

Week 8.

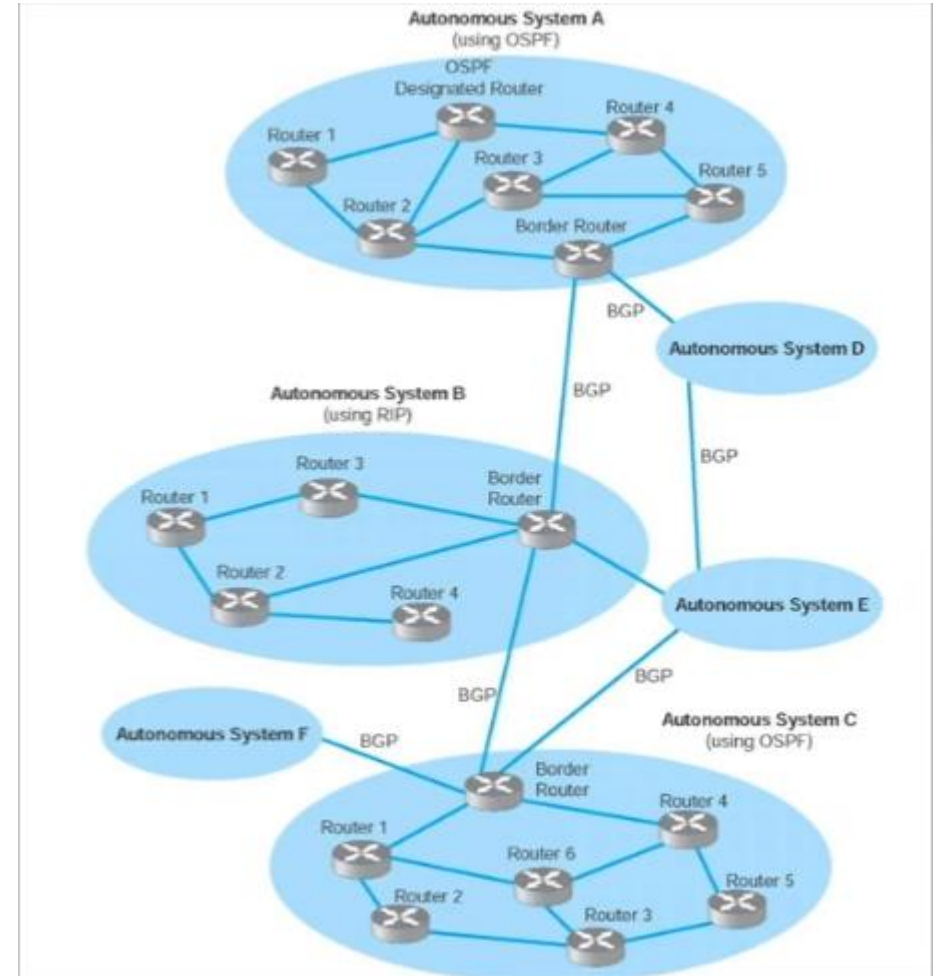
Intra-/Inter- AS Routing and SDN

차례

- Intra AS Routing OSPF
- Inter AS Routing BGP 라우팅
- SDN (Software Defined Networking)
- OpenFlow
- ICMP (Internet Control Message Protocol)

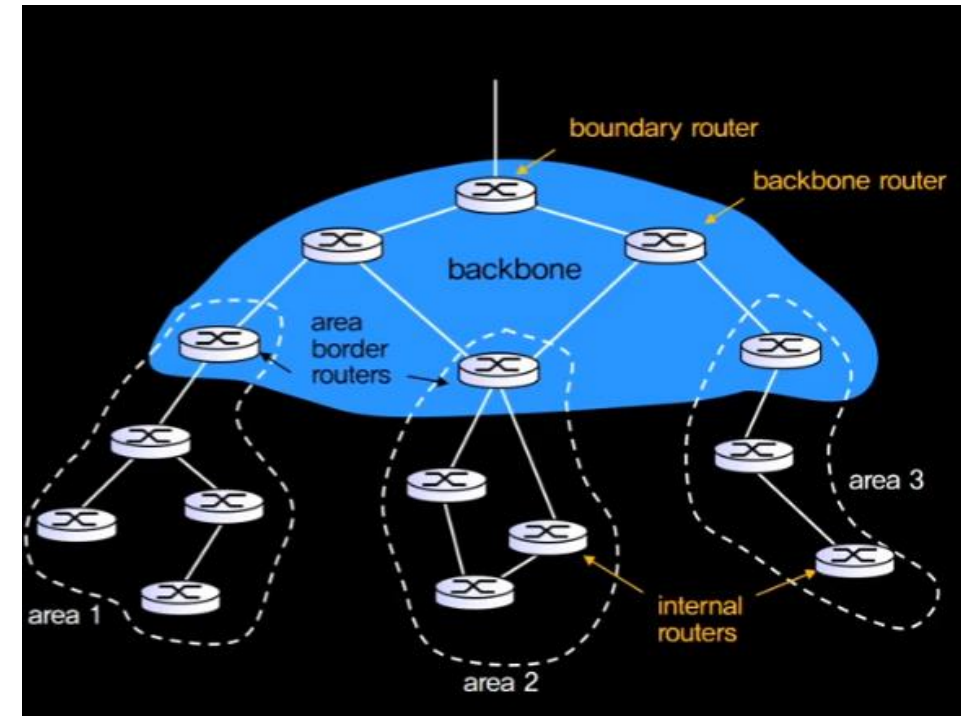
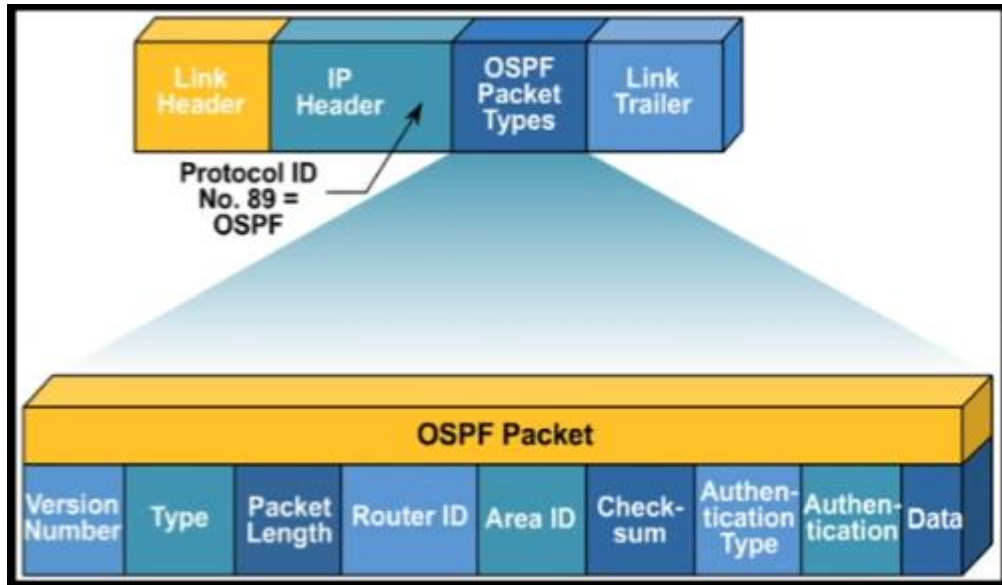
Intra AS Routing OSPF

- 인터넷은 flat한 환경에 있는게 아니라 어떤 중앙 집중 중심이 있고 여기서 가지치기 형태로 나오게 됨.
즉, 하위계층을 가지는 형태의 네트워크
- flat하다면 scalability issue가 있을 것.
- AS : Autonomous System, 한 기관에서 관리하는 네트워크 라우터들의 집합 또는 동일한 prefix를 갖는 라우터들의 집합.
- OSPF : Open Shortest Path First
- 표준 문서가 공개 되어있고, 그 표준을 따르기만 하면 다른 라우터들과 통신하는데 문제 없음.
- 단거리를 우선시하는 프로토콜



Intra AS Routing OSPF

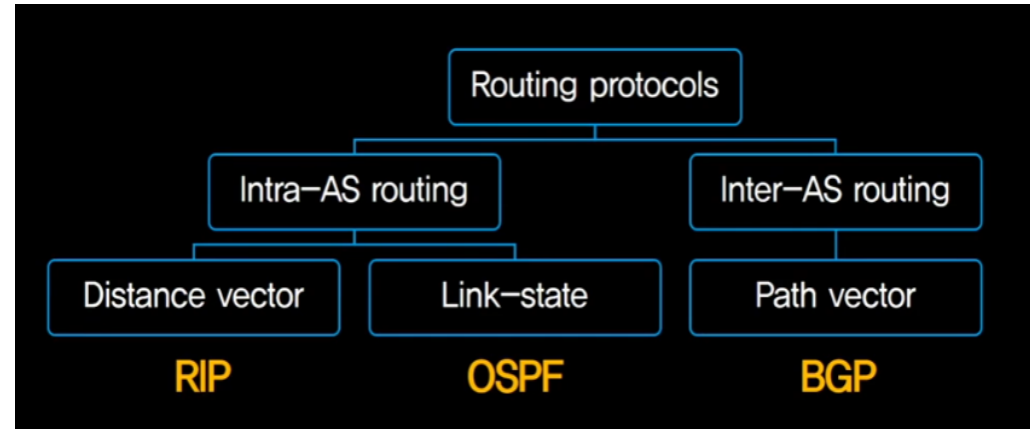
- OSPF : Open Shortest Path First
- 표준 문서가 공개 되어있고, 그 표준을 따르기만 하면 다른 라우터들과 통신하는데 문제 없음.
- 단거리를 우선시하는 프로토콜



Hierarchical OSPF

Intra AS Routing OSPF

- 라우팅 프로토콜 종류



- Intra (내부) / Inter(외부) AS routing

- 대표적인 Intra AS 라우팅은 Link-state 프로토콜, OSPF, RIP 등이 있다.
- 하나의 AS 내에서는 동일한 OSPF나 동일한 RIP를 써야 함.
- '다른 AS가 무슨 프로토콜을 쓰는지' 는 상관 없음
- 외부의 다른 AS 와 링크를 갖는 라우터는 경계 라우터(border router) 또는 게이트웨이 라우터 라고 한다.
- 외부 AS끼리 라우팅을 할 때는 gateway router가 inter-domain routing을 동작 시켜야 함.

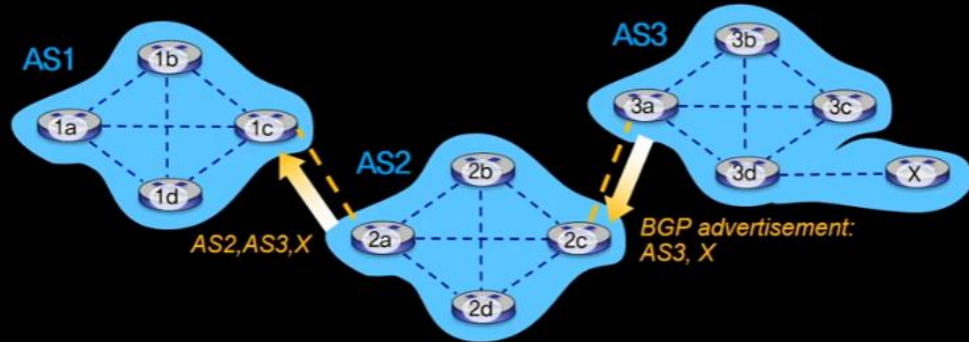
- 경계 라우터 (border router)

- Area border router : 하나의 Area를 요약해서 전달함.
- Backbone Router : 이런 area들을 연결시켜주는 역할을 하고 있음.

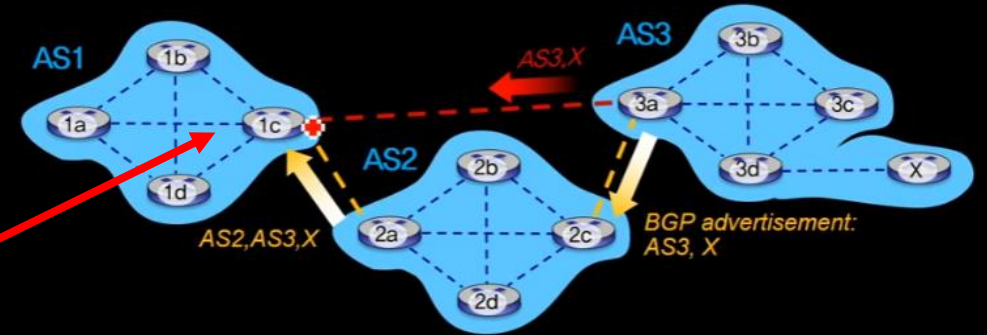
Inter AS Routing BGP 라우팅

- BGP는 AS-PATH, NEXT-HOP 이라는 방식이 있다.
 - AS-PATH : 앞에 봤던 Path vector
 - NEXT-HOP : 그 as로 전달 하기 위해서 다음에 선택 해야 될 next router
- BGP 뿐만 아니라 inter domain routing protocol은 분류를 하자면 policy -based routing 이다.
- Intra domain routing protocol은 되도록이면 모든 노드들이 정확하게 공유하고 있는 것이 중요했음.
그렇지만 BGP는 policy가 더 중요.
- Policy라 함은. 'A경로는 accept, B경로는 decline 하는 등의 정책'을 의미함.
설령 B 경로가 효율적이라 하더라도 policy에 B경로가 decline 되어 있으면 그것을 절대적으로 지킴.

Inter AS Routing BGP 라우팅



- AS2 router 2c receives path advertisement **AS3, X** (via eBGP) from AS3 router 3a
- Based on AS2 policy, AS2 router 2c accepts path **AS3, X**, propagates (via iBGP) to all AS2 routers
- Based on AS2 policy, AS2 router 2a advertises (via eBGP) path **AS2, AS3, X** to AS1 router 1c



Gateway router may learn about multiple paths to destination:

- AS1 gateway router 1c learns path **AS2, AS3, X** from 2a
- AS1 gateway router 1c learns path **AS3, X** from 3a
- Based on policy, AS1 gateway router 1c chooses path **AS3, X**, and advertises path within AS1 via iBGP

1c에서 판단을 내림.

AS1 관리자의 정책에 따라 두 경로 중 하나를 선택해서 X로 가는 데이터그램을 전달 할 수 있다.

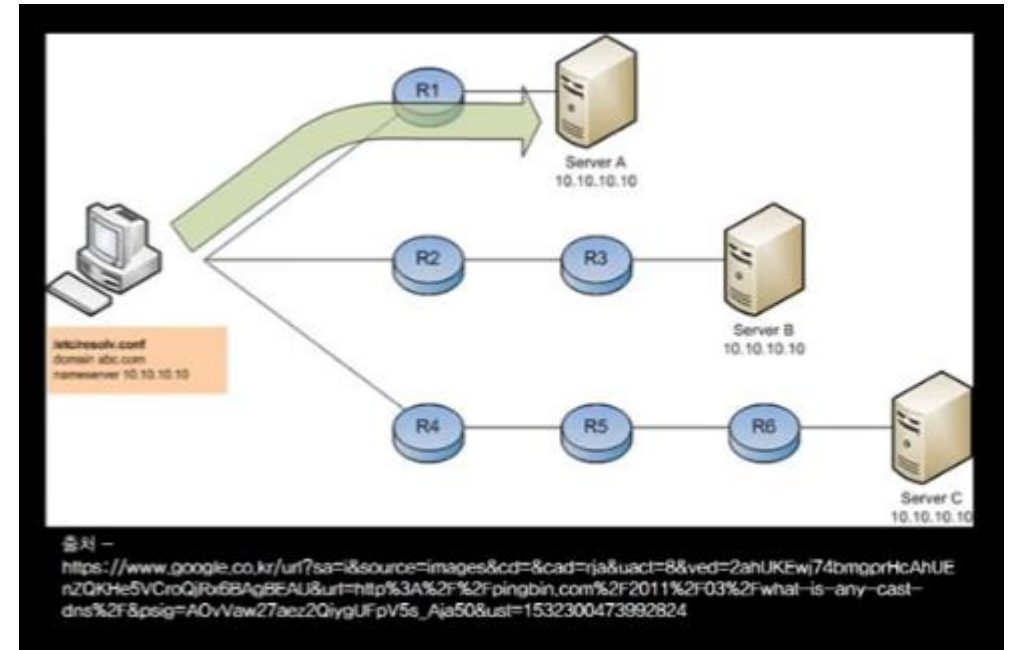
Inter AS Routing BGP 라우팅

- Anycast

전세계에 루트 서버가 한 17개 정도 있는데, 내 시스템이 어느 루트 서버한테 갈 것인지가 명확하지 않을 때, Anycast가 가장 가까운 애를 찾게 해줄 수 있음.

클래스 E 주소를 사용함. (1111 28bits 짜리 주소)

1. 모든 DNS에게 동일한 IP주소를 할당한다.
그러면 할당한 주소가 BGP기능을 통해서 AS 정보가 전달이 될 것이다.
2. BGP 메시지가 사용자한테 오면 호스트는 이 DNS들이 물리적으로 다른 서버라는 것을 인지하지 못하고 같은 서버지만 AS1~n을 통하는 길이 있다고 생각을 함.
3. 그 중에 가장 적은 AS를 거치는 AS1과 통하는 길을 선택하게 되면, 결국은 여러 DNS 시스템 중에 가장 가까운 시스템에게 쿼리를 전달하게 되는 결과를 가져옴.



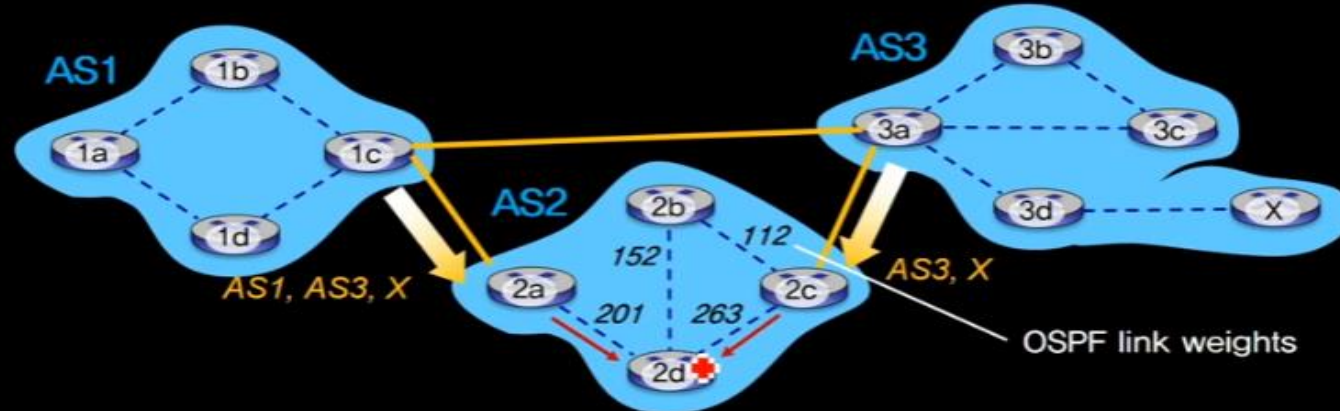
Inter AS Routing BGP 라우팅

- Hot potato routing

이기적인 라우팅. 그냥 더 가까운 border router에게 데이터그램을 무조건 전달하는 방식.

2d의 입장에서서는 2c보다 비용이 적게 드는 2a로 보내게 됨. (AS2의 관리자가 선택할 문제긴 함.)

- 2d learns (via iBGP) it can route to X via 2a or 2c
- **Hot potato routing**: choose local gateway that has least intra-domain cost (e.g., 2d chooses 2a, even though more AS hops to X): don't worry about inter-domain cost!
 - the idea behind is to get packets out of its AS as quickly as possible without worrying about the cost of the remaining portions of the path outside its AS to destination



Inter AS Routing BGP 라우팅

- AS간 경로를 선택하는 방법
 1. Policy 우선 결정법
 2. 최단 AS Path 결정법
 3. 가장 가까운 Next – HOP 라우터 결정법
 4. 기타 관리자가 원하는 기준으로 경로 결정
- Intra-AS routing, Inter-AS routing 비교
 - Why different? → Scalability 때문이다.
모든 라우터가 같은 레벨에 놓여져 있다고 하면, 각 라우터끼리 주고 받아야하는 라우팅 테이블 업데이트 트래픽 때문에 전체 네트워크가 마비가 되는 사태가 벌어질 수 있다.
 - 그래서 같은 AS 내에서는 Intra, 서로 다른 AS 간에는 Inter domain routing protocol을 가지고 연결하는 방식으로 동작한다.

Inter-AS routing은 policy based routing이고, Intra-AS routing은 퍼포먼스를 가장 우선시하는 라우팅 프로토콜이다.

SDN (Software Defined Networking)

- 특정한 데만 링크가 몰림으로써 전체적으로 지연시간이 길어지거나 이런 문제가 발생할 수 있다.
그래서 데이터 트래픽을 분산시키거나 보내는 사람이 의도하는 길로 가게 만듦.
- 트래픽의 전달을 효율적으로 만드는 것 : Traffic Engineering
- 기존 라우팅 프로토콜로는 데이터 트래픽을 분산해서 보내는게 안된다.
네트워크 자원의 효율적인 사용이 불가능 했다.
그러다 보니 특정 라우터에 데이터가 몰려서 전체 네트워크 서비스 품질이 저하되더라 → SDN 고안

● SDN

- 개별 라우터들이 전체 topology를 기반으로 스스로 라우터를 결정하는 것이 아니라
네트워크 전체 정보를 하나의 중앙 집중 시스템으로 모은 뒤에 모은 정보를 기반으로
각 src로부터 각 dest로의 길을 remote controller가 결정을 내림
- 거기에 기반한 포워딩 테이블을 각각의 라우터에 심어주고, 각 라우터는 테이블을 기반으로 전달만 하게 만듦
- 길을 설정하거나 할 때는 중앙 집중형 프로그램이 더 쉽다.

SDN (Software Defined Networking)

- Generalized forwarding

- 기존 포워딩은 dest based forwarding 근데 트래픽 엔지니어링이 불가능.
그래서 특정한 하나가 아니라 사용자가 정한 어떤 set이나 정보의 집합에 기반해서 포워딩 하는 것이 generalized forwarding 이다.

- SDN 시스템 구성

- 네트워크 제어 어플리케이션

- 라우팅, load balancing, access control 등.

- 중앙 집중 컨트롤러

- 네트워크 상태 정보 유지해야함. 물리적으로는 분산되어있으나, 동기화가 되어있어 논리적으로는 하나의 컨트롤러로 보이게 만든다. ('logically' controller)

- SDN-controlled switches

- 스위치들의 집합.

OpenFlow

- SDN 컨트롤러는 스위치들로부터 link-state 정보나 호스트, 스위치 정보들을 모아서 이런 것을 가지고 network graph를 만들고 라우팅 알고리즘을 적용해서 길을 찾음. 찾은 길을 flow table로 만들어서 그것을 스위치에 전달을 해줌.
- 이 컨트롤러 하부에 OpenFlow가 있음. OpenFlow는 SDN 컨트롤러 하부에 스위치들과 통신을 할 때 쓰는 프로토콜.
- 고안한 라우팅 알고리즘을 기존의 라우터나 이더넷 스위치에 넣어서 동작시키는게 불가능 했었고, 본인들의 생각을 실험하고자 만든게 OpenFlow

● OpenFlow 구성요소

■ Flow table

각 노드들이 보내온 데이터그램을 어디로 전달할 것인가

■ Secure channel

컨트롤러가 주고받는 데이터가 SSL 또는 TLS 서비스를 통해서 전달하게끔 되어있는 것.

■ OpenFlow protocol

채널(스위치)과 컨트롤러 사이에 주고받는 데이터 순서라든지 포맷 등 이런 것을 결정하는게 OpenFlow protocol

OpenFlow

- Open Flow 는 크게 세 개로 구별할 수 있다.

1. controller to 스위치 (Synchronous)

■ controller to 스위치 메시지 종류

- ◆ Configuration : 어떻게 설정이 되어있는지? 어떤 값으로 설정을 해야하는지
- ◆ Read-state : flow table 정보, 데이터그램을 전달할 때 데이터가 잘 전달이 되었는지 되지 않았는지, 성공률은 얼마인지.
이 메시지를 통해서 flow table 갱신
- ◆ Modify-state : 기존에 있는 openflow table에 entry를 추가, 삭제, 수정 한다는 메시지. 이 메시지를 통해서 각 스위치의 flow table 갱신
- ◆ Send-packet : 컨트롤러가 '특정한 패킷은 무조건 특정 포트로 내보내라'라고 단도직입적으로 이야기 하는 것.

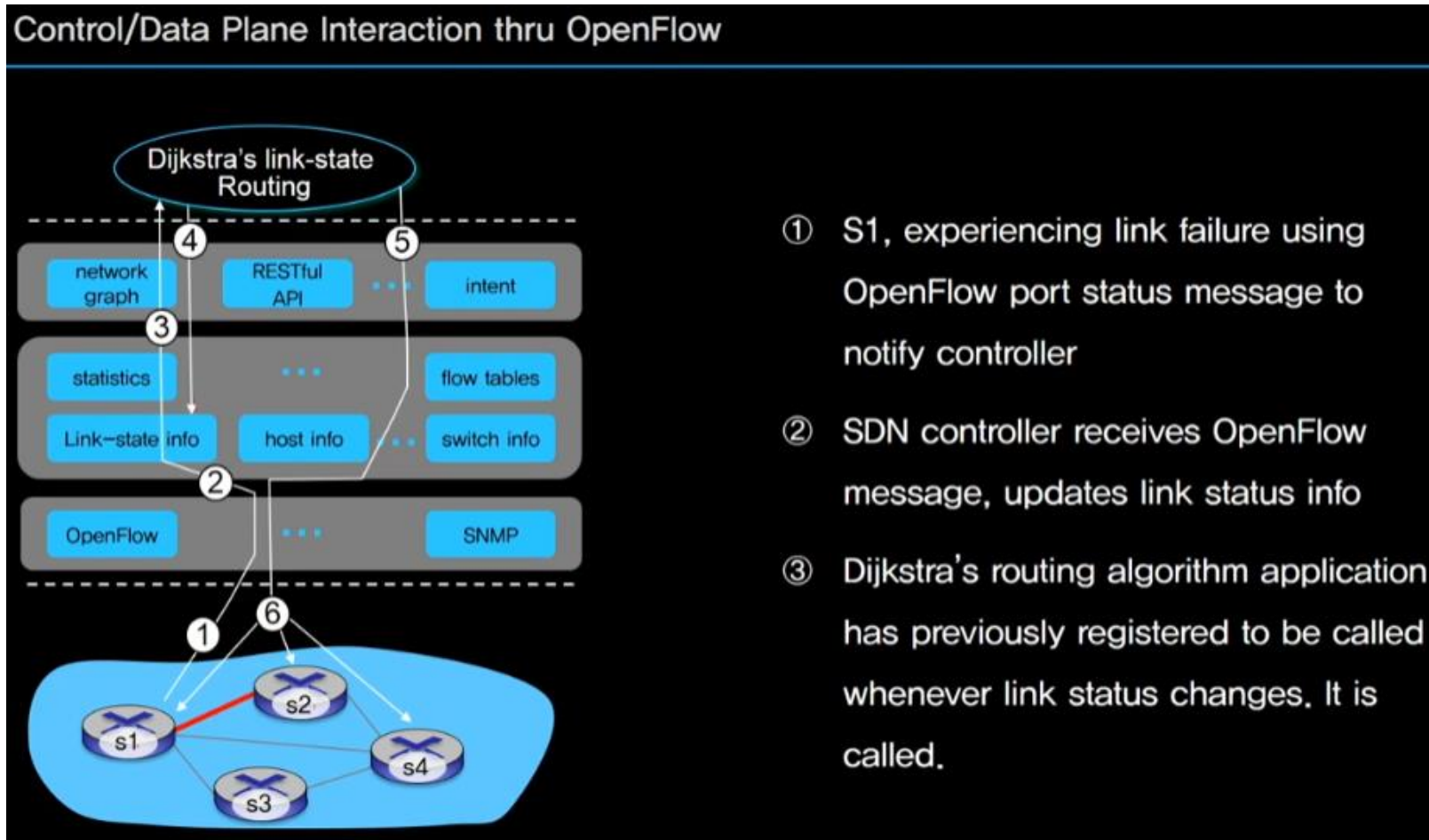
2. 스위치 to controller (Asynchronous)

■ 메시지 종류

- ◆ Flow-removed : 스위치에서 특정 flow가 삭제가 됐다
- ◆ Port status : 스위치가 가지고 있는 특정 포트에 문제가 발생했다든지 등 포트상태
- ◆ Packet-in : 자기가 갖고 있는 flow table 기반으로 패킷을 전달할 수 없는 경우 컨트롤러에게 직접 패킷을 전달함. 새로운 flow의 데이터 그램을 발견했음을 알리는 역할도 함.

3. Symmetric (misc) : TCP 연결을 설립하려고 할 때 사용.

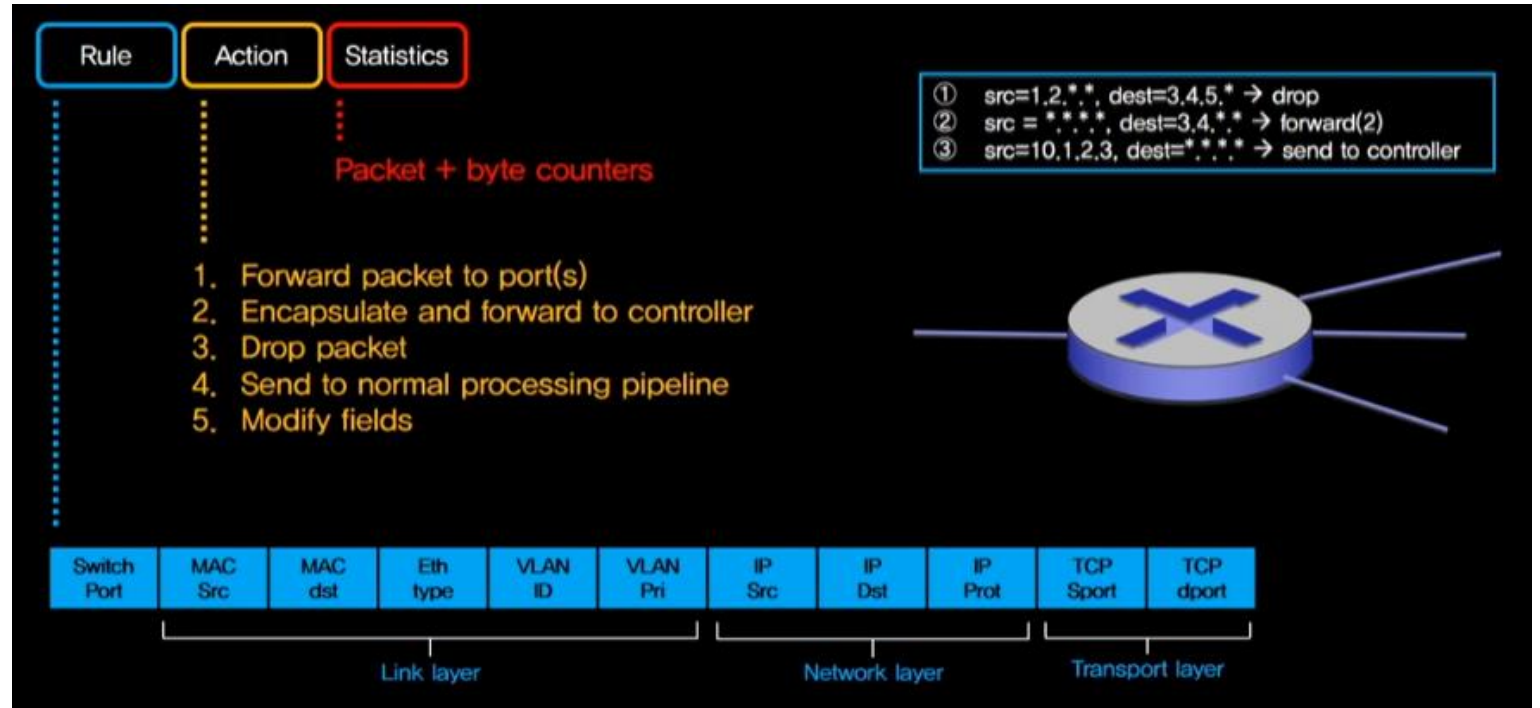
OpenFlow



S1 - S2 연결이 끊긴 경우 링크 스테이트 정보에 알리고 네트워크 그래프 수정하고 라우팅 알고리즘에게 알린다. 그러면 라우팅 알고리즘은 수정된 정보를 기반으로 새로운 길을 찾고 새로운 flow table을 만든 다음, modify state 메시지를 통해 스위치들에게 이를 전달한다.

OpenFlow

- Openflow table entry



- 크게 rule, action, statistics로 이루어져 있음.
- 기존 네트워크로는 dest 주소만 보고 drop이나 특정 포트로 포워딩이 불가능 했는데 openflow에서는 그것이 가능하다.
- 물론 rule에서 어떤 요소를 골라내는지에 따라 destination based forwarding도 가능하다.
Ex) MAC 주소를 가지고 destination based forwarding이 가능하다.
- 그리고 firewall 장비 없이도 firewall이 구현 가능하다. 예를 들어 특정 포트 번호를 쓰는 프로세스가 굉장히 중요해서 외부의 접근을 막아야 한다면, 그 번호로 접근하는 모든 데이터그램을 drop 한다든지 할 수 있다

OpenFlow

- Match plus Action

Device	Match	Action
Router	longest matching with destination IP prefix	forward packets out of a link
Switch	destination MAC address	forward or flood
Firewall	IP addresses and port numbers	permit or deny
NAT	IP addresses and port numbers	rewrite address and port

이런 기능들이 기존 라우팅에서는 별도로 있었는데 이걸 하나로 만들어서 하나의 SDN 라우터를 가지고 이것들을 수행할 수 있다.

ICMP (Internet Control Message Protocol)

- 데이터를 전달할 때 중간에 여러 호스트와 라우터들을 거치게 되는데, 이런 호스트와 라우터들이 에러 리포팅을 하기 위해 사용함.
- 인터넷 프로토콜은 신뢰성을 보장하는 프로토콜이 아님. 그래서 TCP라는 프로토콜이 만들어졌음.
- ICMP 메시지는 IP 데이터 그램 위에서 동작한다.

■ ICMP message: type, code plus first 8 bytes of IP

datagram causing error



출처 -

<https://www.google.co.kr/url?sa=i&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwijnr5zbTcAHUfKp0KHWyWBGMDjRv68AgBEAU&usq=https%3A%2F%2Fpcisco.com%2Finternet-control-message-protocol%2F&psig=AOvVaw2kzsH0VPsH2cgsKC2thm5&ust=1532413959697084>

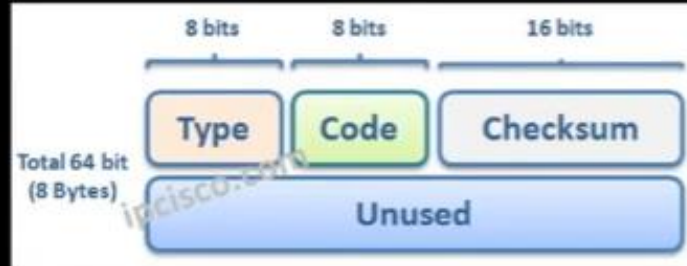
■ ICMPv6: new version of ICMP

- additional message types, e.g. "Packet Too Big"
- multicast group management functions

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control – not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ICMP (Internet Control Message Protocol)

- **ICMP message:** type, code plus first 8 bytes of IP datagram causing error



출처 -
<https://www.google.co.kr/url?sa=i&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwjrcz5zbTcAhUfKpQKHwyWBGMDQjRx68AgBEAU&url=https%3A%2F%2Fpcisco.com%2Finternet-control-message-protocol%2F&psig=AOvVaw2kzsH0VPsh2cgsKC2thm5&ust=1532413959697084>

Type	Code	description
0	0	echo reply (ping)
3	0	dest. network unreachable
3	1	dest host unreachable
3	2	dest protocol unreachable
3	3	dest port unreachable
3	6	dest network unknown
3	7	dest host unknown
4	0	source quench (congestion control – not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

- **ICMPv6:** new version of ICMP
 - additional message types, e.g. “Packet Too Big”
 - multicast group management functions

- ICMPv6 : IPv4에서는 fragmentation, reassembly가 일어났는데, IPv6으로 오면서 fragmentation, reassembly가 중간에 라우터에게 오버헤드가 되었다.
- 그래서 packet too big 이라는 코드를 추가함으로써 라우터가 전달할 수 없을 만큼 큰 패킷이 오면 버려버리고, source에게 메시지를 전달해주면 source가 전달하려는 fragmentation이 크다는 것을 인지하고, 그 다음부터 작게 보내게 된다.

Keyword

- Autonomous System (AS): 명확한 라우팅 정책을 가지고 동일한 네트워크 관리 주체에 의해 관리되는 네트워크, 네트워크의 집합 또는 도메인(domain)
- AS 내부(Intra-AS) 라우팅: 같은 AS 내에 있는 라우터 간의 라우팅
- AS 간(Inter-AS) 라우팅: 서로 다른 AS에 속한 라우터 간의 라우팅으로 사실상 표준인 Border Gateway Protocol (BGP)가 대표적임. 다른 도메인 간에 라우팅이 이루어지므로 성능보다 라우팅 정책이 중요함.
- 소프트웨어 정의 네트워크(Software-Defined Network, SDN): 소프트웨어 기반의 중앙집중형 컨트롤러에 의해 유동적인 라우팅, 부하 균형(load balancing) 및 접근 제어가 가능한 새로운 네트워크 구조
- 오픈플로우(OpenFlow): SDN 컨트롤러와 스위치 간의 통신 규약을 정의하는 대표적 Application Programming Interface (API)
- 인터넷 제어 메시지 프로토콜(Internet Control Message Protocol, ICMP): 라우터를 포함한 네트워크 장비가 제어 정보를 보낼 때 사용하는 프로토콜로서 주로 오류 보고(error reporting) 메시지를 보냄

감사합니다