Groupe de travail Réseau
Request for Comments : 3675
Catégorie : Information

D. Eastlake 3rd, Motorola Laboratories
février 2004
Traduction Claude Brière de L'Isle

.sex considéré comme dangereux

Résumé
Il y a périodiquement des propositions pour rendre obligatoire l'utilisation d'un nom de domaine de niveau supérieur particulier ou d'un bit d'adresse IP pour marquer le matériel "pour adulte" ou "non sûr" ou équivalent. Le présent document explique pourquoi c'est une mauvaise idée du point de vue légal, philosophique, et particulièrement technique.

Table des matières

1. Introduction

Il y a périodiquement des propositions pour rendre obligatoire l'utilisation d'un nom de domaine de niveau supérieur particulier ou d'un bit d'adresse IP pour marquer le matériel "pour adulte" ou "non sûr" ou équivalent. Le présent document explique pourquoi c'est une mauvaise idée du point de vue légal, philosophique, et particulièrement technique.

2. Arrière plan

Le concept d'un domaine de niveau supérieur .sex, .xxx, .adult, ou similaire dans lequel il serait obligatoire de localiser le matériel salace ou similaire est périodiquement suggéré par certains politiciens et commentateurs. D'autres propositions comportent un domaine réservé exclusivement pour du matériel considéré comme approprié pour les mineurs, ou d'utiliser des bits ou gammes d'adresse IP pour mettre à part des contenus.

Dans un rapport d'octobre 1998 qui accompagnait le Décret de protection de l'enfance en ligne, le comité du Commerce domestique disait, "il n'y a aucune barrière technique à la création d'un domaine adulte, et il serait très facile de bloquer

tous les sites de la toile au sein d'un domaine pour adultes". Le rapport disait aussi que le comité se défiait d'une réglementation de l'industrie informatique et que toute décision du gouvernement américain "aurait des conséquences internationales" [HOUSEREPORT].

British Telecom a soutenu la création de domaines de niveau supérieur (TLD, *Top-Level Domain*) pour adultes, disant dans une lettre de 1998 au Ministère américain du Commerce qu'il "soutenait fortement" ce plan. La raison : "Les services explicitement sexuels pourraient alors être obligés par la loi de fonctionner avec des noms de domaine dans ce TLD [ce qui] rendrait beaucoup plus simple et facile de contrôler l'accès à de tels sites..." [BT]. Un des ancêtres de l'ICANN, le comité GTLD, suggérait un domaine de niveau supérieur dans une "zone à feu rouge" dans un appel à commentaires de septembre 1997 [GTLD-MOU].

Certains dirigeants de l'industrie pour adultes ont approuvé le concept. En 1998, Seth Warshavsky, président du groupe Internet Entertainment, déclarait au comité du commerce du Sénat américain qu'il aimerait voir un domaine .adult. "Nous suggérons la création d'un nouveau domaine de niveau supérieur appelé '.adult' où pourrait résider tout le matériel sexuellement explicite de la Toile", disait Warshavsky dans une interview à l'époque [WARSHAVSKY].

Plus récemment, d'autres entrepreneurs de l'industrie ont dit qu'ils n'objectaient pas nécessairement à la création d'un domaine pour adultes pour autant qu'ils pourraient continuer à utiliser .com.

Des groupes conservateurs aux USA disent qu'ils ne sont pas favorables à un tel domaine, et préfèrent que les lois criminelles punissent les éditeurs et distributeurs de matériel sexuellement explicite. Le Centre National Légal pour l'Enfance et la Famille de Fairfax, Virginie, disait en février 2001 qu'il n'était pas en faveur d'une telle proposition. Pour différentes raisons, l'Union Américaine des Libertés Civiles et d'autres groupes de défenseurs des libertés civiles s'y opposent également.

Le sénateur Joseph Lieberman, candidat du parti Démocrate américain à la vice présidence, a adhéré à l'idée lors d'un réunion de juin 2000 de la Commission Fédérale sur la Protection de l'enfance en ligne. Lieberman disait dans une déclaration préparée qu'il "demanderait aux arbitres de l'Internet de simplement respecter les mêmes standard que le propriétaire d'un cinéma classé X ou le propriétaire d'un magasin X qui vend des magazines pornographiques" [LIEBERMAN].

Dans la Loi de 1998 qui crée cette commission, le Congrès américain demande aux membres d'enquêter sur "l'établissement d'un nom de domaine où envoyer tout le matériel qui est dommageable pour les mineurs". La commission a consacré une section de son rapport d'octobre 2000 à ce sujet. Elle concluait qu'un domaine .xxx et un domaine .kids sont tous deux techniquement possibles, mais exigeraient une action de la part de l'ICANN. Le rapport disait qu'un domaine pour adultes pourrait n'être que "modérément efficace" et soulevait des problèmes de confidentialité et de liberté d'expression [COPAREPORT].

La commission explorait aussi la création de zones dites rouge ou verte pour les contenus, au moyen de l'allocation d'un nouvel ensemble d'adresses IP avec IPv6. Tout matériel qui ne serait dans aucune de ces deux zones serait considéré comme entrant dans une zone grise et donc pas nécessairement approprié ou inapproprié pour les mineurs. Les commentaires des membres de la commission étaient largement négatifs : "L'efficacité exigerait des efforts substantiels pour attacher un contenu à des numéros IP spécifiques. Cette approche pourrait réduire la souplesse et entraver une performance optimale du réseau. Elle ne serait pas efficace pour bloquer l'accès à la causette, aux groupes d'informations ou à la messagerie instantanée".

En octobre 2000, l'ICANN a rejeté un domaine .xxx durant sa session initiale d'approbation de domaines de niveau supérieur supplémentaires. Les raisons n'en sont pas entièrement claires, mais l'ancienne présidente de l'ICANN, Esther Dyson, disait que l'industrie pornographique n'estimait pas qu'un tel domaine serait vraiment approprié. Un soutien du .xxx, le registraire ICM de l'Ontario, Canada, a demandé en décembre 2000 à l'ICANN de reconsidérer sa décision [ICM-REGISTRY].

En 2002, le Congrès américain a rendu obligatoire la création d'un domaine kids.us pour le matériel sain pour les enfants. Ceci était le résultat de la conviction que, pour des raisons dont certaines sont décrites dans les paragraphes qui suivent, essayer de légiférer pour le monde entier avec un domaine .kids était inapproprié.

3. Problèmes légaux et philosophiques

Lorsqu'on parle de matériel sexuellement explicite, tout le monde, tribunaux et gouvernements compris a une vision différente de ce qui est acceptable et de ce qui ne l'est pas. Les attitudes changent avec le temps et ce qui était vu comme approprié dans une ville, ou il y a un an, peut déclancher des protestations dans la ville, ou l'année, suivante. Confronté à la

nature fuyante de ce qui dans la description d'activités sexuelles devrait être illégal ou non, un membre de la Cour Suprême américaine a gaiement décrit l'obscénité en disant : "Je sais ce que c'est quand j'en vois".

Aux U.S.A., l'obscénité est définie comme matériel sexuel explicite qui, entre autres choses, viole "les normes communautaires contemporaines" -- en fait, même au niveau national, il n'y a pas de règle communément admise qui gouverne ce qui est illégal et ce qui ne l'est pas. Ce qui complique encore les choses est qu'il y a plus de 200 codes de pays aux Nations Unies, et dans la plupart d'entre eux, des subdivisions politiques peuvent imposer leurs propres restrictions. Même pour des modèles nus légaux, les restrictions d'âge diffèrent. Elles sont généralement de 18 ans, mais seulement de 17 ans dans un pays scandinave. Un photographe qui prendrait ce qui y est considéré comme une photographie légale et tout à fait correcte serait dénoncé comme vicieux et adepte de la pornographie infantile aux U.S.A. Et dans bien d'autres pays et groupes, le seul concept de photographie de nu ou même toute photographie d'une personne de quelque nature qu'elle soit peut être inacceptable du point de vue religieux.

L'Arabie Saoudite, l'Iran, le nord du Nigéria, et la Chine n'ont vraisemblablement pas les mêmes vues libérales que disons, les Pays-Bas ou le Danemark. L'Arabie Saoudite et la Chine, comme quelques autres nations, filtrent de façon extensive leurs connexions Internet et ont créé des agences gouvernementales pour protéger leur société des sites de la Toile que les officiels considèrent comme immoraux. Leurs vues sur ce qui devrait être inclus dans un domaine .sex serait difficilement identiques à celles des nations occidentales libérales.

Ces opinions radicalement différentes sur le matériel sexuel rendent inconcevable qu'un consensus mondial puisse jamais être obtenu sur ces qui est approprié ou inapproprié pour un domaine de niveau supérieur .sex ou .adult. De plus, l'existence d'un tel domaine créerait une irrésistible tentation de la part de législateurs conservateurs d'exiger des éditeurs controversés qu'ils se placent dans ce domaine et de punir ceux qui ne le font pas.

Certains politiciens conservateurs se sont plaint déjà que l'ICANN n'approuve pas .xxx lors de sa réunion d'octobre 2000. Durant une audition de février 2001 à la Chambre des Représentants US, le législateur avertissait qu'il "voulait explorer les raisons pour lesquelles l'ICANN n'a pas approuvé deux noms de domaine de niveau supérieur particuliers -- .kids et .xxx -- comme moyens de protection des enfants contre les abominables saletés qui s'étalent partout dans l'Internet".

Il semble plausible que seuls quelques éditeurs pour adultes, et pas ceux qui ont investi des ressources en construisant une marque autour d'un site .com, vont volontairement abandonner leur nom de domaine actuel. Ils vont plutôt ajouter une variante .xxx et conserver leur adresse première. L'existence de .xxx pourrait amener le législateur aux U.S.A. et dans d'autres pays à exiger d'eux qu'ils publient exclusivement à partir d'un domaine pour adultes, action qui inviterait à une interférence politique avec la gouvernance de l'Internet, et soulèverait des problèmes de liberté d'expression et d'auto définition.

En fait, l'arbitre ultime des noms de domaines génériques de niveau supérieur – au moins actuellement – n'est pas l'ICANN, mais le gouvernement américain. En juillet 2000, le bureau de comptabilité générale du Congrès américain rapportait que le Ministère du Commerce continue d'être responsable des noms de domaines permis par la racine de l'autorité [GAO]. Les auditeurs du GAO concluaient qu'il n'apparaissait pas clairement si le Ministère du Commerce avait "l'autorité requise" sous la loi actuelle pour transférer cette responsabilité à l'ICANN.

L'Union américaine des libertés civiques – et les autres membres de la Campagne internationale pour la liberté de l'Internet mondial – avertissent que les éditeurs qui parlent avec franchise sur le contrôle des naissances, la prévention du SIDA, la sexualité des homosexuels et homosexuelles, le problème social du viol dans les prisons, etc., pourraient être obligés de se déplacer dans un domaine réservé aux adultes. Une fois là, ils seraient stigmatisés et facilement bloqués par les écoles, bibliothèques, entreprises, et autres groupes qui utilisent des logiciels de filtrage. Les éditeurs de telles informations, qui ne se voient pas eux-mêmes comme des pornographes et conserveraient leurs adresses existantes, pourraient faire l'objet de poursuites.

L'existence d'un domaine de niveau supérieur réservé aux adultes ouvrirait vraisemblablement la porte à des initiatives, politiques ou législatives concomitantes. Il y a de nombreux axes différents selon lesquels définir le matériel en cause : sexe, violence, haine, hérésie, subversion, blasphème, drogues illégales, profanation, politiquement correct, glorification du crime, incitation à la violation de la loi, et ainsi de suite. Des groupes d'intérêts particuliers qui ne sont pas concernés par la faisabilité technique et pratique de leur avis font pression avec de telles suggestions sur l'ICANN, le gouvernement des États-Unis, et autres institutions à caractère politique.
Un domaine de niveau supérieur réservé aux adultes pourrait avoir des répercussions légales négatives en mettant en danger la liberté d'expression. La juge à la Cour suprême de justice des U.S.A. Sandra Day O'Connor a suggéré que la présence de "zones adultes" sur l'Internet rendrait vraisemblable qu'on considère comme constitutionnel une future Loi sur la décence dans les communications (CDA, *Communications Decency Act*). Dans son désaccord partiel sur le rejet par la Cour Suprême du CDA en 1997 [CDA], O'Connor dit que "les perspectives d'un zonage final de l'Internet apparaissent prometteuses". (La Cour Suprême a jugé que le CDA violerait les droits de liberté d'expression en faisant un crime de la distribution de matériel "indécent" ou "manifestement choquant" en ligne.)

La confidentialité pourrait être violée par une telle proposition. Il deviendrait plus facile à des gouvernements répressifs et autres institutions de suivre les visites à des sites dans un domaine étiqueté comme "pour adultes" et d'enregistrer des informations d'identification personnelle sur le visiteur. Les gouvernements répressifs auraient instantanément plus de pouvoir pour surveiller les utilisateurs imprudents et les poursuivre à cause de leurs activités. Il est aussi peu vraisemblable qu'un domaine de niveau supérieur réussisse à contrôler l'accès à la causette, à la messagerie électronique, aux groupes de nouvelles, à la messagerie instantanée, et aux nouveaux services non encore inventés.

4. Difficultés techniques

Même en ignorant les difficultés philosophiques et légales soulignées ci-dessus, il y a des difficultés techniques substantielles à tenter d'imposer une classification des contenus par les noms de domaines ou les adresses IP. L'étiquetage obligatoire des contenus est habituellement avancé avec l'idée d'utiliser un nom de domaine de niveau supérieur, discutée au paragraphe 4.1, mais on discute aussi la possibilité d'utiliser les bits ou gammes d'adresse IP au paragraphe 4.2.

Au paragraphe 4.1.4 sont discutées les difficultés avec quelques protocoles de niveau supérieur particuliers. Dans certains cas, ces protocoles utilisent des espaces de noms différents. On devrait garder présent à l'esprit l'idée que des protocoles supplémentaires peuvent à l'avenir être conçus avec des caractéristiques de désignation inimaginables aujourd'hui.

On discute aussi des étiquettes PICS [PICS] comme technologie de remplacement au paragraphe 4.3.

On suppose que le lecteur a une formation technique limitée, aussi quelques informations de base sont incluses ci-dessous. Dans certains cas, les descriptions sont simplifiées et les détails omis.

Cette discussion technique minimise les problèmes de définition. Cependant, il est quand même nécessaire, pour évaluer certaines considérations techniques, d'avoir une estimation de la somme de catégorisations qui serait nécessaire pour un système de censure mondiale réaliste. Il n'y a pas d'espoir d'accord sur ce point. Pour notre propos, nous allons arbitrairement supposer que la population mondiale consiste en approximativement 90 000 communautés qui se recoupent, chacune d'entre elles ayant une catégorisation d'intérêt différente. De plus, on supposera arbitrairement que certains schémas de codage non spécifiés mais habiles permettent une catégorisation mondiale appropriée de toutes les informations avec une étiquette de 300 bits. Certains vont dire qu'une étiquette de 300 bits est trop grosse, d'autres que c'est trop peu. Sans rentrer dans ces considérations, nous l'utiliserons pour certaines évaluations techniques.

4.1 Filtrage des contenus en utilisant des noms

La partie la plus visible à l'utilisateur du système de désignation et d'adressage de l'Internet est le système des noms de domaines [RFC1034], [RFC1035]. Les noms de domaines sont des séquence d'étiquettes séparées par des points, telles que aol.com, world.std.com, www.rosslynchapel.org.uk, ou ftp.gnu.lcs.mit.edu [RFC1035], [RFC1591], [RFC2606]. Les noms de domaines forment une partie importante de la plupart des adresses de la Toile mondiale ou des URL [RFC2396], qui apparaissent couramment après "//". La sécurité pour le système des noms de domaines est normalisée [RFC2535], mais n'a pas encore été déployée de façon significative.

Les noms de domaines désignent des nœuds dans une base de données répartie mondialement par délégation hiérarchisée. Des informations d'une grande variété peuvent être mémorisées à ces nœuds, y compris les adresses IP des machines du réseau (voir au paragraphe 4.2) les informations de livraison de la messagerie, et d'autres types d'informations. Ainsi, les données mémorisées à foo.exemple.com pourraient être les informations sur le numéro pour envoyer des données à une certaine machine, qui sera utilisé si on essaye de joindre <http://foo.exemple.com>, le nom d'un ordinateur (disons mailhost.exemple.com) pour traiter la messagerie adressée à quelqu'un à "@foo.exemple.com", et/ou d'autres informations.

Il y a aussi d'autres systèmes de désignation qui sont utilisés, tels que les noms de groupes de nouvelles et les noms de canal des relais de causette de l'Internet (IRC, *Internet Relay Chat*).

L'idée d'étiquetage habituellement présentée est de réserver un nom de niveau supérieur, tel que .sex ou .xxx pour le matériel "pour adultes" et/ou .kids pour le matériel "sain". Les problèmes techniques et linguistiques que cela pose sont décrits dans les paragraphes qui suivent.

4.1.1 Problèmes linguistiques

Quand on utilise un étiquetage par nom, le premier problème est de quel langage tire t-on les noms à imposer ? Les mots et acronymes peuvent avoir des significations très différentes dans différents langages et la probabilité de confusion est multipliée quand on considère les collisions phonétiques.

À titre d'exemple de possibles problèmes, noter que pendant plusieurs années le gouvernement du Turkménistan a suspendu les enregistrements de nouvelles dans ".tm", qui avait été précédemment une source de revenus, parce que certains des noms de domaine de second niveau enregistrés pouvaient devenir problématiques. En particulier, leur page d'accueil du site <http://www.nic.tm> disait :
Déclaration de .TM NIC
"La demande d'inscription au registre .TM a été époustouflante. Des milliers de noms ont été enregistrés de partout dans le monde. Cependant, certains des noms enregistres pourraient être obscènes selon les lois du Turkménistan, et il en résulte que le registre .TM NIC révise sa politique de noms pour les futurs enregistrements. .TM NIC a suspendu les enregistrements jusqu'à ce qu'une nouvelle politique puisse être mise en œuvre. Nous espérons rouvrir bientôt."

Il y a approximativement 6 000 langages utilisés dans le monde d'aujourd'hui, mais on estime que cela devrait décliner aux environs de 3 000 vers l'an 2100.

4.1.2 Explosion du nombre des noms de domaine de niveau supérieur (TLD)

Un important aspect de la conception du système des noms de domaines (DNS, *Domain Name System*) est la hiérarchie de délégation de la maintenance des données. Le DNS ne fonctionne réellement, et n'a été capable de s'adapter sur les cinq ordres de grandeur de sa croissance depuis son déploiement initial que grâce à cette délégation.

Le premier problème est qu'on s'attend à ce que la plupart des ordinateurs ou sites de la Toile soient un mélange de matériel, dont seulement une partie devraient avoir une classification spéciale. Utiliser des noms de domaine de niveau supérieur (TLD, *top level domain*) multiplie le nombre de zones du DNS dont le site doit se soucier. Par exemple, supposons que le site a déjà plus ou moins trié son matériel en piles "kids", "normal", et "adulte". Sans étiquette de TLD spécial, il peut les mémoriser sous kids.exemple.net, adulte.exemple.net, et autre.exemple.net, par exemple. Cela va seulement exiger la maintenance de la seule zone exemple.net des entrées de la base de données. Avec un étiquetage des TLD spéciaux, au moins exemple.net (pour le matériel normal) exemple.net.sex, et exemple.net.kids devront être entretenus, qui sont dans trois zones séparées, dans des parties différentes de l'arborescence du DNS, sous trois délégations distinctes.

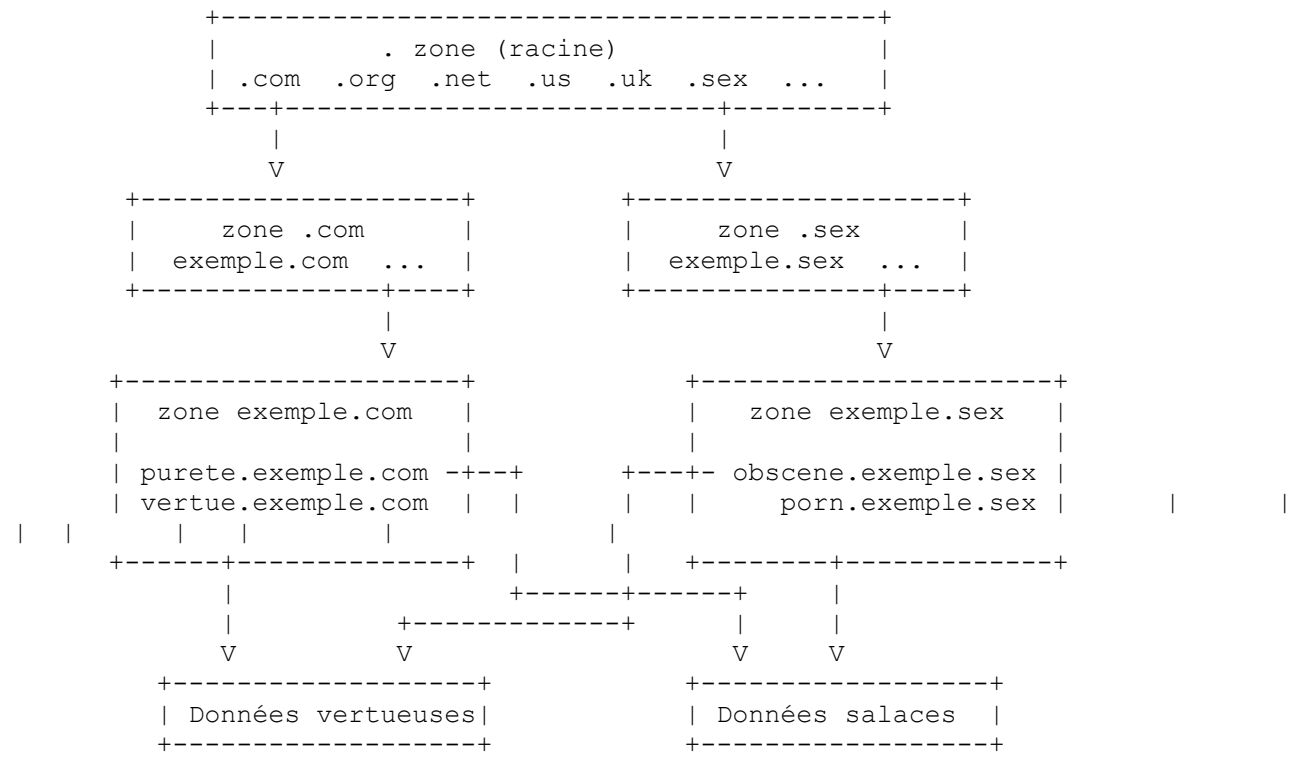
Lorsque le nombre des catégories s'étend, le nombre de combinaisons de catégories explose, et cela devient rapidement complètement ingérable. Si il faut 300 bits d'étiquetage, le système pourrait, en théorie, avoir besoin de 2**300 catégories de noms, ce qui est impossible. Aucun site individuel n'aurait besoin d'utiliser toutes les catégories, et les noms de domaine de catégorie ne devraient pas être tous des noms de niveau supérieur. Mais ce serait quand même un cauchemar ingérable.

4.1.3 On ne peut pas contrôler quels noms pointent sur vous !

Les fournisseurs de données sur l'Internet ne peuvent empêcher personne de créer des noms qui pointent sur l'adresse IP de leur ordinateur avec des noms de domaine trompeurs.

Le système du DNS fonctionne comme une base de données. Il associe certaines données, appelés enregistrements de ressource (RR, *resource record*) à des noms de domaine. En particulier, il peut associer des enregistrements de ressource d'adresse IP à des noms de domaines. Par exemple, quand on clique sur un URL, c'est généralement un nom de domaine qui au sein de cet URL est recherché dans le DNS. L'adresse résultante est alors utilisée pour adresser les paquets envoyés de votre navigateur de la Toile ou autre logiciel au serveur ou à l'homologue.

Rappelons nous ce que nous disions au paragraphe 4.1.1 sur la délégation hiérarchisée. Le contrôle est délégué et toute personne qui contrôle une zone de données du DNS, disons exemple.com, peut insérer des données à ce nom ou à tout nom plus profond (excepté dans la mesure où il a délégué à quelqu'un d'autre une partie de l'espace de noms plus profond). Donc, le contrôleur de exemple.com peut insérer des données de sorte que purete.exemple.com a, associé à lui, la même adresse d'ordinateur, qui est associée à www.obscene.exemple.sex. Cela amène toutes les références à purete.exemple.com à utiliser l'adresse IP associée qui est la même que celle du site de la Toile de www.obscene.exemple.sex. Le gestionnaire de ce site hypothétique de la Toile, qui contrôle la zone obscene.exemple.xxx, n'a pas de contrôle sur la zone DNS exemple.com. Il est techniquement incapable de l'obliger à se conformer à une quelconque loi d'étiquetage ".sex". Autrement, quelqu'un pourrait créer un nom se conformant à une exigence d'étiquetage comme "pour adulte", comme foo.truc.sex, qui pointe en fait sur le site parfaitement anodin de quelqu'un d'autre, peut-être afin de polluer l'étiquetage. Voir le diagramme ci-dessous. Chaque "zone" pourrait être hébergée sur un ensemble différent d'ordinateurs physiques.



4.1.4 Difficultés particulières des protocoles

Des considérations supplémentaires se rapportent à des protocoles particuliers. On n'en examinera ici que quelques uns. Les deux premiers, la messagerie électronique et la Toile mondiale, utilisent l'adressage de nom de domaine. Les deux suivants, les nouvelles du réseau et l'IRC, utilisent des espaces de noms différents et illustrent d'autres problèmes techniques de l'étiquetage fondés sur le nom.

4.1.4.1 Messagerie électronique (SMTP)

Les outils standard de l'Internet ne fournissent aucun moyen d'empêcher les utilisateurs de mettre des noms de domaines arbitraires à l'intérieur de leurs en-têtes de messagerie.
Le protocole standard de messagerie électronique Internet sépare les informations "d'enveloppe" des informations de contenu [RFC2821], [RFC2822]. Les informations d'enveloppe indiquent d'où un message prétend avoir son origine et à qui il devrait être livré. Le contenu a des champs commençant par des étiquettes telles que "From:" et "To:", mais ces champs de contenu n'ont en fait aucun effet et peuvent être arbitrairement falsifiés en utilisant un logiciel simple, normalement disponible, en accédant par telnet à l'accès SMTP sur un serveur de messagerie. Les champs de contenu ne sont pas comparés aux champs de l'enveloppe. Exiger d'eux qu'ils soient les mêmes reviendrait à exiger que les lettres postées dans une boîte aux lettres désignent cette boîte aux lettres comme adresse de retour et à n'admettre que les adresses de retour résidentielles ou d'affaires sur les courriers collectés par le service des postes à partir de cette résidence ou entreprise.

Bien que des clients de messagerie différents affichent différemment les informations d'enveloppe et les en-têtes à partir du contenu des messages électroniques, généralement les champs de contenu principaux reçoivent la préséance. Donc, bien que ce ne soit pas exactement la même chose que l'étiquetage du contenu, on devrait noter qu'il est trivial d'envoyer des messages à quelqu'un avec des noms de domaines arbitraires dans l'adresse de message qui apparaît dans les en-têtes From et To, etc.

Il est aussi facile de programmer un hôte pour envoyer des messages à une adresse de messagerie électronique ou à une liste de diffusion. Les messages envoyés avec les outils de messagerie normaux à ce transmetteur vont automatiquement avoir des en-têtes de contenu qui reflètent le nom du transmetteur, mais celui-ci va changer les informations d'enveloppe et provoquer en fait l'envoi du message à l'adresse de messagerie de destination transmetteuse.

Par exemple, (avec des noms déguisés) supposons qu'il y a une liste de diffusion sociale innocents@foo.exemple.org, et que quelqu'un établit un transmetteur à tortureurs-de-chats@autre.exemple. Les messages envoyés au transmetteur sont transmis et apparaissent sur la liste de diffusion innocents mais avec un en-tête "To: tortureurs-de-chats@autre.exemple" dans son corps, au lieu de son en-tête de contenu usuel "To: innocents@foo.exemple.org". Le logiciel de lecture de la

messagerie affiche alors l'en-tête tortureurs-de-chats. Des choses similaires peuvent être faites en utilisant le dispositif "bcc" *(blind courtesy copy)* ou "copie conforme invisible" de la messagerie Internet.

Il y a des travaux en cours sur la sécurisation de la messagerie électronique ; cependant, de tels efforts ne permettent à présent que de vérifier si une entité particulière est ou non l'auteur réel du message. Lorsque l'authentification est fournie, on ajoute un troisième type d'adresse "From" aux adresses "From" d'enveloppe et de contenu, mais cela n'a rien à voir avec le contrôle ou l'authentification des noms de domaines dans le contenu du message.

4.1.4.2 Accès à la Toile (HTTP)

Comme les serveurs et navigateurs modernes prennent en charge HTTP 1.1 [RFC2616], le nom de domaine utilisé pour accéder au site est disponible. Donc, les sites de la Toile avec des noms de domaines différents peuvent être atteints même si ils sont sur la même machine à la même adresse IP. C'est un petit plus pour l'étiquetage fondé sur le nom car des catégories différentes d'informations sur le même ordinateur peuvent être établies par un accès via des noms de domaines différents. Mais pour un ordinateur avec une diversité de données raisonnable, l'explosion du nombre d'essai de désigner différemment tous les types de données exigerait un nombre de noms ingérable.

Avec la version précédente de HTTP 1.0 [RFC1945], lorsque une demande à la Toile était envoyée à une machine de serveur, le nom de domaine original utilisé dans l'URI n'était pas inclus.

D'un autre côté, la Toile a la transmission automatique. Donc, quand on essaye d'accéder à des données d'un nom de domaine particulier, le serveur peut rediriger votre navigateur, de façon temporaire ou permanente, sur un nom différent, ou il peut vous rediriger sur une adresse IP numérique de façon à contourner le filtrage par nom.

4.1.4.3 Nouvelles (NNTP)

Les nouvelles du réseau [RFC0977], [RFC2980] utilisent des noms de groupes de nouvelles structurés de façon hiérarchique qui sont en apparence similaires aux noms de domaines, sauf que l'étiquette de plus fort poids est sur la gauche et celle de moindre poids est sur la droite, à l'opposé des noms de domaines. Cependant, bien que les noms soient structurés hiérarchiquement, il n'y a pas de contrôle central. Au lieu de cela, les serveurs de nouvelles se connectent périodiquement aux autres serveurs de nouvelles qui ont accepté d'échanger des messages avec eux et ils se mettent à jour les uns les autres avec les seuls groupes de nouvelles avec lesquels ils souhaitent échanger des messages.

Bien que les zones hiérarchiques du système des noms de domaines soient gérées localement, elles doivent être accessibles en commençant aux serveurs racines de niveau supérieur qui sont à leur tour plus ou moins contrôlés par l'ICANN et le Ministère américain du Commerce. Comme il n'y a pas un tel point central ou de tels points centraux dans le monde des nouvelles du réseau, toute paire ou ensemble plus large de serveurs de nouvelles n'importe où dans le monde peut se mettre d'accord pour échanger des messages de nouvelles sous les noms de groupe de nouvelles qu'il leur plait, y compris de dupliquer ceux qui existent ailleurs dans le réseau, rendant un contrôle central ou même toute influence centrale, virtuellement impossible. En fait, au sein de certaines parties de l'espace de noms des groupes de nouvelles sur certains serveurs, n'importe qui peut créer des nouveaux groupes de nouvelles avec n'importe quel nom.

Même si les noms des groupes de nouvelles pouvaient être contrôlés, le contenu des messages est déterminé par les envoyeurs. Bien que certains groupes aient un modérateur, la plupart n'en ont pas. Des messages "Annuler" peuvent être envoyés pour les messages de nouvelles, mais ce mécanisme a fait l'objet d'abus, de sorte que certains serveurs sont configurés pour ignorer les annulations. Dans tous les cas, le message peut avoir été distribué sur un énorme nombre d'ordinateurs dans le monde entier avant l'envoi de n'importe quelle annulation.

Et bien sûr, établir un étiquetage de 300 bits des noms de groupes de nouvelles est tout aussi impossible que pour les noms de domaines.

4.1.4.4 Relais de causette Internet

Le relais de causette Internet [RFC2810] à [RFC2813] est un autre exemple de service qui utilise un espace de noms différent. Il utilise un espace à un seul niveau de "nom de canal" qui est significatif au sein d'un réseau particulier de serveurs IRC. Comme il n'est pas hiérarchique, chaque serveur doit savoir tous les noms, ce qui limite la taille d'un réseau de serveurs.

Comme avec les noms de groupes de nouvelles, le fait que les noms des canaux IRC soit une décision locale, non soumise à, ni accessible par une "racine" mondiale, rend le contrôle politique centralisé virtuellement impossible.

4.2 Filtrage du contenu en utilisant l'adressage IP

Une caractéristique clé du protocole Internet (IP) sur laquelle se fonde l'Internet est qu'il répartit les données en "paquets". Ces paquets sont traités individuellement et acheminés de leur source à leur destination. Chaque paquet porte une adresse numérique pour le point de destination auquel l'Internet va essayer de livrer le paquet.

(L'utilisateur final ne voit normalement pas ces adresses numériques mais a affaire à la place aux "noms de domaines" comme on l'a décrit au paragraphe 4.1 ci-dessus.)

Le système prédominant d'adresses numériques actuellement utilisé est appelé IPv4, ou protocole Internet version 4, qui fournit des adresses de 32 bits[RFC0791]. Il y a une migration croissante vers le plus récent IPv6 [RFC2460], qui fournit des adresses de 128 bits [RFC2373], [RFC2374].

Les paquets peuvent subir des modifications malveillantes pendant le transit mais le résultat le plus courant en est le déni de service.

Un des problèmes de l'utilisation de l'adressage pour le filtrage du contenu est que c'est une technique très grossière. Les adresses IP se réfèrent aux interfaces réseau, qui correspondent habituellement à des systèmes d'ordinateurs entiers qui peuvent héberger de multiples pages de la Toile, des ensembles de fichiers, etc., dont on souhaite bloquer ou activer seulement une petite partie. De plus en plus, une seule adresse IP peut correspondre à une boîte de traducteur d'adresse réseau (NAT, *Network Address Translator*) [RFC2663] qui cache plusieurs ordinateurs derrière elle, bien que dans ce cas, ces ordinateurs ne soient généralement pas des serveurs.

Cependant, même au delà de ce problème de granularité grossière, les contraintes pratiques de l'acheminement hiérarchique rendent impossible l'allocation même d'un seul bit d'adresse IPv4 ou d'un nombre significatif de bits d'adresse IPv6.

4.2.1 Acheminement hiérarchique

Les adresses IP sont techniquement inappropriées pour le filtrage de contenu parce que leur allocation est intimement liée à l'acheminement et à la topologie du réseau.

Comme les paquets de données s'écoulent à travers l'Internet, on doit prendre des décisions sur la façon de les transmettre "vers" leur destination. On fait cela en comparant les bits initiaux de l'adresse de destination du paquet aux entrées d'un "tableau d'acheminement" et en transmettant les paquets comme indiqué par l'entrée du tableau qui a la plus longue correspondance de préfixe.

Bien que l'Internet soit en fait un maillage, si, pour simplifier, on le considère comme ayant un cœur de réseau central à son "sommet", un paquet est normalement acheminé comme suit :

Le code de réseautage local regarde dans son tableau d'acheminement pour déterminer si le paquet devrait être envoyé directement à un autre ordinateur sur le réseau "local", à un routeur pour qu'il le transmette spécialement à un autre réseau du voisinage, ou l'achemine "vers le haut" à un routeur par "défaut" pour qu'il soit transmis au réseau d'un fournisseur de service d'un niveau supérieur. Si la destination du paquet est "assez lointaine", il va finalement être transmis jusqu'à un routeur du cœur de réseau. Un tel routeur ne peut pas envoyer le paquet "vers le haut" car il est au sommet, ou zone "sans défaut", et doit avoir un tableau complet des autres routeurs de niveau supérieur dans lesquels envoyer le paquet. Actuellement, de tels routeurs de niveau supérieur sont des appareils très gros et très coûteux. Ils doivent être capables d'entretenir des tableaux de dizaines de milliers de chemins. Lorsque le paquet arrive au routeur de niveau supérieur de la partie du réseau eu sein de laquelle se trouve la destination, il est transmis "vers le bas" aux routeurs successifs qui sont de plus en plus spécifiques et locaux jusqu'à ce qu'il arrive finalement à un routeur sur le réseau local où se trouve l'adresse de destination. Ce routeur local envoie le paquet directement à l'ordinateur de destination.

Comme toutes ces décisions d'acheminement sont prises sur la base de la plus longue correspondance de préfixe, on peut voir que les adresses IP ne sont pas des noms ou étiquettes générales, mais sont associées de façon critique et intime à la topologie réelle et à la structure d'acheminement du réseau. Si elles étaient allouées au hasard, les routeurs seraient obligés de se souvenir de tellement de chemins spécifiques pour les adresses spécifiques que cela excèderait de loin les capacités techniques actuelles de la conception des routeurs. L'Internet serait condamné à une déroute fatale et ne fonctionnerait plus.

On devrait aussi noter qu'il y a une certaine inefficacité de l'allocation à chaque niveau de la hiérarchie [RFC1715]. Généralement, les allocations sont une puissance de deux adresses et lorsque les exigences augmentent et/ou se réduisent, il n'est pas possible d'utiliser chaque adresse.

(La description simplifiée ci-dessus ignore le multi rattachement et de nombreux autres détails.)

4.2.2 Adresses IP version 4

Il n'y a absolument aucun moyen pratique de réallouer même un seul bit des adresses mondiales Internet IPv4 pour une utilisation de filtrage du contenu. Ces adresses sont en rupture de stock. Une telle allocation diviserait en effet le nombre des adresses disponibles par deux. Il n'y a déjà pas assez d'adresses, même sans parler de l'inefficacité de l'allocation hiérarchisée [RFC1715] et de l'acheminement pour faire cela. Même si ils étaient assez nombreux, les numéros actuels n'ont pas été alloués dans ce but, de sorte que serait nécessaire le dénumérotage de toutes les organisations qui ont des hôtes sur l'Internet, une tâche herculéenne dont le coût se chiffrerait en milliards de dollars.

Même si on surmontait ces problèmes, l'allocation de même un seul bit près du sommet des bits de l'adresse doublerait probablement le nombre de chemins dans la zone sans défaut. Cela excèderait la capacité des routeurs actuels et exigerait la mise au rebut de milliers d'entre eux pour de nouveaux routeurs qui n'existent pas encore, pour un coût gargantuesque. L'allocation d'un bit près du fond des bits d'adresse exigerait une reconfiguration locale à l'échelle mondiale dont la mise en application obligatoire serait impraticable, même si le bit était disponible.

Et tout cela pour l'allocation d'un seul bit à l'étiquetage du contenu, sans parler de plus qu'un seul. Et on a supposé qu'il faudrait en fait 300 bits, soit plus qu'il n'y en a déjà !

L'idée est mort née.

4.2.3 Adresses IP version 6

IPv6 fournit des champs d'adresse de 128 bits [RFC2373], [RFC2374]. De plus, l'allocation des adresses IPv6 est dans son enfance. Donc, l'allocation disons, d'un bit de l'adresse IPv6 pour l'étiquetage est concevable.

Cependant, comme on l'a exposé plus haut (au paragraphe 4.2.1.) chaque bit de poids fort alloué pour l'étiquetage double le coût imposé au système d'acheminement. Allouer un bit doublerait généralement la taille des tableaux d'acheminement.

Allouer deux bits les multiplierait par quatre. Allouer les 300 bits qu'on suppose nécessaire pour un étiquetage mondial réaliste est logiquement impossible pour IPv6, 300 étant bien plus grand que 128, et si c'était possible, il en résulterait des tailles de tableau d'acheminement techniquement irréalisables. Même d'allouer, disons 20 bits, si c'était possible, multiplierait les tailles des tableaux d'un impossible facteur de un million.

Allouer les bits de poids faible pose aussi des problèmes. Il y a des propositions techniques qui utilisent les 64 derniers bits d'une manière incompatible avec leur utilisation pour des étiquettes [RFC2374]. De sorte que ce devra probablement être les "bits du milieu" (en fait les bits de moindre poids de la moitié supérieure). Comme avec IPv4, il serait impossible de mettre cela en application au niveau mondial. Si c'était possible, un ou deux bits pourraient être alloués là, ce qui serait clairement inadéquat.

4.3 Étiquettes PICS

Les étiquettes de la plate-forme pour la sélection des contenus Internet (PICS, *Platform for Internet Content Selection*) sont un système généralisé de fourniture de "classements" pour le matériel accessible par Internet. Les documents PICS [PICS] devraient être consultés pour les détails. En général, PICS suppose un nombre arbitrairement grand de services de classement et de systèmes de classement. Chaque service et système est identifié par un URL.

Il serait assez raisonnable d'avoir plusieurs services PICS qui, en les agrégeant, fournissent 300 bits ou plus d'information d'étiquettes. Il pourrait y avoir un service PICS pour chaque communauté d'intérêt. Cette sorte de technologie est réellement la seule façon raisonnable de faire des catégorisations ou de l'étiquetage des matériels disponibles dans un monde divers et dynamique.

Bien que de tels services d'étiquettes PICS puissent être utilisés pour distribuer des catégories promulguées par la censure gouvernementale, par exemple, on ne sait pas trop si c'est pire que la censure gouvernementale via des pare-feu nationaux.
Un système de classement PICS est essentiellement une définition d'une ou plusieurs dimensions et la gamme numérique des valeurs qui peuvent être allouées dans chaque dimension à un objet classé. Un service est une source d'étiquettes où une étiquette comporte les classements réels. Les classements sont spécifiques ou génériques. Un classement spécifique ne s'applique qu'au matériel d'un URL particulier [RFC2396] et ne couvre pas ce qui est référencé à partir de lui, y compris même les fichiers d'images. Un classement générique s'applique à l'URL spécifié et à tous les URL pour lesquels l'URL en cause est un préfixe

Un exemple simplifié d'étiquette pourrait ressemble à ceci :

(PICS-1.1 "http://service-de-classement-des-films.exemple.net" étiquettes pour "ftp://films.exemple.sex/film-moche" classements (sexe 6 violence 1 langage 8 drogues 2 satanisme 0))

Les descriptions de système de classement lisibles par la machine comportent la gamme de valeurs et l'ensemble des dimensions fournies. Des informations supplémentaires, telles que les débuts et fin de période de validité, peuvent être incorporées dans les étiquettes.

Les étiquettes peuvent actuellement être rendues disponibles de trois façons : (1) incorporées dans HTML, (2) fournies avec les données dans une réponse HTTP, et (3) séparément à partie d'un tiers. Si il faut que le contenu ait les étiquettes incorporées ou transmises par la source lorsque les données sont retournées, comme dans les deux premières façons énumérées ci-dessus, cela soulève les problèmes de la granularité de la catégorisation et du langage forcé. Cependant, si on les utilise de la troisième façon par laquelle un tiers distinct détermine et fournit les étiquettes pour le contenu, et si les utilisateurs sont libres de choisir le ou les tiers qu'ils souhaitent consulter, cela peut prendre en charge une myriade de catégories, éditeurs, et évaluateurs, existant en parallèle.

Les signatures numériques sont disponible pour sécuriser les étiquettes PICS [PICS].

5. Considérations pour la sécurité

Tout schéma d'étiquetage ou de catégorisation doit supposer qu'il y aura des tentatives délibérées pour causer un étiquetage ou catégorisation incorrectes des données. Cela peut être dû à la perception d'un certain avantage d'un étiquetage particulier ou simplement pour enrayer le système. Après tout, si les sources étiquetaient toujours précisément et commodément les informations envoyées, la sécurité serait beaucoup plus facile [RFC3514]. De telles considérations de mise en application sont exposées en conjonction avec les divers mécanismes mentionnés dans ce document.

6. Conclusions

L'idée qu'un seul nom de domaine de niveau supérieur, tel que .sex, ou qu'un seul bit d'adresse IP, pourrait être alloué et devenir l'hébergement obligatoire du matériel "pour adultes" ou "choquant" dans le monde entier est un non-sens légal et technique.

Un accord mondial sur la sorte de matériel qui devrait être enfermé dans un tel ghetto est impossible. Dans le contexte mondial actuel, l'utilisation d'une seule catégorie ou d'un petit nombre de catégories est absurde. La mise en œuvre d'une étiquette de taille raisonnable qui pourrait satisfaire aux critères de nombreuses communautés dans le monde, comme 300 bits, est techniquement impossible au niveau du nom de domaine ou de l'adresse IP et le restera dans l'avenir prévisible. À côté de l'impossibilité technique, une telle obligation serait une entrave illégale à la liberté d'expression devant certaines juridictions, ainsi que la cause de sévères problèmes linguistiques pour les noms de domaines ou autres chaînes de caractères.

Cependant, l'idée d'une cohorte de réviseurs indépendants, dont certains pourraient être des agences gouvernementales, et la capacité que ceux-ci accèdent aux informations pour choisir et utiliser des classements alloués par de tels réviseurs, est possible.

7. Références

[BT] "British Telecom comments to U.S. Commerce Department", 20 février 1998, <http://www.ntia.doc.gov/ntiahome/domainname/130dftmail/BT.htm >

[CDA] "Reno v. American Civil Liberties Union", 117 S.Ct. 2329, 26 juin 1997,

[COPAREPORT] "Final Report of the COPA Commission to the U.S. Congress", 20 ctobre 2000, <http://www.copacommission.org/report/newtopleveldomain.shtml >

[GAO] "GAO Report OGC-00-33R", 7 juillet 2000, < http://www.gao.gov/new.items/og00033r.pdf >

[GTLD-MOU] "GTLD-MOU Policy Oversight committee RFC 97-02", 13 septembre 1997, < http://www.gtld-mou.org/docs/notice-97-02.html >

[HOUSEREPORT] "U.S. House Commerce Committee report", 105th Congress, 5 octobre 1998. <http://www.epic.org/free_speech/censorship/hr3783-report.html >

[ICM-REGISTRY]"Request for reconsideration from ICM Registry to ICANN", 15 décembre 2000, <http://www.icann.org/committees/reconsideration/icm-request-16dec00.htm >

[LIEBERMAN] "Testimony of Senator Joe Lieberman before Children's Online Protection Act Commission", 5 juin 2000 < http://www.senate.gov/~lieberman/press/00/06/2000608958.html >

[PICS] Platform for Internet Content Selection PICS 1.1 "Rating Services and Rating Systems -- and Their Machine Readable Descriptions" à http://www.w3.org/TR/REC-PICS-services >, octobre 1996.

[PICS 1.1] "Label Distribution -- Label Syntax and Communication Protocols" <http://www.w3.org/TR/REC-PICS-labels >, octobre 1996.
PICSRules 1.1 Specification < http://www.w3.org/TR/REC-PICSRules >, décembre 1997.

PICS Signed Labels (DSIG) 1.0 Specification < http://www.w3.org/TR/REC-DSig-label/ >, mai 1998.

[RFC0791] J. Postel, éd., "Protocole Internet - Spécification du protocole du programme Internet DARPA", STD 5, septembre 1981.

[RFC0977] B. Kantor et P. Lapsley, "Protocole de transfert des nouvelles du réseau", février 1986. *(Obsolète, voir la RFC 3977)*

[RFC1034] P. Mockapetris, "Noms de domaines - Concepts et facilités", STD 13, novembre 1987.

[RFC1035] P. Mockapetris, "Noms de domaines – Mise en œuvre et spécification", STD 13, novembre 1987.

[RFC1591] J. Postel, "Structure et délégation du système de noms de domaine", mars 1994. *(Information)*

[RFC1715] C. Huitema, "Ratio H d'efficacité d'allocation d'adresse", novembre 1994. *(Info., MàJ par la RFC 3194)*

[RFC1945] T. Berners-Lee, R. Fielding, H. Frystyk, "Protocole de transfert Hypertext -- HTTP/1.0", mai 1996. *(Information)*

[RFC2373] R. Hinden, S. Deering, "Architecture d'adressage IP version 6", juillet 1998. (*Obsolète, voir* RFC4291) *(PS)*

[RFC2374] R. Hinden, M. O'Dell, S. Deering, "Format mondial d'adresse d'envoi individuel IPv6 agrégeable", juillet 1998. (*Obsolète, voir* RFC3587) *(Historique)*

[RFC2396] T. Berners-Lee, R. Fielding et L. Masinter, "Identifiants de ressource uniformes (URI) : Syntaxe générique", août 1998. (*Obsolète, voir* RFC3986, *STD66*)

[RFC2460] S. Deering et R. Hinden, "Spécification du protocole Internet, version 6 (IPv6) ", décembre 1998. *(MàJ par 5095, D.S)*

[RFC2535] D. Eastlake, 3rd, "Extensions de sécurité du système des noms de domaines", mars 1999. (*Obsolète, voir* RFC4033, RFC4034, RFC4035) *(P.S.)*

[RFC2606] D. Eastlake 3rd et A. Panitz, "Noms réservés de niveau supérieur du DNS", BCP 32, juin 1999.

[RFC2616] R. Fielding et autres, "Protocole de transfert hypertexte -- HTTP/1.1", juin 1999. *(D.S., MàJ par 2817)*

[RFC2663] P. Srisuresh, M. Holdrege, "Terminologie et considérations sur les traducteurs d'adresse réseau IP (NAT)", août 1999. *(Information)*

[RFC2810] C. Kalt, "Relais pour la causette Internet : architecture", avril 2000. *(Information)*

[RFC2811] C. Kalt, "Relais pour la causette Internet : gestion de canal", avril 2000. *(Information)*

[RFC2812] C. Kalt, "Relais pour la causette Internet : protocole client", avril 2000. *(Information)*

[RFC2813] C. Kalt, "Relais pour la causette Internet : protocole serveur", avril 2000. *(Information)*

[RFC2821] J. Klensin, éditeur, "Protocole simple de transfert de messagerie", STD 10, avril 2001. *(Obsolète, voir RFC5321)*

[RFC2822] P. Resnick, "Format de message Internet", avril 2001. *(Remplace la RFC0822, STD 11, Remplacée par RFC5322)*

[RFC2854] D. Connelly et L. Masinter, "Type de support 'text/html'", juin 2000. *(Information)*

[RFC2980] S. Barber, "Extensions communes à NNTP", octobre 2000. (*MàJ par* RFC3977, RFC4643, RFC4644) *(Information)*

[RFC3513] R. Hinden et S. Deering, "Architecture d'adressage du protocole Internet version 6 (IPv6)", avril 2003. (*Obs. voir* RFC4291)

[RFC3514] S. Bellovin, "Le fanion Sécurité dans l'en-tête IPv4", 1er avril 2003. *(Information)*

[WARSHAVSKY] "Congress weighs Net porn bills," article dans CNET, 10 février 1998, < http://news.cnet.com/news/0-1005-200-326435.html >

8. Remerciements

De vifs remerciements pour sa contribution et ses efforts à Declan McCullagh qui a rédigé la substance des sections 2 et 3 du présent document.

9. Adresse de l'auteur

Donald E. Eastlake 3rd
Motorola Laboratories
155 Beaver Street
Milford, MA 01757
USA
téléphone : +1-508-786-7554 (bur)
+1-508-634-2066 (dom)
mél : dee3@torque.pothole.com

10. Déclaration de droits de reproduction

garantie que l'utilisation de l'information ici présente n'enfreindra aucun droit ou aucune garantie implicite de commercialisation ou d'adaptation a un objet particulier.

Remerciement
Le financement de la fonction d’édition des RFC est actuellement fourni par la Internet Society.