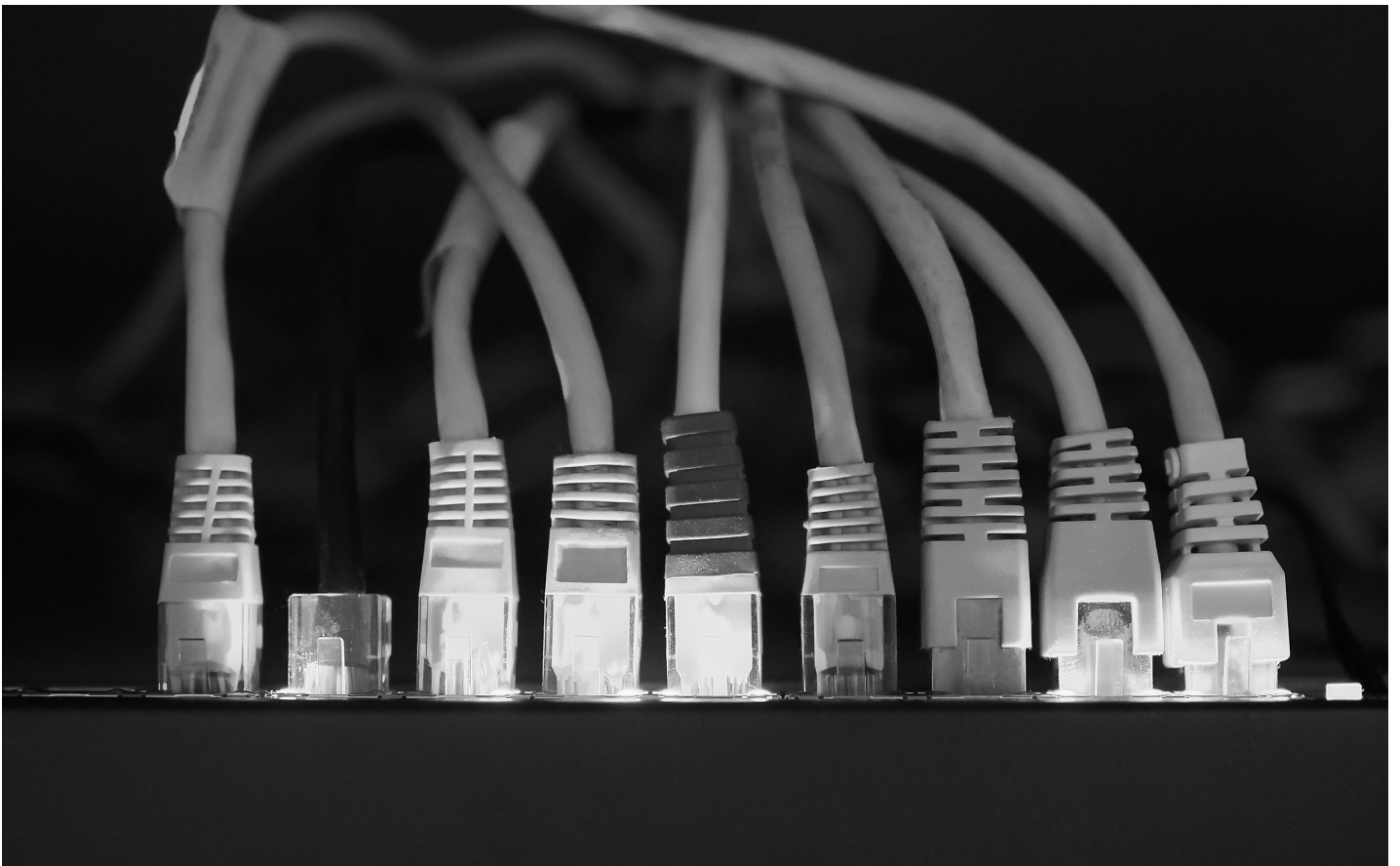


EXTERNAL ROUTING (BGP & MP-BGP)

JUNHO DE 2022

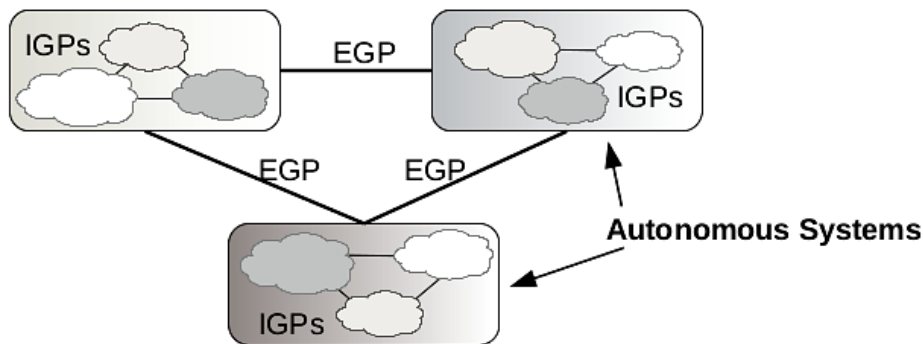
REDES DE COMUNICAÇÕES II



ENGENHARIA DE COMPUTADORES E INFORMÁTICA

BORDER GATEWAY PROTOCOL (BGP) 178

Usado principalmente para routing entre sistemas autónomos.
Autonomous System (AS) é uma rede sob uma única administração.



AS NUMBERS

Um ID alocado pelo InterNIC que é globalmente único.
RFC 4271 define o número AS em 2 bytes.
Há a necessidade de expandir o tamanho de 2 para 4 bytes.
RFC 4893 define suporte para números AS de 4 bytes.

BGP NEIGHBOR RELATIONSHIPS

Também chamada de peering, usualmente é manualmente configurada em routers pelos administradores.
Cada sessão como um vizinho funciona sobre TCP (port 179).
Os pares trocam todas as suas rotas quando a sessão é iniciada.
Atualizações também são enviadas quando a topologia muda na rede ou quando muda a política de routing.
Os pares trocam mensagens de KEEP ALIVE durante a sessão.

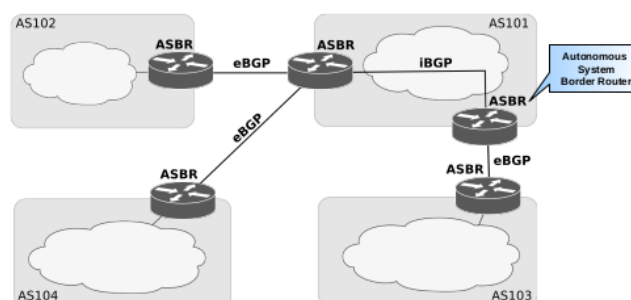
- Evita longos períodos de inatividade.
- Intervalos entre mensagens baixos podem ser configurados de manter uma fail-over rápida.

INTERNAL BGP (IBGP) & EXTERNAL BGP (EBGP)

Relações de vizinhança podem ser estabelecidas entre:

- Routers da mesma AS (iBGP).
- Routers de diferentes AS (eBGP).

Routers que implementam relações de vizinhança são chamados de Autonomous System Border Router (ASBR).

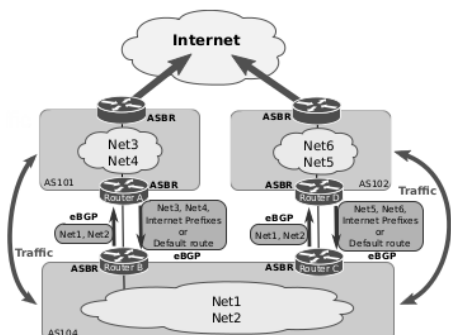
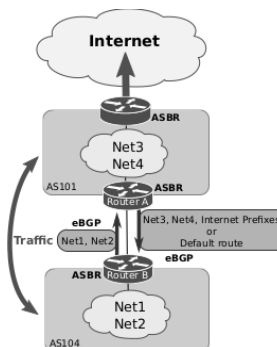


EXTERNAL AND INTERNAL BGP

- eBGP usado entre ASs.
- iBGP usado dentro de uma AS.
- Um router BGP nunca encaminha um caminho aprendido de um iBGP vizinho para outro iBGP vizinho, mesmo que esse caminho seja o melhor.
 - Apenas acontece se o router for configurado como route-reflector.
- Um router BGP encaminha rotas aprendidas de um vizinho eBGP tanto para vizinhos eBGP como iBGP.
 - Podem ser configurados filtros para alterar este comportamento.
- Routers iBGP dentro de uma AS **devem manter uma sessão iBGP como todos os outros iBGP** da mesma AS.
- Permite obter informação das rotas completa para todas as redes externas conhecidas pelo grupo.

SINGLE-HOMED (OR STUB) AS

- A AS tem apenas um router fronteira (ASBR).
 - Apenas um acesso à Internet.
 - Apenas um ISP

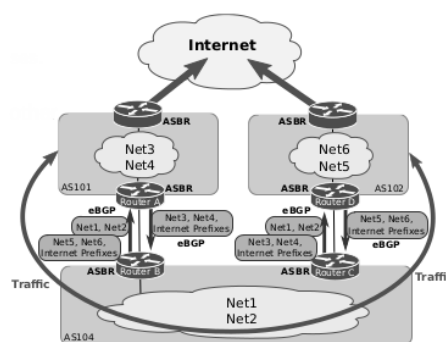


MULTI-HOMED NON-TRANSIT AS

- Tem um ou mais ASBR.
 - Múltiplos acessos à Internet.
 - Múltiplos ISP.
- Não transporta tráfego de outras AS.

MULTI-HOMED TRANSIT ASE 1

- Tem um ou mais ASBR.
 - Múltiplos acessos à Internet.
 - Múltiplos ISP.
- Transporta tráfego de outras AS.

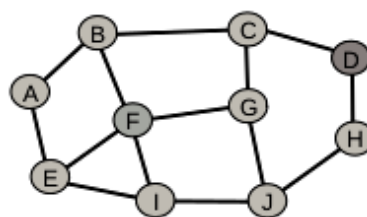


PATH-VECTOR

- BGP é um protocolo path-vector.
- Apesar de ser essencialmente um protocolo de vetor distância que carrega uma lista da AS que atravessam a rota.
- Um EBGp adiciona a sua própria AS á lista antes de encaminhar a rota para um vizinho EBGp.
- Um IBGP não modifica a lista pois está a enviá-la para um vizinho dentro da mesma AS.

EXEMPLO

- **F recebe dos seus vizinhos diferentes caminhos para D.**
- De B -> BCD
- De G -> GCD
- De I -> IFGCD
- De E -> EFGCD



MESSAGES

OPEN

- São usadas para estabelecer uma sessão BGP

UPDATE

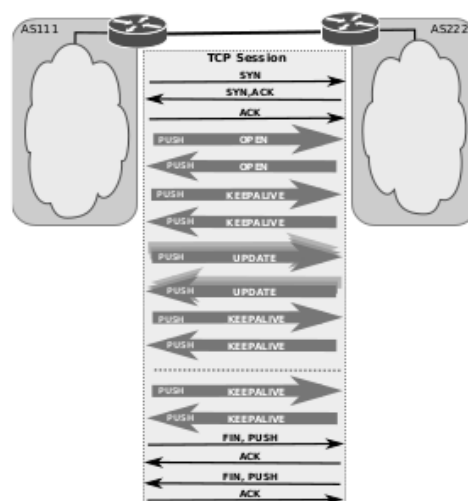
- Usadas para mandar prefixos de routing, assim como atributos de BGP (como o AS-PATH).

KEEPALIVE

- São enviadas sempre que o período de keepalive é excedido e nenhum update foi trocado.

NOTIFICATION

- Mensagens são enviadas sempre que um erro de protocolo é detetado, após isto a sessão BGP é fechada.



UPDATE

Withdraw routes

- Lista de de redes já não alcançáveis.

Path attributes

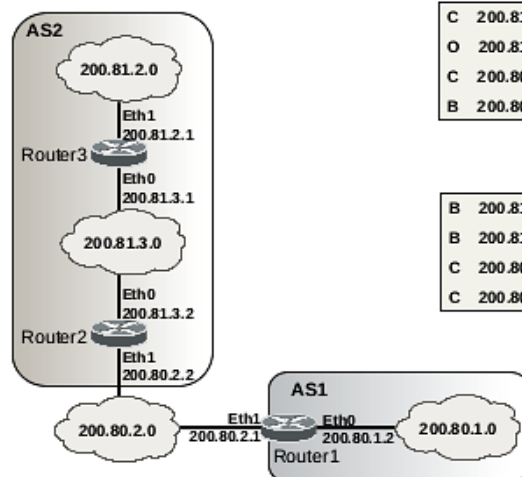
- Parâmetros usados para definir routing e políticas de routing.

Network layer reachability information

- Lista de redes com conectividade.

Withdrawn Routes Length			
Withdrawn Routes			
Length	Prefix		
Length	Prefix		
Total Path Attribute Length			
Path Attributes			
Attr. Flags	Attr. Code	Attr. Length	Attribute Value
Attr. Flags	Attr. Code	Attr. Length	Attribute Value
Network Layer Reachability Information			
Length	Prefix		
Length	Prefix		

EXEMPLOS



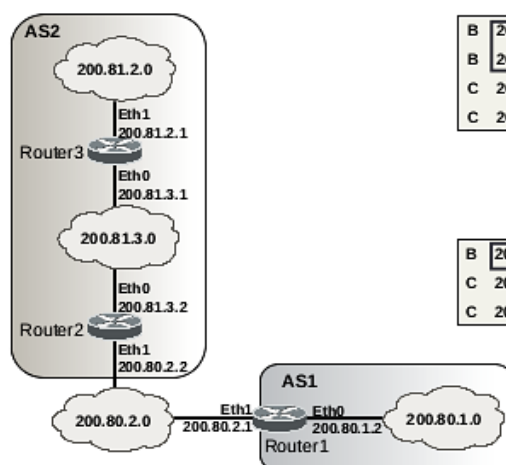
C 200.81.3.0/24 is directly connected, Ethernet0
 O 200.81.2.0/24 [110/20] via 200.81.3.1, 00:01:12
 C 200.80.2.0/24 is directly connected, Ethernet1
 B 200.80.1.0/24 [20/0] via 200.80.2.1, 00:00:29

Router 2's routing table

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's routing table

BGP NETWORKS AGGREGATION



Before aggregation

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

After aggregation

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

BGP ATTRIBUTES

É uma métrica usada para descrever as características de um caminho BGP. Atributos estão contidos dentro de mensagens update passadas entre vizinhos BGP para anunciar rotas.

- **Well-known Mandatory (incluído nas mensagens de BGP)**
 - AS-PATH;
 - Next-hop;
 - Origin;
- **Well-known Discretionary**
 - Preferência local;
 - Atomic Aggregate.
- **Optional Transitive**
 - Aggregator;
 - Community;
 - AS4_Aggregator;
 - AS4_path.
- **Optional Non-transitive**
 - Se o vizinho não suportar este atributo, é eliminado;
 - Multi-exit-discriminator (MED).
- **Cisco-defined (local ao router, não é anunciado)**
 - Peso.

AS-PATH

Quando um anúncio de uma rota passa por um sistema autônomo, o número AS é adicionado a uma lista ordenada que contém os números AS dos sistemas por onde a rota passou.

ORIGIN

Indica como o BGP aprendeu uma rota particular. Podem ser um de três valores:

- **IGP (0)** é utilizado se a rota for interna ao AS, resultante da inclusão explícita no processo de routing de um BGP, de forma manual;
- **INCOMPLETE (2)** é utilizado se a rota é aprendida por outros meios.
- **EGP (1)** é obsoleto.

AS4_PATH & AS4_AGGREGATOR

AS4_PATH tem a mesma semântica de AS_PATH, mas é opcional e transitivo, e o número de AS é de 4 bytes.

AS4_AGGREGATOR tem a mesma semântica de AGGREGATOR, excepto que o AS têm 4 bytes.

Recebendo UPDATES de um novo speaker

- Descodificada cada AS como 4 bytes;
- AS_PATH e AGGREGATOR são afetados.

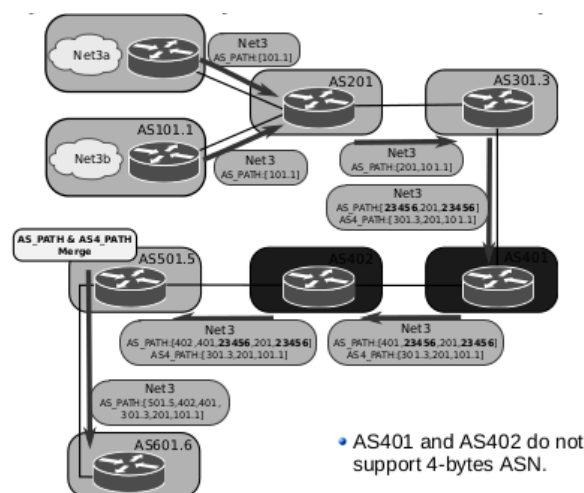
Recebendo UPDATES de um speaker velho.

- AS4_AGGREGATOR vai sobre-escrever AGGREGATOR;
- AS4_PATH e AS_PATH têm de ser combinados para formar um AS_PATH correto.

Combinar AS4_PATH e AS_PATH

- o AS_PATH -> [275 250 225 23456 23456 200 23456 175]
- o AS4_PATH -> [100.1 100.2 200 100.3 175]
- o AS_PATH combinado -> [275 250 225 100.1 100.2 200 100.3 175]

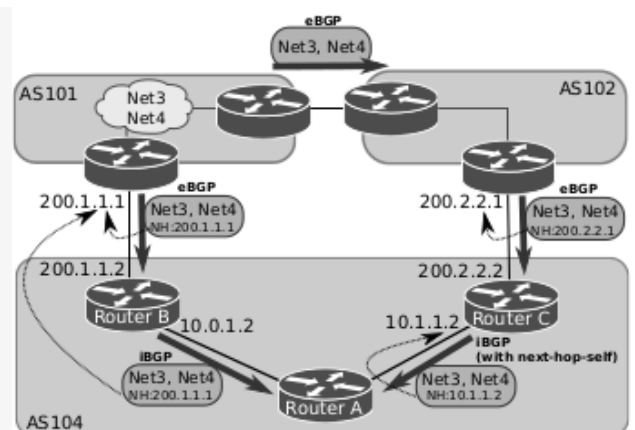
4-BYTES AS OPERATIONAL EXAMPLE



NEXT-HOP

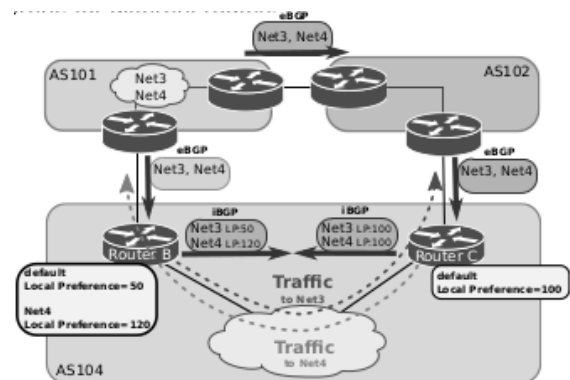
Para o eBGP é o endereço IP que é usado para alcançar o router anunciado. É também o endereço IP da conexão entre pares.

Para o iBGP, o endereço next-hop eBGP é levado para o AS local.



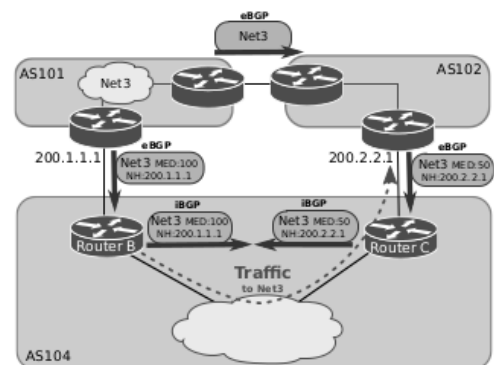
LOCAL PREFERENCE

- É usado para escolher o ponto de saída do sistema autônomo.
 - Valor mais alto é o preferido;
- É propagado pela AS local.
- Pode ser diferente entre routers.

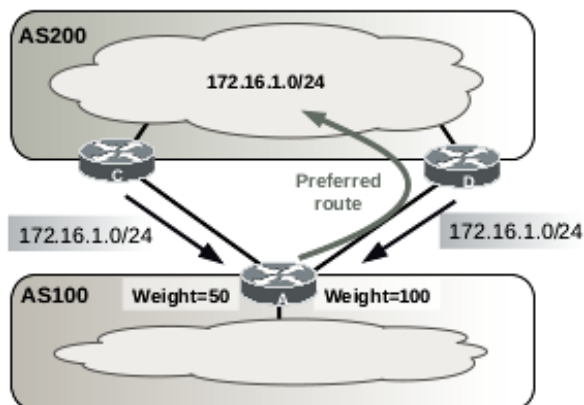


MULTI-EXIT DISCRIMINATOR ATTRIBUTE (MED)

- É usado como uma sugestão para um AS externo.
- O AS externo que recebe o MED pode estar a usar outros atributos BGP para selecionar a rota.
- O **valor mais baixo é o preferido**.
- É desenhado para influenciar o tráfego que chega.



PESO



Se o router aprender de mais que um router no mesmo destino, o router com o valor mais alto será preferido.

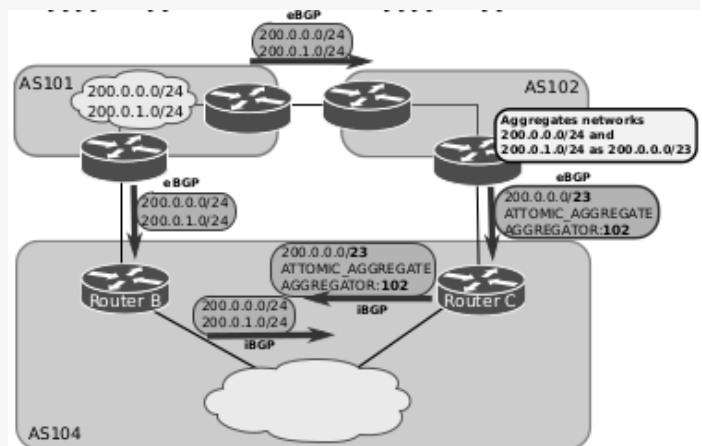
ATOMIC AGGREGATE AND AGGREGATOR ATTRIBUTES

Atomic Aggregate

- Usado para alertar routers de rotas específicas que foram agregadas para uma rota menos específica.
- Rotas mais específicas são perdidas.

Aggregator

- Informação acerca de que AS fez a agregação
- E o IP do router que realizou a agregação.



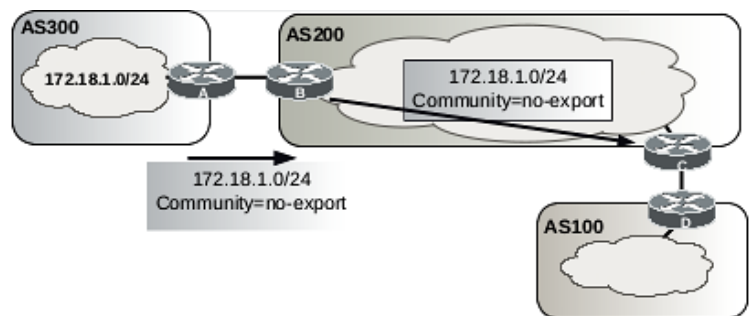
COMMUNITY ATTRIBUTE

Usado para agrupar rotas que pratilhem propriedades para que politicas possam ser aplicada ao nível do grupo.

Valores predefinidos:

- **no-export**: Não anuncia esta rota para EBGp pares.
- **no-advertise**: Não anuncia a rota para nenhum vizinho.
- **internet**: Anuncia esta rota para a internet.

O formato típico é **ASnumber:Cnumber**.



PATH SELECTION

O BGP pode receber múltiplos anúncios para a mesma rota de várias origens.

Apenas seleciona o melhor caminho.

Coloca a rota selecionada na routing table e propaga a rota para os vizinhos. Usa os seguintes critérios por ordem:

1. Maior peso;
2. Maior preferência local;
3. Rota que era originalmente local;
4. Caminho mais curto;
5. Tipo de origem mais baixo;
6. Atributo MED mais baixo;
7. Prefere o caminho externo ao interno;
8. Vizinho IGP mais próximo.