

nic.br

Núcleo de Informação
e Coordenação do
Ponto BR

cgib.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

Roteamento IPv6

ceptro.br nic.br egi.br

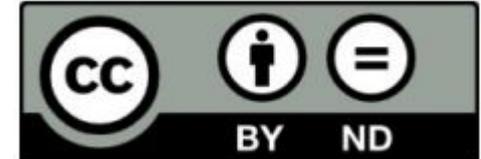
Licença de uso do material

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 CEPTR0.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.

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?

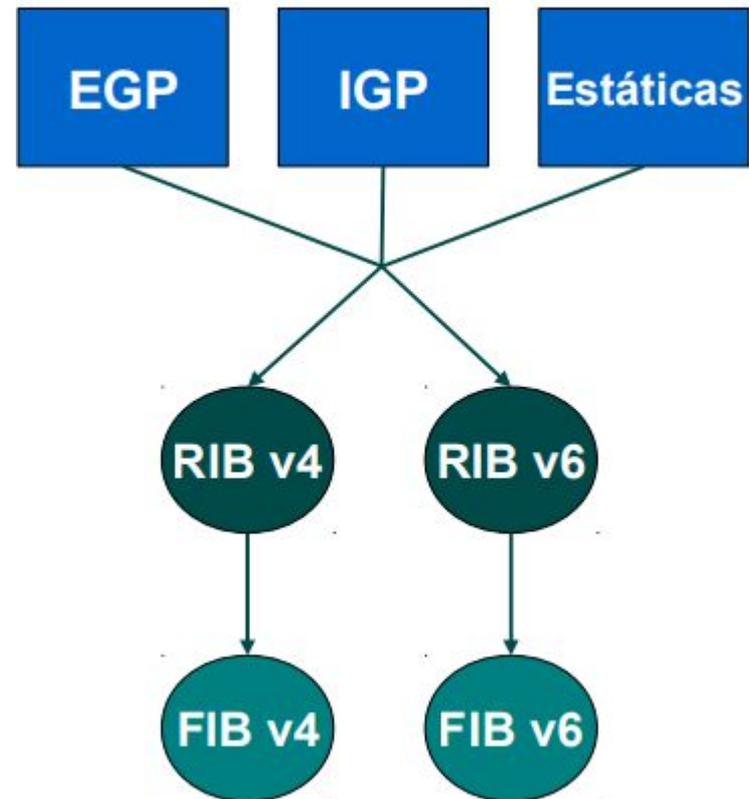
Tabela de Roteamento

- 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.
- ASes 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 multicast **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

8 bits	8 bits	16 bits
Comando	Versão	Reservado
Entrada 1 da tabela de rotas (RTE)		
....		
Entrada n da tabela de rotas		

- RTE

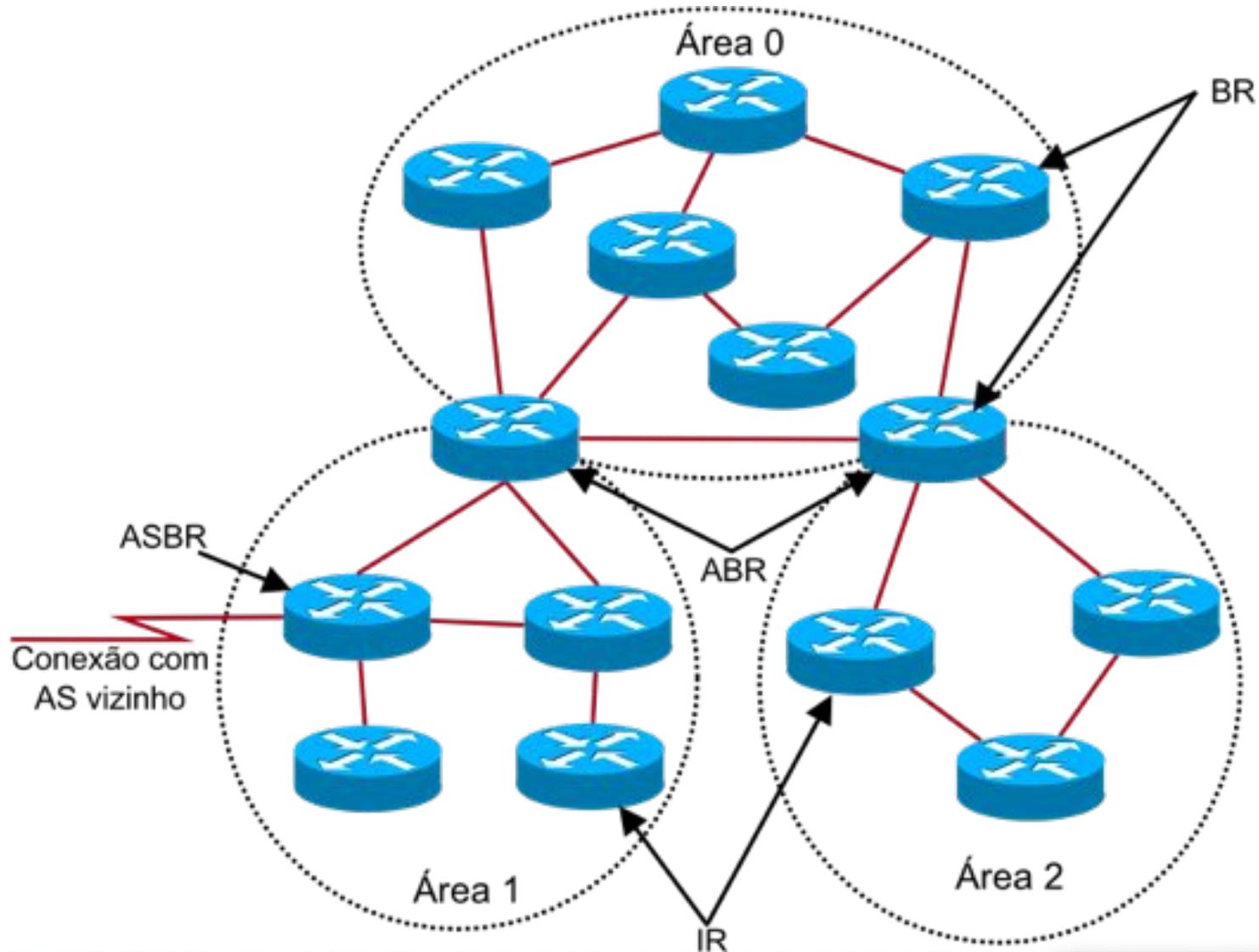
- Prefixo IPv6 (128 bits)
- Identificação da rota (16 bits)
- Tamanho do prefixo (8 bits)
- Métrica (8 bits)

- 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

Laboratório

Experiência 5.1

OSPFv3

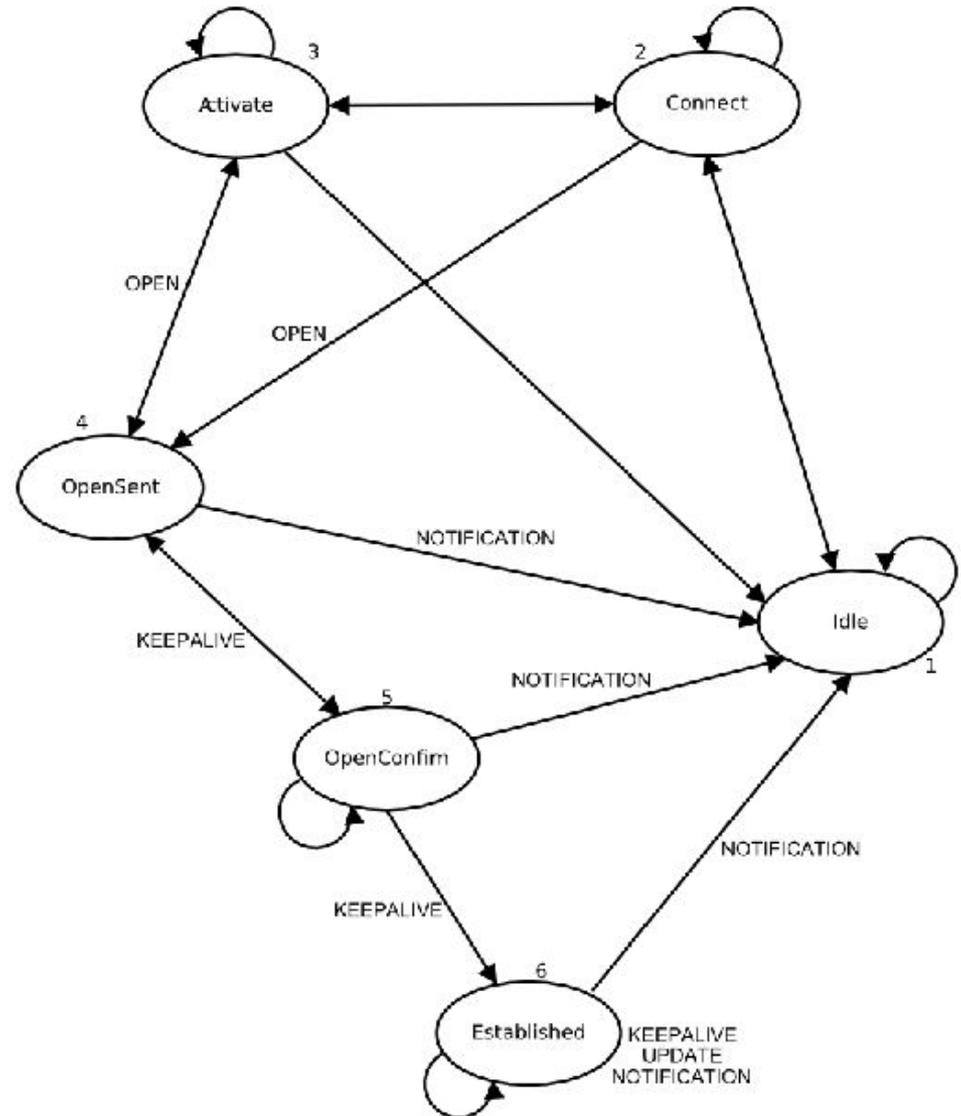
Página 323

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 tipo de mensagem:
 - Open
 - Update
 - Keepalive
 - Notification
- Dois tipos de conexão:
 - eBGP
 - IBGP
- Funcionamento representado por uma Máquina de Estados



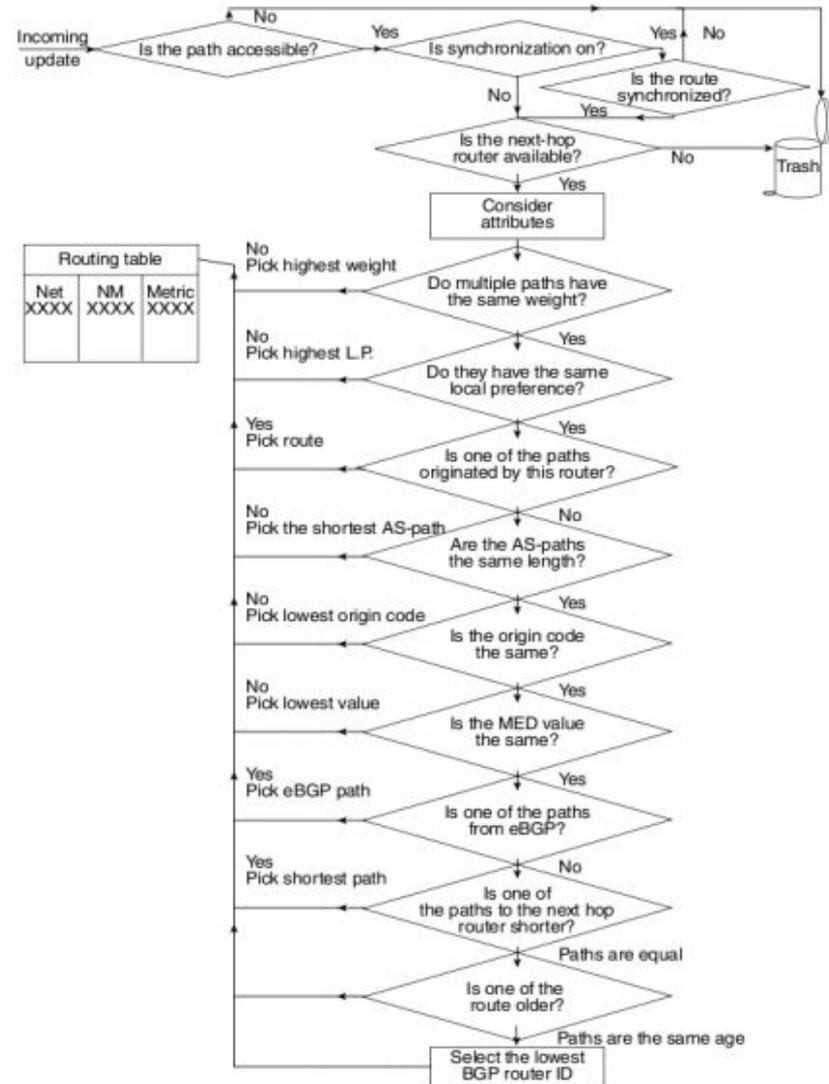
Atributos do BGP

- 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.

<i>ORIGIN</i>	Bem-conhecido	Mandatário
<i>AS_PATH</i>	Bem-conhecido	Mandatário
<i>NEXT_HOP</i>	Bem-conhecido	Mandatário
<i>MULTI_EXIT_DISC</i>	Opcional	Não-transitivo
<i>LOCAL_PREF</i>	Bem-conhecido	Discricionário
<i>ATOMIC_AGGREGATE</i>	Bem-conhecido	Discricionário
<i>AGGREGATOR</i>	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 válido para todo o AS.



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 → ~939.000 entradas
 - Tabela Global IPv6 → ~176.000 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

Laboratório

Experiência 5.2

BGP

Página 334

Dúvidas?



Obrigado !!!

nic.br egi.br

www.nic.br | www.cgi.br