



Web Filtering in the Workplace: Finding the Best Solution

Wavecrest Computing
2006 Vernon Place
Melbourne, FL 32901
Toll-free: 877-442-9346
Voice: 321-953-5351
Fax: 321-953-5350

www.wavecrest.net

Web filtering is typically part of an overall Web-use management program that includes development, communication and enforcement of an Acceptable Usage Policy and users companion software to generate compliance and activity reports.

Introduction

This paper discusses the use of off-the-shelf Web filters in the workplace. It was written by the staff at Wavecrest Computing, a long-time provider of Web filtering and other Internet usage management products and services. Its intended audience includes IT administrators and business managers who are considering the use of a Web filter in their operations. The paper has two purposes:

- To help readers decide if Web filtering is right for their organizations.
- To help those that believe it is right, choose a reliable and effective solution at a reasonable cost.

The authors' *major* – but not only – emphasis is on Web-filter *technology*. However, they do touch briefly on a number of other essential and related subjects, e.g., Acceptable Usage Policies (AUPs), privacy issues, legal considerations, Web-use reporting, online workforce management, and workforce morale.

Corresponding to the two purposes cited above, the paper is divided roughly into two parts:

- The first provides information designed to help the readers decide if Web filtering is right for their organizations, answering questions such as what is Web filtering, why filter, is it fair and reasonable?
- Then, for those who decide that they DO have a valid need, it answers questions such as what are the challenges of implementing a Web-use filtering solution? What are the different types of filters and filtering methods? Just how reliable and effective are they? And how important is vendor selection?

After covering these topics, the authors make a case for their company's filtering products and related services. In doing so, they describe how Wavecrest Computing works closely with individual customers to maximize the reliability and effectiveness of their Internet access filtering programs.

I'll get started with the question, "What is Web Filtering?"

What is Web Filtering?

Most readers will already know the answer to this question, but to cover all possibilities, we include this brief definition.

At its most basic level, Web filtering is the use of software to restrict the (Web) content that one or more computer users can access via the Internet. Filtering can be implemented at various levels: by a government on a nationwide basis, by an ISP to its clients, by a school to its students, by a library to its visitors, by a parent to a child's computer, by an employer to its personnel, etc. Typically, its purpose is to prevent the user(s) from accessing content that the computer's owner(s) or other authorities consider inappropriate or detrimental in some way.

Please note that this paper is concerned solely with the use of Web filters in *organizational* settings, not household situations. Organizational settings include businesses, schools, government agencies, military units and non-profit entities. In these settings, Web filtering is typically *part* of an overall Web-use management program that:

- includes development, communication and enforcement of an Acceptable Usage Policy.
- uses companion software to generate compliance and activity reports. (See note.)

Note: Organizations seldom use Web filtering by itself. Almost all combine it with Web-use reporting. And many use reporting without filtering. While the focus of this paper is on filtering, the importance of Web-use reporting should not be ignored.

Why Filter?

Why should an organization filter its computer users' access to the Web?

There are several reasons. And the major ones – discussed below – are getting more and more critical.

Let's take a brief look at some history.

Use of the Web in workplaces is not new. It started rather simply in the early to mid '90s as a means of obtaining information and advertising products. Soon thereafter, managers and IT personnel noted that the Web brought risks as well as benefits. The earliest risks, which are still around, were:

- loss of productivity caused by workers surfing for personal reasons
- legal liabilities caused by workers downloading pornographic material.

These issues were soon addressed – with reasonable success – through use of early and somewhat rudimentary Web-use monitoring and filtering solutions. (Wavecrest Computing was one of the first in the field, starting in 1996.) Dealing with those early risks entailed some challenges, but as time passed, it became steadily more difficult. That's where we are today.

Let's explore this evolution a bit.

In the 1990s the Web was not an *extremely* 'risky' place. Users could only request and obtain 'static' text and simple graphic information. The concept of online interaction was still in the future, and hacking was in its infancy. It's true that – as mentioned above – organizations were open to the possibility of productivity losses and legal liability, but much worse problems lay ahead.

Starting in the early 2000s, as Web technology grew exponentially in sophistication and usage, organizations of all types and sizes began using the Web in more and more advanced ways. Perhaps the most notable example in recent years was – and still is – the exploding use of interactive Web 2.0 and 'streaming' technologies. The former includes numerous social networking and two-way Web-enabled applications. Facebook alone now has 300 million active users. And 'streaming' includes more and more use of high bandwidth video, audio and Flash transmissions.

Used wisely and with proper safeguards, these remarkable innovations *can* benefit organizations in terms of improved business-related communications, collaboration, automation and efficiency.

Unfortunately though, those same technologies come with significant and ever increasing risks.

And – just as in the early days – these risks originate primarily with personal surfing. Only now they are magnified by the phenomenal increase in (a) the number of hackers (many with criminal intent) and (b) the number of temptations to surf.

Numerous studies bear this out. In addition to the productivity and legal issues mentioned earlier, personal surfing in today's world can lead to serious network performance issues, unnecessary bandwidth costs and – perhaps most significant – business information leaks and network security compromises.

Let's look at some evidence.

First, some statistics provided by International Data Corporation (IDC):

- **70%** of all web traffic to Internet pornography sites occurs during the work hours of 9am-5pm.
- **58%** of industrial espionage is perpetrated by current or former employees.
- **48%** of large companies blame their worst security breaches on employees.

Personal surfing in today's world can lead to serious network performance issues, unnecessary bandwidth costs and – perhaps most significant – business information leaks and network security compromises.

- **64%** of employees say they use the Internet for personal interest during working hours.
- **37%** of workers say they surf the Web constantly at work.
- **27%** of companies say that they've fired employees for misuse of office e-mail or Internet connections, and 65% report some disciplinary measure for those offenses.
- **90%** of employees feel the Internet can be addictive, and 41% admit to personal surfing at work for more than three hours per week.
- **60%** of security breaches occur within the company - behind the firewall.
- **25%** of corporate Internet traffic is considered to be "unrelated to work."
- **30-40%** of lost productivity is accounted for by cyber-slacking.
- **32.6%** of workers surf the net with no specific objective.
- **27%** of Fortune 500 organizations have defended themselves against claims of sexual harassment stemming from inappropriate email.
- **90%** of respondents (primarily large corporations and government agencies) detected computer security breaches within the previous 12 months, 80% acknowledged financial losses due to computer breaches, 44% were willing and/or able to quantify their losses, at more than \$455 million.

The average employee spends over 75 minutes per day using office computers for non-business related activity. That translates into an annual loss of \$6,250 per year, per employee.

More evidence.

According to a recent Gallup poll, the average employee spends over 75 minutes per day using office computers for non-business related activity. That translates into an annual loss of \$6,250 per year, per employee. An average mid-size company of 500 employees could be expected to lose \$3.25 million in lost productivity due to Internet misuse. Put another way, just 20 minutes a day of inappropriate Internet use can cost a 100-employee company over \$8,000 per week (assuming \$50 per hour per employee).

For a quick illustration of how much casual surfing of the Internet could be costing an organization, look at the table below. We used the following assumptions:

- Average hourly cost per employee including overhead: \$20
- Average time spent per week casually surfing the Internet: 3 hours/week

Number of Employees	Cost/Day USD	Cost/Week USD	Cost/Year (47 weeks) USD
5	\$60	\$300	\$14,100
15	\$180	\$900	\$42,300
30	\$360	\$1,800	\$84,600
60	\$720	\$3,600	\$169,200
400	\$4,800	\$24,000	\$1,128,000
1000	\$12,000	\$60,000	\$2,820,000

As bad as this is, productivity is not the only issue.

Security compromises are becoming increasingly costly. Countless numbers of viruses, trojans, worms and other types of malware enter business networks through social networking sites, web-based e-mail accounts and various files downloaded from the Web.

Today's Web 2.0 downloads and interactive processes can consume huge amounts of bandwidth.

Social networking can be particularly troublesome and costly. The FBI states that fraudsters are increasingly hijacking accounts on social networking and other sites and spreading malicious software. One of their favorite techniques involves applications (“apps”) advertised on social networking sites. Appearing legitimate, some of these applications install malicious code or rogue anti-virus software.

Other fraudster techniques involve the use of spam to promote phishing sites. Many of these claim there has been a violation of the terms of agreement or some other type of issue that needs to be resolved. Other spam entices users to download an application or view a video. Others appear to be sent from users' "friends," giving the perception of being legitimate. Once the user responds to the phishing site, downloads the application, or clicks on the video link, their computer, telephone or other digital device becomes infected, the FBI stated.

The examples above point out productivity and security issues, but they are not the only problems.

Bandwidth abuse (aka “hogging”) – with its attendant dollar costs and network performance impact – is another huge issue. Why? Personal surfing at work is no longer limited to the downloading of plain text or simple/static graphics— actions that have minimal impact on the organization’s bandwidth capacity. Today’s Web 2.0 downloads and interactive processes can consume huge amounts of bandwidth. For example, if a user downloads a large number of songs or movies on a regular basis, bandwidth usage may reach unacceptable levels, perhaps 100 times that of an average user, causing the connection of fellow users to suffer and costing the organization considerable money.

For all these reasons and more, increasingly large numbers of organizations are finding it necessary to filter Web usage. From a 2005 study by the American Management Association:

- Sixty-five percent of companies use software to block certain websites, a 27% increase since 2001, according to the AMA.
- Three-quarters of companies monitor employees' website connections in large part due to concern about inappropriate Internet surfing.
- More than a quarter have fired workers for misusing the Internet.

Is It 'Fair' and 'Reasonable' to Filter Web Access?

Considering all the evidence, it would seem obvious that Web filtering is essential in today’s workplaces. Not so. Some argue that Web filtering is an unreasonable invasion of privacy, while others claim that it damages workforce morale. Let’s look briefly at these two assertions:

Privacy. In general, privacy is obviously desirable. But in light of the evidence presented above, should it be a 100% ‘reasonable expectation’ – especially in an organizational setting? After all, by definition, an organization is not a ‘private’ setting, and the resources used for workplace surfing belong to the organization, not the surfer. And courts have held that total privacy in the workplace is *not* a ‘reasonable expectation.’ To the authors of this paper, the ‘privacy’ argument seems very weak.

Morale. Workforce morale is obviously very important, but it leads us to another question, “Should employees let their morale be dependent on personal activity that is not related to their work?” Put another way, “Shouldn’t their morale be more dependent on job satisfaction, relationships with management and co-workers, pay, benefits, workplace atmosphere, etc.?” From the authors’ perspective, it’s difficult to argue against the latter.

The good news in all this is that skilled management – by communicating appropriately with the workforce – can implement reasonable restrictions without damaging morale.

The good news in all this is that skilled management – by communicating appropriately with the workforce – can implement reasonable restrictions without damaging morale.

A key part of the management effort is careful selection of a vendor who will be a helpful partner in the process from beginning to end – from product evaluation to rollout, administration and beyond.

What's Next? For purposes of the remainder of this paper, we *assume* that the reader agrees that Web filtering in the workplace is reasonable and essential. Consequently, under that assumption, the following discussions move ahead, focusing on the managerial, functional and technical aspects of the implementation of a successful Web-filtering implementation program.

Implementation of a Web Filtering Program

General. Implementation of a successful long-term Web filtering program doesn't just 'happen.' Why? The World Wide Web is truly immense, and it grows rapidly every day in sophistication and size. And for purposes of our subject, it is a constantly 'moving target.' Consequently, a successful program requires more than simply installing a computer application and letting it run. It requires competent managerial policy-making, planning, involvement and follow-through.

A key part of the management effort is careful selection of a vendor who will be a helpful partner in the process from beginning to end – from product evaluation to rollout, administration and beyond.

Let's briefly examine the process.

Policy Issues. The foundation of any Web-use management program is an Acceptable Usage Policy (AUP). In this regard, management must decide (or reconfirm):

- what constitutes acceptable and unacceptable usage of the Internet within the organization—and why?
- should the necessary restrictions be applied differently to different groups of users or individuals?
- should restrictions be applied 24x7 or only during certain hours.

Management then needs to incorporate these decisions into a thoughtful, comprehensive AUP. Once satisfied with the wording, management and (if applicable) HR must communicate and explain the AUP and plans for its enforcement clearly and repeatedly to the workforce.

The Outsourcing Question. Having settled the policy issues, or at least gotten them well underway, management now needs to decide whether or not to outsource the technical and operational aspects of the program. At this juncture, they have two options:

1. **Software-as-a-Service (SAAS).** One way to implement a Web-use management program – including filtering – is to engage the services of a SAAS vendor. For a price, a SAAS vendor will offer to handle all the technical aspects of the program. That is, they will provide all hardware and software, configure these resources to support the customer's AUP), provide day-to-day technical administration, furnish reports, etc. This approach can relieve the customer of some technical (IT) workload. But there's a downside. The SAAS approach takes the control of sensitive or proprietary data out of the organization's hands. It can also introduce latency in Internet connections depending on the location of the SAAS vendor's servers; this in turn can impact network performance.
2. **In-House Control.** The more common approach is to have the organization's own IT personnel or consultants manage the selection of a competent vendor and acquisition, implementation and administration of appropriated software (and possibly hardware). Under this option, the installation would be located in the organization's own facility.

Note: Of these two approaches, 'in-house control' is by far the most common. It is also the one that the authors consider the best for most organizations when all factors are considered. Consequently the following discussions and examples are based on 'in-house control' and not on the SAAS approach.

Implementation Assignments. Having settled on the 'in-house' approach, management's next crucial step is to appoint specific individuals to be responsible for the remaining tasks in the overall program. Included are software (or appliance) acquisition and testing, system installation and configuration, and long-term deployment and utilization. Most immediately, the organization will need a responsible project engineer or manager to direct the selection of a Web-use management vendor and product.

Vendor and Product Selection. Selecting a Web filtering product and vendor involves a number of important considerations. Some deal with the product itself, some with the vendor, and some with both.

Product considerations include *functional* and *technical* issues.

1. Functional. Consider these functional issues:

- How does the software keep up with the extremely dynamic nature of the Web?
- How difficult is it to configure the software to reflect, support and help enforce the organization's unique AUP?
- How does the software detect unacceptable content and how well does it do it?
- In addition to filtering, does the product produce useful Web activity reports?

2. Technical. Product *technical* issues include network security, infrastructure compatibility, fail-safe potential, software scalability, initial and follow-on costs, training requirements, maintenance and administration manpower requirements, impact on network performance, etc.

Vendor considerations include qualifications, experience, and technical support capabilities. Also included is knowledge of workforce management and AUP design and enforcement in general. Other vendor considerations include questions such as:

- How willing and able are the potential vendors to partner with the customer for the long haul, i.e., after the sale is made?
- Will they provide updates, upgrades and technical support within the price of the product?
- Will they provide usage tips, advice and help with product customization?

And so on.

And finally, there are the 'bottom line' concerns of reliability and effectiveness which are the focus of this paper.

To begin to answer these questions, it's first necessary to know something about the different types of filters and filtering methods that are available today.

When searching for a filtering product vendor, considerations include qualifications, experience, and technical support capabilities.

Web Filters: Types and Methods

Types of Web Filters. Broadly speaking, from an installation/integration perspective, there are three types of Web filters, and each has several variations.

1. The first type (client-based) requires that the filter software be installed directly on a personal computer. Since this approach is not suitable for organizational use, it will not be discussed here.
2. The second type consists of content-control software that is integrated into the customer's network gateway security infrastructure (e.g., firewalls or proxy servers). Both of these solutions exist in two forms: proxy and passive.

Proxy. Proxy technology prevents traffic from continuing to the targeted destination until inspection has occurred and deemed acceptable under pre-defined rules.

Passive. Passive solutions are by-pass or side-scan filters leveraging content 'signatures.' Passive technology allows packets to pass the filter to the destination. Inspection terminates the traffic by using TCP/IP resets for HTTP protocol and application level knowledge for peering protocols.

3. The third type – generally referred to as an Internet appliance – consists of software packaged in its own hardware. Appliances can usually detect, filter and report on IP protocols other than Web (http).

Web Filtering Methods. With any of these types, filtering Web traffic on the basis of content is a highly complex undertaking with some inherent limitations. To our knowledge, there are only four significant technical methods of doing it. While some come close, none are 100% successful, and each has its advantages, disadvantages, tradeoffs and limitations. The four are:

- **Content scanning:** block a Web page if it contains "bad" words
- **Artificial intelligence:** an improved version of *content scanning*
- **Blacklist:** block sites based on a list (database) of pre-categorized websites
- **Deep Packet Analysis:** block traffic on the basis of content *technology*, e.g., Flash, video streaming, audio streaming, images, Active X and more.

Let's review these briefly.

Method A: Content Scanning. When web pages are scanned for content, they are first downloaded – which costs time and bandwidth. Then the content is scanned for 'bad' (i.e., sexually oriented) words. Depending on the vendor, one or more words trigger the blocking mechanism. The theory sounds good, but in practice many sites are blocked because of word combinations like "sex changes", "breast cancer" etc. This is called *over-blocking*. On the other hand, sites with sexual content that only have pictures (text can also appear in a picture), are not blocked because they don't contain any of the 'bad' words, which is called *under-blocking*. The time that it takes to scan and guess the type of content of a web page varies per page (some pages are *very* large). This method is sufficiently fast for an individual user. However, for 250 or more users, a very fast computer system is required for the proxy server.

Method B: Artificial Intelligence (AI). When web pages are blocked based on artificial intelligence (AI), they are also downloaded first and then scanned, so this method also consumes bandwidth and time for the download process. The various AI methods are more complex versions of method A. To reduce the failures caused by under-blocking and over-blocking, all words in the web page are rated, and some word *combinations* are rated. Some products try to determine if a picture contains nudity by looking at colors and claim a high level of correctness. This improvement of correctness of blocking comes with a large cost: much CPU power. So, for 100 or more users, a very fast computer system is required for the proxy server.

There are four Web filtering methods. They are content scanning, artificial intelligence, blacklist and deep packet analysis.

Method C: Blacklist. When web pages are blocked with the use of a blacklist, they are not downloaded to make a block vs. allow decision. Instead, the URL filter module of the proxy server makes a quick decision based on the URL alone, e.g., www.sex.com is blocked and www.google.com is not. The URL filter makes this decision based on a database ('list') that is structured in content categories (shopping, news, pornography, travel, etc.) that can be designated as 'block' or 'allow.' Such databases are often referred to as a blacklist, URL list or control list. This method is fast since the blocked sites are not downloaded. However, for it to work, a suspicious site must have been previously categorized and entered in the list. This is achievable to a workable extent—but not to 100% perfection.

Note. An interesting variation of the 'blacklist' method is the 'white list.' A white list is simply a group of Web sites (or categories of Web sites) that are authorized to be accessed. When the white list approach is used, all sites *not* on the list are automatically blocked.

Method D: Deep Packet Analysis. As used in this paper, Deep Packet Analysis is an information technology process used by Web-use management software to analyze Web traffic at the individual packet level. It addresses information extracted from the data part of a packet, not just the header. This method allows finer-grained classification of data types than approaches based only on header information, e.g., the Shallow Packet Analysis method. (Note: Data types include Flash, video streaming, audio streaming, images, Active X and more.) Once a classification is made, the software compares it with the customer's predefined 'block' criteria to enable the product to decide what actions (if any) to take on the packet; one of those actions is blocking.

Similarly, the product can use deep packet analysis technology to identify anonymizers (public proxies) for blocking purposes.

So, which of the four methods works best? The authors believe strongly that the best approach combines C and D. Why? Method C can identify 'blockable' traffic on the basis of a page's overall content (i.e., its 'theme'), while Method D enables blocking on the basis of data type—regardless of content or 'theme.' This two-pronged 'double-barrel' approach provides more coverage and better results than any one scheme alone.

Reliability and Effectiveness of Web Filters

As touched on above, no Web filter is absolutely perfect. Methods A and B are highly suspect because of the need for automated interpretation of the content being 'inspected.' Also, they are limited to pornography only. And Methods C and D are continuously challenged by the huge and constantly growing number of websites and data types on the Internet.

With all these impediments it may seem that effective, properly targeted Web filtering is simply a "mission impossible."

Not necessarily so.

Skilled Web filter developers have learned to combine ongoing *personalized customer support processes* with the best technology approaches to provide each individual customer with a *workable* long-term solution—one that blocks upwards of 90% of the content they consider inappropriate.

Put another way, the authors strongly believe that technology alone is not sufficient to successfully manage Web usage in the workplace. It takes a *continuous collaborative partnership arrangement* between customer and vendor to do so. Only then can customers achieve the level of control and 'fine-tuning' needed to get the most out of Internet access while minimizing the associated risks.

It takes a continuous collaborative partnership arrangement between customer and vendor to successfully manage Web usage in the workplace.

Wavecrest Computing

Before discussing Wavecrest's approach to Web filtering, let's take a brief look at the company itself.

Established in 1996, Wavecrest Computing develops and markets 'industrial strength' Internet usage management products for all types and sizes of organizations. It then 'partners' with its customers – providing a spectrum of ongoing support services -- to optimize the customers' individual Internet-usage management programs.

Let's look briefly at Wavecrest products. They're grouped in two 'families.'

1. CyBlock. Particularly germane to this paper is the family known as CyBlock. CyBlock products provide advanced Web filtering capabilities as well as comprehensive Web activity reports. To accommodate different customer requirements, Wavecrest offers three different CyBlock editions:

- CyBlock® ISA – for organizations that have Microsoft ISA Servers.
- CyBlock® Proxy – for organizations that want a standalone software solution that includes its own proxy server.
- CyBlock® Appliance – for organizations that prefer a standalone hardware/software solution.

2. Cyfin. Another family, called Cyfin, consists of products that provide a wide variety of reports that evaluate employees' (or other users') online activity. These products do not include Web filtering capability but are widely used by organizations that prefer to manage Web usage without filtering it. But that's a story for another paper.

The Wavecrest-Customer 'Partnership.' In addition to providing reliable products, Wavecrest supports its customers with a number of long-term support services. Among these are installation and usage support (as needed), training (if needed), product upgrades and URL list updates, documentation support, industry news, and user group forum access. And perhaps most importantly, Wavecrest personnel work one-on-one with individual customers to help them identify (and if necessary block) sites of special local interest.

Note: For more on Wavecrest Computing, visit www.wavecrest.net.

The following sections provide more specific information on the technical and customer service approaches that Wavecrest uses to ensure the success of its customers' Web-use filtering programs.

How Do Wavecrest Filter Products Work?

Before selecting a Web filter, it's important to understand – or at least have a good feel for – how it will work. Oversimplifying a bit, Wavecrest's filter products work as follows. In real time they:

1. Monitor all Web traffic initiated by the workforce.
2. Analyze it several different ways to determine if it violates any policy rules.
3. Block any traffic that violates any of the rules.

Of these tasks, number 2 – determining *what* to block – is the most difficult and challenging.

Let's see how they do it.

Wavecrest personnel work one-on-one with individual customers to help them identify (and if necessary block) sites of special local interest.

Wavecrest's products use RealTimePlus Filtering, a customer-configurable three-layer process.

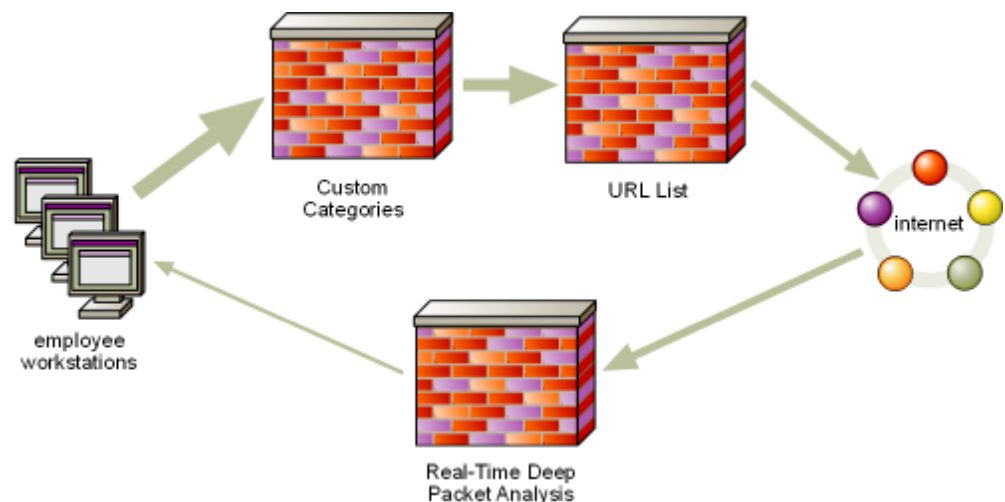
How Do Wavecrest Products Determine *What To Block*?

Wavecrest's overall approach to Web filtering is referred to as *RealTimePlus*. The first part of the name (*RealTime*) refers to its multi-faceted ability to identify and block customer-designated Web sites and data types in real time. The *Plus* refers to a variety of *services* that Wavecrest uses – in partnership with its customers – to continuously improve the reliability and effectiveness of the customers' individual filtering programs.

The following paragraphs provide a brief overview of *RealTimePlus*, its major elements and the steps it employs.

A. *RealTimePlus* Filtering – A Three-Layer Process

As illustrated and explained below, *RealTimePlus* is a customer-configurable three-layer filtering process supported by a collaborative, perpetual, list improvement effort. Configured to support the customer's own filtering policy, it uses three layers of screening. Included are: (1) custom categories, (2) the Wavecrest categorization (control) list and (3) a real-time deep packet analysis process.



Users' requests for visits or downloads must pass through all three layers to be allowed. If they don't, access is denied.

Note. A key component of *RealTimePlus*' overall functionality is its highly effective 'content identification' mechanism. As such, Wavecrest uses it in its Cyfin reporting-only products as well as its CyBlock filtering products.

The remainder of this document is outlined as follows. Section B discusses the three layers themselves. Section C briefly describes how they are used in the overall blocking process. Section D contains supplementary information that we believe is helpful to understanding the subject.

B. The Three Layers

As illustrated above, custom categories constitute the first of the three layers. However, to appreciate the usefulness of custom categories, we believe it is helpful to first understand a bit about the ‘*standard*’ categories found in the Wavecrest URL List, i.e., the second content-identification layer (aka blacklist). Consequently, we discuss it first.

I. The Wavecrest URL List (the “Second Layer”). To accurately identify and categorize the vast majority of Web visits, Wavecrest products use a large, mature categorization control list, often referred to as a blacklist. The list is subdivided into 69 ‘standard’ content-identification categories that Wavecrest populates and updates daily with URLs from around the world. Examples of standard categories include shopping, news, entertainment, sports, social networking, finance, pornography, etc. **Note.** Customers configure the product to specify the categories to be blocked.

Wavecrest developed the list more than ten years ago and has enlarged and improved it continuously ever since. Moreover, the list is used in conjunction with a similarly mature and proven categorization *process*. Wavecrest updates and makes the list available to customers on a daily basis. During initial product/policy setup, customers specify which users or groups of users should be denied access to which categories.

Note: At Wavecrest we are often asked this question, “Given the huge and growing number of sites in the Web, how can your list possibly contain all the sites that might ever be visited by users in your customers’ organizations?” The answer is,

“It doesn’t, and it doesn’t need to. The vast majority of sites are of no interest to our customers’ users. Therefore, for a *given* customer, the list only needs to contain sites that *his* users are likely to visit, and that is only a tiny fraction of the total number of sites in the Web.”

That is the concept on which our products are built. That is, after incorporating *universally* popular sites into the list (by a combination of manual and automated methods), we help individual customers identify other sites that are of particular *local* interest or popularity. We then (a) incorporate these sites into the Wavecrest URL List or (b) help the customers establish special custom categories for them (see next paragraph). Through this two-track collaboration process, we are often able to help customers drive down the percentage of uncategorized sites to below ten percent.

II. Custom Categories (the “First Layer”). Wavecrest products enable customers to create and name ‘custom categories’ to supplement the standard categories discussed above. For example, they can create a custom category to:

- Serve as a “white list” that contains all sites to which visits are allowed (while blocking all others).
- Track and possibly block access to ‘standard’ sites that are not in the Wavecrest URL List but are of special local interest or concern.
- Serve as a “black list” that contains all sites to be blocked (while allowing access to all others).
- Track (but not block) visits to internal servers (intranet sites) and/or partner sites.

Customers can enter as many URLs as they wish into these custom categories. Once established, custom categories function exactly like the Wavecrest ‘control list’ categories discussed earlier. That is, they can be blocked, allowed and/or reported separately.

III. Deep Packet Analysis (the “Third Layer”). Using real-time ‘deep packet analysis,’ the product can determine if the content of a URL is Flash, video streaming, audio streaming, images, or Active X and more. Any or all of these could be considered “inappropriate.” If they so desire, customers can configure the product to block any of them.

Wavecrest personnel help customers identify sites that are of particular local interest or popularity and incorporate these sites into the URL List or help customers establish custom categories.

C. The Blocking Process

Supported by the continuous (and collaborative) list improvement effort discussed above, the blocking process itself works as follows:

1. Customers determine which types of traffic are inappropriate and thus prohibited.
2. Users' URL requests flow through (up to) all three layers.
3. When a request matches any one of the 'prohibited' types of traffic, the CyBlock product automatically blocks access to it.
4. The product responds with a configurable blocking message. Such personalized messages can, for example, point to the organization's own blocking policy.

D. Additional Information

D.1 Special Note re Latency. An in-line technique such as CyBlock's *RealTimePlus* is the most effective way to filter Web use. However, by its very nature, an in-line filter can introduce a certain amount of latency in performance. Wavecrest designed CyBlock with this potential in mind, and when the product is used in conjunction with our minimum recommended requirements for memory allocation and computer speed, the latency should be invisible.

D.2 Use of 'Safe Searching' Feature. To supplement the three layers of filtering discussed above, Wavecrest products can be configured at the server to enable 'safe searching' on Google, Yahoo and Bing. Once enabled, it will block web pages containing explicit sexual content from appearing in search results and prevent users from changing Safe Search controls in their browser.

Summary

Implementation of a successful long-term Web filtering program is a complex management as well as technical endeavor. Several conditions and circumstances make this so. First, the Web is truly immense, it contains many millions of sites, and it's growing extremely fast. Second, the huge variety of content and file types on the Web makes it difficult to identify 'inappropriate' content, and what one organization considers inappropriate, another considers innocent or even essential. Thirdly, countless individuals and organizations are continuously developing techniques for circumventing filters or using Web connectivity to hack into organizational networks.

Despite these obstacles, several competent software firms – with Wavecrest Computing in the forefront – have developed filtering solutions that, while perhaps not 100% perfect, achieve quite acceptable results. And when these solutions are coupled with complementary approaches such as Web activity reporting and carefully implemented AUPs, the results are even better.

Wavecrest's approach to the development of optimum solutions for its customers is to blend well-designed technological innovations with complementary technical support and personalized customer support services.

On the technology side, for example, we augment the original list-based approach to content identification with other innovations such as custom categories, deep packet analysis and 'blocking by file type.'

On the 'softer,' customer support side, Wavecrest is always ready to partner with its customers, providing them with a variety of technical and product services. Of particular note, all customers are invited to participate in a program called OtherWise. As part of this program, Wavecrest personnel work one-on-one with individual customers to help them identify and categorize more and more URLs of important local interest.

The bottom line? The best way for customer organizations to implement a successful and cost-effective Web-use management and filtering program is to partner with the right vendor and work closely and collaboratively with its personnel.

Wavecrest's approach to the development of optimum solutions for its customers is to blend well-designed technological innovations with complementary technical support and personalized customer support services.