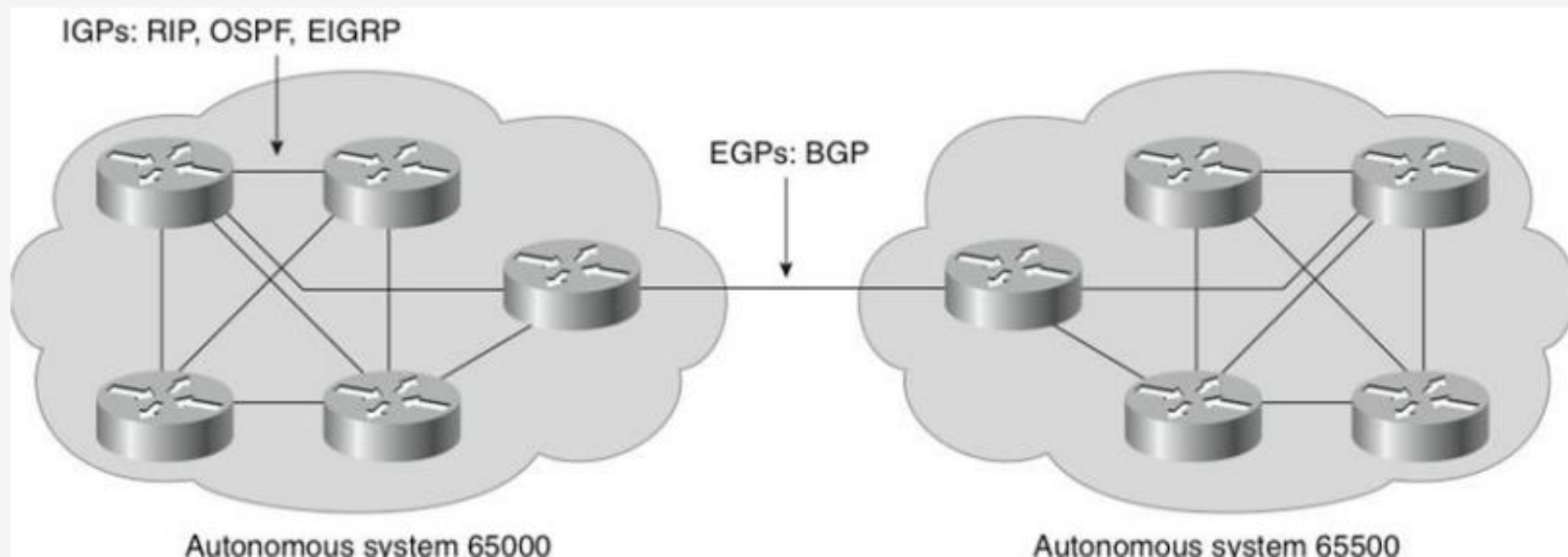




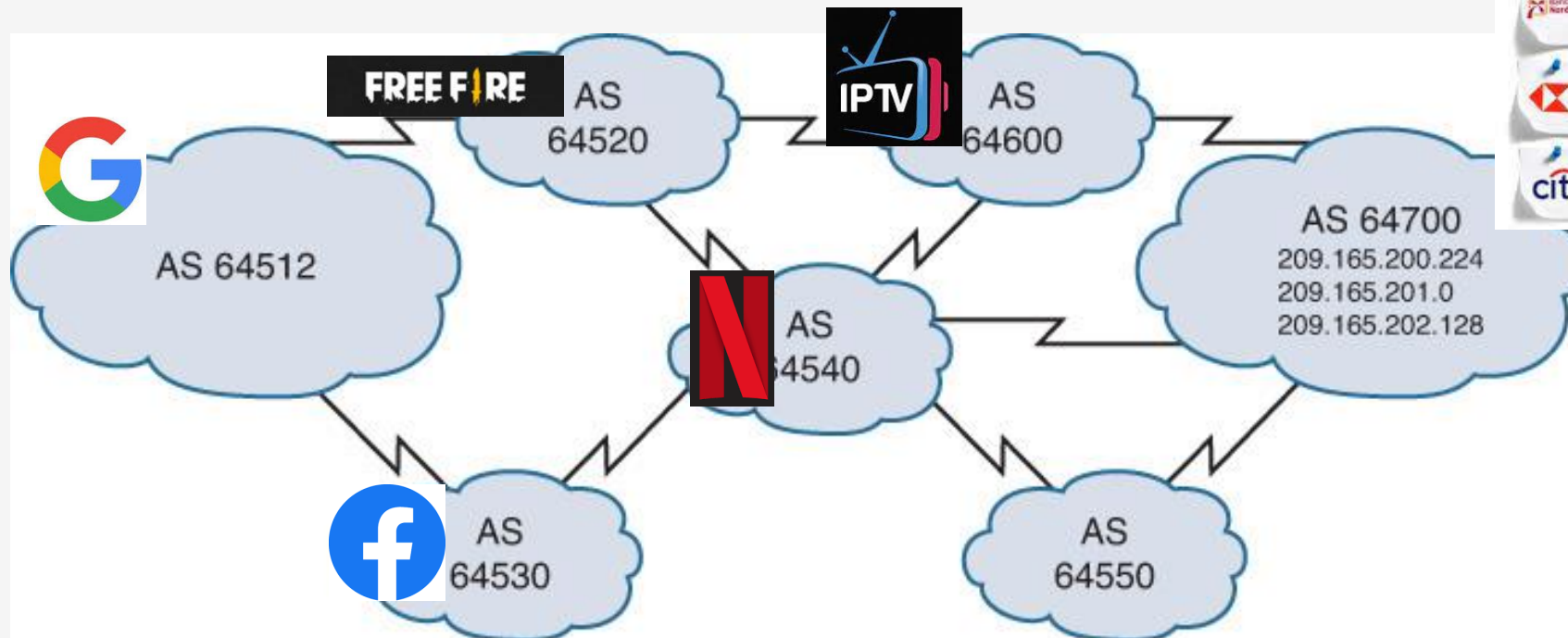
BGP

- No assunto roteadores, podemos dizer que cada roteador utiliza o método hop-by-hop PARA REALIZAR ROTEAMENTO DE PACOTES IP, que consiste em abrir os cabeçalhos dos pacotes que recebe em busca do endereço IP do destinatário, calcula o próximo salto mais próximo do seu destino e entrega o pacote neste próximo salto.
- Este processo se repete até o pacote chegar ao seu destino.



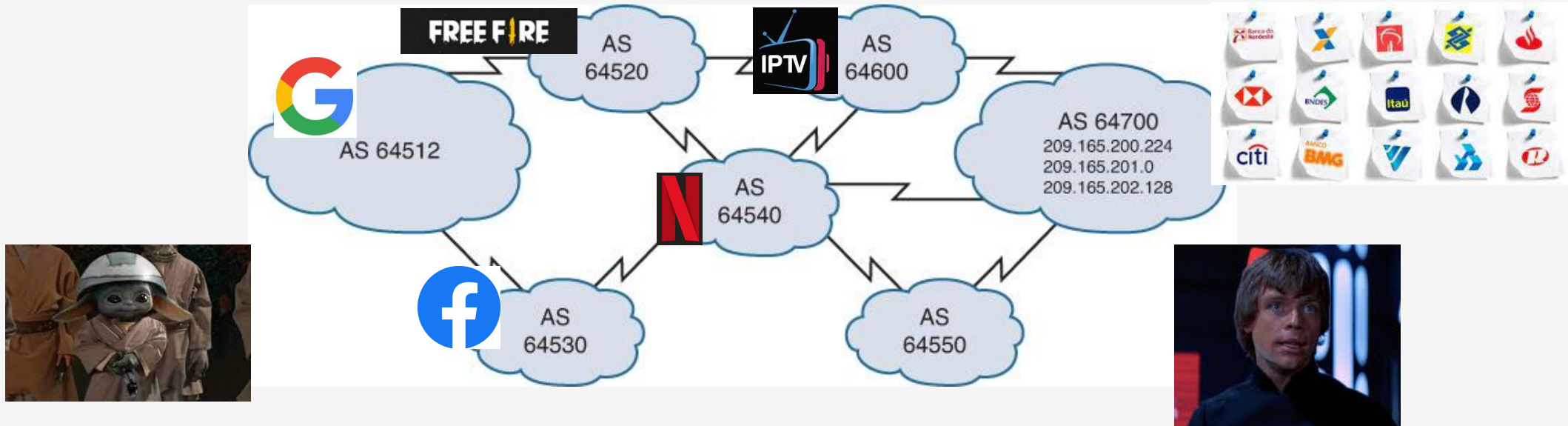
BGP

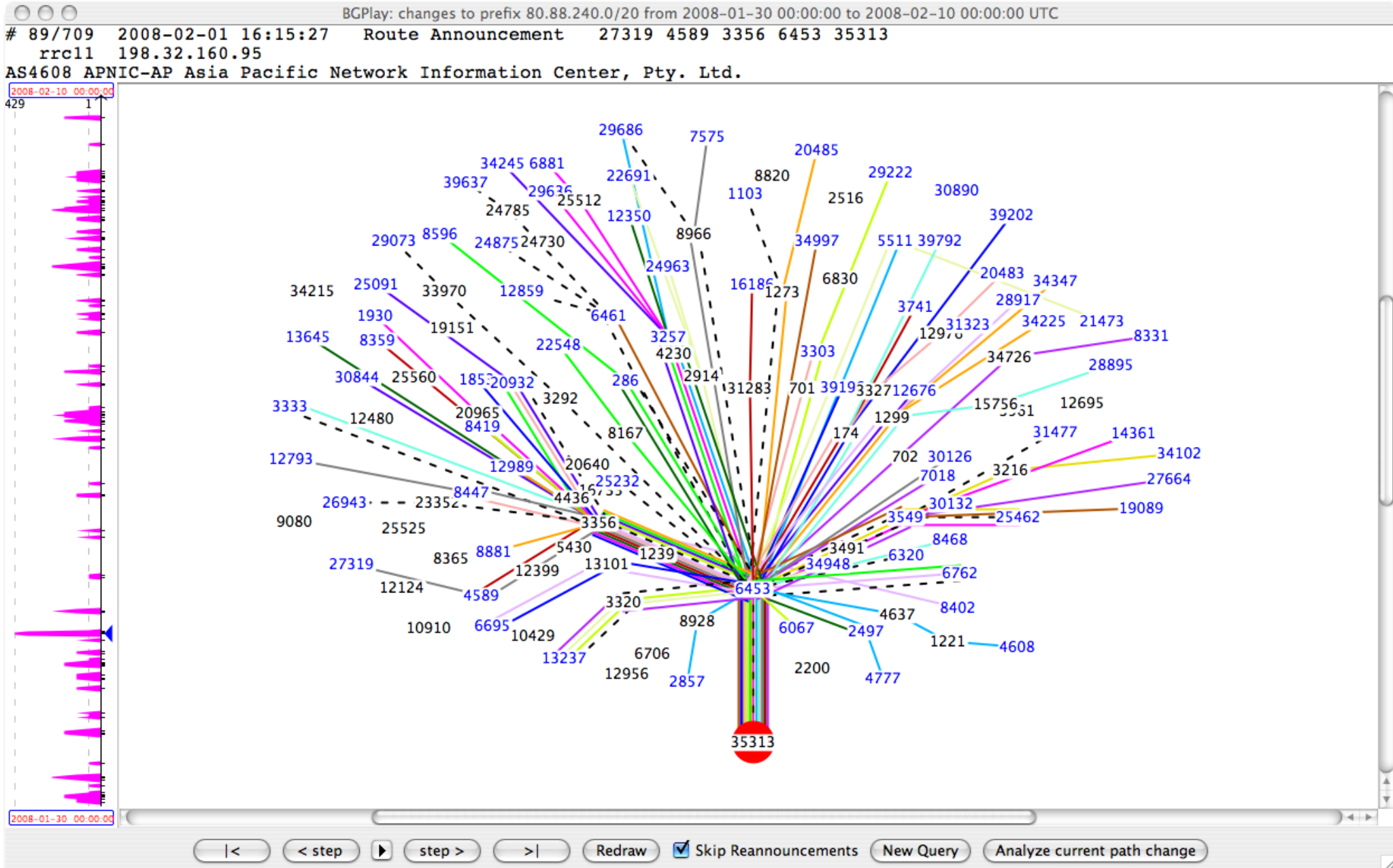
- A Internet é formada por muitas redes diferentes, que se reúnem em uma grande teia global. O IP (Internet Protocol) é a tecnologia usada como conjunto de regras de comunicação, que permite que todas essas diferentes redes operem da forma que conhecemos hoje!



BGP

- ASN é a sigla em inglês para Autonomous System Number, ou Número de Sistema Autônomo.
- A Internet é uma rede de redes. Uma rede formada pela operação conjunta de milhares de redes, de diferentes instituições, com funções distintas: provedores de acesso, provedores de conteúdo, universidades, empresas usuárias da Internet, órgãos do governo, etc. Estas redes que formam a Internet são os Sistemas Autônomos, em inglês Autonomous System (AS).





BGP

- Um sistema autônomo (AS) é uma coleção de roteadores ou redes sob a mesma administração técnica.
- Um protocolo de roteamento interno (IGP) é executado dentro de um sistema autônomo para prover o roteamento interno do AS
- O protocolo de roteamento externo (EGP) é executado entre Sistemas Autônomos para permitir troca de informações de roteamento (NLRI) Network Layer Reachability Information entre eles, esse protocolo é o BGP.
- O ASN é um número de 16 bits ou 32 bits, alocado por um RIR ou NIR para Sistemas Autônomos, e que os identifica de forma única no sistema de roteamento BGP.
- Os números podem estar entre 1 a 65.535.
- RIRs administram os números de AS entre 1 e 64512.
- Os números de AS 64.512 - 65.535 são reservados para uso privado (similarao endereço IP privado) - De uso "proibido": 0, 23456, 65535 e 4294967295;

Regional Internet Registries (RIRs)

RIR Name	Geographic Coverage	Link
AfriNIC	Continent of Africa	www.afrinic.net
APNIC (Asia Pacific Network Information Centre)	Asia Pacific region	www.apnic.net
ARIN (American Registry for Internet Numbers)	Canada, the United States, and several islands in the Caribbean Sea and North Atlantic Ocean	www.arin.net
LACNIC (Latin America and Caribbean Internet Addresses Registry)	Central and South America and portions of the Caribbean	www.lacnic.net
RIPE (Réseaux IP Européens)	Europe, the Middle East, and Central Asia	www.ripe.net

[Sobre Domínios](#) ▾[Tecnologia](#) ▾[Ajuda](#) ▾[Quem Somos](#)[Contato](#)[REGISTRE](#)[Home](#) › [Tecnologia](#) › [Numeração](#) › [Como solicitar](#)

Como solicitar

O Registro.br tratará todas as solicitações de recursos de Numeração Internet por ordem de recepção e processamento sem fazer qualquer distinção. Sob nenhuma hipótese será dado um tratamento especial ou diferenciado nas solicitações.

Considerando o esgotamento dos endereços IPv4 na região, como **noticiado**, não há mais como garantir a alocação desses recursos.

Novas alocações de endereços IPv4 podem ocorrer caso endereços sejam devolvidos ou recuperados. Mas não há como garantir quando isso pode acontecer.

Organizações que assim desejarem, podem solicitar IPv4 e aguardar em uma fila ordenada pela data de solicitação.

A posição de um solicitante por IPv4 na lista de espera está atrelada unicamente ao seu CNPJ não podendo ser transferida a terceiros.

Endereços IPv6 e ASN podem ser solicitados normalmente.

BGP

- Com a previsão de esgotamento números de AS utilizando dois bytes o IETF lançou as RFC 4893 e RFC 5398 para aumentar o número de AS de dois octetos (16 bits), para quatro octetos (32 bits), aumentando o tamanho do conjunto de valores de 65536 a 4294967296.
- A Internet é um conjunto de sistemas autônomos interligados para permitir a comunicação entre redes.
- Uma rede que não é um Sistema Autônomo estará sempre na dependência de uma outra rede que assim seja classificada, como por exemplo, seu provedor de trânsito Internet.
- Por isso, estará dependente também dos endereços IP atribuídos por esse provedor de trânsito e das políticas de roteamento por ele utilizadas.
- Para um provedor Internet, não ser um Sistema Autônomo e depender do seu provedor de trânsito IP é uma grande, enorme, limitação..
- NO MUNDO IDEAL, CADA EMPRESA QUE DESEJA PARTICIPAR DA INTERNET DEVE (OU DEVERIA) OBTER UM NÚMERO DE ASN E FAIXAS DE PREFIXO (IPV4 E IPV6), E A PARTIR DESTA IDEIA SE CONECTAR AS VARIAS REDES INTERCONECTADAS (INTER-AS) PELA REDE GLOBAL, OU CONECTADAS VIA BGP.

BGP

- A primeira menção ao bgp foi no final dos anos 80, sendo ele sucessor do egp.
- criado pela cisco/ibm com objetivo de trocar rotas em asn's usando o protocolo ipv4.
- sua publicação foi na rfc 1105, atualizado posteriormente pelas rfc 1771 e rfc 4271.
- atualmente o bgp esta na versão 4, é um protocolo do tipo " path vector ", utiliza a porta 179 – tcp para estabelecer as " sessões bgp " ou peer inter as.
- por ser um protocolo vetor de caminho, diferente do igp ospf por ex (distance vector), o bgp anuncia os caminhos (paths) e as redes são alcançaveis no final desse caminho
- o bgp descreve o caminho por meio de atributos complexos em muitos casos.

```
BGP-P-AS100#sh ip bgp
BGP table version is 26841439, local router ID is 177.1.1.1
  Network        Next Hop        Metric LocPrf Weight Path
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
               r RIB-failure, S Stale, m multipath, b backup-path, f RT-Filter,
               x best-external, a additional-path, c RIB-compressed,
               t secondary path, L long-lived-stale,
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI validation codes: V valid, I invalid, N Not found

   Network        Next Hop        Metric LocPrf Weight Path
*   0.0.0.0        100.127.0.45      0         0 3389 i
*>  100.127.0.25    0         0 3389 i
*   1.0.0.0/24      100.127.0.45      0 3389 3356 6762 13335 i
*>  100.127.0.25    0 3389 3356 6762 13335 i
*   1.0.4.0/22      100.127.0.45      0 3389 6939 4826 38803 38803 38803 i
i
```

Distance Vector:
Distance = saltos;
Vector = roteador;
Path vector:
Path = Atributos
Vector = ASN.

BGP

- 1989 - RFC 1105 - A Border Gateway Protocol (BGP);
- 1994 - RFC 1654 - A Border Gateway Protocol 4 (BGP-4);
- 1995 - RFC 1771 - A Border Gateway Protocol 4 (BGP-4);
- 2006 - RFC 4271 - A Border Gateway Protocol 4 (BGP-4);
- 2011 - RFC 6286 - Autonomous-System-Wide Unique BGP Identifier for BGP-4;
- 2012 - RFC 6608 - Subcodes for BGP Finite State Machine Error;
- 2012 - RFC 6793 - BGP Support for Four-Octet Autonomous System (AS) Number Space;
- 2015 - RFC 7606 - Revised Error Handling for BGP UPDATE Messages; i



BGP

- Atualizações confiáveis usando o TCP na porta 179.
 - Tabela completa BGP é trocada quando a conexão com o vizinho é estabelecida.
 - todos os vizinhos devem ser manualmente configurados.
 - Mensagens de keepalive periódicas são trocadas para verificar a conectividade TCP.
 - Roteamento Baseado em Política (policy-based routing)
 - O BGP opera trocando informações sobre as redes por mensagens de NLRI (Network Layer Reachability Information).
 - Herda a simplicidade do RIP, mas com operação manual, as rotas são escolhidas pelo menor número de saltos (tamanho do AS-PATH)
 - tempo de convergência depende de cada roteador na Internet
-
- Diferenças entre FIB e RIB
 - Forwarding Information Base (FIB)
 - Quando um router recebe 2 rotas diferentes para o mesmo prefixo os critérios de escolha de rota são analisados. A rota que foi considerada melhor de acordo com os critérios BGP fica na FIB
-
- Router information base (RIB)
 - As rotas que não foram consideradas melhores ficarão guardadas na RIB e utilizadas caso a rota principal fique indisponível
 - O BGP jamais anuncia uma prefixo que não esteja na FIB

BGP

MENSAGEM OPEN

- Mensagem open, é a primeira mensagem a ser trocada assim que estabelece uma sessão BGP. Se uma mensagem open é aceitável, uma mensagem keepalive é enviado como resposta. No cabeçalho da mensagem open possuem seguintes campos:

MENSAGEM UPDATE

- A mensagem update tem a finalidade de trocar informação de roteamento entre os peerings BGP e da mesma forma construir um gráfico em que os prefixos já foi passado.

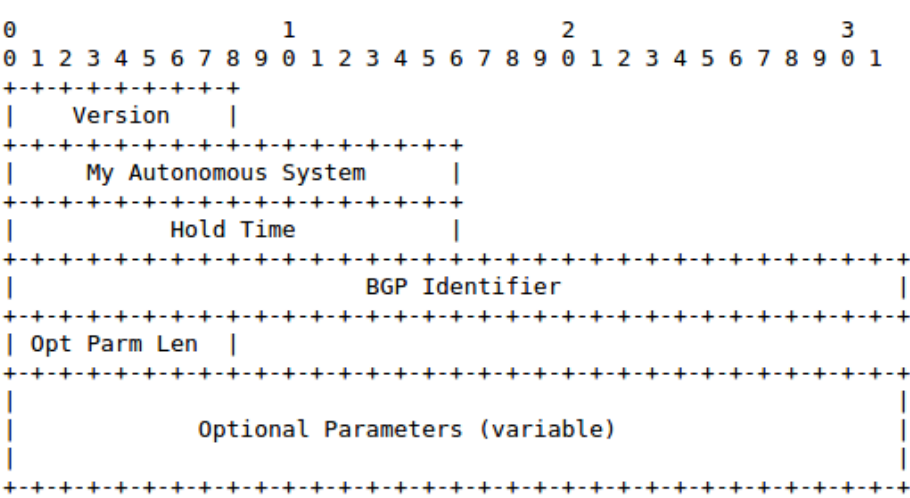
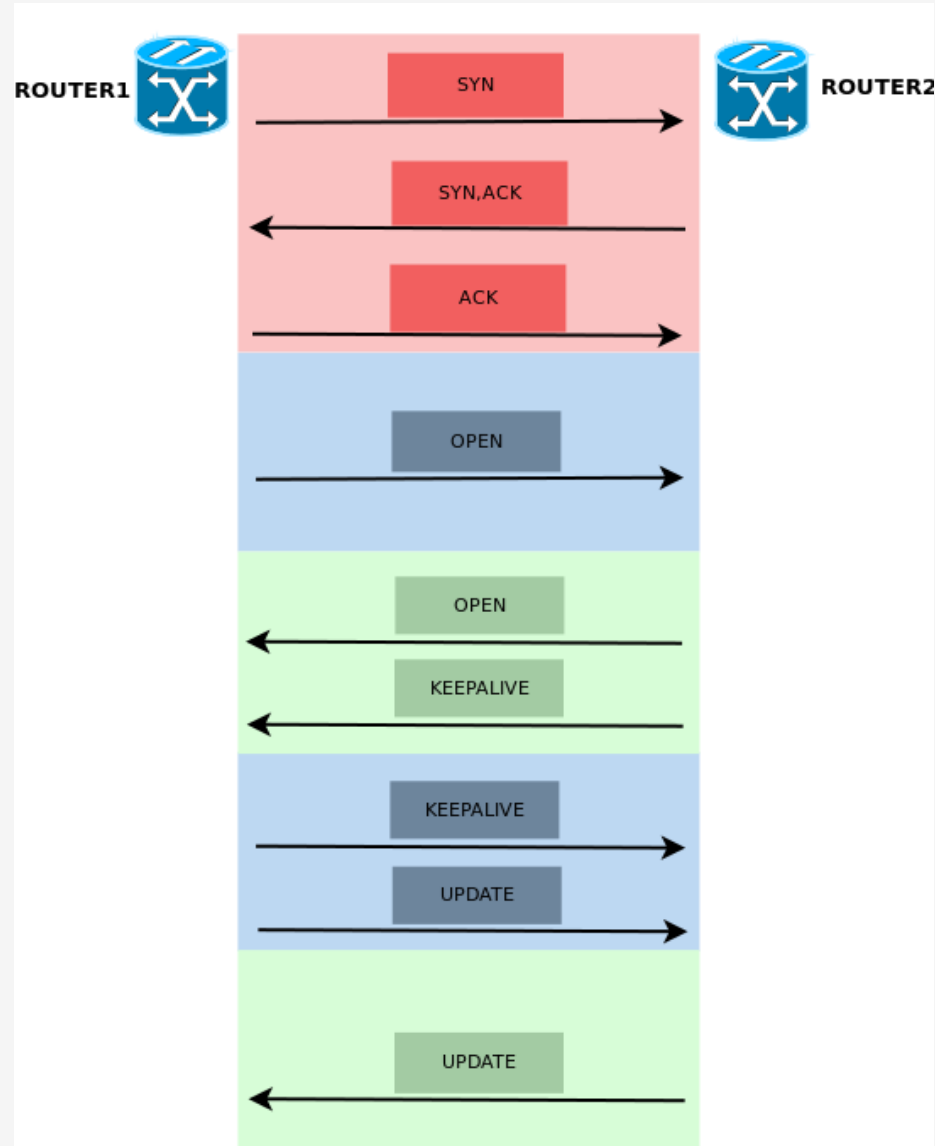
MENSAGEM NOTIFICATION

- É enviado quando há uma condição de erro ou um router BGP restabeleceu a sua sessão TCP e a sessão é encerrada imediatamente. Um exemplo do processo notification:

MENSAGEM KEEPALIVE

- De tempo em tempo cada router envia uma mensagem de keep-alive para que o vizinho saiba que há conectividade IP.
- Caso o keep-alive atrase, o router começa a contagem de hold-time e, se nesse período não for recebido nenhum keep-alive, a sessão é finalizada.
- Os valores padrão são 60 segundos de intervalo para keepalive e 180 segundo de tolerância (hold time)
- Atualmente, temos um downtime de 3 minutos sem internet, acrescido do tempo de propagação das convergências na Internet (~3-10 minutos);

BGP



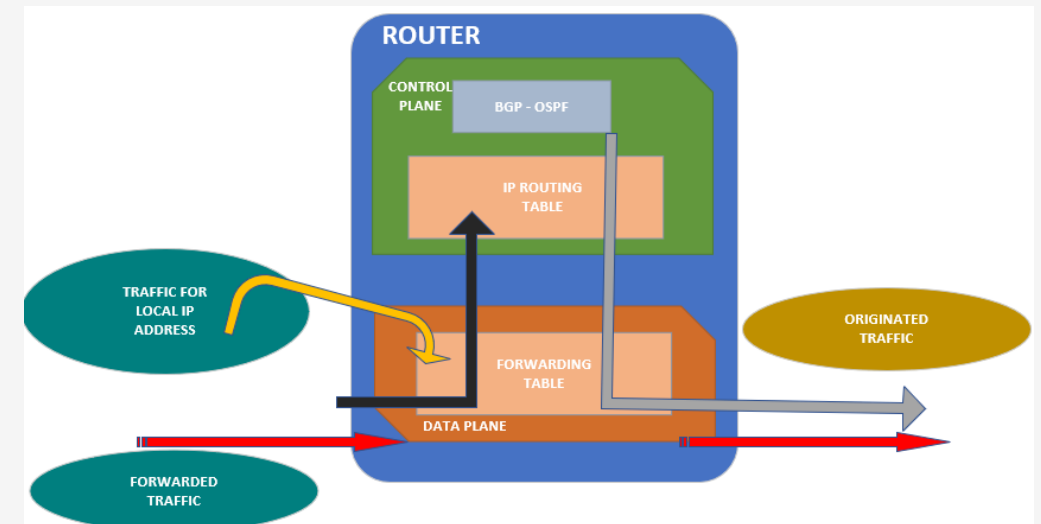
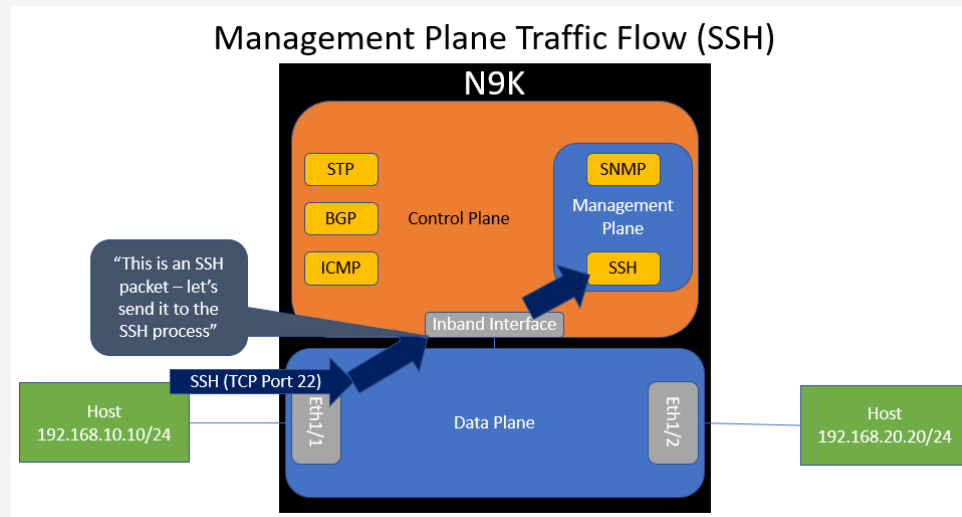
BGP

- TROCA DE ROTAS NLRI
- NLRI é uma lista de redes que são compostas por endereço IP e a máscara, ou prefixo em caso do IPv6
- O NLRI possui diversos atributos BGP ;
- Inicialmente a tabela full routing é trocada entre os peers;
- um speaker BGP só envia NLRI se estiver na IP routing table;
- Um vez alcançado o estágio Established, os peering BGP começam a trocar rotas (prefixos ou NLRI) entre si

- AFI - Address Family Information - : famílias de endereços IPv4, IPv6;
- SAFI: Subsequent Address Family Identifiers (IPv4 Unicast, IPv6 Unicast, IPv4 Multicast, VPNv4)
- Next-hop information - (deve ser do mesmo tipo que o address family, por exemplo AFI ipv6 o next-hop tem que ser um endereço IPv6)
- NLRI - Network Layer Reachability Information
- MP_UNREACH_NLRI - contém as informações de um prefixo que não pode mais ser alcançado é deve ser retirado da tabela BGP;

BGP

- ControlPlane e DataPlane - Soft-Router x Router.
- Em cenários via soft-router todos os recursos do plano de controle (tráfego de rede ex) são processados pelo mesmo recurso do Control Plane, ou seja, se a CPU alcançar 100% irá afetar diretamente o Fluxo da rede.
- Em roteamento via Router os recursos são segmentados em Control-Plane e Data-Plane, onde mesmo se a CPU do plano de controle chegar ao seu limite não irá afetar o fluxo da rede, pois o Plano de Dados foi programado em um Hardware totalmente separado.
- https://www.oreilly.com/library/view/cisco-networks_engineers/9781484208595/9781484208601_Ch11.xhtml..



BGP

- Todo AS de trânsito tem a obrigação de cuidar para que seus clientes, que também são Sistemas Autônomos, não façam anúncios errados para a Internet. Quando isso ocorre e o anúncio passa, acontece o que chamamos de Hijack (sequestro) de IP, derrubando todos os Sistemas Autônomos envolvidos no prefixo anunciado erradamente;
- Route maps servem para o BGP controlar e modificar atributos de roteamento e também definir as condições da propagação de rotas;
- Uma route-map é um conjunto de regras que verificam características dos atributos BGP;
- São utilizadas por todos os clientes bgp para tratar atributos durante os estágios de importação, exportação e transações entre RIBs;
- Route-maps são default-deny, se um atributo for processada por uma route-map e nenhuma regra permitir explicitamente o atributo, ocorre o descarta do fluxo:



BGP

- Route-MAP faz Verificações e Alterações (match):
 - Se o prefixos estão contido em uma prefix-list;
 - Verificar tags, communities, as-path, interface, preferência, mac, rpki;
 - Verificar se outra route-map autoriza a informação;
 - Se constarem múltiplas verificações de match, aplicado lógica AND;
 - Regra sem match submetem todas os atributos a ação da regra, e as próximas regras não serão processadas;
 - (set):
 - Alterar as-path, adicionar/apagar communities, alterar next-hop, métrica;
 - on-match goto - on-match next: jump para uma regra específica no índice da route-map;
- Os prefix-list servem para poder fazer match de redes, ou grupos de redes.
- Aplicar um deny em um prefixo dentro da prefix-list não deleta a informação, apenas sinaliza que o prefixo não participa da prefix-list;
- Aplicar um permit em um prefixo dentro da prefix-list não permite a informação, apenas sinaliza que o prefixo participa da prefix-list;
- A execução é feita do menor ao maior número de sequência, e também existe um deny implícito no final.

BGP

- No protocolo BGP em vez de se anunciar o destino e a distância para alcançar o destino (*distance-vector*) é anunciado uma descrição de todo o percurso (*path-vector*)
- É através dos atributos que se caracteriza o percurso e por isso estes possuem uma importância especial
- Existem quatro tipos de atributos:
 - **Well-Known Mandatory**: são os mais importantes, devem ser incluídos ao anunciar a rota e processados por cada router BGP que os receba
 - **Well-Known Discretionary**: se forem recebidos devem ser reconhecidos e processados pelo router BGP mas podem ou não ser incluídos no anúncio da rota com este tipo de atributo
 - **Optional Transitive**: Podem ou não ser reconhecidos por um router BGP mas devem ser incluídos no anúncio para outros routers BGP
 - **Optional Non-Transitive**: Atributos opcionais que podem ou não ser reconhecidos por um router BGP mas não são anunciados a outros routers BGP

BGP

- Alguns atributos são obrigatórios e automaticamente incluídos em mensagens de update, enquanto outros podem ser configurados manualmente.

Attribute	EBGP	IBGP	
AS_PATH	Well-known Mandatory	Well-known Mandatory	Automatically included in update message
NEXT_HOP	Well-known Mandatory	Well-known Mandatory	
ORIGIN	Well-known Mandatory	Well-known Mandatory	
LOCAL_PREF	Not allowed	Well-known Discretionary	Can be configured to help provide path control.
ATOMIC_AGGREGATE	Well-known Discretionary	Well-known Discretionary	
AGGREGATOR	Optional Transitive	Optional Transitive	
COMMUNITY	Optional Transitive	Optional Transitive	
MULTI_EXIT_DISC	Optional Nontransitive	Optional Nontransitive	

BGP

- 1. Rota com maior valor de WEIGHT (Cisco, FRR, RouterOS, OpenBGPD)
- 2. Rota com maior valor de LOCAL_PREF
- 3. Rota originada localmente (bgp network)
- 4. Rota com o menor AS_PATH.
- 5. Rota com menor tipo de origem. IGP (i) < EGP (e) < INCOMPLETE (?)
- 6. Rota com menor métrica multi-exit discriminator (MED).
- 7. Escolhe a rota recebida por (eBGP) em relação a uma (iBGP).
- 8. Rota com a menor métrica IGP para o nexthop BGP Ex: O next-hop aprendido via OSPF vai "vencer" um next-hop aprendido via IS-IS;
- 9. Rota externa mais antiga
- 10. Rota recebida de um router com menor Router ID
- 11. Rota com o menor tamanho de cluster list. (Ambientes com Route Reflector apenas)
- 12. Rota com o menor endereço de neighbor

BGP

- O BGP apenas analisa os atributos para "desempatar" duas ou mais rotas para o MESMO PREFIXO.
- Caso haja um prefixo mais específico esse sempre será preferido independentemente dos seus atributos "mais favoráveis".
- Sempre que houverem duas ou mais opções de rota para prefixos iguais (mesma máscara de sub-rede) e ambas tiverem sido recebidas via BGP, o protocolo vai escolher a melhor rota de acordo com a ordem:
- A análise segue para o próximo critério apenas quando há empate no critério atual, assim, o tamanho do AS PATH só será analisado caso o valor de weight e local preference sejam os mesmos para as duas rotas.

Padrão Huawei:

External Preference	Internal Preference
Direct 0	Direct 0
OSPF 10	OSPF 10
IS-IS 15	IS-IS Level-1 15
Static 60	IS-IS Level-2 18
RIP 100	Static 60
OSPF ASE 150	RIP 100
OSPF NSSA 150	OSPF ASE 150
IBGP 200	OSPF NSSA 150
EBGP 20	IBGP 200
	EBGP 20

BGP

- COMMUNITIES:
- Atributo opcional e intransitivo e são meios de rotular rotas com o objetivo de assegurar filtros consistentes e políticas de seleção de rotas.
- Qualquer roteador BGP pode rotular os updates de rotas que entram e ou que saem.
- Qualquer roteador BGP pode filtrar rotas que entram e ou que saem ou selecionar rotas preferenciais, baseadas em communities.
- Provedores de Serviço utilizam essas marcações para aplicar políticas de roteamento específicas em suas redes, por exemplo alterando o Local Preference, MED ou Weight.
- Standard: número de 32 bits divididos em 2 partes de 16 bits cada;
- Ex.: 0:65536, 65536:65536, 64888:50101
- Extended: possui um tipo (RT=router-target, SOO=site-of-origin) e um número de 48 bits (16+32 ou 32+16 ou 16+16+16);
- Ex.: RT 0:65536, RT 65536:65536, RT 64888:50101, RT 64888:4200111222;
- Large: community numérica de 96 bits (32+32+32);
- Ex.: 0:65536:999, 65536:65536:1, 64888:50101:666, 4200111222:4200111222:4200111222;
- Communities pré definidas:
- - No-Export : Não faz anúncio aos eBGP peers.
- - No-Advertise : Não faz anúncio para nenhum peer.
- - No-Export-Subconfed : Não faz anúncio para fora do sub-as, usado somente quando se trabalha com confederation.

BGP

- eBGP
- Formado quase sempre por peers diretamente conectados.
- Configuração Multihop é necessária quando os peers não são diretamente conectados.
- Adiciona o AS ao caminho anunciado.
- Por padrão o Next-Hop é mudado para o IP do próprio roteador.
- iBGP
- Para a propagação de rotas consistentes dentro do AS.
- todos os roteadores participantes devem ter sessão BGP entre si (Full Mesh).
- RR são usados para evitar o Full mesh.
- Roteadores não modificam os anúncios recebidos ao reenvia-lo.
- Devido ao princípio acima, o next hop recebido de um eBGP é repassado intacto, sendo que o Roteador de borda não se coloca como next hop (aplicar next-hot-self);
- O next-hop anunciado pelo eBGP é carregado pelo iBGP, que não o modifica

BGP

- eBGP – Peering entre roteadores de diferentes AS's.
- iBGP – Peering entre roteadores do mesmo AS.
- Por padrão no iBGP o "next-hop" não é modificado – Utilize algum IGP (OSPF, ISIS) para garantir a alcançabilidade IP dentro do AS.
- Rotas externas recebidas de um peer iBGP não são passadas para outros peers iBGP, É o que chamamos de "Split Horizon" do BGP;
- Para que isso ocorra é necessário o uso de full mesh ou router reflect.
- Boas práticas estabelecer sessão iBGP via loopback;
- Configuração Multihop é exigida se o peer não estiver diretamente conectado, por padrão o TTL tem valor 1, necessário mudar TTL (16+

BGP

- Possuir IRR - PeeringDB - RPKI
- Antes de estabelecer a sessão BGP com o Route-Server do IX ou Transito IP, o Sistema Autonomo deve possuir uma cadastro atualizado em um Internet Routing Registry (IRR) e informar o AS-MACRO (AS-SET),para que o IX faça a criação/atualização dos filtros de prefixos;
- Deve ainda possuir cadastro no peeringDB.
- Validação de anúncios de rotas via protocolo BGP.



TC IRR

Query:

aut-num:	AS3356
as-name:	Level3
descr:	Level 3 Communications
admin-c:	LV3-LEVEL3
tech-c:	LV3-LEVEL3
import:	from AS3 4.24.88.50 accept rs-peer-4-24-88-50
import:	from AS3 4.78.140.10 accept AS3
import:	from AS3 4.53.48.98 accept AS3^0-24



BGP

- CONFIGURAÇÃO DA POLITICA DE EXPORT
- Anúncios mais específicos para IXs Locais
- Anúncios menos específicos para Operadora de Trânsito
- Anúncios iguais aos de Operadora para IXs Internacionais

- CONFIGURAÇÃO DA POLITICA DE IMPORT
- Priorizar encaminhamento para IXs Locais LOCAL PREFERENCE 500
- Menor Prioridade para Operadora de Trânsito
- LOCAL PREFERENCE 200
- Prioridade iguais aos de Operadora para IXs Internacionais
- LOCAL PREFERENCE 200

- Evite que coisas estranhas sejam enviadas ao PTT:–BDPUs (spanning tree e similares)–CDP (Cisco discover protocol)–MNDP (Mikrotik discover protocol)–IGPs (OSPF, ISIS)–RA IPv6–Proxy ARP habilitado (Cisco principalmente)–Recursos de rádios em bridge no acesso.

BGP

- policies Inbound
 - Descarte (drop) de recebimento de rotas contendo prefixos bogons, martians e não alocados.
 - Descarte (drop) de recebimento de quaisquer prefixos contendo listas de irregulares do atributo AS_PATH.
 - Descarte (drop) de recebimento de prefixos IP dos próprios blocos do Provedor (route security).
 - Descarte (drop) de recebimento de prefixos de clientes do Provedor, assim como de clientes dos clientes do Provedor.
 - Recebimento (accept) de toda tabela rotas da Internet.
- policies Outbound
 - O não repasse (drop) de quaisquer prefixos que tenham sido aprendidos via outras sessões de trânsito.
 - O não repasse (drop) de quaisquer prefixos que tenham aprendidos via sessões de peering.
 - O não repasse (drop) de quaisquer prefixos que tenham sido eventualmente aprendidos por sessões com CDNs.
 - O repasse (anúncio) apenas de prefixos de clientes do Provedor, contemplando também, onde aplicável, os clientes dos clientes do Provedor.
 - Permitir acordos de black hole (por RTBH) com o Provedor e seus upstreams.
 - Permitir mitigação de blocos IPv4 e IPv6 de clientes um upstream de mitigação.



BGP

- BGPmon (<https://www.bgpmon.net/>)
- BGPlay (<https://bgplayjs.com/?section=bgplay>)
- RIPEstat (<https://stat.ripe.net/app/launchpad>)
- Looking Glass e BGP Toolkit (<https://bgp.he.net/>)
- PeeringDB (<https://www.peeringdb.com/>)
- IRR (<https://bgp.net.br/>)
- Team Cymru (<https://team-cymru.com/>)
- BGP Alerter (<https://github.com/nttgin/BGPalerter>)
- Spoofer Caida (<https://www.caida.org/projects/spoofer/>)
- https://www.cisco.com/c/pt_br/support/docs/ip/ip-version-6-ipv6/113489-ipv6-pfl-new-00.pdf
- <https://www.nic.br/media/docs/publicacoes/13/fasciculos-sobre-a-infraestrutura-da-internet-endere%C3%A7os-ip-e-asns-alocacao-para-provedores-internet.pdf>
- [https://semanacap.bcp.nic.br/files/apresentacao/arquivo/1208/Ferramentas%20monitoramento%20\(5\).pdf](https://semanacap.bcp.nic.br/files/apresentacao/arquivo/1208/Ferramentas%20monitoramento%20(5).pdf)

OBRIGADO

