RFC 9490 : Report from the IAB Workshop on Management Techniques in Encrypted Networks (M-TEN)

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 9 mai 2024

Date de publication du RFC : Janvier 2024

https://www.bortzmeyer.org/9490.html

Aujourd'hui, l'essentiel du trafic sur l'Internet est chiffré, et pour d'excellentes raisons. Pas question de revenir là-dessus, mais, ceci dit, il va falloir adapter certaines pratiques de gestion des réseaux. L'IAB a organisé en 2022 un atelier <https://datatracker.ietf.org/group/mtenws/> sur ce sujet, dont ce RFC est le compte-rendu.

Comme les autres ateliers de l'IAB, il s'agit de faire s'exprimer diverses personnes, pas forcément d'arriver à une position commune (surtout sur un sujet...sensible). Tenu entièrement en ligne, cet atelier <https://datatracker.ietf.org/group/mtenws/> commençait par la soumission d'articles, qui devaient être lus d'abord, et discutés lors de l'atelier. La liste des articles figure dans l'annexe A du RFC, et les articles sont disponibles en ligne <https://datatracker.ietf.org/group/mtenws/materials/>.

D'abord, notre RFC réaffirme que la cryptographie est absolument nécessaire à la protection de la vie privée. Les difficultés réelles qu'elle peut poser ne doivent jamais être un prétexte pour réduire ou affaiblir le chiffrement. Qu'il y ait une tension <https://www.bortzmeyer.org/tussle-cyberspace.html> entre l'impératif du chiffrement et certaines méthodes de gestion du réseau, OK, mais la priorité reste la lutte contre la surveillance (RFC 7258 [1]).

Ensuite (section 1), le RFC pose le paysage : le trafic étant largement chiffré, des méthodes traditionnelles de gestion du réseau ne fonctionnent plus. tcpdump ne montre plus les données, on ne peut donc pas distinguer différentes méthodes HTTP, par exemple. Avec QUIC (RFC 9000), c'est même une partie de la couche transport qui est chiffrée donc un observateur extérieur ne peut plus, par exemple,

1. Pour voir le RFC de numéro NNN, https://www.ietf.org/rfc/rfcNNN.txt, par exemple https://www.ietf.org/rfc/rfc7258.txt

évaluer le RTT d'une session qu'il observe (ce qui était possible avec TCP). Il va donc falloir s'adapter notamment en coopérant davantage avec les applications qui, elles, ont accès à tout. Pour l'instant, on a vu surtout passer des récriminations d'acteurs habitués à surveiller le trafic et qui se plaignent que le chiffrement les en empêchent (le RFC 8404 est un bon exemple de ces récriminations). Au contraire, il faut chercher à résoudre une contradiction : permettre aux réseaux d'appliquer leur politique sans céder d'un millimètre sur le principe du chiffrement.

Outre le cas des opérateurs réseau qui veulent examiner le trafic, à des fins louables (améliorer le réseau) ou critiquables (défavoriser certains usages), on peut aussi citer, parmi les intermédiaires qui voudraient observer et/ou interférer avec le trafic réseau, les réseaux d'entreprise qui veulent, par exemple, empêcher les employés d'accéder à certains services, ou bien les mécanismes de contrôle des enfants (appelés à tort « contrôle parental », alors que leur but est justement de remplacer les parents).

Trois thèmes importants étaient discutés à l'atelier : la situation actuelle, les futures techniques et la façon dont on pourrait arriver à déployer ces futures techniques.

Pour la situation actuelle, on sait que, depuis longtemps, les administrateurices réseau comptent sur de l'observation passive, par exemple pour classer le trafic (tant de HTTP, tant de SSH, etc). C'est à la base de techniques comme IPFIX (RFC 7011), qui permet de produire de beaux camemberts. Outre la production de rapports, cette observation passive peut (mais n'est pas forcément) être utilisée pour prévoir les évolutions futures, prioriser certains types de trafic, détecter des comportements malveillants, etc. Les protocoles Internet étaient en effet traditionnellement trop bavards, faisant fuiter des informations vers des observateurs passifs (cf. le concept de « vue depuis le réseau », RFC 8546).

La lutte de l'épée et de la cuirasse étant éternelle, il y a évidemment eu des travaux sur des mécanismes pour empêcher ce genre d'analyses comme l'article soumis pour l'atelier « "*WAN Traffic Obfuscation at Line Rate*" <https://datatracker.ietf.org/doc/slides-interim-2022-mtenws-01-sessa-ditto-wa > ». (Au passage, sur cette idée d'"*obfuscation*", je recommande le livre de Brunton et Nissenbaum <https://www.bortzmeyer.org/obfuscation.html>.) Le RFC mentionne aussi une discussion sur l'idée d'exiger une permission explicite des utilisateurs pour les analyses du trafic (je ne vous dis pas les difficultés pratiques…). Voir l'"*Internet-Draft*" draft-irtf-pearg-safe-internet-measurement.

Dans un monde idéal, le réseau et les applications coopéreraient (et, dans un monde idéal, licornes et bisounours s'embrasseraient tous les jours) mais on voit plutôt une lutte, le réseau essayant d'en savoir plus, les applications de préserver leur intimité. Pourtant, le RFC note qu'une coopération pourrait être dans leur intérêt réciproque. Bien sûr, des applications comme Tor refuseraient certainement toute coopération puisque leur but est de laisser fuiter le moins d'information possible, mais d'autres applications pourraient être tentées, par exemple en échange d'informations du réseau sur la capacité disponible. Après tout, certaines techniques sont justement conçues pour cela, comme ECN (RFC 3168).

OK, maintenant, quelles sont les techniques et méthodes pour résoudre le problème de la gestion des réseaux chiffrés (section 2.2) ? Le RFC exprime bien le fait que le but du chiffrement est d'empêcher **tout** tiers d'accéder aux informations. Donc, sauf faille de sécurité, par exemple suite à une faiblesse du chiffrement, tout accès à l'information impose la coopération d'une des deux parties qui communiquent. Cette question du **consentement** est cruciale. Si on n'a pas le consentement d'une des deux extrémités de la communication, on n'a pas le droit d'accéder aux données, point. Cf. la contribution « "*What[Caractère Unicode non montré [2]]s In It For Me? Revisiting the reasons people collaborate*" <https://www.ietf.org/slides/slides-mtenws-paper-whats-in-it-for-me-revisiting-the-reasons-people-colla

2. Car trop difficile à faire afficher par LaTeX

pdf> ». Comme le répète souvent Eric Rescorla « Sur l'Internet, tout ce qui n'est pas une des extrémités est un ennemi. » (Cela inclut donc tous les équipements réseaux et les organisations qui les contrôlent.) Bref, les utilisateurs doivent pouvoir donner un consentement éclairé, et juger si les bénéfices de la visibilité l'emportent sur les risques. Évidemment, ce principe correct va poser de sérieuses difficultés d'application, car il n'est pas évident d'analyser proprement les bénéfices, et les risques (songez aux bandeaux cookies ou aux permissions des application sur votre ordiphone). Développer des nouveaux mécanismes de communication avec l'utilisateurice sont nécessaires, sinon, comme le note le RFC, « les ingénieurs devront choisir pour les utilisateurs ». Le RFC estime que les utilisateurs donneront toujours la priorité à leur vie privée (notamment parce qu'ils ne comprennent pas l'administration de réseaux et ses exigences), mais cela ne me semble pas si évident que cela. Personnellement, il me semble que le choix par défaut est crucial, car peu d'utilisateurices le modifieront.

Une des techniques qui permettent d'avoir accès à de l'information sans avoir toute l'information est celle des relais (cf. « "*Relying on Relays : The future of secure communication*" <https://www.ericsson.com/en/blog/2022/6/relays-and-online-user-privacy> ». Le principe est de séparer les fonctions : l'utilisateur parle à un relais qui parle au serveur final. Le relais ne connait pas le contenu de la demande, le serveur ne connait pas le demandeur. C'est utilisé dans des techniques comme "*Oblivious HTTP*" (RFC 9458) ou "*Oblivious DNS*" (RFC 9230). La préservation de la vie privée est donc bien meilleure qu'avec, par exemple, un VPN, où l'opérateur du VPN voit tout, le demandeur et la demande (contrairement à ce que prétendent les publicités pour NordVPN que de nombreux youtubeurs transmettent avec leurs vidéos). Le relais permet donc un accès à une information limitée, ce qui permet d'assurer certaines fonctions (comme le filtrage de contenu malveillant) sans avoir un accès complet.

Un exemple souvent cité par les opérateurs de l'intérêt d'accéder à la communication et de la modifier est celui de l'optimisation des flux TCP. Des PEP ("*Performance Enhancing Proxies*") violant le modèle en couches (car, normalement, TCP est de bout en bout) sont souvent déployés, notamment sur les liaisons satellites, dont les performances sont plus mauvaises, et les caractéristiques techniques très spéciales. (Cf. l'exposé de Nicolas Kuhn, « "*QUIC for satellite communications*" <https://www.afnic.fr/wp-media/uploads/2021/09/afnic-jcsa2021_cnes_kuhn.pdf> » à la Journée du Conseil Scientifique de l'Afnic <https://www.afnic.fr/observatoire-ressources/actualites/jcsa21-retour-su
> en 2021.) Le PEP va modifier les en-têtes TCP, voire parfois boucler la connexion TCP et en faire une autre lui-même. Cette technique ne marche évidemment plus avec QUIC, qui chiffre même la couche transport. Et elle mène à l'ossification de l'Internet puisqu'elle rend plus difficile de faire évoluer TCP, car toute évolution risque de perturber les PEP. QUIC avait en partie été explicitement développé pour empêcher ces intermédiaires de bricoler la session. Une autre approche est celle proposée dans « "*The Sidecar : "Opting in" to PEP Functions*" <https://www.ietf.org/slides/slides-mtenws-paper-the-sidecar-op
pdf> ».

Bon, maintenant, comment arriver à déployer de meilleures méthodes ? Dans un environnement fermé ou semi-fermé, cela doit être possible de faire en sorte que, par exemple, toutes les machines mettent en œuvre telle fonction (cf. RFC 8520 pour un exemple). Mais cela ne marche clairement pas pour l'Internet.

Parmi les moyens proposés de résolution de la contradiction entre visibilité par le réseau et protection de la vie privée, le RFC mentionne les "*Zero-Knowledge Middleboxes*". Décrites dans l'article "*du même nom*" <https://www.usenix.org/system/files/sec22fall_grubbs.pdf>, cette méthode utilise les preuves à divulgation nulle de connaissance pour prouver à l'intermédiaire qui fait le filtrage qu'on a bien respecté sa politique de filtrage, sans lui dire ce qu'on a fait. L'article détaille par exemple comment cela peut s'appliquer au DNS, qui est le principal outil de censure de l'Internet dans l'Union européenne. Ces preuves à divulgation nulle de connaissance ayant toujours été mystérieuses pour moi, je ne vous expliquerai pas comment elles marchent, mais notez que les auteurs ont fait un article pédagogique <https://blog.apnic.net/2022/07/08/unpacking-zero-knowledge-middleboxes/
>.

Enfin, le RFC mentionne la proposition "*Red Rover*", qui propose encore un arrangement bisounours <https://www.ietf.org/slides/slides-mtenws-paper-red-rover-a-collaborative-approach-t pdf> où les utilisateurs et les opérateurs collaboreraient pour un filtrage géré en commun. Les auteurs disent que ça marcherait car les utilisateurs « ne veulent probablement pas violer les CGU » (ben si : ils veulent utiliser Sci-Hub même si c'est censuré).

En conclusion, le RFC note en effet qu'on peut être sceptique quant aux chances d'une solution négociée. Mais il met aussi en avant le fait qu'il va être très difficile de trouver une solution qui marche pour toute la variété des réseaux. Et que l'expérience prouve largement que toute nouvelle technique va avoir des effets inattendus et que, par exemple, une solution qui visait à donner aux opérateurs des informations utiles pour la gestion de réseaux, va parfois être détournée pour d'autres buts.

Sinon, sur la question du déboguage des applications dans un monde où tout est chiffré, j'avais fait un exposé à Capitole du Libre <https://www.bortzmeyer.org/capitole-du-libre-2022.html>.