# FIVE THINGS YOU MUST KNOW

## When Selecting an Internet Content Filter

## Contents

## Introduction

Before we even think about what to look for in an internet content filter, we should ponder why would we install a filter in the first place?

In the school or public library setting we would want to protect young minds that could not process all of the information that is freely available on the internet. There would be a concern of letting children see hate literature or pornography when they have not developed an understanding of the context of the information. The London School of Economics released a report a few years ago that stated:

> Nine out of ten children between the ages of 8 and 16 have accessed pornography on the Internet. In most cases, these sites were stumbled upon while searching for information that was homework related. Using the Internet with a filter is using the Internet responsibly**.**

The reality as quoted above led the US government and other governments to implement CIPA or similar legislation. If a school or library is to get public funding they need to ensure that they have adequate internet protection.

In a business setting it becomes a matter of economics and focus. We want our employees to be focused on the task at hand rather than become addicted to gambling or pornography at work. There are a number of statistics about online internet access at work. All of them point to the personal use or abuse of corporate resources.

Even if we are not concerned with what unlimited surfing on the web could do for a person's time or mind there may be a concern for the impact on the network. There are known sites that are infected with spyware or malware. The malware software will record keystrokes including user accounts and passwords. It can also be used to take over a PC and use it to send SPAM. An internet service provider may cut off the internet access until the infected PC has been identified and removed from the network.

In a network setting a content filter could be used to protect the network and to conserve network bandwidth. It does this by stopping certain types of sites (video or audio streaming) so that the limited bandwidth can be used for other internet and network functions.

Depending on the reason for using an internet content filter will give weight to the following five considerations more than others.

## What you Filter

In the past an internet content filter was concerned primarily with the contents of web pages landing on your browser. There is still a concern for this; however, more dangerous content can arrive in your network through other protocols. Peer to peer networking and instant messaging are a group of newer technologies that also can attract undesirable content and software.

Browser based traffic typically comes on one of two internet protocol ports. One port is for the general traffic and another port for secure internet transactions such as online banking. To restrict access to peer to peer or other traffic you could close all other internet protocols and ports except for the ones that you specifically allow. That is a good practice; however, instant messaging or peer to peer file sharing software doesn’t always respect assigned ports. The software will move the content through whatever ports that are left open. So file sharing or instant messaging could be going through port 80 which is normally reserved for regular web surfing.

As worms such as SQL-Slammer, Blaster, or Code-Red hit our networks, transported via protocols that we use on a regular basis, the need for devices that consider the content of packets becomes more and more critical. Deep Packet Inspection is an excellent method to shut down some of the most used attack vectors exploited by malicious content today.

Deep packet content examination is not something new. Antivirus software and network Intrusion Detection systems have been doing it for years. What newer Deep Packet Inspection devices bring to the table are preloaded signatures, similar to those used by an antivirus solution. This way, your firewall is aware of and able to detect and remove malicious content as it arrives at your network. Since the packet's content is being considered at the application layer, traffic anomalies representative of an attack or worm can also be considered and filtered even if a specific signature isn't available for it. For example, if some attack uses a command that is considered nonstandard for a particular protocol, the Deep Packet Inspection would be able to recognize it and drop the malicious content.

To block or restrict peer to peer or instant messaging traffic requires a more complete internet security and content filtering solution than just a pure URL blocker. If all you need is to block web page requests through a browser then the next question of selecting a content filter on understanding a website filtering content technology should answer most of your questions. If you are considering blocking more than “bad” websites then consider a content filter that includes deep packet (or layer seven) inspection.

## How you filter it

Gone are the days when entering ten or twelve bad words in the firewall would block all the bad websites. What is known as "keyword blocking" would give some typically bad over blocking examples. In the attempt to block porn sites the word "breast" may have been used. This would also block websites on breast cancer, chicken recipes, and athletic sites to name a few examples.

**The most common internet content filter is the URL blacklist.** Depending on the source, known "bad" websites would be classified from twelve to seventy two categories. A banned site would be classified under pornography, gambling, hate, violence, time wasting, on line shopping, dating etc. The administrator would then select which categories to block.

There are no internet standards for the categories; each supplier has their own criteria for putting a website into one or more categories. With new websites coming online every day it is critical that the blacklist is kept up to date. With literally millions of "bad" websites on the internet and even with claims by the top internet blacklist suppliers there are bound to be websites not classified in their lists. The question then becomes what does the product do if the website is not on the list? There should be some policy as to what is done with unclassified websites. Do you let them through, do you note them to be classified and then let them through or do you block all unclassified websites?

How "granular" is the content filter. Does it classify the website at the domain level or does it work at a URL level. How do you classify MSN or someone's personal home page which could be something like www.isp.com/myhomehimepage.html?

One way to deal with the problems of the URL blocker is to have a program that analyzes the content of the webpage. **These are known as intelligent content filters.** They will have a number of algorithms to determine if the website is acceptable or not. The content could be blocked on ICRA or similar tags put into the webpage. These are standard website rating tags established by the Internet Content Rating Association. (ICRA). The problem with the tags is that it is a voluntary rating system and the rating is at the discretion of the website creator. As an example, we did a demo of our content filter on a breast cancer site. Everything went through fine until we hit one page. It should have shown us a picture; however, the author rated the picture as "nudity" rather than medical and so the page was blocked. For pages that are not rated there is other logic to determine if the page is acceptable. In the same way that email can be determined "SPAM" there are algorithms that will determine the content of each web page before it is displayed on the screen. The advantage of this technology is that it will catch "bad" content regardless of the source. Since it works on patterns it does not require daily updates in the same way as a URL blocker does. The downside is that the intelligent content filter takes more computing horsepower to determine whether a page is acceptable and may introduce a delay in delivering web content.

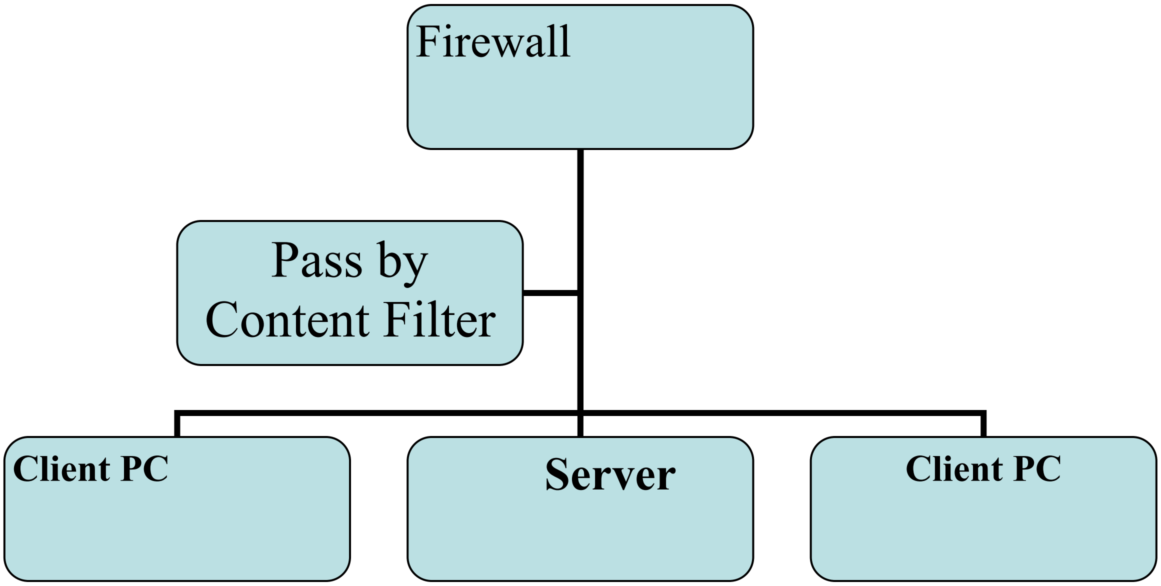
There are short comings with either technology so **one answer is to use a combination filter.** A combination filter will start with the URL blocker. It will check the domain or URL against the various lists of known bad sites and the white list (good sites). If the domain or URL is not found or has been grey listed then the policy would be to use the intelligent content filter to determine if the web content is safe.

This gives the speed of the blacklist yet covers new sites or unclassified sites with the pattern matching logic of the content filter. It is faster than just using the intelligent content filter to determine the outcome of known sites. It quickly blocks the obvious "bad" sites immediately and can block content coming through search engines, proxy servers, and bad content on one page of a social network or other unlisted sites.

## Where you Filter

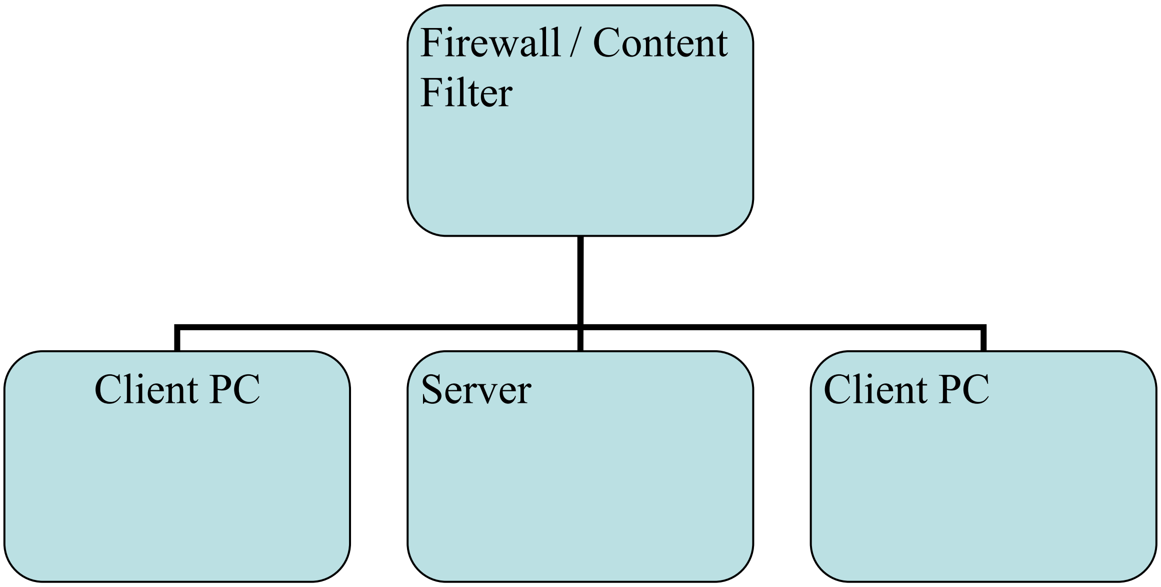
Where you place the content filter is as important as how you filter. The early solutions would place a PC based package on each PC. In a larger network it can get expensive and difficult to administer the policies and updates. A technically proficient user could also bypass the PC based content filter.

Another location for content filter is the Pass-by-method. The software is installed on a separate PC or server and all web requests are sent to this server. This may work for a URL blocker; however, it could be bypassed by routing the request straight through to the firewall. You would have to ensure through group policies that all requests are indeed being checked by the designated server. If you are using an intelligent content filter in a pass by mode it effectively is an internal proxy server. A proxy server could be bypassed.

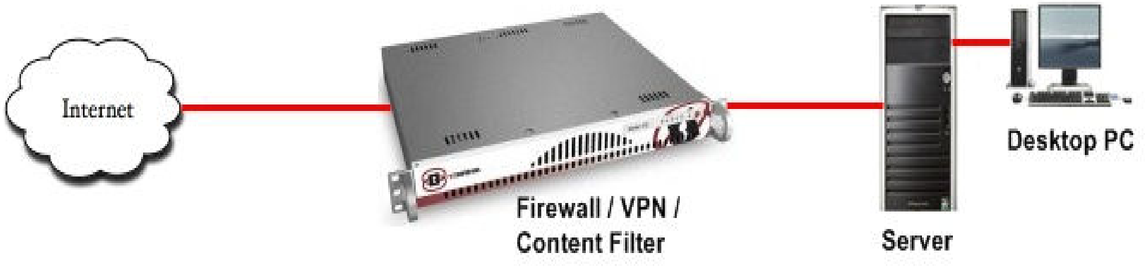


Pass by Content Filter Placement in the Local Area Network

A major disadvantage of content filtering at these locations is that the worm, Trojan horse, a malicious packet, or other undesirable content has already entered your network. Firewalls offering Deep Packet Inspection technology have the ability to detect and drop packets at the entry point of the network. What more appropriate place to stop malicious traffic than at the firewall? If there is only one point of entry and exit then placing the content filter in the firewall or between the firewall and the rest of the network ensures that all traffic must pass through the content filter before it is allowed in the network. This would be the best protection currently available.



Pass through content filter placement in the Local Area Network.



## Who you filter

Who you filter really defines where you place the content filter and is as important as how you filter. Typically there is standard policy put in place and all users must follow the policy. **All users** should be filtered.

In a school setting or library setting the most common exception to this rule is for the librarians and teachers. Students are filtered and the teacher may not be filtered or may have a different level of filtering. In a medical clinic everyone should be filtered to protect from spyware and pop ups; however, medical sites and 'nudity' would be allowed for professional uses.

If there are different levels of content filter, then the user needs to be authenticated against a master profile list to ensure they have the correct internet access. The content filter should be integrated in to an authentication routine so that a valid profile would be logged in the user activity logs. The authentication would be either active directory to a Microsoft Windows server, IDENT or LDAP. If there is no integrated authentication available then the content filter should provide its own internal user authentication system.

## How you manage it

How you manage your network and content filter is a primary concern of the network administrator. The previous considerations have an impact on the end user experience of the Internet. The administration of the content filter has a less direct impact on the user other than it may improve or hinder the administrator's response to user requests.

How many levels of management do you want or need for the content filter? In a school, library or large enterprise situation you would want at least two levels of management. One level would be for the senior network administrator to set-up the filter and provide the global settings. The second level would be for the librarian, teacher, office manager or other administrator to manage the exceptions to the content filter.

The network content filter should record all network activity. It should track and report on all web access for all users. If the access was denied the reason for the denial, and if there was an over ride, should also be logged. It may also be required that you can export the history for archival and reporting purposes.

When a user is blocked the system should display a screen that would make it clear that access was blocked and why it was blocked. Optionally it should allow the user to request an over ride or allow for an over ride pass code. The pass code should be able to be configured to give 20, 30 or 60 minutes of unfiltered internet access. In a library setting it would be preferred to have an override pass code that could be given out. In an enterprise or school setting it may be preferred that a request be sent to permanently override the content filter block.

The internet world is just as heterogeneous as the real world. With the exception of small peer networks you can not expect that everyone will be running the same computer and operating system. Linux, Apple and Microsoft all share a part of the network. A consideration would be if there is any special software that needs to be installed on the network administrator's personal computer. The special software may give a better graphical user interface; however, it does presume or limit the choice of computers that you can use for administration. A web based interface is not dependant on a single operating system platform.

In a school district, library district, municipality or business network there can be many nodes or locations. Consideration should be given for remote management and tools available to monitor and remotely manage multiple points in a network. Tools should be available to co-ordinate blacklists and policies across multiple units. Perhaps a master profile can be made up, backed up and then uploaded to all other units. A facility to export the activity logs would help with archiving the results and to coordinate patterns across multiple locations.

In summary, the administration of the content filter should be intuitive and have at least two levels of administration. The content filter should also allow for secure remote access for administration of one or more units and it should have some capabilities to ensure that there is consistent application and reporting of internet access across the entire network.

## About KDI and the NetSentron™

**Kobelt Development Inc**. (KDI) is a Vancouver Canada based software development and information system support company. KDI started by providing custom written software and International Business Machines Corporation (IBM®) solutions.

We actively seek organizations that we can help benefit from the effective use of information systems. We do this with a team of professionals that understand technology as well as our client's objectives. We work as advocates on our client's behalf to develop and maintain systems designed to meet their goals and expectations.

NetSentron was developed as a natural result of implementing wide area network solutions for our clients. The content filter started as a standard URL blacklist filter and was later developed and enhanced to become a combination content filter as our client base grew to include schools as well as business.

**NETSENTRON** is a registered Trademark of Kobelt Development Inc.

**Kobelt Development Inc. (KDI)**
**#404 – 17768 65A Avenue**
**Surrey, BC**
**Canada**
**V3S 5N4**

**1-604-574-7225**
**Toll Free: 1-800-661-1755**
**www.kdi.ca**
**www.netsentron.com**