ÉCOLE POLYTECHNIQUE DE LOUVAIN

LINFO2142 - COMPUTER NETWORKS : CONFIGURATION AND MANAGEMENT

BGP Communities

Rapport - 2021-2022

Group 03 :
Maximilien LA BARRE : 2045-17-00
Alison LEURQUIN : 2781-17-00
Iris DENAYER : 5581-17-00

Professeur : Olivier BONAVENTURE

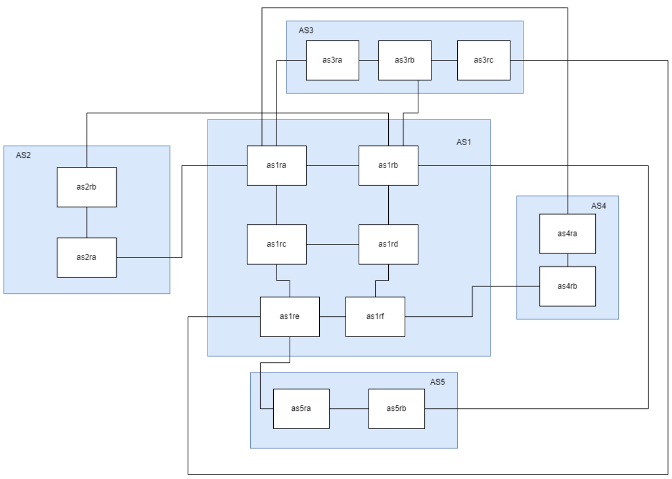
GitHub : https://github.com/Alrq/LINFO2142-Projet

ÉCOLE POLYTECHNIQUE DE LOUVAIN

I. Introduction

As part of the computer networking course, we were asked to analyze how BGP communities are used and then demonstrate and test on the "IPMininet" tool. For that, we will base ourselves on two different implementations that are FRRouting and OpenBGPD. This report will be divided into several parts. To start, we will explain the topology chosen to test the different BGP implementations and how the BGP communities work. Then we will perform an analysis on how the communities are actually used. We will finish with the demonstration and test part on "IPMininet".

II. Description of the topology



For the topology, we were inspired by a minimized real topology of OVH. It is composed of a main AS (the middle one) with 6 routers and a host, then there are several AS that gravitate around it. AS2, AS4 and AS5 have two routers each and one host, however AS3 has three routers and one host. The topology includes eBGP links so that the different hosts can communicate even if they are in different AS. (The hosts have not been put on the diagram for clarity but each router in as1 has a host and then one router per as (asxra) has a host)

Legend :

- Blue square : Represents the different AS
- White square : Represents the different routers, the name is each time composed of the name of the AS of which they are part of, the letter "r" to show that it is a router and a single letter to differentiate them.
- Links : Represent the physical link between the routers and hosts

In this part of the report we see that the topology does not include local-pref, med, etc. This is normal, these will be added during the test part to see how FRRouting and OpenBGPD works in case of failure or if we want to prioritize a link over another.

III. BGP communities

The BGP protocol defines attributes to influence the routing of packets in the network. These attributes include localpref, AS-Path, MED and others. In this section, we focus on community attributes. BGP communities are attributes that allow multiple routes to be grouped into a single entity where the same routing decisions can be applied.

A classical application of the Community attribute is for multi-homing purposes (Multihoming is the practice of connecting a computer network to multiple Internet service providers to improve the reliability of the Internet connection). However, the Community attribute has been used for other purposes like to tag the routes received from a specific peer or at a specific location and to influence the redistribution of specific routes in order to perform some kind of inter-domain traffic engineering. (1)

In this project, we will define different communities, each with a specific strategy. For this we will divide the communities into three distinct categories. (2)

A. Inbound communities (description of a route's entry point)

Inbound communities refer to communities added or used when a route is received by a router on an eBGP session, it can be divided into two categories :

- Route-tagging : The community value can be used by an Autonomous System to indicate the location where the route was received from an external peer. It can be very usefull to know certain information like the type of peers (customer, traffic provider, etc), the geographic location, etc. (3)
  - Type of peer : The AS defines a type of peer (customer, provider, peering partner) and will tag each received route with a community indicating the type from which the route was received.
  - Geographic location : The AS's sometimes need to know the geographic location where a route was received. Often, an AS will rely on a unstructured pairs of key-value to associate a community value with a location (i.e : "13129 :3010" - "frankfurt"). But some AS's use

different way to structure their lists, like one where the value used to tag a received route is based on the phone country code. Another way to do it, is to use a structure baded on the ISO3166 code for country (3561:SRCCC where S is the source (peer or customer), R is the regional code et CCC is the ISO3166 countrycode.
  - Interconnection point : The AS's sometimes need to remember the interconnection point. (13129:2110 — DE-CIX, where DE-CIX is the interconnection point)
  - AS : The AS's sometimes need to remember the AS from which each route were learned, it's useful in confederation (we will see what is it later) or to simplify the configuration of routers, but it becomes redundant with the AS path attribute.
- LocalPref : Can be used to allow customers to modify the localpref of their routes in the network, allowing them to define priority and backup routes (MED can be used for that purpose too but it is placed behind in the BGP priority list so the use of local pref is better, but med can be usefull in certain situations).

B. *Outbound communities (traffic engineering purpose)*

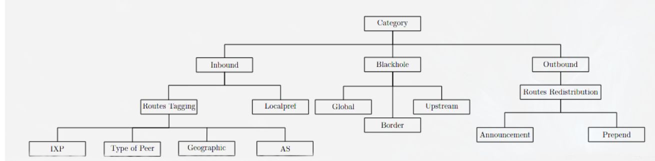
The outbound communities affect thus the redistribution of routes. It can also be divided in two categories :

- AS Prepending : This consists in adding AS numbers, usually its own, to the AS-Path. It will thus allow to prioritize some routes compared to others since the routes with a prepending will have a longer AS-path. It is widely used by ISP (provider of internet access) to allow peers (customers) to request the utilization of AS-Path prepending when announcing some routes to specified external peers, at specified interconnection points or in specified regions. Usually, an AS that provides such communities relies (like the route tagging) on an unstructured set of communities.
- Announcement : The community is attached to a route to indicate whether the route should or should not be announced to a specified peer (customer, traffic provider, etc) or at a specified interconnection point. We will see in the section "Demonstration and test part on IPMininet" different type of announcement community attribute.

C. *Blackhole*

- Security : It allow AS to ask ISP to block certain packets, there is different type of blackholing (global, border and upstream). It can be useful to prevent some DoS attack for example

The BGP communities allow to easily see the relationship between the AS but especially to simplify the traffic engineering since it allows to have more control for the provider (informational tags) as well as for the Customer (Action tags).
So communities can be used for many purpose that we will see later on. But for example, you will be able to set local pref between two AS (can be AS based or community based). The AS based version is done by the provider and is less efficient because of the coordination and configuration required each time the network changes. Setting local pref from communities is therefore much more efficient as it gives more control to the customer. Customers can use these communities to gain more flexibility in the control of its own routing policy.
To have a more schematic look at the different communities and their proportions, this diagram comes from a document made by Benoit Donnet and Olivier Bonaventure. (2)



Category	Category	Type	Proportion (%)	Total (%)
Blackhole			0.22%	0.22%
Inbound	Localpref		1.62%	26.25%
	Routes tagging	Interconnection Point	1.82%	
		Type of Peer	6.33%	
		Geographic Location	9.11%	
		AS	6.08%	
	Unknown		1.26%	
Outbound	Routes redistribution	Announcement	25.9%	73.52%
		AS_PATH prepending	47.43%	
	Unknown		0.16%	

IV. BGP IMPLEMENTATION

To simulate a network, we use IPMininet. We will then use two different implementations of the Border Gateway Protocol on our network in order to compare them.
For that, we will use the daemon OpenBGPD and the network routing software FRRouting. Note : OpenBGPD and frrouting are open source Disclaimer : As the amount of information is extremely limited (open source project), the documentation of the different protocols is scarce.

A. OpenBGPD

OpenBGPD is an implementation of the Border Gateway Protocol. It allows ordinary machines to be used as routers exchanging routes with other systems speaking the BGP protocol. (4)
OpenBGPD is made out of 3 processes : 1 parent process an 2 child processes (SE and RDE).
The Route Decision Engine process (child process, RDE) maintains routing information like the prefix table as well as the AS path. It also calculates the best path per prefix.(5) The parent process starts the Session Engine and the Route Decision Engine. This process doe not drop privileges as it needs to rn on a root port and needs to update the routing table.

B. FRRouting

Unlike OpenBGPD, FRRouting is not a daemon. It is an Internet routing protocol suite, that provides, among other, an implementation of BGP. It is a fork from Quagga and runs on Unix-like platforms (Ubuntu in our case). (6) FRR was created to provide layer-3 connectivity throughout a data center, from the spine and leaf switches all the way down to hosts, virtual machines, and containers. It's designed to streamline the routing protocol stack. Businesses can use FRR for connecting hosts, virtual machines, and containers to the network, advertising network service endpoints, network switching and routing, and internet access and peering routers. (7) FRR provides :

- Simplified, modern data center design
- Subnet freedom and mobility
- Stateless load balancing with anycast
- Enhanced redundancy and flexibility

C. Decision process

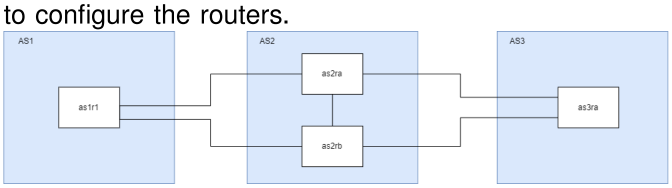
In this sub-section, we will look at the differences in the decision process for these two BGP implementations. All informations of this section has been taken from the official websites of FRrouting and openBGPD. (We will not put the whole decision process of the two implementation here, they are available following this link) (8) (9)
What we can notice directly is that these decisions processes are similar in some points, but differ strongly in others. Indeed, the first point of FRrouting "Weight check" (Prefer higher local weight routes to lower routes) is not ever there in the openBGPD decision process (in the openBGPD official website they called local weight the iBGP).
Then we can see that openBGPD starts its decision process by seeing if there are errors or loops on the network, these paths are then not eligible. Same thing for paths with unreachable nexthop. This check is not done in the decision process of frrouting. We also notice that openBGPD does not do a local route check (Prefer local routes (statics, aggregates, redistributed) to received routes.) but it can do a check with the age of the route (the oldest path is selected) which is not present with FRrouting.
So we can see that even though they are both BGP implementations, they have similarities and differences in the decision process.

V. Redistribution communities

In this section of the report, we will discuss how to improve the structure of communities. Indeed, it suffers from several important drawbacks that limit its widespread utilization. To begin with, an AS can only define 65536 distinct community values which forces AS's to define its own community values in an unstructured manner. In addition, each value must be manually encoded in the BGP configuration of the routers. Then, the AS will advertise these neighbours with its own community semantics but unfortunately there is no standard way to advertise these community values.
This implies that AS's have to manually insert directives into the configuration of BGP routers which is clearly not optimised (e.g. Human Error). A second disadvantage is transitivity. Indeed, It implies that once a community value has been attached to a route, this community is distributed throughout the global Internet.
All this together creates what is called pollution since a large number of community values appear in BGP routing tables.
To solve these problems, an idea (thought by B. QUOITIN, S. UHLIG and O. BONAVENTURE) has been proposed. It consists in attaching a new type of extended community attribute. It will give a more structured and larger space than the community attribute since each extended community value is encoded in an 8 bytes field. This type of community will be non transitive and will allow to encode a set of redistribution actions that are applicable to a set of BGP speakers. Non-transitivity removes the risk of pollution of routing tables by communities, while standardised encoding simplifies the configuration of BGP routers and thus reduces the risk of errors. Moreover, it will allow operators to provide services that go beyond the simple client-provider and peer-to-peer policies currently found on the Internet. But it will also help reduce the impact of DDoS attacks (10)

VI. Demonstration and test part on IPMininet

If you want to see some code, you can check the github link on the report homepage. Throughout this project, we were able to use ipmininet to simulate a network with multiple AS, some of which have multiple routers and of course links between these routers and AS's. To test the BGP communities we decided to take our topology presented at the beginning and to simplify it as much as possible to see how these communities work and in what way / why they are used. Then we tried to use IPmininet to configure the communities but we got a lot of problems due to the poor documentation. After some research and permanence we decided to change the method and configure the routers directly from "xterm", which was quite laborious as each time we restarted the topology, we had to redo the configuration. We did our best to create a small script which, when the topology was launched, allowed us to configure the routers.



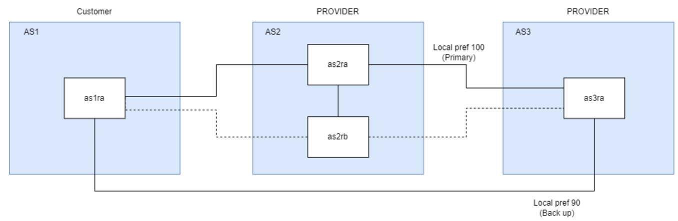
This topology will therefore allow us to test the different uses of this attribute. For starters, it should be noted that the community attribute is a 32-bit value structured as follows : ASN:value.
The first 16 bits generally represent the AS that originated the community. Then the last 16 bits (those behind the " :") are used to represent the value of the community, so its semantics is defined by the AS that will use it.
For example, this table summarizes what we could see when implementing community attributes :

Community	action
701 :120	Set local-pref to 120
701 :400	Do not announce to ...
701 :10x	Prepend X times the AS-Path of ...
65001 :6020	No-export to ...

TABLE I – BGP communities

Warning : when using BGP communities, remember not to exceed 65535 as 16-bit value !(11)
Now, let us illustrate the usage of BGP communities with our topology to give you a better understanding of the subject. (The configuration of communities with IPmininet is quite complicated to understand and to use, for the sake of understanding and readability all the examples presented and all the explanations will be made through explanatory diagrams using topology.)

A. Local-pref



In this configuration, let's say that the dotted link are down for ease of visibility of the schema and that AS1 is a customer of AS2 and AS3 which are therefore providers. If we focus on AS3, thanks to the community, we can allow the network to use one path rather than another. If the community value is set to 90, for example "5:90" for the direct route between AS1 and AS3, and the route that crosses AS2 is tagged with a community value of "5:100" (knowing that 100 is the default local pref, it was not necessary to indicate it, but it was interesting for the example). Having a higher local pref, the link passing through AS2 will be the link used, while the direct link will be a backup link in case the link to AS2 is cut. When we looked at the traceroutes from as1ra to as3ra, we could see that it was taking the route through as2. But when we cut the link (command :"link as1ra as2ra down") the traceroute changed to the backup link.

1) contradiction between community: This subsection will talk about the contradiction that can happen between communities, at the start this question was asked by our teacher asking us to find a way to avoid these contradictions. After a lot of research on the subject, we didn't find much information either in the FRrouting or OpenBGPD implementations. So we thought about a possible solution ourselves.
We therefore hypothesize that the last word should go to the provider. For example if a customer uses 2 communities that contradict each other (like one setting the local pref at 70 and the other one setting it at 100), it is the provider that will choose according to him the behavior to have. So basically it's up to the one who defines the community to choose how it fits in the network.

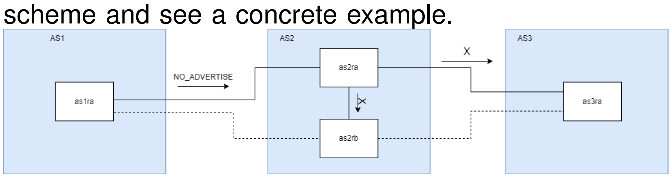
B. Well-known communities

In this section, we will discuss other uses of the community attributes. The standard defines several well-known communities that have local significance and must be supported by any community-aware

BGP router. So we'll see how this works and what it consists of.

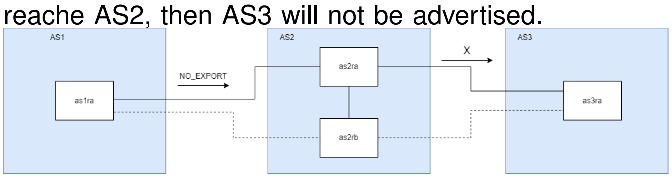
1) NO_ADVERTISE: It indicates that the route must not be advertised to other BGP pairs. so when we add the prefix "no_advertise", the router that receives the packet will store and use the prefix in its table, but will not pass the information on to its neighbours.

To show this, let's go back to our basic scheme and see a concrete example.



How do we test this on mininet ? We can use the ping command to try and see if a router is reachable. In our case, if we ping between router as2rb and as1ra, we will see that as1r1 knows the routes to as2rb but the latter cannot respond to the former's ping because it doesn't know the route. To be sure that everything is working properly, we can simply look at the routing tables of as2rb and see that the route to router as1r1 is not present.

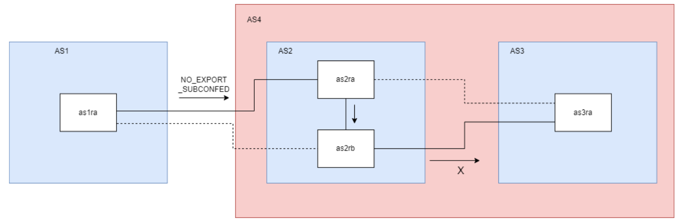
2) NO_EXPORT: It indicates that the route must not be advertised to eBGP peers. So it means that the route will no be exported to external BGP connection. Let us take our example : if the advertisements reache AS2, then AS3 will not be advertised.



In mininet, it can be checked the same way as the NO_ADVERTISE : as3ra will not know the route of as1ra.

3) NO_EXPORT_SUBCONFED (local-as): It indicate that the route must not be advertised to eBgp peers including pairs within a BGP confederation. A BGP confederation is an alternative method to route reflectors, used to reduce the number of BGP peerings within a single AS. Furthermore, it is a single AS that has been subdivided into a number of internal sub-AS's, yet still, is advertised as a single AS to external peers. Within the confederation, each sub-AS (inter-confederation) is assigned a private AS number (64512-65535). Though eBGP is used to peer each of the sub-AS's, the rules of iBGP still apply to each inter-confederation i.e attributes are not altered and an iBGP full-mesh is still required. (12)

We didn't really make a confederation in our topology, but if we had done it, it might have looked like this :



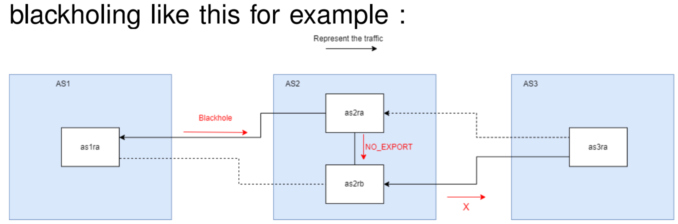
The difference with the NO_EXPORT is that in the NO_EXPORT, it is possible to have confederations. But in these confederations, the advertisement works between the sub AS while in this one, it blocks on the eBGP of the subAS. As for the others, this can be tested in the same way by using pings or by looking at the routing tables.

C. BlackHole

Finally, another use of BGP communities is for blackholing purposes. Blackholing is very useful, it allows an AS to announce a prefix to another (by attaching a blackhole community) and that all traffic towards this prefix has to be dropped.

As explained in the previous sections, blackholing allows among other things to reduce the impact of DDoS attacks. In general, the value 666 is used for these communities (note that this is not true in all cases, some do not use this convention).

In our topology, we can implement blackholing like this for example :



So to explain what is going on, AS1 is receiving traffic from AS2 from AS3 but AS1 wants to protect itself from AS3 as it is suspicious. To do this, AS1 will attach a blackhole community to the announcement it sends to AS2. From there, AS2 understands that it has to drop all packets received from AS3. To make sure it doesn't receive any (eBGP or iBGP) it will send a NO_EXPORT community to the announcement and all this locally (in its AS). Thanks to this, all packets sent by AS3 to AS1 will be discarded and never received by AS1. As for the others, this can be tested by pinging or checking the routing tables of the topology.

D. Conclusion

In this paper we focused on the BGP communities attribute. To understand how it works but also

the observations we have made with the tests on IPmininet. Thanks to this paper, we could see that communities were more and more used but also their strong importance in the networks. We discovered that there are different types or categories of BGP communities (inbound, outbound and blackhole) and that each of them has different goals whether it is for security, traffic engineering or description of the route's entry point.
We could also learn to better use the mininet tool even if it posed a lot of problems to us. Due to close to no experience with the tool and the almost non-existent documentation, we lost a lot of time on small errors. We think it's a pity that a more thorough documentation is not proposed because the tool remains very interesting to understand how the networks work.

RÉFÉRENCES

[1] B. Quoitin and O.Bonaventure, "A survey of the utilization of the BGP community attribute"
[2] Benoit Donnet and O.Bonaventure,"On BGP Communities"
[3] A survey of the utilization of the BGP community attribute of Bruno Quoitin and Olivier Bonaventure, https ://datatracker.ietf.org/doc/html/draft-quoitin-bgp-comm-survey-00
[4] OpenBGPD, https ://www.openbgpd.org/
[5] http://www.openbsd.org/papers/bgpd_sucon04/index.html
[6] FRRouting, https ://frrouting.org/
[7] Nvideo frr, https ://developer.nvidia.com/networking/ethernet-switches/frrouting
[8] openBGPD decision, https ://-docs.frrouting.org/en/latest/bgp.html ?highlight=decision%20proces
[9] frr decision, https ://man.openbsd.org/bgpd.8
[10] Redistribution communities, https ://totem.info.ucl.ac.be/publications/papers-elec-versions/48-quoitin.pdf
[11] BGP communities,R. A. Steenbergen and T. Scholl ,https ://fr.slideshare.net/RichardSteenbergen/bgp-communities-a-guide-for-service-provider-networks
[12] BGP confederation, https ://www.fir3net.com/Networking/Protocols/what-is-a-bgp-confederation.html
[13] Github mininet, https ://github.com/cnp3/ipmininet/tree/master/ipmininet/examples
[14] Networking Note, http ://blog.computer-networking.info/bgp-mininet/
[15] Thomas WIRTGEN O.Bonaventure,The Case for Border Gateway Protocol (BGP)
[16] Routing Policies, https ://www.gin.ntt.net/support-center/policies-procedures/routing/
[17] Using BGP communities, https ://nsrc.org/workshops/2017/apricot2017/bgp/bgp/preso/09-BGP-Communities.pdf
[18] BGP Policy, https ://s-print.net/index.php ?p=policy_bgp
[19] OpenBSD slide, https ://www.openbsd.org/papers/bsdcan2018-mininet.pdf