Border Gateway Protocol (BGP)

6장

6.1 인터넷 라우팅의 이층 구조

- Border gateway: 한 기관의 네트워크에서 인터넷으로 나가는 關門 라우터

- Autonomous system (AS): 한 기관의 네트워크
  - ➢ 자신의 prefix를 전 세계 인터넷에 광고(advertise)하는 네트워
    - ✓ 예) 고려대학교: 163.152.0.0/16
  - ➢ BGP를 통하여 광고
  - ➢ AS number (ASN) 부여 → BGP가 사용



6.1 인터넷 라우팅의 이층 구조

- Tracert 결과가 보여주는 이층 구조
  - ➢ 총 4개 네트워크
    - ✓ KT, HE, FLRNET, UFL
  - ➢ Hurricane Electric (HE) 망 내부
    - ✓ 7개 라우터 경유
  - ➢ KT 망 내부
    - ✓ 7개 라우터 경유

```
C:\Users\사용자>tracert -w 100 www.ufl.edu

최대 30홉 이상의 virtual-l1wvi-prod-core-public.osg.ufl.edu [128.227.9.98](으)로 가는 경로 추적:

1  3 ms  2 ms  1 ms 172.30.1.254
2  8 ms  9 ms  * 125.143.149.126
3  6 ms  8 ms  7 ms 61.78.43.71
4  7 ms  6 ms  6 ms 112.189.15.53
5  * * * 요청 시간이 만료되었습니다.
6  6 ms  6 ms  7 ms 112.174.82.126
7  14 ms  6 ms  8 ms 112.174.84.34
8  143 ms  142 ms  * 112.174.87.250
9  145 ms  151 ms  41 ms  10ge12-17.core1.lax1.he.net [66.220.2.169]
10  * 164 ms  148 ms  100ge14-1.core1.lax2.he.net [72.52.92.122]
11  * 155 ms  152 ms  100ge0-36.core2.phx2.he.net [184.105.81.178]
12  * 186 ms  189 ms  100ge10-1.core1.dal1.he.net [184.105.81.174]
13  186 ms  * 182 ms  100ge8-1.core1.atl1.he.net [184.105.81.169]
14  194 ms  193 ms  * 100ge0-36.core2.jax1.he.net [184.105.213.186]
15  213 ms  208 ms  * florida-lambdarail.10gigabitethernet3-2.core1.jax1.he.net [64.71.158.82]
16  221 ms  * 212 ms con-ufl-gnv-internet-v1804.net.flrnet.org [108.59.29.243]
17  199 ms  202 ms  * ssrb230a-pel-asr9001-1-v16-1.ns.ufl.edu [128.227.236.204]
18  211 ms  * 214 ms ssrb230a-nexus-msfc-1-v15-1.ns.ufl.edu [128.227.236.202]
19  206 ms  * 206 ms ssrb230a-dcpopl3-7009-1-vdc2-v601-1.ns.ufl.edu [128.227.236.155]
20  209 ms  * 209 ms virtual-l1wvi-prod-core-public.osg.ufl.edu [128.227.9.98]

추적을 완료했습니다.
```

```
IPv4주소      : 125.128.0.0 - 125.159.255.255 (/11)
기관명        : 주식회사 케이티

IPv4주소      : 61.78.0.0 - 61.81.255.255 (/14)
기관명        : 주식회사 케이티

IPv4주소      : 112.160.0.0 - 112.191.255.255 (/11)
기관명        : 주식회사 케이티
```

6.1 인터넷 라우팅의 이층 구조

- 망 단위의 라우팅은 BGP + 망 내부의 라우팅은 OSPF/IS-IS
  - BGP는 정책(policy) 기반 라우팅
    - ✓ 망 사이의 연결은 철저하게 거래와 계약에 의함
  - OSPF/IS-IS는 효율 (즉 최단거리) 기반 라우팅



- 이층 구조의 scalability
  - 거대한 인터넷의 크기에도 불구하고 라우팅이 가능케 해 줌
  - 망 내부의 라우팅 변화는 망 단위의 라우팅에 거의 영향을 주지 않음

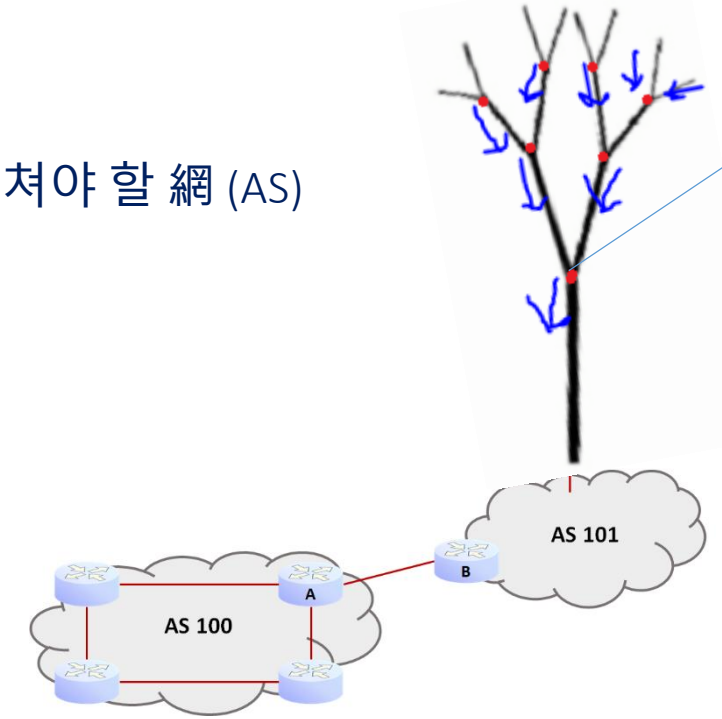
6.2 망 수준의 인터넷 연결 구조

- BGP의 역할: 네트워크의 prefix를 全世界의 core 라우터에 심는 것



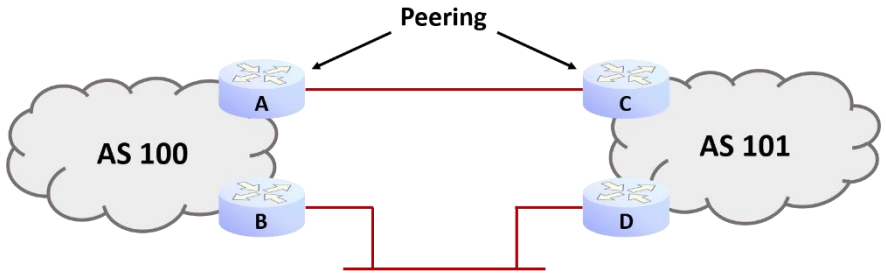
- 방법: 全世界의 core 라우터들로 prefix 정보를 흘려 보냄
  - ➢ 전파의 반대 방향으로 “화살표”를 남김
    - ✓ 직전에 거친 망을 향한 “화살표” : 다음에 거쳐야 할 網 (AS)

- AS의 종류
  - ➢ Stub
  - ➢ Transit

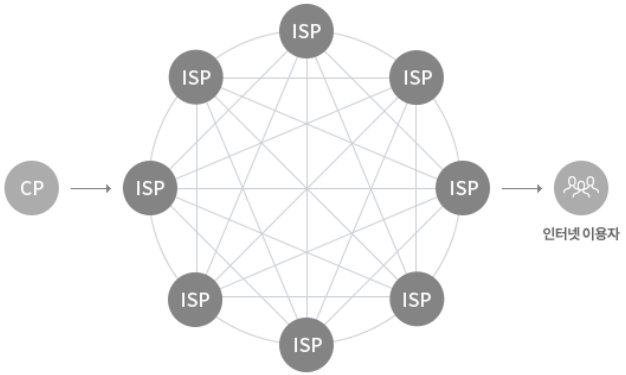


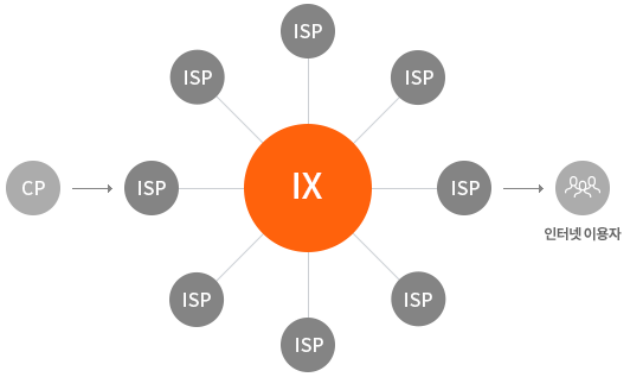
6.2 망 수준의 인터넷 연결 구조

- AS 사이의 관계
  - ➢ 상하 관계: 작은 AS는 큰 AS의 customer
  - ➢ 동등한 연결 (peer)



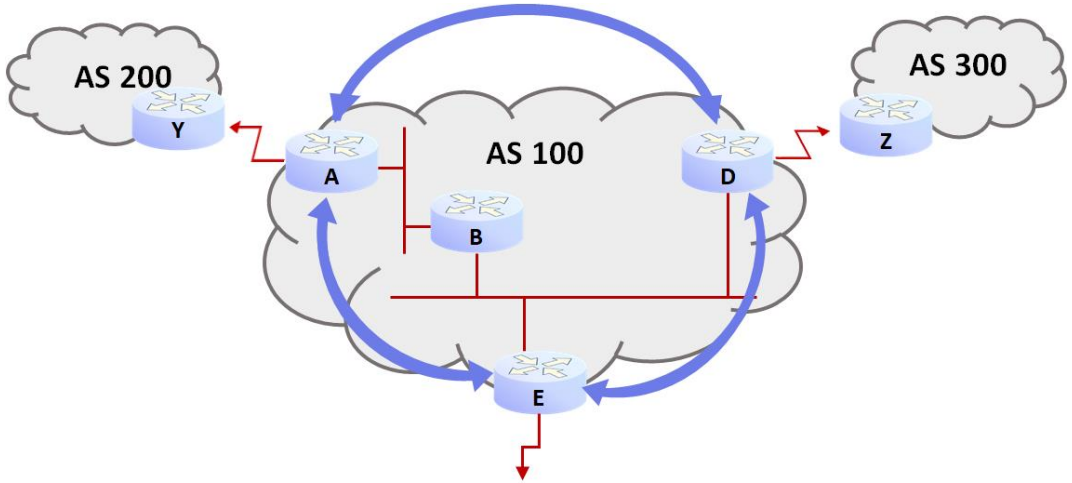
- Peering
  - ➢ Private
  - ➢ Public: Internet Exchange (IX) 사용
    - ✓ 예) KINX





6.2 망 수준의 인터넷 연결 구조

- BGP peering
  - ➢網 단위의 라우팅을 결정하기 위해 인접한 AS 사이에서 BGP로 대화하는 라우터 쌍



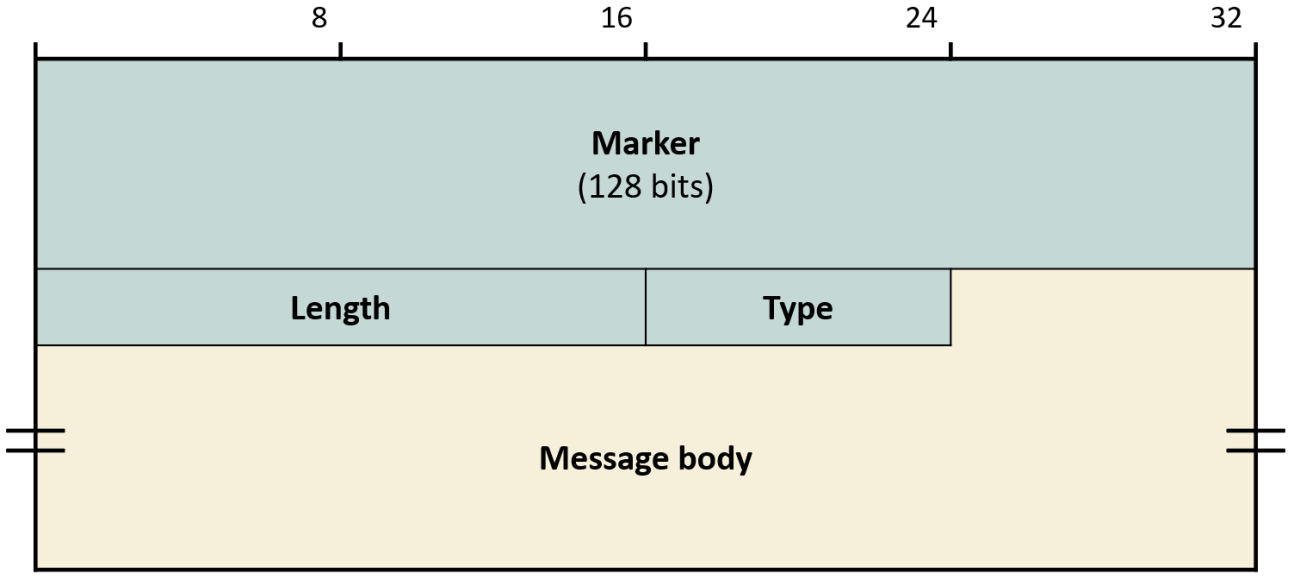
6.3 BGP 프로토콜의 특성

- 정책 (policy) 기반 프로토콜
  - BGP를 통해 전달되어 오는 경로 공고 (announcement)를 받을지 버릴지 정책이 결정
    - ✓ 최단 경로가 무조건 선택되는 것은 아님

- 경로 벡터 (path vector) 프로토콜
  - Announcement 가 거쳐 온 AS의 path를 검사하여 loop 제거
  - AS path의 여러 attribute를 검사하여 정책에 맞는 경로 받아 들임
    - ✓ 예) AS path가 긴 것 보다는 짧은 것 등

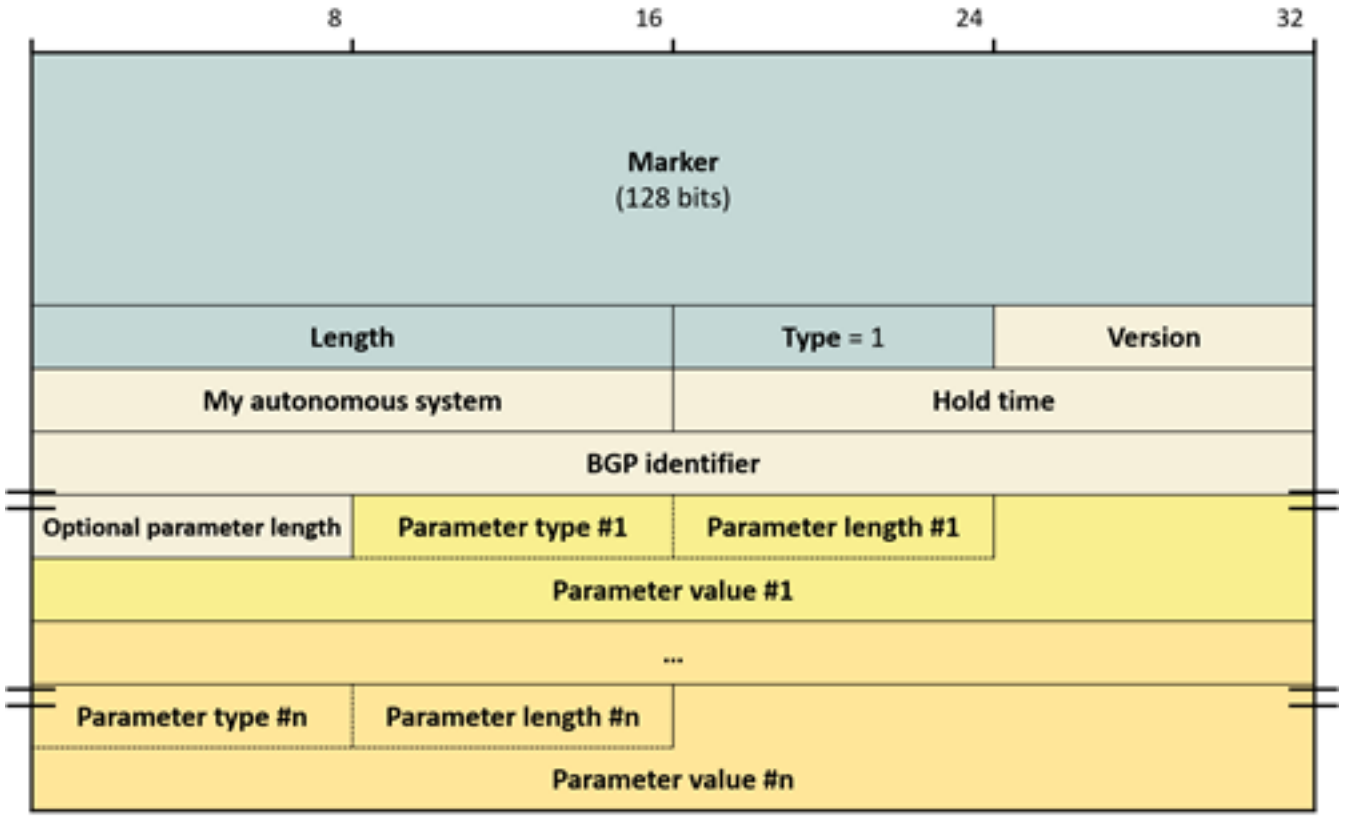
6.4 BGP 메시지 형식

- BGP 메시지 일반 형식
  - ➢ 마커: 인증 용도 (초기값 111111.....1) 또는 BGP 메시지 사이의 경계 표시
    - ✓ 상호 인증 후 고유 패턴으로 변경
    - ✓ BGP 세션 해킹 방지
  - ➢ 길이 = 헤어 + 바디
  - ➢ 타입
    - ✓ 1=연결 (Open)
    - ✓ 2=갱신 (Update)
    - ✓ 3=통고 (Notification)
    - ✓ 4=Keepalive (헤더만)



6.4 BGP 메시지 형식

- Open 메시지 용도
  - BGP 세션을 peer끼리 생성
  - 세션 파라미터 협상
- Open에 대한 응답
  - 긍정= Keepalive
  - 부정= Notification



1	0.000000	4.4.4.4	3.3.3.3	60	11965 → 179 [SYN] Seq=0 Win=16384 Len=0 MSS=536
2	2.025879	3.3.3.3	4.4.4.4	60	179 → 11965 [SYN, ACK] Seq=0 Ack=1 Win=16384 Len=0 MSS=536
3	2.033913	4.4.4.4	3.3.3.3	60	11965 → 179 [ACK] Seq=1 Ack=1 Win=16384 Len=0
4	2.041945	4.4.4.4	3.3.3.3	99	OPEN Message
5	2.049883	3.3.3.3	4.4.4.4	118	OPEN Message, KEEPALIVE Message
6	2.057897	4.4.4.4	3.3.3.3	73	KEEPALIVE Message

6.4 BGP 메시지 형식

- 버전=4
- AS 번호: Open 메시지를 보낸 BGP 라우터가 속한 AS의 번호
- Hold time: default=180 초
  - BGP peer로부터 update를 받지 못한 경우 세션을 포기할 시간
  - Update가 없으나 세션을 유지하고 싶은 경우 1/3 시점에 keepalive 전송
    - ✓ 0= 타임아웃 없음
- BGP 식별자: 발신자인 BGP 라우터의 IP 주소
- 파라미터
  - 예) 발신자가 사용하기 원하는 인증 방법: 마커 필드 계산 방식

6.4 BGP 메시지 형식

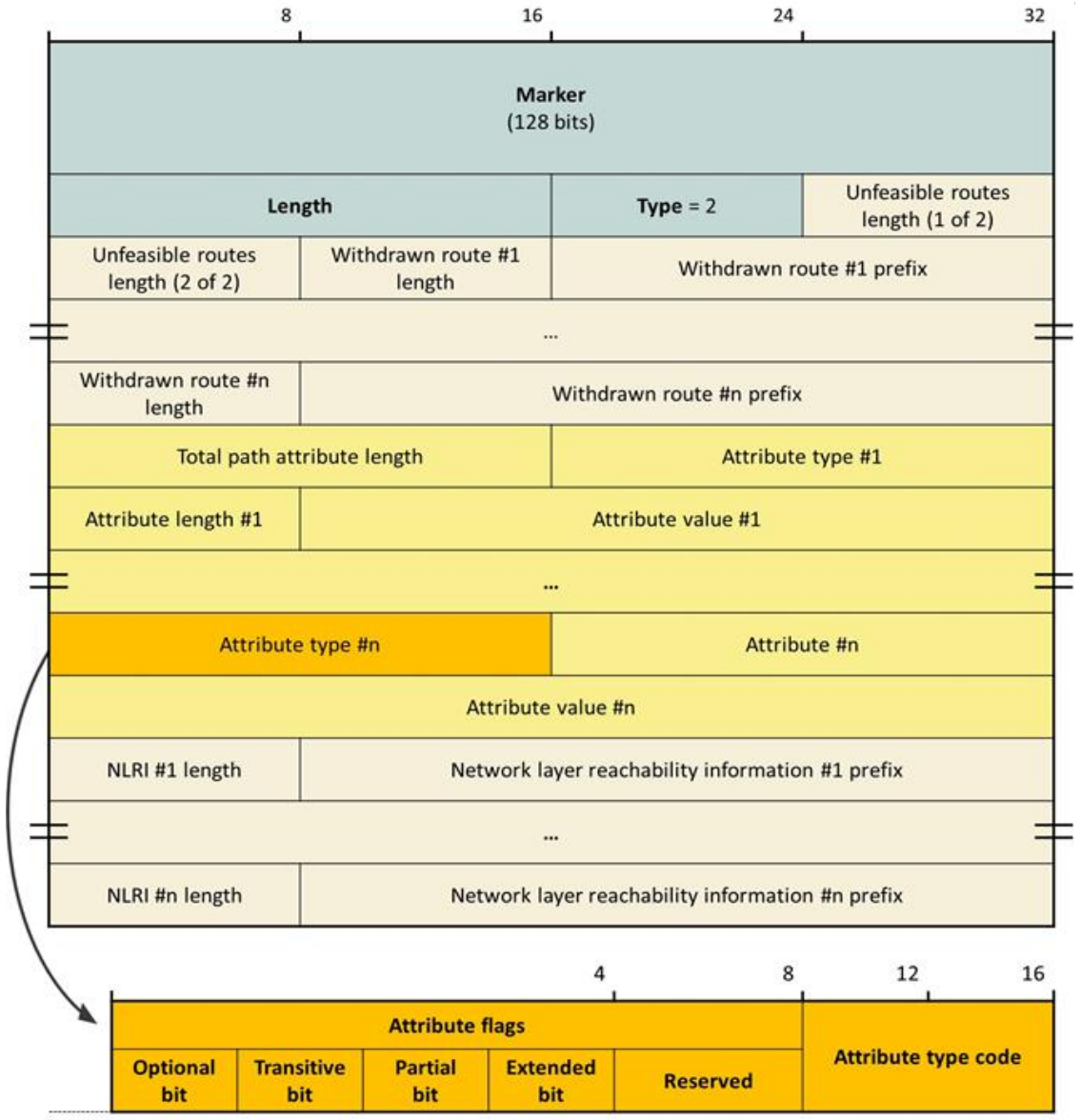
- Open 메시지 예

```
Border Gateway Protocol - OPEN Message
  Marker: ffffffffffffffffffffffffffffffff
  Length: 45
  Type: OPEN Message (1)
  Version: 4
  My AS: 65300
  Hold Time: 180
  BGP Identifier: 4.4.4.4
  Optional Parameters Length: 16
  Optional Parameters
    Optional Parameter: Capability
      Parameter Type: Capability (2)
      Parameter Length: 6
    > Capability: Multiprotocol extensions capability
```

```
Border Gateway Protocol - OPEN Message
  Marker: ffffffffffffffffffffffffffffffff
  Length: 45
  Type: OPEN Message (1)
  Version: 4
  My AS: 65300
  Hold Time: 180
  BGP Identifier: 3.3.3.3
  Optional Parameters Length: 16
  Optional Parameters
    Optional Parameter: Capability
      Parameter Type: Capability (2)
      Parameter Length: 6
    > Capability: Multiprotocol extensions capability
    > Optional Parameter: Capability
    > Optional Parameter: Capability
Border Gateway Protocol - KEEPALIVE Message
  Marker: ffffffffffffffffffffffffffffffff
  Length: 19
  Type: KEEPALIVE Message (4)
```

6.4 BGP 메시지 형식

- Update 메시지의 용도
  - BGP 피어에게 자신이 선택한 경로와 그 애트리뷰트를 알려줌
- 종류
  - 경로 애트리뷰트 (announcement)
  - 도달 불가능 프리픽스 (withdrawal)



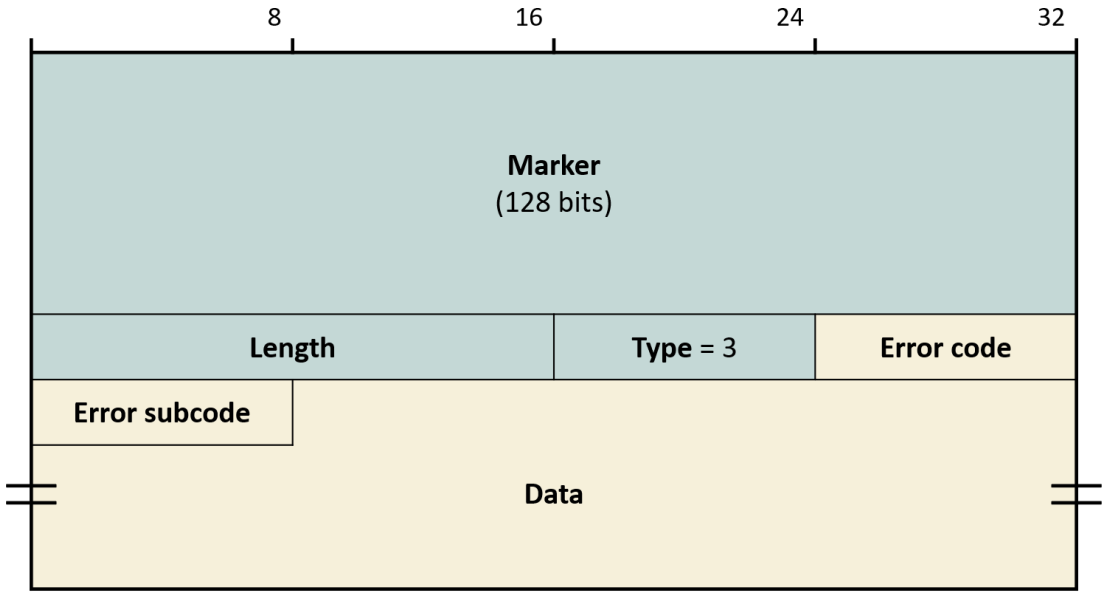
6.4 BGP 메시지 형식

- NLRI의 애트리뷰트 중요
  - ➢ Network Layer Reachability Information

```
Border Gateway Protocol - UPDATE Message
    Marker: ffffffffffffffffffffffffffffffff
    Length: 115
    Type: UPDATE Message (2)
    Withdrawn Routes Length: 0
    Total Path Attribute Length: 92
  Path attributes
    Path Attribute - ORIGIN: INCOMPLETE
    Path Attribute - AS_PATH: empty
    Path Attribute - MULTI_EXIT_DISC: 86
    Path Attribute - LOCAL_PREF: 100
    Path Attribute - EXTENDED_COMMUNITIES
    Path Attribute - MP_REACH_NLRI
      Flags: 0x80, Optional, Non-transitive, Complete
      Type Code: MP_REACH_NLRI (14)
      Length: 33
      Address family identifier (AFI): IPv4 (1)
      Subsequent address family identifier (SAFI): Labeled VPN Unicast (128)
      Next hop network address (12 bytes)
      Number of Subnetwork points of attachment (SNPA): 0
      Network layer reachability information (16 bytes)
        BGP Prefix
            Prefix Length: 120
            Label Stack: 25 (bottom)
            Route Distinguisher: 100:100
            MP Reach NLRI IPv4 prefix: 170.0.0.0
```

6.4 BGP 메시지 형식

- Notification 메시지의 용도
  - 에러
- Keepalive 메시지의 용도
  - Open에 대한 수락
  - 세션의 살아있다는 것을 신호



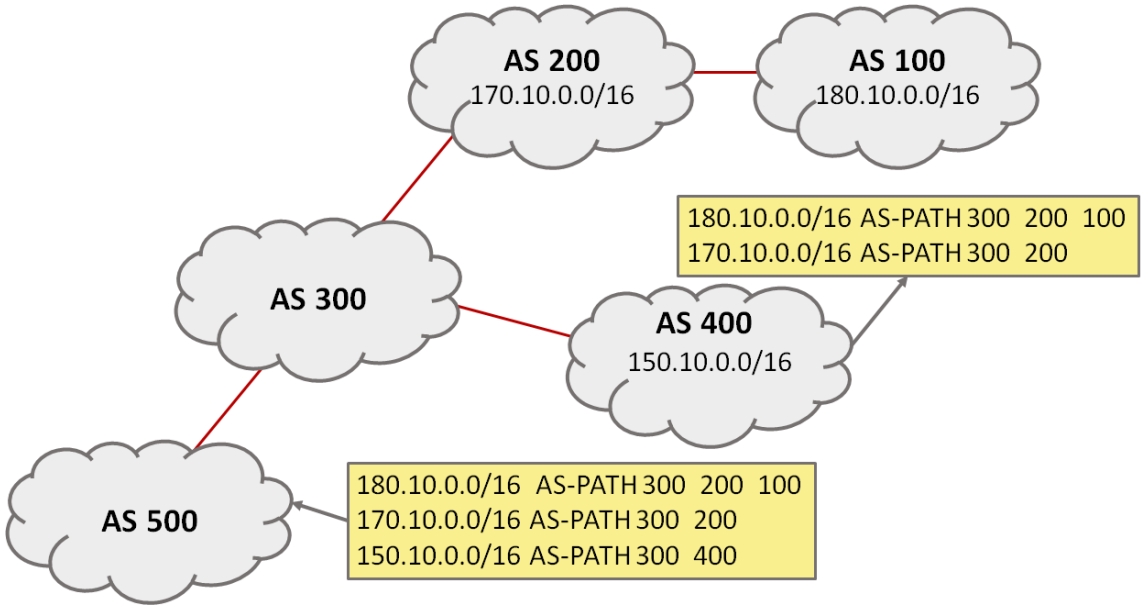

```
Border Gateway Protocol - NOTIFICATION Message
   Marker: ffffffffffffffffffffffffffffffff
   Length: 23
   Type: NOTIFICATION Message (3)
   Major error Code: OPEN Message Error (2)
   Minor error Code (Open Message): Bad Peer AS (2)
   Bad Peer AS: 65200
```

6.5 BGP 경로 애트리뷰트

- 종류
  - ➢ Well-known: 모든 BGP 라우터가 인식 및 구현 의무, Update에 포함할 수 있음
    - ✓ Mandatory: 모든 경로 엔트리에 포함
      - ❖ ORIGIN, AS-PATH, NEXT-HOP
    - ✓ Discretionary: 모든 BGP 라우터가 인식해야, 포함 의무 없음
      - ❖ LOCAL-PREF, ATOMIC-AGGREGATE
  - ➢ Optional: Update에 포함되지 않을 수 있음, 모든 BGP 라우터에 지원 의무 없음
    - ✓ Transitive: 인식 실패시 다른 BGP 라우터들에게 전달은 해야
      - ❖ AGGREGATOR, COMMUNITY
    - ✓ Non-transitive : 인식 못하면 다른 BGP 라우터들에 전파 금지
      - ❖ MULTI-EXIT-DESC

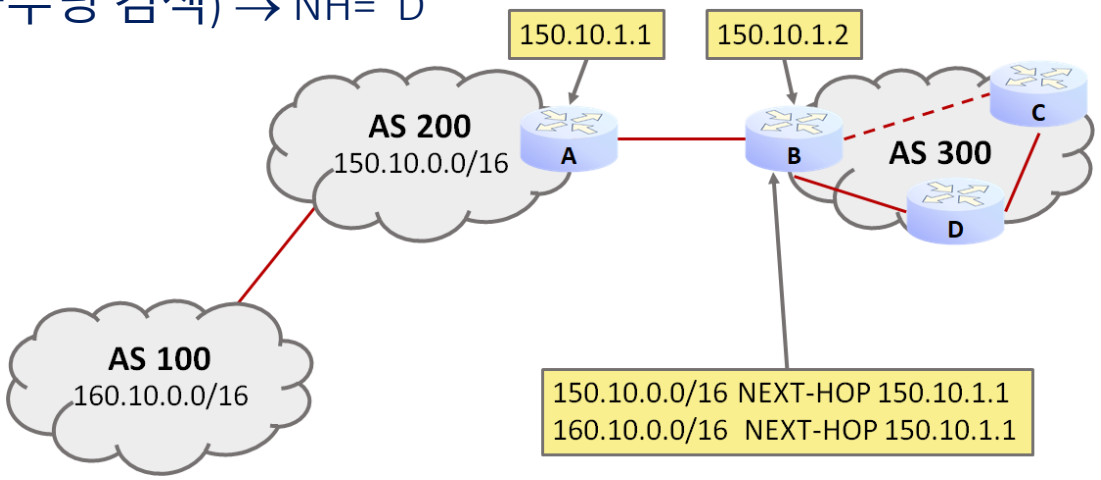
6.5 BGP 경로 애트리뷰트

- AS-PATH
  - ➢ BGP announcement가 전달되어 온 AS 경로
    - ✓ BGP는 “path vector” 프로토콜
  - ➢ 라우팅 loop 탐지



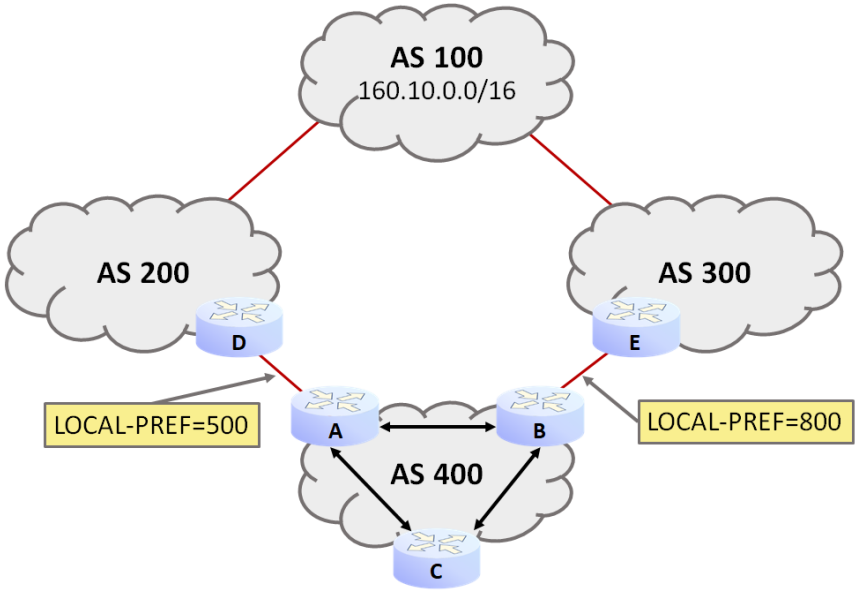
6.5 BGP 경로 애트리뷰트

- NEXT-HOP
  - ➢ 분배된 prefix로 가기 위해 거쳐야 할 다음 홉 라우터의 IP 주소
  - ➢ 특히 eBGP → 같은 link에 있는 eBGP peer의 IP 주소
  - ➢ 다음 홉이 같은 link에 연결되어 있지 않은 경우 → recursive route lookup 사용
    - ✓ C: 160.10.0.0/16 → NH=150.10.1.1 → (IGP 라우팅 검색) → NH="D"



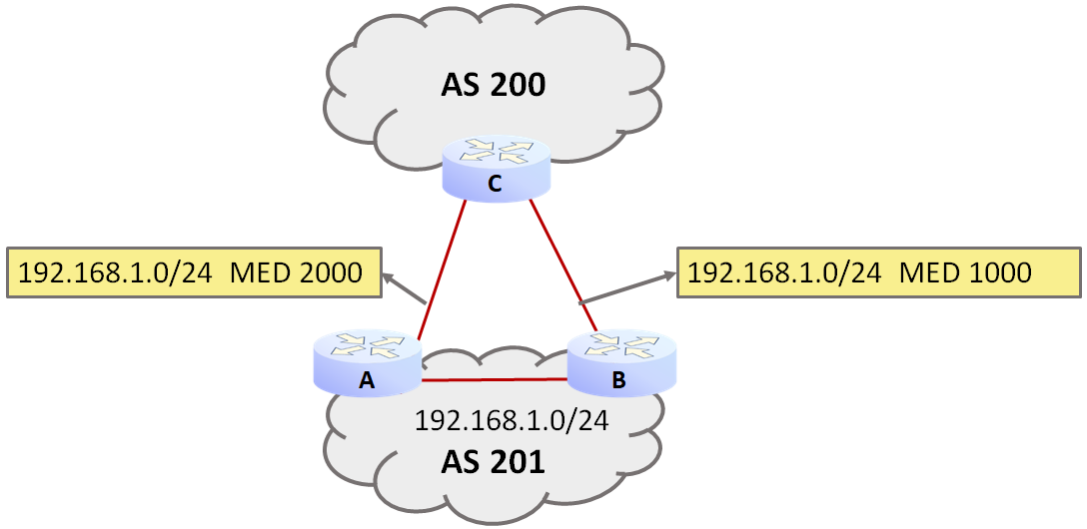
6.5 BGP 경로 애트리뷰트

- LOCAL-PREF
  - ➢ 특정 목적지 prefix를 향해 현 AS를 나가는 경로의 선호도
    - ✓ 높을 수록 좋음
  - ➢ AS 내부의 결정에만 영향을 끼침



6.5 BGP 경로 애트리뷰트

- MULTI-EXIT-DISC (MED)
  - ➢ 다른 AS에게 현 AS로 들어오는 경로 중 선택해 달라고 요청할 때 쓰는 애트리뷰트
    - ✓ 낮을 수록 좋음
    - ✓ IGP 메트릭 값 (7장 참고)이 쓰일 수 있음



6.5 BGP 경로 애트리뷰트

- ORIGIN
  - ➢ Prefix 정보가 애초에 어느 source로부터 얻어졌는지 표현
  - ➢ 종류
    - ✓ IGP 생성
    - ✓ EGP 생성 (더 이상 쓰이지 않음)
    - ✓ 기타
      - ❖ 예) OSPF, 정적 경로로부터 유입

6.6 BGP 경로 선택

- BGP announcement 처리
  - Peer AS로부터 수신 → 내부 라우팅 정보 갱신 → 다른 AS에 전파

- 최적 경로 선택 처리 규칙
  - 높은 LOCAL-PREF > 짧은 AS-PATH > 작은 ORIGIN > 작은 MED > [eBGP 습득 > iBGP 습득] > 가까운 NEXT-HOP > 작은 라우터 ID