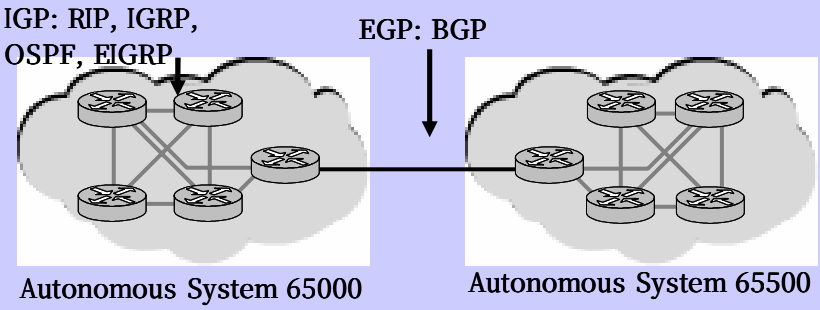
Chapter 8

Configuring Basic Border Gateway Protocol

1

Autonomous System



- AS(Autonomous System)란 기술적으로 단일한 관리체제 안에 있는 네트워크의 집합이다.독자적인 라우팅 정책(Routing Policy)과 IGP를 운영한다.
- IGP는 AS 안에서 작동하는 라우팅 프로토콜이다.
- EGP는 AS와 AS를 연결하는 라우팅 프로토콜이다.

2

AS(Autonomous System)

▪ IANA(Internet Assigned Number Authority)가 AS Number를 할당한다. EIGRP, IGRP 등은 AS 번호를 관리자가 임의로 할당하나 BGP의 경우에는 IANA에서 AS 번호를 할당 받아야 한다.

▪ Autonomous System 지시자는 16-Bit Number로 1-65535의 범위를 갖는다. 64512 – 65535는 Private AS 번호로 예약되어 있다. (65535-1024=64512 ~ 65535)

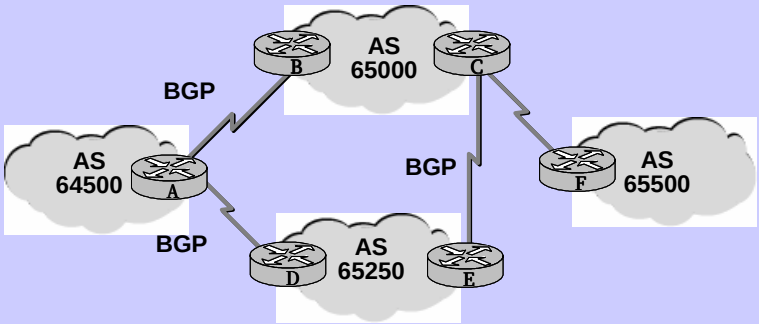
▪ EGP는 AS 사이에 Packet을 Routing 하기 위해서 사용된다. BGP는 EGP 중의 하나로 BGP의 최근 버전은 BGP-4이다.(BGP-1:1989년:RFC 1105, BGP-4:1993년:RFC 1771)

▪ BGP4는 현재의 Internet에서 ISP와 ISP를 연결하기 위해서, Enterprise(기업)를 ISP와 연결하기 위해서 널리 사용된다.

▪ IANA에서 할당한 AS 번호는 BGP와 같은 EGP를 사용하는 데 필요하다.

3

BGP Use



▪ AS간에 BGP가 사용된다. BGP는 AS간에 Routing Information을 교환하는 데 사용되는 Path Vector Protocol이다.
▪ BGP는 Loop-Free한 Routing Information의 교환을 보장한다.

4

Scalable Routing Protocol Comparison

Protocol	Interior or Exterior	DV or LS	Hierarchy Required	Metric
OSPF	Interior	LS	Yes(Area)	Cost
EIGRP	Interior	Advanced DV	No	Composite
BGP	Exterior	Advanced DV	No	Path Attributes (=Path Vectors)

5

When to Use BGP

- Packet이 AS 자신을 경유하여 다른 AS로 가도록 허용하는 경우(Internet Service Provider의 경우)
- 하나의 AS가 다른 AS 들에 Multiple Connection을 갖는 경우
- AS를 경유하는 Traffic의 흐름이 적절하게 통제 되어야 하는 경우
- BGP에 대해서 네트워크 관리자 들이 잘 이해하고 있는 경우
- 현대 네트워크의 Working Guide Line에 의하면 OSPF 등 IGP의 Routing Table의 Limit은 2,000 -3,000으로 이 경우 BGP로 Segmentation 하는 것을 검토해 볼 필요가 있다. (BGP Segmentation 이외에도 IGP Segmentation /Summarization의 방법을 사용할 수도 있다.)

6

Policy Routing Protocol

▪ BGP는 AS Number가 다른 두 네트웍간에 Routing Information을 주고 받을 때 이용될 수 있는 Exterior Gateway Protocol이다.

▪ AS 내부에서는 BGP를 이용할 필요는 없다. RIP, IGRP, OSPF, EIGRP, IS-IS와 같은 Interior Gateway Protocol을 이용하는 것이 바람직하다.

▪ AS Number가 다른 두 네트웍을 연결할 때도 반드시 BGP를 이용할 필요는 없다. AS Number가 다른 두 네트웍을 연결할 때 재분배를 통하여 Routing Information을 주고 받을 수 있도록 할 수 있기 때문이다.

▪ 두개 이상의 AS에 접속하고 또한 독자적인 라우팅 정책을 수립하여 구현할 때 BGP가 필요하다. BGP는 Policy Routing Protocol이라고 하는데 Routing이 라우터에 의해 결정되는 것이 아니라 네트워크 운영자의 정책에 의해 결정되기 때문이다.

7

How Big Is the Internet?

▪ 121,598 Route 이상의 Routing Table을 갖는다.(2003년 4월 21일 기준, telnet route-server.cerf.net 하여 확인할 수 있다.)
▪ 모든 BGP Routing Table을 보유하기 위해서는 최소 64MB의 메모리를 필요로 하며 128MB 이상의 Memory가 권장된다.
▪ ISP로 부터 BGP Route를 AS로 Inject 하는 방법에는 Default Route만을 사용하는 방법, Default Route와 Partial Route를 받는 방법, Full BGP Route를 받는 방법 등이 있을 수 있다.
▪ 6,500개 이상의 AS Number가 사용된다.

```
C:\ >telnet route-server.cerf.net

route-server> sh ip route summary
Route Source        Networks    Subnets     Overhead    Memory (bytes)
connected                0           1           56               144
static                   1           3          224               576
bgp 1838             82493       25109      6025712          15721128
  External: 107602 Internal: 0 Local: 0
internal              1249                                    1453836
Total                87843       33755      6739544          18931992
```

8

When Not to Use BGP

- BGP가 항상 적정한 것은 아니다.
- Internet이나 다른 AS에 Single Connection을 갖는 경우는 BGP를 사용하는 것보다 Default Route또는 Static Route를 사용하는 것이 바람직하다.
- Routing Policy나 Route Selection이 AS의 주요한 관심사가 아닌 경우 BGP의 사용이 적절하지 않다.
- BGP Router는 지속적인 Route Update를 수행해야 하므로 Memory 및 Processor Power가 부족한 라우터에서는 BGP의 사용이 적절하지 않다.
- Network Administrator의 Route Filtering이나 BGP Path Selection Process에 대한 충분한 이해가 없는 경우 BGP의 사용이 적절하지 않다.
- AS 간의 연결에 Bandwidth가 충분하지 않은 경우도 BGP의 사용이 적절하지 않다. => Static Route를 사용한다.

9

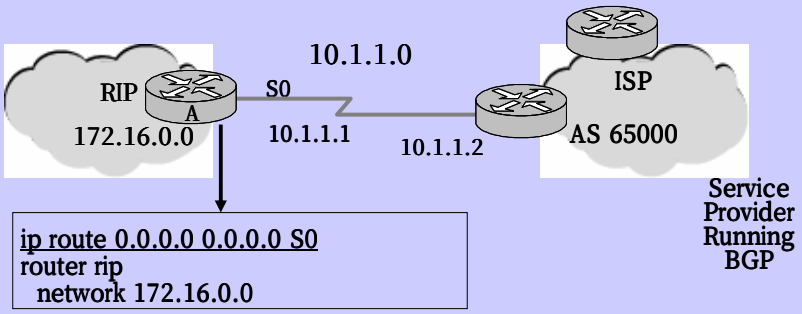
Static Route

Router(config)# ip route *prefix mask* {*address* | *interface* } [distance]

- Static Route의 Administrative Distance는 별도 설정하지 않는 경우, Next-Hop-Address를 사용하는 경우는 1이고, Interface를 사용하는 경우 0이다.
- Multiaccess Media(LAN, Frame Relay, X.25, ISDN 등)를 사용하는 경우에는 반드시 Next-Hop-Address를 사용해야 한다.
- Point to Point Media의 경우는 Interface를 사용할 수 있으며, 특히 Unnumbered Interface를 사용하는 경우에는 Interface를 사용할 수 밖에 없다.
- Floating Static Route는 Static Route의 Administrative Distance값을 Dynamic Routing Protocol의 값보다 크게 잡아, Dynamic Routing Protocol 사용이 불가할 때 Static Route를 사용하게 하는 방법이 된다.

10

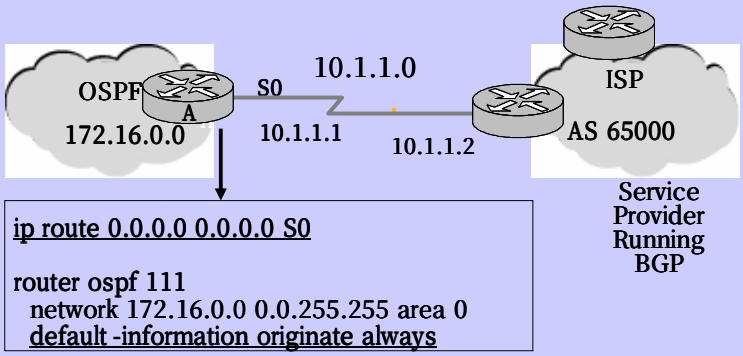
RIP Default Route Example



▪ RIP에서 Default Route는 자동으로 RIP Domain내로 전파된다.

11

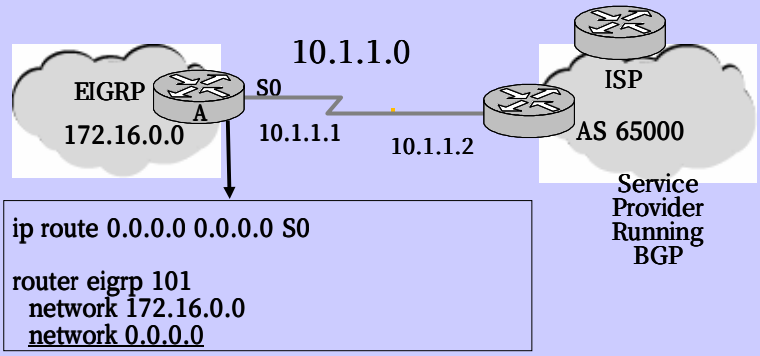
OSPF Default Route Example



▪ default -information originate always 명령은 OSPF가 Default Route를 OSPF Routing Domain에 전파시키도록 한다. Always 키워드는 위의 그림에서 s0가 다운되는 경우에도 Defualt Route를 전파시키라는 의미이다.

12

OSPF Default Route Example




```
ip route 0.0.0.0 0.0.0.0 S0

router eigrp 101
  network 172.16.0.0
  network 0.0.0.0
```

▪ network 0.0.0.0 명령은 EIGRP가 Default Route를 EIGRP Routing Domain에 전파시키도록 한다.

13

BGP Characteristics

▪ BGP는 Enhanced 된 Distance Vector Routing Protocol이다.

- Reliable Update: BGP는 TCP 상에서 작동한다. (TCP Port 179번을 사용한다.)
- BGP Router간에 처음 Connection을 맺을 때는 Full Routing Table을 교환하지만 이후에는 Incremental, Triggered Update 만을 한다.
- TCP Connectivity를 확인하기 위한 주기적인 Keepalive를 주고 받는다.
- 풍부한 Metric(Path Attributes)을 사용한다.
- Loop Free Routing Protocol이다.
- Huge Internetwork(=Internet)에서 사용할 수 있도록 Design 되었다.

14

Loop Free Routing Protocol

▪ AS_Path는 해당 목적지까지 경유하는 AS들의 Number들로 구성된다.

AS100 - AS200 - AS300 - AS400 (130.100.0.0)

• AS400에 있는 라우터는 AS300에 있는 라우터에게 130.100.0.0 네트워크에 대한 정보를 전달할때 AS_path로 400을 전달한다.

AS300 : 130.100.0.0 400

• AS300에 있는 라우터가 AS200에 있는 라우터에게 130.100.0.0에 대한 정보를 전달할때는 자신의 AS Number를 AS_path에 Prepend 시킨 AS_Path 300 400을 전달한다.

AS200 : 130.100.0.0 300 400

• AS200에서 AS100으로 정보가 전달 될 때도 같은 방식이 적용된다.

AS100 : 130.100.0.0 200 300 400

15

Loop Free Routing Protocol

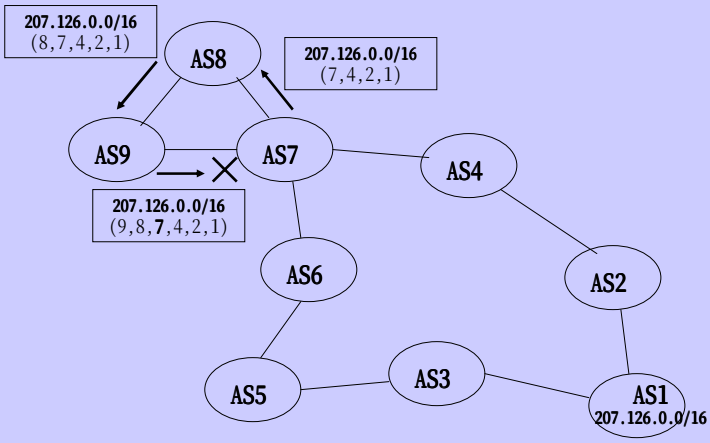
▪ AS_Path Routing : 동일한 목적지에 대한 AS_Path의 길이가 짧을 경우 가까운 곳에 있다고 판단하여 짧은 AS_Path를 가진 경로를 선택한다.

• AS_Path 300 400의 길이는 2, AS_Path 200 300 400의 길이는 3이다.

▪ Loop Free Routing Protocol : AS_Path에 자신의 AS Number가 표시되어 있다면 그것은 자신이 내보낸 정보이므로 해당 정보를 무시한다. 따라서 BGP Routing은 기본적으로 Routing Loop이 발생하지 않는다.

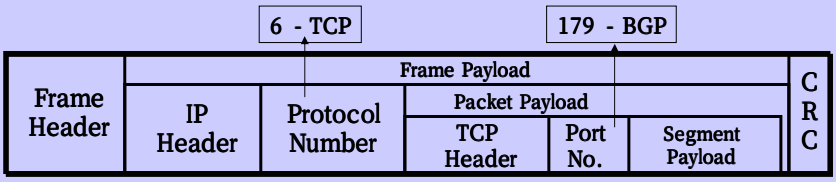
16

Loop Free Routing Protocol



17

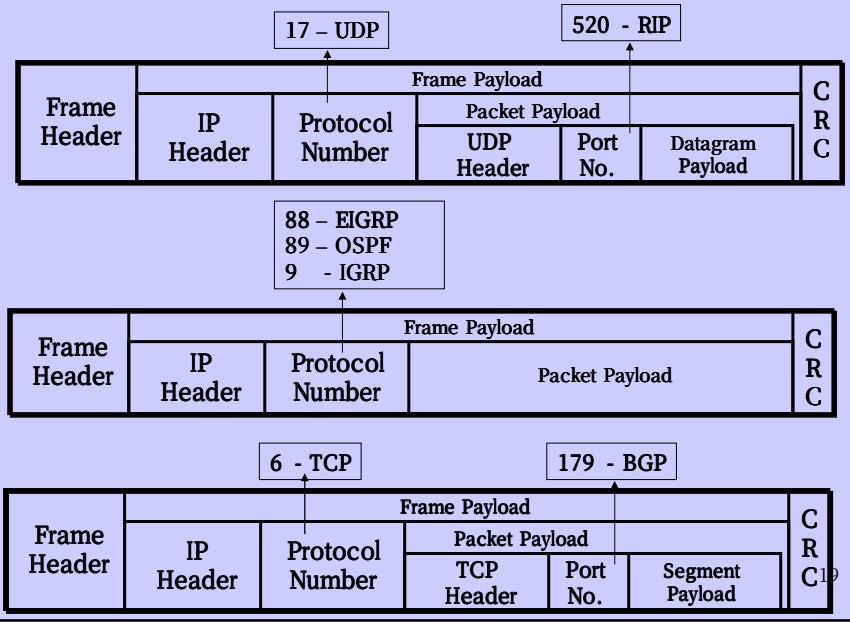
BGP in IP Packets



- BGP는 Enhanced 된 Distance Vector Routing Protocol이다.
- Reliable update: BGP는 TCP 상에서 작동한다. (TCP Port 179번을 사용한다.)

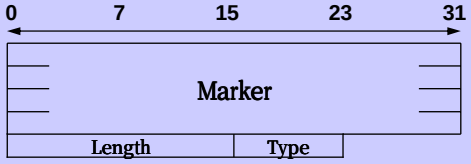
18

RIP, IGRP, OSPF, EIGRP, BGP PDU Diagram



19

BGP Message Header Format



- BGP Message는 16byte Marker와 2byte Length, 1byte Type 필드로 되어 있다.
- Marker : 전달 받은 BGP Message에 대한 인증에 사용되거나, BGP Peer간의 연결이 끊어지는 것은 감지한다.Message Type이 Open인 경우에는 인증에 관련된 내용이 없고 Marker는 모두 1로 채워진다.이 외의 경우에는 인증과정에 사용된다.
- Length : 헤더를 포함한 BGP Message의 길이,가장 작은 크기는 19byte(16byte + 2byte + 1byte) 이며 4096byte를 넘을 수 없다.
- Type : Message의 종류를 구분한다.
  - Open
  - Update
  - Notification
  - Keepalive

20

BGP Message Types

▪ Open :TCP Connection이 맺어진 후 BGP 라우터 간에 처음으로 교환되는 메시지이다. Version, Sender의 AS Number,Hold Time, BGP Identifier (Router ID,OSPF의 Router ID와 동일하다), Optional Parameter (Authentication 등) 정보를 담고 있다. 상호간에 Open Message가 확인되면 (OpenSent 및 OpenConfirm) BGP Session이 맺어지고 이후 Update, Keepalive, Notification Message가 교환된다.

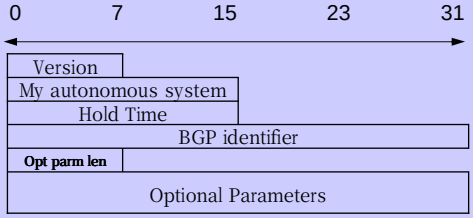
▪ Keepalive : BGP Peer간에 BGP Connection이 Alive한지를 확인하기 위한 Messsage이다. Peer간 Keepalive는 Default로 60초이며, Hold Time은 Default로 180초이다.

▪ Update :단지 하나의 Path에 대한 정보로 Path Attributes 및 Network을 정보로 담고 있다.NLRI(Network Layer Reachability Information), Path Attributes, Withdrawn Routes(Multiple Unfeasible Route)등의 정보를 담고 있다. BGP Peers간에 연결이 이루어지면 처음에는 Full Routing Table을 교환하고 이후에는 Incremental Update를 한다.

▪ Notification :Error가 Detection 되면 BGP Router는 Notification Message를 보낸 후 BGP Connection을 종료한다.(모든 BGP State에서 발생이 가능하다.)

21

OPEN Message Format



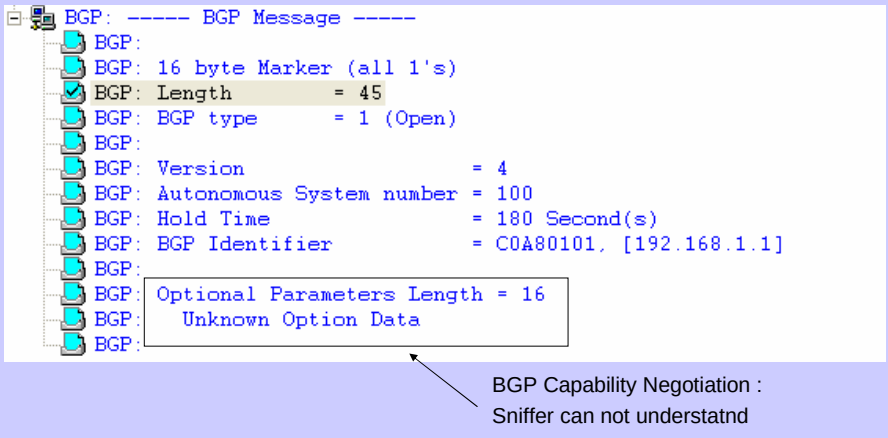
22

OPEN Message Format

- Version : 서로 지원할 수 있는 가장 높은 버젼을 사용하기위해 Negotiation한다. 호환되는 버젼이 정해지기 까지 BGP세션을 reset하고 negotiation한다. Cisco라우터에서는 negotiation을 빠르게 하기 위해 버젼을 정해 놓을 수 있다. (버젼에 대한 Neogotiation과정이 없어진다)
- My Autonomus System : BGP speaker가 속한 AS 번호
- Hold Timer : KEEPALIVE나 UPDATE 메시지를 받은 뒤 다음 메시지를 받기 전까지 허용할 수 있는 최대시간으로 0부터 시작하며 KEEPALIVE나 UPDATE를 받으면 다시 0이 된다.특정 Neighbor가 Hold Time을 넘어서면 더 이상 BGP 정보를 주고 받지 않는 것으로 간주한다.Hold Time 이 0이면 항상 세션이 유지되며 0으로 설정할 수 없는 경우 3초가 가장 낮은 값이다.네이버 간에 설정된 가장 낮은 값으로 정해진다. Version, Hold Time모두 Open message를 주고 받으며 협상을 하는 것은 같지만 Hold Time을 협상 할 때는 Reset되지 않는다. 반면에 Version을 협상할 때는 Reset 된다.
- BGP Identifier : 라우터에 설정된 물리 인터페이스의 가장 높은 IP Address이거나 Loopback 인터페이스의 가장 높은 IP Address가 선택된다. 직접 설정할 수도 있다. (**bgp router -id** *<ip address>*)
- Optional Parameter Legnth (Opt Parm Len) : 0은 Optional Parameter 가 없다는 것을 나타낸다.
- Optional Parameters : <Parameter Type, Parameter Length, Parameter Value> 의 형식이며, 1byte, 1byte, length에서 지정된 값의 길이를 갖는다. Optional Parameter의 예로 Authentication Information (type 1)이 있다.

23

Open Message Example



24

BGP State

- Idle : BGP가 Start 되었으나, Peer간에 Connection이 없는 상태이다.
- Connect : TCP Connection이 완료되기를 기다리는 상태이다.
- Active : TCP Connection Retry를 하고 있는 상태이다.
- OpenSent : Peer로부터 Open Message를 기다리고 있는 상태이다.
- OpenConfirm : Peer로부터 Open Message를 받고, Keepalive Message를 기다리고 있는 상태이다. Keepalive Message를 받게 되면 Established State가 된다.
- Established : Peer로부터 Keepalive Message를 받은 상태로 이 때부터 Update Packet을 주고 받을 수 있다.

- 참고 : TCP 연결이 실패하는 경우 해당 Neighbor에 대한 State는 Idle - Connect -Active 상태를 계속 순환한다.

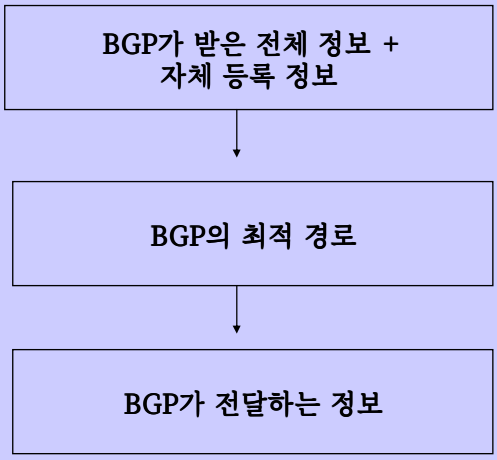
25

BGP의 정보 전달 원칙

- **BGP 가 전달하는 정보**
  - NLRI: Network Layer Reachability Information : Prefix 정보,해당 Prefix에 대한 특성(Attribute) 정보
  - Unfeasible Routes: 자신의 BGP 테이블에서 사라진 Prefix
- **BGP의 정보 전달 원칙**
  - BGP는 자신이 사용하는 정보만을 전달한다.
  - BGP Speaker 입장에서 Best 경로만을 전달한다.
- **정보 전달의 정확성 유지 방법**
  - TCP의 기능(Reliable Transport)을 활용한다.
  - BGP 차원의 ACK는 없다.

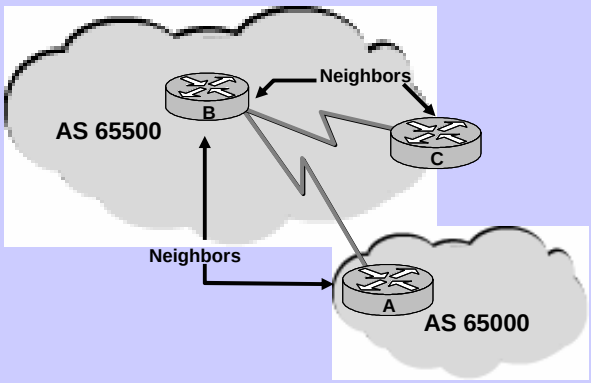
26

BGP 동작 원리



27

Peers = Neighbors



▪ BGP Routing Information을 교환하기 위하여 TCP Connection 을 맺는 두 라우터를 Peer 또는 Neighbor라 한다.

28

Basic BGP Command

▪ BGP Routing Protocol을 Enable 시킨다.

Router(config)#router bgp *autonomous-system*

Router(config-router)#neighbor {*ip-address* | *peer-group-name*} remote-as *autonomous-system*

▪ EBGP는 Default로 TTL 값을 1로 설정하여 Message를 교환한다.(IBGP는 TTL 값 255를 사용한다.) 1홉 이상 떨어져 있는 EBGP Router간의 연결을 위해서는 neighbor {*ip-address* | *peer-group-name*} ebgp-multihop [ttl] 명령을 추가로 사용한다. Loopback Interface를 사용하여 Neighbor를 설정하는 경우도 ebgp-multihop 키워드를 사용하여야 한다. ebgp-multihop 키워드를 사용하면 TTL 값이 255로 설정되며 TTL 값을 명시적으로 한정할 수도 있다.

▪ Neighbor 설정시, Loopback Interface를 사용하는 경우에는 추가로 neighbor {*ip-address* | *peer-group-name*} update-source loopback *interface-number* 명령을 사용하여 neighbor에 대해 알려 주어야 한다.

29

Basic BGP Command

▪ BGP는 BGP Session을 맺기 위해Transport Protocol로 Transmission Control Protocol (TCP)을 이용한다. BGP Session을 맺은 상대측 라우터를 Neighbor 혹은 Peer 라고 하는데, TCP Session을 사용하므로 BGP에서 Neighbor간에 직접 연결되어 있을 필요는 없다.

▪ Neighbor의 IP Address를 지정할 때 유의하여야 할 점은 해당 Neighbor까지 갈 수 있는 경로가 Routing Table에 등록되어 있어야 한다는 점이다. 일반적으로 EBGP Session은 다른 AS에 있는 라우터와 직접 연결한 뒤 맺어지기 때문에 항상 Neighbor가 Routing Table에 등록이 되어 있는 반면, IBGP Session은 직접 연결되어 있지 않은 내부 라우터간에도 맺어지기 때문에 잘못 조작을 하였을 경우 Neighbor가 Routing Table에 등록되어 있지 않을 수도 있다. Neighbor를 Routing Table에 등록하는 방법은 IGP를 정확히 조작해 주거나 Static Route등을 등록하면 된다.

30

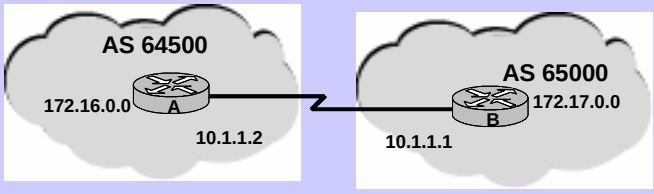
Basic BGP Commands

Router(config-router)# network *network-number* mask *network-mask*

- 라우터에 직접 연결된 네트워크만을 의미하는 것이 아니라, Advertise하려는 모든 네트워크로 자신의 AS에 속한 Network Address를 열거한다. IGP Routing Table에 있는 Route 만을 선언할 수 있다.(참고:Aggregate는 BGP Table에 있는 Specific Route를 Summarize할 수 있다.)
- Prefix는 IP Routing Table의 Entry와 반드시 일치하여야 한다.즉 라우팅 테이블에 존재하는 것과 동일한 주소가 입력되어야 한다. 그렇지 않으면 Injection되지 않는다.(명령은 입력되나 BGP Table에 Inject 되지 않는다.)
- BGP는 Default로 Classful Prefix를 지원하나 no auto-summary 명령을 사용하면 Classless Prefix를 지원하게 된다.
- Subnetting 및 Supernetting 정보를 전달하려면 mask 파라미터를 사용하여야 한다. (Network 명령 뒤에 mask를 사용하지 않은 경우, Auto-summarization이 활성화 되어 있으면 Classful Boundary로 Summarization된 네트워크가 Injection된다. no auto-summary 명령어가 사용된 경우에는 모든 Prefix는 mask를 사용하여 입력하여야 하며 그렇지 않으면 Injection되지 않는다. Auto-summarization이 활성화 되어 있는 경우에도 mask를 사용하면 Subnet을 Inject 시킬 수 있다.)
- IOS 12.0 Version 이하에서는 200개까지의 Network을 선언할 수 있으며 12.0 이상에서는 NVRAM 또는 RAM의 용량에 제한될 뿐이다.

31

BGP Example




```
RtrA(config)#router bgp 64500
RtrA(config-router)#neighbor 10.1.1.1 remote-as 65000
RtrA(config-router)#network 172.16.0.0
```

```
RtrB(config)#router bgp 65000
RtrB(config-router)#neighbor 10.1.1.2 remote-as 64500
RtrB(config-router)#network 172.17.0.0 mask 255.255.0.0
```

32

Resetting BGP

Router#clear ip bgp {* | *address* | *peer-group*} [soft [in | out]]

- BGP Policy Reconfiguration시에 수행해 주어야 한다.
- *은 모든 Neighbor를, address는 특정 Neighbor를 지정한다.
- soft Option 없이 clear ip bgp 명령을 수행하면 BGP Table의 Entry를 제거하고, BGP Session을 Reset 시킨다.
- Soft Option을 사용하면 BGP Session을 Reset하지 않고, 정책변경과 관련된 라우팅 테이블만 다시 Update한다.in은 Inbound, out은 Outbound Update를 의미한다.
- Soft Inbound Reconfiguration은 Memory Overhead를 유발하므로 Neighbor Router에서 Soft Outbound Reconfiguration 하는 것이 바람직하다.

33

BGP Connection Reset이 필요한 BGP Policy Reconfiguration

- BGP 관련 Access List의 변경 또는 추가
- BGP 관련 Weight의 변경
- BGP 관련 Distribution List의 변경
- BGP 관련 Timer 설정의 변경
- BGP Administrative Distance의 변경
- BGP 관련 Route Map의 변경

34

BGP Policy Reconfiguration

▪ BGP는 자신이 결정한 최적 경로 집합(Loc -RIB)에 속하는 정보 만을 전달한다.따라서, BGP의 최적 경로에 변화가 오지 않는다면 Update는 발생하지 않는다.

▪ Outbound Policy 변화 시에도 Update는 기본적으로 발생하지 않는다. outbound soft -reconfiguration을 적용하여야 새로 갱신된 Update가 전달된다.

▪ Inbound Policy를 바꾸어도 BGP의 Decision Process는 새로 구동되지 않는다. 새로운 Neighbor가 추가된 경우는 해당 Neighbor에서 새로운 정보를 받으므로 Decision Process가 적용된다.

▪ Inbound Policy 변화시 이를 적용하려면 BGP Session Clear , inbound soft -reconfiguration, Route Refresh 등을 사용할 수 있다.

35

BGP Soft Reconfig

▪ **clear ip bgp a.b.c.d soft out**
• CPU 적게 차지한다.
• 현재 Advertised된 Route를 취소하고 Route를 재전송한다.

▪ **clear ip bgp a.b.c.d soft in**
• CPU를 많이 차지한다.
• 받은 루트 정보를 모두 유지해야 한다.
• **clear ip bgp a.b.c.d soft in 명령을 수행하기 이전에 neighbor x.x.x.x soft -reconfiguration inbound를 설정해야 한다**.(일반적으로 BGP Configuration 시에 soft -reconfiguration -inbound를 미리 입력시켜 놓는다.)

36

Verifying BGP

Router#show ip bgp [summary | neighbors]

- show ip bgp는 BGP Routing Table을 보여준다.
- show ip bgp summary 명령은 BGP Connection의 Status를 보여준다.
- show ip bgp neighbors는 Neighbor에 대한 TCP 및 BGP Connection을 보여준다.

Router#debug ip bgp

- keepalive, update, events 등의 Option이 있다.

37

show ip bgp

```
RTRA#show ip bgp
BGP table version is 5, local router ID is 192.168.2.1
Status codes:s suppressed,d damped,h history,* valid,> best,i  - internal
Origin codes: i  - IGP, e  - EGP, ?  - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 172.16.0.0       10.1.1.1                               0 65000 i
*> 192.168.1.0      10.1.1.1                 0             0 65000 i
*> 192.168.2.0      0.0.0.0                  0         32768 i
```

Origin codes

- BGP테이블에 변화가 있을 때 마다 테이블 버전이 증가된다. 빠르게 증가하면 네트워크가 불안정하다는 것을 나타낸다.

```
R1# show ip bgp 192.168.0.0
BGP routing table entry for 192.168.0.0/24, version 2
Paths: (1 available, best #1)
  65000
    192.168.1.254 from 192.168.1.254
      Origin IGP, metric 0, localpref 100, valid, external, best
```

38

show ip bgp summary

```
RTRA#show ip bgp sum
BGP table version is 5, main routing table version 5
3 network entries and 3 paths using 363 bytes of memory
3 BGP path attribute entries using 372 bytes of memory
BGP activity 3/0 prefixes, 3/0 paths
0 prefixes revised.

Neighbor  V      AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
10.1.1.1  4 65000      14      13        5    0    0    00:08:03        2
```

39

show ip bgp neighbors

```
RTRA#sh ip bgp nei
BGP neighbor is 10.1.1.1,  remote AS 65000, external link
 Index 1, Offset 0, Mask 0x2
  BGP version 4, remote router ID 172.16.10.1
  BGP state = Established, table version = 5, up for 00:10:47
  Last read 00:00:48, hold time is 180, keepalive interval is 60 seconds
  Minimum time between advertisement runs is 30 seconds
  Received 16 messages, 0 notifications, 0 in queue
  Sent 15 messages, 1 notifications, 0 in queue
  Prefix advertised 1, suppressed 0, withdrawn 0
  Connections established 1; dropped 0
  Last reset 00:16:35, due to Peer closed the session
  2 accepted prefixes consume 64 bytes
  0 history paths consume 0 bytes
 - -More - -
```

- **show ip bgp neighbors**는 Neighor에 대한 BGP Connection정보를 보여준다.

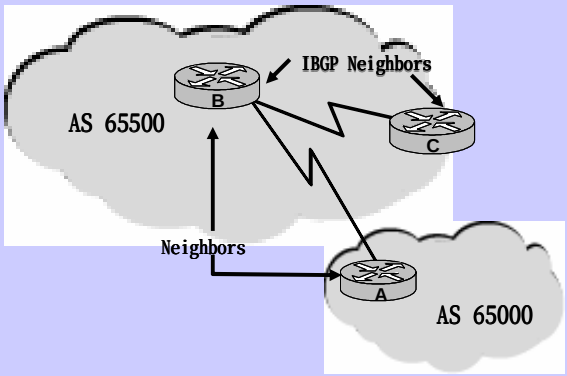
40

debug ip bgp

```
RTRA#debug ip bgp updates
BGP updates debugging is on
RTRA#clear ip bgp *
3w5d: BGP: 10.1.1.1 computing updates, neighbor version 0, table
version 1, starting at 0.0.0.0
3w5d: BGP: 10.1.1.1 update run completed, ran for 0ms, neighbor
version 0, start version 1, throttled to 1, check point net 0.0.0.0
3w5d: BGP: 10.1.1.1 rcv UPDATE w/ attr: nexthop 10.1.1.1, origin i,
aggregated by 65000 172.16.10.1, path 65000
3w5d: BGP: 10.1.1.1 rcv UPDATE about 172.16.0.0/16
3w5d: BGP: nettable_walker 172.16.0.0/16 calling revise_route
3w5d: BGP: revise route installing 172.16.0.0/16  -> 10.1.1.1
3w5d: BGP: 10.1.1.1 rcv UPDATE w/ attr: nexthop 10.1.1.1, origin i,
metric 0, path 65000
3w5d: BGP: 10.1.1.1 rcv UPDATE about 192.168.1.0/24
3w5d: BGP: nettable_walker 192.168.1.0/24 calling revise_route
3w5d: BGP: revise route installing 192.168.1.0/24  -> 10.1.1.1
```

41

IBGP (Internal BGP)

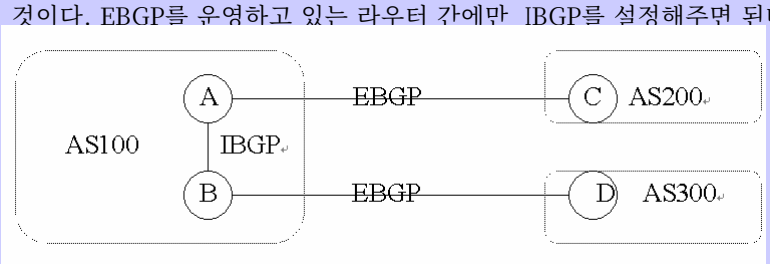


- BGP Neighbor가 동일한 AS에 속한다.
- Neighbor는 직접 연결(Direct Link)이 아니어도 된다.

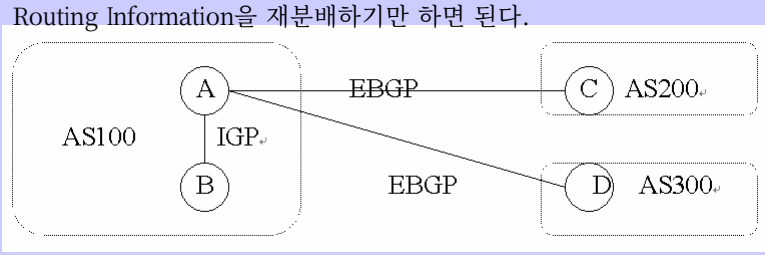
42

IBGP (Internal BGP)

▪ IBGP는 한 AS내에 다른 AS와 연결된 2개 이상의 라우터가 있을 경우에 설정한다. 즉, EBGP를 운영하고 있는 라우터가 2개 이상일 경우일 때만 필요하다는 것이다. EBGP를 운영하고 있는 라우터 간에만 IBGP를 설정해주면 된다.



▪ EBGP를 운영하고 있는 라우터가 1개일 경우 IBGP 설정은 불필요하다. BGP 라우터는 AS 내의 다른 라우터들과는 IGP로 정보를 교환하고, BGP와 IGP간에 Routing Information을 재분배하기만 하면 된다.

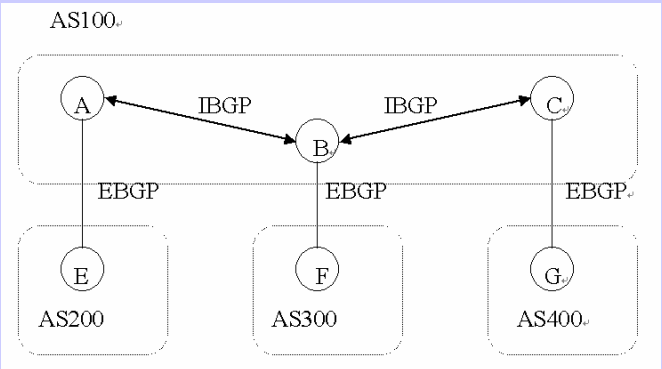


43

IBGP (Internal BGP)

▪ IBGP의 특성중 하나는 IBGP 라우터가 같은 AS내에 있는 IBGP 라우터로부터 네트웍의 변화된 정보를 전달받으면 그 정보를 자신과 IBGP Session을 맺은 다른 IBGP 라우터에게 전달하지 않고 EBGP를 이용하여 다른 AS에 있는 EBGP 라우터에게만 전달한다는 것이다.

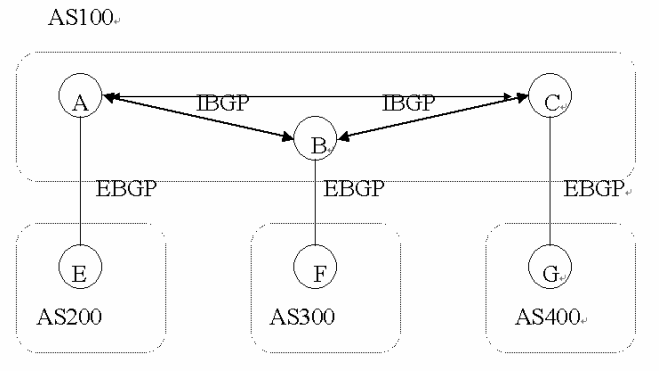
• AS200에서 어떤 변화가 발생되면 라우터 E는 그 사실을 라우터A에게 전달할 것이다. 네트웍의 변화에 대한 정보를 얻은 라우터A는 IBGP Session을 맺은 라우터B에게 전달한다. 라우터B는 정보가 같은 AS에 있는 라우터A로부터 왔으므로 그 정보를 IBGP Session을 맺은 라우터C에게 전달하지 않고 EBGP Session을 맺은 라우터 F에게만 전달한다. 따라서 변화된 정보가 라우터 C를 통해 AS400에까지 전달되지 않는다. 그러므로 라우터A와 라우터C간에 IBGP Session을 맺어 주어야만 한다.



44

IBGP (Internal BGP)

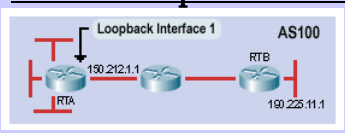
▪ AS내에서 EBGP 라우터간에는 Full Mesh 형태로 IBGP Session을 모두 설정한다. (Route Reflector 또는 Confederation을 사용하면 Full Mesh를 피할 수 있다.)



▪ 참고 : BGP Router가 IBGP로부터 알게 된 경로는 다른 IBGP에 Update 하지 않는 이유 : Route가 IBGP에 의해서 Advertise 되는 경우, 동일 AS 안에서 Advertise 되므로 AS -Path에 AS Number가 추가 되지 않는다. 따라서 EBGP Route와 달리 IBGP Route는 Loop을 근본적으로 방지하는 Mechanism이 없다. Loop 방지를 위해서 BGP는 IBGP Peer로 부터 Advertise 된 경로를 IBGP Peer에 Advertise 하지 않는다.

45

Use of Loopback Interfaces



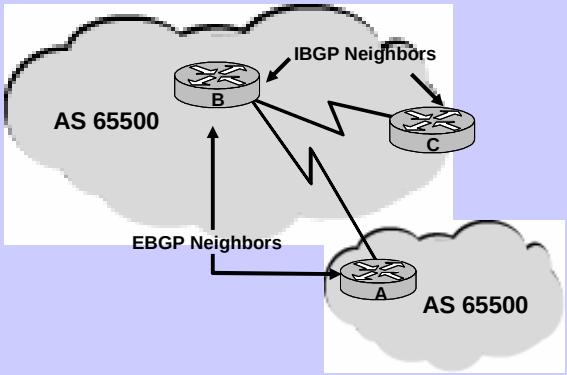

```
RTA#
loopback interface 1
ip address 150.212.1.1 255.255.0.0

router bgp 100
 neighbor 190.225.11.1 remote -as 100
 neighbor 190.225.11.1 update -source loopback 1
```

```
RTB#router bgp 100
 neighbor 150.212.1.1 remote -as 100
```

▪ Neighbor로 설정시 Physical Interface 중 하나를 설정할 수 있는데 이 경우 설정한 Interface가 Unavailable 해지면 BGP Session이 끊어지게 된다. Router A에서 Router B에 설정된 Loopback Interface를 사용하여 Neighbor를 설정하면 이러한 문제를 해결할 수 있다.
▪ Loopback Interface의 사용은 Router A에서 neighbor update -source라는 Router Configuration 명령의 사용을 필요로 한다.

46

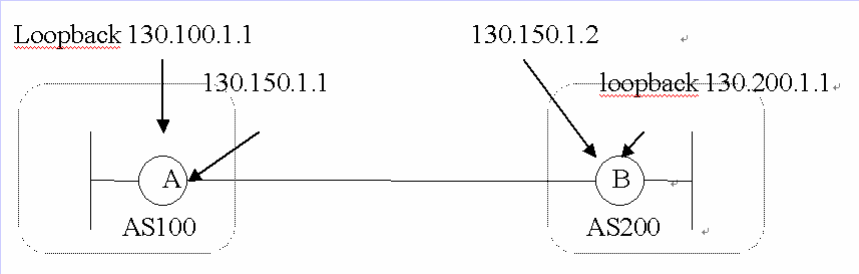
EBGP (External BGP)



- BGP Neighbor가 상이한 AS에 속한다.
- Neighbor는 직접 연결(Directly Connected) 되어야 한다.(직접연결이 아닌 경우는 neighbor ebgp -multihop 명령을 사용한 추가적인 Configuration이 필요하다.)

47

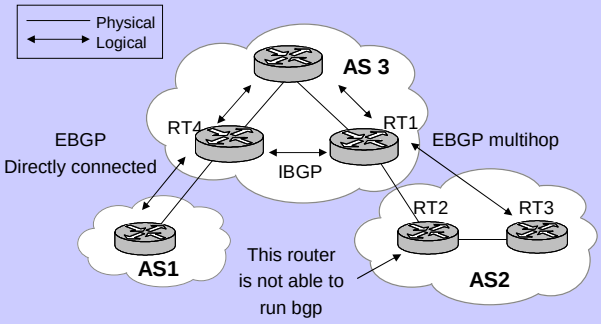
ebgp -multihop



- IBGP 라우터들은 직접 연결되어 있지 않아도 IBGP Session을 맺을 수 있다. EBGP 라우터들은 대개 전용회선과 같은 것으로 직접 연결되어 EBPG Session을 맺는데 어떠한 이유에서 EBGP 라우터들이 직접 연결될 수 없거나 Serial Link에 할당된 IP Address로 Neighbor를 지정할 수 없는 상황이 발생할 수 있다. 이러한 경우에는 **neighbor ebgp -multihop을** 선언해 주어야 한다.
- Frame Relay 망 등에서 PVC의 Serial Inerface IP Address를 사용하여 EBGP Session을 설정한 경우 PVC Fail시에 EBGP Session이 끊기게 된다. 이 경우에 Loopback Interface의 IP Address를 사용하여 EBGP Session을 설정하면 PVC Fail시에 우회 경로를 통하여 EBGP Session이 맺어질 수 있으므로 유용하다.

48

ebgp -multihop



49

EBGP와 Loopback Address

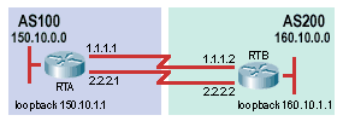
▪ Loopback Interface에 할당된 IP Address로 Neighbor를 선언하는 경우가 있는데(Serial Link가 끊기면 상대 라우터의 Loopback Interface에도 연결이 안되므로 EBGP 설정에 있어 Loopback Interface를 사용하는 것이 큰 의미는 없다.), 이 경우 neighbor의 IP Address를 선언해 주고 **neighbor ebgp -multihop을** 또한 선언해 준다.

▪ 주의하여야 할 사항은 EBGP 라우터가 Loopback Interface의 IP Address에 대한 경로를 알고 있어야 한다는 점이다. AS200에 있는 EBGP 라우터 B가 130.200.0.0(Loopback Network)에 대한 정보를 network 명령에 의하여 AS100의 EBGP 라우터 A에게 전달하는 경우는 130.200.0.0에 대한 정보를 AS100에 Static Route로 등록해 줄 필요는 없으나 그렇지 않은 경우는 , 라우터A에 130.200.0.0에 대한 Static Route 등을 등록하여 주어야 한다. 라우터B측에서도 130.100.0.0에 대한 Static Route 등을 등록하여 주어야 한다.

50

EBGP Load Balancing과 Loopback Interface

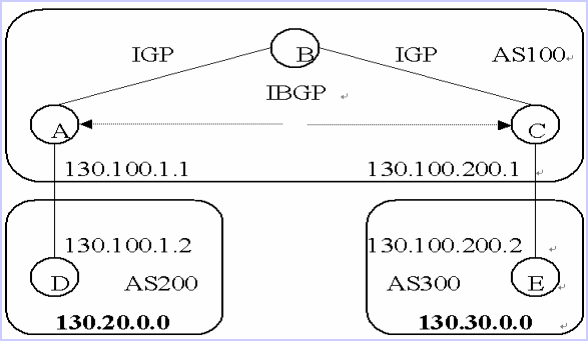
▪ **neighbor ebgp -multihop이라는** router configuration 명령과 loopback interface를 사용하는 것은 두 AS간의 Parallel Link를 통해 Load Balancing을 설정하는데 유용하다. Router A는 1.1.1.2 및 2.2.2.2를 경유하여 160.10.0.0 네트워크에 도달할 수 있다.



RTA	RTB
interface loopback 0 ip address 150.10.1.1 255.255.255.0 router bgp 100 neighbor 160.10.1.1 remote -as 200 neighbor 160.10.1.1 ebgp -multihop neighbor 160.10.1.1 update -source loopback 0 network 150.10.0.0 ip route 160.10.0.0 255.255.0.0 1.1.1.2 ip route 160.10.0.0 255.255.0.0 2.2.2.2	interface loopback 0 ip address 160.10.1.1 255.255.255.0 router bgp 200 neighbor 150.10.1.1 remote -as 100 neighbor 150.10.1.1 ebgp -multihop neighbor 150.10.1.1 update -source loopback 0 network 160.10.0.0 ip route 150.10.0.0 255.255.0.0 1.1.1.1 ip route 150.10.0.0 255.255.0.0 2.2.2.1

51

BGP Synchronization

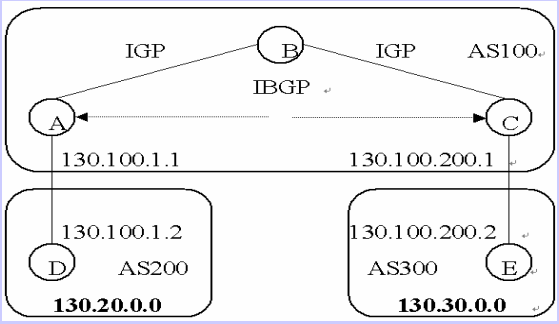


Rule of Synchronization

▪ IBGP Neighbor로 부터 알게 된 경로가 IGP Routing Table에 나타나고, 동시에 BGP Peer에 Advertise 되기 전에 그 경로는 반드시 IGP를 통해서 알게 되어야 한다.

52

BGP Synchronization



▪ 라우터A와 C는 라우터 B를 통하여 IGBP Session을 맺고 있다.라우터A가 라우터D로부터 130.20.0.0에 대한 정보를 전달 받으면 IBGP Neighbor인 라우터C에게 전달하며, 라우터C는 이 정보를 다시 라우터E에게 전달한다.
▪ 이후 AS300에 있는 라우터는 130.20.0.0으로 가려는 트래픽을 라우터C에게 전달할 것이다. 그리고 라우터C는 해당 트래픽을 라우터B에게 전달할 것이다. 그런데 아직 라우터B가 라우터A로부터 130.20.0.0에 대한 정보를 받지 못했다고 가정하면 라우터B는 해당 트래픽을 폐기해 버릴 것이다.
▪ 이러한 문제에 대해 BGP 라우터들은 두 개 이상의 AS를 연결하는 역할을 하는 AS 100이 있을 경우 AS100의 IBGP Router로부터 받은 목적지에 대한 경로를 IGP Router로부터 얻을때까지는 해당 경로를 다른 BGP 라우터에게 알려주지 않는데 이를 Synchronization이라고 한다.
▪ Synchronization이 활성화되어 있다면 라우터C는 라우터A로부터 받은 경로정보를 라우터B로부터 모두 전달 받을 때까지 기다렸다가 이후 라우터E에게 전달한다. 따라서 라우터B에서 트래픽이 폐기되는 것을 방지할 수 있게 된다.

53

Disabling BGP Synchronization

```
Router(config-router)#no synchronization
```

▪ BGP Synchronization 기능을 Disable 시킨다. IGP로 부터 루트를 Update 받기 전이라도 Route를 EBGP Router에 Update 한다.

54

no synchronization

▪ Synchronization은 Routing Information의 전달을 지연시키는 원인이 될 수 있는데, 다음과 같은 상황에서는 Synchronization은 필요하지 않을 수 있다.

자신의 AS가 다른 AS간의 트래픽을 중계하지 않을때 (Stub AS일때)
다른 AS간의 트래픽을 중계하는 라우터들이 모두 IBGP를 운영할때

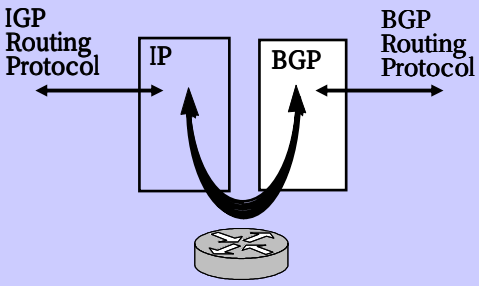
▪ CISCO 라우터의 BGP에서는 기본적으로 Synchronization이 활성화되어 있는데 Synchronization을 비활성화하는 방법은 다음과 같다.

Router(config)# router bgp 65000
Router(config -router)# **no synchronization**

▪ 참고 : 쥬니퍼 라우터의 경우는 BGP에서 Default로 Synchronization이 Disable되어 있다.

55

Routing Tables



▪ BGP는 IGP Routing Table과는 별도로 자신의 Routing Table을 보유한다.
▪ 두 테이블간에 Routing Information은 교환이 가능하다.

56

Advertising Networks *into* BGP

▪ AS 내부의 네트워크를 BGP로 Advertising 하는 방법에는 세가지 가있다.

• Dynamic IGP Route를 Redistributing한다. (Route에 대한 Instability를 초래할 수 있어 Recommended 되지 않는다.)
• Static Route를 만들고 Static Route를 Redistribution 한다.
• IGP Route(Dynamic Route 및 Static Route)를 Redistribution하지 않고 Network 명령을 사용하여 Advertise 한다

▪ Network Address를 등록해 주는 경우의 유의할 점은 예를 들어, 130.100.0.0이 BGP에 의해 외부 AS로 전달이 되려면 130.100.0.0에 대한 경로를 BGP 라우터가 알고 있어야 한다는 점이다. 즉 130.100.0.0에 대한 경로가 BGP 라우터의 Routing Table에 등록되어 있어야 된다.

57

Redistributing Dynamic IGP Routes into BGP

▪ Unstable Route를 유발시킬 수 있으므로 권고되지 않는다.(IGP Route의 모든 변화(Link Down등)가 BGP Update를 유발한다.)

▪ Dynamic Route를 Redistribution 하는 것은 BGP로 부터 받은 네트워크 정보가 다시 BGP로 재 유입되는 Mutual Redistribution을 발생시킬 수 있다. 이를 방지하기 위해서 Local Route만을 포함 시키고, 다른 AS에서의 Route는 Filtering 한다. => Configuration이 복잡해 질 수 있다.

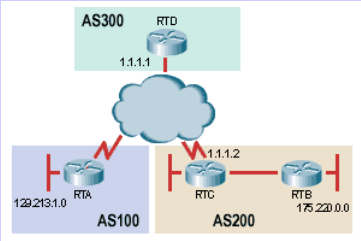
▪ Dynamic Route를 Redistribution 하는 경우, Path Attribute 중 Origin Attribute가 Incomplete 로 나타난다.(**show ip bgp** 명령으로 보면 ?로 나타난다.)

```
RTRA#show ip bgp
BGP table version is 5, local router ID is 192.168.2.1
Status codes:s suppressed,d damped,h history,* valid,> best,i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 172.16.0.0       10.1.1.1                               0 65000 i
*> 192.168.1.0      10.1.1.1                 0             0 65000 i
*> 192.168.2.0      0.0.0.0                  0         32768 i
```

Origin codes

58

Redistributing Dynamic Routes




```
RTC#

router eigrp 10
redistribute bgp 200
redistributed connected
default-metric 1000 100 250 100 1500

router bgp 200
neighbor 1.1.1.1 remote-as 300
neighbor 2.2.2.2 remote-as 200
neighbor 1.1.1.1 distribute-list 1 out
redistribute eigrp 10

access-list 1 permit 175.220.0.0 0.0.255.255
```

▪ **redistribute eigrp 10이라는** router configuration 명령은 EIGRP 10의 Route를 BGP로 Redistribute한다.(BGP를 IGP로 Redistribute 하는 것은 너무 많은 경로가 Inject 되므로 피해야 한다.) **neighbor distribute-list라는** router configuration 명령은 1.1.1.1 Neighbor(Router D)에 Outgoing Advertising을 하는데 있어, Access-List의 1의 적용을 받는다. Access list 1은 network 175.220.0.0만 Advertise 되도록 규정한다.AS 100으로부터 Update 되는 network 129.213.1.0을 포함한 모든 네트워크이 Advertise 되는 것을 금지한다. Dynamic Route를 Redistribution 하는 것은 BGP로 부터 받은 네트워크 정보가 다시 BGP로 재 유입되는 것(Mutual Redistribution)을 Access-List를 사용하여 방지하는 데 있어 세심한 주의를 필요로 한다. 가능하다면 network 명령을 사용한다.

59

Redistributing Static Routes into BGP

```
router bgp 64500
  redistribute static

ip route 192.168.1.0 255.255.0.0 ethernet 0
```

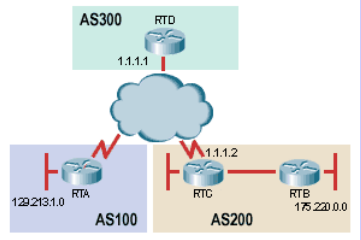
▪ AS 내에 있는 네트워크 또는 Subnet을 Advertise 하는 방법 중의 하나로 Static Route로 잡아 Redistribute 하는 방법이 있다. Static Route를 Redistribute 하는 경우, BGP는 해당 루트에 대한 Origin Attribute로 Incomplete를 사용한다.

▪ Path Attribute 중 Origin Attribute가 Incomplete 로 나타난다.

60

Redistributing Static Routes into BGP

▪ Static Route를 null 0로 잡아 Redistribution : Supernet을 Advertise 하는 최선의 방법은 Static Route를 null 0로 잡아 Redistribution하는 것이다. Static Route를 사용하게 되면 Subnet Down으로 인한 Supernet의 Route Flapping을 방지한다.
▪ 이론적으로 Null 0 Interface를 지정하는 것은 175.220.0.0번 네트워크를 목적지로 하는 Packet을 버린다는 것을 의미하나 실제로는 175.220.0.0 보다 Specific한 Entry Match가 있으므로 라우터는 해당 패킷을 부합되는 Interface로 보내게 된다.

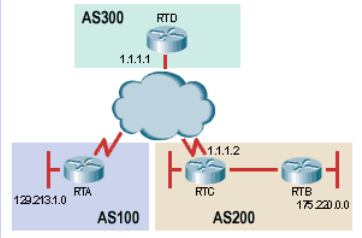



```
RTC#

router bgp 200
 neighbor 1.1.1.1 remote -as 300
 redistribute static

ip route 175.220.0.0 255.255.0.0 null 0
```

61

network 명령의 사용

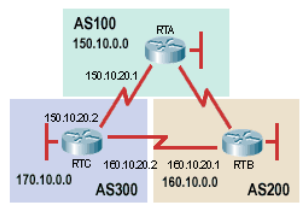
▪ network 명령이 BGP와 함께 사용되면 AS가 Originate 하는 Network 들을 의미한다.(cf.IGP에서의 사용은 IGP가 실행되는 Interface를 나타내는데 사용된다.)
▪ IGP(Dynamic)로 인하여 생성된 Route를 network 명령으로 Advertise 하는 경우 및 Static Route를 설정하고 network 명령으로 Advertise하는 경우로 구분할 수 있다.
▪ network 명령을 사용하여 BGP로 Inject 되는 Route의 BGP Origin Attribute는 IGP로 설정된다.
▪ Cisco Router의 경우 network 명령으로 Advertise 할 수 있는 Maximum network은 200개이다.(IOS 12.0 부터 제한이 없어졌다.) 이를 넘어서는 경우 Redistribution을 사용한다.




```
RTC#

router bgp 200
neighbor 1.1.1.1 remote -as 300
network 175.220.0.0
```

62

network 명령의 사용




```
RTA#
router bgp 100
neighbor 150.10.20.2 remote-as 300
network 150.10.0.0

RTB#
router bgp 200
neighbor 160.10.20.2 remote-as 300
network 160.10.0.0

RTC#
router bgp 300
neighbor 150.10.20.1 remote-as 100
neighbor 160.10.20.1 remote-as 200
network 170.10.0.0
```

63

Route Map

- Route Map은 복합적인 Access List라 할 수 있으며, **Packet 및 Route를 Test**하여 일정한 Action을 줄 수 있다. Redistribution, BGP Implementation (BGP4 Attribute Manipulation), Policy Routing 등에서 사용된다.
- 조건(match)에 대한 Action(set) 형식의 구문으로 되어 있다.
- Route Map Statement는 Number가 부여되며 Statement 단위로 편집될 수 있다.(Route Map에서는 Statement를 추가하거나 제거할 수 있다.)
- Sequence Number를 명시하지 않고 Route Map을 제거하는 경우(ex. no route-map Test 등의 명령을 사용하는 경우) 전체 Route Map이 제거되며, Sequence Number를 명시하지 않고 Statement를 추가하면 10번 Statement로 추가되며 이미 있는 경우 Overwrite 된다.
- Access-list Number ⇒ route-map의 Name에 해당한다.
- Access list의 Line ⇒ route-map에서의 Statement에 해당한다.
- Access List에서의 Addresses와 Mask ⇒ Route Map에서의 Match Statements에 해당한다.
- Access-List와의 가장 큰 차이점은 Route Map은 set 명령을 사용하여 조건(Match)에 대한 Action을 줄 수 있다는 것이다.

64

Route Map의 용도

- Policy Routing
- Redistribution Control
- BGP Update Control
- Route Tagging

65

Route Map Configuration

- Policy Routing을 정의한다.

```
Router(config)#route-map map-tag [permit | deny] [sequence-number]
```

• Permit/Deny의 Default는 Permit이고 Sequence Number의 Default 값은 10이며, Statement 추가시 자동으로 증가하지 않는다.

- Match Condition을 정의한다.

```
Router(config-route-map)# match {conditions}
```

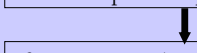
- Match의 경우 취해질 Action을 정의한다.

```
Router(config-route-map)# set {actions}
```

66

Route Map : Match and Set

```
route-map demo permit 10
  match x y z
  match a
  set b
  set c
route-map demo permit 20
  match q
  set r
route-map demo permit 30
```



If {x or y or z) and a match} then {set b and c}
Else
    If q match then set r
    Else
        Set Nothing

67

Route Map : Match and Set

- 첫번째 Statement에 일치하면 나머지 Statement는 체크되지 않는다.(Top Down)
- Deny Action은 Policy Routing의 경우와 Redistribution의 경우가 각각 다르다.
- Route Map이 Ridistribution 등 Routing Update에 사용되는 경우 Route가 Deny Action의 Statement에 Match 되면 해당 Route는 Update 되지 않는다.
- Route Map이 Policy Routing에 사용되는 경우 Packet이 Deny Action의 Statement에 Match 되면 해당 Packet은 Policy Routing 되지 않고 Normal Routing 되어 Forward 된다.
- Route Map 문장에서 Match Statement가 Configuration 되어 있지 않은 경우는 Match Everything(즉 모든 조건)을 의미한다.
- Access -List와 마찬가지로 Route Map의 마지막에는 Implicit Deny All이 적용되므로 이를 방지하기 위해서는 match 및 set이 없는 Statement가 필요하다. match 및 set이 없는 Statement는 모든 경우를 Permit(또는 Deny, route -map이 Permit 설정인지 Deny 설정인지에 따라서) 함을 의미한다.

68

Policy Routing

▪ BGP는 관리자로 하여금 Data가 AS를 경유하는 Data Flow에 대한 Policy, Rule 등을 정의하는 것을 허용한다.

▪ BGP 및 관련 Tool은 Internet상을 경유하는 모든 경로에 대한 Policy를 구현할 수는 없으나 Hop-by-Hop의 Policy 설정은 가능하다. (다른 AS의 Policy를 설정할 수는 없으나 Neighbor AS로 가는 Traffic에 대한 Policy는 정의할 수 있다.)

▪ BGP는 하나의 Destination Route에 대한 Path Attribute에 기준하여 Best Route를 결정한다. 관리자는 Path Attribute Manipulation과 Route Filtering을 통하여 정책기반의 Routing을 구현할 수 있다.

BGP Attributes

▪ BGP metric은 Path Attribute로도 불리운다.

▪ BGP Attributes는 Well-Known Attribute와 Optional Attribute, Mandatory Attribute와 Discretionary Attribute, Transitive Attribute와 Non-Transitive Attribute등으로 구분 된다.

69

Well-known Attributes

▪ Well-Known Attributes : 모든 BGP Router가 인식할 수 있어야 하는 Attribute로 모든 Neighbor에 Propagate 된다.

• Well-Known Mandatory Attributes : 모든 Update Message에 반드시 존재 해야 한다. (AS-Path, Next-Hop, Origin)

• Well-Known Discretionary Attributes : Update Messsage에 있을 수도 있고 없을 수도 있다.(Local Preference, Atomic Aggregate)

Optional Attributes

▪ Optional Attributes : BGP Implementation에 따라 인식 될 수도 있고, 그렇지 않을 수도 있는 Attributed이다.

• Optional Transitive attributes : 만약 인식되지 않으면 Partial로 Mark되고 다른 Neighbor로는 전파되어야 하는 Attribute이다.(Aggregator, Community)

• Optional Non-Transitive Attributes : 만약 인식되지 않으면 무시되고, 다른 Neighbor로도 전파되지 않는 Attribute이다.(MED, Originator_ID, Cluster_List)

70

BGP Attributes

▪ BGP attributes(* 은 Well-known mandatory attribute이다.)

- Origin * (Type Code 1)
- AS-path * (Type Code 2)
- Next-hop * (Type Code 3)
- Multi-Exit-Ddiscriminator (MED) (Type Code 4)
- Local Preference (Type Code 5)
- Atomic Aggregate(Type Code 6)
- Aggregator(Type Code 7)
- Community(Type Code 8)
- Originator-ID (Type Code 9)
- Cluster List (Type Code 10)

71

Origin Attribute

▪ Well-knonw Manadatory Attribute (Type Code 1) 이며 Routing Update 의 Origin을 표시한다.
▪ 어떤 방법을 통해서 BGP Routing Update에 포함되게 되었는 가를 알 수 있다.

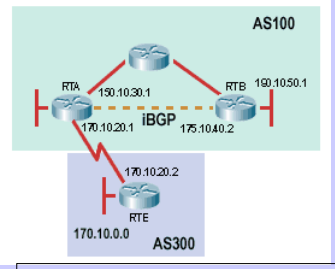
▪ IGP (Attribute Value 0) :어떤 경로가 Originating AS 내부의 IGP Protocol(Static Routing 포함)에 의해 만들어진 Route이다. IGP Routing Table에 있는 경로를 network이라는 Router Configuration 명령으로 BGP로 Inject 시키는 경우에 설정된다.

▪ EGP (Attribute Value 1) :EGP Protocol에 의해서 알게 된 경로가 가지는 Origin Attribute이다. (여기서의 EGP는 BGP이전에 사용되던 프로토콜을 말한다.)

▪ Incomplete (Attribute Value 2) :Redistribution에 의해 Route가 BGP로 유입된 경우의 Origin Attribute이다.

72

Origin Attribute

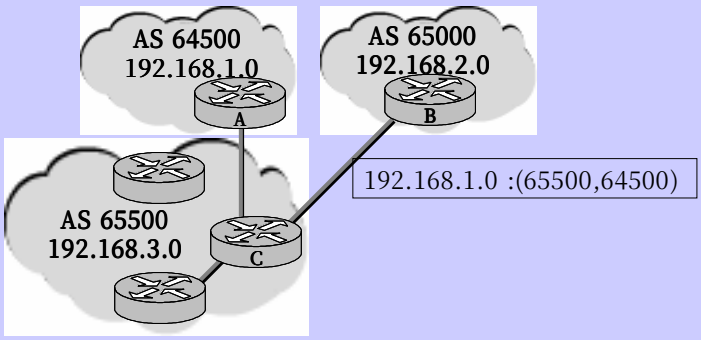


- RTA will reach 170.10.0.0 via: 300 i (Next AS path is 300 and the origin of the route is IGP).
- RTA will reach 190.10.50.0 via: i (Entry is in the same AS and the origin is IGP).
- RTE will reach 150.10.0.0 via: 100 i (Next AS is 100 and the origin is IGP).
- RTE will also reach 190.10.0.0 via: 100 ? (Next AS is 100 and the origin is incomplete "?," coming from a static route).

```
RTA#
router bgp 100
neighbor 190.10.50.1 remote-as 100
neighbor 170.10.20.2 remote-as 300
network 150.10.0.0
redistribute static

ip route 190.10.0.0 255.255.0.0 null0

RTB#

router bgp 100
neighbor 150.10.30.1 remote-as 100
network 190.10.50.0

RTE#

router bgp 300
neighbor 170.10.20.1 remote-as 100
network 170.10.0.0
```

73

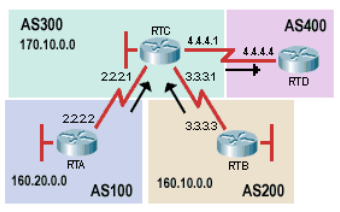
AS-Path Attribute



- Route가 경유하는 AS List를 나타내는 Attribute이다. (Router B에서 192.168.1.0에 대한 Path는 65500 64500 이다.)
- Route가 EBGP Peer간에 Advertise 될 때만 Update 하는 AS의 AS Number가 AS-Path에 추가된다. IBGP Peer간 Update 시에는 추가 되지 않는다.
- **AS-Path는 BGP Router로 하여금 Loop-Free Routing을 보장한다.BGP Router가 자기자신의 AS가 들어간 AS-Path Attribute를 가진 Route를 받으면 BGP Router는 이 Route를 받아 들이지 않는다.**

74

Aggregation and AS -SET

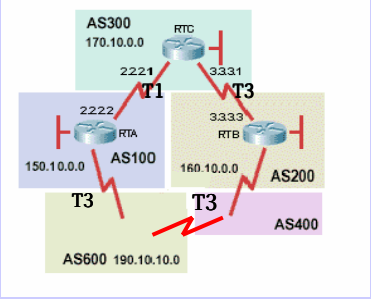



```
RTC#

neighbor 3.3.3.3 remote -as 200
neighbor 2.2.2.2 remote -as 100
neighbor 4.4.4.4 remote -as 400
aggregate -address 160.0.0.0 255.0.0.0 as -set
```

▪ Router C가 Router A와 B로 부터 각각 162.20.0.0 네트워크와 160.10.0.0 네트워크를 Update 받아 160.0.0.0/8 네트워크로 Aggregation하여 Router D에 Update하는 경우, 라우터 D에서의 160.0.0.0 네트워크에 대한 AS_Path는 300으로 AS 300이 160.0.0.0 네트워크의 Origin이 된다.
▪ Router D가 다른 Route를 통하여 AS 100 또는 AS 200에 에 연결되어 있는 경우, 160.0.0.0 네트워크가 Router D로 부터 Router A, B에 네트워크 Origin이 Router C인 것으로 Update 된다면 Router A,B는 이를 받아 들일 수 있어 Routing Loop이 발생할 수 있다.
▪ Routing Loop 문제를 해결하기 위하여 as -set 키워드와 함께 Aggregation을 하게 되면 160.0.0.0 네트워크의 Origin이 AS 100(Router A) 및 AS 200(Router B)으로 설정된다. 이 경우 160.0.0.0이라는 Aggregate의 AS_PATH Information은 300 {100 200}이 된다.

75

AS_Path Manipulation

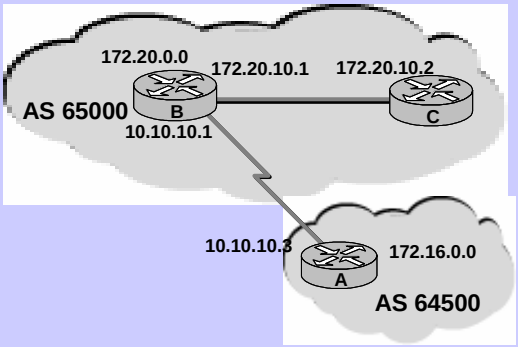



```
RTC#

router bgp 300
network 170.10.0.0
neighbor 2.2.2.2 remote -as 100
neighbor 2.2.2.2 route -map SETPATH out

route -map SETPATH permit 10
set as -path prepend 300 300
```

▪ **prepend** keyword를 사용한 **set as -path** route map 명령은 Router C가 Router A에 Update시에 AS_Path에 추가로 300 300을 추가한다. 170.10.0.0 네트워크의 경우, AS 600에서 Router A를 경유하는 AS_Path가 100.300,300,300으로 되며, AS 400을 경유하는 AS_Path는 400,200,300으로 되어 172.10.0.0 네크워크에 대한 AS 600에서의 Path는 AS400을 경유하는 Route로 설정된다.
▪ Prepending시 주의할 사항으로 Prepend시 사용하는 AS 번호는 반드시 Prepend하는 라우터가 속한 AS 번호를 사용해야 한다. (다른 AS 번호를 사용하게 되면, Prepending된 Route를 다른 AS에서 받아 들일 수 없다.)

76

Next -Hop Attribute



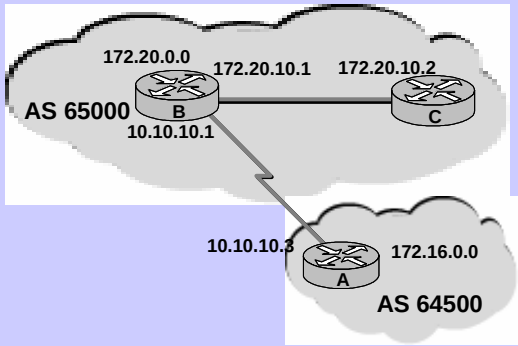
▪ Advertise 되는 Route에 도달하기 위한 Next Hop Router의 IP Address를 나타내는 Attribute이다.

• Advertising Router와 Receiving Router가 다른 AS에 속하는 경우의(EBGP Peer) Next -Hop은 Advertising Router Interface의 IP Address이다.

• Advertising Router와 Receiving Router가 동일 AS에 속하고(IBGP Peer), Update NLRI가 동일 AS에 속하는 Destination인 경우의 Next -Hop은 Advertising Router Interface의 IP Address이다.

• **Advertising Router와 Receiving Router가 동일 AS에 속하고(IBGP Peer), Update NLRI가 다른 AS에 속하는 Destination인 경우의 Next -Hop은 External Peer Router Interface의 IP Address이다.**

Next -Hop Attribute

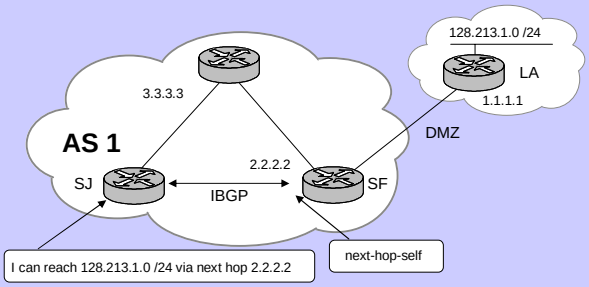


▪ Router A가 network 172.16.0.0에 대해서 EBGP Peer Router인 Router B에게 Next Hop Attribute를 10.10.10.3으로 하여 Advertise 한다.

▪ 이를 받은 라우터 B가 IBGP Peer인 Router C에게 Next Hop Attribute를 10.10.10.3으로 하여 Update 한다.즉 Router B가 Router A(EBGP Neighbor)로 부터 Route를 Update 받아 Router C (IBGP Neighbor)로 Update하는 경우, Next -Hop이 바뀌지 않는다.

78

Use of next-hop-self versus Advertising DMZ

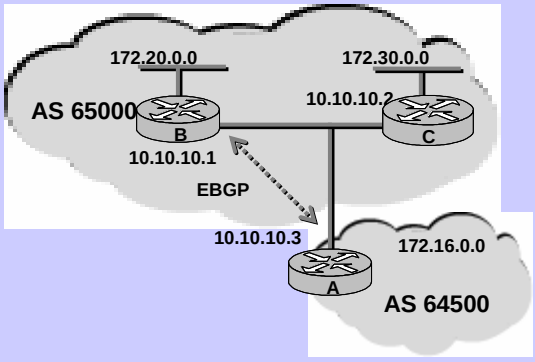


- DMZ 의 정의 : AS간에 공유하고 있는 네트워크를 말한다.(EBGP라우터 간에 직접 연결을 제공하는 네트워크를 말한다.)
- EBGP에서 NEXT_HOP의 전달 :EBGP Peer로 부터 받은 Update의 NEXT_HOP은 IBGP 세션에서는 변경되지 않는다. NEXT_HOP 네트워크를 사용할 수 있어야만 BGP 정보로 사용할 수 있게 된다.

1) DMZ 네트워크를 IGP를 통해서 알린다.
2) next-hop-self를 사용하여 NEXT_HOP을 변경한다.

79

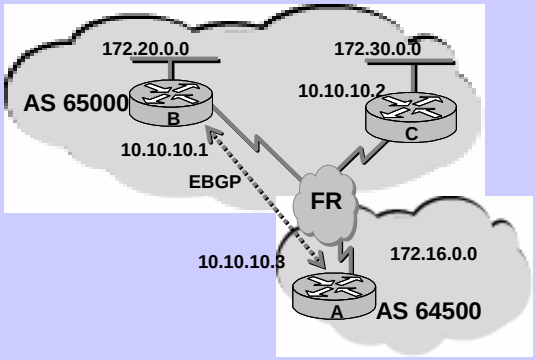
Next-Hop on Multiaccess Network



- Multiaccess Network(LAN 등)에서 Route B는 172.30.0.0 네트워크를 Router A에 Update시 10.10.10.2를 Next Hop으로 하여 Update한다. 이는 172.30.0.0의 Packet이 전달 되는데 있어 불필요한 Hop 수를 줄이는 효과가 있다

80

Next-Hop on an NBMA Network



- NBMA Network에서도 Route B는 172.30.0.0 네트워크를 Router A에 Update시 10.10.10.2를 Next Hop으로 하여 Update한다.
- Router A가 10.10.10.2를 인식할 수 있도록 Frame-Realy Map을 잡던지, Router B가 Next Hop을 자기자신으로 Advertise하도록(neighbor 10.10.10.3 next-hop-self) 설정할 수 있다.

81

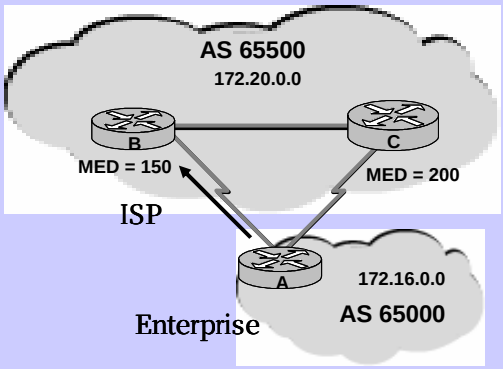
next-hop-self

```
Router(config-router)# neighbor {ip-address | peer-group-name}
next-hop-self
```

- Neighbor Router에 Update 하는데 있어서 이 라우터를 Next Hop으로 설정하도록 강제한다.

82

MED(Multi -Exit Discriminator) Attribute

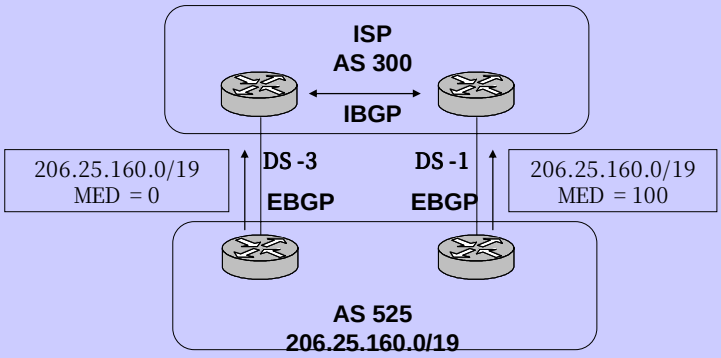


▪ AS 65000에 있는 라우터 A가 AS 65500에 있는 라우터 B와 C에 Multiple Exit으로 연결된 경우, Router B와 Router C가 AS 65000에 있는 라우터 A 에 대해 이 AS(65500)로 오는 데 있어서의 MED Attribute 값을 다르게 주어서 적은 값을 갖는 라우터(B)를 최적의 경로로 선택하게 한다.

▪ Route의 MED Attribute는 다른 AS로는 전달되지 않는다.(Router B와 C가 172.16.0.0 Route에 대한 MED를 비교하며, Router B와 C가 다른 AS로 Route를 Advertise할 때 MED Attribute를 전달하지 않는다.)

83

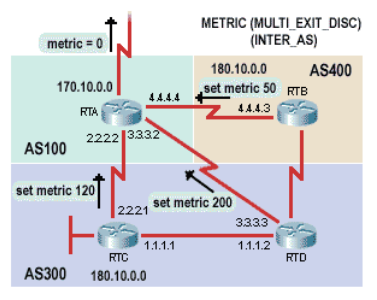
MED(Multi -Exit Discriminator) Attribute



▪ ISP(AS300)에서 206.25.160.0/19 네트워크로의 모든 Traffic은 DS -3 Link를 경유한다.

84

MED(Multi -Exit Discriminator) Attribute




```
RTA#

router bgp 100
neighbor 2.2.2.1 remote-as 300
neighbor 3.3.3.3 remote-as 300
neighbor 4.4.4.3 remote-as 400
bgp always-compare-med
```

▪ Default로 BGP 라우터 A는 어떤 경로에 대한 MED 비교를 동일한 AS 300에 속하는 Router C와 D의 경로만을 비교하므로 Network 180.10.0.0 네트워크에 대한 최적의 경로로 Router C를 택하게 된다.

▪ **bgp always -compare -med 명령을 사용하면 모든 AS의 Router들이 특정 경로에 대한 Metric 값을 비교하는데 있어서의 대상이 되므로 180.10.0.0 네트워크에 대한 Router A에서의 최적의 경로는 MED 값을 50으로 Update한 Router B로 선택된다.**

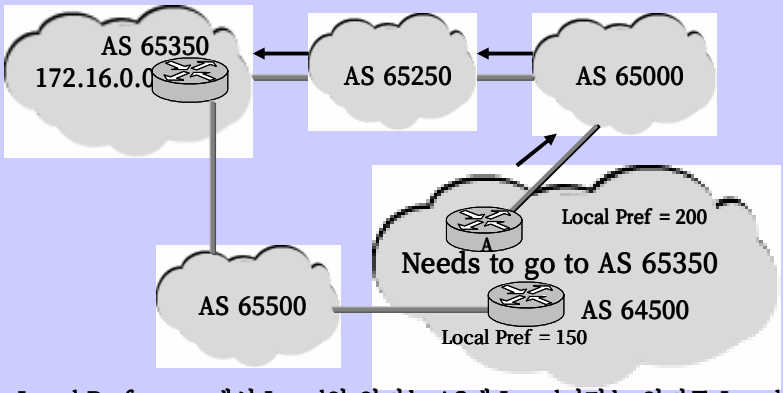
85

MED(Multi -Exit Discriminator) Attribute 설정

```
RTC#
router bgp 300
neighbor 2.2.2.2 remote -as 100
neighbor 2.2.2.2 route -map setmetricout out
neighbor 1.1.1.2 remote -as 300
route -map setmetricout permit 10
set metric 120

RTD#
router bgp 300
neighbor 3.3.3.2 remote -as 100
neighbor 3.3.3.2 route -map setmetricout out
neighbor 1.1.1.1 remote -as 300
route -map setmetricout permit 10
set metric 200

RTB#
router bgp 400
neighbor 4.4.4.4 remote -as 100
neighbor 4.4.4.4 route -map setmetricout out
Route -map setmetricout permit 10
set metric 50
```

86

Local Preference Attribute



▪ Local Preference에서 Local의 의미는 AS에 Local이라는 의미로 Local Preferance Attribute는 Local AS내의 IBGP Neighbor간에만 전달된다.EBGP Neighbor로는 전달되지 않는다.

▪ 동일한 Destination에 대해서 EBGP로 부터 받은 경로를 IBGP Router간에 Update하게 되는 경우, Local Preference Attribute를 비교하여 높은 값을 갖는 경로를 선택하게 된다. Cisco Router에서 IBGP Peer간에 전달되는 Route의 Default Local Preference 값은 100이다.

87

Configuring Weight and Local Preference

Router(config-router)#bgp default local-preference *value*

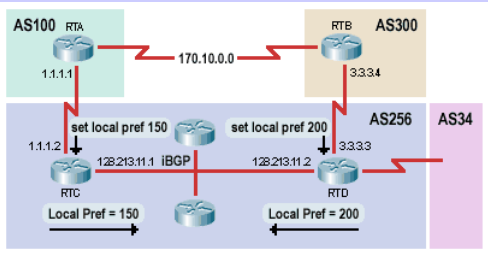
▪ Default Local Preference Value를 바꾼다. 해당 라우터가 모든 Route Update에 적용하는 local-preference 값을 설정한다.

▪ Local Preference는 Weight Attribute가 동일할 때 Tie Breaker로 사용된다.

▪ Default Local Preference는 100이다.

88

Local Preference Attribute

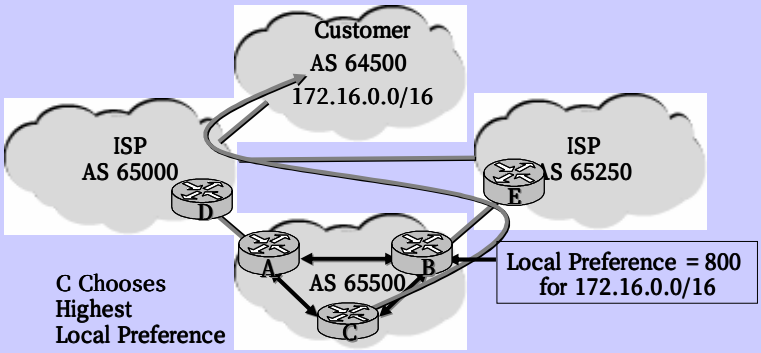



```
RTC#
router bgp 256
neighbor 1.1.1.1 remote -as 100
neighbor 128.213.11.2 remote -as 256
bgp default local -preference 150

RTC#
router bgp 256
neighbor 3.3.3.4 remote -as 300
neighbor 128.213.11.1 remote -as 256
bgp default local -preference 200
```

- Local Preference Attribute는 각각의 Router에서 bgp default local -preference 명령으로 설정되며(Default는 100이다), AS 안의 IBGP Router 간에만 교환된다.
- Routr C는 AS 100으로 부터 172.10.0.0 네트워크를 Update 받으면 그 경로에 대한 Local Preference 값은 150으로 되고, Router D는 AS 300으로 부터 172.10.0.0 네트워크를 Update 받으면 그 경로에 대한 Local Preference 값은 200으로 되어, Router C와 Router D가 172.10.0.0 네트워크에 대한 Route를 교환한다. 결국 172.10.0.0으로 가는 Packet은 모두 라우터 D를 경유하여 AS 300을 통해서 나가게 된다.

89

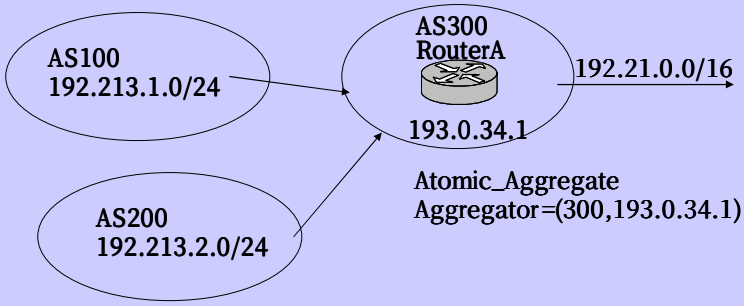
Route별 Local Preference 설정




```
Router B
router bgp  65500
 neighbor 10.1.1.1 remote -as 65250
 neighbor 10.1.1.1  route -map toright in
route -map toright permit 10
 match ip address prefix -list customer
 set local preference 800
ip prefix -list customer permit 172.16.0.0/16
```

90

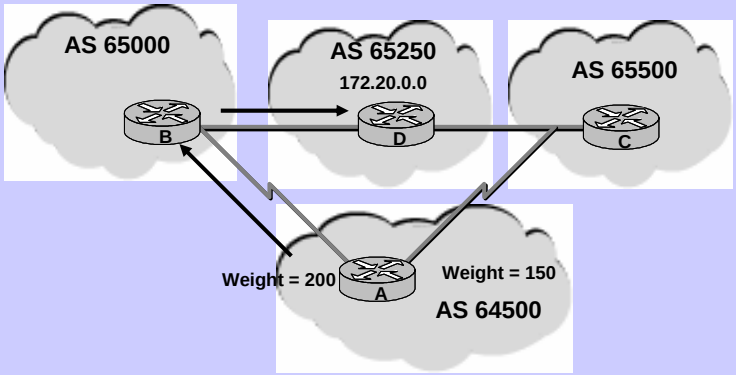
Atomic Aggregate Attribute, Aggregator Attribute



▪ AS300이 AS100으로 부터 192.213.1.0/24에 대한 경로를, AS200으로부터는 192.213.2.0/24에 대한 경로를 Update를 받아 192.21.0.0/16의 경로로 Aggregation하여 다른 AS로 Update하는 경우, 192.21.0.0/16의 Path Attribute는 Atomic Aggregate Attribute와 Aggregator Attribute가 붙게 된다.
▪ Aggregator Attribute는 Aggregation 한 Router(여기서는 RouterA)가 속한 AS 번호와 Aggregation 한 Router의 Router ID로 구성된다.

91

Weight Attribute (Cisco Only)



▪ Weight Attribute는 동일한 목적지에 나가는 경로가 두개 이상인 경우에 사용된다. Cisco Only이며 Weight Attribute는 **Local Router에서만 의미가 있으며**, Routing Update시에 전파되지 않는다.동일한 Destination에 다수의 Route가 있는 경우, 높은 Weight 값을 가진 경로가 선호 된다. Weight Attribute의 Default 값은 0이다.(Peer로 부터 알게된 모든 경로는 Default로 0의 Weight Value를 갖으며, 라우터 자체가 만들어낸 Route는 Default로 32768의 Weight Value를 갖는다.)

92

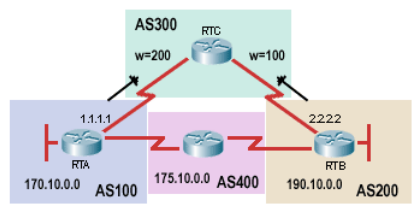
Configuring Weight

```
Router(config-router)#neighbor {ip-address | peer-group-name}
weight weight
```

- Neighbor와의 Connection에 Weight을 할당한다.
- 0-65535의 Range가 가능하며 Defualt는 Local Route에 대해서는 32768 이며 Other Route에는 0 값을 갖는다.

93

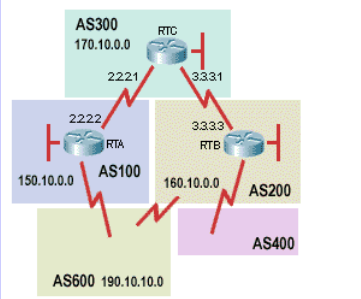
Weight Attribute (Cisco Only)




```
RTC#
router bgp 300
 neighbor 1.1.1.1 remote-as 100
 neighbor 1.1.1.1 weight 200    !!!route to 175.10.0.0 from RTA has
200 weight
 neighbor 2.2.2.2 remote-as 200
 neighbor 2.2.2.2 weight 100    !!!route to 175.10.0.0 from RTB will
have 100 weight
```

94

Route별 Weight 설정




```
RTC#
router bgp 300
network 170.10.0.0
neighbor 3.3.3.3 remote -as 200
neighbor 3.3.3.3 route -map STAMP in

route -map STAMP permit 10
match as -path 1
set weight 20

route -map STAMP permit 20
match as -path 2

route -map STAMP permit 30
set weight 10

ip as -path access -list 1 permit ^200$
ip as -path access -list 2 deny _400_
```

▪ Router B로 부터 Router C로의 Inbound Update에 있어 AS200이 AS_Path로 된 Route는 Weight Value가 20으로 설정되어 Update 되며 AS400을 AS_Path로 포함하는 Route는 Update가 금지 되고, 다른 모든 Route는 Weight Value가 10으로 설정되어 Update 된다.

95

Route Selection Decision Process

▪ 다른 AS로부터의 Update를 받은 BGP는 하나의 Destination에 대해서 오직 하나의 Path를 선택한다.

1. **Route에 AS Loop이 없고, Synchronized된 Route이며, 유효한 Next Hop을 가진 Route에 대해서 다음의 우선 순위로 Route를 Selection 한다.**
2. **Highest Weight (Router내에서만 유효하다.)**
3. **Highest Local Preference (AS내에서만 유효하다.)**
4. **Shortest AS -Path**
5. **Lowest Origin Code (IGP < EGP < Incomplete)**
6. **Lowest MED**
7. **EBGP Path가 IBGP Path보다 선호된다.**
8. **보다 가까운 IGP Neighbor를 거쳐서 도달할 수 있는 Route가 선택된다.**
9. **BGP Router ID가 낮은 Route가 선호된다.**
10. **BGP Router ID가 같은 경우는 Neighbor 설정의 IP Address가 낮은 경로가 선택된다.**

▪ BGP Multipath가 Enable 되어 있는 경우(maximum -paths가 2 이상으로 설정된 경우) 위 8번 까지만을 비교하여 Attribute Value가 동일한 모든 경로를 Routing Table에 Install 한다.)

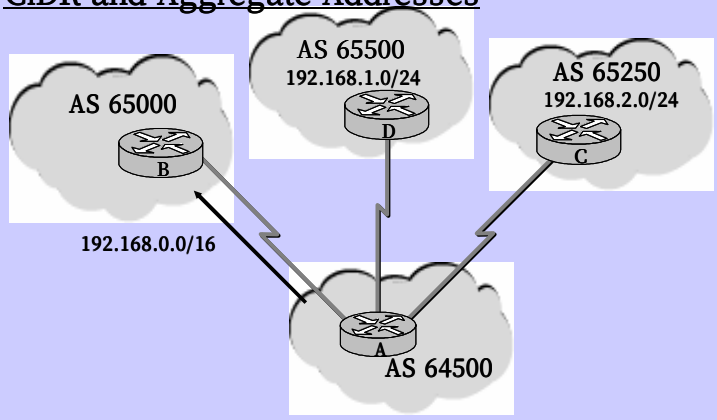
96

Route Selection Decision Process Summary

- BGP는 Destination에 대한 하나의 경로만을 선택한다.여러 개의 Attribute가 Tie Breaker가 되어 하나의 Destination에 대해서는 하나의 경로만을 선택한다.
- 대체로 경로에 대한 Weight, Local Preference등이 동일하기 때문에 경로 결정에 가장 큰 영향을 미치는 요소는 AS_path Attribute가 된다.

97

CIDR and Aggregate Addresses



- Route는 AS를 통과시에 Aggregate 될 수 있다.
- Router C와 D로 부터 각각 192.168.2.0/24, 192.168.1.0/24 네트워크를 Update 받은 Router A는 192.168.0.0/18 네트워크로 Aggregation 하여 Router B로 Update 한다.
- Well-known Discretionary Attribute인 Atomic Aggregate는 Router A가 Route를 Aggregation한 사실을 전달한다.
- Well-known Transitive Attribute인 aggregator는 Route Aggregation을 수행한 Router의 ID와 AS 번호를 전달한다.

98

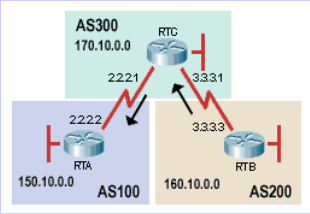
Configuring BGP for Aggregate Addressing

Router(config-router)# aggregate-address *ip-address mask* [summary-only] [as-set]

- BGP Table에 Aggregate (Summary) Entry를 만든다.
- **BGP Table에 있는 Route만 Aggregate 될 수 있다.(Aggregation한 Route의 Specific Route가 최소한 1개 이상 있어야 Aggregation한 Route가 Advertise 된다.)**
- Specific Route가 아닌 Summary Route만을 Advertise 시키려면 summary-only Option 키워드를 사용한다.
- Aggregate 되는 Router가 Route Origin을 포함한 모든 AS_Path를 표시하도록 하려면 as-set 키워드를 사용한다.

99

Aggregation




```
RTC#

router bgp 300
neighbor 3.3.3.3 remote-as 200
neighbor 2.2.2.2 remote-as 100
network 170.10.0.0
aggregate-address 160.0.0.0 255.0.0.0
```

100

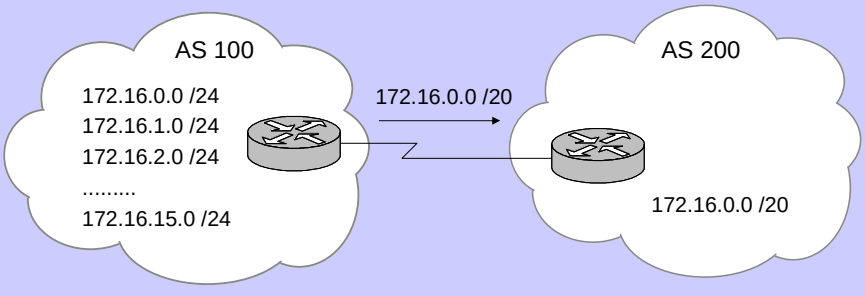
Aggregation

- **BGP Routing Table에 Specific Route를 갖고 있지 않다면 Router는 Network을 Aggregation 할 수 없다.**
- Specific Route는 다른 AS와의 BGP Update, IGP로부터의 Redistribution에 의해서, network 명령에 의해서 BGP Routing Table에 Inject 될 수 있다.
- Router C가 Specific Route는 제외하고 Aggregated Route만을 Update하기를 원하는 경우, aggregate-address 160.0.0.0 255.0.0.0 **summary-only** 명령을 사용한다. 즉 summary-only 키워드를 추가한다.

101

Aggregation Only, Suppressing the More-Specific Routes

router(config-router)# **aggregate** *aggregate-address aggregate-mask* **summary-only**

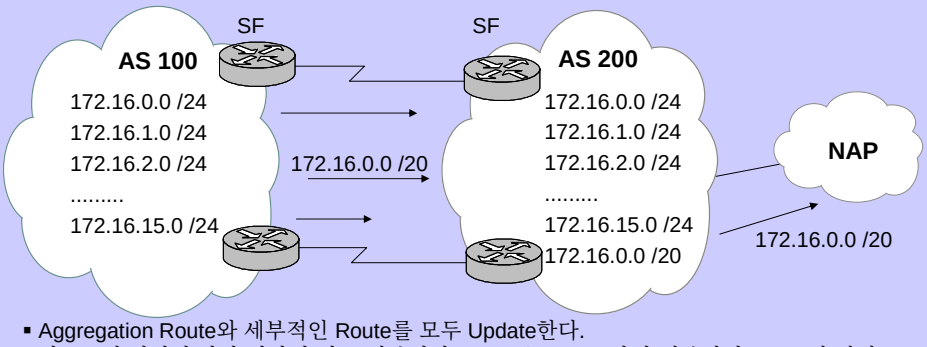


- Aggregation route만을 Update하고 세부적인 Route들은 Update되지 않는다.
- 세부적인 Route를 Update하는 것이 특별한 이점이 없는 경우 사용된다.

102

Aggregate Plus More -Specific Routes

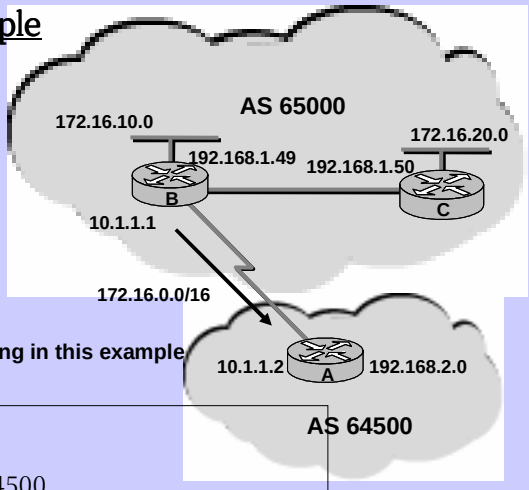
router(config -router)# **aggregate** *aggregate -address aggregate -mask*



- Aggregation Route와 세부적인 Route를 모두 Update한다.
- 한 ISP에 연결이 하나 이상인 경우 세부적인 Route를 Update하여 세부적인 Route에 대해 Best Path를 선택할 수 있도록 한다.
- ISP에서는 Aggregation된 Route만을 NAP로 Update한다.
- 세부적인 Route에 NO_EXPORT Community Attribute를 추가하여 자동적으로 ISP에서 NAP로 Update하지 않게 할 수 있다.

103

BGP Aggregation Example



Note: There is no IGP running in this example

```
Router B
router bgp 65000
neighbor 10.1.1.2 remote -as 64500
neighbor 192.168.1.50 remote -as 65000
network 172.16.10.0 mask 255.255.255.0
network 192.168.1.0 mask 255.255.255.0
no synchronization
neighbor 192.168.1.50 next -hop -self
aggregate -address 172.16.0.0 255.255.0.0 summary -only
```

104

Multihoming

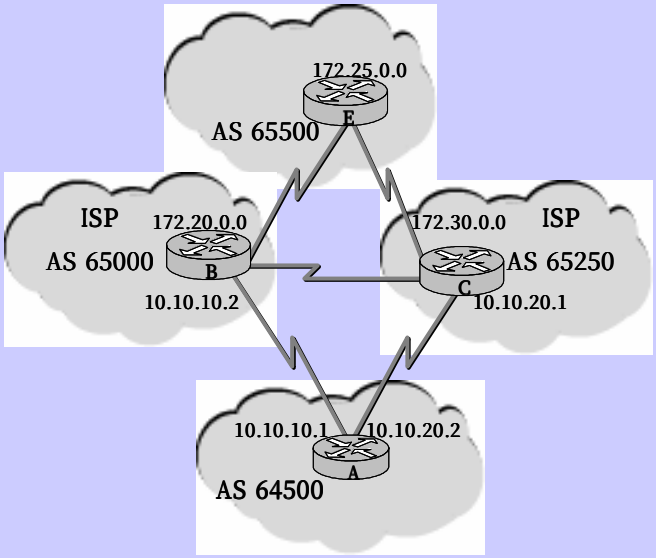
▪ 두개 이상의 ISP에 연결하는 것은 Reliability 및 Performance 향상을 제공한다.

• Reliability : 하나의 ISP와의 Connection이 Fail되는 경우에도 다른 ISP와의 연결을 사용할 수 있다.

• Performance : 두개의 ISP를 통해 경로를 설정할 수 있으므로 Destination Network에 대해서 최적의 Path를 선택하게 할 수 있다.

105

Multihoming Example



106

Configuration(No Tuning)

```
RtrA(config)#router bgp 64500
RtrA(config-router)#network 10.10.10.0 mask 255.255.255.0
RtrA(config-router)#network 10.10.20.0 mask 255.255.255.0
RtrA(config-router)#neighbor 10.10.10.2 remote-as 65000
RtrA(config-router)#neighbor 10.10.20.1 remote-as 65250
```

```
RtrA#show ip bgp
BGP table version is 7, local router ID is 172.16.10.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop          Metric LocPrf Weight Path
*> 10.10.10.0/24    0.0.0.0                0         32768 i
*> 10.10.20.0/24    0.0.0.0                0         32768 i
*  172.20.0.0       10.10.20.1                           0 65250 65000 i
*>                  10.10.10.2             0             0 65000 i
*> 172.25.0.0       10.10.10.2                           0 65000 65500 i
*                   10.10.20.1                           0 65250 65500 i
*  172.30.0.0       10.10.10.2                           0 65000 65250 i
*>                  10.10.20.1             0             0 65250 i
```

▪ 175.25.0.0 Route는 동일한 AS_Path를 갖는 두개의 경로중 BGP Router ID가 낮은 쪽(10.10.10.2의 Router ID는 172.20.0.0이다.)을 택한 것으로 볼 수 있다.

Configuration(Change Weights)

```
RtrA(config)#router bgp 64500
RtrA(config-router)#network 10.10.10.0 mask 255.255.255.0
RtrA(config-router)#network 10.10.20.0 mask 255.255.255.0
RtrA(config-router)#neighbor 10.10.10.2 remote-as 65000
RtrA(config-router)#neighbor 10.10.10.2 weight 100
RtrA(config-router)#neighbor 10.10.20.1 remote-as 65250
RtrA(config-router)#neighbor 10.10.20.1 weight 150
```

```
RtrA#sh ip bgp
BGP table version is 9, local router ID is 172.16.10.1
Status codes: s suppressed, d damped, h history, * valid, > best, i -
internal
Origin codes: i - IGP, e - EGP, ? - incomplete
   Network          Next Hop          Metric LocPrf Weight Path
*> 10.10.10.0/24    0.0.0.0                0         32768 i
*> 10.10.20.0/24    0.0.0.0                0         32768 i
*> 172.20.0.0       10.10.20.1                         150 65250 65000 i
*                   10.10.10.2             0           100 65000 i
*> 172.25.0.0       10.10.20.1                         150 65250 65500 i
*                   10.10.10.2                         100 65000 65500 i
*> 172.30.0.0       10.10.20.1             0           150 65250 i
*                   10.10.10.2                         100 65000 65250 i
```