

실무 네트워크 Design - 3 :

L2 protocol advanced(VLAN, PVST, RSTP, NO STP)

김해중 KBS 보도기술국

이번 시간에는 먼저 router와 switch의 기본적인 동작 방식을 정리한 후, L2 protocol인 VLAN(Virtual LAN), PVST(Per VLAN STP), RSTP(Rapid STP), NO STP(flex-link, RTG)에 대해서 설명하고자 합니다. 그리고 학습한 기술을 switching block에 적용해 보도록 하겠습니다.

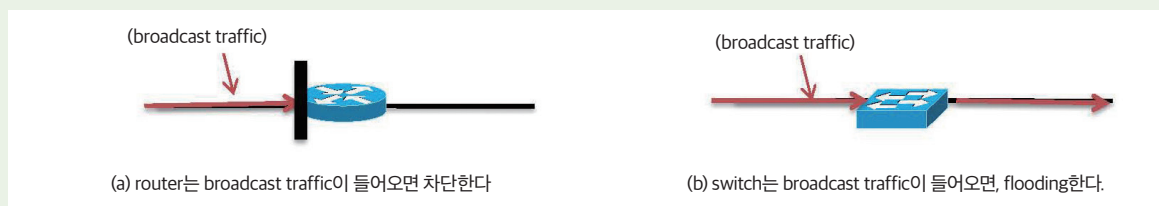
Basic Review

먼저 switch와 router의 기본 동작 방식을 비교해 보겠습니다. “router는 네트워크를 분할한다”라는 말을 많이 들어 보셨을 겁니다. 개념적으로 1개의 선이 1개의 네트워크를 나타낸다고 할 경우, 선 가운데에 router가 들어가면 네트워크는 2개로 나누어집니다. 하지만 switch는 네트워크를 분할하지 못 하기에, 몇 개의 switch가 들어가도 1개의 네트워크입니다.



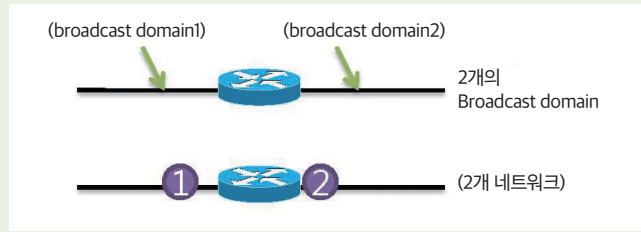
[그림 1] router는 네트워크를 분할하지만, switch는 네트워크를 분할할 수 없다

그리고 라우터는 “broadcast domain을 분리하지만, switch는 불가능하다”라는 말도 많이 들으셨을 겁니다. 즉 router는 broadcast traffic이 들어오면 다른 네트워크로 못 가게 막아 버려서 broadcast domain을 분리하는 것입니다. 하지만 switch는 앞 시간에 배운 것처럼, broadcast traffic이 들어오면, 들어온 port를 제외한 모든 port로 신호를 뿌립니다(이것을 flooding이라고 합니다).



[그림 2] router와 switch 동작방식 차이

즉, router의 2가지 기능인 “네트워크를 분리한다”, “broadcast domain을 분리한다”를 정리하면, 결국 1개의 네트워크가 broadcast domain임을 알 수 있습니다. broadcast traffic은 1개의 네트워크 안에서만 돌아다닐 수 있다는 것입니다.

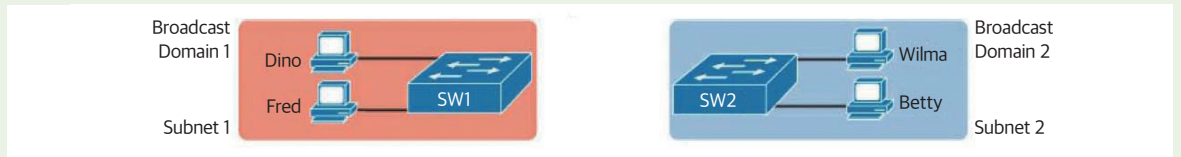


[그림 3] 1개의 네트워크는, 1개의 broadcast domain을 의미한다

VLAN(Virtual LAN)

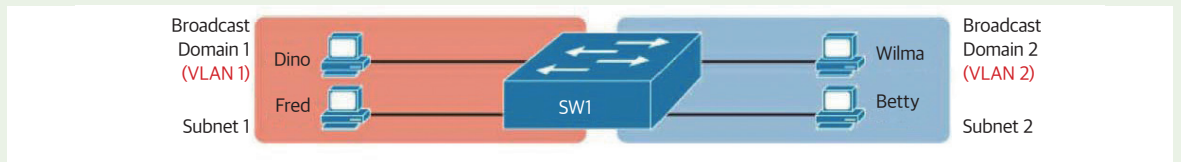
개념

앞에서 설명한 것처럼 네트워크는 router에 의해 나누지만, “switch에 의해서도 네트워크를 나눌 수 있을까”라는 개념에서 VLAN이 출현했습니다. router에 의해 나누어지는 것이 LAN(network)이라면, switch에 의해서도 나누어지는 LAN(network)을 VLAN(Virtual LAN)으로 부릅니다. 개념적으로 표현하니 조금 어려울 수 있는데, 그림을 보면 쉽게 이해할 수 있습니다. [그림 4]를 보면 물리적으로 SW1, SW2는 완전히 분리되어 있습니다. 이렇게 broadcast-domain(network)을 분리하는 것이 LAN 방식입니다.



[그림 4] 2개의 물리적인 switch를 이용해 network를 분리하였다

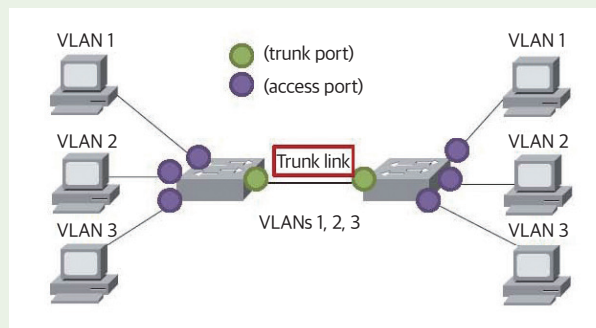
[그림 5]를 보면 1대의 switch를 사용했지만 빨간색, 파란색 영역에 각각 다른 VLAN(VLAN1, VLAN2)을 할당해서 broadcast-domain(network)을 분할했습니다. 그림과 같이 switch에 의해서 network가 구분되어 Virtual LAN(VLAN)이라고 합니다.



[그림 5] VLAN을 사용하면 1개의 switch에서 network를 분할할 수 있다

VLAN Port 종류(2가지)

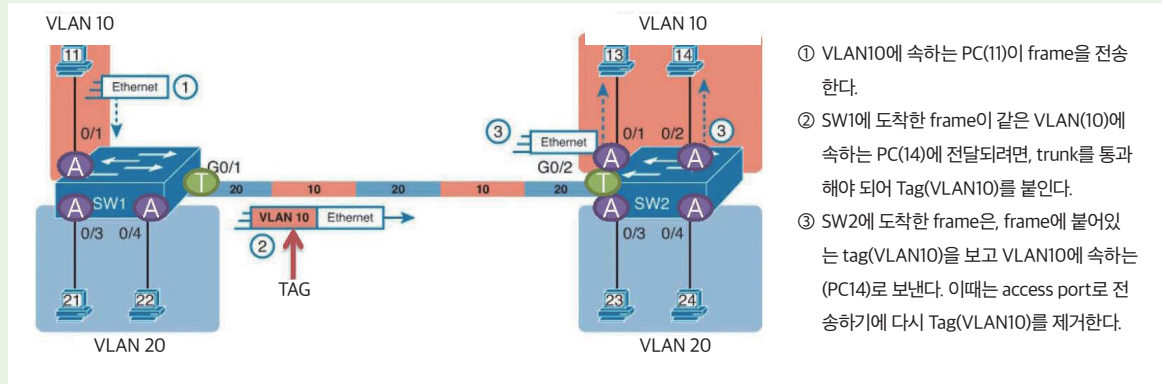
[그림 6]을 보면, 각 PC들이 switch에 연결된 port는 1개의 VLAN만 통과하는데, 1개의 VLAN만 통과하는 port를 access port라고 합니다. 2대의 switch가 연결된 link는 VLAN1, VLAN2, VLAN3 모두 통과하는데, 이렇게 여러 개의 VLAN 정보가 흐르는 link를 trunk라고 합니다.



[그림 6] Access/Trunk port 개념

tagging

trunk port로는 여러 개의 VLAN이 통과하기에, 각 VLAN을 구분할 필요가 있습니다. 그래서 trunk를 통과할 때는 frame에 tag를 붙이는데, 이것을 tagging이라고 합니다. “frame에 너는 10번 VLAN에 속해, 너는 20번 VLAN에 속해” 이렇게 일종의 tag(명찰)를 달아주는 것입니다. [그림 7]을 보면 tagging의 동작 과정을 이해할 수 있습니다.



[그림 7] trunk를 통과할 때 사용하는 Tagging의 동작 원리

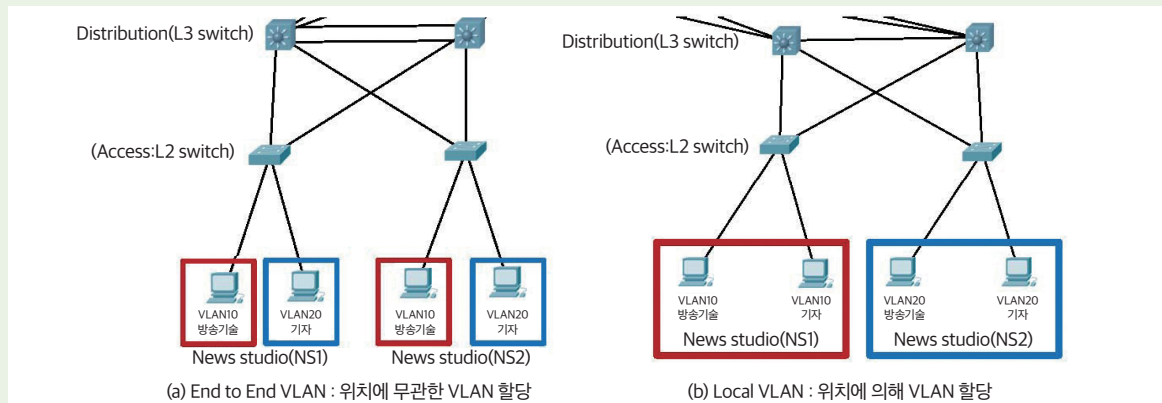
tagging 방식에는 2가지인 802.1Q, ISL이 존재합니다. 2가지의 차이점은 “frame에 어떤 위치에 tag(명찰)를 추가하느냐”입니다. frame의 앞, 뒤에 tag를 추가하는 방식이 ISL이며, frame의 중간에 tag를 추가하는 것이 802.1Q입니다. 현재 ISL은 사실상 폐기된 상태이며, 802.1Q만 사용하고 있습니다.

VLAN의 종류(2가지)

VLAN의 종류에는 Local VLAN과 End to End VLAN 2가지가 있습니다. End to End VLAN은 [그림 8 (a)]처럼 위치와 무관하게, 업무별로 VLAN을 할당하는 방식입니다. 그래서 방송기술은 물리적인 News studio(NS) 위치와 무관한 VLAN10, 기자는 NS 위치와 관계없이 항상 VLAN20이 할당되어 있습니다.

Local VLAN은 [그림 8 (b)]와 같이, 위치에 따라 VLAN을 할당하는 방식으로, NS1의 근무자는 VLAN10, NS2의 근무자는 VLAN20을 할당하는 방식입니다.

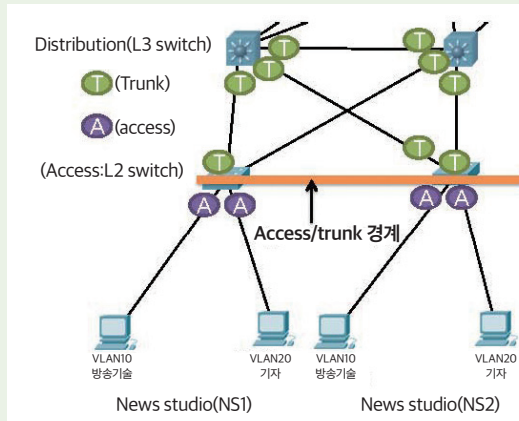
여러분이라면, 어떤 방식을 선호하시겠습니까? 현재 End to End VLAN은 점점 사용하지 않으며, 관리가 용이한 Local VLAN을 선호합니다. End to End VLAN을 사용하면, 같은 공간에 있지만 IP 대역도 달라서 IP 관리도 불편하고, 인사발령 때마다 switch의 VLAN 세팅을 변경해 주어야 하기 때문입니다. 그리고 뒤에서 다루겠지만, STP loop를 해결하는 것에 있어서도 End to End VLAN은 불리합니다.



[그림 8] End to End VLAN과 Local VLAN 비교

VLAN 종류별 일반적인 Port 배치

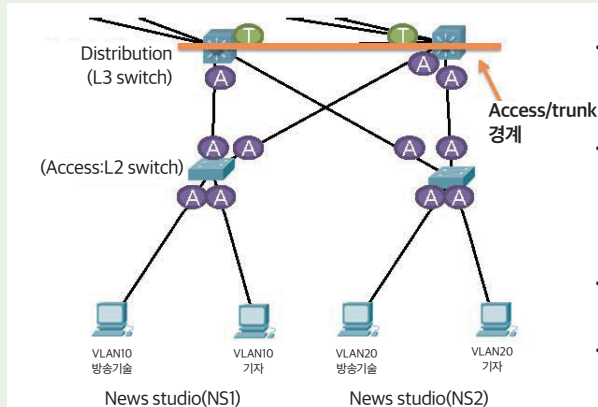
1) End to End VLAN의 port 배치



- access switch 아래쪽은 1대의 VLAN에 속하는 PC만 연결되어 있기에 access port가 된다.
- access switch 위쪽으로는 VLAN10과 VLAN20 모두 통과하기에 trunk port가 된다.
- 결국 access/trunk port 경계는 access switch이다.
- distribution switch 간의 link는 VLAN10, VLAN20이 통과하는 trunk로, 장애발생시 예비 경로 역할을 한다.

[그림 9] End to End VLAN의 port 배치

2) Local VLAN의 port 배치

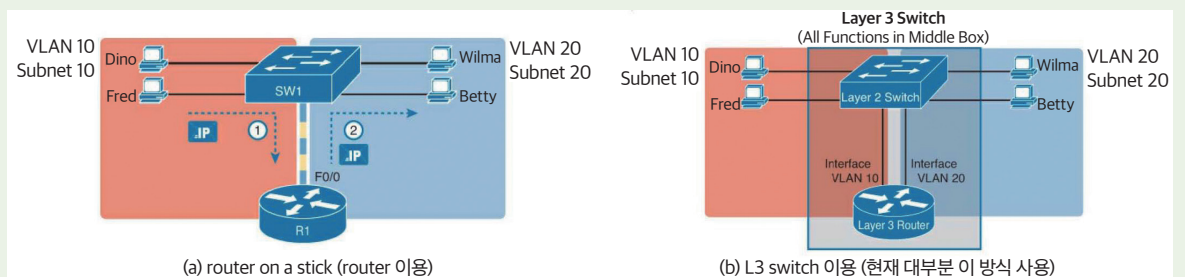


- access switch 아래쪽은 1대의 VLAN에 속하는 PC만 연결되어 있기에 access port가 된다.
- access switch 위쪽도 1대의 VLAN만 통과해서 access로 port가 된다. (이 구간을 trunk로 세팅하는 관리자들도 많지만, broadcast domain을 줄이기 위해 access 세팅을 권장한다)
- 즉 access/trunk port 경계는 distribution switch이다
- distribution switch 간의 link는 VLAN10, VLAN20이 통과하는 trunk로, 장애 발생 시 예비 경로 역할을 한다.

[그림 10] Local VLAN의 port 배치

서로 다른 VLAN과의 통신

다른 LAN 간에 통신하려면, router 같은 L3 장비가 필요합니다. 동일하게 서로 다른 VLAN 간의 통신을 위해서도 L3 장비를 통과해야 합니다. 구현 방식에는 2가지가 있는데, 첫 번째는 router를 이용하는 방식(router on a stick)이 있고, 두 번째는 L3 switch를 이용하는 방식이 있습니다. L3 switch란 L2 기능(switch)과 L3 기능(router)을 하나의 장비 안에서 구현해 놓은 장비로써, 현재 LAN에서는 L3 switch가 router를 거의 대체한 상황입니다. 그래서 서로 다른 VLAN 통신을 할 때는 주로 L3 switch를 사용합니다.

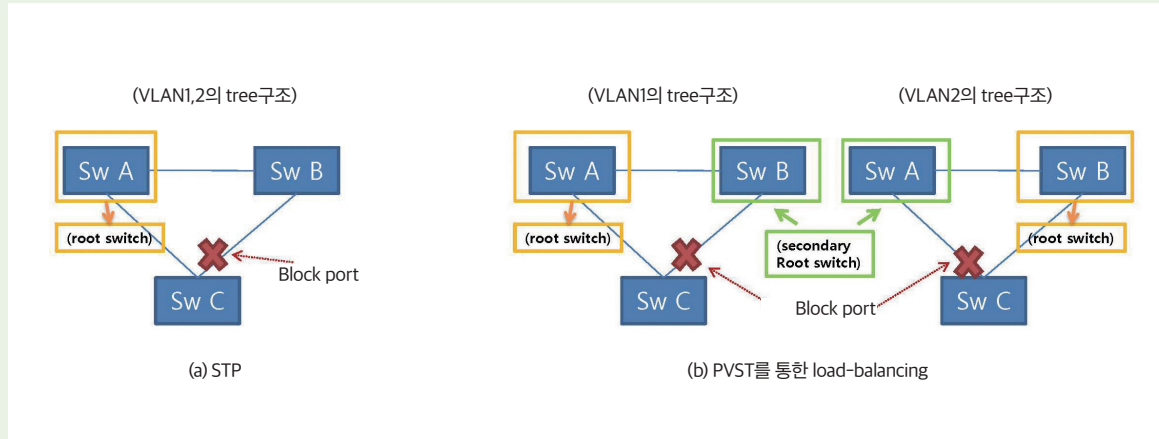


[그림 11] 서로 다른 VLAN 간의 통신 방법 2가지

STP의 진보기술

PVST(Per VLAN STP)

PVST는 STP의 대역폭 낭비 문제를 해결하기 위해 출현한 기술로, VLAN별로 STP를 동작시키는 기술로써, VLAN별로 사용 안 하는 port를 다르게 해서, link 효율성을 높이는 기술입니다.



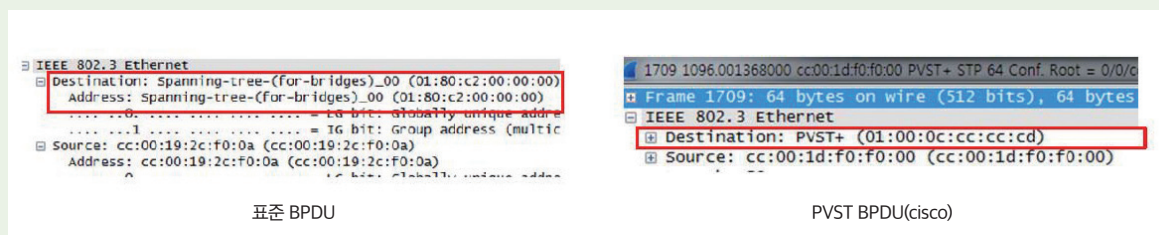
[그림 12] STP/PVST 동작 방식 비교

STP를 사용할 때는 항상 [그림 12 (a)]처럼 switch A가 모든 VLAN의 root switch 역할을 하며, 항상 switch B와 switch C간의 연결 경로를 사용할 수가 없었습니다. 하지만 PVST를 사용하는 경우에는, 각 VLAN별로 기준점인 root switch를 다르게 하여, VLAN별로 tree를 그리게 됩니다. 그래서 [그림 12 (b)]처럼 VLAN 1의 경우는 switch A를 root switch로 만들었고, VLAN 2의 경우에는 switch B를 root switch로 만들었습니다. 이때 secondary root switch도 일반적으로 지정해주며, 이 switch는 root switch에서 장애 발생 시 대체 역할을 수행하는 switch입니다.

그럼 어느 port가 Block이 될까요? 당연히 root switch로부터 가장 먼 곳이면서, root, secondary root switch에 해당하지 않는 port가 됩니다. (그림에서 모든 link 속도는 동일하다고 가정했습니다) 결국 VLAN 1의 통신이 일어나는 경우는 switch B와 switch C 연결 경로를 사용하지 않고, VLAN 2의 통신이 일어나는 경우는 switch A와 switch C경로를 사용 안 하게 되어서, link를 효율적으로 사용하게 됩니다.

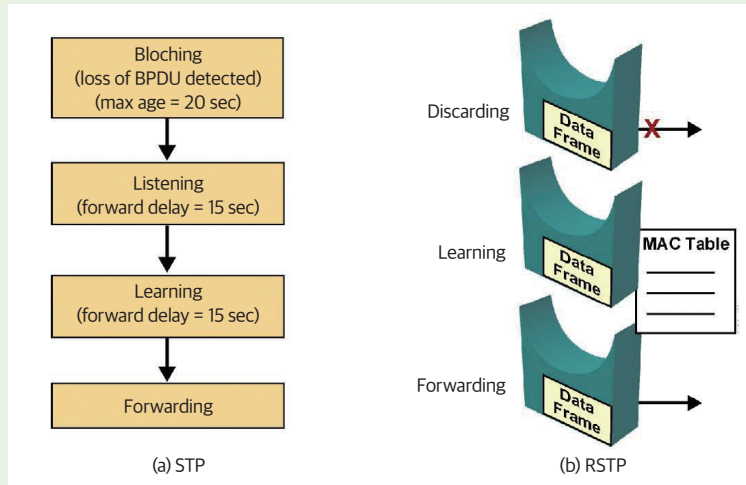
하지만 PVST는 2가지의 문제점을 가지고 있습니다. 첫 번째는 VLAN별로 STP tree 계산을 하기에, switch의 CPU에 더 많은 부담을 주게 된다는 단점이 존재합니다. 두 번째는 PVST가 cisco만이 사용하는 protocol이라는 것입니다. 그래서 wireshark로 패킷 분석을 해보면, cisco 장비는 tag가 안 붙은 VLAN에 대해서는 L2 주소가 (01:80:c1:00:00:00)인 STP BPDU를 전송하고, tagged VLAN에 대해서는 (01:00:0c:cc:cc:cd) L2주소를 가지는 PVST용 BPDU를 동시에 전송함을 알 수 있습니다.

참고로 Juniper의 switch는 cisco의 PVST와 동일한 기능을 하는 VSTP(VLAN STP)라는 고유의 기술을 사용하고 있습니다.



[그림 13] 표준 BPDU와 cisco의 PVST BPDU 비교

RSTP(Rapid STP)



[그림 14] STP/RSTP convergence time 비교

형하였습니다. 사실 RSTP 출현 전에, 장애 시간을 줄이기 위하여 과도기적 기술로 3가지(portfast, uplink fast, backbone fast)가 존재했습니다. RSTP는 이 3가지 기술을 흡수해서 발전시켰을 뿐만 아니라, 다양한 기능을 추가하여서 만들어졌습니다.

RSTP는 IEEE 표준기술이라서, cisco와 juniper 간에 호환해서 사용할 수 있습니다. 하지만 cisco는 RSTP cost 계산을 할 때 16bit만을 사용하는 short 방식을 사용하며, juniper는 32bit를 사용하는 IEEE 표준(long 방식)을 따르고 있습니다. 그래서 cisco 장비에 들어가서 IEEE 표준을 사용하도록 변경을 해주어야 됩니다. 참고로 명령어는 “spanning-tree pathcost method long”입니다.

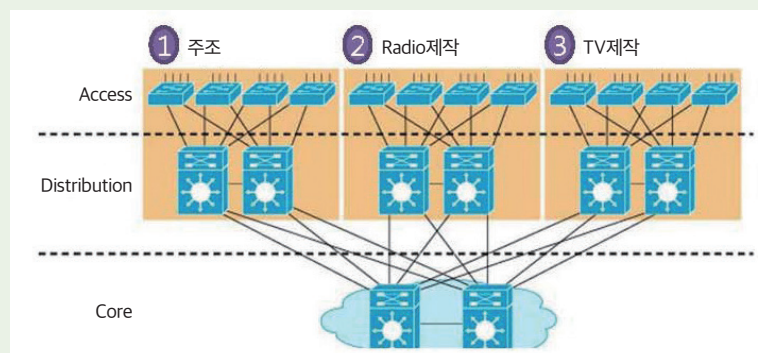
Bandwidth	Short	Long(IEEE표준)
10Mbps	100	2,000,000
100Mbps	19	200,000
1Gbps	4	20,000
10Gbps	5	2,000
100Gbps	N/A	200
1Tbps	N/A	20
10Tbps	N/A	2

[그림 15] RSTP의 2가지 path-cost 방식 비교

PVRST(Per VLAN RSTP)

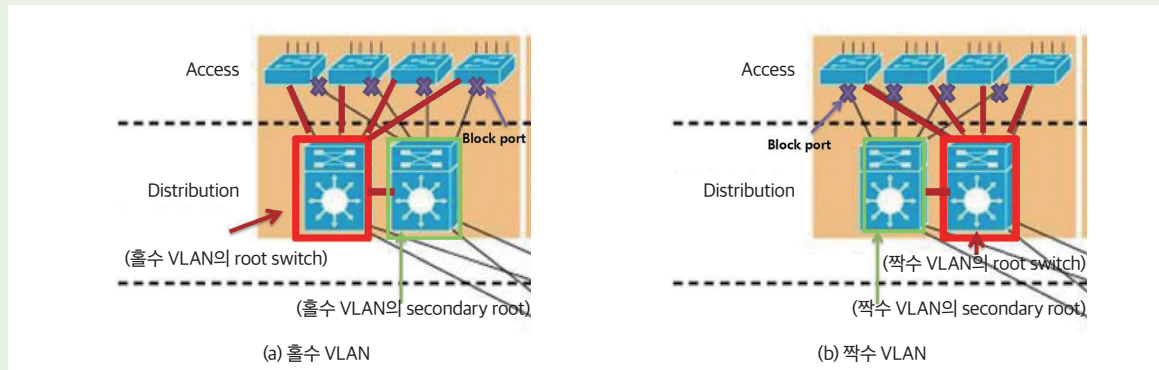
PVST와 RSTP를 결합한 기술이 PVRST이며, STP의 2가지 문제점인 대역폭 낭비와 느린 convergence time을 모두 해결하고 있습니다. 그래서 실제로 cisco 장비를 사용하는 대부분의 회사는 PVRST를 사용하고 있습니다.

Switching block design



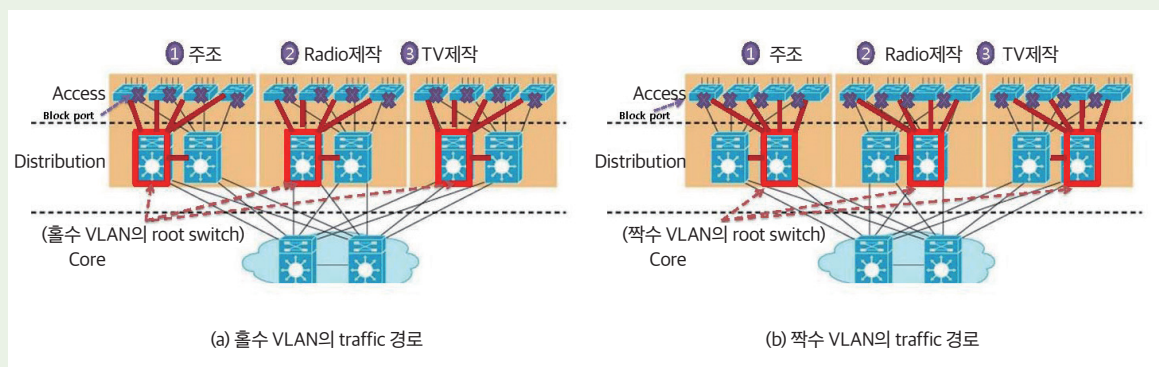
[그림 16] 각 switching block 별로 PVRST를 동작시킨다

이제까지 배운 모든 것을 switching block design에 적용해 보겠습니다. 대부분의 회사는 [그림 16]처럼 3계층 구조로 설계되어 있고, 주황색 영역을 switching block이라고 표현하며, STP는 switching block별로 동작합니다. 즉 3개 부서(①주조 ②Radio제작 ③TV제작)별로 각각 PVRST를 동작시키며, 그래서 개별 switching block별로 root switch를 가집니다.



[그림 17] PVRST를 사용해서 홀수, 짝수 VLAN의 root switch를 다르게 한다

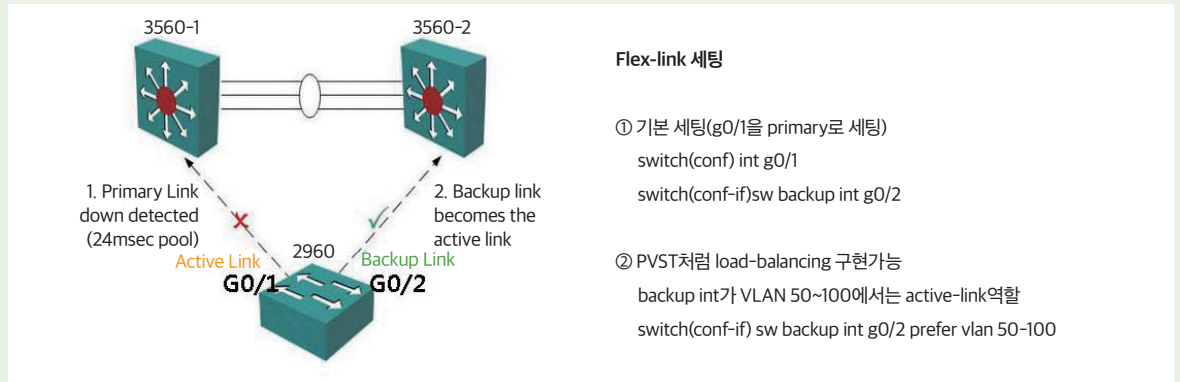
일반적으로 가장 성능이 좋은 distribution layer의 switch를 root switch로 만드는 것을 권장하며, 실제 대부분의 회사는 그 방식을 이용합니다. [그림 17]을 보면, 홀수 VLAN에서는 distribution switch 중에 좌측 switch가 root switch 역할을 하며, 우측 switch는 좌측 switch 장애 시 root 역할을 수행할 수 있도록 secondary root로 세팅을 합니다. 짝수 VLAN에서는 distribution switch 중에 우측 switch가 root switch 역할을 하며, 좌측 switch는 우측 switch 장애 시 root 역할을 수행할 수 있도록 secondary root로 세팅을 합니다. 이렇게 하면, 홀수 VLAN과 짝수 VLAN의 이동 경로를 다르게 하여 load-balancing을 구현할 수 있으며, RSTP로 인해 장애 발생시 1초 내로 줄일 수 있게 되었습니다. [그림 18]에서 빨간 선으로 표시된 곳이 홀수, 짝수 VLAN의 실제 traffic이 이동하는 link이며, 사용 안 하는 port를 X로 표시하였습니다.



[그림 18] 각 switching block별로 홀수, 짝수 VLAN에 대한 root switch를 다르게 하였다

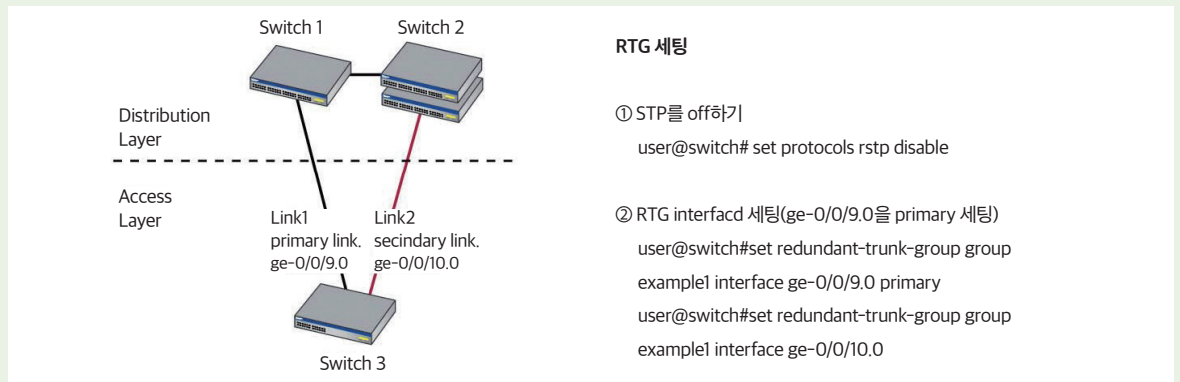
NO STP 기술

switch에서 loop 문제를 해결하려면 STP를 사용해야 합니다. 하지만 STP로 한 번 고통을 겪어본 관리자들은 어떻게든 STP를 사용하지 않으려 합니다. 그래서 STP를 사용하지 않고 loop를 방지하는 기술인 cisco의 Flex-link와 RTG(Redundant trunk groups)를 소개하려 합니다. 2가지 기술은 이름만 다를 뿐, 원리는 동일해 cisco Flex-link 기준으로 설명하겠습니다. [그림 19]과 같이 triangle 구조에서 평상 시에는 항상 primary link를 사용합니다. primary link 장애가 발생하면, 즉시 backup link를 active 되게 하는 기술입니다. flex-link가 동작한 곳은 STP가 자동으로 disable되면서, BPDU를 전송하지 않습니다. Flex-link는 PVST처럼 load balancing을 쉽게 구현이 가능하지만, RTG는 load-balancing 개념이 없기에 local-VLAN에서의 사용을 권장합니다.



[그림 19] NO-STP 기술 : cisco의 flex-link

현재 STP 없이 loop을 방지하는 기타 방법으로는, local-VLAN에서 사용할 수 있는 U구조, cisco의 VSS(Virtual, switching system), VPC(virtual-port channel), TRILL, Fabric-path 등이 있습니다. 언급한 기술은 다음에 자세히 다루도록 하겠습니다.



[그림 20] NO-STP기술 : Juniper의 RTG

요약

이번 시간에 학습한 내용을 아래 표에 정리해 보았습니다.

Protocol	역 할
VLAN	switch에 의해 network를 분리해서 Virtual LAN이라고 하며, End to End VLAN 방식과 Local VLAN 방식이 있다.
PVST	STP의 문제점인 link 낭비를 해결한 기술로써, VLAN별로 사용 안 하는 link를 달리하여 load-balancing을 구현한다. Juniper 장비에서는 동일한 기능을 VSTP(VLAN STP)라고 부른다.
RSTP	STP는 30~50초의 장애 시간을 가질 수 있기에, 그 시간을 1~2초로 줄여주는 기술이다.
PVRST	cisco에서 PVST와 RSTP를 결합한 기술을 의미한다.
NO STP	STP를 사용하지 않고 layer2에서 loop 문제를 해결하는 기술을 의미하며, cisco의 flex-link와 juniper의 RTG 등이 있다.

참고 문헌

CCNA switch
official certification
guide (cisco press)

CCNP switch
official certification
guide (cisco press)

Junos Enterprise
switching (O'Reilly
media)

필자가 switching 공부할 때 가장 어려웠던 개념이 VLAN과 STP였습니다. cisco press 책을 볼 때 개념이 너무 헛갈려서, 구글링을 하며 자료를 찾아 다녔습니다. 대부분의 엔지니어들은 STP 관련 command는 스위치에 빨리 입력해서 사용하지만 정확히 아는 사람이 생각보다 적습니다. 특히 RSTP는 대부분의 운용자들이 “빨리 돼” 이 결과만 알고 있지, 세부적인 STP와 차이점을 아는 사람은 거의 없어서, 공부하면서 힘들었던 기억이 떠오릅니다. 먼저 여기에 적힌 글을 통해 개념 파악하신 후에, cisco press 책을 읽으며 세부동작 방식을 이해하시는 것을 권해드립니다. 다음 시간에는 routing의 기본 개념과 세팅에 대해서 설명하겠습니다. 📖