Aspectos de Segurança para Pontos de Troca de Tráfego

Leandro Márcio Bertholdo – POP/RS

Fernando Krahe – Optiglobe

Liane M. R. Tarouco - PGIE & PGCC/UFRGS

RESUMO

A experiência adquirida na instalação dos pontos de troca de tráfego OptIX-LA e RSIX fez com que registrássemos, neste artigo, aspectos que devem ser considerados em outros pontos de troca de tráfego existentes e futuros, bem como pelos participantes dos mesmos. O intuito deste documento é o de manter a Internet-BR o mais estável possível, alertando para problemas e práticas que podem diretamente violar a segurança e estabilidade de todos os participantes de um PTT.

Palavras-chave: **PTT, NAP, Segurança, Roteamento, BGP**.

ABSTRACT

The experience attained in the rolling out of the OptIX-LA and RSIX network access points, made us register, in this article, aspects that must be considered at other existent and future network access points and by their participants. The objective of this document is to bring more stability and robustness to Internet-BR throwing

Keywords: **NAP, Internet Exchange, Security, Routing, BGP**.

Introdução e definições

A Internet surgiu com a interconexão de redes em um modelo hierárquico onde uma rede central transportava o tráfego entre redes periféricas. Essa estrutura foi substituída por uma arquitetura distribuída operada por provedores comerciais interconectados através de grandes pontos de interconexão de redes. Com o aumento exponencial da Internet mundial esses pontos de interconexão de rede tendem a multiplicar-se pelo Brasil e pelo mundo em relação semelhante a distribuição de novos números de sistemas autônomos e crescimento da tabela BGP [HUS2001] [TEL2002]. Entre as principais vantagens da existência desses pontos de troca de tráfego encontram-se as diminuições de gastos com circuitos de longa distância e a diminuição do tempo de acesso (atraso) entre seus participantes. O tempo de acesso torna-se a cada dia um recurso mais precioso, demanda essa gerada principalmente por aplicações real-time e interativas. Outras vantagens que também devem ser levadas em consideração são aumento da conectividade regional, nacional e até mesmo internacional, que ajudam a minimizar o impacto sobre a conectividade de seus participantes no caso de falhas em equipamentos e circuitos de acesso.

Os pontos de interconexão de redes ou PIR são também conhecidos como pontos de troca de tráfego(PTTs). Também são muito utilizados os termos da lingua inglesa, NAP ("Network Access Point"), EP ("Exchange Point") e IX ("Internet eXchange").

Os PTTs podem ser definidos como redes de alta taxa de transferência ou comutadores (switch), aos quais um número de roteadores é conectado com o objetivo de

trocar tráfego sem o custo do serviço IP. Os PTTs podem ser tão simples como um comutador ethernet ou comutador ATM passando tráfego de uma rede para outra. No entando é recomendado que a rede do PTT opere no mínimo a 100Mbps.

A rede do PTT é o meio de transporte para o tráfego trocado entre os provedores, mas para que isso ocorra é necessário uma conexão lógica entre os mesmos. Essa conexao lógica, chamada de peering, possibilitará que os roteadores troquem informações de alcançabilidade de redes (NLRI) e a maneira mais utilizada para trocar essa informação é utilizando o protocolo de roteamento BGP na sua versão 4.

Cada um desses grandes provedores comerciais e outras grandes redes que passaram a operar a Internet formaram sistemas autônomos ou ASes, i.e., são redes compostas de vários roteadores com uma única política de roteamento e rodando sob uma administração única [HUI1995].

Toda vez que um AS se interconecta a outro AS para trocar tráfego, ocorre uma relação chamada de "peering". Essa relação pode ser classificada em duas categorias bem definidas:

1. Troca de tráfego de interesse mútuo: Onde os dois ASes trocam tráfego próprio ou de seu interesse, ou seja, o tráfego que passa por essa conexão é originado em e destinado a prefixos desses ASes ou prefixos de ASes clientes dos mesmos;
2. Trânsito: Um AS provê trânsito ao outro AS para que esse último chegue a outros ASes através do primeiro. Esse trânsito é uma prestação de serviço normalmente não gratuita.

A relação de trânsito normalmente é estabelecida através de uma conexão dedicada entre roteadores dos dois ASes envolvidos, embora tecnicamente possa ser realizada na rede de um PTT. Não é contudo recomendado o uso da rede compartilhada do PTT para tal prática. Caso o provimento de trânsito seja uma das premissas do PTT é recomendada que seja realizada por uma conexão independente.

Estabelecidas essas definições básicas, pode-se agora diferenciar e enumerar os múltiplos problemas encontrados durante a administração no dia a dia de um ponto de troca de tráfego. Optou-se aqui enumera-los segundo sua freqüência, ou seja, os problemas mais comumente encontrados e outros mais incomuns e complexos, que podem ser criados até intencionalmente por um dos participantes.

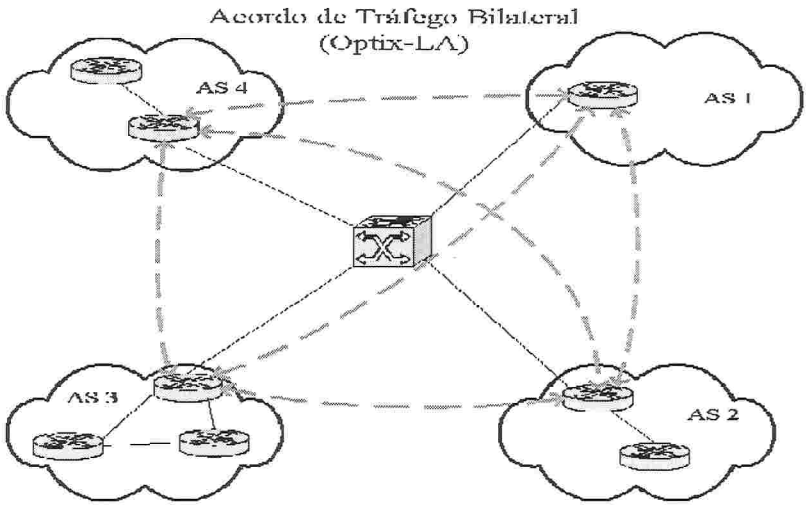
Esse documento procura prover uma definição básica para filtros a serem utilizados pelos participantes e pela própria administração dos pontos de troca de tráfego brasileiros, além de indicar algumas políticas de uso aceitável que podem ser consideradas "best practices" pelos participantes dos PTTs.

Durante a operação de ambos os PTTs e a experiência adquirida na análise de falhas de roteamento da Internet-BR, conseguiu-se enumerar alguns pontos comuns. Estes, em sua maioria ocorrem quando da entrada de um novo participante no PTT.

Problemas de Escalabilidade em um PTT

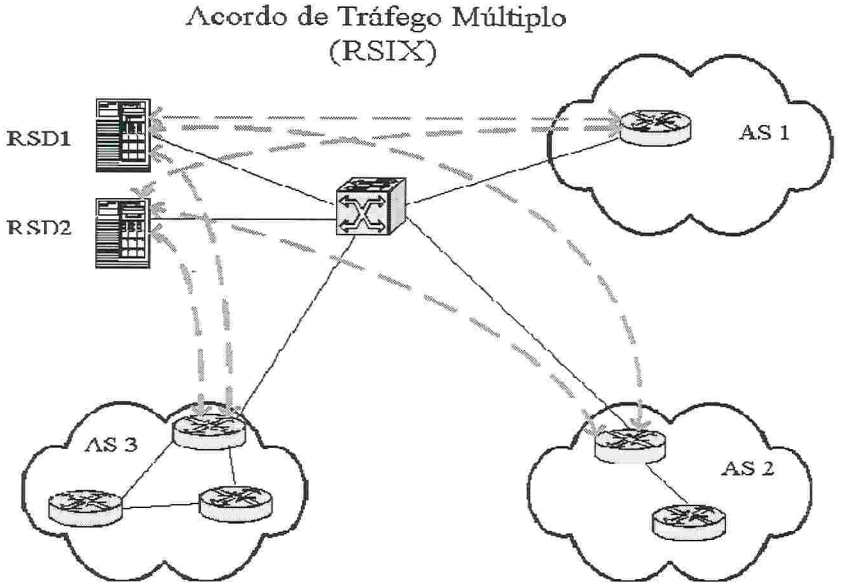
O principal problema que pode ser abordado está diretamente relacionado a forma de funcionamento do PTT, e sua facilidade de crescimento. Cada vez que dois ASes desejarem

estabelecer uma relação de peering, será necessario criar uma sessão BGP entre ambos. Desse modo, quando N roteadores optarem trocar tráfego entre si, serão necessárias N(N-1)/2 sessões BGP entre os mesmos. Ao adicionar um roteador a essa malha, serão necessárias mais N-1 sessões BGP. Esse full-mesh entre os roteadores consome recursos dos mesmos, podendo chegar a um limite tanto em tamanho da configuração como memória para manter o estado das sessões e tabela de roteamento e outras atividades que o roteador realiza.



A alternativa para solucionar esse problema de escalabilidade é a adoção de route-servers. O route-server é um servidor que executa o protocolo BGP e com o qual cada roteador estabelece uma sessão BGP. Agora o número de sessões BGP fica reduzido em N, onde N é o número de roteadores trocando tráfego entre si, e cada nova adição de um roteador significa mais uma sessão BGP.

Neste caso o servidor de rotas está em outro AS e conceitualmente atua como se fosse um AS trânsito entre todos os ASes que trocam tráfego no PTT. Entretanto, técnicamente, o route-server exporta os NLRI sem modificar o atributo next-hop dos mesmos. Como os next-hops dos prefixos aprendidos por um roteador são as interfaces dos próprios roteadores participantes, o encaminhamento dos pacotes não é realizado pelo route-server, e sim diretamente pelos roteadores envolvidos. Todavia, ao resolver o problema de escalabilidade, foi inserido um ponto único de falha. A solução então é utilizar dois route-servers, um primário e um secundário e cada roteador estabelece sessões BGP com ambos os dois servidores. Agora, o número de sessões BGP é 2N, onde N é o número de roteadores.



Os route-server resolvem o problema de escalabilidade nos NAPs, entretanto introduzem um novo problema: a administração dos mesmos. Muitas vezes, os ASes participantes não querem confiar em um deles para administrar o route-server. Assim, uma terceira parte, independente de todos os ASes, é incumbida para administrar os route-server e às vezes até todo o PTT. Nas grandes redes comerciais, peering esta muito mais relacionado a estratégia de negócio do que a engenharia. Portanto, muitos ASes, ao estabelecerem peering, assinam acordos entre si onde as partes se comprometem em não divulgar a informação de que estão fazendo peering. Por causa disso, muitas vezes deixam de utilizar o route-server.

O route-server é normalmente encontrado em PTTs públicos onde se utiliza muito o acordo de peering multilateral, ou seja, onde o acordo de peering é simultâneo entre todos os ASes presentes no PTT. Nesses PTTs normalmente há uma rede compartilhada ou uma malha de comutação ("switch") ao qual todos os roteadores são conectados. Nesses PTTs, muitas vezes a malha de comutação está congestionada ou a conexão de um determinado AS até o PTT está congestionada. Isso pode acontecer quando alguns, mas nem todos os ASes, aumentam sua capacidade até o PTT. Normalmente, o acordo multilateral não obriga uma parte a incrementar seus recursos ao atingir um determinado nível. Visando corrigir esse problema os acordos de tráfego podem estabelecer compromissos aos participantes de manterem a sua utilização média em até 80% do enlace.

Devido ao congestionamento desses PTTs públicos, vários dos grandes provedores, a maioria deles operadoras de telecomunicações ou de propriedade das mesmas, passaram a instalar POPs em grandes data-centers neutros de operadora que se tornaram locais idéias para hospedar PTTs privados. Nesses PTTs privados, os acordos de peering são bilaterias e não se utilizam route-servers. Em algumas vezes, nem mesmo a malha de comutação é compartilhada entre todos. Assim evitam-se os congestionamentos encontrados nos PTTs públicos.

Outro elemento utilizado para melhorar a escalabilidade, principalmente nos PTTs privados, onde muitas vezes não se utiliza um route-server, é o emprego de agrupamento. O agrupamento não interfere no número de sessões BGP, entretanto aborda dois aspectos da

escalabilidade, qual seja, tamanho da configuração e geração e anúncio da mensagem de UPDATE.

O agrupamento é uma ferramenta de configuração utilizada para aplicar os mesmos comandos para múltiplos peers, eliminando assim a necessidade de configurar esses comandos para cada peer. Todos os membros do grupo, ou seja, os peers pertencentes ao grupo, irão receber a mesma mensagem de UPDATE, portanto todos os membros do grupo devem ter a mesma política de roteamento. A tabela de roteamento é encaminhada somente uma vez (para o líder do grupo). Os prefixos são filtrados de acordo com a política de roteamento e a mensagem UPDATE é gerada para o líder do grupo, que se encarrega de replicá-la aos outros membros do grupo que estão sincronizados com o líder. O processo de replicação de uma mensagem de UPDATE é muito mais fácil e rápido que a geração de uma nova mensagem, pois a geração da mensagem requer uma caminhada na tabela de roteamento e avaliação das políticas, enquanto a replicação não.

Um membro do grupo está sincronizado com o líder se todos os UPDATES enviados ao líder também tiverem sidos enviados ao membro do grupo. Quanto mais membros do grupo estiverem sincronizados, maior é o número de UPDATES que o BGP pode replicar. Entretanto, o membro do grupo pode perder sincronia com o lider por algumas razões:

- Baixo throughput TCP;
- Grande número de TCP ACKs preencherem as filas de entrada resultando em descarte de pacotes;
- O peer está ocupado realizando outra tarefa;
- O peer tem uma CPU mais lenta que o do lider

Para solucionar a primeira causa, é possível diminuir o overhead do TCP alcançando um MSS ótimo, qual seja, um MSS baseado no menor MTU dos enlaces entre os dois peers. A segunda causa pode ser solucionada incrementando o tamanho da fila de entrada.

Full rounting e suas implicações nos roteadores conectados a rede do PTT

O roteadores conectados a rede do PTT exportarão NLRI (Network Layer Reachability Information) para os outros roteadores também conectados a mesma rede e com os quais mantenham relações de peering contendo route-objects de seu próprio AS e route-objects dos AS aos quais fornecem trânsito. Cada roteador também irá receber NLRI dos outros roteadores com os quais mantém relação de peering e irá divulgar essa informação para seu AS. Assim, esse roteador podera encaminhar pacotes originados em seu AS ou ASes de seus clientes. Portanto não há necessidade desse roteador possuir uma tabela de roteamento completa ("full-route"), outras rotas ou mesmo uma rota default.

Os roteadores conectados ao PTT que somente conhecem as rotas que exportam e as que apreendem de seus peers, irão automaticamente descartar qualquer pacote destinado a prefixos não constantes em sua tabela de roteamento.

Caso um roteador conectado ao PTT tenha que conhecer, por alguma razão, a tabela de roteamento completa ou uma rota default, e.g, quando o roteador de um determinado AS é, ao mesmo tempo utilizado para peering com alguns ASes e fornecimento de transito para outros, será necessário implementar mecanismos que impossibilitem a divulgação dos

prefixos que não se quer anunciar e filtros de pacotes que descartem pacotes destinados a prefixos não exportados provenientes dos roteadores dos ASes conectados ao PTT.

O filtro de pacotes utilizado quando um roteador conhece prefixos além dos quais exporta aos roteadores dos outros ASes conectados ao PTT é utilizado para se proteger de situações que podem ocorrer no seu próprio roteador ou nos roteadores dos outros ASes.

- O seu próprio roteador, por erro de configuração ou bug no sistema operacional, pode passar a vazar uma rota para a qual não quer encaminhar pacotes. Nesse caso, esse roteador ira causar um "black hole" das redes vazadas para os outros ASes que apreenderem e utilizarem essa rota. É necessária intervensão imediata para corrigir esse problema;
- O roteador de outro AS, por erro de configuração ou bug no sistema operacional, pode passar a enviar pacotes destinados a prefixos não recebidos do roteador ao qual os pacotes estão sendo enviados. Esse caso também gera um "black hole", mas o roteador que estiver recebendo e filtrando esses pacotes não estará exportando uma rede para a qual não pretende encaminhar pacotes;
- Em um caso mais extremo, um AS pode fazer "default traffic" para outro AS de modo a obter transito gratuitamente. O roteador alvo do "default traffic" irá descartar todos os pacotes que nao passarem seus filtros.

É importante salientar que a aplicação de um filtro de pacotes em uma interface faz com que seja verificada a conformidade de todos os pacotes que passam pela mesma. Esse processo exige recursos dos roteadores e pode gerar, entre outros problemas, atrasos e perda de pacotes principalmente quando se tratar de interfaces de alta taxa de transferência, e.g., Gigabit Ethernet, sendo utilizadas a pleno. Alguns roteadores possuem hardware dedicado para realizar tarefas como essa, diminuindo o impacto causado pela analise dos pacotes. Portanto, a escolha de manter no roteador contectado ao PTT somente as rotas exportadas para as quais se encaminham pacotes, além das apreendidas dos roteadores dos outros ASes conectados ao PTT é a solução mais indicada.

Anúncio indevido de trânsito

Uma das falhas comuns encontrada nos PTTs é o anuncio indevido de rotas que ocorre quando um participante divulga rotas de outros ASes, servindo indevidamente de trânsito para esse ASes. Esse caso é uma das situações descritas no item acima, entretanto agora o problema é analisado a partir do ponto de vista de quem recebe tais anúncios. Um AS não necessita full-routing para anunciar prefixos indevidos, basta anunciar prefixos que não sejam seus próprios ou de seus clientes.

Embora, na maioria das vezes seja facilmente detectável, esse anúncio indevido possui grandes implicações que podem prejudicam todos os outros participantes que o estiverem o detectando. Dentre essas implicações, pode-se citar:

- Filtros utilizados pelo AS que está servido de trânsito: Nesse caso, o AS, ao exportar o prefixo, oferece-se como trânsito para outro AS mas filtra no seu backbone determinado protocolo ou endereço, expandindo sua política interna de segurança para os participantes do PTT que estivem aceitando o prefixo.
- Congestionamento do seu enlace: A conseqüência mais comum do anúncio indevido é quando um AS participante do PTT exporta prefixos de outros peers (trânsito ou de

interesse mútuo). Quando isso ocorre, vários ASes que recebem esses anúncios podem optar por mandar tráfego a eles por esse AS, causando uma saturação no enlace desse AS para dentro de seu backbone ou em outros enlaces do mesmo. Esse tipo de erro pode causar um perigoso congestionamento no acesso ao AS que está sendo indevidamente anunciado no PTT. A origem dos anúncios indevidos normalmente está associada a distribuição de rotas entre protocolos ou a filtragem de rotas.

- Roteamento assimétrico: Outra conseqüência que se percebe em "anúncios ilegais" é a possibilidade de instalação de um roteamento assimétrico[1]

A detecção geralmente é realizada através de ferramentas simples como traceroute[2] e análise das tabelas de roteamento do looking-glass[3]. Outra forma é a análise das rotas anunciadas pelos outros peers.Uma das soluções para esses casos é filtrar, no route-server, prefixos cujos AS-PATH indiquem a origem do mesmo nos ASes dos grandes provedores de backbone do país. Filtros como[4]:

```
ip as-path access-list 15 deny  ^[0-9]+_1916    # RNP
ip as-path access-list 15 deny  ^[0-9]+_2716    # Rede Tche
ip as-path access-list 15 deny  ^[0-9]+_11415   # Impsat
ip as-path access-list 15 deny  ^[0-9]+_17379   # Intelig
ip as-path access-list 15 deny  ^[0-9]+_11706   # Terra Networks
ip as-path access-list 15 deny  ^[0-9]+_4230    # Embratel
ip as-path access-list 15 deny  ^[0-9]+_8167    # Brasil Telecom
```

Entretanto, essa prática de filtragem só permite que esses prefixos sejam descartados quando os mesmos não tenham seus AS-PATH alterados[5]. Para o caso de um AS exportar devidamente um destes provedores de backbone, uma exceção é inserida no filtro para aquele AS_PATH específico. O participante do PTT é alertado sobre essa restrição e deve notificar o seu desejo de exportar determinado provedor de backbone. No caso do anúncio realizado por qualquer participante ser barrados por esses filtros, o dono do anúncio é notificado e a sua intenção é confirmada ou o anúncio corrigido.

A alternativa ao uso desses filtros em route-server é configurar, no roteador conectado a rede do PTT, um número máximo de prefixos que esse irá aceitar de seus peers. Esses peers podem ser o route-server ou outros roteadores também conectados a rede do PTT. Essa prática é utilizada nos PTTs que não disponibilizam route-servers, mas é recomendável a qualquer participante de qualquer PTT.

Invariavelmente, todos os ASes conectados a rede do PTT estão, através dessa conexão, trocando tráfego de interesse mútuo, ou seja, estão somente exportando seus blocos CIDR e os blocos de seus clientes. O número total de prefixos raramente é maior que alguns milhares. No caso específico da Internet/BR, a maioria das redes tem menos que algumas dezenas de blocos, com exceção de uma minoria que possui centenas ou até um poucos

---

[1] O roteamento assimétrico caracteriza-se por possuir caminhos diferenciados quando se deseja ir do ponto A→B ou de B→A.

[2] Hoje em dia vários backbones impedem a utilização desta ferramenta através do seus domínio, dificultando a resolução de problemas como esse.

[3] O looking-glass pode ser um roteador com sessões BGP-4 com os route-servers ou um servidor WWW com acesso para consulta as tabelas de rotas dos route-servers. Costumeiramente possui acesso aberto a consultas, permitindo a qualquer pessoa consultar como estão sendo realizados os seus anúncios junto ao PTT.

[4] Essa é uma lista parcial dos filtros.

[5] Caso onde ocorre a injeção de rotas EGP no protocolo IGP e novamente para protocolo EGP.

milhares de blocos, incluindo ai os de seus clientes. Portanto, no caso de sessões BGP estabelecidas diretamente entre os roteadores conectados a rede do PTT, é recomendável aplicar um limite de prefixos para cada peer. Esse limite não precisa ser exatamente o número de blocos que o peer possui, pode haver uma folga para que o peer possa adicionar mais prefixos ou, talvez, anunciar prefixos mais específicos. Se um peer passar a exportar mais prefixos que o limite, a sessão BGP será derrubada. É possível configurar um valor percentual do limite que, quando alcançado, ira enviar mensagens nos logs avisado desse evento. Dessa forma os administradores podem se anteceder a situação. No caso do peer ser o route-server, o valor do limite do número de prefixos deverá comportar a soma de todos os anuncios dos outros participantes.

O uso do limite de prefixos, por não verificar o AS-PATH, protege da situação quando há alteração do AS-PATH no prefixo. Essa situação normalmente ocorre quando ha redistribuição de prefixos entre protocolos de roteamento e será tratada mais adiante.

Outro método que deve ser utilizado juntamente com o limite de prefixos é o filtro baseado em registro de rotas, ou seja, a configuração feita no roteador é baseada em prefixos declarados em repositórios de rotas e não em AS-PATH. Essa configuração pode ser feita manualmente ou através de scripts.Limite de prefixos e filtro de prefixos são duas boas práticas adotadas.

Utilização de filtros demasiadamente restritivos

Uma pequena parte dos detentores de ASNs costumam manter seus próprios filtros contra o PTT, impedindo assim a utilização e divulgação incorretas de AS_PATHs. Entretanto, esse procedimento inconveniente necessita um comunicado da administração do PTT a esse participante quando da entrada de outro ASN, para que esse revise sua política de filtragem. Além disso, como mencionado anteriormente, filtros de AS_PATH não protegem de prefixos cujo AS_PATH foi alterado. O problema aqui ocorre quando um determinado administrador de AS não comunica a existência de uma política de segurança aplicada contra o PTT. O problema toma proporções maiores quando nem toda a equipe de administradores do backbone da empresa tem informação sobre a existência dessa política contra o PTT ou o seu implementador já não mais faz parte da empresa.

Esse é um caso que se manifesta com maior freqüência em acordos de tráfego múltiplo, como é o caso da maioria dos participantes do RSIX. No caso de somente existirem acordos bilaterais, como os praticados no OptiX-LA esse problema praticamente inexiste; quando da entrada de um novo participante no o administrador de determinado AS precisa interagir diretamente, implementando o acordo de troca de tráfego previamente assinado entre as duas instituições.

A conseqüência mais comum dos filtros restritivos é a criação de um caminho assimétrico entre os dois participantes, reduzindo assim os ganhos obtidos com a entrada do novo participante. O novo participante (AS2) acata as rotas recebidas pelo route-server sobre o caminho existente para o AS1, mas o AS1 ignora o anúncio do route-server, não utilizando esse novo caminho (AS1_AS2)

Quando outros tipos de filtros estão aplicados diretamente na interface do roteador de borda do AS1 (filtros por protocolo, serviços TCP ou UDP, ou blocos de endereços), o tráfego enviado de AS2 para AS1 através do PTT é filtrado parcialmente. Esse costuma ser um problema de difícil detecção e envolve ambos os participantes e administração do PTT. Nesse caso específico o alerta é gerado a partir de usuários finais ou clientes de um dos sistemas autônomos. A forma mais simples para determinação desse problema é a utilização de portscanners, devidamente autorizada pelos participantes.

NLRI gerada no roteador conectado ao PTT

Uma situação muito peculiar pode ocorrer quando um AS gera os prefixos que exporta ao PTT no próprio roteador conectado a rede do mesmo. Se esse roteador perder a conectividade com o resto do seu backbone, o AS ficará dividido em partes disjuntas, uma parte composta pelo roteador conectado ao PTT e outra composta pelo resto do backbone com suas outras conexões. O resto do backbone, muito provavelmente, recebera os prefixos dos ASes do PTT através de uma conexão trânsito com outro AS. Assim poderá enviar pacotes para aqueles ASes. Entretanto, o roteador no PTT manterá sua conexão ao mesmo e continuará exportando os prefixos gerados localmente. Os outros ASes conectados ao PTT estarão recebendo os prefixos vindo por dois caminhos distintos, como já estavam antes, mas continuarão preferindo o PTT, assim o tráfego será enviado ao roteador do PTT. Esse roteador não descartará os pacotes por não ter rota mais específica para os mesmos.

Problemas com o atributo NEXT-HOP

O next-hop é um atributo bem-conhecido e mandatário do BGP, e ao qual deve ser dada fundamental atenção. Os prefixos apreendidos em uma conexão de peering somente serão utilizados pelo resto do AS se o next-hop dos mesmos for válido. Algumas situações verificadas nos PTTs é que, muitas vezes, os prefixos apreendidos são divulgados para dentro de um sistema autônomo, mas o next-hop daqueles é inválido devido aos outros roteadores não terem conhecimento deste[6].

Nas conexões eBGP, ou seja, o BGP entre sistemas autônomos, o next-hop default dos prefixos é a interface do roteador que anuncia o prefixo. O atributo next-hop passado aos peers pode ser a própria interface de rede do equipamento (Fast Ethernet) ou uma outra interface interna (Loopback). O roteador que aprende o prefixo deve divulgar dentro de seu backbone como chegar nesse next-hop ou trocá-lo.

A divulgação desse next-hop pode ser feita de duas maneiras:

1. Executar o protocolo de roteamento IGP na interface externa em modo passivo[7]. Assim a rede dessa interface (nesse caso a rede local do PTT) será distribuída pelo IGP por todo o backbone. Devido ao fato do IGP estar rodando em modo passivo nessa interface, não será possível estabelecer adjacências com outros roteadores na mesma.
2. Injetar a rede da interface externa no BGP através da redistribuição de interfaces conectadas controlado por uma politica. Assim, o iBGP irá

[6] O next-hop pertence a uma rede que os roteadores internos ao AS desconhecem totalmente.
[7] No modo passivo o protocolo ignora os anúncios das rotas recebidas.

divulgar a rede local do PTT para o resto do backbone e o next-hop desse prefixo será uma interface interna conhecida dos roteadores internos ao sistema autônomo. Esse método requer uma consulta recursiva para encontrar o next-hop.

Caso não seja desejado divulgar a rede externa dentro do sistema autônomo, pode-se optar por fazer com que o roteador que recebe os prefixos altere o next-hop para uma de suas interfaces internas, conhecidas pelo resto do backbone.

A boa prática normalmente não altera o next-hop, portanto usa uma das duas alternativas citadas. É importante salientar que em muitos PTTs a rede local pertence a um espaço de endereçamento reservado, portanto são redes não roteadas na Internet. Dessa forma, a única maneira de chegar às redes desses PTTs é divulgando-as dentro do próprio AS.

Uso de Local-Preference

Local-Preference é um atributo bem conhecido do BGP que pode ou não estar incluido em uma mensagem de UPDATE. Esse atributo é transitivo dentro de um AS, ou seja, ele é mantido por todo o AS. O local-preference indica a preferência a um prefixo e é o primeiro parâmetro não proprietário a ser avaliado no processo de decisão de qual é a melhor rota para um determinado prefixo.

O local-preference é utilizado para controlar como o tráfego saí de um sistema autônomo e seu emprego deve ser cuidadosamente planejado. É importante que o valor do local-preference para os prefixos dos clientes seja semre maior que o valor do local-preference dos prefixos apreendidos em conexões de peering, tanto de trânsito como de interesse mútuo, assim estará garantido que o tráfego para o cliente irá escoar pela conexão contratada pelo cliente.

Já nas conexões de peering, pode-se utilizar o mesmo valor para trânsito ou troca de interesse mútuo, ou um valor maior para o peering tipo troca de tráfego de interesse mútuo Assim pode-se maximizar os investimentos feitos pelas partes para estabelecer o peering. O local-preference, nesse caso, irá afetar os sistemas autônomos clientes dos ASes que estão realizando peering. Exemplificando: em uma situação onde um AS3 é cliente do AS2 e do AS-Trânsito e o AS1 tem peering com AS2 e é cliente do AS-Trânsito, o AS1 enviará os pacotes para o AS3 através do AS2[8].

É possível criar mecanismos que permitam os clientes manipularem o local-preference de seus prefixos dentro da rede de seu AS fornecedor de trânsito, mas esse mecanismo não deve ser oferecido a qualquer peer. Geralmente essa funcionalidade é implementada através de comunidades BGP.

[8] Nessa situação o ponto de troca de tráfego utiliza um número de AS válido, e todo o tráfego utiliza acrescenta esse AS no AS_PATH recebido.

Não utilização de práticas de segurança na configuração da interface de rede conectada ao PTT.

Em muitos casos os administradores desconhecem ou não utilizam as recomendações de segurança[CIS2001a] para configurar sua interface de conexão com o equipamento do PTT.

Na maioria das vezes o único prejudicado é o próprio participante do PPT. No entanto, todo o PTT também pode tornar-se vulnerável a ser utilizado como origem de ataques do tipo SMURF[SEN1999], caso os seus participantes não configurem corretamente suas interfaces de rede. A boa prática recomenda que as interfaces sejam configuradas :

- Eliminando redirecionamento de pacotes IP: o roteador não deve enviar uma mensagem de redirecionamento quando o router tem que reenviar o mesmo pacote através da interface por onde ele foi recebido.
- Eliminando directed-broadcast: significa que a tradução de um broadcast direto para um broadcast físico está desabilitado. Caso seja habilitado o equipamento pode ser utilizado em um ataque tipo SMURF .
- Eliminando Proxy-ARP: Proxy arp [MAL1995] é utilizado pelo router para auxiliar equipamentos sem capacidade de roteamento a determinar o endereço MAC de outras redes ou sub-redes. Dessa forma o equipamento envia o pacote para o router que o reenvia para o seu destino. Esse comportamento é inadequado por carregar as informações de roteamento para todos os destinos da Internet. Utilizar proxy ARP pode resultar no roteador do PTT tentar manter uma tabela ARP extremamente grande, diminuindo a performance do equipamento.
- Eliminar protocolos, normalmente proprietários de cada fabricante de equipamento, que tentem descobrir outros dispositivos na rede. Esses protocolos, tais como CDP (Cisco Discovery Protocol) ou EDP (Extreme Discovery Protocol), não oferecem uma falha direta de segurança, no entanto podem ser utilizados por um atacante para descobrir informações sobre os equipamentos, versões de software, hardwares, etc. Essas informações podem ser posteriormente utilizadas em um ataque.

Redistribuição entre protocolos

A redistruibuição de NLRI entre protocolos é altamente desaconselhável!

Essa redistruibuição pode ocorrer das seguintes maneiras:

1. Quando se injeta IGP dentro de BGP;
2. Quando se injeta BGP dentro de IGP;

IGP é um tipo de protocolo interno utilizado para transportar informação de estrutura de um backbone, ou seja, é utilizado para transportar informação de next-hop. Assim os roteadores de um AS o utilizam para disseminar a topologia do mesmo. A boa prática faz com que a maioria dos ASes adotem protocolos baseados em estado de enlace como seu IGP. Os mais antigos utilizam IS-IS, mas OSPF também é bastante utilizado. Esses protocolos

baseados em estado de enlace utilizam bastante recursos dos roteadores, pois cada roteador deve manter uma base de dados que mantenha toda ou uma parte da topologia do backbone.

Toda vez que houver um evento que altere a topologia da rede, o IGP irá executar o algorítmo SPF de modo a convergir para a nova topologia. Não é recomendável que o IGP transporte informações desnecessárias que servem somente para sobrecarrega-lo. A boa prática recomenda que o IGP só transporte as interfaces internas de um backbone e suas interfaces de loopback.

BGP é um protocolo externo utilizado para transportar informações de roteamento entre sistemas autônomos, mas a boa prática também o recomenda para transportar essa informação dentro de um sistema autônomo.

A distribuição entre protocolos normalmente ocorre em redes que não implementam BGP em todos os dispositivos de camada 3 de seu backbone. Nessas redes o IGP é quem transporta todas as informações internamente e o BGP é somente utilizado para divulga-las nas interconexões com outras redes, portanto ocorrem tanto distribuições do BGP para o IGP como do IGP para o BGP. Nesse momento são criados os problemas, entre eles:

- Muitas vezes o IGP não suporta a quantidade de rotas que o BGP redistribui no mesmo;
- A instabilidade da(s) outra(s) redes (inclusive de parte da Internet) é transportada para dentro do IGP;
- Perdem-se os atributos do BGP;
- Ao se exportar prefixos indevidos para outro AS, os mesmos são anunciados como se tivessem sidos originados no próprio AS, tornando-os imunes aos filtros de AS-Path.

A solução para esse problema é nunca utilizar redistribuição de um protocolo em outro, a menos que realizada por criteriosos filtros.

Uso de métricas BGP-4 em Múltiplos PTTs

O atributo MED (Multi Exit Discriminator) [REK1995] pode ser utilizado para discriminar entre múltiplos pontos de saída ou entrada para um mesmo AS vizinho, ou seja, o administrador define por qual dos pontos de contato entre 2 backbones chegará o tráfego originado no seu vizinho. Isso é uma prática comum quando existe uma relação de compra de tráfego. No entanto, muitas vezes essa prática pode não ser aceita ou deixar de ser utilizada por desconhecimento.

O maior impacto aqui é a utilização dos links de acesso intra-rede do AS1 serem utilizados para entregar os pacotes destinados ao AS2 no melhor ponto de troca de trafego que convier ao AS2. Isso é uma possibilidade para ASes que trocam tráfego em múltiplos PTTs, embora seja uma prática não muito utilizada.

A utilização de métricas diferenciadas no PPT é um ponto que necessita de concordância por parte de outros participantes. Essa concordância é necessária para alertar sobre a possibilidade ou não de um backbone solicitar ao outro que entregue o tráfego destinado uma determinada rede em um PPT ou outro. Hoje a maioria dos PTTs nacionais aceita o atributo MED.

Aspectos de Segurança dos route-servers

Um ponto bastante sensível em um PTT que utiliza acordos de tráfego múltiplo reside diretamente no servidor de rotas (route-server). A impossibilidade de esse receber ou divulgar ROUTE_UPDATES na tabela BGP poderá levar o PTT a um estado inconsistente ou até mesmo deixá-lo fora de operação. Quedas ou falhas na manutenção das sessões BGP com o route-server não são aceitáveis. Isso transforma o route-server em um ponto único de falha para todo o PTT.

Dentre as medidas de segurança implementadas para evitar falhas nesse sentido podemos citar a necessidade de existirem dois route servers independentes, cada qual mantendo suas próprias sessões BGP com os outros participantes.

Pelo route-server geralmente tratar-se de uma estação de trabalho rodando um software como o GATED [GAT2001] ou ZEBRA [ZEB2001] sob um sistema operacional UNIX, ele está diretamente exposto a falhas de segurança inerente ao sistema operacional e aplicações que nele executam. Dessa forma a necessidade de configuração desta estação para realizar o papel de um bastion host perante a Internet.

Atualmente o RSIX possui duas estações UNIX (FreeBSD) utilizando o software IP-FILTER[xxx]. Visando diminuir ainda mais a alcançabilidade a essa máquina, ela também se encontra sem uma rota default definida[9]. Ou seja, sua alcançabilidade fica restrita aos roteadores e equipamentos diretamente conectados ao PTT. Cabe aqui salientar que as rotas trocadas pelo route-server com os participantes do PTT em nenhum momento farão parte da tabela de roteamento interna da máquina Unix.

No entanto, devido ao route-server possuir um endereço válido na Internet, ele continua vulnerável aos temidos ataques de negação de serviço distribuídos destinados diretamente aos dois hosts..Estima-se a necessidade de um tráfego superior a 50Mbps destinado a cada route-server para que esses deixem de cumprir suas funções.

Bibliografia

[HUS2001] Huston, Geoff; Analyzing the Internet BGP Routing Table (http://www.cisco.com/warp/public/759/ipj_4-1/ipj_4-1_bgp.html)

[TEL2002] Telstra; "BGP Table Report" updated hourly at (http://www.telstra.net/ops/bgp)

[HUI1995] HUITEMA CRHISTIAN Routing in the Internet. Englewood Cliffs, NJ : Prentice Hall,1995.

[IPF2001] IPFilter Software Package, 2001. (http://www.ipfilter.org)

[CIS2001a] Cisco Systems, Essential IOS Features Every ISP Should Consider – Lessons from people who have been operating backbones since the early days of the Net. Abril, 2001 (http://www.cisco.com/public/cons/isp/documents/IOSEssentialsPDF.zip)

[SEN1999] Senie, D.; Changing the Default for Directed Broadcasts in Routers. Request for Comments 2644. August 1999. (http://www.rfc-editor.org/rfc/rfc2644.txt).

[MAL1995] Malkin, G.; ARP extensions – UNARP Request for Comments 1868. November 1995.

[9] Alguns PTTs preferem utilizar endereços privados para o PTT [REK1996].

[REK1995] Rekhter, Y.; (IBM) T.J. Watson Research Center (Cisco). Request for Comment 1771. Março 1995.
[REK1996] Rekhter, Y.; Moscowitz, B; Karrenberg Z. Address Allocation for Private Internets. Request for Comment 1918. Fevereiro 1996.
[GAT2001] Gateway Routing Daemon, 2001. (http://www.gated.org/)
[ZEB2001] Zebra Multi-server Routing Software, 2001. (http://www.zebra.org)