

External Routing
BGP
Resumo Arquiteturas de
Comunicação

Universidade de Aveiro

Sebastian D. González



External Routing BGP Resumo Arquiteturas de Comunicação

Dept. de Eletrónica, Telecomunicações e Informática
Universidade de Aveiro

sebastian.duque@ua.pt(103690)

23 de janeiro de 2024

Conteúdo

1	BGP - Border Gateway Protocol	1
1.1	BGP Neighbor Relationships	2
1.2	Internal BGP (iBGP) & External BGP (eBGP)	2
1.3	iBGP Mesh	2
1.4	Tipo de AS	3
1.4.1	Single-homed (or Stub) AS	3
1.4.2	Multi-homed Non-transit AS	3
1.4.3	Multi-homed Transit AS	4
1.5	Path-vector	4
1.6	Tipos de Mensagem BGP	5
1.7	Atributos BGP	6
1.7.1	AS-PATH	7
1.7.2	Origin	7
1.7.3	Next Hop	7
1.7.4	Local Preference	7
1.7.5	Multi-Exit Discriminator Attribute (MED)	8
1.7.6	Peso	8
1.7.7	Community Attribute	8
1.7.8	Atomic Aggregate and Aggregator Attributes	8
1.7.8.1	Atomic Aggregate	8
1.7.8.2	Aggregator	9
1.8	Seleção do caminho BGP	9
1.9	Multi-Protocol Border Gateway Protocol (MP-BGP)	10
1.9.1	Atributos MP-BGP	10
1.9.2	MP-BGP Negotiation Capabilities	10
1.10	BGP Avançado	11
1.10.1	Sincronização BGP	11
1.10.2	BGP Route Reflectors	11
1.10.3	BGP Route Redistribution	12
1.10.4	BGP Neighborhood Resilience	12
1.10.5	Conflitos BGP e IGP	13
1.10.6	BGP over Tunnels (over IGP)	14
1.10.7	BGP Filtering and Route Maps	14

1.10.8 BGP Case Studies	15
-----------------------------------	----

Lista de Figuras

1.1	Esquema BGP[1]	1
1.2	Single-homed AS example	3
1.3	Multi-homed Non-transit AS example	3
1.4	Multi-homed Transit AS	4
1.5	Exemplo de sessão BGP	5
1.6	AS-PATH [2]	7
1.7	Esquema atributo Peso [3]	8
1.8	Esquema dos atributos Atomic Aggregate e Aggregator Attributes	9
1.9	Esquema Route Reflectors	11
1.10	Esquema Route Redistribution	12
1.11	BGP Neighborhood Resilience	13
1.12	Conflitos BGP e IGP	13
1.13	BGP over Tunnels (over IGP)	14

Glossário

AFI Address Family Information.

AS Autonomous Systems.

BGP Border Gateway Protocol.

IGP Interior gateway protocol.

ISP Internet Service Provider.

MPLS Multi-Protocol Label Switchin.

NLRI Network Layer Reachability Information.

VPN Virtual Private Network.

Warning!!

Isto são apenas uns apontamentos realizados por uma pobre alma de MIECT, feitas a partir dos slides da disciplina e outras fontes 😈. Por favor, não usem apenas estes apontamentos como material de estudo se pretendem obter uma nota minimamente aceitável.

Dito isto, boa sorte a todos e ámen CT 🙏.

Agradecimentos aos Professores Paulo Salvador e Rui Aguiar por todo material fornecido nas aulas.

BGP - Border Gateway Protocol

Border Gateway Protocol (BGP) é usado principalmente para fazer o routing entre Autonomous Systems (AS).

- AS é uma rede ou grupo de redes sob o controle de uma única entidade administrativa.

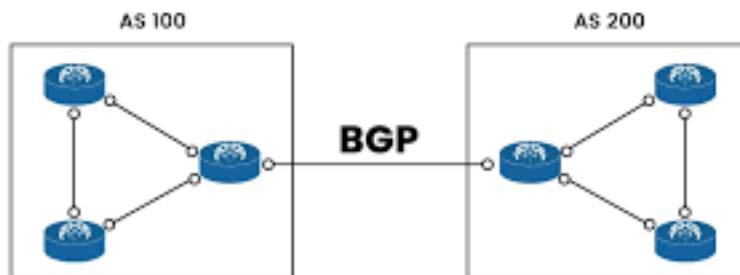


Figura 1.1: Esquema BGP[1]

1.1 BGP Neighbor Relationships

- É configurado manualmente nos routers pelo admin e cada sessão corre sobre TCP (porta 179).
- Na primeira ligação os Peers partilham todas as suas rotas.
- Só existem updates quando há uma mudança de topologia da rede ou na política de encaminhamento
- BGP peers trocam mensagens KEEPALIVE entre si (para prevenir longos períodos de inatividade)

1.2 Internal BGP (iBGP) & External BGP (eBGP)

ibgp - serve para ligar a routers no mesmo AS
ebgp - ligar a routers de diferentes AS

Um router **BGP** não encaminha um caminho de um peer IBGP para outro, a menos que seja considerado o melhor caminho, exceto quando atua como um **route-reflector**.

- Em circunstâncias normais, o **bgp** só partilha o melhor caminho entre peers **ibgp**.

Além disso, o **BGP** encaminha as rotas aprendidas de um **ebgp** para todos os peers **eBGP** e **iBGP**. Esta prática contribui para a disseminação eficiente de informações de roteamento dentro da **AS**.

1.3 IBGP Mesh

IBGP Mesh - routers **ibgp** numa AS devem manter uma sessão **iBGP** com todos os outros routers na mesma **AS**. Essa prática é necessária para obter informações completas sobre as rotas das redes externas, garantindo uma visão abrangente do ambiente de routing.

- A maior parte das redes usam um **IGP -> OSPF**

Existem métodos para reduzir a complexidade do **IBGP Mesh**, como o uso de route reflectors e a implementação de AS privadas ao ligar routers de diferentes AS que são Single-home 1.4.1 (ou Stub). Essas práticas visam otimizar o roteamento e a escalabilidade da rede.

1.4 Tipo de AS

1.4.1 Single-homed (or Stub) AS

São AS que têm um único border router, ou seja, apenas existe um acesso único à Internet. Um único Internet Service Provider (ISP)

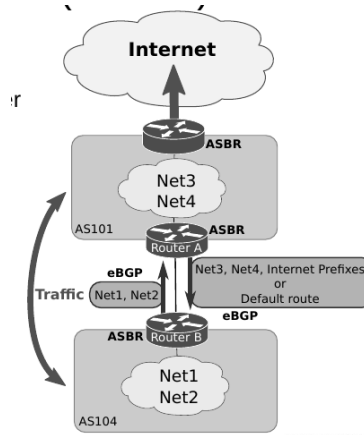


Figura 1.2: Single-homed AS example

1.4.2 Multi-homed Non-transit AS

São AS que têm mais que um border router. Múltiplos acessos únicos à Internet e vários ISP's.

Além disso, não transportam tráfego para outras AS.

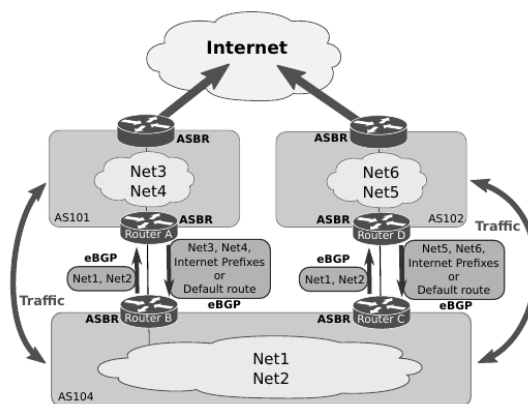


Figura 1.3: Multi-homed Non-transit AS example

1.4.3 Multi-homed Transit AS

Possui as mesmas características que as **Multi-homed Non-transit AS** 1.4.2 mas é capaz de transportar trafego de outras AS.

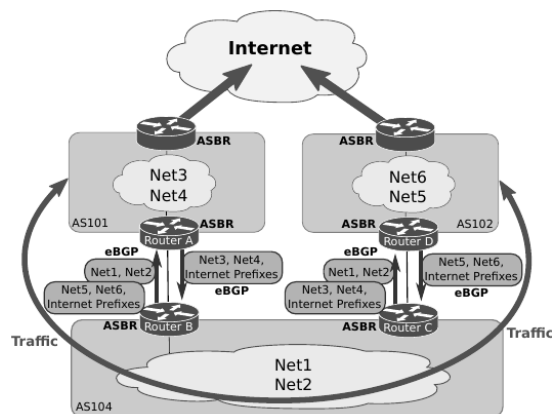


Figura 1.4: Multi-homed Transit AS

1.5 Path-vector

O BGP é um protocolo **Path-vector**, ou seja, utiliza um path-vector para que os routers saibam as rotas de todos os outros.

É essencialmente um distance-vector que transporta uma lista das AS atravessadas pela rota, fornecendo detecção de loops.

- ⇒ Um speaker **EBGP** adiciona as suas AS, a esta lista, antes de encaminhar a rota para outro peer **EBGP**.
- ⇒ Um speaker **IBGP** não modifica a lista porque está a enviar a rota para um peer da mesma AS, por isso a lista das AS não deve ser usada para detectar loops **IBGP**.

1.6 Tipos de Mensagem BGP

- * **OPEN** - serve para estabelecer a sessão BGP.
- * **UPDATE** - para informar atualizações de rotas para outros sistemas BGP, permitindo que os routers possam construir uma visão consistente da topologia da rede, usando o TCP para garantir uma entrega confiável.
- * **KEEPALIVE** - estas mensagens são trocadas quando o período de keepalive é excedido, sem a troca de um UPDATE.
- * **NOTIFICATION** - são enviadas quando é detectado um erro de protocolo. Nesse caso a sessão BGP é fechada.

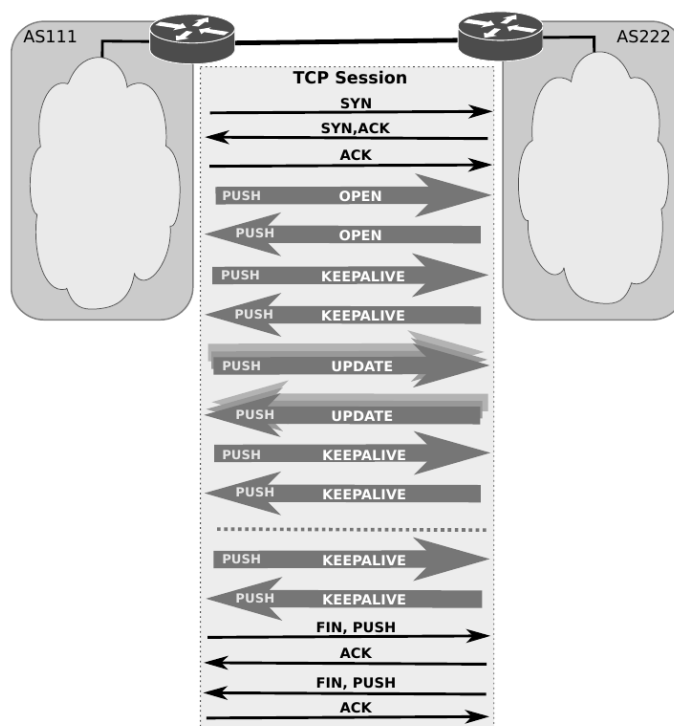


Figura 1.5: Exemplo de sessão BGP

1.7 Atributos BGP

É a métrica usada para descrever as características do caminho BGP. Os atributos estão contidos nas mensagens de update. Existem 4+1 categorias:

1. **Well-known Mandatory** (incluídas em BGP updates)
 - (a) AS-PATH 1.7.1
 - (b) Origin 1.7.2
 - (c) Next-hop 1.7.3
2. **Well-known Discretionary** (podem ou não estar incluídas mas msgs BGP update)
 - (a) Local Preference 1.7.4
 - (b) Atomic Aggregate 1.7.8.1
3. **Optional Transitive** (podem não ser suportadas por algumas implementações BGP)
 - (a) Aggregator 1.7.8.2
 - (b) Community 1.7.7
 - (c) AS4_aggregator
 - (d) AS4_path
4. **Optional Non-Transitive** (podem não ser suportadas por algumas implementações BGP)
 - (a) Se o vizinho não o suportar, o atributo é apagado
 - (b) Multi-exit-discriminator(MED) 1.7.5
5. **Cisco-defined** (local to router, não é anunciado)
 - (a) Peso 1.7.6

1.7.1 AS-PATH

Quando um anúncio de rota passa por uma AS, o n^o do AS é adicionado à lista de AS que a rota de anúncio percorreu.

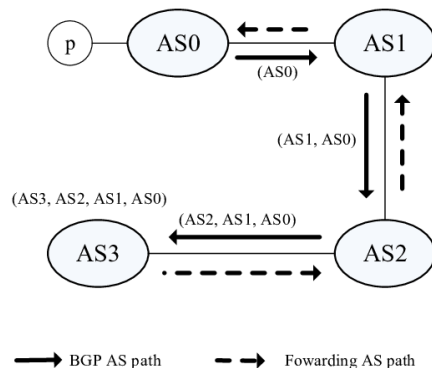


Figura 1.6: AS-PATH [2]

1.7.2 Origin

Indica como é que o BGP aprendeu uma certa rota em particular. Pode ser do tipo: IGP(0), EGP(1), Incomplete((2) origem desconhecida).

1.7.3 Next Hop

O atributo EBGP next-hop é o endereço IP que é usado para atingir a rota anunciada.

⇒ Para o EBGP, o endereço next-hop é o IP da ligação entre os peers.

⇒ Para o IBGP, o endereço next-hop EBGP é carregado na AS local.

1.7.4 Local Preference

É utilizado para escolher um ponto de saída de uma AS local.

* O **Maior Valor** é preferido

Este valor é propagado por toda a AS local. Se existirem múltiplos exit points, o atributo é usado para escolher o exit point de uma rota específica.

1.7.5 Multi-Exit Discriminator Attribute (MED)

É usado como uma sugestão para uma AS externa. A AS externa que recebe o MED pode usar outros atributos **BGP** para selecionar a rota.

* O **Menor Valor** é preferido

MED é desenhado para influenciar chegada de trafego.

1.7.6 Peso

É um definido pela Cisco que não é anunciado aos routers vizinhos. Se o router tiver mais que uma rota pro mesmo destino, a rota com o **maior** peso é preferida.

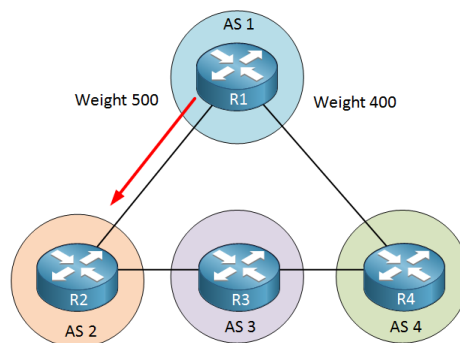


Figura 1.7: Esquema atributo Peso [3]

1.7.7 Community Attribute

Usado para agrupar conjuntos de rotas, que partilham certas propriedades em comum, para poder aplicar certas políticas a nível de cada grupo.

Os atributos predefinidos são:

1. **No-Export** – Não anunciar esta rota aos pares EBGp
2. **No-Advertise** – Não anunciar esta rota a nenhum peer
3. **Internet** – Anunciar a toda a Internet

O formato é Asnumber:Cnumber (300:1, 200:8)

1.7.8 Atomic Aggregate and Aggregator Attributes

1.7.8.1 Atomic Aggregate

É usado para alertar routers que rotas específicas foram agregadas numa rota menos específica.

1.7.8.2 Agregator

Fornece informação sobre qual AS fez a agregação e o IP do router que originou o tal agregado.

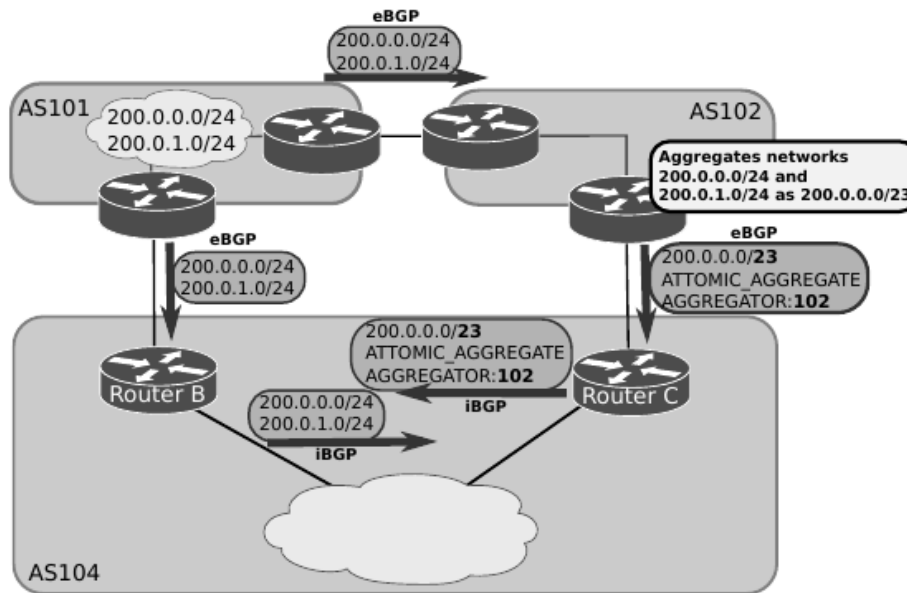


Figura 1.8: Esquema dos atributos Atomic Aggregate e Aggregator Attributes

1.8 Seleção do caminho BGP

O BGP pode receber vários anúncios da mesma rota de múltiplas fontes. Nessa situação, o BGP, seleciona um caminho como o melhor.

O BGP coloca o caminho selecionado na tabela de routing e propaga-o aos vizinhos. O critério utilizado para escolher o melhor caminho é, por ordem:

1. Maior Peso (apenas na Cisco);
2. Maior Local preference;
3. Caminho que é originado localmente;
4. Caminho mais curto;
5. Menor tipo de origem (IGP < EGP < Incompleto);
6. MED mais baixo;
7. Prefere o caminho externo ao caminho interno;
8. Vizinho IGP mais próximo.

1.9 Multi-Protocol Border Gateway Protocol (MP-BGP)

É uma extensão do **BGP** que encaminha informação de routing sobre outros protocolos tais como:

- Ipv6 Unicast;
- Multicast (Ipv4 e Ipv6);
- 6PE – Ipv6 sobre Ipv4 MPLS backbone;
- Multi-Protocol Label Switchin (MPLS) VPN;
- Multi-Protocol Reachability Information NLRI.

1.9.1 Atributos MP-BGP

1. Novos atributos non-transitive e opcionais

(a) MP_REACH_NLRI

⇒ Transporta o conjunto de destinos alcançáveis juntamente com as informações do Next-hop a serem usadas para o encaminhamento para esses destinos.

(b) MP_UNREACH_NLRI

⇒ Transporta o conjunto de destinos inalcançáveis.

2. Atributos que contêm um ou mais triples

(a) Address Family Information (AFI)

⇒ Identifica as informações do protocolo transportadas na NLRI

(b) Informação sobre o Next-hop

⇒ O endereço do Next-hop deve ser da mesma família.

3. Reachability information

1.9.2 MP-BGP Negotiation Capabilities

Os routers MP-BGP estabelecem sessõesBGP com a menssagem OPEN, só que agora este tipo de mensagem contém um novo parâmetro opcional chamado CAPABILITIES que contém:

- Multi-Protocol extensions (AFI/SAFI)
- Route Refresh
- Outbound Route Filtering

1.10 BGP Avançado

Os números de sistemas **AS Privadas** variam de 64512 a 65535 (não é aconselhado o uso de um número de AS privado se se planear conectar a vários ISP no futuro.).

Quando uma rede de cliente é grande, o ISP pode atribuir um número de AS, atribuindo, permanentemente, um número de **AS Público** no intervalo de 1 a 64511. Deve ter um número de AS único para propagar suas rotas BGP para a Internet.

É feito, também, quando uma rede de cliente se conecta a dois ISP's diferentes, como multihoming.

1.10.1 Sincronização BGP

Se uma dada AS encaminha tráfego de outro AS para uma terceiro AS, o BGP não deve anunciar uma rota antes que todos os routers na seu AS tenham aprendido sobre a rota através de um IGP. O BGP espera até que o IGP tenha propagado a rota dentro da AS. Em seguida, o BGP anuncia a rota para os peers externos.

1.10.2 BGP Route Reflectors

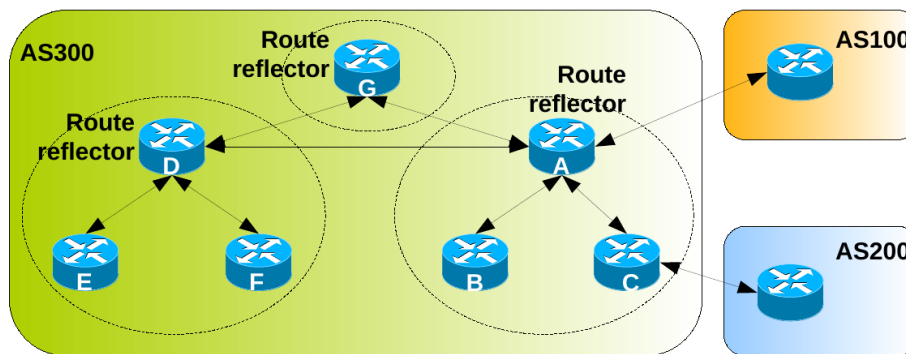


Figura 1.9: Esquema Route Reflectors

Se não fosse pelos Route Reflector G, A e D, a rede, iria de requerer de uma iBGP mesh completa dentro do AS300. O Route Reflector e seus clientes são chamados de **cluster**.

- ⇒ O router A é configurado como um Route Reflector, não é necessário estabelecer uma sessão iBGP entre os routers B e C (e outros).
- ⇒ O router D é configurado como um Route Reflector, não é necessário estabelecer uma sessão iBGP entre os routers E e F (e outros).

1.10.3 BGP Route Redistribution

⇒ A redistribuição de rotas IGP pelo BGP irá:

1. Simplificar a configuração do BGP **Vantagem**
2. O BGP irá anunciar apenas redes internas com conectividade **Vantagem**

⇒ A redistribuição de rotas BGP pelo IGP irá:

1. Tornar as rotas internas conhecedoras de todas as rotas externas **Vantagem/Desvantagem ???**
2. Aumentar o tamanho das tabelas de routing nos routers internos **Desvantagem**
3. Evitar o uso de rotas padrão internas **Vantagem/Desvantagem ???**

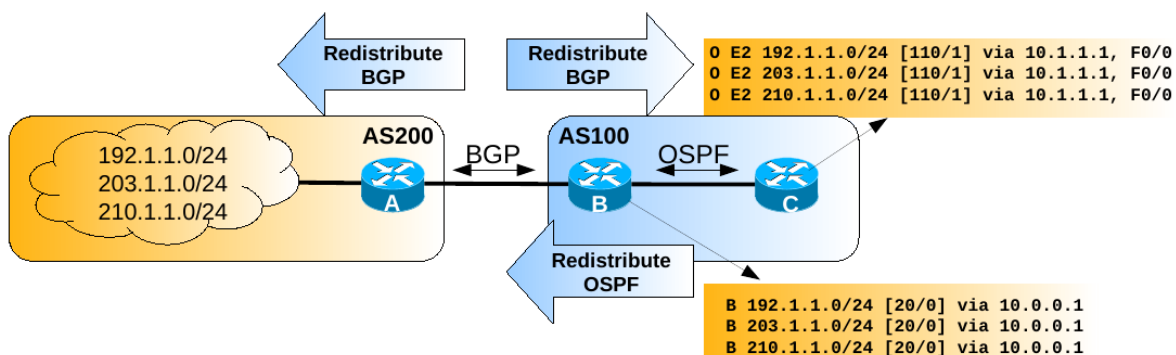


Figura 1.10: Esquema Route Redistribution

1.10.4 BGP Neighborhood Resilience

As relações de vizinhança BGP entre interfaces físicas dependem da estabilidade/estado da interface. As relações de vizinhança (virtuais) que utilizam interfaces/addresses de Loopback são independentes das interfaces físicas.

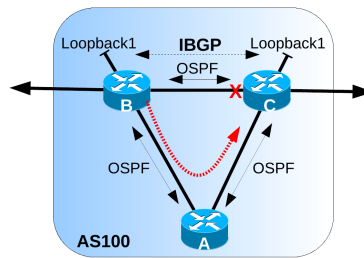


Figura 1.11: BGP Neighborhood Resilience

As interfaces de Loopback são virtuais e baseadas em software. Quando o router está ativo, as interfaces de Loopback estão sempre ativas. A relação de vizinhança é ativa enquanto existir um caminho entre as redes virtuais. O encaminhamento (alternativo) é fornecido pelos IGP's.

1.10.5 Conflitos BGP e IGP

- Conflitos de routing podem surgir com:
 - Routers internos sem BGP
 - Falta de redistribuição de rotas BGP pelo IGP
 - Rotas padrão IGP
 - Rotas preferenciais BGP (sem acordo com rotas padrão IGP)
- Soluções:
 - Ajustar rotas padrão IGP
 - Ajustar rotas preferenciais BGP (por exemplo, com preferência local)
 - Vizinhança BGP e roteamento interno via túneis IP-IP

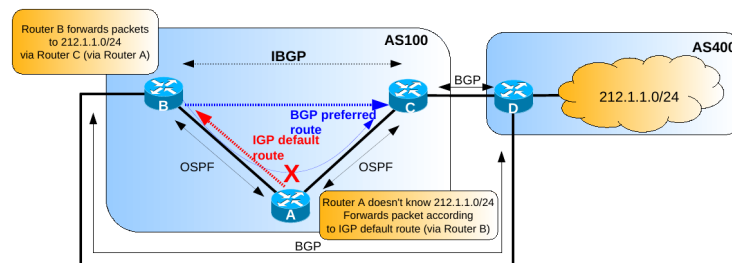


Figura 1.12: Conflitos BGP e IGP

1.10.6 BGP over Tunnels (over IGP)

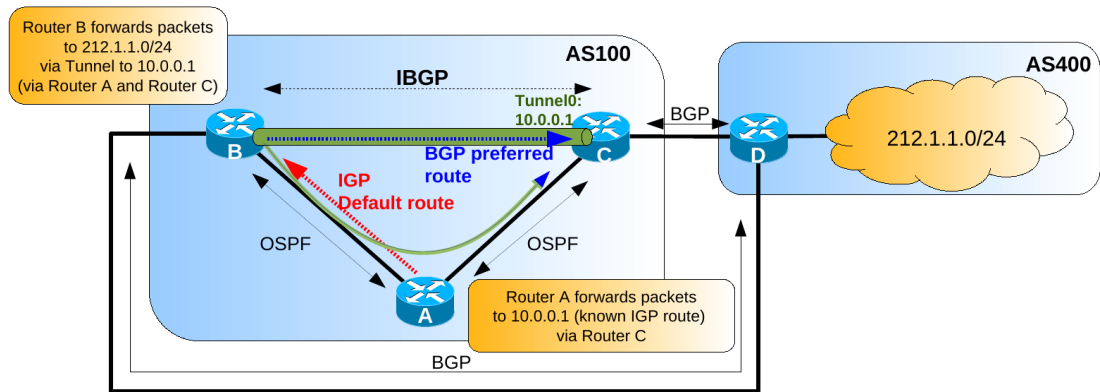


Figura 1.13: BGP over Tunnels (over IGP)

- Utilização de túneis IP-IP para resolver conflitos de roteamento BGP/IGP:
 - Túneis configurados manualmente:
 - * Entre interfaces físicas ou Loopback.
 - Vizinhança BGP via túnel.
 - Rotas BGP aprendidas via túnel (o Next-Hop é o ponto final remoto do túnel).
 - "Rede" do túnel distribuída internamente via IGP.
- No Router A, para qualquer pacote destinado a uma rede externa, é encaminhado via túnel:
 - Um novo cabeçalho IP é adicionado, o novo endereço de destino IP é o ponto final remoto do túnel.
 - Internamente, o pacote é roteado de acordo com o novo cabeçalho IP (endereços IP dos pontos finais do túnel).

1.10.7 BGP Filtering and Route Maps

O envio e recepção de atualizações BGP podem ser controlados através da utilização de vários métodos de filtragem diferentes. As atualizações BGP podem ser filtradas com base em informação de rota, informação de caminho e comunidades.

Os Mapas de rota são utilizados com o BGP para:

- Controlar e modificar informações de encaminhamento.
- Definir as condições pelas quais as rotas são redistribuídas entre domínios de encaminhamento.

1.10.8 BGP Case Studies

Ver nos slides 🐱🐱🐱

Bibliografia

- [1] PyNetLabs, *What is BGP?*, <https://www.pynetlabs.com/bgp-vs-ospf-whats-the-difference/>.
- [2] S. Zhang, *An example of BGP AS path and forwarding AS path*, https://www.researchgate.net/figure/An-example-of-BGP-AS-path-and-forwarding-AS-path_fig1_286582450.
- [3] NetworkLessons.com, *How to Configure BGP Weight Attribute*, <https://networklessons.com/bgp/how-to-configure-bgp-weight-attribute>.