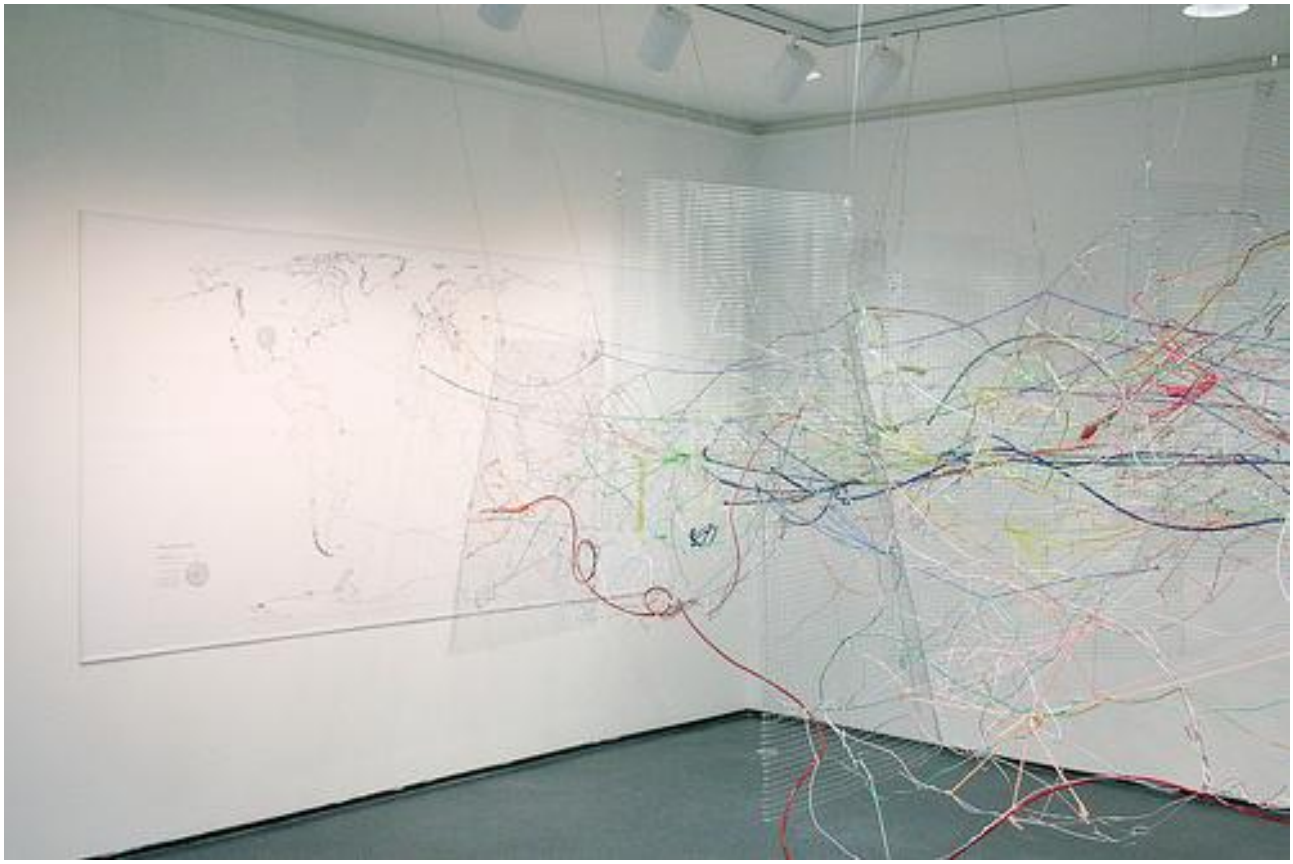
Principle, interests, limitations and risks of hybrid filtering in order to block resources on child pornography hosted on foreign servers.



Main author : Christophe Espern

Thanks to the contributers from IRC channel #fdn and FrnOG mailing-list

Cover picture : Mapping the internet, by Fausto Fernós
http://www.flickr.com/photos/feastoffools/2126692786/

Nota bene

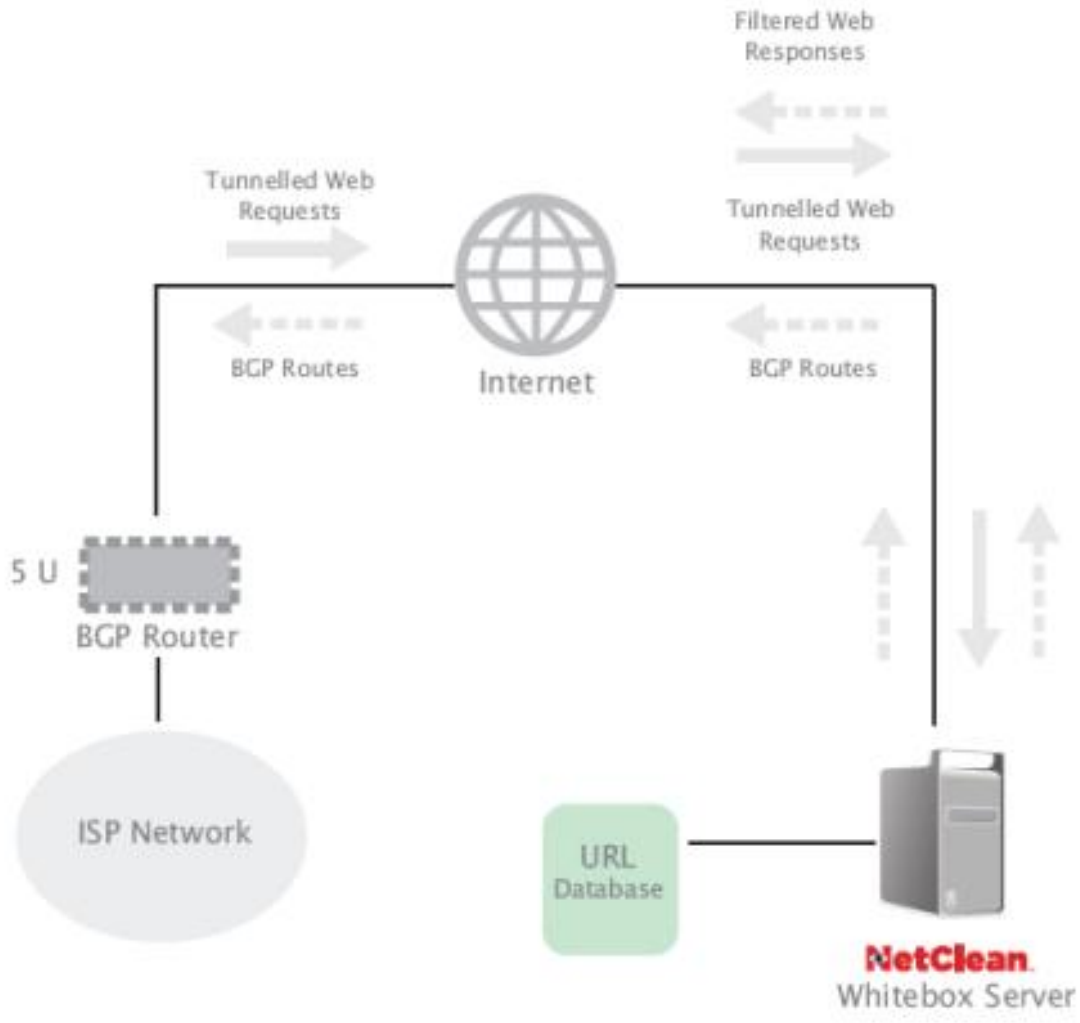
This document does not address content analysis filtering or protocols filtering (except in cases of collateral damage), but only filtering technology on IP, domain name or URL.

Principle

Through investigations or reports from Internet users, police maintain a blacklist of URLs pointing to resources on child pornography. This list is addressed to Internet Service Providers (ISPs) which prevent their subscribers to access these resources.

Specifically, ISPs recover from this blacklist the list of IP addresses corresponding to the domain names hosting the resources to block. They send a command to their routers via the protocol Border Gateway Protocol (BGP) for reconfiguration, so that any request to access a suspicious IP is routed to the filtering platform and not relayed to the server requested by the user.

Thus, when a subscriber requests access to a resource hosted on a website whose IP address was associated by an ISP to a URL listed by the police, the request is rerouted by the ISP's routers to the filtering platform which blocks the transfer if the corresponding resource is in the blacklist, or relays the communication normally otherwise.



Architecture of the Netclean system, typical hybrid filtering system

http://www.netclean.com/EN/documents/NetClean_Whitebox_Tech_EN.pdf

Interests and limitations

This method is called hybrid filtering because it combines several techniques to address the excess blocking issues inherent to IP or DNS filtering, while avoiding the expensive costs of deployment recquired with other filtering techniques based on URL.

It is thus possible to block simply a single picture from a webpage, without the need of such infrastructures as those implemented in countries like China or Saudi Arabia: the platform hosting the filtered data works only on a small portion of total traffic thanks to the pre-sorting by routers on IP addresses.

This method can be circumvented by the users via foreign proxy servers, and can be put in place with only a few clicks. The publisher of the filtered website can also take countermeasures, for example by switching to the https protocol, thus rendering the complete URL indecipherable to the filtering platform. He can also generate unique URLs on demand.

A university study [Clayton, Cambridge, 2005] suggests also that in the United Kingdom where such systems are already deployed, some publishers of filtered content are already using decoy techniques to identify the computers responsible for the blacklisting process, and are therefore able to hide their websites.

Risks

The overall cost depends mainly on the amount of data to analyse, on the operators architecture, and on the consequences in case of overload, configuration errors, abuse or attack of the system, which risks are very real as the following intends to prove.
This damages may involve the responsibility of the French State and have significant disrupting consequences on legitimate, sometimes critical, economical activities

Congestion risks

It's important to have accurate traffic predictions in order to be able to correctly dimension the network platform. But since an IP address can be shared, the traffic of some busy websites not initially targeted can be derived.

A university study [Edelman, Harvard, 2003] stressed that "*more than 87% of active domain names share their IP addresses (ie: web servers) with one or several additional domains, and more than 2/3 of active domain names share their addresses with 50 or more additional domains.*" Since then, this proportion can only have grown bigger.

Therefore, estimating the traffic loading in normal operation is difficult, especially as the publishers of child pornography websites change of IP address frequently, as highlighted by the filters manufacturers. It is able to thwart the countersystems already deployed, but this increases the cost of first level filtering since it involves the monitoring of an increasingly large number of IP.

The traffic to ingest can also suddenly increase drastically in the situation where a suspect IP is the target of a cyberattack designed to saturate it from the filtered network (Denial of Service).

In addition to the risk of enduring attacks targetting the filtered website or other websites sharing the same IP, the filtering system is at risk to be be the direct target of a criminal organisation and a reprisal attack.

Risks associated with the use of the BGP protocol

The use of BGP commands to redefine routes for the filtering of content is not a function for which the BGP protocol was intended.

Fro example, when Pakistan ordere the blocking of the Mahoammed cartoons hosted on You Tube, a pakistani network operator sent a BGP command to equipment that was poorly calibrated. It sent the request to network operators outside of Pakistan. Access to YouTuve was blocked for several hours in a number of countries. This event demonstrates the risks for national security, as some network specialists have highlighted.

«A small group of people could take control of a chain of BGP compatible routers to send BGP prefixes to the entire Internet. The result would not bring down the entire Internet, but would cause serious disruption on a large scale, which is exactly what you would want to do to mount a terrorist attack with maximum impact. In other words, the press coverage on this weakness in the BGP protocol highlights a possible means of attack which could cause serious problems in a period when people have the most need for the Internet ».

(*YouTube Black Hole – What's the real point?* http://www.getit.org/wordpress/?p=82).

Incompatibility with the technical architecture and contractual requirements

To ask network operators to permanently modify their routing configuration is not compatible with the use of standard optimisation techniques, such as route aggregation. This is especially true in France, in respect of the number of peering agreements between operators, and in which the rules of aggregation are the subject of specific contractual clauses.

Furthermore, the fact that criminal organisations use a technique known as Fast flux, (http://en.wikipedia.org/wiki/Fast_flux) – intended to regularly change the association of a domain name with an IP address – implies the frequent reconfiguration of routers, multiplying the risks of problems occurring and increasing the complexity of the configuration to maintain.

Risks of exposing the blacklist

A university study [Clayton, Cambridge, 2005] has demonstrated that the hybrid filtering systems in use in the UK, (CleanFeed, WebMinder) work in such a way that «*the system can be used as an oracle to efficiently locate all illegal websites* ».

The author has established that it is possible for a UK subscriber to obtain anonymously, within 24 hours, the list of all Russian sites on the UK blacklist. [1]

In addition to presenting an enormous risk that this list will circulate on the Internet, or is sold with a notice explaining to users how to circumvent the system, or with a programme that enables users to expand their own list, this weakness can be exploited to facilitate the circumvention by editors of filtered sites, or to maximise a denial of service attack, because it simplifies the observation of the system, and its flaws.

On of the vendors of these systems announced after the publication of the study that the problem had been resolved. The author has demonstrated that this is not the case.

1 The author of this study has not sought to obatin such a list, limiting his investigations to the security issues and evidence of a weakness, notably for legal reasons. His estimate of the time to obtain this list of Russian sites is based on a sample of Russian sites discovered via a network search voluntarily interrupted, and on figures supplied by the IWF which maintains the blacklist (25% of blocked sites are Russian).

Conclusion

The installation of a hybrid filtering system, whilst it appears seductive on paper, presents a number of consequential risks and has a limited effectiveness.

The direct and indirect costs may increase with the growth in usage. Users such as publishers of child pornography can always circumvent it, and also attack it. Its implementation risks reinforcing the techniques used by pedophiles and publishers of child pornography to hide and bury their activity, so it is not found by investigators. It also presents the risk that the blacklist will leak into the public domain.

Network specialists are concerned that this technique is envisaged, in light of its obvious weaknesses and the risks it presents for the entire network. Its implementation constitutes a regression in network terms. They consider that it is irresponsible for the Government to support such a filtering technique and call on it to take responsibility in the case where it is used by a network operator.

Post script : more generally, technical experts have called for the idea of filtering at the heart of the network, to be abandoned altogether. On the one hand, other techniques of this type (see Annexe I)imply either over-blocking or a limited effectiveness (DNS) or very limited effectiveness (IP) ; or an exhorbitant cost, and a limited effectiveness (proxy servers, RST). But overally, such a filtering will run counter to the very architecture of the Internet and its desired development, by recentralising the traffic flows.

In the course of the discussions which contributed to the writing of this paper, many subscribers to the FRnOG list talked of the installation of filtering devices by the ISPs in the connection equipment for subscribers. For some professionals, this technique was the only one they could envisage within the network architecture.

This technique poses several ral problems, of which some were discussed on the FRnOG list. They will be the subject of a further study, such as the legal and political issues inherent in a government project to filter the Internet.

Annexe 1 : References

Academic studies

[Edelman, Harvard, 2003] – Filtering by IP adress

Edelman, B.: Web Sites Sharing IP Addresses: Prevalence and Significance.Berkman Center for Internet and Society at Harvard Law School, 2003.

http://cyber.law.harvard.edu/archived_content/people/edelman/ip-sharing/

[Dornseif, Düsseldorf, 2003] – Filtering by DNS

Dornseif, M.: Government mandated blocking of foreign Web content. In: von Knop, J., Haverkamp, W., Jessen, E. (eds.): Security, E-Learning, E-Services: Proceedings of the 17. DFN-Arbeitstagung Äuber Kommunikationsnetze, Dusseldorf, 2003.

http://md.hudora.de/publications/200306-gi-blocking/200306-gi-blocking.pdf

[Clayton, Cambridge, 2005] – Hybrid Filtering (Cleanfeed, WebMinder, NetClean)

Clayton, Failures in a Hybrid Content Blocking System. University of Cambridge, Computer Laboratory, 2005

http://www.cl.cam.ac.uk/~rnc1/cleanfeed.pdf

[Clayton, Cambridge, 2006] – Filtering by URL using injection of RST packets

Clayton, Murdoch, Watson : Ignoring the Great Firewall of China. University of Cambridge, Computer Laboratory, 2006

http://www.cl.cam.ac.uk/~rnc1/ignoring.pdf

Other resources

Global view

The worst part of censorship is XXXXX : Investigating large-scale Internet content. 23C3, Berlin/Germany, December 29th, 2006

http://events.ccc.de/congress/2006/Fahrplan/events/1473.en.html

Discussions between technical actors on the mailing list FRnOG

Charte sur la confiance en ligne" vers une division de l'inter-net ?

http://www.mail-archive.com/frnog@frnog.org/msg02883.html

Filtrage via BGP shunt : quelle faisabilité ?

http://www.mail-archive.com/frnog@frnog.org/msg02939.html

Ping: il n'y a plus personne ? (à propos du YouTube blackhole)

http://www.mail-archive.com/frnog@frnog.org/msg02441.html

Annexe 2 : Other filtering techniques for ISPs

Filtering by DNS (Domain Name Server)

With this technique, it is not only the illegal content which is filtered, but the entire content of the Internet comain where the content is hosted (ex : geocities.com au lieu de geocities.com/siteperso/pedo.jpg).

In reality, an entire site hosting millions of personal pages could disappear from the Internet just because one image was not deleted within the time demanded by the authorities (some give 24h00, others expect instant removal).

This technique can equally be used to block sub-domains (ex : pagesperso.free.fr) depending on how the request is written. It can prevent communications not specified in the request, for example, it can prevent the sending and reception of other types of communication with the domain, not just access to the web pages hosted on it.

A university study, [Dornseif, Düsseldorf, 2003] which looked at the case of filtering a nazi website ordered by the German authorities, showed that of 27 ISPs all had made at least one error when they configured the filters. They had either not properly blocked the requested site (under-blocking) or they had blocked sites and protocols which weren't specified (over-blocking), or they had managed to do both (under and over-block).

Thus, of 27 ISps, 45% had both under and over-blocked, 55% had only over-blocked, and 27 (59%) had managed to block the communication with several domaines and all had blocked the email address of the target site, even though that wasn't specified by the judge.

The study underlined that « *web content is very volatile. Web servers are re-organised, domains get new owners. This was clearly shown in the case of the requests to block the site www.front14.org: in the autumn of 2001 this site contained a portal of the extreme right, but in the spring of 2002, it had a web catalogue with no political agenda. this underlines the need to identify pages to be blocked, not just by their location but by their current content.* »

the operations necessary to block using DNS are relatively simple, however the complexity of the ensuing maintenance, and hence the overall cost, depends also on the network configurations of the operators. The effectiveness of the technique is very limited. It is sufficient for a trivial manipulation on the user's computer to get around it. The publishers of child pornography only have to put in links using IP addressess instead of domain names, to get around it.

For more information : see Annexe 1 - [Dornseif, Düsseldorf, 2003]

Filtering by IP address

This requires network operators to maintain a list of IP addresses or blocks of IP addresses for which the routers will not transmit packets, but will simply ignore them. Thus, any exchange of content passing through a router applying this type of filtering, is impossible.

This technique blocks all access to a server or group of servers and does not permit different types of content or different web sites to be treated separately. The chances of over-blocking are therefore very high.

A university study [Edelman, Harvard, 2003] underlined this point : « *More than 87% of active domain names are found to share their IP addresses (i.e. their web servers) with one or more additional domains, and more than two third of active domain names share their addresses with fifty or more additional domains. While this IP sharing is typically transparent to ordinary users, it causes complications for those who seek to filter the Internet, restrict users' ability to access certain controversial content on the basis of the IP address used to host that content. With so many sites sharing IP addresses, IP-based filtering efforts are bound to produce "overblocking" -- accidental and often unanticipated denial of access to web sites that abide by the stated filtering rules.* »

Since 2003, the number of web sites exposed in this way, and the techniques of IP address sharing have grown, and thus the risks of over-blocking have increased.

This technique can be circumvented by the user, with the help of proxy servers situated in a foreign country, which the user can access using tunelling techniques, or using dedicated links. These methods cannot be filtered in a democracy, because they use a generic functionality. The publishers of content can circumvent IP filtering using automated re-assigning of new IP addresses to their domain names.

For more information : Annexe 1 - [Edelman, Harvard, 2003]

URL filtering by proxy servers

All requests from users in a country pass through filtering servers which block communications relative to a specified URL. Unlike hybrid filtering, there is no ‘tri’ prefix to the IP address. This technique implies that filtering platforms have sufficient redundancy as all web traffic is filtered.

This technique has been implemented by the national operators in Tunisia and in Saudi Arabia. Le cost of putting it in palce is exhorbitant in a competitive environment such as France where several operators co-exist. The company Noos used a similar technique a few years’ ago to deal with caching. It was abandoned, due to overblocking and the increasing cost of expanding the network. This filtering technique can be circumvented using proxy servers, as described above.

Filtering by URL using injection of RST packets

The URLs of web sites visited are analysed in respect of a list of keywords and a blacklist, and the routers send an RST packet to the client and the server, which results in the closure of the TCP connection. The connection is closed as soon as it is established and no content can be exchanged. This necessitates that all of the traffic to be controlled passed through a network infrastructure controlled by the authorities. It is one of the techniques used in China.

For more information : Annexe 1 - [Clayton, Cambridge, 2006]