LabsHadopi

Livre vert sur les techniques de filtrage

DATE	Mercredi 14 décembre 2011
ÉMETTEUR	Lab Réseaux et Techniques
VERSION	2

La question de la circulation et de l'accès aux contenus en ligne est un enjeu majeur du futur d'Internet. Elle concerne tous les usagers d'Internet qu'ils soient chevronnés ou occasionnels. En effet, on évolue vers une massification de l'accès à Internet et vers une inflation des demandes de connexion. Comme le détaille l'ARCEP :
« Au 30 septembre 2011, le nombre d'abonnements internet à haut et très haut débit sur réseaux fixes atteint 22,4 millions, soit une croissance nette de 340 000 abonnements par rapport à la fin du deuxième trimestre 2011. Sur un an, l'accroissement net s'élève à 1,5 million (+7 %), soit un rythme de croissance stable depuis un an. » [1]

Cette massification du recours à Internet connaît de nombreux avantages ; comme le notent dans leur rapport Mmes les députées Erhel et de la Raudière, Internet est à l'origine d'avancées sociales incontestables : « démocratisation de l'accès au savoir, participation des citoyens aux débats politiques, promotion des idées, diffusion rapide des avancées technologiques, commercialisation des produits et services, coopération et ingénierie institutionnelle, développement économique. »[2]

De fait, il convient de s'interroger sur les conditions de préservation et de pérennisation de ces avancées. Internet touche en effet l'ensemble de la population, toutes classes d'âge confondues et il est indéniable que certains contenus sont, par exemple, malgré leur accès facile, peu appropriés à de jeunes individus ou sont en complète illégalité, dans le cas de la pédopornographie par exemple. Protéger la qualité de l'Internet signifie donc s'interroger sur les modalités de son appropriation. Faut-il en laisser l'entière initiative à l'internaute ? Faut-il privilégier des modalités de filtrages adéquates ?

Poser ces questions aujourd'hui est d'autant plus nécessaire que l'irruption de l'Internet au sein de tous nos échanges, au cœur de notre vie quotidienne amplifie certaines « *dérives* » et conduit généralement à la volonté, somme toute assez naturelle pour certains, de *« vouloir reprendre le contrôle »*. Mais faute de bien comprendre la nature de cette transformation, ses causes et ses fondamentaux radicalement différents de notre monde habituel « régalien », on en vient trop souvent à mal réfléchir par manque d'information, par crainte de devoir apporter une réponse rapide, par peur de ne bientôt plus pouvoir contrôler.

Pendant longtemps, les mesures techniques étaient prises pour des raisons techniques par des gens qui les dominaient encore. Les décisions de filtrage ou de blocage relevaient soit

[1] HTTP://www.arcep.fr/index.php?id=11120
[2] HTTP://www.assemblee-nationale.fr/13/pdf/rap-info/i3336.pdf

de décisions de bons sens (ex : protection du réseau), soit de l'exercice d'une responsabilité citoyenne éclairée (ex : refus d'acheminer des contenus pédopornographiques), soit de décisions judiciaires. Il s'agissait d'un fonctionnement démocratique classique fondé sur le principe de subsidiarité qui, in fine, faisait de la justice l'unique arbitre de l'opportunité des demandes provenant de divers groupes défendant leurs intérêts particuliers. Or l'écosystème du net s'est grandement complexifié. Notre propos n'est pas ici de trancher pour ou contre le filtrage, mais d'apporter des éléments d'aide à la réflexion et, éventuellement à la prise de décision. Nous engagerons donc à la fois une réflexion sur la polysémie du terme « filtrage » et ses spécificités techniques, sans négliger une approche d'ordre plus sociétale sur les enjeux du filtrage dans l'écosystème numérique. Dans ce sens, nous articulerons notre exposé comme suit :

- Eléments de définition : ce que recouvrent les termes de filtrage et de blocage
- L'approche technique du filtrage : comment peut-on filtrer et que veut-on filtrer ?
- Enjeux et risques : la question de l'opportunité du filtrage engage la responsabilité individuelle de l'internaute et ne doit pas être exempte d'une bonne compréhension des risques qui lui sont afférents.

Enfin, une annexe juridique mettra en lumière le rôle éventuel de l'Hadopi dans l'approche du filtrage, comme envisagé par le législateur.

Sommaire

1 | Définitions

Le blocage et le filtrage sont des notions distinctes, qu'il convient de définir.

1.1 Le blocage

Le blocage consiste à interdire le passage d'information d'un point A à un point B. Sur Internet cela peut se traduire par l'acte de rendre impossible toute de connexion d'une machine vers un site Internet, ou bien d'empêcher un contenu d'atteindre un ordinateur personnel ou un poste informatique. L'exemple le plus célèbre de blocage sur Internet est Youtube. En effet, ce site de partage de vidéos a fait l'objet de différents blocages dans différents pays : Bangladesh[3], Chine[4], Maroc[5], Thaïlande[6], Tunisie[7] ou encore la Turquie[8]. Le blocage est une action permettant d'arrêter le trafic. Cela revient à poser un barrage sur une route.

Le blocage peut se faire de différentes manières et peut être opéré par différentes personnes, les techniques utilisées sont fonction de la personne qui souhaitera opérer un blocage ainsi que du but recherché par cette dernière.

1.2 Le filtrage

Le blocage n'est qu'un filtrage strict. Le filtrage se définit comme une limitation sélective d'accès au Réseau[9].

Si le blocage permet d'arrêter tout trafic, sans distinction, le filtrage permet de laisser passer certaines informations, mais seulement après analyses de ces dernières. Ce sont ces analyses qui permettront de déterminer quelles informations sont autorisées à arriver jusqu'à leurs destinataires.

Tout comme pour le blocage, le filtrage peut être opéré de différentes manières, par différentes entités, selon le but recherché.

Néanmoins, des moyens supérieurs à ceux mis en œuvre pour le blocage seront mobilisés, car il s'agit d'une technique plus délicate, dans la mesure où certaines informations vont être acceptées et d'autres vont être expurgées. En effet, il va falloir s'interroger sur les critères permettant de différencier les éléments et installer un mécanisme sophistiqué pour comprendre la nature du trafic. Ce sont les informations se trouvant dans le cœur même de l'infrastructure d'accès au réseau qui seront examinées.

Ainsi, toute question de filtrage ou de blocage se pose forcément en termes de finalité – destination du filtrage – de moyens mis en œuvre pour atteindre cette finalité – efficacité et

[3] HTTP://www.lepoint.fr/monde/le-bangladesh-bloque-facebook-a-cause-des-caricatures-du-prophete-mahomet-30-05-2010-460859_24.php
[4] HTTP://www.infoworld.com/d/security-central/youtube-blocked-in-china-flickr-blogspot-restored-351
[5] HTTP://www.lepetitjournal.com/content/view/15253/312/
[6] HTTP://thailande-fr.com/actu/83-la-thailande-envisage-de-poursuivre-youtube
[7] HTTP://www.lemonde.fr/technologies/article/2011/01/05/la-tunisie-tente-de-reprendre-le-controle-du-web_1461205_651865.html
[8] HTTP://info.france2.fr/sciences-tech/la-turquie-bloque-de-nouveau-youtube-65688009.html
[9] Un réseau est un regroupement de machines interconnectées pouvant communiquer les unes avec les autres.

limites – et de responsabilités de la prise de décision au regard de la destination et des moyens – subsidiarité, qui va prendre la décision de filtrer, et proportionnalité, comment l'entité ayant décidé de filtrer va procéder. Ayant établi que le blocage n'était qu'un sous-ensemble du filtrage, c'est ce dernier qui sera privilégié.

2 | Le filtrage

2.1 Efficacité d'un filtrage

Il s'agit ici de présenter les bases qui vont nous permettre de décider si un filtrage (une technique de filtrage appliquée à un contexte) est efficace ou pas. Il ne s'agit donc pas de déterminer si un filtrage est dangereux (pour les infrastructures réseau par exemple), non opportun, ou d'apporter un quelconque jugement sur celui-ci : nous souhaitons mettre en lumière son efficience.

Habituellement, l'efficacité d'un processus se définit par la capacité à atteindre l'objectif pour lequel le processus a été défini.

Dans le cas du filtrage on ne peut se contenter de cette seule définition.

En effet, lorsque l'on se prononce sur l'efficacité d'un filtrage il faut aussi prendre en considération les effets collatéraux. Ainsi le filtrage d'un site Web ne peut être considéré comme efficace si ce filtrage implique, qu'en plus du site Web que l'on désire filtrer, on empêche l'accès à d'autres sites qui n'étaient pas censés être la cible de ce filtrage.

Lorsque l'on se prononce sur l'efficacité d'un filtrage, il ne faut pas non plus négliger la facilité de contournement. Un filtrage n'est pas efficace si n'importe qui peut facilement le contourner.

L'efficacité d'un filtrage s'évalue donc en fonction de 3 critères :

- Le sous-blocage
- Le sur-blocage
- La facilité de contournement

2.1.1 Sous-blocage

On parle de sous-blocage lorsqu'un filtrage n'arrive pas à atteindre son objectif. Par exemple, on souhaite filtrer un ensemble de fichiers circulant sur un réseau pair-à-pair et on n'arrive qu'à en bloquer la moitié.

Plus le sous-blocage est important, moins le filtrage est efficace.

2.1.2 Sur-blocage

On parle de sur-blocage lorsqu'un filtrage filtre plus que ce qu'il devrait. Par exemple, on souhaite filtrer un ensemble de fichiers circulant sur un réseau pair-à-pair et on bloque la circulation de l'intégralité des fichiers circulant sur ce réseau.

Plus le sur-blocage est important, moins le filtrage est efficace.

2.1.3 Facilité de contournement

On parle de facilité de contournement lorsqu'un filtrage peut être contourné. Par exemple on souhaite filtrer un site Web par l'IP de son serveur, mais il existe un site miroir dont le serveur possède une autre IP et par conséquent le site est toujours accessible.

Plus la facilité de contournement est importante, moins le filtrage est efficace.

2.2 Coût d'un filtrage

Le coût financier d'un filtrage peut difficilement s'estimer.

Car au-delà du coût financier direct qui peut s'évaluer par :

- Le coût d'installation d'équipements/logiciels assurant le filtrage
- Le coût de découverte maintenance et mise à jour des éléments à filtrer

Il y a aussi les coûts financiers indirects difficilement estimables (ex : impact sur l'économie d'un réseau ralenti par une technique de filtrage).

Dans cette partie sur le Filtrage (partie 2 de ce document) nous nous contenterons uniquement d'estimer le coût financier direct d'un filtrage.

2.3 Le filtrage par destination (que veut-on filtrer ?)

2.3.1 Site Web ou partie de site web

Le but est, par exemple, d'empêcher l'accès à http://www.w3c.org**.** On retiendra les 5 critères de filtrage suivants :

2.3.1.1 Filtrage par IP

Le filtrage par IP consiste à interdire l'accès et les connexions au serveur identifié par les IP incriminées.

Efficacité
La facilité de contournement est existante : le site web peut changer d'adresse en quelques minutes et l'utilisateur pourrait utiliser un proxy situé dans un pays non concerné par le filtrage du site Web en question.

Le risque de sous-blocage est existant : le site web peut utiliser plusieurs adresses sans que celles-ci puissent être toutes infailliblement connues.

Le risque de sur-blocage est très important : il est impossible de connaitre précisément la liste des sites Web ou services Internet abrités par une même adresse IP, potentiellement plusieurs centaines voire des milliers selon la solution d'hébergement utilisée.

Conclusion : le blocage d'un site Web par adresses IP n'est que faiblement efficace.

Coût
Faible. Le blocage d'une connexion en fonction de son IP d'origine est quelque chose de couramment employé de nos jours. C'est le principe même d'un pare-feu.

2.3.1.2 Filtrage par nom de domaine

Le filtrage par nom de domaine consiste à interdire l'accès et les connexions aux serveurs

identifiés par le nom de domaine incriminé.

Efficacité
La facilité de contournement est existante : un tel blocage n'empêche aucunement l'accès au site par les adresses IP directement. De plus, changer de nom de domaine est une opération rapide et quasi triviale. Les moyens de communication du nouveau nom sont ensuite nombreux (réseaux sociaux, tchat, forums spécialisés...).

Le risque de sous-blocage est existant : un tel blocage n'empêche aucunement l'accès au site par d'autres noms de domaine.

Le risque de sur-blocage est important : plusieurs milliers de sites Web pourraient être bloqués en même temps que le site visé, par exemple le seul blocage de 'wordpress.com' entrainerait le blocage des centaines de milliers de blogs.

Conclusion : le blocage d'un site Web par nom de domaine du site n'est que faiblement efficace.

Coût
Faible.

2.3.1.3 Filtrage par URL

Le filtrage par URL consiste à inspecter les requêtes HTTP par une technique reposant sur l'utilisation de **DPI** et de filtrer toutes les requêtes vers l'URL incriminée.

L'utilisation d'un proxy par lequel transiteraient toutes les requêtes HTTP permettrait de faciliter un tel filtrage.

Un exemple de ce type de filtrage pourrait être : sur un même site www.site.fr autoriser la page correspondant à l'URL www.site.fr/page1.html et bloquer la page correspondant à l'URL www.site.fr/page2.html.

Efficacité
La facilité de contournement est existante : le filtrage par URL ne peut notamment s'appliquer aux sites auxquels on accède en HTTPS (à moins de casser la sécurité du protocole de chiffrement SSL).

Le risque de sous-blocage est important : il sera difficile de référencer l'intégralité des URL à filtrer.

Le risque de sur-blocage est minime.

Conclusion : le blocage d'un site Web par URL est relativement efficace.

Coût
Coût d'un DPI généralisé, soit extrêmement couteux (voir DPI). Réduction du coût possible par l'utilisation d'un proxy.

2.3.1.4 Filtrage dynamique (par contenu)

Le filtrage dynamique consiste à inspecter lors de l'accès au site Web, le contenu de ses pages et de filtrer selon ce que ces pages contiennent (ex : bloquer toutes les pages qui contiennent le mot « lapin »). Ce type de filtrage repose sur l'utilisation de **DPI**.

Efficacité
La facilité de contournement est existante : le filtrage dynamique ne peut notamment s'appliquer aux sites auxquels on accède en HTTPS (à moins de casser la sécurité du protocole de chiffrement SSL).

Le risque de sous-blocage est important : un tel blocage nécessite la constitution et la mise à jour d'une base de contenu (ou d'identifiant de contenus) et il sera impossible de référencer tout ce que l'on souhaite filtrer.

Le risque de sur-blocage varie de minime à important : bloquer un mot ou un terme ou une expression bloquera l'ensemble des sites comportant ce mot, ce terme ou cette expression, même s'ils ne sont pas directement visés par le filtrage, le sur-blocage sera alors important. Si le filtrage se fait sur un identifiant unique du contenu (une signature ou un hash d'un fichier par exemple) le sur-blocage sera minime.

Conclusion : le blocage dynamique d'un site Web est moyennement efficace.

Coût
Coût d'un DPI généralisé, soit extrêmement couteux (voir DPI).

2.3.1.5 Filtrage hybride IP + dynamique

Le filtrage hybride est un mélange d'un filtrage par IP puis d'un filtrage sur le contenu. Il évite de devoir effectuer du DPI sur la totalité des communications, mais nécessite un ancrage fort dans le réseau du FAI.

Efficacité
La facilité de contournement est existante : le site web peut changer d'adresse IP en quelques minutes et l'utilisateur pourrait utiliser un proxy situé dans un pays non concerné par le filtrage du site Web en question.

Les risques de sous-blocage sont ceux du filtrage par IP soit existant.

Les risques de sur-blocage sont ceux du filtrage dynamique soit variant de minime à important selon le critère de décision (mots clé, signature de fichiers, etc.).

Conclusion : le blocage dynamique d'un site Web est moyennement efficace.

Coût
Coût d'un DPI sélectif, moins couteux qu'un DPI généralisé mais couteux quand même (voir DPI).

2.3.2 Messagerie Instantanée

Le but est, par exemple, de bloquer une messagerie instantanée sur une zone géographique prédéfinie. Il existe 3 critères de filtrage qui peuvent être utilisés pour cela.

2.3.2.1 Filtrage par IP

Le filtrage par IP consiste à interdire l'accès et les connexions au serveur identifié par les IP incriminées.

Efficacité
La facilité de contournement est existante : l'utilisateur pourrait utiliser un proxy situé dans un

pays non concerné par le filtrage du système de messagerie en question.

Le risque de sous-blocage est minime.

Le risque de sur-blocage est important : bloquer par adresse IP revient à interdire totalement l'accès au système de messagerie instantanée concernée.

Conclusion le blocage d'un service de messagerie instantanée par adresses IP n'est que faiblement efficace.

Coût
Faible.

2.3.2.2 Filtrage protocolaire

Le filtrage protocolaire consiste à interdire l'usage du ou des protocoles réseau utilisés par le service de messagerie instantanée.

La facilité de contournement est minime.

Le risque de sous-blocage est minime.

Le risque de sur-blocage est important : la plupart des systèmes de messagerie instantanée utilisent le protocole HTTP, hors c'est le protocole sur lequel repose la majeure partie du Web. Interdire le protocole HTTP revient à bloquer une majeure partie du Web.

Conclusion : le blocage d'un service de messagerie instantanée par adresses IP n'est que faiblement efficace.

Coût
Moyen (certains pare-feu offrent déjà cette fonctionnalité).

2.3.2.3 Filtrage dynamique par zone géographique

Le filtrage par contenu par zone géographique utilise des technologies DPI dans des endroits ciblés du réseau. Il permettrait de filtrer uniquement certains messages en fonction de leur provenance géographique et de leur contenu.

La facilité de contournement est existante : l'utilisateur pourrait utiliser un proxy situé dans un pays non concerné par le filtrage du système de messagerie en question.

Le risque de sous-blocage est important : les modes d'accès aux messageries instantanées sont différents (Web, application mobile...) et les réseaux hétérogènes (nationaux, internationaux, roaming, VPN...) si bien qu'il serait difficile prendre en compte tout aspect pour réussir à filtrer ce que l'on souhaite filtrer.

Le risque de sur-blocage est important : pour les mêmes raisons que ci-dessus, il serait difficile de ne pas filtrer autre chose que ce que l'on souhaite filtrer.

Conclusion le blocage d'un service de messagerie instantanée par adresses IP n'est que faiblement efficace.

Coût
Coût d'un DPI généralisé, soit extrêmement couteux (voir DPI).

2.3.3 Mail

2.3.3.1 Filtrage par IP

Le filtrage par IP consiste à bloquer les mails provenant des serveurs identifiés par les IP incriminées.

Efficacité
La facilité de contournement est faible : le changement d'IP pour un serveur d'envoi de mail n'est pas aisé. Il ne s'agit pas comme dans le cas d'un site Web de déplacer un contenu sur un autre serveur.

Le risque de sous-blocage est existant : si tous les mails provenant des serveurs identifiés par les IP sont bien bloqués, il est possible que d'autres serveurs envoient les mails que l'on désirait bloquer.

Le risque de sur-blocage est très important : tous les mails provenant des serveurs identifiés par les IP seront bloqués y compris ceux en dehors du périmètre de filtrage.

Conclusion : le blocage de mails par les adresses IP des serveurs d'expédition n'est que faiblement efficace.

Coût
Faible.

2.3.3.2 Filtrage dynamique (par contenu)

Le filtrage dynamique consiste à inspecter le contenu des mails et de filtrer selon ce que le mail contient (ex : bloquer toutes les mails qui contiennent le mot « lapin »). Suivant l'endroit où il est réalisé, ce type de filtrage peut reposer sur l'utilisation de DPI.

Efficacité
La facilité de contournement est faible : le moyen de contournement le plus simple consisterait à chiffrer le contenu des mails, mais cela supposerait que seuls les destinataires aient les moyens de déchiffrer ce contenu (utilisation de crypto-asymétrique, échange sur canal privé, etc..). Ce moyen n'est pas envisageable à grande échelle.

Le risque de sous-blocage est existant : un tel blocage nécessite la constitution et la mise à jour d'une base de contenu (ou d'identifiant de contenus) et il sera impossible de référencer tout ce que l'on souhaite filtrer.

Le risque de sur-blocage varie de minime à important : bloquer un mot ou un terme ou une expression bloquera l'ensemble des mails comportant ce mot, ce terme ou cette expression, même s'ils ne sont pas directement visés par le filtrage, le sur-blocage sera alors important. Si le filtrage se fait sur un identifiant unique du contenu (une signature ou un hash de contenu du mail par exemple) le sur-blocage sera minime.

Conclusion : le blocage dynamique de mails est moyennement efficace.

Coût
Suivant la technique utilisée, le coût varie de moyen à élevé.

2.3.4 Usenet

2-3-4-1 Filtrage par groupe

Les informations et contenus sur Usenet sont organisés en groupes. On peut imaginer Usenet comme un arbre dont chaque feuille est un groupe. A l'origine, l'intégralité de l'arborescence est répliquée sur chaque serveur Usenet.

Chaque groupe a un nom qui reflète le type supposé des contenus et informations présents dans le groupe. Il est donc envisageable sur un serveur Usenet d'interdire l'accès à (ou de ne pas y copier) un groupe que l'on voudrait filtrer.

La facilité de contournement est existante : chaque groupe étant en théorie répliqué sur la totalité des serveurs Usenet existants, interdire ou supprimer un groupe revient à l'interdire et le supprimer sur tous les serveurs. Ainsi si un serveur ne joue pas le jeu, le groupe restera accessible via ce serveur et n'importe quel internaute pourra s'y connecter (via un canal chiffré pour éviter toute tentative de filtrage complémentaire par DPI).

Le risque de sous-blocage est existant: les utilisateurs peuvent dissimuler du contenu que l'on souhaiterait interdire dans des groupes ayant des noms ne laissant pas à penser qu'il faille les filtrer.

Le risque de sur-blocage est existant (à moins de considérer la totalité des contenus du groupe comme à filtrer) : ce filtrage risque d'empêcher l'accès à des contenus n'étant pas dans le périmètre de filtrage.

Conclusion le blocage d'une partie de Usenet par le blocage d'un groupe est moyennement efficace.

Coût
Faible.

2-3-4-2 Filtrage par hiérarchie

Comme expliqué précédemment on peut imaginer Usenet comme un arbre dont chaque feuille est un groupe.

S'il est possible sur un serveur Usenet d'interdire l'accès à une feuille de cet arbre (un groupe), il est aussi possible d'interdire l'accès ou de supprimer toute une branche, que l'on appelle une hiérarchie, de cet arbre et ainsi tout un ensemble de groupes.

La facilité de contournement est existante : c'est la même facilité de contournement que pour le filtrage par groupe.

Le risque de sous-blocage est minime : en empêchant l'accès à toute une hiérarchie, on a de fortes chances d'empêcher l'accès aux contenus ciblés.

Le risque de sur-blocage est important (à moins de considérer la totalité des contenus d'une hiérarchie comme étant à filtrer) : ce filtrage empêchera l'accès à des contenus n'étant pas dans le périmètre de filtrage.

Conclusion le blocage d'une partie de Usenet par le blocage d'un groupe est moyennement efficace.

Coût

Faible.

2.3.5 Pair à pair (et autres protocoles applicatifs d'echange de contenus)

2.3.5.1 Filtrage par protocoles

Il est généralement possible de reconnaître qu'un flux réseau est engendré par un logiciel pair-à-pair. Le filtrage par protocoles revient à identifier les flux réseau engendrés par des logiciels pair-à-pair et à les interdire. Cette technique peut nécessiter l'utilisation de DPI.

La facilité de contournement est faible : de nos jours, l'identification de flux est une technique maîtrisée. Même si de nos jours certains protocoles P2P essayent de brouiller et de dissimuler leurs flux, il existe des techniques heuristiques permettant la reconnaissance en dépit du brouillage.

Le risque de sous-blocage est minime : filtrer un protocole équivaut à prévenir la circulation de tous les contenus (et par conséquent ceux que l'on souhaite filtrer).

Le risque de sur-blocage est important (à moins de considérer la totalité des contenus circulant sur le réseau P2P comme étant à filtrer) : ce filtrage empêchera la circulation de contenus n'étant pas dans le périmètre de filtrage.

Conclusion le filtrage d'un réseau P2P par le blocage du protocole qu'il utilise est faiblement efficace.

Coût
Elevé (nécessite du DPI)

2.3.5.2 Filtrage dynamique (par contenus)

Le filtrage dynamique consiste à inspecter le contenu circulant sur le réseau P2P et de filtrer ceux-ci (ex : bloquer toutes les vidéos sur les « lapins »). Ce type de filtrage repose sur l'utilisation de DPI.

Efficacité
La facilité de contournement est existante : le filtrage dynamique ne peut notamment s'appliquer aux sites auxquels on accède en HTTPS (à moins de casser la sécurité du protocole de chiffrement SSL).

Le risque de sous-blocage est important : un tel blocage nécessite la constitution et la mise à jour d'une base de contenu (ou d'identifiant de contenus) et il sera impossible de référencer tout ce que l'on souhaite filtrer.

Le risque de sur-blocage est minime : bloquer un contenu clairement identifié n'entrainera pas le blocage des autres contenus.

Conclusion : le filtrage dynamique d'un réseau P2P est moyennement efficace.

Coût
Coût d'un DPI ciblé sur certains types de flux, très couteux (voir DPI).

2.4 Le filtrage par moyens (comment peut/veut-on filtrer ?)

2.4.1 Introduction

Lorsqu'une information circule, il y a trois acteurs qui interviennent :

- La source de l'information.
- Le milieu par lequel l'information se propage.
- Le destinataire de l'information.

Dans le cas qui nous concerne, la source est généralement le diffuseur d'un contenu sur Internet, le milieu se trouve être le réseau (Internet) et la destination l'utilisateur final (l'internaute).

Afin d'exercer un filtrage sur Internet, le filtrage consistant à bloquer la circulation de certains contenus (site Web, fichier, etc.), il est possible d'agir sur chacun des trois acteurs : source, milieu, destinataire et nous allons voir dans la suite de cette partie comment cela est possible et avec quelles conséquences.

2.4.2 Agir sur la source

On ne parle généralement pas de filtrage ou de blocage dans ce cas. Il est question de retirer le contenu ou le service purement et simplement.

2.4.3 Agir sur le milieu

Il s'agit de tenter de résoudre un problème en s'attaquant au milieu par lequel le problème est propagé. On agit uniquement sur celui qui permet, par la mise à disposition de ses « tuyaux » de « solliciter » ou de « constater » l'infraction en faisant circuler ce que l'on ne veut pas voir utilisé. On ne réprime alors pas l'usage (et on ne risque pas d'entrer dans une pédagogie de responsabilisation) on tente une épuration du milieu.

Pour agir sur le milieu, il y a deux possibilités, que nous allons décrire ici.

2.4.3.1 Agir au niveau du cœur de réseau

Le problème c'est que l'Internet n'a pas vraiment de cœur, de début ou de fin. On ne peut pas créer, à l'instar de l'eau, des stations d'épuration qui se retrouveraient forcément sur le passage de la collecte des eaux usées ou de la distribution des eaux propres. L'Internet est un réseau a-centré.
Par conséquent, il n'y a que deux manières de tenter cette épuration :

• On dispose des mécanismes d'épuration partout ou du moins à chaque croisement possible. Ceci n'est pas imaginable de par le nombre de croisements.

• On réduit le nombre de croisements en recentrant Internet. A l'extrême, on crée de gros points de concentration qui nous permettront d'épurer plus facilement. Ceci se fait au détriment de la décentralisation, entrainant donc une baisse des performances et de la résilience, ou bien à un coût (humain et technique) très élevé de par le nombre d'équipements et de configurations à mettre en production et à maintenir. Cette recentralisation entrainerait Internet loin de ses principes initiaux.

2.4.3.2 Agir au niveau du Nœud de Raccordement Abonné (DSLAM ou autre)

Le Nœud de Raccordement de l'Abonné (ou NRO dans le cas de la fibre) est l'endroit dans lequel aboutissent les lignes de communication des abonnés (fibres, cuivre, hertzien, ...).

C'est généralement le premier endroit ou le trafic est susceptible de prendre ou d'arriver depuis plusieurs voies d'accès. Dans le cas très majoritaire de la France, ces NRA sont équipés de DSLAM (expliquer l'acronyme) permettant de collecter le trafic des lignes ADSL".

Tous les DSLAM en service ne sont cependant pas forcément capables de traiter les paquets IP. Beaucoup ne sont en effet dédiés qu'au simple transport depuis le DSLAM vers un autre équipement et n'embarquent aucune capacité d'analyse, ne serait-ce que de la source ou de la destination des paquets. De manière générale, leur mission ne concerne que le transport de paquets et leurs capacités de traitement sont orientées vers cette mission.

L'ajout d'équipements permettant ce genre de traitement serait envisageable, mais il faut garder à l'esprit qu'il existe environ 13 000 centraux téléphoniques. Cela revient à l'ampleur du travail qui a dû être effectué pour déployer l'ADSL ces dix dernières années

2.4.3.3 Agir au niveau des serveurs de noms (DNS) de l'opérateur

C'est un cas particulier du filtrage qui intervient sur un service et non sur l'infrastructure elle-même. Mais en l'espèce, sans DNS, la plupart des services de l'Internet, utilisés au quotidien ne fonctionneraient plus. Cela n'empêche pas de consulter le contenu, mais limite son accessibilité.

L'analogie la plus proche est celle du réseau routier sur lequel on retire les panneaux indicateurs : vous pouvez toujours vous rendre à votre destination si vous connaissez déjà le chemin. Il est aussi possible d'orienter différemment les utilisateurs en indiquant d'autres directions.

La plupart des utilisateurs utilisent le mécanisme de résolution d'adresses de leur opérateur. Ils seraient alors facilement « trompables ». Mais il serait alors possible d'utiliser d'autres DNS, qui eux, diraient la vérité (et ne filtreraient plus en remplaçant une destination par une autre).

2.4.4 Agir à la destination

2-4-4-1 Sous la maitrise de l'utilisateur

C'est un autre moyen de filtrer et la popularité des logiciels de contrôle parental est là pour nous confirmer la faisabilité technique d'une telle solution. Ces logiciels sont installés à la discrétion des utilisateurs et même s'il existe des paramétrages par défaut c'est toujours l'utilisateur qui peut choisir les paramètres de filtrage appliqués par ces logiciels.

Avec la multiplication des écrans « connectés » domestiques (Télévision, ordinateurs, consoles de jeu, Smartphones … voire demain quasiment tout objet ayant un intérêt à remonter de l'information ou que l'on doit « actionner »), il devient totalement illusoire de penser le « contrôle parental » comme on le pensait avant. Il devient impossible de protéger objet par objet … et d'ailleurs de quoi, de qui et pour qui ? (que se passe-t-il si nous sommes plusieurs à utiliser le même objet ?)

Tout comme, il est absurde de protéger du cambriolage son logement en n'empêchant pas le malfaiteur d'entrer, mais en disposant des cadenas sur chacun des objets contenus présents dans le logement (la première protection utile étant de penser à fermer la porte du logement), il est inconcevable d'imaginer un « contrôle parental » installé sur chacun des équipements connectés.

Si on poursuit le raisonnement avec l'analogie de la porte, se posent d'autres questions :

- Qui (et quoi) sera la porte blindée de notre « chez nous » numérique ?

- Qui la contrôlera ?
- Au-delà d'une grande responsabilité, ne s'agit-il pas d'un grand pouvoir qui se dessine ?

C'est en tous les cas le point de complexité majeure du sujet.

2-4-4-2 Sous la maitrise d'un tiers

Une sécurisation de l'accès à Internet réalisée « à l'insu du plein gré » de l'utilisateur s'apparente à un appartement qui nous est livré avec une porte blindée, mais dont d'autres ont la clef.

Un système « d'épuration », sans la participation de l'utilisateur est une erreur si on considère que l'on ne pourra jamais tout filtrer et que par conséquent le premier rempart doit être l'utilisateur lui-même et non reposer sur l'autre, qu'il s'appelle un parent ou un firewall.

2.5 Le filtrage par type (quelles techniques employer ?)

2.5.1 Blocage sur les noms de domaine (DNS, DPI)

2.5.1.1 Objectif et terrain d'action

Il s'agit d'empêcher, à un niveau ou à un autre, la conversion du nom d'un site en adresse IP pour empêcher les machines des utilisateurs de connaitre cette IP et donc d'accéder au site.

Cette technique se met généralement en place au niveau du DNS des fournisseurs d'accès ; sachant que s'il existe une poignée de très gros représentants, il faut également compter sur plusieurs centaines de petites structures n'ayant parfois qu'une dizaine d'utilisateurs.

2.5.1.2 Efficacité

Le blocage complet d'un nom de domaine entraine un sur-blocage potentiel, par exemple dans le cas où plusieurs sous-domaines ou sous-sites sont hébergés à la même adresse (exemple: site.com/blog1 et site.com/blog2). On peut reprendre l'exemple présenté précédemment : le seul blocage de 'wordpress.com' entrainerait le blocage des centaines de milliers de blogs.

Il est possible pour n'importe quel utilisateur de se servir d'autres serveurs DNS que ceux de son FAI, souvent même en dehors du pays (Google en propose, par exemple) rendant le filtrage sur les DNS du FAI inopérant.

Le filtrage DNS peut alors être couplé avec un filtrage protocolaire n'autorisant les requêtes DNS que sur les serveurs du fournisseur d'accès effectuant le filtrage. Ce filtrage peut, en l'état, être considéré comme une atteinte à la neutralité et peut être contourné au moyen de services DNS utilisant d'autres ports ou protocoles, ou encore en passant par un VPN.

Cette technique peut aussi faire l'objet de DPI, avec tous les travers que cela comporte, directement en cœur de réseau, permettant cette fois d'intercepter la totalité des requêtes DNS effectuées.

De manière générale, le filtrage DNS peut aisément être contourné au moyen d'un VPN permettant de simuler une connexion depuis un autre opérateur, par exemple à l'étranger.

2.5.2 Blocage sur les adresses IP (DNS, BGP)

2.5.2.1 Objectif et terrain d'action

Il s'agit d'empêcher le trafic en provenance ou à destination d'une adresse IP prédéfinie. Concrètement, sur le réseau d'un opérateur, cette mesure peut être implémentée au niveau du DNS, en l'empêchant de répondre l'adresse IP cible du blocage à ses clients. On retrouve les mêmes problématiques de contournement que sur le blocage des noms.

Il peut également être effectué au moyen d'injection BGP, technique consistant à donner, pour l'adresse IP cible, une fausse information à l'ensemble des routeurs du réseau afin que les demandes de connexions ne puissent atteindre le serveur réel. L'emploi de technologies type VPN permet de contourner très aisément ce type de filtrage, en employant d'autres réseaux non soumis aux mesures de blocage.

Il peut enfin être mis en place en inspectant les paquets et en interrompant la connexion une fois qu'elle a été établie. L'inspection de paquets permettant plus de flexibilité, elle ne s'arrête généralement pas à un simple blocage des adresses IP, puisqu'elle permet beaucoup plus.

2.5.2.2 Efficacité

Ce type de filtrage permet donc effectivement de bloquer l'accès à la cible définie depuis le réseau où est effectué le filtrage, mais le sur-blocage est énorme et non estimable au moment où l'ordre de blocage est donné et il est, en fonction de la méthode employée, aisé à contourner ou difficile à mettre en place.

Par ailleurs, il n'est pas dit que la stabilité d'Internet résiste à un grand nombre d'injections d'annonce BGP et que finalement, pour protéger leur réseau, les opérateurs ne se protègent pas de ce genre d'annonces (*route flap dampening*).

2.5.3 Filtrage sur les URL (DPI)

2.5.3.1 Objectif et terrain d'action

Il s'agit de contrôler précisément les URL auxquelles on laisse accès. C'est un filtrage applicable uniquement au Web non chiffré.

Il s'agit, via l'inspection de paquets, d'interrompre une connexion préétablie lorsqu'une URL "interdite" est détectée.

2.5.3.2 Efficacité

Ce type de filtrage présente tous les problèmes du DPI et peut être facilement contourné par l'emploi de connexions chiffrées (HTTPS), par changement d'URL du contenu ou par l'utilisation de VPN.

L'utilisation de ces technologies sur HTTPS est théoriquement possible en ayant des certificats valides pour des sites que l'on ne possède pas. Il suffit pour cela de posséder une autorité de certification reconnue par les navigateurs. C'est le cas de nombreux états ou d'entreprises privées pouvant, le cas échéant, être soumises par un état.

2.5.4 Filtrage sur les contenus (DPI)

2.5.4.1 Objectif et terrain d'action

Suivant le même principe que le filtrage sur des URL, le filtrage sur le contenu inspecte les

paquets transmis à la recherche d'un contenu spécifique : texte, image, son, vidéo, pour en empêcher la transmission.

2.5.4.2 Efficacité

Il ne présente pas le problème du filtrage par URL qui peut être contourné en publiant le même contenu à une autre adresse, mais il oblige alors à un contrôle de l'intégralité du trafic.

La nécessité d'inspecter la totalité du trafic implique de mettre en place des équipements d'inspection de paquets dimensionnés pour supporter l'intégralité de la charge induite sur le réseau par les abonnés. Cela nécessite aussi de centraliser le trafic ou bien de démultiplier les équipements pour en placer plus près des abonnés.

2.5.5 Blocage sur les ports

2.5.5.1 Objectif et terrain d'action

Pour simplifier, considérons que chaque type d'application utilisant Internet le fait par le biais d'un numéro prédéfini pour chacune. Par exemple, le Web utilise massivement le port 80 ou, pour les sites sécurisés, le 443. L'envoi d'email se fait quant à lui par le port 25.

2.5.5.2 Efficacité

Un filtrage par numéro de ports est simple à mettre en place. Presque tous les équipements réseau savent le faire. C'est le principe de base du fonctionnement des firewalls. On pourrait, par exemple, décider que les serveurs Usenet utilisant le protocole NNTP sur le port 119 sont une menace pour le respect du droit d'auteur. Il conviendrait donc de fermer le port 119 pour régler le problème.

Si on met de côté le fait que NNTP n'est qu'un protocole et ne détermine donc pas la licéité des contenus qui y transitent, les personnes souhaitant l'utiliser auront tôt fait d'utiliser un autre port que le 119.

Certaines applications sont mêmes auto-adaptatives, comme beaucoup de logiciels peer-to-peer qui vont tester la connexion de l'internaute à la recherche de ports bloqués et utiliser les premiers disponibles qu'ils trouveront.

Le blocage de ports est donc l'un des plus simples, mais aussi l'un des moins efficaces.

2.5.6 Filtrage hybride

2.5.6.1 Objectif et terrain d'action

Il s'agit d'un mélange du filtrage utilisant BGP et inspection de paquets.

Concrètement, pour éviter le sur-blocage entraîné par l'interdiction totale d'accéder à une IP et pour éviter le coût et les problèmes de redondance engendrés par le DPI, le filtrage hybride propose de ne faire passer au travers du filtre DPI que le trafic à destination des adresses IP distribuant le contenu à bloquer.

2.5.6.2 Efficacité

L'effet sur le reste du trafic est supposé être nul, l'impact sur la globalité du réseau également. C'est bien entendu en supposant que les contenus à filtrer ne font pas partie de grosses plateformes de distribution de contenus. Car si ce genre de plateforme se retrouve

soudain à devoir filtrer l'ensemble du trafic venant d'un site comme MegaUpload, elle aurait tôt fait d'être saturée et d'agir, finalement, comme un filtre simple interdisant l'ensemble du trafic.

Tout comme quand il est utilisé dans le cas du filtrage IP, l'utilisation détournée de BGP n'est pas un acte anodin. De nombreux administrateurs chevronnés commettent souvent des erreurs de configuration ayant plus ou moins d'impact sur le réseau tout en étant des actions de pure routine, on envisage aisément que des cas comme YouTube/Pakistan Telecom pourraient se produire régulièrement si l'utilisation du filtrage via BGP était institutionnalisée.

2.5.7 Filtrage sur les protocoles (DPI – traffic shaping)

2.5.4.3 Objectif et terrain d'action

Le filtrage sur les protocoles inspecte les paquets transmis à la recherche d'une signature de protocole. Une signature de protocole permet d'identifier qu'une suite de paquets transmis entre une source et une destination correspond bien à un type de protocole réseau (et donc si on simplifie : à un usage). Une fois le protocole identifié comme étant à filtrer, la connexion est interrompue.

2.5.4.4 Efficacité

Ce filtrage oblige alors à un contrôle de l'intégralité du trafic.

La nécessité d'inspecter la totalité du trafic implique de mettre en place des équipements d'inspection de paquets dimensionnés ou de « *traffic shapping* » pour supporter l'intégralité de la charge induite sur le réseau par les abonnés. Cela nécessite aussi de centraliser le trafic ou bien de démultiplier les équipements pour en placer au plus près des abonnés.

3 | Les enjeux

3.1 Enjeux du filtrage

3.1.1 Le filtrage à l'intersection des techniques et des usages

Parler des problèmes engendrés par Internet est un abus de langage, car Internet n'est pas mauvais en soi, tout comme les technologies – le P2P (le pair à pair), le streaming – ne sont pas mauvaises par nature, ce qui n'est potentiellement pas le cas de l'usage que nous en faisons.

Ainsi la thématique du filtrage d'Internet ne peut être réduite à une problématique uniquement technique. Elle concerne aussi les usages des internautes. Les usages sont à comprendre comme une combinaison entre un possible technique et des utilisateurs qui se sont emparés des moyens pour en déterminer une utilisation. En surplus des réglementations, il existe également des habitudes et des usages particuliers, souvent regroupés sous l'appellation de Netiquette[10].

Internet est comme un objet qui peut se révéler indispensable à notre quotidien, mais en même temps imprévisible : mal utilisé, Internet, entraîne sur un terrain, non pas de non-droit, mais d'illégalité où généralement toutes les peines ont déjà été largement pensées et prévues.

[10] Voir pour exemple ce site HTTP://netiquette.fr/

Au-delà de la question des usages, se pose également la question de l'espace et du temps d'Internet. Internet impose un devoir de vigilance tant par l'aspect polymorphe de ses contenus que par leur vitesse de diffusion et les conditions de leur accessibilité. La même chose peut être mise en évidence pour les modalités d'accès à un contenu selon la localisation de l'internaute et du contenu. En effet, un usage peut être légal sur un territoire donné et être illégal sur un autre. Il peut également être illégal pendant un laps de temps et devenir légal par la suite. Les jeux d'argent en ligne semblent illustrer cette problématique. En effet, dans le cas de la France, sur le plan de la légalité les jeux d'argent en ligne étaient considérés comme illicites. Suite à des changements législatifs, la mise en place de l'ARJEL[11] et la coopération des sites concernés, les jeux d'argent en ligne sont progressivement entrés dans la légalité au regard de la loi française.

Internet ne peut pas être réduit au Web ou à un service précis, ce n'est pas non plus uniquement un réseau de diffusion. Il peut être utilisé pour cela, mais il n'a pas besoin de diffuser pour exister puisqu'il existait avant ces usages. Internet est un réseau d'échanges. L'expérience du passé et l'humilité de reconnaître que l'on se trompe souvent lorsque l'on tente de prévoir le futur de ces technologies, nous enseignent qu'il est préférable qu'il soit neutre et symétrique et si possible, au plus haut débit envisageable.

3.1.2 Filtrage et responsabilité

Le principe de liberté, souvent confondu avec celui de l'arbitraire dans le langage commun, inclut nécessairement celui de la responsabilité, définie comme le fait d'agir en toute connaissance de cause, en toute conscience. Selon l'entité qui opère le filtrage et selon les moyens utilisés pour le mettre en place, la notion de responsabilité se trouve modifiée. L'internaute serait placé dans une situation de responsabilité moindre si le filtrage d'Internet lui échappait ou s'il n'en avait pas l'initiative. Il serait alors dans une situation de dépendance avec un tiers pour la mise en place d'un système de filtrage. Examinons les enjeux de ces deux possibilités : filtrage par un tiers, filtrage à l'initiative de l'internaute.

3.2 Filtrage par un tiers

3.2.1 Identification des tiers

En effet, il y a différentes entités qui peuvent opérer un filtrage d'Internet : l'autorité judiciaire, l'opérateur de réseau, l'opérateur de services, l'hébergeur d'un site Web, l'employeur, l'administrateur réseau et l'internaute. Ainsi l'autorité judiciaire va par exemple opérer un filtrage afin de répondre à un besoin légitime, notamment celui du respect de la légalité.
Dans le cas de l'opérateur de réseau, l'opérateur de service, de l'hébergeur, de l'administrateur réseau et de l'employeur, il s'agit également de s'assurer du respect de la loi ainsi que de préserver les ressources du système afin de ne pas créer un déséquilibre dans une architecture réseau. En effet, les ressources du système sont partagées entre plusieurs machines : si l'une des machines consomme, de manière excessive, certaines ressources, les autres machines rencontreront des dysfonctionnements, créant une réaction en chaîne. Cet impératif de préservation pourrait justifier qu'un administrateur filtre une partie des éléments arrivant sur les machines composant son réseau afin de ne pas déséquilibrer l'ensemble de la structure.

3.2.2 Une moindre responsabilité et une moindre maîtrise pour l'internaute

Dans l'hypothèse où un filtrage est opéré par un tiers, l'internaute est davantage déresponsabilisé. Il pourrait donc penser que les contenus qu'il consulte sont conformes

[11] Autorité de Régulation des Jeux En Ligne

avec la légalité. Il pourrait donc s'exonérer, au sens juridique du terme, de sa responsabilité. Lorsque les techniques de filtrage ne sont pas mises en place par l'internaute lui-même, il peut y avoir une perte de maîtrise de la part de l'internaute, une perte d'initiative quant aux choix des contenus désirés. Cette perte de moyens pourrait aboutir à une fracture numérique, à savoir un Internet à plusieurs niveaux, constitué d'une hiérarchie fondée sur la possibilité donnée aux individus de contourner les processus de filtrage. Sur le plan technique, lorsque l'internaute est placé dans une situation de consultation de contenus préalablement filtrés par un système non transparent, il n'est pas créateur. Il ne va pas contribuer à l'expansion du réseau, laissant ainsi la place pour un cercle d'initiés.

3.2.3 Un respect de la légalité plus aisé

Toutefois, si le filtrage est à son initiative, le législateur pourra s'assurer plus aisément du respect de la loi et surtout évaluer pleinement la validité des intermédiaires afin de protéger non seulement l'utilisateur, mais aussi de garantir le respect des droits.

Notons néanmoins qu'un filtrage des échanges et des communications pourrait conduire à un excès de protection de la part des internautes, notamment par l'utilisation de technologies de chiffrement. Or, il pourrait y avoir une régulation de chiffrement et de cryptage, comme cela a été le cas en France jusqu'à la loi du 26 juillet 1996[12]. Ce type de régulation est à même de créer des inégalités et peut potentiellement fausser la concurrence. On obtient alors un Internet à deux vitesses avec d'un coté ceux autorisés à utiliser la cryptographie « forte » et ceux n'en ayant pas l'autorisation (et qui par conséquent se trouvent désavantagés par rapport aux premiers).

3.2.4 Un enjeu éthique important

Soulignons que les enjeux en termes d'éthiques et de vie privée sont majeurs dans le cas du filtrage opéré par un tiers sans l'intervention de l'internaute. En effet, dans l'hypothèse où l'internaute n'est pas informé des techniques de filtrage ni des filtres mis en avant, certaines informations le concernant pourraient être analysées sans que son consentement soit recherché.

3.3 Filtrage sous la maitrise de l'internaute

Dans ce cas, l'internaute décide ce qu'il désire consulter et met en œuvre lui-même une solution de filtrage. Cette décision n'appartient qu'à lui.

3.3.1 Une responsabilisation plus forte de l'Internaute

L'internaute va donc chercher à se protéger d'éventuelles agressions, à protéger ses ressources systèmes et à éviter certains contenus, pour des raisons qui lui sont propres, et dont il aura l'initiative. C'est le sens des solutions centrées sur l'utilisateur qui placent l'initiative de ce dernier au cœur du processus de filtrage. L'enjeu est d'initier une dynamique responsable quant à ce que l'on accepte ou non, dans le respect de la position de l'utilisateur. Or, en plaçant l'utilisateur dans une situation de responsabilité, cela permet de donner les moyens au plus grand nombre, de prendre part au monde numérique qui va baigner, faciliter et permettre la plupart des échanges. L'internaute va être créateur, initiateur, concepteur, hébergeur.

3.3.2 Un risque d'inégalité face au filtrage ?

L'inconvénient de ce type de filtrage réside dans le fait que cela suppose une maîtrise

[12] HTTP://legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000733177

technique que tous les internautes peuvent ne pas avoir. Cela implique donc un effort de pédagogie supplémentaire. Par ailleurs, placer entre les mains de l'internaute la question de la responsabilité du filtrage amène à une pleine conscience de cette responsabilité, à une nécessaire connaissance fine de la loi et des différentes réglementations que tout le monde n'a pas : on risquerait une mise en place différentielle du filtrage qui n'impacterait pas les utilisateurs d'une façon uniforme.

Par conséquent, il pourrait être nécessaire de définir des paramétrages de filtrage « par défaut » contournables et documentés dont le sens et la finalité seraient à même d'être appréhendés par tous quelque soit son niveau.

Par ailleurs lorsque l'internaute décide de lui-même de ce qu'il souhaite filtrer, selon des paramètres et des problématiques qui lui sont propres, il prend le risque de sur-bloquer inutilement. Par exemple, s'il a paramétré une liste de mots-clefs dans son moteur de recherche, certaines recherches qu'il mènera pourraient être biaisées du fait d'un filtrage volontairement ou involontairement trop poussé. Cependant, contrairement à un filtrage qui ne serait pas placé sous la maîtrise de l'utilisateur, un mauvais paramétrage n'entrainerait de conséquences que pour l'utilisateur

3.4 Filtrage et avenir d'Internet : les variables risques liées au filtrage

Outre l'initiateur du filtrage, il convient d'évoquer les impacts globaux du filtrage en termes de variables-risques. La procédure n'est pas anodine et a des impacts durables sur la mise en place d'une politique numérique pour le pays. Ces risques variables peuvent assez facilement être étudiés dans des mesures d'impacts qui devraient toujours éclairer des décisions touchant à Internet. Il n'est pas le lieu d'étayer tous les sujets indiqués, mais simplement de les mentionner
La mise en place d'infrastructures – logiciels, réseaux, machines – performantes est un enjeu clair de compétitivité. La question n'est pas seulement de mettre à disposition des personnes un réseau, ni de donner un moyen de distraction supplémentaire, il s'agit de donner les moyens, au plus grand nombre, de prendre part, au monde numérique qui va baigner, faciliter et permettre la plupart de nos échanges. Cette thématique touche directement l'un des principaux enjeux du principe d'accès libre au réseau.

Penser des mécanismes de filtrage va nécessairement entraîner un projet global de management des risques qui lui sont liés. En effet, le filtrage a des conséquences sur le réseau lui–même et sur la circulation des données :

- Qualité et fluidité de nos réseaux : un réseau non filtré aura une meilleure qualité en termes de débit puisqu'il n'y aura aucun point de passage obligatoire à franchir.

- Risque de perte de compétitivité : filtrée, l'infrastructure pourrait être moins performante Ne risquons-nous pas de décourager toute une population d'entrepreneurs, d'inventeurs et de faire fuir les meilleurs à l'échelle internationale ?

- Risque de dénaturation et d'oubli : pour des raisons économiques et pragmatiques, il pourrait être décidé de re-centraliser à l'extrême les points d'échanges et de routage d'Internet. Au lieu d'avoir une myriade de points d'entrée et de sortie, il y aurait un point d'entrée unique et un point de sortie unique. Cette hypothèse conduit nécessairement à une saturation du réseau. Ce type de réseau ne serait dès lors plus réellement Internet et on ne voit pas nettement ce qui le distinguerait des précédents réseaux de diffusion, comme par exemple le Minitel. Les risques potentiels de fragilisations volontaires ou involontaires des architectures et des systèmes seraient alors beaucoup plus élevés qu'aujourd'hui. La conséquence serait

alors une simplification des attaques et des paralysies complètes de ce type de systèmes. A l'heure actuelle, les attaques sont disséminées en différents points d'entrée et de sortie. Canaliser l'ensemble des données en une seule route comportant un point d'entrée et de sortie pourrait amener à la saturation.

- La pérennité des contenus filtrés : enfin, il convient de souligner que le filtrage ne résout pas le problème du contenu. En effet, filtrer un contenu n'implique pas sa disparition du réseau. Le contenu litigieux existera toujours. Il sera simplement invisible pour certaines personnes. Par ailleurs, le fait de filtrer un contenu plutôt que de le supprimer n'exclut pas l'apparition de sites miroirs, qui vont relayer le contenu. L'exemple illustrant le mieux cette hypothèse est Wikileaks[13]. Lorsque le site a été neutralisé, les internautes ont relayé les différents câbles[14], montrant ainsi que ce n'était pas parce qu'une information était bloquée ou filtrée qu'elle n'existait plus. Cet exemple pourrait être la démonstration que le filtrage ne résout pas le problème initial : l'existence d'un contenu litigieux.

4 | Conclusion

La problématique de filtrage d'un contenu litigieux concerne également la notion de persistance des contenus sur Internet. Ceux-ci en effet, même filtrés persistent. Ils constituent une sorte de mémoire de l'Internet qu'il est finalement difficile d'effacer, mais à laquelle il est tout aussi difficile d'accéder. Dès lors, la question du filtrage intéresse aussi l'indexation des données filtrées et leur visibilité dans les moteurs de recherche. Autrement dit, le filtrage ne peut faire l'économie ultérieure d'une réflexion complexe sur ces derniers.

4.1 La question de la sécurité

Avec la problématique de la persistance des données, se pose la question de la sécurité sur Internet, question d'autant plus essentielle lorsqu'elle est mise en rapport avec les publics dits fragiles, notamment les enfants.
De manière naturelle, penser à recourir au filtrage afin de protéger ce public semble légitime.

Néanmoins il a été vu que les contenus filtrés continuent d'exister sur Internet. Ils ne sont pas effacés et restent accessibles aux personnes qui ont des compétences techniques particulières. Cela pose ainsi la question de l'égalité d'accès aux contenus sur Internet entre une sphère composée de personnes qui ont des connaissances techniques permettant de contourner les mesures de filtrages et d'accéder même illégalement à des contenus (films, musique…) et une autre sphère de personnes qui ne les posséderaient pas (et qui continueront donc à payer pour accéder à des contenus que d'autres obtiennent gratuitement). Risque-t-on de se diriger vers un filtrage à deux vitesses, dépendant de la maturité technologique des individus ? On a là une question importante qui touche la démocratisation de l'accès à Internet et l'antinomie possible entre le filtrage et la neutralité du net.

4.2 La question de l'innovation

[13] Voir pour explications et genèse de Wikileaks : **Pirates informatiques ou altermondialistes numériques, Anonymous : peuvent-ils changer le monde ?** Frédéric Bardeau et Nicolas Danet. Editions FYP 2011 – page 38 et de 89 à 112
[14] Voir pour illustration des miroirs de Wikileaks : HTTP://www.multigesture.net/2010/12/09/visualizing-wikileaks-mirrors/#demo

Indéniablement, le filtrage a pour conséquence de rendre Internet davantage centralisé car davantage dépendant de points de passages. D'où une moindre rapidité d'un Internet centralisé par rapport à d'autres Internets (afférents à d'autres pays) ralentissant alors certains protocoles ou rendant plus lents certains services.

Aujourd'hui, les usages ont évolué du fait des possibilités techniques. La fibre optique permet une avancée sans commune mesure avec ce que permettait le cuivre. On ne parle plus de temps de commutation en centaines de millisecondes, mais en nanosecondes. De quoi demain sera fait ? Certains y travaillent déjà[15]. N'allons-nous pas être *de facto* écartés de ces sujets, n'ayant plus les moyens techniques de nous en rapprocher ? Ceci signifie que toute politique de promotion du filtrage va de pair avec un effort technologique important : le temps de l'innovation sur Internet est d'une rapidité importante. Une politique de filtrage doit donc être en phase avec ces évolutions sous peine de risquer une obsolescence forte et donc une in-opérabilité rapide.

Ce risque d'obsolescence est à mettre en rapport avec la proximité des stratégies de filtrage et de certaines infrastructures (réseaux, machines, etc.) : plus les stratégies de filtrage se rapprochent des infrastructures, plus elles appellent la prise en considération de multiples facteurs apparemment étrangers aux finalités de la stratégie initiale envisagée. C'est là qu'intervient la notion de « complexité » du filtrage : celui-ci n'est pas une grille de contrôle applicable délibérément, mais un système complexe au sens d'une mise en mouvement d'une série d'entités en interactions. Ces entités, malgré la précision de plus en plus poussée des ingénieries algorithmiques, ne sont pas à l'abri d'une certaine imprévisibilité (on l'a vu avec le phénomène du sur blocage).

Dès lors, toute stratégie de filtrage nécessite une délimitation des objectifs et une anticipation claire et rigoureuse des conséquences de sa mise en œuvre : c'était là un des objets de ce document, que de mettre en lumière les conséquences d'un choix d'un type de filtrage.

La prudence est donc de mise et tout déploiement d'une stratégie de filtrage est à envisager à bon escient. En effet, l'accumulation de stratégies désordonnées de filtrage pourrait provoquer des dysfonctionnements importants dans les réseaux, susceptibles de pénaliser la compétitivité numérique du pays, c'est-à-dire sa capacité à répondre aux exigences des entreprises, aux aspirations individuelles en matière d'accès au numérique et à la vitalité d'un modèle économique du numérique en pleine maturation. En la matière, c'est une dialectique féconde entre liberté et responsabilité qui est à rechercher.

Avançons alors l'idée que l'utilisateur puisse devenir non pas tant son propre FAI, mais son FAL : Fournisseur d'Accès Local. On considère ici qu'il convient de privilégier une autonomisation locale de l'accès, en même temps qu'un déplacement stratégique des entités centralisées (les grands FAI) vers la périphérie (le réseau domestique). Cette dynamique est intéressante car porteuse d'un double potentiel : la conviction d'une responsabilisation efficace de l'utilisateur en même temps qu'une potentialité d'autonomisation de l'individu par la technologie. »

5 | Glossaire

ADSL Désigne l'acronyme Asymmetric Digital Subscriber Line ou ligne

[15] Exploring Network of the future : HTTP://www.geni.net

	numérique asymétrique. Solution consistant à réutiliser le réseau téléphonique actuel pour transmettre des données à très haut débit.
ARJEL	Acronyme désignant Autorité de Régulation des Jeux En Ligne. Elle est chargée de mettre en place des moyens de régulation, d'information et de contrôle pour protéger les joueurs, prévenir l'addiction au jeu et lutter contre la fraude. Site officiel : HTTP://www.arjel.fr/
BGP	Acronyme désignant Border Gateway Protocol. C'est un protocole d'échange d'informations d'accessibilité de réseaux, conçu pour prendre en charge un volume important de données. Ce protocole a permis la décentralisation du routage d'Internet.
Disque dur	Disque principal d'un ordinateur sur lequel est enregistré le système d'exploitation ainsi que les programmes et les données.
Domaine Internet	Ensemble d'ordinateurs reliés à Internet
DNS	Désigne l'acronyme Domaine Name System (système de noms de domaine). Il s'agit d'un système décentralisé permettant de la correspondance entre une adresse IP et un nom de domaine, dont l'exécution est distribuée sur plusieurs machines. Il est constitué de serveurs.
DPI	Sigle désignant Deep Packet Inspection – Inspection profonde de paquets. Technologie permettant d'analyser le contenu d'un paquet.
FAI	Acronyme désignant Fournisseur d'Accès Internet. Sont regroupés dans cette catégorie les opérateurs de réseaux et les opérateurs de services.
Fibre optique	Mode d'accès à Internet, formé par des fils en silice ou en plastique très fin conduisant une lumière modulée, permettant de transmettre des données à un très haut débit.
FTP	Sigle pour File Transfer Protocol : protocole de transfert de fichiers. Protocole de communication destiné à l'échange de fichiers sur un réseau. Permet de les copier afin d'alimenter un site Web.
Hébergeur	Entité ayant pour vocation de mettre à disposition des internautes des contenus et gérés par des tiers.
ICANN	Acronyme pour Internet Corporation for Assigned Names and Numbers. Société pour l'attribution des noms de domaine et des numéros sur Internet. Depuis 1998, elle est chargée de superviser les règles d'attribution des adresses et des domaines sur Internet. Site officiel : HTTP://www.icann.org/
Internet	Réseau mondial d'ordinateurs relies entre eux par le tissage des voies téléphoniques.
IP	Sigle désignant Internet Protocol. Procédure de communication utilisée sur Internet. Permet de créer les adresses des ordinateurs connectés.
IPv4	Adresses codées sur quatre octets, au début du réseau Internet. Applique la notation décimale.
IPv6	Adresses codées sur six octets, afin de faire face à la pénurie d'adresses. Applique la notion hexadécimale.
Messagerie instantanée	Messagerie électronique servant à communiquer d'ordinateur à ordinateur en instantané et impliquant la présence physique des personnes au moment des échanges. Elle peut être écrite, orale ou par vidéo.

Monitoring	Désigne l'ensemble d'un système de surveillance d'un parc informatique, afin d'en vérifier le bon fonctionnement.
Nom de domaine	Identifiant de domaine Internet.
NRA	Acronyme désignant Nœud de Raccordement d'abonnés au haut débit. C'est un sous-répartiteur téléphonique installé par France Telecom pour couvrir en ADSL les zones trop éloignées du central téléphonique le plus proche et ainsi supprimer les zones d'ombres.
NRO	Acronyme désignant Nœud de Raccordement Optique permettant la convergence des fibres optiques d'abonnés dans une même zone géographique : ville ou quartier.
Paquet	Petits blocs d'information transmis sur Internet, disposant chacun de leurs références propres ainsi que d'une adresse de destination. Ils sont émis indépendamment des uns des autres, le message complet étant reconstitué à la réception.
P2P	Sigle désignant le Peer to Peer, en français, le pair à pair. Il s'agit d'un modèle de réseau informatique sur Internet. Par abus de langage le pair à pair est aussi parfois employé pour des applications réseaux (réseaux organisés en modèle pair à pair) et utilisées pour l'échange des fichiers.
PGP	Signe pour Pretty Good Privacy (confidentialité plutôt bonne). Logiciel gratuit de chiffrement et de signature de données développée par Philip Zimmermann.
Phishing	Terme anglais désignant l'hameçonnage. Technique utilisée pour récupérer frauduleusement des données personnelles comme l'identité, les coordonnées personnelles, les coordonnées bancaires. S'effectue généralement par email, falsifiant un document officiel ou un site Internet connu.
Port	Désigne un point d'accès à un ordinateur.
Protocole	Synonyme de procédure. Séquence à appliquer pour aboutir au résultat désigné.
Proxy	Serveur informatique relayant des requêtes entre une machine et un serveur, utilisé pour des raisons de sécurité. Utilisé pour se rendre anonyme sur Internet.
Opérateur de services	Personne morale proposant des services en rapport avec des réseaux de télécommunication mais n'exploitant pas directement le réseau. Par exemple, en France, les opérateurs de services sont Darty, Bouygues Telecom.
Opérateur de réseaux	Personne morale exploitant un réseau de télécommunications ouvert au public ou fournissant au public un service de télécommunications. En France les opérateurs de réseaux sont Orange, SFR, Free et Numéricable.
Réseau	Un réseau est un regroupement de machines interconnectées pouvant communiquer les unes avec les autres.
Roaming	En français : itinérance. Décrit la faculté de pouvoir appeler et être appeler quelle que soit la position géographique, en utilisant un autre réseau que son réseau d'origine.
Routage	Mécanisme par lequel des chemins sont sélectionnés dans un réseau pour acheminer les données.
RSA	Algorithme de cryptographie asymétrique permettant de chiffrer les communications par deux clefs : l'une publique, l'autre privée.
Serveur	Ordinateur ou groupe d'ordinateur destiné à fournir des services à des utilisateurs connectés.
Site Web	Ensemble de pages utilisant liées entre-elles par des références

	hypertextes et mises en ligne sur internet.
Spam	Désigne des messages non sollicités qui encombrent les boîtes aux lettres électroniques.
Streaming	Diffusion de flux. Lecture en continu de fichiers audio et vidéo, permettant ainsi d'écouter ou de visionner immédiatement ce qui est reçu, sans attendre que l'ensemble du contenu soit téléchargé.
Système d'exploitation	Ensemble des programmes chargés de gouverner un ordinateur et de servir d'interface avec l'utilisateur.
TCP/IP	Signe pour Transmission Control Protocol/Internet Protocol. C'est une collection de protocole régissant Internet.
URL	Sigle désignant Uniform Resource Locator. Adresse Internet exclusive permettant d'atteindre un site précis depuis n'importe quel endroit dans le monde.
Virus	Programme parasite autoreproducteur et autopropagateur qui contamine un ordinateur et peut provoquer des dégâts tels que la destruction de données ou le blocage de la machine. Les virus peuvent être transmis par email, par messagerie instantanée.
VPN	Désigne l'acronyme Virtual Private Network ou réseau virtuel. C'est un mécanisme qui permet de réaliser un échange d'informations avec chiffrement de la communication.
Web	Abbreviation de World Wide Web. Service offert sur Internet et permettant de naviguer sur des sites distribués dans le monde.

6 | Bibliographie

6.1 Pour les définitions de blocage et de filtrage :

Le Point : HTTP://www.lepoint.fr/monde/le-bangladesh-bloque-facebook-a-cause-des-caricatures-du-prophete-mahomet-30-05-2010-460859_24.php
Infoworld : HTTP://www.infoworld.com/d/security-central/youtube-blocked-in-china-flickr-blogspot-restored-351
Le Petit Journal : HTTP://www.lepetitjournal.com/content/view/15253/312/
Thaïlande.fr : HTTP://thailande-fr.com/actu/83-la-thailande-envisage-de-poursuivre-youtube
Le Monde : HTTP://www.lemonde.fr/technologies/article/2011/01/05/la-tunisie-tente-de-reprendre-le-controle-du-web_1461205_651865.html
Info France 2 : HTTP://info.france2.fr/sciences-tech/la-turquie-bloque-de-nouveau-youtube-65688009.html

6.2 Pour le filtrage :

Copyright Protection in the Internet
White Paper
IPoque
HTTP://IPoque.com/sites/default/files/mediafiles/documents/white-paper-copyright-protection-internet.pdf

Internet blocking balancing cybercrime responses in democratic societies
Etude
Cormac Callanan, Marco Gercke, Estelle De Marco, Hein Dries-Ziekenheiner

HTTP://www.aconite.com/sites/default/files/Internet_blocking_and_Democracy.pdf

Filtering the future?: Software filters, porn, pics and the Internet content conundrum
Thèse de doctorat
Christopher D. Hunter
HTTP://md.hudora.de/archiv/sperrung/hunterthesis.pdf

6.3 Pour les enjeux :

Pirates informatiques ou altermondialistes numériques, Anonymous : peuvent-ils changer le monde ? Frédéric Bardeau et Nicolas Danet. Editions FYP 2011 Paris
Multigesture pour l'illustration des miroirs de Wikileaks :
HTTP://www.multigesture.net/2010/12/09/visualizing-wikileaks-mirrors/#demo
Geni : HTTP://www.geni.net
Netiquette : HTTP://netiquette.fr/

6.4 Pour le glossaire :

Dictionnaire informatique et numérique - Henri Lilen – First Interactive 2011 Paris
WikIPedia pour les termes suivants : DPI – Hébergeur – NRO – Roaming – Routage.
HTTP://fr.wikIPedia.org/
ARCEP pour les termes suivants : opérateur de services – opérateurs de réseaux.
HTTP://www.arcep.fr/index.php?id=8055
Hacker's Guide – Eric Charton 4eme édition Pearson Education 2011 Paris.

7 | Annexe juridique : la mission de l'Hadopi en matière de surveillance du développement des technologies de filtrage

L'article L. 331-23 du code de la propriété intellectuelle prévoit que l'Hadopi « évalue (…) les expérimentations conduites dans le domaine des technologies de reconnaissance des contenus et de filtrage par les concepteurs de ces technologies, les titulaires de droits sur les œuvres et objets protégés et les personnes dont l'activité est d'offrir un service de communication au public en ligne. Elle rend compte des principales évolutions constatées en la matière, notamment pour ce qui regarde l'efficacité de telles technologies, dans son rapport annuel prévu à l'article L. 331-14 ».

De cet article, éclairé par les travaux parlementaires qui ont précédé son adoption, il résulte que le législateur a entendu confier à l'Hadopi un rôle de surveillance en matière de développement des technologies de filtrage.

Cette mission découle directement des accords de l'Elysée pris à la suite du rapport Olivennes, dans le cadre desquels les fournisseurs d'accès à Internet et les plateformes d'hébergement de contenu s'étaient engagés à collaborer avec les ayants droit sur les modalités d'expérimentation des technologies de filtrage des réseaux, alors insuffisamment mûres pour en évaluer pleinement le potentiel et les risques.

C'est à l'Hadopi qu'a été confiée cette mission d'évaluation future, qui ne peut se cantonner à un simple rôle de veille technologique. La loi fait obligation à la Haute autorité de s'emparer de la question dès le stade embryonnaire, pour guider la mise en place des mesures de filtrage qu'elle pourrait juger efficaces dans le cadre de sa mission de protection des droits

sur Internet.

Afin que la Haute Autorité puisse pleinement mener cette mission confiée par le législateur, ces expérimentations doivent en toute rigueur être systématiquement portées à sa connaissance pour être évaluées au regard de leurs impact en termes technique, économique, social et sur le plan de la protection des libertés fondamentales. S'agissant d'informations relatives à des technologies expérimentales, qui ne sont qu'exceptionnellement publiques, l'information de l'Hadopi doit en effet résulter d'une communication active et systématique de ces informations par les développeurs. Comme dans toutes les sphères de l'action publique, pouvoirs d'investigation et missions d'évaluation vont ainsi de pair, afin d'assurer la pleine information des autorités publiques préalablement à leurs prises de décisions.