

# Các giao thức định tuyến

Border Gateway Protocol

TS. Trương Diệu Linh

Bộ môn Mạng thông tin & Truyền thông

Viện Công nghệ thông tin

# Mục lục

- Tổng quan về BGP
- Hoạt động của BGP
- Cấu hình BGP

# Tổng quan BGP

- ◆ BGP, viết tắt của từ tiếng Anh Border Gateway Protocol, là giao thức định tuyến đa miền sử dụng trên Internet từ 1994
- ◆ Là giao thức định tuyến liên vùng (giữa các AS)
- ◆ Thuộc loại path vector và định tuyến dựa trên các luật, chính sách của quản trị hơn là các metrics nội vùng
- ◆ Phiên bản BGP hiện nay là phiên bản 4, dựa trên RFC 4271.
- ◆ BGP hỗ trợ định tuyến liên vùng không phân lớp địa chỉ và dùng kỹ thuật kết hợp đường đi để giảm kích thước bảng định tuyến
- ◆ Ví dụ: Một mạng chiếm 255 địa chỉ lớp C từ 203.162.0.0/24 - 203.162.254.0/24 thì chỉ dùng 1 địa chỉ 203.162.0.0/16 để định danh mạng).

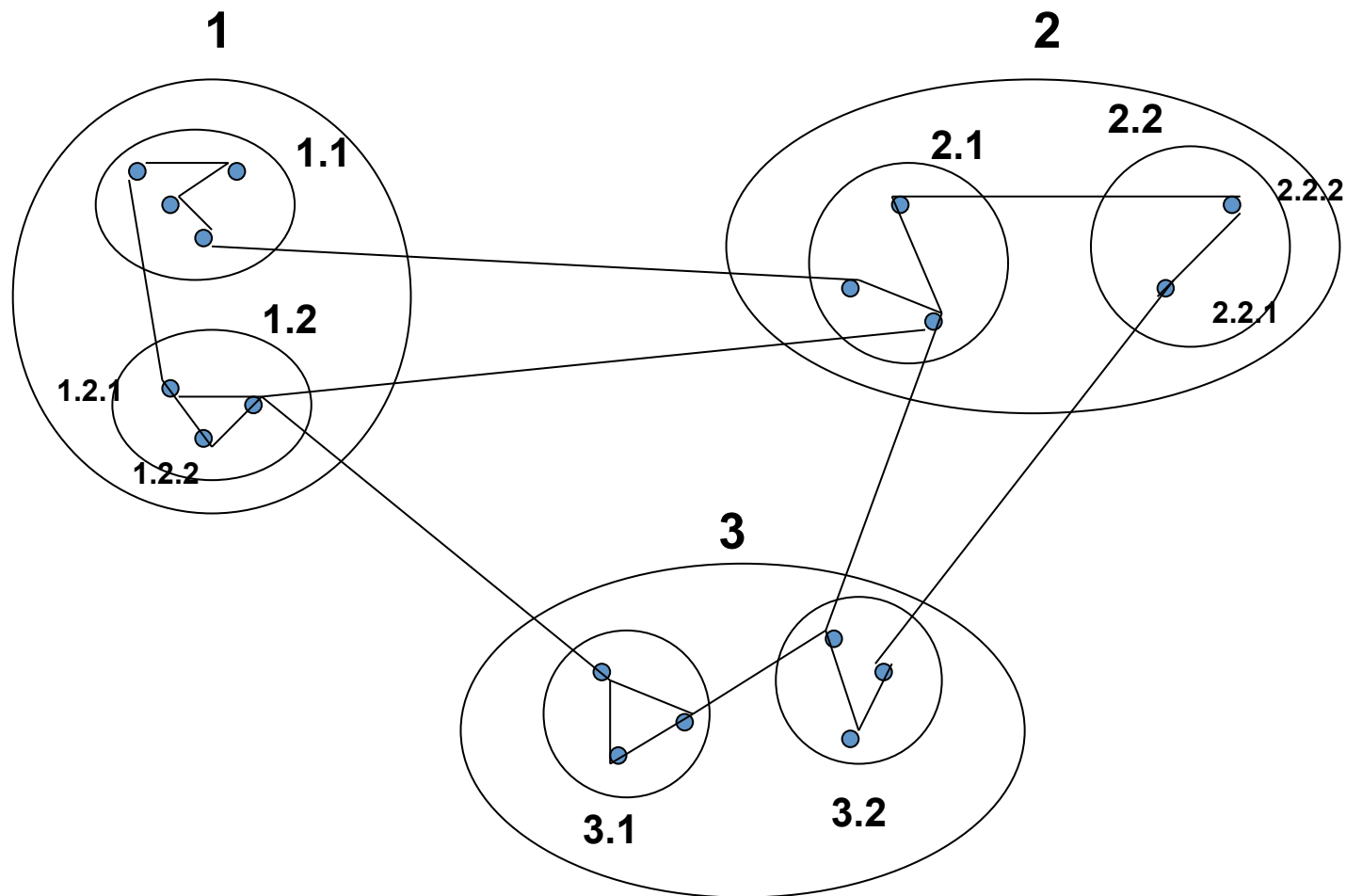
# Tổng quan BGP

- ◆ Ngoài việc sử dụng BGP giữa các AS, BGP cũng có thể được sử dụng trong các mạng riêng quy mô lớn do OSPF không đáp ứng được.
- ◆ Một lý do khác là dùng BGP để hỗ trợ kết nối đến nhiều nhà cung cấp dịch vụ, các nhà cung cấp dịch vụ Internet sử dụng BGP để trao đổi đường đi.

# Vì sao cần BGP

- ◆ Khó có một chính sách và đơn vị chi phí chung (cost) giữa các nhà cung cấp dịch vụ mạng.
- ◆ Cơ sở dữ liệu mạng quá lớn
- ◆ Mạng quá rộng, khó hội tụ
- ◆ → BGP định tuyến theo luật

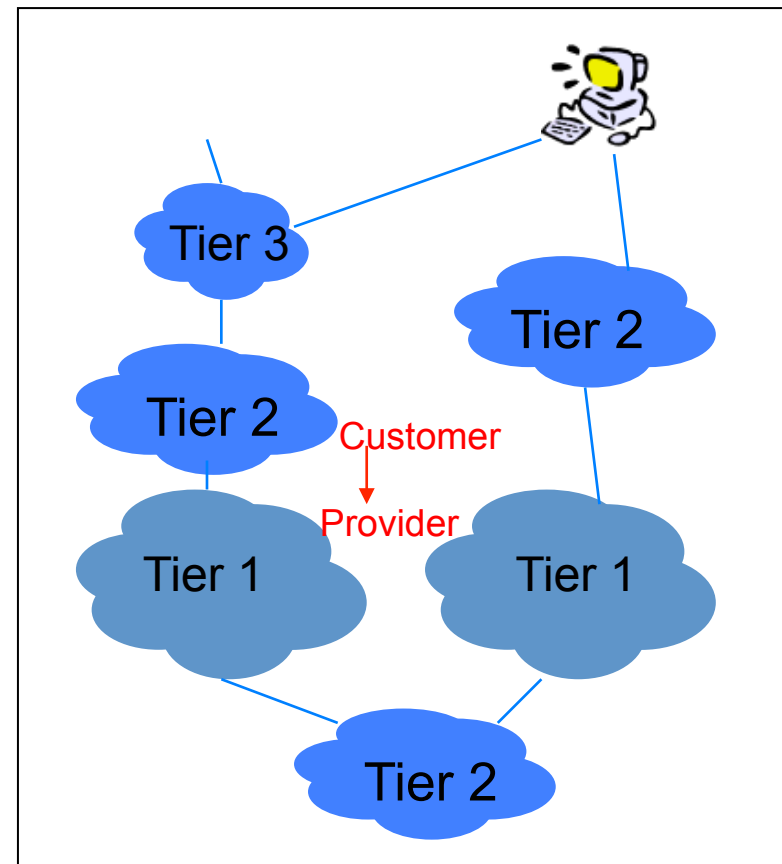
# Tổng quan BGP



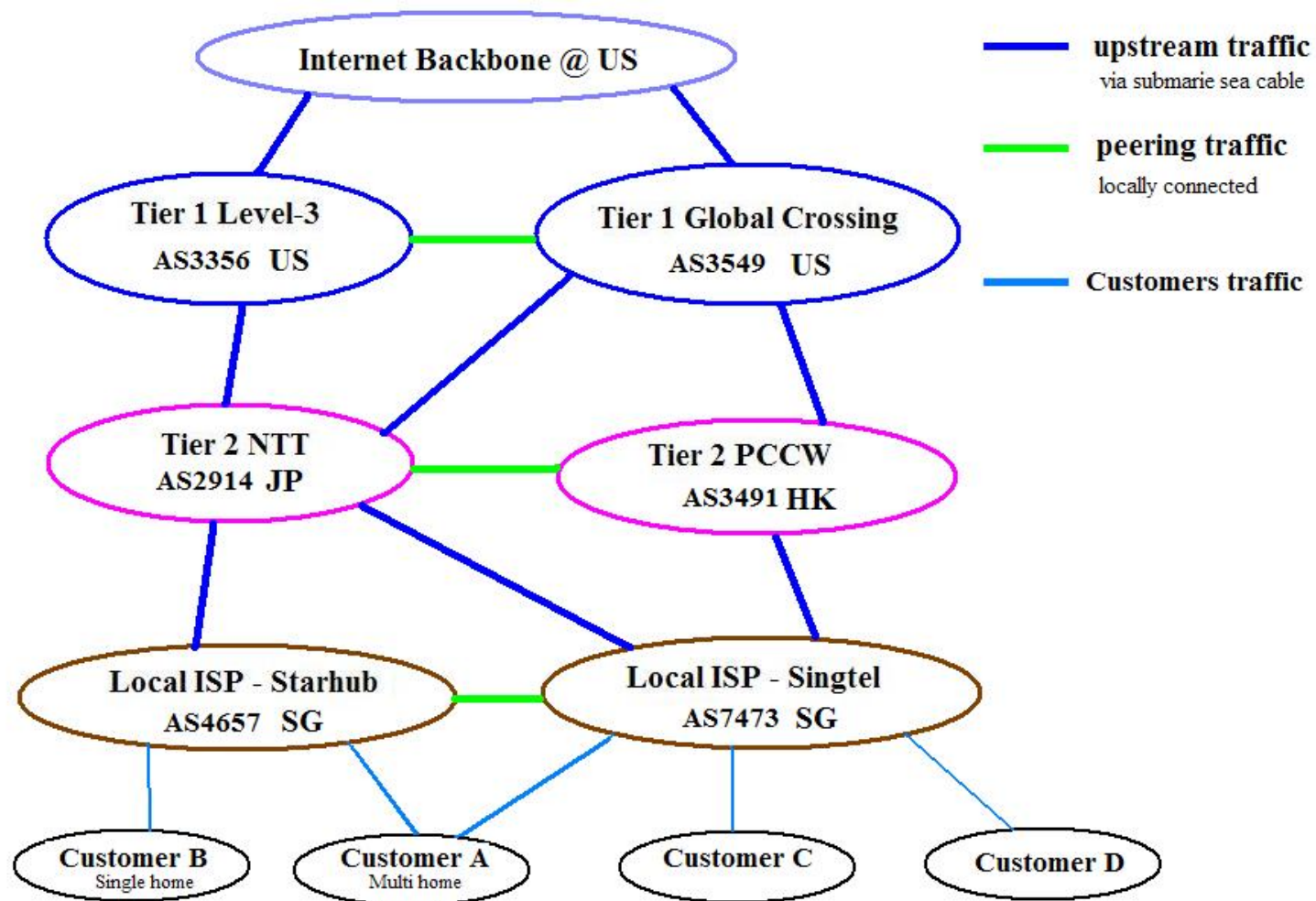
# Tổng quan BGP

## ◆ Kiến trúc logic Internet

- **Tier 1 ISP**
  - **“Default-free” with global reachability info**
- **Tier 2 ISP**
  - **Regional or country-wide**
- **Tier 3 ISP**
  - **Local**

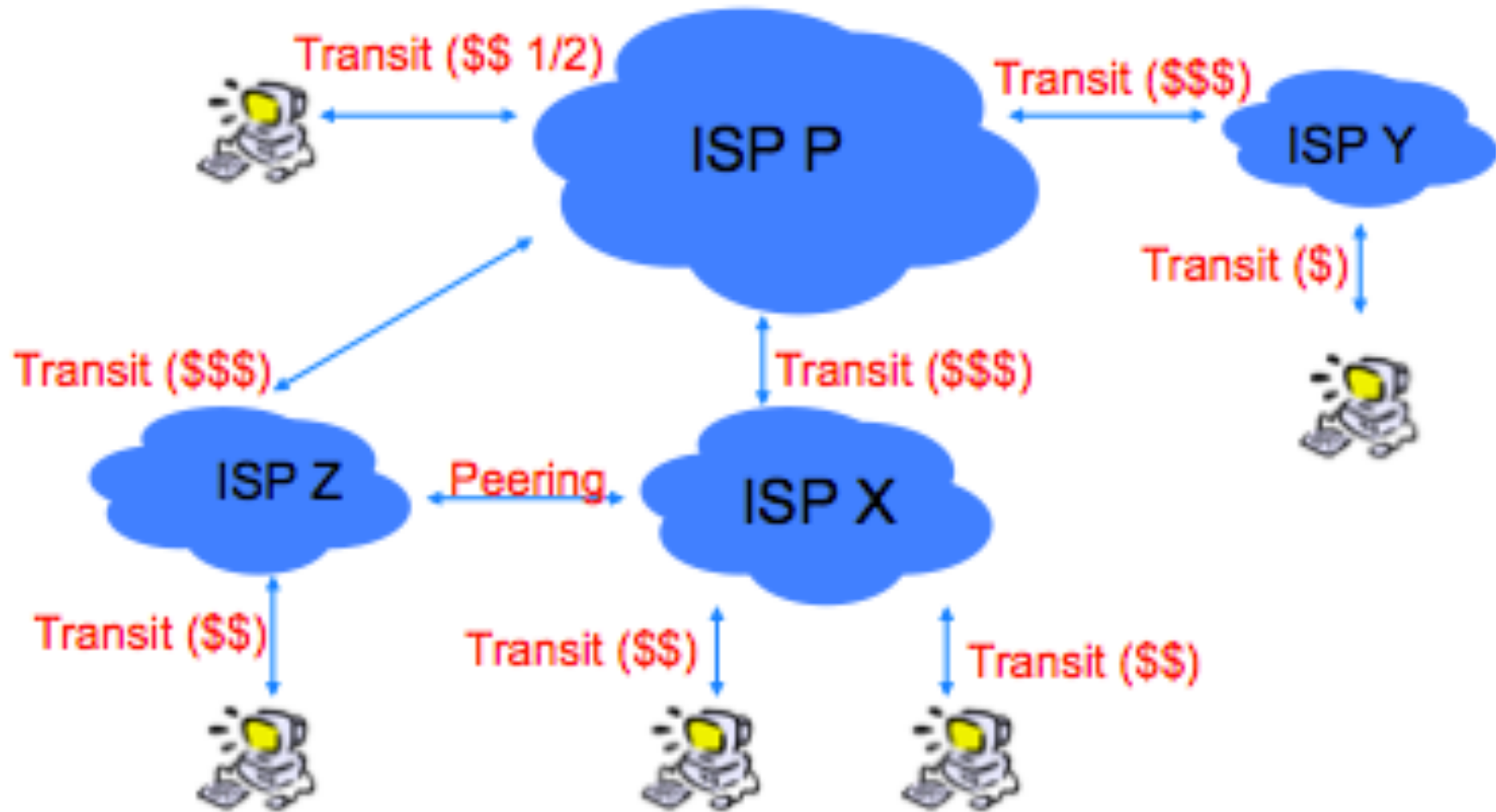


# Tổng quan BGP

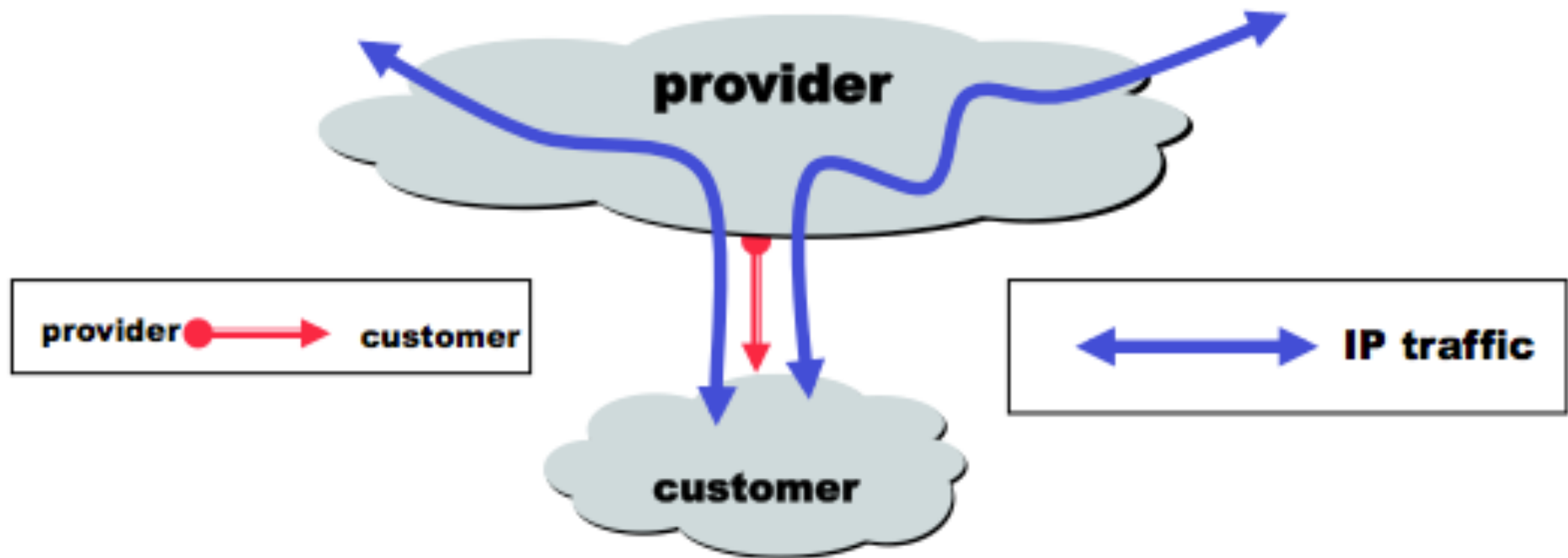




# Tổng quan BGP

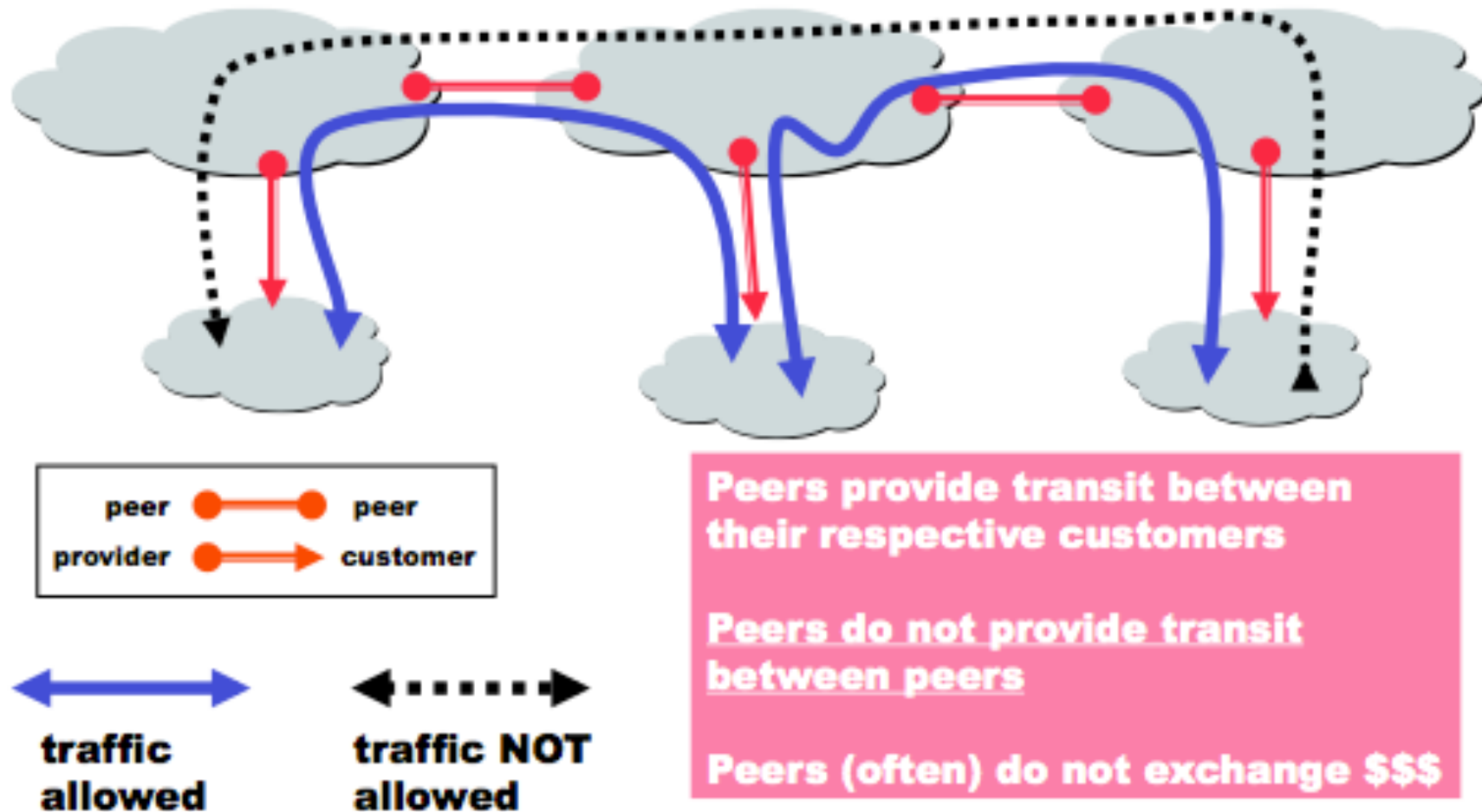


# Tổng quan BGP

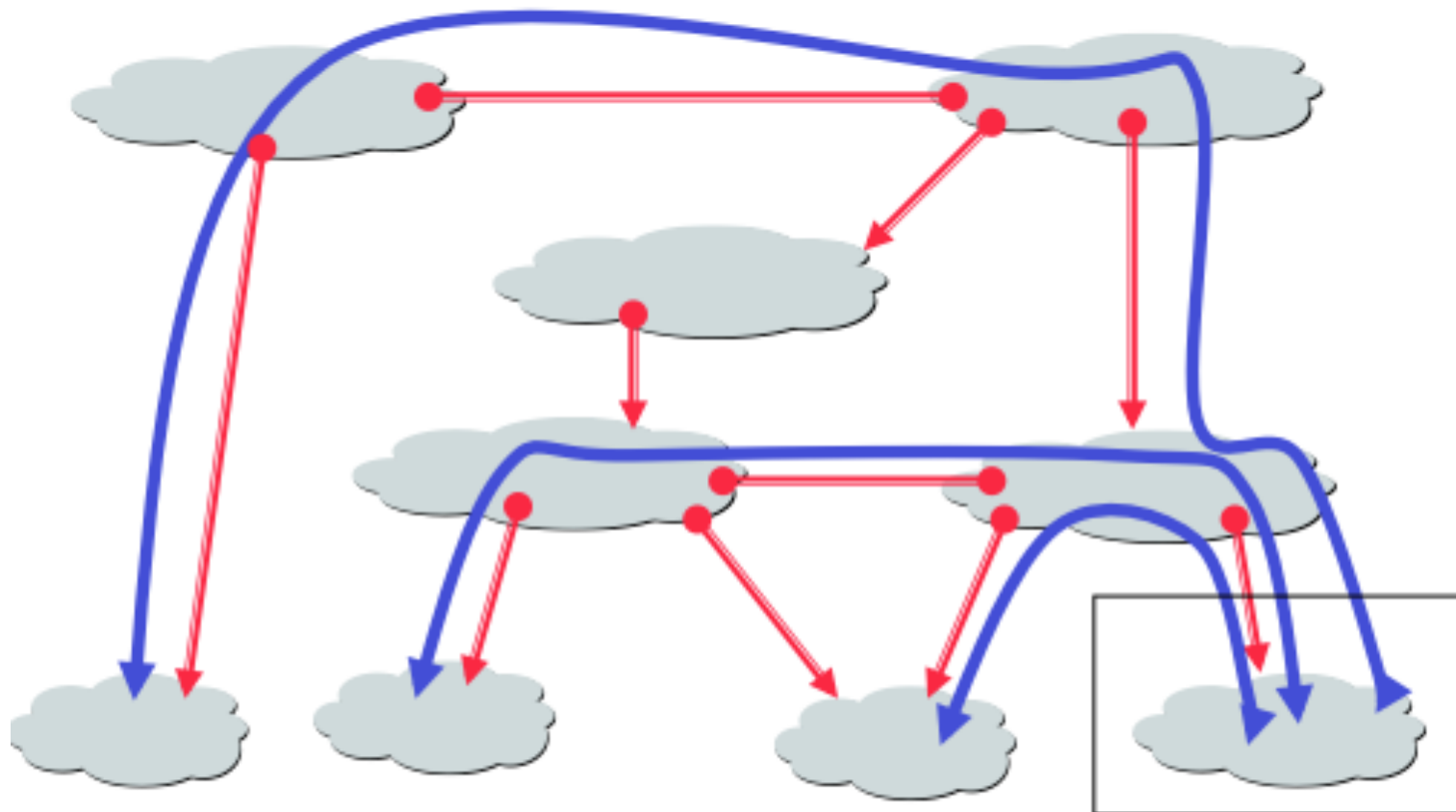


**Customer pays provider for access to the Internet**

# Tổng quan BGP



# Tổng quan BGP



**Peering also allows connectivity between the customers of "Tier 1" providers.**



# Tổng quan BGP

## Peer

- **Reduces upstream transit costs**
- **Can increase end-to-end performance**
- **May be the only way to connect your customers to some part of the Internet ("Tier 1")**

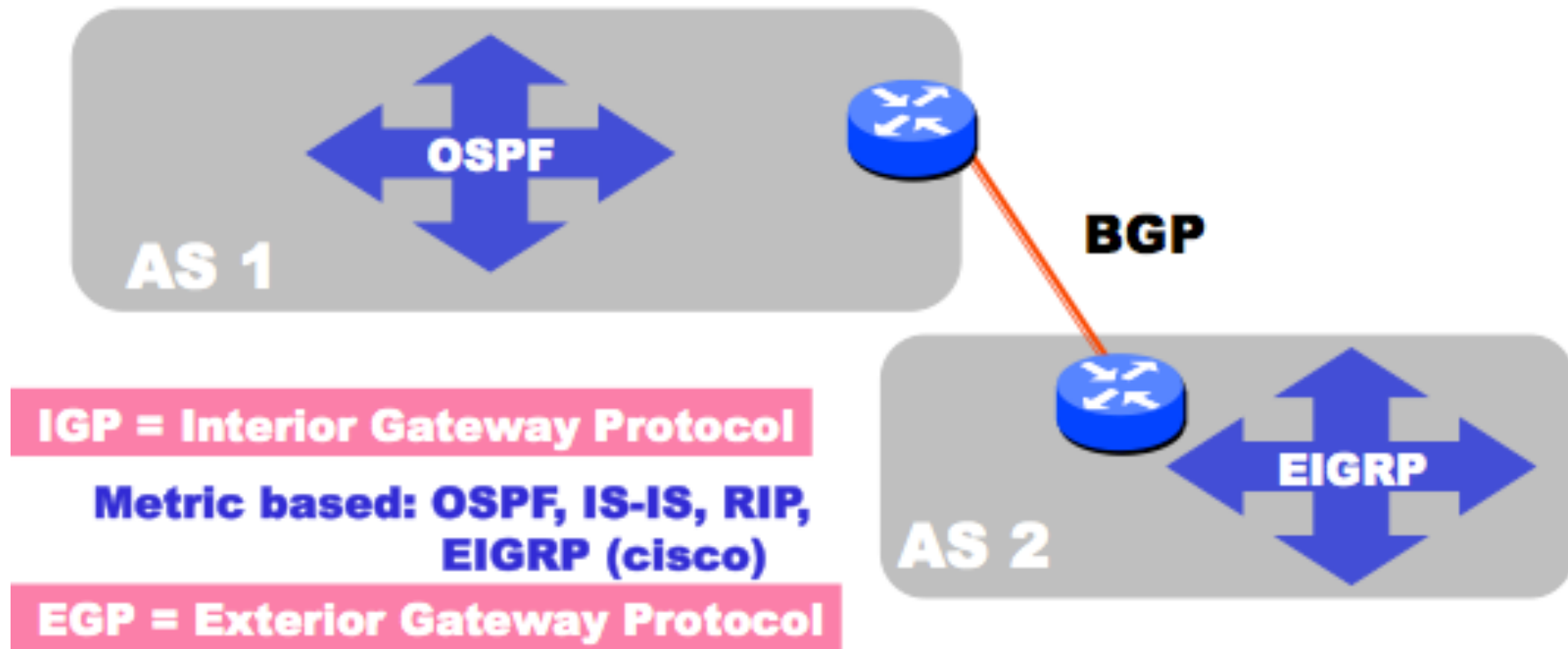
## Don't Peer

- **You would rather have customers**
- **Peers are usually your competition**
- **Peering relationships may require periodic renegotiation**

**Peering struggles are by far the most contentious issues in the ISP world!**  
**Peering agreements are often confidential.**

Peering wars

# Tổng quan BGP



**Policy based: BGP**

**The Routing Domain of BGP is the entire Internet**

Phân loại các giao thức định tuyến động

# Tổng quan BGP

- ◆ Hệ tự trị là một miền định tuyến được cung cấp một chỉ số AS.
- ◆ ASN chỉ số được định danh bởi 16 bit nhị phân. 64512-65535 được quy hoạch riêng.
- ◆ BGP cho phép trao đổi thông tin NLRI (Network Layer Reachability Information)
- ◆ BGP định tuyến bằng cách sử dụng các thuộc tính của các tuyến đường. Mỗi tuyến đường là danh sách các AS cần phải đi qua.
- ◆ BGP thường xuyên trao đổi thông tin về các tuyến đường với router xung quanh và thực hiện lọc, lựa chọn tuyến đường tốt nhất dựa trên các thuộc tính của các tuyến đường.

# Tổng quan BGP

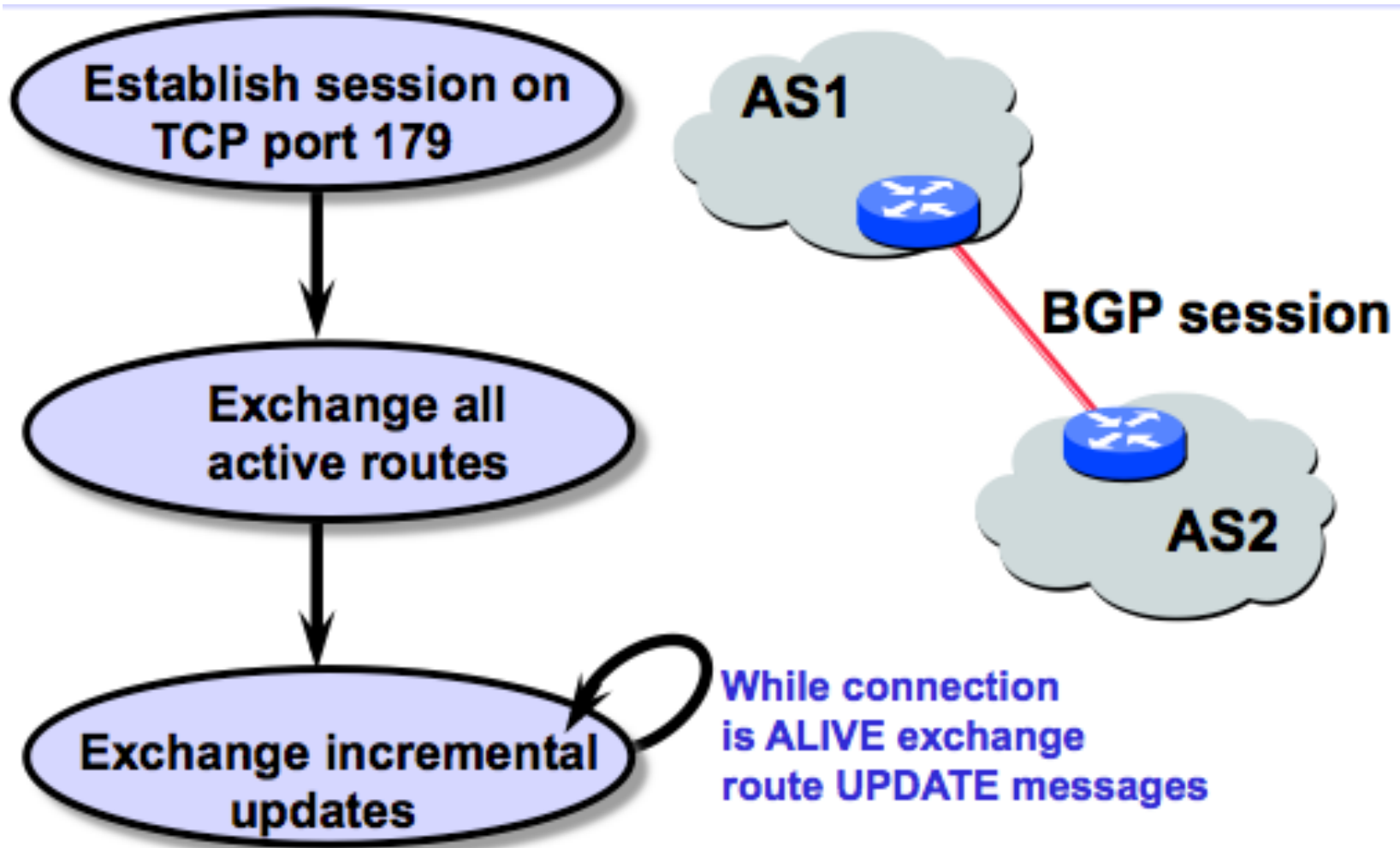
- Hoạt động chính của BGP là trao đổi thông tin reachability giữa các AS
- Thông tin này bao gồm đường đi các AS mà thông tin đi qua
  - Router BGP quảng bá một tuyến đường
  - Một router BGP khi nhận được tuyến đường có thể dùng nó trong bảng định tuyến hoặc không tùy theo các luật ưu tiên của nó.
- BGP thực hiện chuyển tiếp hoàn toàn theo địa chỉ đích và không hỗ trợ các luật không dựa trên địa chỉ đích



# BGP

- Sử dụng TCP để trao đổi dữ liệu giữa các router
- Ban đầu BGP gửi phần của bảng định tuyến tuân theo export policy đến các nút khác
- Sau đó BGP chỉ gửi một phần thay đổi của bảng định tuyến
- Trong BGP không có quá trình update thường xuyên

# BGP



# Các bản tin của BGP

## ◆ Các bản tin chính của BGP:

- ◆ **Open** : Thiết lập kết nối với nút hàng xóm
- ◆ **Keep Alive** : Bắt tay thường xuyên với hàng xóm.
- ◆ **Notification** : Thông báo đóng kết nối với hàng xóm.
- ◆ **Update** : Thông báo tuyến đường mới hoặc hủy một tuyến đường đã quảng bá trước đó.

# Các bản tin của BGP

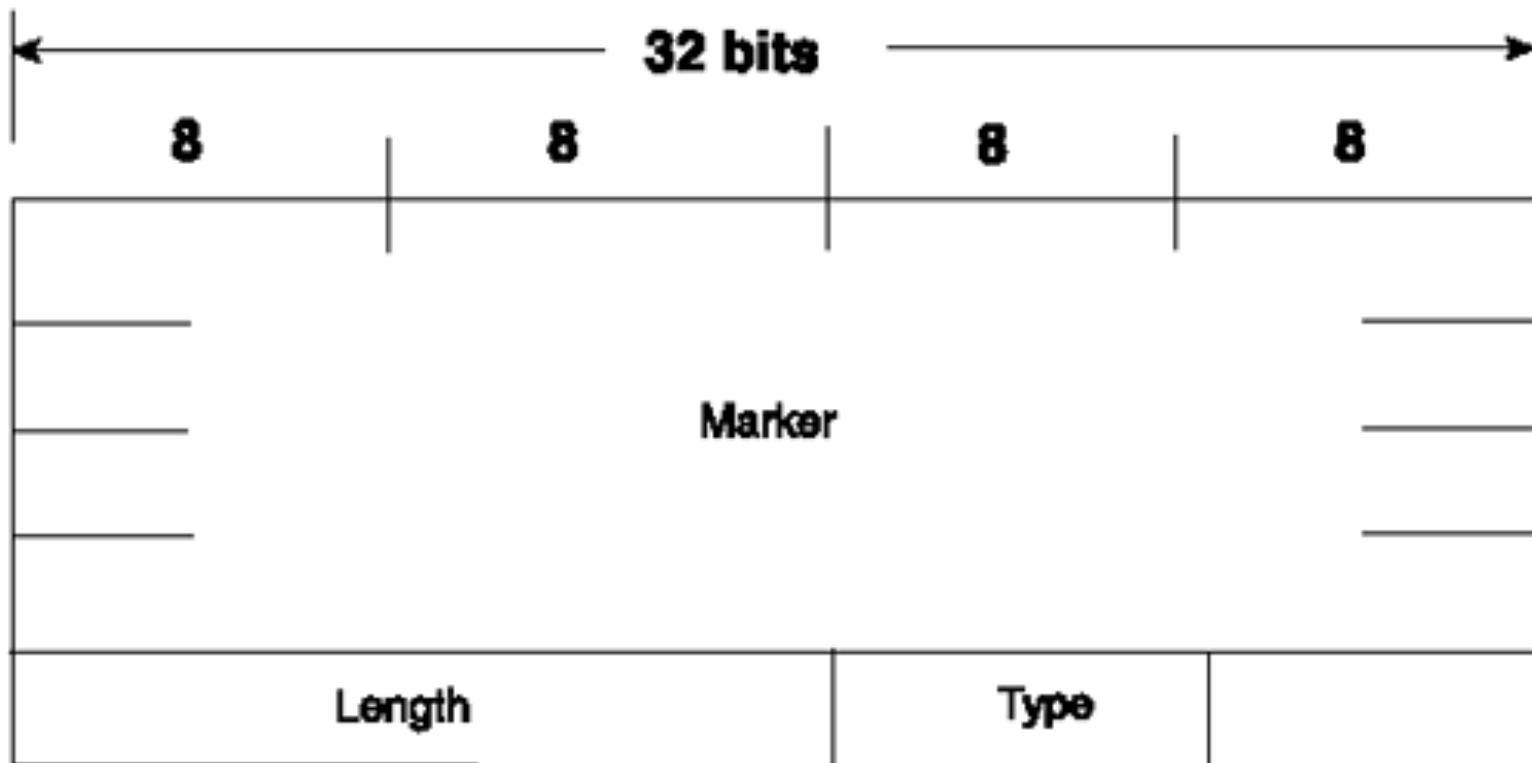
## ◆ Các bản tin của BGP:

- ❖ Được mang bởi các TCP segments sử dụng cổng TCP 179.
- ❖ Bản tin lớn nhất 4096 bytes và nhỏ nhất 19 bytes.
- ❖ Tất cả các bản tin BGP đều có chung một header và phụ thuộc vào các header khác nhau mà trường dữ liệu có thể có hoặc không (sau header)

# Các bản tin của BGP

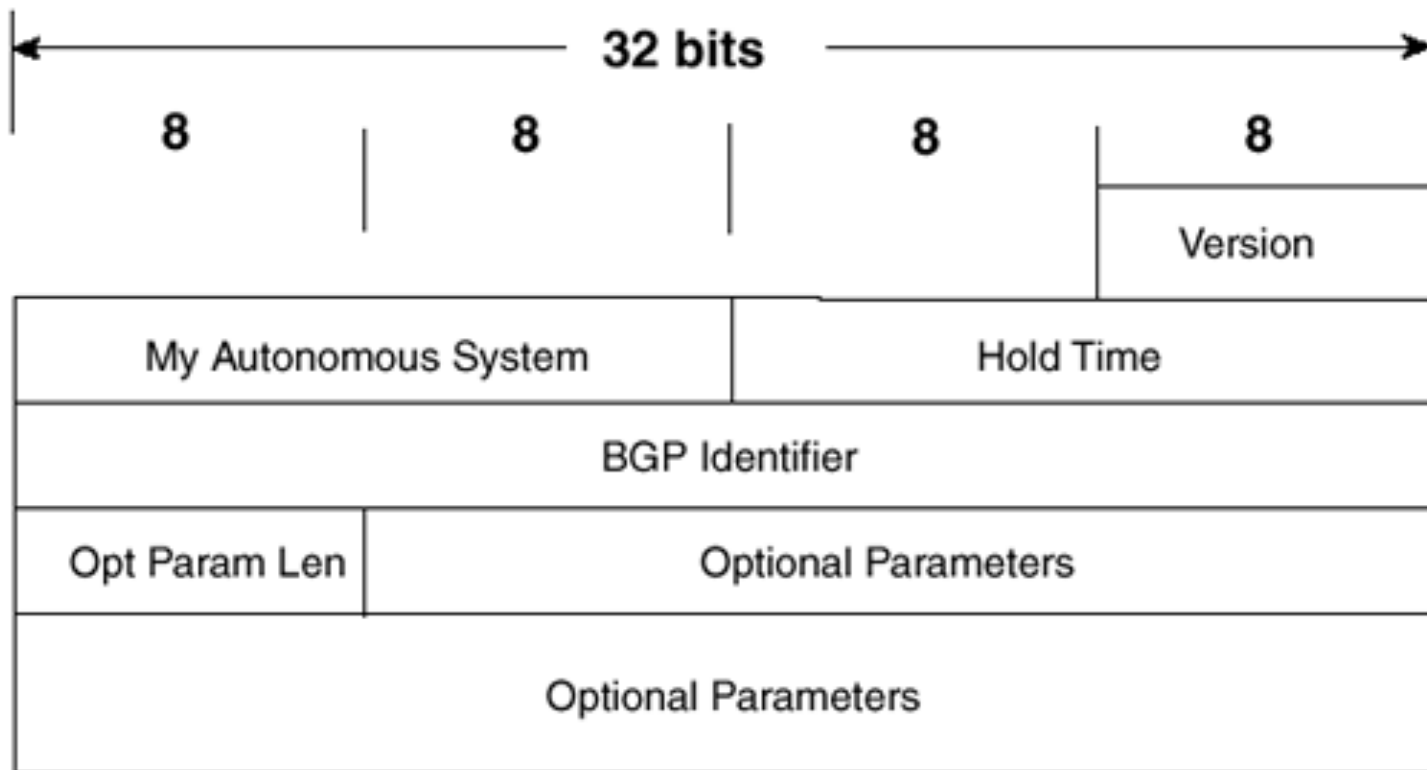
## ◆ Các bản tin của BGP:

- ❖ 1 – Open; 2 – Update; 3 – Notification và 4 – Keepalive
- ❖ Marker được sử dụng vì mục đích tương thích. Gồm toàn 1



# Bản tin Open

- ❖ Sau khi kết nối TCP được mở, một bản in Open được gửi đi từ mỗi bên
- ❖ Để confirm cho Open, 1 bản tin Keepalive được gửi trả lại



# Bản tin OPEN

- My AS: số hiệu của AS gửi
- BGP identifier: IP của BGP gửi
- Hold time: Thời gian tối đa mà nút gửi muốn duy trì liên kết mà không nhận được các gói tin KEEP\_ALIVE hoặc UPDATE hoặc NOTIFICATION.

# Bản tin Update

- ◆ Dùng để gửi thông tin đường đi giữa các nút BGP.
- ◆ Quảng bá các đường đi có cùng tính chất.
- ◆ Withdrawn route bao gồm các bộ <Length, prefix> mô tả danh sách các địa chỉ prefix (với chiều dài prefix trong length) sẽ bị rút khỏi bảng định tuyến.

|   |
|---|
| Withdrawn Routes Length (2 octets)                |
| Withdrawn Routes (variable)                       |
| Total Path Attribute Length (2 octets)            |
| Path Attributes (variable)                        |
| Network Layer Reachability Information (variable) |



# Bản tin update

- Trường Path Attributes
  - ❖ Gồm nhiều bộ 3 : <attribute type code, attribute length, attribute value>
  - ❖ Mỗi Attribute type code gồm 2 byte có cấu trúc như sau:

|   |   |   |   |   |   |   |   |                     |
|---|---|---|---|---|---|---|---|---------------------|
| O | T | P | E | U | U | U | U | Attribute Type Code |
|---|---|---|---|---|---|---|---|---------------------|

## Flag bits

O: Optional bit

0 = Optional

1 = Well-Known

T: Transitive bit

0 = Transitive

1 = Non-Transitive

P: Partial bit

0 = Optional Transitive attribute is partial

1 = Optional Transitive attribute is complete

E: Extended length bit

0 = Attribute Length is one octet

1 = Attribute Length is two octets

U: Unused

# Bản tin Update

## ❖ Ý nghĩa của Attribute Type code và Attribute value

| Attribute Type Code | Attribute Type   | Attribute Value Code | Attribute Value                        |
|---------------------|------------------|----------------------|--|
| 1                   | ORIGIN           |                      | IGP                                    |
|                     |                  | 1                    | EGP                                    |
|                     |                  | 2                    | Incomplete                             |
| 2                   | AS_PATH          | 1                    | AS_SET                                 |
|                     |                  | 2                    | AS_SEQUENCE                            |
|                     |                  | 3                    | AS_CONFED_SET                          |
|                     |                  | 4                    | AS_CONFED_SEQUENCE                     |
| 3                   | NEXT_HOP         |                      | Next-hop IP address                    |
| 4                   | MULTI_EXIT_DISC  |                      | 4-octet MED                            |
| 5                   | LOCAL_PREF       |                      | 4-octet LOCAL_PREF                     |
| 6                   | ATOMIC_AGGREGATE |                      | None                                   |
| 7                   | AGGREGATOR       |                      | AS number and IP address of aggregator |
| 8                   | COMMUNITY        |                      | 4-octet community identifier           |
| 9                   | ORIGINATOR_ID    |                      | 4-octet router ID of originator        |
| 10                  | CLUSTER_LIST     |                      | Variable-length list of cluster IDs    |

# Bản tin Update

- Thuộc tính ORIGIN (Type Code 1)
  - 0 IGP - Network Layer Reachability Information is interior to the originating AS
  - 1 EGP - Network Layer Reachability Information learned via the EGP protocol [RFC904]
  - 2 INCOMPLETE - Network Layer Reachability Information learned by some other means

# Bản tin Update

- Thuộc tính AS\_PATH (Type Code 2)
  - Bao gồm một chuỗi các đoạn mô tả AS path
  - Mỗi đoạn mô tả AS path bao gồm bộ 3:
    - <path segment type, path segment length, path segment value>.
    - Path segment type:
      - 1: AS\_SET: chuỗi AS là không sắp xếp theo thứ tự các AS đi qua
      - 2: AS\_SEQUENCE: chuỗi các AS được sắp xếp theo thứ tự đi qua.
    - Path segment length: số AS trên đường nằm trong path segment value
    - Path segment value: Danh sách các số hiệu AS, mỗi số hiệu 2 byte

# Bản tin Update

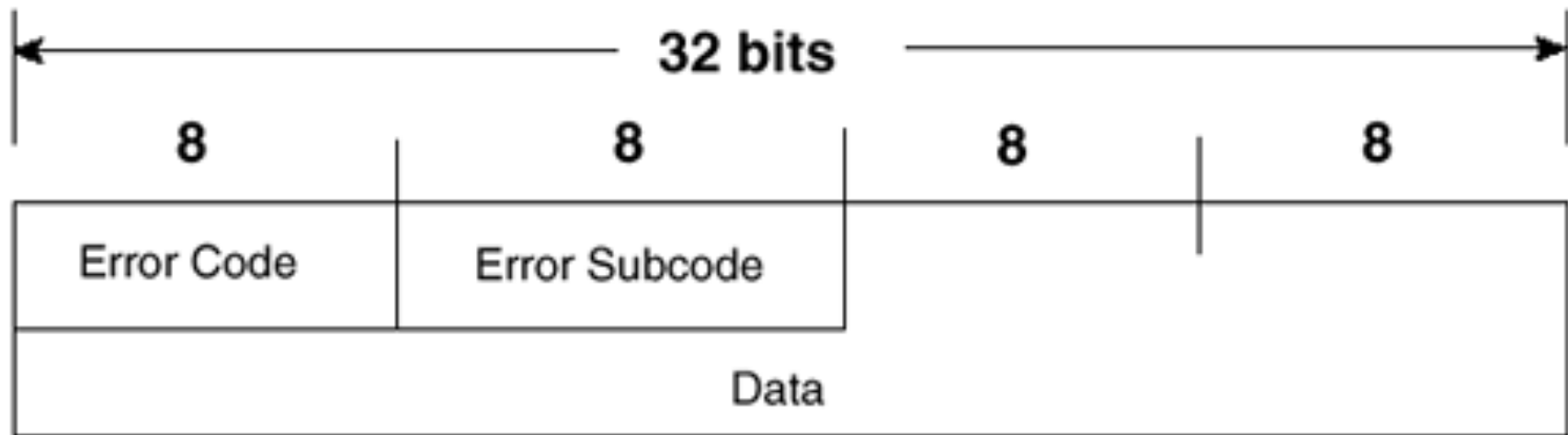
- Thuộc tính NEXT\_HOP (Type Code 3):
  - Attribute value: Địa chỉ IP của nút tiếp theo
  - -----
- Trường Network Layer Reachability Information:
  - Chứa danh sách các bộ <length, prefix>
  - Mô tả bộ prefix các mạng đích tương ứng với path đặc tả phía trên.

# Bản tin UPDATE

- Xử lý bản tin UPDATE:
  - Các tuyến đường có địa chỉ đích đặc tả trong phần prefix của trường WITHDRAWN ROUTES sẽ bị loại.
  - Các tuyến đường mới được đặc tả trong NLRI sẽ được cập nhật
  - Nút BGP thực hiện lại quá trình chọn đường Decision process và áp các chính sách để :
    - Chọn đường sẽ được dùng để định tuyến bởi router này.
    - Chọn đường sẽ được quảng bá.
    - Route aggregation.

# Bản tin NOTIFICATION

- ◆ Một thông báo NOTIFICATION được sử dụng khi có lỗi xuất hiện khiến cho phiên BGP đó kết thúc. Trong quá trình thực hiện BGP có thể có nhiều lỗi xảy ra

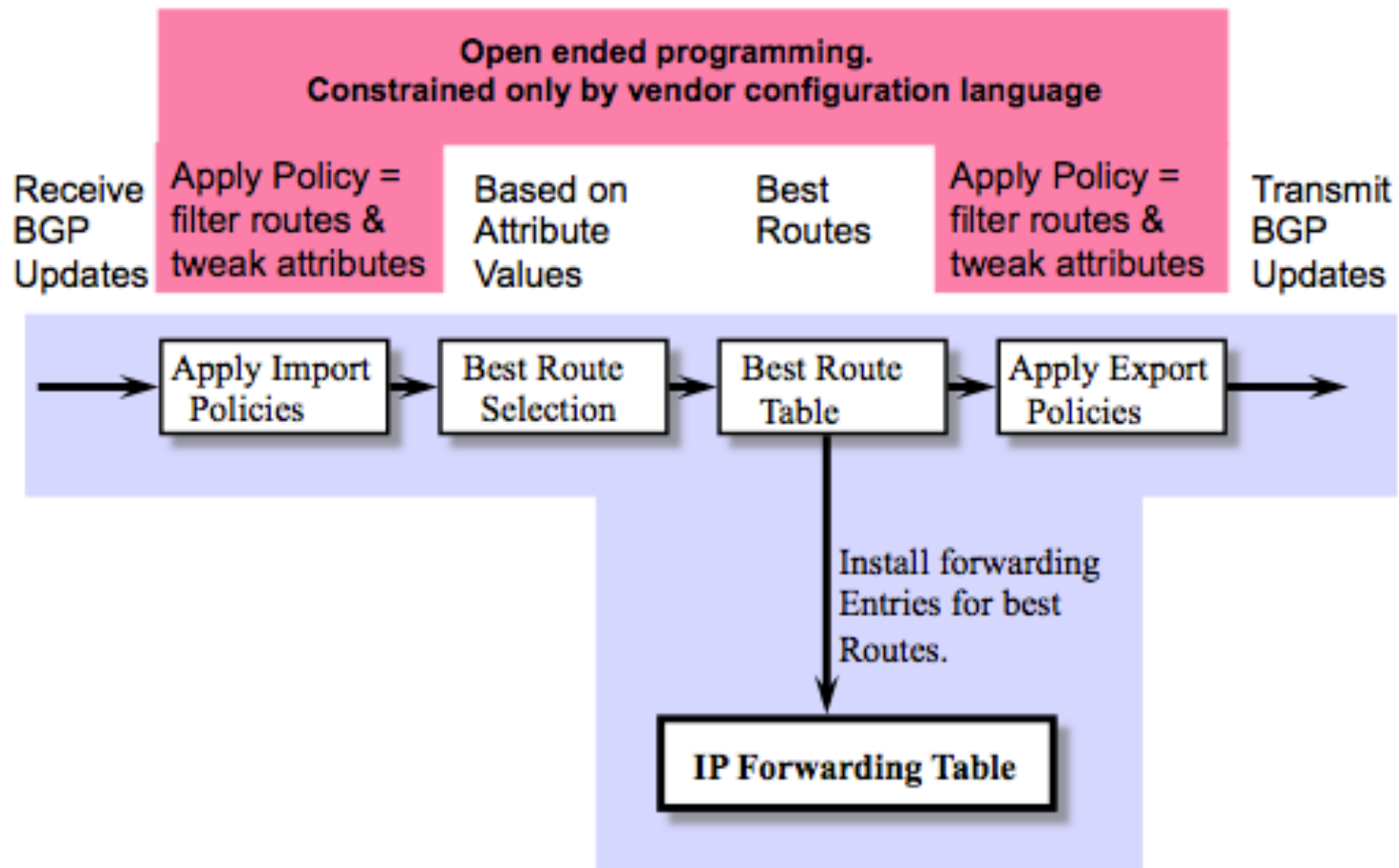


# Bản tin KEEP ALIVE

- ◆ *thông báo KEEPALIVE không chứa dữ liệu mà chỉ chứa phần tiêu đề*



# Hoạt động của BGP



# BGP

- Việc nhận thông tin một tuyến đường tuân theo import policy
- Việc lựa chọn tuyến đường tốt nhất dựa trên các thuộc tính của tuyến đường
- Đưa tuyến tốt nhất vào bảng định tuyến
- Việc quảng bá nội dung bảng định tuyến tuân theo export policy

# Các thuộc tính đường đi

- Các thuộc tính được chia thành 4 nhóm:
- Well-known
  - Mọi bản cài đặt BGP đều phải hiểu
  - well-known mandatory: thuộc tính phải có trong mọi UPDATE
  - well-known discretionary:
- Optional
  - Không nhất thiết được hiểu bởi 1 bản cài đặt BGP.
  - optional transitive: được lan truyền đi tiếp
  - optional non-transitive: không được lan truyền đi tiếp

# Các thuộc tính đường đi

- **Well-known mandatory:** các thuộc tính nhóm này là bắt buộc và được công nhận bởi tất cả các router BGP.
- **Well-known Discretionary:** không yêu cầu các thuộc tính này tồn tại trong các bản tin UPDATE nhưng nếu chúng tồn tại, tất cả các router sẽ công nhận và sẽ có hành động tương ứng dựa trên thông tin được chứa bên trong thuộc tính này.

# Các thuộc tính đường đi

- ◆ **Optional Transitive:** Router có thể không công nhận các thuộc tính này nhưng nếu router nhận được thuộc tính này, nó sẽ đánh dấu và gửi đầy đủ cập nhật này đến router kế tiếp. Các thuộc tính sẽ không thay đổi khi đi qua một router nếu thuộc tính này không được công nhận bởi router này.
- ◆ **Optional Nontransitive:** Các thuộc tính này bị loại bỏ nếu cập nhật mang thuộc tính này đi vào router mà router không hiểu hoặc không công nhận thuộc tính.

# Các thuộc tính đường đi

| Attribute Name  | Attribute Type           | Description   |
|-----------------|--------------------------|---|
| AS_PATH         | Well-known mandatory     | Danh sách các AS mà tuyến được quảng bá thông qua nó  |
| WEIGHT          | Cisco defined attribute  | Được sử dụng để định tuyến trong BGP  |
| LOCAL_PREF      | Well-known discretionary | Được sử dụng để chọn tuyến, thường được sử dụng trong một AS  |
| MULTI_EXIT_DISC | Optional nontransitive   | Được sử dụng để chọn tuyến, thường được sử dụng giữa các AS   |
| ORIGIN          | Well-known mandatory     | Chỉ ra các tuyến BGP được học như thế nào<br>IGP— Từ câu lệnh Network<br>EGP— Từ EBGp hàng xóm<br>Incomplete— Từ redistribution |

# Các thuộc tính đường đi

|           |                      |   |
|-----------|----------------------|---|
| NEXT_HOP  | Well-known mandatory | EBGP— Địa chỉ của giao diện được sử dụng để truyền thông với hàng xóm bên khác<br>IBGP— EBGP next hop được gửi và không thay đổi khi tới các IBGP hàng xóm. |
| Community | Optional transitive  | Một số được sử dụng cho việc áp một chính sách đến một nhóm các tuyến.  |

# Các thuộc tính đường đi

## ◆ Thuộc tính AS\_Path:

❖ AS\_PATH thuộc vào nhóm well-known mandatory, cần được quảng bá và hiểu bởi tất cả các router hàng xóm.

❖ Ví dụ:





# Các thuộc tính đường đi

## ◆ Thuộc tính AS\_Path:

- ❖ AS\_PATH là danh sách các số AS của một tuyến đường được quảng bá.
- ❖ Nếu có nhiều tuyến, thuộc tính AS\_PATH sẽ được sử dụng để lựa chọn tuyến tốt nhất tới đích
- ❖ Ngoài việc AS\_PATH được sử dụng cho việc lựa chọn tuyến tốt nhất đến đích, BGP còn sử dụng thuộc tính AS\_PATH để chống loop

# Các thuộc tính đường đi

## ◆ Thuộc tính AS\_Path:

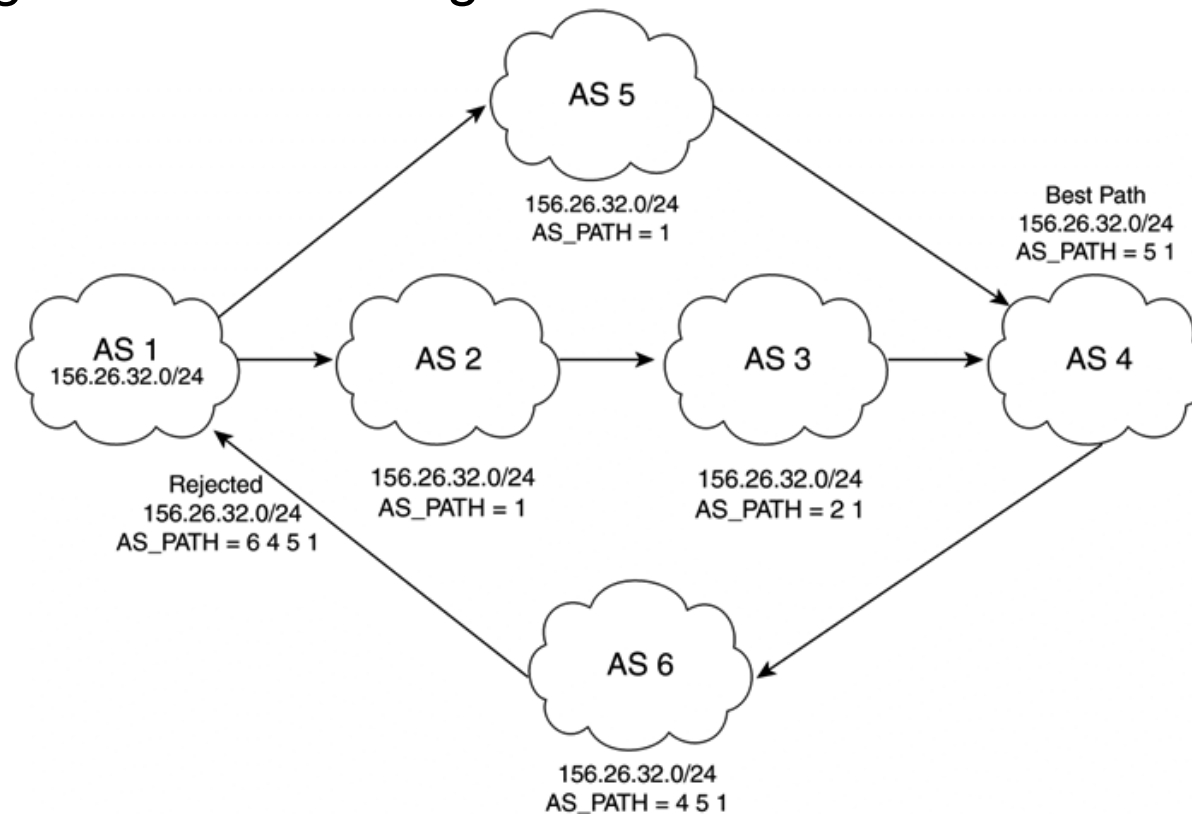
✓ Ví dụ chọn tuyến theo AS\_Path



# Các thuộc tính đường đi

## ◆ Thuộc tính AS\_Path:

- ✓ Các BGP routers từ chối bất kì sự quảng bá nào mà chứa đựng số AS của chúng.



# Các thuộc tính đường đi

## ◆ Thuộc tính Weight:

- ❖ Weight là thuộc tính không được chỉ định trong BGP.
- ❖ Nó là thuộc tính được phát triển trên các thiết bị của Cisco và là thuộc tính do Cisco định nghĩa.
- ❖ Weight là một trong các thuộc tính được sử dụng để quyết định tuyến tốt nhất tới đích.
- ❖ Weight là thuộc tính được biểu diễn bằng 16 bit và giá trị trong khoảng 0 đến 65535.
- ❖ **Chú ý:** Thuộc tính weight chỉ có giá trị trên router đó và không quảng bá ra các router hàng xóm chạy BGP. Các tuyến BGP có Weight mặc định là 32768

# Các thuộc tính đường đi

## ◆ Thuộc tính Weight:

- ✓ Ví dụ chọn tuyến thông qua thuộc tính Weight

Figure 8-6. WEIGHT Attribute



# Các thuộc tính đường đi

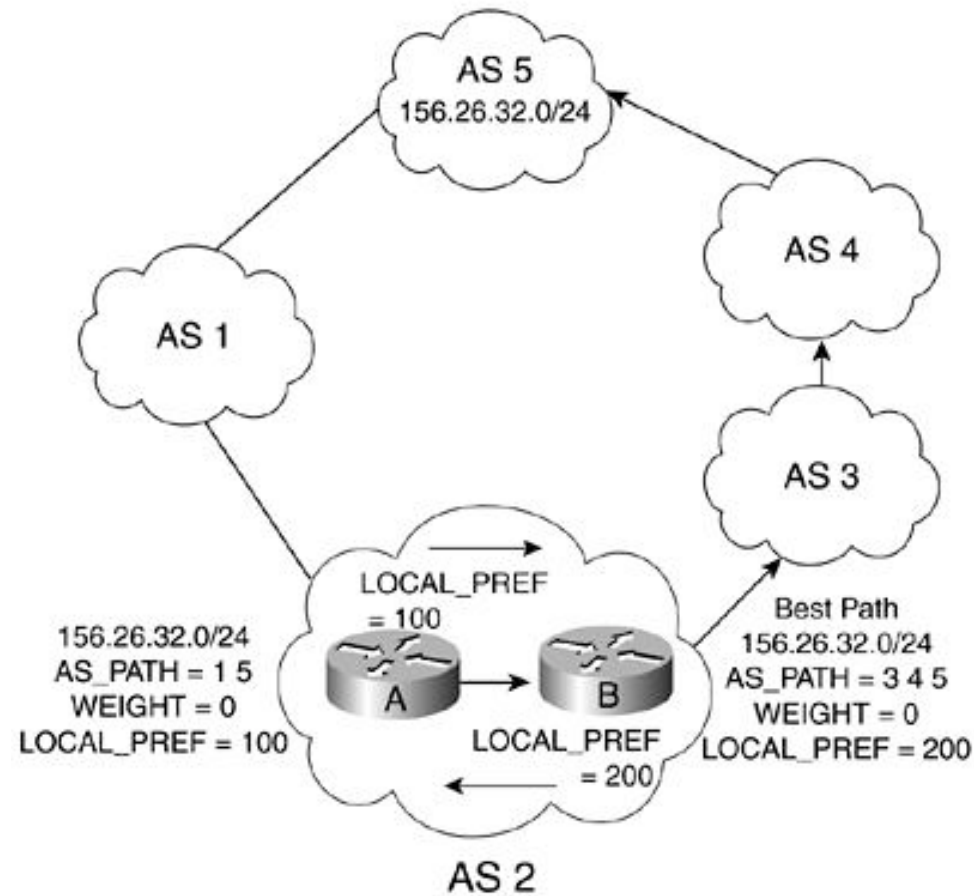
## ◆ Thuộc tính Local\_Pref:

- ❖ Giá trị mà một nút BGP dùng để thông báo cho các nút trong nội bộ AS về mức độ ưu tiên của nút BGP này đối với 1 tuyến đường.
- ❖ LOCAL\_PREF là thuộc tính tương tự như weight, khi có nhiều hơn một tuyến tới một đích, tuyến có thuộc tính LOCAL\_PREF cao nhất (khi các weight bằng nhau) sẽ được lựa chọn là tuyến tốt nhất.
- ❖ Mặc định LOCAL\_PREF có giá trị 100 và giá trị cao hơn sẽ được ưu tiên sử dụng là tham số quyết định tuyến tới đích (khi weight bằng nhau).
- ❖ LOCAL\_PREF là một số 32 bit có giá trị từ 0 – 4294967295

# Các thuộc tính đường đi

## ◆ Thuộc tính Local\_Pref:

✓ Ví dụ:



# Các thuộc tính đường đi

## ◆ Thuộc tính Local\_Pref:

- ❖ Không giống như thuộc tính weight LOCAL\_PREF được quảng bá trong 1 AS.
- ❖ Router A nhận quảng bá cho mạng 156.26.32.0/24 từ Router B với LOCAL\_PREF là 200. Vì thế Router A sử dụng tuyến thông qua AS 3 để đạt tới mạng 156.26.32.0/24



# Các thuộc tính đường đi

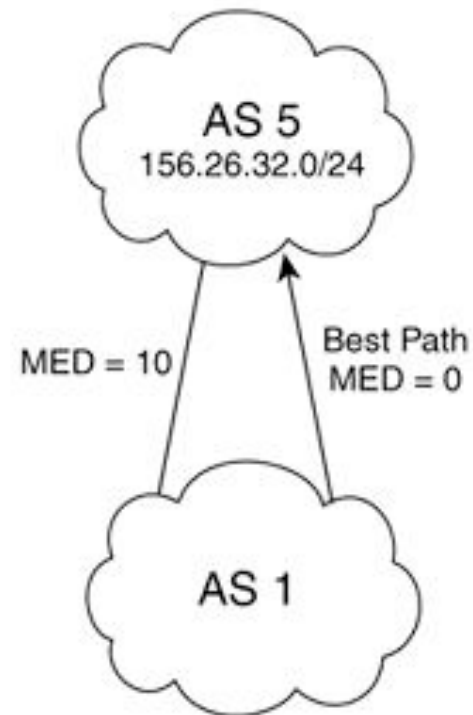
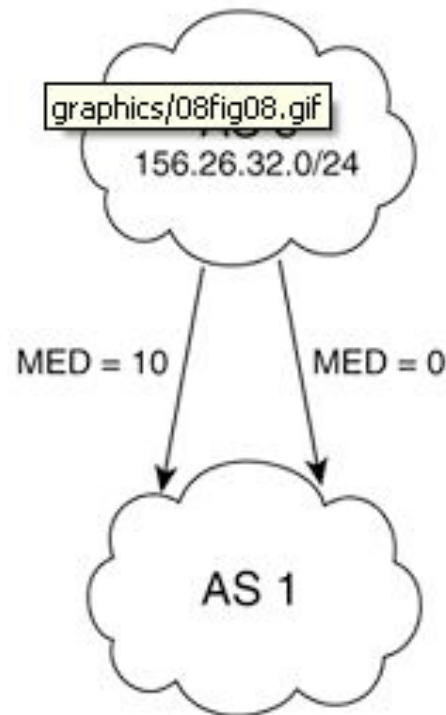
## ◆ Thuộc tính MED (MULTI\_EXIT\_DISC):

- ❖ Thuộc tính Weight được sử dụng để quyết định tuyến đường từ một router.
- ❖ LOCAL\_PREF được sử dụng để lựa chọn tuyến đường rời khỏi 1 autonomous system.
- ❖ Thuộc tính MULTI\_EXIT\_DISC (Multi-exit discriminator hay MED) được sử dụng để quyết định đường đi ra khỏi AS đến các AS khác trong trường hợp có nhiều đường ra.
- ❖ MED cũng là một số có giá trị 32 bit tuy nhiên giá trị thuộc tính thấp hơn sẽ được ưu tiên thay vì giá trị cao sẽ được ưu tiên để chọn tuyến như trong LOCAL\_PREF

# Các thuộc tính đường đi

## ◆ Thuộc tính MED (MULTI\_EXIT\_DISC):

✓ Ví dụ:



# Các thuộc tính đường đi

## ◆ Thuộc tính **ORIGIN**:

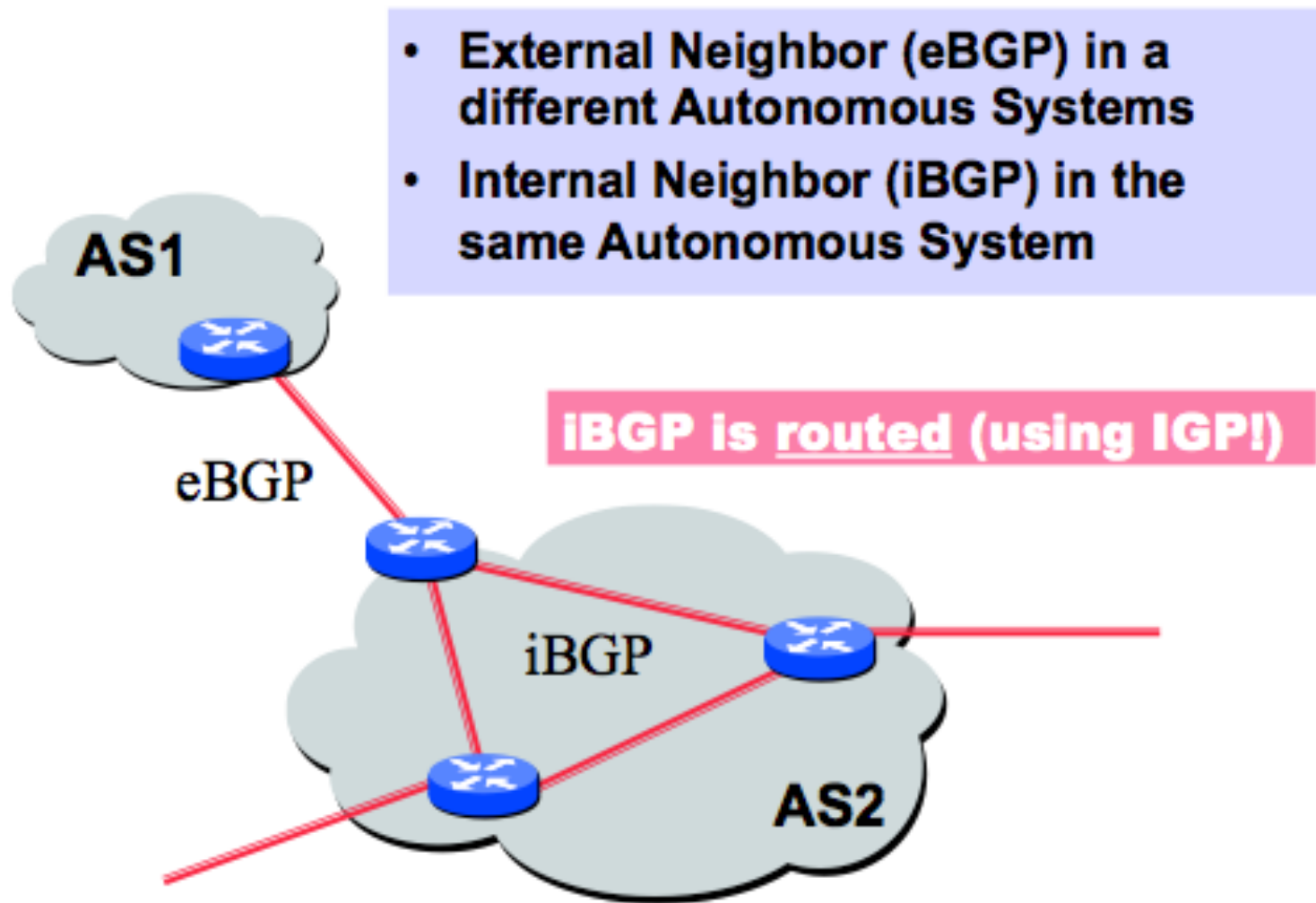
- ❖ Thuộc tính ORIGIN ra gốc của thông tin về một tuyến đường trong bảng BGP routing
- ❖ Thuộc tính ORIGIN có thể là một trong 3 giá trị sau:
  - ✓ IGP: thông tin học được từ bên trong AS gốc
  - ✓ EGP: Thông tin học được từ một giao thức EGP
  - ✓ Incomplete: thông tin học được theo cách khác

# Các thuộc tính đường đi

## ◆ Thuộc tính **ORIGIN**:

- ❖ Có hai kiểu kết nối BGP. Kiểu thứ nhất là kiểu kết nối giữa các router chạy BGP trong cùng 1 AS. Kiểu này được gọi là Internal BGP (IBGP).
- ❖ Kiểu kết nối thứ 2 là kiểu kết nối giữa các autonomous-system khác nhau. Kiểu này được gọi là External BGP (EBGP)
- ❖ Các tuyến được học từ EBGP hàng xóm có thuộc tính ORIGIN là EGP.
- ❖ Các tuyến redistributed trong BGP có thuộc tính ORIGIN là incomplete
- ❖

# Các thuộc tính đường đi



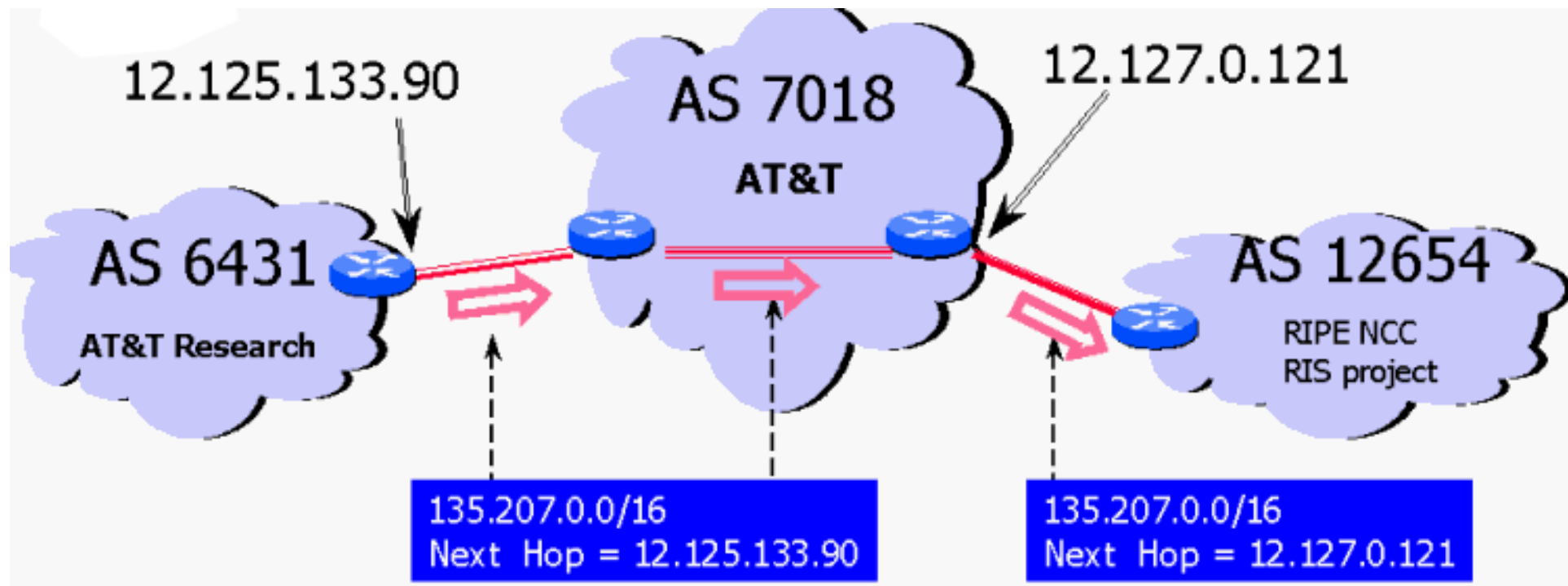
# Các thuộc tính đường đi

## ◆ Thuộc tính Next hop:

- ❖ Khi một router quảng bá một tuyến đến router EBGp hàng xóm, thuộc tính NEXT\_HOP được đặt là địa chỉ IP của giao diện kết nối đến router hàng xóm đó.
  - ❖ Lúc này giao diện router kết nối với hàng xóm sẽ là next hop của đường đi từ nút hàng xóm đến đích theo con đường được quản bá.
- ❖ Nếu router quảng bá một tuyến được học thông qua EBGp đến IBGP hàng xóm, thuộc tính NEXT\_HOP sẽ không được thay đổi.

# Các thuộc tính đường đi

## ◆ Thuộc tính Next hop:



# Các thuộc tính đường đi

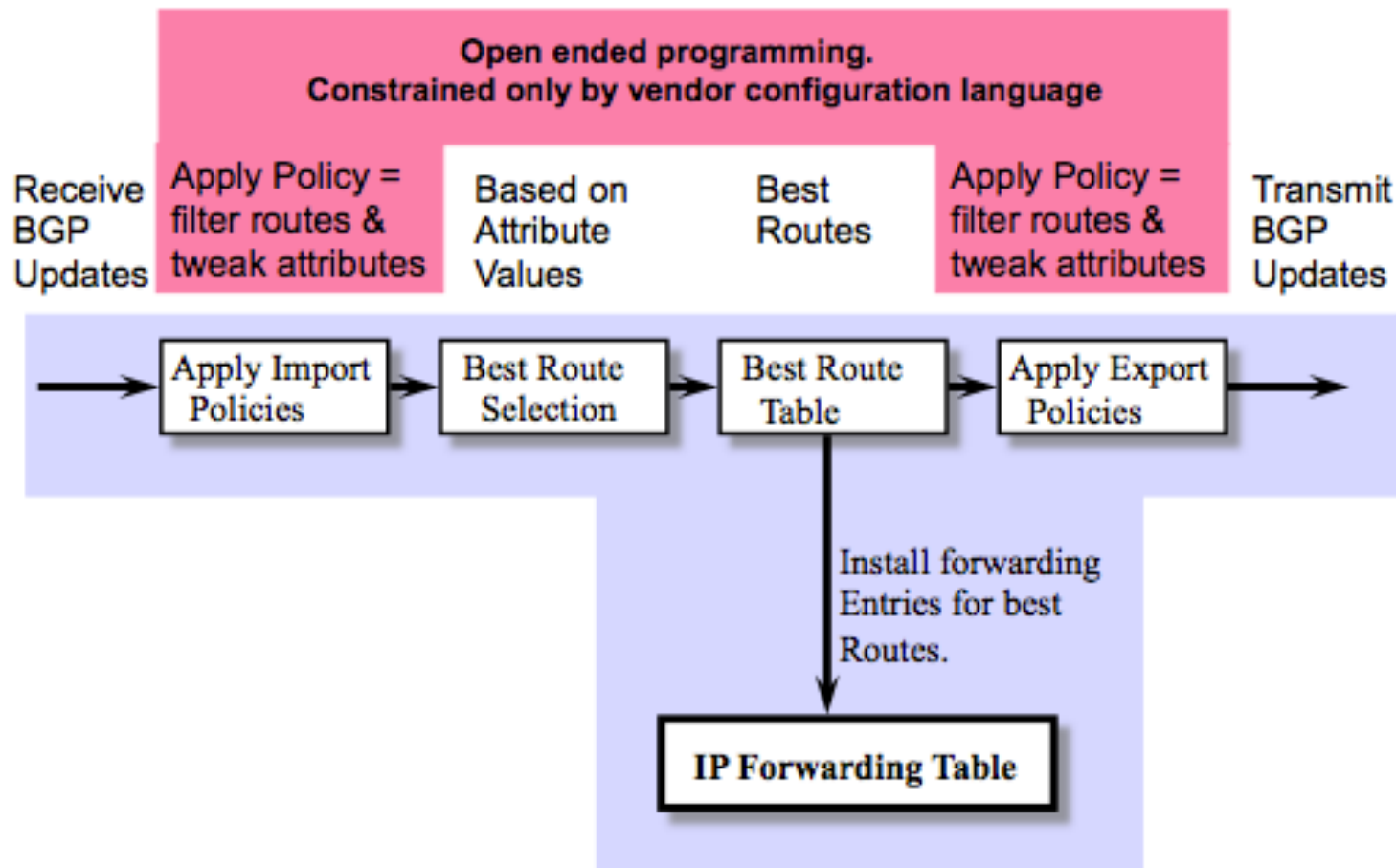
## ◆ Thuộc tính Community:

- ❖ Đây là một thuộc tính không bắt buộc.
- ❖ Thuộc tính này được dùng để cung cấp chính sách cho một nhóm các router đi qua các AS vì các router này có cùng một tính chất và có chung chính sách.
- ❖ Nếu một router nhận được một thông điệp thông tin cập nhật định tuyến với thuộc tính Community được thiết lập, nó sẽ xử lý thông điệp này một cách hợp lý. Nếu nó không hiểu thông điệp này thì sẽ gửi thuộc tính đó cho router hàng xóm xử lý.

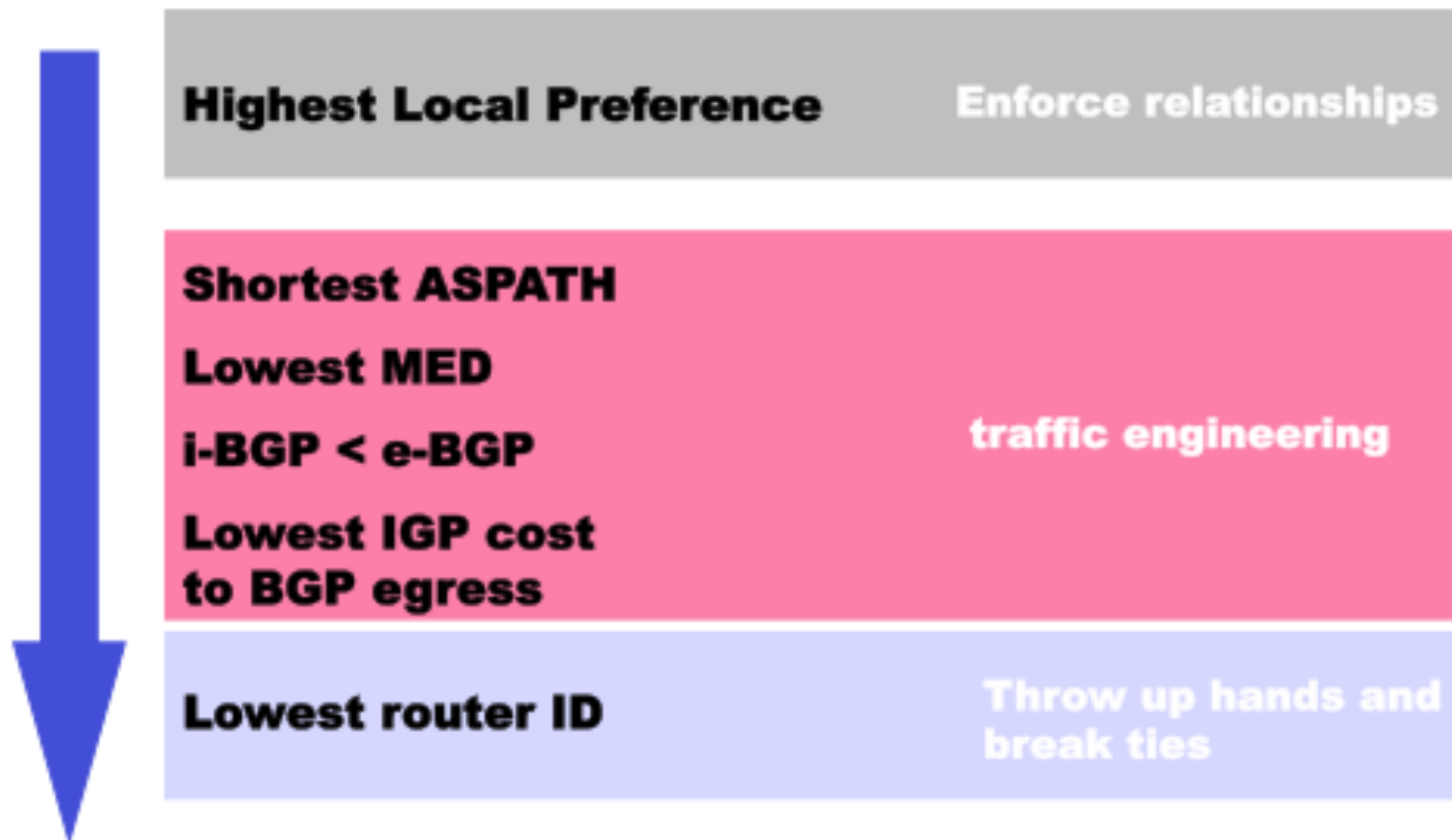


# Hoạt động của BGP

## ◆ Tóm tắt quy trình chọn đường trong BGP:



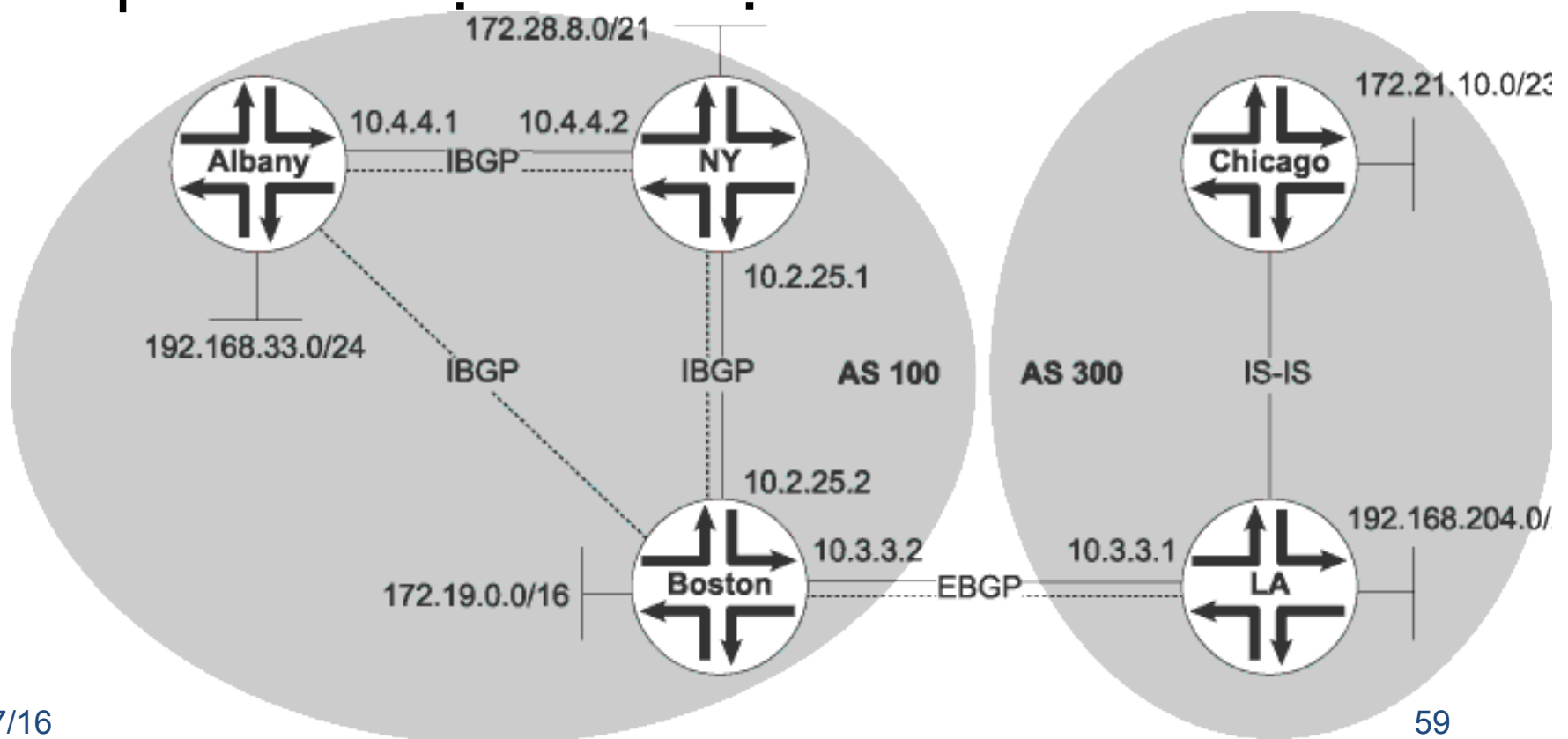
# Chọn đường đi tốt nhất



# Cấu hình BGP trên Cisco router

◆ Chi tiết các lệnh và các bước:

◆ Topo và ví dụ minh họa:



# Cấu hình BGP trên Cisco router

- Cấu hình trên router biên Boston:

```
host1(config)#ip route 172.31.125.100 255.255.255.252
host1(config)#router bgp 100
host1(config-router)#neighbor 10.2.25.1 remote-as 100
host1(config-router)#neighbor 10.4.4.1 remote-as 100
host1(config-router)#neighbor 10.3.3.1 remote-as 300
host1(config-router)#network 172.19.0.0
host1(config-router)#redistribute static
```

# Cấu hình BGP trên Cisco router

- Cấu hình trên router NY:

```
host2(config)#router bgp 100
```

```
host2(config-router)#neighbor 10.4.4.1 remote-as 100
```

```
host2(config-router)#neighbor 10.2.25.2 remote-as 100
```

```
host2(config-router)#network 172.28.8.0 mask 255.255.248.0
```

# Cấu hình BGP trên Cisco router

- Cấu hình trên router LA:

```
host4(config)#router bgp 300
host4(config-router)#neighbor 10.3.3.2 remote-as 100
host4(config-router)#network 192.168.204.0 mask
255.255.252.0
host4(config-router)#redistribute isis
```

# Cấu hình BGP trên Cisco router

- Cấu hình trên router Albany:

```
host3(config)#router bgp 100
host3(config-router)#neighbor 10.4.4.2 remote-as 100
host3(config-router)#neighbor 10.2.25.2 remote-as 100
host3(config-router)#network 192.168.33.0 mask
255.255.255.0
```