김해중
KBS 보도기술국

실무 네트워크 Design - 10 : 네트워크 가상화 2 (VSS, VPC)

지난 시간에 자세히 설명한 cisco의 stackwise와 juniper의 virtual-chassis는 여러 대의 switch를 묶어서 1대의 switch로 통합하는 가상화 기술이었습니다. 이번 시간에 설명할 VSS와 VPC는 2대의 switch만 1대의 switch로 통합할 수 있는 가상화 기술입니다.

현재 VSS와 VPC는 많은 회사의 backbone switch에서 많이 사용되고 있습니다. VSS는 현재 cisco의 catalyst 6800, 6500, 4500X에서 지원을 하고 있으며, VPC는 cisco의 Nexus 5K, 6K, 7K, 9K에서 지원을 합니다. 과거에는 backbone switch로 catalyst 6500이 가장 널리 사용되었으나, 현재는 점차 cisco의 nexus series로 전환하고 있는 추세입니다.

사실 이번 시간에 설명드릴 VSS와 VPC는 많은 공통점을 가지고 있습니다. 그래서 먼저 설명할 VSS는 자세히 다루고, VPC는 VSS와의 차이점만을 간단히 정리해 보겠습니다.

VSS(virtual switching system)

개념

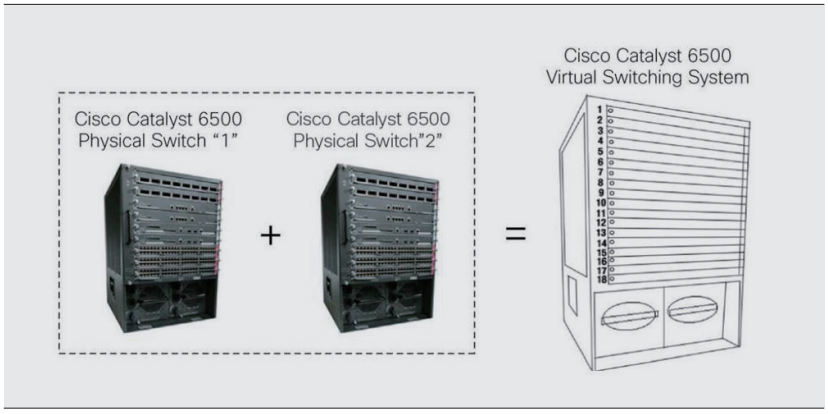


그림 1. VSS는 물리적인 2대의 switch를 1대의 virtual switch로 통합합니다

VSS는 2대의 switch를 1개의 가상 switch로 통합하는 기술로써 이를 통해 loop-free 환경을 구현할 수 있습니다. 현재 VSS는 cisco의 catalyst 6800, 6500, 4500X에서 지원합니다.

특징

VSS pair switch 간의 연결 link를 VSL(virtual switch link)이라 하며, 1 link만 BPDU나 port-channel 같은 control-traffic 전달 용도로 사용되고, 나머지 link는 data 전달 용도로 사용됩니다. 기본적으로 control traffic 용도로 사용하는 link에는 높은 우선순위를 부여하여 처리합니다.

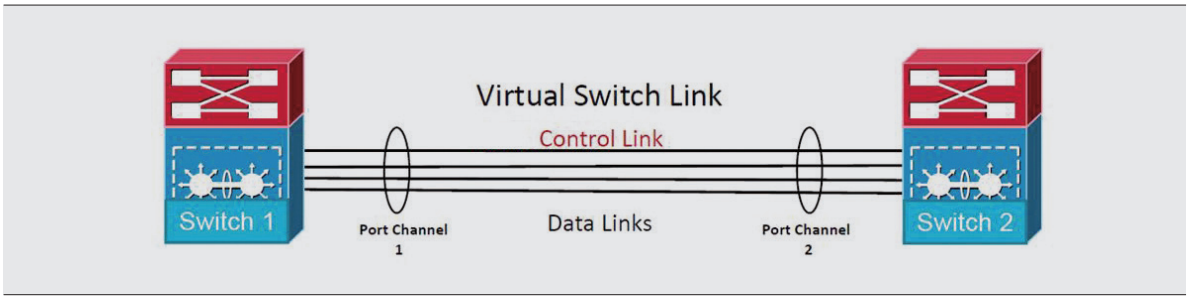


그림 2. VSL(virtual switch lik)는 pair switch 간의 link이다

VSL을 통과하는 모든 traffic에는 32byte VSH(virtual switch header)를 추가해서 보내며, VSH는 COS 정보, VLAN 정보, 입구 switch port, 출구 switch port 정보를 포함합니다.

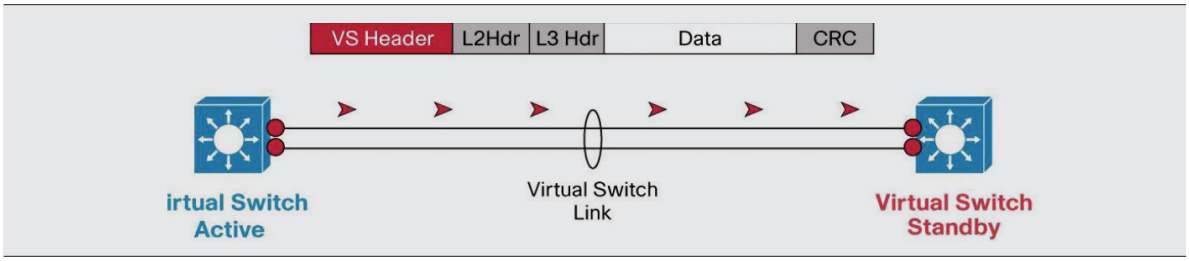


그림 3. VSL을 통과하는 모든 traffic에는 VSH가 추가 됩니다

VSS에 참여하는 2대의 switch가 동시에 VSS를 시작했다면, priority가 높은 장비가 active 역할을 하고, 만약 다른 시간에 VSS를 시작했다면, 먼저 VSS를 시작한 switch가 active 역할을 합니다.

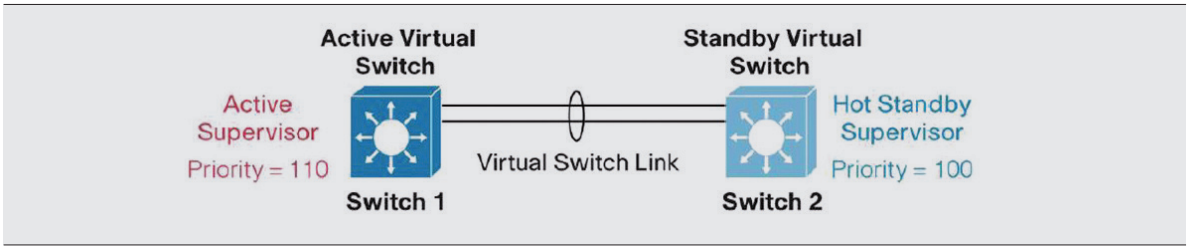


그림 4. priority가 높은 switch가 active virtual switch 역할을 합니다

VSS를 만들려면, [그림 5(a)]처럼 같은 VSS domain 안에 있어야 합니다. 만약 [그림 5(b)]처럼 여러 개의 VSS를 사용하는 환경이라면, 각 domain-ID는 다르게 세팅해야 하며, domain-ID는 1~255값을 가질 수 있습니다.

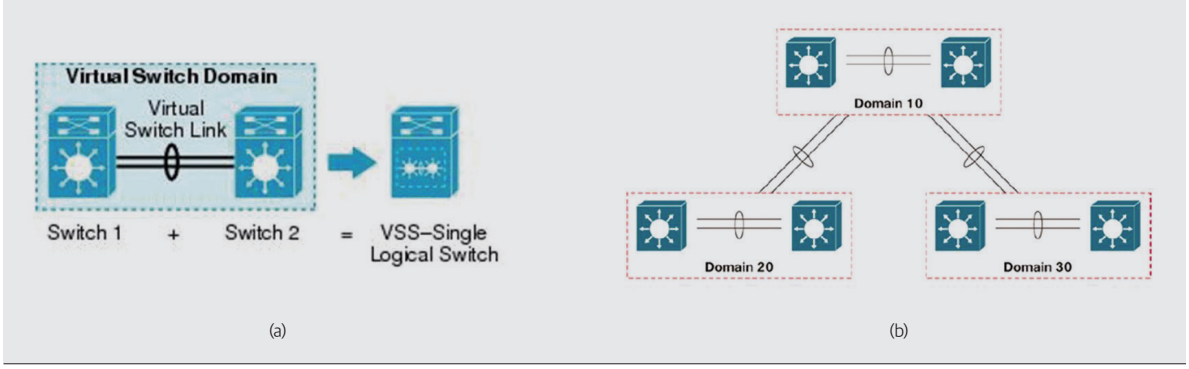


그림 5. VSS를 만드는 pair switch는 동일한 domain-ID를 사용해야 한다

김해중
KBS 보도기술국

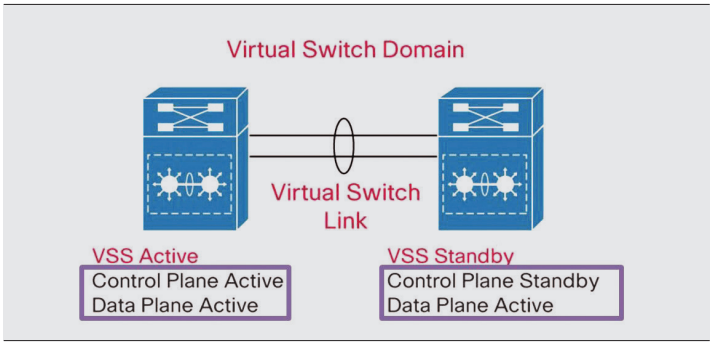


그림 6. control-plane은 active/standby로 동작하고, data-plane은 active/active로 동작합니다

control-plane에서 처리하는 routing protocol 제어 및 관리 제어는 active/standby로 동작을 합니다. 즉 active switch가 처리를 하고, standby switch는 동기화하게 됩니다. standby switch는 active switch에 장애가 발생했을 때만 active 역할을 맡게 됩니다. 하지만 data-plane에서 hardware로 처리하는 패킷 전송은 active/active로 동작합니다.

최초 VSS가 active 상태가 되면, active 역할을 하는 switch의 system mac address를 VSS의 system mac address로 사용하며, 이 주소를 모든 L3 interface의 mac 주소로 사용합니다. 참고로 system mac address는 공장에서 세팅된 EEPROM memory에서 할당 받습니다.

통신의 안전성을 보장하기 위해, active switch에 장애가 발생할지라도, 장애가 발생한 active switch의 mac 주소를 계속 사용하도록 되어 있습니다. 참고로 지난시간을 간단히 복습해보면, cisco stackwise는 "stack-mac persist timer 0"를 입력했을 때만, 장애가 발생한 master switch의 mac 주소를 그대로 사용했습니다.

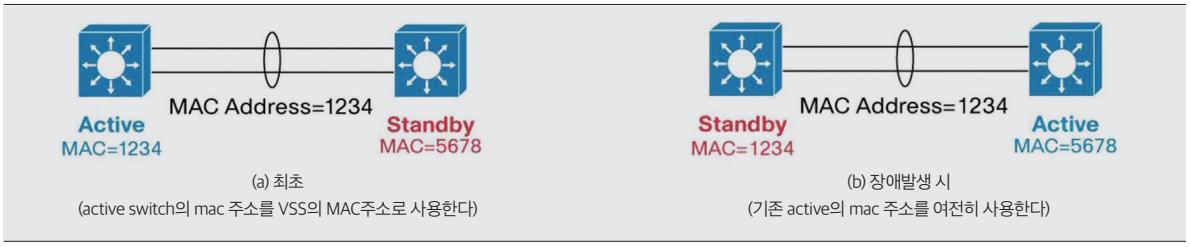


그림 7. active 역할을 하는 switch의 mac address를 VSS의 system mac address로 사용합니다

만약 VSS pair switch 모두 장애가 발생해서, 장애 전 standby switch가 먼저 부팅되어서 active 역할을 맡게 되면, 해당 switch의 mac-address를 VSS의 mac-address로 사용하게 됩니다.

이제는 다른 switch에 이것을 알리기 위해 GARP(gratuitous ARP)를 보내야 할 뿐만 아니라, VSS의 모든 L3 interface의 mac address를 변경해야 되어, 일시적인 traffic 장애가 발생합니다. 이런 문제점을 해결하기 위해, 공장에서 세팅된 mac-address가 아닌, virtual mac address를 사용할 수 있으며 이것을 권장합니다.

virtual mac address는 VSS의 domain-id와 연관성을 가지며, 기본적으로 0008.e3ff.fc00부터 0008.e3ff.ffff 대역을 사용합니다.

```
VSS(config-vs-domain)#switch virtual domain 10
VSS(config-vs-domain)#mac-address use-virtual
Configured Router mac address is different from operational value. Change will
take effect after config is saved and the entire Virtual Switching System
(Active and Standby) is reloaded.
```

그림 8. VSS의 virtual mac address 세팅

```
VSS# show interface vlan 1
Vlan1 is up, line protocol is up
  Hardware is EtherSVI, address is 0008.e3ff.fc0a (bia 0008.e3ff.fc0a)
```

Active sw의 mac

그림 9. 세팅한 virtual mac address가 SVI interface에 할당되어 있다

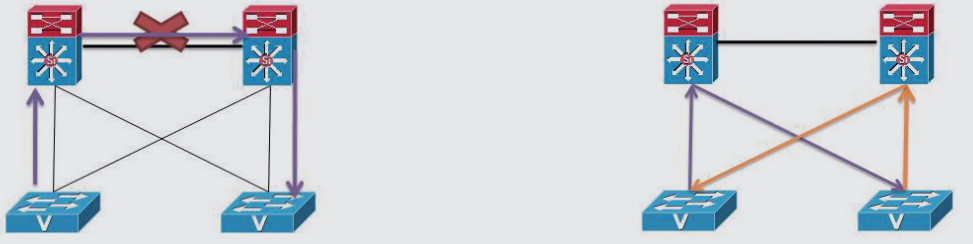
VSS를 사용하면 1개의 router mac 주소와 1개의 interface를 공유하기에, FHRP(HSRP/VRRP)를 사용하지 않아도 되는 장점이 있습니다.

(a) VSS 사용 전(FHRP가 필요하다)

(b) VSS 사용 후(FHRP가 불 필요하다)

그림 10. VSS 사용 전/후의 FHRP비교

평상시에 VSL은 unicast data traffic 전달용도로 사용 안하며, 오로지 control, multicast, broadcast traffic 전달 용도로만 사용합니다. 즉 [그림 11(b)]과 같이, unicast traffic을 수신한 switch가 처리하며, 이것을 local switching이라고 합니다.



(a) VSL link는 unicast data-traffic 전달용도로 사용 안한다.

(b) unicast traffic을 수신한 switch가 처리한다.

그림 11. VSS의 local switching

local switching 구현을 위해, [그림 12]와 같이 기존에 사용하던 이더 채널의 load-balancing 방식을 변형하였습니다. traffic을 수신한 switch가 local switching 하기 위해, hash 결과가 local switch의 link로 되어 있습니다.

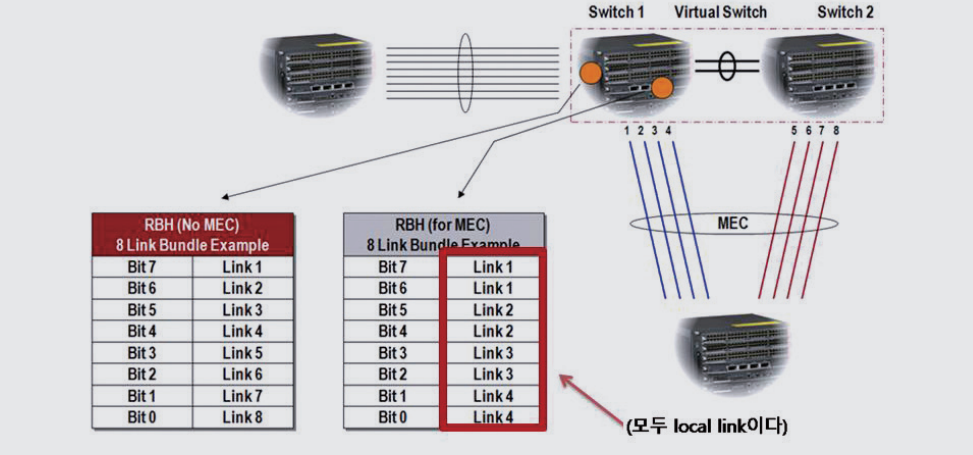


그림 12. VSS는 local switching을 위해, 기존 이더채널의 동작방식을 변형하였다

하지만 link 장애가 발생했을 때는 [그림 13]과 같이 VSL을 사용합니다.

김해중
KBS 보도기술국

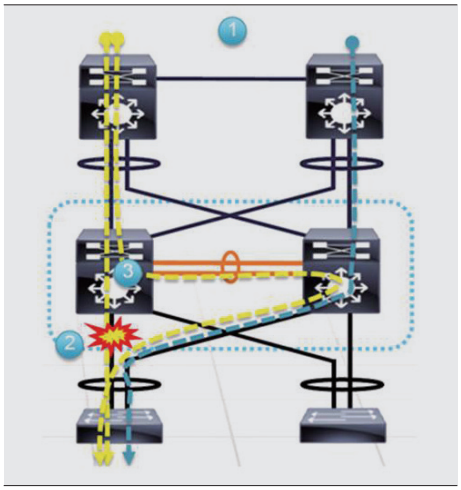


그림 13. link 장애가 발생했을 때만, VSL 사용을 허락한다

만약 switch에 장애는 없는데, VSL에만 장애가 발생하면 어떤 문제가 발생할까요? 이제는 더 이상 VSL로 control 정보를 주고 받지 못하니, VSS에 참여한 스위치들은 상대방 switch에 장애가 발생했다고 생각하고, 자신이 active switch가 됩니다. 이때 양쪽 switch 모두 active가 되기에, 이 현상을 dual-active라고 합니다.

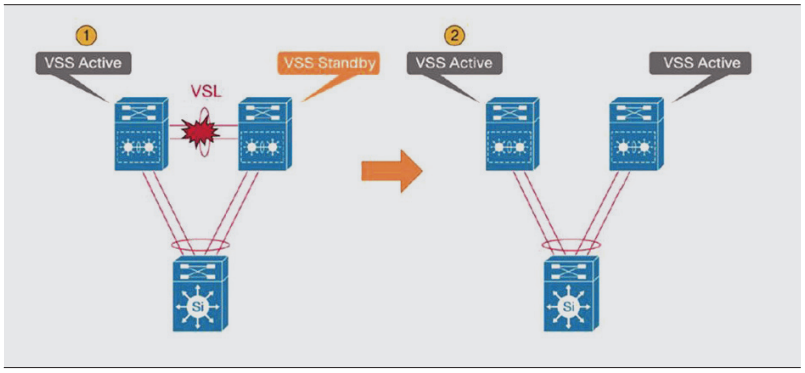


그림 14. VSL 장애 발생시 Dual-active 현상이 발생 합니다

dual-active 현상이 발생했음을 알아내기 위해, 3가지 방법을 사용할 수 있는데, 가능하다면 여러 가지 방법을 같이 사용하는 것을 권장합니다.

첫 번째 방법은 [그림 15]에 표현된 enhanced PAGP 방식입니다. 평상시에 active switch는 자신의 switch id(1)를 PAGP 메시지 안에 넣어서 access switch를 통해 standby switch로 전달합니다.

VSL에 장애가 발생했을 때, standby switch는 active가 되면서 자신의 switch-id(2)를 PAGP 메시지 안에 넣어 access switch로 보냅니다. access switch는 이 메시지를 다시 switch 1에 전달하면, dual active가 발생했음을 알 수 있습니다.

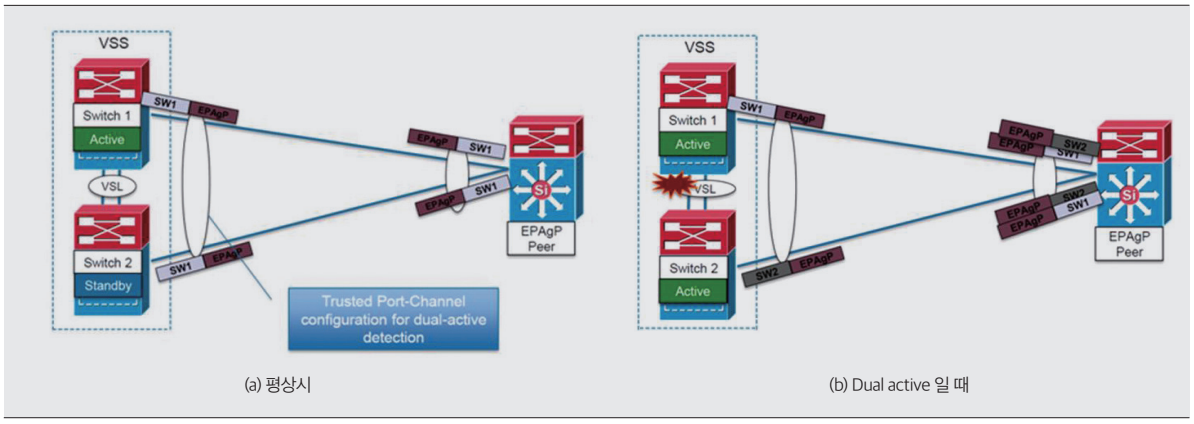


그림 15. Dual-active detection 1 (Enhanced PAGP)

Enhanced PAGP는 기본 enable 상태라서 특별히 세팅하지 않아도 동작합니다. 만약 사용하고 싶지 않다면 [표 1] 에 있는 명령을 입력하면 됩니다.

```
vs-vsl(config)#switch virtual domain 100
vs-vsl(config-vs-domain)#no dual-active detection pagp
```

표 1. enhanced PAGP는 기본적으로 동작하며, disable 하려면 위 명령어를 입력하면 됩니다

두 번째 방법은 "IP BFD"입니다. [그림 16]과 같이 switch 간에 heart-beat 용도로 별도의 layer 3 link를 만드는 방식입니다.

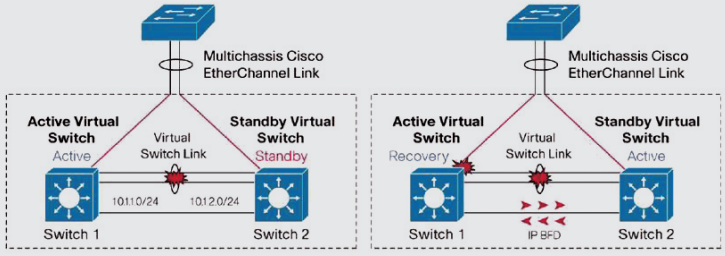


그림 16. Dual-active detection 방법 2 (IP BFD)

"IP BFD"를 세팅할 때는, 먼저 interface에 BFD를 활성화합니다. 그리고 VSS 세팅에서 IP BFD를 적용하면 됩니다. 특이 사항으로는 2 link가 직접 연결되어 있지만, 다른 대역의 IP를 할당하는 것입니다. [그림 17]처럼 switch 1의 IP(10.1.1.1/24), switch 2의 IP(10.1.2.1/24)는 네트워크 대역이 다릅니다.

IP BFD를 세팅하고 나면, 2 switch 간에는 static route가 만들어지는데, 만들어진 static route는 VSL 장애시만 routing table에 올라갑니다. 즉 평상시에는 이 link를 통해 어떤 traffic 교환도 없지만, VSL 장애시만 traffic이 교환되면서 dual active를 탐지할 수 있습니다.

interface 세팅	vss(config-if)#int gig 2/5/1 vss(config-if)#ip address 10.1.2.1 255.255.255.0 vss(config-if)#bfd interval 100 min_rx 100 multiplier 50 vss(config-if)#no shutdown
	vss(config)#int gig 1/5/1 vss(config-if)#ip address 10.1.1.1 255.255.255.0 vss(config-if)#bfd interval 100 min_rx 100 multiplier 50 vss(config-if)#no shutdown
global 세팅	vss(config)#switch virtual domain 10 vss(config-vs-domain)#dual-active detection bfd vss(config-vs-domain)#dual-active pair interface gig1/5/1 interface gig2/5/1 bfd adding a static route 10.1.2.0 255.255.255.0 Gi1/5/1 for this dual-active pair adding a static route 10.1.1.0 255.255.255.0 Gi2/5/1 for this dual-active pair

그림 17. IP BFD 세팅

세 번째 방법은 [그림 18]에 표현된 "VSLP fast hello detection"입니다. "IP BFD"와는 다르게 switch 간에 heart beat 용도로 layer 2 link를 사용합니다.

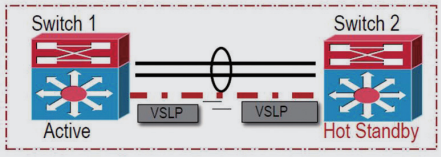


그림 18. Dual-active detection 방법 3 (VSLP fast hello)

interface 세팅	vss(config)#interface f1/2/40 vss(config-if)#dual-active fast hello vss(config-if)#no shutdown
global 세팅	vss(config)#switch virtual domain 100 vss(config)#dual-active detection fast hello

표 2. VSLP fast hello 세팅

김해중
KBS 보도기술국

이 3가지 방법으로 dual active임을 발견하면, 기존의 active switch는 자신의 모든 port를 shut down하면서 recovery mode로 진입하고, 기존의 standby switch는 active switch 역할을 하게 됩니다. recovery 상태에서 VSL link 중에 1개라도 살아나면, switch는 reload를 하게 되며, reload 후에는 standby 역할을 맡게 됩니다. 주의 사항으로 recovery 모드일 때는, 절대로 어떤 VSS 세팅도 하면 안 됩니다.

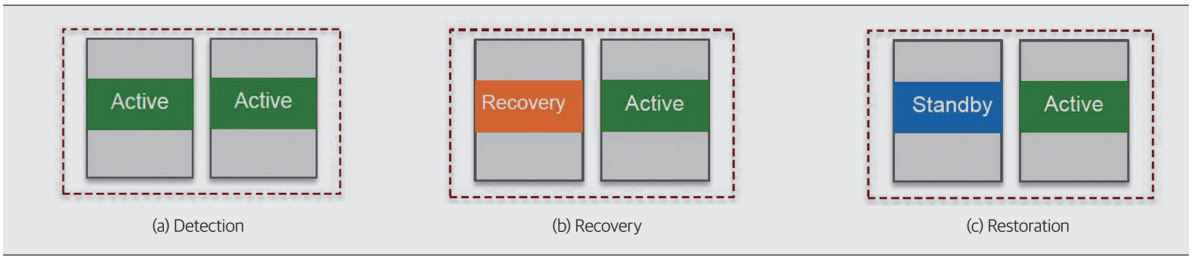


그림 19. Dual-active 발생 시의 3단계

VSS 최초 세팅

[그림 20]을 기준으로 간단히 VSS 세팅을 구현해보겠습니다.

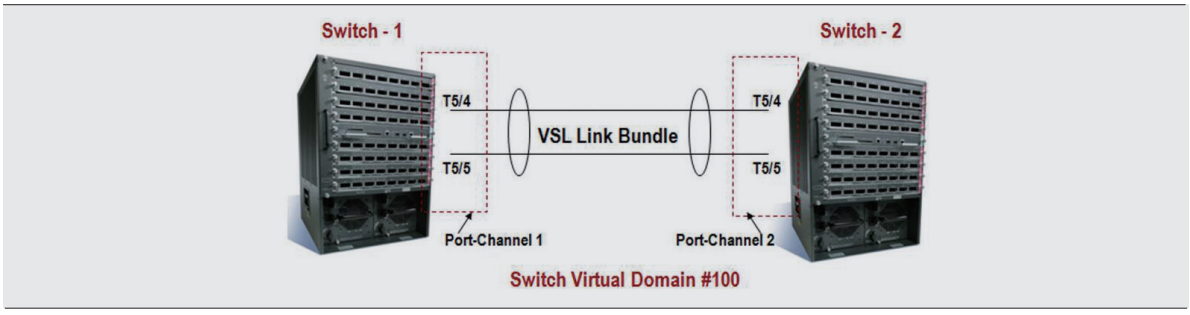


그림 20. VSS 세팅

1) step 1

• 동일한 domain-id(100)를 세팅 후에, 각 switch에 ID(1, 2)를 할당합니다.

• active 역할을 하는 switch에 높은 priority(110)을 할당했습니다.

switch 1	switch 2
VSS(config)#switch virtual domain 100 VSS(config-vs-domain)#switch 1 VSS(config-vs-domain)#switch 1 priority 110 VSS(config-vs-domain)#switch 2 priority 100	VSS(config)#switch virtual domain 100 VSS(config-vs-domain)#switch 2 VSS(config-vs-domain)#switch 1 priority 110 VSS(config-vs-domain)#switch 2 priority 100

2) step 2

• VSL을 이더채널로 세팅합니다.

• VSS 세팅 시, switch 1의 switch-ID, port-channel 번호, VSL ID는 1로 통일하고, switch 2는 2로 통일하면 된다.

switch 1	switch 2
VSS(config)#interface port-channel 1 VSS(config-if)#switch virtual link 1 VSS(config-if)#interface tenG 5/4 VSS(config-if)#channel-group 1 mode on VSS(config-if)#interface tenG 5/5 VSS(config-if)#channel-group 1 mode on	VSS(config-if)#interface port-channel 2 VSS(config-if)#switch virtual link 2 VSS(config-if)#interface tenG 5/4 VSS(config-if)#channel-group 2 mode on VSS(config-if)#interface tenG 5/5 VSS(config-if)#channel-group 2 mode on

3) step 3

• stand-alone mode에서 VSS 모드로 전환합니다.

• [그림 21]처럼 세팅 후에는 자동으로 reboot 됩니다.

switch 1	switch 2
VSS#switch convert mode virtual This command will convert all interface names to naming convention "interface-type switch-number/slot/port", save the running config to startup-config and reload the switch. Do you want proceed? [yes/no]: yes Converting interface names Building configuration... AT THIS POINT THE SWITCH WILL REBOOT	VSS#switch convert mode virtual This command will convert all interface names to naming convention "interface-type switch-number/slot/port", save the running config to startup-config and reload the switch. Do you want proceed? [yes/no]: yes Converting interface names Building configuration... AT THIS POINT THE SWITCH WILL REBOOT

그림 21

4) step 4

• VSS 세팅 결과를 확인합니다.

```
VSS#sh switch virtual role

Switch  Switch Status  Preempt    Priority  Role      Session ID
        Number         Oper(Conf) Oper(Conf)          Local  Remote
-------------------------------------------------------------------
LOCAL    1     UP      FALSE(N)     110(110) ACTIVE    0      0
REMOTE   2     UP      FALSE(N)     100(100) STANDBY   4217   6561
```

switch 1이 priority가 높아서(110) active역할을 하고 있음을 확인할 수 있다.

그림 22

5) interface 표시 변경 확인

• 1대의 장비로 보이게 하기 위해서, switch 1의 interface는 1로 시작하고, switch 2의 interface는 2로 시작합니다.

switch 1	switch 2
SW1-VSS# System detected Virtual Switch configuration... Interface TenGigabitEthernet 1/1/4 is member of PortChannel 1 Interface TenGigabitEthernet 1/1/5 is member of PortChannel 1	SW2-VSS# System detected Virtual Switch configuration... Interface TenGigabitEthernet 2/1/4 is member of PortChannel 2 Interface TenGigabitEthernet 2/1/5 is member of PortChannel 2

그림 23

VPC(virtual port channel)

등장 배경

VPC는 앞에서 설명한 VSS처럼 2대의 switch를 1대의 가상 switch로 통합하는 기술로써, cisco의 NEXUS series(5K, 6K, 7K, 9K)에서 지원합니다. 기존에 backbone switch로 많이 사용되는 catalyst 6500 switch는 IOS 기반이지만, Nexus series는 NX-OS라는 별도의 OS를 사용하고 있어서, 명령어 체계가 조금 다르며, catalyst 6500에 비해서 더욱 많은 feature를 제공합니다.

현재 방송국의 NPS 환경은 LAN, SAN이 물리적으로 구분되어 있어서, 별도의 LAN, SAN switch를 사용하고 있습니다. 하지만 Nexus는 LAN switch 기능뿐만 아니라, SAN switch도 지원을 해서 LAN, SAN 통합이 가능할 뿐만 아니라, LAN을 통해서 SAN을 전달할 수 있는 FCoE(FC over Ethernet)까지도 지원을 합니다.

LAN 환경과 달리 SAN 환경에서는 control-plane을 통합하는 개념이 없습니다. 하지만 VSS는 control-plane을 통합하는 기술이라서, SAN과 FCoE를 지원할 수 없다는 치명적인 약점을 가지고 있었기에, control-plane을 분리하는 VPC를 만들게 되었습니다.

김해중
KBS 보도기술국

VPC/VSS 비교

VPC와 VSS에 대한 간단한 비교를 [표 3]에 정리했습니다.

	VSS	VPC
Control-Plane	Single logical Node	Two independent Node
최대 연결 node 수	2	2
Inter-switch Link	VSL	Peer-link
Loop-free 지원	Yes	Yes
switch configuration	Common config	Combined config
Management	분리되어 있음	통합 관리
virtual-mac address	별도로 세팅	필요기본적으로 지원

표 3. VSS/VPC의 간단한 비교

VPC 기본 용어

VPC에 사용하는 기본적인 용어를 [표 4]에 정리해 보았습니다.

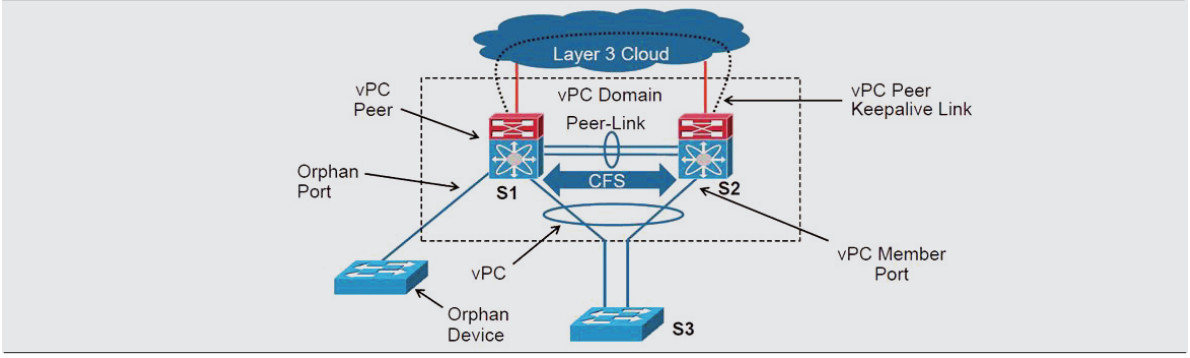


그림 24. VPC

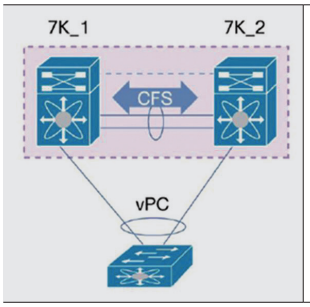
용 어	개 념
VPC domain	VPC를 사용하는 1쌍의 switch
VPC peer	VPC domain을 구성하는 1개의 switch
VPC	downstream device(예:switch, server)에 대한 port-channel
VPC member port	VPC를 만드는 port-channel의 1 port
VPC peer-link	VPC peer간의 control-plane 동기화를 위해서 사용하며, 평상시에는 unicast traffic은 전송하지 않는다.
keepalive-link	Dual-active를 방지하기 위해서, heart-beat 용도로 사용한다.
Orphan device	1개의 VPC peer에만 연결된 장비

표 4. VPC의 기본 용어 정리

control-plane 동작

VPC는 control-plane이 분리되어 있습니다. 분리된 control-plane 간에 정보를 공유하기 위해, 내부적으로 CFS(cisco fabric service)가 동작합니다. 이것은 자동으로 동작하기에, 어떤 세팅도 불필요합니다.

예를 들면, 한쪽 switch에서 mac address learning이 발생하면, CFS를 통해서 반대쪽 switch로 해당 정보를 보냅니다. 그래서 반대쪽 switch도 해당 정보를 가지게 됩니다.



<CFS로 공유되는 정보들>

1. Mac address learning
2. IGMP snooping
3. ARP cache
4. VPC member port status

그림 25. CFS(cisco fabric service)

Data-plane 동작

VPC는 loop를 방지하기 위해, [그림 26]처럼 VPC member port에서 수신한 data는 peer-link를 통해서 다른 VPC member에게 전달 안 되도록 설계되어 있습니다. 결국 VPC는 [그림 27]과 같이 traffic을 수신한 peer switch가 traffic을 전달할 수밖에 없게 됩니다. 즉 peer-link를 사용하지 않습니다.

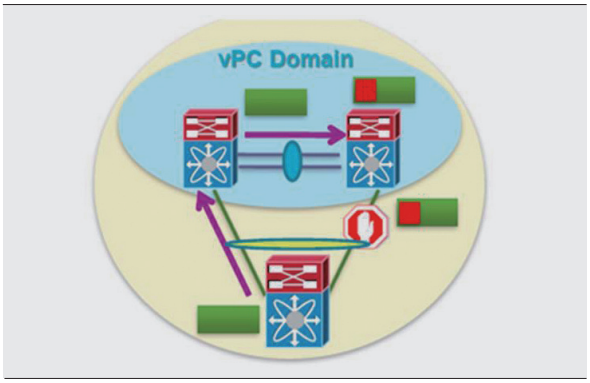


그림 26. VPC의 loop 방지 대책

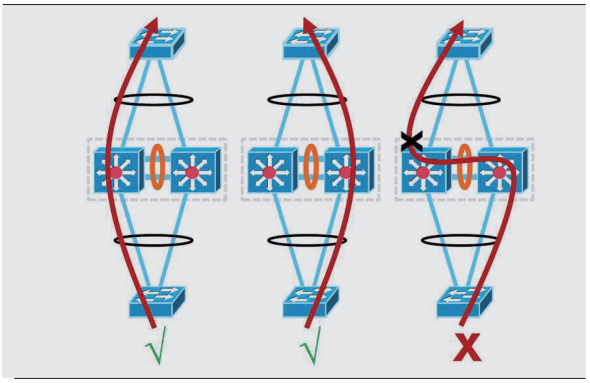
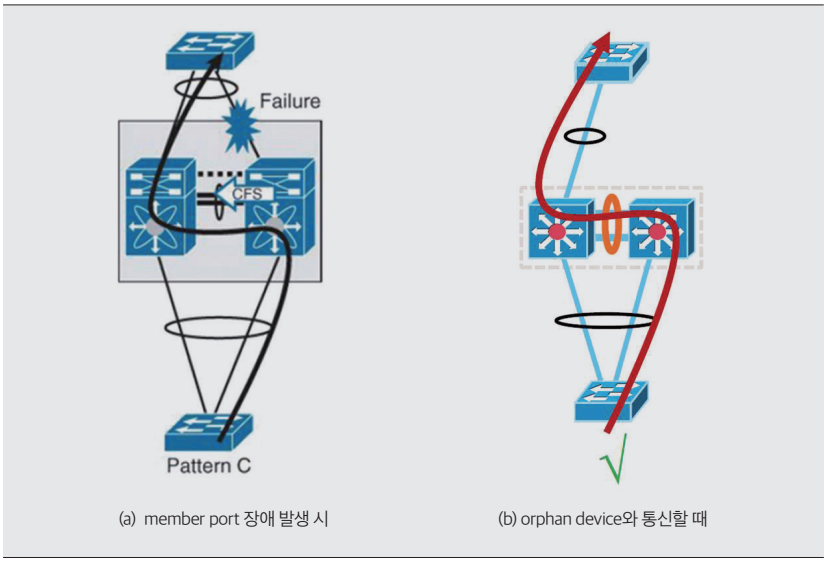


그림 27. VPC의 traffic 흐름: traffic을 수신한 peer가 처리한다

하지만 peer-link로 traffic을 전송하는 2가지 예외가 존재하는데, 첫 번째는 switch의 member port 장애가 발생할 때이며, 두 번째는 VPC를 사용하고 있지 않는 switch(orphan device)에 data를 전송할 때입니다.



(a) member port 장애 발생 시 (b) orphan device와 통신할 때

그림 28. VPC에서 peer-link로 traffic이 전송되는 예외 (2가지)

요약

VSL	VSS pair switch 간의 연결 link로, 평상시에는 control, multicast, broadcast traffic 용도로만 사용합니다.
VSS 동작방식	control-plane은 active/standby로 동작을 하며, data-plane은 active/active로 동작을 한다.
Dual-active	VSL에 장애 발생시, VSS pair switch 모두 active가 되는 현상을 의미합니다.
VSS Dual-active 대책	3가지 방법이 존재하며, 첫 번째는 enhanced PAGP이며, 두 번째는 IP BFD, 세 번째는 VSLP fast hello detection입니다.
VPC 동작방식	control-plane과 data-plane 모두 active-active로 동작을 합니다.
CFS (cisco fabric service)	VPC는 VSS와 달리 control-plane이 분리되어 있어서, control-plane sync를 위해서 CFS를 사용합니다. CFS는 peer-link를 통해 전달합니다.
VPC keepalive-link	VSS는 Dual-active를 방지하기 위해 3가지 방법을 사용했지만, VPC는 keepalive-link를 사용합니다.