Índice

1. IPV6

Os endereços, no IPV6, têm um tamanho de 128 bits.

1.1. Vantagens do IPV6

As vantagens do IPV6 são:

- **Roteamento(Routing)** mais eficiente, uma vez que reduz o tamanho das tabelas de roteamento;
- **Processamento de pacotes mais eficiente**, devido ao cabeçalho(header) simplificado presente no IPV6;
- **Fluxo de dados mais direto**, uma vez que o IPV6 usa multicast em vez de broadcast. O multicast permite que fluxos que precisem de grande largura de banda(fluxos multimédia, por exemplo), sejam enviados para os vários destinos simultaneamente, economizando assim largura de banda;
- **Simplificação da configuração em rede**, uma vez que permite a auto-configuração dos endereços. Um host pode gerar o seu próprio endereço IPV6, combinando-o com o seu endereço MAC;
- **Mais segurança**. Permite a introdução de IPSec no s seus pacotes(ICMPV6), sendo que o IPSec fornece confidencialidade, autenticação e integridade;

Semelhanças entre os protocolos de controlo, ICMPV4 e ICMPv6:

- Ambos são usados para diagnóstico e mensagens de erro;
- Funcionam sobre o ip;
- Ambos ainda são limitados em termos de segurança;

Novas funções do ICMPV6:

- Auto-configuração dos endereços das interfaces;
- Gestão de grupos mulicast(semelhante ao que acontece com o IGMP no IPV4);
- Suporte para a mobilidade dos nós de IPV6;

1.2. A atribuição de endereços IPV6

1.2.1. Auto-configuração (plug-and-play)

- O primeiro endereço a ser obtido é o endereço de local link:

  -Tem um formato genérico do tipo FE80:0:0:0:xxxx:xxxx:xxxx:xxxx;

  -A parte da identificação da interface é obtida com base no seu endereço de nível 2;

  -Este endereço permite uma comunicação imediata com todos os sistemas existentes na rede local;

- A partir daí e através de mensagens ICMPv6 trocadas entre o sistema terminal e o encaminhador da rede local são obtidos os outros endereços.
- O encaminhador anuncia o prefixo que é concatenado com o endereço de nível 2 de forma a obter-se um endereço global único;
- Outra alternativa passa pela utilização do DHCP (autoconfiguração statefull).

1.2.2. Passos da auto-configuração stateless

1. Construir um endereço único de âmbito Link-Local para o interface;
2. Testar se o endereço é único (Enviar uma mensagem ICMPv6 Neighbor-Discovery com endereço origem :: e destino o endereço criado, para verificar se obtém resposta; Se o endereço estiver atribuído recebe Neighbor-Advertisement);
3. Assumir o endereço Link-Local, caso o teste passe;
4. Contactar o router local com mensagem ICMPv6 Router-Discovery(Os router anunciam-se periodicamente com Router-Advertisement, mas essas mensagens podem ser solicitadas com pedido explícito Router-Discovery);
5. Seguir as recomendações do router local(Endereço do servidor DHCPv6 a contactar para prosseguir configuração statefull; Envio de um conjunto de prefixos a juntar à parte local do endereço formando novos endereços IPv6);
6. Configuração de endereços de âmbito Site-Local e Global com base nos endereços fornecidos.

2. Encaminhamento na Internet

2.1. Forwarding

As características do forwarding (reenvio) são:

- Utiliza a tabela de encaminhamento preenchida pelos protocolos de encaminhamento ou pelo administrador;
- Procura na tabela o "próximo salto" e a "interface de saída" ;
- Envia o pacote pelo interface respectivo, encapsulando-o numa trama de acordo com o tipo de interface.

2.1.1. O funcionamento do forwarding

O funcionamento do forwarding é o seguinte:

1. Chegada do pacote IP ao router numa das suas interfaces (ex: Ethernet), encapsulado numa trama; É processado de imediato o cabeçalho de nível 2; O pacote pode ser destruído se não for dirigido ao interface (ou broadcast);

2. Analisa-se de seguida o campo tipo do cabeçalho 2. O valor 0x800 indica que lá dentro vem um datagrama IP. Descarta-se o cabeçalho Ethernet e passa-se ao nível IP;

3. Processa-se o cabeçalho IP, verificando versão, tamanhos e checksum (só no IPv4); se alguma verificação falhar, descarta-se silenciosamente;

4. Senão, verifica se o campo TTL é superior a 1; Decrementa TTL e ajusta checksum (só IPv4) do cabeçalho antes de reenviar; Envia ICMP TTL Exceeded se TTL=0;

5. Pesquisa o endereço de destino IP na tabela de encaminhamento se endereço unicast (multicast e broadcast são tratados de maneira diferente); O endereço de destino é usado como chave de pesquisa na tabela de encaminhamento e obtém-se a rota mais adequada (melhor match);

6. Se não existir rota, o pacote é descartado e devolvido um ICMP Destination unreachable; Se existir, extrai-se da rota a interface de saída e o endereço do próximo router;

7. Se o pacote IP for maior que o MTU da interface de saída, fragmenta-se (só IPv4 e se o bit DF - Don't Fragment o permitir); Envia ICMP Destination unreachable, se não for possível ou permitido;

8. O router acrescenta então um novo cabeçalho de nível 2; Pode ter de usar o ARP (ou equivalente) para obter o endereço de destino nível 2 a partir do endereço IP do próximo salto;

9. Envia pelo interface definido, para o próximo salto;

2.1.2. Alterações ao processo normal de forwarding

2.1.2.1. Multi-path routing ou Load Balancing

Habitualmente só existe uma rota possível (a melhor para cada destino). Mas pode existir mais que uma rota de igual custo para o mesmo destino e o router pode distribuir os pacotes por cada uma dessas rotas, distribuindo o tráfego equitativamente.

A vantagem disto é haver mais largura de banda disponível para tráfego para esse destino.

A pesquisa na tabela devolve várias rotas e a escolha de qual usar depende da implementação.

Por vezes tenta-se manter os fluxos TCP na mesma rota evitando reordenações, sendo escolha da rota é feita usando uma função hash dos endereços origem e destino juntamente com o cabeçalho TCP;

Usando termo "custo" para referir a métrica associada a cada rota, a melhor rota é a de menor custo. A distribuição de carga por múltiplas rotas, pode ser feita de duas formas:

- Distribuição de carga entre rotas de igual custo (O tráfego é distribuído equitativamente por todas as rotas);
- Distribuição de carga entre rotas de custo diferente (Os pacotes são divididos pelas rotas disponíveis, sendo a distribuição de tráfego feita na proporção inversa do custo da rota. Mais tráfego para as rotas de menor custo e vice-versa. Nem todos os protocolos suportam esta opção.);

2.1.2.2. Encaminhamento com base no campo TOS

O campo TOS definia originalmente 5 tipos de serviço (normal, minimizar custo, maximizar fiabilidade, maximizar débito, minimizar o atraso) e tem agora novos usos nos novos modelos de Qualidade de Serviço. Teoricamente o router pode escolher rotas diferentes para diferentes valores do campo TOS.

A pesquisa na tabela é feita usando o par (endereço destino, campo TOS) do pacote como chave de pesquisa:

- A melhor rota para um tipo de serviço pode não ser a melhor para outro;
- Implica um rota por tipo de serviço, na tabela de encaminhamento.

Pouco usado na prática, é ainda um alvo nos novos modelos de Qualidade de Serviço: QoS routing

2.2. Routing

As características do routing (encaminhamento) são:

- Preenche a tabela de encaminhamento com a(s) melhor(es) rotas para as redes de destino (classfull) ou para um conjunto de prefixos de endereços (classless);
- Pode ser um processo manual, feito pelo administrador (encaminhamento estático). Ou um processo automático resultante da operação de um protocolo de encaminhamento, no caso mais comum de encaminhamento dinâmico.

2.2.1. A tabela de encaminhamento (routing table)

Os campos na tabela (destino, interface saída, próximo salto), podem ter milhares de entradas e por isso, a pesquisa tem de ser muito eficiente. É o processo mais crítico.

2.2.1.1. Processamento da pesquisa

A chave de pesquisa são endereços IP. (a porção do endereço que identifica a rede), pode ser obtida usando a máscara de rede.

Portanto, o processo é diferente com endereçamento por classes (classfull) ou sem classes (classless).

2.2.2. Endereçamento Classful

Este tipo de endereçamento acabou por revelar-se totalmente inadequado, devido ao uso ineficiente do espaço de endereçamento (exaustão de espaço), por exemplo, uma classe B aloca 65K hosts mesmo que existam apenas 2K hosts!

2.2.2.1. Pesquisa de rotas

Os seguintes procedimentos ilustram a pesquisa de rotas neste tipo Classful:

1. Examinar os primeiros bits do endereço, determinar a classe (A, B ou C);

2. Extrair o endereço de rede de acordo com a classe (primeiros 8 bit, 16 bit ou 24 bit);

3. Pesquisar na tabela uma entrada correspondente à parte da rede;

4. Se não encontrar, descartar pacote e enviar ICMP Destination Unreachable;

5. Se encontrar, verificar se existem subredes para essa rede;

6. Se existir match com uma das subredes o pacote é encaminhado;

7. Senão existir match com nenhuma subrede, o pacote é descartado e enviado ICMP Destination Unreachable;

2.2.3. Endereçamento Classless

2.2.3.1. Características

As características deste tipo de enderaçamento são:

- Não considera os bits de classe utilizando uma máscara de 32 bits para determinar o endereço de rede;
- Permite encaminhamento mais eficiente por agregação de rotas, designado por CIDR (Classless Internet Domain Routing);
- Tabelas de encaminhamento mais pequenas;
- As rotas são agregadas por grupos de endereços adjacentes;
- Usado pelas tabelas de encaminhamento de ISPs.

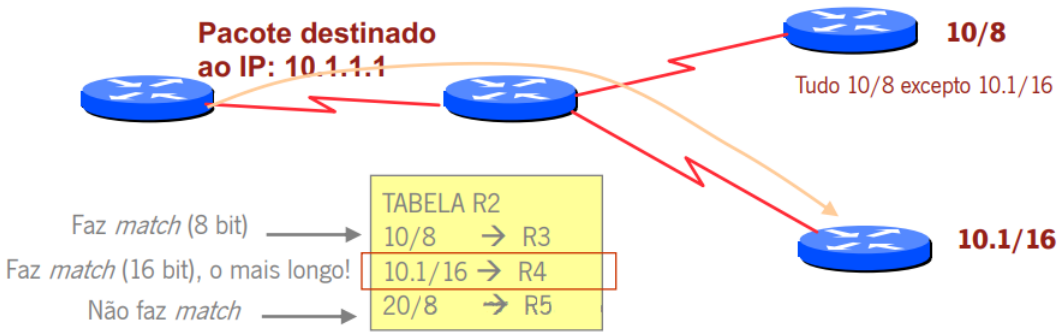
Além disso parte da rede (do endereço da rede) tem comprimento arbitrário (Formato: a.b.c.d/x, em que x é o nº de bits correspondente à parte de rede).

2.2.3.2. Pesquisa de rotas

No Classless, a classe não é relevante, é apenas o melhor match bit a bit:

1. Procurar o prefixo mais longo (mais bits), existente na tabela de encaminhamento que faz match com o endereço;

2. Rota por defeito 0.0.0.0/0 (prefixo mais curto – 0 bit – que faz match com tudo).

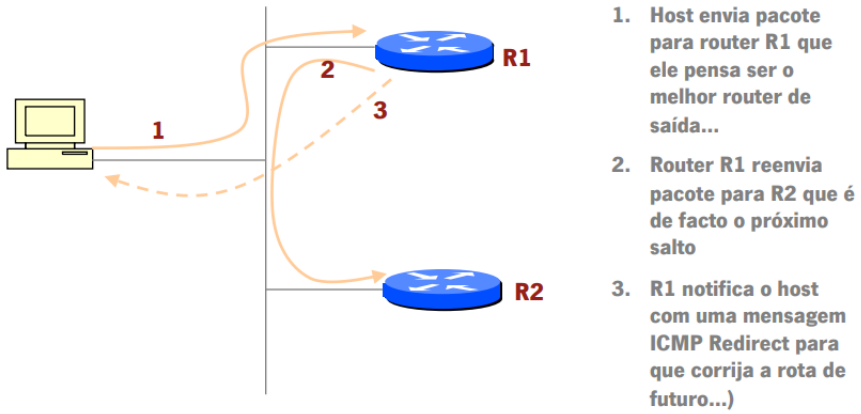


2.3. ICMP

O protocolo ICMP, serve para:

- Reportar erros de encaminhamento (Destination Unreachable, Parameter Problem, Fragmentation Needed, TTL Exceeded);
- Descobrir encaminhadores (Router Discovery, Redirect);
- Testar a acessibilidade (Echo Request, Echo Reply).

Um exemplo do uso do ICMP, pode ser visto na seguinte figura.



2.4. Algoritmos de escalonamento

Dado um conjunto de encaminhadores com ligações de rede a interligá-los, o objectivo do algoritmo de encaminhamento é determinar um "bom" caminho desde a fonte até ao destino.

A topologia de rede é um grafo:

- Os nós do grafo são os encaminhadores;
- Os arcos do grafo são as ligações da rede;
- O custo das ligações pode ser estabelecido em função do atraso, da capacidade, do nível congestão, do custo, da distância, etc;
- Um "bom" caminho significa tipicamente o caminho de "custo mínimo", mas há outras possibilidades.

2.4.1. Classificação dos Algoritmos de Encaminhamento

2.4.1.1. Encaminhamento estático

Neste tipo de encaminhamento, as rotas não mudam c/tempo.

2.4.1.2. Encaminhamento dinâmico

Neste tipo de encaminhamento, as rotas mudam com o tempo (há uma atualização periódica, devido à variação do custo das ligações).

Sendo que a informação pode ser global ou descentralizada.

Se for global:

- Todos os encaminhadores têm um conhecimento completo da topologia e custo das ligações;
- **Algoritmos de estado das ligações (LS- "link state").**

Se for descentralizada:

- Os encaminhadores só conhecem os vizinhos (fisicamente ligados) e o custo das ligações respectivas
- Processo de computação é iterativo, troca de informação com os vizinhos;
- **Algoritmos de vector de distância (DV- "distance vector")**

2.4.2. Algoritmos do Estado das Ligações (LSA)

Neste tipo de algoritmo, todos os nós espalham pela rede o "estado das suas ligações" de forma a construírem a "base de dados topológica":

- Inicialmente necessitam de conhecer apenas os seus vizinhos directos, para quem enviam a identificação de todos os seus vizinhos bem como o custo das ligações que os separam deles;
- Um encaminhador ao receber esta informação actualiza a sua base de dados topológica e re-envia a informação para todos os seus vizinhos;
- Ao fim de algum tempo todos os nós possuem um conhecimento completo da topologia e dos custos de todas as ligações

Sobre esta informação, em cada encaminhador, é utilizado um algoritmo de descoberta dos caminhos de custo mínimo, tipicamente o algoritmo de Dijkstra.

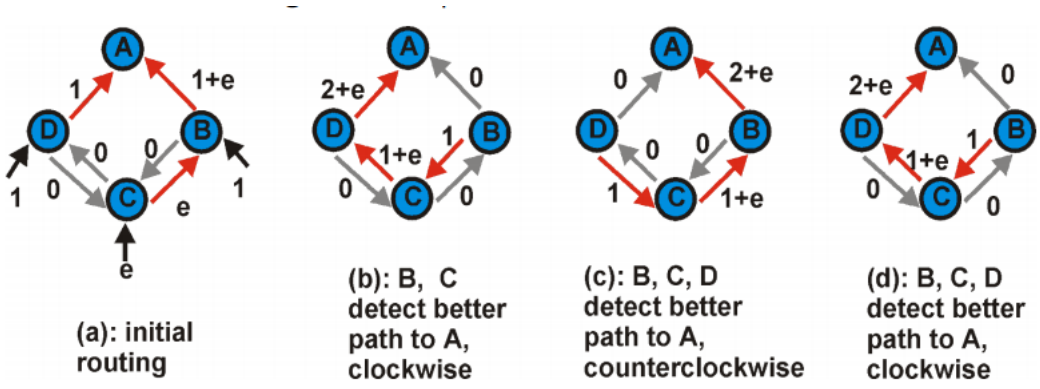
Com o resultado obtido da aplicação do algoritmo de Dijkstra, é preenchida a tabela de encaminhamento.

2.4.2.1. Algoritmo de Dijkstra

Algoritmo Iterativo que ao fim de k iterações consegue descobrir os caminhos de custo mínimo de um determinado nó para k destinos.

No entanto, pode ter falhas:

- Por exemplo, se a métrica reflectir a carga nas ligações, sendo por isso uma métrica assimétrica;
- No exemplo, B e D enviam uma unidade de tráfego para A e C envia e unidades de tráfego também para A.



2.4.3. Algoritmos de Vectores de Distância (DVA)

As características deste tipo de algoritmos são:

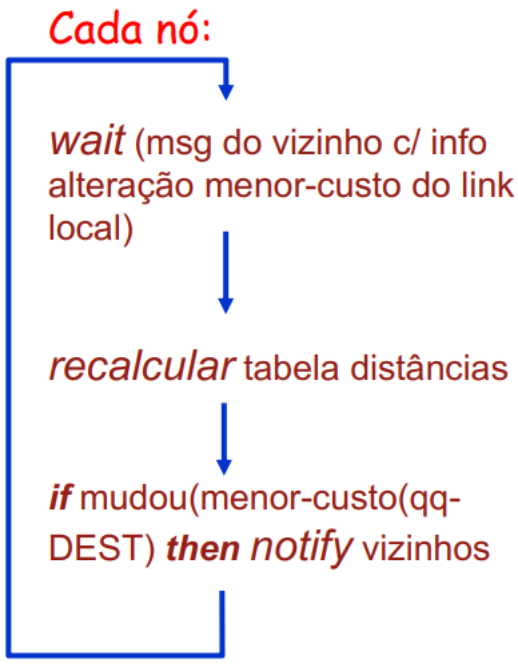
- Cada nó recebe informação de encaminhamento de algum dos seus vizinhos directos, recalcula a tabela de encaminhamento e envia essa informação de volta para os vizinhos;
- O processo continua até que não haja informação de encaminhamento a ser trocada entre nós vizinhos;
- Não exige que os nós estejam sincronizados uns com os outros.

São iterativos e assíncronos, porque cada iteração local é causada por:

- mudança custo link local;
- mensagem do vizinho (vizinho anuncia novo custo).

São distribuídos porque:

- cada nó notifica vizinhos só quando muda o menor custo p/ qq destino;
- vizinhos notificam vizinhos (se necessário).



2.4.4. Estado de Ligação vs Vector de Distância

2.4.4.1. Sobrecarga introduzida pela mensagens de controle

- Nos algoritmos de estado de ligação todos os nós necessitam de conhecer o custo de todas as ligações, por isso sempre que o custo de uma ligação muda, uma mensagem com o novo custo tem que ser enviada para todos os nós.
- Nos algoritmos de vector de distâncias a mudança do custo de uma ligação só provoca o envio de mensagens se resultar na mudança da tabela de encaminhamento

2.4.4.2. Convergência

- Os algoritmos de estado de ligação convergem mais depressa mas, com algumas métricas estão sujeitos a oscilações;
- Em contrapartida os algoritmos de vector de distâncias convergem lentamente, podem apresentar ciclos enquanto não convergem e sofrem do problema da contagem até ao infinito.

2.4.4.3. Robustez

- Nos algoritmos de estado de ligação cada encaminhador calcula a sua tabela de encaminhamento usando a base de dados topológica, de forma independente dos outros encaminhadores. Isso confere a este tipo de algoritmos, robustez
- Nos algoritmos de vector de distâncias se algum encaminhador estiver a calcular mal a sua tabela de encaminhamento, os erros cometidos vão-se propagar aos outros encaminhadores da topologia.

2.4.4.4. Recursos computacionais

- Os algoritmos de estado de ligação são mais exigentes do que os algoritmos de vector de distâncias, quer em termos de memória (base de dados topológica versus tabela de distâncias), quer em termos de capacidade de processamento.

2.5. Sistemas Autónomos

Por razões de escala e de autonomia administrativa, a internet não pode ser encarada como uma topologia de rede onde todos os encaminhadores executam o mesmo algoritmo de encaminhamento para encontrar os melhores caminhos para todos os destinos possíveis. O número de encaminhadores é demasiado grande o que torna a sobrecarga necessária ao cálculo, armazenamento e comunicação da informação de encaminhamento demasiado grande. Idealmente uma organização deveria poder escolher o algoritmo de encaminhamento que deseja utilizar nas suas redes.

Estes problemas são resolvidos agregando os encaminhadores em Sistemas Autónomos (AS – Autonomous Systems):

- Os encaminhadores dentro de um mesmo sistema autónomo utilizam todos o mesmo algoritmo de encaminhamento (LS ou DV) e possuem informação acerca de todos os encaminhadores que fazem parte do sistema autónomo.
- Os protocolos de encaminhamento que se utilizam no interior de um sistema autónomo designam-se por protocolos intra-domínio (intradomain routing protocols) ou internos (IGP - Interior Gateway Protocol

Para interligar os diferentes Sistemas Autónomos entre si é necessário utilizar pelo menos um encaminhador por Sistema Autónomo e com eles constituir uma rede de "nível superior".

Esses encaminhadores além de executarem o protocolo intra-domínio, utilizam um protocolo de encaminhamento inter-domínio (interdomain routing protocol) ou externos (EGP – Exterior Gateway Protocol

As características dos Sistemas Autónomos são:

- Rede ou conjunto de redes que tem uma política de encaminhamento comum, única, clara e bem definida (principalmente para o exterior);
- Normalmente sob controlo administrativo e técnico do mesmo operador;
- Usam normalmente o mesmo protocolo de encaminhamento interno (IGP);
- Identifica-se com o número de sistema autónomo (AS Number).

Os números dos sistemas autónomos:

- Os números de sistema autónomo podem ser privados (AS 64512 até ao AS 65535) ou públicos (atribuídos pelo IANA, ou autoridades regionais, como o RIPE na Europa);
- Usados nas trocas de informação de encaminhamento com os sistemas autónomos vizinhos.

Sistemas autónomos que fazem negócio com a conectividade também se designam por ISP (Internet Service Providers), e estabelecem acordos de parceria entre si (peering agreements):

- Se estão ao mesmo nível, são apenas de troca de rotas ;
- Se não estão ao mesmo nível, o de nível mais baixo (downstream) é cliente, e do nível acima (upstream) é fornecedor.

2.5.1. Protocolos IGP

As características deste tipo de protocolo são:

- Usam processos automáticos de descoberta e troca de informação;
- Todos os encaminhadores são de confiança, sujeitos à mesma administração e às mesmas regras;
- As rotas e outra informação de encaminhamento pode ser difundida livremente entre todos os encaminhadores (todos têm a mesma visão da rede).

Existem neste tipo de protocolo, 2 tipos de categorias:

- **Vector distância** (Bellman-Ford), por exemplo, RIP (só para IP, inspirado no XNS), XNS (Xerox Networking System), CISCO IGRP, EIGRP;
- **Estado da Ligação**, por exemplo, OSPF (só para IP).

2.5.1.1. RIP

As características do RIP, são:

- Componente do código de rede BSD do UNIX: Routed (route management daemon);
- Muito Simples, fácil de implementar, exigindo pouca ou nenhuma configuração;
- Usa uma única métrica: número de saltos (Hop Count);
- Desenhado como IGP para redes de tamanho relativamente limitado (máximo 15 saltos de custo 1);
- Tem que lidar com o problema da contagem até infinito;

- Os endereços são quantidades de 32 bits que podem representar hosts, redes ou sub-redes;
- Não se passa nenhuma informação sobre o tipo de endereço, cabe aos routers analisar os endereços e deduzir que informação é passada.

Cada entidade participante deve manter uma tabela, que não é mais do que uma lista de destinos possíveis e as respectivas distâncias, além de alguma informação de manutenção:

Coluna	**Significado**
Destino	*Endereço IP de destino (32 bit)*
Métrica	*Número de saltos até ao destino*
Interface-saída	*Interface física por onde chegar ao próximo salto*
Próximo-salto	*Endereço do primeiro router no caminho para esse destino*
Flag	*Indicação se a rota foi actualizada há pouco tempo*
Temporizadores	*Conjunto de temporizadores associado a esta entrada*

Procedimento a executar por cada entidade envolvida:

- Manter tabela com uma entrada para cada destino T possível, contendo uma distância D e o primeiro encaminhador G a usar;
- Periodicamente enviar actualizações de rotas a todos os vizinhos: um conjunto de mensagens com todas as entradas T da tabela e respectiva distância D;
- Quando receber uma actualização T de um vizinho G':

  -Desenhado como IGP para redes de tamanho relativamente limitado (máximo 15 saltos de custo 1);

  -Comparar a distância D' com a existente na tabela actual para o mesmo destino T; adoptar a nova rota se D' for menor que a actual;

  -Se o vizinho G' é o mesmo que consta da tabela (G' = G), adoptar a nova rota qualquer que seja a distância.

Temporizadores:

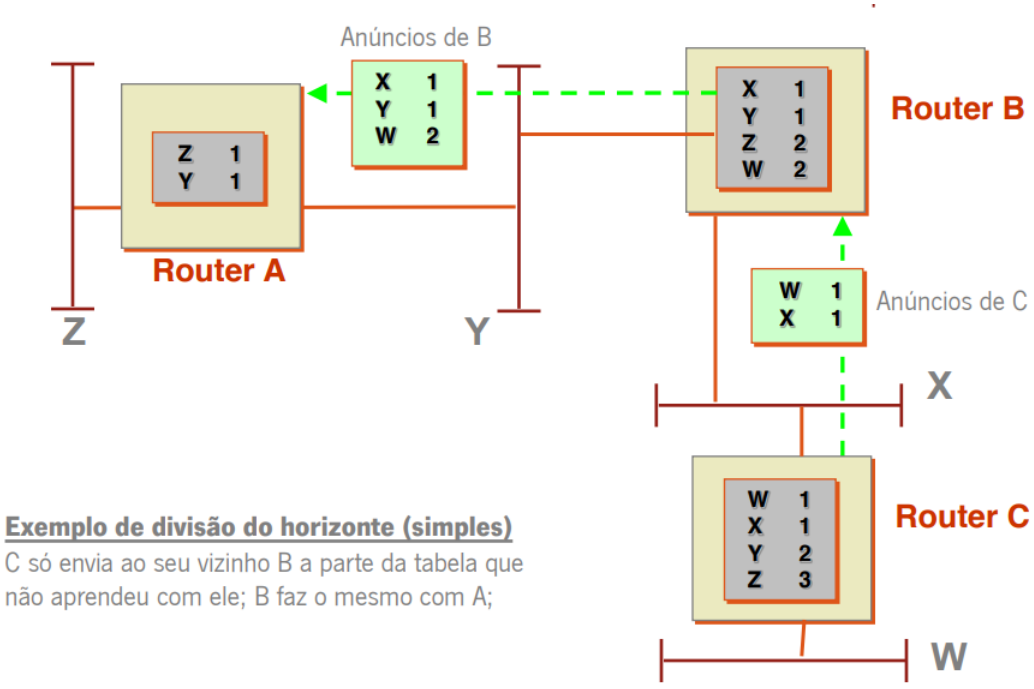
Designação	**Significado**
Temporizador de Actualização (***Update Timer***)	*Marca o intervalo de tempo entre actualizações; normalmente um router envia actualizações aos seus vizinhos a cada 30 segundos*
Temporizador de Invalidação (***Invalid Timer***)	*Intervalo de tempo após o qual a rota é declarada inválida, se não for atualizada; normalmente 180 segundos; deve ser superior a 3 vezes o intervalo de tempo entre atualizações;*
Temporizador de Sustentação (***Holddown Timer***)	*Intervalo de tempo, após invalidação, em que não se aceitam outras rotas para este destino; normalmente 180 segundos;*
Temporizador de Limpeza (***Flush Timer***)	*Tempo que uma rota deve ser mantida na tabela (desde a última atualização), até ser eliminada; deve ser superior a Invalid + Holddown;*

Se um encaminhador não receber nenhum advertisement de um dos vizinhos durante 180 segundos –> vizinho/ligação é declarado morto:

- As rotas via vizinho são tornadas invalidas;
- São enviados novos advertisements aos vizinhos;
- Os vizinhos por sua vez enviam novos advertisements (se as suas tabelas de encaminhamento mudarem);
- Desta forma a falha é propagada rapidamente para toda a rede.

2.5.1.1.1. Divisão do horizonte (método simples)

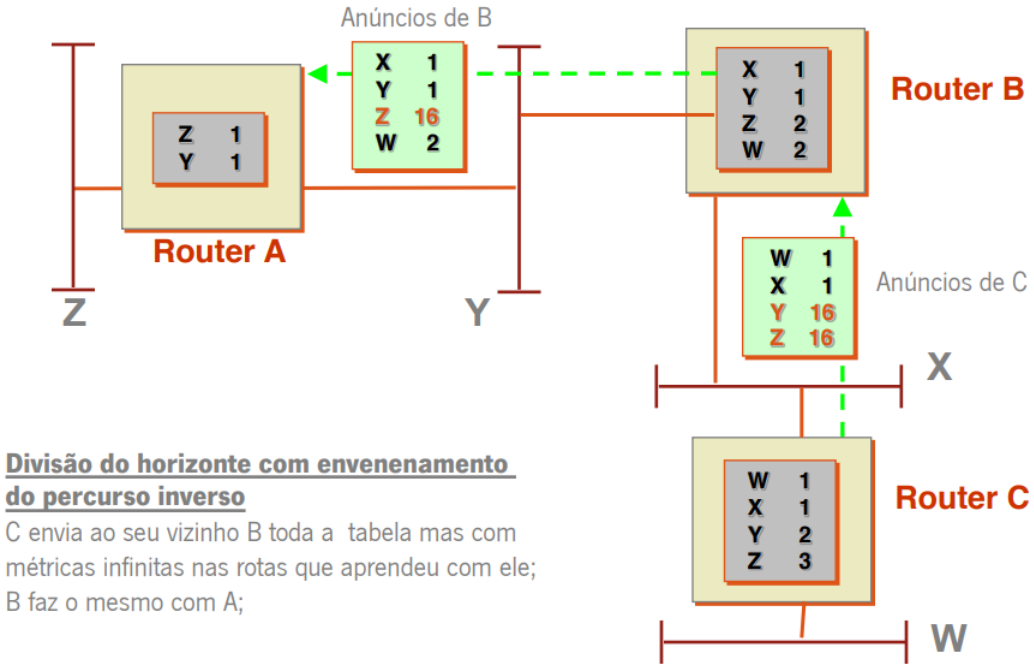
- Não anunciar a um vizinho directo uma rota por ele aprendida;
- Não evita ciclos mais longos.



Exemplo de divisão do horizonte (simples)
C só envia ao seu vizinho B a parte da tabela que não aprendeu com ele; B faz o mesmo com A;

2.5.1.1.2. Divisão do horizonte (com envenenamento do percurso inverso)

- Anunciar aos vizinhos rotas que passam por eles, mas com métrica "infinita".



Divisão do horizonte com envenenamento do percurso inverso

C envia ao seu vizinho B toda a tabela mas com métricas infinitas nas rotas que aprendeu com ele; B faz o mesmo com A;

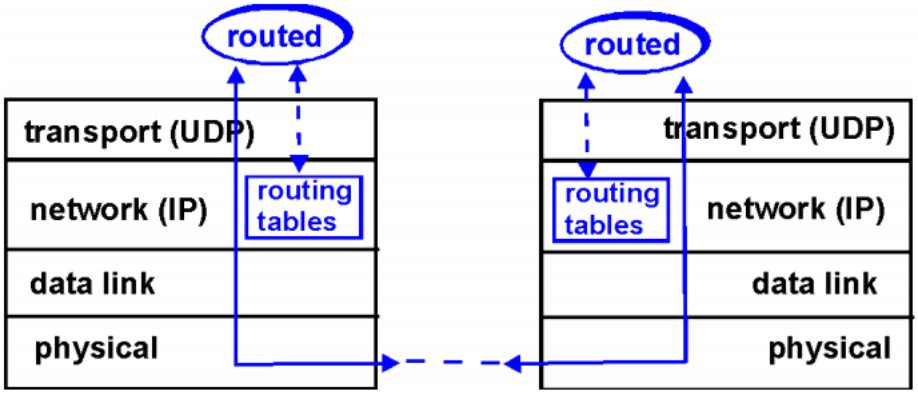
2.5.1.1.3. Actualizações forçadas (triggered) + temporizador de sustentação (hold-down timer)

- Sempre que uma rede fica inactiva, o router dispara o envio daquela rota como estando inacessível.

2.5.1.1.4. Tabela de Encaminhamento no RIP

No RIP de um sistema UNIX as tabelas de encaminhamento são geridas por um processo chamado routed (daemon), ou seja, ao nível da aplicação.

Os advertisements são enviados em pacotes UDP periodicamente.



2.5.1.1.5. Desvantagens do RIP

- Imaginando uma tabela com 100 rotas, enviada de 30 em 30 segundos, cabendo 25 rotas em cada datagrama UDP (max. 512 bytes), são 8 datagramas por minuto.
- As distâncias são medidas em saltos com pouca relação com os custos... e a métrica "infinito" são 16 saltos apenas;
- RIPv1 é classfull, RIPv2 suporta VLSM (Variable-Length Subnet Mask), mas de adoção não generalizada;
- RIPv1 não suporta autenticação (perigo de injeção de rotas falsas)
- Apesar dos métodos de divisão do horizonte, atualizações forçadas etc., a convergência pode ser lenta.

2.5.1.2. OSPF

Os protocolos baseados estado da ligação convergem mais rapidamente e são considerados mais estáveis (menos ciclos) que os baseados no vetor distância.

Cada nó conhece a topologia completa e calcula os caminhos mais curtos.

As características do OSPF são:

- Cada nó coleciona o estado de todas as suas ligações num LSA (Link State Advertisement);
- Envia para todos os outros nós da rede (flooding), de forma fiável, de modo a que todos construam o mapa da rede;
- Com o mapa, cada nó calcula os caminhos mais curtos usando o algoritmo de caminhos mais curtos (SPF) de Dijkstra;
- Usa uma métrica baseada na largura de banda;
- suporta multipath routing: distribuição equitativa de carga no máximo por 4 rotas de igual custo por destino);

2.5.1.2.1. Métrica OSPF

Embora possa ser qualquer valor que o administrador entenda usar, normalmente o custo é proporcional à largura de banda do link:

$$Metrica_{OSPF} = \frac{10^8}{LarguraBanda_{(bps)}}$$

Menor largura de banda corresponde a um maior custo.

64 Kbps serial link –>1562

4 Mbps token ring –>25

Ethernet –>10

2.5.1.2.2. Bases de Dados OSPF

- Base de dados de Adjacências (Adjacencies Database) – tabela com todos os vizinhos com quem um router estabeleceu uma comunicação bidireccional (vizinhança OSPF);
- Base de dados de LSA (Link State Database) – base de dados topológica com o mapa da rede, idêntica em todos os routers;
- Base de dados de reenvio (forwarding database) – com as melhores rotas calculadas para cada destino.

2.5.1.2.3. Protocolos OSPF

- **Hello Protocol** – mantém actualizada a base de dados de adjacências com a troca de pacotes "Hello";
- **Exchange Protocol** – permite a sincronização de um novo vizinho quando ele é descoberto;
- **Reliable Flooding** – difunde de forma fiável os LSA por todos os routers; cada pacote com LSAs deve ser confirmado e enviado por todos os links excepto aquele por onde foi recebido;

Cada router constrói um ou mais LSA (Link State Advertisements), com o estado das suas ligações, que difunde para todos os outros routers de forma fiável:

- O router que origina o LSA fica responsável por ele;
- Todos os routers recebem os mesmos LSA e constroem o mesmo mapa da rede;

Os LSAs são reenviados periodicamente (30 em 30 minutos) ou quando há alterações, usando o protocolo RELIABLE FLOODING;

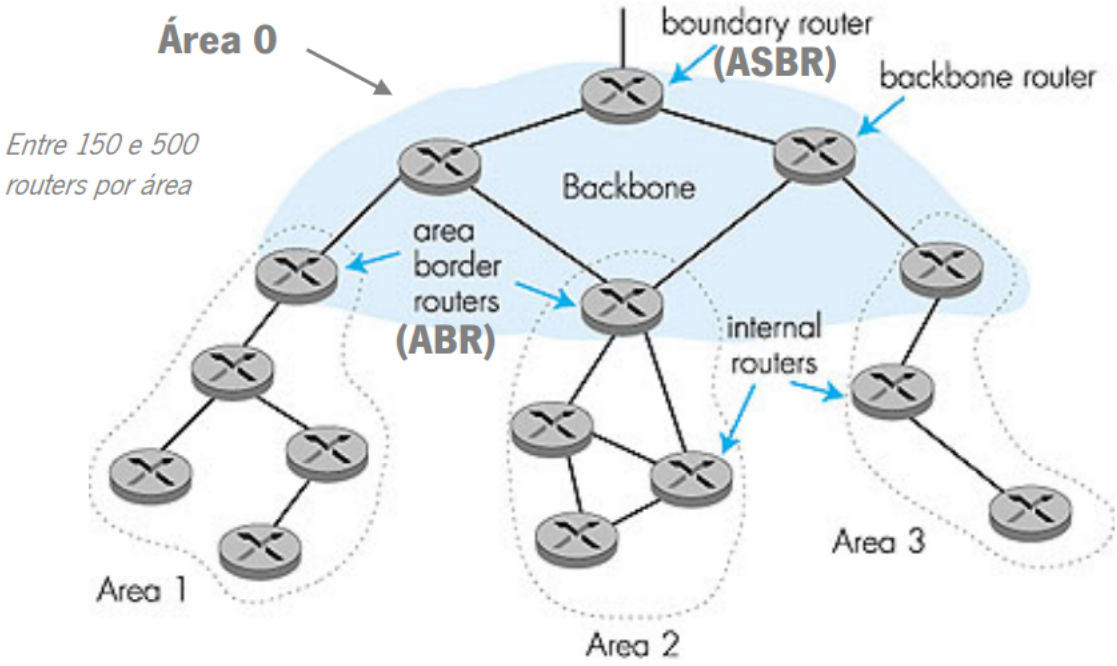
Para reduzir o tamanho das tabelas e o consumo de memória e CPU, os domínios de encaminhamento OSPF podem dividir-se em áreas, numa hierarquia de dois níveis(todas as áreas estão ligadas a uma área de backbone: área 0.0.0.0 (32bits)).

Os routers ligados a mais que uma área designam-se por ABR (area border routers) e podem trocar informação sumariada de uma área para a outra (summary-LSAs).

2.5.1.2.4. Tipos de pacotes OSPF

- Type 1: Hello; pacote de 64 byte enviado a intervalos regulares (5s) para manter uma relação activa;
- Type 2: DBD (Database Description); sumário do conteúdo da base dados com os LSA, enviado a novo router descoberto;
- Type 3: LSR (Link-State Request); Pedido de informação mais específica sobre um link da base de dados de LSA do vizinho;
- Type 4: LSU (Link-State Update); transporta LSA para os vizinhos; por exemplo como resposta a um LSR;
- Type 5: LSAck (Link-State Acknowledgement); confirma a recepção de um LSA; todas as actualizações tem de ser confirmadas;

- Encaminhamento Hierárquico a 2 níveis dentro de um Sistema Autónomo:



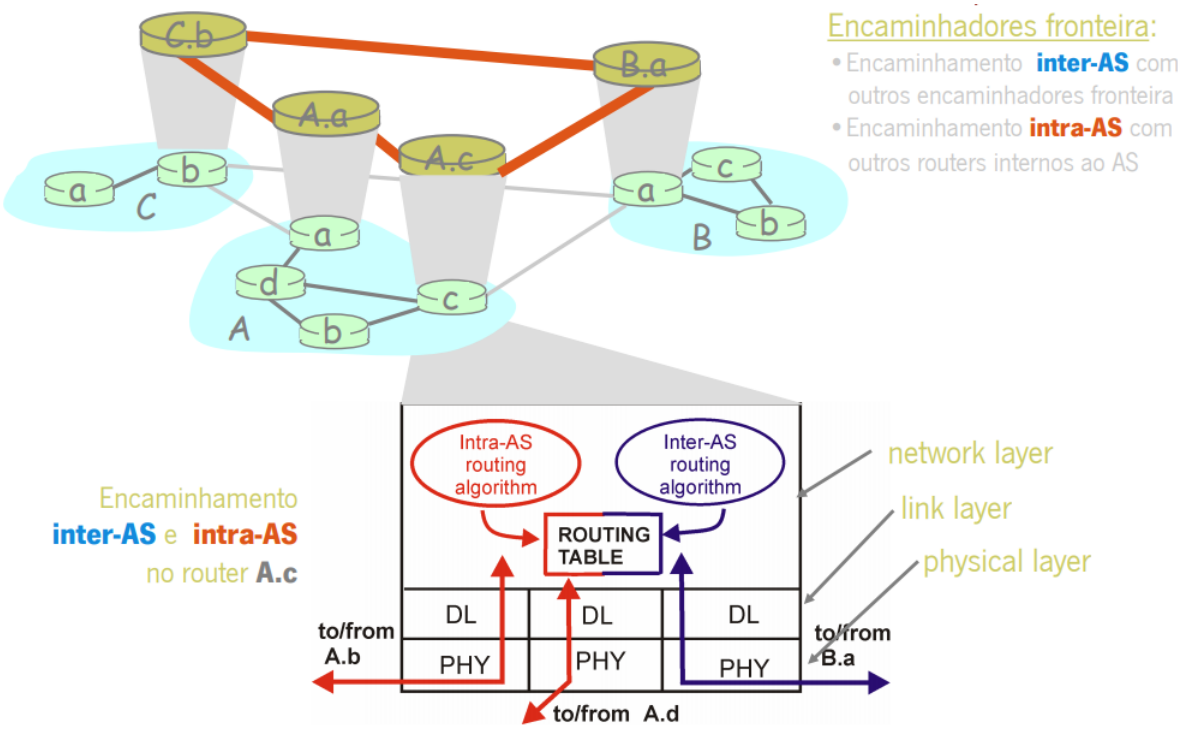
2.5.1.2.5. Tipos de Routers OSPF

- ABR (Area Border Routers) – ligam a área ao backbone (originam e recebem LSA sumariados);
- BR (Backbone Routers) e IR (Internal Routers) – só trocam LSAs dentro da área;
- ASBR (Autonomous System Boundary Routers) – trocam informação com outros sistemas externos (originam LSA externos).

2.5.1.2.6. Vantagens da abordagem hierárquica

- Link-state advertisements difundidos apenas no interior de cada área;
- Cada nó dentro de uma área possui apenas um conhecimento detalhado da topologia da sua área;
- Para as redes de outras áreas é conhecida apenas a direção a tomar (via caminho mais curto).

2.5.2. Encaminhamento Intra-AS e Inter-AS

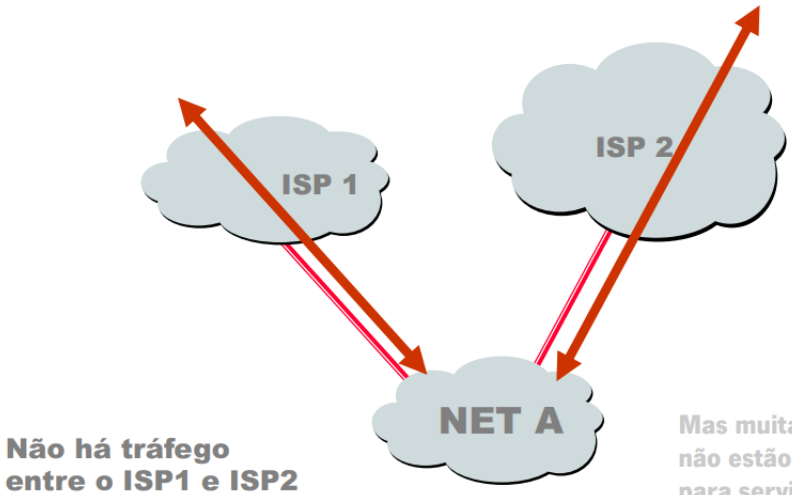


2.5.2.1. Encaminhamento inter-AS

Requisitos para o encaminhamento inter-AS:

- Deve ser escalável para o tamanho da Internet;
- Foco na conectividade e não a optimização;
- Deve usar técnicas de agregação de endereços para minimizar tamanho das tabelas e volume de tráfego de controlo;
- Deve permitir encaminhamento baseado em políticas entre sistemas autónomos;
- Política refere-se à escolha arbitrária de uma rota num menu de rotas (olhando para os atributos dessas rotas);

- **AS de trânsito *versus* AS multi-homed**

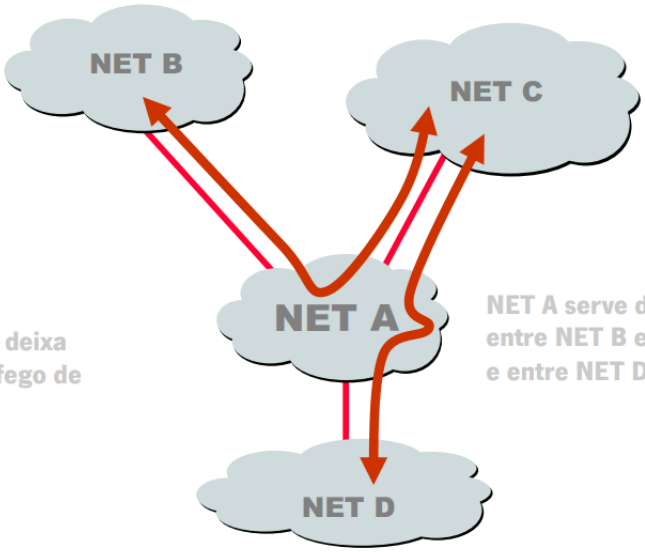


Os ISP possuem redes de **trânsito**

Não há tráfego entre o ISP1 e ISP2 que passe através de NETA

Mas muitas organizações não estão vocacionadas para servir de trânsito para terceiros...

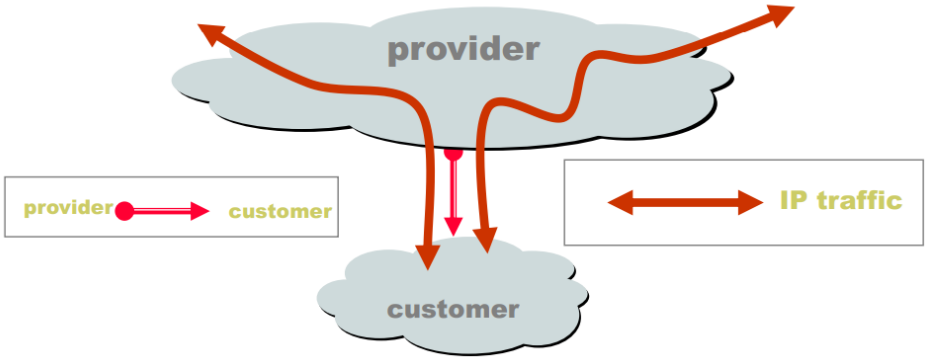
- **Muitos AS são *selectivamente* de trânsito**

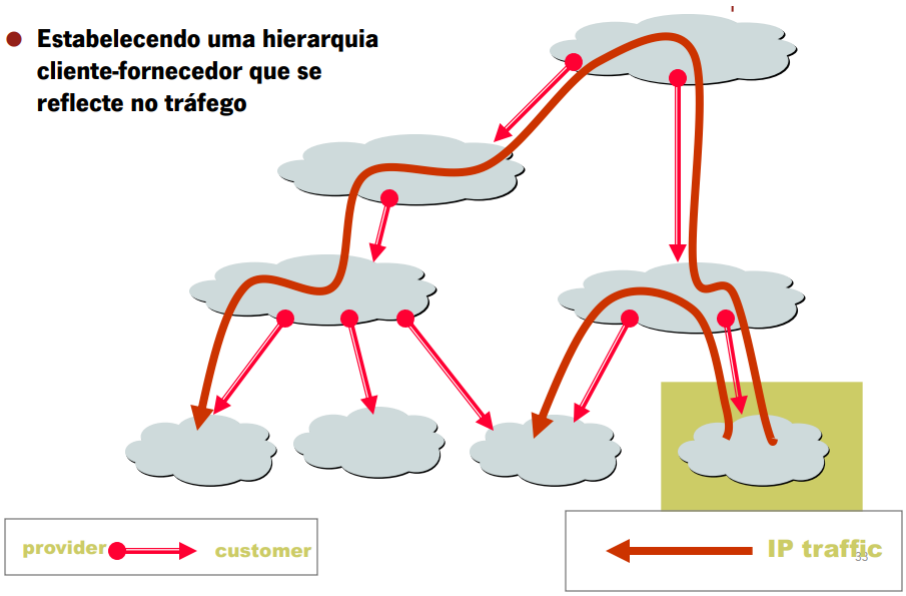


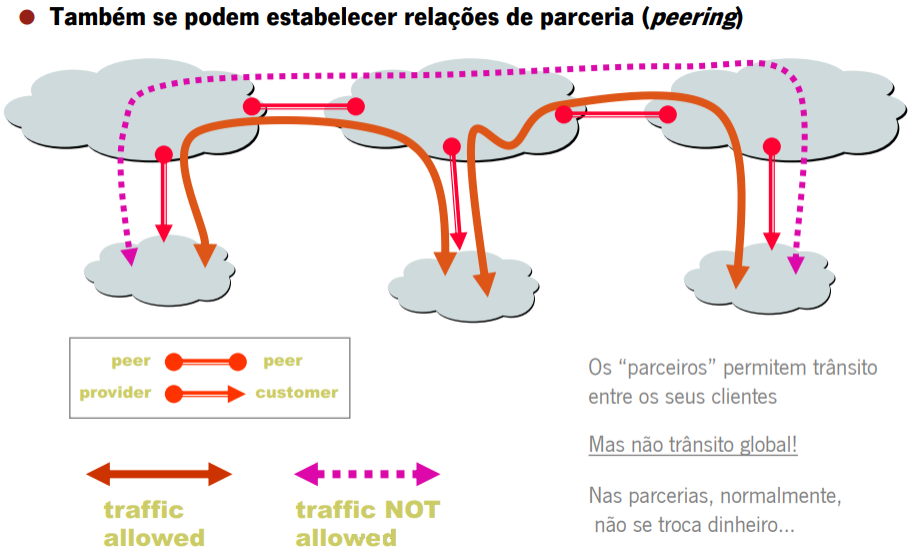
NET A não deixa passar tráfego de B para D

NET A serve de trânsito entre NET B e NET C e entre NET D e NET C

- **Os clientes pagam acesso aos fornecedores...**





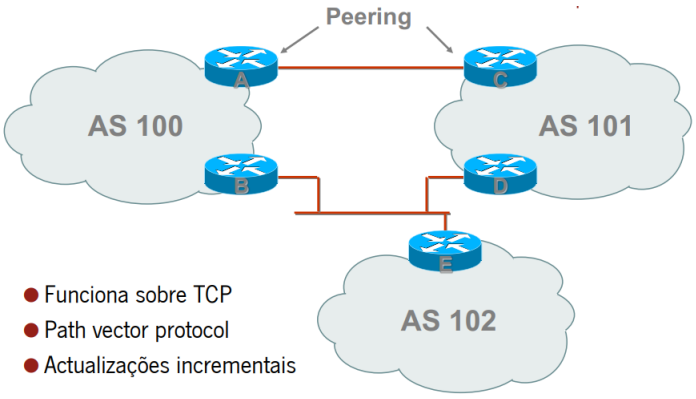


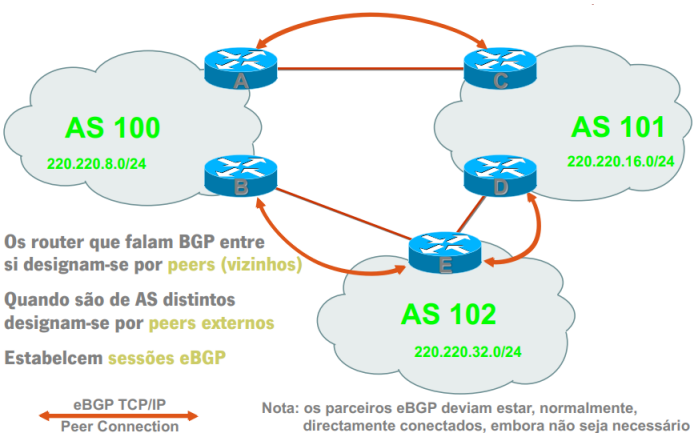
2.5.3. Protocolo BGP

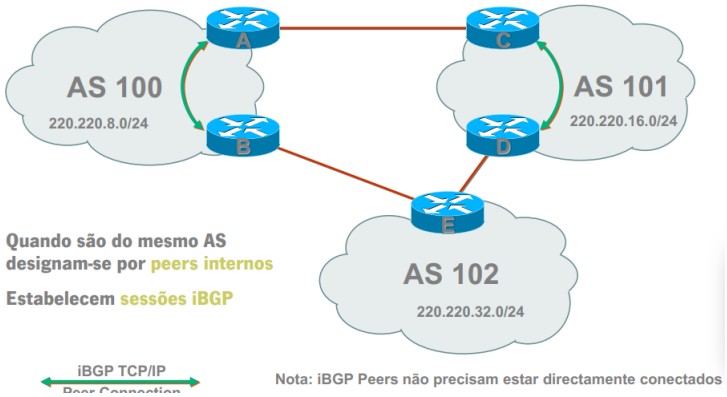
- É um protocolo de encaminhamento baseado em políticas;
- É o standard de facto da Internet no que toca a protocolos EGP;
- Relativamente simples, mas de configuração complexa, sendo que os erros cometidos podem ser vistos e mesmo causar problemas ao mundo inteiro;

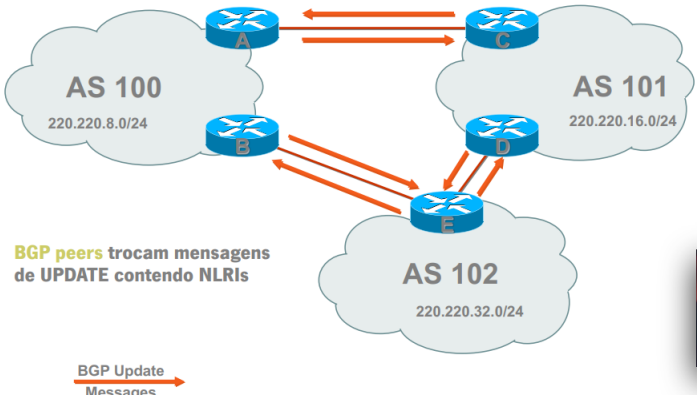
2.5.3.1. Funcionamento básico

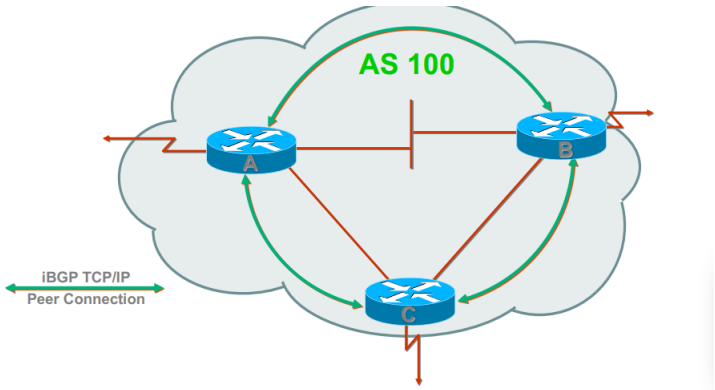
- Cada AS gera um conjunto de NLRI (network layer reachability information), por exemplo prefixos IPv4 e respectivas máscaras;
- Os routers fronteira (border routers) trocam os NLRI do seu AS com os routers fronteira parceiros (BGP peers) dos AS vizinhos;
- Podem ter mais do que um caminho por prefixo... mas escolhem e instalam o melhor na tabela de reenvio.
- Cada rede de SAs é representada por um grafo (BGP Tree);
- O BGP utiliza um algoritmo DV mas todas as alternativas são armazenadas numa tabela BGP;
- Cada par de SAs conectados diretamente através de uma ligação BGP formam uma relação de vizinhança BGP;
- O estabelecimento de uma ligação de vizinhança BGP ativa é sempre manual, por deliberação do administrador de rede;
- Suporta utilização de rotas alternativas para um mesmo destino;
- Suporta Políticas de Encaminhamento (Policy Routing), muito úteis para encaminhamento entre ISPs, por exemplo.



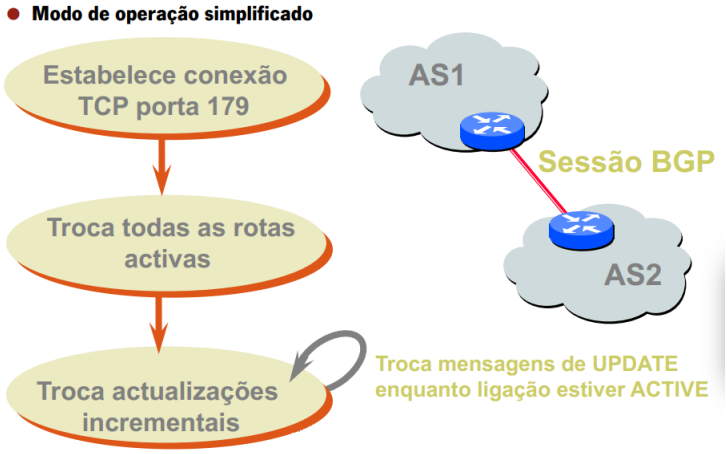

AS 100
220.220.8.0/24
AS 101
220.220.16.0/24
AS 102
220.220.32.0/24
A
B
C
D
E
Os router que falam BGP entre si designam-se por peers (vizinhos)
Quando são de AS distintos designam-se por peers externos
Estabelcem sessões eBGP
eBGP TCP/IP Peer Connection
Nota: os parceiros eBGP deviam estar, normalmente, directamente conectados, embora não seja necessário


AS 100
220.220.8.0/24
AS 101
220.220.16.0/24
AS 102
220.220.32.0/24
A
B
C
D
E
Quando são do mesmo AS designam-se por peers internos
Estabelecem sessões iBGP
iBGP TCP/IP
Nota: iBGP Peers não precisam estar directamente conectados


AS 100
220.220.8.0/24
AS 101
220.220.16.0/24
AS 102
220.220.32.0/24
A
B
C
D
E
BGP peers trocam mensagens de UPDATE contendo NLRIs
BGP Update Messages



Cada router iBGP tem de estabelecer uma sessão iBGP com todos os outros routers iBGP num sistema autónomo!!! (*BGP Mesh*)



Quatro tipos de mensagens:

- **Open**: Estabelece uma conexão entre dois routers BGP;
- **Keep Alive** : Trocada regularmente para manter conexão válida;
- **Notification** : Termina uma sessão BGP estabelecida;
- **Update** : Anuncia novas rotas ou retira (withdraw) rotas anunciadas previamente.

$$\text{anúncio} = \text{prefixo} + \text{atributos}$$

Utiliza o conceito de vector de caminhos (path-vector) que permite evitar ciclos em topologias complexas.

Cada rota tem o caminho completo definido como a sequência de números dos AS que é preciso atravessar (AS Path).

No nível inter-AS, o caminho mais curto pode não ser o preferido por razões políticas, de segurança ou de custo:

- Routers diferentes podem ter diferentes preferências (políticas) => os pacotes encontram diferentes políticas no seu percurso;
- BGP permite a existência de atributos para AS e para caminhos que permitem implementar políticas (policy-based routing).

2.5.3.2. Atributos das rotas BGP

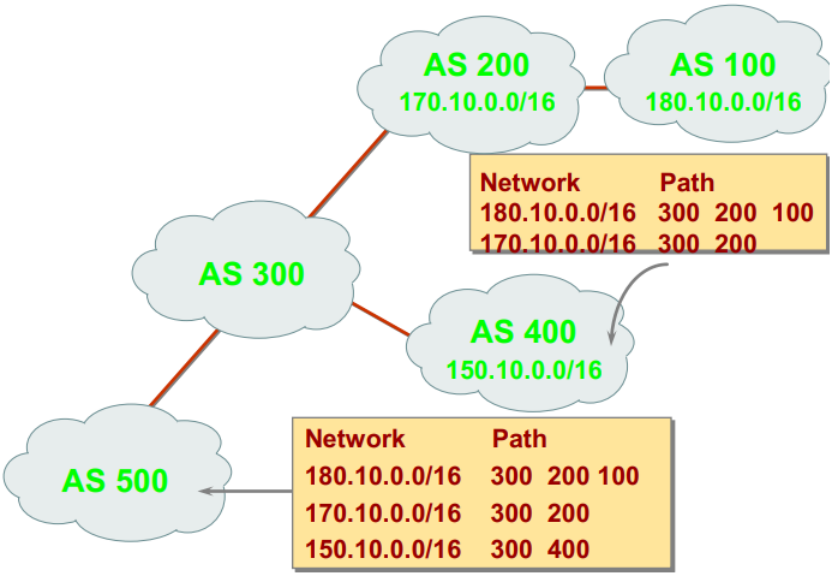
2.5.3.2.1. Atributo ORIGIN (bem conhecido, obrigatório)

Caracteriza a origem da rota, ou seja, como foi injectada no BGP:

- IGP – internas, introduzidas pelo comando "network X.Y.W.Z";
- EGP – externas, redistribuídas pelo protocolo EGP;
- Incomplete – internas, redistribuídas dos protocolos intra-domínio (redistribute OSPF);
- IGP < EGP < Incomplete

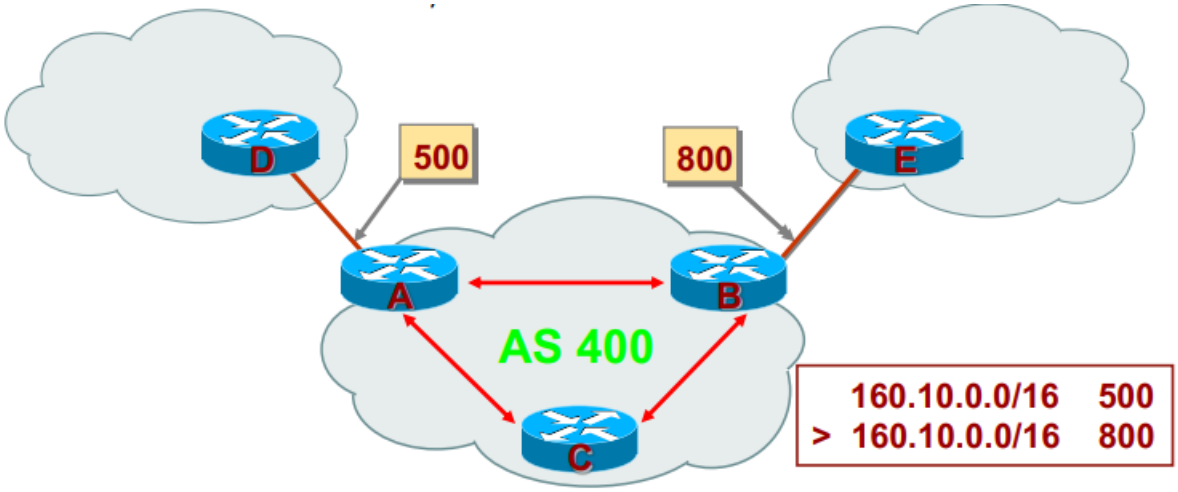
2.5.3.2.2. Atributo AS_PATH (bem conhecido, obrigatório)

- Enumeração completa dos nºs de sistema autónomos no percurso até ao destino;
- São usados na selecção,
- Permitem detectar ciclos.



2.5.3.2.3. Atributo LOCAL_PREF

- Só tem significado dentro de um sistema autónomo;
- É definido no router fronteira de entrada no domínio, como forma de definir uma preferência quando há múltiplos pontos de saída;
- Prefere-se o valor maior, sendo o default=100.



2.5.3.2.4. Atributo WEIGHT

- Definido pela CISCO (só para routers CISCO);
- Define um grau de preferência por rotas (valor maior) com significado apenas dentro do próprio router;
- Só influencia o algoritmo de selecção do próprio router.

2.5.3.3. Funções dos encaminhadores BGP

- Recebe e filtra anuncios de rotas dos vizinhos directamente ligados a ele;
- Selecciona rotas;
- Envia anuncios de rotas para os seus vizinhos.

2.5.3.4. Algoritmo BGP de seleção da melhor rota:

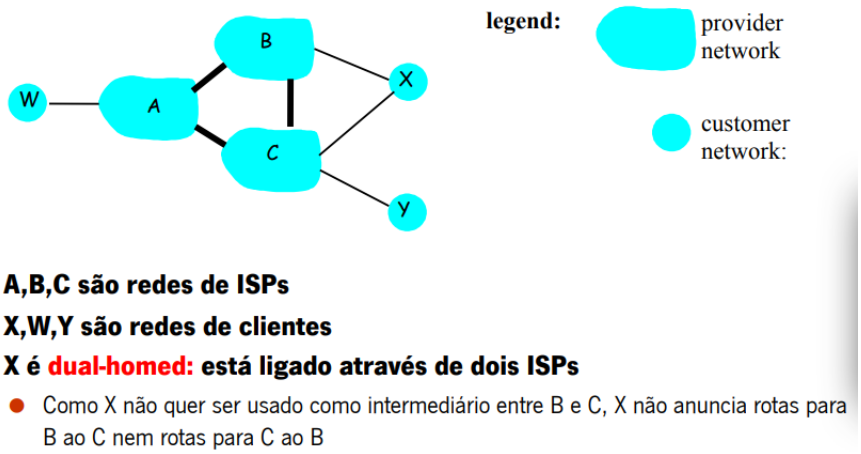
1. Verificar se NEX_HOP está acessível;
2. Verificar se o AS não consta do AS_PATH (deteção de ciclos);
3. Escolher a rota com valor WEIGHT mais alto (só nos CISCO);
4. Se WEIGHT igual, escolher rota com maior LOCAL_PREF);
5. Se LOCAL_PREF igual, escolher rota com AS_PATH mais curto;

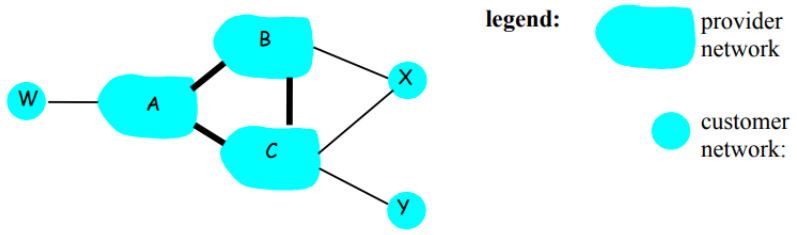
6. Se todas são externas, com o mesmo AS_PATH length, preferir a rota de menor ORIGIN (IGP < EGP < INCOMPLETE);
7. Se ORIGIN iguais, preferir o de menor MED;
8. Se MED igual, preferir a rota cujo NEXT_HOP esteja mais próximo;
9. Preferir rotas aprendidas por eBGP às iBGP;
10. Se ainda houver empate, escolher a rota recebida do router com menor identificador (ROUTER-ID – que é um endereço IP - menor);

2.5.3.5. Encaminhamento Inter-domínio na Internet: BGP

Supondo que o encaminhador X envia os seus caminhos para o seu par encaminhador W:

- W pode ou não selecionar os caminhos divulgados por X dependendo de custo, políticas (não usar determinados ASs para encaminhar tráfego), prevenção de ciclos, etc;
- Se W selecionar os caminhos anunciados por X, então: Path (W,Z) = W, Path (X,Z)
- Nota: X pode controlar o tráfego de entrada controlando as rotas que anuncia aos seus pares (Por exemplo, se não quiser encaminhar tráfego do encaminhador Z não divulga rotas para o encaminhador Z).





- **A anuncia para B o caminho AW**
- **B anuncia para X o caminho BAW**
- **B não anuncia para C o caminho BAW**
  - Nem W, nem C são clientes de B
  - B pretende forçar C a encaminhar para W, via A
  - B pretende encaminhar apenas para/de os seus clientes!

2.5.3.6. Encaminhamento Intra-Domínio versus Encaminhamento Inter-domínio

2.5.3.6.1. Políticas

- No encaminhamento inter-domínio é fundamental ter o controle sobre a forma como o encaminhamento é efectuado. Por exemplo, a decisão de não encaminhar determinado tipo de tráfego através de um AS, tem de ser possível;
- No encaminhamento intra-domínio as decisões "políticas" de encaminhamento assumem pouca importância, uma vez que todos os nós estão sob a mesma autoridade administrativa.

2.5.3.6.2. Escala

- O encaminhamento hierárquico reduz o tamanho das tabelas de encaminhamento e a quantidade e tamanho das mensagens de actualização da informação de encaminhamento.

2.5.3.6.3. Desempenho

- No encaminhamento intra-domínio o desempenho é a preocupação principal, ao passo que no encaminhamento inter-domínio tem um papel secundário, sendo ultrapassado pelas políticas.