RFC 8335 : PROBE: A Utility For Probing Interfaces

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 23 février 2018

Date de publication du RFC : Février 2018

https://www.bortzmeyer.org/8335.html

Pour tester qu'une machine est bien joignable, vous utilisez ping ou, plus rigoureusement, vous envoyez un message ICMP de type `echo`, auquel la machine visée va répondre avec un message ICMP `echo reply`. Ce test convient souvent mais il a plusieurs limites. L'une des limites de ce test est qu'il ne teste qu'une seule interface réseau de la machine, celle par laquelle vous lui parlez (deux interfaces, dans certains cas de routage asymétrique). Si la machine visée est un gros routeur avec plein d'interfaces réseau, le test ne vous dira pas si toutes fonctionnent. D'où cette extension aux messages ICMP permettant de spécifier l'interface qu'on veut vérifier.

A priori, ce RFC ne s'intéresse qu'aux routeurs, les serveurs n'ayant souvent qu'une seule interface réseau. La nouvelle technique, nommée `PROBE`, n'a pas de vocation générale, contrairement à ping, et concernera surtout les administrateurs réseau. D'autant plus que, comme elle est assez indiscrète, elle ne sera a priori pas ouverte au public. Notez qu'elle permet non seulement de tester une autre interface du routeur, mais également une interface d'une machine directement connectée au routeur. Les scénarios d'usage proposés sont exposés dans la section 5, une liste non limitative de cas où ping ne suffit pas :

- Interface réseau non numérotée (pas d'adresse, ce qui est relativement courant sur les routeurs),
- Interface réseau numérotée de manière purement locale (par exemple adresse IPv6 "*link-local*"),
- Absence de route vers l'interface testée (si on veut tester l'interface d'un routeur qui fait face à un point d'échange, et que le préfixe du point d'échange n'est pas annoncé par le protocole de routage, ce qui est fréquent).

En théorie, SNMP pourrait servir au moins partiellement à ces tests mais, en pratique, c'est compliqué.

ping, la technique classique, est très sommairement décrit dans le RFC 2151 [1], section 3.2, mais sans indiquer comment il fonctionne. La méthodologie est simple : la machine de test envoie un message ICMP `Echo` (type 8 en IPv4 et 128 en IPv6) à la machine visée (l'amer <https://www.bortzmeyer.org/amer-mire.html>). L'amer répond avec un `Echo Reply` (type 0 en IPv4 et 129 en IPv6). La réception de cet `Echo Reply` indique que la liaison marche bien dans les deux sens. La non-réception indique d'un problème s'est produit, mais on n'en sait pas plus (notamment, on ne sait pas si le problème était à l'aller ou bien au retour). Ici, on voit le test effectué par une sonde Atlas <https://atlas.ripe.net/> sur l'amer `2605:4500:2:245b::42` (l'un des serveurs hébergeant ce blog), vu par tshark :

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc2151.txt`

```
13.013422 2a02:1811:c13:1902:1ad6:c7ff:fe2a:6ac → 2605:4500:2:245b::42 ICMPv6 126 Echo (ping) request id=0x
13.013500 2605:4500:2:245b::42 → 2a02:1811:c13:1902:1ad6:c7ff:fe2a:6ac ICMPv6 126 Echo (ping) reply id=0x05
```

ICMP est normalisé dans les RFC 792 pour IPv4 et RFC 4443 pour IPv6. L'exemple ci-dessus montre un test classique, avec une requête et une réponse.

Notre RFC parle d'« interface testée » ("*probed interface*") et d'« interface testante » ("*probing interface*"). Dans l'exemple ci-dessus, l'interface Ethernet de l'Atlas était la testante et celle du serveur était la testée. Le succès du test montre que les deux interfaces sont actives et peuvent se parler.

Au contraire de ping, `PROBE` va envoyer le message, non pas à l'interface testée mais à une interface « relais » ("*proxy*"). Celle-ci répondra si l'interface testée fonctionne bien (état `oper-status`, cf. RFC 7223). Si l'interface testée n'est pas sur le nœud qui sert de relais, ce dernier détermine l'état de cette interface en regardant la table ARP (RFC 826) ou NDP (RFC 4861). Aucun test actif n'est effectué, l'interface est considérée comme active si on lui a parlé récemment (et donc si l'adresse IP est dans un cache). `PROBE` utilise, comme ping, ICMP. Il se sert des messages ICMP structurés du RFC 4884. Une des parties du message structuré sert à identifier l'interface testée.

L'extension à ICMP "*Extended Echo*" est décrite en section 2 du RFC. Le type de la requête est 42 pour IPv4 et 160 pour IPv6 (enregistré à l'IANA, pour IPv4 <`https://www.iana.org/assignments/icmp-parameters/icmp-parameters.xml#icmp-parameters-types`> et IPv6 <`https://www.iana.org/assignments/icmpv6-parameters/icmpv6-parameters.xml#icmpv6-parameters-2`>). Parmi les champs qu'elle comprend (les deux premiers existent aussi pour l'ICMP "*Echo*" traditionnel) :

- Un identificateur qui peut servir à faire correspondre une réponse à une requête (ICMP, comme IP, n'a pas de connexion ou de session, chaque paquet est indépendant), 0x0545 dans l'exemple vu plus haut avec tshark,
- Un numéro de séquence, qui peut indiquer les paquets successifs d'un même test, 1 dans l'exemple vu plus haut avec tshark
- Un bit nommé L (local) qui indique si l'interface testée est sur le nœud visé par le test ou non,
- Une structure qui indique l'interface testée.

Cette structure suit la forme décrite dans la section 7 du RFC 4884. Elle contient un objet d'identification de l'interface. L'interface qu'on teste peut être désignée par son adresse IP (si elle n'est pas locale - bit L à zéro, c'est la seule méthode acceptée), son nom ou son index. Notez que l'adresse IP identifiant l'adresse testée n'est pas forcément de la même famille que celle du message ICMP. On peut envoyer en IPv4 un message ICMP demandant à la machine distante de tester une interface IPv6.

Plus précisément, l'objet d'identification de l'interface est composé, comme tous les objets du RFC 4884, d'un en-tête et d'une charge utile. L'en-tête contient les champs :

- Un numéro de classe, 3, stocké à l'IANA <`https://www.iana.org/assignments/icmp-parameters/icmp-parameters.xml#icmp-parameters-ext-classes`>,
- Un numéro indiquant comment l'interface testée est désignée (1 = par nom, cf. RFC 7223, 2 = par index, le même qu'en SNMP, voir aussi le RFC 7223, section 3, le concept de `if-index`, et enfin 3 = par adresse),
- La longueur des données.

L'adresse est représentée elle-même par une structure à trois champs, la famille (4 pour IPv4 et 6 pour IPv6), la longueur et la valeur de l'adresse. Notez que le RFC 5837 a un mécanisme de description de l'interface, portant le numéro de classe 2, et utilisé dans un contexte proche.

La réponse à ces requêtes a le type 43 en IPv4 et 161 en IPv6 (section 3 du RFC). Elle comprend :

— Un code (il valait toujours 0 pour la requête) qui indique le résultat du test : 0 est un succès, 1 signale que la requête était malformée, 2 que l'interface à tester n'existe pas, 3 qu'il n'y a pas de telle entrée dans la table des interfaces, et 4 que plusieurs interfaces correspondent à la demande (liste complète dans un registre IANA <https://www.iana.org/assignments/icmpv6-parameters/icmpv6-parameters.xml#icmpv6-parameters-codes-type-161>),
— Un identificateur, copié depuis la requête,
— Un numéro de séquence, copié depuis la requête,
— Un état qui donne des détails en cas de code 0 (autrement, pas besoin de détails) : si l'interface testée n'est pas locale, l'état vaut 2 si l'entrée dans le cache ARP ou NDP est active, 1 si elle est incomplète (ce qui indique typiquement que l'interface testée n'a pas répondu), 3 si elle n'est plus à jour, etc <https://www.iana.org/assignments/icmp-parameters/icmp-parameters.xml#icmp-parameters-ext-classes>,
— Un bit A ("*active*") qui est mis à un si l'interface testée est locale et fonctionne,
— Un bit nommé 4 (pour IPv4) qui indique si IPv4 tourne sur l'interface testée,
— Un bit nommé 6 (pour IPv6) qui indique si IPv6 tourne sur l'interface testée.

La section 4 du RFC détaille le traitement que doit faire la machine qui reçoit l'ICMP "*Extended Echo*". D'abord, elle doit jeter le paquet (sans répondre) si ICMP "*Extended Echo*" n'est pas explicitement activé (rappelez-vous que ce service est assez indiscret, cf. section 8 du RFC) ou bien si l'adresse IP de la machine testante n'est pas autorisée (même remarque). Si les tests sont passés et que la requête est acceptée, la machine réceptrice fabrique une réponse : le code est mis à 1 si la requête est anormale (pas de partie structurée par exemple), 2 si l'interface testée n'existe pas, 3 si elle n'est pas locale et n'apparait pas dans les tables (caches) ARP ou NDP. Si on trouve l'interface, on la teste et on remplit les bits A, 4, 6 et l'état, en fonction de ce qu'on trouve sur l'interface testée.

Reste la question de la sécurité (section 8 du RFC). Comme beaucoup de mécanismes, PROBE peut être utilisé pour le bien (l'administrateur réseaux qui détermine l'état d'une interface d'un routeur dont il s'occupe), mais aussi pour le mal (chercher à récolter des informations sur un réseau avant une attaque, par exemple, d'autant plus que les noms d'interfaces dans les routeurs peuvent être assez parlants, révélant le type de réseau, le modèle de routeur…) Le RFC exige donc que le mécanisme ICMP "*Extended Echo*" ne soit pas activé par défaut, et soit configurable (liste blanche d'adresses IP autorisées, permission - ou non - de tester des interfaces non locales, protection des différents réseaux les uns contre les autres, si on y accueille des clients différents…) Et, bien sûr, il faut pouvoir limiter le nombre de messages.

Ne comptez pas utilise PROBE tout de suite. Il n'existe apparemment pas de mise en œuvre de ce mécanisme publiée. Juniper en a réalisé une mais elle n'apparait encore dans aucune version de JunOS.