

RFC 8092 : BGP Large Communities Attribute

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 19 février 2017

Date de publication du RFC : Février 2017

<https://www.bortzmeyer.org/8092.html>

Ce RFC normalise un nouvel attribut des annonces BGP, « *Large Communities* ». Les « communautés » BGP sont des courtes données collées aux annonces BGP et qui permettent d'indiquer certaines caractéristiques des routes. Les décisions des routeurs peuvent utiliser ces caractéristiques. Mais les communautés originales étaient trop courtes (seulement quatre octets) : le nouvel attribut permet des communautés de douze octets.

Les « communautés » sont définies dans le RFC 1997¹. On les apprend via les documentations des opérateurs ou des points d'échange. Par exemple, celle du point d'échange irlandais <<https://www.inex.ie/technical/route-servers/>> (section « *Community based prefix filtering* »). Un attribut COMMUNITY dans une annonce BGP peut comporter plusieurs communautés. Traditionnellement, les quatre octets des communautés initiales sont utilisées pour représenter le numéro d'AS dans les deux premiers octets (ainsi, une communauté est mondialement unique, par ce système d'allocation à deux niveaux), et des données spécifiques à l'AS dans les deux suivants. Évidemment, avec les numéros d'AS de quatre octets du RFC 6793, ça ne marchait plus. D'où cet attribut LARGE_COMMUNITY, désormais stocké dans le registre IANA <<https://www.iana.org/assignments/bgp-parameters/bgp-parameters.xml#bgp-parameters-2>> sous le numéro ("*type code*") 32. (Il y a bien eu une autre tentative d'augmenter la taille des communautés, dans le RFC 4360, mais pas suffisamment pour que les AS à quatre octets puissent être utilisés partout.) Comme pour les « petites » communautés, ces grandes communautés dans une annonce forment un ensemble (donc, non ordonné) : plusieurs routeurs auront pu ajouter une communauté à cet ensemble.

Les communautés sont importantes car elles sont utilisées dans la politique de routage. BGP ne cherche pas à trouver le meilleur chemin : il fait du routage politique, où les décisions sont prises en fonction de choix faits par les opérateurs (privilégier tel ou tel lien pour le trafic entrant, par exemple).

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc1997.txt>

Les informations contenues dans une annonce BGP (section 4.3 du RFC 4271) habituelle ne sont pas toujours suffisantes, et c'est pour cela que les communautés ont été introduites par le RFC 1997, pour ajouter des informations utiles, comme l'endroit où telle route a été apprise. L'attribut `COMMUNITY` (numéro 8) est transitif (section 5 du RFC 4271), ce qui veut dire qu'après réception d'une annonce, il est transmis aux autres routeurs (d'où l'importance de marquer la communauté avec un numéro d'AS, pour que les communautés soient uniques au niveau mondial, sans qu'il existe un registre central des communautés).

Le nouvel attribut `LARGE_COMMUNITY` (numéro 32) est également optionnel et transitif (section 2 de notre RFC). Il se compose d'un ensemble de grandes communautés, chacune étant stockée sur douze octets. L'idée est qu'on utilise les quatre premiers octets pour identifier l'AS (ce qui va bien avec les grands AS du RFC 6793), ce qui va garantir l'unicité des communautés. Le nombre de communautés dans un attribut `LARGE_COMMUNITY` est donné par le champ Longueur de l'attribut, les attributs BGP étant encodés en TLV (cf. RFC 4271, section 4.3).

En cas d'agrégation de routes (section 3 du RFC), il est recommandé d'utiliser comme communautés l'union des ensembles de communautés des différentes annonces.

Et comment on va représenter ces grandes communautés sous forme texte? (Sur le câble, entre les deux routeurs, c'est du binaire, en gros boutien, cf. RFC 4271, section 4.) On note trois groupes de quatre octets, séparés par un deux-points, par exemple `2914:65400:38016` (section 4 de notre RFC), où le premier champ est presque toujours l'AS. (On trouve plein d'exemples dans le RFC 8195.)

Comme toutes les grandes communautés font exactement douze octets, si le champ Longueur de l'attribut n'est pas un multiple de douze, l'attribut est invalide, et le routeur qui reçoit cette annonce doit la gérer comme étant un retrait de la route (RFC 7606).

Un point de sécurité important en section 6 du RFC; en gros, les grandes communautés ont quasiment les mêmes propriétés de sécurité que les anciennes petites communautés. Notamment, elles ne sont pas protégées contre une manipulation en transit : tout AS dans le chemin peut ajouter des communautés (même « mensongères », c'est-à-dire indiquant un autre AS que le sien) ou retirer des communautés existantes. La section 11 du RFC 7454 donne quelques détails à ce sujet. Ce problème n'est pas spécifique aux communautés, c'est un problème général de BGP. L'Internet n'a pas de chef et il est donc difficile de concevoir un mécanisme permettant de garantir l'authenticité des annonces.

Il existe déjà de nombreuses mises en œuvre de BGP qui gèrent ces grandes communautés. Par exemple IOS XR, ExaBGP <<https://github.com/Exa-Networks/exabgp/wiki>>, BIRD, OpenBGPD, GoBGP <<http://osrg.github.io/gobgp/>>, Quagga, bgpdump depuis la version 1.5 <<https://bitbucket.org/ripenc/bgpdump/commits/fad519504b30fala34c673ce7d1e03532234be05>>, pmacct <<http://www.pmacct.net/>>... Une liste plus complète figure sur le Wiki <<https://trac.tools.ietf.org/wg/idr/trac/wiki/draft-ietf-idr-large-community%20implementations>>. Mais il y a aussi le site Web du projet <<http://largebgpcommunities.net/>>, où vous trouverez plein de choses. Si vous avez accès à un routeur BGP, ou à un "looking glass" qui affiche les grandes communautés (c'est le cas de celui du Ring de NLnog <<http://lg.ring.nlnog.net/>>), les deux préfixes `2001:67c:208c::/48` et `192.147.168.0/24` ont une grande communauté (`15562:1:1`). Si vous essayez sur un routeur qui a un vieux logiciel, ne comprenant pas ces grandes communautés, vous verrez sans doute quelque chose du genre « "unknown attribute" ». Ici sur IOS à Route Views <<http://www.routeviews.org/>> :

```
% telnet route-views.oregon-ix.net
...
Username: rviews
route-views> show ip bgp 192.147.168.0
BGP routing table entry for 192.147.168.0/24, version 37389686
```

```

Paths: (41 available, best #21, table default)
  Not advertised to any peer
  Refresh Epoch 1
  3333 1273 2914 15562
    193.0.0.56 from 193.0.0.56 (193.0.0.56)
      Origin IGP, localpref 100, valid, external
      Community: 1273:22000 2914:410 2914:1206 2914:2203 2914:3200
      unknown transitive attribute: flag 0xE0 type 0x20 length 0xC
        value 0000 3CCA 0000 0001 0000 0001
  ...

```

Ici sur un vieux IOS-XR (le test a été fait à l'époque où l'attribut avait le numéro 30 et pas 32, d'où le 0x1e) :

```

RP/0/RSP0/CPU0:Router# show bgp ipv6 unicast 2001:67c:208c::/48 unknown-attributes
  BGP routing table entry for 2001:67c:208c::/48
  Community: 2914:370 2914:1206 2914:2203 2914:3200
  Unknown attributes have size 15
  Raw value:
  e0 1e 0c 00 00 3c ca 00 00 00 01 00 00 00 01

```

Et ici sur JunOS :

```

user at JunOS-re6> show route 2001:67c:208c::/48 detail
  2001:67c:208c::/48 (1 entry, 1 announced)
  AS path: 15562 I
  Unrecognized Attributes: 15 bytes
  Attr flags e0 code 1e: 00 00 3c ca 00 00 00 01 00 00 00 01

```

Notez que certaines configurations (parfois activées par défaut) du routeur peuvent supprimer l'attribut « grandes communautés ». Pour empêcher cela, il faut, sur JunOS :

```

[edit protocols bgp]
user at junos# delete drop-path-attributes 32

```

Et sur IOS-XR :

```

configure
router bgp YourASN
  attribute-filter group ReallyBadIdea ! avoid creating bogons
  no attribute 32
!
!

```

Trois lectures pour finir :

- Un bon article sur l'acceptation <<https://labs.ripe.net/Members/emileaben/large-bgp-communities-des-grandes-communautés-dans-la-nature-et>>
- Un outil (fondé sur Docker) pour expérimenter avec ces grandes communautés <https://labs.ripe.net/Members/pier_carlo_chioldi/the-large-bgp-communities-playground>.
- Le RFC 8195 qui donne de nombreux exemples d'usage de ces grandes communautés. <https://www.bortzmeier.org/8092.html>