RFC 8404 : Effects of Pervasive Encryption on Operators

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 23 novembre 2018

Date de publication du RFC : Juillet 2018

https://www.bortzmeyer.org/8404.html

La vie privée sur l'Internet fait aujourd'hui l'objet d'innombrables attaques et l'une des techniques de défense les plus efficaces contre ces attaques est le chiffrement des données. Il protège également contre une autre menace, la modification des données en transit, comme le font certaines FAI (par exemple Orange Tunisie <https://twitter.com/SarhanTN/status/1022462456482996224>). Il y a de nombreuses campagnes de sensibilisation pour promouvoir le chiffrement (voir par exemple le RFC 7258 [1]), avec des bons résultats. Évidemment, ce chiffrement gène ceux qui voudraient espionner et modifier le trafic, et il fallait donc s'attendre à voir une réaction. Ce RFC est l'expression de cette réaction, du côté de certains opérateurs réseau, qui regrettent que le chiffrement empêche leurs mauvaises pratiques.

Dès le début, ce RFC était basé sur une hypocrisie : prétendre être purement descriptif (une liste de pratiques que certains opérateurs réseau utilisent, et qui sont impactées par le chiffrement), sans forcément en conclure que le chiffrement était mauvais. Mais, en réalité, le RFC porte un message souvent répété : le chiffrement gène, et il ne faudrait pas en abuser. On note par exemple que le RFC ne s'indigne pas de certaines des pratiques citées, alors que beaucoup sont scandaleuses. Le discours est « il faut trouver un équilibre entre la protection de la vie privée et la capacité des opérateurs à gérer leurs réseaux ». Une telle référence à l'équilibre m'a toujours énervé : non, il ne faut pas d'équilibre entre le bien et le mal, il faut faire ce qu'on peut pour gêner la surveillance. Certaines des pratiques citées dans le RFC sont mauvaises et tant mieux si le chiffrement les rend difficiles. Certaines autres sont plus neutres, mais devront quand même s'adapter, les révélations de Snowden ont largement montré que la surveillance de masse est un fait, et qu'elle justifie des mesures radicales de protection. (Notez que les premières versions de ce RFC étaient bien pires, et qu'il a été affadi par les discussions successives, et suite à l'opposition rencontrée.)

Hypocritement, les textes sacrés comme le RFC 1958, RFC 1984, RFC 2804, et bien sûr les RFC 7258 et RFC 7624 sont cités, hommage du vice à la vertu. Mais tout en le faisant, ce nouveau RFC 8404

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

déforme ces textes. Quand le RFC 7258 dit qu'évidemment, il faut que les opérateurs puissent continuer à gérer leurs réseaux, ce RFC 8404 lui fait dire qu'il faut donc limiter le chiffrement, comme si « gérer un réseau » voulait forcément dire « accéder aux communications ». De même, on trouve une référence au principe de bout en bout (RFC 2775, RFC 3724, RFC 7754), alors même que le reste de ce RFC ne fait que citer des pratiques violant ce principe.

Ce RFC 8404 se veut, on l'a dit, une description de pratiques actuelles, et prétend hypocritement « ne pas forcément soutenir toutes les pratiques présentées ici ». En effet, tous les exemples cités par la suite sont réellement utilisées mais pas forcément par tous les opérateurs. C'est une des choses les plus désagréables de ce RFC que de laisser entendre que tous les opérateurs seraient unanimes dans leur désir de lire les communications de leurs clients, voire de les modifier. (Par exemple en utilisant le terme tribal de "*community*" pour parler des opérateurs réseau, comme s'ils étaient un clan unique, d'accord sur l'essentiel.)

Il est exact qu'au début de l'Internet, rien ou presque n'était chiffré. Tout circulait en clair, en partie parce que la perte de performances semblait excessive, en partie parce que les États faisaient tout pour décourager le chiffrement et limiter son usage, en partie parce que la menace semblait lointaine (même les plus paranoïaques des experts en sécurité ne se doutaient pas que la surveillance atteignait les proportions révélées par Snowden.) Certains opérateurs ont alors pris de mauvaises habitudes, en examinant en détail les communications, à des fins plus ou moins innocentes. Les choses changent aujourd'hui, avec un déploiement plus important du chiffrement, et ces mauvaises pratiques doivent donc disparaitre, ce que les auteurs du RFC ont du mal à avaler.

Des premiers exemples de ces mauvaises pratiques sont données en section 1.2. Par exemple, le RFC cite un rapport de l'EFF <https://www.eff.org/deeplinks/2014/11/starttls-downgrade-attacks> montrant que certains opérateurs inséraient des données dans les flux SMTP, notamment à des fins de suivi des utilisateurs, et, pour empêcher ces flux d'être protégés par TLS, retiraient la commande `STARTTLS` de la négociation (ce qu'on nomme le "*SSL stripping*", cf. RFC 7525, section 3.2). Le RFC ose même dire que pour certains, c'était « considéré comme une attaque » comme si cela n'était pas évident que c'est une attaque ! Notant que le chiffrement systématique empêcherait cette attaque, le RFC demande qu'on fournisse à ces opérateurs réseau malhonnêtes une alternative à leurs pratiques de piratage des sessions !

Un autre exemple donné est celui des réseaux d'entreprise où il y a, dit le RFC, un « accord » pour que le contenu du trafic des employés soit surveillé par le patron (dans la réalité, il n'y a pas d'accord, cette surveillance est une condition de l'embauche). Ce serait donc un cas radicalement différent de celui des réseaux publics, via le FAI, où il n'y a pas de tel « accord » et, pourtant, déplore le RFC, le chiffrement rend cette surveillance plus difficile. (Ce point particulier a fait l'objet de nombreux débats lors de la mise au point de la version 1.3 du protocole TLS, qui visait en effet à rendre la surveillance plus difficile. Cf. RFC 8446.)

À partir de la section 2, commence le gros du RFC, les études de cas. Le RFC est très détaillé et je ne vais pas tout mentionner ici. Notons que la section 2 commence par du chantage : si on ne donne pas accès aux opérateurs au contenu des communications, ils « vont déployer des méthodes regrettables du point de vue de la sécurité », et il faut donc leur donner les données, pour éviter qu'ils n'utilisent des moyens encore pires.

D'abord, ce sont les chercheurs qui effectuent des mesures sur l'Internet (par exemple CAIDA <http://www.caida.org/>) qui sont mentionnés pour expliquer que le chiffrement va rendre ces études bien plus difficiles. (Au passage, puisque le RFC ne le dit pas, signalons que la science ne justifie pas tout et que ces études doivent se conformer à l'éthique, comme tout le monde <https://labs.ripe.net/Members/mirjam/ethics-in-network-measurements>.) Le RFC oublie de dire que beaucoup

de ces études n'utilisent que l'en-tête IP, et parfois les en-têtes TCP et UDP, qui ne sont pas affectés par le chiffrement. Comme cette réthorique est souvent présente dans ce RFC, cela vaut la peine de préciser : les techniques de chiffrement les plus déployées aujourd'hui (TLS et, loin derrière, SSH) laissent intacts les en-têtes IP et TCP, que les chercheurs peuvent donc regarder comme avant. IPsec, lui, masque l'en-tête TCP (et, dans certains cas, une partie de l'information venue de l'en-tête IP), mais il est très peu utilisé dans l'Internet public (à part une partie de l'accès au VPN de l'entreprise). QUIC, s'il sera un jour déployé massivement, dissimulera également une partie de l'information de couche 4. Bref, aujourd'hui, la partie en clair du trafic donne déjà plein d'information.

Après les chercheurs, le RFC cite les opérateurs qui regardent en détail un flux réseau pour déboguer des problèmes applicatifs. Là encore, l'exemple est très malhonnête :

— Comme indiqué ci-dessus, la partie non chiffrée du paquet reste accessible. Si on veut déboguer des problèmes TCP subtils, du genre la fenêtre qui ne s'ouvre pas assez, même si TLS est utilisé, l'information TCP complète reste disponible, avec la taille de la fenêtre, les numéros de séquence et tout le toutim.
— Mais, surtout, quel opérateur réseau débogue ainsi les problèmes applicatifs de ces clients ? Vous croyez vraiment que M. Michu, client d'un FAI grand public, va pouvoir solliciter l'aide d'un expert qui va lancer Wireshark pour résoudre les problèmes WebRTC de M. Michu ? (C'est l'exemple donné par le RFC : « ma vidéo haute définition est hachée ». S'il existe un FAI dont le support accepte de traiter ce genre de problèmes, je veux bien son nom.)
— Dans des discussions privées, on m'a parfois dit qu'il ne s'agissait pas de FAI grand public, qui ne passeront en effet jamais du temps sur ces problèmes, mais d'opérateurs ayant une clientèle de grandes entreprises, pour lesquelles ils jouent un rôle de conseil et d'assistance sur les problèmes réseaux. Mais, si le client est volontaire, et veut vraiment qu'on l'aide, rien ne l'empêche de couper le chiffrement ou de demander aux applications d'afficher de l'information de déboguage détaillée.

Le RFC ne mentionne pas assez que le développement du chiffrement est en bonne partie le résultat des pratiques déplorables de certains opérateurs. Si on fait tout passer sur le port 443 (celui de HTTPS), et en chiffré, c'est précisément parce que des opérateurs se permettaient, en violation du principe de neutralité <https://www.bortzmeyer.org/neutralite.html>, de traiter différemment les applications, se fiant au port TCP utilisé pour reconnaitre les différentes applications. Refusant de comprendre ce problème, le RFC déplore au contraire que « tout le monde étant sur le port 443, on ne peut plus détecter certains applications afin de les prioriser », et affirme que cela justifie le DPI. Pas un seul questionnement sur la légitimité de ce traitement différencié. (Notez également un truc réthorique très malhonnête : parler de prioriser certaines applications. Comme la capacité <https://www.bortzmeyer.org/capacite.html> du réseau est finie, si on en priorise certaines applications, on en ralentit forcément d'autres. Mais c'est plus vendeur de dire qu'on priorise que d'admettre qu'on discrimine.) Pour éviter de dire ouvertement qu'on viole la neutralité du réseau, le RFC la redéfinit en disant qu'il n'y a pas violation de la neutralité si on différencie entre applications, seulement si on différencie entre utilisateurs.

Il n'y a évidemment pas que les opérateurs réseau qui sont responsables de la surveillance et de l'interférence avec les communications. La section 2.4 du RFC rappelle à juste titre que les opérateurs travaillent dans un certain cadre légal et que l'État les oblige souvent à espionner (ce qui s'appelle en novlangue « interception légale ») et à censurer. Par exemple, en France, les opérateurs sont censés empêcher l'accès aux sites Web présents sur la liste noire (secrète) du ministère de l'intérieur. Le chiffrement rend évidemment plus difficile ces activités (c'est bien son but !) Le RFC note par exemple que le filtrage par DNS menteur, la technique de censure la plus commune en Europe, est plus difficile si on utilise le chiffrement du RFC 7858 pour parler à un résolveur DNS externe, et traite cela comme si c'était un problème, alors que c'est au contraire le résultat attendu (s'il n'y avait pas de surveillance, et pas d'interférence avec le trafic réseau, le chiffrement serait inutile).

Mais le RFC mélange tout par la suite, en citant comme exemple de filtrage légitime le contrôle parental. Il n'y a nul besoin qu'il soit fait par le FAI, puisqu'il est souhaité par le client, il peut être fait

sur les machines terminales, ou dans le routeur de la maison. C'est par exemple ce qui est fait par les bloqueurs de publicité, qui ne sont pas une violation de la neutralité du réseau puisqu'ils ne sont pas dans le réseau, mais sur la machine de l'utilisateur, contrôlée par elle ou lui.

Un autre exemple où le RFC déplore que le chiffrement empêche une pratique considérée par l'opérateur comme utile est celui du "*zero rating*", cette pratique où l'accès à certains services fait l'objet d'une tarification réduite ou nulle. Le chiffrement, note le RFC, peut servir à dissimuler les services auxquels on accède. Là encore, on peut se demander si c'est vraiment un problème : le "*zero rating*" est également une violation de la neutralité du réseau, et ce n'est pas forcément mauvais qu'il devienne plus difficile à déployer.

Mais l'exemple suivant est bien pire, et beaucoup plus net. La section 2.3.4 du RFC décrit la pratique de modification des flux HTTP, notamment l'insertion d'en-têtes. Plusieurs vendeurs de matériel fournissent des équipements permettant d'« enrichir » les en-têtes, en ajoutant des informations personnelles <https://www.bortzmeyer.org/enrichir-qui.html>, comme le numéro de téléphone, permettant un meilleur ciblage par les publicitaires. Et ils n'en ont même pas honte, annonçant cyniquement cette capacité dans leurs brochures. Et le RFC décrit cette pratique en termes neutres, sans préciser qu'elle est mauvaise, devrait être interdite, et que le chiffrement permet heureusement de la rendre plus difficile ! La soi-disant neutralité de la description mène donc à décrire des pratiques scandaleuses au même niveau que l'activité normale d'un opérateur réseaux. Le RFC cite bien le RFC 8165, qui condamne cette pratique, mais n'en tire pas les conséquences, demandant, là encore, qu'on fournisse aux opérateurs malhonnêtes un moyen de continuer ces opérations (par exemple par une insertion des données personnelles directement par le logiciel sur la machine du client, ce qui est possible dans le monde du mobile, beaucoup plus fermé et contrôlé), et ce malgré le chiffrement.

La section 3.1.2 du RFC revient sur les buts de la surveillance, lorsqu'on gère un réseau. Elle mentionne plusieurs de ces buts, par exemple :

— Détection de logiciel malveillant. À noter que lui, de toute façon, ne va pas se laisser arrêter par une interdiction du chiffrement.
— Conformité avec des règles imposant la surveillance, qui existent par exemple dans le secteur bancaire. On pourrait garder le chiffrement de bout en bout tout en respectant ces règles, si chaque application enregistrait toutes ses communications. C'est évidemment la bonne solution technique (section 3.2.1 du RFC) mais, à l'heure actuelle, beaucoup de ces applications sont fermées, n'enregistrent pas, et ne permettent pas de traçabilité. (Notons aussi que des normes comme PCI imposent la traçabilité, mais n'imposent pas d'utiliser du logiciel libre, ce qui est assez incohérent. La quadrature du cercle - sécuriser les communications contre l'écoute, tout en permettant l'audit - n'est soluble qu'avec des applications qu'on peut examiner, et qui enregistrent ce qu'elles font.)
— Opposition à la fuite de données. La « sécurité » souvent mentionnée est en fait un terme vague recouvrant beaucoup de choses. La sécurité de qui et de quoi, faut-il se demander. Ici, il est clair que le patronat et l'État voudraient rendre plus difficile le travail de lanceurs d'alerte, par exemple WikiLeaks. En France, une loi sur « le secret des affaires <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000037262111&dateTexte=&categorieLien=id> » fournit aux entreprises qui veulent dissimuler leurs mauvaises actions une défense juridique contre la vérité. La limitation du chiffrement vise à ajouter une défense technique, pour empêcher le ou la lanceur d'alerte d'envoyer des informations discrètement au Canard enchaîné.

La section 3.2 du RFC mentionne d'autres exemples d'usages actuels que le chiffrement massif peut gêner. Par exemple, elle cite la lutte contre le spam en notant que le chiffrement du courrier (par exemple avec PGP ou DarkMail) empêche un intermémdiaire (comme le serveur de messagerie) de lutter contre le spam. (La bonne solution est évidemment de faire l'analyse du spam sur la machine finale, par exemple avec bogofilter.)

Les révélations de Snowden ont montré que l'espionnage des liaisons **internes** aux organisations était une pratique courante de la NSA. Pendant longtemps, beaucoup d'administrateurs réseau ont

considéré que « pas besoin d'utiliser HTTPS, c'est un site Web purement interne ». C'était une grosse erreur technique (l'attaquant est souvent interne, et les attaques contre le DNS ou le routage peuvent donner accès aux communications censées être internes) et elle a donc logiquement été exploitée. Aujourd'hui, les organisations sérieuses chiffrent systématiquement, même en interne.

Mais en même temps, beaucoup d'entreprises ont en interne des règles qui autorisent la surveillance de toutes les communications des employés. (En France, il existe des limites à cette surveillance <https://www.numerama.com/politique/286579-la-cedh-fixe-les-conditions-a-respecter-pour-surveill html>, mais pas aux États-Unis. Le RFC rappelle juste que certaines politiques définies par le patron autorisent des accès protégés contre la surveillance, par exemple pour se connecter à sa banque. Écrit avec un point de vue très états-unien, le RFC ne mentionne que les politiques de la direction, pas les limites légales.) Techniquement, cela passe, par exemple, par des relais qui terminent la session TLS, examinent le contenu de la communication, puis démarrent une nouvelle session avec le serveur visé. Cela permet par exemple de détecter le logiciel malveillant que la machine Windows télécharge via une page Web piégée. Ou bien une attaque XSS que l'utilisateur n'a pas vue. (Notez que, par une incohérence fréquente en entreprise, la sécurité est présentée comme absolument essentielle, mais les postes de travail tournent sur Windows, certainement la cible la plus fréquente de ces logiciels malveillants.) Mais cela permet aussi de surveiller les communications des employés.

Si le but est de se protéger contre le logiciel malveillant, une solution simple est de faire l'examen des messages suspects sur la machine terminale. Cela peut être plus compliqué à déployer mais, normalement, c'est faisable à la fois juridiquement puisque cette machine appartient à l'entreprise, et techniquement puisqu'elle est a priori gérée à distance.

La section 5 du RFC détaille certaines techniques de surveillance pour des systèmes particuliers. Par exemple, les fournisseurs de messagerie examinent les messages, notamment pour déterminer s'il s'agit de spam. SMTP sur TLS ne gêne pas cette technique puisqu'il est de serveur à serveur, pas de bout en bout. Même PGP ou S/MIME n'empêchent pas l'examen des métadonnées, qui peuvent être suffisante pour détecter le spam.

De même (section 6 du RFC), TLS ne masque pas tout et permet quand même certaines activités de surveillance. Ainsi, le nom du serveur contacté, indiqué dans l'extension TLS SNI ("*Server Name Indication*", section 3 du RFC 6066) n'est pas chiffré et peut donc être vu. (Des efforts sont en cours actuellement à l'IETF pour boucher cette faille dans la protection de la vie privée, et semblent bien avancer.)

En résumé, ce RFC, qui a commencé comme très « anti-chiffrement » a été un peu modifié lors des très vives discussions qu'il a généré à l'IETF, et essaie désormais de ménager la chèvre et le chou. La tonalité demeure quand même hostile au chiffrement sérieux, mais on y trouve suffisamment d'éléments techniques précis pour que tout le monde puisse y trouver une lecture intéressante.