IX.br BGP Communities Policy

Version: 3.0
Last modified date: May 13, 2019

1. **Objectives**

This document presents the usage policy of the *communities* in the IX.br IXs, which aims to:

1. adapt the functioning of the IX.br route servers to good practices documented in RFCs 7947 [1] and 7948 [2];

2. decrease the operational work in IX.br and concomitantly provide agility and independence to the participants in the configuration of filters, implementing for this purpose the treatment of *communities* for filters in the *route servers*, so that the functionality provided by such *communities* can replace the manual configuration of specific filters that is currently made;

3. to enable IX.br participants to negotiate and use *communities* among themselves to facilitate the implementation of filters, making the route servers transparent to them;

4. allow the BGP MED attribute to be used for traffic engineering, making the route servers transparent to it;

5. to allow participants to filter incoming prefixes more easily by maintaining the existing tagging of the source AS today, but also by adding source validation markings based on the registry.br and also using the RIRs and RPKI;

6. enable participants to more easily implement certain routing policies, offering *communities* the addition of *prepends* to specific destinations.

2. **The routes servers**

The Internet Exchanges, such as the IX.br localities, provide the infrastructure to enable the exchange of IP traffic between its participants, usually using a shared network layer 2, such as Ethernet. The Border Gateway Protocol (BGP) is typically used to facilitate the exchange of information about the addresses present on each network in that infrastructure.

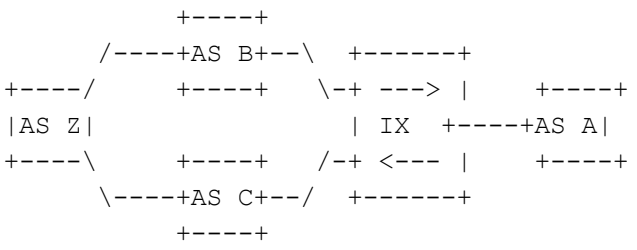
Historically, IX participants closed bilateral BGP sessions with each other to update their route tables. However, for many-party IXs, this approach generates enormous operational work.

In the IXs of IX.br, the route servers allow the technically facilitated deployment of the Multilateral Traffic Exchange Agreement (MPA), where each participant agrees to exchange traffic with the others. Its function is that of a *broker*, an intermediary, facilitator or concentrator: each participant of the multilateral agreement closes a session with the route server. The prefixes received from each will be passed on to the others, with the exception of the participant who originated the advertisement, but without adding the ASN of the route server, and without modifying the attribute *next_hop*.

Currently, IX.br offers the filter service on the route servers and uses as a protection mechanism the limitation of the number of ads per participant.

In addition to the changes in the Community use policy discussed in this document, there are two changes proposed here in the operation of the route servers:

1. **Path hiding mitigation:**



Consider the situation shown in the figure where Autonomous Systems A, B and C are participants in a IX of IX.br and AS Z is a transit client of both AS B and AS C.

The route server of IX.br receives the prefixes of AS Z both via AS B and via AS C. As it works the same way as a normal BGP router, chooses the best route and only passes it to the AS A. AS A receives the AS Z prefixes only via AS B, or only via AS C, not both.

If AS A and AS B have filters implemented in IX so that they do not receive routes from each other, and if the route server chooses AS B as the best path to AS Z, AS A does not receive prefixes from AS Z .

This situation is different than it would if the aS a closed bilateral BGP sessions with both the aS B, as with the aS C. in this case receive the aS Z prefixes for both ways.

The situation exemplified above is what we call *path hiding*, or concealment of paths. The IXs of IX.br, currently with the exception of IX.br of São Paulo/SP and Curitiba/PR, currently operate in this situation.

There are several ways to mitigate this, described in RFC 7947 [1], which allow all different possible paths to reach participants, better mimicking the same behavior that would occur if there were bilateral sessions among all IX participants.

In the route servers of IX.br, two approaches are implemented: multiple RIBs, in which there is a separate RIB for each client (peer) and ADD-PATH (Additional Paths), in which the route server negotiates with the client (peer) sending a negotiated amount of additional routes to the best route.

2. Transparency to the MED attribute (MULTI_EXIT_DISC)

The MED attribute is an optional non-transitive attribute that is used to interconnect different ASs with multiple points of entry or exit to differentiate them. As the route server aims to be an 'invisible' element in the network, from the BGP point of view, RFC 7947 [1] specifies that it must be propagated.

Currently in IX.br not all route servers are transparent to the MED attribute. **This document specifies the implementation of transparency to the MED attribute on all route servers of IXs of IX.br.**

3. *Communities*

The *communities* were added to BGPv4 protocol with the aim of creating a mechanism for grouping prefixes, so that the routing decision may also take place based on the identity of a group, providing greater flexibility in making heuristics decision for the formation of the BGP route table. That is, *communities* allow you to mark groups of prefixes, and decisions can be made based on that markup.

The attribute community was defined by means of RFC1997 [3], which describes its inclusion in the BGPv4 protocol. This in turn was initially defined by RFC 1271 [4] and is currently described in RFC 4271 [5].

Examples of use include control over prefix propagation, filtering, and mitigation of DDoS attacks.

The communities are widely disseminated and used on the Internet. Currently, in IX.br, they are treated in the route servers according to the following policies:

1. All the communities received from one participant are ignored and are not propagated to the others.

2. **Source ASN:** Each prefix received and propagated by the route server is marked with a community to identify the originating originator in the 26162:ASN format. If the participant's ASN is 32 bits, a static conversion is done, using documentation and private ASNs mapped (static and manually) in a table. The main objective of this marking is the use of these communities in the configuration of manual filters among the participants internally, in the route server itself of IX.br.

3. **Filters**: will be defined Communities (in case 32-bit ASNs will be used *extended communities* with the same syntax) [6] specific to filters, which will be processed by the route servers, in order to avoid propagation of the marked prefixes to certain participants, or to indicate that they should only be propagated to specific participants. These communities will not be propagated to other participants.

4. ***Prepends*:** will be defined Communities (in case 32-bitbe used *extended communities* with the same syntax) [6] specific for applying prepends on prefixes sent to a particular destination. These *communities* will not be propagated to other participants.

5. **ASN of origin:** Each prefix received and propagated by the route server is marked with two communities, one community to identify the participant that originated it, in the format 26162:ASN and another to identify the location where the prefix was received, in the format 26162:DDD. If the participant's ASN is 32 bits,will be used extended community.

6. **Transparency to *communities*:** Incoming communities that are not destined to the route server will be passed on to other clients (peers) without changes, thus reinforcing the transparency feature of the route servers.

7. **Source validation:** Prefixes received and propagated will be marked with communities that will identify whether the origin of the prefix was successfully validated, unsuccessful or not validated, and by which means validation was performed. Validations will be used initially with queries based on Registro.br, RIRs and via RPKI.

4. **Details on policies for communities in IX.br**

The following describes in more detail the policies for communities:

Policy 1: community for target ASN filter

The community to specify the filter will have the following format, with possible use of *extended communities* (in the two-octet format AS specific *extended community*, as defined in RFC 4360) [5] to specify 32-bit ASNs. You can specify multiple different ASs so that they do not receive a particular ad, marking the same with the *communities* appropriate.

65000: <ASN> - NOT export the prefix to the specified AS

It is important to note that this community will work as a one-way filter in the route server. Therefore the AS interested in the filter should also discard the AS routes to be filtered in its input policy.

In addition to this community, the following is also accepted and processed an auxiliary community, whose purpose is to mark a prefix that will be exported only to the specified AS. It is intended to allow participants to exchange traffic with or implement specific policies for only one or a subset of the ASs present in the multilateral agreement. You can specify multiple different ASs, so that they only receive a certain ad, marking the same with the communities appropriate.

65001: <ASN> - exports the prefix ONLY to the specified AS

The <ASN> specifies the target participant. In the interpretation of these two *communities* (65000: <ASN>, 65001: <ASN>) by the route server, the second is more priority.

Examples of use:

1. The route server receives from AS 64496 the prefix 203.0.113.0/24 **without *communities***, or with any *community* other than 65001: * or 65000: *:
   Action: **the prefix 203.0.113.0/24 will be exported to all ASs, with the exception of AS 64496 itself, which originated it.**

2. The route server receives from the AS 64496 the prefix 203.0.113.0/24 **marked with *community* 65000: 65551:**
   Action: The prefix 203.0.113.0/24 will be exported to all ASs, with the exception of AS 64496 itself, which originated it, and AS 65551, which was specified in *community* 65000:65551.

3. The route server receives from AS 64496 the prefix 203.0.113.0/24 **marked with *community* 65000:65551, and with community 65000:64500:**
   Action: the prefix 203.0.113.0/24 will be exported to all ASs, except for AS 64496, and AS 65551 and 64500, which were specified in *communities* 65000:65551 and 65000:64500.

4. The route server receives from the AS 64496 the prefix 203.0.113.0/24 **marked with community 65000:64496:**

Action: The prefix 203.0.113.0/24 will be exported to all ASs, with the exception of the AS 64496 itself, which originated it. A community 65000:64496 in this case has no practical effect. The same behavior will occur if the community is 65000:26162, 65000:0, or if the ASN specified at 65000:<ASN> is not in the Multilateral Peering Agreement.

5. The route server receives from the AS 64496 the prefix 203.0.113.0/24 **marked with community 65001:65551:**
   Action: The prefix 203.0.113.0/24 will be exported only for AS 65551, specified in *community* 65001:65551.

6. The route server receives from the AS 64496 the prefix 203.0.113.0/24 **marked with communities 65001:65551 and 65001:64500:**
   Action: The prefix 203.0.113.0/24 will be exported only for the AS 65551 and 64500, specified in communities 65001:65551 and 65001:64500.

7. The route server receives from AS 64496 the prefix 203.0.113.0/24 **marked with *community* 65001:64496:**
   Action: The prefix 203.0.113.0/24 will not be exported to any AS. A community 65001:64496 specific should be exported only to AS 64496, but since this is the AS itself that originated the advertisement, the route server will not export it. The same behavior will occur if the community is 65001:26162, 65001:0, or if the ASN specified in 65001:<ASN> is not in the Multilateral Peering Agreement.

8. The route server receives the prefix 203.0.113.0/24 from AS 64496 **with communities 65000:65551 and 65001:64500:**
   Action: The prefix 203.0.113.0/24 is only exported to AS 64500, specified in community 65001:64500 . A community 65000:65551 in this case has no practical effect. The community 65001:64500 has preference on filter implementation and determines that the prefix will not be exported to any other AS, but the 64500. Add a community 65000:65551 specifying the prefix should not be exported to AS 65551 is redundant and not has effect.

9. The route server receives from the AS 64496 the prefix 203.0.113.0/24 **marked with communities 65000:65551 and 65001:65551:**
   Action: The prefix 203.0.113.0/24 will be exported only for AS 65551. Note that using both communities simultaneously to a same target AS does not make sense, since they specify opposite actions. In this case, community 65001:65551 has priority and is the action specified by it to be performed.

Since these communities are specified for actions on the route servers, **they will not be exported.**

Policy 2: *community* for adding *prepends* to a target ASN

In some cases it may be desirable for a participant to add prepends in prefixes sent to a particular target AS as a traffic engineering tool.

For example, a participant A present in two IXs of IX.br, say São Paulo and Rio de Janeiro, exchanges traffic with a participant B, also present in these two IXs, through ATM. Participant A may prefer that for some prefixes the IX of São Paulo has preference for participant B in

relation to that of Rio de Janeiro. Thus, in the IX of Rio de Janeiro, in sending the prefixes to participant B, participant A would include prepends.

On many occasions the use of unbundling as a traffic engineering tool is preferable to using *prepends*, but in some cases it is not even possible. For example, when the participant has only one /24 IPv4 or /48 IPv6 block. In these cases, the use of prepends, with parsimony, may be a viable alternative.

The community to specify the *prepends* (maximum 3) has the following format, being possible to use *extended communities* (in the two-octet format AS specific *extended community*, as defined in RFC 4360) [6] to specify 32-bit ASNs. It will be possible to specify several different ASs, so that they receive a certain ad with *prepends*, marking the same with the *communities* appropriate.

64601: <ASN> - adds 1 *prepend* on sending the prefix to the specified AS
64602: <ASN> - adds 2 *prepend* on sending the prefix to the specified AS
64603: <ASN> - adds 3 *prepend* on sending the prefix to the AS specified

Examples of use:

1. The route server receives from AS 64496 the prefix 203.0.113.0/24 **without *communities***, or with any community other than 64601:*, 64602:* or 64603:*:
Action: **the prefix 203.0.113.0/24 will be exported normally, without the addition of *prepends*, without AS PATH change.**

2. The route server receives from AS 64496 the prefix 203.0.113.0/24 **marked with community 64603:65551, whose original AS PATH is**
203.0.113.0/24 64496 i:
Action: the prefix 203.0.113.0/24 will be exported:
(a) for AS 65551, with 3 *prepends* in AS PATH:
203.0.113.0/24 64496 64496 64496 64496 i
(b) for all other ASs, with the exception of AS 64496 itself, and AS 65551, specified in the *community*, the prefix will be exported without addition of *prepends* or change of AS PATH:
203.0.113.0/24 64496 i

Policy 3: Source identification

Currently, each prefix received and propagated by the route server is marked with a community to identify the participant that originated it, in the format:

26162:<ASN>

The prefix will also be marked with a second *community*, to identify the IX (location) of IX.br, in the format:

26162:65XXX

where XXX represents the IX of origin of the prefix, according to the following table:

IX (locality)	XXX

Aracaju, SE	079
Belém, PA	091
Belo Horizonte, MG	031
Brasília, DF	061
Campina Grande, PB	083
Campinas, SP	019
Cuiabá, MT	065
Caxias do Sul, RS	054
Curitiba, PR	041
Florianópolis, SC	048
Fortaleza, CE	085
Foz do Iguaçu, PR	045
Goiânia, GO	062
João Pessoa, PB	083
Lajeado, RS	051
Londrina, PR	043
Maceió, AL	082
Manaus, AM	092
Maringá, PR	044
Natal, RN	084
Porto Alegre, RS	051
Recife, PE	081
Rio de Janeiro , RJ	021
Salvador, BA	071
Santa Maria, RS	055
SJ dos Campos, SP	012
SJ Rio Preto, SP	017
São Luís, MA	098
São Paulo, SP	011
Teresina, PI	086
Vitória, ES	027

Each prefix received and propagated by the route serveris marked with a community to identify the participant who originated it, with support to *extended communities* for 32-bit ASNs. In addition a second community to identify the IX (locality) of origin of the advertisement will also be used, according to the table presented above.

Policy 4: Transparency to *communities*

All communities or *extended communities* received, except those defined in Policy 1 (65000:<ASN> and 65001:<ASN>) are ignored by routing servers, and propagated to the other participants.

This policy has the function of allowing the IX.br participants to negotiate and use *communities* among themselves, to facilitate the implementation of filters.

An ongoing discussion in the community of IX.br participants is that it could offer mechanisms to mitigate DDoS attacks, such as lack hole filters. While this is technically possible, there is an inherent risk that can have administrative and legal implications, for example if a failure event occurs by directing an unwanted route to the black hole. This would make the protection tool the very cause of a DoS.

An interesting alternative is that some participants, such as large hosting or transit providers can offer the black hole mechanism in their own networks, through specific communities, and disseminated to other participants. Maybe communities standardized. The black hole, in this way, would be closer to the origin of the attack than if implemented in the IX itself. Blocked traffic would not even reach the IX network.

In addition, this reinforces the transparency feature of the route servers, which is a working premise. This is in line with what has been discussed in the technical community [1] in the process of developing a BCP on the subject.

Policy 5: Prefix Validation

BGP is an unsafe protocol. Configuration errors, by entering the wrong numbers, or an intentionally malicious configuration can result in the capture of prefixes from one network, on the other. That is, an AS may, intentionally or by accident, advertise routes from another, diverting traffic, which may allow the obtaining of sensitive information, or cause unavailability of the victim's service.

The RPKI structure that is slowly being deployed on the Internet is intended to be a solution to partially mitigate this problem. In RPKI the entity that owns a certain prefix can specify, in a database that is offered by the RIR, which Autonomous Systems can advertise the prefix in question. BGP routers have mechanisms to automatically and securely measure ads on this basis by validating their source.

The risk inherent in *peering can be* reduced by reducing the number of *peers*. For example, instead of participating in the multilateral agreement, one can be made peering only with few ASs, among which there is a good trust. However, it is a dubious practice, since even if it can reduce the risk related to bad intentions, it is practically impossible to avoid occasional errors. Moreover, it is a practice that goes against the very design of *Internet Exchanges*, which aim to stimulate the exchange of traffic.

The prefix validation service will be offered as an additional tool to reduce the risk of third-party prefixes being captured. Consists of consistency analysis of each ad, based on the comparison with external bases, and marking them with communities appropriate.

Based on the analysis of these communities a participant may know that a prefix whose advertisement was originated by a particular AS was actually assigned to it by the Registro.br or an RIR. In addition to Registro.br database, will be the validation on the basis of the IRRs, public services for the registration of routing policies, and the database of RPKI.

For each new route injected into the route server, an agent searches the various bases for information on the link between those prefixes and the ASN that formed the advertisement. Three states are possible, for each of the bases:

invalid prefix: the prefix is in the base and does not correspond to the ASN;

valid prefix: the prefix is in the base and corresponds to the ASN;
Unknown prefix: The prefix does not appear in the base, or the base is not available for query.

The route server will mark each prefix with *communities* corresponding to the validations available, according to the values below:

Registro.br database (one of the three communities will mark the prefix):

- **26162:65110 - invalid in the Registro.br**
- **26162:65111 - valid in the Registro.br**
- **26162:65112 - unknown in Registro.br**

basis of IRRs (one of the three communities will mark the prefix):

- **26162:65120 - invalid in IRR**
- **26162:65121 - valid in IRR**
- **26162:65122 - unknown in the IRR**

RPKI (one of the three communities will mark the prefix):

- **26162:65130 - invalid in RPKI**
- **26162:65131 - valid in RPKI**
- **26162:65132 - unknown in RPKI**

In addition to the validation of the source presented above, other validations will be made in order to increase security. If the ASN that originated the announcement has registered the identification of its AS-SET in the PeeringDB, it will be validated and marked with one of the following *communities*:

AS-SET (one of the two communities will mark the prefix):

- **26162:65150 – AS-SET Invalid**
- **26162:65151 – Valid AS-SET**

Prefixes will be marked as Stub BR when they are prefixes from the Registro.br and whose associated ASN is directly connected to IX.br:

Stub BR (one of two *communities* will mark the prefix):

- **26162:65180 - is a Stub BR**
- **26162:65181 - is not a Stub BR**

The classification of the ASN as Stub BR will have implications in the way the Route Servers will handle the ads received as well as it can be used for use of special services, like the use of *community* for Blackhole.

Ads from an ASN Stub BR should have only ASN in the AS-PATH itself.

Prefixes size invalid (</8 or> /24 for IPv4, </3 or> /48 for IPv6) or BOGON (using invalid Internet address block):

- **26162:65190 - prefix size is invalid**
- **26162:65191 - prefix BOGON**

Ads containing one or more ASNs in the AS-PATH BOGONs:

- **26162:65192 - ASN BOGON**

Ads containing one or more ASNs classified as being of free transit (Tier-1):

- **26162:65193 - Transit Free (Tier-1)**

More information on the source validation process can be found in the document "Safer Internet Program - Actions in IX.br" [7].

5. **Implications for security and stability of the platform**

Regarding security, **changes in the behavior of route servers, regarding mitigation of path concealment, and transparency to the MED**, as well as the **policies regarding communities 1 (filters) , 2 (*prepends*) and 3 (origin identification)** do not cause changes in the current situation.

Politics 4 (transparency) can have positive effects the medium term, if the community implement the treatment of communities to black hole, received via IX.br.

There is, however, the risk that some participants have now configured the treatment for certain communities for advertisements received via IX.br, as well known communities and black hole. Currently there is no risk because the route servers prevent the propagation of any community. The transparency proposed in policy 3 would open the possibility of an attack perpetuated by an AS X, involving the announcement of AS A prefixes, marked with the *community* of *black hole* AS B. This would effectively disrupt communication between ASs A and B, consisting of a DOS attack. To mitigate this risk, the participant who implements the *black hole* can make use of the *communities* source validation, applying it only to trusted ads.

Politics 5 (source validation) depend for their success in their effective use by the community of participants. The validation information of the implicit prefixes in these *communities* will be useful only if used to implement filters, and have the potential to increase the security and stability of IX.br. However, if participants do not implement the filters, the effect will be null. And if the filters are poorly implemented by the participants, the effect may be the opposite, with security risks and platform stability. Errors in the validation process can also pose a risk if the filters are implemented in a very restrictive way, accepting only valid prefixes, for example.

7. **References**

[1] https://tools.ietf.org/html/rfc7947
[2] https://tools.ietf.org/html/rfc7948
[3] https://tools.ietf.org/html/rfc1997
[4] https://tools.ietf.org/html/rfc1271
[5] https://tools.ietf.org/html/rfc4271
[6] https://tools.ietf.org/html/rfc4360
[7] http://www.ix.br/doc/acoes-seguranca-ix-br-20180927.pdf