RFC 7422 : Deterministic Address Mapping to Reduce Logging in Carrier Grade NAT Deployments

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 26 décembre 2014

Date de publication du RFC : Décembre 2014

https://www.bortzmeyer.org/7422.html

Un nouveau RFC pour Big Brother : quand un FAI veut savoir quel abonné utilisait telle adresse IP à tel moment, c'est simple, non, il lui suffit de regarder les journaux du système d'allocation d'adresses ? En fait, non, c'est simple en théorie, mais cela a beaucoup été compliqué par le développement du **partage d'adresses**, dont l'exemple le plus connu est le CGN. Si la police ou les ayants-droit disent à un FAI « on voudrait savoir **qui** utilisait 192.0.2.199 le mercredi 24 décembre à 08 :20 », le FAI va se rendre compte que des dizaines d'abonnés utilisaient cette adresse IP à ce moment. Trouver l'abonné exact va nécessiter d'examiner d'énormes journaux. Ce RFC propose donc une méthode pour réduire la taille de ces journaux, en attribuant les ports du CGN de manière partiellement déterministe. L'examen des journaux pourra donc être plus efficace et on trouvera plus vite le méchant abonné qui a osé commettre des délits impardonnables, comme de partager des œuvres culturelles.

Le problème de l'identification d'un abonné précis, en présence de partage d'adresses IP est difficile. (Cela concerne IPv4 seulement car IPv6 n'a pas ce problème et son déploiement natif complet résoudrait le problème pour moins cher - cf. RFC 7021 [1] ; mais, comme disaient les Shadoks, « pourquoi faire simple quand on peut faire compliqué ? ») Je recommande fortement la lecture du RFC 6269 pour en saisir tous les aspects. Notez déjà un point important : si on n'a pas noté, au moment de l'observation du comportement illégal, non seulement l'adresse IP source, mais également le port source, on n'a guère de chance de remonter jusqu'à un abonné individuel, dès qu'il y a utilisation du CGN. Un premier pré-requis est donc que les serveurs journalisent le port source (comme demandé par le RFC 6302) et que les systèmes d'observation du réseau enregistrent ce port. Ensuite, muni de l'adresse IP source, du port source, **et** d'une heure exacte (ce qui suppose que tout le monde utilise NTP ou équivalent, cf. RFC 6269, section 12), on peut aller voir le FAI et lui poser la question « on voudrait savoir **qui** utilisait 192.0.2.199 :5347 le mercredi 24 décembre à 08 :20 » (notez la nouveauté, la mention du port, ici 5347). Si

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc7021.txt

le FAI alloue dynamiquement adresses et ports en sortie, et journalise ces allocations, il lui « suffira » de faire un grep dans les journaux pour donner la réponse. Mais, et ce mais est à l'origine de notre nouveau RFC, ces journaux peuvent être de très grande taille.

Le partage massif d'adresses, tel que pratiqué dans les CGN, est motivé par l'épuisement des stocks d'adresses IPv4 <https://www.bortzmeyer.org/epuisement-adresses-ipv4.html>. Plusieurs RFC décrivent des variantes du concept de CGN (RFC 6264, RFC 6333...) mais tous ont un point commun : une adresse IP est partagée, non pas seulement entre les membres d'un même foyer, mais entre des abonnés d'un même FAI, abonnés qui ne se connaissent pas, et ne forment pas légalement une entité unique. Pas question donc, dans un état de droit, de punir tous les utilisateurs de 192.0.2.199 parce que l'un d'entre eux a fait quelque chose d'illégal, comme de distribuer des fichiers pris dans l'entreprise qui pirate les ordinateurs de ses clients.

A priori, l'information sur qui utilisait quel couple {adresse, port} à un moment donné est connue du CGN. Il l'enregistre, quelque chose du genre (syntaxe imaginaire) :

```
2014-12-24T08:20:00 Outgoing connection from 10.4.8.2:6234, allocate 192.0.2.199:5347
2014-12-24T08:21:15 192.0.2.199:5347 is now free (used for 10.4.8.2:6234)
```

Le FAI sait donc que l'adresse interne correspondant à 192.0.2.199 :5347 était 10.4.8.2 et il peut alors consulter son plan d'adressage interne pour savoir de quel abonné il s'agit (ou bien consulter un autre journal si ces adresses internes sont elle-mêmes dynamiques). Au fait, si vous voulez un exemple réel des journaux d'un routeur CGN, voici un extrait de la documentation de Juniper :

```
Jun 28 15:29:20 cypher (FPC Slot 5, PIC Slot 0) {sset2}[FWNAT]: ASP_SFW_CREATE_ACCEPT_FLOW: proto 6 (TCP) ap
Jun 28 15:29:23 cypher (FPC Slot 5, PIC Slot 0) {sset2}[FWNAT]:ASP_NAT_POOL_RELEASE: natpool release 129.0.0
```

Comme dans mon exemple imaginaire, on y voit que deux entrées sont enregistrées, pour l'allocation du port et pour sa libération (sur les routeurs A10, on peut configurer séparement les formats de journalisation pour ces deux événements, avec les mots-clés `fixed-nat-allocated` et `fixed-nat-freed`).

Mais quelle est la taille de ces journaux de ces CGN ? L'expérience montre des tailles de 150 à 175 octets par entrée, ce qui est la taille de mon exemple imaginaire ci-dessus, qui journalise en mode texte (cela peut bien sûr se faire de manière structurée dans une base de données, ou bien cela peut se comprimer, les journaux texte se réduisent facilement d'un facteur 2 ou 3 à la compression). Mais à quelle rythme les entrées sont-elles créées ? Il ne semble pas y avoir beaucoup d'études précises sur ce sujet mais des observations chez des FAI états-uniens indiquent des moyennes par abonné autour de 33 000 connexions par jour. Cela ferait plus de 5 mégaoctets par jour et par abonné. Avec un million d'abonnés, le FAI devrait stocker 150 téraoctets par mois. Et il faut la capacité d'acheminer ces journaux : avec seulement 50 000 abonnés, il faut dédier 23 Mb/s entre le routeur CGN et le serveur de journalisation.

Et, une fois stockées les données, il reste à les fouiller. Il faut trouver deux évenements (le début de l'allocation et sa fin) au milieu de ces énormes fichiers, ce qui va prendre du temps.

Une solution élégante, préconisée par notre RFC, est d'avoir un mécanisme **déterministe** d'allocation des ports en sortie, de manière à ne pas avoir à journaliser l'allocation. Il suffira alors de faire tourner l'algorithme à l'envers pour savoir qui avait tel couple {adresse IP publique, port}.

Quelle est la dynamique d'allocation des ports ? Même si un abonné utilise des milliers de connexions par jour, à un instant donné, sa consommation est bien plus faible. Si le rapport entre le nombre d'abonnés et le nombre d'adresses IP publiques est faible (mettons de l'ordre de 10), chaque abonné pourra utiliser des milliers de ports sans gêner les autres. On peut donc allouer des intervalles entiers de ports, sans avoir besoin de journaliser chaque allocation d'un port donné. Il « suffit » donc d'avoir une fonction déterministe, qui associe à chaque adresse IP interne une adresse IP externe et un intervalle de ports externes. Un exemple trivial d'une telle fonction serait l'allocation de l'intervalle 1024-2999 au premier (dans l'ordre des adresses IP internes) abonné, de 3000-4999 au deuxième, etc. Lorsqu'on a épuisé les numéros de port, on passe à la deuxième adresse IP publique et on recommence. En sens inverse, lorsqu'on recevra la requête « qui utiisait 192.0.2.1 :4219 le mercredi 24 décembre à 08 :20 ? », on saura, sans consulter le journal, que c'était le deuxième de nos abonnés (deuxième intervalle de ports de la première adresse publique). Cette fonction n'est qu'un exemple, la décision d'utiliser telle ou telle méthode est une décision purement locale. (Attention à ne pas allouer les ports séquentiellement dans l'intervalle donné, afin de limiter les risques pour la vie privée de l'abonné. La section 5 du RFC détaille ce risque pour la vie privée, et suggère des mesures.)

Pour cela, le routeur CGN a besoin de connaitre la liste des adresses internes (avec certains CGN comme DS-Lite, technique de coexistence temporaire IPv4/IPv6, ce seront des adresses IPv6, cf. la section 4 de notre RFC), celles des adresses externes disponibles pour sa fonction de CGN, le nombre total d'abonnés (pour calculer le rapport avec le nombre d'adresses publiques), le nombre de ports par utilisateur, la liste des ports à ne pas utiliser, et, bien sûr, la fonction déterministe de correspondance entre une adresse interne et un couple {adresse externe, port externe}. Parmi les fonctions possibles :

— Allocation séquentielle (comme dans l'exemple proposée plus haut, ou dans l'exemple plus détaillé de la section 2.3 du RFC),
— Entrelacement : si on a un facteur de 10 entre le nombre d'abonnés et le nombre d'adresses publiques, on alloue à chaque abonné un port sur dix. Le premier abonné a les ports 1024, 1034, 1044, etc,le deuxième 1025, 1035, etc,
— Alternance sur l'adresse : un abonné utilise toujours le même port externe mais avec une adresse IP publique différente par connexion. Cela ne marche que si on a autant d'adresses IP publiques que de connexions par abonné mais cela simplifie beaucoup la recherche,
— Méthode cryptographique, inspirée de la section 2.2 du RFC 6431 : on chiffre une concaténation d'une clé et d'autres informations et le résultat chiffré nous donne le port externe à utiliser. Attention, il faudra la clé pour inverser la fonction et donc retrouver l'abonné, il ne faut pas la jeter si on veut pouvoir répondre à des demandes concernant un passé un peu lointain, alors que les clés ont été changées.

Outre les ports « système » (RFC 6335), exclus de l'allocation, il est prudent de garder en réserve un intervalle de ports pour des allocations dynamiques traditionnelles, avec journalisation de l'allocation. Cela permet, par exemple, de gérer les utilisateurs avancés qui utilisent tellement de connexions sortantes qu'ils épuisent l'intervalle des ports. Avec cette réserve, on pourra toujours les satisfaire. Il faudra certes enregistrer ces allocations mais on aura quand même gagné en taille, en ne journalisant que moins d'information.

Comme on n'a rien sans rien, la plupart des méthodes d'allocation déterministe sont moins « efficaces » qu'une allocation purement dynamique, au sens où elles sous-utilisent les ports (pour les utilisateurs peu gourmands). En outre, elles imposent davantage de travail aux équipes opérationnelles (choix d'un algorithme, réglage de ses paramètres...). D'autre part, il ne faut pas s'imaginer qu'un CGN dissimule sérieusement l'abonné final, l'algorithme d'allocation des ports peut toujours être rétroingénierié (cf. section 6, sur la sécurité).

À noter qu'il faut aussi, pour que tout se passe bien, noter la configuration du CGN et ses éventuels changements. Pas question d'appliquer l'algorithme d'aujourd'hui à une requête judiciaire qui concernerait le mois précédent, si l'algorithme ou ses paramètres ont changé. Il faut donc que le routeur CGN journalise également les changements de paramètres (comme le nombre de ports par abonné).