

Experiência de Transmissão no XIX SBRC Multicast Inter-Domínio

Elvis Melo Vieira
elvis@npd.ufsc.br
RMAV - Florianópolis
NPD-UFSC
Campus Universitário – Trindade
Florianópolis 88040-9000

Engenharia de Rede: Multicast, IP, Gerência de Redes, MSDP, PIM, MBGP

Resumo

A experiência de transmissão multicast no IX Simpósio de Redes de Computadores tem por objetivo testar a transmissão multicast IP, em tempo real, entre domínios de multicast diferentes na Rede Nacional de Pesquisas. Estes domínios multicast podem estar definidos no mesmo AS (Autonomous Systems), ou AS separados. Isto difere do que tradicionalmente vem sendo realizado com a utilização de túneis multicast ou mtunnels. Entretanto, esta solução além da baixa performance, não apresenta uma boa escalabilidade. A experiência relatada a seguir envolve a definição e configuração dos protocolos de roteamento multicast (MBGP, MSDP, PIM e DVMRP), gerência da infraestrutura multicast e a aplicações a serem utilizadas. Esta experiência está sendo implementada para acontecer durante o IX SBRC em Florianópolis, através de uma parceria entre a RMAV-Florianópolis e a Rede Nacional de Pesquisa.

1 Introdução

O desenvolvimento da transmissão *multicast* na Internet teve o seu começo impulsionado pelas experiências do Mbone. Basicamente, o Mbone é um *backbone* composto por máquinas que utilizavam o DVMRP (RFC1075) como protocolo de roteamento *multicast* [1]. Entretanto, quando havia intermediárias máquinas que não suportavam roteamento *multicast*, eram utilizados túneis *multicast* ou *mtunnels*. Contudo, esta solução além da baixa performance, não apresenta uma boa escalabilidade. Com o desenvolvimento da infraestrutura de redes, as novas máquinas começaram a suportar novas capacidades, incluindo roteamento multicast e principalmente, o novo protocolo de roteamento denominado Protocol Independent Multicast (PIM) [2][3]. Este protocolo, na sua versão Sparse-Mode (RFC2362) é mais escalável e voltado para transmissões onde:

- Há poucos receptores em um grupo.
- Fontes e Receptores são separados por links WANs.

Desta forma, o roteamento *multicast* envolvendo transmissões situadas em um mesmo AS está bem resolvida. Entretanto, transmissões *multicast* envolvendo nós de rede em AS diferentes ainda está sendo desenvolvida e testada, principalmente no projeto da Internet2 (<http://www.internet2.edu/multicast>).

2 Domínios de Multicast

O IP *multicast* trafega de um nó fonte para os membros de um grupo *multicast* através de um fluxo que segue a topologia de uma árvore de distribuição, conectando todos os nós fontes a todos os nós receptores do grupo. Esta árvore pode ser compartilhada por todos os nós fontes ou, uma árvore de distribuição separada, construída para cada nó fonte. Adicionalmente, árvores de distribuição compartilhada são formadas ao redor de um roteador central, chamado rendezvous point (RP), do qual o tráfego é distribuído desconsiderando a localização da fonte. O PIM-SparseMode inicia com uma árvore compartilhada, podendo utilizar uma árvore para cada nó fonte se o tráfego for elevado.

3 Domínios Multicast

Um domínio *multicast* pode abranger uma ou mais árvores de distribuição. Dentro de um domínio, um protocolo de roteamento *multicast*, como o PIM, é usado para manter atualizada as rotas da(s) árvore(s) de distribuição. Adicionalmente, a idéia de domínio *multicast* é desassociada da idéia de AS, ou seja, dois ou mais domínios *multicast* podem estar definidos em AS diferentes ou não.

Para que os membros de um grupo *multicast* possam receber tráfego de fontes em outros domínios, é necessário um protocolo que troque informações de fontes ativas e seus grupos *multicast* correspondentes em cada árvore (mensagens SA) . O protocolo definido para realizar esta tarefa é o Multicast Source Discovery Protocol (MSDP) [1][4], o qual conecta múltiplos domínios PIM-SparseMode. Cada domínio PIM-SparseMode usa seu próprio RP(s) independente(s) de RPs em outros domínios.

Quando a troca de informações MSDP acontece entre domínios definidos em diferentes AS, então além da troca de fontes e grupos ativos efetuados pelo MSDP, há a necessidade dos roteadores BGP-4 trocarem informações específicas sobre roteamento *multicast*. Para isto, o BGP-4 foi estendido e definido o Multiprotocol Extensions for BGP-4 (MBGP). O MBGP (RFC 2283) fornece um método para os roteadores determinarem os caminhos que as árvores de distribuição *multicast* irão seguir, e utilizar, para a entrega dos pacotes *multicast*, da fonte para o receptores apreendidos pelo MSDP.

4 Descrição da Topologia da Experiência

A experiência de transmissão *multicast* no XIX Simpósio de Redes de Computadores tem por objetivo testar a transmissão *multicast* IP, em tempo real, entre domínios de *multicast* diferentes. Os domínios *multicast*, assim definidos, podem estar no mesmo AS (Figura 1) ou, em AS separados.

Na Figura 1, os roteadores da RNP representados são MBGP peers e MSDP peers, ou seja, roteadores que trocam informação de MBGP e MSDP. A idéia é que cada PoP da RNP tenha um roteador deste tipo, definindo um domínio *multicast*. Adicionalmente, o roteador pode ser RP da árvore de distribuição do PoP, se não houver disponibilidade de outro roteador. Idealmente, cada PoP deve ter um roteador RP para grupos *multicast* cuja transmissão inter-domínio e outro roteador RP para transmissões intra-domínio ou locais.

5 Configuração dos Roteadores

Os roteadores Cisco 7200 e 7500 foram configurados para trabalhar com MBGP, MSDP e PIM-SparseMode. A experiência está restrita ao AS 1916 (da RNP) mas nada impede a inclusão de outros ASes. Resumidamente foram efetuados os seguintes procedimentos, usando comandos do Cisco IOS 12.0:

1. Habilitação do roteamento *multicast* globalmente
2. Configuração das interfaces para trabalhar com PIM-SparseMode
3. Configuração para os roteadores aceitarem mensagens BSR, facilitando a negociação do roteador que irá ser RP, quando há mais de um roteador candidato, dentro de um domínio *multicast*.
4. Reconfiguração do BGP-4, habilitando o MBGP e as rotas de rede habilitadas para troca de informação *multicast* entre os BGP Peers.
5. Habilitação do protocolo MSDP e configuração dos Peers MSDP.
6. Configuração da interoperabilidade DVMRP/PIM, permitindo que o protocolo DVMRP seja também usado internamente no domínio *multicast*.

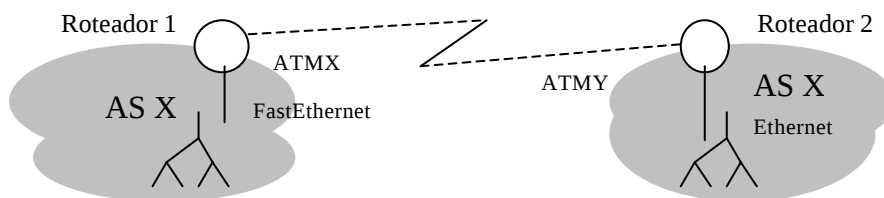


Figura 1 – Experiência Multicast entre AS diferente

6 Interoperabilidade DVMRP/PIM nos Roteadores

O protocolo desejável de roteamento interno *multicast* é o PIM. Entretanto, nem todos os equipamentos possuem implementação disponível. Entre eles está o IBM Multiprotocol Switched Service (MSS) ou IBM 8210 utilizados nas RMAVs. Nestes equipamentos está disponível o protocolo DVMRP. Portanto, foi necessário usar a capacidade dos roteadores Cisco dos PoPs para realizarem a interoperabilidade entre o protocolo PIM e DVMRP.

7 Gerenciamento

Resumidamente, o gerenciamento pode ser dividido em dois tipos: gerenciamento SNMP e não-SNMP.

7.1 Gerenciamento SNMP

O gerenciamento SNMP, sendo o gerenciamento preferido, é feito tendo o Tivoli Netview como software de gerenciamento. As tarefas de gerenciamento envolvem consultas à MIB PIM da Cisco e a MIB MSDP, em fase *draft* no IETF, ambas implementadas nos roteadores Cisco. A MIB proprietária PIM da Cisco é utilizada para conseguir informações do estado e performance do roteamento PIM como:

- Informações sobre o nó RP.
- Informações sobre as mensagens Join/Prune do PIM.
- Notificações sobre mudanças nas interfaces PIM ou na árvore de distribuição.

Com a MIB MSDP, algumas informações utilizadas são:

- Informações sobre a performance das mensagens AS.
- A tabela (cache) de SA ativas.
- Informações sobre os peers MSDP.

Adicionalmente, a Intel vem trabalhando em uma MIB para gerenciar o protocolo RTP (<http://developer.intel.com/ial/mib>). Para isto, ela disponibilizou um subagente para ambientes Unix, o qual interage com o agente SNMP de domínio público UCD-SNMP. A utilização deste subagente está sendo testada no momento. Pretende-se conseguir com este agente, informações sobre o tráfego RTP nas interfaces dos equipamentos com as aplicações RTP (como VIC, VAT e RAT).

7.2 Gerenciamento Não-SNMP

O gerenciamento não-SNMP visa suprir certas necessidades de gerenciamento ainda não cobertas com o gerenciamento SNMP. Para tanto, as seguintes aplicações são utilizadas:

- **MHealt** : Permite a descoberta e visualização gráfica dos membros de um grupo *multicast*.
- **RTPMon**: Permite obter informações de cada membro de um grupo *multicast*. Na experiência esta sendo usado para obter informações sobre os descartes e a variação de atraso dos pacotes RTP.

Outras ferramentas ainda estão sendo testadas e em breve se pretende chegar a conclusão da sua aplicabilidade na experiência. Um resumo destas ferramentas pode ser conseguido em <http://www.caida.org/tools/taxonomy/multicast.xml>.

8 Conclusão

Esta experiência tem permitido avançar os conhecimentos sobre roteamento *multicast* e seu gerenciamento. Com os testes preliminares realizados em laboratório, pode-se obter algumas conclusões sobre a aplicabilidade dos protocolos testados, principalmente sobre sua performance, onde foram encontrados alguns problemas de demora dos anúncios de sessões. O foco atual da experiência está em repassar a configuração feita no laboratório para os demais PoPs da RNP. Para isto estamos trabalhando em parceria do pessoal técnico da RNP, representado principalmente por Adenilson Raniery. Os resultados da experiência serão divulgados no site da RMAV-FLN (www.rmav-fln.ufsc.br).

9 Referências

- [1] An Overview of Inter-Domain Multicast Routing – White Paper, disponível em <http://www.microsoft.com/windows2000/library/technologies/communications/default.asp>
- [2] PIM-SM Protocol Improves Multicast Service Availability, Cisco Beyond Basic IP Newsletter V1.11. Disponível na Cisco Systems.
- [3] PIM-SM Multicast Routing Protocol – White Paper, disponível em <http://www.microsoft.com/windows2000/library/technologies/communications/default.asp>.
- [4] MSDP Feature Broadens Reach of Multicast Services, Cisco Beyond Basic IP Newsletter V1.2.2. Disponível na Cisco Systems.