RFC 2330 : Framework for IP Performance Metrics

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 4 janvier 2009

Date de publication du RFC : Mai 1998

https://www.bortzmeyer.org/2330.html

La mesure des performances sur Internet est un vaste sujet, rendu plus difficile par l'utilisation massive de mots mal définis ou, pire, utilisés à contre sens (« débit », « bande passante » ou, pire, « vitesse » étant de bons exemples). C'est pour cela que le travail du groupe IPPM <http://tools.ietf.org/wg/ippm> ("*IP Performance Metrics*") a commencé par ce RFC 2330 [1] qui essaie de définir un cadre solide pour parler de mesures de performances.

Prenons l'exemple d'un cas qui semble trivial : un liaison Internet est saturée et on n'arrive plus à y faire passer tous les films qu'on veut regarder. Pour la remplacer, on hésite entre deux liaisons dont on voudrait bien mesurer la capacité, afin de choisir la « meilleure ». Mais que veut dire exactement « cette liaison permet un débit de 20 Mb/s » ? Ce chiffre est mesuré dans quelle couche ? Est-il valable en permanence ou bien est-ce une moyenne ? Dépend-il de la taille (ou d'autres caractéristiques) des paquets ? Sans réponse claire à ces questions, le chiffre de 20 Mb/s ne signifie pas grand'chose. D'autant plus que les vendeurs ont tout intérêt à brouiller les pistes. (Au fait, cette notion, nommée « capacité », est définie dans le RFC 5136.)

La section 3 du RFC résume les buts : définir des **métriques**, c'est-à-dire des grandeurs **mesurables** et définies de manière rigoureuse. Pour cela, notre RFC spécifie le cadre commun à toutes les métriques, qui seront développées dans des RFC ultérieurs (comme le RFC 2678 pour la métrique de connectivité ou bien le RFC 2681 pour la métrique de temps d'aller-retour entre deux points).

Quels sont les critères pour une « bonne » métrique ? C'est la question que traite la section 4. Parmi ces critères :

— Pouvoir être définie rigoureusement,
— Pouvoir être mesurée de façon reproductible,
— Être utile en pratique, pour les utilisateurs ou les administrateurs du réseau.

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

On voit que les questions de vocabulaire vont être cruciales. Pour rompre avec le flou des données, il faut que les grandeurs mesurées soient définies d'une manière rigoureuse et précise. Ainsi, la section 5 donne des définitions pour un certain nombre d'éléments qui seront utilisés dans les métriques. Par exemple, "*host*", machine connectée au réseau, inclus explicitement les routeurs, "*link*" est un lien de couche 2, "*path*" un ensemble contigu de "*links*", etc.

Une fois ces bases de langage établies, le RFC s'attaque aux concepts (section 6). Il s'attache à bien établir la différence entre la **définition** d'une métrique (qui doit avant tout être complète et rigoureuse) et la **mesure** effective (qui n'est pas toujours facile, comme illustré plus loin par la discussion sur la mesure du "*wire time*").

Cette section est aussi l'occasion de poser certaines conventions comme le fait que les temps se mesurent toujours en UTC et que les kilos valent toujours 1000 (suivant la règle des télécommunications, celle du stockage informatique étant plutôt qu'un kilo vaut 1024).

Une fois une métrique définie, il faut la mesurer : c'est l'objet de la section 6.2. On peut le faire directement en mesurant la grandeur qui nous intéresse ou indirectement, par exemple en calculant ou en déduisant cette grandeur d'autres mesures. Les particularités de la mesure doivent être soigneusement étudiées et publiées avec les résultats. Par exemple, certaines mesures modifient la grandeur mesurée (des pings répétés chargent le réseau et modifient donc le temps d'aller-retour que ping est censé mesurer). La science de la mesure est aussi ancienne que la physique et ces problèmes, tout comme ceux des erreurs ou incertitudes (section 6.3) sont bien connus, mais souvent oubliés lorsqu'il s'agit de réseaux informatiques. Le RFC a donc raison de les rappeler et d'insister sur l'importance de les documenter lorsqu'on publie.

Les métriques ne vivent pas seules : il est souvent utile de les **composer**, ce que traite la section 9. Elle décrit en 9.1 les compositions **spatiales** où les mesures sur des sous-chemins ("*subpath*") s'additionnent. Par exemple, le délai de propagation d'un paquet le long d'un chemin ("*path*") est proche de la somme des délais de propagation sur les sous-chemins. En 9.2, on trouve les compositions **temporelles**, où le passé permet de prédire le futur. Par example, si on a mesuré le débit sur un lien depuis vingt-quatre heures, et que sa variation selon l'heure est bien comprise, on va pouvoir déduire le débit futur pendant les prochaines vingt-quatre heures.

La plupart des mesures, en matière de réseau, font intervenir le temps. Le concept est souvent mal compris et une section entière, la 10, est dédiée à Chronos. Certes, l'Internet dispose depuis longtemps d'un protocole standard pour la gestion d'horloges, NTP (RFC 5905). Mais ses buts sont différents : il vise à garder des horloges synchronisées sur une longue période. Pour la mesure, on s'intéresse à des périodes plus courtes pour lesquelles NTP n'est pas forcément la solution.

Quels sont les problèmes avec les horloges ? Une horloge peut avoir un **écart** ("*offset*") avec le « vrai » temps, UTC (ou bien avec une autre horloge, si on s'intéresse juste au fait qu'elles soient synchronisées). Si cet écart est inférieur à une certaine valeur, on dit que l'horloge est **correcte** ("*accurate*"). Et l'horloge peut avoir un **décalage** ("*skew*"), elle peut avancer plus ou moins vite que la normale (et donc cesser au bout d'un moment d'être correcte). Pire, la dérive peut être variable, auquel cas on mesure la dérivée seconde du décalage sous le nom de **dérive** ("*drift*"). Enfin, l'horloge a une certaine **résolution** ("*resolution*"), la plus petite unité de temps qu'elle peut mesurer.

Pour minimiser l'effet de ces limites sur la mesure, le RFC recommande donc que les horloges des différentes machines impliquées soient synchronisées, idéalement à partir d'une source extérieure commune (comme le GPS). Le RFC illustre l'importance de la synchronisation par un simple exemple. Si on

mesure le délai de transmission d'un paquet sur un lien transaméricain, une valeur de 50 ms est raisonnable. Si le décalage est de 0,01 %, après seulement dix minutes, l'écart atteindra 60 ms soit davantage que ce qu'on veut mesurer.

NTP n'est pas forcément la bonne solution car son exactitude dépend des liens Internet, exactement ceux dont on veut mesurer les caractéristiques. D'autre part, NTP donne la priorité à la correction, pas à la limitation du décalage. NTP peut délibèrément accélérer ou ralentir une horloge (augmenter ou diminuer le décalage) pour la rapprocher du temps correct. On n'a pas envie que cela survienne pendant une mesure !

Un autre problème de mesure lié au temps est celui du "*wire time*" (section 10.2). Les mesures sont souvent effectuées sur les machines elle-mêmes, et pas via un équipement spécial connecté au câble. Les machines ne sont pas forcément optimisées pour la mesure et n'ont pas forcément un noyau temps réel. Elles peuvent donc introduire leurs propres inexactitudes. Pour séparer les délais internes à la machine des délais du réseau, le RFC recommande d'utiliser dans les métriques le « temps du câble » ("*wire time*"), pas le temps de la machine. Le premier reflète le moment où le paquet est effectivement présent sur le câble. Plus rigoureux, il est par contre plus difficile à mesurer. Ainsi, pour un programme comme ping, le temps de la machine ("*host time*") est facilement connu grâce à `gettimeofday` alors qu'il n'y a pas de moyen de mesurer le temps du câble, depuis une application Unix ordinaire. C'est donc une instance particulière du problème discuté plus haut, où le souci de donner une définition correcte de la métrique a priorité sur la facilité de mesure.

Pour réaliser une mesure du temps du câble, le RFC recommande d'utiliser un filtre à paquets comme le BPF qui associe à chaque paquet une estampille temporelle en général proche du temps du câble. Il recommande également de ne **pas** faire tourner ce filtre sur la machine qu'on veut observer, pour limiter les perturbations.

Les mesures impliquent des statistiques et la section 11 quitte les rivages de la physique pour revenir à ceux de la mathématique. Elle est consacrée à la différence entre les mesures uniques ("*singleton*") et les échantillons ("*samples*"), ensembles de mesures uniques. Par exemple, un échantillon de mesures consiste en « une heure de mesures uniques, chacune prise à intervalles de Poisson avec un écart moyen d'une seconde entre les mesures ». L'échantillonnage est un art délicat et la section 11.1 expose différents moyens de le réaliser. Par exemple, le plus simple est la mesure périodique, à intervalles fixes. Mais elle a des défauts :

— Si la grandeur mesurée est elle-même périodique, la mesure risque de se faire toujours au même moment, par rapport à la période du phénomène observé. On peut alors ne pas percevoir la périodicité de celui-ci.
— Si la mesure est active, l'injection de données à intervalles périodiques entraîne parfois des effets de synchronisation non désirés (cf. un article fameux, S. Floyd et V. Jacobson, "*The Synchronization of Periodic Routing Messages*" <`http://ee.lbl.gov/papers/sync_94.pdf`>, "*IEEE/ACM Transactions on Networking*", avril 1994).
— Enfin, l'échantillonage périodique est prévisible, ce qui peut être gênant si un malhonnête essaie d'influencer le résultat.

Le RFC recommande donc d'échantilloner à intervalles aléatoires, suivant une certaine distribution. Quelques exemples de telles distributions suivent comme celle de Poisson en section 11.1.1.

Autre piège de la mathématique, les probabilités (section 12). Le RFC met en garde contre leur abus lors de la définition d'une métrique. Dire que 34 paquets sur 100 ont été perdus (ce que ping affiche sous l'étiquette "*packet loss*") est un fait, dire que « la probabilité de perte d'un paquet est de 0,34 » est une interprétation, qui suppose que le phénomène soit aléatoire (ce qui est faux, la perte des paquets, contrairement à la désintégration d'un atome radioactif est très déterministe).

Outre les questions physiques et mathématiques, la mesure cache des pièges liés à l'informatique. Pour plusieurs raisons, les différents paquets ne reçoivent pas le même traitement. Par exemple, certains FAI ralentissent délibérement les paquets considérés comme liés à un protocole pair à pair. La mesure dépend donc du type de paquets, ce que la section 13 formalise avec la notion de « paquets de type P ». Lorsqu'on dit « la machine `2001:db8:42::bad:cafe` est joignable », cela n'est pas toujours vrai pour tous les protocoles, en raison notamment des coupe-feux. Il faut donc dire « la machine `2001:db8:42::bad:cafe` est joignable pour un certain type P », où P est défini comme, mettons « paquets TCP à destination du port 80 ».