UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANA
DEPARTAMENTO ACADÊMICO DE ELETRÔNICA
CURSO DE ESPECIALIZACÃO EM CONFIGURAÇÃO E GERENCIAMENTO DE SERVIDORES E EQUIPAMENTOS DE REDES

MAURO BORDINHÃO JUNIOR

DESCRIÇÃO DAS CARACTERÍSTICAS E FUNÇÕES DO PROTOCOLO BGPv4

MONOGRAFIA

CURITIBA
2012

MAURO BORDINHÃO JUNIOR

DESCRIÇÃO DAS CARACTERÍSTICAS E FUNÇÕES DO PROTOCOLO BGPv4

Monografia apresentada como requisito parcial para obtenção do grau de especialista em Configuração e Gerenciamento de Servidores e Equipamentos de Redes, do Departamento Acadêmico de Eletrônica da Universidade Tecnológica Federal do Paraná.
Orientador: Prof. Dr. Augusto Foronda

CURITIBA

2012

Dedico esse trabalho ao meu pai Mauro Bordinhão, que sempre me apoiou e é exemplo de humildade e perseverança nas situações adversas e favoráveis.

E a minha mãe Valdinéia Aparecida de Souza Bordinhão, que concedeu todo amor que um filho necessita e me inspira a ser uma pessoa melhor.

RESUMO

BORDINHÃO, Mauro. **Descrição das características e funções do protocolo BGPv4.** 2012. 32 f. Monografia (Especialização em Configuração e Gerenciamento de Servidores e Equipamentos de Redes). Universidade Tecnológica Federal do Paraná, Curitiba, 2012.

Esta monografia visa descrever as características e funções do *Border Gateway Protocol* (BGP) na versão 4. A pesquisa apresenta os componentes necessários para o funcionamento de conexões BGP entre os sistemas autônomos, são descritas as etapas de uma conexão BGP até ser atingido o pleno estabelecimento da sessão. Também são detalhados os atributos e métricas usados pelo BGP para gerenciar o tráfego de link, políticas de roteamento e técnicas para evitar *loops* de roteamento.

Palavras-chave: BGP. EGP. ASN. *Peer*. *Neighbor*. *Upstream*. Redes.

SUMÁRIO

LISTA DE FIGURAS

1 INTRODUÇÃO

1.1 TEMA

O funcionamento atual da internet é baseado na comunicação entre sistemas autônomos (*Autonomous System* ou somente AS), cada AS é um domínio com políticas administrativas de roteamento.

O *Border Gateway Protocol* (BGP) é o protocolo padrão utilizado na configuração de roteadores que interconectam redes de AS´s, possibilitando que as políticas de roteamento sejam aplicadas.

O BGP é um protocolo do tipo *External Gateway Protocol* (EGP), devido ao fato de manter conexão entre redes de AS´s distintos, no entanto ele pode atuar na comunicação interna entre roteadores usando um mesmo AS, esse procedimento é conhecido como *Internal Border Gateway Protocol* (IBGP).

O BGP também pode atuar em conjunto com protocolos do tipo *Internal Gateway Protocol* (IGP), protocolos de classificação IGP trocam informações dentro de um único AS, alguns exemplos de protocolos IGP são: *Open Shortest Path First* (OSPF), *Routing Information Protocol* (RIP), protocolos de classificação do tipo EGP que é o caso do BGP trocam informações de roteamento tanto dentro de um mesmo AS e/ou entre AS´s distintos.

1.2 DELIMITAÇÃO DA PESQUISA

No decorrer desse documento serão apresentadas as etapas para adquirir um numero de sistema autônomo através das entidades responsáveis e como usar o BGP para efetuar conexões com outros *peers* de modo que seja possível gerenciar o tráfego de rede.

Serão detalhados conceitos de Transito, *Peering* e Rota *Default* que são usados pelos sistemas autônomos.

Atributos BGP vão ser apresentados nos próximos capítulos com imagens e detalhes simulando ambientes reais de operadoras e provedores de acesso.

Também serão apresentados detalhes dos diferentes tipos de tabela de roteamento como a Tabela *Full Routing* que contém todas as rotas de ips da internet ou uma Tabela *Partial Routing* que recebe anúncios limitados do seu P*eer* e dessa forma fica suscetível a ter redução da gerência de tráfego.

1.3 PROBLEMA

Houve um protocolo anterior ao BGP chamado *Exterior Gateway Protocol* (EGP) conforme a *Request For Comments* 904 – Requisição para comentários 904 (RFC), no entanto o EGP possuía muitas limitações, pois ele mantinha uma topologia do tipo árvore a toda internet ou seja em um único *backbone* os sistemas autônomos são conectados como troncos e não em pares. O EGP seguiu até a versão 3 e não foi mais implementado, o BGP surge com a inovação de reconhecer a internet como vários números autônomos interconectados ou seja várias redes de *backbones* interconectadas.

.

1.4 OBJETIVOS

Serão mostrados o objetivo geral e específico que a monografia pretende atingir.

1.4.1 Objetivo Geral

Explicar as funções do protocolo BGP, visando um esclarecimento maior de políticas de roteamento com o BGP.

1.4.2 Objetivos Específicos

Os objetivos específicos são:

- Mostrar a função dos sistemas autônomos;
- Descrever os recursos do protocolo BGP;
- Detalhar como aplicar políticas de roteamento BGP;
- Explicar formas para evitar *loops* de roteamento;

1.5 JUSTIFICATIVA

Ao concluir a elaboração desse documento, o objetivo é que o mesmo sirva para fins de consulta para estudantes que se interessarem em pesquisar protocolos interdomínio, sistemas autônomos, para compreensão das políticas de roteamento que são efetuadas em sistemas autônomos de pequeno e grande porte.

1.6 PROCEDIMENTOS METODOLÓGICOS

Ao desenvolver esse documento, utilizar-se-á das referencias bibliográficas que envolvem o assunto, materiais on-line e dispositivos de redes interconectados tais como *Looking Glasses* localizados em pontos de troca de trafego (PTT).

O estudo explicará as diferenças e vantagens entre o IBGP e o EBGP, as funções dos atributos, boas práticas na utilização do BGP serão evidenciadas visando evitar *loops* de roteamento e incidentes na configuração da rede.

1.7 FUNDAMENTAÇÃO TEÓRICA

O conteúdo referente aos conceitos, utilidades e técnicas que serão abordados nessa monografia, foram coletados de revisão de literatura nas áreas de redes de computadores, obras de Gough (2002), Paquet (2003), Teare (2003), Peterson (2003), Davie (2003) e artigos web, para a parte prática um cenário *Single Multi Home* BGP básico é criado através do software *opensource* sob a licença GNU, chamado GNS3 / *Dynagen* / *Dynamips*, através do mesmo é possível virtualizar imagens de sistemas operacionais de roteadores, o roteador Cisco modelo 3604 é utilizado no cenário *Single Multi Home* apresentado.

Para elucidar algumas nuances do protocolo BGP acessos a *route serves* denominados *Looking Glasses* que estão presentes nos PTT´s nacionais são efetuados no intuito de visualizar resultados de comandos do BGP, tais *Looking Glasses* consistem em servidores que espelham rotas BGP dos PTT´s, na maioria dos casos esses servidores utilizam Linux ou Free-BSD em conjunto com o software *opensource* Quagga/Zebra.

2 BGP

Neste capitulo serão descritas as características do protocolo BGP necessários para compreender seu funcionamento.

2.1 SISTEMAS AUTÔNOMOS

Um sistema autônomo é a identificação única e global de uma entidade possuidora de blocos de endereços ipv4 e/ou ipv6 que usa políticas de roteamento interdomínio através do protocolo BGP. Os sistemas autônomos (*Autonomous System*, ou somente AS) são obtidos geralmente por organizações que demandam grandes serviços de internet ou prestam serviços de internet.

De acordo com PETERSON e DAVIE (2003, p 223): “A ideia básica por trás dos sistemas autônomos é oferecer um modo adicional de agregar a informação de roteamento hierarquicamente em uma grande inter-rede, melhorando assim a escalabilidade”.

O AS é concedido por um *Regional Internet Registries* (RIR), o RIR que atende a América Latina é o *Latin America and Caribbean Network Information Centre* (LACNIC). Para solicitar um AS é necessário preencher um formulário exigido pelo LACNIC: http://lacnic.net/templates/asn-template-pt.txt e aguardar o andamento do processo, alternativamente é possível solicitar um AS através de um *National Internet Registry* (NIR). O NIR do Brasil é o registro.br que também exige um formulário devidamente preenchido: http://registro.br/provedor/numeracao/pedido-form.txt. O mesmo processo é feito para outras requisições de numeração tais como endereços ipv4 ou ipv6.

Os AS´s são identificados numericamente por um numero inteiro de 16 *bits* (2 octetos) variando de 0 até 65535, no entanto devido a supressão de AS´s de 16 bits, foi necessário a criação de novos AS´s que são identificados por um numero inteiro de 32 bits.

O uso de AS´s de 16 bits são organizados da seguinte forma, o AS 0 e 65535 são reservados, o *range* do AS 1 até o AS 64495 são destinados ao uso público. Também há um intervalo que se inicia no AS 64496 até o AS 64511 que são reservados para documentação conforme a RFC 5398 e por fim o intervalo de AS´s de uso privado que se inicia a partir do AS 64512 até o AS 65534.

O uso de AS´s de 32 bits varia de 0 até 4294967295. O intervalo de uso público está no *range* 65552 até 4294967295 e o intervalo de 65536 até 65551 está reservado conforme a RFC 5398.

Um AS funcional tem agregado a si blocos de ipv4 e/ou ipv6 e dessa forma é possível se comunicar com diferentes AS´s e administrar o tráfego de dados recebido e enviado, usando políticas de roteamento.

Uma lista de AS´s nacionais pode ser obtida no endereço:

http://www-public.int-evry.fr/~maigron/RIR_Stats/RIR_Delegations/Delegations/ASN/BR.html#LACNIC

2.2 CONCEITOS DE TRANSITO E *PEERING*

A figura 1 exemplifica os conceitos de Transito e *Peering*.

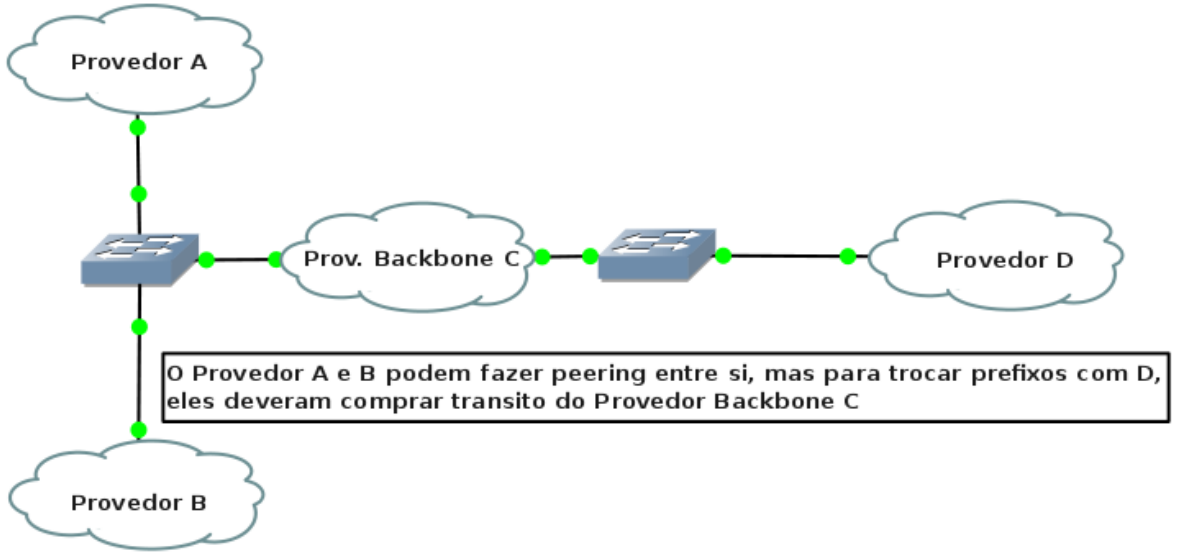

Figura 1: Transito e Peering
Autoria própria, 2012

O conceito de trânsito é a situação em que o roteador *Upstream* estabelece sessão BGP com outro roteador e permite que esse roteador alcance outros roteadores interconectados ao *Upstream*, ou a própria internet. Conforme PETERSON e DAVIE (2003, p 225) "AS de trânsito: um AS que possui conexões com mais de um outro AS e que foi projetada para transportar o tráfego de trânsito e local"

O *Peering* se trata de sessão BGP entre dois roteadores ou mais roteadores para troca de trafego, de acordo com PETERSON e DAVIE (2003, p 225) "AS multiconectado: um AS que possui conexões com mais de um outro AS, mas se recusa a transportar tráfego em trânsito"

2.3 TABELA *FULL ROUTING* E TABELA *PARTIAL ROUTING*

O provedor pode enviar as tabelas de roteamento de diferentes formas, variando conforme uso de recursos do roteador e demanda de gerencia de roteamento por parte do AS que está recebendo a tabela de rotas.

Segundo GOUGH (2003, p 402) “Ao conectar-se a algo amplo como a internet, é necessário que se faça um planejamento e que se tenha prudência. Principalmente, é essencial decidir quais atualizações serão enviadas para o mundo exterior e como os roteadores dentro do sistema autônomo terão conhecimento da necessidade da configuração adicional”.

Pode se optar por receber apenas uma rota *default* com o BGP, e a gerencia de trafego fica a cargo somente do roteador *Upstream*.

A tabela *Full Routing* atualmente contém mais de 400 mil rotas, conforme a figura 2, dessa forma os roteadores em ambas extremidades devem conter *hardware*, especialmente memória e processamento suficiente para suportar tal necessidade.

A tabela *Partial Routing* contém os prefixos definidos pelo *Peer Upstream* conectado e uma rota *default*, dessa forma o numero de rotas recebido é reduzido, os recursos de *hardware* são menos utilizados e consequentemente a flexibilidade para implementar políticas de roteamento para quem recebe a tabela *Partial Routing* é menor.

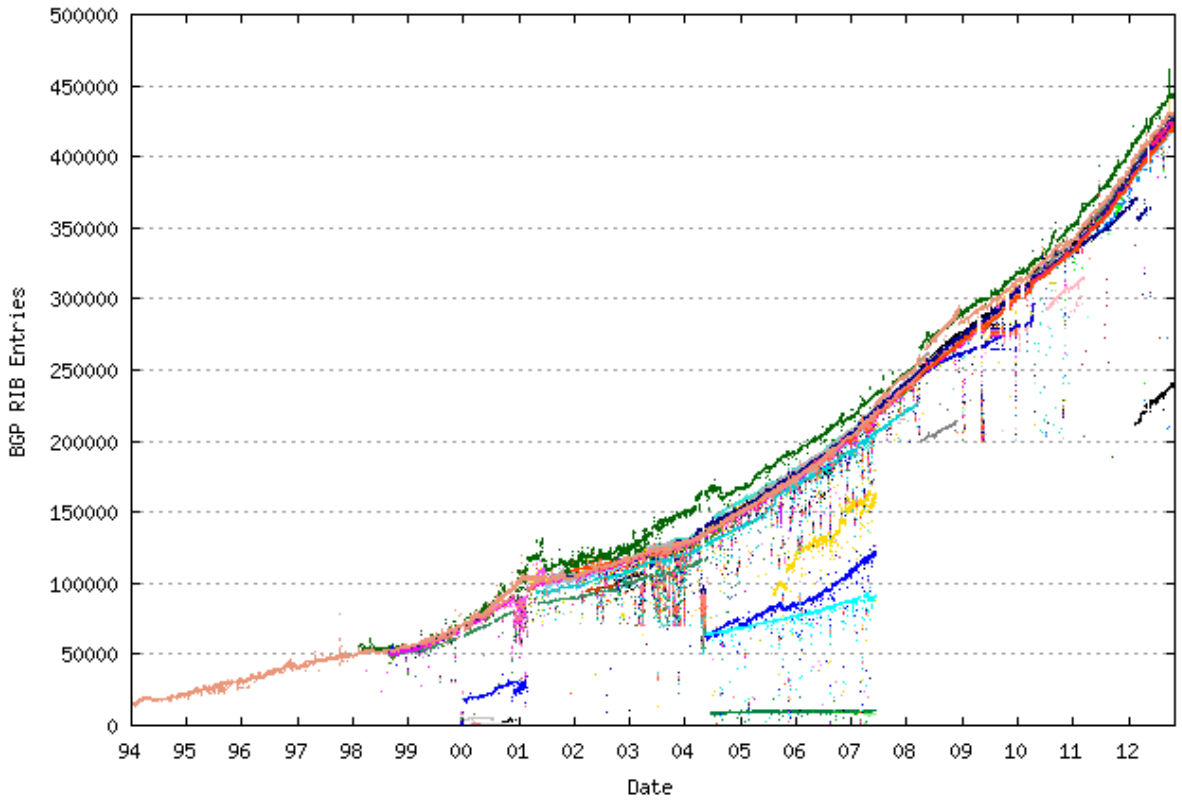

Figura 2: Growth of the BGP Table - 1994 to Present

Fonte: http://bgp.potaroo.net/, 2012

2.4 FUNDAMENTOS DO BGP

O BGP aprende, armazena, anuncia rotas e escolhe o melhor caminho através da internet de um modo global, é o protocolo escolhido para interligar provedores de acesso, operadoras de telecomunicações e empresas que também possuem AS. A versão 4 do BGP é especificada nas RFC´s 1771 e 1772. O BGP é um protocolo de roteamento do tipo *Path Vector*, pois utiliza como um de seus atributos a contagem de AS´s ou AS_*PATH*, esse atributo é muito similar a contagem de saltos utilizada pelo protocolo IGP RIP, no entanto o protocolo RIP é do tipo *Distance Vector*.

Segundo GOUGH (2003, p 341) "Quando um vizinho é visto, uma sessão de formação de pares do TCP é estabelecida e mantida. Os testes do BGP-4 são enviados periodicamente para sustentar o enlace e manter a sessão".

O BGP estabelece uma vizinhança entre os roteadores para efetuar as trocas de informações de roteamento e para garantir a confiabilidade o BGP usa o protocolo TCP e a porta 179 para que as mensagens *OPEN*, *KEEPALIVE*, *UPDATE*, *NOTIFICATION* sejam encaminhadas entre os roteadores.

A mensagem do tipo *OPEN* é utilizada para estabelecer a conexão entre os *peers*. Nela contém os campos de negociação entre os *Peers*, como versão de protocolo BGP, numero AS, campo *HOLD TIME* que delimita o tempo em segundos para a troca de mensagens entre os *peers* e parâmetros adicionais.

A mensagem *KEEPALIVE* é enviada periodicamente para que o tempo especificado pelo *HOLD TIME* não expire, visando manter a conexão ativa entre os *Peers*.

Mensagens do tipo *Update* levam informações para atualizar as tabelas de roteamento, prefixos são carregados por essa mensagem, rotas inalcançáveis assim como atributos que serão posteriormente detalhados.

As mensagens do tipo *NOTIFICATION* são encaminhadas caso ocorrerem erros. Elas contém o código de erro, sub-código de erro e dados relacionados para fins de depuração do problema.

2.5 ESTADOS DA CONEXÃO BGP

A figura 2 monstra o fluxo de etapas até atingir o estado *Established*.

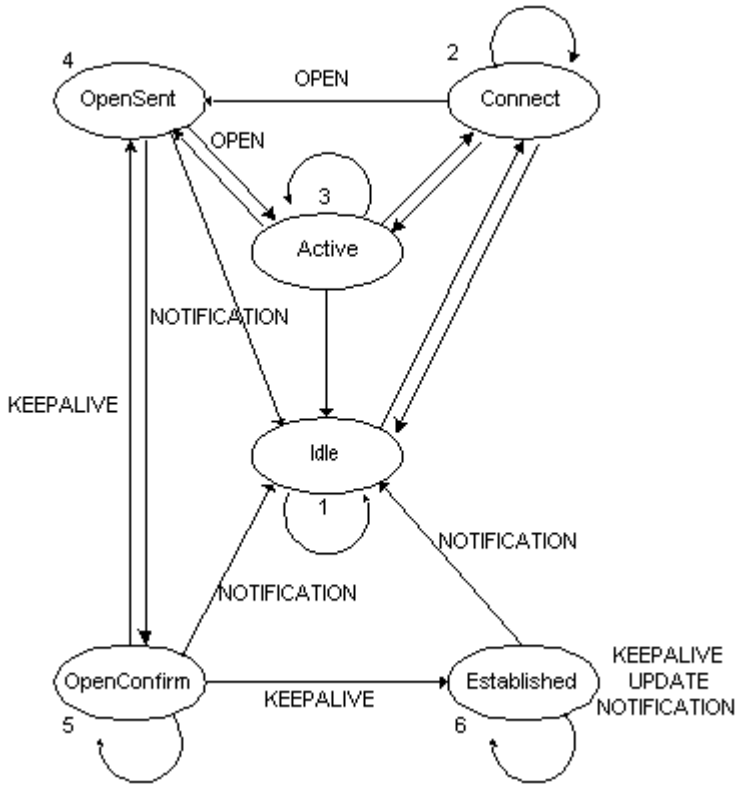

Figura 3: Máquina de estados finitos para sessões BGP.

Fonte: http://www.gtrh.tche.br/ovni/roteamento3/bgp_3.htm, 2003

A conexão BGP passa pelas seguintes etapas para atingir o estado *Established*.

- *Idle*: é o estagio inicial da conexão BGP, o *peer* local deve conter as configurações adequadas e aguarda-se que o *peer* remoto esteja devidamente configurado para avançar até a segunda etapa chamada *Connect*. O estado *Idle* se persistir por muito tempo pode significar um problema remoto ou até mesmo uma conexão interrompida.

- *Connect*: nessa etapa o BGP espera a conexão da camada de transporte do TCP destinando a porta 179, assim que esse processo for executado com o recebimento da mensagem *OPEN*, o próximo estágio é o *OPENSENT*. Se ocorrer algum problema na conexão TCP a próxima etapa será *ACTIVE*, em casos diferentes dos citados a conexão volta para o estagio inicial *Idle*.

- *Active:* no modo *ACTIVE* o *peer* faz a tentativa de conectar através de uma conexão TCP, em caso de sucesso o próximo passo é o *OPENSENT*, em caso de falha o processo voltará para a etapa *CONNECT*.
- *OpenSent:* nesse estado é esperado a mensagem *OPEN*, assim é feito uma verificação e caso exista alguma inconsistência como numero de versão BGP errado ou mesmo o AS errado a mensagem *NOTIFICATION* será encaminhada e o estado retornará para o inicio do processo que no caso é o modo *Idle*, nos casos em que não ocorrerem os problemas citados a conexão iniciará e os pacotes *KEEPALIVE* serão encaminhados, o *Holdtime* será negociado e a escolha será definida pelo menor *Holdtime* entre os *Peers*. Nesse mesmo processo é feito a checagem quanto ao AS afim de identificar se é um mesmo AS em ambos *Peers* se tratando de BGP interno ou em caso de AS´s distintos se caracterizando como uma conexão BGP externa.
- *OpenConfirm:* nessa etapa o *OpenConfirm* aguarda o recebimento do *Keepalive* para prosseguir para a próxima etapa *Established*
- *Established:* É o estado que a conexão foi efetuada com sucesso e os *Peers* podem trocar informações de roteamento através de mensagens *UPDATE*, para manter a conexão ativa a mensagem *KEEPALIVE* será constantemente enviada entre os *Peers*.

2.6 ATRIBUTOS BGP

Os Atributos são propriedades inseridas na mensagem *UPDATE* e a principal função dos atributos é influenciar a entrada ou saída de tráfego.

Conforme GOUGH (2003, p 350) “Os atributos do BGP-4 são usados para determinar o melhor caminho a ser selecionado. Essencialmente, eles são a métrica do BGP-4”

Há dois tipos de atributos definidos como *Well-Known* (Bem Conhecido) e Opcional, os atributos do tipo *Well-Known* são aqueles que todas as implementações do BGP devem reconhecer e consequentemente são repassados aos vizinhos BGP. O atributo definido como *Well-Known* mandatório deve aparecer na descrição da rota, já o atributo Well-Known arbitrário não precisa aparecer na descrição de uma rota.

Um atributos Opcional não é suportado em todas as implementações BGP, o atributo Opcional definido como transitivo que não é implementado em um roteador deve ser passado inalterado aos outros roteadores BGP, em casos como esse o atributo é identificado como

Parcial. Atributos opcionais intransitivos ou somente não transitivos devem ser excluídos por um roteador que não implementou o atributo.

Abaixo alguns exemplos de atributos *Well-Known* mandatórios.

- Atributo AS-*PATH*

O Atributo AS-*PATH* é do tipo *Well-Known* mandatório. Esse atributo registra o caminho percorrido para alcançar uma rota, esse caminho é definido através da contagem de AS´s, quando uma atualização de rota atravessa determinado AS, o numero AS é indexado naquela atualização, o AS-*PATH*, tem uma função importante na prevenção de *loops* de roteamento em redes que usam BGP externo, pois todo o caminho percorrido é registrado e cada salto possui um AS.

A figura 4 abaixo exemplifica o atributo AS-*PATH*.

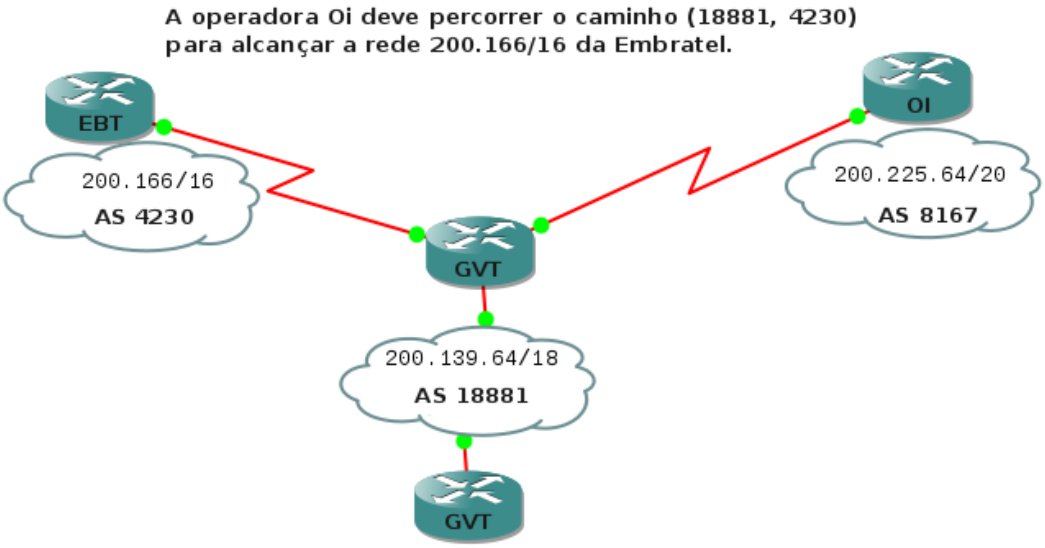

Figura 4: Atributo *AS-PATH*

Fonte: Autoria própria, 2012.

- Atributo *Next-hop*

O Atributo *Next-hop* também é um atributo Well-Known mandatório, ele indica um próximo endereço IP necessário para alcançar uma rede destino. Há algumas diferenças quanto ao uso desse atributo dependendo do modo externo ou interno que o BGP está operando.

Em conexões BGP externas o *Next-Hop* é o endereço IP do *neighbor* que enviou a atualização de roteamento e na figura 4 abaixo o roteador da operadora OI anuncia a rede 200.225.64/20 para o roteador de borda da operadora GVT, dessa forma o Next_Hop que o roteador da GVT usará para alcançar a rede anunciada pela OI será o endereço IP: 172.25.25.2.

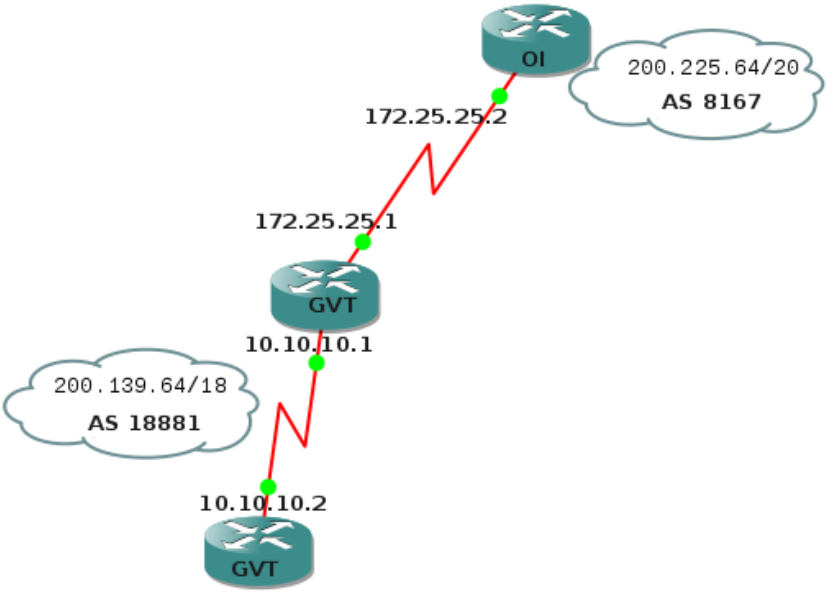

Figura 5: Atributo *Next-Hop*

Fonte: Autoria própria, 2012.

- *Origin*

O Atributo *Origin* é um atributo *Well-Known* mandatório que informa a origem da rota, descrevendo se a rota vem do mesmo AS, geralmente por protocolos IGP, e é descrito com um “i” na tabela BGP, se o AS é distinto se tratando de conexão BGP externa um “e” é informado na tabela BGP.

O atributo *Origin* também identifica se a atualização de roteamento vêm de origem desconhecida, como em casos de redistribuição para o BGP, em situações como essa o atributo *Origin* vai utilizar o valor *Incomplete*.

Abaixo exemplos de atributos Well-Known arbitrários.

- *Local Preference* (*LOCAL_PREF*)

O atributo *Local Preference* é do tipo *Well Known* arbitrário. Ele influencia o tráfego de saída, ou seja ele identifica um melhor caminho para as saidas do AS. O atributo é identificado com um valor numérico que varia entre 0 até 4294967295, um caminho com *Local Preference* maior é escolhido para alcançar redes fora do AS. A figura 5 abaixo descreve um cenário que o atributo *Local Preference* é aplicado.

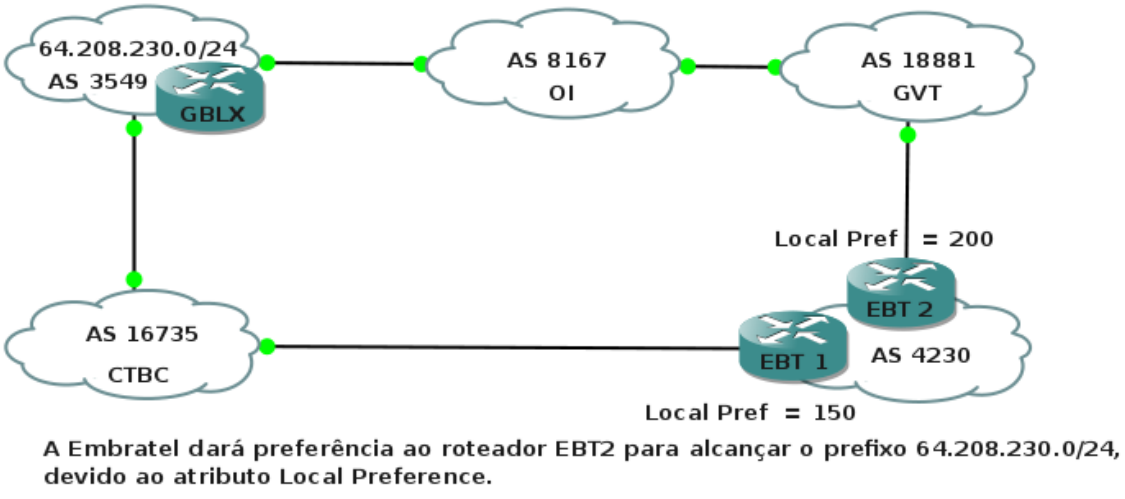

Figura 6: Atributo Local Preference
Fonte: Autoria própria, 2012.

Atributos Opcionais Transitivos

- *Community*

O atributo *Community* é do tipo Opcional Transitivo, devido a isso se um roteador não entender o conceito de *Community* ele não tratará tal informação, somente repassará a informação para o próximo roteador. O conceito de *Community* permite que seja possível filtrar as rotas recebidas e encaminhadas, essa forma é aplicada ao marcar as rotas com um indicador da *Community*, dessa forma os roteadores podem tomar decisões conforme o indicador da *Community*.

Atributo Opcional Não Transitivo

- *Multi Exit Discriminator* (MED)

O Atributo MED é do tipo Opcional não transitivo, ele influencia o trafego de entrada, e o valor mais baixo é reconhecido como melhor e é preferido para entrar no AS. O MED é transportado para um AS e somente nele usado. A figura 6 mostra o atributo MED sendo aplicado de forma que a operadora GVT vai preferir receber atualizações de roteamento do prefixo 200.166/16 através do roteador EBT2 que contém o atributo MED menor.

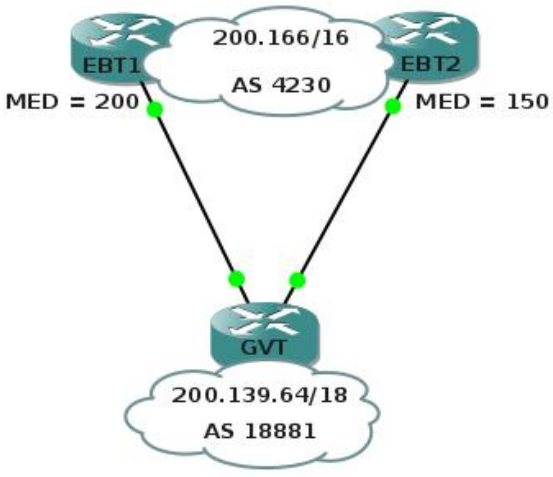

Figura 7: Atributo MED
Fonte: Autoria própria, 2012.

O Atributo MED compara os valores atribuídos apenas quando se trata de um mesmo AS e para que ele considere comparar AS´s distintos, o comando bgp *always-compare-med* deverá estar aplicado no roteador.

2.7 BGP INTERNO E BGP EXTERNO

O BGP interno ou IBGP consiste na utilização do BGP dentro de um mesmo AS, ou seja, o *peering* será feito entre roteadores com um mesmo AS, essa forma de utilização é praticada com o intuito de carregar prefixos de rotas entre os roteadores do AS local usando a eficiência do protocolo BGP, no entanto o IBGP se difere do EBGP devido ao fato que não há como utilizar o atributo AS-*PATH* visto que o AS não mudará ao passar pelos saltos pois se trata sempre do mesmo AS.

Sendo assim cria-se a possibilidade de *loop* e para evitar tal problema uma boa prática é interligar todos os roteadores em forma de malha completa (*full-mesh)*, pois com todos os roteadores interligados é possível que cada um saiba o caminho absoluto para chegar a determinado destino e isso evitará a situação de loop.

Segundo GOUGH (2003, p 388) “A questão com a grande capacidade do BGP-4 é que a escala é limitada à do BGP-4 interno, onde uma configuração totalmente interconectada se faz necessária para garantir total conectividade. Lembre-se de que, para evitar *loops* de roteamento, o protocolo deve seguir a regra de que nenhuma atualização aprendida dos pares internos pode ser enviada a outros pares internos. É semelhante à regra do *split horizon* do IGP”.

O lado negativo de usar uma topologia do tipo *full-mesh* é que conforme a rede se torna maior as interligações expandem a ponto de afetar o desempenho da rede principalmente durante as atualizações trocadas entre os *peers.* Portanto uma rede *full-mesh* é de alto custo e não recomendada em redes com muitas interconexões. Como alternativa existe uma forma segura de evitar *loops* de roteamento com menos custo e com eficiência similar que é chamada de roteador refletor (*router-reflector*). Um *router-reflector* devidamente configurado se trata de um roteador que fica responsável por mandar as atualizações de rotas aos demais *peers* dentro de um mesmo AS.

Conforme GOUGH(2003, p 390) “O refletor de rota desafia a regra do corte horizontal, que afirma que o roteador do IBGP-4 não propagará uma rota que foi aprendida de um par dentro do mesmo sistema autônomo (um par do IBGP4).

No entanto, quando um roteador for configurado como um refletor de rota, ele enviará caminhos aprendidos dos pares do IBGP-4 para outros pares do IBGP-4. Ele fará o encaminhamento somente para os roteadores que tiverem sido identificados como refletores de rota e para clientes dos vizinhos do IBGP/EBGP”.

O BGP externo ou EBGP faz conexão de AS´s distintos visando a troca de prefixos com um único *peer* ou com vários *peers*, também possibilitando a prática de ser um AS multiconectado e fornecer trânsito a outros AS´s. O atributo AS-*PATH* é adicionado em cada salto que a rota percorrer e também auxilia para que *loops* de roteamento não aconteçam.

2.8 PROCESSO DE ESCOLHA DA MELHOR ROTA

Segundo PAQUET e TEARE (2003, p 312) “Após o BGP receber as atualizações sobre os diferentes destinos dos diversos sistemas autônomos (ASs), o protocolo resolve qual caminho deve ser escolhido para atingir um destino específico”.

Caso exista um prefixo mais especifico, esse será preferido, como por exemplo o prefixo 200.146.64.0/24 é mais especifico que o prefixo 200.146.0.0/16. Quando existir varias rotas para um mesmo destino, o processo de escolha deve ser baseado nas informações contidas nos atributos e segue a ordem abaixo:

- *Next-Hop* – Caso não exista rota para o endereço de *Next-Hop*, ou o mesmo esteja inalcançável por qualquer outra razão, essa rota é descartada.
- *Weight* – O atributo *Weight* é proprietário em soluções Cisco, e caso seja essa a arquitetura utilizada a rota preferida é escolhida baseada no valor maior do atributo *Weight*.
- Se houver empate no atributo *Weight*, ou se simplesmente for uma arquitetura diferente da Cisco que não utiliza tal atributo, o próximo que será avaliado é o *Local Preference*, que será preferido conforme o maior *Local Preference*.
- Em caso de empate do atributo *Local Preference*, a rota gerada pelo roteador local deve ser escolhida.

- Se a rota não foi gerada localmente deve se escolher a rota com o AS-*PATH* mais curto.
- Se o atributo AS-*PATH* for igual, o atributo *Origin* deve ser avaliado respeitando a seguinte ordem: IGP < EGP < *incomplete*.
- Se todos os códigos do atributo *Origin* forem iguais, o atributo MED será o próximo a ser analisado e a preferência será definida pelo MED menor.
- No caso de empate no atributo MED, a preferência será dada para rotas EBGP ao invés de IBGP.
- Quando não há rotas EBGP, a preferência é dada ao roteador IGP mais próximo.
- E por ultimo, o desempate para a escolha da melhor rota será definido através do roteador com o menor Router-ID BGP.

3 LABORATÓRIO BGP

Nesse capitulo será feito a simulação de um cenário simples conhecido como *SINGLE MULTI HOME*, o software GNS3/*Dynamips*/*Dynagen* será utilizado para a virtualização do sistema operacional *INTERNETWORK OPERATING SYSTEM* ou simplesmente IOS da companhia Cisco no roteador modelo 3604 também do fabricante Cisco. Um cenário do tipo *SINGLE MULTI HOME* consiste em um roteador que faz conexões BGP externas com dois ou mais provedores estabelecendo sessões BGP e assim podendo anunciar e receber prefixos ip´s desses dois ou mais roteadores.

3.1 LABORATÓRIO BGP EXTERNO *SINGLE MULTI HOME*

Nessa simulação de ambiente de BGP a empresa fictícia chamada *Networking-Data* possui o AS 500, são estabelecidas duas sessões BGP, uma sessão com o provedor A de AS 700 e outra sessão com o provedor B de AS 900.

Os provedores A e B possuem uma interface *loopback* configurada em cada roteador de borda, as interfaces *loopback* estão configuradas com um endereço ip da rede que é anunciada por cada provedor através do protocolo BGP para a empresa fictícia *Networking-Data*.

Esse cenário é básico e o intuito é visualizar o estabelecimento de sessão BGP entre os roteadores e resultados de comandos de consulta e manipulação BGP.

No exemplo a empresa *NETWORKING-DATA* possui apenas um roteador de borda e assim pode garantir redundância de acesso de provedores e possibilidade de gerência de tráfego caso seja necessário.

A figura 8 descreve o cenário proposto:

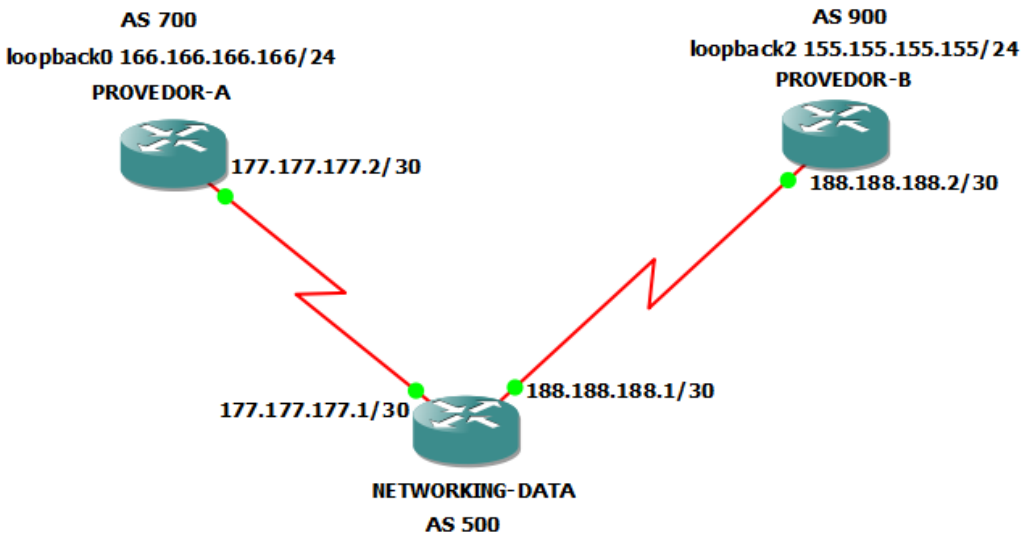

Figura 8: Cenário da empresa fictícia *NETWORKING-DATA*.
Fonte: Autoria própria, 2012.

3.2 CONFIGURAÇÃO DOS ROTEADORES

Roteador de borda NETWORKING-DATA:

Primeiro acesso, modo *enable*, modo *configure-terminal*, adicionando *hostname*:

```
R1$ enable
R1# configure terminal
R1(config)# hostname NETWORKING-DATA
```

Configuração ip da interface serial 0/1 conectada ao provedor A:

```
NETWORKING-DATA(config)#
NETWORKING-DATA(config)#interface serial 0/1
NETWORKING-DATA(config-if)#
NETWORKING-DATA(config-if)#ip address 177.177.177.1 255.255.255.252
NETWORKING-DATA(config-if)#no shutdown
```

Configuração ip da interface serial 0/0 conectada ao provedor B:

```
NETWORKING-DATA(config)#
NETWORKING-DATA(config)#interface serial 0/0
NETWORKING-DATA(config-if)#
NETWORKING-DATA(config-if)#ip address 188.188.188.1 255.255.255.252
NETWORKING-DATA(config-if)#no shutdown
```

Configuração para estabelecimento de sessão BGP com o provedor A:

NETWORKING-DATA(config)
NETWORKING-DATA(config)#router bgp 500
NETWORKING-DATA(config-router)#neighbor 177.177.177.2 remote-as 700

Descrição para o *neighbor*:

NETWORKING-DATA(config-router)#neighbor 177.177.177.2 description eBGP com o provedor A

Configuração para estabelecimento de sessão BGP com o provedor B:

NETWORKING-DATA(config)
NETWORKING-DATA(config)#router bgp 500
NETWORKING-DATA(config-router)#neighbor 188.188.188.2 remote-as 900

Descrição para o *neighbor*:

NETWORKING-DATA(config-router)# neighbor 188.188.188.2 description eBGP com o provedor B

Configuração do roteador de borda no PROVEDOR A:

Primeiro acesso, modo *enable*, modo *configure-terminal*, adicionando *hostname*:

R2$ enable
R2# configure terminal
R2(config)# hostname PROVEDOR-A

Configuração ip da interface serial 0/0 conectada ao cliente NETWORKING-DATA:

PROVEDOR-A(config)#interface serial 0/0
PROVEDOR-A(config-if)#ip address 177.177.177.2 255.255.255.252

Adicionando banda e habilitando a interface serial 0/0:

PROVEDOR-A(config-if)#clock rate 2015232
PROVEDOR-A(config-if)#no shutdown

Configuração de interface *loopback* e endereçamento da mesma:

PROVEDOR-A(config)#interface loopback0
PROVEDOR-A(config-if)#ip address 166.166.166.166 255.255.255.0
PROVEDOR-A(config-if)#no shutdown

Configuração para estabelecimento de sessão BGP com o cliente NETWORKING-DATA:

PROVEDOR-A(config)#router bgp 700
PROVEDOR-A(config-router)#neighbor 177.177.177.1 remote-as 500

Descrição para o *neighbor*

PROVEDOR-A(config-router)# neighbor 177.177.177.1 description eBGP com NETWORKING-DATA

Anuncio da rede 166.166.166.0/24 para o cliente NETWORKING-DATA:

PROVEDOR-A(config-router)#network 166.166.166.0 mask 255.255.255.0

Configuração do roteador de borda no PROVEDOR B:

Primeiro acesso, modo *enable*, modo *configure-terminal*, adicionando *hostname*:

R3$ enable
R3# configure terminal
R3(config)# hostname PROVEDOR-B

Configuração ip da interface serial 0/0 conectada ao cliente NETWORKING-DATA:

PROVEDOR-B(config)#interface serial 0/0
PROVEDOR-B(config-if)#ip address 188.188.188.2 255.255.255.252

Adicionando banda e habilitando a interface serial 0/0:

PROVEDOR-B(config-if)#clock rate 2015232
PROVEDOR-B(config-if)#no shutdown

Configuração de interface *loopback* e endereçamento da mesma:

PROVEDOR-B(config)#interface loopback2
PROVEDOR-B(config-if)#ip address 155.155.155.155 255.255.255.0
PROVEDOR-B(config-if)#no shutdown

Configuração para estabelecimento de sessão BGP com o cliente NETWORKING-DATA:

PROVEDOR-B(config)#router bgp 900
PROVEDOR-B(config-router)#neighbor 188.188.188.1 remote-as 500

Descrição para o *neighbor*

PROVEDOR-B(config-router)# neighbor 188.188.188.1 description eBGP com NETWORKING-DATA

Anuncio da rede 155.155.155.0/24 para o roteador de borda NETWORKING-DATA:

PROVEDOR-B(config-router)#network 155.155.155.0 mask 255.255.255.0

A figura 9 abaixo, detalha a topologia, especificando as interfaces seriais conectadas entre os roteadores.

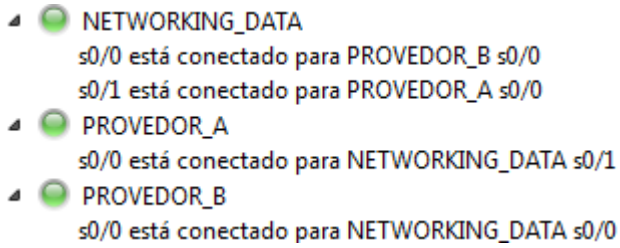

Figura 9: Resumo da topologia.

Fonte: Autoria própria, 2012.

3.3 EFETUANDO TESTES PARA CHECAR A CONECTIVIDADE.

O comando *ping* pode ser utilizado para checar a conectividade entre os *peers*, a partir da empresa *NETWORKING-DATA* deve-se alcançar os prefixos anunciados e ip´s diretamente conectados, no entanto para que seja possível utilizar o comando *ping* entre os provedores A e B, será necessário adicionar como *source* as interfaces *loopback* conforme descrito nos comandos efetuados no provedor A e provedor B nas respectivas figuras 10 e 11 abaixo:

```
PROVEDOR-A#ping 155.155.155.155 source loopback0

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 155.155.155.155, timeout is 2 seconds:
Packet sent with a source address of 166.166.166.166
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 260/268/276 ms
PROVEDOR-A#
```

Figura 10: Comando *ping* 1.
Fonte: Autoria própria, 2012.

```
PROVEDOR-B#ping 166.166.166.166 source loopback2

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 166.166.166.166, timeout is 2 seconds:
Packet sent with a source address of 155.155.155.155
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 272/315/344 ms
PROVEDOR-B#
```

Figura 11: Comando *ping* 2.
Fonte: Autoria própria, 2012.

3.4 COMANDOS DE CONSULTA BGP

Comando: show ip bgp

Após a configuração efetuada, para consultar as informações dos *peers* / *neighbors* conectados, prefixos ip´s recebidos no roteador de borda *NETWORKING-DATA* o comando abaixo pode ser emitido no IOS, conforme a figura 12:

```
NETWORKING-DATA#show ip bgp
BGP table version is 3, local router ID is 188.188.188.1
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
              r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 155.155.155.0/24 188.188.188.2            0             0 900 i
*> 166.166.166.0/24 177.177.177.2            0             0 700 i
NETWORKING-DATA#
```

Figura 12: Comandos 1.
Fonte: Autoria própria, 2012

O resultado obtido contém informações referentes ao endereço ip *NEXT HOP*, endereço de *router* ID. O caminho de AS *PATH* pode ser consultado na coluna *PATH*, outros atributos como *Local Preference* e o atributo somente utilizado em dispositivos Cisco chamado *Weight* podem ser consultados.

Comando: show ip bgp summary

Retorna informações extremamente úteis conforme a figura 13 abaixo:

```
NETWORKING-DATA#show ip bgp summary
BGP router identifier 188.188.188.1, local AS number 500
BGP table version is 3, main routing table version 3
2 network entries using 202 bytes of memory
2 path entries using 96 bytes of memory
2 BGP path attribute entries using 120 bytes of memory
2 BGP AS-PATH entries using 48 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 466 total bytes of memory
BGP activity 2/0 prefixes, 2/0 paths, scan interval 60 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
177.177.177.2   4   700       6       6        3    0    0 00:01:16        1
188.188.188.2   4   900       6       6        3    0    0 00:01:07        1
```

Figura 13: Comandos 2.
Fonte: Autoria própria, 2012

Neighbor	V	AS	MsgRcvd	MsgSent	TblVer	InQ	OutQ	Up/Down	State/PfxRcd
177.177.177.2	4	700	6	6	3	0	0	00:01:16	1
188.188.188.2	4	900	6	6	3	0	0	00:01:07	1

No trecho do resultado do comando informado acima, a coluna *Neighbor* disponibiliza os ips dos *peers* remotos, a coluna V informa a versão corrente do protocolo BGP que no caso é a 4, a coluna AS informa os AS´s remotos, nas colunas MsgRcvd/MsgSent são informadas o numero de mensagens trocadas entre os *peers*, entre essas mensagens estão as mensagens de *update*, *keepalive*, etc.

A coluna TblVer é sempre alterada sucessivamente quando uma nova rota é inserida, ou seja ela é a versão da tabela BGP.

Os itens InQ e OutQ são atualizados conforme a fila de entrada e saída, já os itens da coluna UP/DOWN informam o tempo que a sessão está operante ou não.

O ultimo item State/PfxRcd é muito importante para identificar o estado da conexão BGP, nessa coluna os resultados *IDLE*, OPEN, OPENSENT, ACTIVE, ESTABLISHED podem ocorrer dependendo do estado da conexão, nessa mesma coluna em caso de sucesso na conexão os números de prefixos são informados, na figura 14 abaixo o resultado do comando informa que a sessão BGP com o *peer* 188.188.188.2 está no estado *ACTIVE*.

```
NETWORKING-DATA#show ip bgp summary
BGP router identifier 188.188.188.1, local AS number 500
BGP table version is 2, main routing table version 2
1 network entries using 101 bytes of memory
1 path entries using 48 bytes of memory
1 BGP path attribute entries using 60 bytes of memory
1 BGP AS-PATH entries using 24 bytes of memory
0 BGP route-map cache entries using 0 bytes of memory
0 BGP filter-list cache entries using 0 bytes of memory
BGP using 233 total bytes of memory
BGP activity 1/0 prefixes, 1/0 paths, scan interval 60 secs

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
177.177.177.2   4   700      11      10        2    0    0 00:06:02        1
188.188.188.2   4   900       0       0        0    0    0 never    Active
```

Figura 14: Comandos 3.
Fonte: Autoria própria, 2012.

Na figura 15 abaixo, é aplicado novamente o comando: show ip bgp summary no *Looking Glass* do PTT no Rio de Janeiro pertencente ao nic.br, podemos visualizar que a coluna *State*/PfxRcd tem vários estados diferentes.

```
lg.rj.ptt.br$ show ip bgp summary
BGP router identifier 200.219.138.252, local AS number 20121
RIB entries 192361, using 18 MiB of memory
Peers 29, using 129 KiB of memory

Neighbor        V    AS MsgRcvd MsgSent   TblVer  InQ OutQ Up/Down  State/PfxRcd
200.219.138.3   4 16397      15      16        0    0    0 never    Active
200.219.138.5   4 28338  382307  365828        0    0    0 04w5d13h      112
200.219.138.6   4 16735 1073478  730584        0    0    0 1d21h48m     1834
200.219.138.9   4 28573  661867  352785        0    0    0 02w4d23h      217
200.219.138.10  4 14026       0       0        0    0    0 never    Connect
200.219.138.11  4 28347  135092  132655        0    0    0 23w1d05h Connect
200.219.138.15  4 28338  383168  365774        0    0    0 05w0d03h      103
200.219.138.16  4 262534  389935  354416        0    0    0 03w0d12h        1
200.219.138.17  4 28604  821631  337049        0    0    0 4d07h52m    11803
200.219.138.18  4 28338  310250  297584        0    0    0 06w3d08h      114
200.219.138.19  4 52868  226829  208090        0    0    0 08w5d03h        2
200.219.138.22  4 26615  156442  142728        0    0    0 03w1d20h      201
200.219.138.23  4 52868  227663  208420        0    0    0 10w2d08h        2
200.219.138.101 4  1916 5130643  732153        0    0    0 22w0d13h    35668
200.219.138.104 4 28301  225164  233512        0    0    0 06w4d06h        1
200.219.138.107 4 14026  651754  365667        0    0    0 10w2d16h     6967
200.219.138.108 4 53157  402948  366301        0    0    0 04w1d21h        3
200.219.138.109 4 28301  284037  267629        0    0    0 03w0d13h        1
200.219.138.113 4 18881   63146   41539        0    0    0 02w1d14h     3414
200.219.138.250 4 26162    8535  179350        0    0    0 36w1d12h Active
200.219.138.253 4 26162  697329  365882        0    0    0 10w2d16h     5795
200.219.138.254 4 26162  738465  366480        0    0    0 31w3d15h     5701

Total number of neighbors 22
```

Figura 15: Comandos 4.
Fonte: Autoria própria, 2012.

4 CONCLUSÃO

Nesse capitulo serão descritos os objetivos alcançados nesse trabalho.

4.1 RESULTADOS OBTIDOS

A pesquisa feita sobre o protocolo BGP mostrou a importância do mesmo para que a internet seja mantida em funcionamento, o BGP é complexo e exige profissionais altamente capacitados para que ambientes de alta disponibilidade possam usufruir de todas as vantagens que esse protocolo disponibiliza como por exemplo: redundância, políticas de roteamento, gerenciamento de tráfego, balanceamento de carga, etc

O controle rigoroso na liberação de AS´s se faz necessário visto o grau de complexidade, conhecimento necessário e também obviamente o risco que um AS mal configurado pode causar a toda a rede.

O cenário implementado representou em pequena escala e com pouca complexidade parte do escopo de como a grande rede internet funciona.

Não há indícios de um sucessor para o BGP tão cedo, pois o mesmo atende toda a demanda com flexibilidade e já trabalha em conjunto com novas tecnologias como o endereçamento ipv6.

Atualmente órgãos gestores de internet como o NIR do Brasil nic.br tem mantido pontos de troca de tráfego em diversas regiões, os PTT´s são responsáveis por facilitar o trafego entre AS´s efetuando sessões BGP em um único local para fins de troca de tráfego viabilizando o fluxo de dados e se tornando uma ferramenta útil para provedores e operadoras que se interligam através desse recurso usando o protocolo BGP.

4.2 REFERÊNCIAS

GOUGH, Clare. **CCNP Routing – Guia de Certificação do Exame:** Alta Books, 2002.

PAQUET, Catherine.; TEARE, Diane. **Construindo Redes Cisco Escaláveis.** São Paulo: Cisco Press, 2003.

PETERSON, Larry L.; DAVIE, Bruce S. **Redes de Computadores – Uma abordagem de Sistemas**. 3 ª ed. Rio de Janeiro: Elvesier, 2004.