

# RFC 8963 : Evaluation of a Sample of RFCs Produced in 2018

Stéphane Bortzmeyer  
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 23 janvier 2021

Date de publication du RFC : Janvier 2021

<https://www.bortzmeyer.org/8963.html>

---

On raconte souvent que la publication d'un RFC prend longtemps, très longtemps, trop longtemps, et on dit parfois que cela est dû à une bureaucratie excessive dans le processus de publication. Peut-on objectiver cette opinion ? Dans ce RFC, l'auteur étudie le trajet de plusieurs RFC (pris au hasard) publiés en 2018 et regarde où étaient les goulets d'étranglement, et ce qui a pris du temps. Les RFC étudiés ont mis en moyenne trois ans et quatre mois à être publiés, depuis le premier document écrit jusqu'à la publication finale. (Mais il y a de grandes variations.) La majorité du temps est passée dans le groupe de travail, ce ne sont donc pas forcément les étapes « bureaucratiques » qui retardent.

Heureusement pour cette étude, le "*DataTracker*" <<https://datatracker.ietf.org/>> enregistre les différentes étapes de la vie d'un document et permet donc au chercheur de travailler sur des données solides. Mais attention, si les données sont fiables, leur interprétation est toujours délicate. Et, si une métrique devient un objectif, elle cesse d'être une bonne métrique. Si l'IETF ne cherchait qu'à raccourcir le délai de publication, ce serait simple, il suffirait d'approuver sans discussion et immédiatement toutes les propositions, même mauvaises. (Ce qu'on nomme le "*rubber stamping*" en anglais.)

La section 2 expose la méthodologie suivie. D'abord, déterminer les dates importantes dans la vie d'un RFC. Tous les RFC ne suivent pas le même parcours, loin de là. Le cas typique est celui d'un document individuel, qui est ensuite adopté par un groupe de travail, approuvé par l'IESG, puis publié par le "*RFC Editor*" <<https://www.rfc-editor.org/>>. Dans cette analyse, trois périodes sont prises en compte :

- Le temps de traitement dans le groupe de travail, qui part de la première publication et s'arrête au dernier appel à l'IETF,
- Le temps de traitement par l'IETF, qui part du dernier appel à l'IETF et se termine par l'approbation par l'IESG,
- Le temps de traitement par le "*RFC Editor*", qui part de l'approbation par l'IESG et se termine par la publication en RFC.

Ainsi, pour le dernier RFC publié au moment où j'écris cet article (et qui n'est donc pas pris en compte dans l'analyse, qui couvre des RFC de 2018), pour ce RFC 9003<sup>1</sup>, on voit sur le DataTracker <<https://datatracker.ietf.org/doc/rfc9003/>> qu'il a été documenté pour la première fois en décembre 2017, adopté par un groupe de travail en avril 2018, soumis à l'IESG en avril 2020, a eu un "IETF Last Call" en juin 2020, a été approuvé par l'IESG en novembre 2020 et enfin publié en RFC début janvier 2021. Mais beaucoup d'autres RFC suivent un autre parcours. Pour la voie indépendante (décrite dans le RFC 4846), l'analyse de ce RFC 8963 considère l'approbation par l'ISE ("Independent Submission Editor") comme terminant la première période et l'envoi au "RFC Editor" comme terminant la deuxième.

Les vingt RFC qui sont utilisés dans l'étude ont été pris au hasard parmi ceux publiés en 2018. Le RFC reconnaît que c'est un échantillon bien petit mais c'est parce que les données disponibles ne fournissaient pas toutes les informations nécessaires et qu'il fallait étudier « manuellement » chaque RFC, ce qui limitait forcément leur nombre. Je ne vais pas lister ici tous les RFC étudiés, juste en citer quelques uns que je trouve intéressants à un titre ou l'autre :

- Le RFC 8446 est un très gros morceau. C'est la nouvelle version d'un protocole absolument critique, TLS. Non seulement le RFC est gros et complexe, mais le souci de sécurité a entraîné des études particulièrement poussées. Et, pour couronner le tout, ce RFC a été au cœur de polémiques politiques, certains reprochant à cette version de TLS d'être trop efficace, d'empêcher d'accéder aux données. La phase finale chez le "RFC Editor", dite « AUTH48 » (car elle se fait normalement en 48 heures...) n'a pas pris moins de deux mois ! C'est en partie dû à une nouveauté de l'époque, utiliser GitHub pour suivre les changements à apporter. Pour chaque RFC étudié, l'auteur note également son succès ou insuccès en terme de déploiement et TLS 1.3 est un succès massif, avec des dizaines de mises en œuvre différentes et une présence très forte sur l'Internet.
- Le RFC 8324 est un cas très différent. Consacré à une série d'opinions sur le DNS et ses évolutions récentes, il ne s'agit pas, contrairement au précédent, d'une norme IETF, mais d'une contribution individuelle. N'engageant que son auteur, elle ne nécessite pas le même niveau d'examen. C'est donc ce RFC qui est le titulaire du record dans l'étude : neuf mois seulement entre la première publication et le RFC.
- Le RFC 8429 avait pour but de documenter l'abandon de plusieurs protocoles de cryptographie désormais dépassés. A priori très simple et très consensuel, il a été retardé par une discussion plus bureaucratique à propos de ses conséquences exactes sur le RFC 4757, puis par un délai très long d'un des auteurs à répondre lors de la phase finale (« AUTH48 » a pris trois mois).
- Le RFC 8312 portait aussi sur un sujet consensuel, l'algorithme de lutte contre la congestion CUBIC étant aujourd'hui largement adopté. Mais il a quand même pris des années, en partie car, justement, CUBIC étant consensuel et largement déployé, il ne semblait pas crucial de le documenter. (Rappelez-vous que le travail à l'IETF est fait par des volontaires.)
- Le RFC 8492 a également pris longtemps car c'était un projet individuel, dont l'auteur n'avait pas le temps de s'occuper, ce qui explique des longs délais à certains points. En outre, il a été retardé par la nécessité d'attendre un changement de politique d'enregistrements dans certains registres à l'IANA, changements qui dépendaient de TLS 1.3, qui prenait du temps.
- Un autre RFC qui a souffert du fait que même ses auteurs ne le considéraient pas comme urgent a été le RFC 8378, un projet très expérimental et dont la publication en RFC a pris plus de quatre ans.
- Un des buts de l'étude était de voir s'il y avait une corrélation entre le délai total et certaines caractéristiques du RFC, par exemple sa longueur. Mais le RFC 8472, quoique simple, a pris plus de trois ans, en partie pour un problème bureaucratique sur le meilleur groupe de travail à utiliser.
- En revanche, le RFC 8362 n'a pas de mal à expliquer le temps qu'il a pris : il modifie sérieusement le protocole OSPF et nécessite des mécanismes de transition depuis les anciennes versions, qui ont eux-même pris du temps.

---

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc9003.txt>

Bien, après cette sélection d'un échantillon de RFC, qu'observe-t-on (section 4 de notre RFC)? Le délai moyen entre le premier "*Internet Draft*" et la publication en RFC est de trois ans et trois mois. Mais les variations sont telles que cette moyenne n'a guère de sens (c'est souvent le cas avec la moyenne <<https://www.bortzmeyer.org/mediane-et-moyenne.html>>). Le plus malheureux est le RFC 8492, six ans et demi au total, mais ce cas n'est pas forcément généralisable. Pour les RFC de l'IETF, où le délai moyen est presque le même, le temps passé dans le groupe de travail est de deux ans et neuf mois, à l'IESG de trois à quatre mois et à peu près pareil chez le "*RFC Editor*". Bref, 80 % du temps se passe en « vrai » travail, dans le groupe, les passages plus « bureaucratiques » n'occupent que 20 % du temps et ne sont donc pas les principaux responsables des délais.

L'étude regarde aussi des échantillons de RFC publiés en 2008 et en 1998 (en utilisant cette fois, à juste titre, des valeurs médianes) et note que le délai était à peu près le même en 2008 mais qu'il était trois à quatre fois plus court en 1998.

Et la partie purement éditoriale, chez le "*RFC Editor*" ? On a vu qu'elle prenait de trois à quatre mois. C'est quatre en moyenne mais si on exclut l'extrême RFC 8492, c'est plutôt trois mois, ce que je trouve plutôt court pour relire et préparer des documents longs et complexes, où une coquille peut être très gênante (ce ne sont pas des romans!) Quelles sont les caractéristiques des RFC que l'on pourrait corrélérer à ce délai chez le "*RFC Editor*" ? L'étude montre une corrélation entre le délai et la longueur du document, ce qui n'est pas surprenant. Même chose pour la corrélation entre le nombre de changements faits depuis l'approbation par l'IESG et la publication. (Une des tâches du "*RFC Editor*" est de faire en sorte que les RFC soient en bon anglais alors que beaucoup d'auteurs ne sont pas anglophones.) Plus étonnant, le nombre d'auteurs ne diminue pas le délai, au contraire, il l'augmente. Un nouveau cas de « quand on ajoute des gens au projet, le projet n'avance pas plus vite » ? Il ne faut pas oublier qu'il existe une zone grise entre « changement purement éditorial », où le "*RFC Editor*" devrait décider, et « changement qui a des conséquences techniques » et où l'approbation des auteurs est cruciale. Il n'est pas toujours facile de s'assurer de la nature d'un changement proposé par le "*RFC Editor*". Et, si le changement a des conséquences techniques, il faut mettre les auteurs d'accord entre eux, ce qui est long (ils peuvent être dans des fuseaux horaires différents, avec des obligations professionnelles différentes) et peut soulever des désaccords de fond.

Et pour la voie indépendante ? Trois des RFC de l'échantillon étaient sur cette voie. Ils ont été publiés plus vite que sur la voie IETF, ce qui est logique puisqu'ils ne nécessitent pas le consensus de l'IETF (voir le RFC 8789). En revanche, le fait que la fonction d'ISE ("*Independant Stream Editor*", cf. RFC 8730) soit une tâche assurée par un seul volontaire peut entraîner des retards ce qui a été le cas, ironiquement, de ce RFC 8963.

L'étude se penche aussi sur le nombre de citations dont bénéficie chaque RFC (section 5) car elle peut donner une idée sommaire de leur impact. Elle a utilisé l'API de Semantic Scholar. Par exemple, pour le RFC 8483, publié l'année de l'étude mais pas inclus dans l'échantillon :

```
% curl -s https://api.semanticscholar.org/v1/paper/10.17487/rfc8483?include_unknown_references=true \
jq .
```

On voit qu'il est cité une fois, précisément par ce RFC. En revanche, le RFC 8446, sur TLS 1.3 est cité pas moins de 773 fois :

```
% curl -s https://api.semanticscholar.org/v1/paper/10.17487/rfc8446?include_unknown_references=true | \
jq .citations|length
773
```

C'est de très loin le plus « populaire » dans le corpus (l'échantillon étudié). C'est bien sûr en partie en raison de l'importance critique de TLS, mais c'est aussi parce que TLS 1.3 a été étudié par des chercheurs et fait l'objet d'une validation formelle, un sujet qui génère beaucoup d'articles dans le monde académique. C'est aussi le cas du RFC 8312, sur Cubic, car la congestion est également un sujet de recherche reconnu. Notez que le passage du temps est impitoyable : les RFC des échantillons de 2008 et 1998 ne sont quasiment plus cités (sauf le RFC 2267, qui résiste). Est-ce que le nombre de citations est corrélé à l'ampleur du déploiement ? Dans le cas des RFC 8446 et RFC 8312, on pourrait le croire, mais le RFC 8441 a également un très vaste déploiement, et quasiment aucune citation. À l'inverse, un RFC qui a un nombre de citations record, le RFC 5326, n'a quasiment pas eu de déploiement, mais a excité beaucoup de chercheurs académiques (et même médiatique, vu son sujet cool), ce qui explique le nombre de citations. Bref, le monde de la recherche et celui des réseaux n'ont pas des jugements identiques.

On avait noté que le RFC 2267 était beaucoup cité. Or, il n'est plus d'actualité, ayant officiellement été remplacé par le RFC 2827. Mais les citations privilégient la première occurrence, et ne reflètent donc pas l'état actuel de la normalisation.

Bon, on a vu que les citations dans le monde de la recherche ne reflétaient pas réellement l'« importance » d'un RFC. Et le nombre d'occurrences dans un moteur de recherche, c'est mieux ? L'auteur a donc compté le nombre de références aux RFC dans Google et Bing. Il y a quelques pièges. Par exemple, en cherchant « RFC8441 », Bing trouve beaucoup de références à ce RFC 8441, qui parle de WebSockets, mais c'est parce qu'il compte des numéros de téléphone se terminant par 8441, ou des adresses postales « 8441 Main Street ». Google n'a pas cette amusante bogue. Ceci dit, il est noyé par le nombre très élevé de miroirs des RFC, puisque l'IETF permet et encourage ces copies multiples (contrairement aux organisations de normalisation fermées comme l'AFNOR). Ceci dit, les RFC 8441 et RFC 8471 sont plus souvent cités, notamment car ils apparaissent fréquemment dans les commentaires du code source d'un logiciel qui les met en œuvre. Le « test Google » est donc assez pertinent pour évaluer le déploiement effectif.

Quelles conclusions en tirer (section 6 du RFC) ? D'abord, si on trouve le délai de publication trop long, il faut se focaliser sur le temps passé dans le groupe de travail, car c'est clairement le principal contributeur au délai. [En même temps, c'est logique, c'est là où a lieu le « vrai » travail.] Malheureusement, c'est là où le *"DataTracker"* contient le moins d'information alors que les phases « bureaucratiques » comme le temps passé à l'IESG bénéficient d'une traçabilité plus détaillée. Ainsi, on ne voit pas les derniers appels (*"Last Calls"*) internes aux groupes de travail. Bien sûr, on pourrait rajouter ces informations mais attention, cela ferait une charge de travail supplémentaire pour les présidents des groupes de travail (des volontaires). Récolter des données pour faire des études, c'est bien, mais cela ne doit pas mener à encore plus de bureaucratie.

Du fait que l'étude ne porte que sur des RFC effectivement publiés, elle pourrait être accusée de sensibilité au biais du survivant. Peut-être que les RFC qui n'ont pas été publiés auraient aussi des choses à nous dire ? L'abandon d'un projet de normalisation est fréquent (pensons par exemple à DBOUND <<https://www.bortzmeyer.org/fin-dbound.html>> ou RPZ <<https://www.bortzmeyer.org/echec-rpz.html>>) et parfaitement normal et souhaitable ; toutes les idées ne sont pas bonnes et le travail de tri est une bonne chose. Mais, en n'étudiant que les RFC publiés, on ne peut pas savoir si le tri n'a pas été trop violent et si de bonnes idées n'en ont pas été victimes. L'IETF ne documente pas les abandons et une étude de *"drafts"* sans successeur pourrait être intéressante.

Pour terminer, je me suis amusé à regarder ce qu'il en était des trois RFC que j'ai écrit :

- Le RFC 7626 a été publié moins de deux ans après le premier document <<https://datatracker.ietf.org/doc/rfc7626/>> ce qui, on l'a vu, est inférieur à la moyenne. Fortement poussé par l'engagement de l'IETF <<https://www.bortzmeyer.org/ietf-securite-espionnage-bis.html>> à agir après les révélations de Snowden, il n'a guère connu d'obstacles.

- Le RFC 7816 a pris deux ans. Il a passé à peine plus d'un mois chez le "*RFC Editor*", ce qui est très peu. Comme le précédent, il faisait partie d'un projet porteur, visant à améliorer le respect de la vie privée dans le DNS. Il a connu des objections techniques mais son statut « *Expérimental* » ne nécessite pas une solution parfaite. Une partie du temps a été passée dans un problème dont le RFC 8963 ne parle pas : une histoire de brevets et de négociations juridico-politiques.
- Le RFC 8020 est le seul des trois à être sur le chemin des normes, en théorie bien plus exigeant. Mais il bat tous les records en n'ayant mis qu'un an à être publié. C'est en partie parce qu'il est simple, et relativement consensuel (il y a quand même eu des objections, par exemple sur le danger qu'il pouvait poser si la zone n'était pas signée avec DNSSEC).