RFC 7536 : Large-Scale Broadband Measurement Use Cases

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 8 mai 2015

Date de publication du RFC : Mai 2015

https://www.bortzmeyer.org/7536.html

Mesurer les performances d'un réseau, c'est crucial. Cela permet de savoir si le réseau a bien les caractéristiques promises et cela permet de comparer les réseaux entre eux. D'où l'existence du groupe de travail LMAP <https://tools.ietf.org/wg/lmap> de l'IETF, dont voici le premier RFC. Il décrit deux études de cas où la mesure est nécessaire : un opérateur qui veut s'assurer de la qualité du service qu'il fournit, et un régulateur des télécommunications qui veut vérifier que les opérateurs livrent bien ce qu'ils promettent, et ne discriminent pas certaines utilisations. Il y a bien sûr d'autres cas possibles (l'utilisateur final qui veut mesurer les performances de **sa** connexion...)

Ce RFC concerne aussi bien les mesures de l'accès fixe (ADSL, par exemple) que de l'accès par un mobile. Il n'y a aucune différence dans les problèmatiques métrologiques entre ces deux accès.

La section 2 de notre RFC décrit les deux scénarios d'usage, repris ensuite en détail dans la section 3 pour le cas du FAI et dans la section 4 pour le cas du régulateur. Commençons par le FAI. La mesure de son réseau l'aide à :

- Identifier et isoler des problèmes (par exemple un lien saturé),
- Vérifier que des SLA sont respectés,
- Concevoir les extensions à son réseau (où faut-il déployer le nouveau matériel et les nouvelles fibres, en priorité),
- Comprendre le vécu des clients (ce qu'on nomme parfois la QoE - "*Quality of Experience*", qui ne dépend pas seulement de données brutes, comme la capacité des câbles, mais aussi de services comme les résolveurs DNS ou comme les CDN),
- Comprendre (voire prédire) l'impact de nouvelles technologies déployées dans le réseau. (Le RFC prend l'exemple d'IPv6, pour les opérateurs attardés qui ne l'ont pas encore déployé.) Cela peut aussi concerner des problèmes bien plus triviaux comme la mise à jour du "*firmware*" d'une machine qui entraine tout à coup une chute des performances <http://www.ietf.org/proceedings/85/slides/slides-85-iesg-opsandtech-7.pdf>.

Il y a plein de détails techniques à prendre en compte pour ce genre de mesures et le problème est bien plus complexe <http://www.afnic.fr/fr/ressources/blog/mesurer-la-qualite-de-l-acces-a-l-i
html> que ne le croit l'amateur qui débarque dans le domaine. Par exemple, le RFC note que la mesure du vécu de l'utilisateur nécessite des mesures de bout en bout, jusqu'à la machine de M. Michu, mais que le réseau de ce dernier, dans sa maison, n'est pas sous le contrôle du FAI (le Wi-Fi peut être peu efficace en raison de problèmes radio). Le FAI tend donc plutôt à mesurer jusqu'au CPE, ce qui est plus fiable mais moins représentatif.

La meilleure façon de mesurer les caractéristiques du réseau est de faire des mesures actives, avec une machine qui envoie des requêtes et mesure le résultat. Pour ne pas interférer avec le trafic normal de l'utilisateur, il vaut mieux faire ces mesures actives lorsque l'utilisateur ne fait rien. Cela implique que le dispositif de mesure puisse savoir quand l'utilisateur est inactif (la documentation de SamKnows <https://www.bortzmeyer.org/samknows.html> demande que l'engin soit placé en coupure sur le réseau local, lui permettant ainsi d'observer tout le trafic).

Des mesures actives lancées à la demande peuvent même être utilisées en réponse à une plainte d'un utilisateur. Il appelle le support en disant « Ça rame », l'employé au support clique sur un bouton sur son interface Web et cela lance des tests de performance depuis la "*box*" de cet abonné et, quelques secondes plus tard, leurs résultats s'affichent sur l'écran du support.

Une autre solution est de faire uniquement des mesures passives mais elles ne permettent d'observer que le débit, pas la capacité <https://www.bortzmeyer.org/capacite.html> (qui est le débit maximum possible). Si on mesure 2 Mb/s de trafic, est-ce parce que le réseau ne permet pas plus ou bien simplement parce que c'est ce que fait l'utilisateur en ce moment ? En outre, les mesures passives soulèvent des problèmes liés à la protection de la vie privée.

Le régulateur, lui, veut évaluer les performances de plusieurs opérateurs et, typiquement, les publier (ce qui nécessite de la rigueur dans les faits et de la pédagogie dans les explications). Cela peut servir par exemple à mesurer le déploiement de l'accès Internet à « haut débit » <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52010DC0245&from=EN>, ou bien à mesurer les pratiques de « gestion du trafic » (terme qui, dans le code des FAI, désigne les discriminations contre tel ou tel type de trafic). Dans tous les cas, vu les conséquences de la publication de ces mesures, le régulateur va devoir s'assurer qu'elles sont précises, significatives et reproductibles.

Il existe plusieurs façons de faire ces mesures : on peut utiliser par exemple un "*panel*" d'utilisateurs supposés représentatifs, et installer chez eux un équipement de mesure active, qui signale ce qu'il a obtenu à un serveur central. C'est ainsi que fonctionne SamKnows <https://www.bortzmeyer.org/samknows.html> (ou bien un autre système, non utilisé par le régulateur, les sondes Atlas <https://atlas.ripe.net/>). Une autre solution, non mentionnée par notre RFC, est de mettre les sondes de mesure, non pas chez les utilisateurs (où il est très difficile de s'assurer qu'elles sont branchées correctement, et où elles peuvent rentrer en concurrence avec le trafic normal de l'utilisateur) mais dans des locaux spécialisés, loués à cette fin. C'est ce que fait l'ARCEP pour ses mesures de l'accès à l'Internet en France <http://www.arcep.fr/index.php?id=10606>.

Une des motivations, pour le régulateur, est de déterminer les violations de la neutralité de l'Internet, afin de savoir s'il est utile de prendre des mesures légales ou autres, pour rétablir cette neutralité. Déterminer, par des mesures actives, si certains services réseau ou certaines destinations sont favorisés ou au contraire défavorisés n'est pas facile (cf. le projet Glasnost <http://mlab-live.appspot.com/tools/glasnost> et sa publication « "*Enabling End Users to Detect Traffic Differentiation*" <http://www.measurementlab.net/download/AMIfv945ljiJXzG-fgUrZSTu2hs1xRl5Oh-rpGQMWL305BNQh-BS
HRrnYU-aFd0rv332RDReRfOYkJuagysstN3GZ__lQHTS8_UHJTWkrwyqIUjffVeDxQ/> »). À l'inverse, détecter un blocage complet (par exemple la fermeture du port 25) est plus simple. Le sujet de

la neutralité du réseau est, à juste titre, très chaud, et cette demande d'une mesure du phénomène restera sans doute élevée (cf. les règles du BEREC <http://berec.europa.eu/eng/document_register/subject_matter/berec/download/0/1101-berec-guidelines-for-quality-of-service-_0.pdf>).

La section 5 de notre RFC se penche ensuite sur des détails pratiques dans la mise en œuvre de ces mesures. On va donc avoir N sondes qui font faire des mesures et un serveur central qui collectera les mesures (ce qui nécessitera un protocole de communication, de préférence sécurisé, entre les sondes et ce serveur). Les mesures seront répétées automatiquement et/ou pourront être activées à la demande (par exemple pour déboguer un problème précis). Une fois téléversés dans le serveur, les résultats des mesure devront être analysés (et ce sont parfois d'assez gros fichiers). L'analyse, quoique non couverte dans ce RFC, est également une source de risques. Comme dit l'adage « les chiffres, on leur fait dire ce qu'on veut ». Par exemple, le comité technique de l'ARCEP qui était chargé de superviser les décisions techniques pour le programme de mesure a eu à étudier des questions délicates comme la suppression de mesures aberrantes.

De nombreux choix pratiques vont ensuite se poser. Par exemple, la sonde active doit-elle être faite avec du matériel spécialisé (c'est le cas de SamKnows, de la mesure ARCEP, ou des RIPE Atlas) ou bien avec du logiciel que les utilisateurs téléchargeront (comme Grenouille <http://www.grenouille.com/> ou Netalyzr <http://netalyzr.icsi.berkeley.edu/>). Ces sondes spécialisées sont plus coûteuses, compliquées à déployer (on ne peut pas en mettre partout, il faut des critères de sélection, etc) mais plus fiables que la machine de M. Michu, probablement bourrée de virus qui la ralentissent, et qui n'est pas forcément allumée en permanence. (Une autre différence vient du fait que la sonde spécialisée peut être branchée immédiatement après le routeur, s'épargnant ainsi d'utiliser un LAN peu fiable et irrégulier.) Dans les deux cas, il faudra se demander si l'échantillon est vraiment représentatif ou si l'auto-sélection plus ou moins grande des utilisateurs a trop biaisé les résultats.

Enfin, la section 7 de notre RFC se penche sur les problèmes de sécurité, notamment de protection de la vie privée (RFC 6973 [1]). Parmi les risques, si on crée une infrastructure de mesures actives, on a fabriqué un gentil "*botnet*". Si un méchant peut en prendre le contrôle, il peut l'utiliser pour des attaques par déni de service réparties. (C'est une des raisons pour lesquelles l'exécution de mesures par les sondes Atlas, ouverte au public, est soumise à la dépense d'un certain nombre de crédits.) Toujours en cas de mesures actives, les sondes peuvent potentiellement balayer des réseaux internes qui ne seraient normalement pas accessibles de l'Internet (c'est pour cela que les Atlas ne peuvent pas tester les adresses du RFC 1918, car elles pourraient alors être utilisées pour la reconnaissance du réseau local où elles sont situées). Si les sondes ont des capacités de collecte passive, un méchant peut les utiliser comme « boîtes noires » d'espionnage des réseaux sur lesquels elles sont connectées. Il est donc nécessaire de ne pas fournir un accès illimité à ces sondes. Elle doivent n'être accessibles qu'à leur contrôleur, après authentification forte, et, si elle peuvent exécuter des mesures choisies par les utilisateurs (cas des sondes Atlas), ces mesures doivent être limitées (ainsi, les 1es Atlas ne peuvent pas lancer de téléchargement HTTP depuis un URL quelconque).

Dans le cas de mesures comparatives (« le FAI X est 20 % plus rapide que le FAI Y »), un autre risque de sécurité est celui de triche. Un attaquant qui contrôlerait, au moins partiellement, le système, risquerait de pouvoir fausser les résultats et favoriser un FAI par rapport aux autres. Là encore, de bonnes pratiques de sécurité sont nécessaires. Il y a enfin des attaques plus subtiles, comme d'identifier les lignes où des mesures sont faites, et les favoriser, obtenant ainsi de meilleurs résultats.

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc6973.txt