

Xebra: Commodity HW 기반 가상 라우터

이신형

고려대학교 운영체제연구실

Contents

- Xebra 소개
 - Why Xebra
 - XEN
 - SR-IOV
- Xebra 구성
 - Network Isolation
 - Performance Isolation
 - V-Channel
- Conclusion

Xebra 소개

Motivation

- Search for Future Internet architecture
- The center piece - router
- How to virtualize router
- Xebra 의 탄생
 - 연구재단 과제 1년차 결과

Xebra approach

- 일반 PC 하드웨어 활용
- SW 플랫폼 기반
- 가상화 기술 적용
 - Open Source 가상화 플랫폼 Xen 사용
- I/O 성능 문제 해결
 - SR-IOV: Single Root IO Virtualization

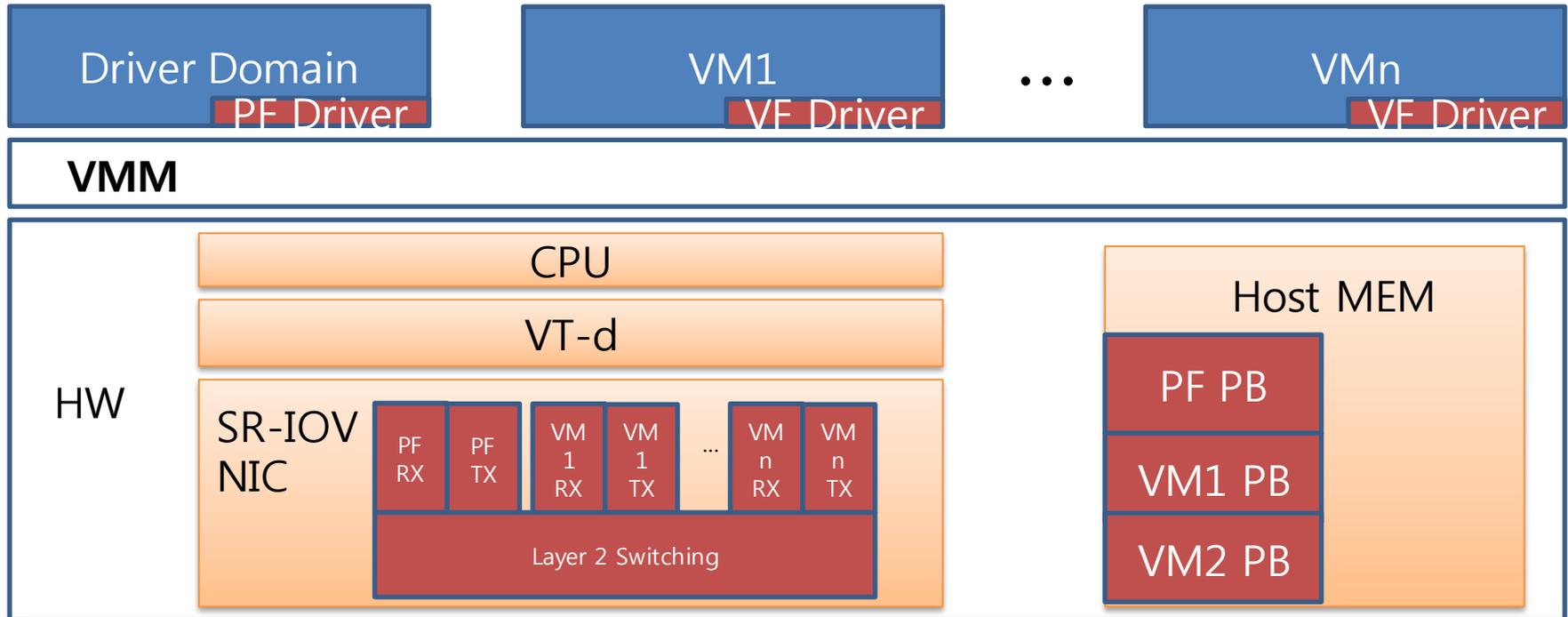
Why Xen

- Open Source Virtual Machine Monitor
 - 자유로운 사용 및 방대한 전문 인력의 지원이 가능
- Hypervisor 기반
 - 커널 기반 가상화는 Isolation을 제대로 지원하지 못함
 - OS 기반 가상화에 비해 성능 저하가 적음
- Para Virtualization
 - 드라이버 성능 최적화로 고성능
- HVM 지원
 - 최신 가상화 기술을 사용하는 HW 가상화 지원
- Commodity Guest OS 지원
 - Windows, Linux, BSD, etc.

SR-IOV 소개

- Single Root IO Virtualization
- PCI I/O 기기의 오버헤드를 최소화 하여 가상화 하기 위한 기술
- Vt-d 기술과 SR-IOV를 지원하는 PCI 주변기기 필요
 - 2009년 이후 출시된 최신의 메인보드는 Vt-d 기술 모두 지원
 - 고성능을 요구하는 I/O 기기(10Gb NIC 등)들의 SR-IOV 지원 확대

SR-IOV 구조



- **PF driver** : Physical Function의 자원을 관리하고 VF를 총괄한다.
- **VF driver** : OS와 디바이스 드라이버간의 인터페이스로서 Virtual Function driver를 설정하고, I/O 명령을 내린다.
- **VT-d** : DMA Remapping 및 Interrupt Remapping 담당

SR-IOV 성능 실험

- 실험 환경
 - Intel Xeon Quad X5650 (2.66Ghz)
 - Xen 4.0, VCPU 4개
 - GuestOS: Linux 2.6.37.1
 - XEN Para Virtualized 1G NIC
 - SR-IOV 1G NIC Virtual Function

Configuration	Pkt bytes	Throughput(pps)	Throughput(Mbps)
Xen PV	64 Bytes	49,910 pps	25 Mbps
SR-IOV	64 Bytes	1,123,831 pps	575 Mbps

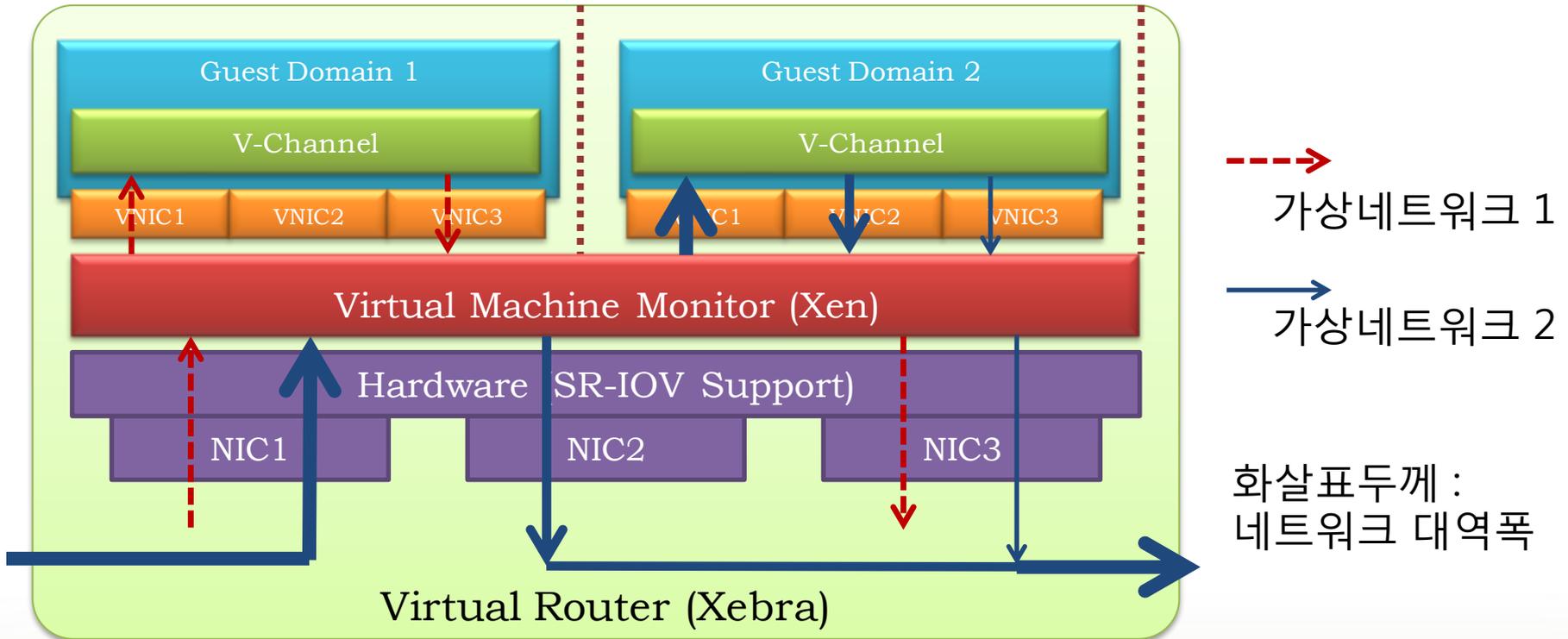
- 25배 이상 성능 향상

Xebra 구성

Xebra의 기능

- Network Isolation으로 네트워크간 간섭 최소화
 - 다양한 종류의 네트워크 동시에 지원 가능
 - 보안에 대한 우려 해소
- Performance Isolation으로 가상 네트워크간 대역폭 보장
 - QoS에 대한 지원
 - 다른 가상 네트워크의 과도한 대역폭 사용에 의한 트래픽 독점 방지
- V-Channel을 통해 링크 본딩 기술 제공
 - 다수의 낮은 대역폭 링크를 모아 높은 대역폭 지원

Xebra 구성도

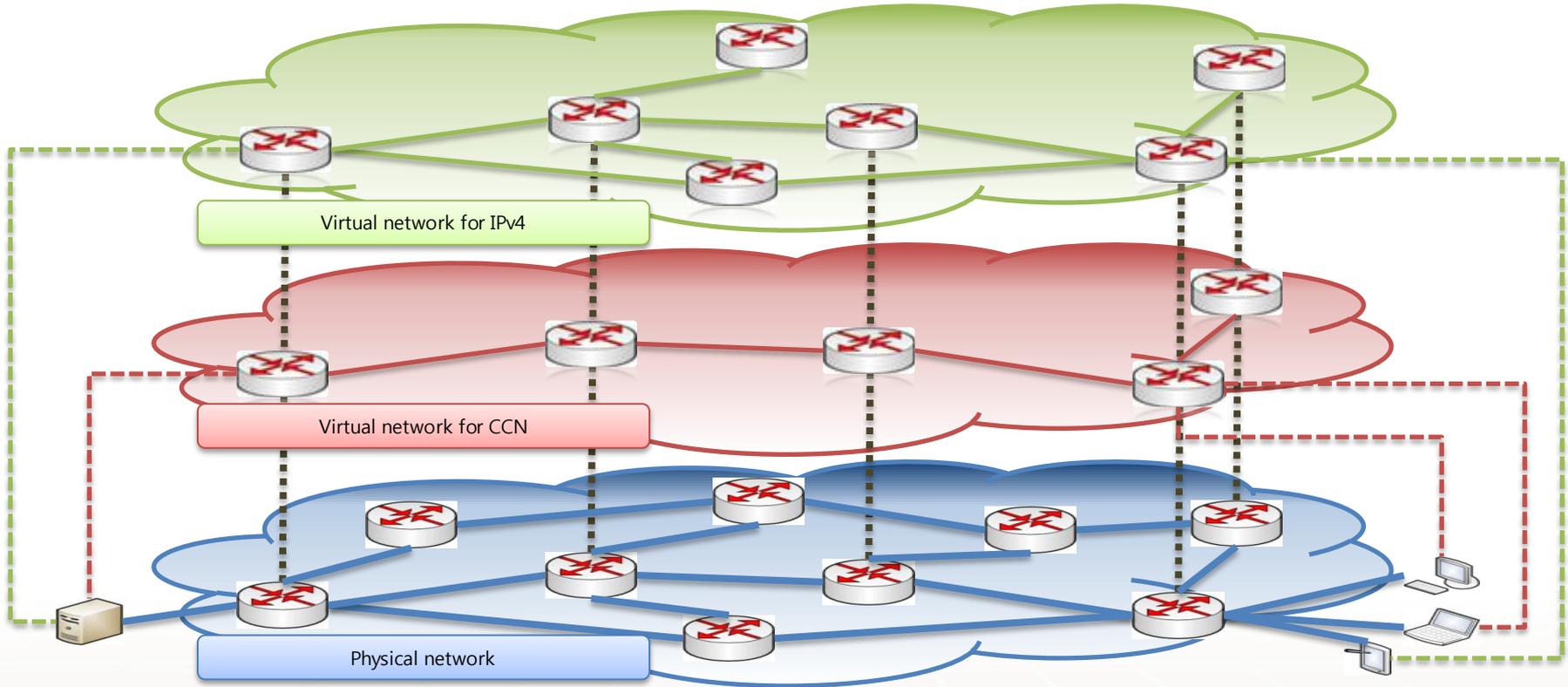


- 각 Guest Domain이 각각 별도의 가상 네트워크 지원 (Network Isolation)
- 각 도메인간 네트워크 대역폭 보장 (Performance Isolation)
- 필요시 다수의 링크를 모아 하나의 링크처럼 사용 가능 (V-Channel)

Network Isolation

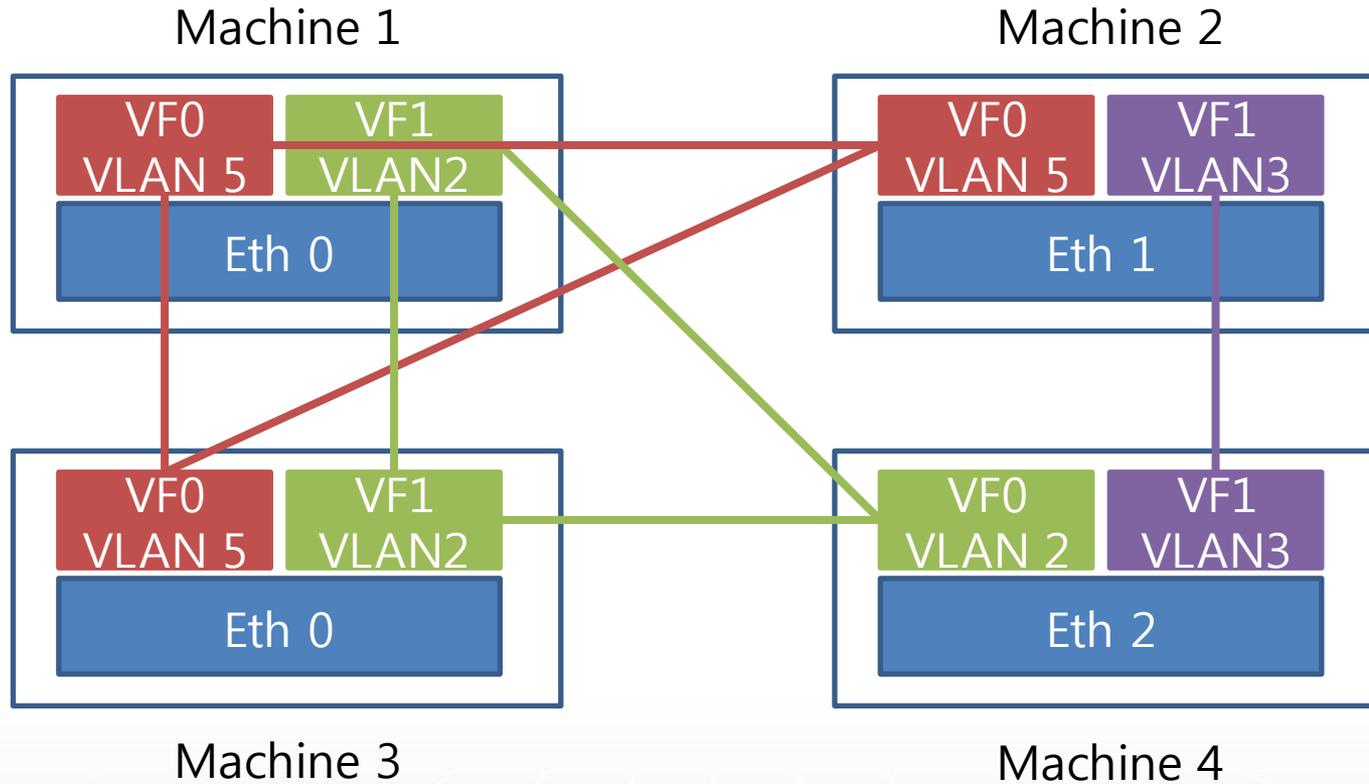
- 가상 네트워크간 완전히 고립되어 서로 영향을 주지 않도록 지원
- 서로 다른 주소체계 및 라우팅 프로토콜 지원
- 가상네트워크 ID를 통해 구분
 - VLAN ID(802.1q)를 가상네트워크 ID로 사용

Network Isolation



하나의 Physical Network를 가상화하여 IPv4를 지원하는 가상 네트워크와 CCN을 지원하는 가상 네트워크를 생성하여 통신 가능

VLAN ID를 이용한 가상 네트워크 구축

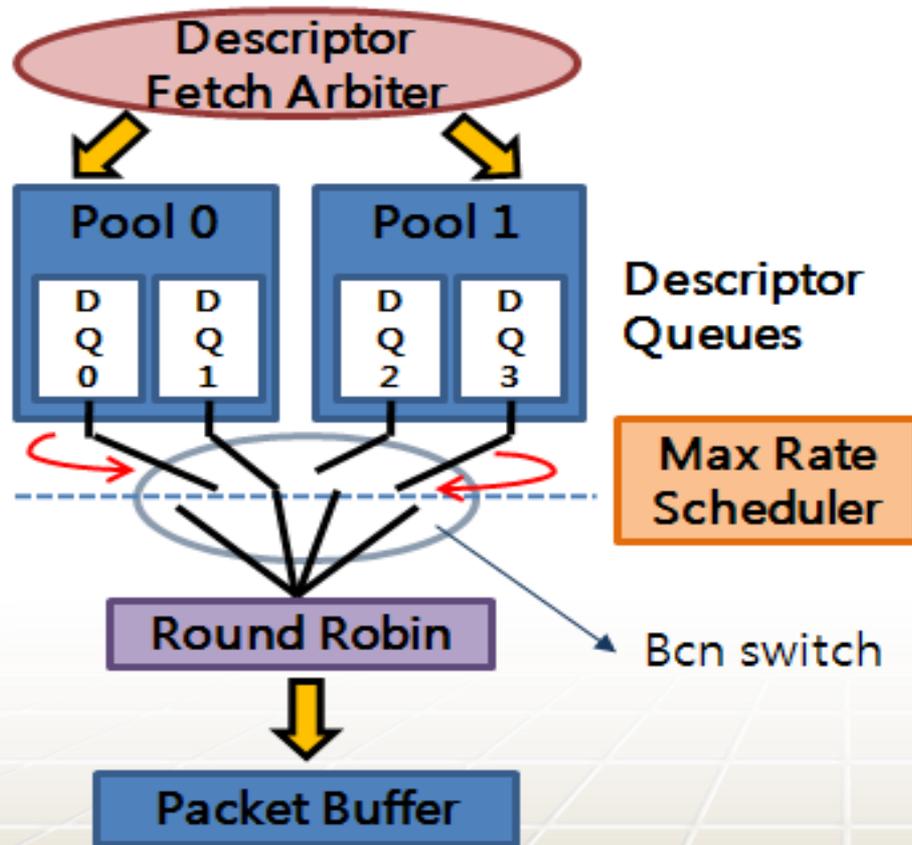


- VLAN ID를 설정하여 VF 도메인간 Isolation을 지원
- VLAN ID 2, 3, 5의 가상 네트워크 3개 구축

Performance Isolation

- Performance Isolation
 - 각각의 네트워크 토폴로지가 다양한 Contents를 제공
 - 따라서 이를 지원하기 위해서 Network 사이에 대역폭을 동적으로 조절하는 기술이 필요.
- Performance Isolation 지원 기법
 - BCN : Backward Congestion Notification
 - VF에게 할당된 Cap을 기준으로 대역폭을 조절
 - VF가 할당된 Cap을 넘게 되면 BCN Switch를 off
 - 할당된 Cap 이상으로 Bandwidth를 사용할 수 없다.

BCN 스케줄링 구조



- Descriptor Fetch Arbiter : 시스템의 메모리의 링버퍼로부터 전송할 패킷을 지정된 Pool에 할당하는 역할
- Pool : NIC상에 존재하는 VF
- Descriptor Queue : VF내에 Packet을 전송하는 Queue,
- Max Rate Scheduler : 할당해준 Cap 이하로 대역폭을 할당, 할당된 값 이상이 될 경우 Switch off

BCN 실험결과

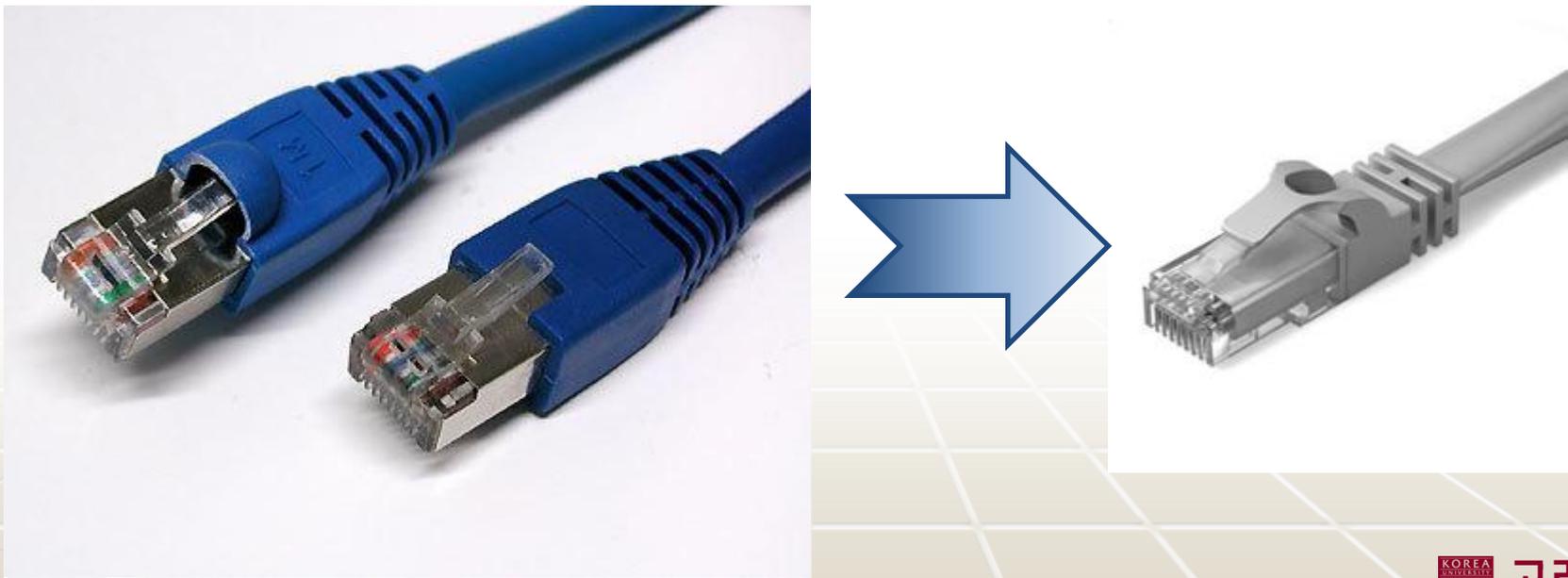
총합: 10Gbps 단위 : Mbps

VM1 bcn Rate	VM2 bcn Rate	VM1 BW	VM2 BW	BW Sum
-	-	4826	4925	9801
3000	7000	2986	6891	9877
4000	6000	3873	5987	9860
7000	7000	4825	4968	9763
8000	6000	6151	4009	10160

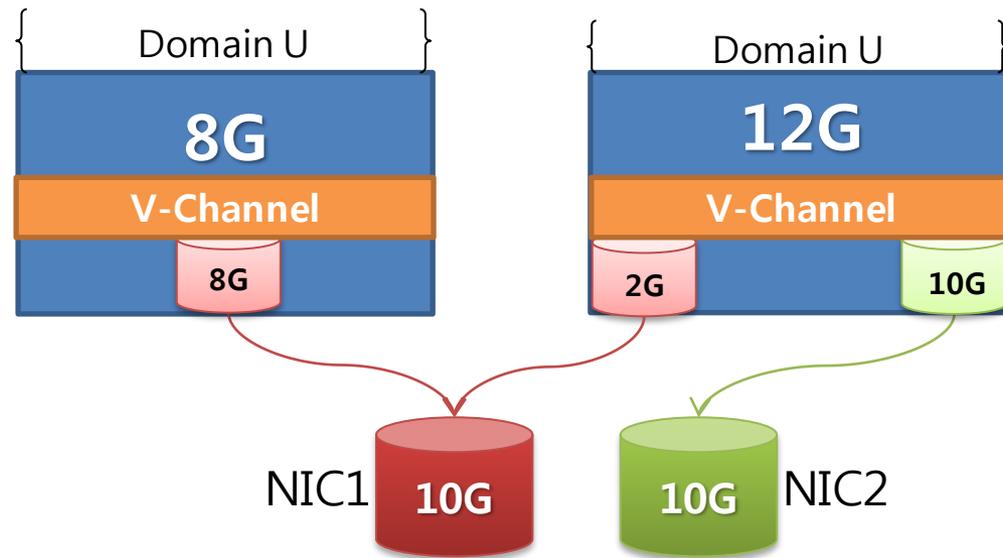
- 10Gb 이더넷 카드 사용 (총합: 10000Mbps)
- 두개의 가상머신에서 BCN을 사용하여 BW 설정
- 각 가상머신에서 설정한 BW대로 실제 측정됨
- 설정한 BW의 합이 10Gb를 넘어갈 경우 각 비율에 맞게 BW 분배

V-Channel 소개

- Linux의 Bonding Module
 - Multiple NIC을 하나의 Virtual NIC으로 사용
 - Fault-tolerant, Load Balancing 기능 제공

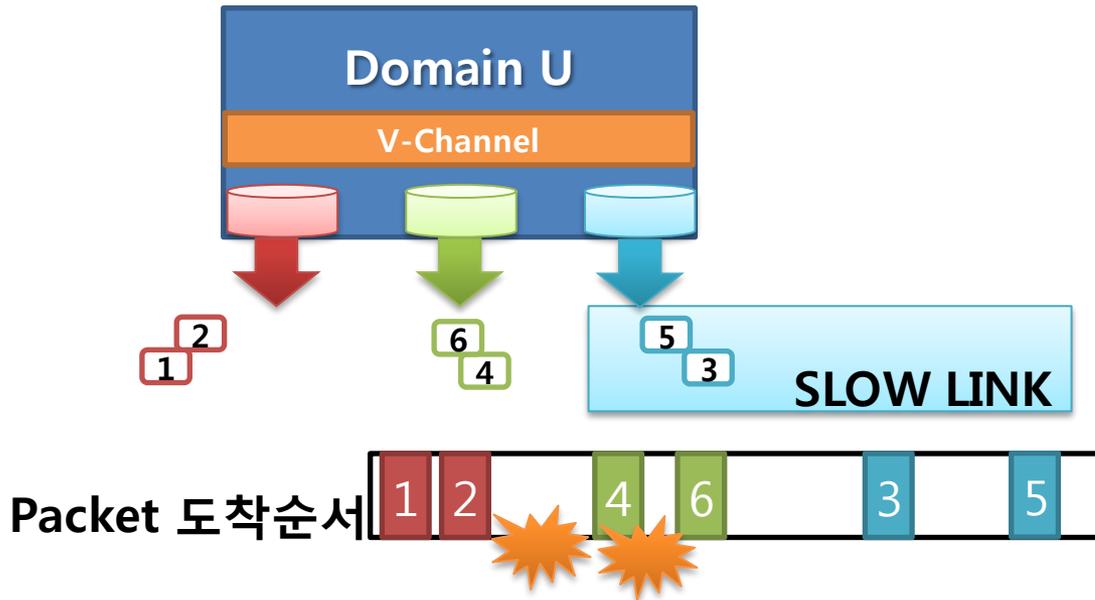


V-Channel 동작 방식



- 1개의 NIC이 수용할 수 있는 대역폭을 초과
 - 다른 NIC의 으로 Balancing 하여 통합 관리 필요
- 2가지의 문제점
 - Out-of-order
 - Load balancing

Out-of-order

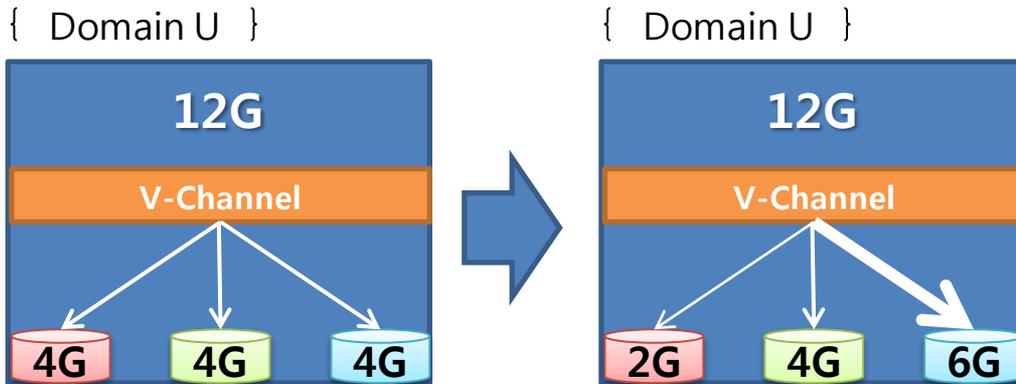


- V-Channel 사용시 링크 속도의 차이에 따라 패킷의 도착 순서가 바뀔 수 있음
- 패킷 도착 순서가 바뀌므로 인해 필요 없는 재전송 등이 일어남

• 해결 방법

- Source 및 Destination의 Address와 Port를 전송할 NIC 선택에 반영
- 동일한 Flow는 동일한 NIC에서 처리함

Load balancing



- 기존 기법은 동일한 대역폭으로 패킷 분배
- 각 NIC마다 할당된 대역폭을 고려한 패킷 분배가 필요

- 해결 방법

- NIC 선택 시 NIC의 Bandwidth 비율 만큼의 Flow 할당

Conclusion

Conclusion

- 미래 인터넷의 가상 네트워크를 지원하기 위한 가상 라우터 구조 제안
- Xen을 사용하여 성능 향상 및 신기술을 빠르게 접목
- SR-IOV 기술을 이용하여 성능 한계 극복
- Network Isolation 기능을 VLAN ID를 통해 구현
- BCN 기술을 사용하여 Performance Isolation 기능 구현
- 다수의 링크를 하나로 묶어 효율적으로 사용하기 위해 V-Channel 기능 구현
- 홈페이지 : <http://os.korea.ac.kr/xebra>
 - 소스코드 다운로드 가능
 - 매뉴얼 제공

Xebra HW 사양

- Intel Vt-d가 지원되는 X86 호환 CPU (i5 이상 모두 지원, Core2Duo중 일부 지원)
- Intel Vt-d가 지원되는 메인보드 (i5 이상 CPU 지원하는 모든 Intel 메인보드, Core2Duo CPU 지원하는 일부 Intel 메인보드)
- 메모리 4GB 이상
- HDD 100GB 이상
- 인텔 82599 10Gb 이더넷 카드

Future Work

- Linux 네트워크 프로토콜 스택의 성능 저하
→ 네트워크 성능 향상을 위한 Linux 최적화
- 다양한 네트워크 기술을 사용하는 가상 네트워크 실제 동작
- GPGPU를 사용하여 가상 라우터 성능 극대화
- 이동 단말을 위한 mobile router 가상화 지원

Q/A