RFC 7799 : Active and Passive Metrics and Methods (with Hybrid Types In-Between)

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 30 mai 2016

Date de publication du RFC : Mai 2016

https://www.bortzmeyer.org/7799.html

Dans le monde de la métrologie, il existe deux classes de mesures effectuées sur l'Internet : les mesures **actives** et les mesures **passives**. La signification de ces termes semble assez intuitive mais ce nouveau RFC fournit des définitions rigoureuses, qui n'étaient pas encore publiées. (Il y a aussi des mesures hybrides.)

Cela fait pourtant de très nombreuses années que ces deux termes sont largement utilisés. Par exemple, en 2000 s'est tenue une conférence « "*Passive and Active Measurement (PAM)*" <https://www.ripe.net/participate/meetings/pam-2001> ». Et les termes étaient certainement plus anciens.

En première approximation, on peut caractériser ces deux classes ainsi (section 1 du RFC) :

— Une mesure **active** nécessite un envoi de paquets dédiés à cette mesure.
— Une mesure **passive** ne dépend que des paquets qui seraient envoyés de toute façon, même si la mesure n'avait pas lieu.

Du point de vue des outils, on peut donc noter que les ultra-traditionnels ping et traceroute font des mesures actives et que tcpdump permet des mesures passives.

Mais, les choses étant plus compliquées dans la vraie vie des vrais réseaux, notre RFC va un peu plus loin que cela. Sa section 3 définit plus précisement les termes. Elle s'appuie sur des RFC précédents comme les RFC 2330[1] et RFC 6390 qui définissaient, par exemple, la notion de métrique (une grandeur qu'on mesure), ou comme le RFC 7011, qui définissait la notion de point d'observation. On s'appuie aussi sur des documents qui ne sont pas des RFC comme la norme UIT Y.1540 <http://www.itu.int/rec/T-REC-Y.1540/en> qui définit la notion de « population [de paquets] qui nous intéresse » ("*stream of interest*") ou comme la norme UIT Y.1731 <http://www.itu.int/rec/T-REC-Y.1731/en>.

Notez que les adjectifs « actif » et « passif » peuvent s'appliquer aux méthodes de mesure ou bien aux métriques.

Armé de ces définitions, on peut dire que les méthodes de mesure actives ont les caractéristiques suivantes :

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc2330.txt

— La population de paquets qui nous intéresse est générée (synthétisée). Parfois un autre flux de paquets est généré, même si leur sort ne nous intéresse pas directement, par exemple pour générer de la charge.
— Les paquets générés ont un contenu qui est spécifique à la mesure, qui n'a pas d'intérêt autrement. C'est par exemple le numéro de séquence dans les paquets ICMP (RFC 792) que génère ping, numéros qui permettent de faire correspondre une réponse à une question.
— Source et destination des paquets sont connus a priori.
— D'ailleurs, toutes les caractéristiques de la population sont connues à l'avance, ce qui facilite nettement l'observation. Parfois, certaines caractéristiques seront changées en route (comme le champ TTL dans l'en-tête IPv4).

Comme la mesure active introduit du trafic réseau qui n'existerait pas sans elle, elle perturbe le réseau qu'on veut mesurer. Un des défis importants des mesures actives est de minimiser cette perturbation ou, au minimum, de pouvoir quantifier cette perturbation pour l'indiquer comme marge d'erreur dans les résultats (section 4.1).

Une métrique active est simplement une métrique dont la définition inclut des méthodes actives. Par exemple, le type des paquets qu'on a injecté dans le réseau.

Les méthodes passives, par contre, ont ces propriétés :

— Elles sont fondées sur l'observation de paquets non modifiés et non perturbés. Une méthode passive ne laisse aucune trace sur le réseau.
— Elles ne sont donc utiles que s'il existe déjà des paquets intéressants, puisqu'elles n'en créent pas et n'en modifient pas.

Il faut faire attention en observant les paquets : comme on ne choisit pas leurs caractéristiques, on risque de ne pas voir le vrai phénomène, car on n'observait que le phénomène supposé. (Il m'est arrivé d'utiliser tcpdump pour observer un trafic LDAP et de filtrer sur l'adresse du serveur LDAP. Je ne comprenais pas pourquoi aucun trafic n'était visible. En fait, je ne capturais que les paquets IPv4 alors que le trafic se faisait en IPv6.)

Une observation passive peut se faire en un seul point, ou en plusieurs. Un exemple d'observation passive est ce que fait un routeur IPFIX (RFC 7011).

Une métrique passive est une métrique fondée uniquement sur l'observation passive. Par exemple, le temps de trajet entre deux points peut être formalisée dans une métrique passive : on observe le paquet à un point et à l'autre, et on en déduit le temps de son voyage.

Évidemment, dans le monde réel, les classifications ne sont pas toujours très simples. Par exemple, il existe des méthodes hybrides. Les « hybrides de type I » partent d'un flux de paquets existant mais le modifient, ou bien modifient le traitement qu'il subit. Par exemple, si on observe un flux existant mais qu'on envoie en même temps des données dans le réseau pour le charger, on est en présence d'une méthode hybride de type I. Les métriques spatiales du RFC 5644 sont un cas de métriques hybrides.

Et les hybrides de type II ? Le terme indique les méthodes où il y a plusieurs populations intéressantes générées activement (et d'autres qu'on observe passivement).

La section 4 de notre RFC discute de certains problèmes de métrologie et compare l'effet des méthodes et métriques actives ou passives. Par exemple, si on veut mesurer la capacité <https://www.bortzmeyer.org/capacite.html> d'un lien, une méthode active serait de tenter d'envoyer le plus de données possible, une méthode passive d'observer le trafic et de regarder le débit maximum atteint. La première méthode va remplir le tuyau, gênant les applications en cours. La seconde risque de sous-estimer la capacité (si les machines sont lentes, elles peuvent ne pas réussir à remplir le tuyau.)

Au passage, l'importance des mesures est telle que l'IETF envisage un mécanisme de mesure instrumentant les paquets eux-mêmes : c'est l'option IPv6 PDM ("*Performance and Diagnostic Measurements*", actuellement à l'état de projet), qui consiste à ajouter un nouvel en-tête d'extension aux paquets IP pour y noter les informations nécessaires aux mesures. Les mesures ainsi faites seraient plutôt des hybrides de type I. (On ajoute des données à des paquets existants.)