

Communities BGP no IX.br

IX Fórum Fortaleza, 03/2026

Julio Sirota (jsirota@nic.br)

ix.br nic.br cgi.br

Anúncio BGP ou *route*:

<Prefixo> <atributos BGP>

<Prefixo> <ORIGIN, AS_PATH, NEXT_HOP, MED, LOCAL_PREF,
COMMUNITY,>

COMMUNITY

- Atributo transitivo
- Uma *route* pode ter múltiplas communities
- Tipos:
 - Well-known (RFC 1977) (0xFFFF0000 - 0xFFFFFFFF) (65535:0 a 65535:65535)
 - Privadas AS : <número>

COMMUNITY

- Tamanhos:
 - Standard (32-bit): (0-65535):(0-65535) ASN:<valor>
 - Extended (64-bit) RFC 4360: Type:Administrator:<valor>
 - Large (96-bit) RFC 8092: Global Administrator:LocalData1:LocalData2

- Para que são utilizadas:
 - Inserir informações adicionais no anúncio
 - Solicitar alguma ação por parte do *Peer* (AS que está recebendo o anúncio)
 - As ações são previamente divulgadas e conhecidas entre os AS

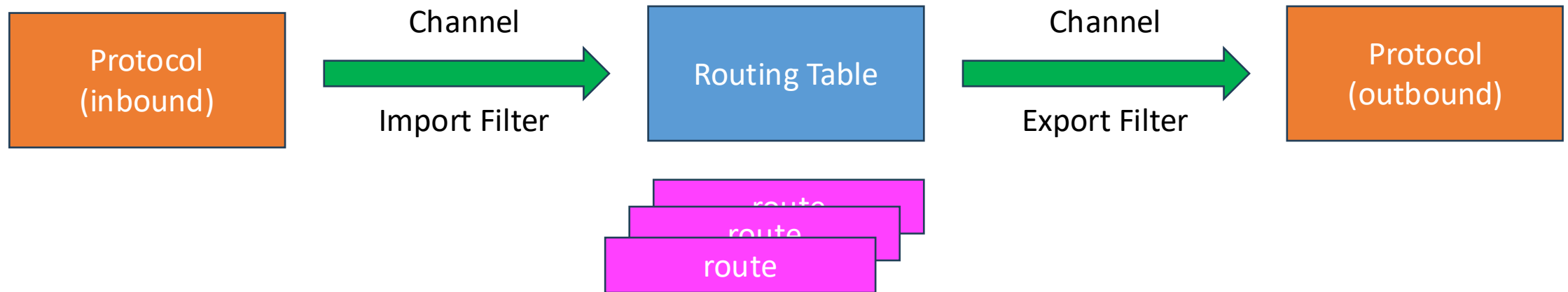
COMMUNITY

- Tamanhos:
 - Standard (32-bit): (0-65535):(0-65535) ASN:<valor>
 - ~~Extended (64-bit) RFC 4360: Type:Administrator:<valor>~~ A partir de 08/2026
 - Large (96-bit) RFC 8092: Global Administrator:LocalData1:LocalData2
- Para que são utilizadas:
 - Inserir informações adicionais no anúncio
 - Solicitar alguma ação por parte do *Peer* (AS que está recebendo o anúncio)
 - As ações são previamente divulgadas e conhecidas entre os AS

Route Server

- Fecha uma sessão BGP com cada participante do Acordo de Troca de Tráfego Multilateral
- Recebe os anúncios do participantes
- Devolve tabela com as rotas de todos os outros participantes

Arquitetura (usamos atualmente o software BIRD)



Entrada dos anúncios (*routes*)

- Inserção de Communities informativas
 - ASN do participante que fez o anúncio: 26162:peer-asn
Ex: 26162:**22548**
 - IXP de origem: 26162:65<id>
Ex: 26162:**65061** (IX.br de Brasília)

26162 é o ASN dos
RS de todas as
localidades do IX.br

IXP location	ID
Aracaju/SE	079
Belém/PA	092
Belo Horizonte/MG	031
Boa Vista/RR	095
Brasília/DF	061
Campina Grande/PB	183
Campinas/SP	019
Campo Grande/MS	067
Caruaru/PE	181
Cascavel/PR	145
Caxias do Sul/RS	054
Cuiabá/MT	065
Curitiba/PR	041

·
·
·

Entrada dos anúncios (*routes*)

- Inserção de Communities Informativas

- RTT até o participante (*next-hop*): 26162:6466n 26162:660:n

Peer RTT 0.001ms < 10ms	26162:64661	26162:660:1
Peer RTT 10ms < 50ms	26162:64662	26162:660:2
Peer RTT 50ms < 100ms	26162:64663	26162:660:3
Peer RTT 100ms < 150ms	26162:64664	26162:660:4
Peer RTT 150ms < 200ms	26162:64665	26162:660:5
Peer RTT 200ms < 250ms	26162:64666	26162:660:6
Peer RTT > 250ms	26162:64667	26162:660:7
Peer RTT = unknown	26162:64669	26162:660:9

- O RTT é medido em período de tráfego significativo durante 5 intervalos de medição, sendo considerada a média sobre 60% dos melhores resultados. A configuração dos RS é atualizada uma vez ao dia.
 - São feitas interações entre o servidor do RS até o *next-hop*

Entrada dos anúncios (*routes*)

- Inserção de Communities Informativas
 - Perda de pacotes até o participante (*next-hop*): 26162:6467n 26162:670:n

Peer LOSS 0%	26162:64671	26162:670:1
Peer LOSS 0.001% < 2%	26162:64672	26162:670:2
Peer LOSS 2% < 10%	26162:64673	26162:670:3
Peer LOSS > 10%	26162:64674	26162:670:4
Peer LOSS = unknown/100%	26162:64679	26162:670:9

- O *Packet LOSS* é medido 5 vezes por dia em período de tráfego significativo, sendo considerada a média de 60% dos melhores resultados.
- As medidas são feitas entre o servidor do RS até o *next-hop*

Entrada dos anúncios (*routes*)

- Inserção de Communities Informativas
 - GEO/RIR do AS que originou o anúncio: 26162:6468n 26162:680:n

Origin-asn of Afrinic	26162:64680	26162:680:0
Origin-asn of Apnic	26162:64681	26162:680:1
Origin-asn of Arin	26162:64682	26162:680:2
Origin-asn of Lacnic	26162:64683	26162:680:3
Origin-asn of Ripe	26162:64684	26162:680:4
Origin-asn of Brazil	26162:64685	26162:680:5

Entrada dos anúncios (*routes*)

- Inserção de Communities de Validação
 - Confirmação de anúncio de BH (Black Hole)

Confirmed BH	26162:666	26162:616:666
--------------	-----------	---------------

- Quando o *Route Server* recebe e aceita o anúncio ele irá:
 - Alterar o *next-hop* do anúncio apontando para um MAC e IP pré-determinados no IXP, onde descartaremos o fluxo.
 - Uma community informando que é um BH confirmado pelos Route Servers é inserida no anúncio

Entrada dos anúncios (*routes*)

- Inserção de Communities de Validação
 - Validação de origem para ASNs brasileiros, usando os dados do registro.br

registro.br invalid	26162:65110	26162:100:0
registro.br valid	26162:65111	26162:100:1
registro.br unknown	26162:65112	26162:100:2

- Utilizada como fonte de dados o arquivo publicado diariamente pelo registro.br, disponível via FTP. Um ASN será considerado brasileiro, desde que listado no arquivo, assim como os blocos IP válidos para um ASN, desde que listados no arquivo e associados ao ASN.
- Diariamente os RS recebem estas informações através de sua configuração, tendo como validar os anúncios recebidos. Não é um processo dinâmico.

Entrada dos anúncios (*routes*)

- Inserção de Communities de Validação
 - Validação de origem usando os AS-SETs informados no PeeringDB

IRR - prefix present in AS-SET	26162:65121	26162:200:1
IRR - prefix not present in AS-SET	26162:65122	26162:200:2

- A resolução dos AS-SET e a coleta do que é anunciado por cada AS participante (peer-asn) é feita com dados dos IRR (Internet Routing Registry). São consultados: *afrinic, apnic, arin, jpirr, lacnic, level3, nttcom, radb, ripe e tc*.
- Diariamente os RS recebem estas informações através de sua configuração, tendo como validar os anúncios recebidos. Não é um processo dinâmico.

Entrada dos anúncios (*routes*)

- Inserção de Communities de Validação
 - Validação de origem usando a base de dados de RPKI (*Resource Public Key Infrastructure*)

RPKI invalid	26162:65130	26162:300:0
RPKI valid	26162:65131	26162:300:1
RPKI unknown	26162:65132	26162:300:2

- A cada horas as bases de dados de RPKI disponíveis são coletadas para consultas locais.
- Este é um processo dinâmico.

Entrada dos anúncios (*routes*)

- Inserção de Communities de Validação
 - Validação de boas práticas

Tamanho de prefixo inválido	26162:65190	26162:190:0
Prefixo Bogon	26162:65191	26162:190:1
ASNs Bogon	26162:65192	26162:190:2
Transit free (Tier-1)	26162:65193	26162:190:3
Never via RS (registrado no PeeringDB)	26162:65194	26162:190:4
Prefixo de IXPs (registrados no PeeringDB)	26162:65195	26162:190:5

- Anúncios com estas Communities não serão exportados pelos RS, mas aparecem no LG

Entrada dos anúncios (*routes*)

- Proteção dos IXPs e dos participantes

Prefixos de redes ATM/MLPA do IX.br
Re-enviando de volta rotas esportadas
Rotas não originadas por BGP
Tamanho do AS-PATH < 1 ou > 64 ASNs
Anúncio contém ASN do IX.br
Primeiro ASN do AS-PATH não é o peer-asn
next_hop != peer-ip

- Anúncios com os problemas reportados acima são descartados na entrada do RS, sem marcação de erro por Community e nem presença em consulta pelo LG.

Saída dos anúncios (*export routes*)

- Communities são enviadas pelo participante (*peer*) para controlar o Route Server e fazer Engenharia de Tráfego
- São tratadas no filtro de saída (*export*) do RS
- Quem deve ou não receber os anúncios

Não anunciar para um peer-asn	65000:peer-asn	(ro rt):65000:peer-asn	65000:0:peer-asn
Anunciar apenas para o peer-asn	65001:peer-asn	(ro rt):65001:peer-asn	65001:0:peer-asn

- Adição de prepend para aumentar o tamanho do AS-PATH

Inserir um prepend no anúncio para o peer-asn	64601:peer-asn	(ro rt):64601:peer-asn	64601:0:peer-asn
Inserir dois prepend no anúncio para o peer-asn	64602:peer-asn	(ro rt):64602:peer-asn	64602:0:peer-asn
Inserir três prepend no anúncio para o peer-asn	64603:peer-asn	(ro rt):64603:peer-asn	64603:0:peer-asn

Saída dos anúncios (*export routes*)

- Controle de anúncios por região geográfica

Não anuncia para peer-asn do Afrinic	65002:0	65002:0:0
Não anuncia para peer-asn do Apnic	65002:1	65002:0:1
Não anuncia para peer-asn do Arin	65002:2	65002:0:2
Não anuncia para peer-asn do Lacnic	65002:3	65002:0:3
Não anuncia para peer-asn do Ripe	65002:4	65002:0:4
Não anuncia para peer-asn do Brasil	65002:5	65002:0:5

Anuncia apenas para peer-asn do Afrinic	65003:0	65003:0:0
Anuncia apenas para peer-asn do Apnic	65003:1	65003:0:1
Anuncia apenas para peer-asn do Arin	65003:2	65003:0:2
Anuncia apenas para peer-asn do Lacnic	65003:3	65003:0:3
Anuncia apenas para peer-asn do Ripe	65003:4	65003:0:4
Anuncia apenas para peer-asn do Brasil	65003:5	65003:0:5

Saída dos anúncios (*export routes*)

- Controle de anúncios pelo valor do *RTT* até o *peer*

Não anuncia quanto o peer rtt > 10ms	65010:10	65010:0:10
Não anuncia quanto o peer rtt > 50ms	65010:50	65010:0:50
Não anuncia quanto o peer rtt > 100ms	65010:100	65010:0:100
Não anuncia quanto o peer rtt > 150ms	65010:150	65010:0:150
Não anuncia quanto o peer rtt > 200ms	65010:200	65010:0:200
Não anuncia quanto o peer rtt > 250ms	65010:250	65010:0:250
Não anuncia quanto o peer rtt = unknown	65010:999	65010:0:999

Adiciona um prepend quando o peer rtt > 10ms	64611:10	64611:0:10
Adiciona um prepend quando o peer rtt > 50ms	64611:50	64611:0:50
Adiciona um prepend quando o peer rtt > 100ms	64611:100	64611:0:100
Adiciona um prepend quando o peer rtt > 150ms	64611:150	64611:0:150
Adiciona um prepend quando o peer rtt > 200ms	64611:200	64611:0:200
Adiciona um prepend quando o peer rtt > 250ms	64611:250	64611:0:250
Adiciona um prepend quando o peer rtt = unknown	64611:999	64611:0:999

Adiciona dois prepend quando o peer rtt > 10ms	64612:10	64612:0:10
Adiciona dois prepend quando o peer rtt > 50ms	64612:50	64612:0:50
Adiciona dois prepend quando o peer rtt > 100ms	64612:100	64612:0:100
Adiciona dois prepend quando o peer rtt > 150ms	64612:150	64612:0:150
Adiciona dois prepend quando o peer rtt > 200ms	64612:200	64612:0:200
Adiciona dois prepend quando o peer rtt > 250ms	64612:250	64612:0:250
Adiciona dois prepend quando o peer rtt = unknown	64612:999	64612:0:999

Adiciona três prepend quando o peer rtt > 10ms	64613:10	64613:0:10
Adiciona três prepend quando o peer rtt > 50ms	64613:50	64613:0:50
Adiciona três prepend quando o peer rtt > 100ms	64613:100	64613:0:100
Adiciona três prepend quando o peer rtt > 150ms	64613:150	64613:0:150
Adiciona três prepend quando o peer rtt > 200ms	64613:200	64613:0:200
Adiciona três prepend quando o peer rtt > 250ms	64613:250	64613:0:250
Adiciona três prepend quando o peer rtt = unknown	64613:999	64613:0:999

Saída dos anúncios (*export routes*)

- Controle de anúncios pelo valor do *LOSS* até o *peer*

Não anuncia quando o peer loss > 2%	65011:2	65011:0:2
Não anuncia quando o peer loss > 10%	65011:10	65011:0:10
Não anuncia quando o peer loss = unknown	65011:999	65011:0:999

Adiciona um prepend quando o peer loss > 2%	64621:2	64621:0:2
Adiciona um prepend quando o peer loss > 10%	64621:10	64621:0:10
Adiciona um prepend quando o peer loss = unknown	64621:999	64621:0:999

Adiciona dois prepend quando o peer loss > 2%	64622:2	64622:0:2
Adiciona dois prepend quando o peer loss > 10%	64622:10	64622:0:10
Adiciona dois prepend quando o peer loss = unknown	64622:999	64622:0:999

Adiciona três prepend quando o peer loss > 2%	64623:2	64623:0:2
Adiciona três prepend quando o peer loss > 10%	64623:10	64623:0:10
Adiciona três prepend quando o peer loss = unknown	64623:999	64623:0:999

Saída dos anúncios (*export routes*)

- Solicita a desconexão da sessão BGP (*Graceful Shutdown*)

graceful shutdown	65535:0	-
-------------------	---------	---

- Anúncio de *Black Hole* (BH)

Anúncio de Black Hole (BH)	65535:666	65535:616:666
----------------------------	-----------	---------------

- Não anuncia para um PTT do IX.br. Usa a tabela de localidade já apresentada.

Não anuncia para um PTT do IX.br	65004:65XXX	65004:0:65XXX
----------------------------------	-------------	---------------

Thanks !

www.ix.br

© jsirota@nic.br

Março, 2026

nic.br egi.br

www.nic.br | www.cgi.br