

# Content Security Guide

**© Copyright Webdefender AG**

All rights reserved. Reproduction, processing and translation beyond the scope provided for in copyright law is prohibited without written approval.

Release July 2001

**Warranty**

The manufacturer reserves the right to make changes to this publication. Webdefender offers no warranty for the information contained in this document. Webdefender accepts no liability for indirect, direct, ancillary, collateral or other damages relating to the delivery, provision or use of this material. This content security guide is intended solely for information purposes. The information contained in this document represents the topics dealt with from the viewpoint of Webdefender at the time of publication. Because Webdefender must react to changing market requirements, this shall not be construed as any obligation on the part of Webdefender and Webdefender cannot guarantee the correctness of the information set out here after the time of publication.

**Trade marks**

Webdefender and LanShield are registered trade marks of the Webdefender company. Other product or company names used in this document may be protected trade marks of their relevant owners.

# Contents

<b>I. Introduction</b> .....	<b>5</b>
<b>II. Conventional firewalls vs. Content Security</b> .....	<b>8</b>
Definition of Content Security .....	8
Software and hardware systems .....	8
Systems with filter lists vs. lexical analysis.....	9
Areas of application – security and productivity .....	11
<b>III. Mail content filtering</b> .....	<b>15</b>
Security risks during data transmission.....	15
Classic solutions .....	17
Content security solutions .....	17
Summary mail content filtering .....	18
<b>IV. Web Content Filtering</b> .....	<b>19</b>
Security gaps in the Internet.....	19
Classic method .....	20

Content security methods .....	21
Summary web content filtering .....	23
<b>V. Mobile Malicious Code.....</b>	<b>24</b>
The problem of viruses .....	24
Conventional solutions .....	25
Content security solutions .....	26
Summary malicious code filtering.....	27
<b>VI. Summary .....</b>	<b>28</b>
<b>VII. Glossary .....</b>	<b>29</b>
<b>VIII. Contacts .....</b>	<b>29</b>
<b>IX. Impressum .....</b>	<b>37</b>

## **I. Introduction**

### **Expansive development in the use of Internet technology**

E-mail and the Internet have become an indispensable part of our daily business routines. According to IDC, around 320 million PC users worldwide will have access to the Internet by the year 2002 and around 7.9 billion e-mails will be sent via the Internet every day. In Germany alone 4 million Internet addresses are currently registered, with up to 9,000 new addresses added every day. According to the INTERNET SOFTWARE CONSORTIUM this yields a figure of approx. 100 million Internet addresses worldwide, which in turn contain a large number of Internet pages.

### **New dangers**

The explosive growth of Internet communications mean that secure solutions are gaining increasing importance. Virus attacks are becoming more frequent and can cause damage that is quantified in billions. Many users will probably still remember viruses like "I-love-you" and "Kournikova". Many companies are also at risk from within: confidential data such as business reports or development results can fall into the wrong hands by e-mail – whether accidentally or intentionally.

### **Cyberslacking**

Private surfing at the workplace can also cost companies millions. Studies show that several working days per year can quickly be lost in this way. A study commissioned by Sterling Commerce shows that private surfing can result in a loss of about 17 working days per employee per year. You can calculate the figures for your own business by entering the relevant values in the following formula:

Private Internet use (hours per month) x Employee costs (per hour) x Number of users  
= Costs per month

### **Classic solutions**

Until now, it has been attempted to combat these threats mainly by using classic firewalls and anti-virus programs. These methods operate almost exclusively with comparison lists that define what may get through a firewall and what may not. To protect against viruses, anti-virus programs use evaluation of file extensions as well as lists. Files in \*.jpg format are never allowed into the company. This means that a lot of helpful information is also blocked. With millions of Internet addresses and more and more new viruses all the time, it is impossible in practice to keep the lists up-to-date. This is why classic solutions often feature an error quota of up to 80 percent.

### **Content Security**

While in the past the point was still to protect data against unauthorized external users and to keep viruses at bay, the need now is to form an effective barrier that prevents or controls all unauthorized access and forwarding of data either externally or internally, together with the tired-and-tested functions of a firewall. Content security solutions thus check the contents of the data packages sent, not just at the packaging. This new technology can reduce the error quota to two to three percent and can carry out the following tasks for increased security:

- Undesirable e-mails (Spam mail, junk mail, etc.) is kept off the company network
- Protection for confidential data (e.g. development results, profit and loss statements) against unauthorized transmission by e-mail
- Protection against uncontrolled surfing on undesirable web sites (e.g. with pornographic or radical right-wing content)
- Internet access management (regulation of private surfing to specified content and periods)

The content security guide points up the dangers to modern office communication posed through the Internet, analyzes the methodology of classic security technologies in contrast with content security solutions, and evaluates the results.

## II. Conventional firewalls vs. Content Security

### Definition of Content Security

Unlike classic firewall solutions in which the point is to protect networks and data from unauthorized external users, content security solutions focus on the need to create an effective barrier against all unauthorized access and to prevent the transmission of sensitive data externally. In technical terms, the techniques used in classic firewalls and content security solutions differ in the fact that in the classic firewall only the header information in a transmission (like the address information on an envelope) is evaluated. In contrast, the content of transmissions is checked in content security solutions (i.e. the envelope is opened and the letter is read).

### Software and hardware systems

Depending on the provider, content security solutions operate on the basis of software or hardware. They are installed either on the computer at the employee's desk or on a so-called proxy server.

- Software-based solutions should generally be adapted to the existing IT system (operating system, browser settings, etc.) In addition to the higher configuration and administration effort involved, gaps in security can arise due to implementation errors – however small.
- All software contains errors, so-called bugs. Errors are detected over and over again as a result of continuous "updating" and these can be used specifically as alternative channels for attacks.



- Dangerous code, which can be concealed in Java applets or ActiveX applications, does not stop at the software of the content security solutions or its environment and can threaten the internal network.

In contrast, hardware-based solutions offer the following advantages:

- Easy installation, configuration and administration of the system, irrespective of the operating system, browser settings, etc.
- Examination of the complete data traffic (e-mails; www) through the central installation of the system (gateway) at the interface between the internal network and the Internet
- Exclusion of security gaps in the system environment or in the software
- Secure analysis of unknown application programs with possible viruses (codes) in a hardware-based environment (sandbox)

### **Systems with filter lists vs. lexical analysis**

Depending on the method of analysis used in the content security solution, the content is examined on the basis of individual keywords or by means of a lexical analysis. For the most part, content checking in previous content security solutions depends on the simple cataloging of undesirable web pages (URL addresses or IP addresses) and the blocking of certain file extensions or e-mail senders. Cataloging prevents the requested transmission from being called up by suppressing the downloading of the required content. The actual content of the requested web page is not checked. This gives rise to a wide range of problems. Because of the immense growth of the Internet, it is not

possible to keep an up-to-date "blacklist" for various areas (e.g. pornography, drugs, right-wing radicalism, etc.) Many providers have therefore changed over to simply including popular sites on their lists and to checking these on a continuous basis.

In addition, more and more web sites operate with dynamic content, i.e. the content is administered in a database and is attuned to the relevant user. Automatic search mechanisms cannot react here, so that undesirable web sites are not detected and cannot be added to the "blacklist". Unauthorized web sites are identified on the basis of "blacklists" through a comparison of the entered URL or IP address. This blocks the entire web site (in some cases for more than one user) – even if only parts of the content fall into the undesirable category. Useful information sources, such as stern.de or yahoo.com are completely blocked because of "prohibited" subpages.

In order to ensure a procedure for checking the content of data that is largely independent of the growth of the Internet, content security solutions such as LanShield analyze the data traffic (eMails, WWW) between an internal network and the Internet on a binary level and compare these with the filter categories on the basis of logical textual contexts and algorithms. This means that there is no need to create or update "blacklists". Likewise, file attachments are also analyzed on binary level and compared with the filter categories. The file extension, which may be falsified, has no bearing on the check. Thus, web site content is checked on the basis of the user settings. This means that not only dynamic data is found. In addition, only web sites with unauthorized content are blocked. Useful web addresses are preserved – only unauthorized areas are blocked. So-called policies are drawn up for users and user groups under Internet access management and these determine which user groups are permitted to view, send or download which content. This makes it possible to optimize Internet usage and increases productivity in relation to the Internet.

Providers of content security are faced with a particular problem when it comes to the checking of eMail content. Often it is not the text of the eMail that represents the sensitive security factor, but rather the file attachment. Previous solutions mostly check the attachment on the basis of the file ending (e.g. .mp3, .exe). This criterion can be by-passed by simply renaming the file ending, so that confidential, sensitive or undesirable file attachments are not identified and filtered out. This danger can also be avoided by performing binary analysis on the content of both the eMail and its attachment.

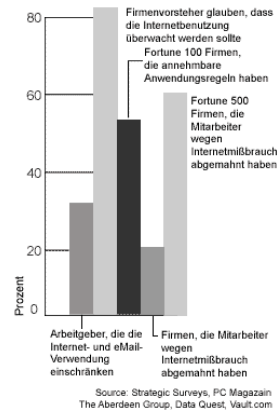
### **Areas of application – security and productivity**

Content security solutions are used wherever data is transferred between the internal network and the Internet. They complement classic firewalls and close existing gaps in security:

- Keeping undesirable eMails (Spam mail, Junk mail, etc.) off the company network
- Protection of confidential information (e.g. development results, profit and loss accounts) against unauthorized transmission by eMail
- Prevention of uncontrolled surfing at the workplace (cyberslacking)
- Internet Access Management: Regulation of the handling of the Internet through the definition of policies for each employee or group

Studies already indicate the major demand for content security solutions:

Wie Arbeitgeber darauf reagieren



The term content security focuses particularly on three areas: the scanning of eMail and Web content and the analysis of mobile malicious code (MMC).

Dangerous mobile malicious code (viruses, worms and Trojans) is mainly hidden in useful applications such as Java applets, ActiveX and executables which are transmitted as eMail attachments and Web applications. Mobile code can cause a high level of damage.

The assignment of rules represents a key aspect of the content security approach. This means that options are defined depending on persons or groups. Thus, for example, the finance manager might be allowed to access the stock market pages because this is part of his work, while IT employees would not be permitted to access the same stock market pages. Also, information sent by eMail is analyzed by the content security solution in terms of the sender.

In other words, content security solutions scan all incoming and outgoing data traffic. This means that they can be usefully employed wherever eMail and Internet are used. They can be used particularly effectively in the research sector to prevent the unauthorized disclosure of important research results, development data and patents. However, content security solutions also play a role in the areas of human resources and controlling because they can be used to prevent confidential employee data, applications or financial data from being transmitted and to combat the manipulation of accounting data. Content security solutions are also useful in the record industry, where they can prevent the distribution of illegal MP3 files.

In all areas of business, content security solutions can be used to prevent the transmission of eMails with defamatory content and to keep mass mails and junk mail at bay. Likewise, content security solutions provide a mechanism that prevents uncontrolled surfing in the workplace. Surfing does not have to be completely prohibited, but can be regulated. For example, access can be permitted to certain topics during lunch hour or after regular working hours. Other pages, for example with pornographic or violent content, can be permanently blocked. This also yields interesting applications for protecting young people, for example through use in schools. Pages with content that could be damaging to children, for example pornography, right-wing radicalism or criminal content, can be blocked, keeping them well away from school children..

However, content security solutions are being used with increasing frequency in the private sector. Here, for example, visits to web sites by surfing children can be regulated and a secure protection against viruses, worms and Trojans can be put in place.

### **III. Mail content filtering**

#### **Security risks during data transmission**

As a rule, general access is permitted to eMail within companies and eMail has now established itself as the most commonly used method for file transfer in the Internet. This opens up a whole new range of communication and advertising options, which are, however, also open to abuse. There is nothing easier than sending a hidden attachment by eMail. In addition, important data is left lying around open and untended in the network. Fraud, passing on of customer and order data, manipulation of accounting data or disclosure of development and research results to the competition have thus become a genuine danger that cannot be quelled by means of firewalls.

Attachments can contain mobile malicious code (viruses, worms and Trojans) which can cause a lot of damage in a company's network. Junk mail or Spam mail which is sent unsolicited from every corner of the world, as well as the sending of enormous data packages, such as \*.mp3 files, tie up line resources and computer capacity, in turn slowing down and impeding internal processes.

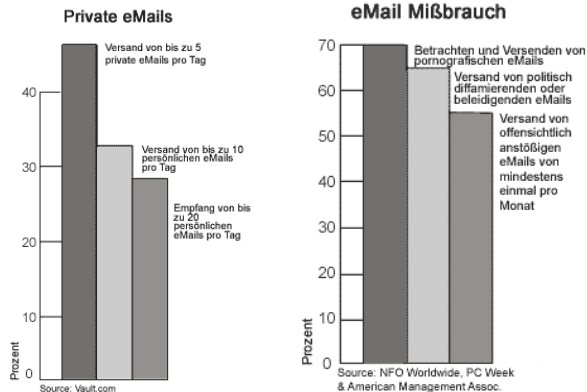
In addition, many companies are unaware that they can be held liable for the e-mails sent by employees. It is not just in the current telecommunications laws in Germany and the last paper drawn up by the European Committee against cybercrime dated 19 November 2000 (which was changed in relation to new media in its latest version) that marked contradictions arise.

Conventional forms of communication are still being compared with the new media, creating the basis for a dubious flood of legal cases due to ambiguous and unsuitable interpretations of statutes. Judgements have already been reached in other European countries in which companies have been held liable because their employees sent defamatory e-mails over the company's mail

server. In this case, the court classified these e-mails as publications, identifying the employee as the author and the company that supplied the mail service as the publisher.

This is why it should be assumed that the ambiguous legal situation will lead to surprising legal cases. It is obvious that the rapid growth of the Internet worldwide will lead to similar, if not identical, problems in other parts of the world.

Studies point to the growth in private use of company networks at the workplace:





## **Classic solutions**

Classic firewalls only check eMails in terms of sender addresses and the file endings of the attachments. Only eMails whose senders or protocols (e.g. "smtp") have the "green light" can get past the firewall. When the file endings are checked, certain formats, such as \*.exe or \*.mp3, are generally blocked. This block means that even some useful files are not allowed past the firewall. In general, classic solutions do not involve a content check. In addition to blocking incoming desirable content, this can also mean that eMails coming from your own network can generally be transmitted unless a file ending in the attachment prevents this. However, confidential information contained in the eMail itself can be transmitted to the outside without hindrance.

## **Content security solutions**

Many mail content filtering products are based on "proxy" technology, in other words, all transmitted and received eMails are forced to detour to a "scanner". The eMails are evaluated and forwarded or placed in quarantine. This method involves a number of major disadvantages. On the one hand, no web-based applications are recorded by these scanners, while on the other the "forwarding rules" on the client can be manipulated easily. Another key disadvantage is the effort involved in implementation because every connected client must be reconfigured.

Other content security solutions are capable, for example, of using lexical analysis to keep e-mails with particular content (e.g. right-wing radicalism, libelous content, sexist material, etc.) at bay or to block content relating to a company. All associated raw data packages are intercepted in the datastream, temporarily stored in the cache and then scanned, along with any attachments, for critical content. After this, the system-inherent "event handler" decides whether the eMail is to be

forwarded or blocked and whether the system administrator needs to be notified. In addition, standard disclaimers can be attached to all outgoing e-mails, such as "the author bears sole responsibility for the content of this e-mail; the content and opinion expressed do not necessarily reflect the views of the XY company." If a company uses software of this kind, thereby attempting to prevent misuse, then it is always free from such liability.

### **Summary Mail Content Filtering**

Only those eMails whose content has actually been checked are in fact secure communication routes. Content security solutions ensure that eMails do not leave or enter a company without being checked first. The possibility of scanning attachments means that simply renaming file endings is ineffective because content security solutions check the content of the attachment at binary level. The option for assigning rules means that companies can steer their eMail traffic in secure routes. Sensitive data remains inside the company. Only certain persons or groups can send individual files, depending on their area of responsibility.

Content security solutions reliably recognize Spam and junk mail. Unknown code which might contain mobile malicious code is first executed in a quarantine area. If irregularities occur in the quarantine area, then transmission is blocked. This means that mobile malicious code can no longer get into the company.

Hardware-based content security solutions which scan mails on the basis of a lexical analysis have the lowest error quota, cannot be by-passed by means of simple manipulations and are easy to install.

## IV. Web Content Filtering

### Security gaps in the Internet

A study commissioned by Sterling Commerce indicates that about DM 104 billion is lost by German business every year due to private surfing in the workplace, also known as cyberslacking. On average, employees with Internet access spend about 3.2 hours per week online for private purposes. With average labor costs of DM 49.23 per hours and 16.2 million jobs with Internet access in Germany, this yields costs of around DM 104 billion, not counting network charges. In addition, an average of 17 working days per year are lost per employee. A study conducted by the University of the Saarland indicates that pages with pornographic content (now accounting for about 20% of the over 150 million web sites) are particularly popular. However, it is not just erotic material that tempts hobby surfers - sport and leisure topics are also very popular.

The now commonly used term cookie is generally used to refer to usually harmless commercial Internet packages. By setting a cookie consisting of a simple text file and no longer than 4 kb, user behavior can be influenced in an indirect way, for example a different view will be displayed the next time a user visits a particular Internet site. The reading of a cookie, which can only be carried out by the web server that actually set it, cannot cause an application to be started. This means that a cookie does not contain malicious code, even as a virus. This means that the question of whether cookies are damaging is purely subjective, depending on whether it bothers you to be analyzed, evaluated and influenced by modern techniques. These cookies are referred to as server-side cookies

Unlike the previous types of cookies mentioned, there are also client-side cookies which, unlike their cousins, contain actual script code which is executed on the client side and, for example, can

read address books and e-mail accounts and transmit these without detection. Here too a check of the actual content can provide a remedy.

However, cyberwoozles represent a very particular kind of danger. Cyberwoozling is a process in which unauthorized parties tap complete files from the computer during the time they spend on the Internet, possibly also manipulating these files. Thus, for example, complete address databases, research results, personnel files or other important company-internal data can be tapped by the competition at any time. The term "cyberwoozling" goes back to the figure from English legend known as the Woozle, who was reputed to nibble the hairs on people's legs while they were sleeping..

### **Classic method**

A classic firewall consists of hardware or software that monitors all network traffic from the Internet to the internal network and that only allows "good" data packages to pass. The list-based filters used to extend firewalls produce a very similar error. Because each firewall identifies the communication partner on the basis of IP addresses, Internet addresses are automatically converted into IP addresses ([www.webdefender.de](http://www.webdefender.de) = 212.184.109.193). However, an IP address generally fronts not just one, but up to 1,000 home pages with a wide variety of content. A single Internet page, for example with pornographic content, could cause all the other home pages to be unjustly placed on the index, thereby blocking them. The best-known example of this is the erotic section of STERN magazine which caused the STERN home page to be put on the index, preventing access to all pages.

Current studies indicate that the error quota of classic products is up to 80 percent. This lack of accuracy disqualifies all commercial use and therefore makes no meaningful contribution to a

reduction in costs. License prices, which are usually user-dependent, bear little relation to the benefits that can be hoped to be achieved.

### **Content security methods**

This frequently used procedure is an attempt to control access to critical content in the Internet by means of a simple cataloging of undesirable content providers (URL addresses or IP addresses). The user is prevented from calling the actual transmission because the downloading of the unwanted content is prevented. The actual content of the web pages is not checked.

Web content filtering with so called "blacklists" only contains a fraction of the actual unwanted web sites. In addition, the maintenance of these lists represents an insoluble problem because no organization is capable of viewing and evaluating 150 million web sites per day. Thus, list-based content filters are subject to an error quota of up to 80% - this means that commercial use can only be justified by the viewpoint "at least we're doing something".

The lists are generated using crawler-like search engines which search the Internet for possible critical content on a day-by-day basis. These crawlers operate very similarly to conventional search engines. It's not surprising therefore to learn that this scheme produces a wide range of results that have nothing to do with what is being searched for. A further manual check cannot be made for reasons of time constraints alone. For this reason, addresses are also added that are obviously not worth blocking. How else could one explain the fact that the Vatican web site ended up on the index because the Pope published a report into abortion. The American organization Peacefire offers detailed information on an almost daily basis, reporting not only the faults in filter software, but also providing a program for bypassing the filter software of some well-known manufacturers.

The detection of critical content by means of crawlers takes place automatically by means of linked references to Internet pages. This works until this crawler encounters a dead-end (dead link) or this link refers to the content of a database. This means that Internet pages that are dynamically generated and individually created remain hidden from these crawlers. This would be like looking for an article in a magazine archive which was also located in a no-access area. Producers of conventional filter software therefore react by completely blocking all the pages located behind this address, thereby also blocking web material with no defamatory content

However, this could quickly become a full-time job for administrators, also tying up computer capacity, particularly because "erotic" web sites frequently change their links (URLs). Even if several thousand pages were to be defined as inaccessible, two problems would still remain:

- Constant database lookups would impede access to the network.
- Access to an "anonymous" web site that would act like a proxy, concealing the true destination address from the firewall would be an easy way to bypass this restriction.

This is why it is only possible to provide comprehensive security control by means of a lexical analysis of the content. Unlike classic firewalls, content security solutions carry out this analysis. Content security solutions are not intended generally to suppress private surfing in the workplace, but rather to provide management with a tool for ensuring that the Internet is used effectively and productively in the company. Pages with pornographic or radical right-wing content can generally be blocked, for example. Other areas, such as investment or travel, can be released at certain times. This means that employees can check stock market prices or book their next vacation during their lunch hour or after regular business hours.

In addition, the use of content security solutions means that certain areas can be released for particular persons or groups of persons. For example, the finance manager can access the Investment area any time because this is part of his work, while all other employees can only access this area at certain approved times. This option for assigning rules provides companies with a useful tool for regulating Internet traffic in a rational way and for avoiding the blocking of valuable sites.

### **Summary Web Content Filtering**

Content security solutions provide an opportunity for effective Internet access management. They protect against uncontrolled surfing at the workplace and unwanted cookies. The fact that rules can be assigned means that supervisors have a mechanism that enables them to regulate the use of the Internet and to protect the network from dangers from the Internet.

## **V. Mobile Malicious Code**

### **The problem of viruses**

The fact that browsers and mail programs are becoming increasingly easy to use and even more sophisticated in their functions means that any user can easily load damaging content – referred to as "mobile malicious code" – from remote Internet pages or as an e-mail attachment onto his own computer or into the company's network without being aware of the extent of the potential damage that could be caused.

Mobile malicious code is mostly transported with applications like Visual Basic scripts, Java applets or ActiveX controls. Paradoxically it is precisely these file types that enable technologies to be used which also facilitate useful functions, such as Internet banking, animated and interactive Web presentations or shopping functions. From a business viewpoint, the restriction of these technologies leads to a drastic reduction in the scope of functions of the new media, impacting considerably on the profitable business processes of Internet and e-mail.

The Java programming language, which was developed by SUN Microsystems, contains applets which were specially developed for the Internet and which can be executed as part of an Internet page. These pages are usually started unnoticed when an Internet page is called. Although the security precautions in the current version have been radically improved, nonetheless, JAVA still contains functions that can cause computers to crash if misused and which can provide hackers with an opportunity for IP spoofing.

JAVA applets have become well-known through Internet banking and entail undoubted advantages due to their non-platform-dependence and special display options in the Internet. Microsoft has



developed an add-on for JAVA in the shape of ActiveX, which is also executed by calling an Internet site. Among other things, Microsoft's ActiveX enables interactive applications to be implemented in the Internet.

However, much greater security problems arise because of the close link with the Microsoft operating system and with the Internet programs from Microsoft. Thus, although Microsoft's Internet Explorer offers a wide variety of settings for limiting ActiveX Controls, a layperson is hopelessly over-burdened due to the complexity of the system. The very close integration of Internet with the Windows operating system therefore entails a major area of attack which has already led to enormous damage, up to and including complete system failures. As already explained in detail, the development of new functions has always led to new security problems.

Such precautions in companies are still a rarity. People prefer to down-play latent worries about security with the argument that the precautions are unproductive and expensive, acting out of ignorance about which measures are to be recommended.

### **Conventional solutions**

Conventional firewalls and virus scanners do not offer adequate protection here because they cannot automatically recognize damaging content as such and allow it to get past. The protection provided by firewalls and anti-virus programs to date is mostly dependent on how up-to-date the virus comparison list is. Until the anti-virus program is updated, the occurrence of a new virus always poses a disproportionately high residual risk.

Until now, no central solution has been available which carries out a classification of data as a transparent firewall, blocking damaging content before it can wreak havoc in the network. In order

to reduce the potential risk, network administrators react by excluding certain file types, such as \*.vbs, Java applets or ActiveX controls by means of the firewall. However, this also means that useful functions such as Internet banking and various web presentations are also excluded.

### **Content security solutions**

This makes it necessary to find a solution that can monitor all network traffic with the Internet on the basis of content as a firewall extension. This means examining all requested web pages, e-mails and other packages, including all attachments, in order to reduce them to the raw form of the data and to test their content.

The analysis of mobile code on the basis of content facilitates for the first time a contemporary protection against even unknown script code. Independently of the updating of comparison lists, all Internet content is already scanned at the gateway for damaging content in http and eMail connections and forwarding is prevented as necessary. New and unknown kinds of code hidden in web sites, attachments or downloads are detected, exported to a quarantine area and their effects are tested. Depending on the action triggered within this quarantine area, the content security solution decides whether the called web site, file or eMail should be forwarded or discarded. Useful code is allowed to pass through the content security firewall, while damaging code is blocked.

To avoid the consequences of mobile malicious code, hardware-based content security solutions of the quarantine area itself exist in the form of an independent flash memory. This means that content security solutions do not allow malicious code to spread to conventional mass storage media, thereby causing permanent damage.

## **Summary Malicious Code Filtering**

Only content security solutions provide reliable protection against malicious code. An actual content check ensures that no undesirable content or damaging application leave or enter the company. Classic firewalls cannot provide this function because they only check for file endings and therefore many useful applications and files are also blocked.

## VI. Summary

The explosive growth in the use of Internet technology worldwide brings with it many benefits and dangers. In addition to virus attacks, cyberslacking and data theft also threaten companies and cause damage that is measured in billions. New technologies are needed in order to offer companies effective protection against these new dangers. Classic security measures such as firewalls and anti-virus programs can only prevent these dangers to a limited extent because the filter mechanisms are not enough to react to new technologies and dangers. On the other hand, error quotas of up to 80 percent means that these security mechanisms cannot be considered for commercial use.

In contrast, content security solutions can react effectively because they examine the content of the data packages sent via the Internet. Hardware-based content security solutions are particularly effective because they are both easy to install and provide good protection against attack. The method of content checking enables hidden files, such as malicious code, to be analyzed. Effective content security solutions check the content of data packages according to logical textual relations, not just on the basis of keywords or blacklists. This makes it possible to guarantee that the highest possible level of security is achieved and that the productivity of the network is significantly enhanced. Useful applications are no longer blocked. Access to Internet sites which are visited out of private interest is prevented. Confidential company data can no longer reach the outside world.

## **VII. Glossary**

### **ActiveX**

This is a technology developed by Microsoft. It is purely an MS Windows application, i.e. ActiveX controls are not executed on other platforms. Among other things, it is used to create executable programs on web sites. ActiveX permits far more intervention on a Windows system than would be possible using script languages.

This technology, like Java, makes interactive operations such as eCommerce or eBanking possible on the Internet. Older, mostly CGI-based techniques are not capable of offering the same level of security for transactions of this kind.

### **Algorithm**

A methodical, repetitive procedure which operates according to a certain scheme, e.g. the find/replace function in a word processing program.

### **Anti-virus programs**

Also referred to as virus scanners, anti-virus programs are installed on the desktop PC or on the server. They filter viruses on the basis of comparative lists.

### **Content security**

Unlike classic firewall solutions which concentrate on protecting networks and data from unauthorized external users, content security solutions focus on a content-based examination of Internet communication. From a technical viewpoint, these two technologies differ in the fact that only the "header information" of a transmission is evaluated in a classic firewall (similar to the information on an envelope). In content security solutions however, the content of the transmission is checked (in other words the envelope is opened and the letter is read).

Content security solutions are used wherever data is exchanged between the internal network (LAN) and the Internet. Every time the Internet is accessed, the data is first checked by the content security solution and then passed on or, if necessary, blocked. This takes place on a hardware or software basis, depending on the provider.

### **Cookies**

These are short texts that can be stored and read by the information provider on the user's hard disk. They were developed by Netscape in JavaScript as meta-tags. Evaluation takes place via CGI, Java, etc. These are (still) essential for interactive use, e.g. in eCommerce the shopping basket is stored in cookies until an order is made. However, cookies are also used to "spy" on the user's online behavior.

### **Cybercrime**

When used generally, this term covers all criminal acts committed by means of computer networks such as the Internet. The problem here is that different legal standards apply worldwide. What is permitted in one country (e.g. the distribution of radical right-wing literature) may be illegal in other countries. However the Internet recognizes no boundaries. It is difficult to prosecute the criminals as long as there are no minimum international standards.

### **Cyberslacking**

Private, unwanted surfing at the workplace. According to a study commissioned by Sterling Commerce, German business loses about DM 104 billion per year through private surfing at the workplace. On average, employees with Internet access spend about 3.2 hours per week online for private purposes. With average labor costs of DM 49.23 per hour and 16.2 million workplaces with Internet access in Germany, costs of about DM 104 billion are yielded, not counting network

charges. Private surfing can result in a loss of an average of 17 working days per employee per year.

### **Cyberwoozles**

This refers to the practice of drawing data from the user's PC while he is surfing the Internet. This technique is used to find out eMail addresses, passwords, visited Internet sites, etc.

### **\*.exe (executables)**

Files that run independently on commonly used operating systems as programs. They are referred to as compiled program data if they need a runtime processor to run. These files can contain viruses and are dangerous if the recipient opens them without knowing their content.

### **Filter categories**

Categories in the content security solution on the basis of which data traffic between the LAN and the Internet is scanned, e.g. pornography, investment, sport, travel etc.

### **Firewall**

This is the name for software that runs on the connecting systems between the Internet and Intranet.

It is used to prevent external access to IP addresses in the Intranet. It protects internal data. If configured appropriately, it can be used to exclude URLs with unwanted content from being called. The information about the source and target address contained in a package is used by the firewall to decide whether it can be allowed to pass or whether it should be rejected.

### **http (Hypertext Transfer Protocol)**

A protocol which is used to transmit HTML objects and other objects to Intranet browsers.

### **Internet access management**

This regulates access to the Internet. This does not completely rule out access through the assignment of rules, but permits it during the lunch hours or after business hours. Sites with pornographic content, for example, can be permanently blocked. In addition, rules can be set up on the basis of user groups.

### **IP address**

From IP Version 4 onwards this is a number consisting of 4 octets ( $4 * 8$  bits, or quads) which is required in order to identify an Internet user (or node) uniquely. Internet servers are assigned a fixed IP address; they can be found using both the IP address and the "meaningful" name (URL) managed by the DNS. For simplicity, the 32-bit entry is written in 4 decimal values, e.g. 127.215.205.156 instead of 01111111.11010111.11001101.10011100.

### **Java applet**

This is the name given to a program created in Java.

The program was translated into byte code by the Java compiler and this code is then executed by the Java virtual machine.

### **LAN (Local Area Network)**

This is a network housed on a site or in a building that does not use any public lines.

### **Lexical analysis**

The analysis of http, smtp, POP3 and attachments at binary level in the quarantine area



### **Mobile Malicious Code**

Malicious mobile code consists of applications that can be transmitted through the downloading of files from the Internet or through eMails. They contain executable web content (active content) such as ActiveX controls, Java applets or Java script and can damage or destroy both the operating system as well as application programs and data structures. Desktop firewalls that use sandbox technology form a security environment, the sand box, around the web browser. Applications and active content from the Internet can run within these walls without needing to access resources outside of the sandbox.

### **POP3**

A protocol used to arrange stored eMails. A POP3 server accepts SMTP eMails on behalf of a user. The user's client uses SMTP in order to send an eMail and POP3 calls up the eMail from the server.

### **Proxy server**

This is a software that is installed on a system at the interface to the Internet and Intranet or on the Internet Service Provider (ISP) system.

Among other things, it serves to temporarily store called files in a cache so that these are available faster the next time they are called.

### **Public Key Infrastructure (PKI)**

The controlling of encryption by a hierarchy of "authorities" which administers the distribution and validation of public codes.

**Quarantine area**

An area in which eMails or http files that may contain dangerous content, can be isolated from users and analyzed by the content security solution.

**Sandbox**

An environment that allows no access to operating system routines.

**SMTP (Simple Mail Transfer Protocol)**

A protocol used for e-mail transmissions on the Internet.

**Trojan horse**

Trojan for short

These are damaging programs that use a harmless and often useful host program to make their way onto a system. This means that damaging routines can make their way onto the computer, enabling passwords to be cracked, for example.

**User**

Anyone who uses a program, software or application, including programs and media that enable access to the Internet or to a mailbox.

**Viruses**

These are software items that have the potential to cause damage and that proliferate, unknown to the user, by being attached to program or document files (macro viruses).

When the program is called up or when the file is opened, the virus is activated and infects other current programs or open documents. Internet technology, e.g. Email or downloading, make such

distribution very easy. The damage caused by these programs is now estimated at several billion DM per year worldwide.

### **Worms**

These are a type of virus that proliferate automatically by means of network connections.

They do not affect other files. The damage caused by this type of virus consists in the overloading of servers and routers, usually leading to a crash.

Macro viruses have recently become known which use the address book under Windows to send themselves to other users by means of attachments in Spam and junk mail, causing eMail servers to crash under the burden.

The first virus on the Internet was a worm virus. On 2 November 1988, this disabled about 10 percent of the 60,000 routers and servers then in existence for one day.

## **VIII. Contacts**

### **Europe**

Webdefender AG  
Am Alten Viehmarkt  
84028 Landshut  
Germany  
Phone: +49.871.94244.0  
Fax: +49.871.94244.11  
email: [info@webdefender.com](mailto:info@webdefender.com)

### **Asia / Pacific**

Webdefender Asia/Pacific Co.Ltd.  
18/F One International Finance Centre  
1 Harbour View Street  
Central, Hong Kong  
Phone. +852.2.1668.243  
Fax. +852.2.1668.999  
email: [pohlueke@webdefender.com](mailto:pohlueke@webdefender.com)

### **USA**

Webdefender Inc.  
Pruneyard Center  
1999.S.Bascom.Ave.Suite.700  
Campbell.CA.95008  
United States of America  
Phone.+1.408.879.2343  
Fax.+1.408.879.2635  
email: [menn@webdefender.com](mailto:menn@webdefender.com)

## IX. Impressum

*Editorial team:*

Webdefender AG

Am Alten Viehmarkt 3

84028 Landshut

Germany

Phone: +49.871.94244.0

[info@webdefender.com](mailto:info@webdefender.com)

[www.webdefender.com](http://www.webdefender.com)

# GUIDE

© Copyright reserved Webdefender AG, Landshut, Germany, 2001

**Webdefender AG**  
Am Alten Viehmarkt 3  
D-84028 Landshut - Germany  
Telefon +49 871 942440  
info@webdefender.de  
www.webdefender.de

**WEBDEFENDER**  
READY FOR CONTENT SECURITY