  DvsPortset-0         1.18    0.00
  50331662      46777:VM-2.eth0   all(2)  DvsPortset-0         0.98    0.00
  50331663      48692:VM-3.eth0   all(2)  DvsPortset-0         1.18    0.00
  50331664        LACP_MgmtPort     n/a DvsPortset-0         0.00    0.00
```

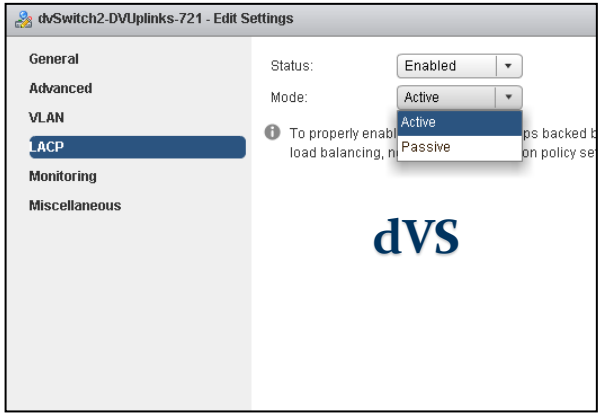
```
SV-5K-1# show mac address-table | grep aaa
* 511      aaaa.cafe.0001    dynamic   0         F    F    Po300
* 511      aaaa.face.0002    dynamic   0         F    F    Po300
* 511      aaaa.feed.0003    dynamic   0         F    F    Po300
SV-5K-1#
```

```
-------------------------------------------------------------------------------
Group Port-        Type     Protocol  Member Ports
      Channel
-------------------------------------------------------------------------------
300   Po300(SU)    Eth      LACP      Eth101/1/31(P)   Eth101/1/32(P)
SV-5K-1#
```

VMware Switch Dependent

Route Based on IP Hash + LACP(vDS) – Switch 설정방법






```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300 mode active
```

```
SV-5K-1(config-if)# channel-group 300 mode ?
  active     Set channeling mode to ACTIVE
  on         Set channeling mode to ON
  passive    Set channeling mode to PASSIVE
```

```
SV-5K-1(config-if)# channel-group 300 mode
```

VMware 구성시의 결론

- **vSphere Standard Switch (vSS)**
  - Switch independent [포트채널(x)]
    - Route based on originating virtual port
  - Switch dependent [포트채널(O)]
    - Route based on IP hash

- **vSphere Distributed Switch (vDS)**
  - Switch independent [포트채널(x)]
    - Route based on physical NIC load
  - Switch dependent [포트채널(O)]
    - Route based on IP hash + LACP

UCS 는?

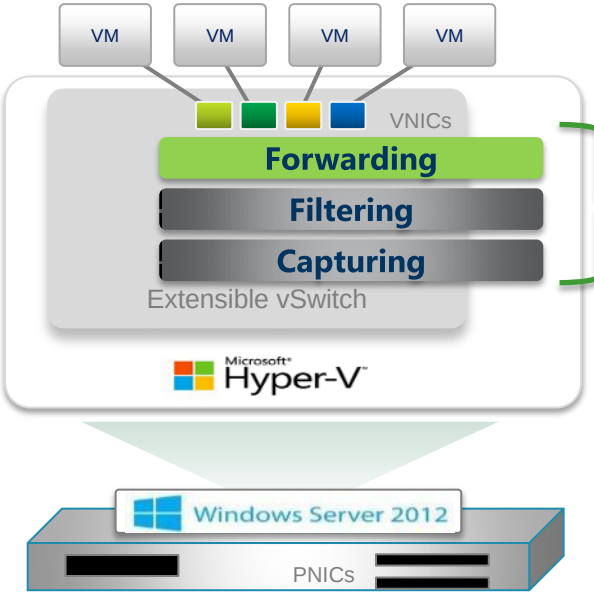
= Route based on physical NIC load [vDS]

Port ID [vSS]

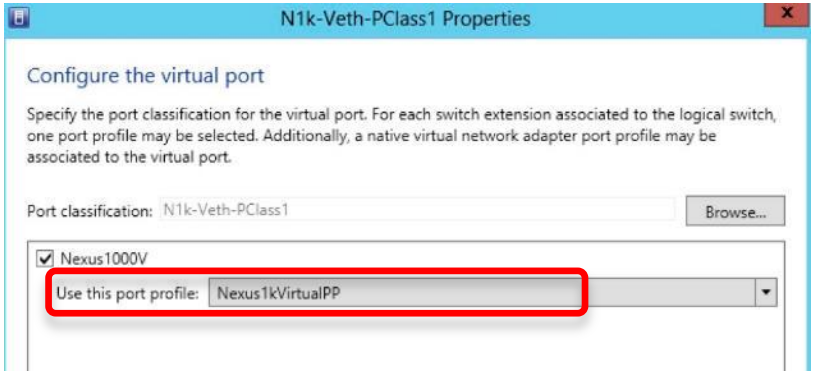
Hyper Visor – Microsoft Windows 2012 R2

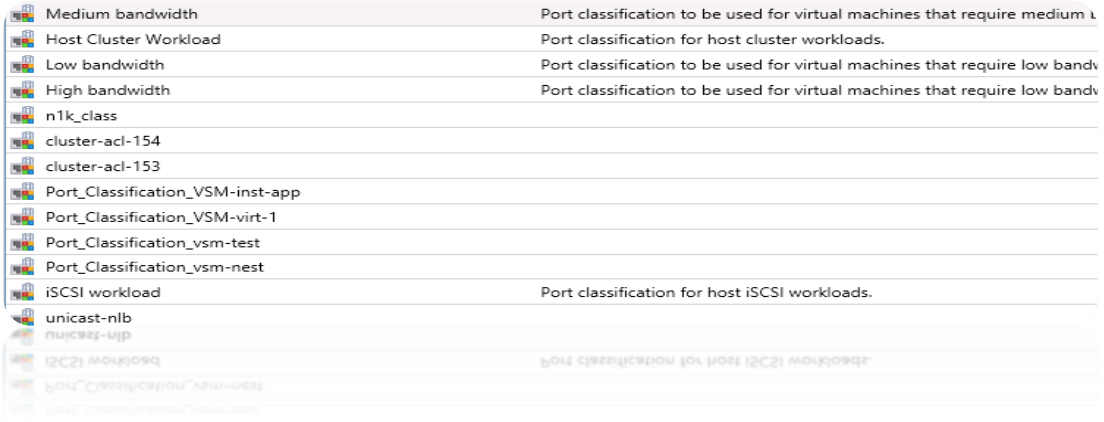
Cisco Connect
Seoul, Korea
April 29-30, 2014

Microsoft SCVMM 네트워킹 컨셉트

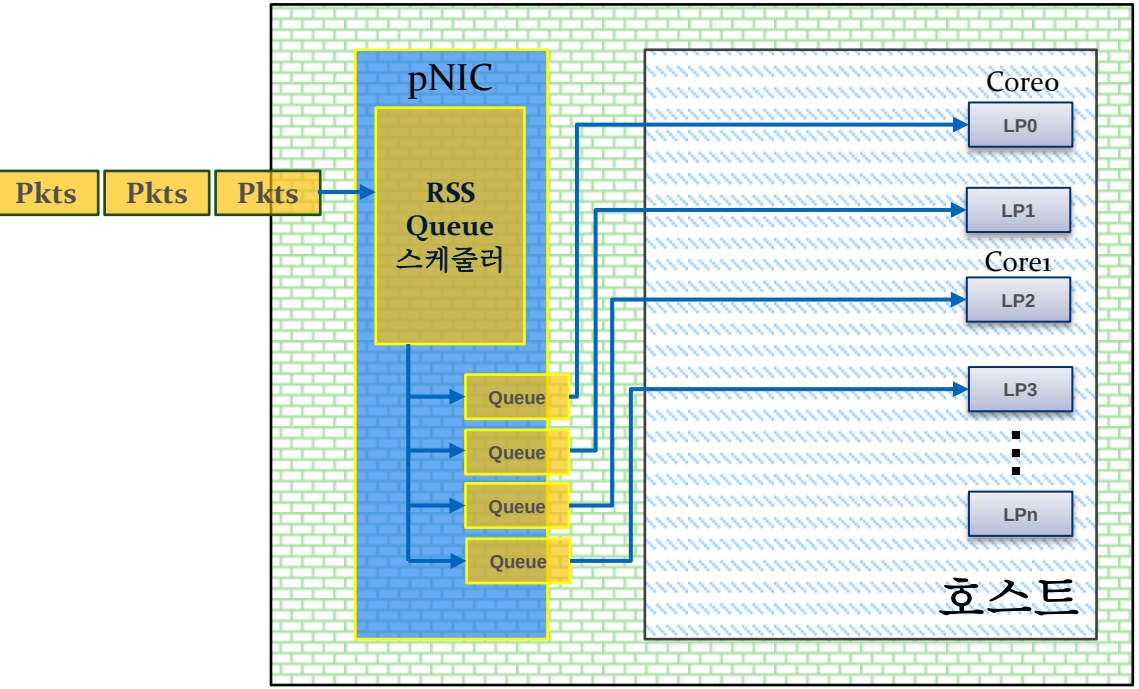


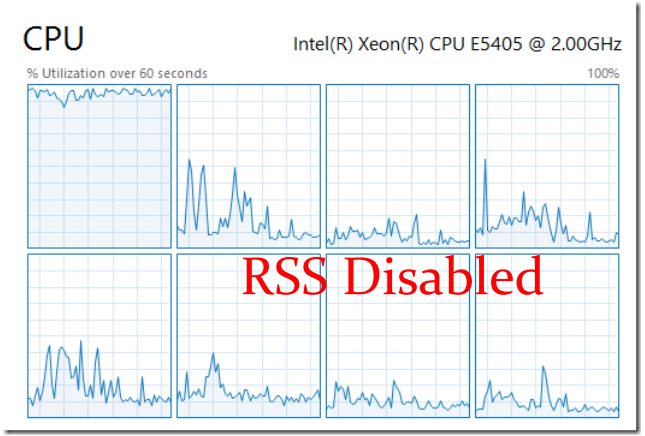
확장[Extensible] 모듈들의 정책이 모여 가상네트워킹 포트의 프로파일 구성

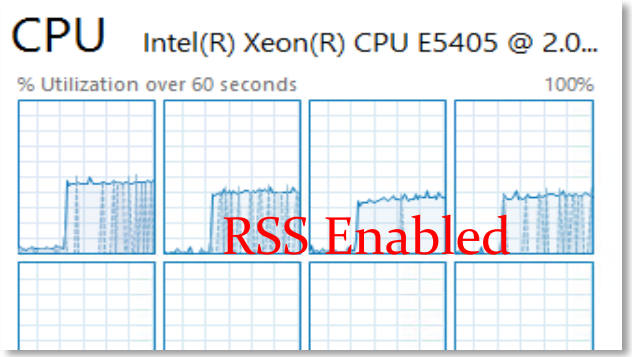




RSS - Receive Side Scaling

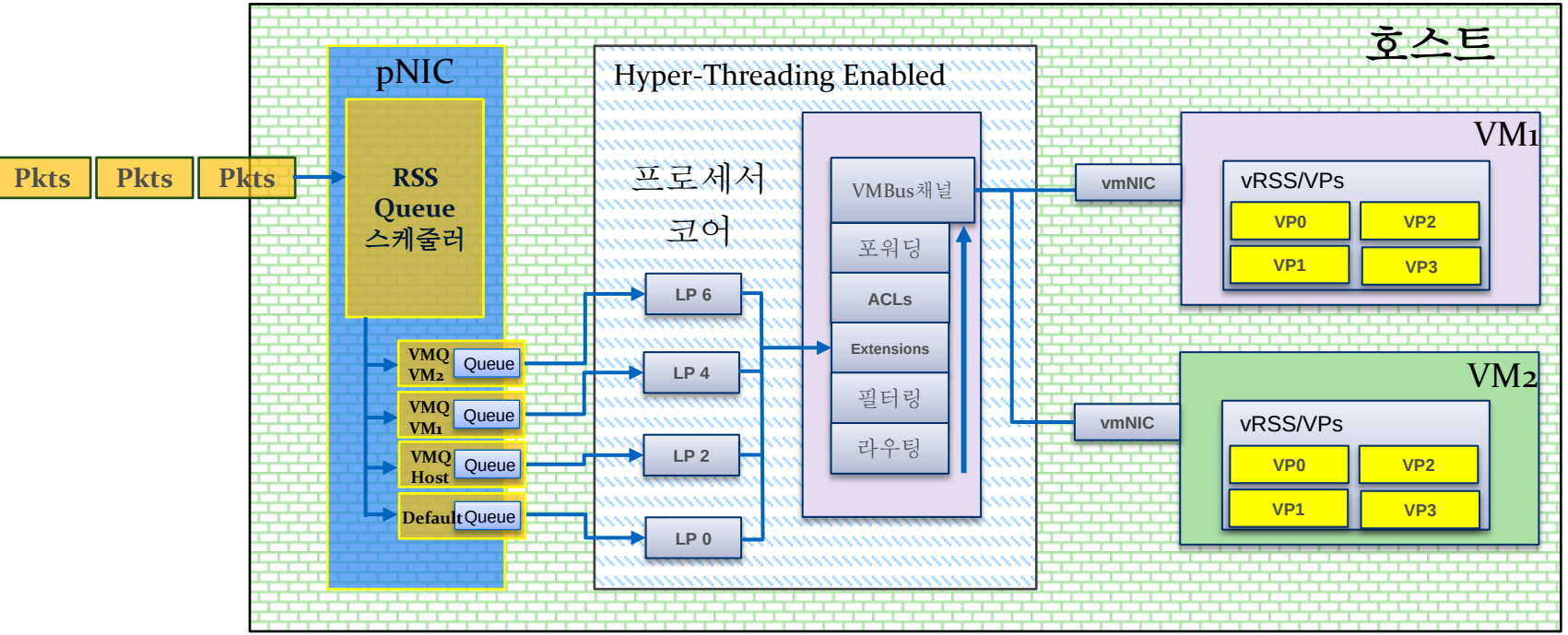






Enabled: Windows 2008/R2 Disabled : Windows 2012/R2

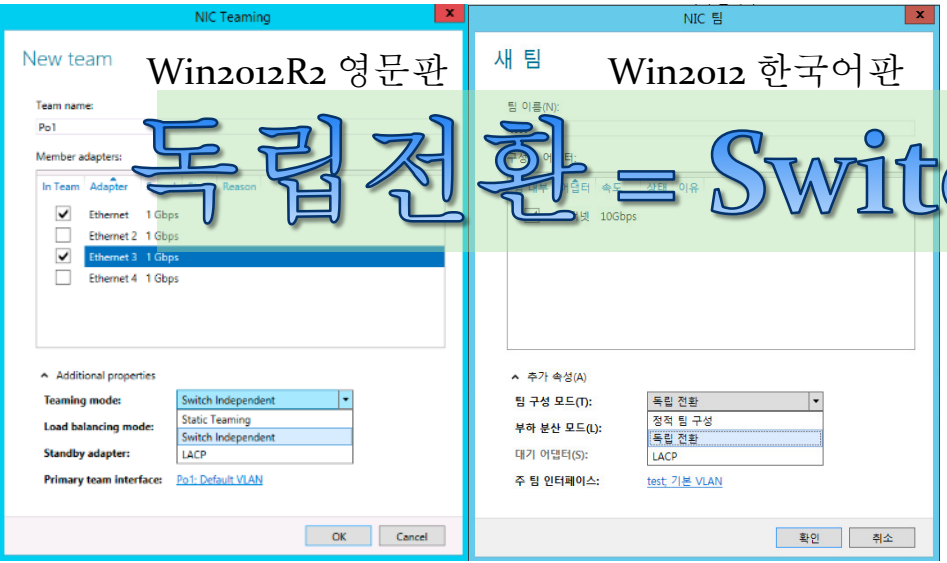
VMQ - Virtual Machine Queue



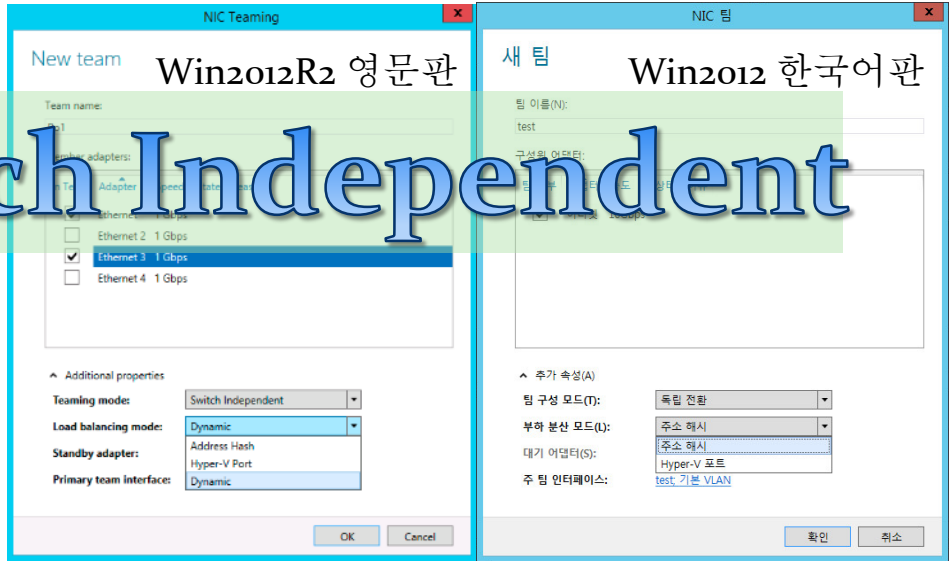
vSwitch 가 생성되면, RSS 는 Disable 되고, VMQ 가 Enable 됨

Teaming 및 로드분산 옵션

팀 구성모드 옵션



부하분산모드 옵션



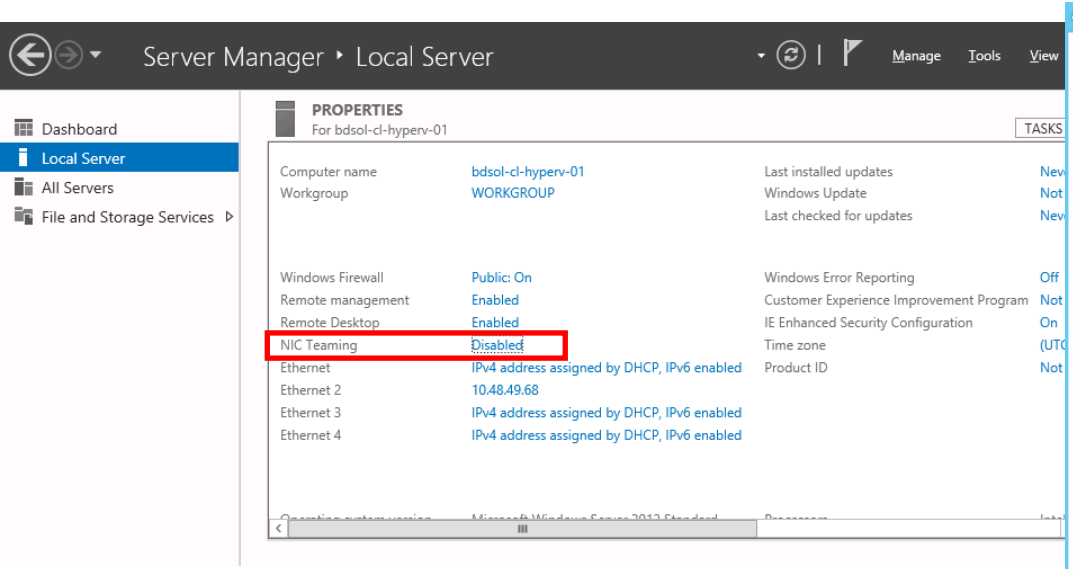
독립전환 = Switch Independent

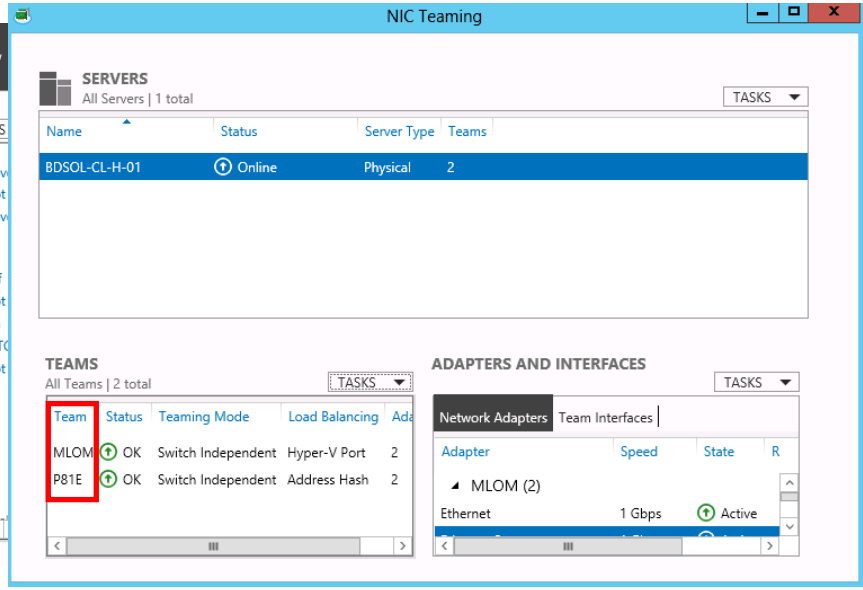
- 팀 구성모드: Switch Independent [독립전환]
  Switch Dependent [Static /LACP]
  [정적 팀구성 /LACP]

- 부하 분산모드: Address Hash [주소해시]
  Hyper-V Port [Hyper-V포트]
  Dynamic

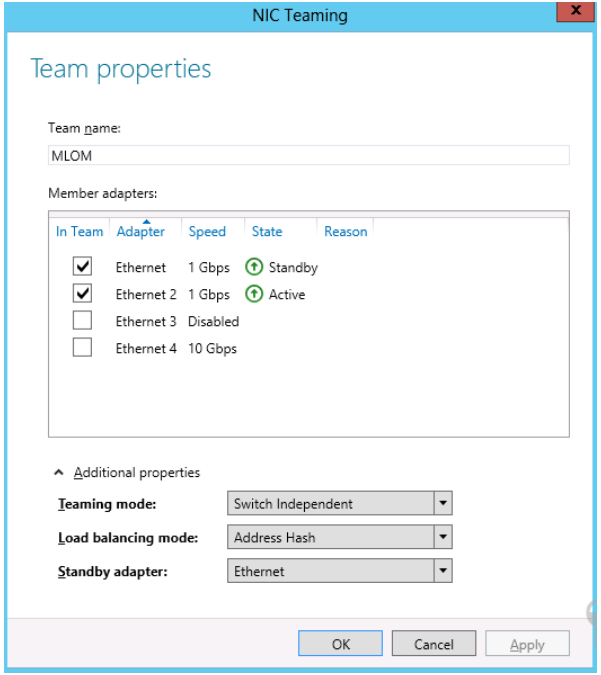
NIC Teaming 설정

- Windows 2012 이전 버전 - NIC 제조사의 Teaming 소프트 웨어 사용
- Windows 2012 이후 버전 - Windows 자체 Teaming 기능 제공





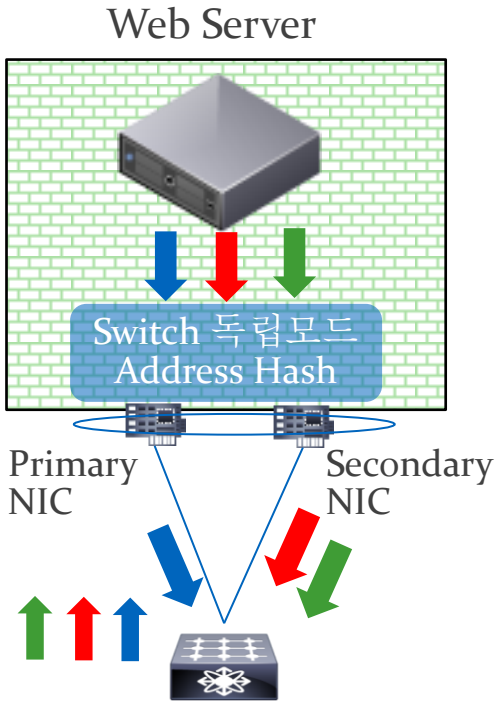
Active / Standby의 의미

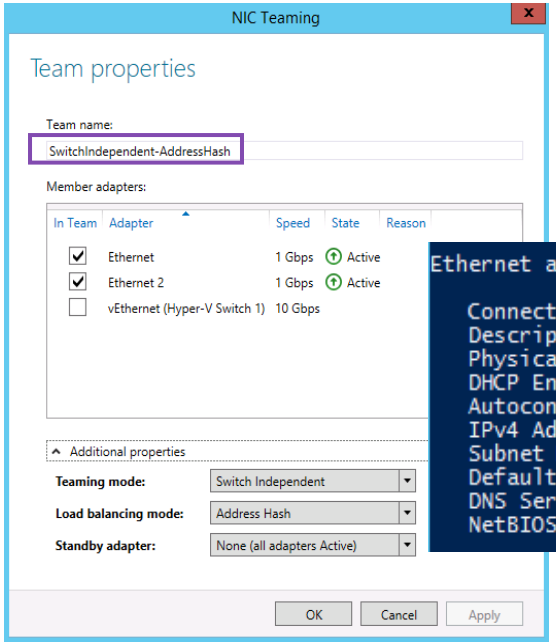


Nexus 5k 확인

```
SV-5K-1# show int e101/1/31-32 | grep "is up"
Ethernet101/1/31 is up
Ethernet101/1/32 is up
SV-5K-1# show mac address-table | grep aaa
* 511      aaaa.cafe.0001    dynamic   10         F    F  Eth101/1/31
* 511      aaaa.face.0002    dynamic   10         F    F  Eth101/1/31
* 511      aaaa.feed.0003    dynamic   10         F    F  Eth101/1/31
SV-5K-1# show mac address-table interface e101/1/32
SV-5K-1#
```

Switch Independent - Address Hash [Hyper-V 에는 비추]





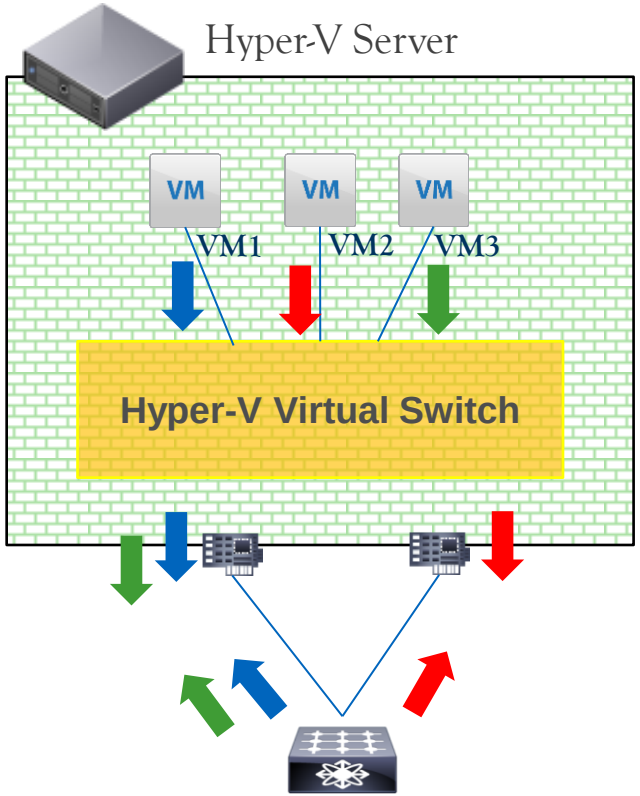
- **Outbound : 고른 트래픽 분산**
- **Inbound : 언제나 Primary NIC사**

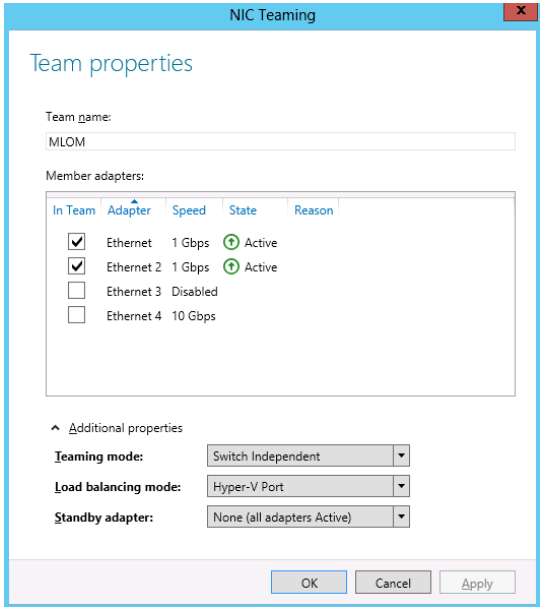
```
Ethernet adapter SwitchIndependent-AddressHash:

   Connection-specific DNS Suffix  . :
   Description . . . . . . . . . . . : Microsoft Network Adapte
   Physical Address. . . . . . . . . : 50-3D-E5-9D-32-ED
   DHCP Enabled. . . . . . . . . . . : No
   Autoconfiguration Enabled . . . . : Yes
   IPv4 Address. . . . . . . . . . . : 8(Preferred)
   Subnet Mask . . . . . . . . . . . : .0
   Default Gatewa
   DNS Servers . .                   : 245
   NetBIOS over Tcpip. . . . . . . . : Enabled
```

Teaming 인터페이스
: Primary NIC MAC사용

```
SV-5K-2# show mac address-table vlan 511 | egrep Eth101/1/31|Eth101/1/32
* 511      503d.e59d.32ed    dynamic   10          F    F  Eth101/1/31
* 511      503d.e59d.32ef    dynamic   10          F    F  Eth101/1/32
SV-5K-2#
```

Switch Independent - Hyper-V Port

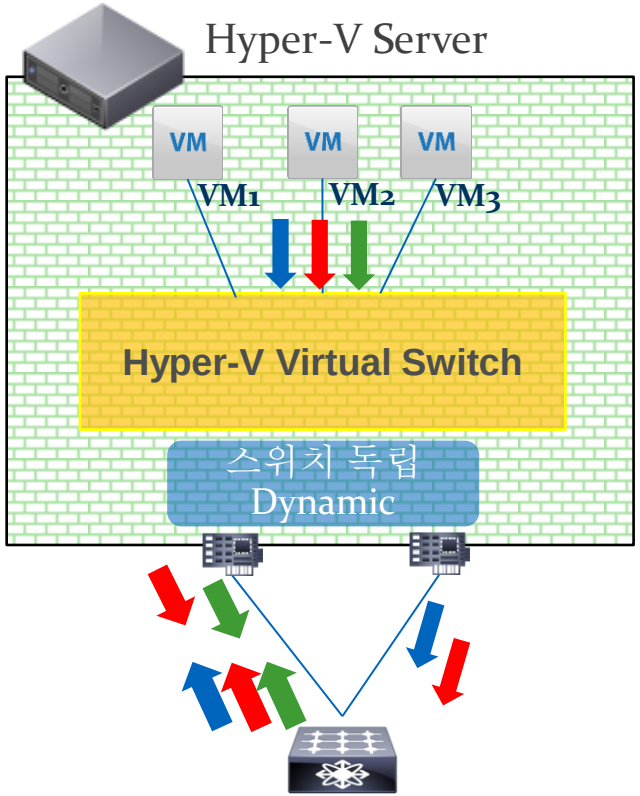




- Hyper-V Port-ID 에따른 vNIC Pinning
- VM의 가용대역폭이 한쪽 아답터로 제한됨

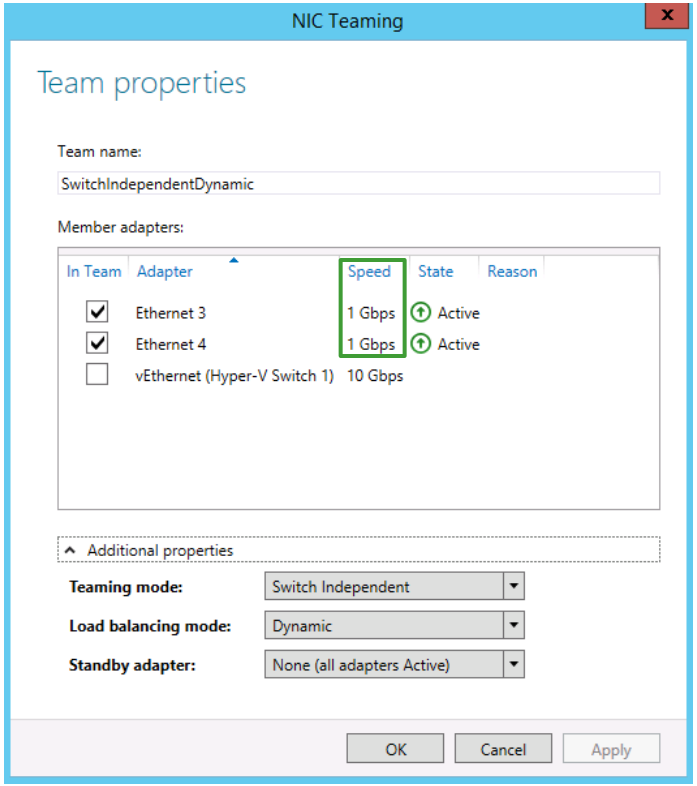
```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/31
* 511       aaaa.face.0002    dynamic   0          F    F  Eth101/1/31
* 511       aaaa.feed.0003    dynamic   0          F    F  Eth101/1/32
SV-5K-1#
```

Switch Independent - Dynamic



- Windows 2012 R2 부터 지원
- Outbound 트래픽은 pNIC의 대역폭 현황에 따라 재 분배됨
- pNIC간 패킷의 재조합[Reordering]없이 트래픽 이동가능 (flowlets)
- 단일 Outbound 플로우가 여러 개의 Source MAC 을 사용
- Inbound 트래픽은 Hyper-V Port mode 와같이 동작

Switch Independent - Dynamic




```
SV-5K-1# show mac address-table | egrep Eth101/1/31|Eth101/1/32
* 511       503d.e59d.32ec     dynamic   180         F     F  Eth101/1/31
* 511       503d.e59d.32ee     dynamic   10          F     F  Eth101/1/32
* 511       aaaa.cafe.0001     dynamic   0           F     F  Eth101/1/31
* 511       aaaa.face.0002     dynamic   0           F     F  Eth101/1/31
* 511       aaaa.feed.0003     dynamic   0           F     F  Eth101/1/32
SV-5K-1#
```

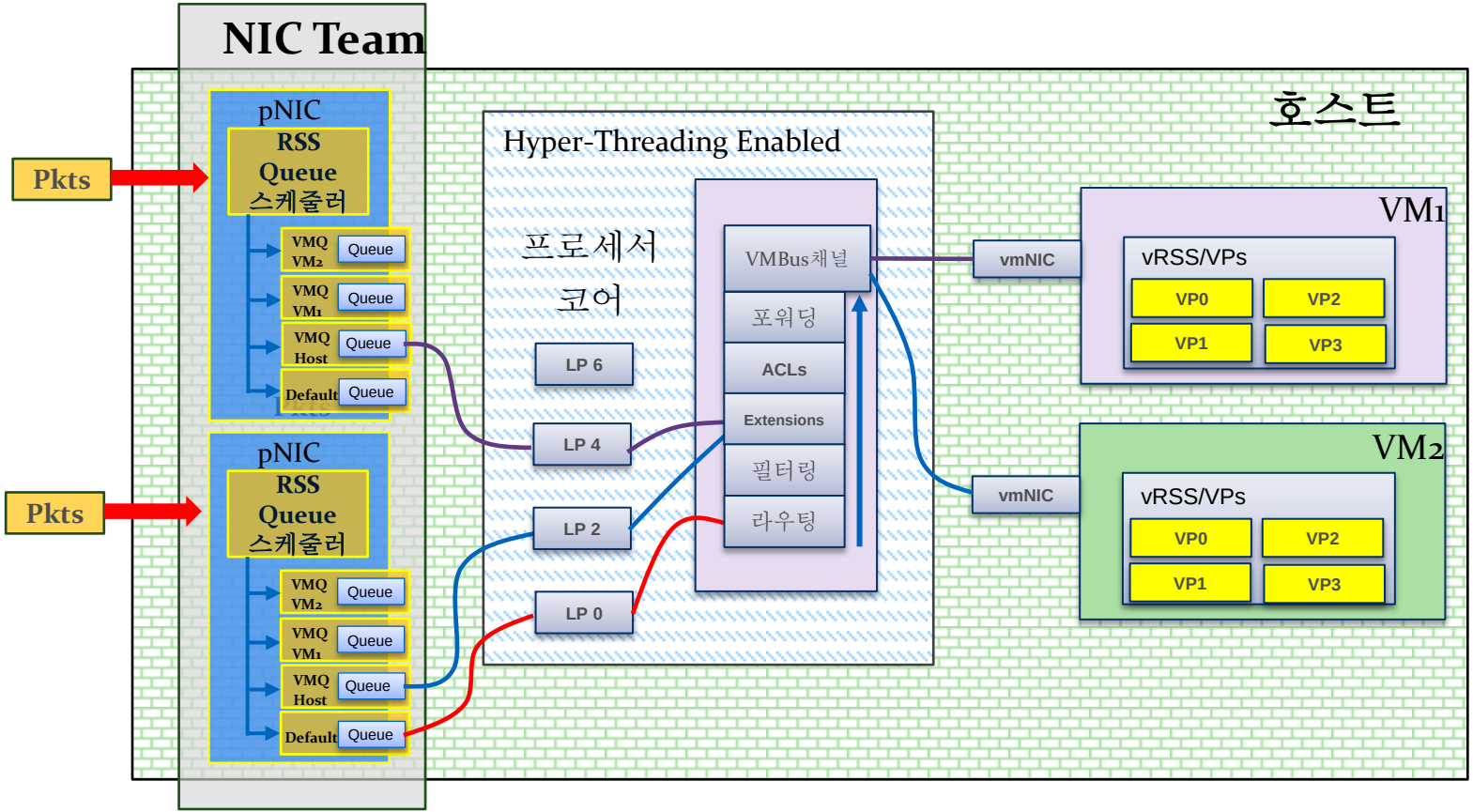
VM 2에서 iPerf 테스트 : Outbound 1.82Gb

```
[240]   5.0- 6.0 sec  5.59 MBytes  46.9 Mbits/sec
[232]   5.0- 6.0 sec  4.06 MBytes  34.1 Mbits/sec
[224]   5.0- 6.0 sec  32.8 MBytes   275 Mbits/sec
[216]   5.0- 6.0 sec  34.8 MBytes   292 Mbits/sec
[208]   5.0- 6.0 sec  10.7 MBytes  89.6 Mbits/sec
[200]   5.0- 6.0 sec  86.9 MBytes   729 Mbits/sec
[192]   5.0- 6.0 sec  36.0 MBytes   302 Mbits/sec
[248]   5.0- 6.0 sec  6.41 MBytes  53.7 Mbits/sec
[SUM]   5.0- 6.0 sec   217 MBytes  1.82 Gbits/sec
```

Teaming 된 NIC의 최대 B/W 사용가능

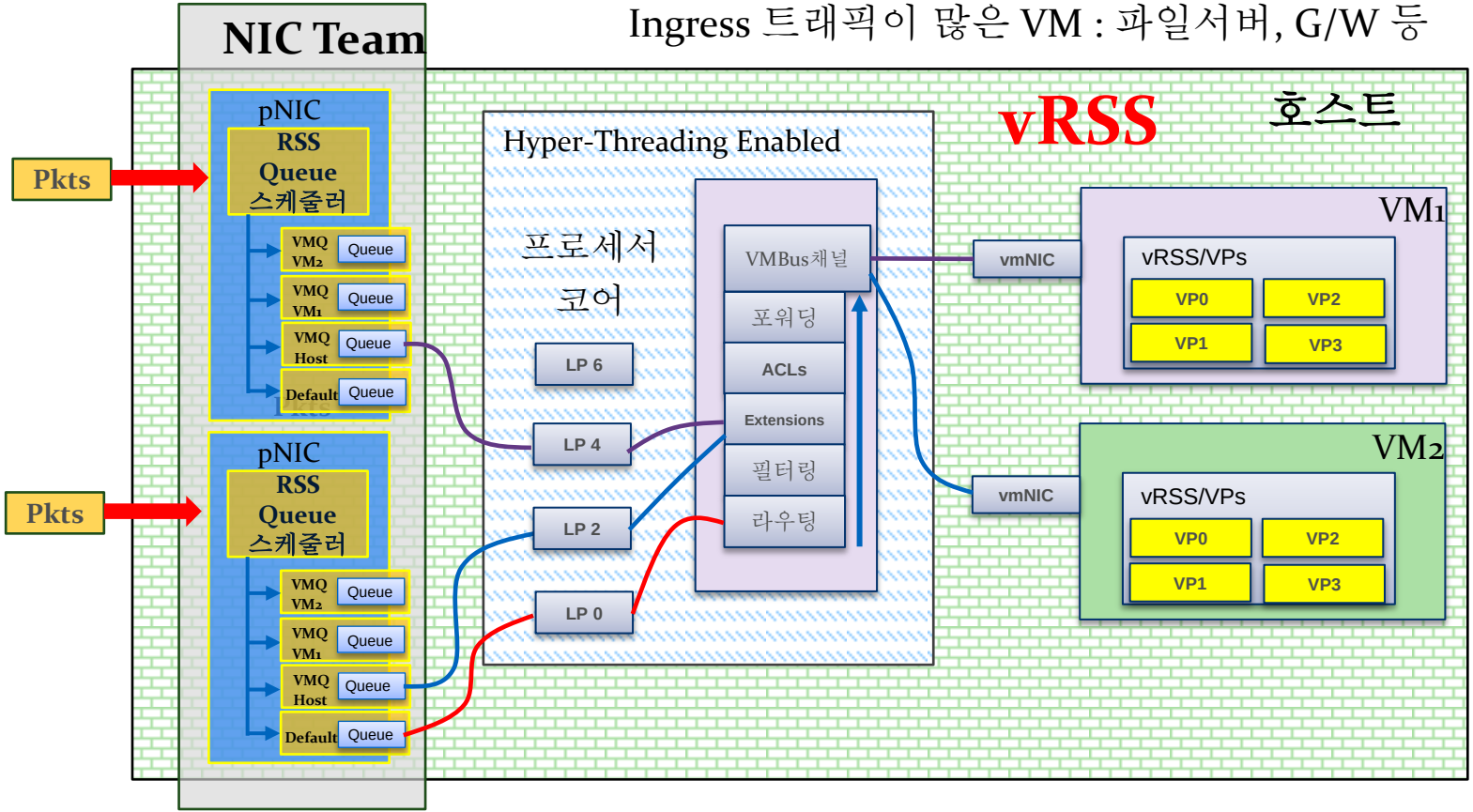
Switch Independent - VMQ 리포팅 모드 - Sum of Queue

Hyper-V Port and Dynamic

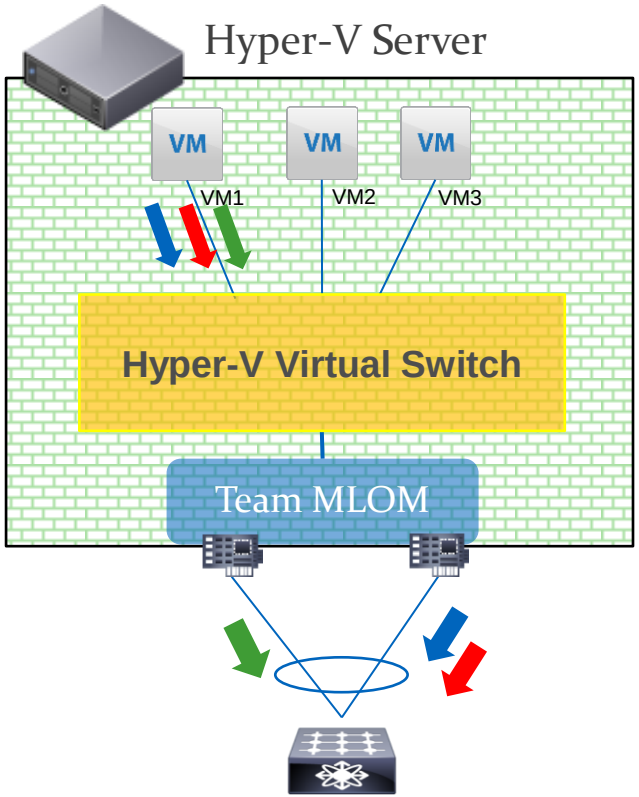


VMQ + vRSS

Windows 2012 R2 추가기능

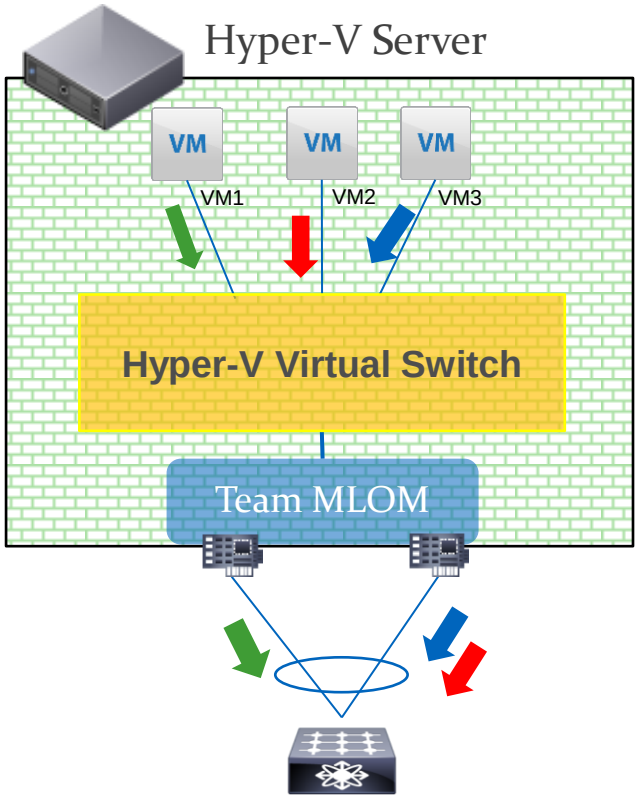


Switch Dependent - Address Hash



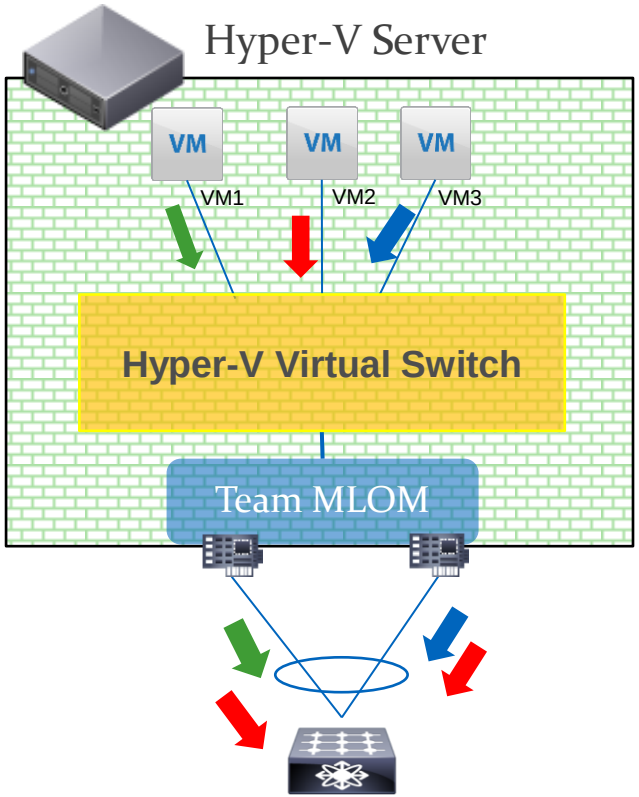
- Outbound 트래픽은 Address Hash에 기반하여 Teaming멤버pNIC들로 분산
- Inbound 트래픽은 스위치의 LoadBalancing 정책에따라 분산
- 스위치에서 Static / LACP 포트채널 필요
- 사용사례 : 다양한 Apps 를 구동하는 VM을 사용하며 Traffic의 분산이 필요시

Switch Dependent - Hyper-V Port



- vNIC 들은 Teaming 멤버중 하나의 pNIC에 Pinning
- Inbound 트래픽은 스위치의 LoadBalancing 정책에 따라 분산
- Static / LACP 포트채널 필요
- 사용사례 : 여러 VM 들이 같은 Application 을 구동하는 경우

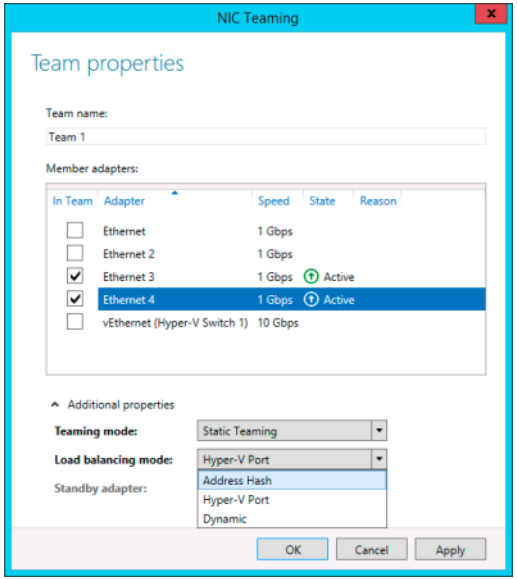
Switch Dependent - Dynamic

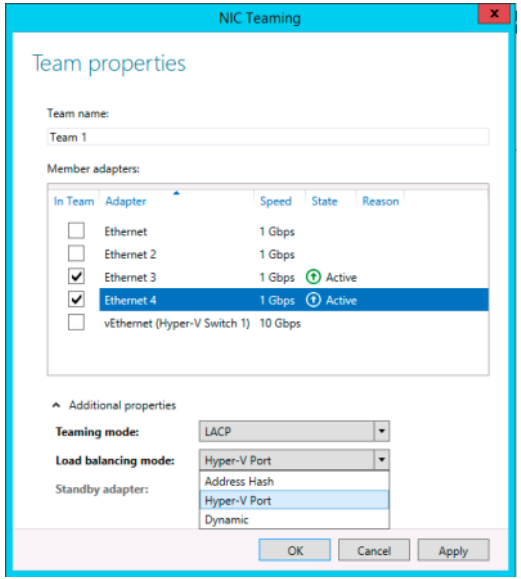


- Windows 2012 R2 부터 지원
- Outbound 트래픽은 pNIC의 대역폭 현황에 따라 재 분배됨
- Inbound 트래픽은 스위치의 LoadBalancing 정책에따라 분산
- Static / LACP 포트채널 필요

Switch Dependent

Load Balancing Mode: Address Hash, Hyper-V Port and Dynamic





Teaming 모드:

Static Teaming/LACP

Nexus Switch 설정

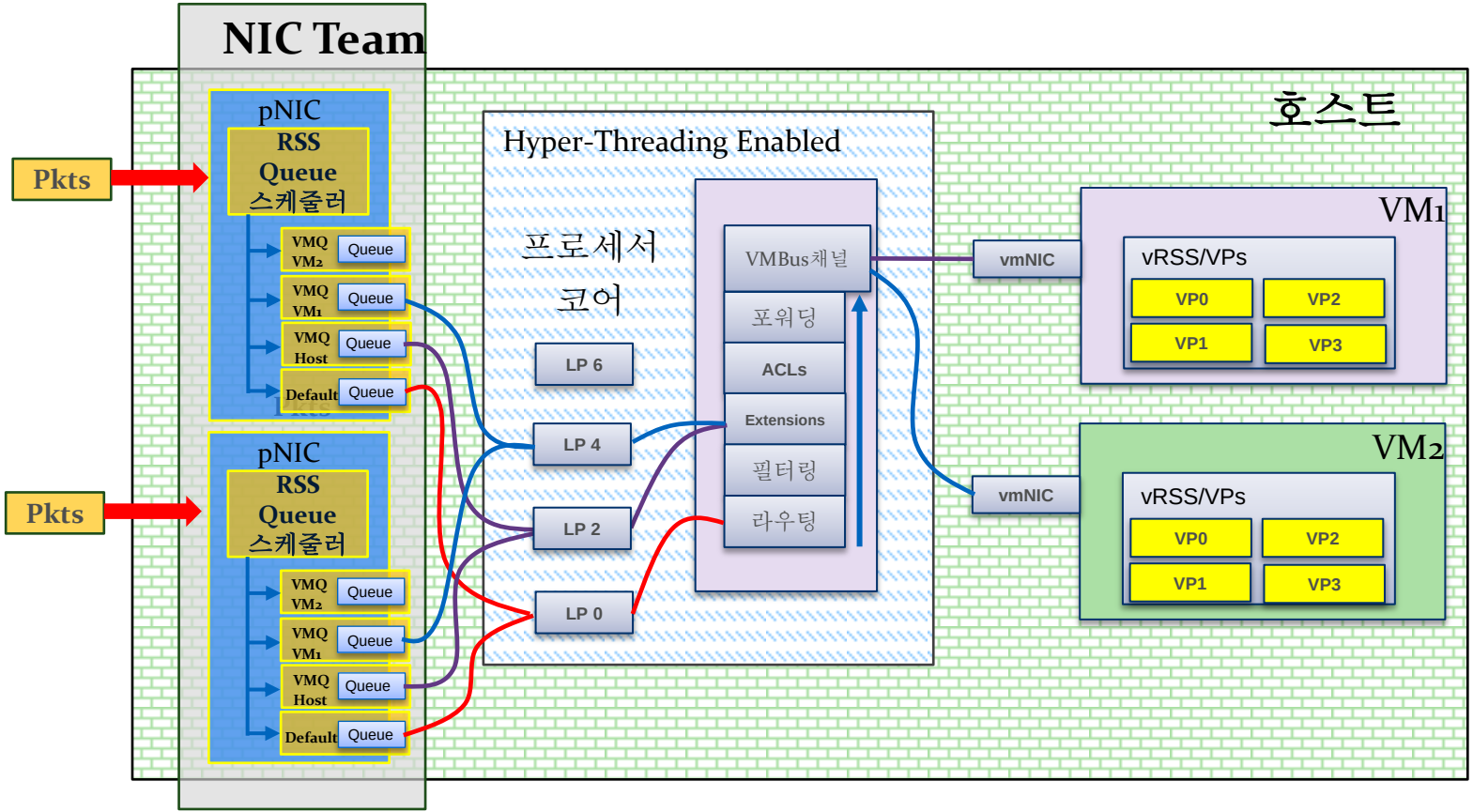
VMware 설정시와 동일

```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic   0         F    F   Po300
* 511       aaaa.face.0002    dynamic   0         F    F   Po300
* 511       aaaa.feed.0003    dynamic   0         F    F   Po300
SV-5K-1#
```

모든 MAC은 Portchannel 에서 Learning

Switch Dependent - VMQ 의 리포팅모드 - Min Queue

Switch Dependent + Independent [Address Hash]



VMQ 의 리포팅모드 설정 Best Practice

	Address Hash	Hyper-V Port	Dynamic
Switch Dependent	Min Queues	Min Queues	Min Queues
Switch Independent t	Min Queues	Sum of Queues	Sum of Queues

[설정예시] 20 Core [HT=40 LP] , 2x 10G NIC

Minimum Queue :

- Power Shell : Set-NetAdapterVMQ –Name Eth1, Eth2 –BaseProcessorNumber 2 – MaxProcessors 19

Sum of Queue :

- Power Shell : Set-NetAdapterVMQ –Name Eth1 –BaseProcessorNumber 2 –MaxProcessors 10
  Eth1 포트 트래픽은 Core 2 - 10 사용

- Power Shell : Set-NetAdapterVMQ –Name Eth2 –BaseProcessorNumber 22 –MaxProcessors 9
  Eth2 포트 트래픽은 Core 11 - 20 사용

Windows 2012/[R2] 구성시의 결론

- **Switch Independent**
  - Teaming 모드: Switch Independent
  - LB 모드 : Dynamic [Win2012 R2]
  - LB 모드 : Hyper-V Port [Win2012]

- **Switch Dependent**
  - Teaming 모드: LACP
  - LB 모드 : Dynamic [Win2012 R2]
  - LB 모드 : Hyper-V Port [Win2012]

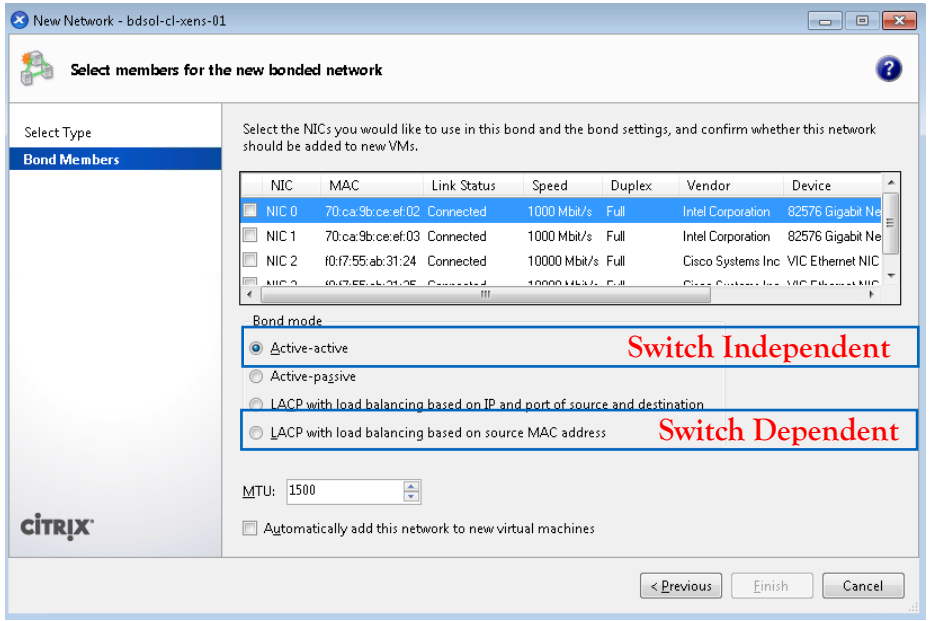
= Dynamic [Win2012R2]/Hyper-V Port[Win2012]

Hyper Visor - Xen Server 6.2

Cisco Connect
Seoul, Korea
April 29-30, 2014

Uplink Options

Xen 6.1 이후 LACP 옵션추가



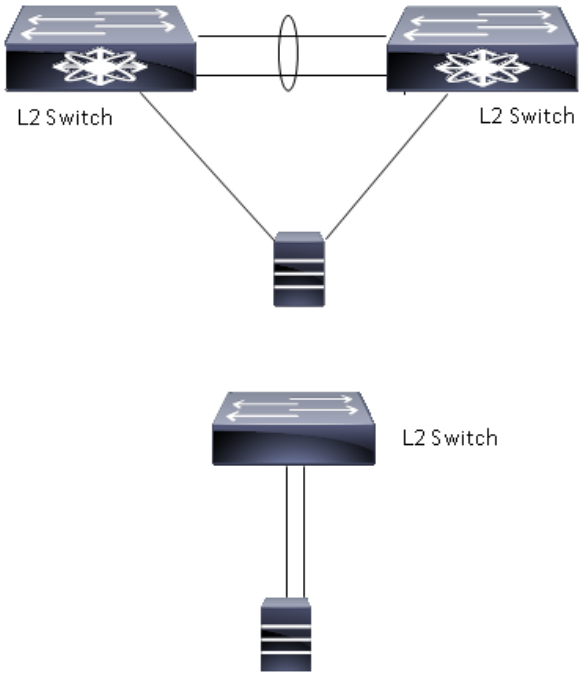
- Switch Independent - Active/Active
- Switch dependent - LACP LB w/ Src MAC

Outbound 트래픽은 SRC MAC Hashing 을 통해vNIC Pinning 되지만, 10초마다 Load에 따른 재분배

Hypervisor - Network 토폴로지 정리

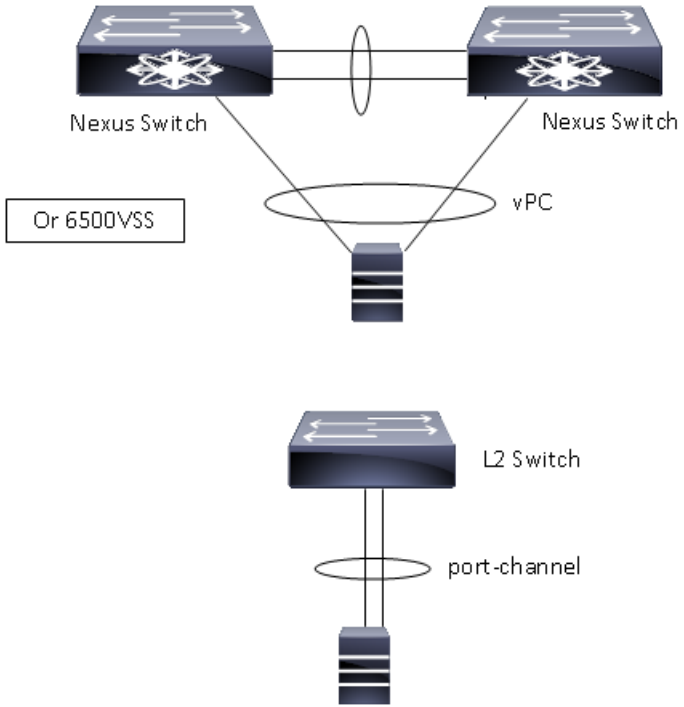
Cisco Connect
Seoul, Korea
April 29-30, 2014

Switch Independent



- **VMware**
  - Route based on originating virtual port (vSS)
  - Route based on source MAC hash
  - Route based on physical NIC load (vDS)
  - Use explicit failover order
- **Hyper-V**
  - Switch Independent - Address hash
  - Switch Independent - Hyper-V Port mode
  - Switch Independent – Dynamic
- **XenServer**
  - Active-active
  - Active-passive

Switch Dependent

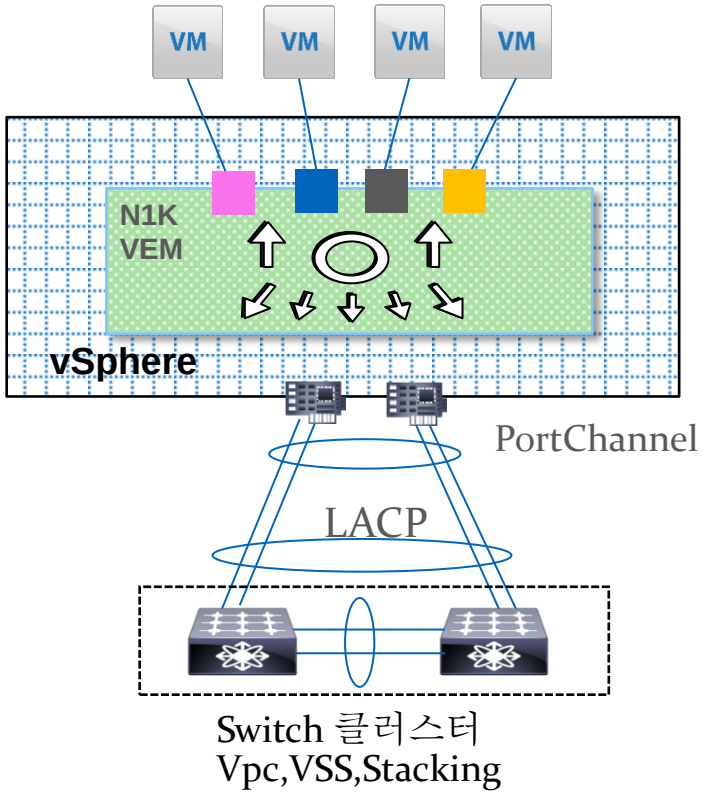


- **VMware**
  - Route based on IP hash
  - Route based on IP hash + LACP (vDS)
- **Hyper-V**
  - Switch Dependent - All Address hash modes
  - Switch Dependent - Hyper-V Port mode [Win2012]
  - Switch Dependent – Dynamic [Win2012 R2]
- **XenServer**
  - LACP with load balancing based on IP and port of source and destination
  - LACP with load balancing based on source MAC address

Nexus 1000V Uplink 구성

Cisco Connect
Seoul, Korea
April 29-30, 2014

Switch dependent - LACP



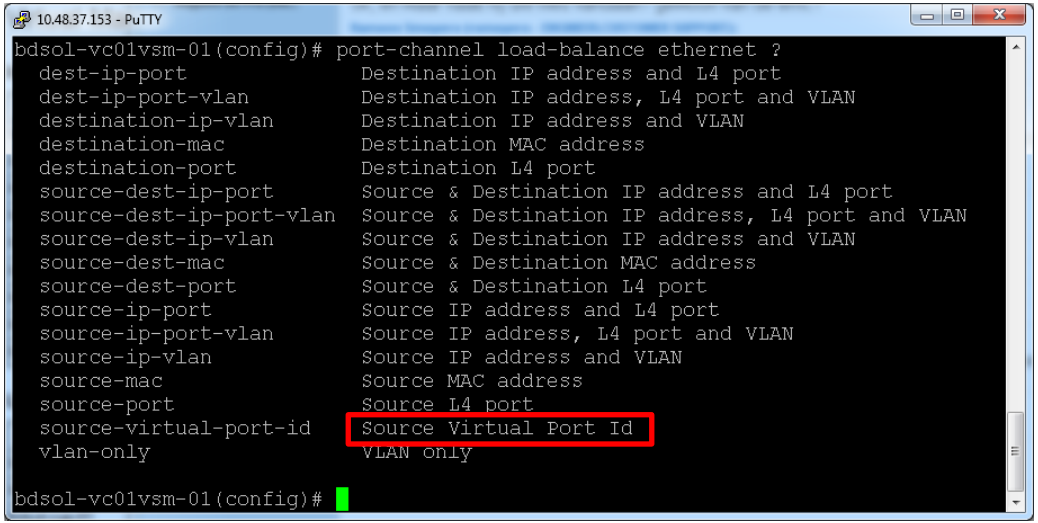
- Switch dependent 구성
- N1K 에서 표준 LACP Port Channel 구성
- 총 17가지의 Hashing 알고리즘
  - Flow-based 해싱
  - Source Based 해싱

    Source Mac Address 해싱 [Default]

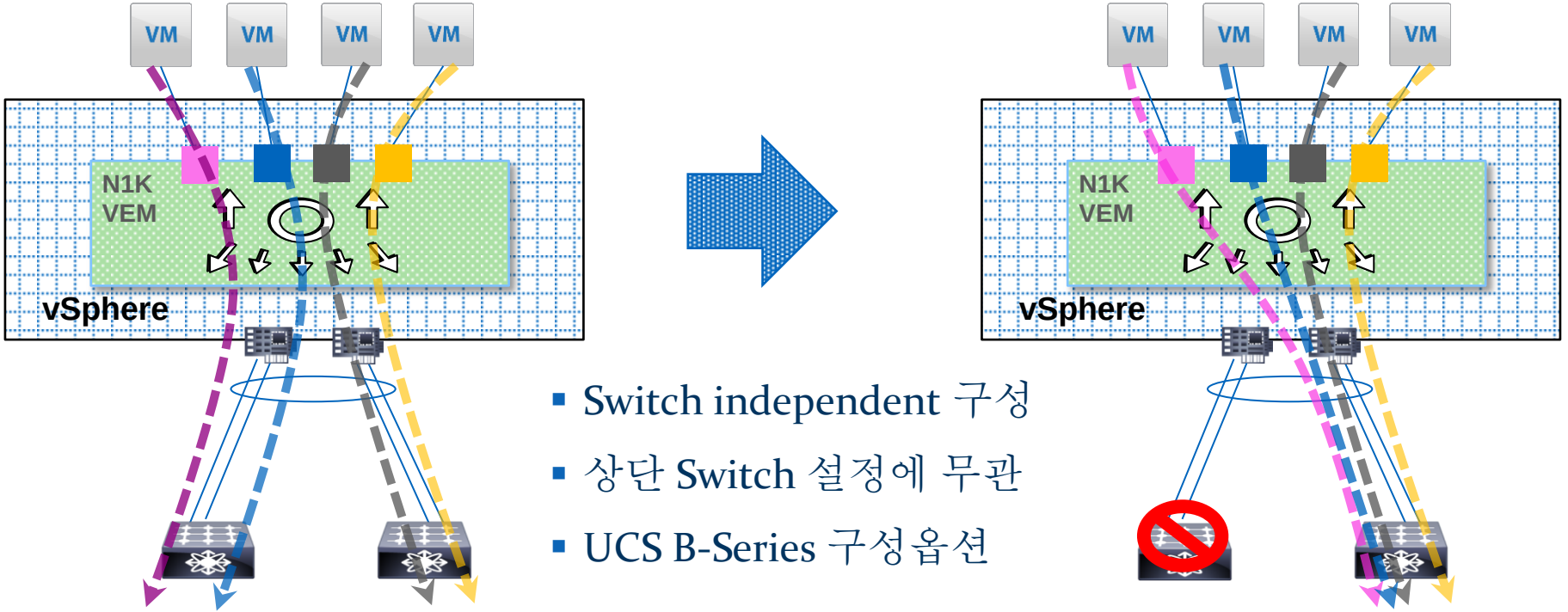
N1k-VSM(Config)#
channel-group auto mode active

Switch dependent - LACP - Load-balancing 옵션

환경에 맞는 다양한 **Load Balancing** 설정옵션을 제공합니다.



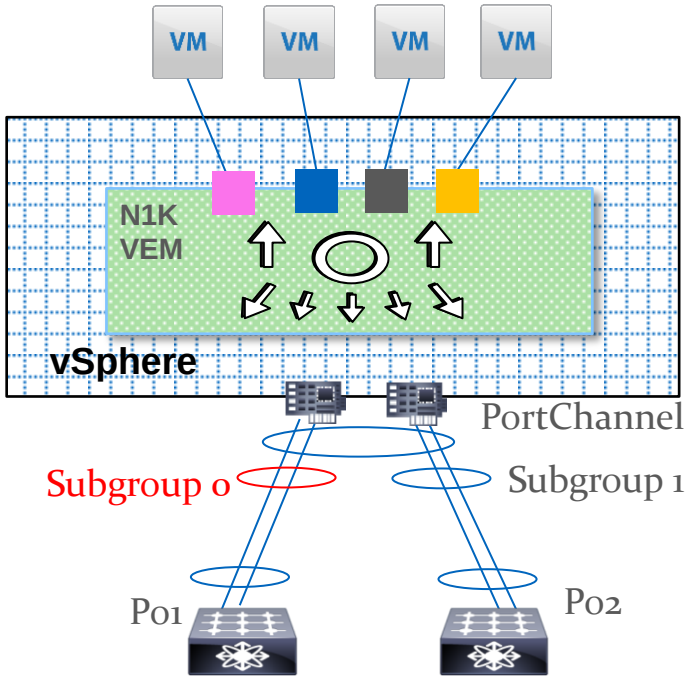
Switch Independent - vPC-HM MAC Address Pinning



- Switch independent 구성
- 상단 Switch 설정에 무관
- UCS B-Series 구성옵션

N1k-VSM(Config)#
channel-group auto mode mac-pinning

vPC-HM Subgroups [S/W Independent/Dependent]



- 두 개의 서로 연결되지 않은 Uplink 스
  위치 구성시 사용

- 하나의 스위치로 가는 Uplink 가 2개
  이상일때, 포트채널 구성필요

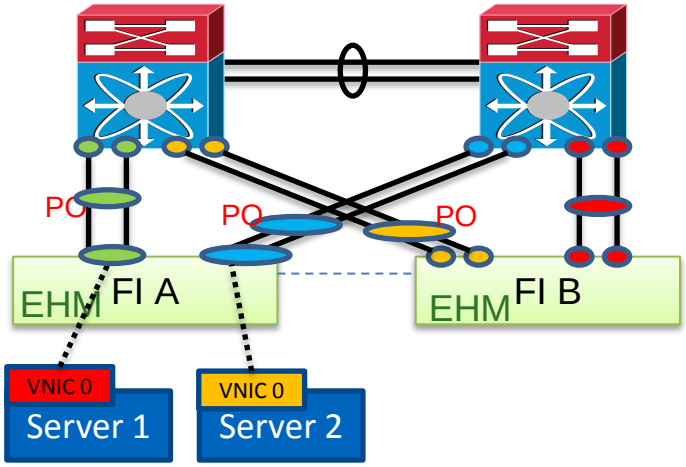
N1k-VSM(Config)#
channel-group auto mode on sub-group cdp

UCS Networking 구성 TIPs

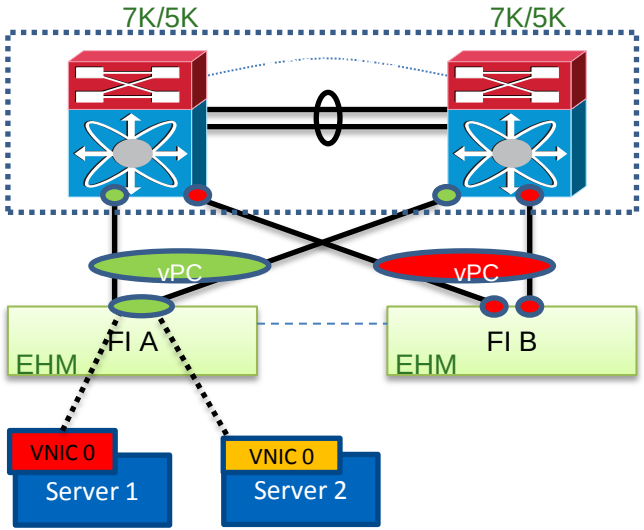
Cisco Connect
Seoul, Korea
April 29-30, 2014

UCS 와 Uplink 스위치 연결

- 상단 vPC 또는 VSS 등의 구성 불가시 개별 업링크 연결 구성
- 상단 스위치에서 vPC 또는 VSS 등의 구성 가능시 포트채널 구성

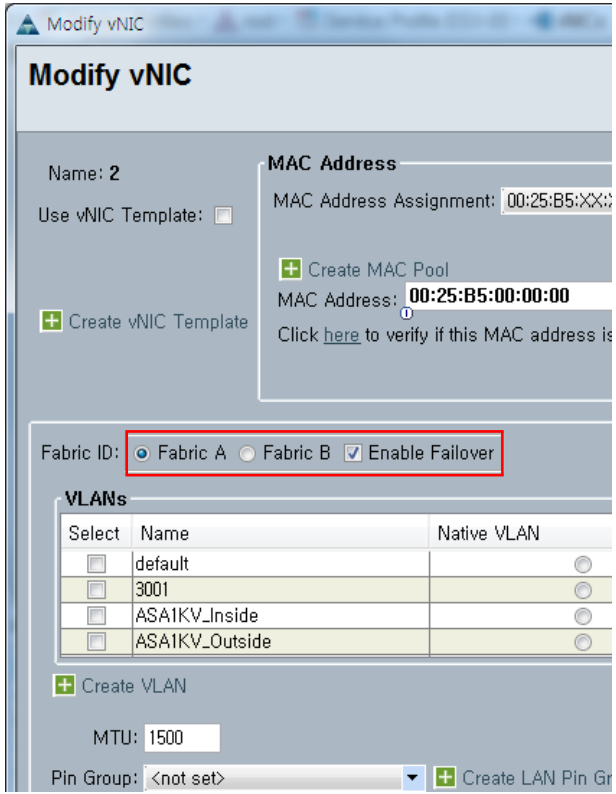


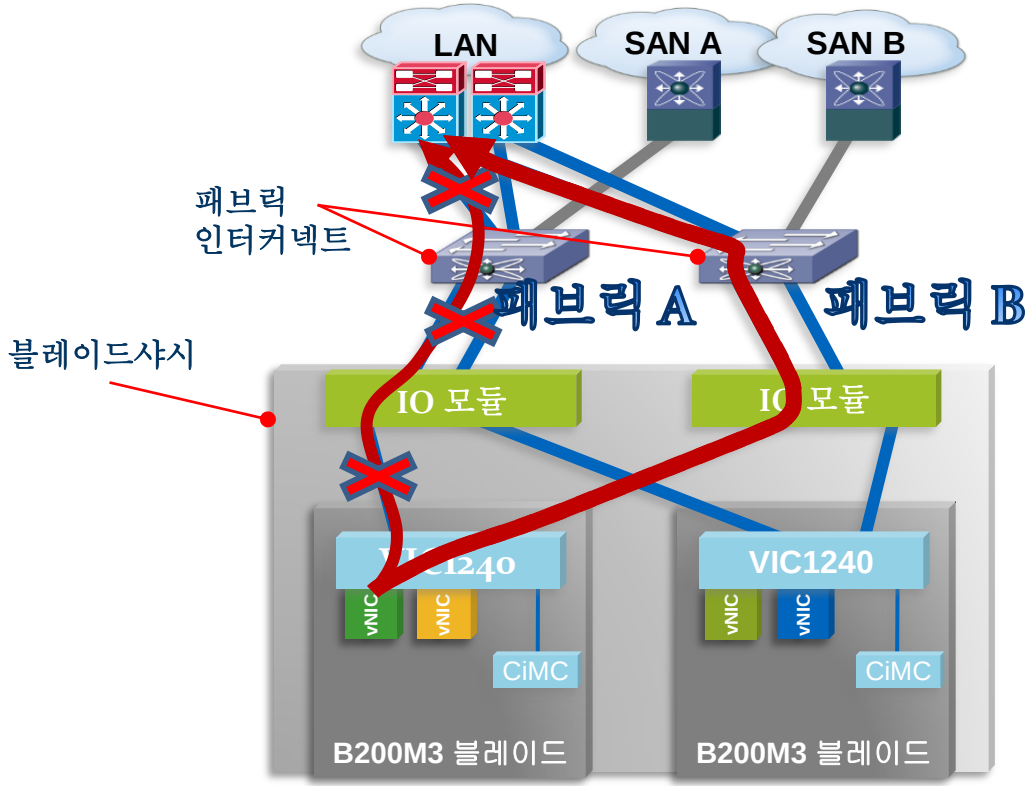
Without vPC/VSS



With vPC/VSS

UCS 서버의 Fabric - Failover 기능





Bare Metal에 운영체제 설치시 사용 추천

UCS FI 에 연결되는 상단 스위치의 Configuration

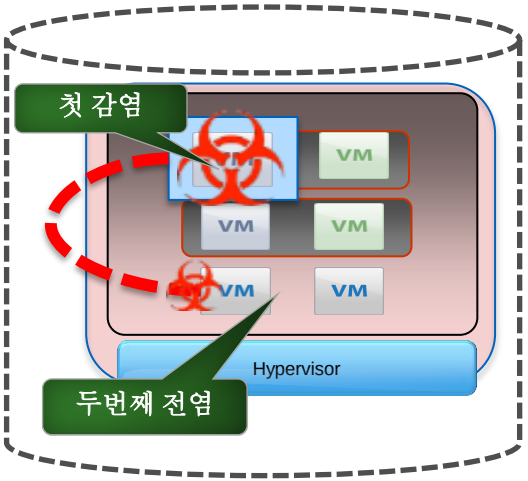
- switchport mode trunk
- switchport trunk allowed vlan (All 혹은 특정 VLAN)
- spanning-tree port type edge trunk (NXOS 사용시)
- spanning-tree portfast trunk (IOS 스위치 사용시)
- spanning-tree bpduguard enable
- spanning-tree bpdufilter enable
- switchport trunk native vlan X (FI 에서 설정된 Native VLAN 과 동일)

가상 환경의 보안

Cisco Connect
Seoul, Korea
April 29-30, 2014

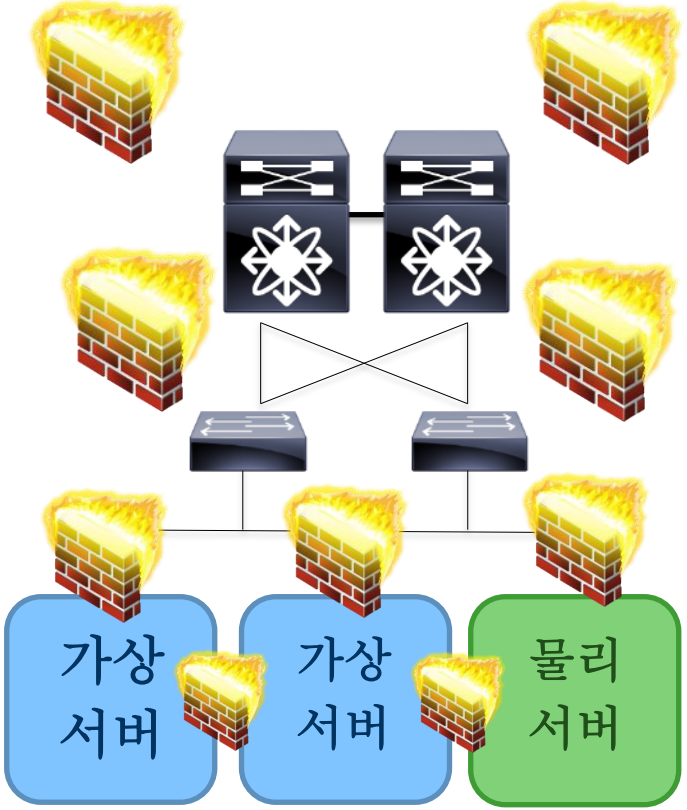
가상환경 보안의 어려움
Policy, Workflow, Operations

- **통합 보안정책 적용**
  - 물리 서버 – 가상서버에 동일 보안정책 적용불가
  - 이동성을 가진 VM 에 정책 적용 불가
- **운영과 관리**
  - 가상머신의 가시성, 일관성 등의 부재
  - 관리 및 효율적인 장애 해결의 어려움
- 역할에 대한 불 명확성
  - 가상 인프라에 관리에 대한 모호한 Ownership 문제
  - 중복된 조직으로 인한 표준화의 어려움
- 인프라와 어플리케이션의 구분 불명확
  - 동일 서버 내에서, 가상서버 및 어플리케이션의 구분
  - 표준 및 비 표준 시스템 간의 구분이 어려움



중앙집중식 혹은 분산형? 아니면 둘다? - FW에 대한 고민

- 방화벽을 어디에 놓을 것인지에 대한 고민
- 가상 호스트 보안을 위한 중앙 집중식 방화벽 구성 -> 가상화 도입초기
- 가장 큰 어려움은 확장성
- L2 레벨에서의 가상 호스트 분리에 대한 요구사항은 어떻게 수용할것인가?
- 가상 호스트의 이동에 따른 보안적용방안은?



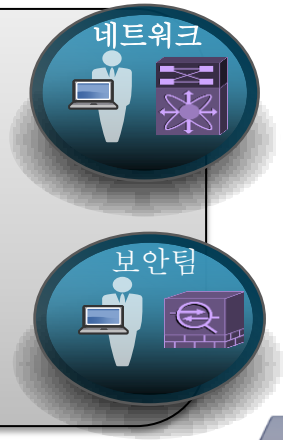
가상 인프라 네트워킹의 핵심
Virtual Switches: Nexus 1000V

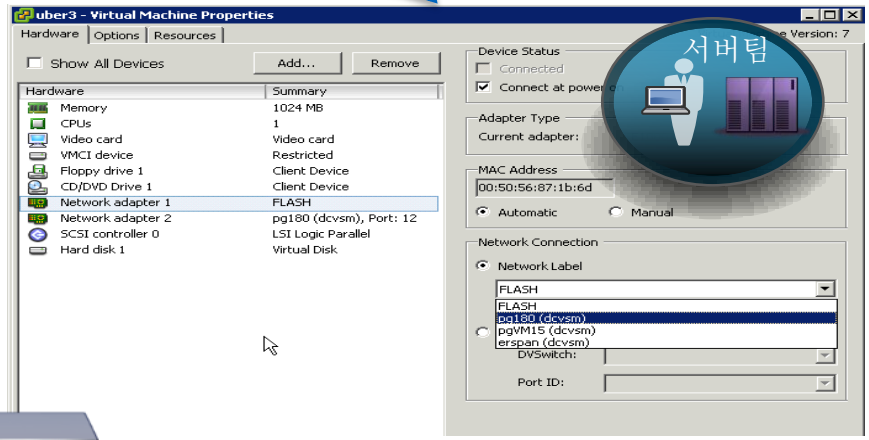
- Port Profile 을 통한 정책 설정/적용으로 기존과 동일한 인프라 운영모델 유지
- VLANs, Private VLANs, Port-based ACL 등을통한 네트워크 분리 및 보안정책 유지
- 기존 네트워크 기능인 ERSPAN 및 NetFlow 기능을 활용하여 가상머신 트래픽 흐름에 대한 가시성 확보

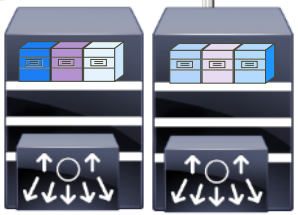
포트 프로파일 [Port Profile]

vCenter API

```
port-profile vm180
  vmware port-group pg180
  switchport mode access
  switchport access vlan 180
  ip flow monitor ESE-flow input
  ip flow monitor ESE-flow output
  no shutdown
  state enabled

interface Vethernet9
  inherit port-profile vm180

interface Vethernet10
  inherit port-profile vm180
```







Nexus 1000V 지원기능:

- ✓ ACLs
- ✓ Quality of Service (QoS)
- ✓ PVLANs
- ✓ Port channels
- ✓ SPAN ports

vMotion
정책의 일관성유지

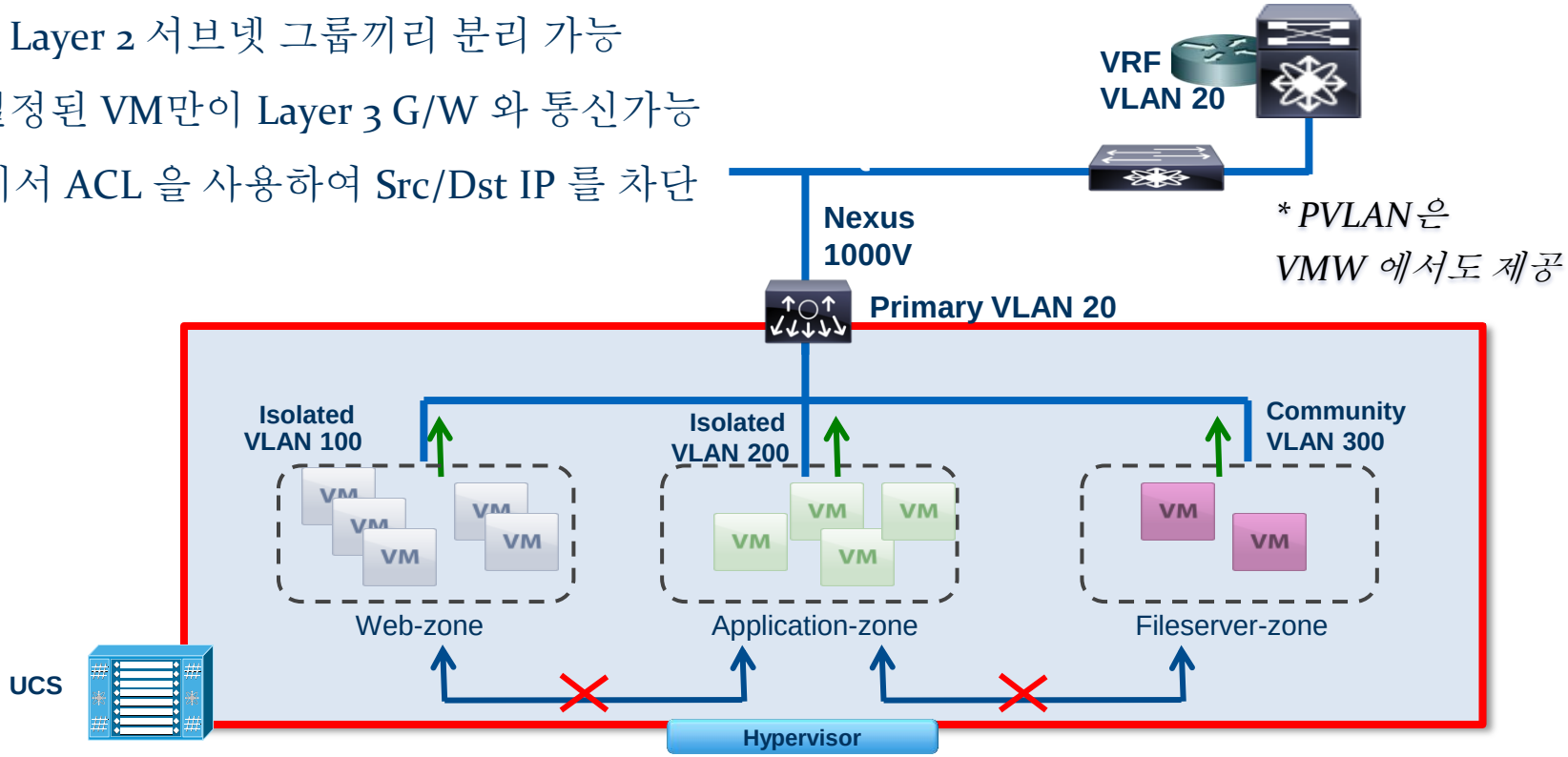
Nexus 1000V 스위치 제공 기능들

1	스위칭	▪ L2 Switching, 802.1Q Tagging, VLAN Segmentation, Rate Limiting (TX) ▪ IGMP Snooping, QoS Marking (COS & DSCP)
2	보안	▪ Virtual Service Domain, Private VLANs w/ local PVLAN Enforcement ▪ Access Control Lists (L2–4 w/ Redirect), Port Security ▪ Dynamic ARP inspection, IP Source Guard, DHCP Snooping
3	프로비져닝	▪ Automated vSwitch Config, Port Profiles, Virtual Center Integration ▪ Optimized NIC Teaming with Virtual Port Channel – Host Mode
4	가시성	▪ VMotion Tracking, ERSPAN, NetFlow v.9, CDP v.2 ▪ VM-Level Interface Statistics
5	관리	▪ Virtual Center VM Provisioning, Cisco Network Provisioning ▪ Cisco CLI, Radius, TACACs, Syslog, SNMP (v.1, 2, 3)

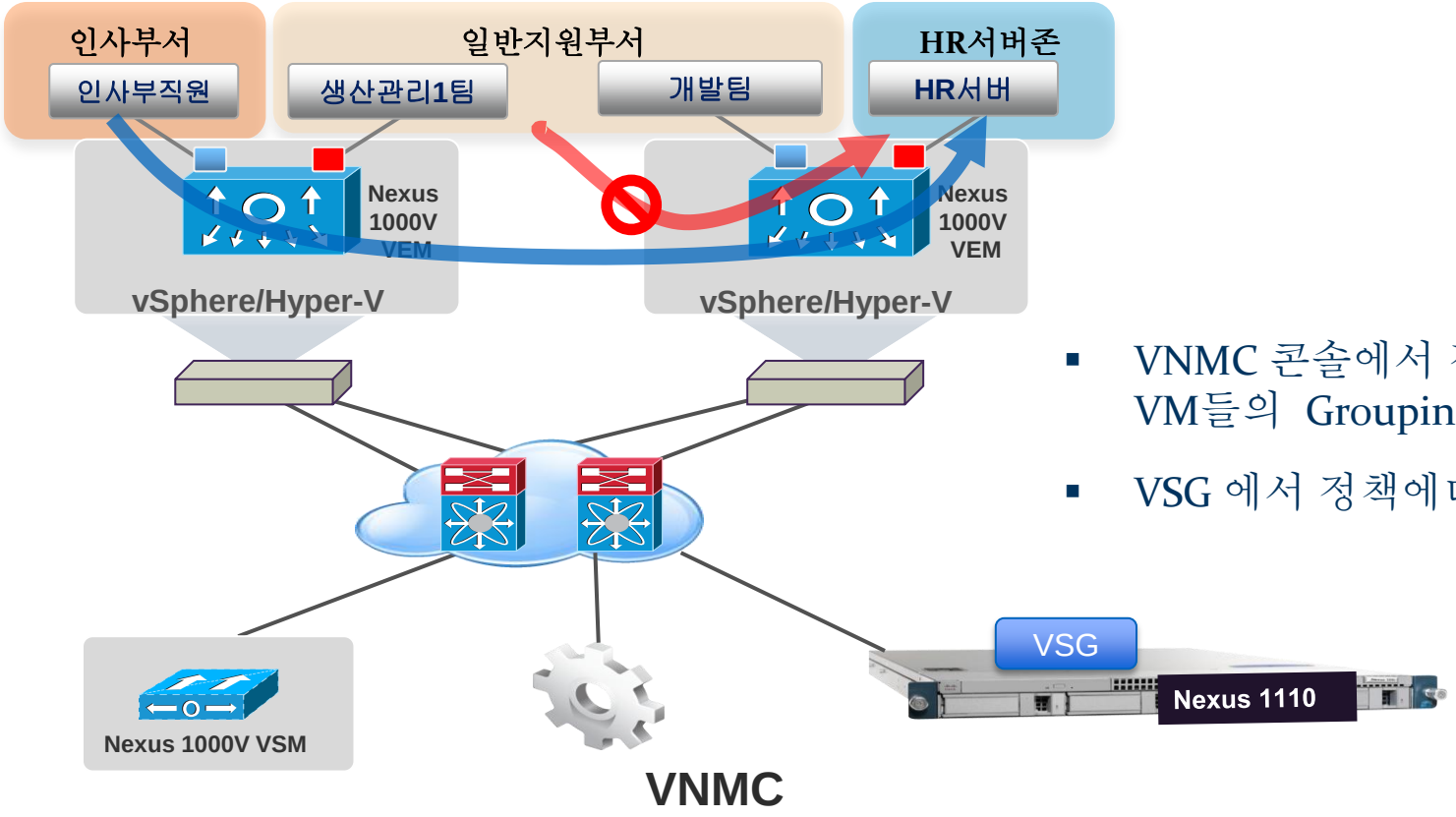
Layer 2 분리

VM 간 분리를 위한 PVLAN

- 동일한 Layer 2 서브넷 그룹끼리 분리 가능
- 오직 설정된 VM만이 Layer 3 G/W 와 통신가능
- G/W 에서 ACL 을 사용하여 Src/Dst IP 를 차단



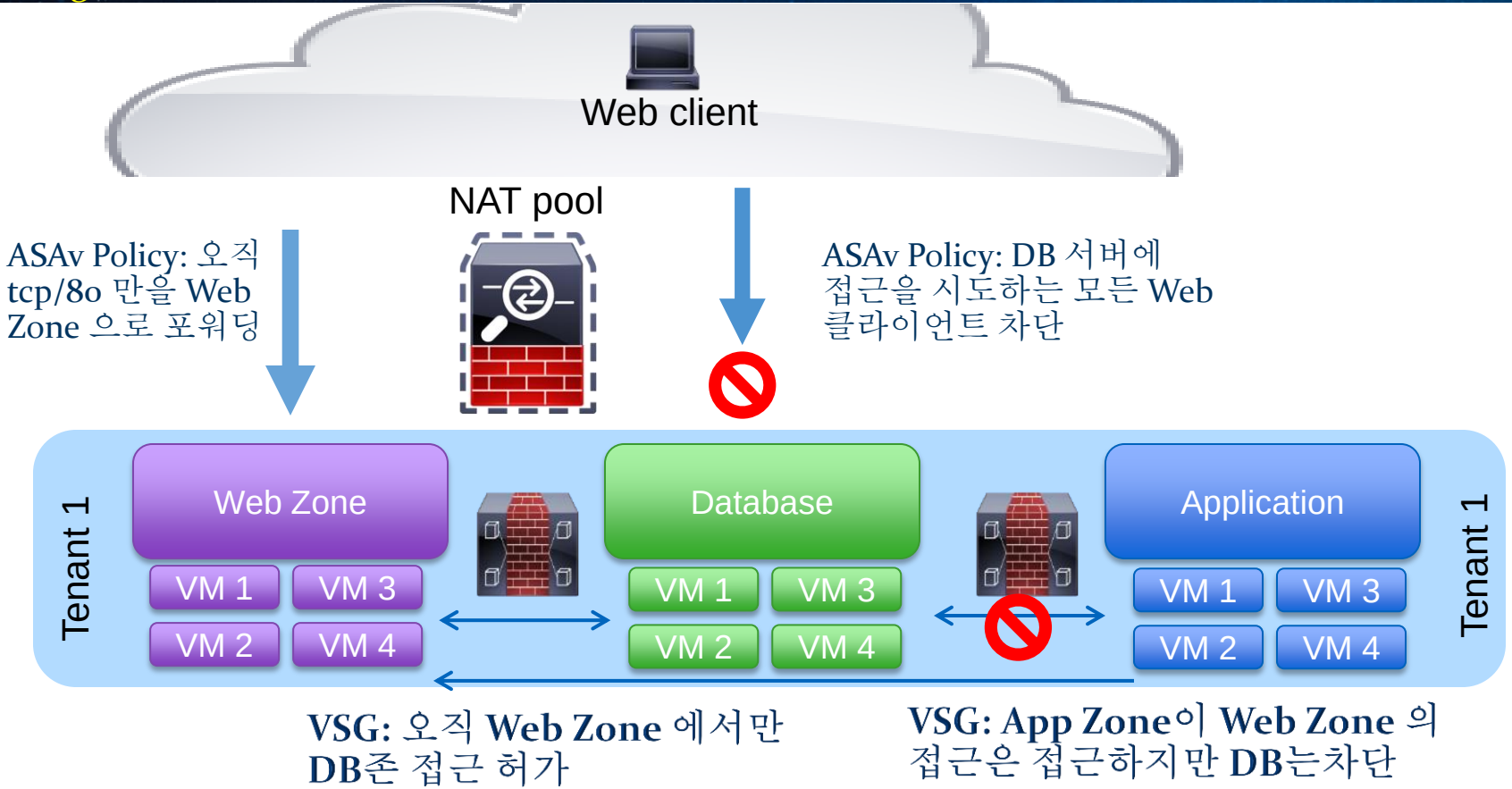
VSG를 이용한 Zone 기반 보안 적용



- VNMC 콘솔에서 같은 보안등급 VM들의 Grouping을 위한 Zone 설정
- VSG 에서 정책에대한 검사 수행

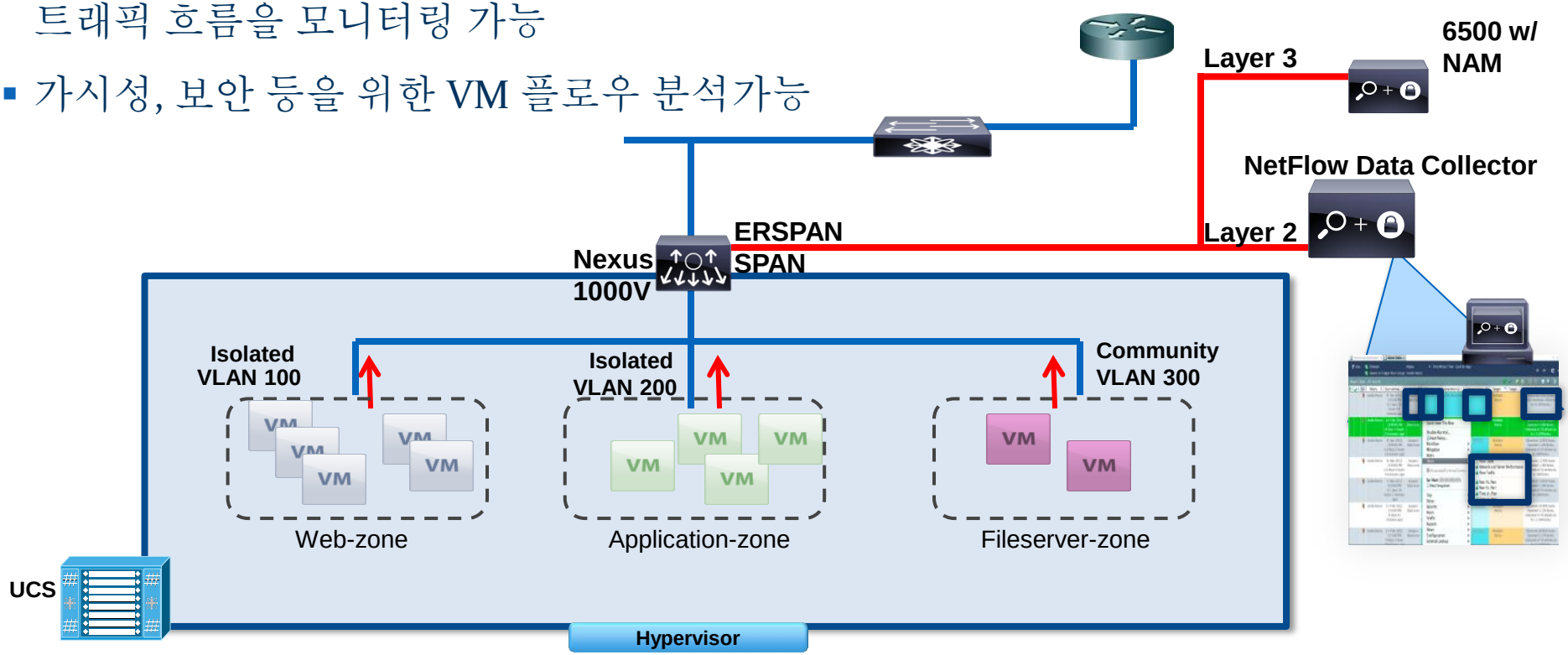
ASAv 와 VSG - 3 Tier 서비스 예시

Edge 방화벽[ASAv] 과 Zone 방화벽[VSG]을 통한 가상 네트워킹 보안



VM 가시성
가상네트워크의 트래픽 관리및 분석을 위한 Netflow

- 가상 스위치의 SPAN포트를 통해 가상머신(VM)의 트래픽 흐름을 모니터링 가능
- 가시성, 보안 등을 위한 VM 플로우 분석가능



마치며..

Cisco Connect
Seoul, Korea
April 29-30, 2014

Summary

- 개별 하이퍼 바이져의 Teaming구성 옵션 및 동작방식
- 하이퍼 바이져의 가상 네트워킹 설정에 따른 Cisco 스위치 설정 방법
  - ✓ Switch Independent
  - ✓ Switch Dependent
- Nexus 1000V의 업링크 설정방식 비교
- UCS 의 가상 네트워킹 구성 Best Practice
- 가상환경에서의 보안


CISCO
TM