nic.br

Núcleo de Informação e Coordenação do Ponto BR

cgi.br

Comitê Gestor da Internet no Brasil

registro.br cert.br cetic.br ceptro.br ceweb.br ix.br


nic.br cgi.br
ceptro.br

Curso BCOP

Boas Práticas BGP

ceptro.br nic.br cgi.br

Licença de uso do material

Esta apresentação está disponível sob a licença

Creative Commons

Atribuição – Não a Obras Derivadas (by-nd)

http://creativecommons.org/licenses/by-nd/3.0/br/legalcode

Você pode:

- **Compartilhar** — copiar, distribuir e transmitir a obra.
- **Fazer uso comercial da obra.**
- Sob as seguintes condições:

Atribuição — Ao distribuir essa apresentação, você deve deixar claro que ela faz parte do Curso de Formação para Sistemas Autônomos do CEPTRO.br/NIC.br, e que os originais podem ser obtidos em http://ceptro.br. Você deve fazer isso sem sugerir que nós damos algum aval à sua instituição, empresa, site ou curso.

Vedada a criação de obras derivadas — Você não pode modificar essa apresentação, nem criar apresentações ou outras obras baseadas nela..

Se tiver dúvidas, ou quiser obter permissão para utilizar o material de outra forma, entre em contato pelo e-mail: info@nic.br.

Conceitos de BGP

- BGP é um protocolo fofoqueiro!

Minha Rota
2001:db8::/32

Rota Recebida
2001:db8::/32 65537 65536 i

Conceitos de BGP

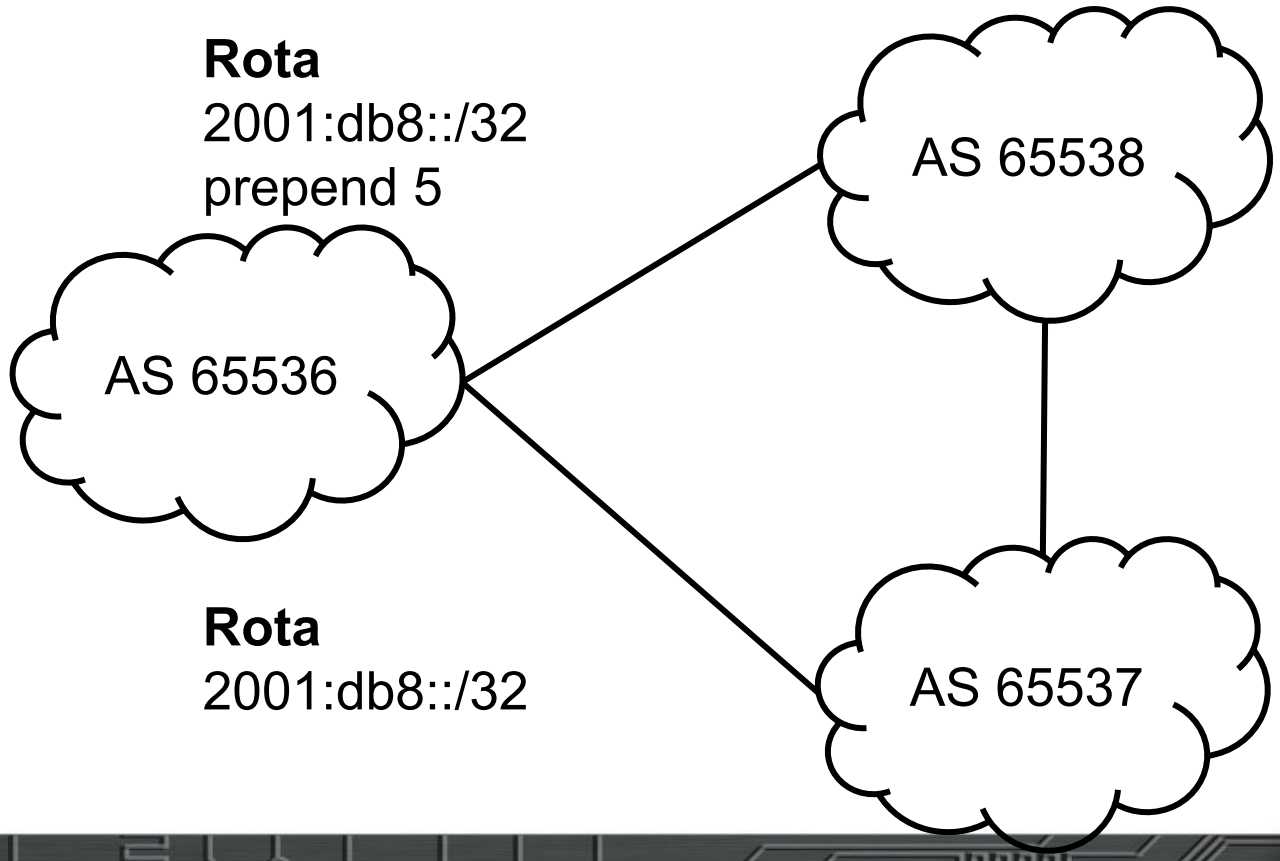
- Quem conta o conto pode aumentar um ponto!

Minha Rota
2001:db8::/32

Rota Recebida
2001:db8::/16 65537 i

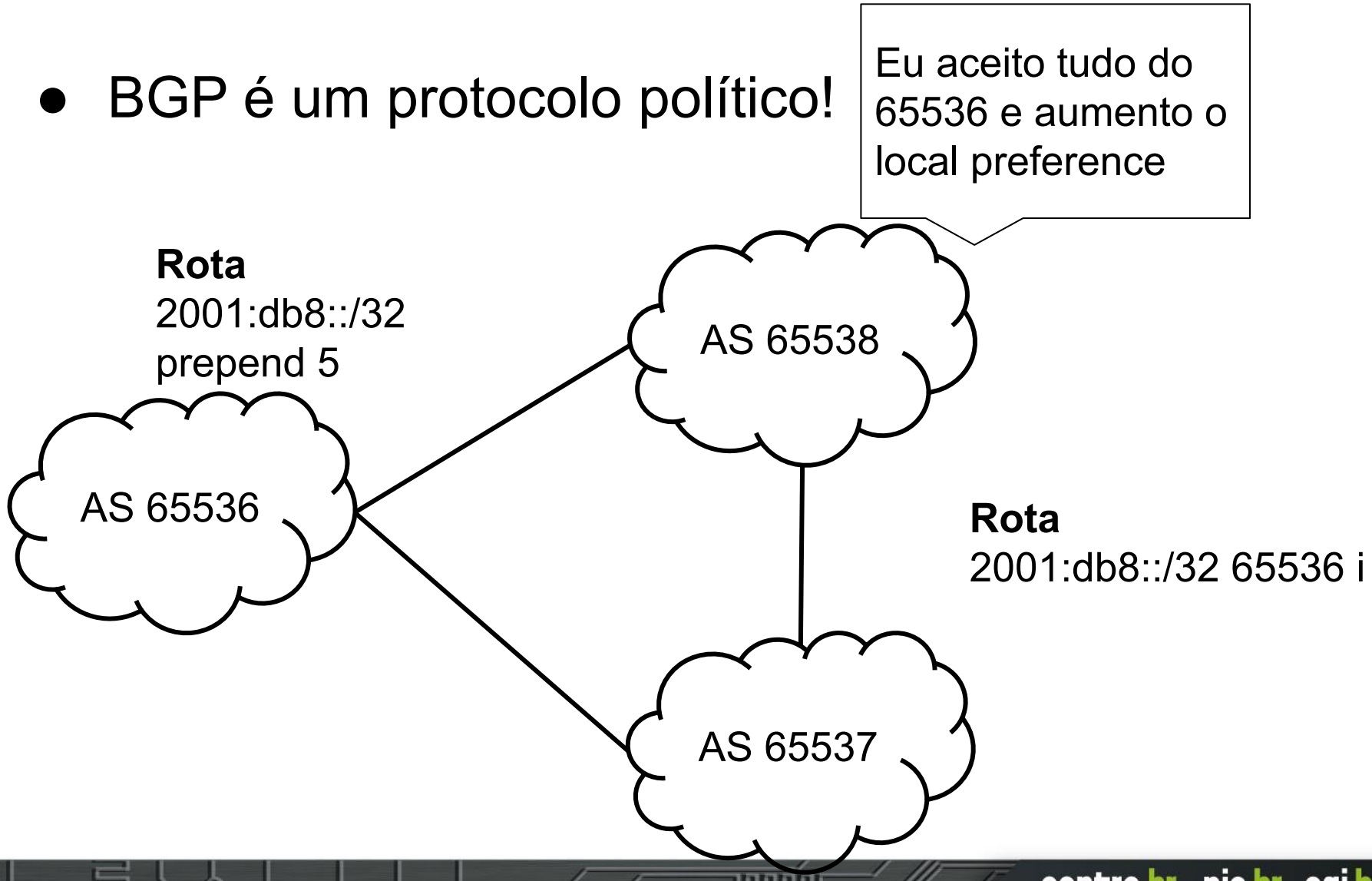
Conceitos de BGP

- BGP é um protocolo político!



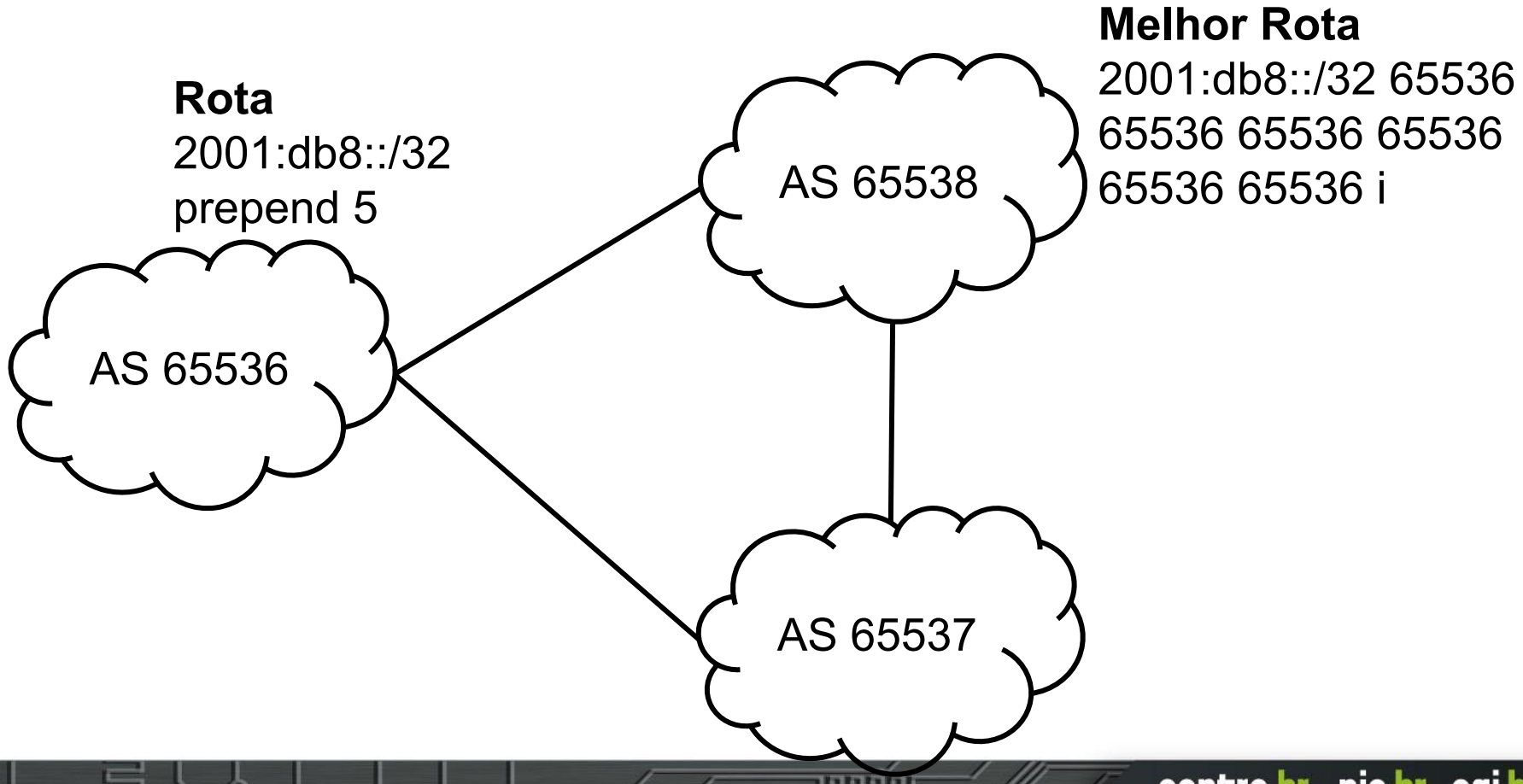
Conceitos de BGP

- BGP é um protocolo político!



Conceitos de BGP

- BGP é um protocolo político!



Entendendo o BGP

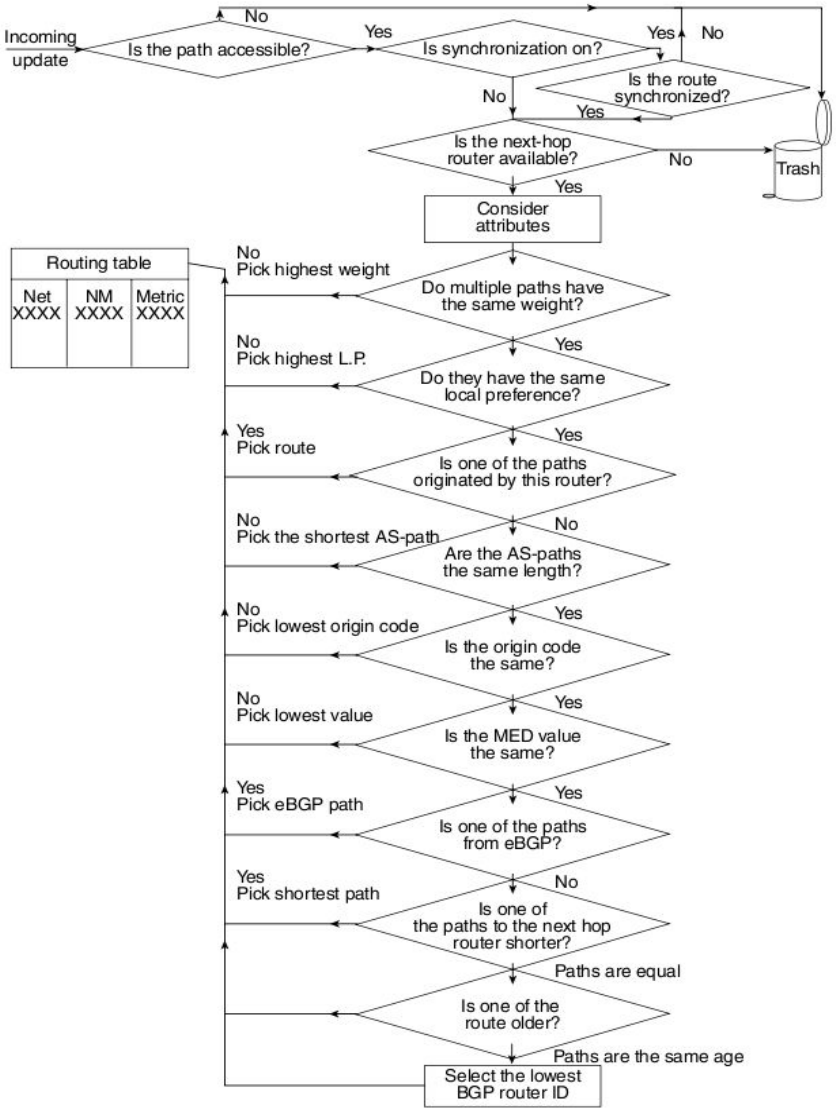
- É importante sempre conversar com outros sistemas autônomos.

- Não adianta configurar e largar!

- Precisa monitorar e entender que seu funcionamento depende de todos.

Atributos BGP

- Os atributos são considerados na seleção dos caminhos
  - Se este for conhecido, acessível e se o next hop estiver disponível
  - A forma de seleção pode variar com a implementação do BGP



Atributos BGP

- Bem conhecidos
  - Todas as implementações BGP os reconhecem
    - Mandatórios: sempre estão presentes nos updates que carregam informações de prefixos (NLRI Network Layer Reachability Information)
    - Discricionário: não estão em todos os updates
- Opcionais
  - Não são suportados por todas as implementações BGP
    - Transitivos: devem ser repassados pelos updates, mesmo que não sejam suportados
    - Não transitivos: não devem ser repassados pelos updates

Atributos novos no MP-BGP

- O **MP-BGP** é **necessário** para suportar **IPv6**
- **Dois atributos novos**:
  - **Multiprotocol Reachable NLRI** (MP_REACH_NLRI): carrega o conjunto de destinos alcançáveis junto com as informações do next-hop;
  - **Multiprotocol Unreachable NLRI** (MP_UNREACH_NLRI): carrega o conjunto de destinos inalcançáveis;
  - Estes atributos são **Opcionais e Não-Transitivos**.

Atributos BGP (AS Path)

- **Bem conhecido e mandatório**

- **Indica o caminho para se chegar a um destino**, incluindo todos os ASes intermediários

- É **usado** para:
  - **Detectar loops**
  - **Aplicar políticas**

Atributos BGP (MED)

- **M**ulti-**E**xit **D**iscriminator
- **Opcional** e **não transitivo**
- Indica para os **vizinhos BGP externos qual o melhor caminho** para uma determinada rota do AS, influenciando o tráfego de entrada
- O **menor MED ganha**
- **Ausência** de MED implica **MED=zero**

Atributos BGP (Local Preference)

- **Bem conhecido e discricionário**
- O **valor** pode ser **associado a uma rota**, indicando o caminho preferencial de saída.
- O caminho com a **maior Local Preference ganha**.
- Só vale **dentro do AS**

Weight

- Não é um atributo (**é local para o roteador**)
- O **maior weight ganha**
- Pode ser aplicado as rotas aprendidas de um dado vizinho, ou por meio de filtros
- Influencia o **tráfego de saída**

Estabelecendo uma sessão BGP

- As **sessões BGP** utilizam **TCP**, na **porta 179**
- Pode-se utilizar **IPv4 ou IPv6**
- Recomenda-se o uso de **interfaces loopback**
  - **Loopbacks são interfaces lógicas**
  - **Elas não "caem"**. São **independentes** de **problemas com links**.

Loopback no iBGP

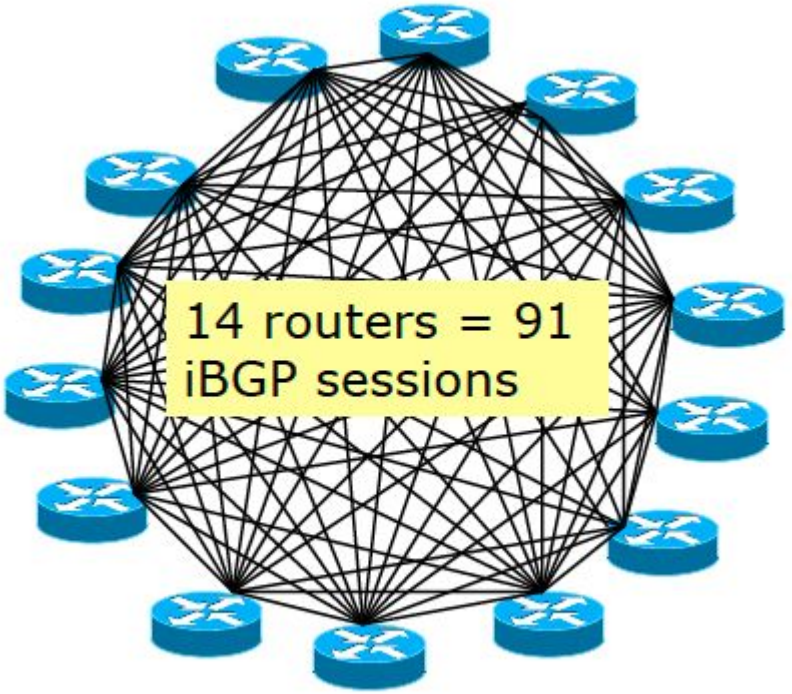
- No **iBGP** devemos **sempre usar interfaces loopback**
  - **Usando interfaces físicas**, **se o link for interrompido**, a **sessão BGP** também **será**
  - Usando **loopbacks** temos **uma estabilidade maior.**
    - Como as rotas para os **IPs das loopbacks são aprendidos via IGP**, se um enlace for interrompido, a sessão contínua estabelecida, com os pacotes fazendo um caminho alternativo.

Uma loopback por serviço

- **Um dos benefícios do uso das loopbacks é a possibilidade de separar os serviços e protocolos** em um dado roteador:
  - Cada qual usa uma **loopback e IP próprios**
  - A prática pode **facilitar a migração de serviços** entre diferentes roteadores
  - O **lado negativo é o maior consumo de IPs** na infraestrutura

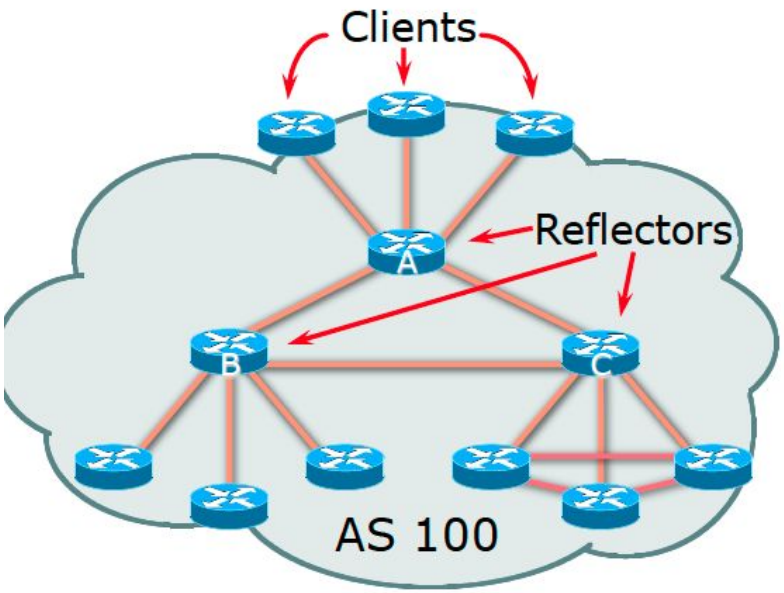
Escalando o iBGP

- **iBGP operando em mesh**
  - Quantidade de sessões = n(n-1)/2 onde n é número de roteadores.
  - Use Route Reflector
    - Clientes
    - Refletores
  - Confederation
    - Muito complexo



Escalando o iBGP

- Route Reflector
  - Refletor recebe rotas de todo mundo
    - Seleciona o melhor caminho
      - Se o melhor caminho for de um cliente, reflete para todos
      - Se o melhor caminho for de um refletor, reflete somente para os clientes



Autenticando sessões BGP com MD5

- **É recomendável usar autenticação MD5 para as sessões BGP**
  - A **configuração é simples**: os roteadores vizinhos compartilham **uma mesma chave (uma senha)**
  - **A cada pacote é adicionado um checksum codificado**, que o outro roteador pode verificar utilizando sua chave MD5, **ajudando a garantir sua autenticidade e integridade**
  - **A técnica dificulta ataques**

```
- neighbor “ip-address ou peer-group-name” password
“senha” (Cisco)
- authentication-key “senha” (Juniper)
```

TTL Security Check

- **Por padrão**, os pacotes das **sessões eBGP** são enviados com valor de **TTL/Hop-Limit igual a 1**, buscando garantir que quem está enviando o pacote é um vizinho diretamente conectado. **Porém um atacante externo pode facilmente forjar um pacote com TTL/Hop-Limit igual a 1 no enlace**.

- O TTL Security Check é uma ideia bastante simples e engenhosa:
  - O roteador envia pacotes com **TTL/Hop-Limit igual a 255** (**valor máximo desse campo**).

  - **No próximo roteador**, o valor será decrementado, **e igual a 254** (255-1).

    - **Um atacante em outra rede não conseguirá** inserir um pacote com **TTL/Hop-Limit igual a 255** no enlace. Exemplo Cisco:

```
neighbor 2001:DB8:200:FFFF::255 ttl-security hops 1
```

Desabilitando serviços e protocolos

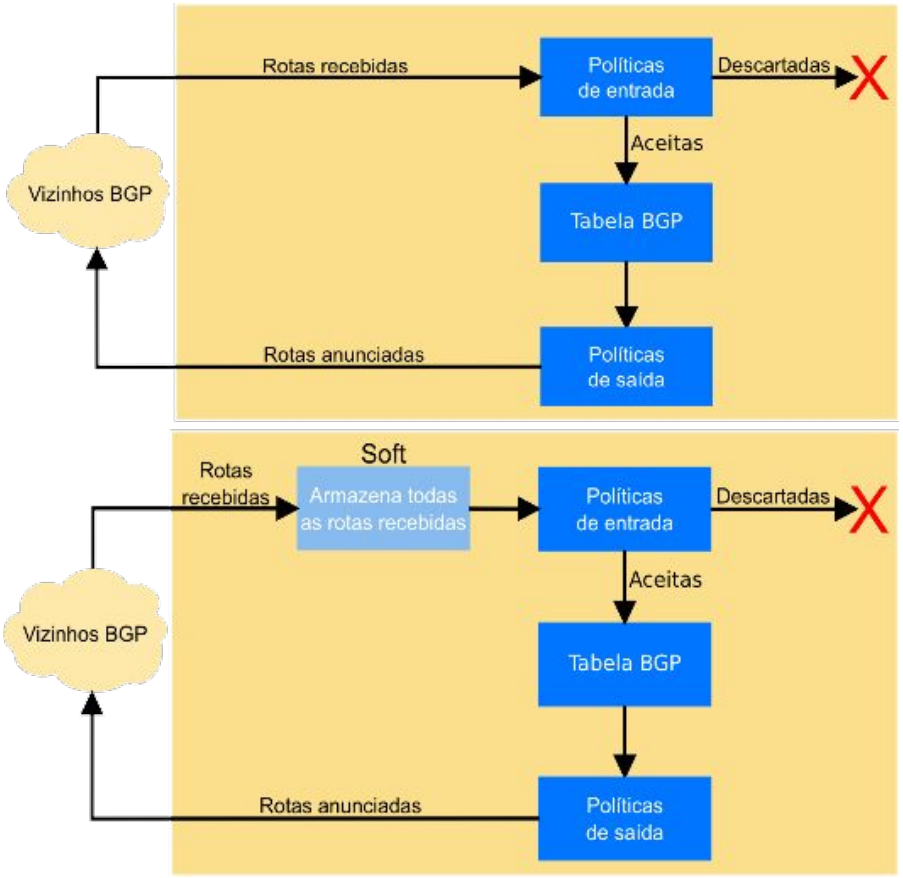
- **Nas interfaces** onde são estabelecidas **sessões eBGP** é fundamentalmente que todos os **serviços e protocolos desnecessários estejam desabilitados**, de forma particular:
  - **IGP (OSPF / IS-IS)**
  - **Router Advertisement (RA) no IPv6**

Route refresh

- **Solicita que o vizinho cujas rotas são afetadas por uma mudança de política, reenvie toda a informação pertinente.**
  - Isso se chama Route Refresh
  - Não usa memória
  - Não necessita de configuração extra
  - Maioria dos roteadores suportam
- **Essa capacidade é informada no estabelecimento de uma sessão BGP e é possível verificá-la olhando as informações do vizinho.**
- **Após uma mudança em um filtro é preciso solicitar o refresh para o roteador vizinho, com um comando. Isso não é automático!**
  - /routing bgp peer refresh <numbers>

Soft Reconfiguration Inbound

- Antigamente era uma boa prática!
- Habilitando “soft reconfiguration”, é criada uma nova tabela, com a informação original.
  - Isso consome mais memória
  - Permite que filtros sejam modificados facilmente
- Serve para troubleshooting
  - É possível saber o que foi enviado antes de se aplicar os filtros



Filtros

- **Alguns roteadores são permissivos e, se nenhum filtro for aplicado, aceitam tudo que os vizinhos enviam.**

- **É uma boa prática aplicar filtros de entrada e saída para cada vizinho, ANTES de estabelecer qualquer sessão eBGP.**

Filtros de entrada

- **Clientes**
  - Deve-se aceitar apenas os prefixos que foram designados (por você mesmo) ao cliente, ou alocados a ele pelo NIC.br ou por um RIR.
- **Fornecedores de trânsito (upstreams)**
  - Você paga seu fornecedor de trânsito para que ele forneça acesso à toda a Internet (full routing ou rota default).
- **Peers (com quem realizamos troca de tráfego)**
  - Deve-se combinar antes que prefixos serão anunciados ou aceitos.
  - No caso sessões BGP, em um acordo ATM no IX, deve-se receber todos os prefixos anunciados, com as seguintes exceções:
    - **Se você têm clientes de trânsito no IX**, deve-se filtrar os prefixos deles. Assim evita-se que o tráfego na direção do cliente passe pelo IX, no lugar de passar no link de trânsito
    - **Se você têm upstreams no IX**, pode ser desejável filtrá-los, forçando o tráfego a fluir pelo link de trânsito em ambas as direções, e evitando assimetrias.

Filtros de entrada

- Verifique a lista do bogons (prefixos que não deveriam aparecer no BGP), do Team Cymru:
  - www.team-cymru.org/Services/Bogons/http.html
- **Para IPv4**
  - É preciso lembrar que **não há mais endereços reservados para alocações futuras**. Deve-se remover todos os filtros baseados no status dos blocos nos RIRs. Ver:
  - http://tools.ietf.org/html/rfc6441
- **Para IPv6**
  - **Você pode bloquear tudo por padrão e permitir apenas o 2000::/3, ou os prefixos mais específicos /12 e /23 sob responsabilidade de cada RIR.** Alguns bogons podem estar dentro do espaço dos RIRs, então também devem ser bloqueados explicitamente.
- **Feed automático de bogons**:
  - http://www.team-cymru.org/Services/Bogons/routeserver.html

Filtros de entrada

- Aplicando **corretamente os filtros**, você ajuda a:
  - **Garantir a integridade da sua própria rede**
  - **Garantir a integridade de toda a Internet**
- **É responsabilidade** de cada **Sistema Autônomos** ser um **bom cidadão da Internet**!!!

Prefixos no iBGP e eBGP

- O iBGP deve ser usado para transportar os prefixos de seus clientes/usuários. Não use OSPF ou outro IGP.
  - Crie uma rota estática para a interface do cliente (ou agregador).
  - Use “bgp network” para originar o prefixo no iBGP
  - O prefixo existirá enquanto a rota estática existir e a interface estiver ativa.

- Esses prefixos não são exportados no eBGP. No eBGP devem estar presentes apenas os prefixos agregados, mais aquele necessários para engenharia de tráfego.
  - Os prefixos usados para engenharia de tráfego não dependem daqueles presentes no iBGP. Os prefixos presentes no iBGP não devem ser exportados para o eBGP.
  - Os prefixos usados para engenharia de tráfego devem ser gerados na borda da rede, com rotas estáticas para null e comandos do tipo “bgp network”.

Communities

- Descritas na RFC 1997
  - São um atributo Opcional e Transitivo
  - Cada community é um número inteiro de 32 bits, representada por dois inteiros de 16 bits (RFC 1998)
    - Um formato comum de representação é <ASN>:nn
    - 0:0 até 0:65535 e 65535:0 até 65535:65535 são valores reservados
- Communities são usadas para agrupar destinos
  - Pode-se marcar um grupo de caminhos aprendidos, ou a exportar, com uma determinada community, de acordo com filtros
  - Pode-se filtrar rotas, ou modificar outros atributos, segundo às communities a qual a rota pertence
- São úteis para aplicar políticas tanto dentro do AS, quanto entre diferentes ASes

Communities

- www.iana.org/assignments/bgp-well-known-communities
  - no-export 65535:65281
    - Não anuncie essa rota para nenhum peer e-BGP
  - no-advertise 65535:65282
    - Não anuncie para nenhum peer BGP
  - no-peer 65535:65284
    - Não anuncie para tráfego bilateral

Dúvidas?

Obrigado !!!

nic.br cgi.br

www.nic.br | www.cgi.br