

Camada de Rede (Encaminhamento) (Protocolos RIP, OSPF e BGP)

Redes e Serviços

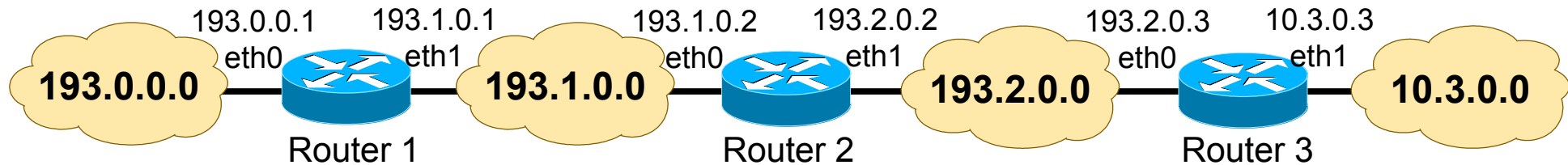
**Licenciatura em Engenharia Informática
DETI-UA**

Encaminhamento Estático vs. Dinâmico

- O encaminhamento estático é predefinido por configuração estática.
 - Define-se o *next-hop* para atingir uma determinada rede.
 - O *next-hop* deverá ser o endereço do próximo *router* do caminho e deverá pertencer a uma rede que o *router* já conhece.
 - É necessária a definição de uma rota estática para cada rede para a qual se pretende conectividade.
 - ➔ Pode-se usar a rede 0.0.0.0/0 que representa todas as redes.
 - ➔ Quando se define uma rota estática usado a rede 0.0.0.0/0 chama-se rota por omissão (*default route*).
- O encaminhamento dinâmico pressupõe o uso de um protocolo de comunicação entre os nós da rede de modo a determinar as redes existentes e os melhores caminhos (nexp-hops) para as atingir.
- O encaminhamento dinâmico é preferível no entanto em cenários simples o encaminhamento estático pode ser aceitável.



Rotas Estáticas



- Sem qualquer encaminhamento definido os *routers* apenas conhecem as redes a que estão diretamente ligados.
- Exemplo 1
 - ♦ Router2 não conhece as redes 193.0.0.0/24 e 10.3.0.0/24
 - ♦ Rotas estáticas necessárias:
 - 193.0.0.0/24 acessível via 193.1.0.1 (Router1)
 - 10.3.0.0/24 acessível via 193.2.0.3 (Router3)
- Exemplo 2
 - ♦ Router1 não conhece as redes 193.2.0.0/24 and 10.3.0.0/24
 - ♦ Rotas estáticas necessárias:
 - 193.2.0.0/24 acessível via 193.1.0.2 (Router2)
 - 10.3.0.0/24 acessível via 193.1.0.2 (Router2)
 - OR
 - Usando uma rota por omissão (*default route*): 0.0.0.0/0 acessível via 193.1.0.2 (Router2)



Encaminhamento Dinâmico

- O encaminhamento dinâmico permite à rede adaptar-se automaticamente a mudanças na topologia, sem envolvimento do administrador.
- Os routers trocam informação sobre as redes existentes e estado dos diferentes links/redes.
 - ♦ Os routers trocam informação apenas com routers a correr o mesmo protocolo de encaminhamento.
 - ♦ Quando a topologia da rede muda, a nova informação é dinamicamente propagada pela rede, e cada router atualiza a sua tabela de encaminhamento de modo a refletir as alterações.



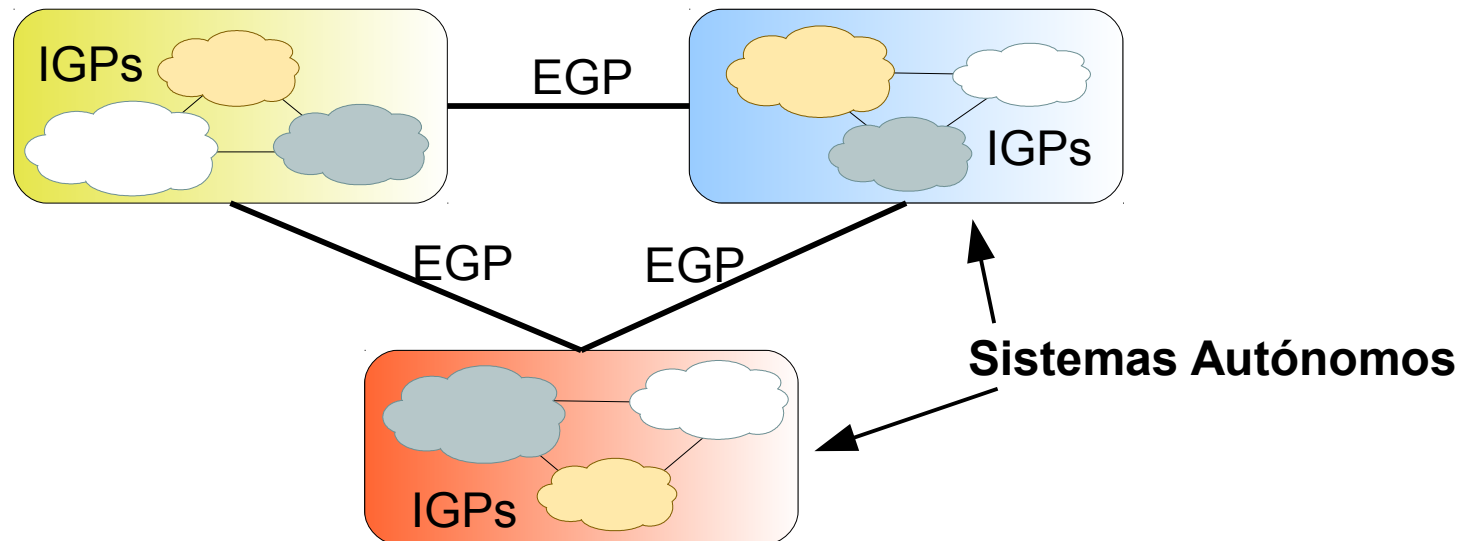
Distância Administrativa

- Os protocolos de encaminhamento possuem estruturas de métricas e algoritmos incompatíveis com outros protocolos.
- Numa rede com múltiplos protocolos de encaminhamento é crítico a existência de um mecanismo que permita selecionar o melhor caminho de rede de um conjunto fornecido por vários protocolos.
- Os routers possuem para cada caminho um valor chamado de distância administrativa para selecionar o melhor caminho para um destino (mesmo prefixo e máscara de rede) aprendido de vários protocolos.



Sistemas autónomos

- AS (Autonomous System) - conjunto de routers com uma política de encaminhamento própria, sob a responsabilidade de uma única administração
- Cada AS é identificado por um endereço único de 16 bits atribuído por um Internet Registry ou por um ISP
- Encaminhamento no interior dos ASs - realizado por IGPs (Interior Gateway Protocols) tais como o RIP, o OSPF, e o IS-IS
- Encaminhamento entre ASs - realizado por EGPs (Exterior Gateway Protocols) tais como o BGP
- Os IGPs e os EGPs têm diferentes objectivos
 - ♦ IGPs: otimizar o desempenho
 - ♦ EGPs: atender a questões de ordem política, económica e de segurança (para além do desempenho)



Distance Vector vs. Link State

- *Distance vector*

- ◆ Cada router conhece a informação que os routers vizinhos que enviam periodicamente
- ◆ Cada router determina os percursos de custo mínimo para todos os destinos baseando-se na versão distribuída e assíncrona do algoritmo de Bellman-Ford
- ◆ Exemplos: RIP, IGRP, EIGRP

- *Link state*

- ◆ Os routers conhecem a topologia completa da rede e usam um algoritmo centralizado para determinar os percursos de custo mínimo para todos os destinos
- ◆ A informação necessária para construir e manter em cada router a topologia da rede é obtida por um processo de inundação (flooding)
- ◆ Exemplos: OSPF, IS-IS



RIP (Routing Information Protocol)

- É um protocolo do tipo *distance vector*
 - ♦ cada router mantém um vetor distância constituído por uma lista das redes IP que conhece e, para cada rede, a sua estimativa atual do melhor custo
 - ♦ cada router envia periodicamente o seu vetor distância para os seus vizinhos
 - ♦ cada router usa a informação enviada pelos vizinhos para atualizar o seu vector distância
- O custo de um percurso de um router para uma rede é dado pelo número de routers intermédios
 - ♦ Considera percursos com um custo máximo 15 (16 = infinito)
- Cada router determina as entradas da sua tabela com base nos vetores distância recebidos dos vizinhos
 - ♦ Para cada rede existente, o router inclui na sua tabela de encaminhamento uma entrada para cada um dos vizinhos que lhe proporcionam um percurso de custo mínimo



RIP Version 1

- A versão 1 do RIP (RIPv1) é um protocolo do tipo *classful*.
 - ♦ Não anuncia as (sub-)máscaras das redes.
- RIPv1 usa o endereço 255.255.255.255 para enviar os seus anúncios.
 - ♦ Todos os equipamentos, incluindo PC e servidores, têm de processar os pacotes.
- Não suporta autenticação das mensagens

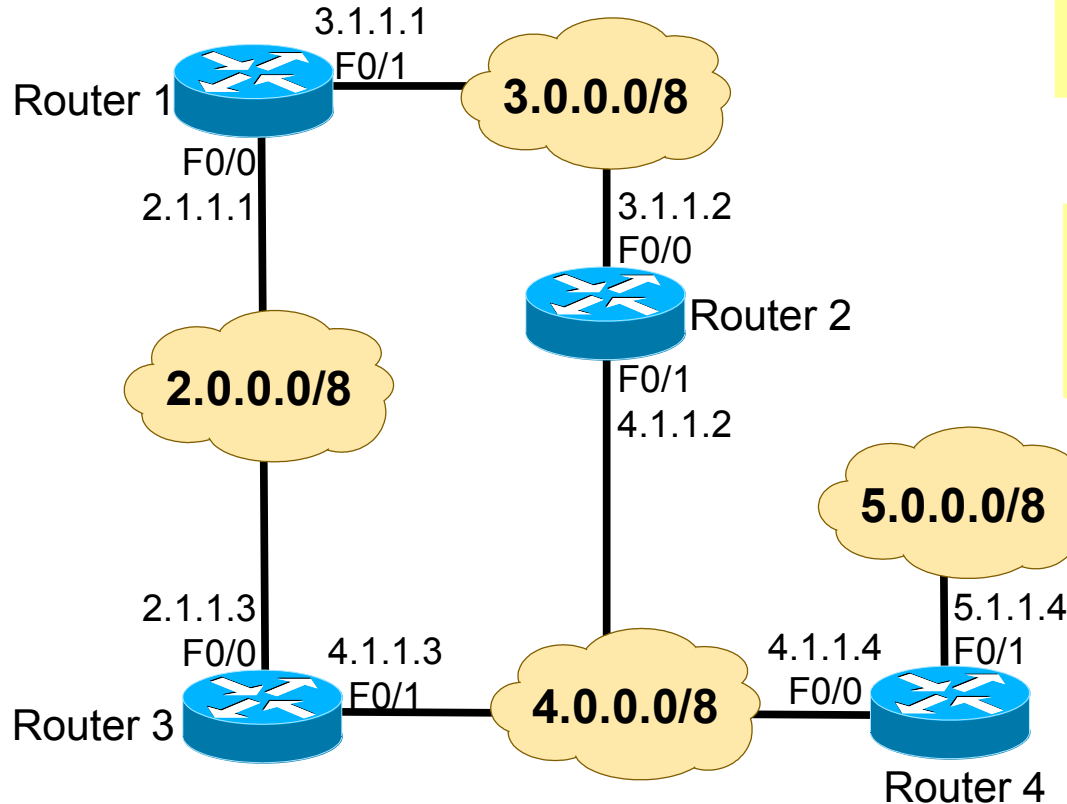


RIP Version 2

- A versão 2 do RIP (RIPv2) é um protocolo do tipo *classless*.
 - ♦ Os anúncios RIPv2 incluem o prefixo e (sub-)máscara da rede.
 - ♦ Suporta máscara de tamanho variável.
- RIPv2 usa o endereço multicast 224.0.0.9 para enviar os seus anúncios apenas para os routers a correr o protocolo RIPv2.
- RIPv2 suporta autenticação de mensagens usando *message-digest* ou autenticação em texto aberto.



Tabelas de Encaminhamento com RIPv1



```

C    2.0.0.0/8 is directly connected, FastEthernet0/0
C    3.0.0.0/8 is directly connected, FastEthernet0/1
R    4.0.0.0/8 [120/1] via 3.1.1.2, 00:00:10, FastEthernet0/1
      [120/1] via 2.1.1.3, 00:00:00, FastEthernet0/0
R    5.0.0.0/8 [120/2] via 3.1.1.2, 00:00:08, FastEthernet0/1
      [120/2] via 2.1.1.3, 00:00:00, FastEthernet0/0

```

Router 1

```

R    2.0.0.0/8 [120/1] via 4.1.1.3, 00:00:15, FastEthernet0/1
      [120/1] via 3.1.1.1, 00:00:07, FastEthernet0/0
C    3.0.0.0/8 is directly connected, FastEthernet0/0
C    4.0.0.0/8 is directly connected, FastEthernet0/1
R    5.0.0.0/8 [120/1] via 4.1.1.4, 00:00:22, FastEthernet0/1

```

Router 2

```

R    2.0.0.0/8 [120/1] via 4.1.1.3, 00:00:26, FastEthernet0/0
R    3.0.0.0/8 [120/1] via 4.1.1.2, 00:00:25, FastEthernet0/0
C    4.0.0.0/8 is directly connected, FastEthernet0/0
C    5.0.0.0/8 is directly connected, FastEthernet0/1

```

Router 4

```

C    2.0.0.0/8 is directly connected, FastEthernet0/0
R    3.0.0.0/8 [120/1] via 4.1.1.2, 00:00:09, FastEthernet0/1
      [120/1] via 2.1.1.1, 00:00:08, FastEthernet0/0
C    4.0.0.0/8 is directly connected, FastEthernet0/1
R    5.0.0.0/8 [120/1] via 4.1.1.4, 00:00:20, FastEthernet0/1

```

Router 3



RIP - Pacotes trocados entre routers

- RIP Request (Opcional)

- Enviado por um router que foi ligado recentemente ou quando a validade da informação relativamente a determinado destino expira (timeout = 180 seg.)
- Pode pedir informação relativamente a um destino específico ou a todos os destinos

- RIP Response

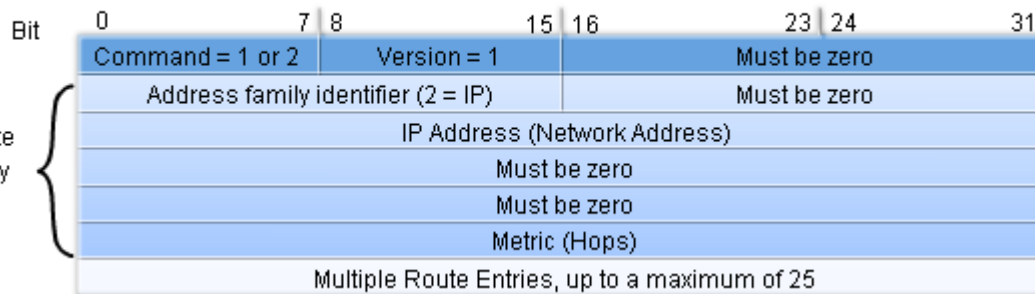
- Contém um vector distância
- É enviado:
 - ➔ 1º – Periodicamente (30 segundos no máximo)
 - ➔ 2º – Opcionalmente, quando a informação é alterada (triggered updates)
 - ➔ 3º – Em resposta a um RIP Request
- ➔ Nos dois primeiros casos
 - Na versão 1 do RIP, é enviado para o endereço de broadcast
 - Na versão 2 do RIP, é enviado para o endereço multicast 224.0.0.9 (Routers com RIP)
- ➔ No terceiro caso, é enviado apenas para quem enviou o RIP Request



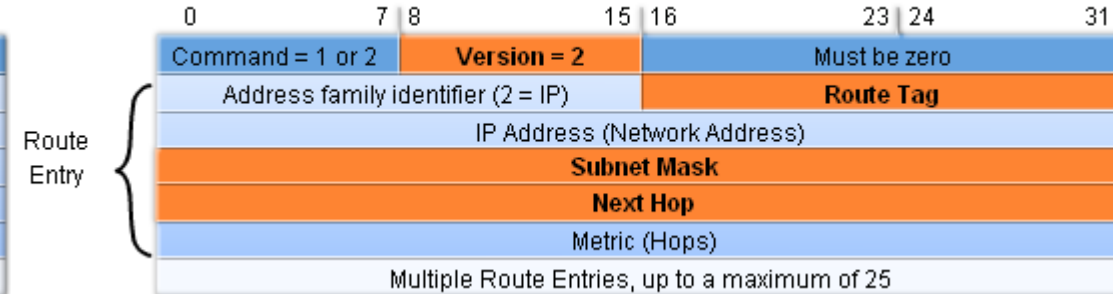
RIPv1 vs. RIPv2 Responses

- Novos campos RIPv2 nas mensagens Response:
 - Subnet mask
 - ➔ Suporta máscaras de tamanho variável.
 - ➔ Torna o RIPv2 um protocolo de encaminhamento *classless*.
 - Route tag
 - ➔ Atributo atribuído a uma rota que deve ser preservado e re-anunciado com a rota.
 - ➔ Fornece um método de separar rotas para redes dentro do domínio de encaminhamento RIP de rotas importadas do exterior.
 - Next hop
 - ➔ Endereço IP para onde os pacotes deverão ser enviados.
 - ➔ 0.0.0.0 neste campo indica que os pacotes deverão ser enviados para o router que enviou esta mensagem.

RIPv1



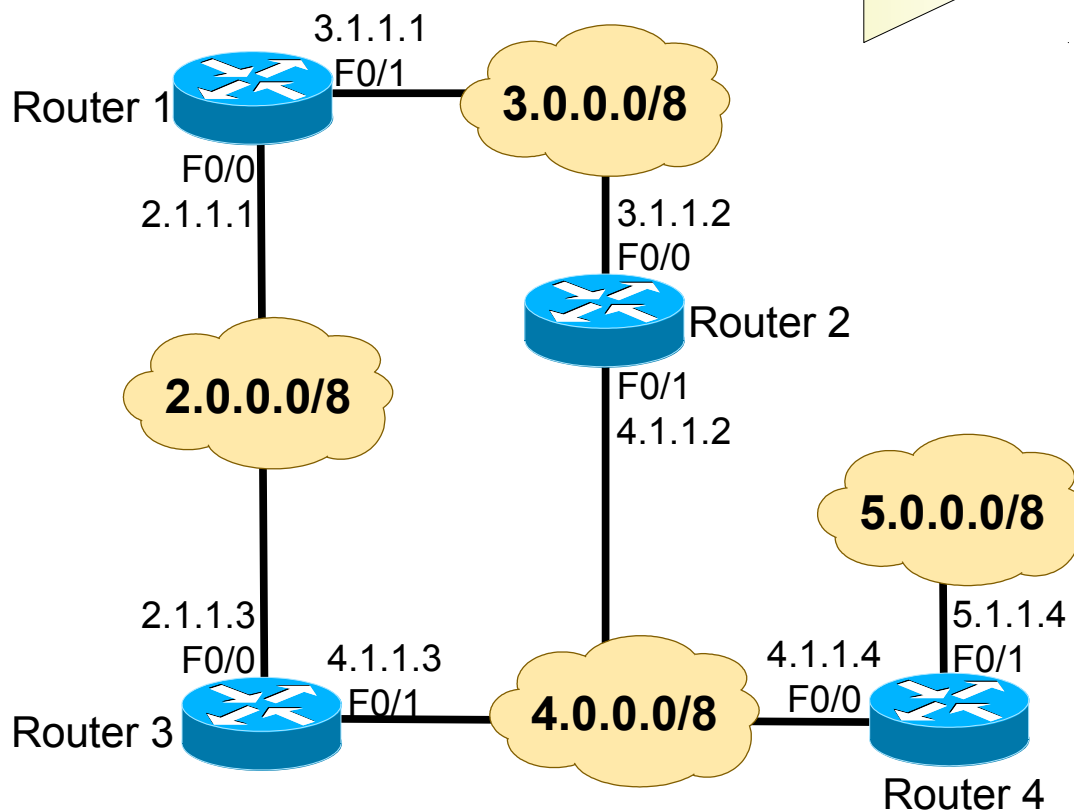
RIPv2



Distance Vector & *Split Horizon*

- Sem split horizon: cada router anuncia o vetor distância completo por todas as interfaces.
- Com split horizon: em cada interface, o router anuncia apenas as redes destino para as quais essa interface não é usada no encaminhamento dos pacotes de dados.
- O split horizon diminui o tempo de convergência das tabelas de encaminhamento quando há alterações de topologia.
- Existe no RIPv1 e RIPv2.

Pacotes RIPv1 enviados pelo Router 1 para a rede 3.0.0.0



Sem split horizon (RIPv1)

```
▸ Internet Protocol, Src: 3.1.1.1 (3.1.1.1), Dst: 255.255.255.255 (255.255.255.255)
▸ User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
▾ Routing Information Protocol
  Command: Response (2)
  Version: RIPv1 (1)
  ▾ IP Address: 2.0.0.0, Metric: 1
    Address Family: IP (2)
    IP Address: 2.0.0.0 (2.0.0.0)
    Metric: 1
  ▾ IP Address: 3.0.0.0, Metric: 1
    Address Family: IP (2)
    IP Address: 3.0.0.0 (3.0.0.0)
    Metric: 1
  ▾ IP Address: 4.0.0.0, Metric: 2
    Address Family: IP (2)
    IP Address: 4.0.0.0 (4.0.0.0)
    Metric: 2
  ▾ IP Address: 5.0.0.0, Metric: 3
    Address Family: IP (2)
    IP Address: 5.0.0.0 (5.0.0.0)
    Metric: 3
```

Com split horizon (RIPv1)

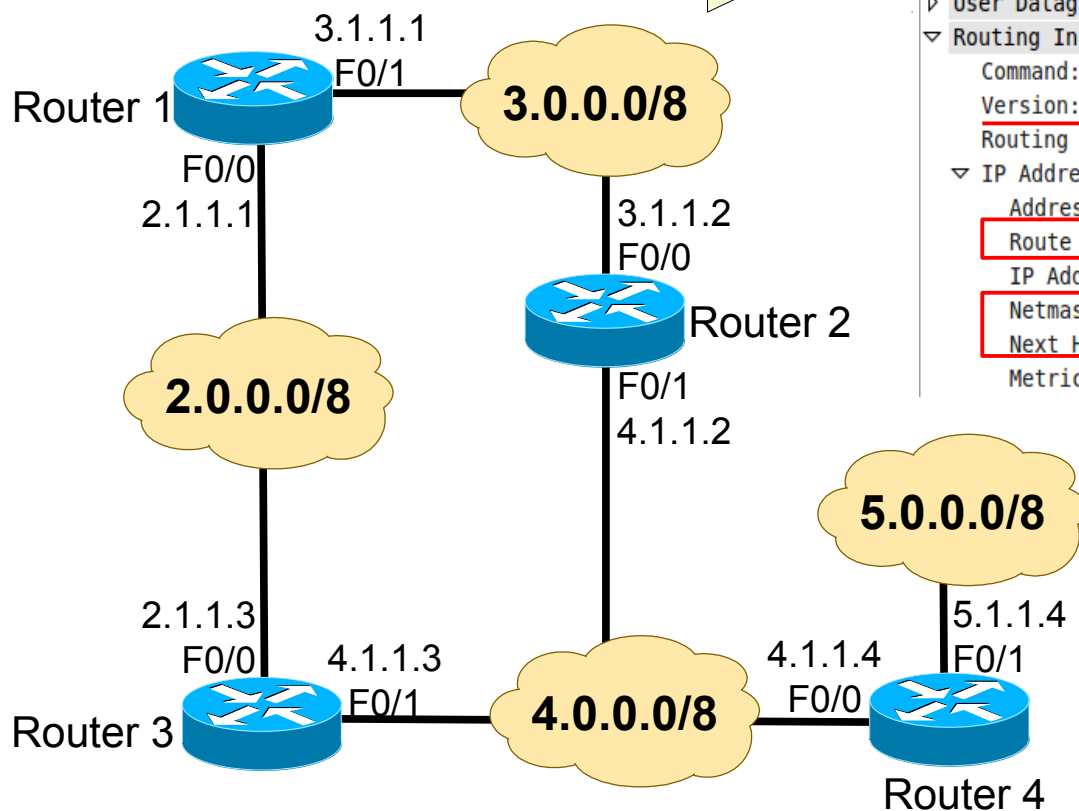
```
▸ Internet Protocol, Src: 3.1.1.1 (3.1.1.1), Dst: 255.255.255.255 (255.255.255.255)
▸ User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
▾ Routing Information Protocol
  Command: Response (2)
  Version: RIPv1 (1)
  ▾ IP Address: 2.0.0.0, Metric: 1
    Address Family: IP (2)
    IP Address: 2.0.0.0 (2.0.0.0)
    Metric: 1
```


RIPv2 Response

Com *split horizon*

1 0.000000	3.1.1.1	224.0.0.9	RIPv2	Response
3 7.936722	3.1.1.2	224.0.0.9	RIPv2	Response
6 27.017157	3.1.1.1	224.0.0.9	RIPv2	Response
8 36.912644	3.1.1.2	224.0.0.9	RIPv2	Response
13 56.403552	3.1.1.1	224.0.0.9	RIPv2	Response
14 64.328612	3.1.1.2	224.0.0.9	RIPv2	Response
17 85.258230	3.1.1.1	224.0.0.9	RIPv2	Response
19 92.773090	3.1.1.2	224.0.0.9	RIPv2	Response

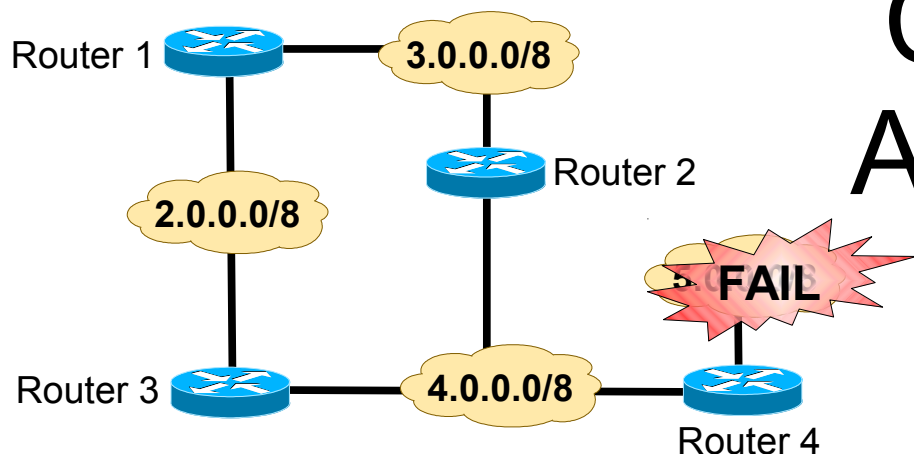
Pacotes RIPv2 enviados pelo
Router 1 para a rede 3.0.0.0



```
Frame 1: 66 bytes on wire (528 bits), 66 bytes captured (528 bits) on interface 0
Ethernet II, Src: c2:00:10:2a:00:01 (c2:00:10:2a:00:01), Dst: IPv4mcast_00:00:09 (01:00:5e:00:00:09)
Internet Protocol, Src: 3.1.1.1 (3.1.1.1), Dst: 224.0.0.9 (224.0.0.9)
User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
Routing Information Protocol
  Command: Response (2)
  Version: RIPv2 (2)
  Routing Domain: 0
  IP Address: 2.0.0.0, Metric: 1
    Address Family: IP (2)
    Route Tag: 0
    IP Address: 2.0.0.0 (2.0.0.0)
    Netmask: 255.0.0.0 (255.0.0.0)
    Next Hop: 0.0.0.0 (0.0.0.0)
    Metric: 1
```



Comportamento do RIP Após a Falha de um Link



1 0.000000	4.1.1.3	255.255.255.255	RIPv1	Response
6 8.912414	4.1.1.2	255.255.255.255	RIPv1	Response
7 9.153108	4.1.1.4	255.255.255.255	RIPv1	Response
10 29.471758	4.1.1.3	255.255.255.255	RIPv1	Response
12 37.902389	4.1.1.2	255.255.255.255	RIPv1	Response
13 38.501693	4.1.1.4	255.255.255.255	RIPv1	Response
17 58.210507	4.1.1.4	255.255.255.255	RIPv1	Response
18 58.264449	4.1.1.3	255.255.255.255	RIPv1	Response
19 60.217652	4.1.1.3	255.255.255.255	RIPv1	Response
20 60.220680	4.1.1.2	255.255.255.255	RIPv1	Response
24 66.876540	4.1.1.4	255.255.255.255	RIPv1	Response
25 66.812377	4.1.1.3	255.255.255.255	RIPv1	Response

▷ Frame 17: 66 bytes on wire (528 bits), 66 bytes captured (528 bits)
▷ Ethernet II, Src: c2:03:10:2a:00:00 (c2:03:10:2a:00:00), Dst: Broadcast (ff:ff:ff:ff:ff:ff)
▷ Internet Protocol, Src: 4.1.1.4 (4.1.1.4), Dst: 255.255.255.255 (255.255.255.255)
▷ User Datagram Protocol, Src Port: router (520), Dst Port: router (520)
▽ Routing Information Protocol
Command: Response (2)
Version: RIPv1 (1)
▽ IP Address: 5.0.0.0, Metric: 16
Address Family: IP (2)
IP Address: 5.0.0.0 (5.0.0.0)
Metric: 16

O comportamento do RIPv1 e RIPv2
é idêntico!



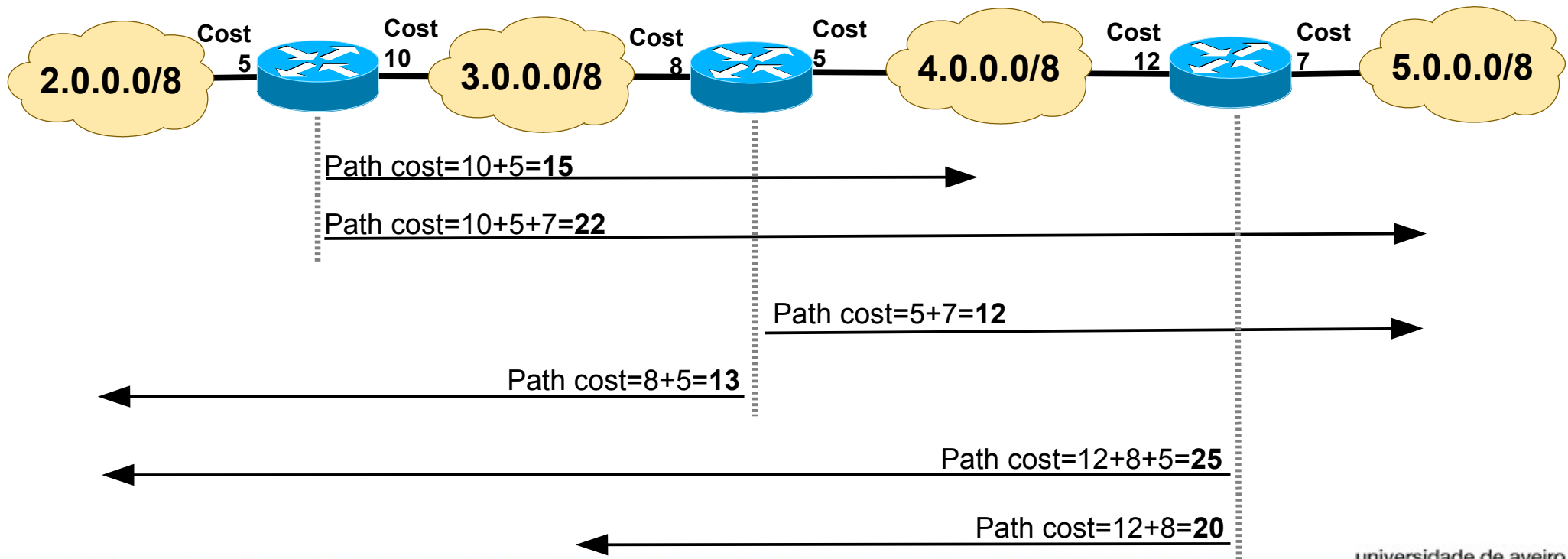
OSPF (Open Shortest Path First)

- É um protocolo do tipo link state
- Cada router contém uma base de dados com a topologia da rede
- A informação necessária à construção das bases de dados é trocada através de flooding
- Usa o algoritmo de Dijkstra para calcular os percursos de custo mínimo
 - ♦ É associado um custo a cada interface de cada router.
 - ♦ O custo de um percurso de um router para uma rede é dado pela soma dos custos das interfaces de saída no sentido do router para a rede.
- A informação topológica é enviada em Link State Advertisements (LSA), contidos nos pacotes Link State Update.

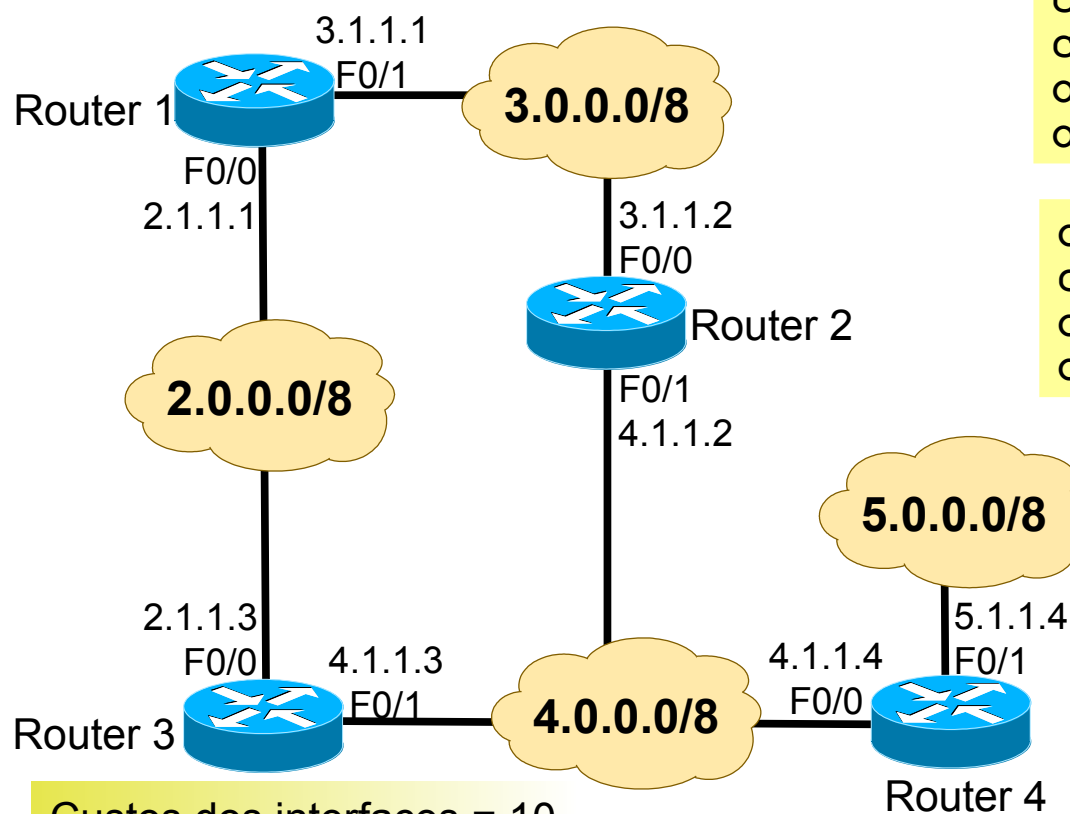


OSPF Path Cost (Custo de um Caminho)

- Cada interface de um router tem associado um custo OSPF.
- O custo de um percurso de um router para uma rede é dado pela soma dos custos das interfaces de saída no sentido do router para a rede.
 - Os routers para aceder a redes diretamente ligadas nunca usam caminhos OSPF.



Exemplo - percursos mínimos no OSPF



```
C 2.0.0.0/8 is directly connected, F0/0
C 3.0.0.0/8 is directly connected, F0/1
O 4.0.0.0/8 [110/20] via 2.1.1.3, 00:01:13, F0/0
O 5.0.0.0/8 [110/30] via 2.1.1.3, 00:01:10, F0/0
```

```
O 2.0.0.0/8 [110/20] via 3.1.1.1, 00:01:13, F0/0
C 3.0.0.0/8 is directly connected, F0/0
O 4.0.0.0/8 [110/30] via 3.1.1.1, 00:01:13, F0/0
O 5.0.0.0/8 [110/40] via 3.1.1.1, 00:01:10, F0/0
```

Router 1 e Router 2 depois de desligar a F0/1 do Router 2

```
C 2.0.0.0/8 is directly connected, F0/0
C 3.0.0.0/8 is directly connected, F0/1
O 4.0.0.0/8 [110/15] via 3.1.1.2, 00:01:13, F0/1
O 5.0.0.0/8 [110/25] via 3.1.1.2, 00:01:10, F0/1
```

Router 1, com o custo da F0/1 do Router 2 colocado a 5

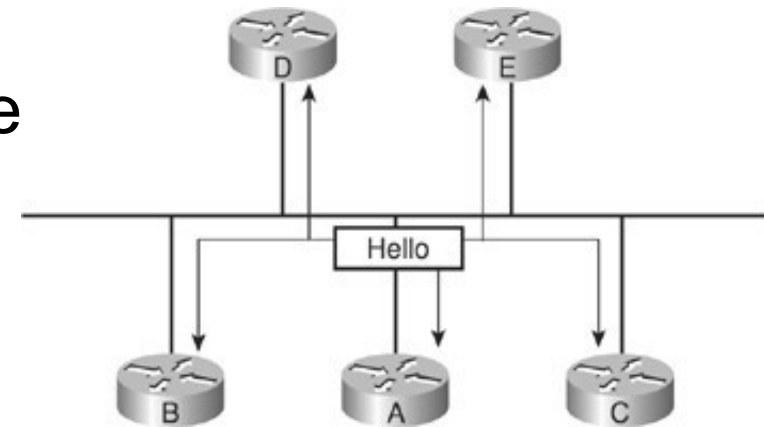
```
C 2.0.0.0/8 is directly connected, F0/0
C 3.0.0.0/8 is directly connected, F0/1
O 4.0.0.0/8 [110/20] via 3.1.1.2, 00:01:13, F0/1
[110/20] via 2.1.1.3, 00:01:31, F0/0
O 5.0.0.0/8 [110/30] via 3.1.1.2, 00:01:10, F0/1
[110/30] via 2.1.1.3, 00:01:10, F0/0
```

Tabela inicial do Router 1



Adjacências OSPF

- Um router a correr OSPF tem primeiro de estabelecer adjacências com os seus vizinhos.
 - Trocado mensagens OSPF Hello.
- Dois routers OSPF numa ligação ponto-a-ponto série formam adjacências entre si.
- No entanto, routers OSPF em redes broadcast (como LAN Ethernet), elegem um router com Designated Router (DR) e outro como Backup Designated Router (BDR).
 - Todos os outros routers formam adjacências com estes dois routers.



Eleição do DR e do BDR

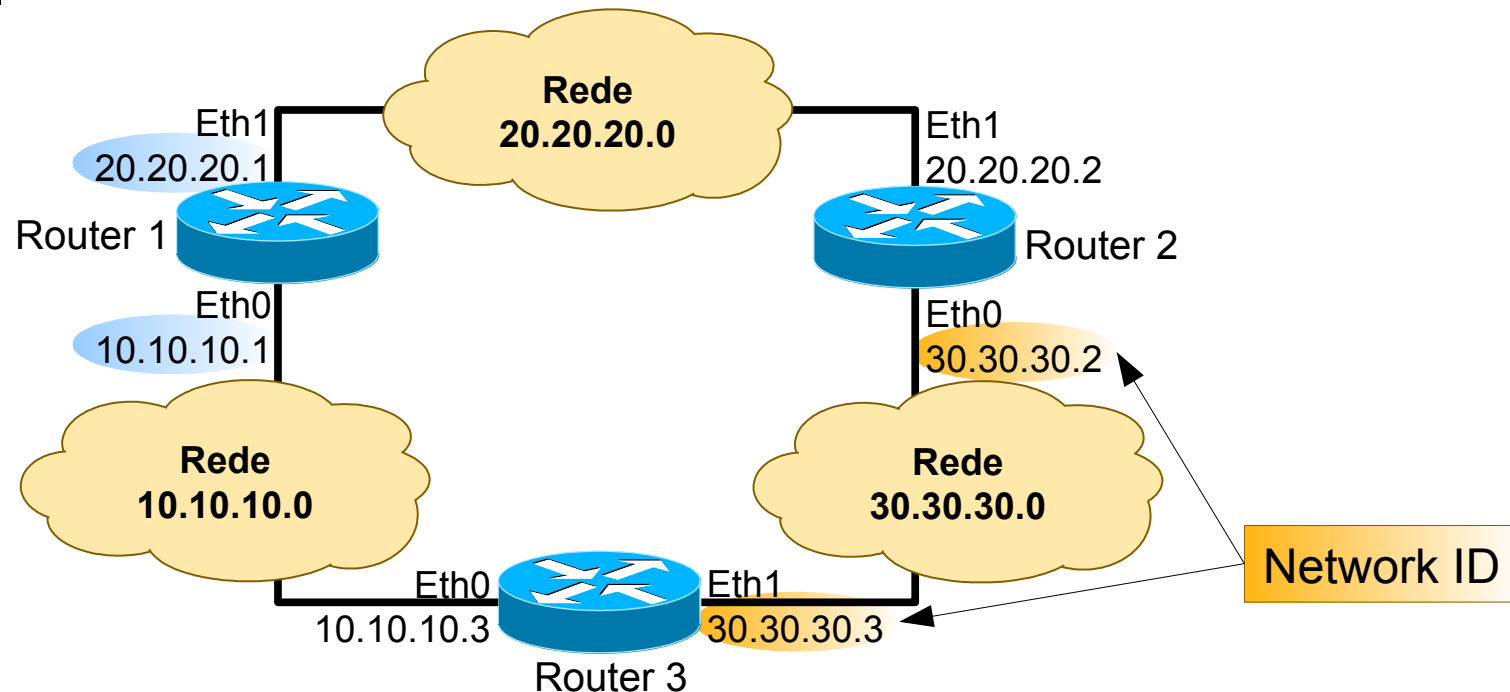
- O primeiro router a ser ligado torna-se o DR e o segundo o BDR
- Caso múltiplos routers arranquem em simultâneo:
 - O DR será o router que, de entre os ligados à LAN, tiver a maior prioridade.
 - ➔ A prioridade OSPF é um valor definido administrativamente
 - Torna possível que seja o gestor da rede a determinar qual o DR e o BDR.
 - Em caso de empate será escolhido o router com maior RID
 - ➔ Router ID - identifica um router e corresponde (em alternativa)
 - Um valor definido administrativamente.
 - Ao endereço IPv4 mais elevado de todas as interfaces do router no momento da ativação do protocolo OSPF.
- Depois de um router ser eleito DR, nenhum outro router pode usurpar essa posição.
- Se o DR avariar, o BDR passará a ser o novo DR e será eleito um novo BDR.



Bases de dados OSPF:

Identificação dos Routers e das Redes

- OSPF Router ID: endereço IP de uma das suas interfaces (o maior no instante de ativação do protocolo) ou um valor definido administrativamente.
- OSPF Network/Link ID: endereço IP da interface do seu Designated Router (DR).



Bases de Dados OSPF

- As bases de dados estão organizadas em duas tabelas. Uma com informação relativa a todos os routers (Router Link States) e outra com informação relativa a todas as redes intermédias (Net Link States);
- As redes são identificadas pelo Network ID, os routers são identificados pelo Router ID;

OSPF Router with ID (20.20.20.1) (Process ID 1)

Router Link States (Area 0)

Link ID	ADV Router	Age	Seq#	Checksum	Link count
20.20.20.1	20.20.20.1	40	0x8000000A	0x00E7FB	2
30.30.30.2	30.30.30.2	69	0x80000006	0x002906	2
30.30.30.3	30.30.30.3	41	0x80000007	0x00283D	2

Net Link States (Area 0)

Link ID	ADV Router	Age	Seq#	Checksum
10.10.10.3	30.30.30.3	41	0x80000001	0x00051C
20.20.20.2	30.30.30.2	70	0x80000001	0x00A164
30.30.30.3	30.30.30.3	154	0x80000001	0x00A91C



Bases de Dados OSPF (exemplo)

- Relativamente a cada router é armazenada informação relativa às várias redes que lhe estão directamente ligadas

```
LS age: 321
Options: (No TOS-capability, DC)
LS Type: Router Links
Link State ID: 20.20.20.1 ← Identifica o router
Advertising Router: 20.20.20.1
LS Seq Number: 8000000A
Checksum: 0xE7FB
Length: 48
Number of Links: 2 ← Número de interfaces

  Link connected to: a Transit Network ← Rede intermédia
    (Link ID) Designated Router address: 20.20.20.2 ← Network ID
    (Link Data) Router Interface address: 20.20.20.1 ← Endereço da Interface
    Number of TOS metrics: 0
    TOS 0 Metrics: 1 ← Custo da Interface

  Link connected to: a Transit Network
    (Link ID) Designated Router address: 10.10.10.3
    (Link Data) Router Interface address: 10.10.10.1
    Number of TOS metrics: 0
    TOS 0 Metrics: 1
```



Bases de Dados OSPF (exemplo)

- Relativamente a cada rede intermédia é armazenada informação relativa aos vários routers que lhe estão directamente ligados

Routing Bit Set on this LSA

LS age: 483

Options: (No TOS-capability, DC)

LS Type: Network Links

Link State ID: 10.10.10.3 (address of Designated Router) ← Network ID

Advertising Router: 30.30.30.3

LS Seq Number: 80000001

Checksum: 0x51C

Length: 32

Network Mask: /24

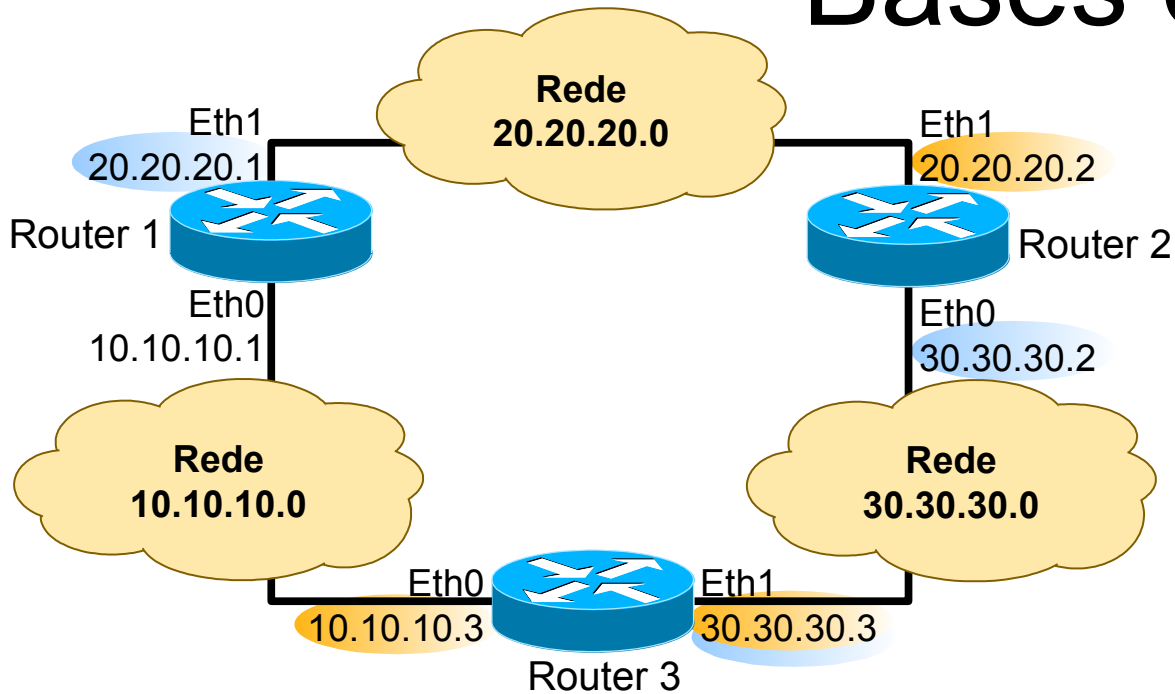
Attached Router: 30.30.30.3 }

Attached Router: 20.20.20.1 }

← Routers directamente ligados (ID)



Bases de dados OSPF



Routing Bit Set on this LSA

LS age: 208

Options: (No TOS-capability, DC)

LS Type: Network Links

Link State ID: 20.20.20.2 (address of Designated Router)

Advertising Router: 30.30.30.2

LS Seq Number: 80000001

Checksum: 0xA164

Length: 32

Network Mask: /24

Attached Router: 30.30.30.2

Attached Router: 20.20.20.1

Network Link State relativo à rede 20.20.20.0

LS age: 321

Options: (No TOS-capability, DC)

LS Type: Router Links

Link State ID: 20.20.20.1

Advertising Router: 20.20.20.1

LS Seq Number: 8000000A

Checksum: 0xE7FB

Length: 48

Number of Links: 2

Link connected to: a Transit Network

(Link ID) Designated Router address: 20.20.20.2

(Link Data) Router Interface address: 20.20.20.1

Number of TOS metrics: 0

TOS 0 Metrics: 1

Link connected to: a Transit Network

(Link ID) Designated Router address: 10.10.10.3

(Link Data) Router Interface address: 10.10.10.1

Number of TOS metrics: 0

TOS 0 Metrics: 1

Router Link State relativo ao Router 1



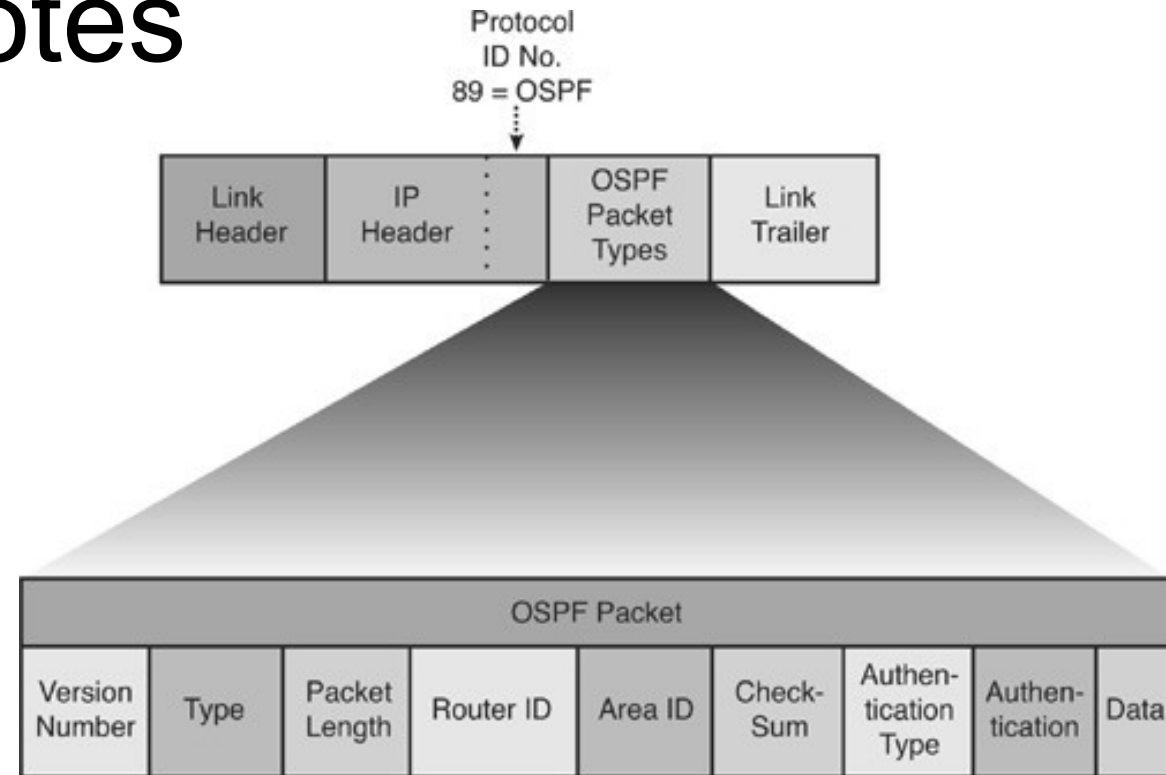
Pacotes OSPF

- Hello – Para descoberta de vizinhos, construção de adjacências com eles e eleição do DR/BDR.
- Database Description (DBD) - Usado para a verificação do conteúdo e sincronização das bases de dados.
- Link-State Request (LSR) – Usado para pedir um Link-State Advertisements (LSA) específico a outro router.
- Link-State Update (LSU) – Usando para enviar os Link-State Advertisements (LSA).
- LSAck – Confirma a correta receção dos pacotes.



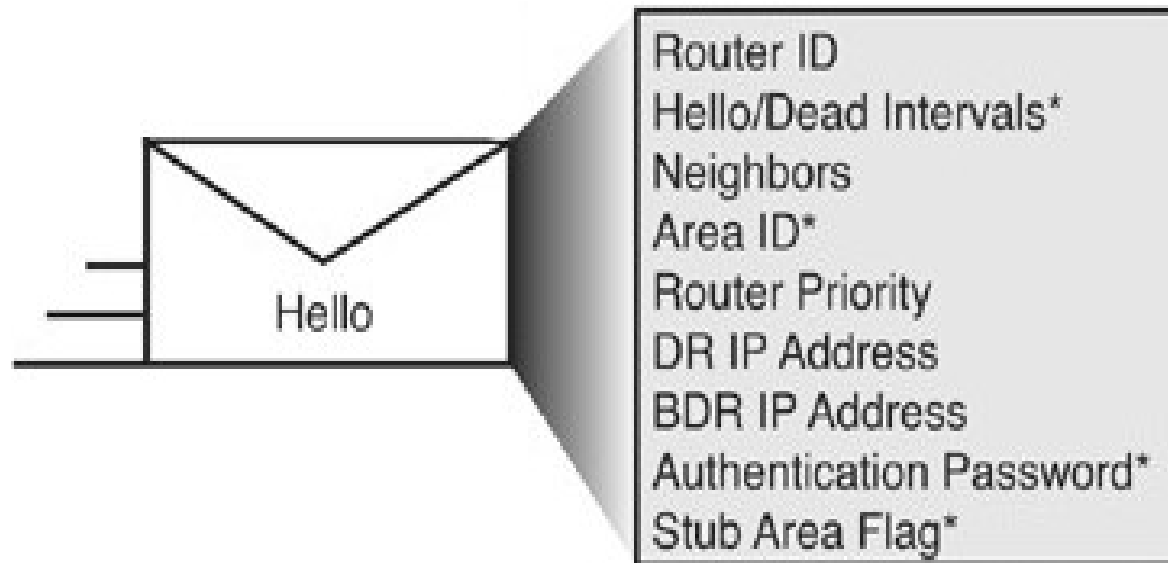
Formato dos Pacotes OSPF

- Version Number
 - ◆ 2 para OSPF Version 2 (OSPF para IPv4).
 - ◆ 3 para OSPF Version 3 (OSPF para IPv6).
- Type
 - ◆ Tipo do pacote OSPF.
- Packet Length
- Router ID
 - ◆ Identifica o router emissor.
- Area ID
 - ◆ Identifica a área OSPF onde o pacote teve origem.
- Checksum
 - ◆ Usado para verificar a integridade do pacote.
- Authentication Type
 - ◆ Define o tipo de autenticação.
- Authentication
 - ◆ Dados de autenticação.
- Data
 - ◆ Contem os dados concretos
 - ◆ Hello – Lista de vizinhos.
 - ◆ DBD – Sumário da base de dados.
 - ◆ LSR – Tipo de LSA pretendido e router ID do destinatário do pedido.
 - ◆ LSU – LSA completos. Pode conter múltiplos LSA.
 - ◆ LSAck – Sem dados.



Pacotes Hello

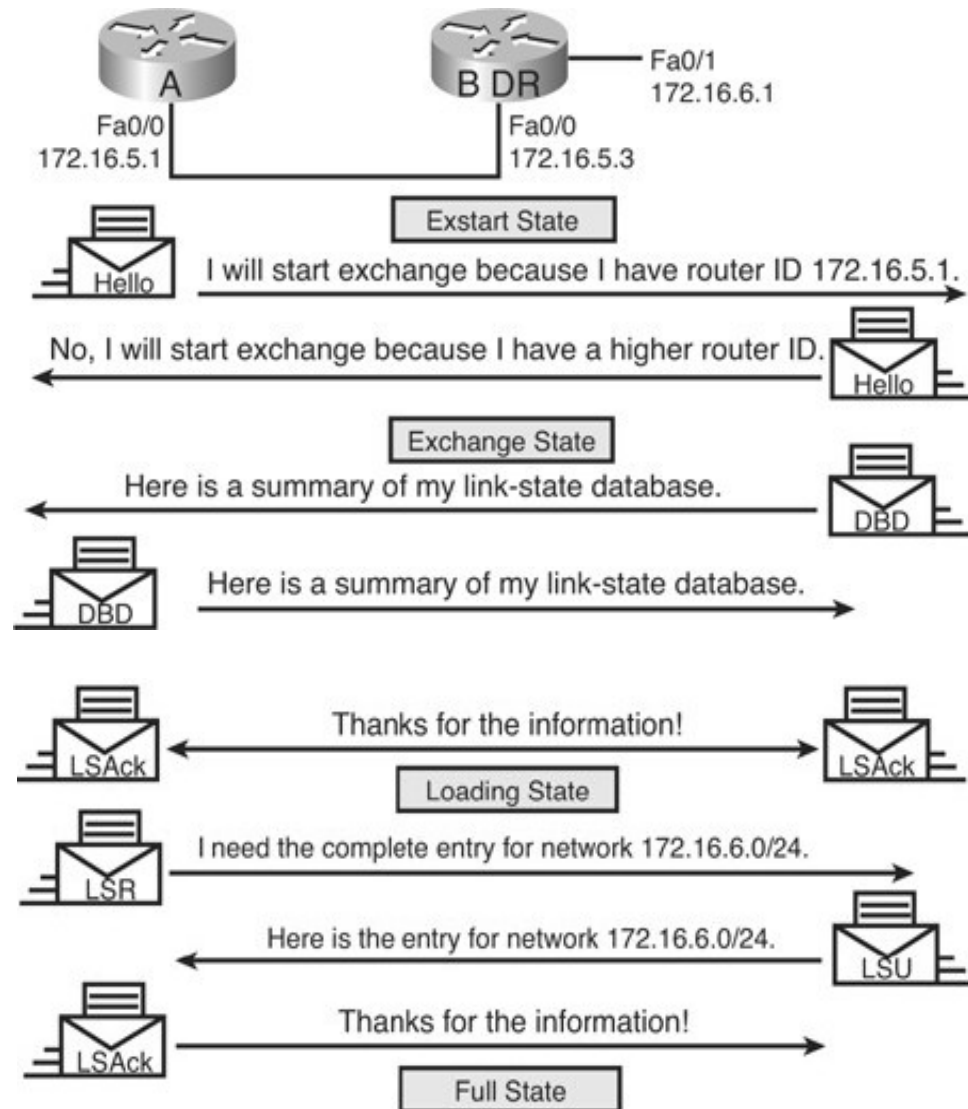
- Estes pacotes são utilizados:
 - ♦ Para descobrir quem são os vizinhos em cada interface
 - ♦ Para eleger e descobrir o DR e o BDR de cada rede
- Por omissão, são enviados de 10 em 10 segundos
 - ♦ Objectivo: detectar falha de conectividade
- Não transportam informação de encaminhamento
 - ♦ A sua recepção não altera as bases de dados pelo que não requer processamento das tabelas de encaminhamento



*Entry Must Match on Neighboring Routers



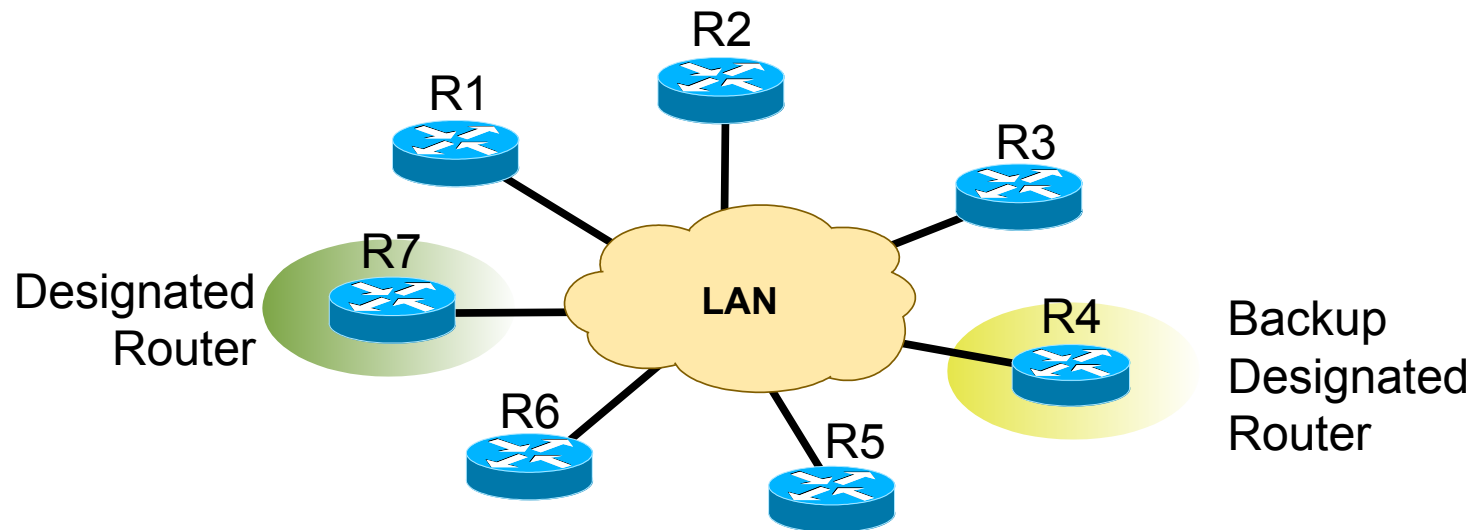
Troca de Informação de Topologia



- Uma relação mestre/escravo é criada entre cada router e o DR (e BDR) adjacente.
 - Apenas o DR troca e sincroniza informação de link-states com os routers.
- Os routers trocam um ou mais pacotes DBD.
 - Os DBD incluem um sumário da base de dados OSPF.
 - Os routers usam um número de sequência para determinar qual a informação de link-state mais recente.
- Usando LSAck os routers confirmam a receção dos DBD.
- Se os DBD indicam que há informação mais recente, o router envia um LSR a pedir essa nova informação.
- O outro router responde com um LSU que contem toda a informação (LSA) pedida.
- Usando LSAck os routers confirmam a receção da informação.



Propagação de LSA em LANs



Objectivo: Assegurar que o envio de um LSA numa LAN é efectivamente recebido por todos os routers vizinhos; assegurar que as bases de dados de todos os routers estão sincronizadas

Método:

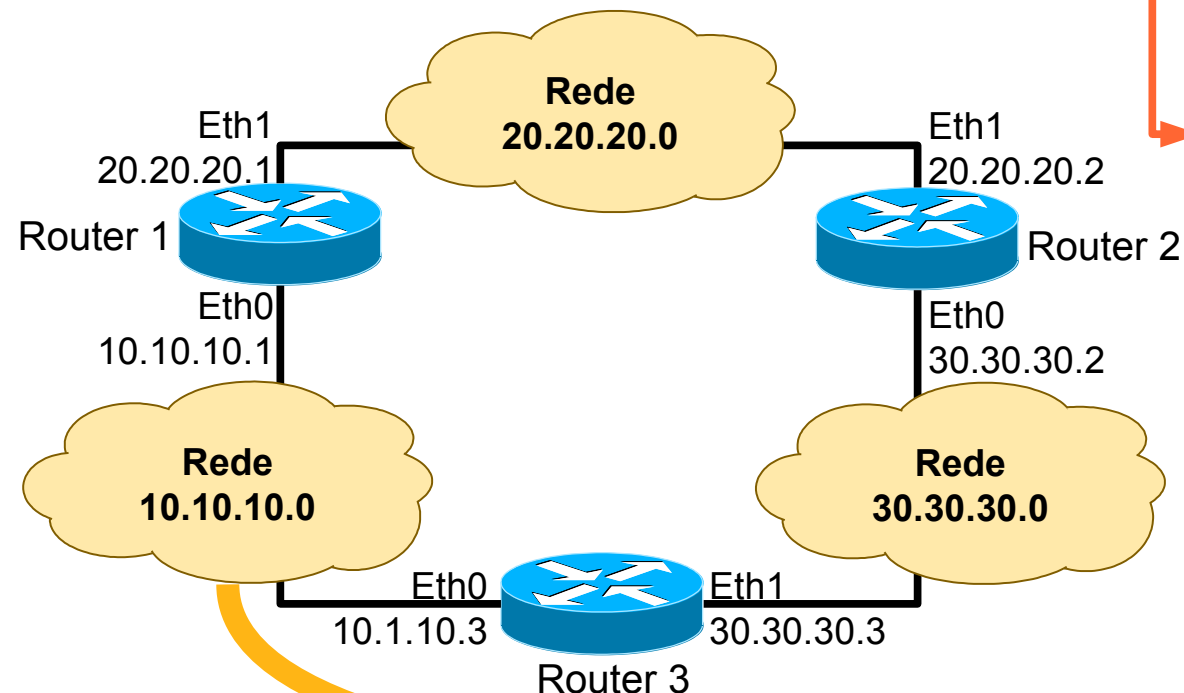
- R3 tem um LSA para ser enviado para a LAN
- R3 envia o LSA usando o endereço de multicast allRouters (224.0.0.6)
- DR e BDR recebem o LSA
- DR envia o LSA usando o endereço de multicast allOSPFrouters (224.0.0.5)
- Se todos os routers recebem o LSA correctamente, BDR, R1, R2, ..., R6 enviam um LSACK usando o endereço multicast allRouters
- Se o DR não receber ACKs de R2, R5 e BDR dentro de um timeout, o DR retransmite o LSA 3 vezes, para R2, R5 e BDR

Exemplo pacotes OSPF

OSPF activado no Router 1

OSPF activado no Router 3

OSPF activado no Router 2



Time	Source	Destination	Protocol	Info
0.000000	10.10.10.1	224.0.0.5	OSPF	Hello Packet
10.002318	10.10.10.1	224.0.0.5	OSPF	Hello Packet
20.003116	10.10.10.1	224.0.0.5	OSPF	Hello Packet

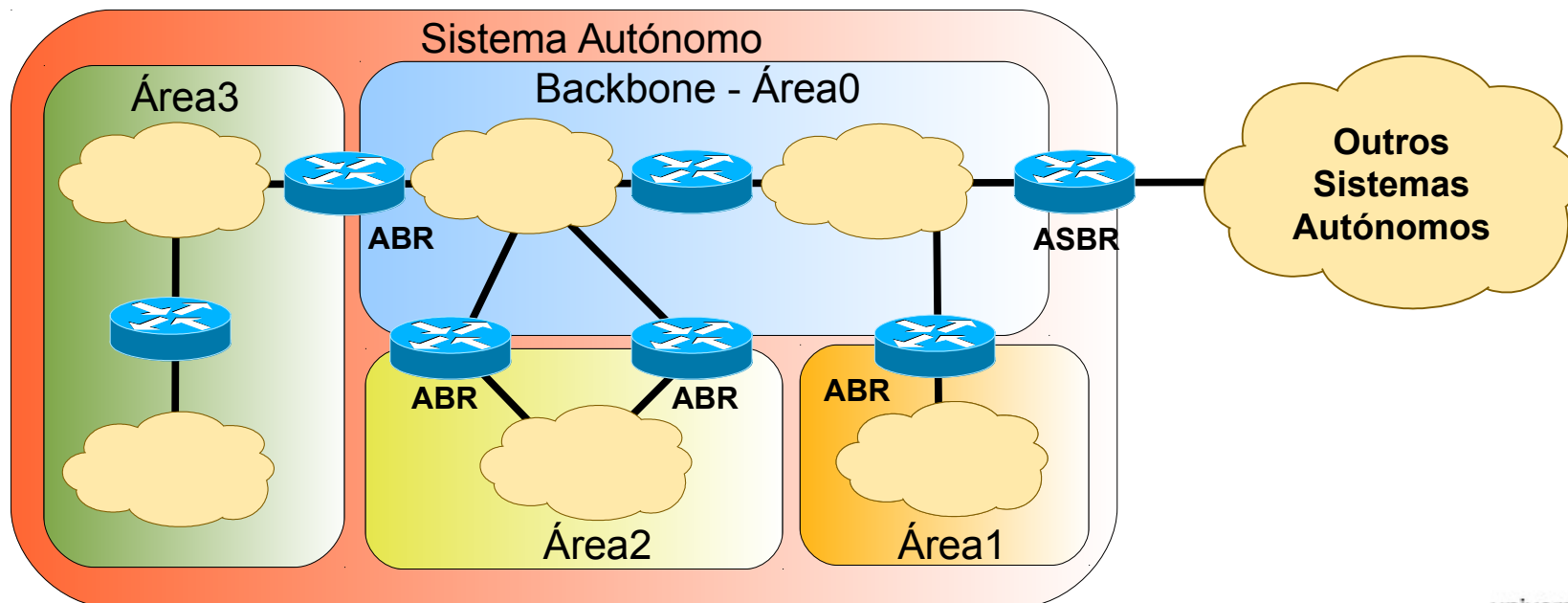
80.000000	10.10.10.3	224.0.0.5	OSPF	Hello Packet
83.683033	10.10.10.3	224.0.0.5	OSPF	LS Update
83.715683	10.10.10.3	224.0.0.5	OSPF	Hello Packet
83.717864	10.10.10.1	10.10.10.3	OSPF	Hello Packet
83.726166	10.10.10.3	10.10.10.1	OSPF	DB Descr.
83.726258	10.10.10.3	10.10.10.1	OSPF	Hello Packet
83.728433	10.10.10.1	10.10.10.3	OSPF	DB Descr.
83.732590	10.10.10.3	10.10.10.1	OSPF	DB Descr.
83.734733	10.10.10.1	10.10.10.3	OSPF	DB Descr.
83.738942	10.10.10.3	10.10.10.1	OSPF	LS Request
83.741083	10.10.10.1	10.10.10.3	OSPF	LS Update
84.240362	10.10.10.3	224.0.0.5	OSPF	LS Update
86.245792	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
86.380876	10.10.10.1	224.0.0.5	OSPF	Hello Packet
86.741036	10.10.10.1	224.0.0.5	OSPF	LS Acknowledge
93.721376	10.10.10.3	224.0.0.5	OSPF	Hello Packet
96.380005	10.10.10.1	224.0.0.5	OSPF	Hello Packet

213.780338	10.10.10.3	224.0.0.5	OSPF	Hello Packet
216.542473	10.10.10.1	224.0.0.5	OSPF	Hello Packet
216.568852	10.10.10.1	224.0.0.5	OSPF	LS Update
217.048427	10.10.10.1	224.0.0.5	OSPF	LS Update
217.084909	10.10.10.1	224.0.0.5	OSPF	LS Update
219.067748	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
219.650308	10.10.10.1	224.0.0.5	OSPF	LS Update
222.150349	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
223.779492	10.10.10.3	224.0.0.5	OSPF	Hello Packet
224.284149	10.10.10.3	224.0.0.5	OSPF	LS Update
224.789598	10.10.10.1	224.0.0.5	OSPF	LS Update
224.789775	10.10.10.3	224.0.0.5	OSPF	LS Update
226.545718	10.10.10.1	224.0.0.5	OSPF	Hello Packet
226.785254	10.10.10.1	224.0.0.5	OSPF	LS Acknowledge
227.294756	10.10.10.3	224.0.0.5	OSPF	LS Acknowledge
233.779863	10.10.10.3	224.0.0.5	OSPF	Hello Packet
236.544658	10.10.10.1	224.0.0.5	OSPF	Hello Packet

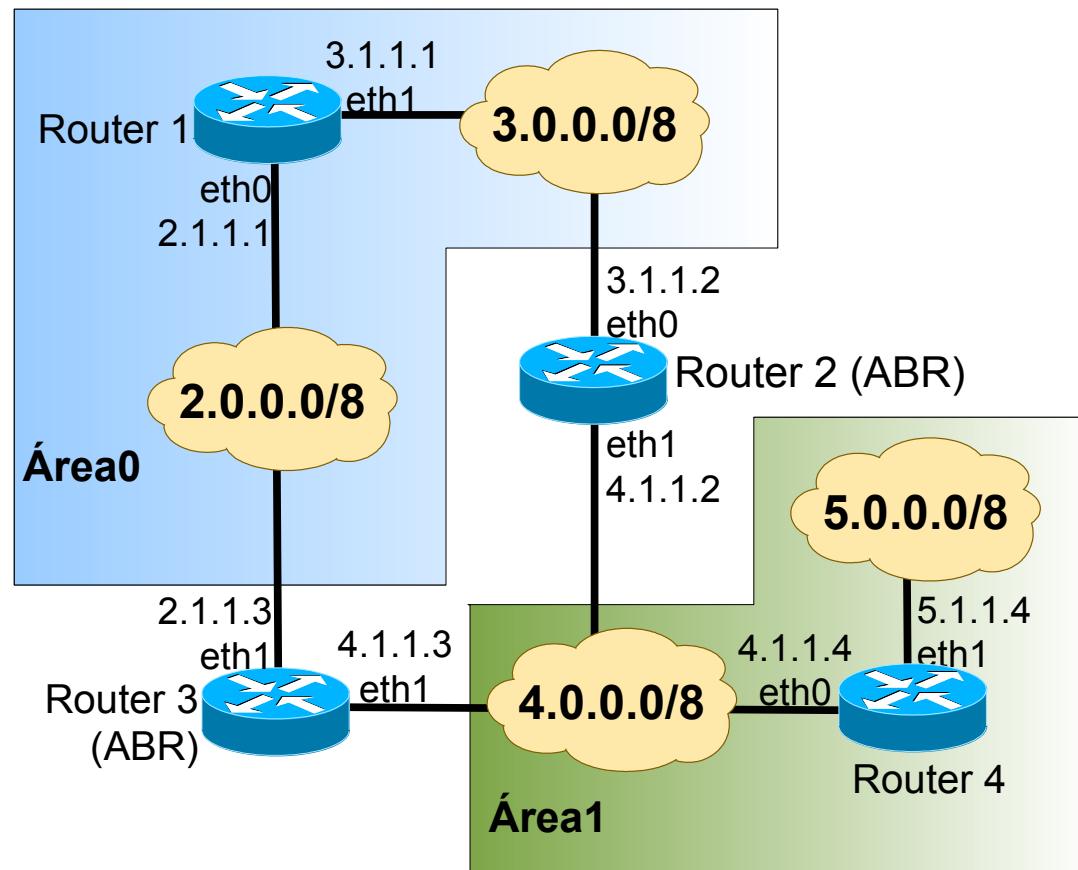


Áreas no OSPF

- Objetivo: tornar o protocolo escalável para sistemas autónomos grandes.
- Um router interior conhece apenas a topologia da sua área.
- Todas as áreas tem de estar (fisicamente ou virtualmente) ligadas à Área0.
- Um router fronteira entre áreas é chamado de Area Border Router (ABR).
 - ♦ Um ABR conhece a topologia das áreas a que está ligado.
- Um router fronteira do domínio OSPF é chamado de Autonomous System Border Router (ASBR).



Encaminhamento hierárquico OSPF



Advertising Router: 3.1.1.1	Advertising Router: 4.1.1.2
Number of Links: 2	Number of Links: 1
Router Interface address: 3.1.1.1	Router Interface address: 3.1.1.2
TOS 0 Metrics: 10	TOS 0 Metrics: 10
Router Interface address: 2.1.1.1	Advertising Router: 4.1.1.3
TOS 0 Metrics: 10	Number of Links: 1
	Router Interface address: 2.1.1.3
	TOS 0 Metrics: 10

Router Link States do Router 1

Link State ID: 2.1.1.3	Link State ID: 3.1.1.2
Network Mask: /8	Network Mask: /8
Attached Router: 3.1.1.1	Attached Router: 3.1.1.1
Attached Router: 4.1.1.3	Attached Router: 4.1.1.2

Net Link States do Router 1

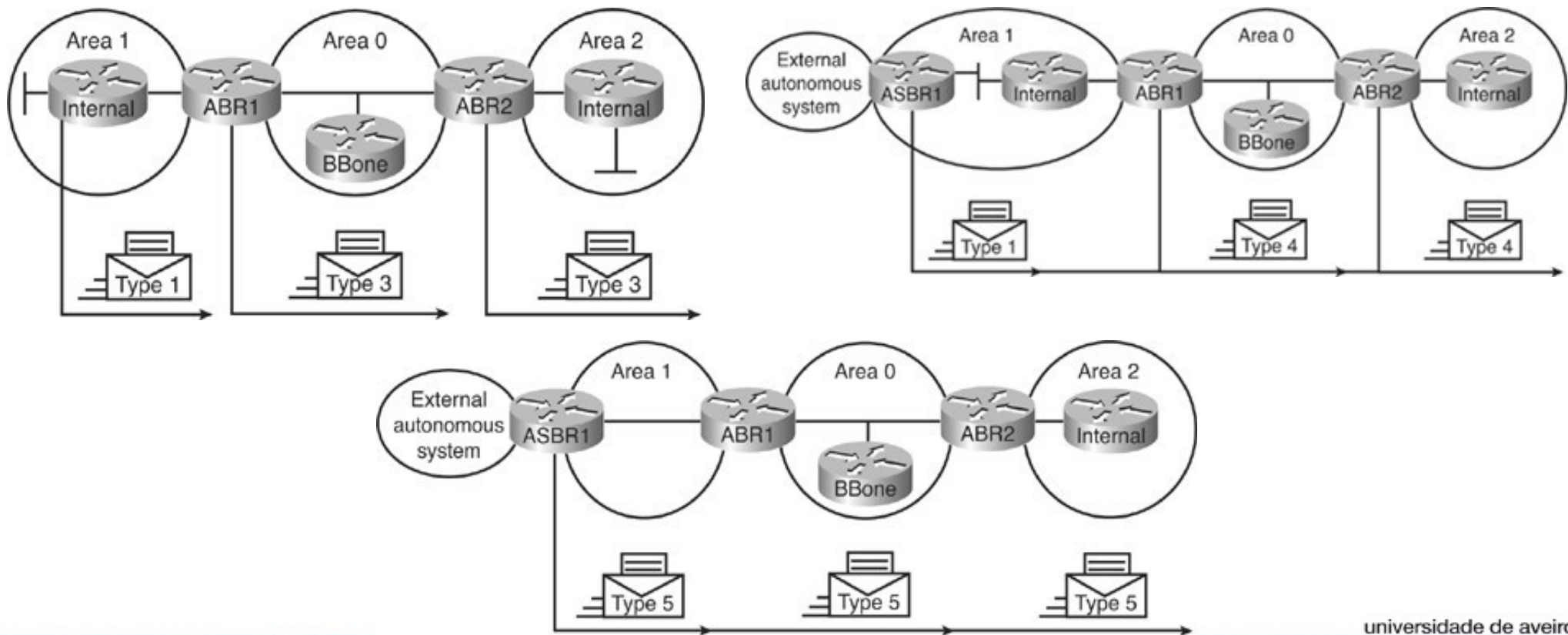
Link State ID: 4.0.0.0	Link State ID: 5.0.0.0
Advertising Router: 4.1.1.2	Advertising Router: 4.1.1.2
Network Mask: /8	Network Mask: /8
TOS: 0 Metric: 10	TOS: 0 Metric: 20
Link State ID: 4.0.0.0	Link State ID: 5.0.0.0
Advertising Router: 4.1.1.3	Advertising Router: 4.1.1.3
Network Mask: /8	Network Mask: /8
TOS: 0 Metric: 10	TOS: 0 Metric: 20

Summary Net Link States do Router 1



Tipos de OSPF LSA

- Type 1 (Router LSA) – Informação sobre routers.
- Type 2 (Network LSA) – Informação sobre redes.
- Types 3 (Summary LSA) – Sumários de redes de outras áreas.
- Types 4 (Summary LSA) – Sumário de rotas para os routers fronteira do domínio OSPF (os ASBR).
- Type 5 (AS external LSA) – Informação sumariada de redes externas ao domínio OSPF.



RIP vs. OSPF

- RIP

- ♦ Protocolo mais simples.
- ♦ Encaminhamento baseado em número de saltos.
- ♦ Não é escalável (infinito = 16).
- ♦ Processamento contínuo de tabelas de encaminhamento.

- OSPF

- ♦ Protocolo complexo (para sincronizar bases de dados distribuídas).
- ♦ Escalável (para redes grandes, a solução é o encaminhamento hierárquico).
- ♦ Maior flexibilidade de encaminhamento (baseado em custos configuráveis).
- ♦ Processamento pontual das tabelas de encaminhamento.
- ♦ Utilização intensa da rede apenas quando há alterações da topologia da rede (processo de flooding).
- ♦ Processo de convergência das tabelas de encaminhamento mais rápido.



Encaminhamento IPv6

RIPng

- Similar ao RIPv2,
 - ♦ É um protocolo distance-vector, máximo de 15 saltos, split-horizon, etc.
- Diferenças entre o RIPv2 e RIPng,
 - ♦ Usa IPv6 para o transporte.
 - ♦ Anuncia prefixos de rede IPv6 e o next-hop é um endereço IPv6.
 - ♦ Usa o endereço multicast IPv6 FF02::9 (all-RIP-routers) como endereço de destino dos anúncios RIP.



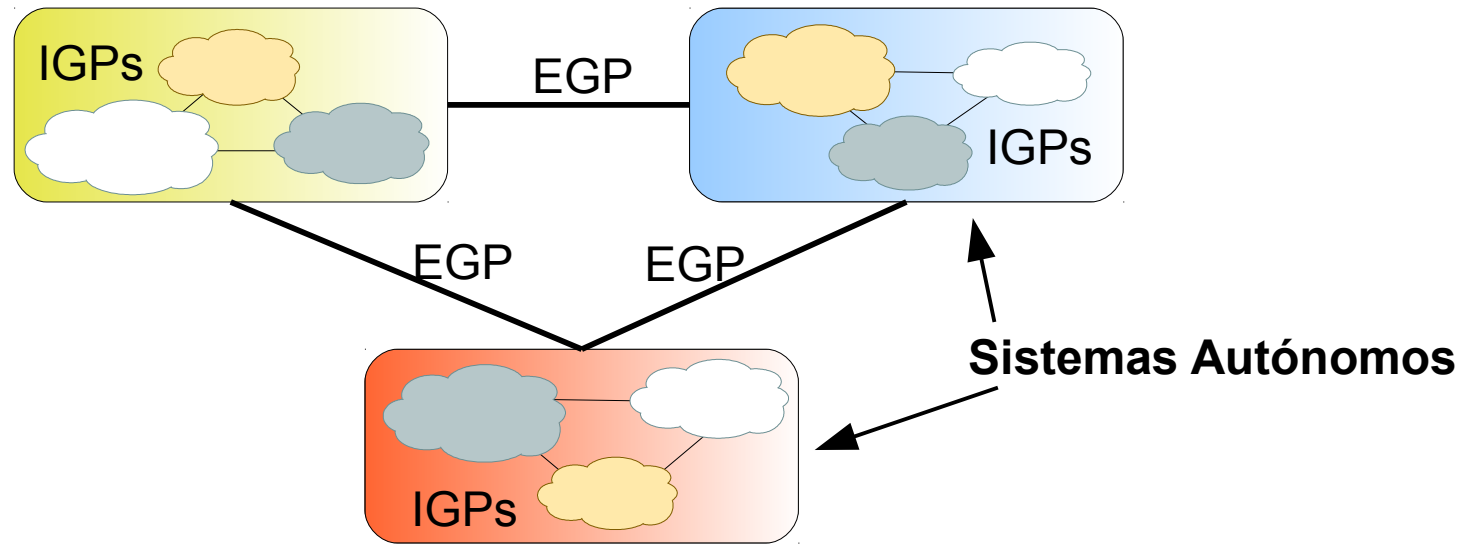
OSPFv3

- Baseado no OSPFv2, com melhoramentos,
 - ♦ Usa IPv6 para o transporte.
 - ♦ Distribui prefixos de rede IPv6.
 - ♦ Usa os endereços IPv6 multicast FF02::5 (OSPF IGP) e FF02::6 (OSPF IGP Designated Routers).
- A base de dados com a descrição da topologia não é específica do IPv6.
 - ♦ Router ID, Area ID, Link ID continua a ser um identificador de 4 bytes.
 - ♦ Os vizinhos são identificados por um Router ID de 4 bytes.
 - ♦ Apenas existe uma tabela adicional na base de dados que mapeia um identificador de rede (link) num determinado prefixo e máscara de rede IPv6
- São usados os endereços IPv6 Link-Local com origem dos anúncios.
- Novos Tipos de LSA:
 - ♦ Link LSA (Type 8)
 - ➔ Contem informação sobre os endereços IPv6 Link-Local e prefixos IPv6.
 - ♦ Intra-Area Prefix LSA (Type 9)
 - ➔ Associa um prefixo IPv6 com uma rede ou router.



BGP

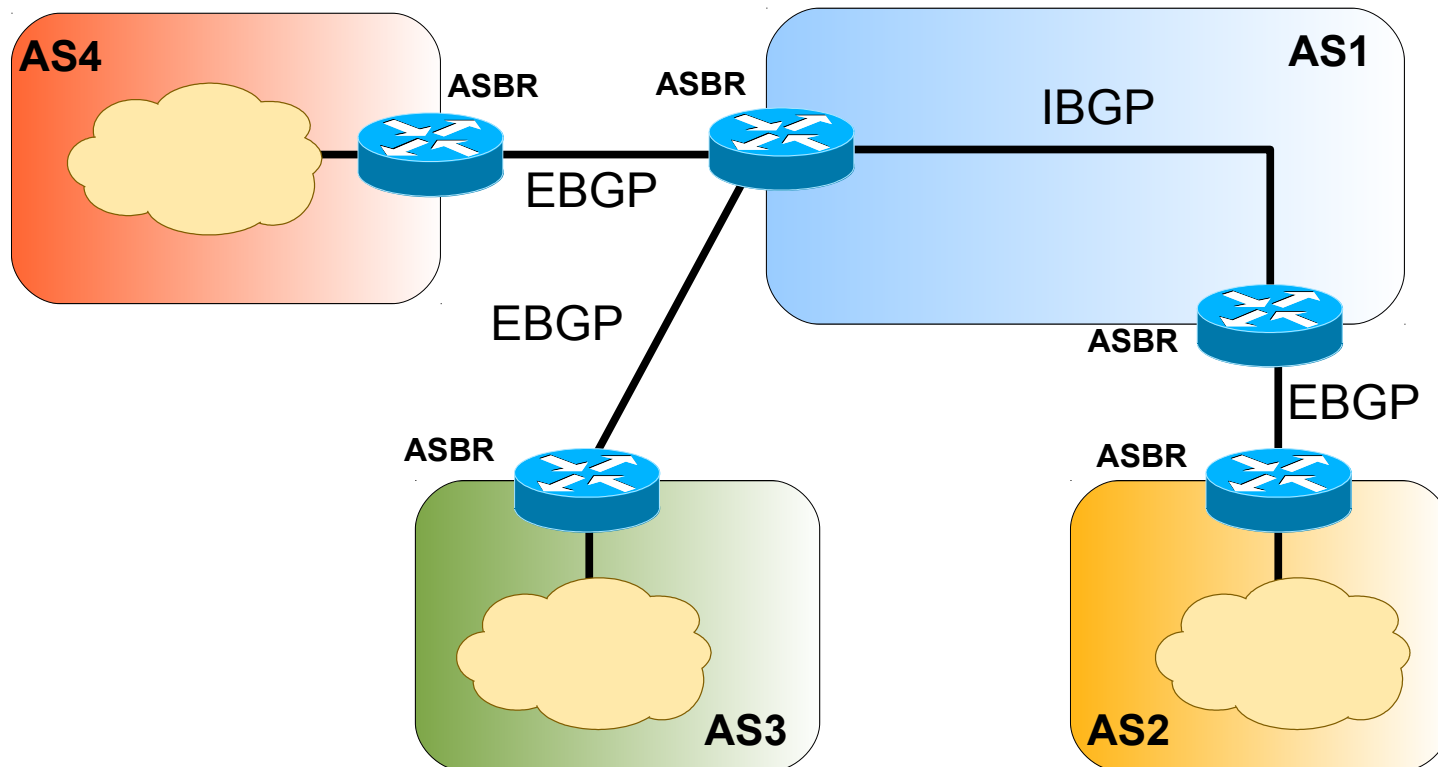
Border Gateway Protocol (BGP)



- Protocolo para o encaminhamento entre diferentes Sistemas Autónomos (do tipo EGP)
- O BGP usa como protocolo de transporte o TCP e o número de porto 179
- Número AS definido com 2-bytes
 - ◆ AS Públicos= 1 a 64511
 - ◆ AS Privados= 64512 a 65535
 - ◆ Os números de AS de 2-bytes disponíveis estão quase esgotados
- O número do AS passou de 2-bytes para 4-bytes
 - ◆ Permite 4.294.967.295 AS
 - ◆ Transição gradual

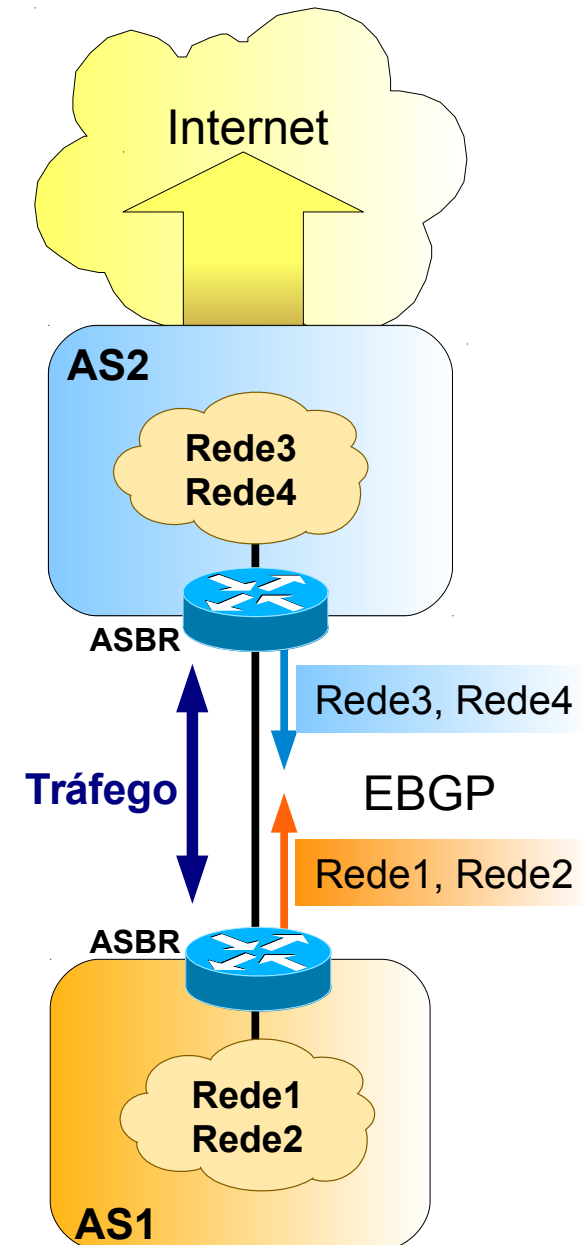
BGP Interno (IBGP) e BGP Externo (EBGP)

- As relações de vizinhança podem ser estabelecidas entre routers de um mesmo SA (Internal BGP - IBGP) ou de diferentes SA (External BGP - EBGP)



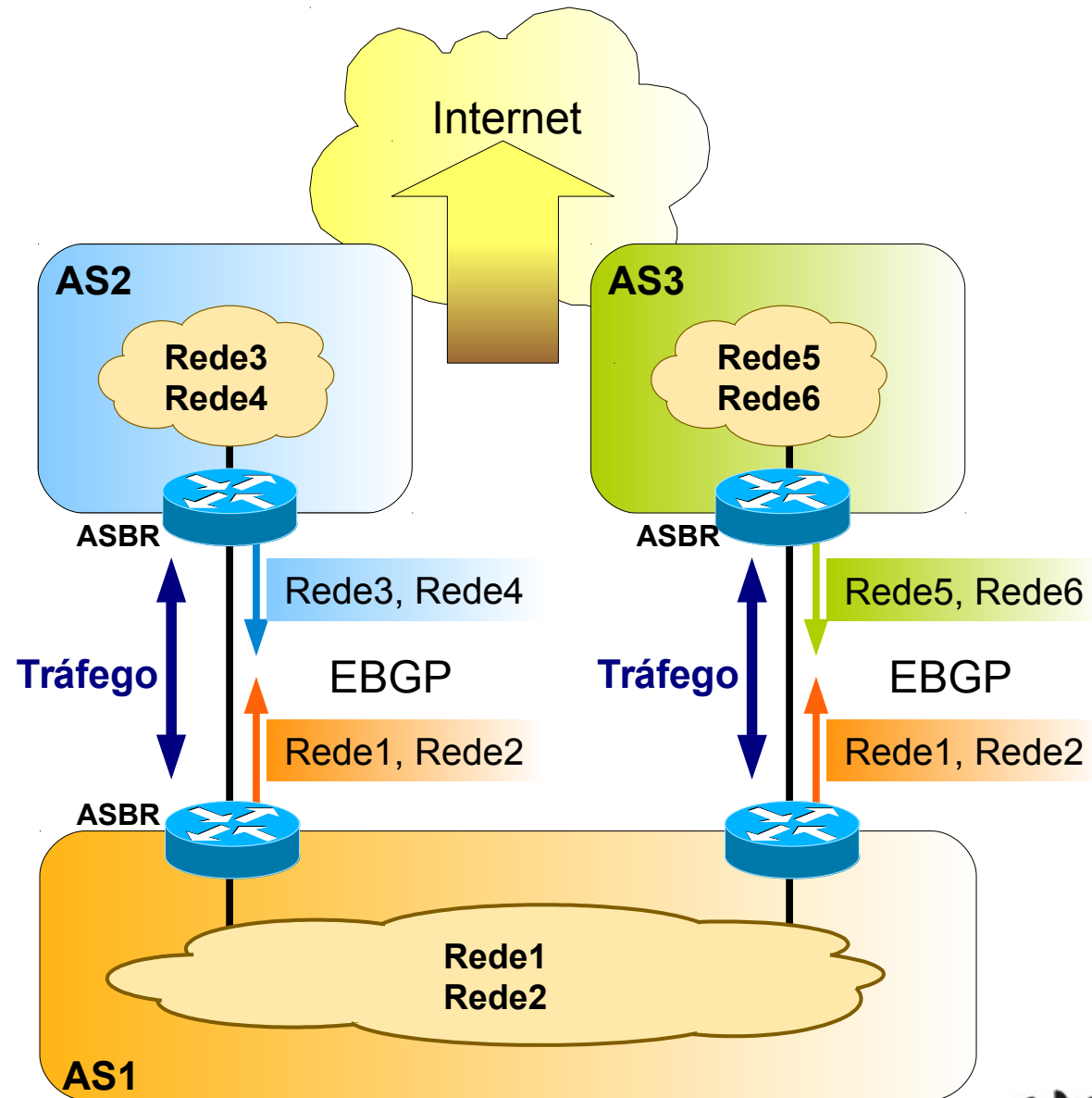
Sistema Autónomo single-homed

- Possui apenas um *router* fronteira para atingir redes fora do seu domínio administrativo



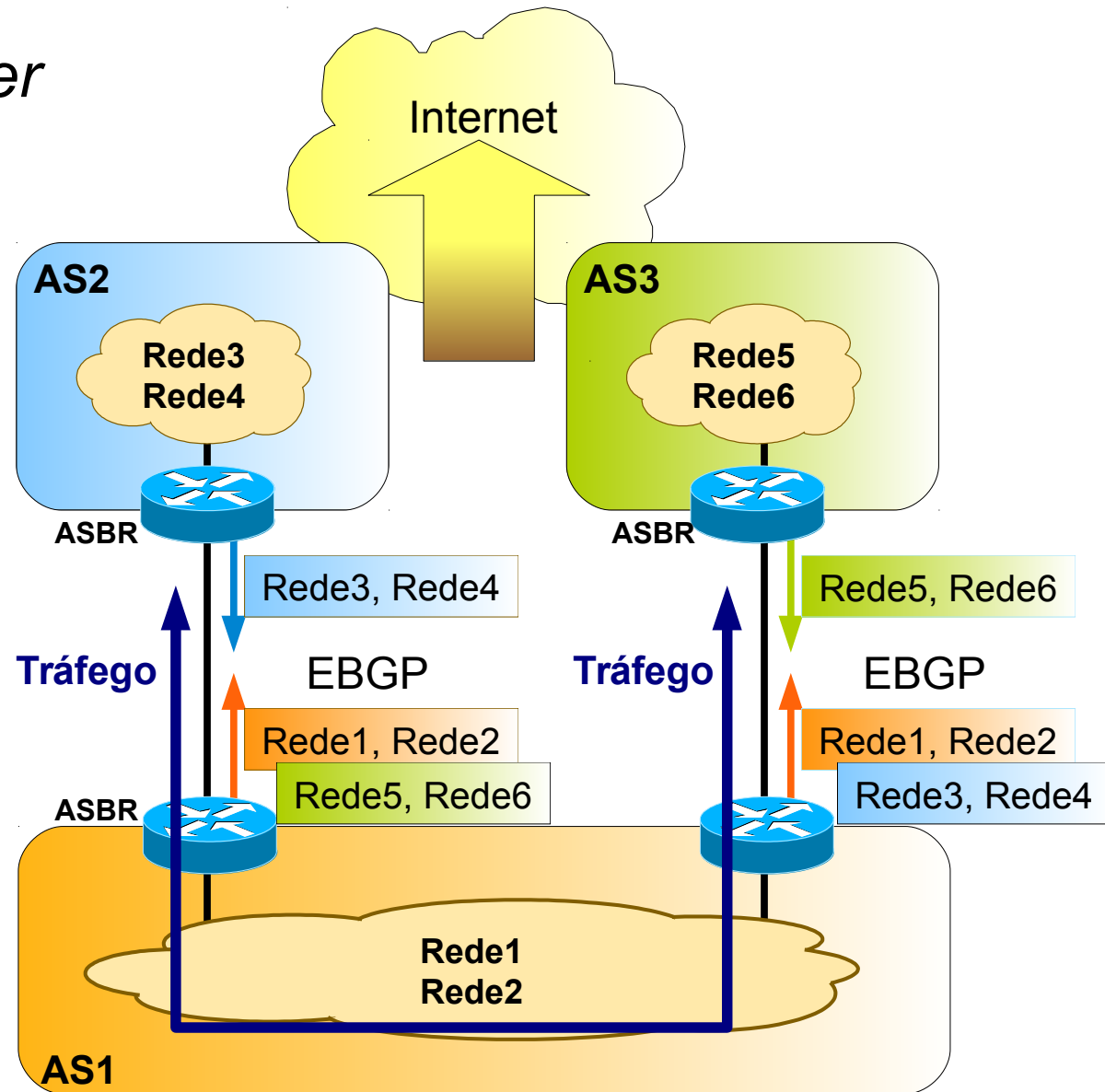
Sistema Autônomo non-transit multi-homed

- Possui mais do que um *router* fronteira para o mundo exterior.
- Não permite que o tráfego de trânsito o atravessasse.



Sistema Autónomo transit multi-homed

- Possui mais do que um *router* fronteira para o mundo exterior.
- Suporta tráfego de trânsito.

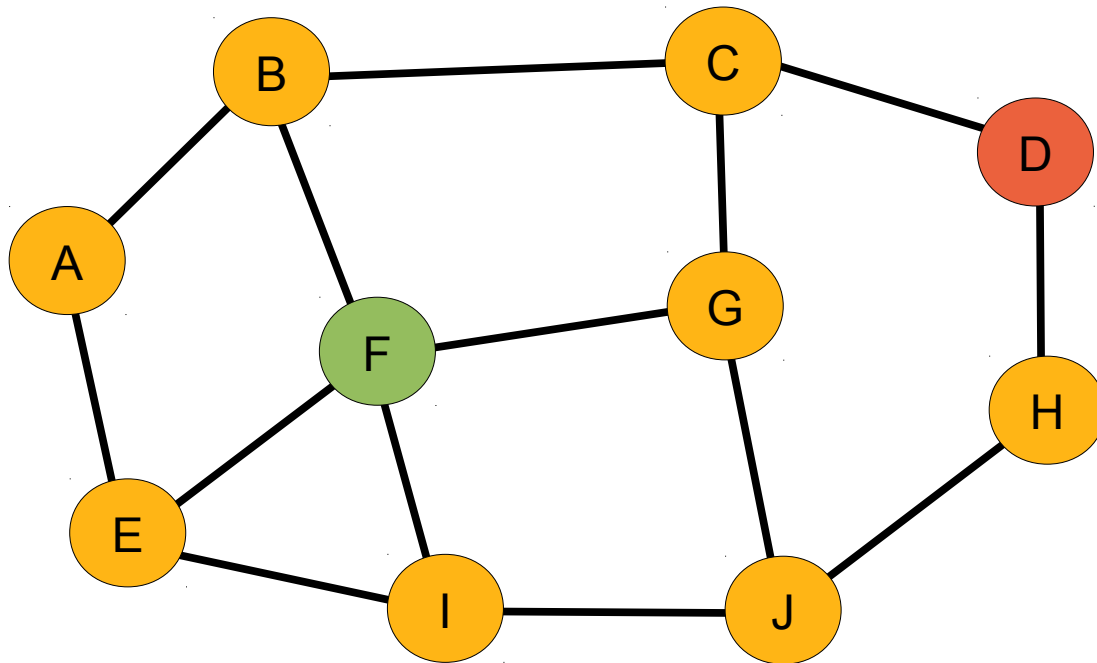


Pacotes BGP

- Os routers BGP começam por estabelecer relações de vizinhança usando mensagens **Open** (declaram, por exemplo, o número do seu AS)
- Inicialmente são trocadas todas as rotas BGP; a informação de encaminhamento é sempre transportada nas mensagens **Update**
- Não havendo qualquer alteração de rotas, os routers trocam apenas mensagens **Keepalive** (são enviadas periodicamente entre vizinhos BGP para manter a relação de vizinhança)
- As mensagens **Notification** são transmitidas para reportar situações de erro e para terminar uma ligação



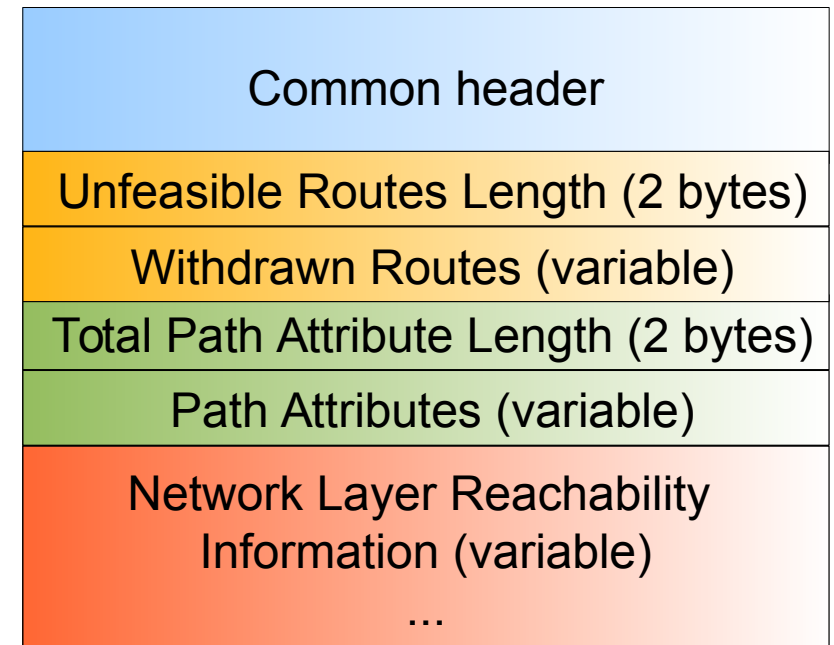
Path vectors



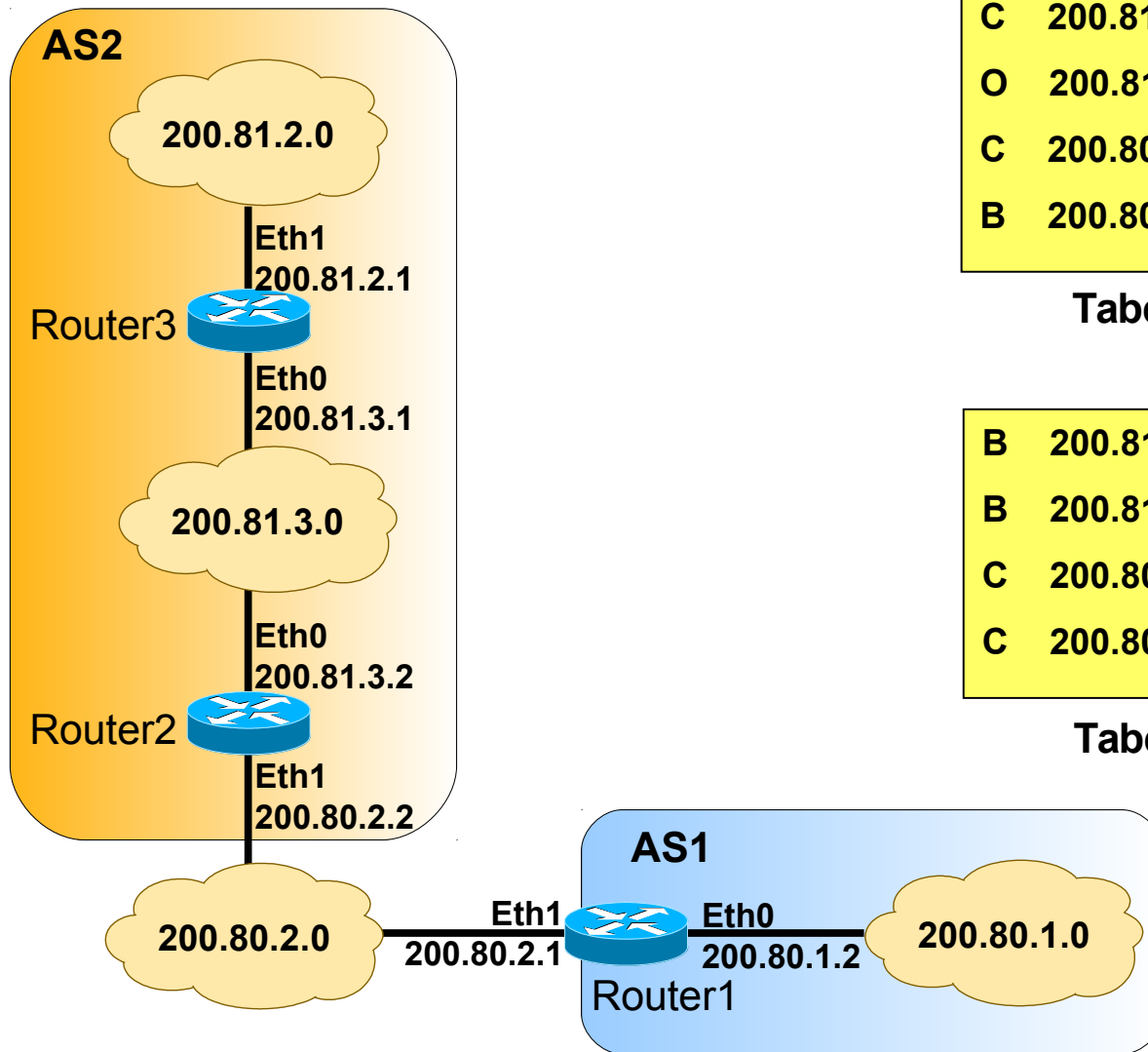
- F recebe dos seus vizinhos o caminho para as redes de D:
 - ♦ De B: “Eu uso BCD”
 - ♦ De G: “Eu uso GCD”
 - ♦ De I: “Eu uso IFGCD”
 - ♦ De E: “Eu uso EFGCD”

Mensagens *Update*

- Withdrawn routes – lista de redes IP que já não podem ser atingidas
- Path attributes – permite implementar políticas de encaminhamento
 - AS_PATH: enumeração dos SAs no percurso para o destino; o SA é identificado por um número com dois octetos
- Network layer reachability information – listagem das redes destino anunciadas



Exemplo



C 200.81.3.0/24 is directly connected, Ethernet0
O 200.81.2.0/24 [110/20] via 200.81.3.1, 00:01:12
C 200.80.2.0/24 is directly connected, Ethernet1
B 200.80.1.0/24 [20/0] via 200.80.2.1, 00:00:29

Tabela de encaminhamento do Router 2

B 200.81.3.0/24 [20/0] via 200.80.2.2, 00:01:58
B 200.81.2.0/24 [20/0] via 200.80.2.2, 00:01:57
C 200.80.2.0/24 is directly connected, Ethernet1
C 200.80.1.0/24 is directly connected, Ethernet0

Tabela de encaminhamento do Router 1

Exemplo - agregação BGP

Antes da AGREGAÇÃO

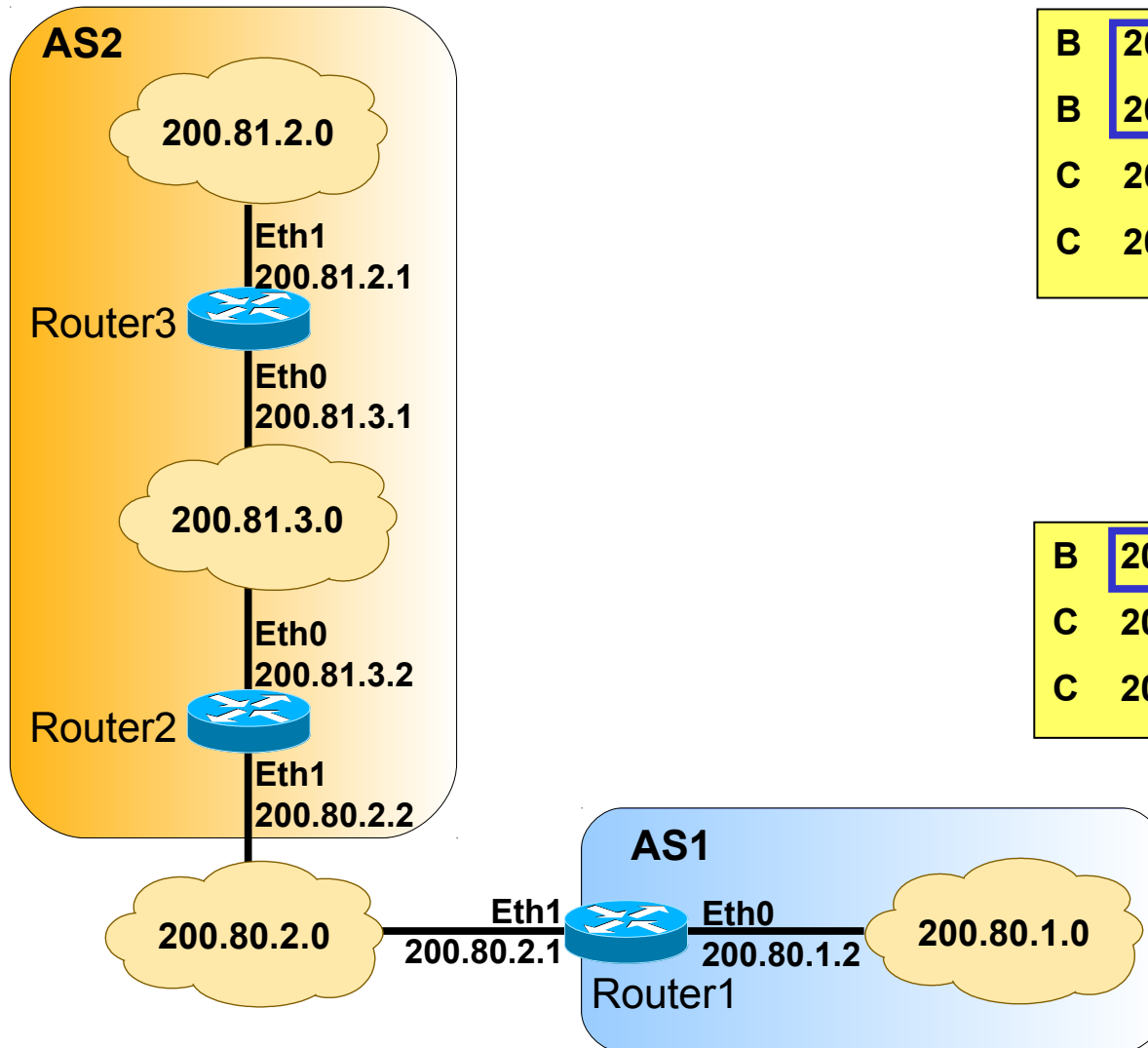
B 200.81.3.0/24 [20/0] via 200.80.2.2, 00:01:58
B 200.81.2.0/24 [20/0] via 200.80.2.2, 00:01:57
C 200.80.2.0/24 is directly connected, Ethernet1
C 200.80.1.0/24 is directly connected, Ethernet0

Router 1

Depois da AGREGAÇÃO

B 200.81.2.0/23 [20/0] via 200.80.2.2, 00:01:06
C 200.80.2.0/24 is directly connected, Ethernet1
C 200.80.1.0/24 is directly connected, Ethernet0

Router 1



Atributos BGP

- Um atributo BGP, ou path attribute, é uma métrica usada para descrever as características de um caminho BGP.
- Os atributos estão incluídos nas mensagens update que são trocadas pelos peers BGP para anunciar rotas. Existem 4+1 categorias de atributos BGP.
 - ♦ Well-known Mandatory (incluídas nos updates BGP)
 - AS-path, Next-hop, Origin.
 - ♦ Well-known Discretionary (podem ou não estar incluídas nos updates BGP)
 - Local Preference, Atomic Aggregate.
 - ♦ Optional Transitive (podem não ser suportadas por todas as implementações BGP)
 - Aggregator, Community.
 - AS4_Aggregator, AS4_path
 - Usados na transição de números de AS de 2-bytes para 4-bytes
 - ♦ Optional Non-transitive (podem não ser suportadas por todas as implementações BGP)
 - Se o vizinho não suportar este atributo, ele é apagado
 - Multi-exit-discriminator (MED).
 - ♦ Cisco-defined (local ao router, não é anunciado)
 - Weight



Atributos AS-path e Origin

- AS-path

- ♦ Quando o anúncio de uma rota passa através de um Sistema Autónomo, o número do SA é adicionado a uma lista ordenada de números de SAs correspondente aos sistemas que o anúncio já atravessou.

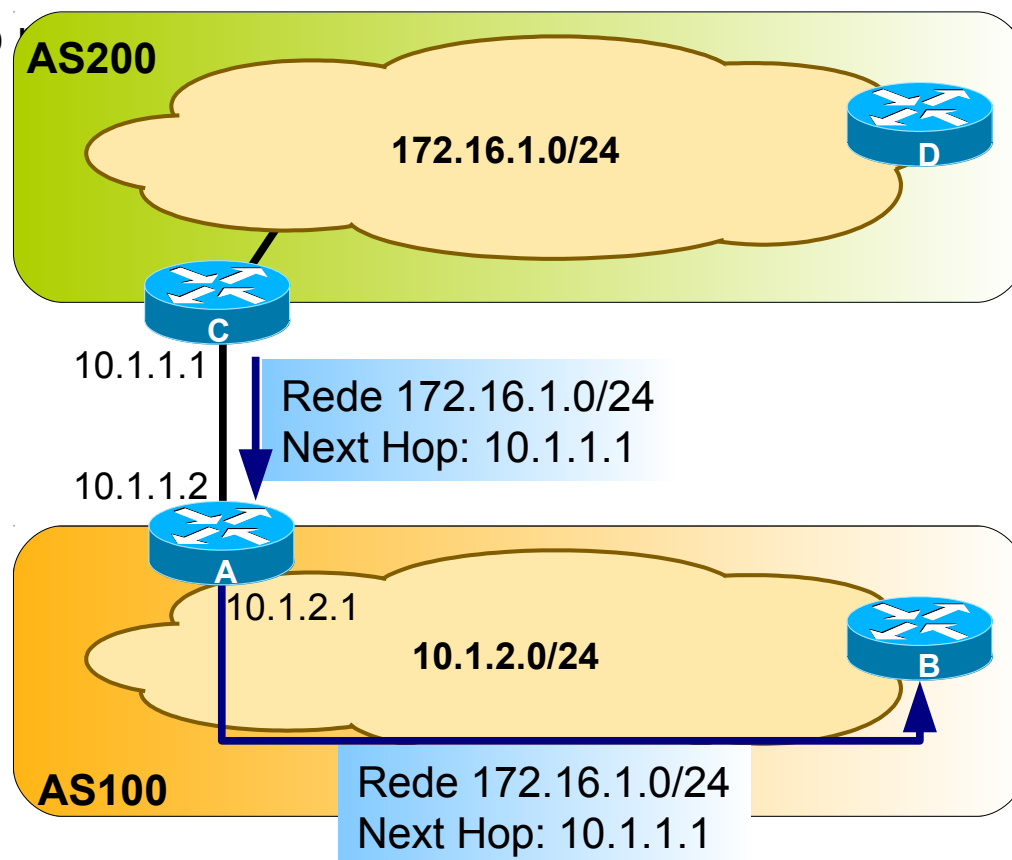
- Origin

- ♦ Indica como é que o BGP aprendeu a informação relativa a uma determinada rota.
 - ➔ IGP — A rota é interior ao SA de origem.
 - ➔ EGP — A rota é aprendida através do Exterior Border Gateway Protocol (EBGP).
 - ➔ Incomplete — A origem da rota é desconhecida ou foi aprendida de qualquer outra forma.



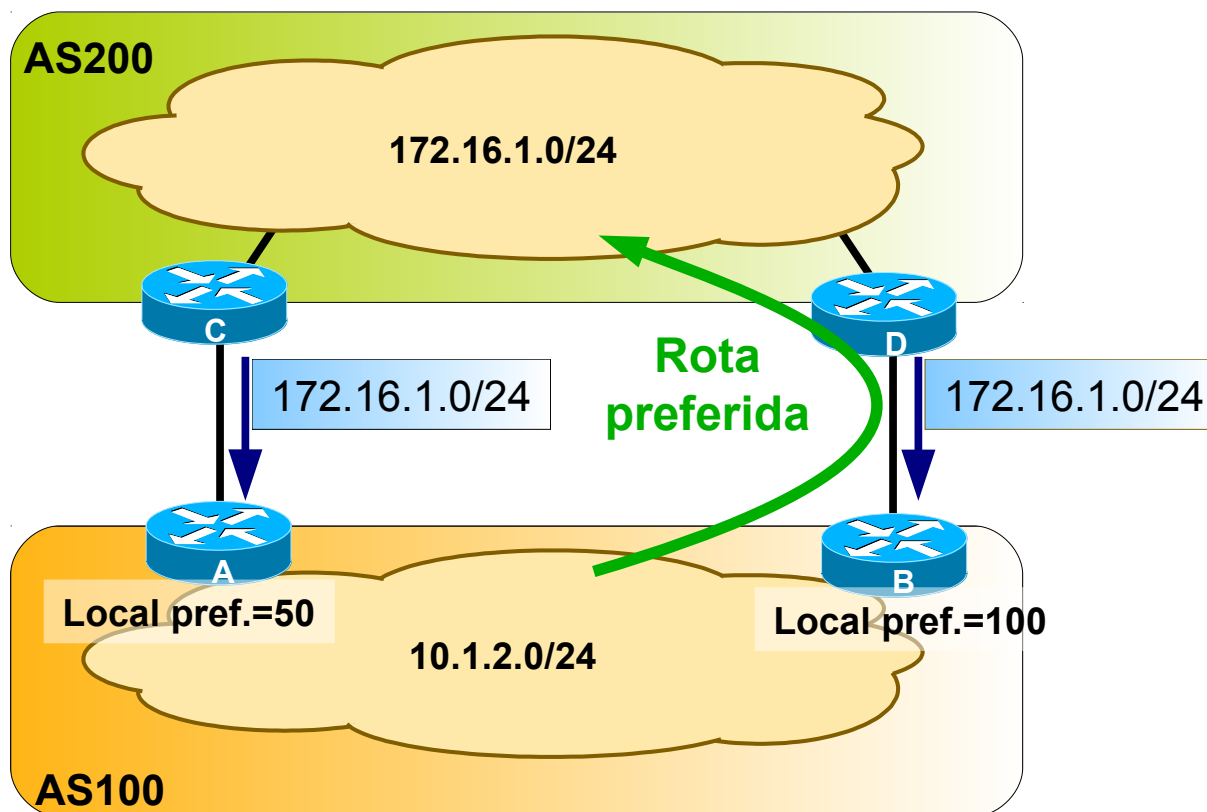
Atributo Next-Hop

- O atributo next-hop do EBGP é o endereço IP que é usado para alcançar o router anunciante.
- Para o EBGP, o endereço next-hop é o endereço IP da ligação entre os peers.
- Para o IBGP, o endereço next-hop EBGP é transportado para dentro do SA local
 - ♦ Por configuração, o



Atributo Local Preference

- Este atributo é usado para escolher um ponto de saída do SA local.
- Este atributo é propagado por todo o SA local.
- Se existirem múltiplos pontos de saída do SA, o atributo local preference é usado para selecionar o ponto de saída para uma rota específica.



Atributos Atomic Aggregate e Aggregator

- Atomic Aggregate

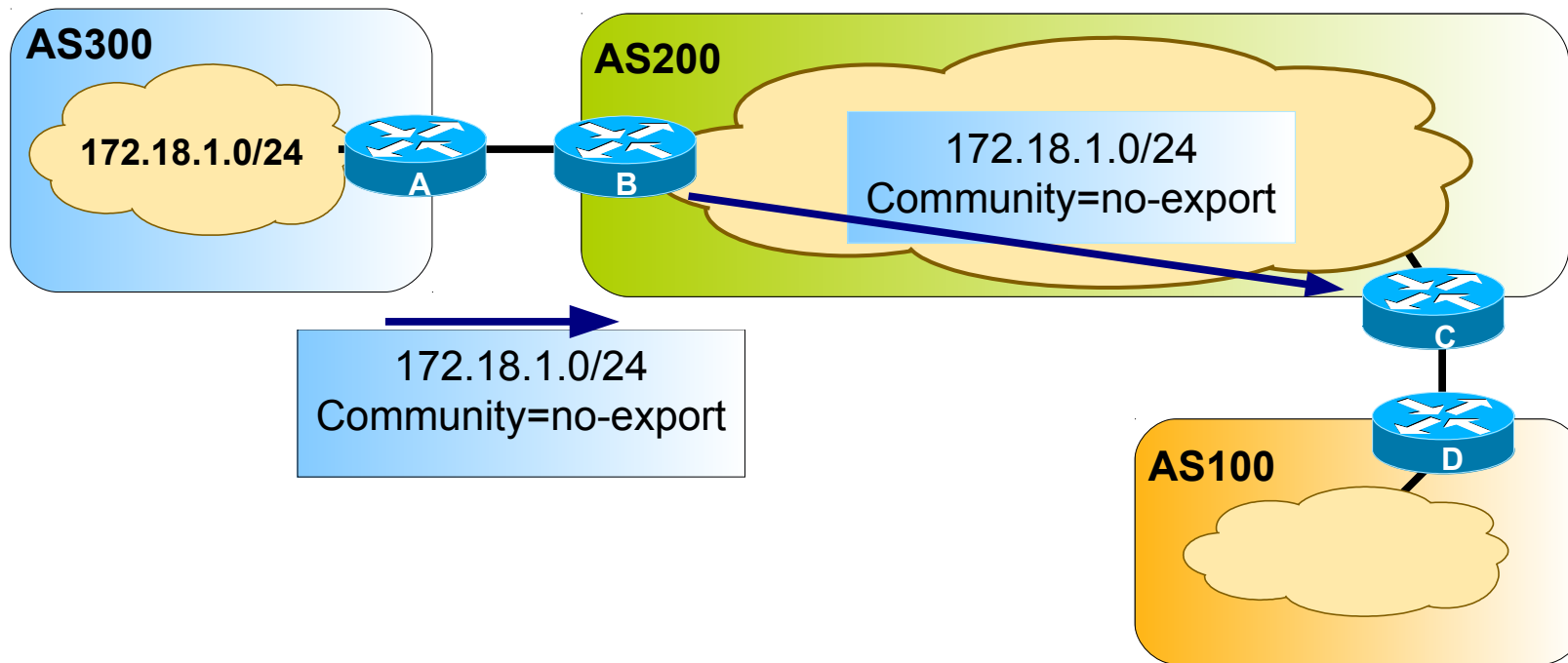
- ♦ É usado para alertar os routers que certas rotas específicas foram agregadas numa rota menos específica.
- ♦ Quando este tipo de agregação acontece, são perdidas as rotas mais específicas.

- Aggregator

- ♦ Fornece informação relativamente ao SA que realizou a agregação.
- ♦ É o endereço IP do router que originou o agregado.



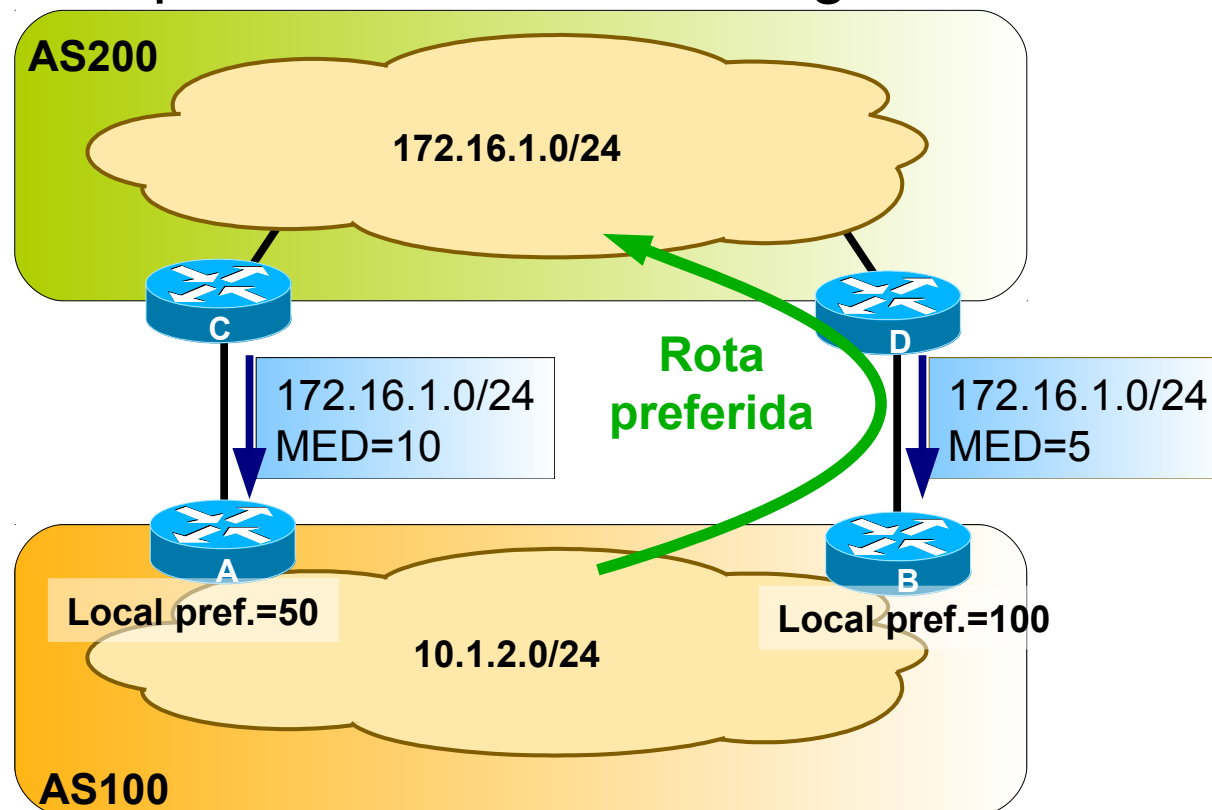
Atributo Community



- Usado para agregar rotas que partilham propriedades comuns de tal forma que possam ser aplicadas políticas ao nível do agregado
- Os atributos community pré-definidos são:
 - ♦ no-export – não anunciar esta rota aos peers EBGP.
 - ♦ no-advertise - não anunciar esta rota a nenhum peer.
 - ♦ internet – anunciar esta rota à comunidade Internet; todos os routers da rede pertencem a ela.

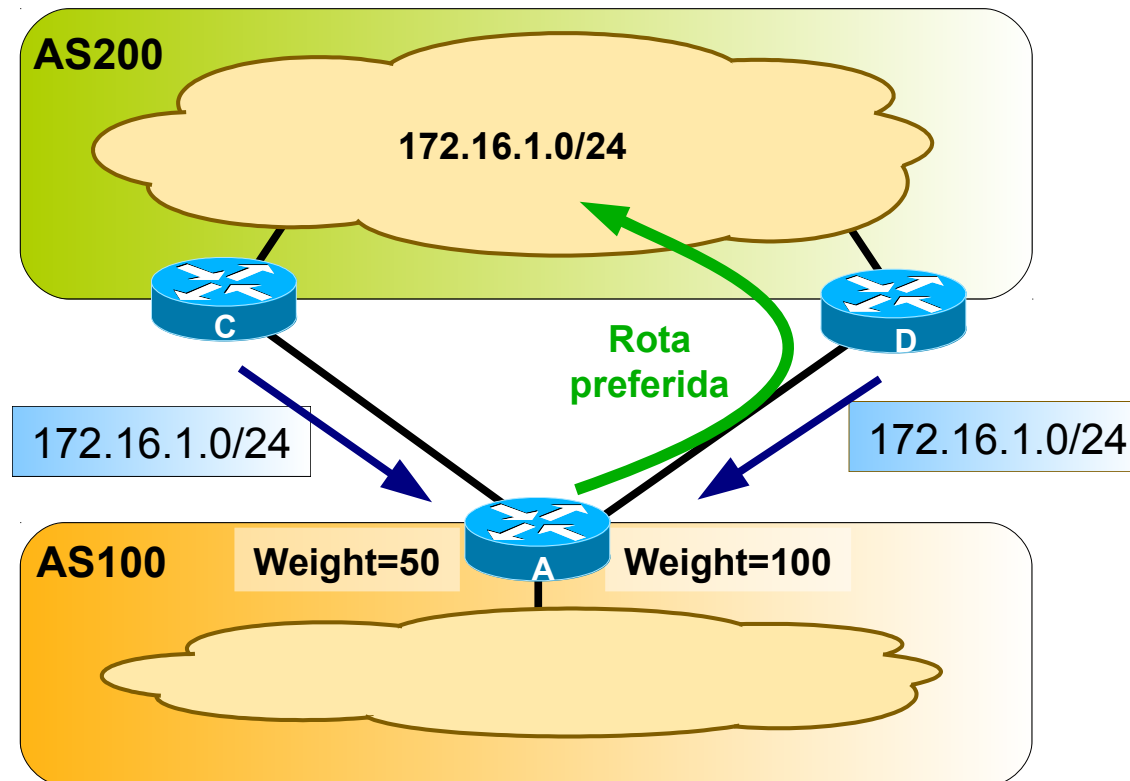
Atributo Multi-Exit Discriminator (MED)

- O atributo MED é usado como sugestão a dar a um SA externo.
- O SA externo que está a receber os MEDs pode estar a usar outros atributos BGP para a seleção das rotas.
- O valor mais baixo da métrica é o preferido.
- MED foi idealizado para influenciar o tráfego de entrada.



Atributo Weight

- Weight é um atributo definido pela Cisco e é local a um router.
- Não é anunciado aos routers vizinhos.
- Se o router aprender mais do que uma rota para o mesmo destino, a rota com o weight mais elevado é a preferida.



Seleção de caminhos BGP

- BGP pode receber múltiplos anúncios para uma mesma rota a partir de múltiplas fontes.
- BGP seleciona apenas um caminho como o melhor.
- BGP coloca o caminho selecionado na tabela de encaminhamento IP e propaga-o aos seus vizinhos.
- BGP utiliza os seguintes critérios, por ordem:
 - Maior weight (apenas para a Cisco)
 - Maior local preference
 - Caminho que foi originado localmente
 - Shortest path
 - Menor tipo de origem (IGP menor do que EGP, EGP menor do que incompleto)
 - Menor atributo MED
 - Prefere o caminho externo sobre o caminho interno
 - Vizinho IGP mais próximo



Filtragem BGP e Route Maps

- O envio e a recepção de updates BGP pode ser controlado usando diferentes métodos de filtragem.
- Os updates BGP podem ser filtrados com base em
 - ♦ Informação da rota
 - ♦ Informação do caminho
 - ♦ Comunidades
- Route maps são usados com BGP para
 - ♦ Controlar e modificar a informação de encaminhamento
 - ♦ Definir as condições de acordo com as quais as rotas são redistribuídas entre domínios de encaminhamento



Sincronização BGP

- A sincronização diz que, se o seu SA encaminha tráfego de outro SA para um terceiro SA, o BGP não deve anunciar uma rota antes que todos os routers do seu SA tenham aprendido essa rota via IGP.
- O BGP espera até que o IGP tenha propagado a rota no interior do SA. Então, anuncia a rota aos peers externos.



MP-BGP

- Extension to the BGP protocol
- Carries routing information about other protocols:
 - IPv6 Unicast
 - Multicast (IPv4 and IPv6)
 - 6PE - IPv6 over IPv4 MPLS backbone
 - Multi-Protocol Label Switching (MPLS) VPN (IPv4 and IPv6)
- Exchange of Multi-Protocol Reachability Information (NLRI)
 - New non-transitive and optional attributes
 - ➔ MP_REACH_NLRI
 - Carry the set of reachable destinations together with the next- hop information to be used for forwarding to these destinations
 - ➔ MP_UNREACH_NLRI
 - Carry the set of unreachable destinations
 - Attribute contains one or more triples
 - ➔ Address Family Information (AFI) with Sub-AFI
 - Identifies protocol information carried in the Network Layer Reachability Information
 - ➔ Next-hop information
 - Next-hop address must be of the same family

