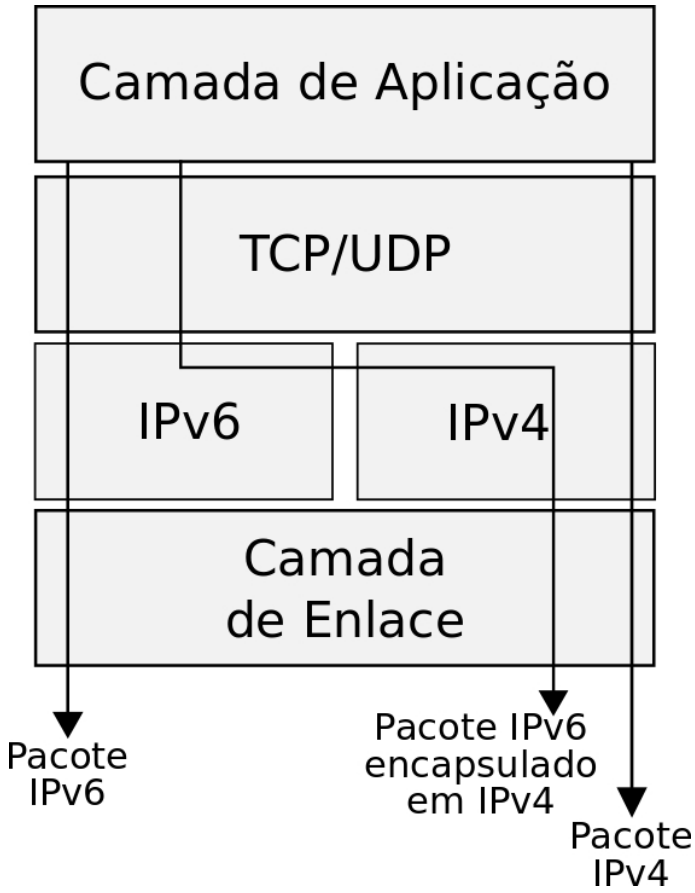
# IPv6.br

## A Nova Geração do Protocolo Internet

# Roteamento IPv6

# Considerações Importantes

- IPv4 e IPv6 → Camada de Rede
- Duas redes distintas
  - Planejamento
  - Suporte
  - *Troubleshooting*
  - Arquitetura dos equipamentos
  - ...



# Considerações Importantes

Características Fundamentais do Endereço IP

- Identificação
  - Unívoca
  - Comandos: `host`, `nslookup`, `dig...`
- Localização
  - Roteamento e encaminhamento entre a origem e o destino
  - Comandos: `mtr -4/-6`, `traceroute(6)`, `tracert(6)`...

Semântica Sobrecarregada

- Dificulta a mobilidade
- Desagregação de rotas

# Considerações Importantes

Separar as funções de localização e identificação.

- LISP (*Locator/Identifier Separation Protocol*).
  - Permite uma implementação de forma gradual.
    - não exige nenhuma alterações nas pilhas dos *host* e nem grandes mudanças na infraestrutura existente.
  - EID (*Endpoint Identifiers*).
  - RLOC (Routing Locators).
  - ITR (*Ingress Tunnel Router*) / ETR (*Egress Tunnel Router*).
    - Fazem o mapeamento entre EID e RLOC.
  - Utiliza tanto IPv4 quanto IPv6.

# Considerações Importantes

Prefixo IP

- O recurso alocado pelo Registro.br ao AS é um bloco IP.
- O bloco IP não é roteável.
  - bloco é um grupo de IPs.
- O prefixo IP é roteável.
  - número de bits que identifica a rede;
  - você pode criar um prefixo /32 igual ao bloco /32 IPv6 recebido do Registro.br;
  - pode criar um prefixo /33, /34,... /48.
- Esta nomenclatura é importante.
  - ativação de sessões de transito com outras operadoras;
  - *troubleshooting*.

# Como o roteador trabalha?

Ex.:

1. O roteador recebe um quadro Ethernet;
2. Verifica a informação do Ethertype que indica que o protocolo da camada superior transportado é IPv6;
3. O cabeçalho IPv6 é processado e o endereço de destino é analisado;
4. O roteador procura na tabela de roteamento *unicast* (RIB - *Router Information Base*) se há alguma entrada para a rede de destino;
   - Visualizando a RIB:
     `show ip(v6) route` → Cisco/Quagga
     `show route (table inet6)` → Juniper

# Como o roteador trabalha?

5. *Longest Match* - procura a entrada mais específica. Ex.:

- O IP de destino é 2001:0DB8:0010:0010::0010
- O roteador possui as seguintes informações em sua tabela de rotas:
    - 2001:DB8::/32 via interface A
    - 2001:DB8::/40 via interface B
    - 2001:DB8:10::/48 via interface C
- Os três prefixos englobam o endereço de destino, porém o roteador sempre irá preferir o mais específico, neste caso, o /48;
- Qual é a entrada mais específica IPv4 e IPv6?

6. Uma vez identificado o prefixo mais específico, o roteador decrementa o *Hop-Limit*, monta o quadro Ethernet de acordo a interface, e envia o pacote.

# Como o roteador trabalha?

E se houver mais de um caminho para o mesmo prefixo?

- Utiliza-se uma tabela predefinida de preferências.
  - número inteiro entre 0 e 255 associado a cada rota, sendo que, quanto menor o valor mais confiável é a rota;
  - avalia se está diretamente conectado, se a rota foi aprendida através do protocolo de roteamento externo ou interno;
  - tem significado local, não pode ser anunciado pelos protocolos de roteamento;
  - seu valor pode ser alterado caso seja necessário priorizar um determinado protocolo.

E se o valor na tabela de preferências também for o mesmo?

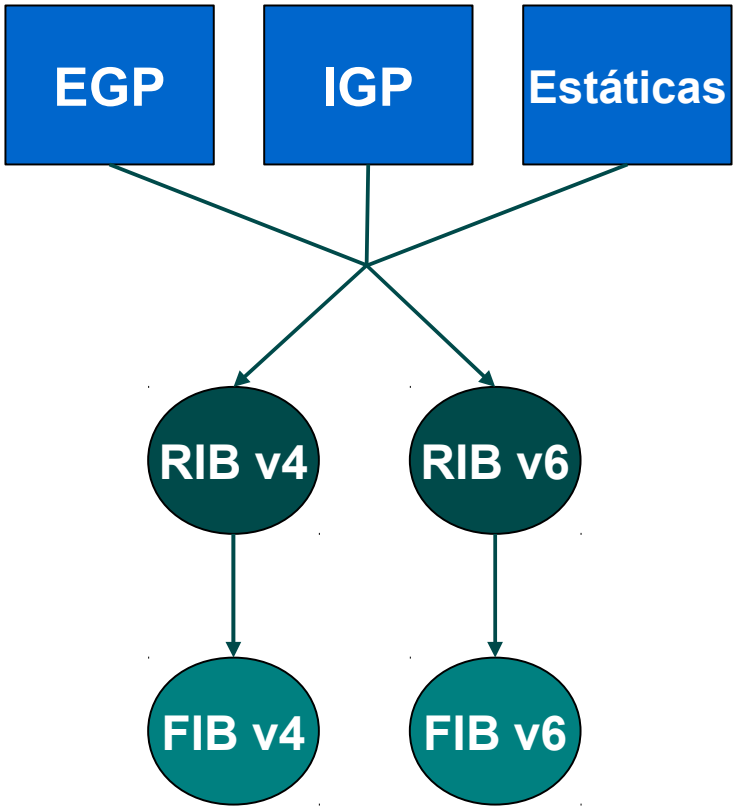
# Tabela de Roteamento

- O processo de escolha das rotas é idêntico em IPv4 e IPv6, porém, as tabelas de rotas são independentes.
    - Há uma RIB IPv4 e outra IPv6.
- Através de mecanismos de otimização as melhores rotas são adicionadas à tabela de encaminhamento
    - FIB - *Forwarding Information Base;*
    - A FIB é criada a partir da RIB;
    - Assim como a RIB, a FIB também é duplicada.
- Em roteadores que possuem arquitetura distribuída o processo de seleção das rotas e o encaminhamento dos pacotes são funções distintas.

# Tabela de Roteamento

- São as informações recebidas pelos protocolos de roteamento que “alimentam” a RIB que por sua vez “alimenta” a FIB.

- Os Protocolos de Roteamento se dividem em dois grupos:
  - **Interno (IGP)** - protocolos que distribuem as informações dos roteadores dentro de Sistemas Autônomos. Ex.: OSPF; IS-IS; RIP.
  - **Externo (EGP)** - protocolos que distribuem as informações entre Sistemas Autônomos. Ex.: BGP-4.



# Rota Default

- Quando um roteador não encontra uma entrada na tabela de rotas para um determinado endereço, ele utiliza uma rota *default*.
- Servidores, estações de trabalho, *firewalls*, etc., só conhecem as redes diretamente conectadas em uma interface.
  - Para alcançar alguém que não esteja diretamente conectado, eles terão que usar rota *default* para um outro que conheça.
- Todo mundo precisa ter rota *default*?

# Rota Default

- DFZ (*Default Free Zone*) - conceito existente entre as operadoras. É uma região da Internet livre de rota *default*.
- Roteadores DFZ não possuem rota *default*, possuem tabela BGP completa.
- ASs que possuem tabela completa precisam ter rota *default*?
- A tabela completa, mostra todas as entradas de rede do mundo.
  - roteadores têm que processar informações do mundo inteiro em tempo real;
  - problemas de escalabilidade futura.

# Rota Default

- Se houver tabela completa e rota *default*, neste caso, a rota *default* vai ser usada?
- Ex.:
  - Imagine uma rede comprometida pela infecção de um *malware*;
  - A máquina contaminada irá “varrer” a Internet tentando contaminar outras máquinas, inclusive IPs que não estão alocados, e não estão na tabela completa;
  - Se houver rota *default*, o seu roteador vai encaminhar esse tráfego não válido para frente;
  - Essa é uma das razões de se utilizar DFZ;
  - Sugestão: criar uma rota *default* e apontar para Null0 ou DevNull, e desabilitar o envio das mensagens '*ICMP unreachable*'.
- A rota *default* em IPv4 é 0.0.0.0/0 e em IPv6 ::/0.

# Protocolos de Roteamento Interno

- Há duas principais opções para se trabalhar com roteamento interno:
  - OSPF
  - IS-IS
    - protocolos do tipo *Link-State*;
    - consideram as informações de estado e mandam atualizações de forma otimizada;
    - trabalham com estrutura hierárquica.
- Terceira opção
  - RIP
- O protocolo de roteamento interno deve ser habilitado apenas nas interfaces necessárias.

# RIPng

- *Routing Information Protocol next generation* (RIPng) - protocolo IGP simples e de fácil implantação e configuração.
- Protocolo do tipo Vetor de Distância (Bellman-Ford).
- Baseado no RIPv2 (IPv4).
- Protocolo específico para IPv6.
  - Suporte ao novo formato de endereço;
  - Utiliza o endereço *multcast* **FF02::9** (*All RIP Routers*) como destino;
  - O endereço do próximo salto deve ser um endereço *link local;*
  - Em um ambiente IPv4+IPv6 é necessário usar RIP (IPv4) e RIPng (IPv6).

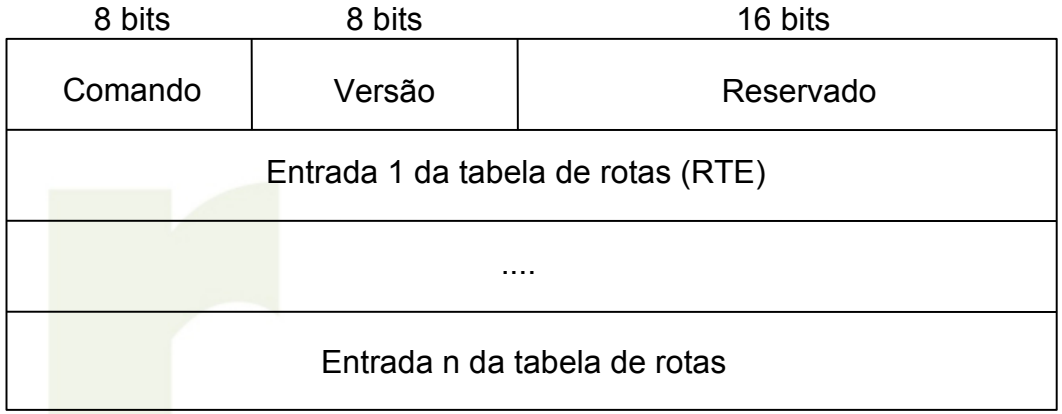
# RIPng

- Limitações:
  - Diâmetro máximo da rede é de 15 saltos;
  - Utiliza apenas a distância para determinar o melhor caminho;
  - *Loops* de roteamento e contagem até o infinito.

- Atualização da tabelas de rotas:
  - Envio automático a cada 30 segundos - independente de mudanças ou não.
  - Quando detecta mudanças na topologia da rede - envia apenas a linha afetada pela mudança)
  - Quando recebem uma mensagem do tipo *Request*

# RIPng

- Mensagens *Request* e *Response*

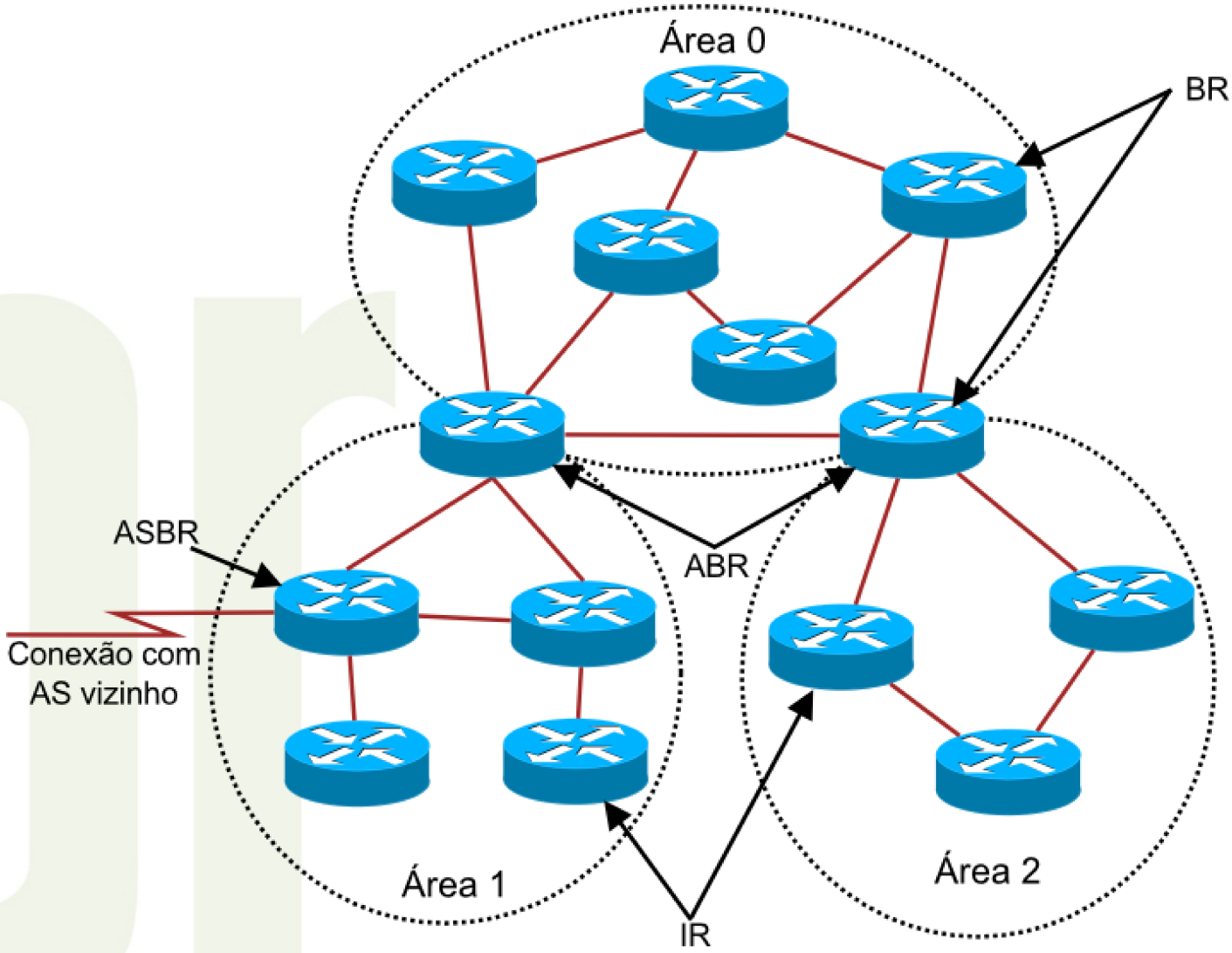



- RTE
  - Prefixo IPv6 (128 bits)
  - Identificação da rota (16 bits)
  - Tamanho do prefixo (8 bits)
  - Métrica (8 bits)
- Diferente do RIPv2, o endereço do próximo salto aparece apenas uma vez, seguido de todas as entradas que devem utilizá-lo.

# OSPFv3

- *Open Shortest Path First version 3* (OSPFv3) - protocolo IGP do tipo *link-state*
  - Roteadores descrevem seu estado atual ao longo do AS enviando LSAs (*flooding*)
- Utiliza o algoritmo de caminho mínimo de Dijkstra
- Agrupa roteadores em áreas
- Baseado no OSPFv2
- Protocolo específico para IPv6
  - Em um ambiente IPv4+IPv6 é necessário usar OSPFv2 (IPv4) e OSPFv3 (IPv6)

# Roteadores OSPFv3



# OSPFv3

Semelhanças entre OSPFv2 e OSPFv3

- Tipos básicos de pacotes
  - Hello, DBD, LSR, LSU, LSA
- Mecanismos para descoberta de vizinhos e formação de adjacências
- Tipos de interfaces
  - *point-to-point*, *broadcast*, NBMA, *point-to-multipoint* e *links* virtuais
- A lista de estados e eventos das interfaces
- O algoritmo de escolha do *Designated Router* e do *Backup Designated Router*
- Envio e idade das LSAs
- AREA_ID e ROUTER_ID continuam com 32 bits

# OSPFv3

Diferenças entre OSPFv2 e OSPFv3

- OSPFv3 roda por enlace e não mais por sub-rede
- Foram removidas as informações de endereçamento
- Adição de escopo para *flooding*
- Suporte explícito a múltipla instâncias por enlace
- Uso de endereços *link-local*
- Mudanças na autenticação
- Mudanças no formato do pacote
- Mudanças no formato do cabeçalho LSA
- Tratamento de tipos de LSA desconhecidos
- Suporte a áreas Stub/NSSA
- Identificação de vizinhos pelo Router IDs
- Utiliza endereços *multicast* (*AllSPFRouters* **FF02::5** e *AllDRouters* **FF02::6**)

# IS-IS

- *Intermediate System to Intermediate System* (IS-IS) - protocolo IGP do tipo *link-state*
- Desenvolvido originalmente para funcionar sobre o protocolo CLNS
  - *Integrated IS-IS* permite rotear tanto IP quanto OSI
  - Utiliza NLPID para identificar o protocolo de rede utilizado
- Trabalha em dois níveis
  - L2 = Backbone
  - L1 = Stub
  - L2/L1= Interligação L2 e L1

# IS-IS

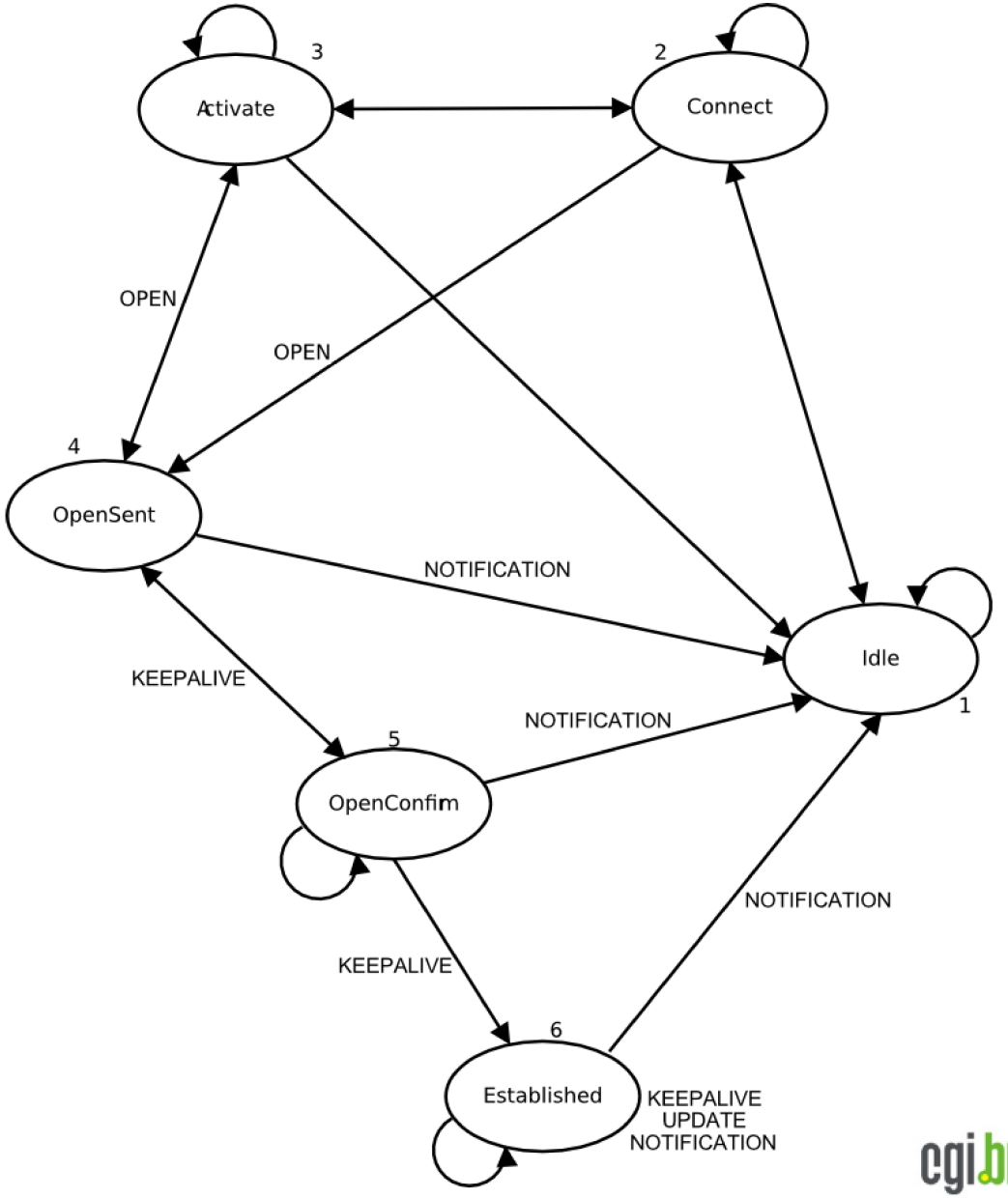
- Não há uma nova versão desenvolvida para trabalhar com o IPv6. Apenas adicionaram-se novas funcionalidades à versão já existente
- Dois novos TLVs para
  - IPv6 Reachability
  - IPv6 Interface Address
- Novo identificador da camada de rede
  - IPv6 NLPID
- Processo de estabelecimento de vizinhanças não muda

# Protocolo de Roteamento Externo

- O protocolo de roteamento externo padrão hoje, é o *Border Gateway Protocol* versão 4 (BGP-4).
    - protocolo do tipo *path vector.*
- Roteadores BGP trocam informações de roteamento entre ASs vizinhos.
    - com essas informações, desenham um grafo de conectividade entre os ASs.

# BGP

- Porta TCP 179
- Quatro tipos de mensagem:
  - *Open*
  - *Update*
  - *Keepalive*
  - *Notification*
- Dois tipos de conexão:
  - eBGP
  - iBGP
- Funcionamento representado por uma Máquina de Estados.



# Atributos do BGP

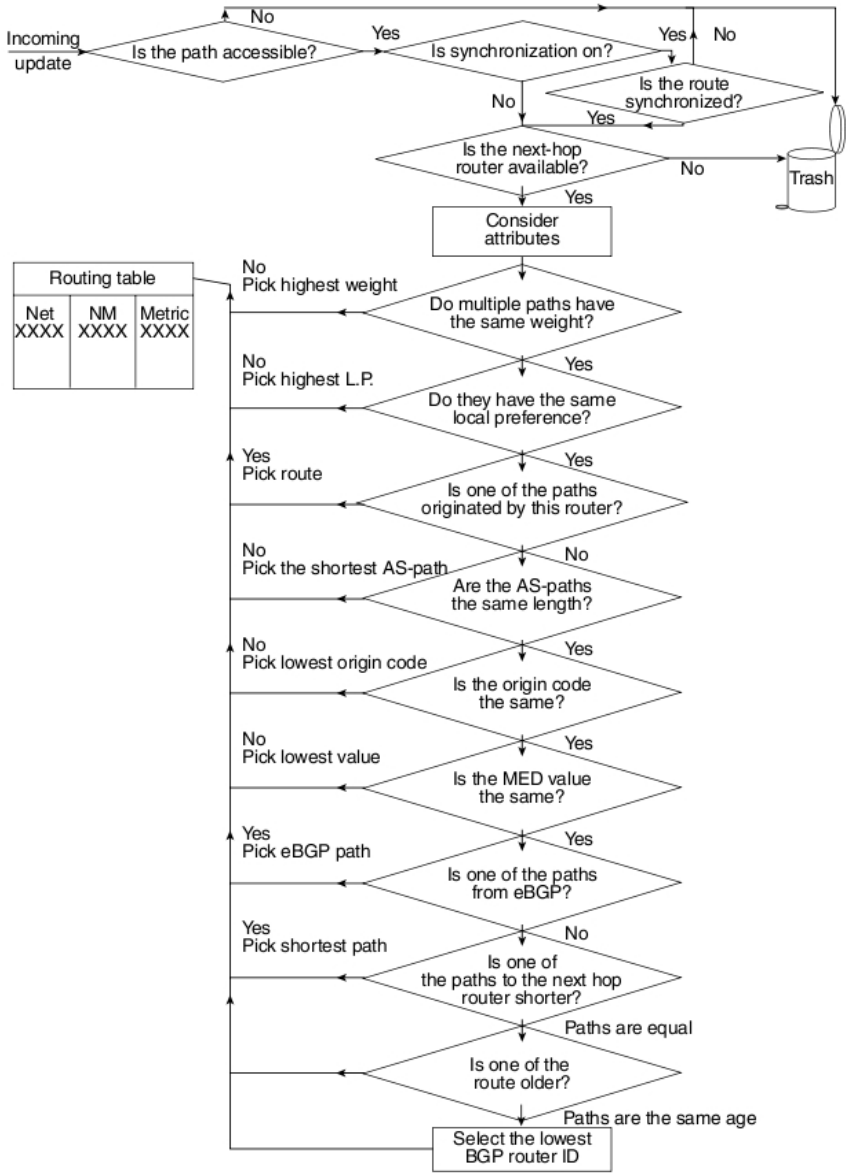
- O critério de seleção entre diferentes atributos do BGP varia de implementação para implementação.

- Os atributos BGP são divididos em algumas categorias e sub-categorias.

| | | |
|---|---|---|
| *ORIGIN* | Bem-conhecido | Mandatório |
| *AS_PATH* | Bem-conhecido | Mandatório |
| *NEXT_HOP* | Bem-conhecido | Mandatório |
| *MULTI_EXIT_DISC* | Opcional | Não-transitivo |
| *LOCAL_PREF* | Bem-conhecido | Discricionário |
| *ATOMIC_AGGREGATE* | Bem-conhecido | Discricionário |
| *AGGREGATOR* | Opcional | Transitivo |

# Atributos do BGP

- Os atributos são considerados se o caminho for conhecido, se houver conectividade, se for acessível e se o *next hop* estiver disponível.
- A forma de seleção pode variar de acordo com a implementação.
- O *LOCAL_PREFERENCE* é um atributo extremamente poderoso para influenciar o tráfego de saída.
- O valor do *LOCAL_PREFERENCE* é valido para todo o AS.




Seleção do caminho no BGP (CISCO).

# Multiprotocolo BGP

- *Multiprotocol* BGP (MP-BGP) - extensão do BGP para suportar múltiplos protocolos de rede ou famílias de endereços.
    - Para se realizar o roteamento externo IPv6 é essencial o suporte ao MP-BGP, visto que não há uma versão específica de BGP para tratar esta tarefa.
- Dois novos atributos foram inseridos:
    - *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.

# Multiprotocolo BGP

- MP_REACH_NLRI
  - *Address Family Identifier* (2 Bytes)
  - *Subsequent Address Family Identifier* (1 Byte)
  - *Length of Next Hop Network Address* (1 Byte)
  - *Network Address of Next Hop* (variável)
  - *Reserved* (1 Byte)
  - *Network Layer Reachability Information* (variável)

- MP_UNREACH_NLRI
  - *Address Family Identifier* (2 Bytes)
  - *Subsequent Address Family Identifier* (1 Byte)
  - *Withdrawn Routes* (variável)

# Tabela BGP

- As informações sobre as rotas da Internet encontram-se na tabela BGP.
- Em roteadores de borda, essas informações são replicadas para a RIB e para a FIB, IPv4 e IPv6.
    - Tabela Global IPv4 → ~430.000 entradas
    - Tabela Global IPv6 → ~10.000 entradas
- A duplicidade dessas informações implica em mais espaço, memória, e processamento.
    - Agregação de rotas
    - Evitar anúncio de rotas desnecessários
    - Limitar a quantidade de rotas recebidas de outros ASs
        - Importante em IPv4
        - Fundamental em IPv6

# IPv6.br

## A Nova Geração do Protocolo Internet

# Boas Práticas de BGP

# Estabelecendo sessões BGP

- Uma sessão BGP é estabelecida entre dois roteadores baseada numa conexão TCP.
  - porta TCP 179;
  - conexão IPv4 ou IPv6.
- Interface de *Loopback*
  - interface lógica;
  - não “caem”.

# Estabelecendo sessões BGP

iBGP entre *loopbacks*

- É fundamental estabelecer sessões iBGP utilizando a interface de *loopback*.
  - via IP da interface real:
    - se o *link* for interrompido, a sessão também será.
  - via IP da interface de *loopback*:
    - mais estabilidade;
    - os IPs das interfaces de *loopback* serão aprendidos via protocolo IGP.
    - se o *link* for interrompido, a sessão pode ser estabelecida por outro caminho.

# Estabelecendo sessões BGP


eBGP entre *loopbacks*

- Balanceamento
  - Ex.:
    - Há dois roteadores e cada roteador representa um AS;
    - Eles estão conectados por dois *links;*
    - Utilizando o IP das interfaces reais:
      - Serão necessárias duas sessões BGP;
      - Eventualmente com políticas diferentes.
    - Utilizando o IP das interfaces de *loopbacks*
      - É estabelecida uma única sessão BGP;
      - Cria-se uma rota estática para o IP da interface *loopback* do vizinho através de cada *link.*

# Estabelecendo sessões BGP

eBGP entre *loopbacks*


- Essa rota estática deve ser via interface ou via IP?
  - Se for uma interface serial pode-se apontar a rota para a interface;
  - Se for uma interface Ethernet deve-se apontar para o IP.
- Em *link* serial, o tamanho de rede IPv4 normalmente utilizado é /30.
  - Um /30 possui 4 IP; rede; *broadcast;* e os dois lados;
  - Em *links* seriais pode-se utilizar /31.
- Qual o equivalente ao /31 em IPv6?
- Em IPv6 pode-se trabalhar com redes /64 em *links* seriais.
- Uma boa opção é trabalhar com /112.

# Estabelecendo sessões BGP

eBGP entre *loopbacks*


- Normalmente utilizam-se na interface de *loopback* prefixo /32 IPv4 ou /128 IPv6.
- O IP da *loopback* é de responsabilidade do próprio AS.
  - Não se deve utilizar IP privado.
- O IP do *link* de trânsito é da responsabilidade do Provedor de Trânsito.
  - Esse IP deve ou pode ser roteável?
    - Se for uma relação IP interna com a operadora, pode ser um IP válido, da operadora e não roteável, ex. conexão MPLS;
    - Se for um serviço Internet, o IP DEVE ser roteável.

# Estabelecendo sessões BGP

eBGP entre *loopbacks*


- Segurança
  - A utilização de interfaces *loopbacks* em sessões eBGP não é necessária apenas para garantir balanceamento.
  - Estabelecer sessões eBGP utilizando o IP da interface, facilita muito ataques contra a infraestrutura.
  - É recomendável trabalhar eBGP entre *loopbacks* mesmo que só haja um *link*.

# Estabelecendo sessões BGP

eBGP entre *loopbacks*


- Segurança
  - Ex.:
    - Para estabelecer uma sessão eBGP sobre TCP são necessárias 4 informações básicas:
      - 2 IPs e 2 portas TCP (179 e >1024).
    - Se a sessão eBGP for estabelecida utilizando o IP da interface:
      - normalmente identifica-se um dos IP utilizando `traceroute`;
      - descobrindo o primeiro, descobre-se o segundo, visto que normalmente utiliza-se /30;
      - a terceira informação é uma porta padrão, a 179.
    - Ou seja, de 4 variáveis 3 podem ser descobertas de forma relativamente fácil.

# Estabelecendo sessões BGP

eBGP entre *loopbacks*


- Segurança
  - Uma das formas de derrubar um As ou um destino, é derrubar o AS que provê conectividade para ele.
  - Estabelecendo uma sessão eBGP utilizando *loopbacks:*
    - os IPs são das redes internas, não tendo relações entre eles;
    - dificulta a descoberta via `traceroute`.

# Estabelecendo sessões BGP

- Também recomenda-se trabalhar com uma *loopback* por função e não uma por roteador:
  - pode-se configurar uma *loopback* para o Router ID, uma para o iBGP e uma para o eBGP;
  - facilita a migração de serviços;
  - traz flexibilidade, porém, consome mais endereços IP.

# Utilizando MD5

- Uma importante técnica de proteção é a utilização de MD5 para autenticação das sessões BGP.
- Garante que apenas roteadores confiáveis estabeleçam sessões BGP com o AS.
- O algorítimo MD5 cria um *chechsum* codificado que é incluído no pacote transmitido.
- O roteador que recebe o pacote utiliza uma chave de autenticação para verificar o *checksum.*

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

# TTL-Security Check

- Trabalhar com TTL ou *Hop-Limit* igual a 1 auxilia na segurança
  - Permite que apenas se receba mensagens eBGP de quem estiver diretamente conectado;
  - Porém isto é facilmente burlado.
- RFC5082 recomenda o uso de TTL ou *Hop-Limit* igual a 255.
- Ex.:

```
router-R13(config-router)# neighbor 2001:DB8:200:FFFF::255
ttl-security hops 1
```

  - Define o valor mínimo esperado para o *Hop-Limit* de entrada para pelo menos 254 (255 - 1).
  - O roteador aceitará a sessão a partir de 2001:DB8:200:FFFF::255 se este estiver a 1 salto de distância.

# TTL-Security Check

- Esse é o terceiro mecanismo de proteção do eBGP apresentado até o momento:

  1º – estabelecer a sessão entre *loopbacks*;

  2º – Usar MD5;

  3º – Usar *TTL-Security Check*.

- O *TTL-Security Check* é pouco utilizado, mas é extremamente útil.
- Apenas enviar o pacote com TTL 255 não é suficiente. Também é preciso configurar o vizinho, senão...
  - ...a sessão eBGP poderá ser estabelecida por um *link* diferente do correto;
  - ...dificultará a detecção da origem de problemas.
- Sessões entre *loopbacks* use `ttl-security hops 2`.

# Desabilitando a Descoberta de Vizinhança

- Há roteadores que trazem o anúncio de mensagens RA habilitado por padrão.
- Se for utilizado na interface do roteador um endereço /64 vai haver descoberta de vizinhança, mesmo entre roteadores.
  - Com isso o roteador pode anunciar que ele é o *gateway* padrão;
  - Pode gerar *looping*
- Não há problemas em *links* para estações de trabalho.
- Em *links* entre roteadores deve-se desabilitar o envio de RA.

  - `ipv6 nd ra suppress` (Cisco)
  - `ipv6 nd suppres-ra` (Cisco / Quagga / Juniper)

# Verificando Configurações

- Verificando os protocolos configurados:
  - `show ip protocols` (Cisco)
  - `show ipv6 protocols` (Cisco)
  - No Quagga existe um *daemon* específico para cada protocolo de roteamento, tratado como um processo separado.
- Verificando o status e os endereços das interfaces:
  - `show ip interface brief` (Cisco)
  - `show ipv6 interface brief` (Cisco)
  - `show interface terse` (Juniper v4 e v6)
  - Note que, no caso de se trabalhar com sub-interfaces, o endereço *link-local* IPv6 será o mesmo. São interfaces lógicas distintas, mas o endereço é composto pelo MAC da física.

# Conferindo as configurações do eBGP e do iBGP

- Visualizando a configuração corrente a partir do BGP (Cisco):

```
router-R13#show running-config | begin bgp
router bgp 64501
 no bgp default ipv4-unicast
 bgp log-neighbor-changes
 neighbor 2001:DB8:21:FFFF::254 remote-as 64501
 neighbor 2001:DB8:21:FFFF::254 description R12
 neighbor 2001:DB8:21:FFFF::254 update-source Loopback20
 neighbor 2001:DB8:21:FFFF::254 version 4
 neighbor 2001:DB8:21:FFFF::255 remote-as 64501
 neighbor 2001:DB8:21:FFFF::255 description R11
 neighbor 2001:DB8:21:FFFF::255 update-source Loopback20
 neighbor 2001:DB8:21:FFFF::255 version 4
 neighbor 2001:DB8:200:FFFF::255 remote-as 64512
 neighbor 2001:DB8:200:FFFF::255 description R03
 neighbor 2001:DB8:200:FFFF::255 ebgp-multihop 2
 neighbor 2001:DB8:200:FFFF::255 update-source Loopback30
 neighbor 2001:DB8:200:FFFF::255 version 4
 ...
```

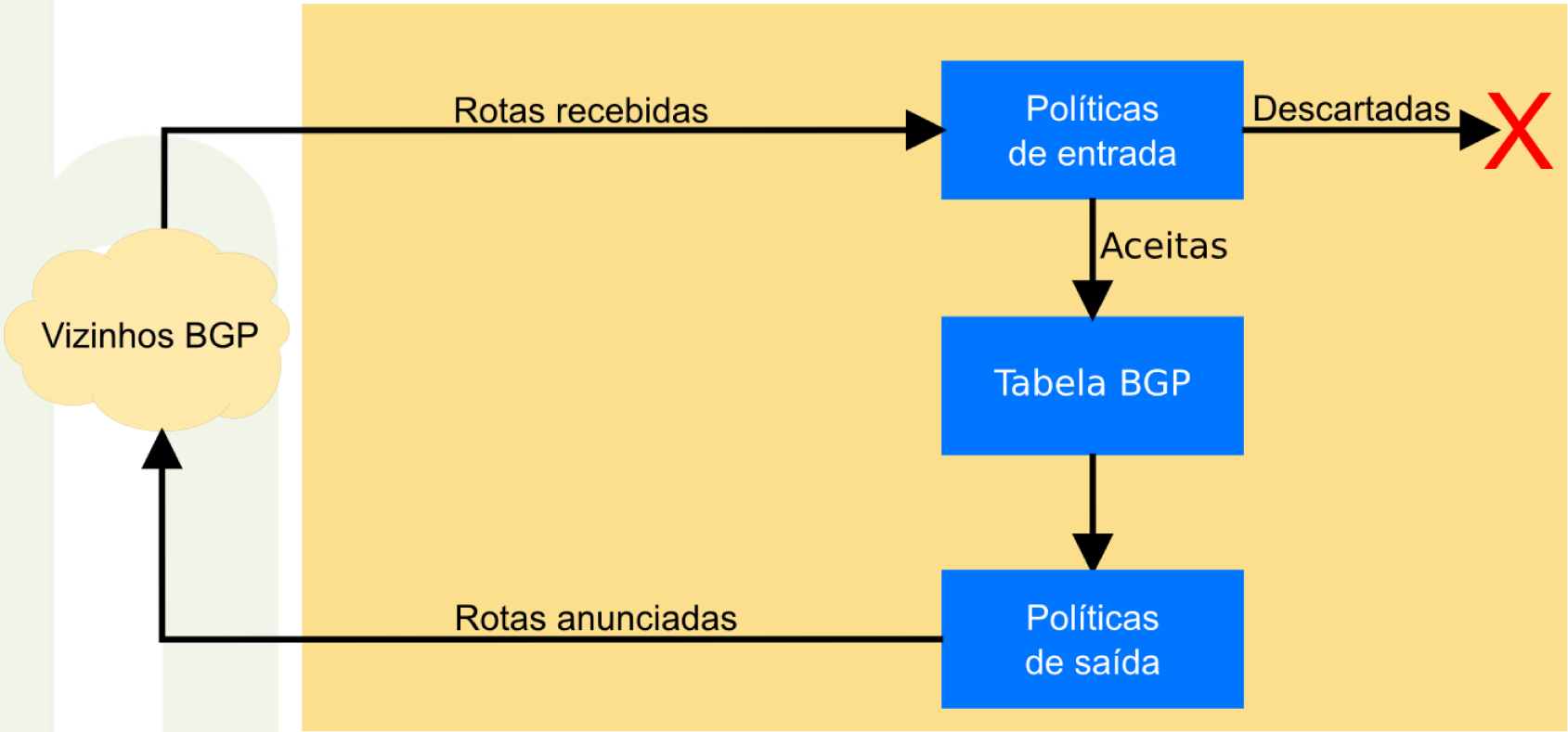
# Configurações do *address-family*

- Em roteadores Cisco e Quagga, para utilizar IPv6 é preciso especificar a família de endereços com a qual se está trabalhando.
- Aplicar as configurações específicas de cada família para cada vizinho.

```
router-cisco# show running-config | begin address-family ipv6
address-family ipv6
 neighbor 2001:DB8:21:FFFF::254 activate
 neighbor 2001:DB8:21:FFFF::254 next-hop-self
 neighbor 2001:DB8:21:FFFF::254 soft-reconfiguration inbound
 neighbor 2001:DB8:21:FFFF::255 activate
 neighbor 2001:DB8:21:FFFF::255 next-hop-self
 neighbor 2001:DB8:21:FFFF::255 soft-reconfiguration inbound
 neighbor 2001:DB8:200:FFFF::255 activate
 neighbor 2001:DB8:200:FFFF::255 soft-reconfiguration inbound
 neighbor 2001:DB8:200:FFFF::255 route-map BGPin-IPv6-AS64512 in
 neighbor 2001:DB8:200:FFFF::255 route-map BGPout-IPv6-AS64512 out
 network 2001:DB8:21::/48
 network 2001:DB8:21:8000::/49
exit-address-family
```

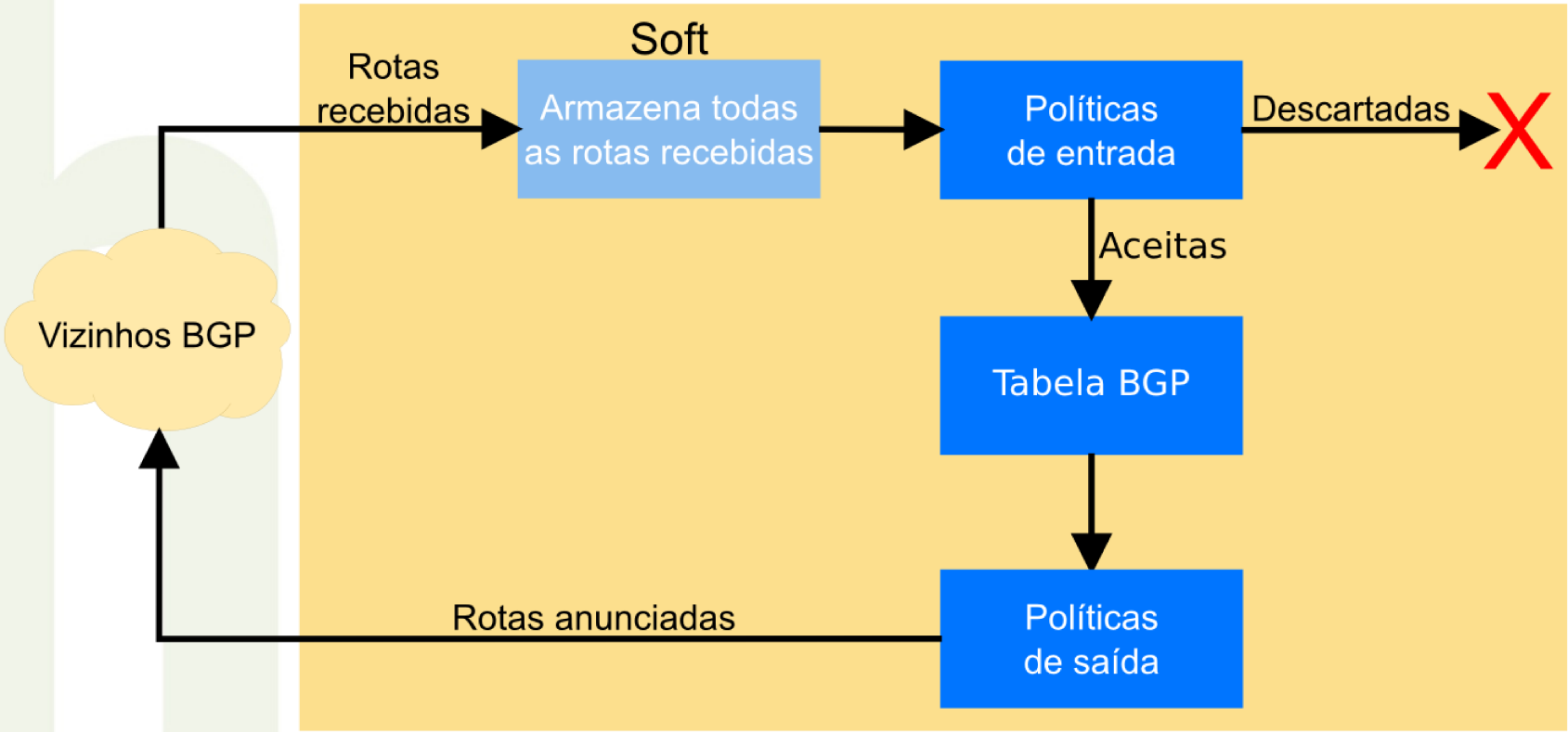
# Configurações do *address-family*

- *Soft-Reconfiguration Inbound*



# Configurações do *address-family*

- *Soft-Reconfiguration Inbound*




- `router-R13# clear bgp ipv6 unicast 2001:DB8:200:FFFF::255 soft in`

# Configurações do *address-family*

- *Route Refresh*
  - Quando os roteadores iniciam uma sessão BGP, cada roteador passa uma série de informações sobre os recursos que ele conhece, como: quais *capabilities* ele suporta.
  - Uma delas é o *route-refresh*.
  - Permite recuperar as informações originais da tabela de rotas sem "derrubar" a sessão BGP e sem criar tabelas adicionais.
    - Solicita ao vizinho o reenvio da tabela de rotas.
  - Para saber se o roteador suporta *route-refresh* use o comando:
    - `show ipv6 bgp neighbor 2001:DB8:200:FFFF::255`

# Conferindo as configurações do eBGP e do iBGP

- Visualizando a configuração corrente a partir do BGP (Juniper):

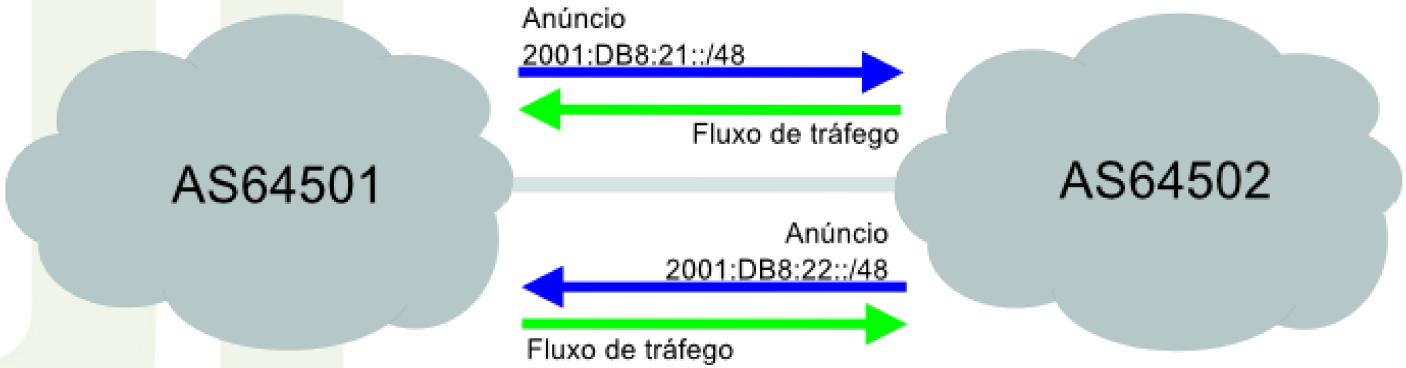
```
juniper@R11> show configuration protocols bgp
        protocols {
    bgp {
        group iBGPv6 {
            type internal;
            local-address 2001:DB8:21:FFFF::255;
            export next-hop-self;
            neighbor 2001:DB8:21:FFFF::252;
            neighbor 2001:DB8:21:FFFF::254;
        }
        group eBGP-AS64511v6 {
            type external;
            neighbor 2001:db8:100:1::1 {
                import nh-BGPin-IPv6-AS64511;
                export nh-BGPout-IPv6-AS64511;
                peer-as 64511;
            }
        }
    }
```

# Decisão de Roteamento

- Os roteadores tomam decisões de acordo com as informações que eles conhecem.
- Essas informações são recebidas e passadas aos outros roteadores através dos protocolos de roteamento interno e externo.
  - Os roteadores só anunciam a melhor rota que eles conhecem para um determinado destino.
- Essas informações serão utilizadas para influenciar o tráfego de entrada e o de saída do AS.

# Influenciando o Tráfego

- Os prefixos que um AS anuncia, interferem no tráfego de entrada ou saída?
  - Os prefixos anunciados interferem na forma como os outros conhecem o AS.
    - tráfego de entrada.
  - Os prefixos recebidos de outras redes interferem no tráfego de saída.



# Influenciando o Tráfego

- O que é mais fácil, influenciar o tráfego de entrada ou de saída?
- Ex.:
  - Um AS possui um bloco IPv4 /20;
    - Este AS pode gerar para a Internet anúncios de prefixos até um /24, o prefixo IPv4 mais específico normalmente aceito pelas operadoras;
    - Quantos prefixos /24 podem ser gerados a partir de um /20?
    - E quantos prefixos podem ser gerados entre /20 e um /24?
    - E entre um /32 e um /48 IPv6?

# Influenciando o Tráfego

- A Internet sabe chegar até um AS por até 31 prefixos IPv4.
- E quantas entradas IPv4 um AS conhece da Internet?
- Portanto há muito mais poder para trabalhar com o tráfego de saída.
  - Maior quantidade de informações;
  - Nos prefixos é que são baseadas as decisões de roteamento.
    - Balanceamento de tráfego;
    - Contabilidade de tráfego;
    - ....
- A influência do tráfego de entrada e de saída está associada à política de roteamento a ser implementada.
  - Há duas frentes: a de entrada e a de saída, chamadas de AS-IN e AS-OUT.
- Da mesma forma para IPv4 e IPv6.

# Plano de Endereçamento

- Distribuição dos serviços, servidores, etc., entre partes distintas do bloco IP.
  - facilita a influência do tráfego de entrada e saída de seu AS;
  - não adianta concentrar todo o tráfego principal atrás do mesmo prefixo /24 ou /48 anunciado na Internet.
- Esta má distribuição irá restringir a influência do tráfego de entrada ou de saída?

# Plano de Endereçamento

- AS-OUT
  - É o que será anunciado para a Internet;
  - Interfere com o tráfego de entrada.
  - Ex.:
    - O AS64501 possui um /48 IPv6;
    - para fazer o balanceamento do tráfego, influenciaremos para que metade deste tráfego entre por um *link* e a outra metade entre por outro;
    - divide-se o /48 em dois /49, anunciando o primeiro /49 em um *link* e o segundo /49 em outro *link*.

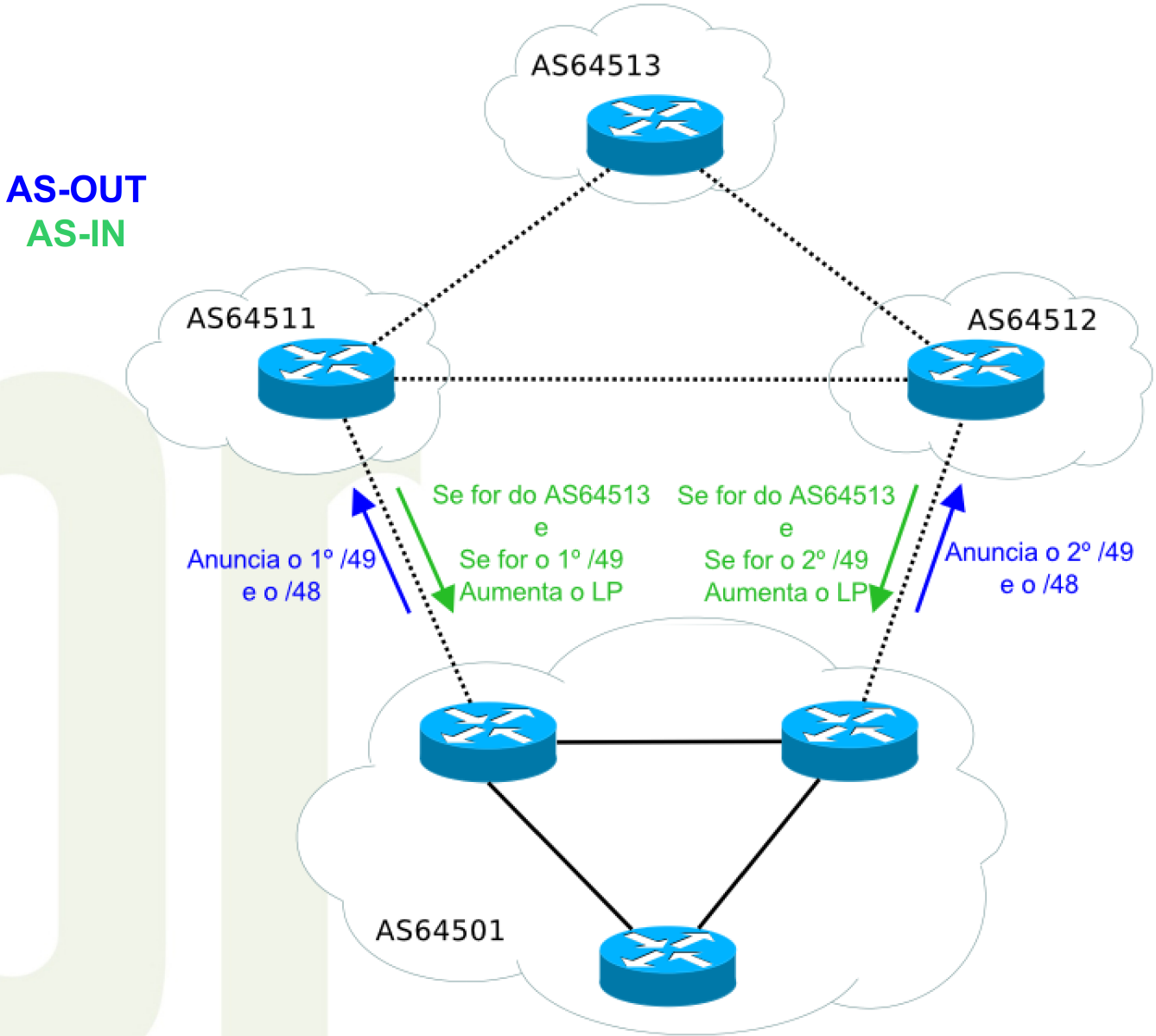
# Plano de Endereçamento

- AS-IN
  - Depende dos anúncios recebidos da Internet
    - normalmente a tabela completa.
  - Interfere com o tráfego de saída.
  - Pode-se influenciar o tráfego de saída alterando o valor do *LOCAL_PREFERENCE* de acordo com determinadas condições.
  - *LOCAL_PREFERENCE* é o atributo com maior força para influenciar o tráfego de saída.
    - Ex.: O AS64501 precisa influenciar seu tráfego de saída, de modo que o tráfego com destino ao primeiro /49 do AS64513 saia preferencialmente pelo *link* com o AS64511 e o tráfego com destino ao segundo /49 do AS64513 saia preferencialmente pelo *link* com o AS64512.
  - Preferencialmente é uma palavra chave para o BGP.

# Plano de Endereçamento

- Redundância
  - Cada /49 IPv6 é conhecido pelo mundo por apenas um *link*
    - Se um desses *links* cair, o /49 anunciado por ele ficará inacessível.
  - Para termos redundância, deve-se anunciar também o /48 nos dois *links*.
  - Como a preferência é pelo prefixo mais específico, se os dois *links* estiverem ativos, a Internet vai preferir os /49.
  - Quando um dos *links* cair e um dos /49 deixará ser anunciado, porém a Internet ainda terá a opção do /48 anunciado no outro *link*, garantindo a redundância.
  - Deve-se distribuir o tráfego entre os dois, colocando metade dos consumidores de tráfego de entrada em um /49 e metade no outro.

# Plano de Endereçamento



# Políticas de Roteamento

- Uma função importante do BGP está associada à manipulação dos atributos e os testes condicionais:
  - Cisco
    - *route-map* – define as condições para a redistribuição de rotas e permite controlar e modificar informações de políticas de roteamento;
    - *prefix-list* – mecanismo de filtragem de prefixos muito poderoso. Permite trabalhar com notação de prefixo, adicionar descrição e trabalhar com sequencia;
  - Juniper
    - *route-filter* – utilizado para comparar rotas individualmente ou em grupos.

# Estabelecendo a Política de Saída

- Cisco

```
ipv6 prefix-list BGPout-IPv6-AS64512 description Prefixos para AS64512
ipv6 prefix-list BGPout-IPv6-AS64512 seq 10 permit 2001:DB8:21::/48
ipv6 prefix-list BGPout-IPv6-AS64512 seq 20 permit 2001:DB8:21:8000::/49
```

- Juniper

```
policy-statement BGPout-IPv6-AS64511 {
    term term-1 {
        from {
            route-filter 2001:db8:21::/48 exact;
            route-filter 2001:db8:21::/49 exact;
        }
        then accept;
    }
    term implicit-deny {
        then reject;
    }
}
```

# Aplicando a Política de Saída

- Cisco

```
route-map BGPout-IPv6-AS64512 permit 10
 match ipv6 address prefix-list BGPout-IPv6-AS64512
```

- Juniper

```
    policy-statement nh-BGPout-IPv6-AS64511 {
        term term-1 {
            from policy BGPout-IPv6-AS64511;
            then accept;
        }
        term implicit-deny {
            then reject;
        }
    }
```

# Estabelecendo a Política de Entrada

- AS-PATH - Atributo fundamental do BGP. Consiste no ASN das redes pelas quais o pacote passará até chegar ao destino.
- Análise do AS-PATH com expressões regulares:

- Cisco / Quagga

```
ip as-path access-list 32 permit .*
ip as-path access-list 69 deny .*
ip as-path access-list 300 permit (_64513)+$
```

- Juniper

```
as-path ALL .*;
as-path AS64513 ".*( 64513)+$";
```

# Estabelecendo a Política de Entrada

- Cisco

```
ipv6 prefix-list BGPin-IPv6-AS64513 description Prefixos Preferidos do AS64513
ipv6 prefix-list BGPin-IPv6-AS64513 seq 10 permit 2001:DB8:300:8000::/49
```

- Juniper

```
policy-statement BGPin-IPv6-AS64513 {
    term term-1 {
        from {
            route-filter 2001:db8:300::/49 exact;
        }
        then accept;
    }
    term implicit-deny {
        then reject;
    }
}
```

# Estabelecendo a Política de Entrada

- Filtros de proteção
  - Cisco

```
ipv6 prefix-list IPv6-AS64501-all description Todos Blocos IPv6
ipv6 prefix-list IPv6-AS64501-all seq 10 permit 2001:DB8:21::/48 le 128
```

  - Juniper

```
    policy-statement IPv6-AS64501-all {
        term term-1 {
            from {
                route-filter 2001:db8:21::/48 orlonger;
            }
            then accept;
        }
        term implicit-deny {
            then reject;
        }
    }
```

# Estabelecendo a Política de Entrada

- Filtros de proteção
  - Cisco

```
ipv6 prefix-list IPv6-block-deny description Prefixos Gerais Bloqueados
ipv6 prefix-list IPv6-block-deny seq 10 permit ::/0
ipv6 prefix-list IPv6-block-deny seq 20 permit ::/8 le 128
ipv6 prefix-list IPv6-block-deny seq 30 permit 3ffe::/16 le 128
ipv6 prefix-list IPv6-block-deny seq 40 permit 2001:db8::/32 le 128
ipv6 prefix-list IPv6-block-deny seq 50 permit 2001::/33 le 128
ipv6 prefix-list IPv6-block-deny seq 60 permit 2002::/17 le 128
ipv6 prefix-list IPv6-block-deny seq 70 permit fe00::/9 le 128
ipv6 prefix-list IPv6-block-deny seq 80 permit ff00::/8 le 128
ipv6 prefix-list IPv6-block-deny seq 90 permit fc00::/7 le 128
```

# Estabelecendo a Política de Entrada

- Filtros de proteção
  - Juniper

```
policy-statement IPv6-block-deny {
    term term-1 {
        from {
          route-filter ::/0 exact;
          route-filter ::/8 orlonger;
          route-filter 3ffe::/16 orlonger;
          route-filter 2001:db8::/32 orlonger;
          route-filter 2001::/32 longer;
          route-filter 2002::/16 longer;
          route-filter fe00::/9 orlonger;
          route-filter ff00::/8 orlonger;
          route-filter fc00::/7 orlonger;
        }
        then accept;
    }
    term implicit-deny {
        then reject;
    }
}
```

# Estabelecendo a Política de Entrada

- Filtros de permissão
  - Cisco

```
ipv6 prefix-list IPv6-block-permit description Prefixos Gerais Permitidos
ipv6 prefix-list IPv6-block-permit seq 10 permit 2000::/3 le 48
```

  - Juniper

```
policy-statement IPv6-block-permit {
    term term-1 {
        from {
           route-filter 2000::/3 prefix-length-range /3-/48
        }
        then accept;
    }
    term implicit-deny {
        then reject;
    }
}
```

# Aplicando a Política de Entrada

- Cisco

```
route-map BGPin-IPv6-AS64512 deny 10
 match ipv6 address prefix-list IPv6-AS64501-all
!
route-map BGPin-IPv6-AS64512 deny 20
 match ipv6 address prefix-list IPv6-block-deny
!
route-map BGPin-IPv6-AS64512 permit 30
 match ipv6 address prefix-list BGPin-IPv6-AS64513
 match as-path 300
 set local-preference 150
!
route-map BGPin-IPv6-AS64512 permit 40
 match ipv6 address prefix-list IPv6-block-permit
```

# Aplicando a Política de Entrada

- Juniper

```
policy-statement nh-BGPin-IPv6-AS64511 {
    term term-1 {
        from policy IPv6-AS64501-all;
        then reject;
    }
    term term-2 {
        from policy IPv6-block-deny;
        then reject;
    }
    term term-3 {
        from {
            as-path AS64513;
            policy BGPin-IPv6-AS64513;
        }
        then {
            local-preference 150;
            accept;
        }
    }
    term term-4 {
        from policy IPv6-block-permit;
        then accept;
    }
 term implicit-deny {
        then reject;
    }
}
```

# Verificando a Vizinhança BGP

- Mostrando todos os vizinhos BGP IPv4:
  - `show ip bgp summary` (Cisco / Quagga)
  - `show bgp summary` (Juniper)
- Mostrando todos os vizinhos BGP de ambas as famílias:
  - `show bgp ipv4 unicast summary` (Cisco / Quagga)
  - `show bgp ipv6 unicast summary` (Cisco / Quagga)
  - `show bgp all summary` (Cisco / Quagga)

# Verificando a Vizinhança BGP

- Cisco

```
router-R13#show bgp ipv6 unicast summary
BGP router identifier 172.21.15.253, local AS number 64501
BGP table version is 45, main routing table version 45
28 network entries using 4368 bytes of memory
54 path entries using 4104 bytes of memory
45/17 BGP path/bestpath attribute entries using 7560 bytes of memory
34 BGP AS-PATH entries using 848 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
Bitfield cache entries: current 2 (at peak 3) using 64 bytes of memory
BGP using 16944 total bytes of memory
26 received paths for inbound soft reconfiguration
BGP activity 49/1 prefixes, 96/21 paths, scan interval 60 secs

Neighbor          V      AS MsgRcvd MsgSent  TblVer InQ OutQ Up/Down State/PfxRcd
2001:DB8:21::254  4   64501    1867    1856       45   0    0 1w0d    Active
2001:DB8:21::255  4   64501    4136    3642       45   0    0 1d07h           26
2001:DB8:20::255  4   64512    1896    1876       45   0    0 1d07h            0
```

# Verificando a Vizinhança BGP

- Juniper

```
juniper@R11> show bgp summary
Groups: 4 Peers: 5 Down peers: 1
Table          Tot Paths  Act Paths Suppressed    History Damp State    Pending
inet.0                19         17          0          0          0          0
inet6.0               56         27          0          0          0          0
Peer                     AS  InPkt   OutPkt  OutQ Flaps Last Up/Dwn  State|#Active/Received/Accepted/Damped
10.1.8.1              64511   3785    4127    0      0  1d 7:26:53 17/17/17/0
172.28.15.252         64508   3776    4135    0      0  1d 7:26:38 0/2/2/0
172.28.15.254         64508   3775    4136    0      0  1d 7:26:46 Connect
2001:db8:28::252      64508   3794    4147    0      0  1d 7:26:40 Establ inet6.0: 0/29/29/0
2001:db8:28::254      64508   3775    4149    0      0  1d 7:26:46 Establ inet6.0: 0/0/0/0
2001:db8:10::1        64511   3810    4128    0      0  1d 7:26:57 Establ inet6.0: 27/27/27/0
```

# Looking Glass

- É importante verificar através de *Looking Glasses* remotos como as operadoras e toda a Internet recebem os anúncios do AS.
- Cisco

```
bgpd-R01> show bgp regexp _64501$
BGP table version is 0, local router ID is 10.3.255.255
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale, R Removed
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*  2001:db8:21::/48 2001:db8:300:11::2                       0 64511 64501 i
*>                  2001:db8:300:12::2                       0 64512 64501 i
*  2001:db8:21::/49 2001:db8:300:11::2                       0 64511 64512 64501 i
*>                  2001:db8:300:12::2                       0 64512 64501 i
*  2001:db8:21:8000::/49
                    2001:db8:300:12::2                       0 64512 64511 64501 i
*>                  2001:db8:300:11::2                       0 64511 64501 i
Total number of prefixes 3
```

# Looking Glass

- Juniper

```
juniper@R11> show route table inet6.0 aspath-regex .64513$

inet6.0: 59 destinations, 84 routes (59 active, 0 holddown, 0 hidden)
+ = Active Route, - = Last Active, * = Both

2001:db8:300::/48  *[BGP/170] 01:44:51, localpref 100
                      AS path: 64511 64513 I
                    > to 2001:db8:100:1::1 via ge-0/0/0.2105
                    [BGP/170] 01:44:13, MED 0, localpref 100, from 2001:db8:21:ffff::252
                      AS path: 64512 64513 I
                    > to fe80::224:97ff:fec1:c8bd via ge-0/0/0.2101
2001:db8:300::/49  *[BGP/170] 01:44:13, MED 0, localpref 150, from 2001:db8:21:ffff::252
                      AS path: 64512 64513 I
                    > to fe80::224:97ff:fec1:c8bd via ge-0/0/0.2101
                    [BGP/170] 01:44:51, localpref 100
                      AS path: 64511 64513 I
                    > to 2001:db8:100:1::1 via ge-0/0/0.2105
2001:db8:300:8000::/49
                   *[BGP/170] 01:44:51, localpref 150
                      AS path: 64511 64513 I
                    > to 2001:db8:100:1::1 via ge-0/0/0.2105
```