

The Open Proxy Problem

**Internet2 Members Meeting
Arlington VA, Wednesday, April 9th, 2003**

**Joe St Sauver, Ph.D. (joe@oregon.uoregon.edu)
Director, User Services and
Network Applications
University of Oregon Computing Center**

<http://darkwing.uoregon.edu/~joe/proxies/>

I. Introduction

My interest in proxy servers

- My interest in proxy servers goes back many years now.
- For example, I brought up the first Squid box at the University of Oregon (then a Sparc 5, wow! :-)), and I also encouraged deployment of caching web proxies at other Oregon University System schools and K12 sites statewide served by Oregon's OWEN/NERO network.

Early exposure to casual attitudes about web cache security

- I've also done beta testing of commercial cache boxes. My interest in proxy server security really dates from that testing work.
- While testing one particular commercial cache appliance, I noted it had **no** access controls at all; my feedback on that point to the vendor was blown off, and I was told "don't worry, our caches will always be deployed behind a firewall." No, they weren't.

What was old became new again

- My interest in open proxy security issues was rekindled this last year when it became clear that spammers were exploiting insecure proxy servers to inject unsolicited commercial email.
- Examples of bulk email software products touting their use of proxies for sending bulk email includes: G-Lock's EasyMail, List Sorcerer, Send-Safe, and many others.

Questions I had...

- Clearly abuse of open proxies for sending spam had become a systematic/structural phenomenon. I became intrigued, and decided I should study the open proxies that were being abused. Questions I wanted to be able to answer included:
 - Where were all these open proxies located? (Put another way, what ISPs seemed least competent when it came to dealing with abused boxes?)

Questions I had...

- -- How many open proxies were out there?
(I'd assumed that there were at most a few hundred, or maybe a couple of thousand, but I was off by orders of magnitude)
 - Which proxy blacklists worked best?
 - I also wanted to test a theory I had that when publicly identified, insecure proxies tended to get fixed, or crushed into unusability by massive worldwide demand.
- This talk is the result of my investigation into open proxies and those topics.

"Is this talk relevant to me?"

- Because this talk introduces a security topic which hasn't been talked about at previous Internet2 meetings, you may wonder, "Is this talk relevant to me?"
- I suppose that depends...
 - If you're an end user who's wondered how spammers anonymously shovel unsolicited commercial email at you, yes, it will be relevant.

"Is this talk relevant to me?" (2)

- -- If you're a sysadmin attempting to develop a strategy to cope with spam, attempting to understand an attack vector you may be confronting, or attempting to understand why it is important to secure your own proxy, it's definitely relevant.
- If you are an engineer responsible for your network's security, it definitely will be relevant.

"Is this talk relevant to me?" (3)

- -- If you are a policy person, concerned with acceptable use issues, privacy and anonymity issues, bandwidth management policies, maintaining Internet2/non-Internet-2 network traffic separation, etc., it will be relevant.
- The rest of you can hit the bar early. :-)

Talk format

- Just as we've done for other Internet2 talks, this presentation has been prepared with sufficient detail to allow for *post hoc* use as a tutorial, so that folks who may not be here can still work through what was covered.
- We've attempted to keep the presentation at an intermediate level of technical detail, with "something for everyone." Some may find it to be more technical than they might like, others may find it rehashes what they already know in spots -- sorry about that.

What this talk is NOT about...

- This talk is NOT about eliminating open proxies as a way of facilitating censorship.
- Nor is this a primer on "how to be a cracker/hacker"; all the security issues mentioned are already publicly known and well documented.
- Lastly, this talk is not meant to dictate how you should run your network or how to configure your servers -- that's a decision for you to make after considering the totality of all applicable circumstances (but I do have some suggestions)

II. A Brief Tutorial on Caching Proxy Servers

What's a caching proxy server?

Why would anyone run one?

- Caching web proxy servers are **NOT** intrinsically evil (*malum in se*).
- For instance, consider a computer lab being used by a class. The instructor may say, "Okay class, let's all look at the Smithsonian's web site. Please go to <http://www.si.edu/>"
- The thirty or forty students in that class then (all more-or-less simultaneously) retrieve a copy of the Smithsonian's home page (and its associated images) over the Internet.

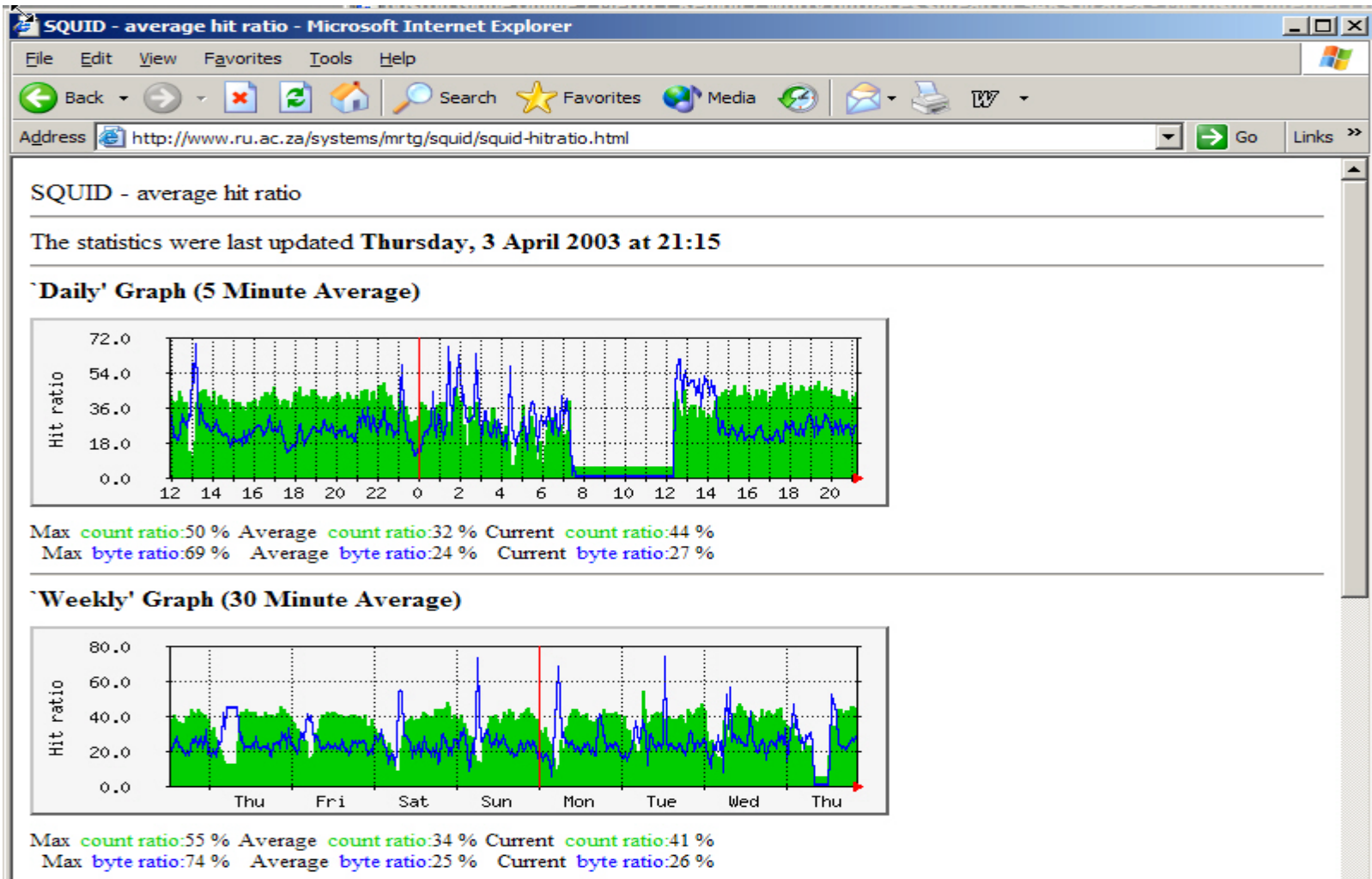
Redundancy Department, Department of

- Think about what just happened -- why should *each* person in that class retrieve their *own* copy of the Smithsonian's web page via the Internet? Why not just let the *first* person to ask for that page retrieve a copy over the Internet, saving and (locally) sharing that recent copy with other local users who are also interested in that same page? It turns out that that's precisely what caching web proxy servers actually do....

Bandwidth savings associated with doing web proxy caching...

- It is common to see cache vendors claim that a properly deployed cache can typically serve 1/3 to 1/2 of all end user page requests from a local web cache, thereby reducing bandwidth usage by up to 25% or more.
- You can see some publicly available proxy cache stat reports by searching google for calamaris "Proxy Report"
(Calamaris is one of the more popular web proxy cache log parsers).

Some folks even use MRTG to track cache hit ratios...



Improving the "Internet experience"

- Caching can also improve the user's "Internet experience," since document retrievals "feels faster" (and large documents are delivered faster, considering bandwidth-delay product issues) when served from a local, lightly loaded, properly engineered cache box connected via gigabit ethernet.

There are many web caching proxy server products which one could use...

- Squid (free): <http://www.squid-cache.org/>
- Blue Coat (formerly CacheFlow):
<http://www.bluecoat.com/>
- NetApp: <http://www.netapp.com/products/netcache/>
- Volera: <http://www.volera.com/>
- ... and many others (including "big names" like Cisco, IBM, Microsoft, Sun, etc.)

Do ISPs actually use web proxy caching?

- You betcha. Notwithstanding arguments for network transparency (e.g., RFC 2775), and notwithstanding the ready availability of cheap commodity transit bandwidth (and the importance of non-proxy-enabled P2P applications in determining ISP bandwidth usage), caching is still common at many large ISPs such as AOL, Comcast, Cox, Road Runner, etc., as well as at large universities (e.g., <http://www.cites.uiuc.edu/webcache/>)₂₀

Both ends of the spectrum...

- One of the (many) ironies of web caching is that web caches tend to be deployed by two completely dissimilar types of sites:
 - a) huge ISPs (such as RBOCs, cable modem providers, and large universities) offering broadband connectivity to 10s or 100s of thousands of users, and
 - b) small sites that are thinly connected to the Internet (such as foreign sites paying outrageous fees for connectivity).

Are all web pages cacheable?

- It is comparatively easy to intentionally (or accidentally) create non-cacheable pages:
 - https (secure web pages), or pages protected with HTTP authentication
 - pages with dynamic content (e.g., URLs including .cgi, .asp, a ? or a ; are often not cached), or pages using cookies
 - pages explicitly marked as non-cacheable
- To check the cacheability of a given page, see <http://www.ircache.net/cgi-bin/cacheability.py>

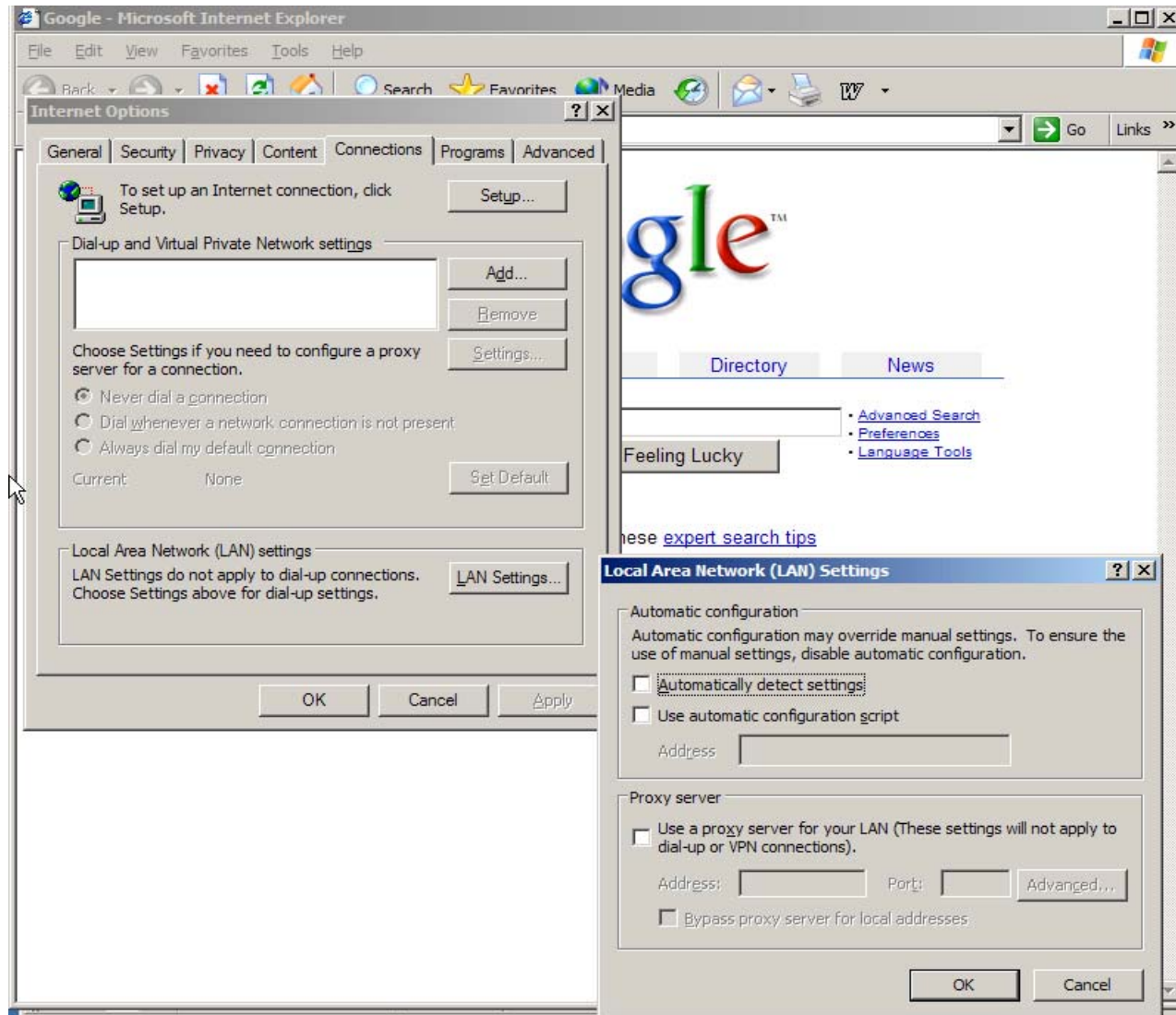
CacheNow! webpages

- One of the most influential pages encouraging both cache deployment and cache-friendly web page design is the CacheNow! Site at <http://vancouver-webpages.com/CacheNow/>
- So assuming an ISP wanted to deploy a web proxy cache, how might they do it?

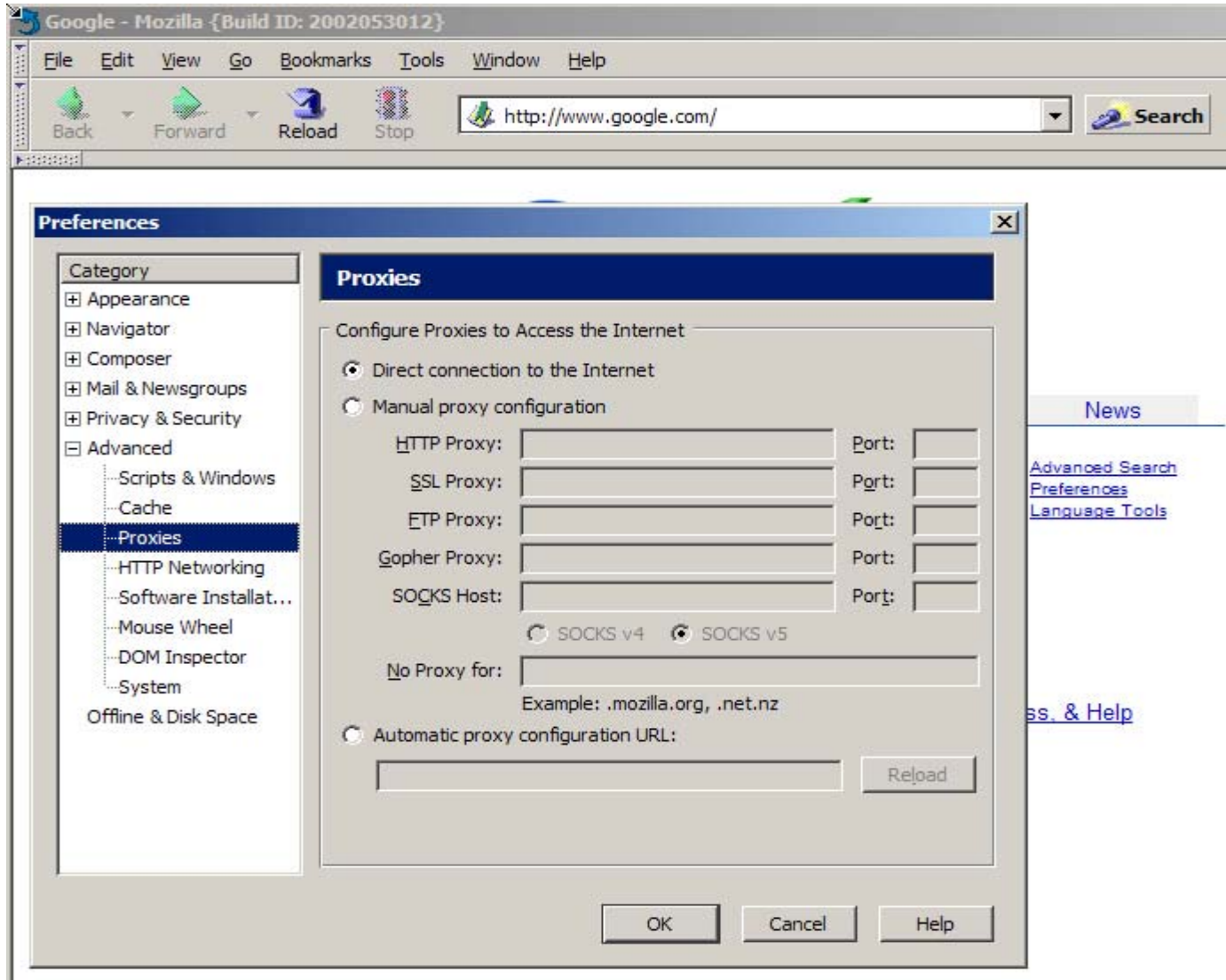
Voluntary use of a web cache

- There are three different ways an ISP or other site could deploy & use a caching proxy:
 - 1) A site can offer a caching web proxy and allow users to manually configure their browser to use it (or not use it) as they personally see fit. This approach assumes that users will be willing and able to manually configure their web browser's to use the proxy server. [Doing that configuration isn't all that hard, but it isn't particularly intuitive, either, requiring entry of a host name & port number²⁴]

Manually configuring IE



Manually configuring Mozilla



ISPs "incenting" the voluntary use of a web cache

- Why anyone would bother to use a non-mandatory web cache?
- At least some sites may offer "incentives" to encourage web cache use, such as exempting traffic flowing through that cache from per-byte traffic charges, excluding traffic flowing through the cache from per-user traffic quotas, or excluding traffic flowing through the cache box from traffic shaping rulesets (making it faster).

Examples of sites incenting use of a proxy server

- http://rcn.oregonstate.edu/bandwidth_faq
"Any traffic you use through the proxy server does not count against your inbound traffic limits."
- http://www.ucs.uwa.edu/web/info/access/netusage_faqs/traffic
"If the item is already in the cache there is no charge."

Another approach: WPAD

- 2) A site could also exploit WPAD (Web Proxy Auto-Discovery Protocol) to auto-direct most browsers (including IE) to a suitable local web cache. This assumes:
 - users have left "Automatically detect settings" checked in their Internet Explorer Preferences (see the "Manually configuring IE" slide earlier in this talk)
 - your web cache has a suitable name (e.g., wpad.<domain> (or WPAD info is being passed via DHCP at address assignment time))

Some WPAD references

- -- (expired draft) <http://www.wrec.org/Drafts/draft-ietf-wrec-wpad-01.txt>
- (expired draft) <http://www.wrec.org/Drafts/draft-cooper-webi-wpad-00.txt>
- (03/1996) <http://wp.netscape.com/eng/mozilla/2.0/relnotes/demo/proxy-live.html>
- <http://www.microsoft.com/windows2000/en/datacenter/help/autodis.htm> (see also the MS IE 5.X Resource Kit, Chapter 21)

Security sidebar: wpad.<domain> is a magic/important hostname

- Because many web browsers automatically look for wpad.<domain>, uh, some security conscious folks might want to insure that that address is pointed at an, uh, trustworthy host. Empirically checking 211 Internet2 members to see if wpad.<domain> was in fact defined, I found that only six domains (bradley.edu, brandeis.edu, orst.edu, swmed.edu, ucsd.edu, uoregon.edu) had bothered to define wpad.<domain>. Hmm.₃₁

Another approach: transparent caching

- 3) A site can transparently ("passively") route all web traffic through a cache box, either by using Web Cache Communication Protocol (WCCP) on a router or layer 4 ethernet switch, or by physically forcing all traffic through an inline network gateway device which includes proxy server functionality.

Transparent caching with WCCP

- For more information on WCCP, see:
 - <http://www.cisco.com/warp/public/732/Tech/switching/wccp/index.html>
- An example of configuring a cache box using WCCP is available at:
 - <http://www.cacheflow.com/support/config/transparent/wccp.cfm>
- Before considering using WCCP, see also:
<http://www.ciac.org/ciac/bulletins/i-054.shtml>

III. Inline Proxy Servers Aren't Just Cache Boxes Anymore

... they also include a corkscrew, a screwdriver, a nail file, a can opener, a magnifying glass, a tiny pair of little scissors, a toothpick....

Transparent caching using an inline gateway device

- The primary alternative to steering traffic via WCCP for inline transparent caching is forcing web traffic through a network "choke point" -- an inline gateway device functioning as a proxy (the gateway device may also act as a web content filter/traffic monitor, a firewall, anti-virus scanner, etc.)
- Customary downsides to single points of failure, and problems going really fast through an appliance, are hereby stipulated.³⁵

Despite single points of failure issues and capacity issues...

- ... inline transparent cache boxes are still quite popular because of all the additional stuff that can be done in addition to the proxy server's basic caching functionality.
- Put another way, the availability of a single centralized possible point of control is just "too sweet" for many admins to forgo, which is why web content filtering software is perhaps the most common add-on....

Content filtering via an inline web proxy

- Some examples of web proxy filtering ("censorware") products deployed via inline transparent proxy boxes include:
 - Bess (<http://www.n2h2.com/>)
 - BlueCoat (http://www.bluecoat.com/solutions/content_filtering.html)
 - SquidGuard (<http://www.squidguard.org/>)
- [A critique of the merits of "censorware" is available at <http://censorware.net/> see also <http://www.sethf.com/anticensensorware/>]

Advertising content filters deployed via an inline proxy

- It is worth mentioning that besides the semi-controversial "censorware" products targeting "objectionable"/"recreational" web content, there are proxy filtering products which target cruft such as ads, popups, and a host of other obnoxious advertising-related stuff. <http://internet.junkbuster.com/> and many others are listed at http://dmoz.org/Computers/Software/Internet/Servers/Proxy/Filtering/Ad_Filters/

Anti-viral filtering via an inline web proxy server

- Sites may also combine web proxies with anti-viral filtering at a gateway box.
- Examples of products doing this sort of thing include:
 - Trend Micro's InterScan VirusWall
 - McAfee WebShield
 - Symantec AntiVirus Gateway

But hey, you're running desktop antivirus software, and SMTP executable attachment defanging with procmail already, right?

Proxy servers for privacy enhancement

- Some people believe that proxy servers will give them "enhanced privacy;" maybe... but don't forget X-Forwarded-For: headers!
- Various browser anonymity checking web sites will let you see what your browser is revealing when you connect, including:
<http://www.all-nettools.com/pr.htm>
<http://privacy.net/analyze/>
<http://www.samair.ru/proxy/proxychecker/>

If you really need privacy...

- There are some companies that offer privacy enhancement services via proxy servers such as allconfidential.com, primedius.com, anonymizer.com, freedom.net, guardster.com, etc.
- Curious? You can test drive an anonymizer: <http://anon.free.anonymizer.com/http://cnn.com/>
- *Note:* I'm not qualified to assess the quality of the privacy delivered by these or any other service, but there are analyses out there you should see. For example...

<http://cs.bu.edu/techreports/pdf/2002-003-deanonymizing-safeweb.pdf>

DEANONYMIZING USERS OF THE SAFEWEB ANONYMIZING SERVICE *

David Martin
Research Assistant Professor
Computer Science Department
Boston University
dm@cs.bu.edu

Andrew Schulman
Chief Researcher
Workplace Surveillance Project
Privacy Foundation
undoc@sonic.net

February 11, 2002

Abstract. The SafeWeb anonymizing system has been lauded by the press and loved by its users; self-described as “the most widely used online privacy service in the world,” it served over 3,000,000 page views per day at its peak. SafeWeb was designed to defeat content blocking by firewalls and to defeat Web server attempts to identify users, all without degrading Web site behavior or requiring users to install specialized software. In this article we describe how these fundamentally incompatible requirements were realized in SafeWeb’s architecture, resulting in spectacular failure modes under simple JavaScript attacks. These exploits allow adversaries to turn SafeWeb into a weapon against its users, inflicting more damage on them than would have been possible if they had never relied on SafeWeb technology. By bringing these problems to light, we hope to remind readers of the chasm that continues to separate popular and technical notions of security.

Windows connection sharing

- Some entities run Windows host-based proxy servers as a way of sharing a single Internet connection. Examples include:
 - ICS (integrated in Windows itself...)
 - AnalogX Proxy
 - Avirt Spaghetti
 - Deerfield WinGate
 - Grok Developments NetProxy
 - Ingetic Proxy+
 - Kerio WinRoute Pro
 - Youngzsoft CCProxy, etc., etc., etc.

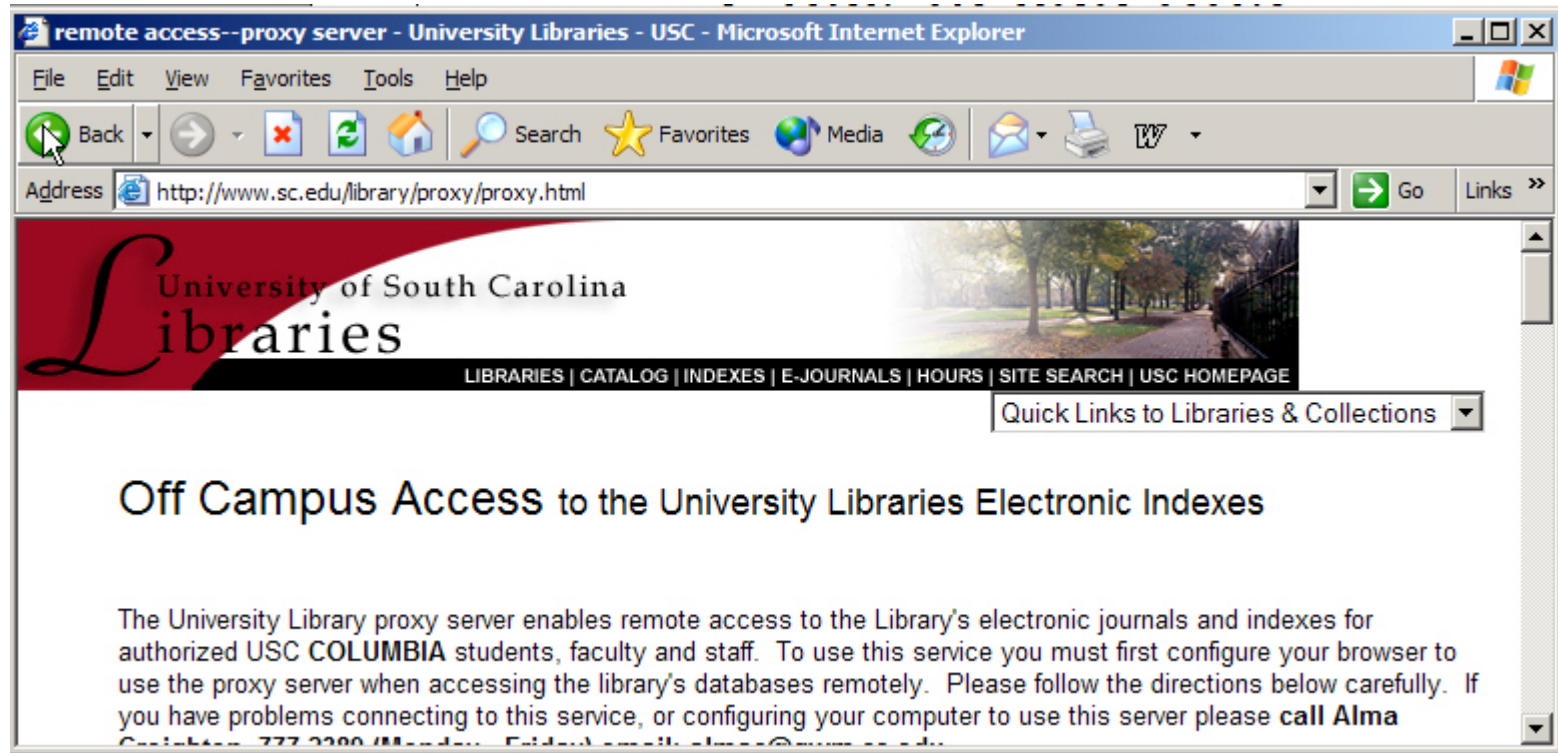
Windows connection sharing insecurity

- While some of those connection sharing products go to great pains to do that sharing securely, other Windows connection sharing products are quite "casual" about security.
- Many of the open proxies we'll talk about later are actually associated with Windows connection sharing software installed by technically unsophisticated users who have no idea just what they've done...

Reverse proxies

- A final category of proxy server is the reverse proxy server. Reverse proxy servers are commonly deployed to allow remote users to do username and password authentication and gain access to domain-name- or ip-address-range-limited resources such as proprietary online databases. Reverse proxies are commonly deployed by academic libraries; a better alternative is to deploy a VPN offering authentication and encryption.

A typical academic library reverse proxy server



IV. Open Proxies

From benign to...

- Now that you understand a little about how proxy servers are supposed to work, let's buckle down and talk about the true subject of this talk: open proxies.

What is an "open proxy?"

- An open proxy is a computer that accepts connections from anyone, anywhere, and forwards the traffic from those connections as if it had originated locally from that host.
- In some cases, the proxied connection may only allow access to the world wide web, but in many cases the open proxy may also be used to ftp files, read and post Usenet news, send email (including spam), do IRC or instant messaging, launch a DOS attack, etc.

Open proxies are NOT the same as open SMTP relays

- Folks sometimes confuse open SMTP relays (which most folks now have pretty well under control) with open proxy servers.
- Open proxies are NOT the same as open SMTP relays -- open proxies are a far, far more serious problem, since they allow traffic for virtually ANY network service to be "bounced through" that host (although open proxies can and do also act as spam conduits).

Open proxies have been the subject of security bulletins...

The screenshot shows a Microsoft Internet Explorer browser window with the title bar "CERT/CC Vulnerability Note VU#150227 - Microsoft Internet Explorer". The address bar shows the URL "http://www.kb.cert.org/vuls/id/150227". The page content is from the Carnegie Mellon Software Engineering Institute, CERT Coordination Center. The main heading is "Vulnerability Note VU#150227" followed by the subheading "Multiple vendors' HTTP proxy default configurations allow arbitrary TCP connections via HTTP CONNECT method". The "Overview" section states: "Multiple vendors' HTTP proxy services use insecure default configurations that could allow an attacker to make arbitrary TCP connections to internal hosts or to external third-party hosts." The "I. Description" section begins with: "HTTP proxy services commonly support the HTTP CONNECT method, which is designed to create a TCP connection that bypasses the normal application layer functionality of the proxy". A left sidebar contains navigation links: "Vulnerability Notes Database", "Search Vulnerability Notes", "Vulnerability Notes Help Information", and a "View Notes By" section with links for "Name", "ID Number", "CVE Name", and "Date Public".

CERT/CC Vulnerability Note VU#150227 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print W

Address <http://www.kb.cert.org/vuls/id/150227> Go Links >>

Carnegie Mellon
Software Engineering Institute
CERT Coordination Center

Home Site Index Search Contact FAQ
vulnerabilities, incidents & fixes *security practices & evaluations* *survivability research & analysis* *training & education*

[Vulnerability Notes Database](#)

[Search Vulnerability Notes](#)

[Vulnerability Notes Help Information](#)

Vulnerability Note VU#150227

Multiple vendors' HTTP proxy default configurations allow arbitrary TCP connections via HTTP CONNECT method

Overview

Multiple vendors' HTTP proxy services use insecure default configurations that could allow an attacker to make arbitrary TCP connections to internal hosts or to external third-party hosts.

I. Description

HTTP proxy services commonly support the HTTP CONNECT method, which is designed to create a TCP connection that bypasses the normal application layer functionality of the proxy

View Notes By

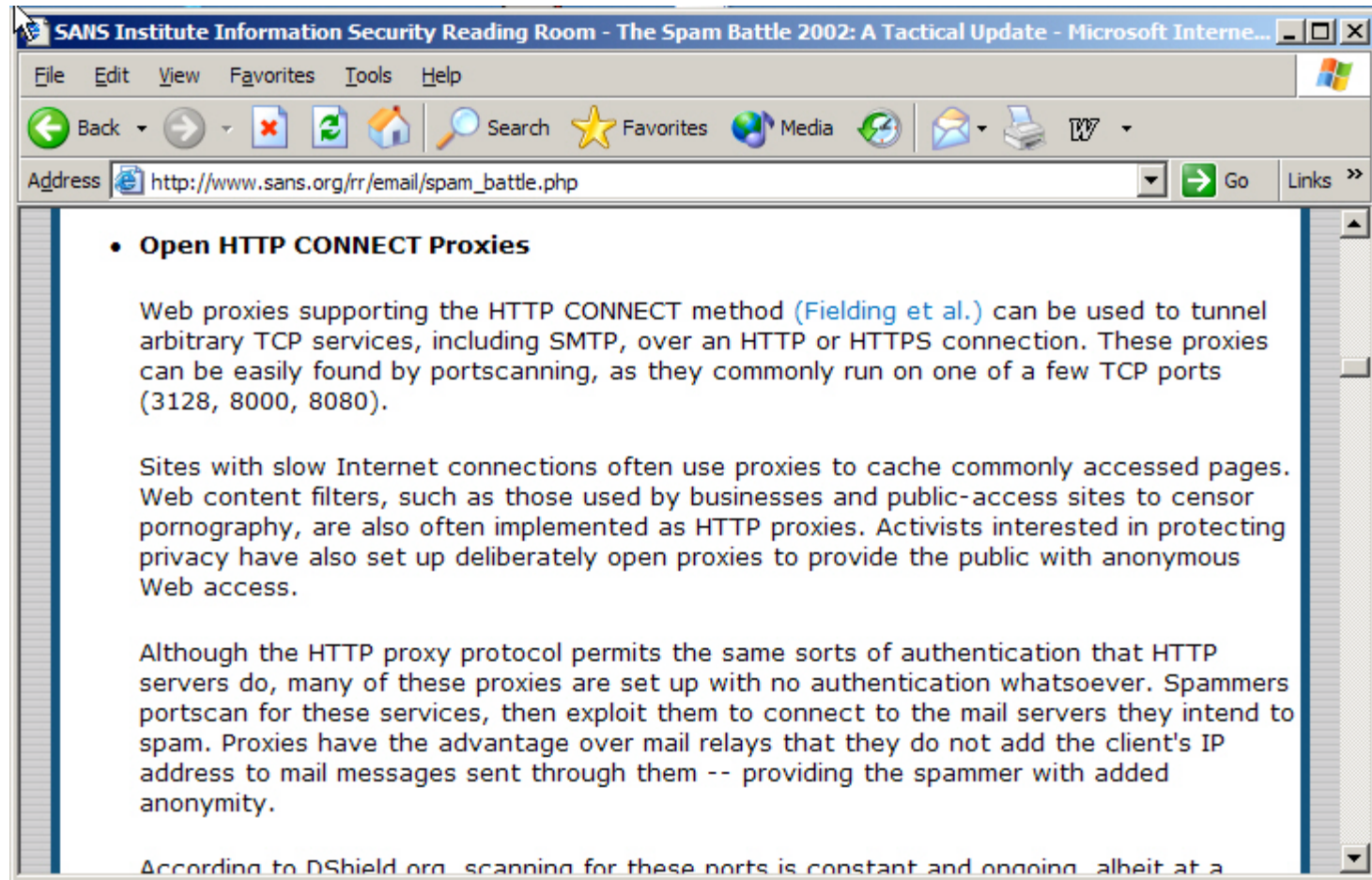
[Name](#)

[ID Number](#)

[CVE Name](#)

[Date Public](#)

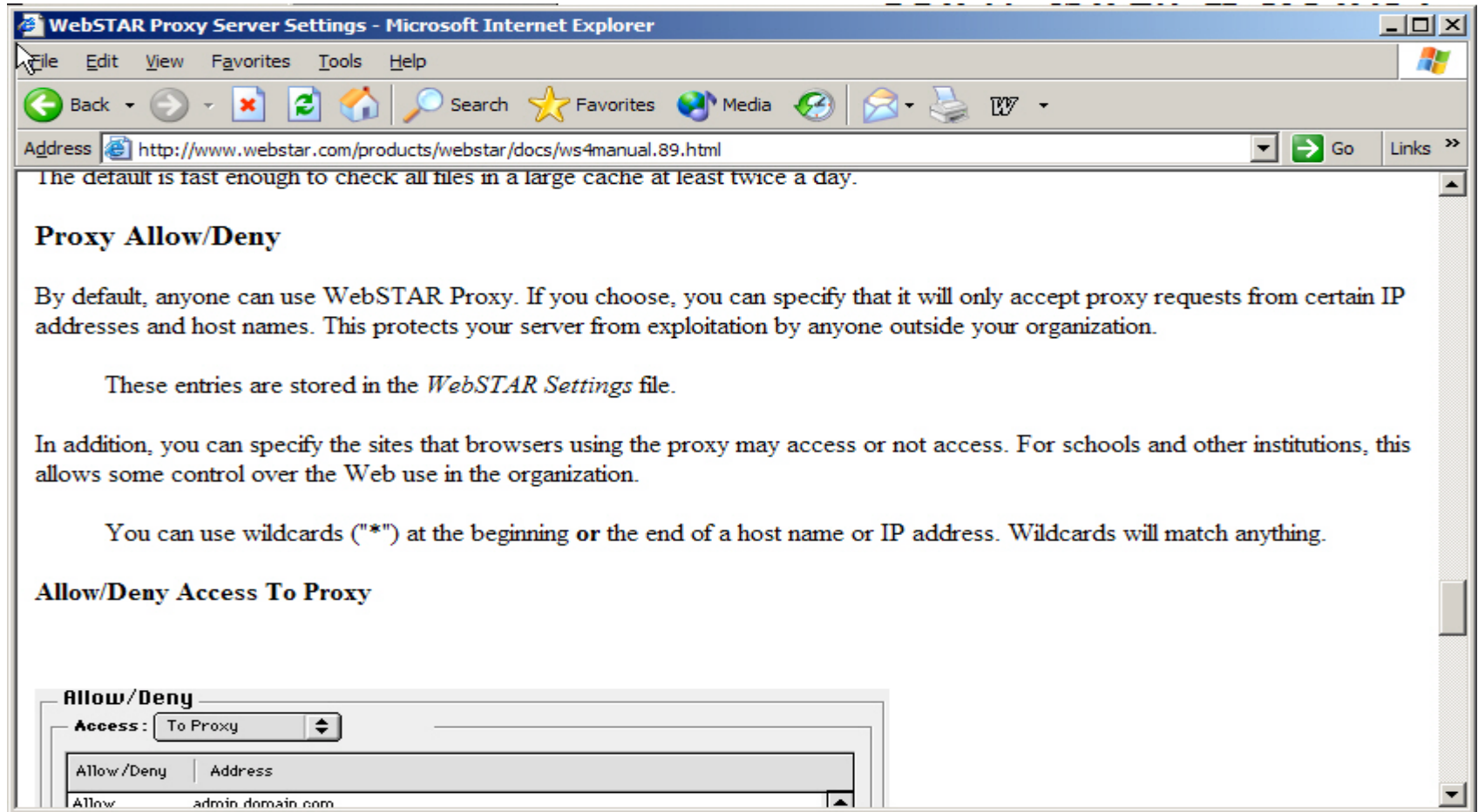
And excellent narrative discussions...



So how does a proxy server become open and abusable?

- A proxy server becomes open due to:
 - misconfiguration/lack of configuration by the administrator (e.g., a proxy server may ship "open by default," and access control lists may never be installed, or if they are installed, they may have been mis-set)
 - inherent protocol/app deficiencies
 - a conscious decision on the part of the party installing the proxy to run it open (Owned boxes, political motivations, etc.)

Example of shipping "open by default"



Trojan'd proxy servers

- Other users may be running a proxy server which was installed by a hacker/cracker via a trojan horse
- Canonical example: jeem.mail.pv
Jeem creates an open SMTP relay plus two open proxy ports on odd high numbered ports. See, for example:
<http://securityresponse.symantec.com/avcenter/venc/data/backdoor.jeem.html>
- As the pool of "regular" open proxies get secured, we'll all see more Jeem'd machines⁵⁵...

V. Why Are Open Proxies of Interest to "Bad Guys"?

Are bad guys really interested in open proxies?

- Yes!
- I believe open proxies are of exceptional interest to various and sundry "bad guys" for many reasons.
- To understand why, it helps to think about things from their point of view for a bit...

(a) "I don't want folks to know where I'm *really* coming from"

- Connections made via an open proxy are often non-accountable, since the proxy may be doing no logging, or if logging is being done, logs may be unavailable to those investigating network incidents.

What if proxy server log files might be available?

- In the case of bad guys who are exploiting proxy servers with the goal of trying to "cover their tracks," proxy server logs files *might* sometimes be obtainable. The accepted "bad guy solution" to that problem is to simply chain multiple proxy servers together, either manually or using a product such as <http://proxychains.sourceforge.net/>
- Doing explicit traffic routing via multiple indirect hops is not really a brand new idea.

Remember "blueboxes"?

- In 1971, (a long, *long* time ago by Internet standards), a popular activity with some "telephone hobbyists" was something called "tandem stacking." Someone engaged in tandem stacking might use a special device to chain a phone call from one central office switch to another, with the most audacious striving to build a path which would route a simple intra-city call thru switches spanning the globe. (*Esquire*, 10/1971)

Flash forward to 2003...

- Thirty two years later, people are *still* routing traffic in unexpected ways -- but now the oddly routed traffic is network data traffic, not voice telephony traffic.
- For example, any technically inclined person will have wondered, "Why am I getting spammed (or why is my firewall getting probed) from odd places in Asia, Africa, and South America?"
- Concise answers: open proxies (of course).

(b) "I want to attack you from many odd locations at once!"

- Open proxies allow a single entity to launch attacks/send traffic from multiple provider-diverse sources at the same time, thereby complicating the problem of blocking spam or firewalling an attack. Dealing with multiple parallel (potentially changing) attack sources is one of several reasons why distributed denial of service network attacks are potentially so tough to deal with.

(c) "I want to try misleading naïve users by forging random garbage into mail headers!"

- Unlike spam sent via an open SMTP relay, spam sent via an open proxy server can be constructed so as to have arbitrary Received: headers, thereby inhibiting efforts at backtracking spam to its source.

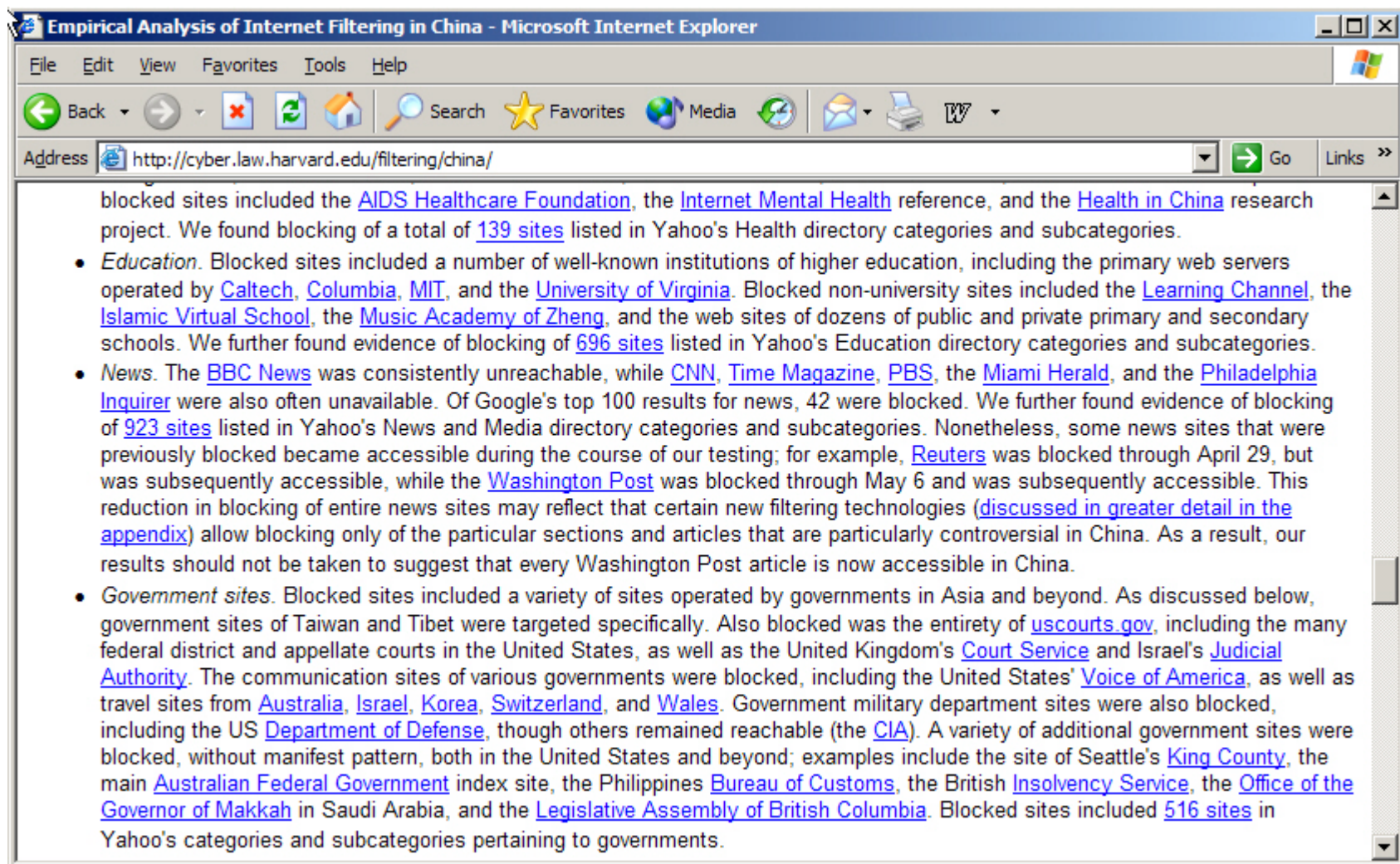
"Falsification of routing data"

- It is interesting that many of the latest generation of state anti-spam laws (see <http://spamlaws.com/>) prohibit spammer "falsification of message routing data"
- Use of open proxies is pretty much the best/only "message routing falsification" trick spammers have available once you get users to the "could you please turn on full headers?" level of spam analysis/reporting (<http://micro.uoregon.edu/fullheaders/>)

(d) "How dare you try to censor me!"

- By using an open proxy server, a user may be able to overcome local connection filtering. For example, if your local network disallows connections to recreational web sites, but allows you to connect to an open proxy, you can access a recreational web site of interest by connecting to it indirectly, via the open proxy. Open proxy servers are thus particularly popular with subjects of totalitarian regimes, and K12 students.

For example: filtering in CN...

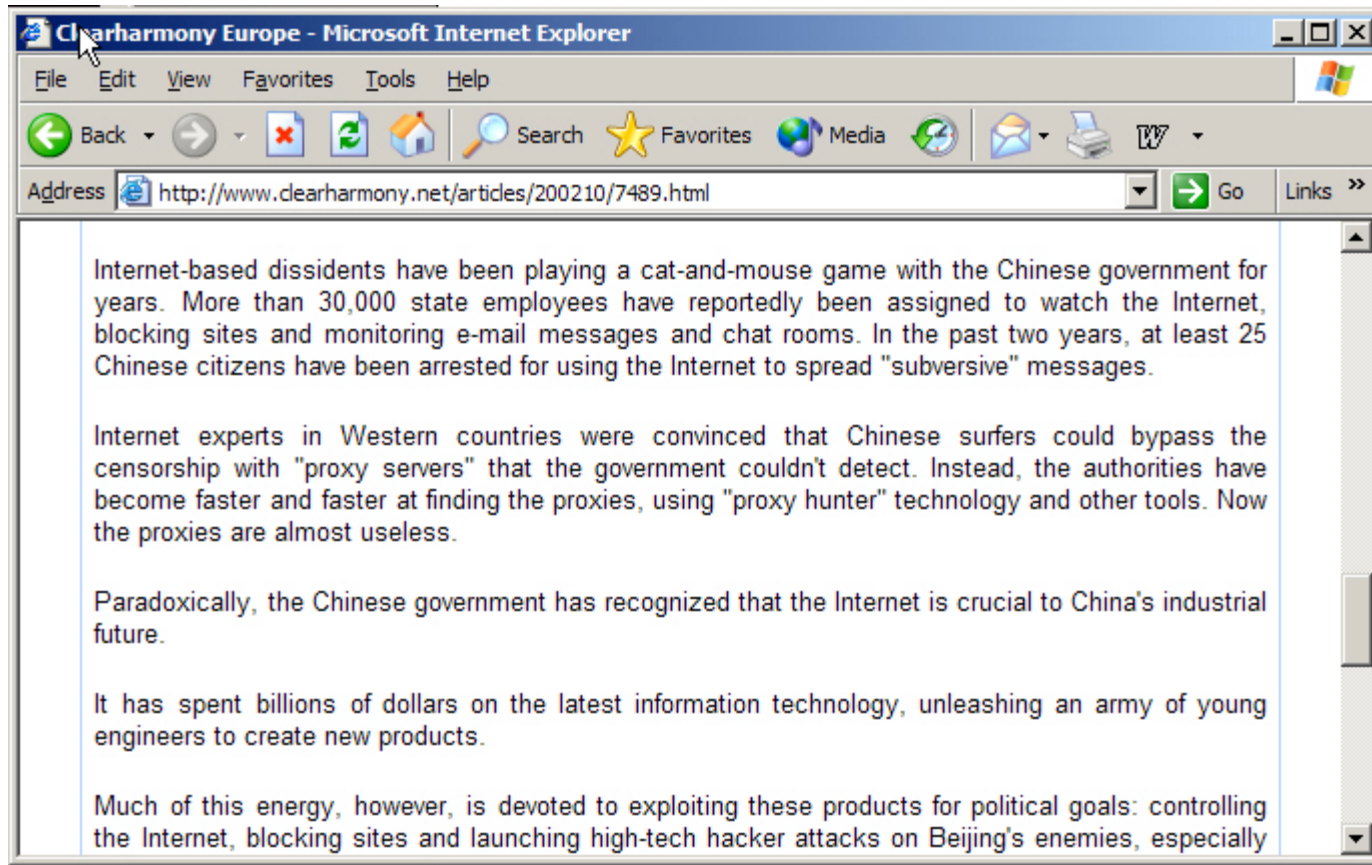


The screenshot shows a Microsoft Internet Explorer browser window. The title bar reads "Empirical Analysis of Internet Filtering in China - Microsoft Internet Explorer". The address bar shows the URL "http://cyber.law.harvard.edu/filtering/china/". The main content area displays a webpage with the following text:

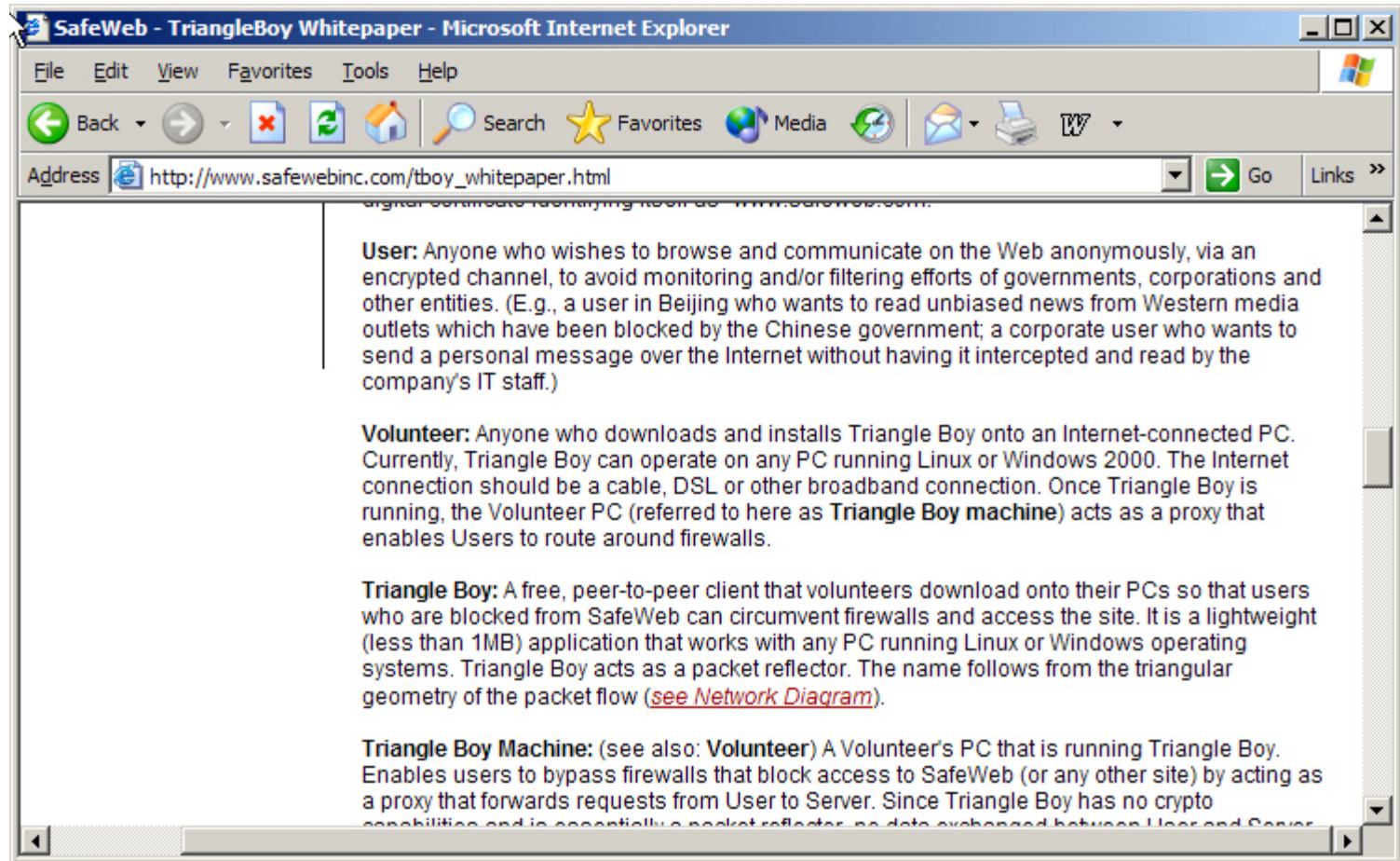
blocked sites included the [AIDS Healthcare Foundation](#), the [Internet Mental Health](#) reference, and the [Health in China](#) research project. We found blocking of a total of [139 sites](#) listed in Yahoo's Health directory categories and subcategories.

- *Education.* Blocked sites included a number of well-known institutions of higher education, including the primary web servers operated by [Caltech](#), [Columbia](#), [MIT](#), and the [University of Virginia](#). Blocked non-university sites included the [Learning Channel](#), the [Islamic Virtual School](#), the [Music Academy of Zheng](#), and the web sites of dozens of public and private primary and secondary schools. We further found evidence of blocking of [696 sites](#) listed in Yahoo's Education directory categories and subcategories.
- *News.* The [BBC News](#) was consistently unreachable, while [CNN](#), [Time Magazine](#), [PBS](#), the [Miami Herald](#), and the [Philadelphia Inquirer](#) were also often unavailable. Of Google's top 100 results for news, 42 were blocked. We further found evidence of blocking of [923 sites](#) listed in Yahoo's News and Media directory categories and subcategories. Nonetheless, some news sites that were previously blocked became accessible during the course of our testing; for example, [Reuters](#) was blocked through April 29, but was subsequently accessible, while the [Washington Post](#) was blocked through May 6 and was subsequently accessible. This reduction in blocking of entire news sites may reflect that certain new filtering technologies ([discussed in greater detail in the appendix](#)) allow blocking only of the particular sections and articles that are particularly controversial in China. As a result, our results should not be taken to suggest that every Washington Post article is now accessible in China.
- *Government sites.* Blocked sites included a variety of sites operated by governments in Asia and beyond. As discussed below, government sites of Taiwan and Tibet were targeted specifically. Also blocked was the entirety of [uscourts.gov](#), including the many federal district and appellate courts in the United States, as well as the United Kingdom's [Court Service](#) and Israel's [Judicial Authority](#). The communication sites of various governments were blocked, including the United States' [Voice of America](#), as well as travel sites from [Australia](#), [Israel](#), [Korea](#), [Switzerland](#), and [Wales](#). Government military department sites were also blocked, including the US [Department of Defense](#), though others remained reachable (the [CIA](#)). A variety of additional government sites were blocked, without manifest pattern, both in the United States and beyond; examples include the site of Seattle's [King County](#), the main [Australian Federal Government](#) index site, the Philippines [Bureau of Customs](#), the British [Insolvency Service](#), the [Office of the Governor of Makkah](#) in Saudi Arabia, and the [Legislative Assembly of British Columbia](#). Blocked sites included [516 sites](#) in Yahoo's categories and subcategories pertaining to governments.

And it is clear the Chinese are aware of open proxy servers



"Triangleboy"



(e) "Ack! They're blocking common P2P ports..."

- While there is substantial interest among users in accessing web content via proxies, and spammers certainly like to use proxies to send email, administrators may not recognize that even non-proxified peer-to-peer applications such as Kazaa, Edonkey, Grokster, Morpheus, etc. can **also** use proxy servers via 3rd party proxy tunnelling applications such as ProxyCap
(<http://proxylabs.netwu.com/proxycap/>)

"My ISP is blocking outbound traffic sent direct to port 25..."

- Some bad guys may also be interested in open proxy servers as a way of getting past provider-installed filters on any outbound SMTP traffic which isn't being sent via the provider's designated SMTP servers.
- Providers who filter outbound port 25 traffic should also be smart enough to filter common proxy server ports, but maybe not.

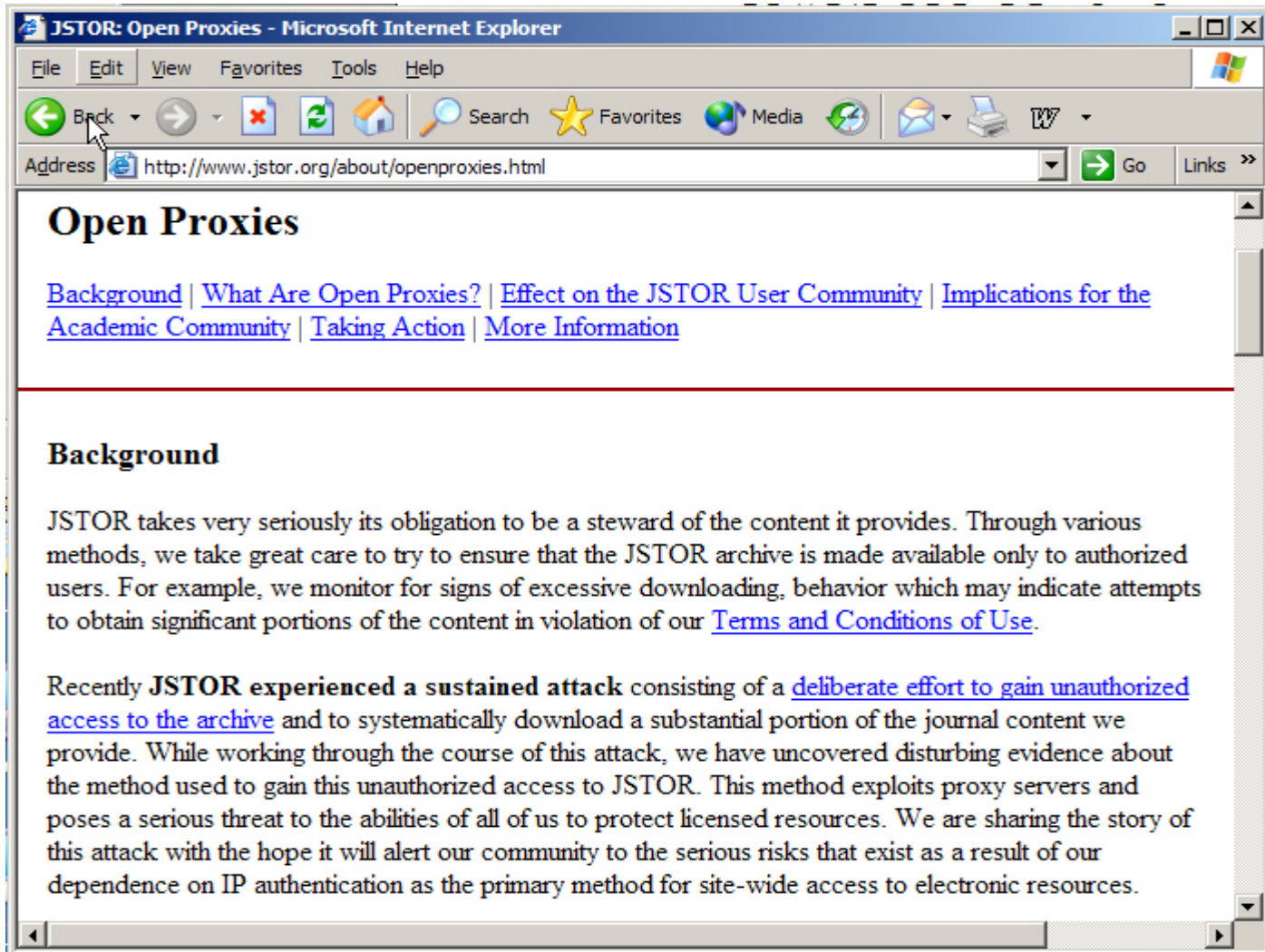
(f) "Hey, *I know* how we can get access to Internet2..."

- Particularly relevant to this audience, you should note that open proxy servers running at Internet2-connected sites may grant access to resources which might otherwise not be available, such as network access to Abilene, or network access to a federal government high performance mission network such as DREN, ESNet, NISN, etc.

**(g) "Limited just to their site?
Nah, it's open to the world..."**

- More than just access to high performance networks is at risk from open proxies. Other assets which are vulnerable to the existence of local open proxies include:
 - Usenet News servers
 - site-licensed software distribution servers, and
 - online proprietary databases.
- For example...

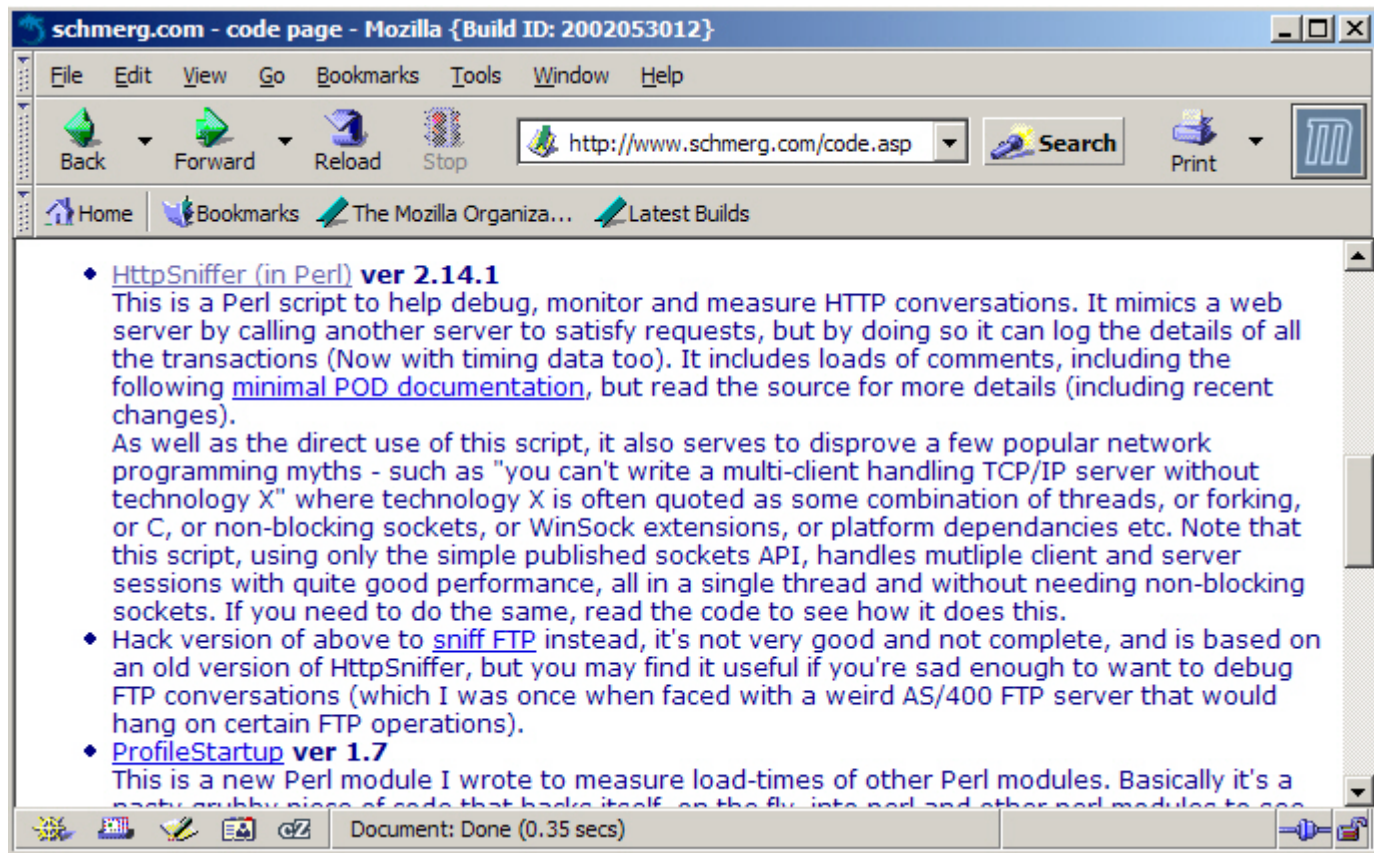
JSTOR and open proxies



(h) "I know a way we can get all sorts of traffic to sniff..."

- Open proxy servers may (or may not) offer you some level of privacy -- a proxy server may be logging nothing about a transaction that occurs via it, or, on the other hand, the proxy server may be undetectably sniffing every character that passes through it (and the origin of those transmissions), snagging unencrypted usernames and passwords, or other confidential info....

HttpSniffer



(i) "I'm *not* making enough on clickthroughs right now..."

- Open proxies may also be exploited by those who are trying to artificially generate inflated "hits" on revenue-generating web site links. (Pay-per-hit revenue programs typically limit payments made on a per-unique-address basis, so to artificially inflate pay-per-hit revenues, you need lots of addresses from which to generate "hits")
<http://www.securiteam.com/securitynews/6M00B2A0KQ.html>

(j) "Do you really suppose we could..."

- And of course, open proxy servers allow bored people to try random network experiments such as routing web traffic from a local workstation to a local server via a chain of proxies spanning the world, just like blueboxers from the early 1970's.
- I'm just waiting for network researchers to start exploiting open proxies as volunteer endpoints for measurement projects. :-;

VI. Open Proxies (From the Point of View of the Intended Users of That Proxy)

"I don't like this place at all
Makes me wonder what I'm here for
Someone take this pain away..."

*Yet Another Day (Riva Remix),
from Touched (George Acosta)*

Problems associated with hosting an open proxy

- In addition to being a "public nuisance" or a security risk to the Internet at large for all the reasons outlined above, open proxy servers really do a disservice to "innocent parties" who sit behind them, too.

(a) Firewall? What firewall?

- Open proxy servers may serve as a conduit for inbound attacks, completely bypassing a site's firewall architecture.

This has happened to some prominent sites....

BW Online | February 27, 2002 | New York Times Internal Network Hacked - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print W

Address http://www.businessweek.com/technology/content/feb2002/tc20020227_4161.htm

BusinessWeek online

BW MAGAZINE DAILY BRIEFING INVESTING GLOBAL BUSINESS TECHNOLOGY SMALL BUSINESS B-SCHOOL

FEBRUARY 27, 2002

TECHNOLOGY

[Recent Tech Features](#)
[Tech Special Reports](#)
[CNet News & Reviews](#)
[Product Reviews](#)
[Internet Security](#)
[Science & Health](#)
[Tech Columns](#)
[Tech Videos](#)
[E Biz](#)
[Newsletter Sign-Up](#)

LIFESTYLE
COLUMNS
FORUMS & CHATS
NEWSLETTERS
PERSONAL FINANCE
SEARCH & BROWSE
SPECIAL REPORTS
TOOLS & SCOREBOARDS
VIDEO VIEWS

Put product at the center

SECURITY FOCUS

New York Times Internal Network Hacked

How open proxies and default passwords led to Adrian Lamo padding his rolodex with 3,000 op-ed writers, from William F. Buckley Jr. to Jimmy Carter

Security holes in the New York Times internal network left sensitive databases exposed to hackers, including a file containing Social Security numbers and home phone numbers for contributors to the Times op-ed page, SecurityFocus Online has learned.

In a two-minute scan performed on a whim, twenty-one-year-old hacker and sometimes-security consultant Adrian Lamo discovered no less than seven misconfigured proxy servers acting as doorways between the public Internet and the Times' private intranet, making the latter accessible to anyone capable of properly configuring their Web browser.

Provided By

SecurityFocus

[Printer-Friendly Version](#)
[E-Mail This Story](#)

RELATED ITEMS

[Security Focus Archive](#)
[Find More Stories Like This](#)

(b) Sharing your pipe with a 100,000 of your closest friends

- Because anyone, anywhere, can freely access the Internet from an open proxy server, unauthorized users will often completely saturate the bandwidth available to that server. This typically results in extremely poor performance for the proxy server's intended users (often folks located in remote parts of the world where bandwidth is scarce or expensive).

Hey, its only money...

- ISPs hosting lots of open proxies also tend to need bigger routers and more commodity transit bandwidth than normal since their open proxy-running customers will be running at unusually high network traffic levels.
- I guess open-proxy-friendly Internet service providers must just love to buy bandwidth (or maybe this is an intentional attempt to look busy enough to justify settlement-free peering? Yeah, that's it, that must be it...) 83

(c) Warrants, subpoenas, and writs, oh my!

- If you host open proxy servers, you should not be surprised if you see a steady stream of warrants, subpoenas and writs seeking customer information, copies of server contents (or the servers themselves).
- I would assert that it is better to buy network engineers and/or security staff to deal with open proxies rather than lawyers to deal with warrants, but each to their own.

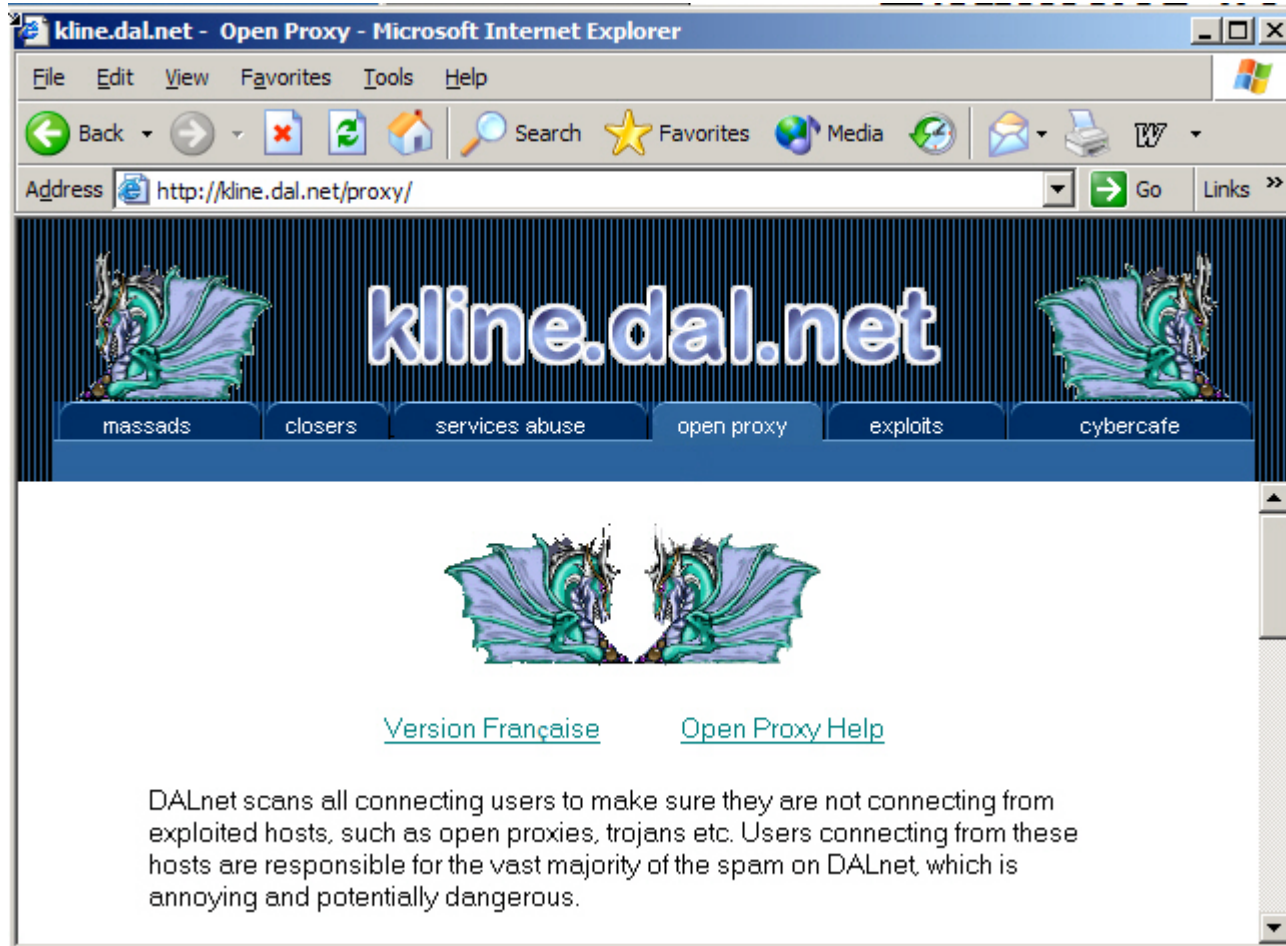
(d) Open proxies may attract probes for other vulnerabilities

- Hosting persistently open proxies may result in an increased risk of that host (and its network) getting scanned for other vulnerabilities, presumably because persistent open proxies serves as an indicator that no one cares/no one is paying attention. This is much like the association between graffiti and crime rates in decaying urban areas. [Customers of some RBOCs must be seeing *incredible* levels of scans...]⁵

(e) Anti-open proxy DNSBLs may block legitimate users

- As open proxy servers become identified and added to open proxy blacklists, legitimate users of those proxy servers may suddenly find that they are blocked by DNSBLs from accessing Internet resources (such as IRC servers) because they are connecting from an open proxy server.

Example of an IRC network blocking open proxies



"Compared to the locusts, the frogs weren't really that bad"

- While having an open proxy DNSBL list a particular /32 can be admittedly inconvenient if you are a user of that open proxy server, it is far LESS inconvenient than having your entire country blocked!
- Yes, there ARE country-wide blacklists in use by people who are completely fed up with spam from some parts of the world that just don't seem to care about network abuse. (I discourage use of country-wide DNSBLs⁸)

Some examples of country-wide blacklists

- <http://www.blackholes.us/> (DNSBLs for network blocks assigned to ISPs in AR, BR, CN, HK, JP, KR, MY, NG, RU, SG, TW, TH; also has blackhole DNSBLs for selected large US/international ISPs)
- <http://www.ocean.com/asianspamblocks.html>
- See also: "Not All Asian E-Mail is Spam" (<http://www.wired.com/news/politics/0,1283,50455,00.html>)

(f) "Semi-innocent" local users may get targeted by inept local bandwidth witch hunts

- When connections get saturated and local performance becomes awful, rather than suspecting that users from all over the world are connecting to an open proxy and gobbling up bandwidth, many folks will just say "AHAH! Someone is <fill in relatively trivial unacceptable local network behavior here>..." with predictable results: a local inquisition and bandwidth crackdown.

(g) More joy of open proxies: getting LOTS of complaints

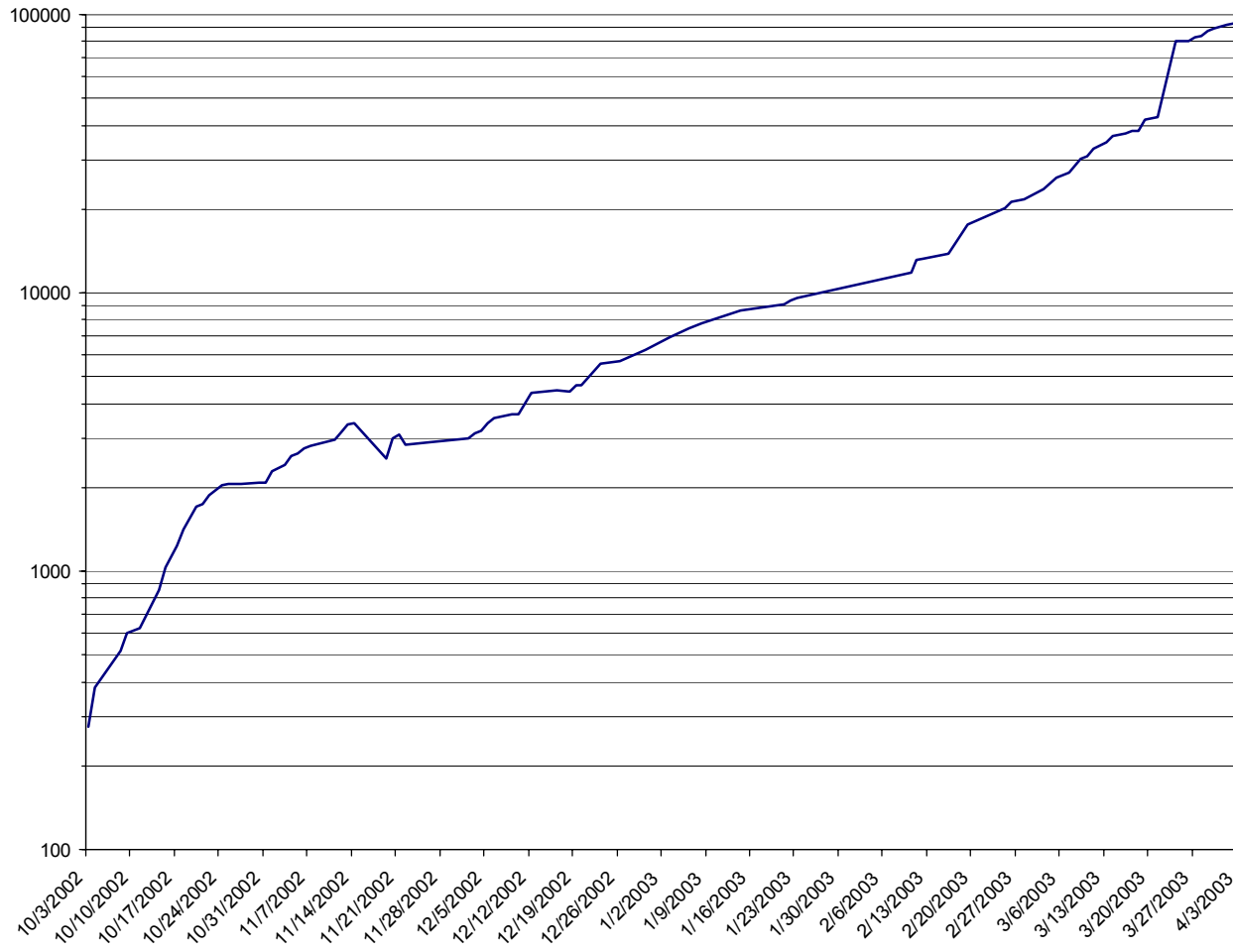
- The parties of record responsible for your network will get LOTS of complaints from angry users who've gotten spammed or otherwise abused via a local open proxy. Parties who will get complaints include whois-listed contacts for your domain and your network address block, your post master and security staff, (c) management, etc. If left undealt with, complaint volume can cause a abuse response "death spiral."

Okay, so having an open proxy really isn't that much fun...

- 100% correct. Having an open proxy server really can be miserable.

What's amazing to us is that despite the substantial pain associated with hosting an open proxy server, and the fact that an open proxy server can exist only if BOTH the system owner/sysadmin AND their ISP don't take steps to deal with the problem, there are LOTS of open proxies out there.

VII. How Many Open Proxies Are Out There?



A serious epidemic, or one person with sniffles?

- The severity of the open proxy problem, like many problems, is largely a function of its size.
- Obviously, if there are only a few hundred open proxies, the problem is a different one than if there are thousands or tens of thousands of open proxies.

Bounding the immeasurable

- No one can authoritatively tell you the total number of open proxies in existence on the Internet today -- that number is constantly changing, and is fundamentally unknowable without systematically probing all possible proxy server ports on all possible addresses.
- Put another way, while we may know how many we've seen so far, we don't know (yet) how many more open proxies are still lurking out there undetected, ripe for abuse.

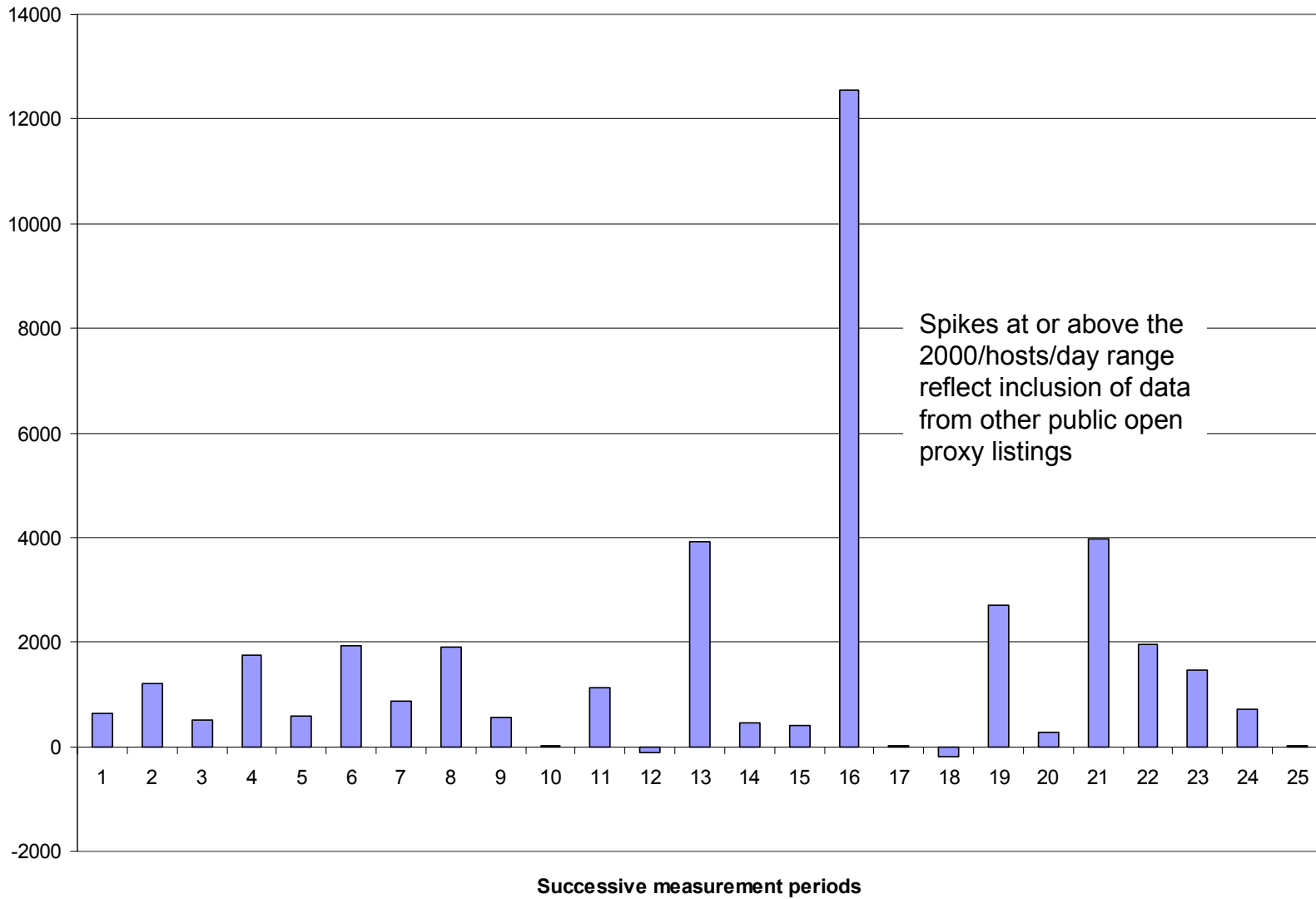
Working toward a number...

- There are, however, some ways we can work towards an estimate of the number of open proxies.
- For example, the reported size of some publicly available open proxy lists tends to run in the tens of thousands to hundreds of thousands of unique addresses.
- Obviously, just from that indicator alone, we know we're talking about an epidemic, not one person with a head cold.

Or we could look at the rate of discovery of new open proxies

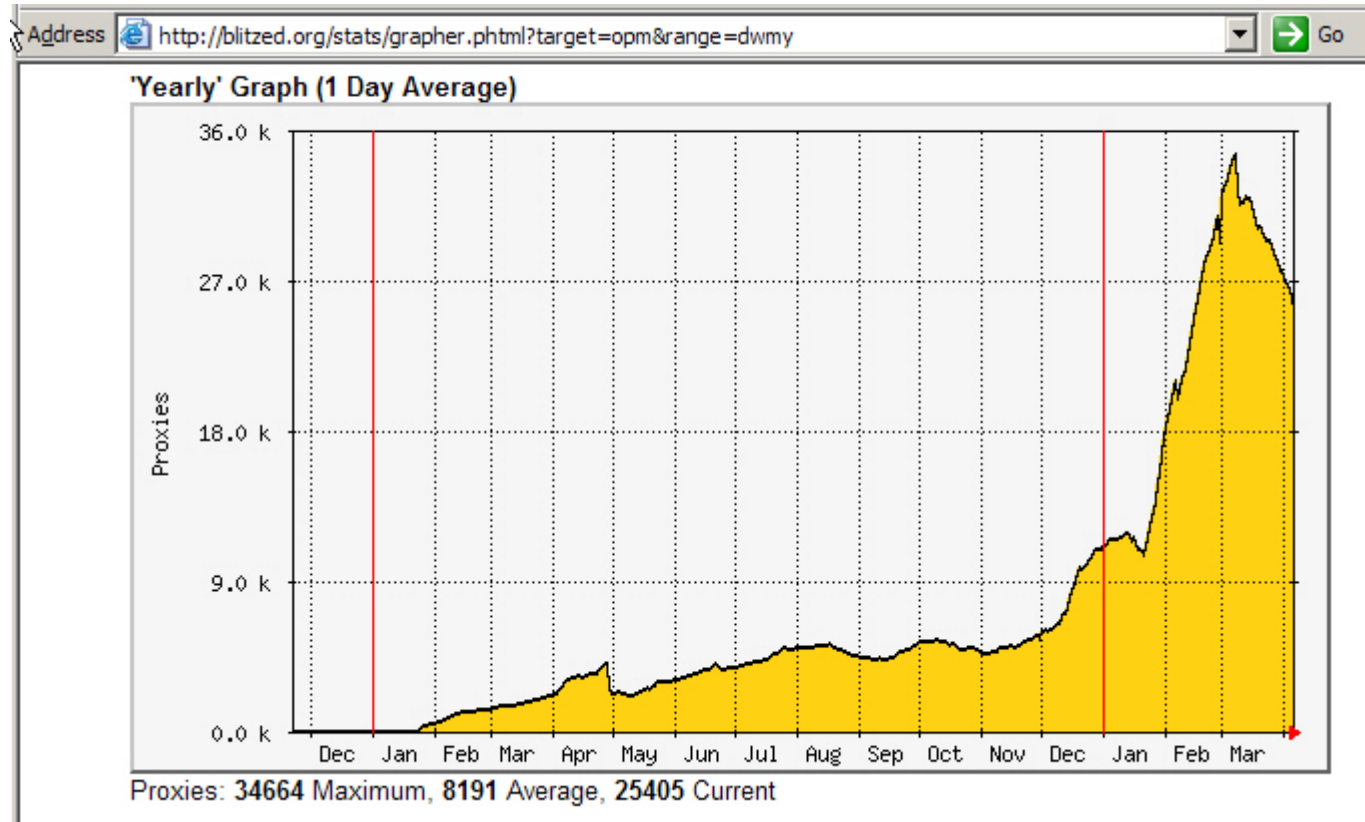
- Let's assume spammers are aggressively looking for new open proxies. As they begin to have problems finding new one, the number of newly abused open proxies we see per day should decrease, and our estimate of the true number of open proxies should begin to asymptotically approach the true number of open proxies. Unfortunately, we're nowhere near asymptotic yet...

**Average new open proxy hosts/day
March 2003 through early April 2003**



One (possible) positive sign...

- We have noted one positive sign: the number of open proxy hosts listed by Blitzed has actually begun to decline:



VIII. Sorting the Sheep from the Goats

How do we know if a host is an open proxy server?

- There are five main ways whereby you can determine if a particular IP address is now or has formerly been an open proxy server:
 - you can check <http://openrbl.org/>
 - you can query open proxy DNS blacklists
 - you can use a fully functional open proxy tester such as <http://hatcheck.org/proxy>
 - you can scan the dotted quad in question for common open proxy ports, or
 - we may be able to watch MRTG graphs.

http://openrbl.org/

Back Forward Reload Stop Search

Home Bookmarks The Mozilla Organiza... Latest Builds

[dnsbl](#) [stats](#) [whois](#) [zones](#) **DNSBL Lookup:** (Dotted-IP, Long or Hostname)

Lookup 4.3.157.36 (lsanca1-157-036.biz.dsl.gtei.net) in 11+20 Zones
AS: 4.0.0.0/8 [AS1](#) GENUITY Cambridge/Massachusetts
Net 4/8 [SATNET](#) Cambridge, Massachusetts [@genuity.net](#)
Results: Positive=11, Negative=20 (2003-03-26 18:57:58 UTC)

- [@SPAM/spamsource](#): PERMBLOCK Open Proxy - <http://openrbl.org/4.3.157.36>;
- [DSBL/dsbl.org](#): DSBL Insecure host <http://dsbl.org/listing?4.3.157.36>;
DSBL Unconfirmed submission <http://dsbl.org/listing?4.3.157.36>;
- [WIREHUB/permblock](#): PERMBLOCK Open Proxy - <http://openrbl.org/4.3.157.36>;
- [SORBS/dnsbl.sorbs.net](#): Open Server [socks/1080] See:
<http://www.dnsbl.sorbs.net/cgi-bin/lookup?IP=4.3.157.36>;
- [UPL/proxies.monkeys.com](#): BLOCKED: See
<http://www.monkeys.com/upl/listed-ip-0.cgi?ip=4.3.157.36>;
- [OSIRU/relays.osirusoft.com](#): This entry was last confirmed open on 3/1/2003;
- [SPAMCOP/bl.spamcop.net](#): Blocked - see <http://spamcop.net/bl.shtml?4.3.157.36>;
- [NJABL/dnsbl.njabl.org](#): open proxy -- 1046457305;
- [BOPM/opm.blitzed.org](#): open proxy - see <http://blitzed.org/proxy/?ip=4.3.157.36>;
- [XBL/gte](#): GOLDPOCKET-235-25; NETBLK-VOLCOM-209-29; NETBLK-SATURNFIVE-105;
spam from dialinx.net; NETBLK-EXCALNET-210-31; burlee; relaytests;
www.echomail.com;
- [REYNOLDS/t1.bl.reynolds.net.au](#): see <http://bl.reynolds.net.au/t1/>;
- Negative 20: [@COUNTRY](#) [@DYNAMIC](#) [@ISP](#) [ASS](#) [BLARS](#) [DORKS](#) [DRBL](#) [FIVETEN](#) [INTERSIL](#) [J](#)
[NOMORE](#) [ORDB](#) [RFC_IPWH](#) [SBL](#) [SPAMBAG](#) [SPAMSITE](#) [SPEWS](#) [WYTNJTO](#) [ZTL](#)

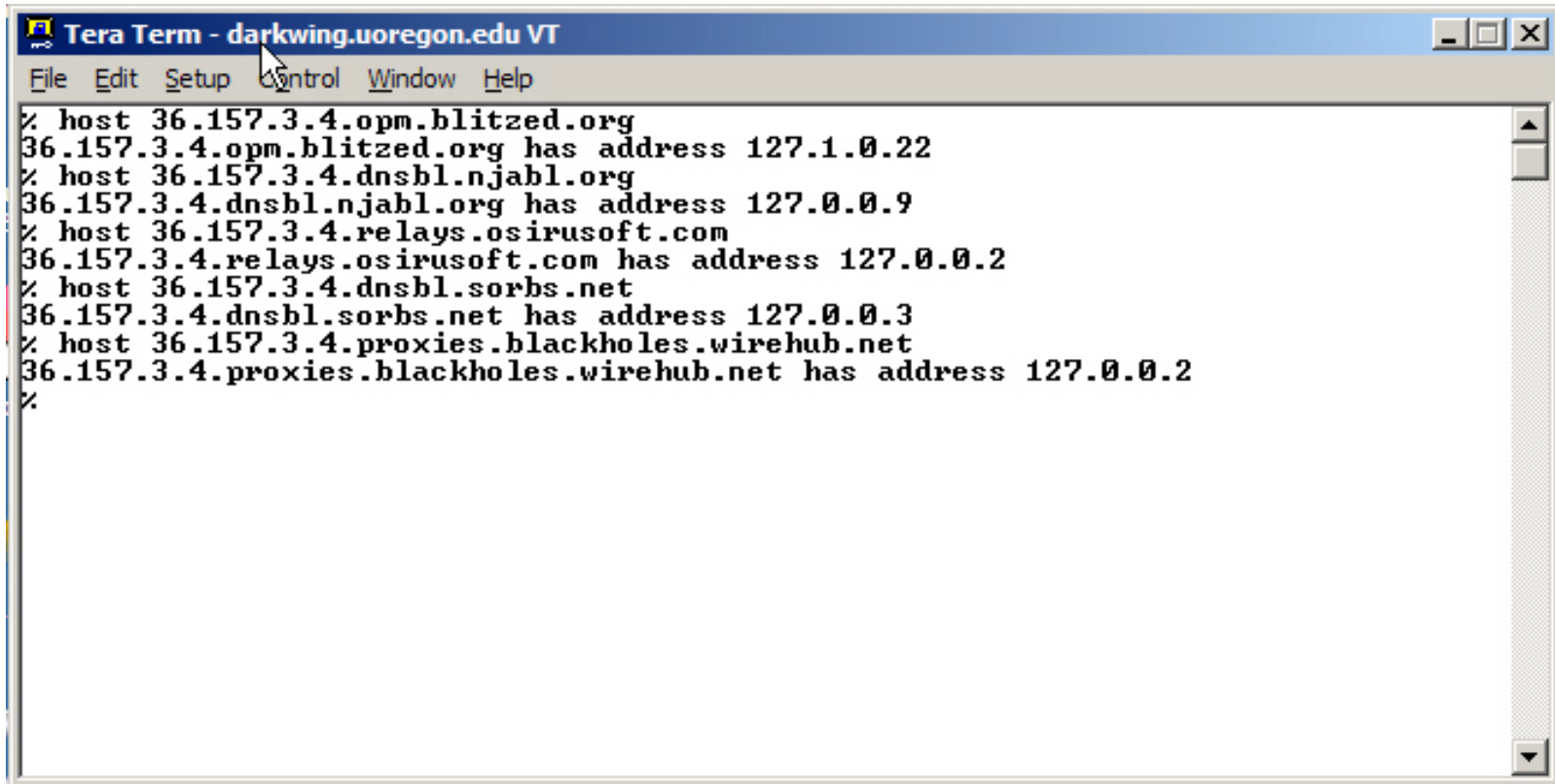
Hints for 4.3.157.36: ([external](#), use BACK or ALT-LEFT when done)

- Track "lsanca1-157-036.biz.dsl.gtei.net" at [[Whois](#) & [Abuse](#)|[SpamCop](#)]
- Search "4.3.157.36" at [[Google](#)|[SenderBase](#)] [[MAPS](#)|[Schlund](#)]
- **CHECK**: Nominate Open-Relay-Tests at: [[DEVNULL](#)|[ORDB](#)]
- Proxy-Test at [[hatcheck.org](#)] [[Add Comment](#)] [[Cached /24](#)]

About OpenRBL

- OpenRBL is a very convenient way for a naïve user to query a comparatively small number of hosts, but it really isn't designed for bulk queries:
 - it is relatively slow
 - it permits a limited number of queries/day
 - it has anti-scripting features built-in
- If you're doing many queries, you'll probably want to do those queries directly.

Querying DNS blacklists



A screenshot of a Tera Term terminal window titled "Tera Term - darkwing.uoregon.edu VT". The window has a menu bar with "File", "Edit", "Setup", "Control", "Window", and "Help". The terminal displays a series of DNS queries and responses for various IP addresses. The queries are prefixed with a prompt character "%". The responses indicate the IP address and the associated DNS blacklist entry.

```
% host 36.157.3.4.opm.blitzed.org
36.157.3.4.opm.blitzed.org has address 127.1.0.22
% host 36.157.3.4.dnsbl.njabl.org
36.157.3.4.dnsbl.njabl.org has address 127.0.0.9
% host 36.157.3.4.relays.osirusoft.com
36.157.3.4.relays.osirusoft.com has address 127.0.0.2
% host 36.157.3.4.dnsbl.sorbs.net
36.157.3.4.dnsbl.sorbs.net has address 127.0.0.3
% host 36.157.3.4.proxies.blackholes.wirehub.net
36.157.3.4.proxies.blackholes.wirehub.net has address 127.0.0.2
%
```


Understanding DNSBLs

- While domain name servers normally are simply used to translate domain names to numeric IP addresses, DNS servers can also be used as an efficient way to convey other info (usually in the form of a "coded" network address from the 127.0.0.0 block), such as whether a network address is known to be an open proxy server. For reasons relating to maintenance of the DNSBL listings, DNSBLs usually use reversed IPs.

Understanding DNSBLs (2)

- For example, if you wanted to query the fictitious DNSBL zone `badhost.foo.bar` to see if `123.45.6.78` was listed, you'd use `host` or `nslookup` or `dig` to check to see if `78.6.45.123.badhost.foo.bar` was defined.
- Note that DNSBL's are "opaque" -- unless the operator chooses to make a copy of that zone publicly available (e.g., via zone transfers), one can only tell if an entry is defined by explicitly checking that address.

Some notes on DNS blacklists

- Some things to note when querying DNS blacklists:
 - (1) Open proxies exist which aren't in any blacklists (duh); conversely some listed dotted quads may no longer be open proxies
 - (2) Some DNSBLs list open proxies AND open relays AND spam-tolerant hosts AND virus-infested hosts AND ... pay close attention to the addresses each DNSBL returns if you only care about open proxies.

Some notes on DNS blacklists (2)

- (3) Some DNSBLs may have restrictive legalistic terms and conditions that are trivial to accidentally violate. I would urge you to respect those terms and conditions, and simply avoid DNSBLs with restrictive T&C's -- there are others w/o tight T&C's.
- (4) Because DNSBLs are remote databases delivered via DNS, recognize that DNS queries **may** sometimes fail (e.g., if all servers delivering DNSBL 'foo' are offline).

Some notes on DNS blacklists (3)

- (5) If you do lots of DNSBL queries, your local name server infrastructure may suddenly become even more important than normal to you, and may need watching to avoid performance issues.
- [Note to self: *time for DNS server benchmarking work?*]
- [Second note to self: *after looking at open proxies problem, is it time to look at the issue of open recursive DNS servers?*]

Some notes on DNS blacklists (4)

- (6) It is (sort of) trivial to automate the DNS queries using shell scripts/small programs.

Note from the trenches: forget about assuming it will be feasible for someone to manually deal with open proxies -- you really **MUST** automate this process due to the transaction volume. Also note that you are (potentially) talking about a LOT of DNS queries, so automate intelligently.

And of course...

- If you decide to automatically block email traffic from open proxies, you WILL end up using a DNSBL since that's basically the only scalable approach. :-)
- Some nice introductions to using DNSBL's with sendmail is available at <http://mail-abuse.org/rbl/usage.html>

http://hatcheck.org/proxy

Back Forward Reload Stop

Home Bookmarks The Mozilla Organiza... Latest Builds

Proxy ☐ Verbose [Clear](#)

Ports

Testing lsanca1-157-036.biz.dsl.gtei.net [4.3.157.36] - please wait...

PERMBLOCK Open Proxy - <http://openrbl.org/4.3.157.36>

DSBL Insecure host <http://dsbl.org/listing?4.3.157.36>; DSBL Unconfirmed subm

Test complete - no proxies found (0)

```
0: Using mail server: 198.78.66.238 (hatcheck.org)
1: Trying addr '4.3.157.36' port '80' proto 'http' ... cannot connect
2: Trying addr '4.3.157.36' port '3128' proto 'http' ... cannot connect
3: Trying addr '4.3.157.36' port '8080' proto 'http' ... cannot connect
4: Trying addr '4.3.157.36' port '8081' proto 'http' ... cannot connect
5: Trying addr '4.3.157.36' port '1080' proto 'socks4' ... connected
6: >>> binary message: 4 1 0 25 198 78 66 238 0
7: <<< TIMEOUT: timeout waiting for response
8: Trying addr '4.3.157.36' port '1080' proto 'socks5' ... connected
9: >>> binary message: 5 1 0
10: <<< TIMEOUT: timeout waiting for response
11: Trying addr '4.3.157.36' port '23' proto 'telnet' ... cannot connect
12: Trying addr '4.3.157.36' port '23' proto 'cisco' ... cannot connect
13: Trying addr '4.3.157.36' port '23' proto 'wingate' ... cannot connect
14: Test complete - no proxies found
```

Related links [hatcheck.org | openrbl.org] 74fb

Sometimes black is white (or grey, or red, or ...)

- Note this example used the same dotted quad, tested the same day as the DNSBL tests
- It can be disturbing to find that doing a fully functional test of a dotted quad listed in a DNSBL sometimes doesn't result in consistent results...
- Surely, in an ideal world, DNSBLs and active open proxy testers would concur in calling a host an open proxy (or not an open proxy) -- but we don't live in an ideal world.

Sources of inconsistency

- Some possible sources of inconsistency between DNSBL's and proxy testers include
 - 1) a formerly open proxy may truly no longer be open, but no one has gotten that dotted quad delisted from all the various DNSBLs out there right now.
 - 2) the open proxy may still be open, but may only be intermittently available (e.g., an open proxy running on a desktop that is only powered up 8-5 local time).

Sources of inconsistency (2)

- 3) The fully functional open proxy tester may be getting firewalled by the open proxy operator, their Internet service provider, or their ISP's upstream provider, even though the open proxy itself may still be available from other locations on the Internet.
- 4) The open proxy may be running on an uncommon port, or may be periodically changing the port(s) it is using to hinder detection (or to evade upstream filtering of common open proxy ports by the ISP).

Sources of inconsistency (3)

- 5) The open proxy may only be open for a limited range of services (e.g., web browsing, but not SMTP traffic transmission, for example), and the proxy tester might be checking the proxy only for some service it doesn't offer (like SMTP).
- 6) The open proxy server may have been running on a dynamically allocated address, and its lease may have expired (allowing that address to be recycled for use by some other innocent/secure host).

Sources of inconsistency (4)

- 7) An actively abused open proxy server may be completely saturated, resulting in TCP timeouts or other odd errors.
- 8) Proxy servers may accept incoming connections on one address and create outgoing connections on a completely different address. Testing an output ("apparent source") interface rather than an input interface may result in incorrect inferences being made.

Sources of inconsistency (5)

- 9) The putative open proxy may NEVER have been truly open, although it may have exhibited suspicious behaviors (e.g., it may have open ports on numbers strongly associated with open proxies, e.g., 1080 or 6588, etc.), or a host may have been maliciously nominated as an act of retribution (a so-called "Joe-job"), etc. [Most DNSBL's require evidence and validate user submissions, but there are exceptions; know your BL's listing criteria!]

Scanning via NMAP or specialized proxy hunting tools

- Administrators may use a general purpose scanning tool such as NMAP (<http://www.insecure.org/nmap/>) to scan their own hosts or own networks to identify potential open proxies; there are also specialized proxy detection and analysis tools in widespread circulation such as Proxy Hunter, Proxy Sniper, etc. (see: <http://www.proxys4all.com/tools.shtml>)

And speaking of scanning...

- Scanning someone else's host(s) or someone else's network(s) without their permission may be/is unlawful (at least in some states) and is not recommended (although we empirically know it is a common practice).
- The open proxy delisting paradox: "If one believes a host to be an open proxy, how is one to learn that that host is no longer an open proxy if the owner doesn't know of your belief and active scans are unlawful?"

Common proxy ports

- If you are scanning for proxies, it is helpful to know the ports to watch for, although of course any server (including an open proxy server) can be bound to any arbitrary TCP/IP port. Some proxies may be running on well known ports such as 80 (http) or 443 (https):
 - SOCKS 4/5: 1080
 - HTTP: 3128, 8080, 6588, 80, 81, 4480
 - Wingate: 23
 - Peekabooby/Triangleboy/etc.: 443

To manually test a connect mode open proxy

- Telnet to the open proxy port then enter:

```
CONNECT foo.bar.baz:25 HTTP/1.0  
<return>  
<return>
```

If you see `200 Connected` you know that you've found an open proxy that's willing to channel SMTP traffic to server `foo.bar.baz`

MRTG as an open proxy spotting tool

- Yet another way of spotting a *possible* open proxy server is by watching traffic graphs for individual switch ports where outgoing traffic closely mirrors incoming traffic. This technique is mentioned (and nicely illustrated) at:

<http://www.rsc-london.ac.uk/technical/network/monitoring/> (see the "spotting open proxy servers" section)

Or you can just wait for the complaints to pour in...

- The final way to identify open proxies on your own network is to do nothing, and simply wait for the complaints to come pouring in.
- **At a minimum EVERY DOMAIN should have a monitored abuse@<domain> address! See RFC 2142 at section 4!**
- <http://www.abuse.net/>
<http://www.rfc-ignorant.org/>

IX. Our Open Proxy List

The use-it-and-lose-it paradox

- One of the most delightful things about spammers using open proxies is that when a spammer sends spam through an open proxy, that act advertises the existence of that open proxy, facilitating its closure.
- Thus, when we'd see a "hit" against one or more of the open proxy DNS blacklists, or notice a new open proxy spamming us directly, we'd add an entry for that host to:
<http://darkwing.uoregon.edu/~joe/open-proxies-used-to-send-spam.html>

Tracking open proxies

- We began doing that in September 2002, systematically looking at all IP addresses associated with spam which slipped through our filters and which were reported to us, as well as at the IP addresses of all mail which had been rejected by filtering rulesets running on our shared systems. [You *could* just scrutinize ALL SMTP relay addresses seen in your logs, but you'll waste a lot of time and do a lot of pointless queries.]

You won't notice open proxies if you're drowning in other spam...

- If you're interested in identifying open proxies by tracking their appearance in spam, as we were, the first step is to carve off all the *other* sources of spam, e.g.:
 - direct-from-dialup spam
 - spam sent via open SMTP relays
 - spam sent via vulnerable formmail cgi's
 - spam sent from so-called "bulletproof" dedicated spamhouses

Blocking most non-proxy spam sources via DNSBLs

- While there are many ways of blocking the spam from those other sources, one combination that works fairly well is the mail-abuse.org RBL+ (not free, but quite affordable in zone transfer mode for universities), plus the free SBL from spamhaus.org. That combo will kill most spam not coming in from open proxies (although you may still need some supplemental local blocks).

Add an open proxy DNSBL

- After you've blocked those other spam sources, most of what's left will be spam from open proxies.
- You can either let that spam come in, wait for it to get reported, and then manually snag the relevant IP address from each spam's headers, or you can add one or more open proxy DNSBLs. If you add an open proxy DNSBL, you can then snag the addresses of potential open proxies (along with other spam sources) from your logs.

Pointers to some popular open proxy DNSBLs

- Blitzed: <http://www.blitzed.org/bopm/>
- Osirusoft:
<http://relays.osirusoft.com/faq.html>
- SORBS:
<http://www.dnsbl.sorbs.net/using.html>
- Wirehub:
<http://basic.wirehub.nl/blackholes.html>
- NJABL: <http://njabl.org/>
- ... and there are others.

Format of the open proxy list

- There are currently ~100K entries in a format that looks like:

```
[snip]
200.149.218.155 (02/25/2003) [PE218155.user.veloxzone.com.br] --OSW
200.149.218.156 (03/09/2003) [PE218156.user.veloxzone.com.br] ----N
200.149.218.161 (02/19/2003) [PE218161.user.veloxzone.com.br] B--SW
200.149.218.164 (03/20/2003) [PE218164.user.veloxzone.com.br] --OSWN
200.149.218.173 (12/22/2002) [PE218173.user.veloxzone.com.br] --OS
200.149.218.177 (02/25/2003) [PE218177.user.veloxzone.com.br] B-O-W
200.149.218.180 (04/03/2003) [PE218180.user.veloxzone.com.br] ---SW-
200.149.218.187 (02/27/2003) [PE218187.user.veloxzone.com.br] ----W
200.149.218.188 (02/19/2003) [PE218188.user.veloxzone.com.br] B-O-W
200.149.218.206 (03/20/2003) [PE218206.user.veloxzone.com.br] ---SW-
200.149.218.208 (01/23/2003) [PE218208.user.veloxzone.com.br] B--SW
200.149.218.229 (03/09/2003) [PE218229.user.veloxzone.com.br] ---SWN
200.149.218.238 (03/20/2003) [PE218238.user.veloxzone.com.br] ---SW-
200.149.218.245 (03/20/2003) [PE218245.user.veloxzone.com.br] ---SW-
200.149.219.15 (01/10/2003) [PE219015.user.veloxzone.com.br] open on 1080 and 3128
200.149.219.37 (03/20/2003) [PE219037.user.veloxzone.com.br] ---SW-
200.149.219.41 (04/03/2003) [PE219041.user.veloxzone.com.br] ---SWN
[snip]
```

Format of the open proxy list (2)

- The entries are maintained in numeric order by dotted quad, one entry per line.
- Each line shows the dotted quad in question, the date DNSBLs were checked for that address, the hostname associated with the dotted quad (or "no reverse DNS" if applicable), and a mask showing which open proxy DNSBLs listed the address at the time it was checked/listed (and possibly information about the ports the proxy used)

Coding of DNSBL proxy entries

- The three to six character mask at the end of each entry is encoded using the scheme:

B	opm. b litzed.org
-	[used to show a now-omitted DNSBL]
O	relays. o sirusoft.com (127.0.0.9)
S	dnsbl. s orbs.net (127.0.0.2, 127.0.0.3, and 127.0.0.4)
W	proxies.blackholes. w irehub.net
N	dnsbl. n jabl.org (127.0.0.9)

- When a host isn't listed on a given DNSBL, a dash is entered as a placeholder

"Wait a minute! By publishing that kind of list, you're just making the problem worse!"

- No. There are already plenty of open proxy lists in existence, and those lists routinely include information (such as port numbers) that amateur/bulk proxy abusers need. My list *only* includes port numbers in limited circumstances (for example, when I'm documenting a proxy that isn't otherwise listed on a DNSBL we use, or I've personally received spam via that proxy).

Hardcore proxy abusers don't use hosts from public lists...

- Known open proxies tend to be saturated and slow, so professional open proxy abusers tend to scan for their own "fresh" proxies, buy private lists of open proxies from proxy scanning specialists, or trade open proxies among themselves. (For some sense of that activity, albeit on a hobbyist/casual scale, search for proxy or proxies in groups.yahoo.com or groups.msn.com)

Don't shoot the messenger

- The first step to fixing any problem is dragging it out from the shadows into the light of day. If you refuse to talk about a problem, it will never get fixed. The open proxy problem NEEDS to get fixed.
- Unless you can document and detail a problem, many ISPs are unwilling to take action to fix that problem.
- People need to see the full extent of the problem to appreciate the need for large scale corrective action.

Besides...

- *Anyone* who gets spammed and has access to sendmail logs, web server logs, firewall logs, etc. could build a similar list; I'm not doing something magic here...
- On the other hand, we do know that our list gets retrieved LOTS of times every day, sometimes via open proxies (which we dutifully add to the list). :-)

What domains are seen on the open proxy list?

- 27.3% no reverse DNS
- 5.2% telesp.net.br
- 3.4% prodigy.net.mx
- 3.3% veloxzone.com.br
- 2.6% rr.com
- 2.6% wanadoo.fr
- 2.1% interbusiness.it
- 2.0% dsl-verizon.net
- 1.8% telecom.net.ar
- 1.7% hinet.net
- 1.6% pacbell.net
- 1.5% attbi.com
- 1.4% swbell.net

What domains are seen on the open proxy list? (2)

- 1.0% ameritech.net
- 1.0% skynet.be
- 0.9% comcast.net
- 0.9% btopenworld.com
- 0.8% rima-tde.net
- 0.8% bellsouth.net
- 0.8% speedyterra.com.br
- 0.8% brasildtelecom.net.br
- 0.8% ntl.com
- 0.7% tpnet.pl
- 0.6% advancedsl.com.ar
- 0.6% virtua.com.br
- 0.6% rogers.com

What domains are seen on the open proxy list? (3)

- 0.5% shawcable.net
- 0.5% ttd.es
- 0.5% sympatico.ca
- 0.5% prima.net.ar
- 0.5% bigpond.net.au
- 0.4% charter.com
- 0.4% adelphia.com
- 0.4% t-net.net.ve
- 0.4% t-dialin.net
- 0.4% chello.nl
- 0.3% videotron.ca
- 0.3% brdterra.com.br
- 0.3% terra.com.br

What domains are seen on the open proxy list? (4)

- 0.3% `cox.net`
- 0.3% `telepar.net.br`
- 0.3% `speedy.com.ar`
- 0.3% `vtr.net`
- 0.3% `wanadoo.net`
- 0.3% `tele.dk`
- 0.3% `a2000.nl`
- 0.3% `optonline.net`
- 0.3% `papalegua.com.br`
- 0.3% `bezeqint.net`
- 0.2% `blueyonder.co.uk`
- 0.2% `verizon.net`

What domains are seen on the open proxy list? (5)

- 0.2% infoweb.ne.jp
- 0.2% ono.com
- 0.2% surfer.at
- 0.2% ethome.net.tw
- 0.2% telekom.at
- 0.2% telus.net
- 0.2% proxad.net
- 0.2% hispeed.ch
- 0.2% att.net
- 0.2% planet.nl
- 0.2% arcor-ip.net
- 0.2% hansenet.de
- 0.2% axelero.hu

What domains are seen on the open proxy list? (6)

- 0.2% `Ajato.com.br`
- 0.2% `uswest.net`
- 0.2% `mindspring.com`
- 0.2% `fibertel.com.ar`
- 0.2% `hkcable.com.hk`
- 0.2% `charter-stl.com`
- [all others contributed less than 0.2%]

The no reverse DNS folks

- The same people who can't securely configure their proxies obviously also don't give a damn about inaddr's. :-)
- In some cases, the lack of reverse DNS may be due to domain names not being "relevant" (e.g., at sites that use non-roman languages), but other ISPs may *intentionally* not provide a reverse address in an effort to reduce the number of complaints they receive... That's okay, we 'll soon be mapping those dotted quads to responsible parties via other means₁₄₅

Too big to block?

- If you meditate on the country code distribution shown in that list, you can see why some use country-wide blocks, even if they do inflict lots of collateral damage.
- There are some folks on that list who should (and do) know better than to ignore open proxies on their network. They may have apparently come to believe "we're too big to get blocked," or "we don't want to cut off *any* paying customer, even if they are insecure -- we'll just ignore the complaints."¹⁴⁶

Fast connections (except from higher education) are beloved

- Clearly, there is an association between connection speed and open proxy presence; fast connections are more likely to be trying to do connection sharing, and because those connections are fast, they tend to be attractive to abusers.
- For the most part, higher education sites do NOT tend to show up much, which is excellent news (and contrary to some commonly articulated popular perceptions).

And yes, some open proxies have been listed "forever"

- It is absolutely true that there are some proxies on the list that have been listed for a REALLY long time, e.g., since October 2002 in some cases. What can I say? Some people simply may not care if they have an open proxy; in other cases, the proxies may be secured, but the system owner may not know how to get off a DNSBL we use, or may not care to bother.

If a host is already on the list...

- If you do keep a local list like ours, it is easy to forget (as you process your logs, looking for new open proxies to add), that you can skip any address you've already got listed -- you don't need to requery all the open proxy blacklists if you've got that address already locally listed. Grep your local list first! [duh]

Taking entries off the list

- Periodically we recheck the blacklists for all the entries on our list and remove the dotted quads that are no longer listed on any of the five used.
- Retesting can become, um, tricky, when you're talking about doing half a million queries (~100K hosts X 5 DNSBLs).
- It currently takes roughly half a day to do half a million retests... yes, we could make the rechecks faster/more aggressive, but we need to be careful of our impact on DNS servers...

Effect of adding and deleting DNSBLs

- During the time we've been maintaining this list, we've added and deleted a number of DNSBLs from coverage, and have done off-list testing of some additional DNSBLs (some of which we have elected to not add)
- When DNSBLs are added or deleted, you may see a noticeable jump or drop in the number of entries listed; this is not a sign of data problems. (see the graph on slide 93)

IX. "What Can I Do?"

Chip in...

- The most important step, if you see spam from an open proxy that isn't already listed at sites such as OpenRBL, is to report it. Open proxy DNSBL's develop better coverage and work better for all of us as more people use and contribute to them.
- One of the best ways to report spam you may receive is via <http://spamcop.net/>
- Be sure to also train your end users how to report spam they may receive!

Make sure you aren't part of the problem...

- If you run a proxy server, review your config and your log files for problems.
- If you are responsible for your campus' network, make sure it isn't infested with open proxy servers.
- Review your acceptable use policy to insure that you've disallowed open proxy servers, either by name, or via general prohibitions on "unauthorized resource sharing"
- Make sure you've got an abuse@ address

Protect your own mail servers

- Use an open proxy DNSBL to protect your own mail servers, just as you may already reject mail from open SMTP relays.

Blocking traffic from open proxies is a basic step that a growing number of major ISPs are already doing. For example:

- <http://postmaster.info.aol.com/ops.html>
- http://security.rr.com/mail_blocks.htm
- <http://help.yahoo.com/help/us/mail/defer/defer-02.html>

Which open proxy DNSBL should you use?

- Picking a DNSBL to use in production (e.g., to protect a mail server) involves evaluating a variety of different factors, including:
 - what's the maintainer's reputation?
 - does the DNSBL catch the open proxies you're seeing? (how inclusive is it?)
 - does the DNSBL remove open proxies when they are no longer open? (and does it do so automatically? upon request? or?)
 - is the DNSBL fast? reliably available?

Some of those factors are hard to evaluate objectively...

- ... but it is easy to evaluate "coverage" or "inclusivity". Looking just at what's currently being "caught," the current ranking looks like:

Sorbs:	98.1%
Wirehub:	90.1%
NJABL:	47.1%
Blitzed:	19.0%
Osirusoft:	14.0%

If you want to use just one DNSBL...

- We'd suggest picking either Sorbs or Wirehub (they tend to be fairly congruent) :

DNSBL COVERAGE PATTERNS WHICH INCLUDE BOTH SORBS AND WIREHUB:

34135	36.7%	---SW-
20871	22.4%	---SWN
8954	9.6%	B--SWN
5191	5.6%	--OSW-
3067	3.3%	B--SW-
2410	2.6%	--OSWN
2233	2.4%	B-OSWN
692	0.7%	B-OSW-

- But what should you pick if you wanted to use a second complementary DNSBL?

A second DNSBL...

- If you decided to pick Wirehub as a first DNSBL, a good 2nd choice might be Sorbs or NJABL. If you added Sorbs to Wirehub:

5097	5.5%	---S--	
1183	1.3%	---S-N	
684	0.7%	--OS--	
468	0.5%	----W-	
216	0.2%	B--S-N	
198	0.2%	B--S--	
121	0.1%	--OS-N	(no remaining pattern with >100 hits)

If you added NJABL to Wirehub:

4232	4.5%	-----N
1183	1.3%	---S-N
468	0.5%	----W-
446	0.5%	B-----N
216	0.2%	B--S-N
121	0.1%	--OS-N

A second DNSBL... (2)

- If you decided to pick SORBS as your first DNSBL, a good 2nd choice might be NJABL:

5097	5.5%	---S--	
4232	4.5%	-----N	
1183	1.3%	---S-N	
684	0.7%	--OS--	
446	0.5%	B-----N	
216	0.2%	B--S-N	
198	0.2%	B--S--	
121	0.1%	--OS-N	(remaining patterns each <100 hits)

Educate downstream partners

- Some I2 sites/state networks are already aware of the open proxy issue, and are doing a good job getting the word out to their downstream partners.

For example, see:

<http://www.more.net/security/advisories/2002/020304.html>

You could do likewise...

Educate the carriers

- If you buy transit bandwidth or negotiate peering with carriers, don't miss that opportunity to beat the drum about the problem of open proxies. Carriers are NEVER more receptive to your feedback than when they're trying to make a sale. Insist that they describe the steps they take to deal with open proxy abuse, and spam in general, before you sign that P.O.

Finally...

- You may want to become involved at the state level in promoting anti-spam laws which address open proxy server abuse.
- Twenty six states have some sort of anti-spam law as of the start of this year -- how about yours? (see <http://spamlaws.com/>)
- If you don't have one, work with your state Attorney General's office to get one passed, or volunteer to provide technical assistance.

X. Conclusion

Outcomes in a nutshell

- We believe we have a steadily improving understanding of how many open proxies are out there, where they are located, and how they can be efficiently blocked using one or more DNS blacklists.
- Many insecure open proxies are getting closed, although the problem is far from over.
- It is becoming steadily harder for spammers degrade your email by sending "untraceable" spam.

Some future work

- Open proxies without reverse addresses need to be attributed to responsible entities.
- Correlation of open proxies with spamvertised URLs needs to be done to establish what spamvertisers are using open proxies to send spam (this will be increasingly important as states pass legislation making that behavior unlawful).
- Our data should be provided in formats other than just a flat IP-sorted web page.

Acknowledgments

- While I am solely responsible for the content and opinions expressed in this document, I would like to thank a number of people who have provided invaluable support and/or technical assistance on this project, including Joanne Hugi, my boss and the Associate VP for Information Service; Steve VanDevender and Bob Jones of the Computing Center Systems group; Jon Miyake, Computing Center Acceptable Use Officer (and Perl expert); the whole Computing Center Network Services DNS crew (particularly John Kemp and Jason Edmiston); all the people who offer DNSBLs or other antispam tools to the net; and my family, which has patiently put up with my latest obsession.

Thank you!

- Thanks for your patience with this long talk so late in the day.
- Questions?