

Protocolos em Redes de Dados

Aula 11

Difusão Inter-Domínio (PIM-SM/MSDP/MBGP e BGMP)

Luís Rodrigues

FCUL

2005-2006



Multicast Internet Exchange (MIX)

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



Sumário

- ▶ Aplicação de difusão entre domínios.



Inter-Domain IP Multicast

- ▶ Problema:
 - ▶ Como expandir a arquitectura baseada no MBONE (mais MOSPF ou PIM)?
- ▶ Solução intermédia:
 - ▶ Tentar usar a infra-estrutura existente com pequenas extensões.
- ▶ Solução de longo prazo:
 - ▶ Definir um protocolo semelhante ao BGP para multicast.
- ▶ Na prática a solução de longo prazo nunca foi adoptada e a solução intermédia tornou-se a norma de facto no IPV4.



Problema da descoberta de fontes

- ▶ Os protocolos esparsos usam o conceito de *rendez-vous point* (RP).
- ▶ Como extender 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 dedicadas.
- ▶ 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.



Utilização intra-domínio

- ▶ Apesar de concebido para resolver o problema da difusão inter-domínio, o MSDP é também usado para coordenar múltiplos rendezvous internos a um mesmo domínio.
- ▶ Em conjugação com o anycast rendezvous (ver aula anterior) permite resolver de forma pragmática o problema da tolerância a falhas de um rendezvous único.



MSDP Mesh-Group

- ▶ Sub-conjunto dos rendezvous que está completamente ligado.
- ▶ Neste sub-conjunto não é necessário re-encaminhar, para os outros membros do grupo, os anúncios recebidos de um membro do grupo.
- ▶ Facilita a configuração pois não existe o risco de gerar ciclos na propagação dos anúncios.



Fragilidades

- ▶ 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).
- ▶ Solução de emergência: limitar administrativamente a frequência máxima dos anúncios:
 - ▶ Favorece ataques de negação de serviço.



Rotas para as fontes

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



Multiprotocol extensions to BGP

- ▶ Existe uma extensão ao BGP que permite trocar rotas referentes a vários protocolos.
 - ▶ BGP+ (RFC 2283).
- ▶ 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

- ▶ *Não* troca informação acerca de endereços de difusão!
- ▶ Os anúncios BGP trocam informação acerca do endereço das fontes e do *next hop* para chegar a esse endereço.
- ▶ 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

- ▶ 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

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



Source-Specific Multicast

- ▶ 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).
- ▶ Elimina a necessidade de possuir um rendez-vous point.
- ▶ O endereço de grupo serve apenas para distinguir diferentes fluxos de dados com origem na mesma fonte.



Border Gateway Multicast Protocol (BGMP)

- ▶ Protocolo em estudo para suportar a difusão.
 - ▶ Último draft: 19 de Janeiro de 2004.
 - ▶ A proposta nunca obteve suficiente apoio para progredir no processo de normalização.
 - ▶ Na prática não substituiu a solução descrita anteriormente baseada no PIM-SM/MSDP/MBGP.
- ▶ 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

- ▶ 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

- ▶ Tal como o BGP, baseado em ligações TCP entre os encaminhadores fronteira dos domínios.
- ▶ 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.
- ▶ É criada uma árvore partilhada bi-direccional:
 - ▶ Esta árvore optimiza o encaminhamento caso existam emissores fora do domínio raiz.



Operação do BGMP

- ▶ Tipicamente, a árvore não mantém informação específica para cada fonte.
- ▶ 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).
- ▶ 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.



Dificuldades

- ▶ Auto-suficiência:
 - ▶ Não é obvio que se consiga usar sem suporte da rede para realizar o transporte de modo eficiente.
- ▶ Modelo desadequado:
 - ▶ Os emissores querem controlar quem pode enviar para o grupo.
 - ▶ É necessário assegurar que os endereços não são partilhados.



Operação do BGMP

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



Controlo de acesso

- ▶ É necessário assegurar que só as fontes autorizadas enviam para o grupo.
- ▶ É necessário assegurar que só os receptores autorizados recebem do grupo.
- ▶ É necessário assegurar que só se ligam à árvore domínios autorizados.



Dificuldades

- ▶ Os fornecedores, tendem a migrar os encaminhadores antigos, do centro para a periferia da rede à medida que vão adquirindo equipamentos novos.
 - ▶ Pode demorar até possuírem toda a rede com suporte para difusão.
- ▶ O custo de gerir o suporte para difusão só se justifica quando o serviço for muito solicitado, algo que não se verifica.



Resumo

- ▶ PIM/MSDP/MBGP.
- ▶ BGMP.

