RFC 6179 : The Internet Routing Overlay Network (IRON)

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 21 mars 2011

Date de publication du RFC : Mars 2011

https://www.bortzmeyer.org/6179.html

Sous la houlette du groupe de travail RRG ("*Routing Research Group*" <http://www.irtf.org/charter?gtype=rg&group=rrg>) de l'IRTF, un gros effort est en cours pour définir une nouvelle architecture de routage pour l'Internet. Cette future architecture doit permettre de poursuivre la croissance de l'Internet sans se heurter aux limites quantitatives des routeurs, et en autorisant des configurations qui sont actuellement peu pratiques comme le "*multi-homing*". Cet effort a été synthétisé dans les recommandations du RFC 6115[1]. La plupart des propositions d'architecture résumées dans le RFC 6115 seront sans doute publiées, chacune dans son RFC. Notre RFC 6179 ouvre la voie en décrivant le système **IRON** ("*Internet Routing Overlay Network*").

IRON fait partie de la famille des architectures "*Map-and-Encap*" où les réseaux ont un identificateur indépendant de leur position dans le réseau et où les routeurs doivent utiliser un mécanisme de "*mapping*", de correspondance entre l'identificateur connu et le localisateur cherché, avant d'encapsuler les paquets, pour qu'ils puissent atteindre leur destination via un tunnel. L'originalité d'IRON est de combiner "*mapping*" et routage pour distribuer l'information nécessaire aux routeurs. Comme les autres architectures "*Map-and-Encap*" (par exemple LISP), IRON ne nécessite pas de changement dans les machines terminales, ni dans la plupart des routeurs, mais seulement dans les routeurs d'entrée et de sortie des tunnels.

Comme IRON utilise une terminologie très spéciale, qu'on ne trouve pas dans les autres protocoles, cela vaut la peine, si on veut lire ce RFC, de bien apprendre la section 2, qui décrit le vocabulaire. IRON est donc un système d'"*overlay*", c'est-à-dire de réseau virtuel fonctionnant au dessus d'un réseau « physique », l'Internet. Il part de RANGER RFC 5720 en le généralisant à tout l'Internet. En partant du bas (la description du RFC, c'est peut-être symptomatique, part du haut), chaque organisation connectée à IRON (un **EUN** pour "*End User Network*") aurait un préfixe à elle, un identificateur, nommé le **EP** ("*EUN*

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc6115.txt

Prefix"). Ces préfixes viennent de préfixes plus larges, les **VP** ("*Virtual Prefix*"), qui sont alloués par de nouvelles organisations, les **VPC** ("*Virtual Prefix Companies*"). Ces organisations n'ont pas d'équivalent dans l'Internet actuel : les VPC pourraient être des FAI mais ce n'est pas obligatoire. Le RFC ne donne aucun détail sur les points non-techniques de cette architecture (par exemple, comment une VPC gagnera sa vie ? Si on change de VPC, pourra t-on garder son préfixe ?) considérant que ces questions sont hors-sujet pour une spécification technique.

Les routeurs IRON se connectent entre eux par des liens NBMA qui sont simplement des tunnels au dessus de l'Internet existant.

Toujours dans le vocabulaire original (IRON n'utilise pas les termes d'ITR - "*Ingress Tunnel Router*" - et ETR - "*Egress Tunnel Router*" - probablement parce que son architecture ne colle pas bien à ce modèle), IRON repose sur trois types de routeurs (appelés collectivement **agents**) :

— Le **client** est le routeur chez l'EUN, l'organisation terminale (entreprise, association, université, etc). Il connait les EP de l'organisation et se connecte au « réseau » IRON (éventuellement via du NAT). La section 3.1 détaille ces clients.
— Le **serveur** appartient à la VPC et il reçoit les « connexions » des EUN qui hébergent les clients. La section 3.2 le décrit complètement.
— Le **relais** appartient également à la VPC et assure l'interconnexion avec l'Internet actuel, pour les correspondants qui ne sont pas dans IRON. Il va donc parler BGP avec les routeurs actuels et utiliser les protocoles IRON avec les autres relais. Il est présenté complètement en section 3.3.

Les détails de l'architecture sont ensuite détaillés en sections 3 à 6. Les agents IRON (les trois sortes de routeurs présentées plus haut) forment des tunnels entre eux, en utilisant la technique VET ("*Virtual Enterprise Traversal*", RFC 5558) et encapsulent selon le format SEAL ("*Subnetwork Encapsulation and Adaptation Layer*", RFC 5320). Curieusement, l'auteur prend soin de préciser que SEAL permet de tunneler d'autres protocoles qu'IP et il cite CLNP, ce qui donne une idée de l'âge des idées qui sont derrière SEAL...

SEAL dispose d'un protocole d'échange d'informations sur les routes, les MTU, etc, SCMP ("*SEAL Control Message Protocol*") et c'est lui qui permet à l'"*overlay*" IRON de fonctionner.

Chaque VPC représente donc un morceau d'un "*patchwork*" global. Les morceaux sont connectés entre eux via l'Internet public. Le client final, lui, doit se connecter à l'Internet **puis** acquérir une connectivité IRON auprès d'une VPC, et s'équiper d'un routeur client qui va relier son EUN (son réseau local) à la VPC.

Pour amorcer la pompe (section 5), chaque relais est configuré avec la liste des VP qu'il va servir, la liste des serveurs (les routeurs des VPC), et avec celle de ses voisins BGP. Avec ces derniers, il établit des liens comme aujourd'hui et annonce sur ces liens les VP. Avec les serveurs, le relais découvre la correspondance entre les EP, les identificateurs des réseaux IRON et les serveurs. (Rappelons que les identificateurs, EP et VP, ont la même forme que des adresses IP et peuvent donc être annoncés comme tels aux routeurs BGP actuels.)

Les serveurs, eux, sont nourris avec la liste des relais, auxquels ils vont transmettre les VP qu'il servent. Ils attendent ensuite les connexions des clients, apprenant ainsi petit à petit les EP connectés, que les serveurs transmettront aux relais.

Les clients, eux, auront besoin de connaître les serveurs de leur VPC, ainsi que leurs EP, et un moyen de s'authentifier.

Une fois tout ceci fait, les routeurs n'ont plus qu'à router les paquets comme aujourd'hui (section 6, qui décrit les différents cas, notamment selon que le destinataire du paquet est dans IRON ou pas ; par exemple, le cas du "*multi-homing*" est en section 6.5.2, qui précise qu'un client peut avoir plusieurs localisateurs pour un même EP).

Et d'une manière plus générale, comment cela va se passer en pratique ? La section 7 se lance dans certains détails concrets. Par exemple, comme le routage à l'intérieur d'IRON est souvent sous-optimal, mais qu'un mécanisme d'optimisation des routes permet de les raccourcir par la suite, le premier paquet de chaque flot aura sans doute un temps de trajet bien supérieur à celui de ses copains.

Et, comme toute solution utilisant des tunnels, IRON aura sans doute à faire face à des problèmes de MTU, que cette section mentionne, mais jette un peu rapidement à la poubelle.

Toujours pour ceux qui s'intéressent aux problèmes concrets, l'annexe B est une très bonne étude des propriétés de passage à l'échelle d'IRON.

IRON ne semble pas avoir été testé, même en laboratoire, même tout simplement dans un simulateur. Et mon avis personnel ? Je dois avouer n'avoir pas tout compris et avoir la sensation inquiétante que ce n'est pas uniquement une limite de mon cerveau, mais que l'architecture d'IRON reste encore insuffisamment spécifiée.