

Développement de protocoles à l'IETF

Stéphane Bortzmeyer
AFNIC
Immeuble International
78181 Saint-Quentin-en-Yvelines
bortzmeyer@nic.fr

Résumé

L'IETF (Internet Engineering Task Force) est une organisation de normalisation qui est responsable de la plupart des protocoles utilisés dans l'Internet, à partir de la couche 3. Ces normes sont publiées dans les fameux RFC, documents très bien écrits et librement accessibles et distribuables.

Comme les autres organisations équivalentes, elle a une forte culture spécifique, et de nombreuses particularités intéressantes.

Outre une description générale de l'IETF, cet article explique l'élaboration de plusieurs de ces normes, comme les étiquettes de langue, le futur protocole NCP (Name server Control Protocol), ou la tentative de normaliser SPF. L'accent sera mis sur des projets auxquels l'auteur a personnellement participé, qui ne sont pas forcément représentatifs des travaux typiques de l'IETF.

Mots clefs

IETF normalisation RFC TCP/IP protocole Internet

1 Introduction

L'IETF (Internet Engineering Task Force) est une organisation de normalisation qui est responsable d'une bonne partie des protocoles qui font fonctionner Internet. Elle se charge d'une partie de la couche 7 (avec des protocoles comme SMTP, pour le courrier, ou HTTP, pour le Web) et de protocoles qui sont formellement associés à la couche 7 mais qui sont plutôt des protocoles d'infrastructure que d'application (le DNS étant le plus connu). Mais l'IETF se charge surtout de la couche 4 (transport, comme TCP et ses cousins moins connus comme SCTP) et de la couche 3 (routage, avec évidemment IP mais aussi les protocoles de contrôle comme BGP et OSPF).

L'IETF s'étend aussi vers les formats de données (comme le format de publication et de

syndication Web Atom¹ ou comme JSON, qui est une norme IETF) et vers les couches inférieures à la couche 3, comme MPLS, un de ses protocoles vedettes. Beaucoup de textes de l'IETF servent aussi au travail interne, en décrivant les processus et l'organisation, ou bien en établissant des « normes pour écrire des normes » comme ABNF (p. 6) pour les grammaires formelles.

S'il existe un très grand nombre de SDO (Standard Development Organization) dans le domaine des réseaux, l'IETF a des caractéristiques originales. Elle est particulièrement ouverte (vous pouvez y « adhérer » dès aujourd'hui, sans formalités, ni paiement) et le style de travail y est peu formel. On s'inscrit sur une liste de diffusion et on est « membre de l'IETF » et on peut écrire des normes.

L'IETF a beaucoup changé depuis sa création, en général dans le sens d'une plus grande formalisation, provoquée à la fois par l'augmentation de sa taille et par celle de ses responsabilités.

L'IETF est surtout connue par sa production, les fameux RFC² (Request For Comments). Si tous les RFC ne proviennent pas de l'IETF, elle a par contre la tâche de produire ceux qui sont sur le chemin des normes³, c'est-à-dire les RFC qui décrivent les règles que doivent respecter les logiciels pour qu'Internet fonctionne.

Distribués gratuitement et redistribuables aux collègues et aux étudiants, les RFC ont également l'avantage d'être publiés, depuis toujours, dans un format ouvert, accessible à tous. En outre, bien qu'on ne puisse pas dire qu'on lise un RFC comme un roman, ils sont en général assez lisibles pour le programmeur ou l'administrateur réseau normal.

Les RFC sont ainsi largement responsables du succès de l'Internet, par rapport à certaines

¹Il n'existe pas de méta-organisme de normalisation qui se chargerait de distribuer le travail entre les différents organismes. Atom aurait pu, et cela aurait été plus logique, être développé par le W3C. Mais ce dernier a préféré suivre une autre voie.

²Tous sont disponibles en <http://www.ietf.org/rfc.html>

³Standards track, en anglais

propositions concurrentes à une époque, propositions mieux soutenues par les autorités mais nettement moins accessibles.

2 Comment fonctionne l'IETF

2.1 Organisation

L'IETF a une structure très particulière. Ce n'est pas une association, ni une entreprise, c'est... rien. L'IETF n'existe pas, elle n'est pas une personne morale et on la chercherait en vain dans un annuaire des organisations.

Puisqu'elle n'existe pas, l'IETF n'a donc pas de membres ou, plus exactement, tout le monde peut se dire membre puisqu'il n'y a pas de mécanisme d'adhésion.

Cette structuration est issue des débuts de l'IETF, à une époque où tout était informel. Elle a perduré par suite de la difficulté qu'il y a à créer des structures internationales (il n'existe aucun mécanisme simple aujourd'hui pour créer une ONG internationale et avoir une IETF enregistrée dans un seul pays en mettrait beaucoup mal à l'aise).

Mais dans le monde où nous vivons, il est difficile de ne pas être enregistré, de ne pas avoir de personnalité morale. L'IETF est donc coiffée par l'ISOC (Internet Society), qui lui fournit son parapluie organisationnel et légal. Ainsi, si un avocat états-unien veut faire un procès car un RFC lui déplait, c'est l'ISOC qu'il doit assigner devant le tribunal.

La propriété intellectuelle de l'IETF, elle, est gérée par un organisme conjoint entre l'ISOC et CNRI, l'IETF trust.

On ne peut donc pas « adhérer à l'IETF ». Un membre de l'IETF est simplement quelqu'un qui prend part aux activités de ladite IETF.

Le fait de ne pas avoir de personnalité morale n'empêche pas d'avoir une structure. Celle de l'IETF est également originale car elle n'est pas strictement hiérarchique.

La partie administrative de l'IETF (organisation des réunions) est confiée à l'IAOC (IETF Administrative Oversight Committee), un comité qui supervise le travail quotidien du directeur opérationnel, l'IAD (IETF Administrative Director).

Le financement de l'IETF, quatre millions de dollars par an, est assuré à moitié par l'ISOC, à moitié par les participations aux meetings physiques⁴. Comme il n'y a pas de membres, il n'y a pas de cotisation.

La partie technique du travail est découpée en secteurs⁵ comme la « Routing area » qui se charge des protocoles de routage (par exemple BGP) ou la « Security Area » qui se charge de trouver des solutions aux problèmes de sécurité de plus en plus nombreux. Le travail de l'auteur de cet article se fait en grande partie dans l'« Application area » qui prend en charge la couche 7. Le secteur a deux directeurs⁶ et, ensemble, ces directeurs forment l'IESG (Internet Engineering Steering Group), la structure qui suit le travail quotidien de l'IETF et doit examiner chaque document.

Chaque secteur est découpé en groupes de travail (« Working group ») qui sont le cœur de l'IETF. Par exemple, l'« Application area » compte actuellement un groupe de travail sur les étiquettes de langue (LTRU, cf. p. 6), un sur le format de publication Web Atom, un sur l'internationalisation des adresses de courrier électronique, un sur la nouvelle version du langage de filtrage du courrier Sieve, etc.

C'est dans ces groupes de travail que se fait l'essentiel de l'activité de l'IETF. Un groupe de travail a deux présidents⁷. La création d'un groupe est un travail de longue haleine et très formalisé. Il est essentiel que le groupe ait des objectifs clairs et limités. Sa charte est écrite de manière très étroite, pour éviter les groupes « éternels », qui continueraient une vague activité bien au delà de ce qui serait utile.

Un peu à part se situe l'IAB (Internet Architecture Board). Contrairement à l'IESG, qui est située en plein milieu du torrent, l'IAB est sur la rive et supervise l'évolution à moyen et long terme des normes Internet. C'est un peu la statue du commandeur de l'IETF, chargée par exemple des rappels d'architecture solennels (RFC 1984 ou RFC 4924) mais aussi d'impulser des nouveaux travaux (RFC 3869 ou 4948).

Comment sont choisies les personnes qui forment l'IAOC, l'IESG ou l'IAB ? Pour l'IAOC, qui supervise les tâches administratives, c'est décrit dans le RFC 4333.

Pour l'IESG et l'IAB, le RFC pertinent est le 3777. C'est un document complexe et pas très amusant à lire, mais on peut le résumer de la façon suivante :

1. Un Nomcom (Nominating Committee) est constitué. Ses membres sont des volontaires qui doivent avoir participé à au moins trois des cinq dernières réunions physiques. Le Nomcom est tiré au sort parmi ces volontaires, selon la procédure publiquement vérifiable du

⁵Area en anglais

⁶Area director en anglais

⁷Chair en anglais

⁴<http://iaoc.ietf.org/budget.html>

RFC 3797. Le président du Nomcom est désigné par l'ISOC.

2. N'importe qui peut nommer des personnes pour les postes libres.
3. Pour chaque poste, le Nomcom choisit un candidat parmi les nominés.
4. Pour chaque poste, il existe un organisme de confirmation qui a un droit de veto sur la désignation (l'ISOC pour les postes à l'IAB, par exemple).

En pratique, les directeurs de secteurs et les présidents de groupes viennent souvent des grands acteurs du domaine, éditeurs de logiciel, constructeurs de matériel, opérateurs et (pour le DNS) registres de noms de domaines. Contrairement à certaines organisations de normalisation, tous les participants sont en effet censés être des praticiens actifs du domaine concerné.

Le résultat le plus connu de l'activité de l'IETF est l'ensemble des RFC⁸. Ces documents, très variés en taille et en importance, sont un peu les « textes sacrés » de l'Internet. Certains ont le statut de normes (il existe trois étapes sur le « chemin des normes », proposition, projet et norme⁹), d'autres sont publiés pour information, d'autres enfin sont qualifiés d'expérimentaux car l'IETF n'est pas sûre du résultat mais tous ont en commun d'être très bien écrits, faciles à lire pour l'ingénieur et distribués librement sur l'Internet, un point qui a beaucoup contribué au succès des protocoles IETF.

Par exemple, le caractère très concret et relativement lisible des RFC fait que les prendre comme sujet de TP est tout à fait réaliste. Quel que soit le niveau de l'étudiant, il y a au moins un RFC qu'il peut implémenter.

L'accent mis sur l'utilité concrète des normes fait que l'IETF a toujours résisté à la tentation d'adopter le dernier format de fichiers à la mode. Certains RFC en activité ont plus de vingt-cinq ans et aucun format de fichier n'aurait tenu aussi longtemps. Les RFC sont donc toujours en texte brut, avec le jeu de caractères ASCII, comme à l'époque où les précurseurs des dinosaures communiquaient avec des modems à 300 bauds. Cela permet de les afficher, de les indexer et de les chercher facilement.

On l'a dit, les RFC n'ont pas tous le même statut. Une place à part doit être faite aux RFC du 1^{er} avril, publiés avec le même sérieux que les autres, comme le célèbre RFC 1149 ou

⁸Certains RFC, une minorité, ont été développés en dehors de l'IETF.

⁹*Proposed standard, draft standard et standard*, en anglais.

comme le RFC 2324 sur le protocole de contrôle des machines à café.

2.2 Fonctionnement

L'essentiel du travail à l'IETF se fait dans deux cadres : au sein des listes de diffusion des groupes de travail et lors des réunions physiques.

Les groupes de travail fonctionnent surtout par liste de diffusion. Il n'est pas exagéré de dire qu'un groupe de travail, c'est surtout une liste de diffusion. Ces listes sont toujours publiques (tout le monde peut s'y inscrire, rappelez-vous que l'IETF n'a pas de procédure d'adhésion) et surtout archivées publiquement. Il est essentiel que les décisions soient publiquement documentées¹⁰.

La plupart des RFC ont commencé leur vie comme Internet-Draft (I-D). Un I-D peut être écrit par n'importe qui, membre ou pas de l'IETF. Il est publié sur les serveurs de l'IETF et la discussion peut alors commencer. L'insistance mise sur les documents écrits est forte. Les critiques sont souvent reçues par un « Proposez donc un texte alternatif ». Tout est ainsi documenté.

L'I-D est parfois adopté officiellement par un des groupes de travail (et c'est quasiment indispensable si on veut qu'il aille sur le chemin des normes), sinon, il reste proposition individuelle.

Une fois que, après moult itérations, l'Internet-Draft a atteint un stade qui semble satisfaisant aux présidents du groupe, ils peuvent lancer un « Working group last call » pour permettre à toutes les objections d'être soulevées. Les présidents doivent alors juger si le groupe de travail est parvenu à un consensus. Il n'existe en effet pas de vote dans les groupes de travail¹¹, puisqu'il n'y a pas de processus d'adhésion donc pas de moyen de délimiter le corps électoral.

Si le consensus est jugé suffisant (rappelez-vous que tout est public donc il est difficile pour un président de prétendre qu'il n'y a pas eu d'objection s'il y en a eu), et si l'I-D est destiné au chemin des normes, il est alors transmis à l'IESG, qui va l'examiner, éventuellement faire un « IETF Last call » permettant aux non-membres du groupe de donner leur

¹⁰Il paraît que la France a voté non à la norme ISO 639-3 sur les langues. Impossible de le vérifier, même les votes ne sont pas accessibles publiquement à l'ISO, sans même parler des discussions. Au contraire, à l'IETF, on peut toujours expliquer un choix en donnant l'URI du message où ce choix était justifié.

¹¹Une formule courante du folklore IETF est « We reject presidents, kings and voting. We believe in rough consensus and running code. »

avis, avant de l'approuver et de le transmettre au « RFC editor », l'organisation (indépendante de l'IETF) qui publie les RFC, avec lenteur et précaution¹².

D'autres outils que la messagerie ont également pris davantage d'importance récemment : les gestionnaires de tâches notamment, qui permettent de s'assurer que rien n'a été oublié, que les points soulevés sont effectivement traités.

La messagerie instantanée, via le protocole Jabber/XMPP (RFC 3920), est également de plus en plus utilisée.

Les réunions physiques, quant à elles, ont perdu de leur importance technique depuis le développement des NTIC mais elles restent indispensables si on veut réellement influencer le processus de production des normes, si on veut connaître les autres participants autrement que par leur adresse électronique, ou tout simplement si on veut bien s'amuser et faire admirer son nouveau portable. Elles rassemblent plus d'un millier de participants et se tiennent trois fois par an, dans différentes villes¹³. Le Tao (RFC 4677) dit que le code d'habillement est strict : comme le port du badge est obligatoire, il faut forcément au moins une chemise ou un T-shirt.

Qui participe à l'IETF ? Bien sûr, il y a toutes sortes de personnes. Mais les biais sont nombreux. Il faut parler anglais, avoir de l'argent et du temps, être plutôt « grande gueule » si on veut avoir des chances d'être entendu. On ne s'étonnera donc pas s'il y a une grande majorité d'hommes blancs, venus des pays développés. Bien que, officiellement, on participe toujours à l'IETF personnellement, quel que soit son employeur, en pratique, les grands équipementiers réseau (Cisco, Juniper), les grands opérateurs (France Télécom, Verisign), les grands éditeurs de logiciel (Sun, Microsoft, IBM) sont très bien représentés.

2.3 Problèmes et questions ouvertes

Le RFC 3774 est une lecture intéressante. C'est un document de l'IETF critiquant l'IETF. Il existe peu de textes ou de documents sur l'IETF (le livre de Christian Huitema cité dans la bibliographie est un rare contre-exemple, plein d'anecdotes amusantes sur l'IAB) et les

¹²L'IANA (Internet Assigned Numbers Authority), une activité actuellement assurée par l'ICANN, est également dans la boucle, si le RFC nécessite d'allouer des ressources depuis un registre, par exemple s'il faut réserver un nouveau numéro de port TCP ou UDP.

¹³En général, plutôt Minneapolis (Minnesota) que Hawaï ou Maurice.

RFC tendent évidemment à présenter une vue favorable de l'organisation qui les a écrits. Or, comme toute organisation humaine, l'IETF a sa théorie et sa pratique...

Le RFC 3774 est donc consacré à l'examen des problèmes et limites de l'IETF. La question de la monoculture linguistique¹⁴, par exemple, y est longuement documentée (sans qu'une solution soit offerte).

Le résultat du travail de l'IETF, les RFC, pourraient aussi évoluer. Les RFC ne sont pas structurés, donc les analyser automatiquement, par exemple pour faire des analyses bibliographiques, est excessivement pénible. Ainsi, le RFC 2629 qui tentait de résoudre ce problème est toujours resté au milieu du gué : beaucoup l'ont adopté mais pas suffisamment pour que le RFC editor en fasse un format officiel.

3 Le futur NCP

Voyons maintenant quelques exemples de travaux concrets faits par l'IETF.

Le DNS (Domain Name System) est un exemple de protocole qui est officiellement en couche 7 mais, en pratique, fait partie de l'infrastructure de l'Internet et n'est typiquement pas vu par les utilisateurs. Actuellement, deux groupes de travail se penchent sur le DNS, dnsop, qui a en charge l'aspect opérationnel du DNS (par exemple la mise en garde contre les serveurs récursifs ouverts) et dnsex, plus connu par le nom de sa liste de diffusion, « namedroppers », qui a en charge les extensions du protocole (nouveaux types d'enregistrement ou bien grands changements comme DNSSEC, qui a occupé l'essentiel des activités de namedroppers depuis des années).

Comme toujours à l'IETF, les nouveaux RFC ne sont développés que si des volontaires s'en occupent. Il y a plus d'idées que de temps pour les réaliser et on voit certaines idées réapparaître de temps en temps, tout le monde les trouvant excellentes, mais personne n'ayant le temps et la volonté de les faire aboutir dans un RFC.

Par exemple, autant le protocole DNS est normalisé, autant les interactions entre les ser-

¹⁴À la réunion de Prague en 2007, Ted Hardie, alors directeur de l'« Application Area » avait fait son exposé en chinois pendant la séance plénière. Ses transparents, également en chinois, étaient entrecoupés d'extraits du film « Charade », de Stanley Donen, pour que les francophones ou les italo-phones puissent saisir quelques mots. Sa conclusion, en anglais, était qu'il n'est pas facile de travailler dans une autre langue que la sienne.

veurs de noms (à part le transfert de zones entre un maître et un esclave) ou bien entre les serveurs de noms et leurs administrateurs, sont actuellement le monde exclusif des protocoles privés. Par exemple, les systèmes d'Infoblox ou le logiciel libre PowerDNS ont un moyen pour un serveur esclave de connaître les zones DNS servies par le maître, de façon à se reconfigurer automatiquement.

À plusieurs reprises, des propositions ont été faites pour sortir de cette situation. Il y a même eu des Internet-Drafts d'écrits. Mais l'enthousiasme est retombé à chaque fois.

La proposition actuelle se nomme NCP (Name-server Control Protocol). NCP a été proposé sous forme d'un I-D en décembre 2006. Cet I-D a été discuté à Prague en mars au sein de dnsop. La création d'un «design team», un groupe restreint d'experts qui doit déblayer le terrain et avancer sur le cahier des charges, laissant ensuite le groupe de travail chercher une solution, a été décidée. Quelques mois après, à la réunion de Chicago, ledit «design team» n'avait toujours rien fait. Oh, vanité des œuvres humaines !

Un autre «design team» a alors été mis en place, sous la direction, cette fois, d'une personnalité reconnue, Jaap Akkerhuis. On va voir si le travail avance mieux. Mais la volonté des présidents de dnsop de bien définir le cahier des charges avant de se lancer dans la définition du protocole se heurte au désir de beaucoup de brûler les étapes et de spécifier la solution d'abord, avant de chercher quel était le problème.

4 Cosmogol décollera t-il ?

Selon un adage, «le cordonnier est toujours le plus mal chaussé». Et c'est vrai que l'IETF, organisation d'experts en réseaux informatiques, normalisant les techniques les plus complexes, a souvent un fonctionnement très traditionnel, voire archaïque. C'est ainsi que les documents produits, les RFC, sont majoritairement écrits en langue naturelle et que les langages formels ne s'y glissent que très lentement.

Écrire une norme qui marche est un art. La langue naturelle n'est pas toujours le meilleur outil pour cet art. Elle est ambiguë et elle ne peut pas être analysée et traduite automatiquement. Les chercheurs en informatique ont beaucoup travaillé pour produire des langages de spécification formelle qui n'auraient pas cet inconvénient. Les résultats sont mitigés et dépendent beaucoup du domaine. Par exemple,

dans le champ des langages de programmation, si les grammaires formelles en BNF ou similaire ont été un grand succès, la sémantique des langages de programmation courants reste toujours décrite en langage naturel.

L'IETF n'a donc pas de normes entièrement en langage formel. En dehors de l'anglais, les langages les plus utilisés sont SMI pour les MIB (Management Information Base, les schémas de l'information gérée via le protocole SNMP) et ABNF (p. 6) pour les formats. Jusqu'à la sortie du RFC 4997 en juillet 2007, l'IETF n'avait même pas de langage formel pour décrire les en-têtes des paquets, pourtant son cœur de métier. Un autre cas où l'IETF n'a traditionnellement pas de langage formel est celui des automates à états finis.

De tels automates sont fréquents dans les RFC. Ils sont typiquement exprimés par un graphique en ASCII-art, parfois par un tableau de transitions et parfois encore par une liste de transitions. N'étant pas formalisés, ils ne peuvent pas être analysés automatiquement et c'est ainsi que des RFC ont pu être publiés avec des automates incomplets ou erronés.

Un effort a donc commencé en juin 2006 pour examiner la question. L'idée était de décrire les automates à états finis sous forme d'un langage texte (contrairement à des systèmes comme UML), pouvant être utilisés dans les RFC, comme ABNF. Ce sera le langage Cosmogol.

Le projet n'a pas rencontré un enthousiasme délirant. Certaines personnes doutaient que le jeu en vaille la chandelle, vues la petite taille et la simplicité des automates qu'on trouve typiquement dans les RFC. D'autres pensaient préférable d'utiliser un langage déjà existant, comme le Z.100/SDL de l'IUT (qui n'a jamais été utilisé dans aucun RFC).

Le projet a surtout été porté par une seule personne, avec quelques participants qui aidaient, mais sans assurer de responsabilités. Un draft a été écrit et implémenté, un [site Web](#) a été développé, une liste de diffusion créée par l'IETF et une BOF (Birds Of a Feather, une réunion semi-informelle lors des réunions IETF) s'est tenue à Prague en mars 2007. Plusieurs automates publiés dans des RFC ont été traduits dans le langage proposé, Cosmogol. Voici un exemple de code Cosmogol, inspiré du RFC 3931 ; on indique d'abord l'état actuel de l'automate, le message reçu puis, après la flèche, l'état suivant (et une action) :

```
idle : "Local open request" ->
      wait-ctl-reply : "Send SCCRQ";
```

```

idle : "Receive SCCRQ, acceptable" ->
      wait-ctl-conn : "Send SCCRQ";

wait-ctl-reply :
  "Receive SCCRQ, acceptable" ->
    established : "Send SCCRQ";

wait-ctl-reply :
  "Receive SCCRQ, not acceptable" ->
    idle : "Send StopCCN, clean up";

```

Si la BOF de Prague a été un certain succès, avec une large participation, aucun Internet-draft n'a encore utilisé le nouveau langage et aucun groupe de travail n'a été créé.

Le projet est actuellement suspendu de fait.

On notera qu'aucune décision formelle n'a été prise par l'IETF. Même des langages très établis comme ABNF n'ont jamais été officiellement « approuvés » et encore moins imposés. Chaque auteur de RFC est libre des techniques qu'il utilise pour décrire son protocole ou format. Cosmogol n'a donc pas été refusé, comme il aurait pu l'être dans une organisation plus formaliste, il a simplement échoué, pour l'instant, à susciter suffisamment d'intérêt.

5 ABNF, norme complète ou pas ?

Cette histoire, pas encore terminée, illustre deux questions fréquentes à l'IETF, celle des langages formels et celle de la progression sur le chemin des normes.

Beaucoup de RFC contiennent des grammaires formelles, pour décrire un format ou les messages échangés dans un protocole. Les étiquettes de langue, p. 6, sont un exemple typique d'un mini-langage, décrit dans une telle grammaire. Ces grammaires formelles ont de nombreux avantages sur les descriptions en langue naturelle : elles sont sans ambiguïté et elles peuvent être traitées par des outils comme [Eustathius](#), que ce soit pour les vérifier ou pour produire automatiquement des analyseurs syntaxiques.

À une époque, chaque RFC définissait approximativement sa grammaire formelle et l'utilisait ensuite. ABNF est issue d'un effort de Dave Crocker de regroupement de la définition de la grammaire en un seul RFC, le 2234 en 1997, devenu ensuite le 4234 en 2006.

Notons qu'il n'y a pas grand'chose d'obligatoire à l'IETF. Aucune autorité ne peut dire aux auteurs de RFC « Désormais, vous êtes obligés

d'utiliser ABNF ». Ceux qui le font le choisissent volontairement. La principale « dissidence » vient du RFC 2616, qui normalise HTTP, et qui utilise un langage différent.

Notons aussi que, bien que l'un des principaux avantages des langages formels soit la possibilité de vérification automatique, cette possibilité n'est pas utilisée systématiquement. C'est ainsi que certains RFC ont été publiés avec des bogues dans leur grammaire ABNF...

Le RFC 2234 était « Proposition de norme », le 4234 est « Projet de norme ». Pour franchir l'étape supplémentaire menant au statut de « Norme » tout court, ABNF a dû subir tout un processus, qui n'est pas encore terminé, malgré l'extrême ancienneté de cette norme.

Notamment, ABNF a dû subir le test de l'implémentation. Un membre de l'IETF doit faire un rapport sur les mises en œuvre existantes et sur leur interopérabilité. L'IETF n'aime pas les normes purement théoriques¹⁵. Cela ne sert à rien de spécifier un beau format si personne n'arrive à le programmer complètement¹⁶.

ABNF a passé ce test (il existe au moins deux implémentations indépendantes, qui interopèrent ou, dans ce cas, qui traitent les grammaires de la même façon).

Il a en revanche été bloqué, peut-être temporairement, par un point soulevé assez tardivement. Le RFC 4234 normalise non seulement le format des grammaires mais aussi une bibliothèque de productions prêtes à l'emploi comme HEXDIG pour les caractères utilisables pour écrire de l'hexadécimal ou ALPHA pour les lettres de l'alphabet latin. Une de ces productions, LWSP (espaces en début de ligne), est en effet critiquée comme trop puissante : des auteurs de RFC l'ont utilisé sans comprendre ses conséquences.

Bien que le problème ne concerne pas le langage ABNF lui-même, ce problème bloque la norme actuellement.

6 Les étiquettes de langue

Un des problèmes classiques de l'Internet est l'internationalisation, c'est-à-dire la capacité à fournir le même service pour des utilisateurs de différentes langues et cultures. Le problème est très riche, très complexe, et n'est que partiellement de la responsabilité de l'IETF. Celle-ci s'occupant plutôt des couches basses, la majorité des problèmes d'internationalisation lui passent, littéralement, au dessus.

¹⁵C'est une application du principe « Rough consensus and running code »

¹⁶Comme c'est fréquent avec les normes ISO

Il y a bien sûr les questions de processus, la langue utilisée à l'IETF – exclusivement l'anglais, les formes de la discussion, qui peuvent dérouter les participants issus de certaines cultures, etc. Ces questions sont discutées dans le RFC 3774 mais sans que des solutions soient proposées.

Pour ce qui concerne les protocoles, la politique de l'IETF en matière d'internationalisation est décrite dans le RFC 2277. En gros, ce RFC pose comme principe que les éléments du protocole (comme GET ou POST en HTTP) n'ont pas à être traduits mais que les messages affichés aux utilisateurs peuvent l'être.

Cette règle fort simple ne traite pas tous les cas. Le plus sensible est celui des noms de domaines, qui sont à la fois éléments de protocole et textes ayant une signification, et maniés par les utilisateurs.

Il reste que l'IETF a certaines activités directement liées à l'internationalisation. Une des plus visibles concerne les étiquettes de langue, les chaînes de caractères qui permettent d'identifier une langue particulière. Ces étiquettes, normalisées dans le RFC 4646, sont utilisées par de nombreux protocoles et formats, XML et HTTP étant les plus communs.

Par exemple, un client HTTP indique la langue qu'il souhaite avec l'en-tête `Accept-language` de sa requête. Un document XML indique sa langue avec l'attribut `xml:lang`. Cet attribut prend comme valeur une étiquette de langue, étiquette qui indique la langue mais aussi d'autres caractéristiques comme l'écriture utilisée ou les variantes dialectales. C'est ainsi que `fr` désigne le français, `br` le breton, `uk` l'ukrainien mais aussi `az-Latn` l'azéri en caractères latins (en Iran, il est en général écrit avec l'alphabet arabe), `en-scotland`, la variante de l'anglais qu'on parle en Écosse, etc.

L'établissement de la syntaxe de ces étiquettes ne pose que des problèmes techniques. En revanche, la liste des langues soulève des problèmes linguistiques et politiques complexes, comme le fait que certains valenciens nient que le valencien soit un dialecte du catalan, ou comme l'attitude des autorités de Pékin qui ont du mal à admettre que le terme de « chinois » recouvre en fait plusieurs langues. L'IETF n'a pas les compétences linguistiques ou la légitimité politique pour intervenir dans ce domaine. D'où le choix de s'appuyer sur des normes existantes comme ISO 15924 pour les écritures.

Le travail sur ce sujet a commencé à l'IETF il y a fort longtemps. Le premier RFC était le RFC 1766, écrit par l'expert Harald Alvestrand en mars 1995. Il a été révisé dans le RFC 3066, en

janvier 2001. Un des principes de ce travail était la création d'un groupe informel séparé, représenté par la liste de diffusion `ietf-languages`, qui était chargé de maintenir la liste des étiquettes valables, le RFC décrivant la syntaxe.

Un groupe de travail, avec de nouveaux auteurs, LTRU (Language Tag Registry Update), a été formé en mars 2005 pour réviser le RFC 3066. Les points les plus importants étaient l'intégration des écritures, une grammaire plus souple et surtout un registre stable (les normes ISO ne sont pas stables, un identificateur peut disparaître ou, pire, être réaffecté).

Le résultat du travail de ce groupe a été la publication du RFC 4646 en septembre 2006. La syntaxe est compatible (les anciennes étiquettes restent toutes valables) mais elle permet désormais d'exprimer bien plus de choses. Un registre stable a été créé à l'IANA. Un [site Web](#) a été mis en place pour présenter le travail du groupe, le registre et les outils qui permettent de l'utiliser.

Compte-tenu de l'extrême sensibilité de ces sujets, il n'est pas étonnant que ce groupe ait connu une histoire plus mouvementée que ceux qui s'occupent des zéros et des uns circulant sur le câble. Dans certains cercles, l'IETF a été accusée de vouloir « contrôler les langues mondiales » ou, crime suprême, de vouloir remettre en cause le travail de l'ISO. Ainsi, contrairement à la plupart des RFC, ceux de LTRU ont fait l'objet d'une procédure d'appel auprès de l'IAB (l'appel a été repoussé).

Le travail ne s'arrête pas là. La norme ISO 639 qui identifie les langages a connu deux itérations, ISO 639-1 et ISO 639-2, toutes les deux gérées par la Bibliothèque du Congrès états-unien. Une troisième version, ISO 639-3 a été confiée à une organisation missionnaire, SIL, et multiplie par douze le nombre de langues (les premières versions ne prenaient en compte que les langues bien établies), tout en introduisant un nouveau concept, celui de macrolangue, une langue qui peut être vue comme une seule langue sous certains aspects et comme une famille de langues sous d'autres aspects. Un exemple est l'arabe, macrolangue constituée de l'arabe standard, de l'arabe égyptien, de l'arabe marocain, etc.

LTRU a donc vu sa charte modifiée pour intégrer ce travail. Rappelons qu'un groupe de travail IETF a une charte très limitée, pour éviter les groupes éternels, sans but bien défini. L'intégration de ISO 639-3 nécessitait donc un processus de modification de la charte du groupe en août 2006. Une fois celle-ci effectuée, le travail a commencé sur le RFC 4646-

bis. Ce travail est toujours en cours, l'intégration du concept de macrolangue s'avérant plus délicate que prévu.

7 L'échec de MARID

Si certains travaux à l'IETF se concluent brillamment par l'adoption de beaux RFC largement applaudis, si d'autres aboutissent avec douleur mais la satisfaction d'arriver au but¹⁷, si certains meurent tranquillement et discrètement par manque d'impetus (la grande majorité), d'autres connaissent une histoire pleine de bruit et de fureur, n'aboutissent pas à grand-chose et laissent tout le monde déçu et frustré, voire hargneux.

C'est ce qui s'est produit avec le groupe de travail MARID (MTA Authorization Records in DNS), créé en avril 2004 dans la ferveur générale, pour traiter d'un grave problème, le spam et plus spécialement l'absence d'authentification du courrier électronique qui rend la lutte contre le spam si difficile.

MARID semblait bien parti. La technologie, SPF (*Sender Policy Framework*) était déjà prête, tous les acteurs de l'industrie, convaincus de la nécessité d'une lutte vigoureuse contre le spam, participaient,

Mais le groupe a éclaté très rapidement, déchiré notamment par le débat autour des brevets détenus par Microsoft sur une des technologies possibles. L'IETF n'a pas de politique uniforme en matière de brevets¹⁸, à part l'obligation pour les participants de révéler les brevets dont ils peuvent avoir connaissance (RFC 3979).

D'autres problèmes ont été soulevés dans MARID (les questions de sécurité informatique sont souvent passionnelles). MARID a finalement été fermé d'autorité par l'IESG en septembre 2004, ne laissant que des RFC requalifiés d'expérimentaux, notamment les RFC 4406 et 4408.

8 Conclusion

Il existe de nombreuses SDO très différentes dans leur degré d'ouverture, leur politique quant aux brevets, leur attention portée aux problèmes concrets. L'IETF est sans doute une des plus informelles (malgré l'incontestable formalisation qui l'envahit depuis quelques années) et des plus ouvertes et elle reste le principal lieu de normalisation pour les proto-

¹⁷C'est ce qui est arrivé au groupe LTRU, avec les étiquettes de langue (p. 6).

¹⁸Ni en matière de mise en œuvre (logiciel libre ou pas) puisque l'IETF n'a pas d'activité de développement logiciel.

coles de couche 3 et 4, tout en prenant sa part dans la couche 7.

L'une des difficultés futures sera de gérer les effets de décisions prises par les acteurs, en l'absence de tout processus de normalisation, comme le déploiement massif du NAT (Network Address Translation), ou celui des réseaux pair-à-pair, ou l'usage massif de techniques anti-spam à forts dommages collatéraux. Les décideurs dans ces domaines ne semblent pas accorder beaucoup de poids aux normes IETF, ni aux mises en garde de l'IAB.

9 Bibliographie

- [1] *IETF Home Page* <http://www.ietf.org/>
- [2] Christian Huitema, *Et Dieu créa l'Internet*. Eyrolles, 1995.
- [3] Stéphane Bortzmeyer. *IETF Trust: l'IETF perd officiellement sa propriété intellectuelle*. <http://www.bortzmeyer.org/ietf-trust.html> 2005
- [4] RFC 1149 *Standard for the transmission of IP datagrams on avian carriers*. D. Waitzman. 1990
- [5] RFC 1984 *IAB and IESG Statement on Cryptographic Technology and the Internet*. IAB, IESG. 1996.
- [6] RFC 2026 *The Internet Standards Process -- Revision 3*. S. Bradner. 1996
- [7] RFC 2277 *IETF Policy on Character Sets and Languages*. H. Alvestrand. 1998
- [8] RFC 2616 *Hypertext Transfer Protocol -- HTTP/1.1*. R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee. 1999
- [9] RFC 2629 *Writing I-Ds and RFCs using XML*. M. Rose. 1999.
- [10] RFC 2850 *Charter of the Internet Architecture Board (IAB)*. B. Carpenter, 2000
- [11] RFC 3774 *IETF Problem Statement*. E. Davies. 2004
- [12] RFC 3777 *IAB and IESG Selection, Confirmation, and Recall Process: Operation of the Nominating and Recall Committees*. J. Galvin. 2004

- [13]RFC 3797 *Publicly Verifiable Nominations Committee (NomCom) Random Selection*. D. Eastlake 3rd. 2004
- [14]RFC 3920 *Extensible Messaging and Presence Protocol (XMPP): Core*. P. Saint-Andre, 2004.
- [15]RFC 3935 *A Mission Statement for the IETF*. H. Alvestrand. 2004
- [16]RFC 3979 *Intellectual Property Rights in IETF Technology*. S. Bradner, 2005
- [17] RFC 4071 *Structure of the IETF Administrative Support Activity (IASA)*. R. Austein, Ed., B. Wijnen, 2005
- [18]RFC 4234 *Augmented BNF for Syntax Specifications: ABNF*. D. Crocker, P. Overell. 2005
- [19] RFC 4333 *The IETF Administrative Oversight Committee (IAOC) Member Selection Guidelines and Process*. G. Huston, Ed., B. Wijnen, 2005.
- [20]RFC 4406 *Sender ID: Authenticating E-Mail*. J. Lyon, M. Wong. 2006.
- [21] RFC 4408 *Sender Policy Framework (SPF) for Authorizing Use of Domains in E-Mail, Version 1*. M. Wong, W. Schlitt. 2006.
- [22]RFC 4646 *Tags for Identifying Languages*. A. Phillips, M. Davis. 2006
- [23]RFC 4677 *The Tao of IETF - A Novice's Guide to the Internet Engineering Task Force*. P. Hoffman, S. Harris. 2006
- [24]RFC 4748 *RFC 3978 Update to Recognize the IETF Trust*. S. Bradner, 2006
- [25] RFC 4924 *Reflections on Internet Transparency*. B. Aboba, E. Davies. 2007
- [26]RFC 4997 *Formal Notation for RObust Header Compression (ROHC-FN)*. R. Finking, G. Pelletier. 2007