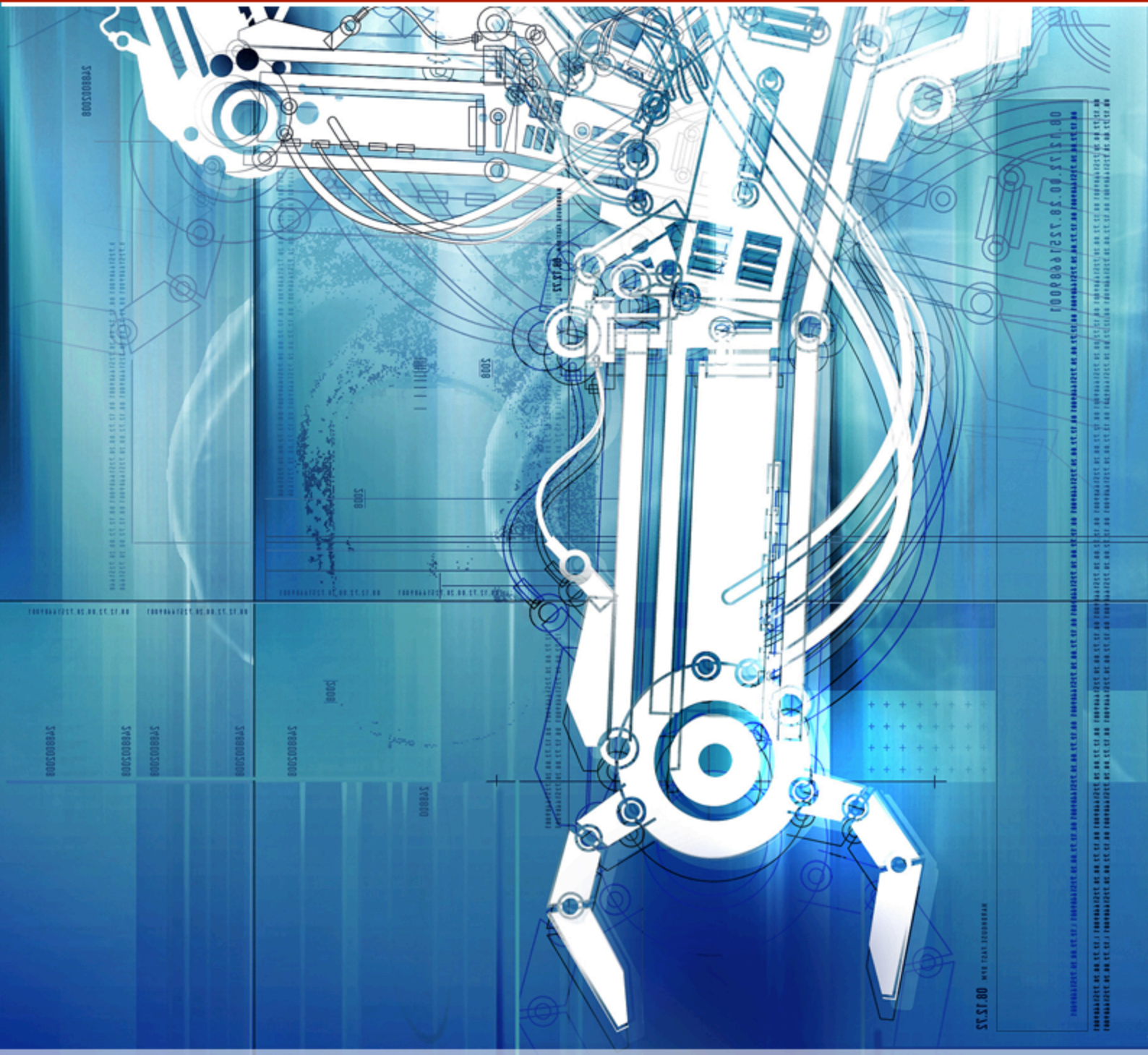
[IN]SECURE

OPEN. INFORMATIVE. TO THE POINT. Issue 3 - August 2005



SECURITY VULNERABILITIES, EXPLOITS AND PATCHES
by Dr. Gerhard Eschelbeck, Qualys CTO

PDA ATTACKS: PALM SIZED DEVICES - PC SIZED THREATS
by Seth Fogie, Airscanner VP

12 MONTHS OF PROGRESS FOR THE MICROSOFT SECURITY RESPONSE CENTRE
by Stephen Toulouse, Security Program Manager of the MSRC

TABLE OF CONTENTS

A lot has happened in the security world in the past 2 months since the last issue of (IN)SECURE was released. It's always a pleasure to hear from the readers and we're happy to report that interest has been very high with many news and comments rolling in. Keep it up!

This time we have some well-known authors writing on very different topics. We're sure everyone will find something they find to be interesting. To top it all, we're running a book contest whose winner will be the people that send in the most interesting suggestions related to (IN)SECURE. Point your browser to www.insecuremag.com/contest and be creative!

The editorial team:
Mirko Zorz
Berislav Kucan

Visit the magazine website at www.insecuremag.com

(IN)SECURE Magazine contacts

Feedback and contributions: editors@insecuremag.com

Advertising and marketing: marketing@insecuremag.com

Distribution

Combat The Rise In Web Attacks With Acunetix Web Vulnerability Scanner 2

Acunetix released Acunetix Web Vulnerability Scanner: a tool to automatically audit website security. Acunetix WVS 2 crawls an entire website, launches popular web attacks (SQL Injection, Cross Site scripting etc.) and identifies vulnerabilities that need to be fixed.

Acunetix WVS is available as an enterprise or as a consultant version. A subscription based license can be purchased for as little as $395, whereas a perpetual license starts at $2995. For more information visit www.acunetix.com

SmoothWall Launches School Guardian

SmoothWall has launched School Guardian, an integrated Internet gateway firewall and web content filter designed specifically for educational establishments. School Guardian addresses the three major challenges faced by school and college networks - blocking unsuitable web content, controlling access to Internet services and preventing attackers from compromising private systems and critical data. To learn more visit www.smoothwall.net.

Criston Releases Precision 5.2

Criston, the leading European provider of systems, patch and vulnerability management announces that Criston Precision 5.2, an integrated software suite dedicated to systems and security management on a company-wide scale, is now available. Precision 5.2 brings a revolutionary autonomic approach to customer station management. Based on intelligent agent technology, the solution provides users with a tool that is proven for carrying out the following applications: automatic IT assets inventory and management, large-scale software distribution, remote administration, security patch management, OS deployment and migration, mobile device management, self-healing and supervision for systems, etc. More information about Criston and their products can be found at www.criston.com

CyberGuard SG580 Appliance With Unified Threat Management Functionality

CyberGuard Corporation announced a new all-in-one, centrally managed desktop security appliance to protect small and mid-sized enterprises against external and internal threats to their network. A robust network security solution that unifies layers of defense and response mechanisms, the CyberGuard SG580 provides Unified Threat Management functionality, including a powerful stateful-inspection firewall, service-based intrusion detection blocking, Anti-Virus protection, and threat containment through Security Policy Enforcement. For more information visit www.cyberguard.com.

Trend Micro Offers Three New Anti-Spyware Solutions

Trend Micro, Inc. announced three new solutions that detect and remove evasive spyware. Trend Micro Anti-Spyware represents the first significant offering resulting from the company's May 2005 acquisition of InterMute, Inc., a leading developer of anti-spyware products, and offers immediate protection against some of today's most insidious spyware programs.

New products include Trend Micro Anti-Spyware 3.0, Trend Micro Anti-Spyware Enterprise Edition 3.0 and Trend Micro Anti-Spyware for Small and Medium Businesses 3.0. In North America, the suggested retail price for Trend Micro Anti-Spyware 3.0 is $29.95 with renewal pricing at $14.95. The price for Trend Micro Anti-Spyware for Small and Medium Businesses 3.0 ranges from $20/seat at the 5-25 user level, to $11/seat at the 501-1000 user level. For more information visit www.trendmicro.com.

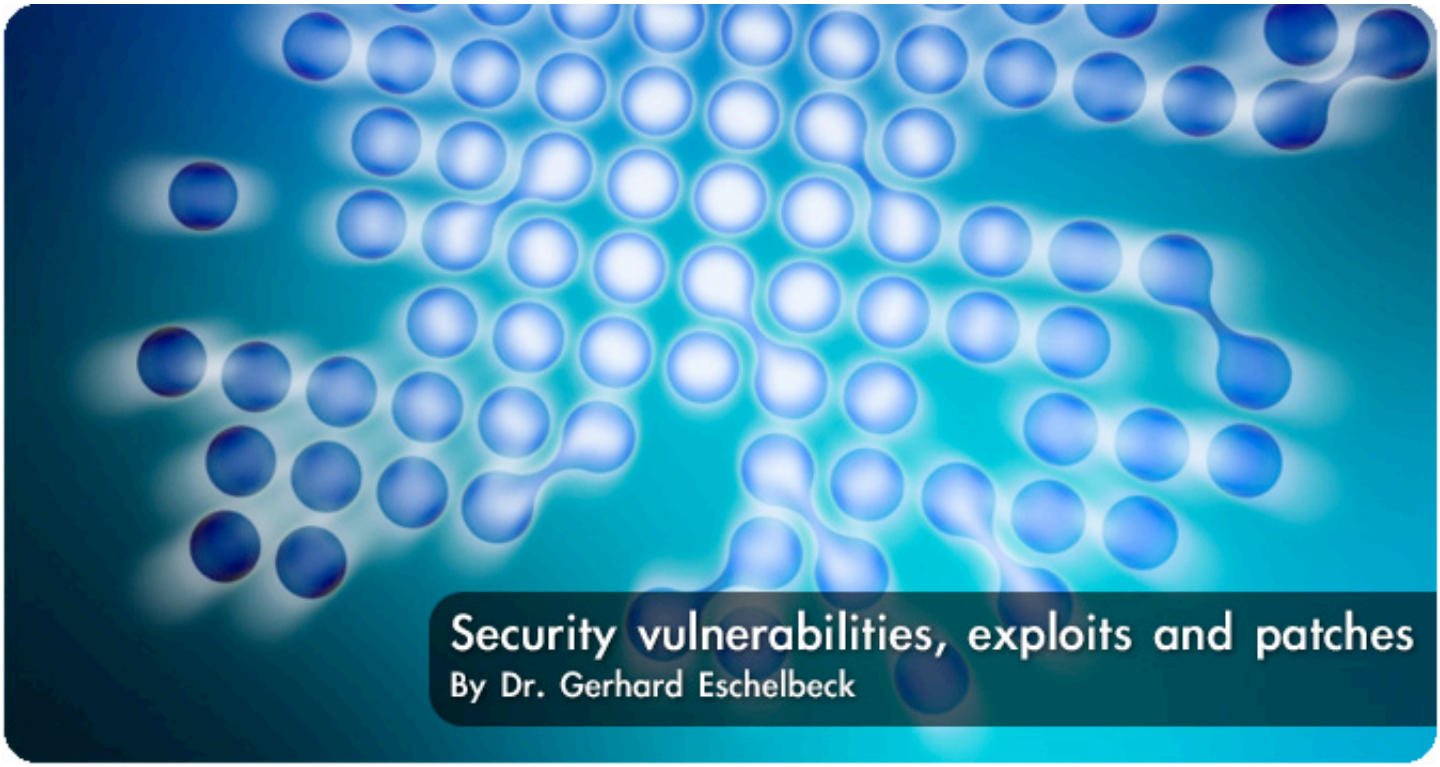
BeCrypt Launches Disk Protect 3.0

BeCrypt, the leading UK encryption security company, has unveiled version 3.0 of its DISK Protect full disk encryption security solution for laptop and desktop computers. DISK Protect 3.0 offers features specifically tailored to the needs of business users, including Single Sign-On, secure hibernation, removable media encryption and extended smart card support.

DISK Protect is easy to install using standard network deployment tools. Once the user has entered his or her DISK Protect password and logged in to Windows, encryption is transparent. Everything written to the hard disk is automatically encrypted and everything read from the hard disk is automatically decrypted, while the user is unaware that anything extra is happening. Find out more at www.becrypt.com.

Cost Of Sarbanes-Oxley Compliance Is At The Expense Of Other Security Spending

A new report published by the Information Security Forum (ISF) warns that the cost of complying with the Sarbanes-Oxley legislation is diverting spending away from addressing other security threats. The global not-for-profit organisation with over 260 Members including half of the Fortune 100, says that many of its members expect to spend more than $10m on information security controls for Sarbanes-Oxley. The business imperative to comply also means that in many cases the true cost of compliance is unknown. For more information about the ISF and a list of their members, visit www.securityforum.org.



With the growing reliance and dependence on our interconnected world, information security is a subject of interest to nearly everybody. Information security - with its focus on confidentiality, integrity and availability - is frequently undermined by security vulnerabilities.

One of the most drastic demonstrations of security vulnerabilities and exploits was shown during a recently conducted experiment "Time to Live on the Network"[1]. During this test, default-installed, and not fully patched systems were connected to the Internet and were monitored for activity. Not surprisingly, within a few minutes the first system came under attack and was completely compromised a few minutes later. Security vulnerabilities and exploits are a real world issue requiring focus and attention for enterprises as well as home users.

Security Vulnerabilities – The path to security breaches

Security vulnerabilities linger and consequently create a breeding ground for exploits, leading to security breaches. Security vulnerabilities originate from many different areas - incorrectly configured systems, unchanged default passwords, product flaws, missing security patches, or even users being tricked into opening a malicious email. All these vulnerabilities have one issue in common – they cause security exposure to an individual system, or even a whole organization.

The most prevalent and widely exploited security vulnerabilities are caused by programming flaws in various software products and applications. Typical examples of such programming flaws are so called buffer overflows in computer programs. Buffer overflows trigger accidental overwriting of sections of memory, which can be compared with people filling out handwritten forms, which provide one space for each character of a person's name. If there are not enough spaces for a person's name you have the equivalent of a buffer overflow in a computer program. The result of a buffer overflow could crash the affected program, or even allow an attacker to execute arbitrary code, and therefore take over the system.

[1] "Time to Live on the Network", Avantgarde, www.avantgarde.com/xxxxttln.pdf

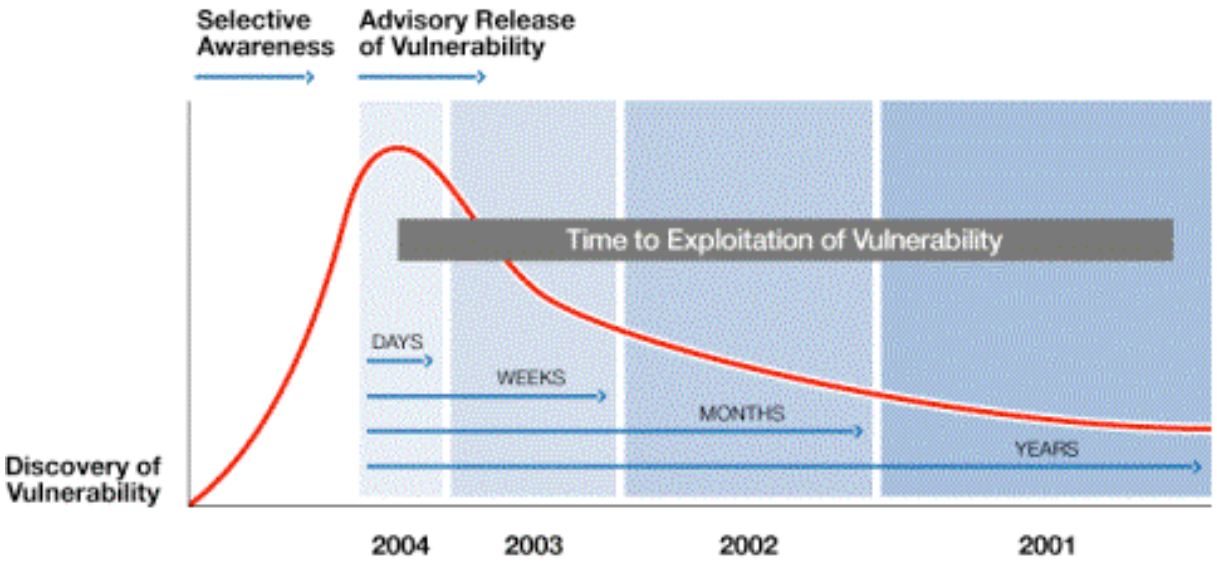
The security research community as well as vendors identify and publish on average 40 new security vulnerabilities per week in various products, from operating systems, databases, applications to even networking devices. Upon release of new vulnerabilities they are being assigned unique CVE (Common Vulnerabilities and Exposures) identifiers [2] to uniquely distinguish and reference them throughout their life-cycle. Depending on the severity of an individual security vulnerability, it allows an attacker to bring down, access confidential data, or take control of a vulnerable system. The software development community is responding with proactive steps. Improved software development processes as well as continuous education of software engineers is one of the steps taken to enhance product security. Focused testing for security flaws helps to prevent and to identify such flaws before products are released to the market.

Exploits and Attacks

Exploits are specifically crafted malicious programs which take advantage of security vulnerabilities and their systems. Browsing the web, reading email, or just being connected to the Internet can lead to exploitation of security vulnerabilities. Exploit programs are utilized by individual attackers to target and take over an individual system – mostly with a specific motive in mind, such as access to confidential information or for financial benefit. Furthermore, exploits are published widely and serve as building blocks for worms and automated attacks. Such malicious programs replicate and circulate automatically on networks identifying unpatched systems. One of the first examples of such automated attacks was the Morris worm in 1988 [3], followed by many more recent examples such as Slammer, Blaster, Sasser, and other worms. Depending on the specific payload worms are carrying, in some cases the victims may be able to recover from the attack, but mostly it is necessary to fully rebuild compromised systems to ensure system integrity.

A critical factor for the impact of an exploit is the timing – how quickly the exploit code is created and available for a specific vulnerability. Recent automated attacks shrank the time-to-exploit window from months to days and happened faster than any possible human response. Rapid availability of exploits creates significant windows of exposure for organizations until they remedy their critical systems. SQL Slammer happened six months after discovery, Nimda was four months, Slapper was six weeks, Blaster came just three weeks after news of the vulnerability, and the Witty worm struck the day after announcement of the vulnerability.

The diagram below illustrates compression of the discovery/attack life-cycle.



[2] "Common Vulnerabilities and Exposures", MITRE Corporation, cve.mitre.org

The most forceful scenario was the Witty worm, which on March 19, 2004 struck about 12,000 computers running firewalls from Internet Security Systems. Witty reached its peak after about 45 minutes. At that time it had infected most of the vulnerable hosts. According to an analysis by CAIDA and UCSD [3], Witty earned several exploitation "firsts": widespread, destructive payload; spread in an organized manner with more ground-zero hosts; had shortest interval between vulnerability disclosure and worm release (one day); attacked only hosts running security software; and proved that applications in a niche market are as vulnerable as those from a software monopoly.

Security Patches – Protecting from Exploits

The timely installation of security patches or other workarounds to every vulnerable system is a necessary defense mechanism to prevent exploits from attacking and compromising the system. In a perfect world a vendor provides security patches at the time of the release of a vulnerability, providing users the ability to protect their systems from exploits. Unfortunately, this is not always the case, and sometimes a vulnerability becomes known before remedies are available. In some instances, even exploits had been circulating before patches were available. These so called zero-day exploits pose a significant danger putting users at risk from exploitation. Applying security patches is not the only solution to the problem. Workarounds exist to mitigate risk and prevent exploitation. Intrusion prevention technologies and other filtering capabilities help to prevent attacks without the immediate need of installing patches.

One of the key issues for every organization is to identify the perfect timing for patching vulnerable production systems. Sometimes patching requires downtime and is disruptive. On the other side lingering exploits require urgent attention. Patch strategies within organizations have matured significantly over the past two years, and organizations are building metrics to measure the severity and criticality of vulnerabilities to determine their urgency. Also, the implementation of predictable patch release schedules (i.e. monthly, weekly) from various vendors helps to eliminate the patch-of-the-day syndrome many organizations were struggling with in the past.

[3] See "The Spread of the Witty Worm," Cooperative Association for Internet Data Analysis and University of California at San Diego, www.caida.org/analysis/security/witty

The Vulnerability Management Process

Successful defenses against network vulnerabilities require utmost understanding the nature of the risk they pose. Vulnerability Management involves the process of identification, prioritization, and remediation of security vulnerabilities. Following the principle of "You can't manage what you can't measure", many enterprises have successfully implemented a systematic vulnerability management process involving the following six steps:

1) Discovery: Identification and discovery of devices, systems, and network topology to keep track of constantly changing networks.
2) Asset Prioritization: Assigning business values and priorities to individual systems and applications. Network security teams should correspondingly prioritize their remediation efforts based on asset value and criticality to an organization.
3) Assessment & Analysis: Comprehensive analysis of systems as well as identifying criticality and severity of security vulnerabilities and security exposure. This information helps to determine what business resources are at risk and what needs attention first.
4) Remediation: Eliminating identified security vulnerabilities by reconfiguring, updating or patching systems. Sometimes workarounds can be applied as a temporary solution.
5) Verification: Validation of patches and workarounds to confirm proper remediation.
6) Policy Compliance: Measuring and reporting against security policies and compliance requirements, such as HIPAA, Sarbanes Oxley, as well as industry specific regulations.

Driven by an organization's security policy, the vulnerability management process should be implemented as a global effort within an enterprise. The level of threat a vulnerability poses to an organization should determine the level of action. The following questions should guide the process. Is the vulnerability exploitable from any system on the network, or does it depend on a user account on the target? Has exploit code already been released? What business resources are affected by the vulnerability? Those factors are unique for every circumstance and determine the threat level of a vulnerability within a specific environment. Vulnerability benchmarks, such as the SANS/FBI Top 20 [4] are frequently adopted to measure specific vulnerability exposure of an environment.

The Business Side of Vulnerabilities and Exploits

Vulnerabilities have a measurable impact on organizations of any size. When critical systems are unavailable and data is not accessible due to an attack, organizations are loosing valuable business. Many enterprises implement vulnerability management as a proactive process closely linked into a broader risk management strategy. Business owners within an organization need to be involved to establish the required support. Information about security exposure is reported to the executive level and in some organizations even board level on an ongoing basis. In particular, tracking vulnerability information over time is a very valuable tool to justify security investment and to proof the return of investment. Other driving forces for implementing a consistent vulnerability management process are regulatory requirements. In particular, industries where confidentiality and integrity of information (such as financial, health care, or other critical sectors) are most critical requirements, organizations perform regular vulnerability audits to verify and report regulatory compliance. Also, the popular trend of outsourcing IT systems and operations drives the implementation of security service level agreements, whereby the outsourcing provider has to conform to specifically defined metrics in terms of patching security vulnerabilities. Vulnerability management processes and security audits are a vital part to measure and enforce such service level agreements.

[4] "The Twenty Most Critical Internet Security Vulnerabilities", www.sans.org/top20

Actions to be Taken

In summary, security attacks on networks and data are increasing in number and sophistication. A new generation of automated security threats is exploiting vulnerabilities faster than any possible human response effort. The timely and complete detection of security vulnerabilities and rapid application of remedies is the most effective preventive measure network managers can use to thwart automated attacks and preserve data security.

Best practices can guide vulnerability management and remediation, helping CIOs, chief security officers, network and IT managers, and security specialists to strengthen and prioritize the protection of internal and external networks. Protection strategies include:

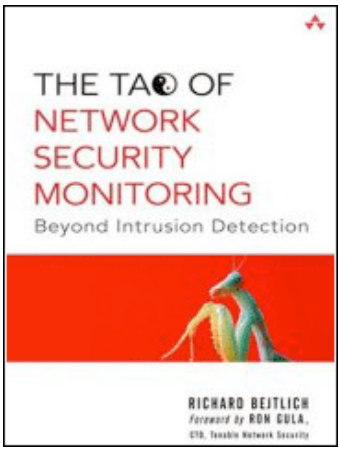
- Education and Awareness. Providing users with actionable information about threats and remedies is a crucial success factor.
- Regular Audits of Security Systems. New automated audit solutions discover everything susceptible to attack, identify, and prioritize vulnerabilities, and provide appropriate remedies.
- Timely Patch Management. This critical process frequently requires manual support with automated solutions to remedy systems in need of urgent care.
- Implementation of Real-Time Threat Prevention. Firewalls and intrusion prevention systems can help stop attacks before penetration.
- Ongoing Evaluation of Security Policy. Trend analysis provides data for enforcing policy and ensures that security systems meet the ever-changing nature of attack threats.

Gerhard Eschelbeck is chief technology officer and vice president of engineering for Qualys, Inc. He published the industry's first research derived from a statistical analysis of millions of critical vulnerabilities over a multi-year period. Eschelbeck presented his findings before Congress, and is a significant contributor to the SANS Top 20 expert consensus identifying the most critical security vulnerabilities. He holds several patents in the field of managed network security and earned Masters and Ph.D. degrees in computer science from the University of Linz, Austria. Eschelbeck can be reached at ge@qualys.com.

The Tao of Network Security Monitoring: Beyond Intrusion Detection

by Richard Bejtlich

Addison-Wesley Professional, ISBN: 0321246772



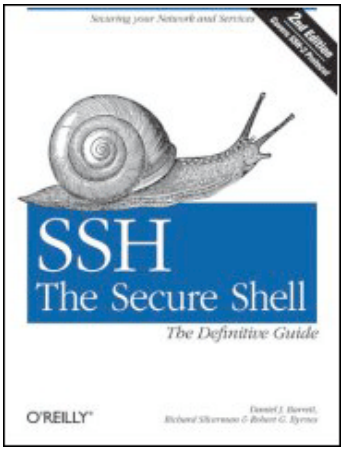
By combining a couple of facts about Richard Bejtlich, such as being one of the the biggest names in the information security community and being an avid reviewer of technical books at Amazon (currently over 180 reviews posted), you can be certain that his book must be a perfect read.

In "Tao of Network Security Monitoring" Bejtlich provides a wealth of information on the NSM concept, where he especially focuses on products, process and human interaction with NSM (covering both sides of the coin, security analysts, as well as attackers). It was also very interesting to read about all the tools Bejtlich uses and recommends. Bottom line, great book for network security professionals.

SSH, The Secure Shell : The Definitive Guide, Second Edition

by Daniel J. Barrett, Richard E. Silverman and Robert G. Byrnes

O'Reilly, ISBN: 0596008953



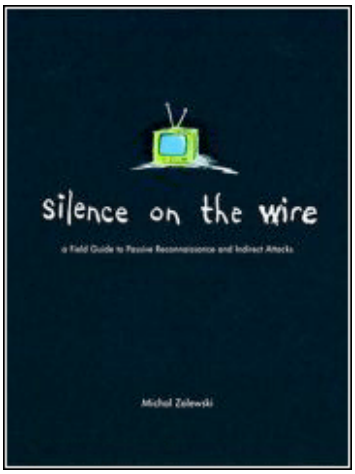
In the second edition of this book, the authors cover the steps of SSH evolution in the past four years (the first edition was releases in early February 2001).

No doubt, the title of this book is true - this is the definitive guide into wonders and possibilities of SSH. There is a lot of newly added stuff: deep understanding of the SSH-2 protocol, coverage of the latest version of OpenSSH and SSH Tectia, new chapters on running OpenSSH on Mac OS X and Microsoft Windows, interesting case study on running a single sign-on between Windows and Linux with Kerberos, etc. A must read for anyone that wants to know SSH.

Silence on the Wire: A Field Guide to Passive Reconnaissance and Indirect Attacks

by Michal Zalewski

No Starch Press, ISBN: 1593270461



This is a book aimed at an audience very much into computer security. It offers a glimpse into some of the aspects of security that the majority of the consultants didn't consider to be a threat at all. The point here is on how we are all exposing ourselves to risks without even being aware of it.

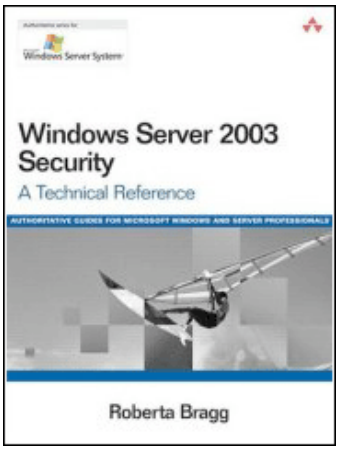
The amount of detail is stunning for such a small volume and the examples are amazing. Many have praised this book for bringing innovative thinking into the world of security. You will definitely think different after reading this title.

The author is a well known security expert and you can read an interview with him in this issue of (IN)SECURE.

Windows Server 2003 Security: A Technical Reference

by Roberta Bragg

Addison-Wesley Professional, ISBN: 0321305019



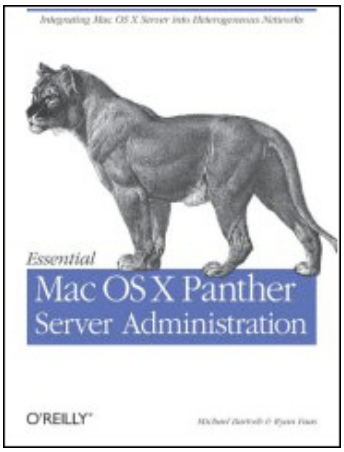
As the author notes at the beginning, if you read this book (never mind if you are a Microsoft or Linux buff), you may find that Windows Server 2003 should have a place in your network.

Security was always a word that lot of people couldn't associate with the Redmond based giant, but publications like the one I am taking a look this moment are here to prove them wrong. The book is planned to be a technical reference for the security related part of Windows Server 2003 administration. Spreading over 1100 pages, it covers an extreme amount of useful information that is well suited for a long list of different knowledge level administrators.

Essential Mac OS X Panther Server Administration

by Michael Bartosh and Ryan Faas

O'Reilly, ISBN: 0596006357



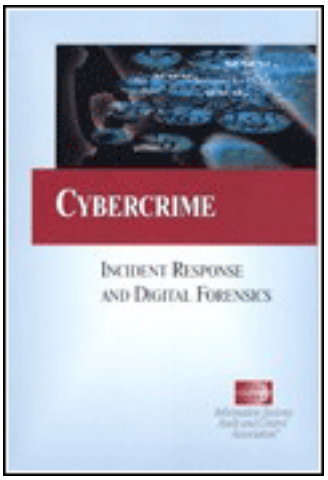
There are tons of Mac users out there that are religiously using Apple products, so using Mac OS X Panther as a server solution shouldn't be a bad thing, right?

This book will guide you through all the details you need to be familiar with to call yourself a good Mac OS X administration. From the security perspective, there are a couple of good info-packed chapters on setting up the firewall service and a virtual private network.

Overall the authors provide a powerful in-depth walkthrough covering all the aspects of configuring and using Mac OS X Panther as a server solution.

Cybercrime: Incident Response & Digital Forensics

by the Information Systems Audit and Control Association
ISACA, ISBN: 1893209687



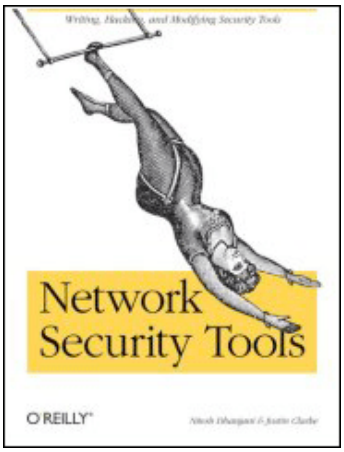
This publication serves its content to the executives, giving them both information as well as advices on the ever present cyber security topics such as: regulatory guidance, addressing cyber risks, creating incident response program as well as basics on the digital forensic investigations.

The book is accompanied by a a rather lengthy set of step by step incident response questionnaires.

“Cybercrime: Incident Response & Digital Forensics” is published by Information Systems Audit and Control Association (ISACA) and it is not widely available. You can get your copy from the ISACA web site - www.isaca.org.

Network Security Tools

by Nitesh Dhanjani and Justin Clarke
O'Reilly, ISBN: 0596007949

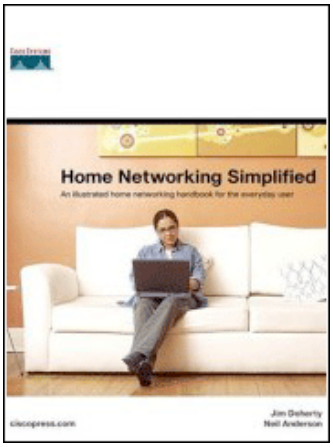


The title of this book implies that it contains information on popular network security tools. While its main topic are this kind of tools, the content delivered superseded my expectations and definitely placed this book among the most interesting titles I recently came across. It did that because, while the authors mention some of the popular security tools, they focus on extending the tool's reach, by manually coding new signatures, detection algorithms and attack mechanisms.

Some of the tools in the spotlight include Nessus, Ettercap, Hydra, Nmap, Nikto and Metasploit network. As you progress towards the middle, you are hit by some fantastic info on writing your own network sniffers (based on libppcap) and packed injection tools.

Home Networking Simplified

by Jim Doherty and Neil Anderson
Cisco Press, ISBN: 1587201364



I was very surprised when I took a glimpse at "Home Networking Simplified". We are used to receiving Cisco Press books, which are almost always very technical and cover specific network technologies. This time Cisco targets the average user by delivering a worthy publication discussing networking basics.

This 300+ pages guide is written in the way that makes it possible for the networking beginners understand what they need to do and how to do it in a fast and efficient manner. All of the book's examples are accompanied by either photos or illustrations, which makes the book pretty easy to apprehend. A quite large portion of the book covers networking security where the authors especially deal with wireless (in)security. If each average computer literate household would own this book, the online world would be a much safer place.

PDA attacks: palm sized devices - PC sized threats

By Seth Fogie

A PDA is generally considered a simple computerized device that is handy for email/task updates, listening to music, and playing games. Unfortunately, this 'simple device' mindset has created a loophole in many corporate networks that can be exploited by attackers.

Since most companies barely recognize the PDA in their policies or control systems, they often have no idea how they are used on the network or at remote sites. This is bad. Not only can the PDA itself be compromised to the point where an attacker can access everything held in its memory, but in many cases, the PDA itself can be made into a trojan that can turn a network inside out. This article will examine just some of the ways that a PDA can be owned by an attacker and what can happen as a result.

Defining the PDA

For the sake of this article, a PDA is any digital device that is used to assist the owner with their daily task management. This includes, for the most part, all Pocket PC devices, Palm based devices, Smartphones, and related products. In addition to these more obvious categories, the content in this article also applies to handheld computers that use the CE .NET platform.

For example, many inventory scanners now use embedded operating systems that can run client/server based programs operating over a wireless link. Please note that while much of this article will focus on Windows Mobile devices, the general attacks can apply to a Treo just as easy as a Pocket PC.

Obstacles for the PDA Attacker

One of the obstacles a PDA attacker has to overcome is the fact that the devices operating system files are often stored in ROM. This can be a challenging obstacle because trying to debug a program that runs from ROM is not always possible. As a result, an attacker either has to target a 3rd party program, or probe a flaw via blind testing (AKA blackbox testing).

Second, almost every PDA is unique. For example, just because Dell and HP both use Windows Mobile, doesn't mean it is the same version. Both of these companies install their own OEM software in addition to the core OS files, which can really make writing malware challenging. As if this isn't enough, ROM upgrades also have a huge impact on how a PDA's file structure is laid out and how the core files are loaded. As a result, there are major differences between any two PDA's memory addressing.

Third, attacking a PDA is relatively new. As a result, there is little in the way of support or documentation. An attacker has to be dedicated to the point of obsessed, or else have some exposure to the processor or platform used on the target device.

The point is this: attacking a PDA is not as easy as attacking a PC, at least on the surface. However, once you accept the limitations, a handheld can be a very easy platform to attack. The rest of this article will demonstrate just some of the vectors and methods we discovered that can leave a PDA wide open for attack.

The CAB File

Cabinet files have long been used by Microsoft to package up a group of files that are needed during the installation of a software product. As a result, it is no surprise that the same CAB format is used on Windows Mobile devices to install programs that are either directly downloaded to the PDA, or passed over from the connected Active Sync connection. The problem isn't found in the fact that the PDA uses these files, but more in how these files are processed and then subsequently deleted.

It was discovered early on that CAB files self delete themselves after they are executed. While this attribute is probably an attempt to help PDA users keep their memory clean from clutter, it ironically behaves similar to self deleting trojan files that can be found on the PC. This alone has very little in the way of substance when it comes to security, but it is a very useful way to hide the installation of a backdoors into a PDA. Since the execution of the CAB file can create new files and registry entries, as well as update or delete existing ones, a user can easily be tricked into installing malware that is hidden away inside a seemingly valid packaged file. For example, how do you really know that CAB file you downloaded off the Pocket PC warez site really contains just the game you thought was there? The good news is that Windows Mobile 5 will not automatically delete executed CAB files. This won't stop you from infecting yourself, but it will allow you to investigate the files contents post install to see what kind of unwanted files and settings might be lurking.

The point is this: attacking a PDA is not as easy as attacking a PC, at least on the surface.

Autorun

Another minor, but very significant threat to Pocket PC users is the fact that inserted media cards have the power to execute code via a relatively unknown Autorun feature. As designed by Microsoft, if a file named autorun.exe is placed inside a folder named '2577' (on ARM based devices) it will be copied to the local /Windows directory and executed. If the autorun.exe is a trojan or some other form of malware, then it will be executed unknowingly by the PDA user. To illustrate just one of the possible abuses of the autorun feature, imagine the consequences if the brador.exe trojan (discussed later) was renamed to autorun.exe and placed into the 2577 folder on a dedicated SD Atari game card. This card could easily be passed around the workplace and infect every PDA user in the company. Of course, it would be just as easy for an attacker to install a hidden FTP server, or cause the PDA to instantly hard reset.

Soft and Hard Resets

The PDA is a unique device in that it stores its data on internal RAM/ROM memory chips. In general, the core OS files are stored onto a flashable ROM chip, with all additional programs and files being stored on RAM chips. The latest models have started to include more ROM based memory, and even drives, but this still not where most documents, tasks, calendars, etc. are stored.

The side effect of this design is that the device must maintain a constant source of power or it will loose all this RAM based data. Most people are aware of the side effect of letting their batteries die, but what most PDA users are not aware of is that a hard reset can be initiated with only a few lines of code, as illustrated in listing 1 below.

If you happen to download an executable, or perhaps insert a media card with a hard reset program setup for autorun, your PDA will be wiped.

```
#include <windows.h>
#include <winioctl.h>
#define IOCTL_HAL_REBOOT CTL_CODE(FILE_DEVICE_HAL, 15, METHOD_BUFFERED,
FILE_ANY_ACCESS)
extern "C" __declspec(dllimport)void SetCleanRebootFlag(void);
extern "C" __declspec(dllimport) BOOL KernelIoControl(
     DWORD dwIoControlCode,
     LPVOID lpInBuf,
     DWORD nInBufSize,
     LPVOID lpOutBuf,
     DWORD nOutBufSize,
     LPDWORD lpBytesReturned);

int WINAPI WinMain(HINSTANCE hInstance, HINSTANCE hPrevInstance,
     LPTSTR lpCmdLine, int nCmdShow)
{
     SetCleanRebootFlag();
     KernelIoControl(IOCTL_HAL_REBOOT, NULL, 0, NULL, 0, NULL);
     return 0;
}
```

Listing 1: The Hard Reset Code

If you happen to download an executable, or perhaps insert a media card with a hard reset program setup for autorun, your PDA will be wiped.

The problem isn't so much that this code exists. Instead, the problem is that the KernelIOControl command responsible for a hard reset is easy to access by any program running on the PDA.

In other words, the lack of security of the kernel level command can turn a useful function into a potential denial of service program.

Pocket IE

In the PC world, Internet Explorer is the target of many exploits. Fortunately, most of these attacks are impossible against the Pocket PC because Pocket IE (PIE) is a much stripped down version of its big brother. However, this does not mean PIE is free from attacks. The following outlines several of the potential ways PIE can be attacked.

DoS Attack

Although PIE comes with a stripped down version of JavaScript, it is still possible to perform various annoying script based attacks that will force the end user to reboot their PDA. These are only worth mentioning as they are trivial. However, there are several other problems within Pocket IE that cause the browser to instantly close. One error in this program can be found in the way it handles HTML lists that use CSS formatting, as illustrated in listing 2 on the following page.

```
<style>
#layer1 div.sublayer1 { width:50%; margin:0 1 2 3; padding:4;
float:right; }
#layer1 div.sublayer2 { width:50%; margin:0 1 2 3; padding:4;
float:right; }
</style>
<div id="layer1">
<div class="sublayer1">
<ul>111</ul>
</div>
<div class="sublayer2">
<ul>222</ul>
</div>
</div>
```

Listing 2: HTML Code that will Crash Pocket IE

URL Obfuscation

Several years ago, it was discovered that Internet Explorer would process Unicode encoded URLs. When this flaw is combined with special allowances within the HTTP protocol for user:pass@domain requests, you have a means to trick people into thinking they are clicking on a valid link. Unfortunately, they could be directed to a spoofed site. For example, the following link will take you to www.airscanner.com, not www.paypal.com as one may think. Given the limited size of the PIE address bar, it is easy to imagine that someone would not suspect they have become the victim of an attack.

```
http://www.paypal.com&login.ran
d-%00%01AE67D12EF9090AB933@%36%
39%2E%30%2E%32%30%30%2E%31%30%3
6/
```

PIE Domain Object Model

Frames have long been used by web developers to help them present a collection of web pages all under one main frame. The catch to this is that code in one frame cannot read from or write to the content in another frame. The reason for this is that an attacker could hide malicious code in an invisible frame and alter the content in another frame such that an end user could be tricked into giving up personal information. The following demonstrates just one way this could be exploited. Notice the URL which apparently points to Johnny.ihackstuff.com, but actually points to a framed webpage on the Airscanner web server. Once the page loads, the content on the main page of Johnny's website is altered via hidden code running in an invisible frame (figure 1).

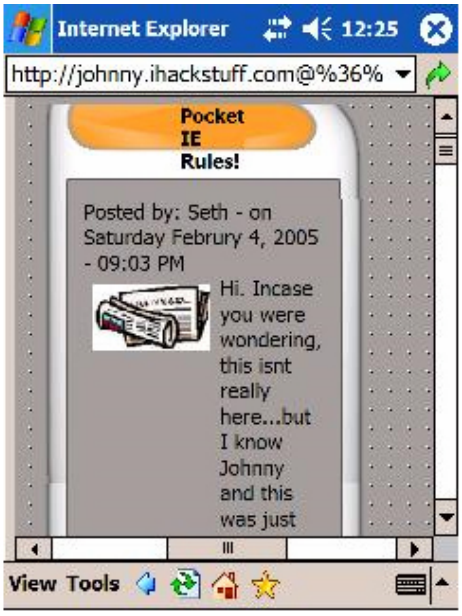


Figure 1: Using XFS to Alter Web Content

PIE Local File Access

Another quirk with PIE is that it can access local files on the PDA. Combine this with the previously described problem with the domain object model security and there is a risk for exploitation. The following is just a small list of the types of files that can be loaded into a PIE frame:

- xls
- htc & htp (in IE)
- cpl (Control Panel) items
- ini files (in IE)
- 2bp images (in IE)
- go to any folder
- go to 00 (root) folder

One potential vector for abuse was found in the fact that PIE can be tricked into loading up previously cached webpages, which can then be altered by code in a hidden frame. For example, we learned that it was fairly simple to pull up a cached version of the Paypal.com home page, alter the Form field properties to point to our own server, which would then capture any passed user credentials when the submit button was pressed.

Keyboard Logger

The PC world has long been wary about keyboard loggers because they have been used by attackers for many years. However, when it comes to the Pocket PC, not much thought is given to the possibility of a logger that can capture all your keystrokes because the concept of a keyboard is completely different.

Instead of a physical button based keyboard, the Pocket PC uses a SIP (Soft Input Panel) that combines a graphical keyboard with some code to emulate keypresses and output a character into the part of the screen that has the focus. While it is true that this makes it a little bit harder to capture keystrokes, it is by no means impossible.

When you PDA boots up, it first scans the registry for all installed input panels. By default, this includes the Block Recognizer, Keyboard, Letter Recognizer, and Transcriber. Each of these methods is controlled by a ROM based DLL file that holds both the graphic for the input panel and the code needed to convert input from the stylus into characters. For example, the main Keyboard option is all controlled by a file called MSIM.DLL.

Thanks to the way the input panel was designed, it is rather easy to create a new keyboard and install it onto the PDA. In fact, there are several keyboard packages you can buy and install. With this in mind, we created our own keyboard with a little bit of extra code that captured the inputted character and placed it into a text file. Next we figured out what registry settings needed to be altered to remove the existing 'Keyboard' entry. Finally, we worked out the necessary registry keys that would allow our new keyboard to be loaded with the name of 'Keyboard', and packaged it all up into a CAB file. After a few tests and some trouble shooting, we had a working keyboard logger that looked and acted like the default keyboard that comes with our version of Windows Mobile, except it outputted a copy of each character into a hidden file.

The following lists the registry settings we had to update:

- IsSIPInputMethod disabled for real keyboard
    - CLSID: 42429667-ae04-11d0-a4f8-00aa00a749b9 (set 1 to 0)
- 'Keyboard' name & icon borrowed by keylogger.dll
- New keyboard has own CLSID with settings
- HKCU\ControlPanel\SIP\DefaultIM\{CLSID}
- IsSIPInputMethod enabled for fake keyboard

Once installed, the end user will have no idea their keyboard was swapped out for a newer improved model. In addition, the only way to tell that a new DLL is being loaded with the system is if you use a debugger to view the DLL files associated with the executable responsible for handling the keyboard and other hardware on the PDA.

Forensics/ROM Rips

The PDA has a very limited amount of data storage ability, with the exception of the 1+ gig memory cards you can now buy for them. However, if you want quick access to that data you would have to take the PDA back to your computer, hook it up, and manually copy over all the data stored on the device. At least, that is unless you take advantage of the power of a forensics or ROM/RAM data dumper.

The law enforcement community has started to notice PDA's and now examine them when they are found at a crime scene. To help with this, there are several forensics programs available that can make short work of pulling across the entire contents of the PDA. For example, Paraben software makes just such a program that temporarily installs a client on the PDA and uses that to create a connection with a PC to download the contents of the PDA. In addition to this, there are several tools available at sites like xda-developers that can extract the entire contents of your PDA into an inserted SD card via the autorun.exe method we discussed earlier. What this means to the PDA owner is that their PDA only has to be in the wrong hands for a short period of time for it to have its contents complete ripped.

The law enforcement community has started to notice PDA's and now examine them when they are found at a crime scene.

Disassembling Binaries

The Pocket PC/Windows Mobile platform is typically run on an ARM processor, which is common in low heat and reduced power environments.

This small fact combined with the limited number of instructions (ARM - Advanced RISC Multiprocessor //RISC – Reduced Instruction Set) makes learning and understanding the ARM instructions relatively easy, especially when compared to other processors.

The first thing you would need to understand to disassemble binary files is how a processor works. In general, it read the opcodes at a specific location in memory (program counter) and performs the very specific command. Depending on the instruction, the processor will add, subtract, branch to a different part of the code in memory, or even do nothing at all. To help the processor do its job, it has 32 registers, which are nothing more than 8-byte memory storage sites located in the processor. This onsite storage simply helps the processor move faster by keeping memory access local, instead of reading and writing directly to and from the RAM.

To get inside the binary files on a Pocket PC, you will need a few programs. First, you will need a disassembler to convert the raw hex code into a more readable opcode format. We generally use IDA Pro, which understands Windows CE files. Second, we use the debugger included with Embedded CE++ from Microsoft.

This is needed to watch and change the operation of the code on the fly and to generally see what is happening live. Finally, you will need some sort of hex editor to make changes to the code for on system testing. We use UltraEdit, but there are many other options. Once you have all these tools, you can start peeking around inside your Windows Mobile device to see what is really happening.

The only other item to deal with is that you have to understand the opcodes and instructions used by the processor.

In other words, you have to speaking the language of ARM assembly. Fortunately, by understanding o the following few instructions, it is possible for you to read and even write ARM based assembly code that can run on a Pocket PC.

Move (MOV) - Moves a value into a register. This could be a hard coded value or the data in another register.

Compare (CMP) - Compares the value in a register with a hard coded value or the data in another register.

Branch (B) - Branches the execution to a specified memory location.

Branch Link (BL) - Branches the code execution, and then returns to original memory address.

Load Register (LDR) / Store Register (STR) - Stores data or loads data to or from the RAM into the processors registers.

There are other opcodes that you will come across. If you really want to dig into a processor at this level, we recommend you visit www.arm.com and download their processor manuals.

Backdoor FTP

Once you understand how a program works at the processor level, you can do all kinds of things to change its functionality.

In this specific case, we were able to turn a visible and perfectly legit FTP server (figure 2) into a hidden and undetectable backdoor.

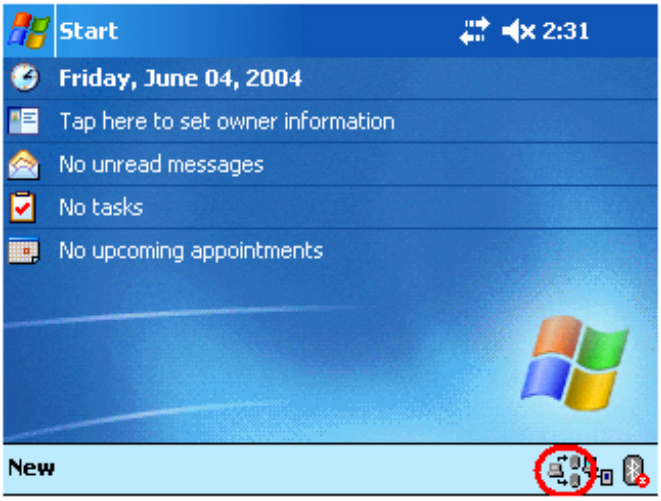


Figure 2: Visible FTP Server Icon

To do this, we first locate the code responsible for the icon that appears on the menu bar. This took a few minutes, but thanks to the function name listing in IDA Pro, we traced this part of the code back to a function named Shell_NotifyIcon. Once found, we then changed the code as follows:

```
MOV Shell_NotifyIcon to MOV R0,
R03A 01 00 EB to 00 00 A0 E1
```

The left side shows the original command, below which is the original hex values. To remove the icon, we updated the code to never show the icon. In its place we inserted a virtual non-operation (NOP) by writing in the command to move R0 into R0.

We also changed the default port number from the standard port 21 to a higher port number that would probably not attract much attention. This was as simple as locating the hard coded port and changing it to the hex equivalent of our preferred port.

Hidden Remote Control

For our next example, we found another program that gave us remote control access to the Windows Mobile device; vRemote.

This program again had a very visible window that pops up on startup, which we wanted to remove (figure 3 on the following page).

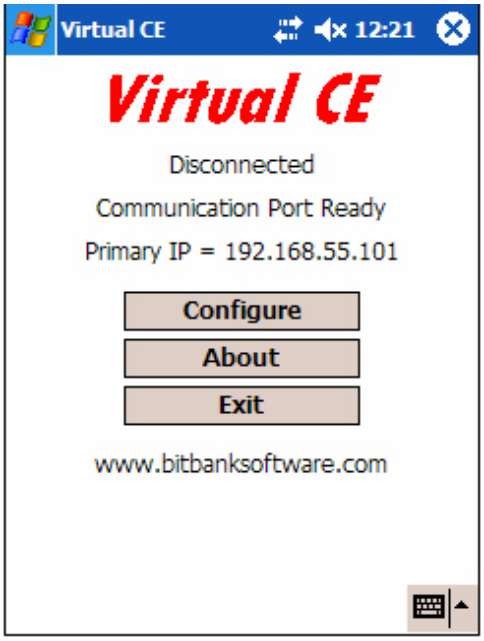


Figure 3: vRemote Window

The following shows the changes we made, which complete removed all visible indication that the vRemote program was running on the device.

```
BL ShowWindow to MOV R0, R0 (Virtual NOP)
A6 15 00 EB to 00 00 A0 E1
```

Again, we were able to do this by targeting the ShowWindow function in IDA and overwriting it with a simple NOP function. In this case, we also had to overwrite two other parts of the program that made reference to this code.

The point of these two examples is that Windows Mobile programs can easily be altered to create malicious versions that can create havoc or assist an attacker in gaining access to sensitive data stored on a PDA. If this low level modification is interesting, you can read more about it in "Aggressive Network Self Defense". Chapter one of the book goes into great detail about how two PDA owners fight a digital war by modifying and re-modifying code at the ASM level.

Known Malware

Windows Mobile has not yet attracted much attention when it comes to publicly released viruses and Trojans. In fact, to this date, there are only two known pieces of malware for Windows Mobile device. Although the number of malware is small, their technical sophistication is great, as we will discuss next month in part two of this series.

Network Based Attacks

The Windows Mobile platform is designed for wireless networking. As a result, it is vulnerable to many of the same types of problems that regular PC's must deal with on an hourly basis. The following outlines several types of attacks that can be used against a PDA over a network.

Denial of Service

A PDA generally uses a low power wireless network card to access a local Wi-Fi network. From here it can use the web, download email, etc. Since this device is lower power, its wireless signal strength is not that strong. As a result, it is rather easy to create enough interference using a rouge access point or laptop to keep a PDA from getting a good connection. In addition, it is sometimes possible to perform a rapid ping DoS against a PDA, which simple takes up enough resources that the PDA slows to a crawl. Finally, an attacker can target services on regular PC's, such as the Active Sync port 901, which will keep an associated PDA from ever establishing a connection.

Since DoS attacks are a dime a dozen we will not spend any more time on them. Just know that they can be used by attackers.

Bluetooth Issues

Bluetooth attacks have long been making headlines. However, when it comes to the details of the attacks, most of them are targeted for the cell phone, not the PDA. Despite this, there are a couple issues that we should address. Note that these problems may or may not affect your device depending on the Bluetooth software installed by the vendor/OS maker.

The first is the 'attack' of bluejacking, which is really nothing more than sending messages to a Bluetooth enabled PDA or phone. There are numerous sites online, such as bluejackq.com where you can learn more about this issue.

The next issue that can affect PDA owners is Bluetooth denial of service attacks. One method for this is to send a 'ping' packet to the Bluetooth enabled devices, which ties up the system resources and ensures no one else can connect to the device.

There are also other related attacks that can disable or disrupt service.

PIN cracking is the final issue we would like to mention. Since Bluetooth is only protected by a 4-digit pin number, it is fairly easy to brute for attack the pin on a target device. Generally, this can be accomplished in several hours, depending on how many client devices are guessing the pin and if the right programs are used. It is also possible to find non-discoverable devices via a brute force approach that scans for MAC addresses used by Bluetooth devices.

Since Bluetooth is only protected by a 4-digit pin number, it is fairly easy to brute for attack the pin on a target device.

Wi-Fi Issues

Wireless networking attacks against a PDA are not specifically a PDA threat. However, given the fact that PDA owners generally use Wi-Fi connections to view email, connect back to the office, and surf the internet, we have to at least mention this vector of attack.

In short, a PDA owner can be attacked in the following ways: WEP key cracking, WPA key cracking, MiTM attacks, web content injection, sniffing related issues, rouge AP's, and denial of service attacks.

Using these types of attacks, an malicious hacker can capture and view emails, obtain access to corporate networks, learn password information, trick PDA users into accessing spoofed content online, and much more.

The point is, every thing that affects regular PC users when it comes to wireless connectivity, also affects PDA owners.

Buffer Overflow Attacks

It seems as if every day brings a new buffer overflow exploit against some PC based program. As a result, most people are familiar with the term. However, how many people look at a PDA and think they could be attacked via the same type of weakness?

As it turns out, Windows Mobile devices are vulnerable to buffer overflow attacks in much the same way as their big brothers. In fact, it only took us about 10 minutes to find one in the same FTP server program we previously discussed.

Specially, the ftpsvr.exe file uses the general FTP commands much like any other FTP server. However, it fails to properly check the parameters of the command, which means an attacker can overflow the memory stack with raw code with a long and padded hex string that contains ARM assembly code.

While the details vary from device to device and OS version to OS version, we discovered it was possible to inject the code responsible for a hard reset over a network and cause a PDA to hard reset.

Fortunately, buffer overflow attacks are somewhat limited on a PDA. Since there is no real command shell, an attacker would have to essentially upload an entire program to gain access to the Pocket PC.

On the other hand, brador is a small trojan and could be used in just such an attack.

PDA Attack Tool

We demonstrated several ways that a PDA can be abused and attacked. However, this is just one side of the story because a PDA can also be used to facilitate attacks against other computer users.

This section will look at just some of the ways that a Pocket PC can be used by an attacker to penetrate a network from the palm of their hand.

Sniffer

A sniffer provides the user with a look into the data that is traversing the network. In today's switched network, sniffing generally requires ARP manipulation or some other type of network trickery. The exception to this is the wireless network, which is where most PDA's operate.

A wireless network operates much like a hub, when looking at it from a sniffers perspective. All the data is available to any device that is close enough to see the traffic. Since PDA's often go unnoticed and can be used inconspicuously, it provides an attacker with the perfect device to get inside a building and collect wireless traffic for future analysis.

While Windows Mobile based sniffers generally require the user to be associated with a network before it will detect any traffic, Linux based Pocket PC's can capture data easily on any selected channel using popular tools like Kismet or tcpdump. Figure 4 is a screen shot of a vxSniffer capturing the data from another wireless user as they view Google.com.

It is also important to note that a Pocket PC can include an Ethernet card as well as a wireless card.

With this optional component installed, it is possible to plug a PDA into a switch or hub and directly access the data passing over the wire.

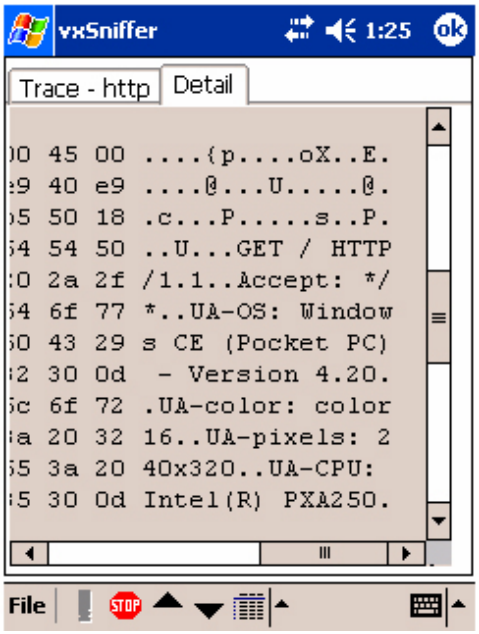


Figure 4: vxSniffer in Action

Vulnerability Scanner

Once a PDA has access to a network, it can perform the same kind of scans and probes that typically are ran from a regular computer. Ping scans and sweeps, probes for popular services, and even exploits can be execute from a Pocket PC in order to penetrate a network. For example, we use an iPAQ running Linux routinely to perform nmap scans and more inside target networks. In addition, with Linux installed, it is fairly easy to build perl scripts that can do all kinds of automated penetration testing.

Drop and Go Backdoor

A PDA is a fairly cheap device. You can pick an iPAQ off Ebay for under $100. Combine this with an extension sleeve, a wireless NIC and Ethernet card and you have the perfect drop and go backdoor into most any network (figure 5).

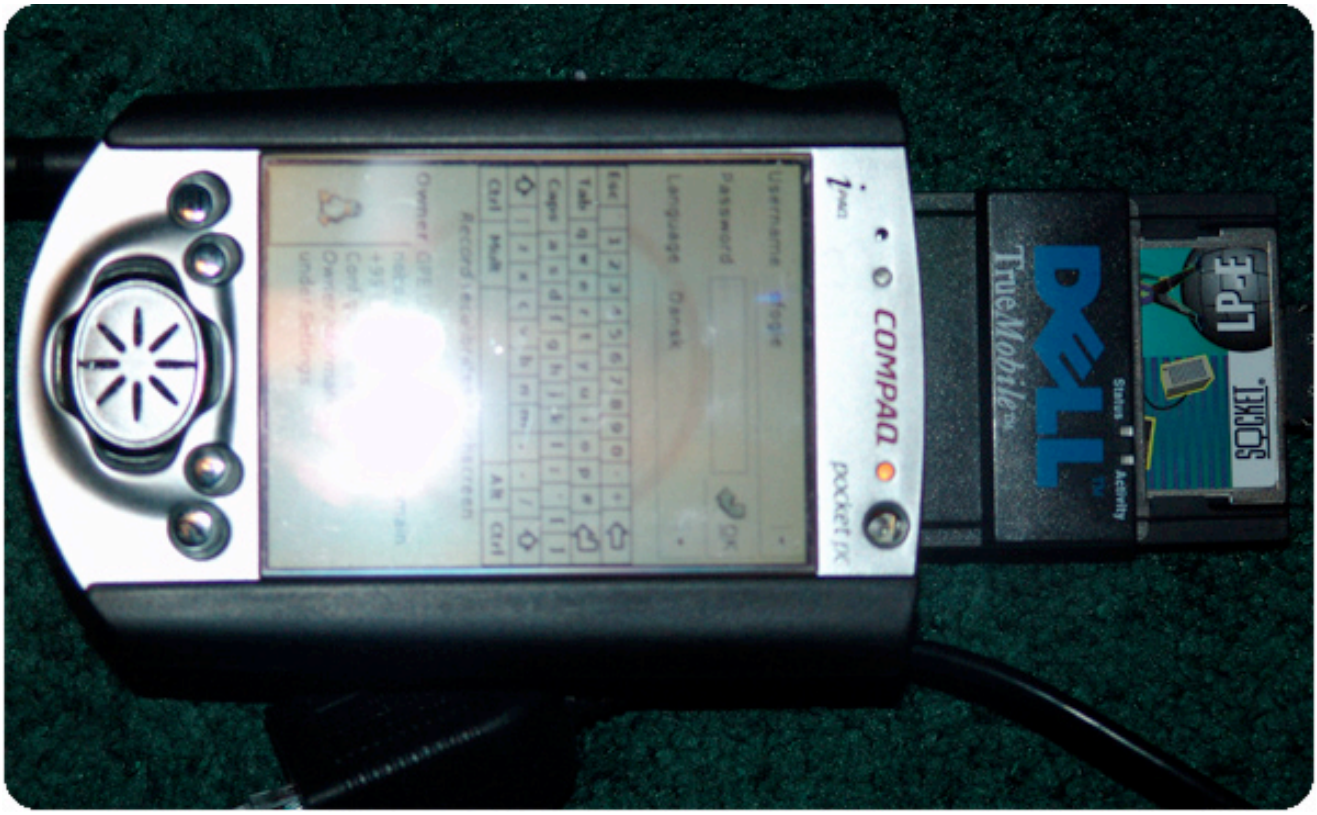


Figure 5: An example iPaq Backdoor with Linux, WNIC, and Ethernet NIC

The following outlines the general steps to having a disposable hardware based backdoor that can be easily dropped into a wiring closet and remotely accessed via a wireless link to provide anonymous access into a companies network.

1. Install familiar Linux onto iPAQ.
2. Install drivers as needed for wireless card and Ethernet card.
3. Create a script to configure the wireless card upon boot up in ad-hoc mode, with optional encryption, on a free wireless channel.
4. Install ssh, tcpdump and nmap on the iPAQ.
5. Take device to site and find a free port behind a printer, under a desk, or in a closet and plug it in. If necessary, you can use a passive hub to create a free port between an existing computer and the wall.
6. Turn on PDA and walk away. Then from laptop or another PDA, establish a connection with the PDA using the pre-established IP address, and connect via an ssh tunnel.
7. You can now run tcpdump to establish correct internal IP scheme, and configure Ethernet card with corresponding IP address information.

The rest is up to you to figure out...

If you don't think this is a reality, we recently used this exact setup in a site test and were able to connect to their network remotely via the P2P wireless link. In addition, one of the Stealing the Continent books from Syngress outlines a similar type of device they call a 'creeper'. The point is, a PDA makes for a great disposable device that an attacker can drop in behind a firewall and use to remotely own your network.

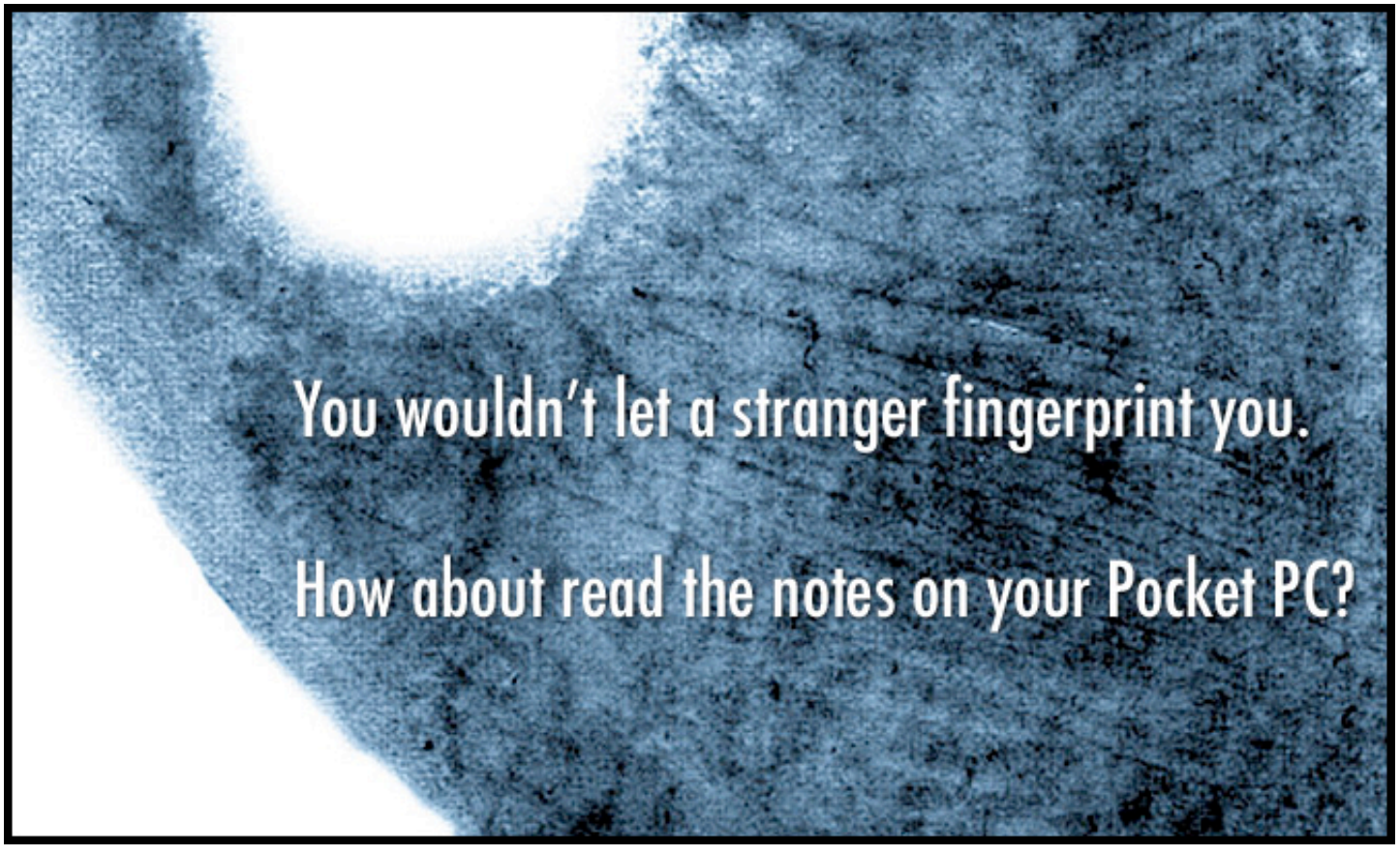
Securing the PDA

By this point you should be able to recognize many of the dangers involved with owning and using a PDA. For the SOHO user, most of these problems can be mitigated through the same general care shown when using regular PC's; don't download unknown files, use antivirus and firewall as needed, etc. However, the enterprise environment is a different issue. Here the company is responsible for monitoring and controlling PDA use, which requires a more complex prevention/protection strategy. To address problems for enterprise users, we will take a closer look at the details behind implementing an enterprise wide PDA security policy in part three of this series.

Summary of PDA Attacks

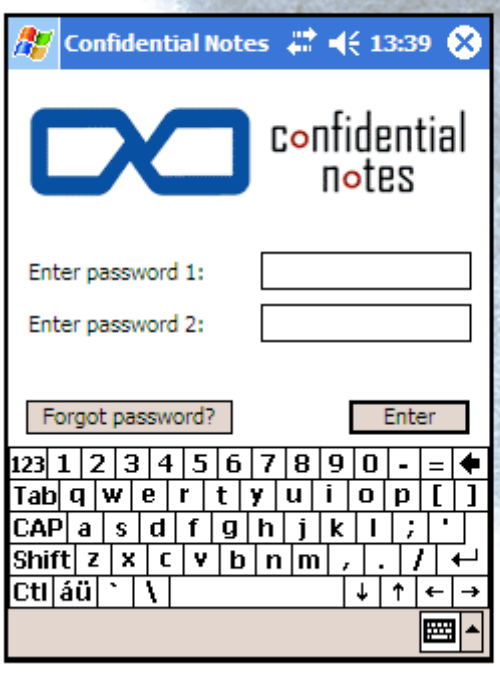
A PDA can provide its user with games, productivity tools, web browsing, email, document writing, and more. However, along with these valuable tools comes a risk that can not be denied – a Pocket PC can become the target for attack. Be it a malicious piece of code, or a Pocket IE trick, it is important to recognize that that PDA is not free from harm. In addition to this, it is also important to see the PDA as a potential threat to your own security. These devices are easy to smuggle into a business and can be used to propagate an attack against network devices. Don't make the mistake of assuming a PDA is a simple date keeper. As the cliché goes, it isn't the size the counts... it is how you use it that matters.

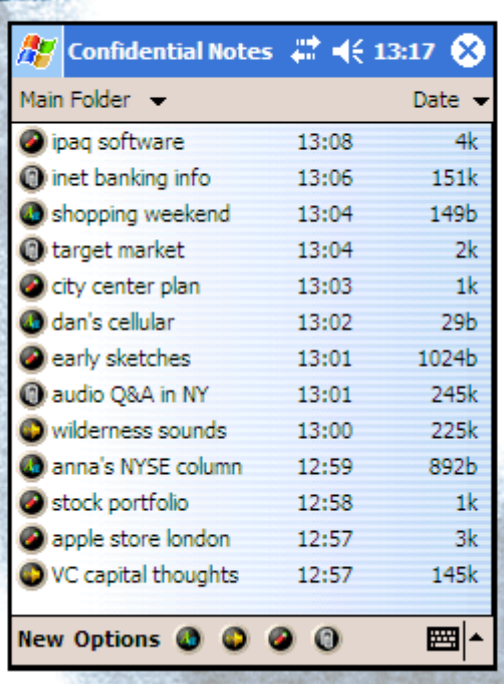
Seth Fogie is the VP for Airscanner.com, a mobile device security company, where he is responsible for the product testing, research and development. Seth is also a regular speaker at conferences such as BlackHat and Defcon, and has authored numerous articles and books related to information security, the latest which is "Aggressive Network Self Defense" from Syngress.

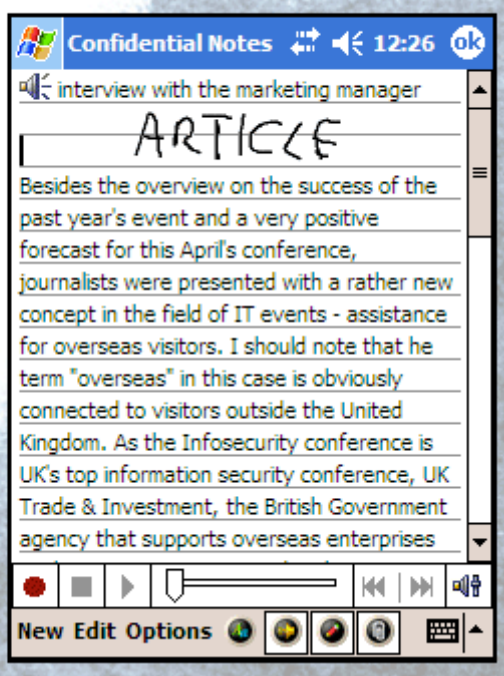


Confidential Notes is a practical and easy to use solution that instantly provides you with a high level of security for your mobile data.

For more information on Confidential Notes visit www.pocketpcsecurity.com







8th Information Security Conference (ISC'05)
21 September-23 September 2005 - Singapore
http://isc05.i2r.a-star.edu.sg/

The 4th International Workshop for Applied PKI (IWAP'05)
21 September-23 September 2005 - Singapore
http://iwap05.i2r.a-star.edu.sg

IT Security World 2005 Conference & Expo
26 September-1 October 2005 - Hyatt Regency San Francisco, USA
http://www.misti.com/

HealthSec 2005 Conference & Expo
28 September-30 September 2005 - Hyatt Regency San Francisco, USA
http://www.misti.com/virtprogHS05/program.asp

RSA Conference Europe 2005
17 October-19 October 2005 - Austria Center, Vienna, Austria
http://2005.rsaconference.com/europe

CNIS 2005: IASTED International Conference on Communication, Network and Information Security
14 November-16 November 2005 - Phoenix, USA
http://www.iasted.org/conferences/2005/phoenix/cnis.htm

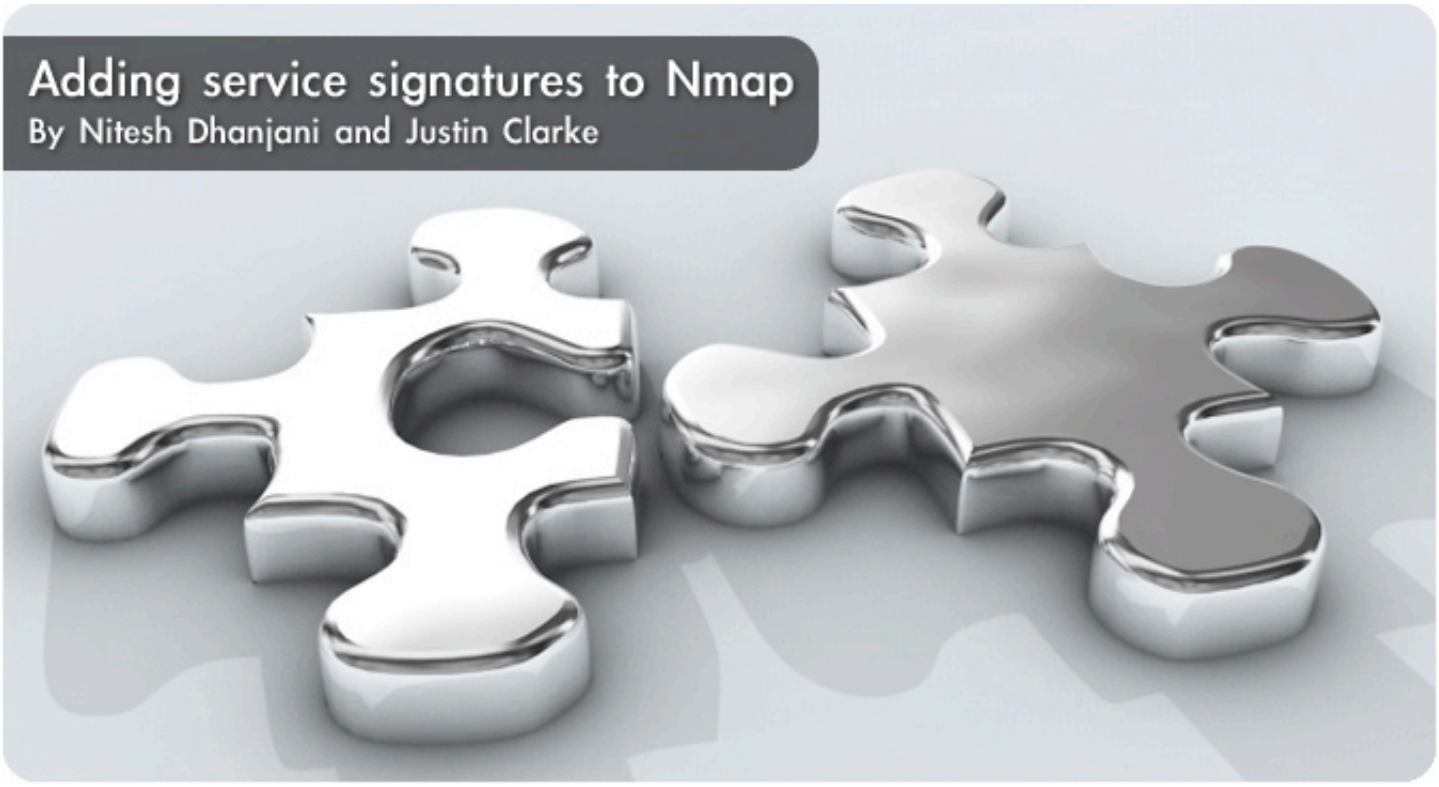
Asiacrypt 2005
1 December-4 December 2005 - Chenna, Madras
http://www.iacr.org/conferences/asiacrypt2005/

3rd International IEEE Security in Storage Workshop
13 December-13 December 2005 - Golden Gate Holiday Inn, San Francisco, USA
http://www.ieeeia.org/sisw/2005

RSA Conference 2006
13 February-17 February 2006 - McEnery Convention Center, San Jose, CA, USA
http://2005.rsaconference.com/us/C4P06/

Adding service signatures to Nmap

By Nitesh Dhanjani and Justin Clarke



Recent versions of the popular port scanner Nmap can detect the type and version of services running on a network. This is illustrated in example 1 below:

Example 1. Example Nmap version scan

```
>nmap -sV 127.0.0.1
Starting nmap 3.50 ( http://www.insecure.org/nmap/ ) at 2003-07-05 17:12 EDT
Interesting ports on localhost (127.0.0.1):
(The 1658 ports scanned but not shown below are in state: closed)
PORT STATE SERVICE VERSION
22/tcp open ssh OpenSSH 3.8.1p1 (protocol 2.0)
Nmap run completed -- 1 IP address (1 host up) scanned in 1.104 seconds
```

This scan is implemented as a series of probes and responses in the file `nmap-service-probes`.

This file defines the probes that will be sent to the service to elicit some response, as well as a series of regular expressions against which to match responses to determine which services are running and, where possible, their versions.

At a high level, the version-scanning methodology follows this process:

- If the port is a TCP port, connect to it and listen. This is called the NULL probe. Many services will return a banner on connection. If a match is made, processing stops.

- If no match is given, or if the protocol is UDP, probes defined in the `nmap-service-probes` file will be attempted if the protocol and the port ranges in the file match. If a response matching a probe is found, processing stops. If a soft match occurs (whereby a service is recognized, but not its type or version), follow-on probes will be limited to relevant ones.

- If no match is found, each probe in the `nmap-service-probes` file will be tried, regardless of the ports on which the service usually runs. This will be limited where a soft match has already occurred.

- If SSL was found, Nmap will connect using SSL (if available) to run the version-detection process again.

If a service responds to a probe sent during this process, but Nmap does not recognize the response, Nmap prints a fingerprint for the service that you can use to report the signature to the Nmap developers, as shown in example 2.

You can use this, together with the version and service information, to include a signature that recognizes this service in the `nmap-service-probes` file in the future.

Example 2. Nmap unrecognized service

```
>nmap -sV -p 4738 127.0.0.1
Starting nmap 3.50 ( http://www.insecure.org/nmap/ ) at 2003-07-05 17:39 EDT
Interesting ports on localhost (127.0.0.1):
PORT STATE SERVICE VERSION
4738/tcp open unknown
1 service unrecognized despite returning data. If you know the service/version, please
submit the following fingerprint at http://www.insecure.org/cgi-bin/servicefp-
submit.cgi :
SF-Port4738-TCP:V=3.50%D=7/5%Time=40E9CA80%P=i686-pc-linux-gnu%r(NULL,59,"
SF:Login\x20with\x20USER\x20<name>\x20followed\x20by\x20PASS\x20<password>
SF:\x20or\x20ANON\r\nCheck\x20privileges\x20with\x20PRIVS\r\n")%r(GenericL
SF:ines,59,"Login\x20with\x20USER\x20<name>\x20followed\x20by\x20PASS\x20<
SF:password>\x20or\x20ANON\r\nCheck\x20privileges\x20with\x20PRIVS\r\n")%r
SF:(GetRequest,59,"Login\x20with\x20USER\x20<name>\x20followed\x20by\x20PA
SF:SS\x20<password>\x20or\x20ANON\r\nCheck\x20privileges\x20with\x20PRIVS\
SF:r\n")%r(HTTPOptions,59,"Login\x20with\x20USER\x20<name>\x20followed\x20
SF:by\x20PASS\x20<password>\x20or\x20ANON\r\nCheck\x20privileges\x20with\x
<cut>
Nmap run completed -- 1 IP address (1 host up) scanned in 75.504 seconds
```

At this point we have several options:

- Submit the signature to the URL provided and wait for the next version of Nmap. If responses were received from the probes sent, and the service is something that could be expected to be running on someone else's environment, this might be the best choice.

- Create a working match and/or probe statement, and submit that to Fyodor at fyodor@insecure.org.

For services that require a custom probe and can be expected to be found in another environment, this might be the best choice.

- Create a working match and/or probe statement for your own use. You might choose this option if your environment contains custom-written software running proprietary services or protocols. In this case it is necessary to know how to write the probes and matches to detect these proprietary services running on the environment being tested.

Regardless of which option you choose, it is very useful to know how to write your own probe and match signatures.

The nmap-service-probes File

The keywords contained in the `nmap-service-probes` file are listed in Table 1.

Table 1. `nmap-service-probes` keywords

Keyword	Format
`Probe`	`Probe <protocol> <probe name> <probe string>`
`match`	`match <service> <pattern> [version info]`
`softmatch`	`softmatch <service> <pattern>`
`ports`	`ports <portlist>`

Table 1. `nmap-service-probes` keywords (continued)

Keyword	Format
`sslports`	`sslports <portlist>`
`Totalwaitms`	`totalwaitms <milliseconds>`

Probes

A probe entry consists of the values shown in Table 2 below:

Parameter	Description
Protocol	TCP or UDP.
Probe name	Name of the probe (human-readable).
Probe string	String starting with a q, then a delimiter that will start and end the string sent. The string can consist of printable characters, as well as quoted unprintable characters and control characters in standard C or Perl notation.

Here are some example probe strings:

`Probe TCP NULL q||`
Send nothing, waiting the amount of time specified in totalwaitms.

`Probe TCP GenericLines q|\r\n\r\n|`
Send carriage return, newline, carriage return, newline.

`Probe UDP DNSStatusRequest q|\0\0\x10\0\0\0\0\0\0\0\0\0|`
Send the binary string 0x00 0x00 0x10 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00 0x00.

Matches

A match entry consists of the values defined in Table 3 below:

Parameter	Description
Service	Name of the service the pattern matches.
Pattern	A Perl-compatible regular expression to match the expected response for this service. This is of the format `m/regex/opts`.
Version info	A field specifying additional version information. This is of the format `v/product name/version/info/`. This can contain variables matched from the matching pattern, such as `$1`, `$2`, where the matching pattern contains `( )` matches. Any or all entries can be empty.

Nmap uses the Perl Compatible Regular Expressions `(libpcre)` library for evaluating regular expressions.

Perl regular expressions are documented at `www.perldoc.com/perl5.8.0/pod/perlre.html`

Here are some example match strings:

`match ssh m/^SSH-([.\d]+)-OpenSSH[_-](\S+)/ v/OpenSSH/$2/protocol $1/`
Match strings such as `SSH-1.5-OpenSSH-3.4p1`, reading the version string (3.4p1) and protocol (1.5) into the `$2` and `$1` variables, respectively.

`match ftp m/^220[- ].*FTP server \(Version (wu-[-.\w]+)/s v/WU-FTPD/$1//`
Match strings such as 220 FTP server (Version `wu-2.6.0`) and extract the version wu-2.6.0.

`match mysql m/^.\0\0\0\n(4\.[-.\w]+)\0...\0/s v/MySQL/$1//`
Match the version of MySQL 4.x from the binary response.

Soft matches

A soft match occurs when a service can be identified, but no additional information can be derived. A soft-match entry consists of the values defined in Table 4 below:

Parameter	Description
Service	Name of the service the pattern matches.
Pattern	A Perl-compatible regular expression to match the expected response for this service. This is of the format `m/regex/opts`.

Here are some example soft-match strings:

- `softmatch ftp m/^220[- ].*ftp server.*\r\n/i`
- `softmatch imap m/^\* OK [-.\w,:+ ]+imap[-.\w,:+ ]+\r\n$/i`

ports

`ports` is a comma-separated list of ports, as well as port ranges (e.g., 35067–35090) on which the service will commonly run. This is used to ensure that probing is done efficiently, and therefore the `ports` entry should follow the `Probe` entry in `nmap-service-probes`.

sslports

`sslports` is a comma-separated list of ports, as well as port ranges (e.g., 55522–55525) on which the service will commonly run over SSL. This is used to ensure that probing is done efficiently, and therefore the `sslports` entry should follow the `Probe` and `ports` entries in `nmap-service-probes`.

totalwaitms

`totalwaitms` is used to specify the timeout for a `Probe`. It is not needed unless the service you are probing does not respond immediately. If it is used, it should follow the `Probe` entry.

CSO and CISO - perception vs. reality in the security kingdom

By Melisa LaBancz-Bleasdale

The enterprise environment, like any large collection of individual roles and responsibilities, is based on a hierarchical structure that has pretty much gone unchanged since the first corporation opened it's doors.

There is the CEO, there may be a President - although often the CEO and the President are one in the same – and then there is a coterie of high level positions that are or aren't there depending on the structure of the company and their target market.

At any given time you can have a CFO, a CTO, a COO and a CIO. Add security to that in the form of a CSO or, in more specific cases a CISO, and that's a lot of salary, expertise and ego to keep track of. The job of keeping them all in line usually falls to the Board of Directors and on some level, the shareholders, but the people that make up the board, and the majority shareholders have decision making powers that defy easy explanation.

The blurred line between a security executive's role and their executive responsibility has been the recent focus of numerous high profile breaches that involved the downfall of corporations and their reputations. The minute all hell breaks loose in the papers, people are pointing fingers. Conspiracy theories abound - if you're an executive with the words "Security" or "Information" in your title, you're basically responsible for just about everything as far as the media and most of the industry is concerned.

After every major security breach comes an eventual corporate disclosure and these disclosures fuel endless speculation about who is at fault.

Perception can be a wonderful thing when taken into context with reality.

Rich Baich, author of "Winning As a CISO", is the Managing Director of PricewaterhouseCooper's Security and Data Management Practice in North Carolina. He was the CISO of ChoicePoint during the apex of that company's most challenging security situation.

A living bull's eye for the venomous press and media, he calmly navigated his way through hearsay, conjecture and finger pointing to arrive at a life-changing career decision. No longer comfortable with his vision being out of alignment with his reality, Rich, like many other security executives before and after him, realized that he could make a bigger contribution to security in the industry by leaving the top of the food chain. In an effort to make an impact, effect results and see the fruits of their labor realized, it's a decision more high level security executives are making nowadays.

Mike Assante, CSO of American Electric Power and a former U.S. Navy intelligence officer will be leaving his executive title behind at the end of July. In his new capacity, he'll be taking on a broader leadership role that will allow him to help both private industry and government discover protective measures and solutions whether through technology or process changes for a wide range of critical infrastructure protection needs.

"What was interesting about Rich [Baich] being the target of negative press is that it really caused a lot of CSO's to pause. It's always kind of been an assumption that non-traditional security events such as the one at ChoicePoint, is the crisis response and incident response role that we play today, and it's one that we should be playing. It made me wonder how I would I pre-empt those hard to define events. I felt that I should talk to the CEO and say, 'you are depending on me to handle these types of events right?' A lot of us stopped and asked ourselves those questions because of Rich's situation. It made a lot of us think," says Assante.

So who decides what a CSO and CISO should be doing with their time if not keeping the company's information secure?

As an active duty Cryptographer for the U.S. Navy, Baich is accustomed to dissecting impossible puzzles and making order out of chaos. True to his analytical nature, his is a viewpoint based on practicality, " When it comes to a major security breach, I don't think it's a matter of who's at fault because you really need to look at how people mitigate risk today.

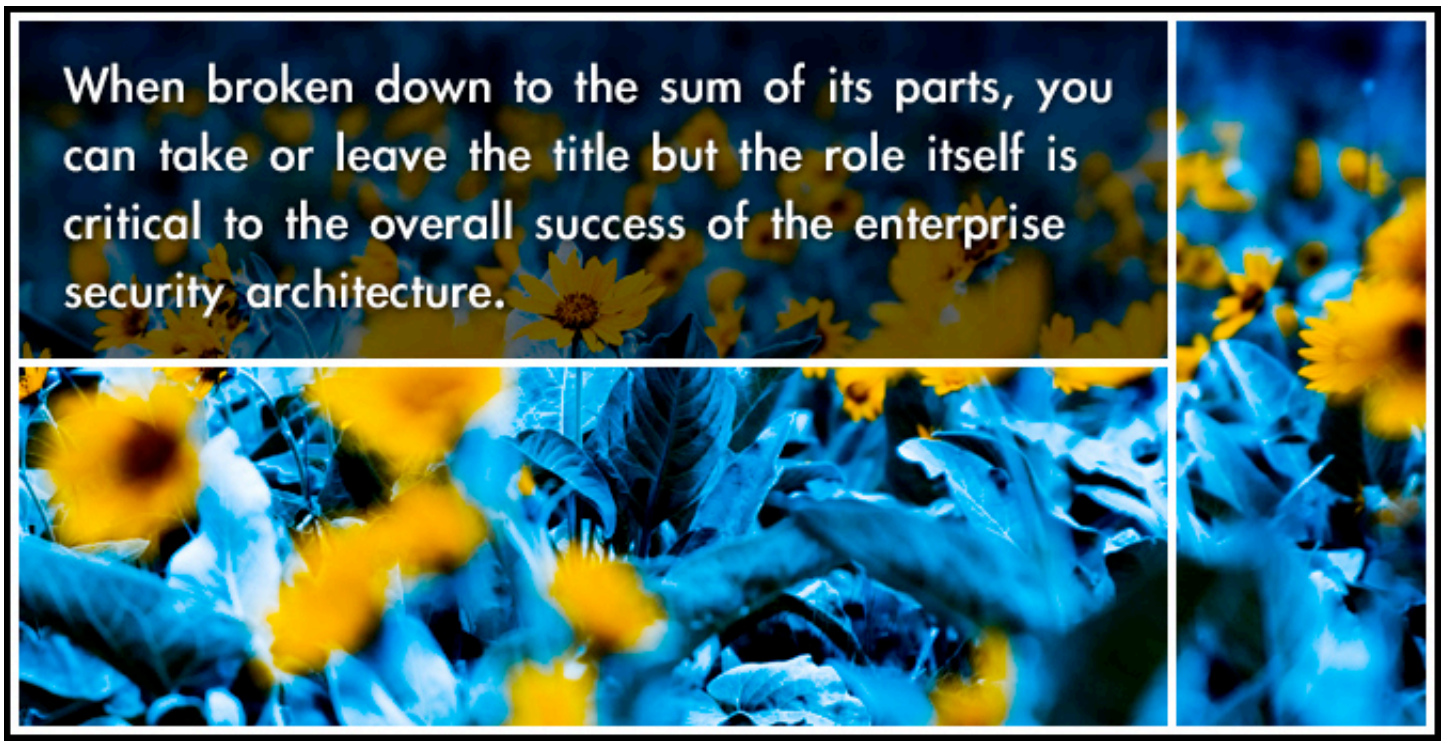
When you look at the role of the CISO, or any upper level security executive and how that role is evolving, you need to take into account their job description, empowerment and budget. It's questionable that the role should be reporting up under the CIO because the CIO's job is IT uptime and security is part of uptime but unfortunately the CIO has to make some very hard choices at times and managing risk isn't necessarily something that they're compensated for," Baich explains.

So who decides what a CSO and CISO should be doing with their time if not keeping the company's information secure? Who writes the job description? "Ironically, I think the way the job description has evolved is that either the CIO or the executive recruiters are asked to define it, at least as far as what functionality the CISO has. It's challenging because it's such an evolving role. Everyone knows what people do in sales, everyone knows what the job of marketing is, most people know what the job of the CIO is, but the tough part is that there's not a generally accepted job description for the CISO, at least not that I'm aware of. It's defined differently just about everywhere," says Baich.

It becomes apparent that those in the role of CISO and CSO are challenged by the ambiguities surrounding their role. Highly educated, they are often current or one time practitioners of complex security disciplines. They are hands on executives who are actively immersed in various forms of security. They have high ideals and work best when they're able to analyze and predict, recommend and implement.

However perfect for the jobs they're been hired to do, they face a difficult obstacle. Each has a direction they should be going, and each has constraints on where they're allowed to go.

"As a CSO, you should get a broad definition of your role up front, which is something I'm going to do if I ever walk into another CSO position. The role needs to have a much higher report-in, maybe to the CEO or CFO. These are the types of questions that I'm now prepared to ask because the role of the CSO is in a pioneering stage relative to how companies treat it. I think there's a lot of misunderstanding on both sides. In some cases, when somebody takes the role of CSO they make assumptions. They think, 'Hey, I'm here to provide leadership across the spectrum,' and unfortunately those assumptions don't pan out to be true when an event takes place," explains Assante.



"I believe that the appropriate role for a CISO - and it depends on different organizations in different industries – is one in which operational security is evolving. People are starting to see that it's a business function. If it's a business function then it should report to a business leader. What's a business leader? My interpretation is a COO or the President of the company. To help the future leaders step into a security executive role it's important to understand that this is not a technology issue. Technology is only 10% of it. Reporting under the CTO is not the right thing as that person is responsible for technology. The CIO is responsible for information operations. Well what is security and what is risk? It should be viewed as mitigation of operational risk. Although a lot of people will say that a lot of time and press is being wasted on defining organizational issues when we should be worried about fixing the security incident, I push back heavily on that opinion and say that it's tough to fix things if you don't have the right reporting structure and the right budget line to get it done," adds Baich.

It begs to be asked whether this role is really necessary within an organization that already has a CTO or a CIO. It brings into question whether this person is just a figurehead for a fully functioning IT group that may lack executive pomp and presence.

When broken down to the sum of its parts, you can take or leave the title but the role itself is critical to the overall success of the enterprise security architecture.

Baich says, “I want to make it clear that I don’t necessarily think that there needs to be a CISO, but there needs to be one senior executive understanding security, whether you call that a CISO, a CSO, a Chief Risk Officer or an Operational Risk Officer, it really needs to be more clearly defined in the industry. I do believe that the integration of the services that fall under that person’s title needs to functionally report in through the business leader and not a non-business entity.”

The need for a CSO/CISO may be greater than the role itself. To understand the role, it needs to be adequately defined and this depends on whether an organization is inclined to be proactive or reactive when it comes to security risk.

Assante explains, “After talking to several CSO’s, I got the sense that we’re all in the same boat because the operational responsibilities of the CSO or CISO are really ambiguous when it starts to deviate outside of normal security events. These positions were created in response to reoccurring security events that are having an impact on companies, especially in the CISO space.

These re-occurrences were definable, we were able to draw a circle around malicious code attacks, major virus attacks, DOS attacks, and companies decided they needed someone to take a leadership role to coordinate what happens across the organization and the business units to make sure the company could respond and quickly recover to get systems back on line. That was a well-defined thing. When people become CSO’s it’s very clear that they’ll lead in the defined area. Companies don’t respond to dynamic environments well because they’re always lagging behind. When you look at security risks - and I believe that’s a very dynamic environment - companies have a hard time saying ‘Yeah that’s a security risk too and therefore you should be the one driving that.’ Operationally, outside of the traditional security events, I believe the CSO and CISO’s responsibilities are very ill defined.”

Some companies are choosing to do away with the role altogether. This mindset strips away the singular security executive and spreads the defined duties amongst other functional groups. Dispersion often causes dilution. What happens when these functional groups are unable to form a unified opinion of what is and isn’t a risk to the company? Who underwrites the overall security process? Who makes recommendations based on risk assessment, monitors the changing threatscape, and meets with the business leaders to discuss security objectives? In a situation where the security of the company is dispersed, how can everyone keep track of what the other is doing?

The need for a CSO/CISO may be greater than the role itself.

Even though he’s left his title in the past, Baich believes that it’s critical to have such a role, especially within large, multi-national organizations, “A company has a responsibility to protect their employees. There are things like work place violence, background checks on employees and every company has computers so that creates access control requirements. Any company of any size is going to have an internet presence so there needs to be network security, policies, a process for dealing with customers who might utilize the company information in a way that causes embarrassment in the media or loss of data. You need a team that thinks along the lines of the adversaries. That’s how you get better at security. You can implement a methodology, the ISO standards, SAT 70’s and adhere to the current and future laws, but in the end what you want is individuals that spend their time looking at trends, understanding what criminals and potential adversaries are doing. This will allow the organization the ability to be proactive in mitigating future trends. You’ll be more prepared than if you didn’t have anyone in this post doing the lookout. “

Increasingly we see the effects of a reactive security plan. Savvy criminals harvesting hundreds of thousands of accounts for valuable information, companies whose internal employees have been selling them out, thwarting traditional security methods by virtue of being an insider. Security is a never-ending race of threat and response and no technology is going to solve every problem within the organization. Keeping on top of current threats and being prepared for future threats is a constant challenge.

With security threats on the constant rise, it would be easy to assume that all companies are behind the notion that security is a priority. The fact is that "priority" is directly tied to budget in all too many circumstances. Critical security decisions often compete with disparate organizational needs for funding. The theme of "budgetary constraint" comes up time and again but where is all the money going? Offers Assante, "Realistically a lot of CSO's have authority around a defined budget such as keeping our firewalls updated or installing IDS and things like that. Those are well-defined buckets and the CISO or the CSO has the ability to spend their budget as necessary. When you start getting into assessing risk proactively, whether assets or business processes, that's an area that's not really funded.

There isn't a flexible amount set aside for managing security, the money is there to manage ongoing security operations costs for protective measures that are currently in your arsenal, whether you have a sensor architecture for IDS or a guard force. That money maintains those and slowly enhances those if necessary. When you look at the broad definition of risk, it's calling for different protective measures and that's a whole new area that we don't have a budget for. The only time we get that kind of budget is when the government steps in and makes a requirement that says, 'You will do this', and then all of the sudden the money is set aside. But still, that's reactive. We find ourselves in an odd kind of place. "

At the end of the day, who is really responsible for security?

"I never really saw myself as a security executive, I see myself as a business executive responsible for security risk. In that sense, I want to understand all the business processes and I want to be able to contribute positive value to the organizations by being much more proactive. At the end of the day it is not my job to make risk decisions. It's the business leader's job to decide on acceptable or unacceptable risk.

If someone thinks something is an acceptable, then our job's over. If something does materialize after the fact, then we're called upon to be purely reactive and take care of it. If someone looks at a risk that we've identified and defined and they decide it's unacceptable, then my job is to help identify protective measures or options or maybe even re-engineer a business process so it would be less vulnerable to risk. That's all fine, but ultimately it is the business unit leader, whether it's an operating unit VP or the CEO, to make the acceptable risk decision. I always thought that as an executive, you didn't want me looking at yesterday's risk and restating the obvious, my job as someone senior in the organization is to look ahead and align where the company is going with how I think the threatscape will be changing," notes Assante.

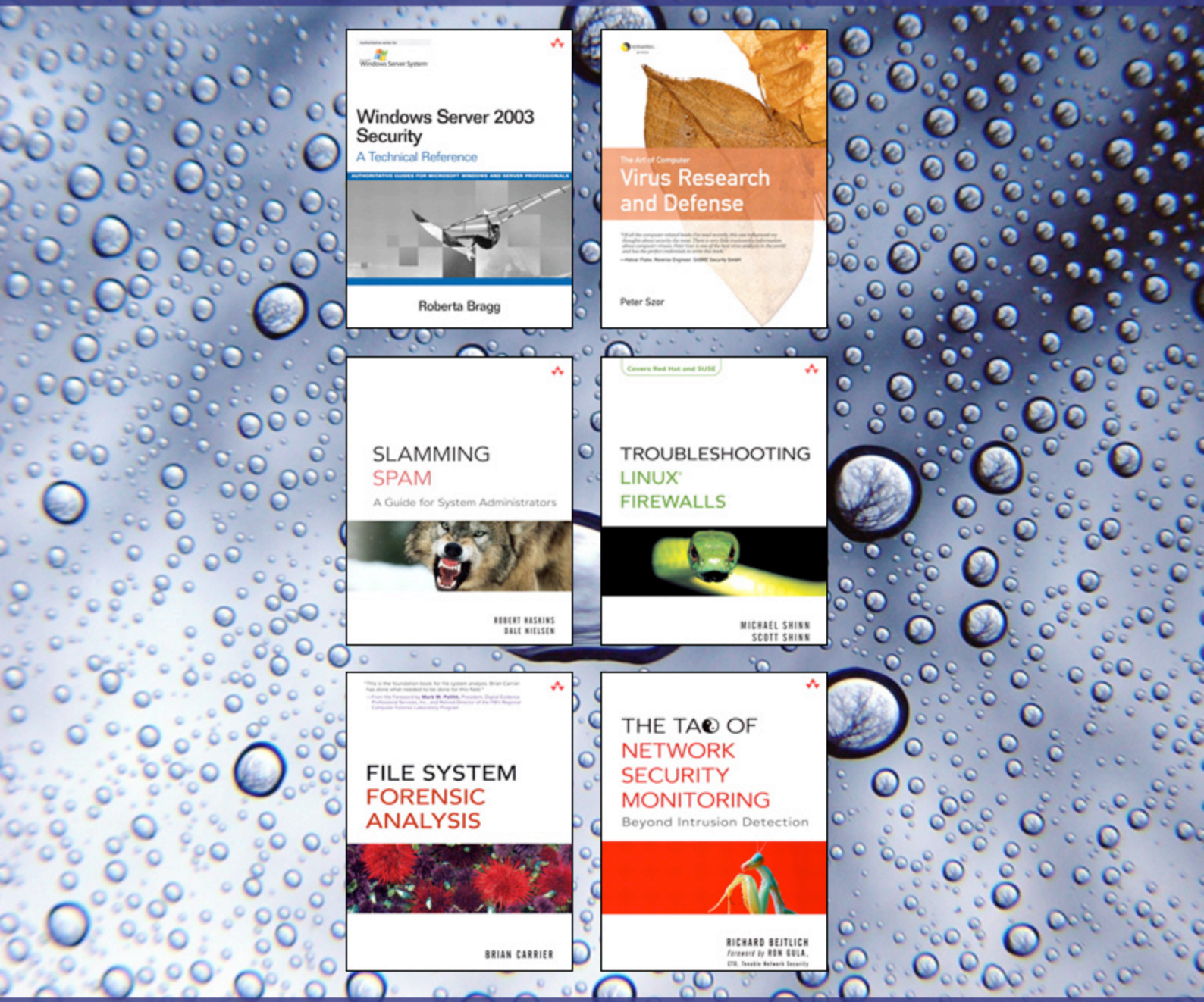
"You know you're winning in the role of CSO/CISO if you have employees sending attachments to the IT department saying 'I don't know who this is from or what to do with it.' That's creating a corporate culture but that cultural shift is a very difficult thing for an organization to understand and invest in. I really hope that the industry will take a hard look at this and find the appropriate place for a leader that is empowered to make a difference in a security executive role," ends Baich.

Melisa LaBancz-Bleasdale is a San Francisco area communications consultant and strategist specializing in the security industry. Her focus is on emerging and innovative security technology, current events and market concerns. Visit www.superheated.com to find out more about Melisa.

Want some knowledge? Enter the (IN)SECURE book contest!

ADDISON-WESLEY

We are giving away a copy of the following titles:



What do you have to do? You have to be creative and send us suggestions on what you would like to see in (IN)SECURE. The best submissions will be awarded with a book.

Go to www.insecuremag.com/contest to enter.

The contest ends September 26th 2005.

TaoSecurity Blog

http://taosecurity.blogspot.com

This blog, run by Richard Bejtlich is full of useful information about many security topics including forensics, incident response and network security monitoring. Bejtlich is the author of "The Tao of Network Security Monitoring: Beyond Intrusion Detection" and contributing author to a few other books. This blog is recommended reading for anyone looking for thoughts on specific problems.

Schneier on Security

http://www.schneier.com/blog

If you're looking for a great resource that covers general security information, this is definitely worth bookmarking. The blog, as you may have guessed from the title, is run by Bruce Schneier, an internationally renowned security expert. He is the author of "Applied Cryptography", "Secrets and Lies" and "Beyond Fear". Schneier also publishes a free monthly newsletter that you may be interested in, Crypto-Gram, with over 100,000 readers.

Martin McKeay's Network Security Blog

http://www.mckeay.net/secure

Another excellent blog comes from the mind of Martin McKeay, a network security professional. Packed with interesting insights into the world of computer security it delivers relevant information with some very good comments. What differs this blog from others is that also people with less technical knowledge can get some good pointers. A good example of that is the short but to-the-point "Security Primer for the non-technical" (http://www.mckeay.net/secure/archives/000013.html).

Security Awareness for Ma, Pa and the Corporate Clueless

http://securityawareness.blogspot.com

As the authors say themselves: "This blog gives computer security tips and tricks to government, corporations and home users." One of the people that contribute to this blog is Winn Schwartau, author of several computer security books and a leading expert on information security, infrastructure protection and electronic privacy. What you can find in this blog is posts covering all aspects of computer security and geared towards that beginner to intermediate audience.

When you are immersed in IT security (or IT in general for that matter), things hardly look simple. This is especially true for perimeter security gateway solutions targeting the enterprise market.

Answers here are rarely simple, and technologies are fragmented into separate solutions requiring more money and administration - despite marketing brochures claiming the opposite. New threats lead to new security technologies which in turn lead to increased administrative overhead and spending, while integration with existing infrastructure remains only a wish.

For example, firewalls are used for access control, gateway antivirus for virus protection, intrusion detection for various network and DoS attacks, URL filtering to filter unwanted web content, anti-spam to reduce junk mail, and so on. Often the solutions are located on different boxes or appliances and sometimes require network topology change. In addition, they usually have their own separate management consoles, the result being separate maintenance and administration, especially if different vendors are involved.

Perhaps it is because of this confusion that I'm caught by surprise when confronted with simple questions. I heard one such question from a colleague who's become only recently involved with IT security - it goes something like this: "Why can't you put a box inline with network traffic in order to transparently protect from all kinds of network based threats?" It's a simple question that deserves an answer.

The answer could be a new generation of firewall technology called Unified Threat Management (UTM). The term was apparently coined by IDC (http://tinyurl.com/eyqn2) and has recently become an often recurring buzzword in IT security.

In a nutshell, UTM appliances unify firewall, gateway antivirus, and intrusion detection and prevention capabilities into a single platform to protect from common network threats. The idea is to put all the functionality on a box (hardware unit) and insert it into the existing traffic flow without tinkering too much with the infrastructure in place. When you look at it, it seems the most natural and simple thing to do – the elegant solution.

UTM and network based threats

The history of IT security closely follows the evolution of network-based threats. In fact, IT security has followed an evolutionary path whereby each new threat would trigger a countermeasure. As threats have indeed multiplied in the past decade or so, so have the corresponding countermeasures, that is, technologies designed to address them. A sample list of those technologies would include:

- packet filtering and stateful inspection firewalls,
- Intrusion Detection (IDS) and signature-based Intrusion Prevention Systems with deep packet inspection,
- user authentication,
- IPsec VPN,
- clientless SSL VPN,
- gateway antivirus for mail and web traffic,
- anti-spam and e-mail content filtering,
- URL filtering (web access content filtering).

Once upon a time, things were much simpler. Firewall was virtually the only technology you use to solve perimeter access problems: limit accessible services, separate the public network from the internal segments, create a DMZ, and so on. Of course, the traditional firewall is now a mature technology. Its features typically include packet and stateful inspection, DHCP, NAT, PAT, PPPoE, etc.

The appearance of network layer attacks, such as IP spoofing, SYN flood, Ping of Death and other related DoS attacks, led to the next development: network intrusion detection and prevention systems (IDS/IPS). These systems would check IP packets against signatures of known attacks or traffic anomalies. Today, specialized IDS/IPS appliances are still offered in some cases, but are now largely an integrated part of firewall solutions.

Another technology that's become an industry standard is IPsec based VPN, used for both branch office and for mobile user encrypted connectivity. Transfer of confidential corporate data across the public network certainly posed a severe threat. And so, almost from its inception, VPN support had become a standard feature of firewall appliances.

Firewall and IPS technologies operate on the network layer.

THE HISTORY OF IT SECURITY CLOSELY FOLLOWS THE EVOLUTION OF NETWORK-BASED THREATS

However, the major proliferation of new threats in the last several years has been concentrated largely on the application layer of the protocol stack (also known as OSI Layer 7 – see Figure 2). These new threats are using most popular application protocols such as SMTP (e-mail), HTTP (web), Instant Messaging and others as vectors to spread and infect systems, degrade computer performance, steal sensitive data, consume network resources or simply reduce employee productivity. Examples of these threats include e-mail worms, application specific worms targeting vulnerable systems, spam, phishing attacks, inappropriate web content and spyware. The standard way to solve this new class of problems is to introduce gateway AV protection and content scanning by breaking the traffic flow and inserting specialized application proxies – for example for HTTP or SMTP traffic. Of course, since full traffic scanning on the application level is much more resource intensive than simple layer 3 analysis done by traditional firewalls, gateway content scanning for a particular protocol would have to be assigned to separate machines (or appliances). Precisely because of the performance bottlenecks and throughput issues, application level scanning has seen slow integration with other network security technologies.
However, thanks to Moore's law and the related advances in processing power, latest market developments indicate all the mentioned technologies are now being integrated into a single hardware platform: UTM (see Figure 1).

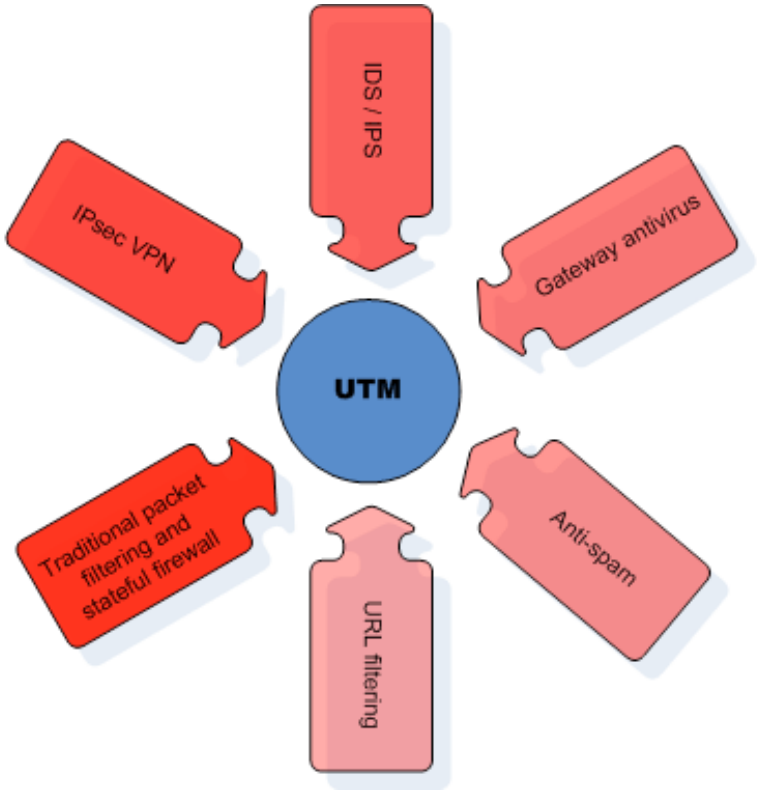


Figure 1

How does UTM work?

The best way to understand how the whole thing works is to look at the Open Systems Interconnecton model, better known as the OSI Reference model. The TCP/IP protocol suite roughly resembles this theoretical model, as shown in Figure 2.

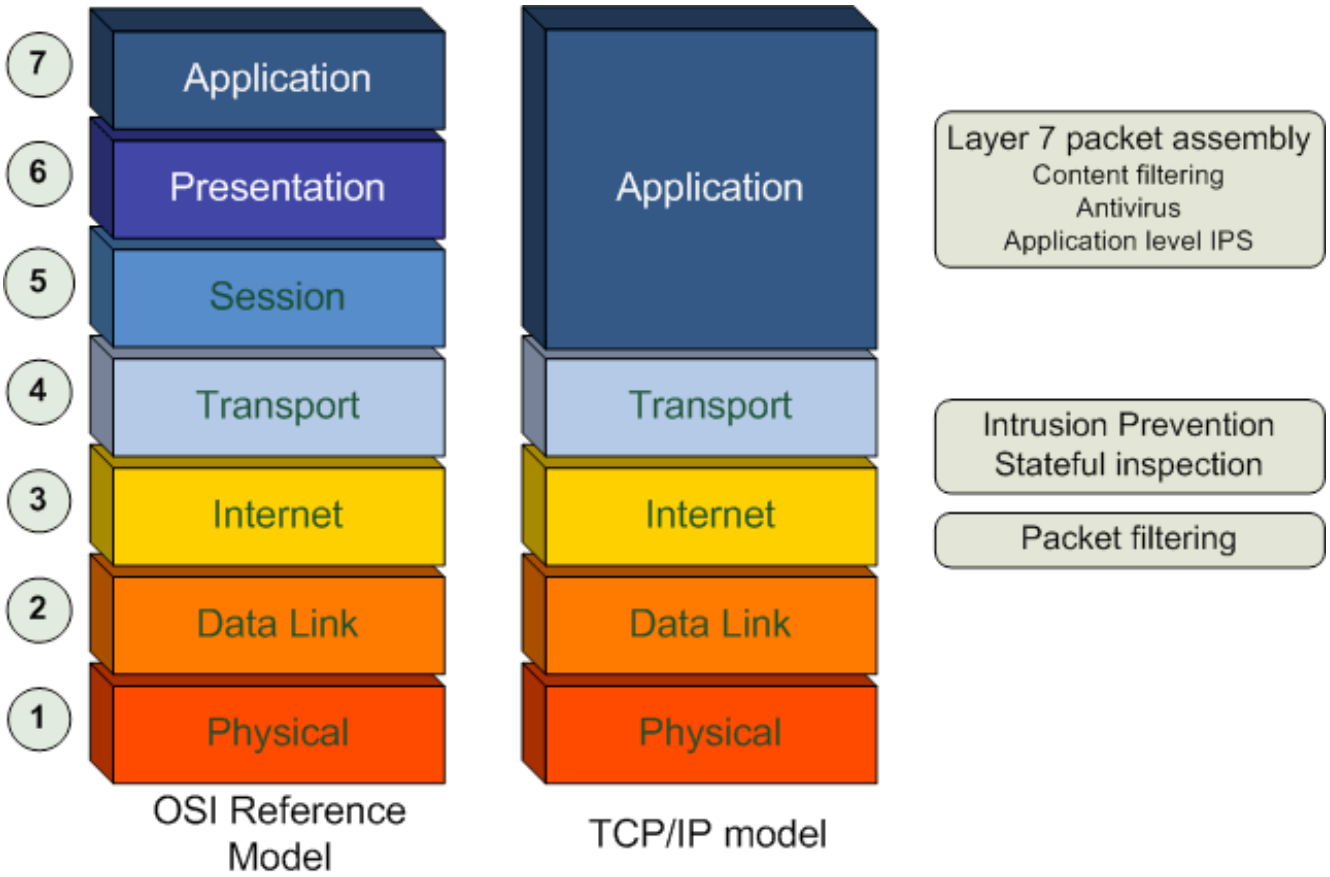


Figure 2

Traditional firewall defences work on the Internet layer, or OSI Layer 3. That's where packet filtering, stateful inspection and basic intrusion detection takes place. Essentially, Layer 3 scanning means analysis of IP packet headers only, without looking into packet contents – this is why it's not too much resource intensive.

But, as has been noted, most recent threats exploit weaknesses not on the network layer, but up the protocol stack to the Application layer, commonly referred to as OSI layer 7. Application protocols such as HTTP or SMTP are the most common vectors for threats such as viruses, worms, spyware, application level vulnerability exploits, spam e-mail, unwanted content, and so on.

However, application layer scanning is an entirely different process than simple Layer 3 filtering. Layer 7 threats such as an e-mail worm are often spread across many thousands of ip packets. What is needed is reconstruction of packets into application level objects. The whole process is roughly illustrated in Figure 3.

Let's suppose a connection has passed all the layer 3 filters. In order to do full application level scanning, traffic has to be proxied. It means packets associated with this connection are terminated, collected and then assembled into application level objects (for ex. an attachment or web page).

This involves data duplication into a memory or disk buffer and only then scanning for malicious content. If the resulting action is to deliver the content, the appliance deconstructs the application objects and rebuilds the original packets which are then proxied to the destination host inside the protected network. The result of this process is transparent traffic interception and in-depth scanning.

It is easy to see why performance drops dramatically and latency increases when attempting to do even simple application layer filtering. To compensate, one needs high-end hardware and optimised operating system environment, and this is something UTM appliances are based on.

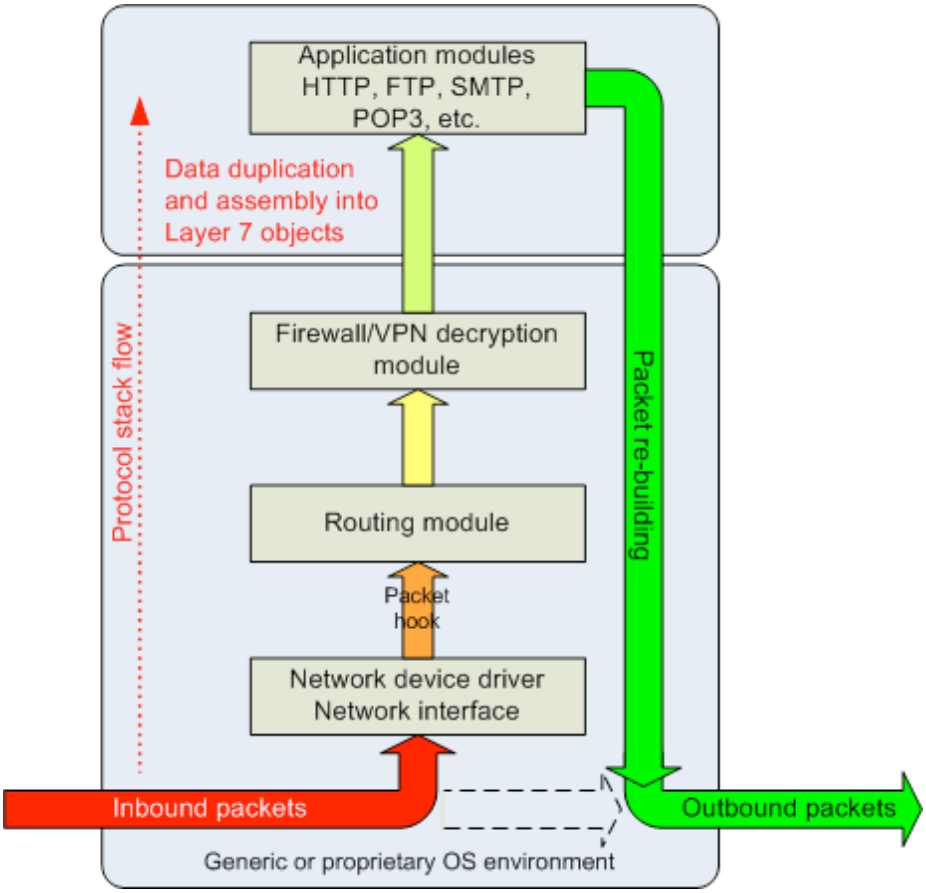


Figure 3

The Industry

The growing UTM market is currently dominated by a number of different vendors, each coming from a different background and core business. In addition, new vendors are entering the market continually. What follows is a brief summary of key participants in the UTM market.

Fortinet is a relative newcomer to IT security arena. The company was founded by the former NetScreen CEO in 2000 with a UTM vision almost from the very start (they claim to have the largest market share in this respect). Since mid 2002, Fortinet has been delivering its core product: FortiGate antivirus firewall. This appliance also offers VPN, IPS, content filtering and anti-spam services, fitting it into the UTM class. As expected, a wide range of models is offered ranging from SOHO&SMB users (FortiGate 60 or 100) to large enterprises and service providers (FortiGate 1000-5000 high-end appliances).

Internet Security Systems (ISS) is well known for its expertise in intrusion detection/prevention and vulnerability assessment. ISS is managing X-Force, the leading security research and development team responsible for a major part of new vulnerability discoveries. X-Force research results in security advisories on high-risk threats and immediate product updates to protect against the latest threats and vulnerabilities.

ISS has expanded its firewall and advanced IPS/IDS systems with application level inspection including antivirus, web content filtering, anti-spam, vulnerability detection, and so on – the end result being an integrated security UTM appliance called Proventia. It currently ships in three basic variants: M10 for small businesses and remote offices, M30 for branch office locations and medium-sized businesses, and finally M50 targeting large enterprises.

Secure Computing is specialized in building secure application-level proxies and firewalls. Its core product is the Sidewinder G2 firewall appliance which includes advanced proxy security scanning, antivirus, anti-spam (from Cloudmark), URL filtering (Smartfilter), IDS/IPS and IPsec VPN. Thanks to its advanced proxy inspection, Sidewinder has a reputation of a very stable and reliable firewall appliance.

Symantec is best known for its best-of-breed antivirus technology, although it has significantly expanded its portfolio making the company a one-stop shop for IT security solutions. Through Brightmail acquisition in May 2004, Symantec has also integrated superior spam filtering into its product line.

Symantec offers UTM functionality via the Symantec Gateway Security 5400 series which targets the enterprise market. The solution integrates full inspection firewall technology, protocol anomaly based intrusion prevention and intrusion detection engines, antivirus protection, URL-based content filtering, anti-spam, and IPsec VPN. Although Symantec offers similar appliances for the SMB market (Gateway Security 300 series), they have reduced functionality in terms of content filtering and anti-spam.

Other UTM vendors include SonicWall, ServGate, NetASQ (mostly on the European market) and of course Cisco, which recently entered the market with its Cisco ASA 5500 Series Adaptive Security Appliances.

No doubt there will be other interesting developments as the market matures in the near future - including strategic alliances between various specialized vendors, or even mergers and takeovers.

Final remarks

The convergence of gateway security solutions into a single unified appliance is certainly a positive step.

However, there are still some drawbacks to overcome.

UTM vendors can hardly provide "best-of-breed" solutions for all the technologies involved (antivirus, anti-spam, IPS, etc.).

When using separate solutions from different vendors, one can select the best product for the particular problem. However, when implementing a single UTM vendor, there is no other way but to accept the quality (for better or worse) of the integrated products in the single device. It will require some time before UTM grows into a solution that's able to seamlessly integrate with the existing infrastructure and provide the same range of features currently implemented in specialized products focused on a particular threat class.

In addition, it is difficult to make a fair assessment and comparative analysis of all the UTM vendors and products precisely because so many technologies are involved. To test them all against throughput, antivirus detection, vulnerability exploit detection, content filtering, logging and reporting capabilities would require extensive resources. This is probably the reason why there are not many comparative tests available for the general public. However, the reader may consult a recent test by Secure Enterprise Magazine (http://tinyurl.com/bqapw). Performance is another major concern. As mentioned earlier, application level filtering has put substantial load on processing hardware and the result is both reduced throughput and increased latency. Although this can be offset by deploying in clustered scenarios, the associated costs of the UTM could rise significantly. However, based on anticipated market developments as well as on increasing hardware capabilities, the benefits will ultimately compensate for all the drawbacks in the very near future. And so here's why UTM delivers:

- traditionally separated technologies are unified on a single box;
- less administration overhead with only one management system to worry about;
- transparent inline deployment with no topology and/or routing changes;
- centralized logging of all gateway security events.

UTM may not be the silver bullet of IT security, but it certainly is a better way to solve perimeter security than the currently fragmented security technologies.

Robert Buljevic is a security consultant working for Mack IT (www.mack.hr), a best-of-breed security solutions distributor in Croatia.

WINDOWS - Tor

http://www.net-security.org/software.php?id=253

Tor is a toolset for a wide range of organizations and people that want to improve their safety and security on the Internet. Using Tor can help you anonymize web browsing and publishing, instant messaging, IRC, SSH, and other applications that use the TCP protocol. Tor also provides a platform on which developers can build new applications with built-in anonymity, safety, and privacy features.

LINUX - Shorewall

http://www.net-security.org/software.php?id=40

Shorewall is a high-level tool for configuring Netfilter. You describe your firewall/gateway requirements using entries in a set of configuration files. Shorewall reads those configuration files and with the help of the iptables utility, Shorewall configures Netfilter to match your requirements.

MAC OS X - JellyfiSSH

http://www.net-security.org/software.php?id=605

JellyfiSSH is a simple bookmark manager for connecting to *NIX boxes like BSD/Linux etc via Telnet or SSH 1 or 2. You can set preferences for each bookmark including the terminal colours, fonts, window size, transparency and default login.

POCKET PC - Airscanner Mobile Firewall

http://www.net-security.org/software.php?id=573

Airscanner Mobile Firewall is a low-level, bi-directional, packet filtering firewall that examines all incoming and outgoing traffic to ensure it is permitted based on access control lists that are selected from a set of predefined filters, or from filters manually created by a user. The firewall parses packets as they come in (or go out) on the wire and matches the data against a ruleset of ports or IP addresses, URLs, etc.

If you want your software title included in the HNS Software Database e-mail us at software@net-security.org

As web application security gets more attention, it's rare these days to meet someone in the Information Security world who hasn't heard of SQL Injection. Unfortunately, not everyone who has simply heard of it or understands the basic premise of it really understands it to the point they should.

What's worse is that often times the risk is assumed to be limited to the data used on the website. In reality, the risk of SQL Injection can be quite extreme, and goes well beyond the data to the database server itself – and in some cases even beyond. You can fight complacency in an organization by demonstrating the true risk. In this article, Matthew Fisher will walk you through the basics of SQL injection in detail, and then move on to researching and demonstrating the true potential to SQL Injection.

Beginner's SQL Injection

SQL Injection is perhaps the most dangerous vulnerability in the world of web application security. It takes advantage of a mistake that's easy for a novice web programmer to make, is often extremely exploitable, and is quite common. SQL Injection relies on three factors:
a) a dynamic query, meaning one that changes based on input from the client.
b) a concatenated string query versus a parameritized query.
c) poor or non-existent validation of the input from the client used in the query.

An example of a vulnerable query would be this one:

```
sSql = "select ErrorMessage from
ErrorMessages where ErrorCode = "
& Request("ErrorCode")
```

This query has two portions to it - the static portion which never changes (marked in orange), and the portion that changes based on the error code passed to the page in the URL which is marked in blue. Since we see the ASP Request object right in the query, we know that it's using the information sent from the browser without validating it first. If the query simply used a variable for the ErrorCode that could not be presumed, since the validation could have been performed before that variable was used in the query.

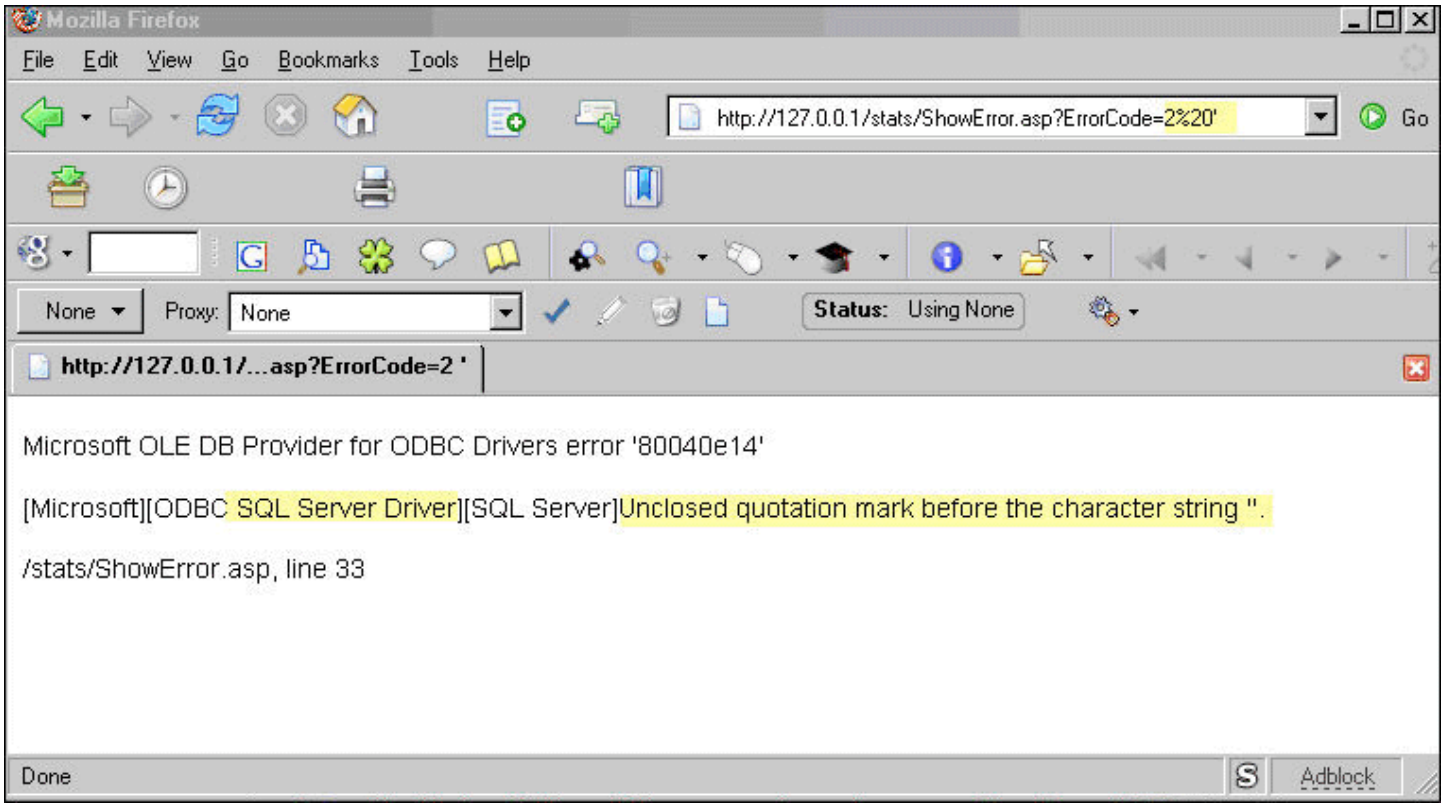
When the page that contains this query is executed, the actual query executed against the database will vary depending on the input received from the browser. For instance, with this request:

```
http://127.0.0.1/stats/
ShowError.asp?ErrorCode=2
```

The database query then becomes:

```
"select ErrorMessage from
ErrorMessages where ErrorCode = 2"
```

If the ErrorCode in the URL were changed to 3, then the query would accordingly change to become "... where ErrorCode =3". And that is exactly where the problem lies. The query simple changes according to whatever is in the browser. That's not much of a problem for valid input (such as "1","2", or "3") but it's a serious problem with invalid input. By throwing "garbage" into the url we can see how the query handles it. Here we enter a few basic reserved SQL characters or keywords and see the result:



The character entered, a simple single-quote (') is used by SQL to denote strings, so it expects them in pairs. The resulting ODBC error tells us quite a bit.

First of all, the mere existence of the ODBC error tells us that the programmer did not validate against the single quote (which is a very important character in SQL Injection). Instead, it simply appended it to the end of the query and ran it against the database. When the database complained about the query, the web application didn't even handle the exception but merely printed the results to the screen. In essence, the ErrorCode parameter is now a two-way pipe to the database server.

Of course, the information the ODBC error presents is quite helpful as well, telling you instantly the make of database server being used, and giving clues to the programmer's query. In some rare cases, you may actually even get pieces of the complete back-end query presented to you.

Now that we know the application isn't validating it's input very well, it's simply a measure of determining what is and isn't validate. We could try testing individual characters and keywords, or we can simply go for an attack and see if it makes it through. Attacking integers with SQL Injection is by far the easiest. Since the application is expecting a number – and not characters – the web page isn't going to wrap the input with single quotes.

This means that we don't have to try to escape out of the single quotes with out injection ; we can simply piggyback our query on top of the programmers with a UNION attack. The word UNION is reserved for SQL use.

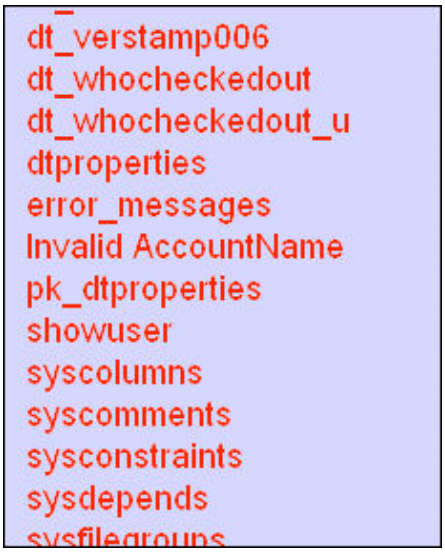
It's a special keyword that tells the database that you're going to give it two separate queries and you want it to combine – or 'union' – the two recordsets back into one. It essentially lets us directly piggy-back our query on top of the web programmers like this:

```
http://127.0.0.1/stats/
ShowError.asp?ErrorCode=2 union
select name from sysobjects
```

Remember that the vulnerable query just takes whatever is in the ErrorCode parameter and slaps it onto the query. Thus in this example, the query actually executed against the database becomes:

```
select ErrorMessage from
ErrorMessages where ErrorCode = 2
union select name from sysobjects
```

The database runs the two separate queries and returns one recordset, and with a little luck the results are printed to the screen. In our example page, the script writes the complete recordset back all at once. You can even see the appropriate error message that corresponds to ErrorCode 2 in it:



What essentially happened is that while the ErrorCode input was only supposed to be a parameter to a query, it ended up modifying the base query itself.

A UNION attack against an integer is perhaps the simplest SQL Injection attack there is. If the input were a string, the attack becomes more complex because now we have to deal with escaping out of the string properly and closing it properly – an issue further aggravated by things like MAGIC_QUOTES (don't rely purely on magic quotes though! Do your own input validation!) There are other situations that make it even more complex –for instance, the browser input going into the middle of the queries, etc. The more complex the underlying SQL query is, the more complex the injection. The point of this article, however, isn't to teach you everything there is to know about performing a SQL Injection attack; it's to raise your awareness of the risk.

The True Risk of SQL Injection

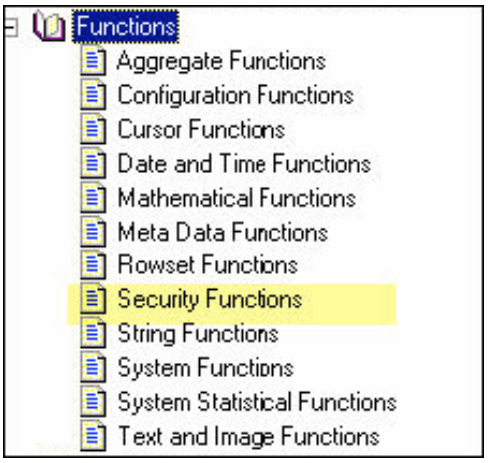
While almost every demo of SQL Injection you'll see shows selecting data from the database – which is indeed scary enough – the truth of the matter is that SQL Injection is actually much scarier than that.

Microsoft SQL Servers uses Transact-SQL, which is a remarkably robust language, and is made even more functional with the addition of system stored procedures and extended stored procedures. Data Manipulation Language – which is composed of commands like SELECT and INSERT - are a mere fraction of what's actually available to a DBA. The SQL language also contained an entire genre of commands called Data Definition Language which manage the objects in the database (such as tables, columns, etc.) DDL is used to create and manage the objects, and DML is used to populate the objects. In fact, a really good DBA will create, load, and manage their database entirely through SQL scripts if they want. The capabilities of SQL go beyond that though; there are Functions, System Stored Procedures, Extended Stored Procedures, Statements and Enhanced Statements, and more. The bottom line is that a DBA can almost completely manage then entire database server from t-sql scripts, and if they can, then so can anyone who can perform SQL Injection.

Researching SQL Injection

There are several resources available to learn the capabilities of Transact-SQL.

Of course, there are plenty of books you can buy, but there's also plenty of documentation that comes with SQL itself. Books online has a large reference section on the T-SQL language along with the various functions and procedures shipped.



The reference section is conveniently broken down into categories such as "Security."

You can also just query for system stored procedures and extended stored procedures, since they're maintained in sysobjects. Use wildcards to make your queries more effective, like this search for registry related procedures which returns a list of procedures for manipulating the registry (Just the top 10 are shown to keep the screenshot manageable).

```
1> select top 10 name from sysobjects
  where name like '%reg%'
2> go
 name

 ---------------------------------

 sp_MScopyregvalue

 sp_MSregistersubscription

 sp_MSunregistersubscription

 sp_vupgrade_registry

 xp_instance_regaddmultistring

 xp_instance_regdeletekey

 xp_instance_regdeletevalue

 xp_instance_regenumkeys

 xp_instance_regenumvalues

 xp_instance_regread

(10 rows affected)
```

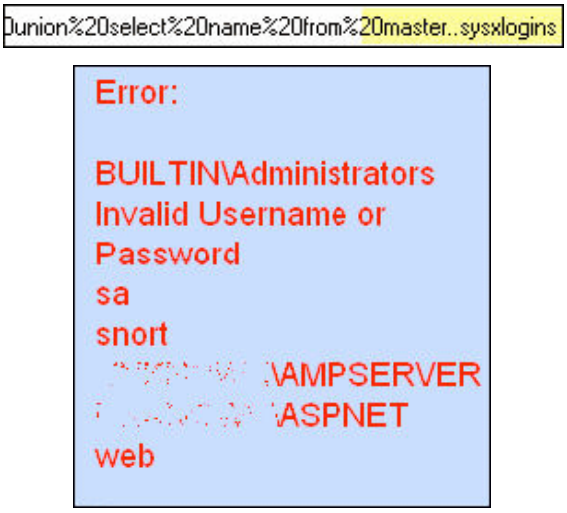
Some of The Possibilities

Just in the prior examples we see great promise as to the fun that can be had with SQL Injection... Not many folks realize that you can actually manipulate the Windows registry through SQL statements. Once you start researching the capabilities, you'll be amazed at what you can do with SQL Injection.

For instance, the security section of SQL Books Online describes the variable USER, which returns the user executing the command. This makes a great way to identify your access level on the database. Too many sites simply setup a DSN with full SA access for the web application to use, meaning that the web application – and thus your injections – run as the database owner. Selecting the USER variable tells you exactly who your injections are on the site:

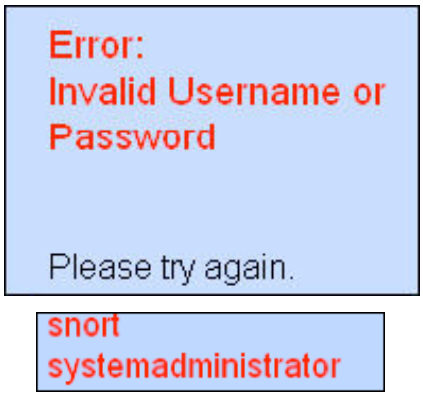
Of course 'DBO' is simply the role that the account is logged into – it's not the actual login ID itself. If you want to see the actual login id, in fact all the login ids for the server, you can select them from the SysXLogins table located in the master database. Since it's doubtful that the web application itself defaults to the master database (well, at least it's hopeful that it doesn't) you'll need to specify the master database in your query by prefacing the table name with "master.." and the site politely returns all the user logins to the database server. The password hashes are also displayed in this table although they are binary. Of course, you can also just fish for the connection string in the registry or text files; more often than not

the userid and password will be stored in plain text.



If you'd like to add your own login to the database or database server, you can do with with SP_ADDLOGIN. With this system stored procedure, you get to define the username, the password, and even the default database and language!

```
ErrorCode=1;
exec%20sp_addlogin%20'systemadmini
strator'%20,%20'mypassword','maste
r'
```



Notice that since stored procedures are executed, instead of performing a UNION here we simply finish off the script's hardcoded query with a semicolon and we don't get any feedback from the site. To confirm that the userid addition was successful, however, we simply query sysxlogins again.

You can also shutdown a database with the SHUTDOWN command: `;shutdown ;` The shutdown command isn't even a query, so it doesn't have to be preceded by a UNION, a SELECT, or any of the other commonly known (and validated) keywords used for SQL Injection. If it were injected into a string (versus an integer value) than one would have to do the usual escaping, but when injected into an integer it's remarkably stealthy (and short! Which makes it easy to use even when the script checks the length of the input). As it's name implies, it quickly and quietly shuts down the database. Of course, shutting down the database is only quasi-evil. If you really feel malicious you could use the DDL DROP command instead, which will remove the database from the server: `;drop database production_database`. As an added bonus if will physically delete it from the operating system's file system as well.

One area of SQL that's particularly fascinating is the ability to force a MS-SQL database to query another database. This can be done in several manners:

1. Pre-Defining a Linked Database with all necessary connection info. In this case, the connection is defined once in the server, then any actions using that connection simply reference it. There is also additional functionality available when using a defined Linked Database than when using other techniques.

2. Using Ad-Hoc Names. The OPENROWSET and OPENDATASOURCE command accept full connection settings in the query itself and can be used for ad-hoc queries of remote databases. As their name implies, they're intended for ad-hoc use only, hence the need to define the full connection. Each one is able to work with a variety of network provides, which include providers for MS-SQL database, providers for JET (to subquery Access databases), Excel and more, and each provider accepts it's connection string in different ways.

The website connectionstrings.com does a fantastic job of showing different possible connection strings for multiple database servers, providers and network types. Each network type will work for different scenarios – the most powerful, however is the Win32 TCP/IP library; this allows the creation of an ad-hoc connection to a remote database. The possibilities for abusing this are endless, ranging from dumping the entire injected database to a

remote database, to ports canning the back end network. While the database could be enumerated and dumped directly through the injection and resulting web responses, this could potentially generate lots of suspicious inbound traffic. This inbound traffic could potentially be reduced by dumping the database through an ad-hoc query to a hacker's own database. This could create lots more outbound traffic, but the odds are that the victim's egress filtering and monitoring is not as strong an their ingress monitoring. Additionally, ad-hoc names can be used to port scan from the injected database simply by specifying the IP and port to scan and examining the resulting query. The following injection tells the database to attempt a connection to 192.168.0.1 on port 80:

```
http://127.0.0.1/stats/
ShowError.asp?
ErrorCode=1UNION%20select%20*%20fr
om%20OPENROWSET('SQLoledb','uid=sa
;pwd=;Network=DBMSSOCN;
Address=192.168.0.1,80;
timeout=5','select blah from blah)
```

which results in this error message displayed in the web page:

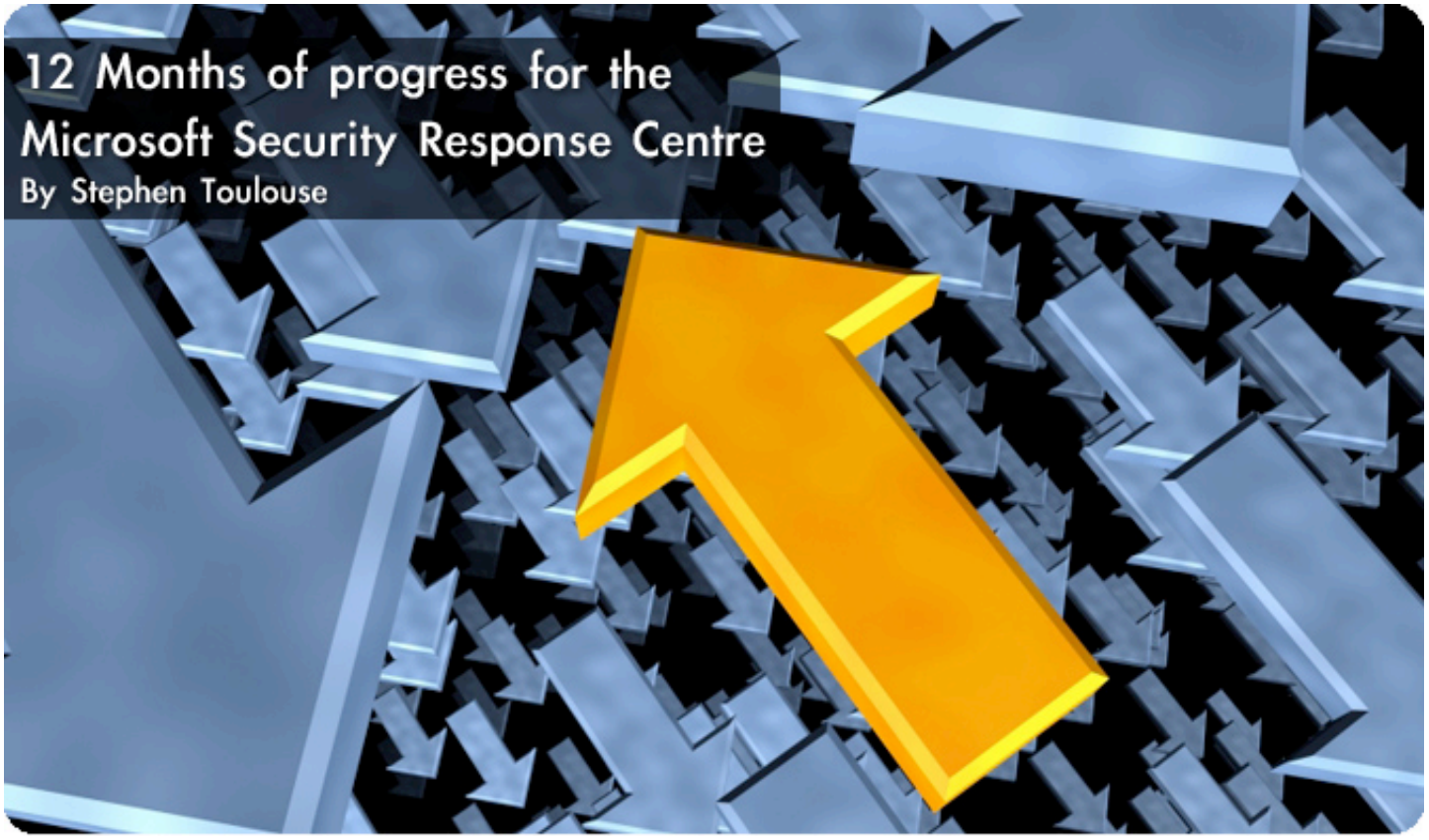
Attempting the same injected query, but now specifying port 22222, which is known to be closed on this system, we get the following error message:

The fact that two different error messages are returned makes it trivial to thus script out the injected port scans.

Summary

SQL Injection is a well known attack, and the general IT community is quickly becoming aware of the potential risk to their data. Few truly understand how dangerous it is, however, and some have even argued with me that they don't need to worry about SQL Injection if they don't store any private data. While at a minimum your data could be exposed, it's quite possible that the complete database server itself is owned and further network attacks will be facilitated. Defending against SQL Injection is easy with parameterized queries and good input validation. Unfortunately, it takes a while for "word to get out" and practices to improve – look at how long buffer overflow attacks have been known! By demonstrating just how malicious SQL Injection can really be, hopefully your clients' or your own organization will respond to this threat more aggressively.

Matthew Fisher is a Senior Security Engineer for SPI Dynamics, the expert in Web application security assessment and testing. He has held multiple certifications and has spoken on the topic of Web application security at numerous conferences for the Department of Defense, civilian Federal agencies, as well as the commercial sector.



12 Months of progress for the Microsoft Security Response Centre

By Stephen Toulouse

As the Internet has grown in popularity so too have threats against computer users; making it critical for individuals and companies to employ effective security strategies to protect their critical information. Microsoft created the Microsoft Security Response Centre (MSRC) to investigate, fix and learn about security vulnerabilities and to help keep customers protected from malicious attacks.

The MSRC is comprised of individuals, teams and entire groups around Microsoft; all dedicated to analysing, developing and delivering quality security updates, tools and prescriptive guidance to customers to help protect customers from security threats.

The last 12 months have been a particularly busy time for the MSRC, and, upon reflection, there are two activities that stand out to me. These were the releases of two major operating system service packs: Windows XP Service Pack 2 and Windows Server 2003 Service Pack 1. Windows XP SP2 was released in August 2004, and we are very pleased with the results so far. One of the key goals around this release was to get enhanced security features for Windows XP into the hands of consumers and enterprises, and so far more than 218 million copies have been distributed worldwide. This was an important security milestone for us. Many people put a lot of effort into this service pack and features like the firewall being on by default and the hardening changes made to Internet Explorer are already paying off and helping customers become more secure.

In Service Pack 1 for Windows Server 2003, the great features and security enhancements I mention above for Windows XP SP2 were also incorporated into this product, along with many other changes. We're particularly excited about the Security Configuration Wizard feature, which reduces the attack surface by querying users about the role their servers fill and then stopping all services and blocking ports that are not needed.

There is very significant work going on behind the scenes in the development cycle of current and all future software releases coming from Microsoft. Now, certain categories of software released from Microsoft now must go through the Security Development Lifecycle process which aims to provide customers with high quality software that is meticulously engineered and rigorously tested to help withstand malicious attack. We've published a lengthy whitepaper about this which is available at msdn.microsoft.com/library/default.asp?url=/library/en-us/dnsecure/html/sdl.asp. Essentially the SDL is a mandatory process that certain categories of Microsoft software must go through before it is released publicly. It helps us make sure that the software coming from Microsoft today has the latest security engineering advancements included in the code for the benefit of customers. It's a huge step forward for us to have this now as a formal process for our software. So far, we have used the SDL on Windows Server 2003, SQL Server 2000 SP3, and Microsoft Exchange Server SP3. Windows Server 2003 was the first operating released at Microsoft that implemented large portions of the SDL, and compared to Windows 2000, it had 63 percent fewer vulnerabilities in the first year.

While these developments cover significant activity on the product development side at Microsoft as a whole, the Microsoft Security Response Center has also made available a number of free tools and special guidance that can help customers become more secure.

Customers have told us that they want more prescriptive and timely guidance on security issues and Microsoft has responded to that feedback by continuously improving the security communications we deliver to customers.

This spring, we announced a pilot of a new offering, Microsoft Security Advisories, which aim to provide guidance and information about security related software changes or software updates. Microsoft Security Advisories, a supplement to the Microsoft Security Bulletins, address security changes that may not require a security bulletin but that may still impact customers' overall security.

In addition to the Microsoft Security Advisories, Microsoft has recently made available the Advanced Notification Program to help IT professionals plan their resources appropriately for deploying security updates. Three business days before the bulletins are released, general information is provided about the maximum number and severity of the bulletins. We've also enabled a Security Notification Service to alert customers to new bulletins and advisories as well as an RSS feed and MSN Messenger Alerts for security bulletins.

The MSRC also hosts monthly technical webcasts to offer customers additional support and guidance when deploying security updates and a regular Security360 webcasts to make prescriptive security guidance, education and training available to customers.

One of my favorite new things we've launched this year is the MSRC blog which provides insight directly from those working in the MSRC on recent security related news, announcements, activities and threat issues. This is a great way to get to know those folks that are working behind the scenes night and day to help protect customers. You can read all about at blogs.technet.com/msrc/default.aspx.

Another new tool released this year is the Malicious Software Removal Tool. This tool is updated each month to remove the most common malware threats that may be present on a user's machine. To be clear, this tool is not meant to be a substitute for good anti-virus software. However, it can help customers get back on their feet if they have been affected by any of the threats the tool is designed to remove. We have had a good response to this so far and look forward to continuing to update it each month to help customers.

In addition, Microsoft has come to offer customers a consistent and integrated set of new technologies that reduce the complexity and help customers better manage the update process for Microsoft software.

In June we announced the immediate availability of Windows Server Update Services (WSUS) and Microsoft Update (MU). WSUS is the update management component of Windows Server that enables mid-sized and enterprise companies to more easily assess, control and automate the deployment of Microsoft software updates. MU is a new service offered at no charge that gives customers everything they get through Windows Update (WU), plus high priority updates for more recent versions of Office and other Microsoft applications. It's a one-stop destination for updates that help make your computer more secure, up-to-date, and performing at its best.

ANOTHER NEW TOOL RELEASED THIS YEAR IS THE MALICIOUS SOFTWARE REMOVAL TOOL. THIS TOOL IS UPDATED EACH MONTH TO REMOVE THE MOST COMMON MALWARE THREATS THAT MAY BE PRESENT ON A USER'S MACHINE.

Only recently in July, we released the Microsoft Baseline Security Analyzer (MBSA) 2.0 which helps improve the security management process by detecting common security misconfigurations and missing security updates on your computer systems.

We also released the SMS 2003 Inventory Tool (SMS). This tool enables the detection and deployment of the latest security updates, update rollups and service packs from Microsoft; improved patch management through a more comprehensive and more widely-supported detection technology; broader detection support for more Microsoft products; and consistent product support across multiple detection technologies including parity with Automatic Updates.

The next 12 months will be as busy as these last 12 months have been. The security of our customers' computers and networks will remain a top priority for Microsoft, and Microsoft remains committed to building software and services that will help better protect our customers and the industry. It may never be possible to completely "cure" the security problem, but Microsoft and the MSRC is hard at work every single day, working in conjunction with the industry, with law enforcement, and with experts in government, academia and the private sector around the world to make the impact of malicious hackers as manageable as humanly possible. By building trust in computing our technology can be experienced in the way it was intended: to help customers accomplish what they need and want to do.

Stephen Toulouse is the security program manager of the Microsoft Security Response Centre (MSRC).

Michal Zalewski is a 24 years old computer enthusiast rather well-known among his InfoSec peers for finding a couple of noteworthy vulnerabilities, releasing several interesting research papers, and coding some small UNIX utilities for security researchers and system administrators. His hobbies include robotics, photography and mathematics.

You have been active when it comes to vulnerability research. What process do you go through when searching for a bug? Is it a planned activity where you search for something precise following certain rules or is the discovery accidental while you're using the software in question?

Zalewski: A bit of both. There are formal rules you can follow to perform a security audit of an application for a customer, but still, many vulnerabilities disclosed by researchers, including some most interesting, prominent and unique ones, are found because someone was bored, experimented in random ways with a particular application (or aimlessly browsed through the source code), and had some luck.
As they say, greatest discoveries are not announced by "eureka!" but "hmm... that's funny..." :-)

My public activity in the security field is purely a result of a hobby; publishing material of course helps with InfoSec employment - it provides essential training and good credentials - but I'm luckily not in a position where I have to force myself to find problems just to gain recognition. As such, I often pursue issues that are subtle, new and unique, and require unconventional approaches.

There's been a heated debate going on for years around the full disclosure of vulnerabilities. What do you see as the pros and cons when it comes to full disclosure? Is the disclosure being handled properly?

Limited- or non-disclosure policies are almost exclusively advocated by circles that benefit from withholding information from customers - vendors who are usually

negatively impacted by vulnerability-related PR buzz, and companies or individuals who make money or gain status by having advance knowledge of problems (such as commercial security software vendors or solution providers).

Secrecy never seemed to stop determined researchers with malicious intent from reverse engineering patches or obtaining leaked details from the "trusted" sources; it does, however, affect customer's ability to detect attacks, test susceptibility of his systems, or implement workarounds. Furthermore, it limits vendor's accountability for a failure to address problems in a timely manner (and multi-billion dollar companies like Microsoft or Oracle are known for sometimes taking months or years to fix trivial problems unless a considerable public pressure mounts).

I don't think there's much argument going on between informed customers with past experience in dealing with mainstream vendors, and bona fide researchers - open disclosure, like democracy, is bad, but we have no better options.

You've developed several tools over the years. The most popular to date is certainly p0f, a versatile passive OS fingerprinting tool. What features do you plan to add to p0f in the future?

I focus on features that are requested by users; as such, I cannot really tell in advance what is going to appear in next version, other than mentioning small improvements such as better database integration or support for more usage scenarios.

P0f would most certainly be a good starting point to develop a fully-fledged network mapper (correlate data from various passive fingerprinting mechanisms, map out network topology, record and accumulate data about commonly established connections; produce a detailed visual analysis of the network) - but that's a long shot.

You are the author of "Silence On The Wire". Many agree that the book contains uncommon computer security challenges not present in other titles. You have been dubbed the security expert that filled a gap in the security field. How do you feel about that?

I'm glad that the book was welcomed quite warmly by critics and other readers. I thought I might have something to add to the status quo, and that was the driving force behind the book. I'm far from believing that SotW was revolutionary or brilliant, though. It has its shortcomings, and is just a book, after all :-)

In "Silence On The Wire" you note that computer security for you is not a single problem but an exercise in seeing the entire ecosystem and understanding its every component. What advice do you have for the aspiring security professional that may be bound by formal education and certifications? What can they do to effectively expand their knowledge?

Just explore and learn on your own, plain and simple. Infosec knowledge devaluates and becomes obsolete alarmingly fast. You need to understand the fundamentals and think creatively to stay on top. Otherwise, regardless of certifications, you'd merely stay afloat :-)

I'm not denouncing formal education or certifications, though. Random exploration and self-teaching has its pitfalls: you are likely to focus on what you enjoy, and be ignorant of other problems; when you focus on interesting problems only, you might lack the ability to pay attention to mundane details, lack the patience, lack the business skills.

Based on what you know at the moment, do you expect more computer security awareness and better defense mechanisms in the future, or are the years to come only about to bring more problems?

I'm always expecting the worst :-)

The photo used with this interview was made by Michal Zalewski. You can find out more about him by visiting lcamtuf.coredump.cx.

Who's guarding your Exchange Server?

Fifi = a single anti-virus engine!

Buster = the real thing!

Get the leading email content security & anti-virus solution!

GFI MailSecurity

Email content/exploit checking, anti-Trojan & anti-virus

If you are serious about mail server protection, get the leading email content security, anti-Trojan and anti-virus solution, **GFI MailSecurity for Exchange/SMTP**, the only product to offer these unique features:

- **Multiple virus engines** – For better security
- **Email content & attachment checking** – Quarantine dangerous attachments and content
- **Email exploit protection** – Perform email intrusion detection and defense
- **HTML threats analysis** – Disable HTML scripts
- **Trojan & Executable Scanner** – Detect potentially malicious executables
- **Server-based anti-spam** – with the GFI MailEssentials bundle!

– Used by customers like NASA, Caterpillar, European Central Bank, MG Rover Group, Toyota & many more

Download your FREE trial from www.gfi.com/insec

GFI NETWORK SECURITY
CONTENT SECURITY
MESSAGING

tel: +1 (888) 243 4329 / +1 (919) 379 3402 | fax: +1 (919) 379 3402 | www.gfi.com/insec

OpenSSH is supplied with Macintosh OS X and runs much like it does for other Unix-like operating systems.

The primary differences and distinguishing features are:

- Some extra setup before the OpenSSH server, sshd, can be accessed by the outside world
- The software, which is a modified version of OpenSSH maintained by Apple
- Some important differences in the way sshd is configured by default, such as invocation and Kerberos support

Using the SSH Clients

The usual OpenSSH clients, ssh, scp, and sftp, work normally without any extra effort on your part:

```
# Log into server.example.com as user
smith
$ ssh -l smith server.example.com

# Copy myfile from your local machine to
server.example.com
$ scp myfile server.example.com:

# Run an interactive file-copy session
with sftp
$ sftp server.example.com
```

Using the OpenSSH Server

Before you can use `sshd` on Mac OS X, you'll need to enable the server and possibly open up the Mac's firewall.

In addition, you'll want to know about some configuration differences as compared to most other OpenSSH installations.

Enabling the Server

SSH server startup is controlled from the Sharing pane in System Preferences, under Services, as in Figure 1.

To enable sshd, select Remote Login and click the Start button.

Opening the Firewall

By default, the Mac OS X personal firewall will block SSH connections from the outside world. If you have this firewall enabled, you must manually permit SSH traffic through it. This is done from the Sharing pane in System Preferences, under Firewall, as in Figure 2.

Figure 1. Enabling the SSH server in System Preferences

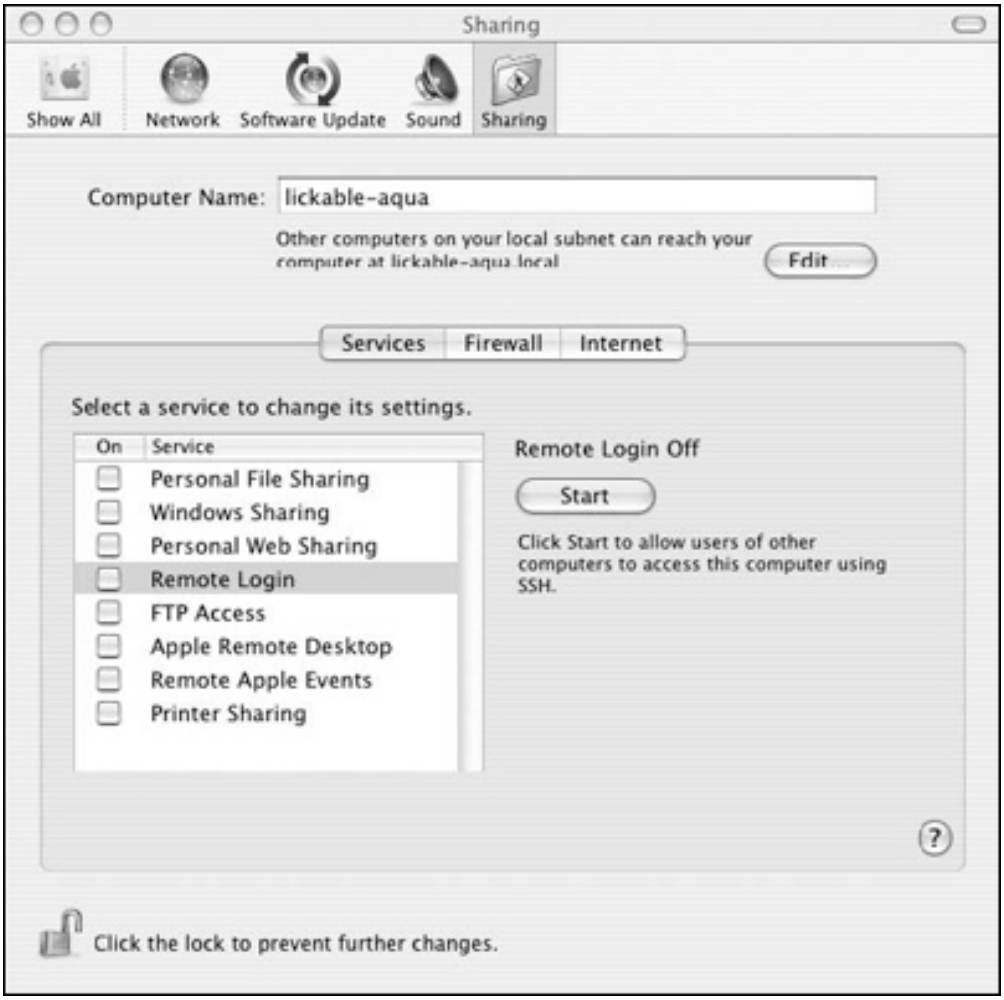
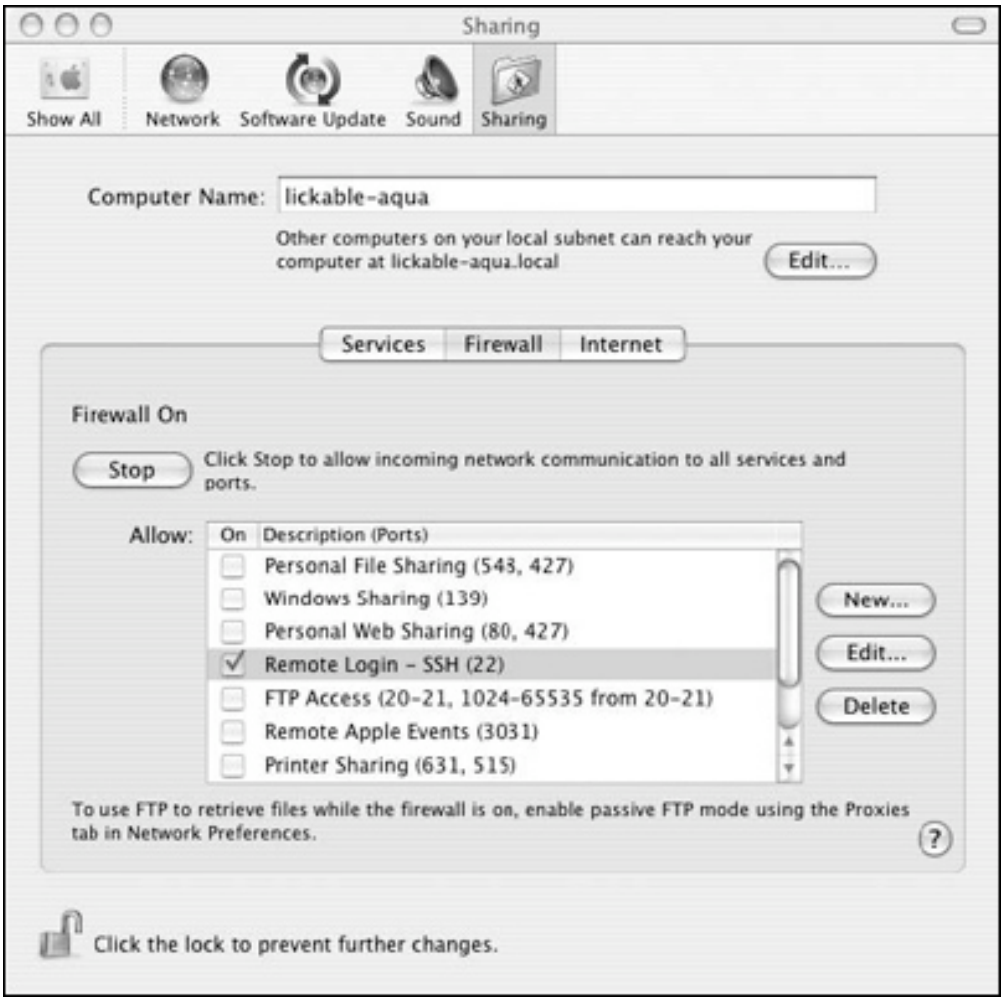


Figure 2. Opening a firewall hole for SSH in System Preferences



Control by xinetd

In most Unix-like operating systems, the OpenSSH server runs as a daemon, listening for SSH connections. On Mac OS X, however, `sshd` is controlled by the superserver daemon, `xinetd`. [5.3.3.2] Whenever an SSH client attempts to contact sshd on TCP port 22, `xinetd` notices the attempt and invokes a single instance of `sshd` (specifically, `sshd -i`) to serve that connection.

The `xinetd` configuration file for `sshd` is `/etc/xinetd.d/ssh`:

```
# /etc/xinetd.d/ssh:
service ssh
{
disable = no

socket_type = stream
wait = no
user = root
server = /usr/libexec/sshd-keygen-wrapper
server_args = -i
groups = yes
flags = REUSE IPv6
session_create = yes
}
```

Note the use of the wrapper script `sshd-keygen-wrapper`: it will generate new host keys if they are missing, as after a fresh OS install.

Server Configuration Details

On Mac OS X, the serverwide configuration files are found in the `/etc` directory instead of the more common `/etc/ssh`: for example, the serverwide configuration file is `/etc/sshd_config` rather than `/etc/ssh/sshd_config`.

The SSH software is a modified version of OpenSSH maintained by Apple; they backport security fixes to it whenever required.

Kerberos Support

The OS X OpenSSH build has protocol 2 Kerberos support for both user and server authentication, following the major Internet-Drafts on these (`draft-ietf-secsh-gsskex` and `draft-ietf-galb-secsh-gssapi`). It implements user authentication via the `gssapi` and `external-keyx` methods; it does not yet have the improved `gssapi-with-mic` method. In case a Kerberos-secured key exchange has been used for server authentication, the `external-keyx` method allows the `userauth` protocol to refer back to the previous Kerberos exchange for user authentication, skipping an unnecessary extra authentication phase.

This Kerberos support is also fully DNS-enabled, meaning it will find Kerberos authentication servers from information in the DNS if it is available. In a network of compatible and correctly configured Kerberos and OpenSSH servers, no extra configuration is needed for a plain OS X host newly attached to the network to use Kerberos for secure, single-signon client SSH connections. All that is required is to run:

```
$ kinit user@REALM
Please enter the password for user@REALM:
********
$ ssh user@host
```

Place the following lines into `/etc/krb5.conf` to relieve the user from having to specify the realm—and if the Kerberos principal and OS X account usernames are the same, then a simple `kinit` will suffice:

```
[libdefaults]
default_realm         = REALM
```

Instead of the command-line utility `kinit`, you can use the OS X GUI Kerberos utility: `/System/Library/CoreServices/Kerberos.app`

e-Secure Malaysia 2005
Conference & Exhibition

28 Sept - 1 Oct 2005
Putra World Trade Centre,
Kuala Lumpur

"Building e-Trust for National Competitiveness"

www.esecuremalaysia.org.my

Concurrent events

MyCrypt 2005
International Conference on Cryptology

Certified Information System Security Professional
BOOT CAMP

The International Platform for Information Security Professionals!

Your opportunity to meet and network with world class experts, corporate leaders, policy makers and security professionals

Trade Exhibition

- A diverse showcase of security technologies and latest products by more than 60 companies, including MSC Trustgate, Info Trek, IMS Asia, NTA-Monitor, Symantec Corporation, Advancenet Technology, Scan Associates, e-Lock Corporation, Computer Associates and more!
- Listen to free Product Presentations and build business network

Conference

- A comprehensive 6-track conference focusing on security technologies, CERT and incident response, network & application security, information security management
- Renowned speakers from U.S.A., Japan, and leading corporations in Asia!
- Free talks on 'Positive Use of the Internet' covering topics such as online safety guidelines, parenting the net generation, cyber crime, chat rooms, and more

MyCrypt 2005

- An international Conference on Crytology (MyCrypt 2005) presenting 19 best papers on Cryptology by researchers from Germany, USA, Switzerland, France, Belgium, Norway, Japan, Taiwan, Korea, Hong Kong, Australia, Canada and Malaysia

And More! Free product presentations and Conference Free Tracks – Open to the Public!

Enquiries

Exhibition : Ms Karen Dass
Tel: 03-7727 2828 • Fax: 03-7727 2566
Email: exhibition@esecuremalaysia.org.my

Conference : Ms Andrea Samuel
Tel: 03-7727 0619 • Fax: 03-7727 0614
Email: conference@esecuremalaysia.org.my

Jointly-Organised by: Ministry of Science, Technology & Innovation Malaysia; Malaysian Communications & Multimedia Commission; National ICT Security & Emergency Response Centre

Supporting Organisations: Ministry of Energy, Water & Communications Malaysia; Malaysian National Computer Confederation; Malaysian Administrative Modernisation & Management Planning Unit

Education Affiliate: (ISC)² SECURITY TRANSCENDS TECHNOLOGY™

MyCrypt Co-organisers

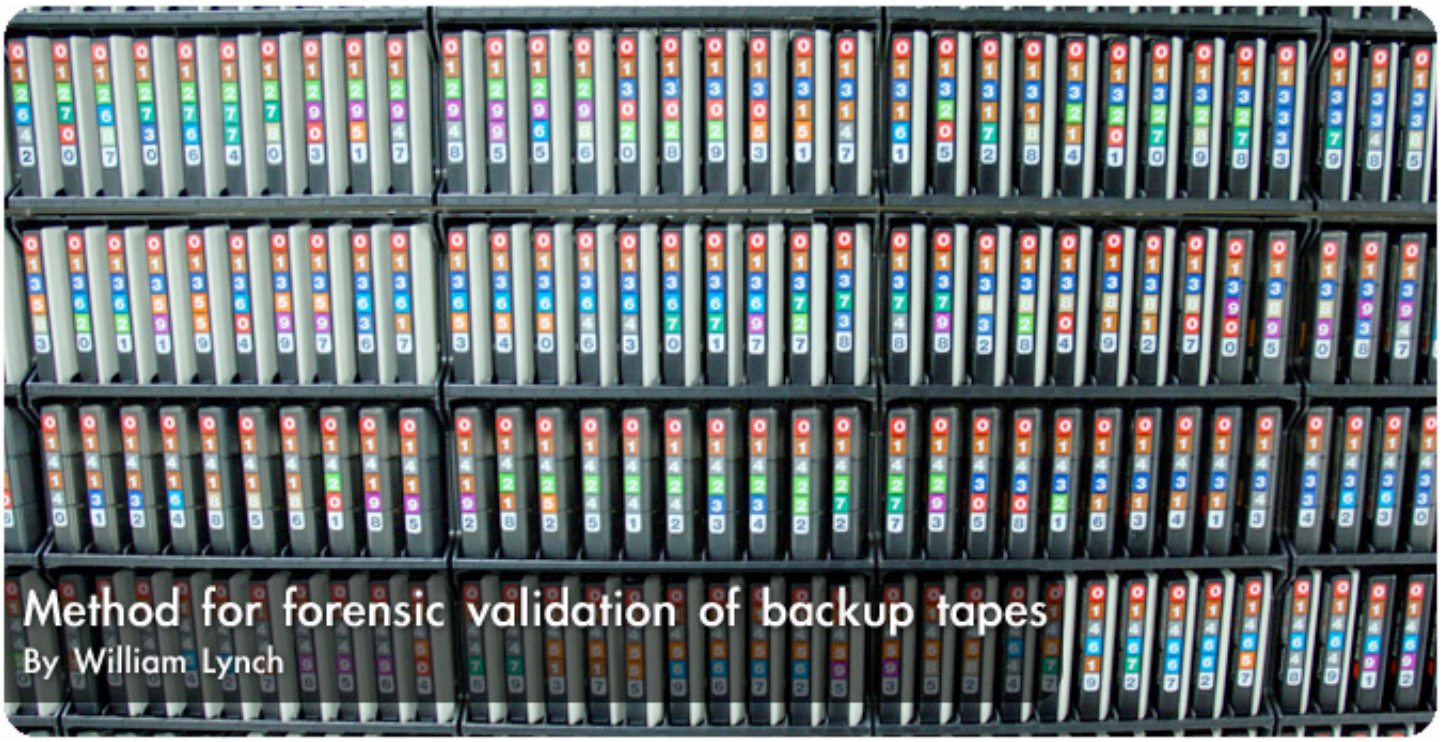
Gold Sponsors: Computer Associates

Silver Sponsors: symantec.

Lanyard Sponsor: NTA

Official Online Media Partners: CNET Asia; ZDNet Asia

Media Partners: virus bulletin; PC; ConnectWorld; (IN)SECURE



Method for forensic validation of backup tapes

By William Lynch

Computer forensic investigations usually involve the recovery of data from mundane data storage devices such as hard drives, but often involve recovery from less common devices such as magnetic tapes.

Regardless of the media, the information recovered cannot be considered forensically sound unless the media is left unchanged by the recovery method. Validating that the media remains unmodified is usually handled by some sort of checksum verification process.

This article describes a method for verifying that a magnetic backup tape remains unchanged after a data extraction.

The Scenario

Data needed to be extracted from several users' Exchange mailboxes as part of a fraud investigation. The mailbox data was to be extracted from restores of the Exchange data stored on archived monthly backup tapes created with ArcServe.

The actual restore and recovery methods are beyond the scope of this article, which is limited to describing a method for integrity validation of the backup tapes. The data extracted from the tapes may be used in a court of law, and as such there needs to be a way to validate that the data on the tapes remains unchanged after extraction, such that the extraction process could be repeated with identical results.

Integrity Validation Process

In theory, validating a backup tape is identical to validating a single file. The process involves using a one-way hashing function such as MD5 or SHA-1 to calculate a fixed-length checksum of file.

The hash generated is intrinsically sensitive to minute changes, such that changing a single character in a multi-megabyte text file will result in a completely different checksum. Thus, by creating a fingerprint checksum of the backup tape before the restoration process begins and another after the restoration process is complete, the two can be compared to see if the tape data has been altered. If the checksums are the same, then the data on the tape remained unchanged, but if the checksums are different, then the data on the tape may has been altered and may be tainted for forensic use.

How is the validation process changed if the item under inspection is not simply a single file, but instead a whole device, such as a hard drive or magnetic tape?

Under Linux, these devices are represented as single files, and can be treated as such to some degree.

For example, using the `md5sum` and `dd` commands against the /dev/hdd device would calculate the MD5 checksum of the entire /dev/hdd hard disk. In its simplest form, the command might look like:

```
dd if=/dev/hdd | md5sum
```

The `dd` command is used to collect raw data from the /dev/hdd device (until the end of the device is reached), which is then piped to the `md5sum` command for checksum calculation. If any data on the disk were to change, repeating the process would yield a different checksum. Unfortunately, when working with a tape device, the process is not as simple.

Complications from Tape Devices

Data is read from tape devices differently than it is read from hard disks. With a hard disk, data can be read in a random fashion, such that any address on the disk can be read at any given time.

However, tapes must be read sequentially from the beginning of the tape to the end of the tape. The caveat of this is that the `dd` command will only read data until if reaches and end-of-file (EOF) marker.

A tape consists of a variable number of data "chunks", each of which ends with an EOF marker. Thus, without modification, the `dd` command will only read the first chunk of data on the tape, which is typically only the tape header.

Validating that the header remains unchanged after the extraction process is useless because it says nothing about the data on the remaining 99+% of the tape.

Just because it isn't possible to validate the entire tape in a single pass, doesn't mean that the entire tape can't be validated.

Work-Around Solution

Just because it isn't possible to validate the entire tape in a single pass, doesn't mean that the entire tape can't be validated.

One way to work-around this problem is view the tape as a collection of chunks. If each chunk can be validated, then the entire tape can be validated. The `mt` (magnetic tape) command will help in positioning the tape at the appropriate chunk.

If the chunks are to be validated programmatically, then there must be a way to identify the total number of chunks contained on the tape in order to identify the end of the tape.

Unfortunately, there doesn't seem to be an easy way to handle this at the onset, so another work-around is required such that if the `dd` command reports an error, the end of the tape has likely been reached, so the program will need to check on the fly if the next chunk is the end of the tape. If it is, the pass is complete, but if not the tape needs to wind back to that start of that chunk, again using the mt command.

Validating the Validation Process Itself

What's the assurance that the validation process itself doesn't alter any data on the tape? The only way to handle this is with a two-pass process.

If the data is unchanged during the validation process, then the checksums of each chunk should match for each pass. If they do, then we can be reasonably certain that our validation process isn't altering any of the data on the tape.

A secondary validation process can be obtained by utilizing the number of records read out by the `dd` command. Although the md5sum validation routines are more than adequate, the nature of forensics work makes it desirable to have multiple validation checks. In this case, should the number of records read out per chunk change between passes, this would be noted in the script output, which would flag the validation as a failure. Note that this is simply a backup check process, as the md5sum collections would undoubtedly be different if the number of records read by `dd` changed between passes.

Collecting the Results

Because there could be dozens of chunks of data on each tape, it would be undesirable to hand-validate each of these for two passes. As a short-cut solution, the checksum of each chunk of data could be stored in a file and a super-checksum can be made of this file, which would be representative of a checksum of the entire tape. The super-checksum created this way is an abstraction and not a true checksum of the entire tape, but fulfills the same purpose with a manageable size.

Practical Implementation

A reference script, tapeverify.sh, written as a bash shell script is included at the end of this article. This script was developed to handle the Exchange recovery scenario described before. It was developed on Knoppix (which easily identified the SCSI adapter and tape drive in the system). This script successfully validates DLT tapes written by ArcServe NT and has not been tested against other media types or backup sessions. When run, the output of the overall verification is displayed at the terminal and an archive is created in the current directory which holds the output and all intermediary files. For a 40 GB tape, the verification process takes approximately 6 hrs on a PIII-600.

The practical implementation example was only tested against DLT tapes containing Exchange data created by ArcServe. While it should be equally effective against other data and media, it has not been tested for such purposes.

The output created is a compressed tar file located in the current directory. The contents of the archive are as follows where [pid] indicates the process ID of the script and [pass] indicates which pass (1 or 2) to which the item corresponds:

- tapedata.[pid].[pass].records - The results of the records out from dd
- tapedata.[pid].[pass].md5sum - The results of the md5sums from each chunk
- tapedata.[pid].[pass].log - The logged output of all commands during execution with timestamps
- tapedata.[pid].verified - The final results comparing the records out and md5sums for each pass

Upon completion, the program displays the contents of tapedata.[pid].verified, which looks like the following:

```
#####  Individual MD5 Verification  #####
99be51e3f2be8936a9bf9f64308f34b0  /tmp/tapedata.23981.1.md5sum
99be51e3f2be8936a9bf9f64308f34b0  /tmp/tapedata.23981.2.md5sum
#####  Individual Record Count Verification  #####
f3d2db40509ed0e0e4bc513bdcc4026a  /tmp/tapedata.23981.1.records
f3d2db40509ed0e0e4bc513bdcc4026a  /tmp/tapedata.23981.2.records
```

So long as each pair matches, the tape is validated. The same process can be used to re-validate the tape after a data extraction process, i.e. restoring the data on the tape to a system. If the post-restore validation matches the pre-restore validation, then the tape can safely be said to have been unchanged by the restore process.

Summary

Computer forensics investigations are somewhat useless if the process of investigation changes the original evidence media. Typically, a hashing checksum is used to validate the media before and after the data extraction procedure to ensure the media has not been modified. However, standard verification techniques do not lend well to magnetic tapes, and different techniques are required for the validation process.

This article covered a process in which a tape is separated into an arbitrary number of chunks and each chunk is independently validated.

The reference script, tapeverify.sh, is included on the following page.

William Lynch is Senior Consultant for Computer Task Group's Information Securty Services Practice.

```
#!/bin/bash
# This is a poor man's tape data verification script
# Written 20050511 by FWL (bill.lynch@ctg.com)
#
# BSD License Follows
#
# Copyright (c) 2005, William Lynch, CTG
# All rights reserved.
#
# Redistribution and use in source and binary forms, with or without modification,
# are permitted provided that the following conditions are met:
#
#  * Redistributions of source code must retain the above copyright notice, this list
#  of conditions and the following disclaimer.
#  * Redistributions in binary form must reproduce the above copyright notice, this
#    list of conditions and the following disclaimer in the documentation and/or
#    other materials provided with the distribution.
#  * Neither the name of Computer Task Group (CTG) nor the names of its contributors
#    may be used to endorse or promote products derived from this software without
#    specific prior written permission.
#
# THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND
# ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED
# WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED.
# IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT,
# INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT
# NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR
# PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY,
# WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE)
# ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY
# OF SUCH DAMAGE.
#
# This script was created against DLT tapes holding data written by
#  Arcserve NT.  I cannot speak to its viability against other media or
#  data, though I suspect it would work.
# Loop through the tape twice calculating the MD5 sums of each recordset
for i in `seq 1 2`; do
  echo "Beginning Pass $i at `date`"

  # Write a header for the logfile
  echo "#####  Tape Verification Logfile Pass $i  ####" > /tmp/tapedata.$$.$i.log

  # Prepare the tape and display the tape details
  mt -f /dev/nst0 rewind >> /tmp/tapedata.$$.$i.log 2>&1
  mt -f /dev/nst0 status >> /tmp/tapedata.$$.$i.log 2>&1

  # Initialize variables
  ENDTAPE=0

  # Process the MD5 sums of each backup session

  while [ $ENDTAPE -ne 1 ]; do

   # Timestamp
   date >> /tmp/tapedata.$$.$i.log 2>&1

   # Collect and calculate the MD5 sum of this backup session
   dd if=/dev/nst0 bs=64k 2>> /tmp/tapedata.$$.$i.log | md5sum >> /tmp/tapedata.$$.$i.log

   # Display the tape address
   mt -f /dev/nst0 tell >> /tmp/tapedata.$$.$i.log 2>&1
   LOCATION=`mt -f /dev/nst0 tell | awk '{ print $3 }' | awk -F. '{ print $1 }'`

   # Check to see if we are at the end of the tape
   dd if=/dev/nst0 bs=64k count=1 > /dev/null 2>&1
   ENDTAPE=$?

   # Return to the last location
   mt -f /dev/nst0 seek $LOCATION
```

```
  done

  # Extract the MD5 sums
  echo "#####  Individual MD5 Extracts  ####" > /tmp/tapedata.$$.$i.md5sum
  grep '\-$' /tmp/tapedata.$$.$i.log | sort -u >> /tmp/tapedata.$$.$i.md5sum

  # Extract the record count
  echo "#####  Individual Record Log  ####" > /tmp/tapedata.$$.$i.records
  grep 'records out' /tmp/tapedata.$$.$i.log | sort -u >> /tmp/tapedata.$$.$i.records

done

# Rewind the tape
mt -f /dev/nst0 rewind > /dev/null 2>&1

# Calculate the final MD5 sums
echo "#####  Individual MD5 Verification  #####" > /tmp/tapedata.$$.verified
md5sum /tmp/tapedata.$$.1.md5sum >> /tmp/tapedata.$$.verified
md5sum /tmp/tapedata.$$.2.md5sum >> /tmp/tapedata.$$.verified
echo "#####  Individual Record Count Verification  #####" >> /tmp/tapedata.$$.verified
md5sum /tmp/tapedata.$$.1.records >> /tmp/tapedata.$$.verified
md5sum /tmp/tapedata.$$.2.records >> /tmp/tapedata.$$.verified

# Display the results to the console
clear
echo "#####  Tape Verification Results  #####"
cat /tmp/tapedata.$$.verified

# Package up all the working files
OLDDIR=`pwd`
cd /tmp
for i in `seq 1 2`; do
 echo "tapedata.$$.$i.records" >> /tmp/tapedata.$$.tarfiles
 echo "tapedata.$$.$i.md5sum" >> /tmp/tapedata.$$.tarfiles
 echo "tapedata.$$.$i.log" >> /tmp/tapedata.$$.tarfiles
done
echo "tapedata.$$.verified" >> /tmp/tapedata.$$.tarfiles
tar cvzf $OLDDIR/tapedata.$$.tar.gz -T /tmp/tapedata.$$.tarfiles > /dev/null
rm `cat /tmp/tapedata.$$.tarfiles`
rm /tmp/tapedata.$$.tarfiles
cd $OLDDIR

exit 0
```