

Protocolos em Redes de Dados

- Aula 12 -

MOSPF **Difusão Inter-Domínio (MSDP, MBGP e BGMP)**

Luís Rodrigues

ler@di.fc.ul.pt

DI/FCUL

- ⑥ Extensões OSPF para difusão.
- ⑥ Aplicação de difusão em grande escala.

MOSPF

- ⑥ Extensões difusão para o OSPF.
- ⑥ Usado internamente a organizações.
- ⑥ Potencialmente mais estável por ser baseado no “estado-dos-elos” e não em vectores de distância.
- ⑥ Baseia-se no cálculo de árvores com raiz na fonte.
- ⑥ Utiliza um novo anúncio, o `group-membership-LSA`.

Optimização na procura por difusão

- ⑥ A difusão pode ser usada para procurar servidores:
 - △ Procurar o membro mais próximo de um grupo em difusão.
- ⑥ Ir realizando pedidos em difusão, usando um TTL sucessivamente maior até obter uma resposta.
 - △ MOSPF otimiza esta procura, só encaminhando os pacotes se o TTL permitir chegar ao destino.

Reconfiguração

- ⑥ 1. S1 começa a emitir:
 - △ O encaminhador tenta calcular uma árvore.
 - △ Como não encontra membros descarta o pacote.
- ⑥ 2. Aparece um receptor:
 - △ “GMP Host Membership report”; “designated router” envia `group-membership LSA`; todos ficam a saber que há um membro; próximo pacote já encontra membros; uma árvore é calculada.
 - △ Se um encaminhador fizer interface com um “stub network” faz um anúncio agregado.

Reconfiguração (cont.)

- ⑥ 3. Quando um elo falha as árvores são re-calculadas quando se encaminha de um novo pacote.
- ⑥ 4. Alguém sai:
 - △ Descoberto com um “leave” ou ausência de resposta a um inquérito.
 - △ Novo anúncio `group-membership-LSA`.

Anúncios “group-membership-LSA”

LSA AGE		
Opcoes	Tipo=6	group-membership LSA
Link State ID		endereço multicast
Encaminhador que faz o anuncio		
Numero de sequencia do anuncio		
Comprimento		
Referenced LS Type		Tipo de anuncio associado
Anuncio de rede associado		Por exemplo, anuncio feito pelo de

Cálculo de árvores de encaminhamento

- ⑥ 1. Quantas árvores existem?
 - △ Uma para cada grupo/fonte.
- ⑥ 2. Quando são calculadas?
 - △ A pedido, quando chega um pacote e a rota não existe.
- ⑥ 3. O que é guardado em cada router?
- ⑥ 4. Qual a diferença para o ponto-a-ponto?

A “cache” de encaminhamento para difusão

- 6 Cada entrada guarda:
 - ▲ O encaminhador de onde os pacotes devem vir.
 - ▲ As interfaces para as quais o pacote é re-encaminhado.
- 6 Manutenção da cache:
 - ▲ Quando há alterações, a entrada é apagada.
- 6 Cálculo das árvores:
 - ▲ Enquanto no ponto-a-ponto a raiz é o próprio encaminhador na difusão a raiz é sempre a fonte.
 - ▲ Regras para desempate em caminhos de custo semelhante.

Encaminhamento hierárquico em difusão

- 6 Os anúncios de filiação são recebidos pela “área espinha” (backbone) mas não são propagados para as restantes áreas.
- 6 Os encaminhadores fronteira de uma área não necessitam de saber se existem membros noutras áreas:
 - ▲ Simplesmente recebem todos os pacotes em difusão e enviam-nos para a “área espinha” (backbone).
- 6 Por sua vez, os encaminhadores da espinha enviam os pacotes para as áreas onde existem membros.
- 6 Dentro de cada área seguem a árvore.

Multicast Internet Exchange (MIX)

- 6 Pontos onde fornecedores de serviço com suporte para multicast podem trocar tráfego em difusão, assim como sinalização de suporte à difusão.
- 6 Permitem manter uma malha com suporte a difusão sem recorrer ao encapsulamento em tráfego ponto-a-ponto.

Inter-Domain IP Multicast

- 6 Problema:
 - ▲ Como expandir a arquitectura baseada no MBONE (mais MOSPF)?
- 6 Solução intermédia:
 - ▲ Tentar usar a infra-estrutura existente com pequenas extensões.
- 6 Solução de longo prazo:
 - ▲ Definir um protocolo semelhante ao BGP para multicast.

Problema da descoberta de fontes

- Os protocolos esparsos usam o conceito de *rendez-vous point* (RP).
- Como estender este conceito para um sistema contendo vários A.S.?

Problema da descoberta de fontes

- Um único RP possui várias desvantagens:
 - Quem não possui o RP fica dependente de outros A.S.
 - Quem tem pode ser obrigado a fazer trânsito sem ter emissores nem receptores.
- Solução: possuir um RP em cada A.S.

Multicast Source Discovery Protocol (MSDP)

- Protocolo que permite aos RPs de cada A.S. trocarem informação acerca das fontes de informação.
- Uma fonte contacta o RP local e este distribui essa informação por uma árvore de RPs usando ligações TCP dedicada para o efeito.
- Os RPs que tiverem membros no seu domínio registam-se na fonte de modo a participarem na árvore de escoamento.

MSDP

- Cada *rendezvous point* (RP) é configurado com a identificação dos RPs com os quais vai estabelecer trocas de informação.
- Quando uma nova fonte se regista, esta informação é disseminada para os restantes RPs numa mensagem designada por “Source-Active” (SA).
- As mensagens SAs são propagadas deste modo por uma árvore de escoamento estabelecida entre os RPs.
 - Esta árvore é estabelecida, seguindo regras de RPF para decidir quais os anúncios que devem ser inundados e quais os que devem ser descartados.

Fragilidades

- ⑥ Ramen worm:
 - △ *Worm* que atacava a rede fazendo *scan* a todos os endereços IP.
 - △ Um *bug* no código fazia com que os endereços IP multicast também fossem testados.
 - △ Cada teste gerava uma mensagem de “Source Active” (podendo original mais de 4.000 por minuto).
- ⑥ Solução de emergência: limitar administrativamente a frequência máxima dos anúncios:
 - △ Favorece ataques de negação de serviço.

Rotas para as fontes

- ⑥ Os receptores podem optar por tentar juntar-se a uma árvore de escoamento com raiz na fonte.
- ⑥ Para isso irá enviar mensagens de “Join” directamente para fonte.
- ⑥ Tal obriga a que os encaminhadores no percurso tenham rotas para essa fonte *por* encaminhadores que suportam difusão.

Multiprotocol extensions to BGP

- ⑥ Existe uma extensão ao BGP que permite trocar rotas referentes a vários protocolos.
 - △ BGP+ (RFC 2283).
- ⑥ Estas extensões são usadas para transmitir informação de encaminhamento para as fontes de difusão distinta da usada para tráfego unicast.
 - △ Indicando percursos com suporte para difusão.

Multiprotocol extensions to BGP

- ⑥ *Não* troca informação acerca de endereços de difusão!
- ⑥ Os anúncios BGP trocam informação acerca do endereço das fontes e do *next hop* para chegar a esse endereço.
- ⑥ Permite que cada sistema autónomo mantenha duas tabelas de rotas distintas:
 - △ Uma usada para encaminhar tráfego ponto-a-ponto.
 - △ Outra usada para manter a tabela RPF para encaminhamento em difusão.

Múltiplas tabelas dentro do sistema autónomo

- 6 Em sistemas baseados em IS-IS, é possível usar as extensões multi-protocol para manter duas tabelas distintas dentro do sistema autónomo.
 - △ M-ISIS.

Problemas da combinação PIM-SM/MBGP/MSDP

- 6 Tempo de criar a árvore é muito lento.
- 6 Os primeiros pacotes podem perder-se por a árvore ainda não estar estabelecida.
- 6 Se a fonte gerar uma pequena rajada, todos os pacotes podem ser perdidos.

Source-Specific Multicast

- 6 Solução pragmática que assume que a subscrição é feita para um fonte específica.
 - △ As mensagens de *subscribe (join)/unsubscribe (leave)* identificam um par (S,G).
- 6 Elimina a necessidade de possuir um rendez-vous point.
- 6 O endereço de grupo serve apenas para distinguir diferentes fluxos de dados com origem na mesma fonte.

Border Gateway Multicast Protocol (BGMP)

- 6 Protocolo em estudo para suportar a difusão.
 - △ Último draft: 19 de Janeiro de 2004.
- 6 Cria uma árvore de domínios.
 - △ Tipicamente o domínio que inicia a sessão e que é responsável pelo endereço de difusão respectivo.

Endereços de difusão no IPv6

- ⑥ Unicast Prefix Based Multicast
 - △ Um endereço de difusão é constituído por dois componentes, um prefixo de rede e um identificador de grupo.
 - △ O prefixo de rede implicitamente identifica um domínio que é responsável pelo endereço.
 - △ Evita conflitos de endereços e a necessidade de executar protocolos de reserva de endereços em difusão.

Operação do BGMP

- ⑥ Tal como o BGP, baseado em ligações TCP entre os encaminhadores fronteira dos domínios.
- ⑥ As mensagens de *join* e *prune* são enviadas, de encaminhador fronteira em encaminhador fronteira, até ao domínio de raiz.
 - △ Este domínio está implícito no endereço de difusão.
- ⑥ É criada uma árvore partilhada bi-direccional:
 - △ Esta árvore optimiza o encaminhamento caso existam emissores fora do domínio raiz.

Operação do BGMP

- ⑥ Tipicamente, a árvore não mantém informação específica para cada fonte.
- ⑥ No entanto, os encaminhadores BGMP podem também manter estado referente a árvores com raiz na fonte, para compatibilidade com o protocolo usado no sistema interior (por exemplo, DVMRP, PIM-DM, PIM-SM).
- ⑥ A árvore partilhada construída pelo BGMP pode não corresponder à árvore com raiz no emissor.
 - △ Nesse caso, é necessário assegurar que o pacote é injectado dentro do sistema autónomo pelo encaminhador fronteira previsto pelo protocolo interior.

Operação do BGMP

- ⑥ Alterações ao encaminhamento ponto-a-ponto podem obrigar à reconfiguração da árvore partilhada.
 - △ Nomeadamente, quando o *next hop* para o domínio raiz é alterado.
- ⑥ Neste caso, é necessário enviar mensagens de *prune* para eliminar os ramos desactualizados, e novas mensagens de *join* para criar a nova árvore.

Resumo

- ⑥ MOSPF.
- ⑥ MSDP.
- ⑥ MBGP.
- ⑥ SSM.
- ⑥ BGMP.