What Remains To Be Done in Cyber Security

Joe St Sauver, PhD
joe@uoregon.edu or joe@internet2.edu
Manager, Internet2 Nationwide Security Programs
and Manager, InCommon Certificate Service

University of Michigan SUMIT_2012
Ann Arbor, MI, 9:30-10:20 AM, 19 Oct 2012

http://pages.uoregon.edu/joe/to-be-done/

Disclaimer: all opinions expressed in this talk are those of the author, and do not necessarily represent the opinion of any other entity.

I. Introduction

"One never notices what has been done; one can only see what remains to be done."

Marie Curie (the first woman to receive a Nobel Prize), from a 1984 letter to her brother.

Good Morning!

- Let me begin by saying what a pleasure it is to be here with you in Ann Arbor this morning.
- I'd like to thank **Paul Howell**, my friend/colleague (and the University of Michigan's Chief IT Security Officer), for the invitation to talk with you today.
- I'd also like to thank **Kim Wheeler**, who works with Paul, for doing a terrific job on the logistics for my visit.
- You may have suspected that there's an existing connection between Paul Howell and myself, and if so, you're right. Paul and I both are part of the **Higher Education Information Security Council (HEISC) Leadership Team**, as well as working on a variety of other info-security work together. I appreciate Paul extending an invitation for me to talk today *in spite* of all he knows about me.

My Somewhat Odd Slide Style

- In particular, Paul knows that **my slide style is odd**.
- **I know, I really do, how I'm supposed to do PowerPoint** talks (lots of graphics, 3 to 5 bullets per slide, using hidden speaker notes, etc.). I just tend to do things a bit differently, as you'll notice by the time we're done today.
- I use this different style intentionally for multiple reasons:
  -- I take my accessibility obligations for the **deaf and hard of hearing** very seriously. Some members of today's audience may also not be native English speakers
  -- I know that while there are hundreds in this audience today, **others will look at these slides later, online**
  -- If I don't carefully script my remarks, I know that **I'll run out of time when I still have much to say**
- Oh yes: don't assume you need to "read along" with the slides as we go through them – I won't be!

The Level of This Talk

- It's always hard to get the right level for security talks.
- Many audiences are mixed, with some listeners up to their eyeballs working security issues, but with other folks starting from ground zero; similarly, some are technical, while others are more managerial.
- **Events open to the public, like this one, can be particularly tricky when it comes to hitting the right level.**
- My plan today is to try to talk about some pressing issues, including a little background for context as we go along. Even if you don't "get all the details," hopefully you'll still get a sense of some of the threats we currently face.
- On the other hand, if some or all parts of this talk are just a "rehash" for you, I apologize in advance for "shooting too low."

Why I'm Here With You Today

- I'm here today to talk, but **I'm actually here to ask for your help. We need your help when it comes to tackling the cybersecurity work that remains to be done.**
- Some of this work is very applied and pragmatic and "hands on," while in other cases there's fundamental "clean slate" research that still needs to happen. There's work enough for everyone (and then some!)
- Given that UM is #1 in research among public universities for the second straight year,[1] I'm really hoping that we can turn some of your school's formidable research prowess toward unsolved issues in cyber security.
- There are certainly many pressing issues out there...

[1] http://www.ns.umich.edu/new/releases/20311-new-federal-rankings-u-m-again-tops-in-research-spending-at-public-universities

DHS Cybersecurity Roadmap's "Hard Problems in Information Security Research"

- If you ask "What are our biggest system and network cybersecurity issues?" you may get different answers from different people.
- For example, a few years back, Douglas Maughan of the Department of Homeland Security (DHS) lead a series of community workshops in an effort to identify the toughest cybersecurity problems faced by the community.
- The result was the November 2009 DHS report entitled "A Roadmap for Cybersecurity Research," www.cyber.st.dhs.gov/docs/DHS-Cybersecurity-Roadmap.pdf
- That report identified <u>eleven hard problems</u> in information security research as of that time.

1) **Scalable trustworthy systems** (including system architectures and requisite development methodology)
2) **Enterprise-level metrics** (including measures of overall system trustworthiness)
3) **System evaluation life cycle** (including approaches for sufficient assurance)
4) **Combatting insider threats**
5) **Combatting malware and botnets**
6) **Global-scale identity management**
7) **Survivability of time-critical systems**
8) **Situational understanding and attack attribution**
9) **Provenance** (relating to information, systems, and hardware)
10) **Privacy-aware security**
11) **Usable security**

That's An Interesting List, I Think

- Some of those issues are unquestionably hard problems in cybersecurity, and eminently worthy of attention. I'm delighted that some of those issues are currently the focus of substantial work by the community.

- **However, what I found particularly striking (at least as a participant in a couple of those workshops) is one issue that *didn't* make that list: the problem of distributed denial of service ("DDoS") attacks.**

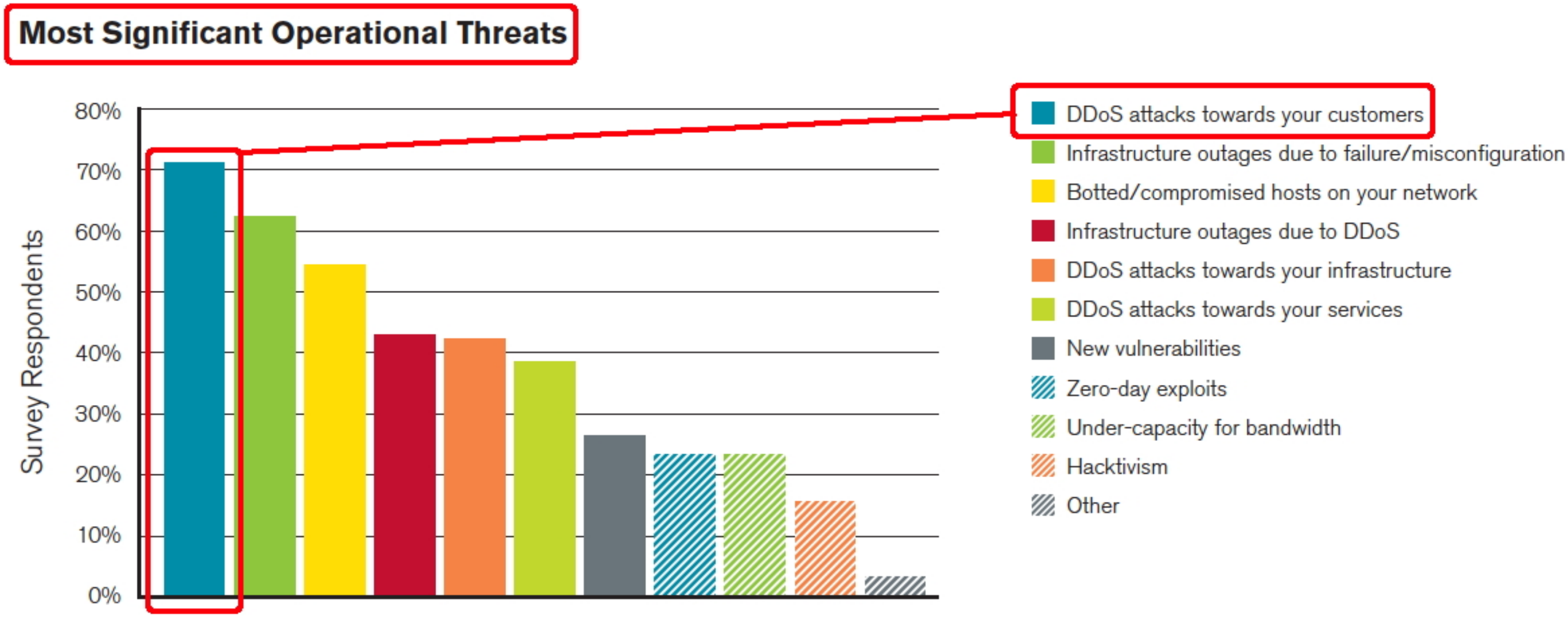
- **At least from the perspective of network operators, DDoS attacks are their #1 operational worry.**

- Let's start there.

II. Distributed Denial of Service Attacks (DDoS)

May 2005: "Explaining Distributed Denial of Service Attacks to Campus Leaders,"
http://pages.uoregon.edu/joe/ddos-exec/ddos-exec.pdf

The Most Significant Operational Threat: "DDOS Attack Towards Your Customers"



Source: Arbor Networks, Inc.

Worldwide Infrastructure Security Report, 2011 Volume VII
http://ddos.arbornetworks.com/report/ at page 12

What *IS* a DDoS Attack?

- In a DDoS, attackers attempt to render a targeted site inaccessible (or unusably slow) by flooding it with traffic or otherwise 'using up' its available capacity. For example:
  -- Sending **floods of spoofed UDP traffic**, perhaps by leveraging DNS amplification to multiply the volume of attack traffic, saturating a connections inbound bandwidth
  -- **SYN flooding** a site, starting (but not completing) TCP connection after TCP connection after TCP connection, until there's no capacity for new connections to be accepted
  -- **HTTP GET attacks**, whereby thousands or tens of thousands of systems collaborate to repeatedly request large files from a system, thereby saturating the system or its outbound bandwidth
- Note that these are just examples. There are many other types of DDoS attacks that you may also encounter.

The Objective of Denial of Service Attacks

- Regardless of HOW someone executes a DoS attack, the OBJECTIVE is the same: the attacker is trying to keep others from being able to access a site or use a resource.
- This is effectively the online version of:
  -- calling in repeated bomb threats against a site
  -- autodialing someone's personal phone number 1,000's of times a day (or getting others to call for you, perhaps by offering "free pizza" to the first 1,000 people who call)
  -- squirting glue in a keyhole, thereby keeping users from being able to open the door until they replace the lock.
- "DOS" attacks are obviously not just an "online thing"
- Motivations for DoS'ing a site may include revenge, political objectives, attempts at extortion ("if you don't want me to keep on DoSing you, pay me $X"), etc.

Examples of Some Recent DoS Attacks

- "**AT&T** hit by DDoS attack, suffers DNS outage"
  www.pcworld.com/article/260940/atandt_hit_by_ddos_attack_suffers_dns_outage.html

  > *"A distributed denial-of-service attack aimed at AT&T's DNS (Domain Name System) servers has disrupted data traffic for some of the company's customers.* ***The multi-hour attack*** *began Wednesday morning West Coast time and* ***at the time of this writing, eight hours later, does not appear to have been mitigated."***

- "**Major banks** hit with biggest cyberattacks in history"
  money.cnn.com/2012/09/27/technology/bank-cyberattacks/
- "Hackers Take Down **Stock Exchange Sites** As Mideast Conflict Heats Up Online"
  www.forbes.com/sites/calebmelby/2012/01/20/hackers-take-down-stock-exchange-sites-as-mideast-conflict-heats-up-online/
- "DDoS attack of rare power behind **WikiLeaks** take-down"
  gcn.com/articles/2012/08/13/wikileaks-ddos-attack-trapwire.aspx
- "**Burma** hit by massive net attack ahead of election"
  www.bbc.co.uk/news/technology-11693214

DoS Attacks by the Numbers

- "The company reports that it has fended off more malicious traffic in the first three months of 2012 than it did in all of 2011 – **9.5 petabytes of raw data and 408 trillion network packets.**"
  arstechnica.com/business/2012/04/bad-bots-ddos-attacks-spike-in-first-quarter-outdoing-all-of-2011/
- "DDoS attacks: **150Gb per second and rising**"
  www.zdnet.com/ddos-attacks-150gb-per-second-and-rising-7000005075/
- "Distributed Denial of Service – Deep Dive"
  wwwns.akamai.com/rsa_2011/RSA_NOCC_DDoS.pdf

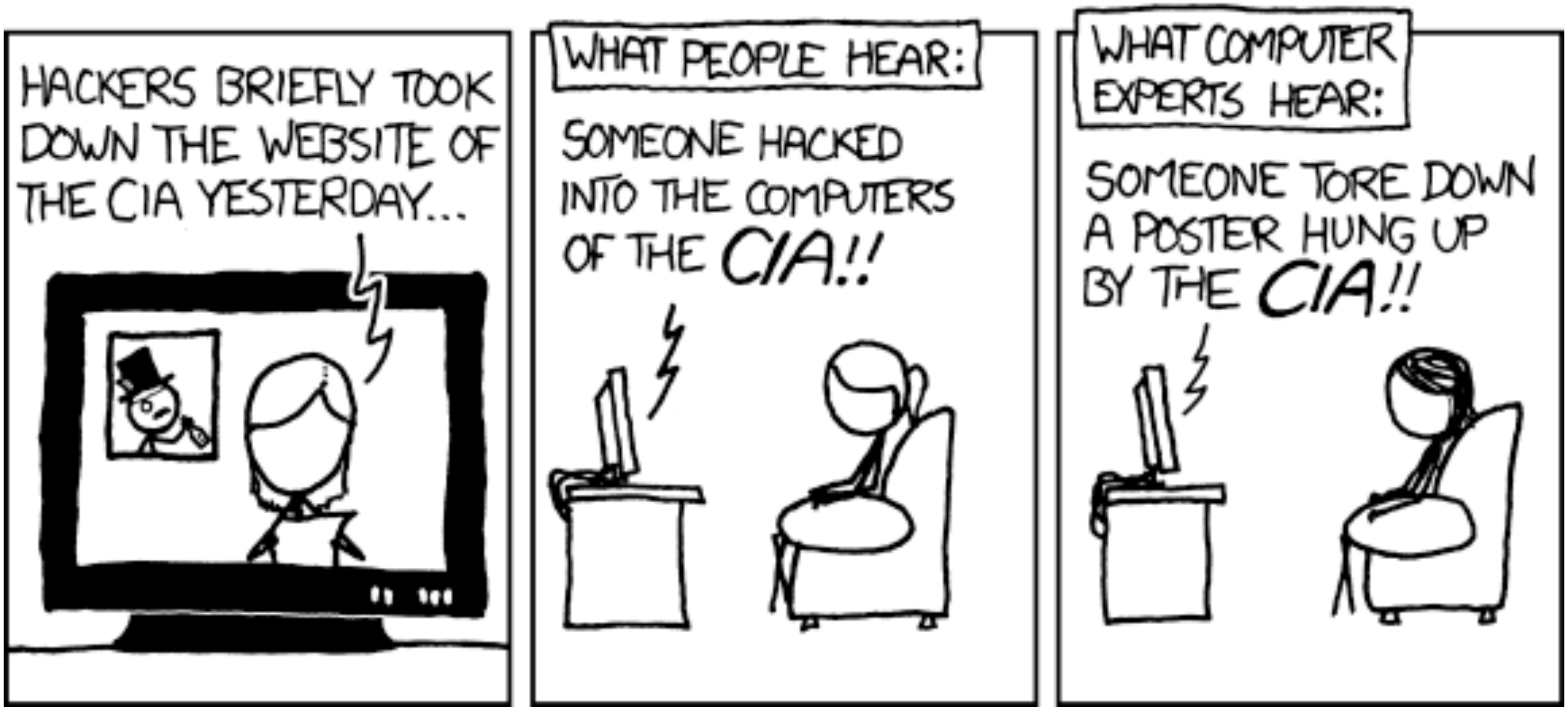
  "The Largest DDoS Ever Recorded: July 4th 2009 [...]
  795,000 page views a second
- "DDoS attacks in H2 2011"
  www.securelist.com/en/analysis/204792221/DDoS_attacks_in_H2_2011

  *The **longest DDoS attack** in the second half of the year lasted for **80 days, 19 hours, 13 minutes** [...]*
- These are some BIG attacks IMHO.

DoS Attacks vs. Site Breaches

- It's probably worth emphasizing that a DDoS attack is *not* the same as a site compromise.
- When a site gets compromised, a cracker gets unauthorized access to a system, perhaps by guessing or intercepting the administrator's password, or dropping malware that installs a backdoor. When that happens, the hacker/cracker may have complete control of the system. Securing it will normally require rebuilding that system from scratch. That rebuilding/reinstalling process may take the site offline for a bit, but that's NOT how a site is normally DDoS'd, that down-for-a-while-for-a-fixup is just a "side effect."
- On the other hand, if a site is "just" being DoS'd, as soon as the DoS stops, the site will be back to normal (perhaps modulo any backlogs that may have accumulated).

An Example of How People Confound Intrusions with DDoS Attacks


HACKERS BRIEFLY TOOK DOWN THE WEBSITE OF THE CIA YESTERDAY...
WHAT PEOPLE HEAR:
SOMEONE HACKED INTO THE COMPUTERS OF THE CIA!!
WHAT COMPUTER EXPERTS HEAR:
SOMEONE TORE DOWN A POSTER HUNG UP BY THE CIA!!

http://xkcd.com/932/
Creative Commons Attribution-NonCommercial 2.5 License

Sometimes DDoS Attacks MAY, In Fact, Actually Just Be An Irritation

- For example, it's nice when researchers can access the CIA World Factbook on the CIA's web site, but if the CIA's public web site is temporarily unusable due to a DDoS, it's unlikely that having that site down would impact the CIA's meat-and-potato intelligence collection and analysis activities in any substantive way.

- Presumably those core agency mission activities are not co-hosted with, and do not rely on the availability of, that agency's public website :-;

Other Times, DDoS Attacks Against Public Websites Can Be Hugely Problematic

- Imagine a site that does virtually all of its business online
- Assume that the business has concentrated periods of time that are "make or break" periods for that company (hypothetically, think about things like the run up to Christmas for retailers, or the period right before big sporting events for online gambling businesses)
- Let's assume that a targeted business has many competitors who'd love to have a shot at serving (and permanently stealing!) their customers, and that customers have little or no "lock in" with their current vendor (e.g., switching costs are low or non-existent)
- In cases like that, DDoS attacks against the organization's public web site can truly be a huge issue.

But We Don't REALLY Know Just How Bad The DDoS Problem Currently Is

- Unlike breaches involving personally identifiable information (PII), which are a category of security incident that must be reported in most states, **there's no requirement that sites report DoS attacks.**
- Because there's no requirement to report, not surprisingly, many DoS attacks never get reported. The ones that you do see in the media are usually just the most awful ones, e.g., the ones that are "too big to hide," or the ones that have been heavily advertised by the "perps" themselves
- **We need better data collected about DDoS attacks, including their frequency and duration, their magnitude, the mechanism(s) employed to effect the attacks, the targets, the approach used to mitigate the attack, etc.**

DoS Attacks, Collateral Damage & Mitigation

- Imagine that you have the bad luck to be on the same shared server as the target of a DDoS attack. When your online neighbor, the target of a DDoS, gets hit and taken offline, so do you, simply because you have the bad luck to share resources with the actual DDoS target.
- This is obviously a case of "collateral damage."
- Collateral damage is one of the biggest problems associated with distributed denial of service attacks, and minimizing collateral damage is a prime reason why some sites "mitigate" DDoS attacks by taking the target of the attack offline.
- While this serves to protect other innocent customers, it also effectively "perfects" or "completes" the DDoS against the DDoS target. Great "mitigation" technique, eh?

"We'll Just Call the FBI and Let Them Sort It All Out"

- The FBI and other law enforcement officials will often be interested in major DDoS attacks, however **their attention will not provide symptomatic relief when a DoS occurs**, nor is it a guarantee of a successful investigation and eventual prosecution – DDoS cases can be hard to work.
- You should also understand that many times denial of service attacks are transnational, which introduces special investigatory issues, and requires FBI coordination with foreign LE counterparts, which can introduce substantial investigative delays. Denial of service attacks committed by individuals overseas (and attacks made by minors whether here in the US or abroad), may also result in disappointingly short sentences. This may dampen LEA enthusiasm for proceeding with a potentially hard-to-investigate, hard-to-prosecute, low-payoff case.

"Why Can't DDoS Attack Traffic Just Be Filtered At Our Site's Firewall?"

- Assume the attack is a packet flooding attack. Even if that traffic gets blocked at the site's perimeter firewall, it will already have traversed the site's Internet link, filling it to overflowing. Discarding that traffic at the firewall won't do anything to help (it's too late at that point). To actually help, the packet flooding traffic would need to be **blocked upstream**, <u>before</u> it hits the site's Internet link.
- Some network operators DO allow customers to announce so-called "blackhole routes" that discard any traffic intended for a specific IP address (typically a DDoS attack target).
- For example, Internet2 offers its connectors this option...

BGP Discard Routing

"Internet2 Network Connectors can now advertise routes to the Internet2 Network via BGP for which all traffic to those routes will be discarded by the Internet2 Network routers. [...] if a more specific route is tagged with the BGP Community 11537:911 and the mask length is between /24 and /32, the route advertisement will be accepted and the NEXT-HOP will be set to the discard interface causing all packets destined to that route to be discarded by Internet2 Network router(s)."

https://wiki.internet2.edu/confluence/download/attachments/17383/Internet2+Network+Transit+Security+Policy.pdf

But Remember, While This Protects A Shared Link, It Finishes The Attacker's Job For Them

- When you announce a blackhole route for the target of a DDoS, the attacker who was DDoS'ing you gets exactly what they wanted: the targeted site is down. Everyone else may be back to normal, but the targeted site is offline.
- At this point, some people may ask, "Well, why not filter the unwanted traffic based on the *source* of the attack traffic? That way you could block just the systems that are originating the attack traffic, while still letting legitimate traffic through to the site!"
- That would be terrific, but the problem is this: **imagine an attack that appears to come from tens or even hundreds of thousands of different source addresses**.

Many Networks Are Poorly Instrumented

- If you're going to try to filter attack traffic based on source addresses, you need to know "What are the addresses that I should ACL (e.g., filter)?" Unfortunately, many networks are poorly instrumented. In those cases, operators may have no ability to see the traffic targeting them whatsoever. This is particularly true as network speeds increase to 10Gbps or even 100Gbps.
- In other cases, the operator may at least have Netflow data exported from their routers, but to keep the load on those routers reasonable (e.g., to allow the routers to have enough capacity to actually do its primary job of routing packets), Netflow data may only be "sampled" – that is, you might see every hundredth flow or every thousandth flow. That might be good enough for billing purposes, but may not be good enough for attack mitigation purposes.

Network Filtering Capacity May Also Be Limited

- There are limits to the amount of filtering that a router or other network device can reasonably be expected to perform.

- Even if you are able to successfully identify the traffic you want to filter, your network devices may simply not have the capacity to do large-scale fine-grained filtering of hundreds of thousands of source IP addresses (admittedly, some devices are better when it comes to doing this than others, but all network devices have practical limits when it comes to scrubbing unwanted traffic)

Other Things That Some People Try

- You can overprovision your bandwidth and system capacity (but that's no guarantee that the bad guys won't be able to simply outscale you, or use a more subtle DDoS attack vector rather than just flooding you with packets).
- You can also use a content delivery network (not cheap), or purchase DDoS-resistant hosting from a company that specializes in that sort of thing (but again, that sort of specialized hosting usually isn't cheap, and you may find yourself hanging out with lots of other sites that are also "DDoS magnets," perhaps not what you want...)
- You can try using specialized architectures (such as architectures built around reverse proxies, or anycast infrastructures), or you can try site agility (e.g., use of domain names with short TTLs, allowing you to move your site around from one targeted IP address to another)...

Bottom Line, When It Comes to DoS Attacks

- A determined adversary **can** take your site down, or pretty much ANY site down, and keep it down as long as they want, or at least they can make you struggle very hard to keep it up.
- Current "solutions" to this problem are either inherently problematic (blackholing target addresses, thereby completing the DoS), or quite expensive, and often critically rely on limited pools of network engineering talent.
- Given how integral and irreplaceable the Internet has become, I think DoS attacks are a problem that merits more/better ongoing attention by the community.
- **Documenting and working on improved mitigation strategies for DoS attacks needs to become a material focus of ongoing cybersecurity research attention – the system or network you protect may be your own.**

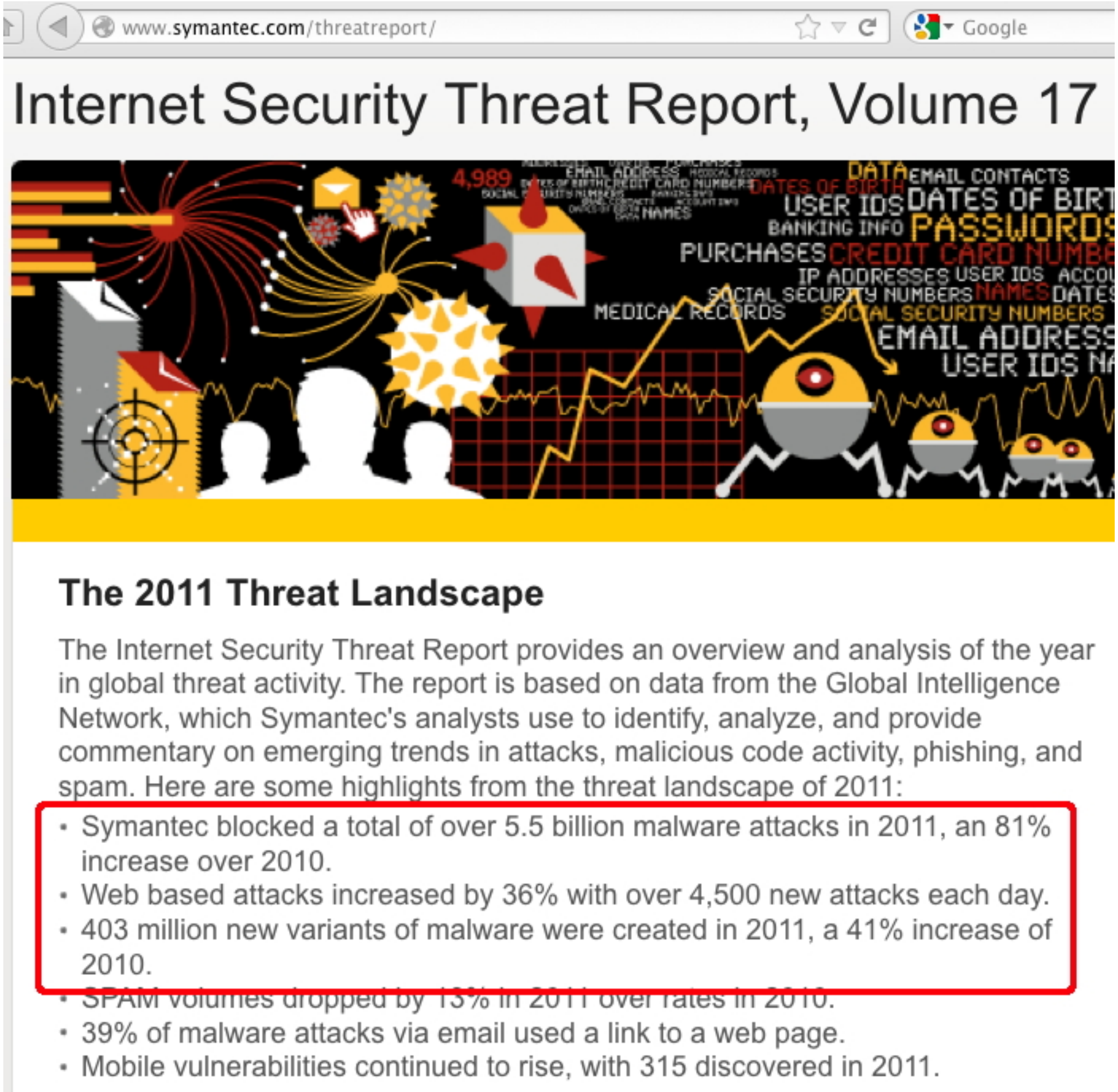
III. Malware and Botnets

March 2005: "Spam Zombies and Inbound Flows to Compromised Customer Systems," http://pages.uoregon.edu/joe/zombies.pdf

May 2007: "We Need a Cyber CDC or Cyber World Health Organization," http://pages.uoregon.edu/joe/ecrime-summit/ecrime-summit.pdf

October 2011: "Malware Analysis for Neophytes: A MAAWG Training Seminar," http://pages.uoregon.edu/joe/malware-analysis-paris/malware-analysis-paris.pdf

April 2012: "Botnets, the FCC CSRIC Working Group, and Opportunities for Internet2 Industry Partners and Researchers," http://pages.uoregon.edu/joe/i2mm-csric-wg7/i2mm-csric-wg7.pdf



Internet Security Threat Report, Volume 17

The 2011 Threat Landscape

The Internet Security Threat Report provides an overview and analysis of the year in global threat activity. The report is based on data from the Global Intelligence Network, which Symantec's analysts use to identify, analyze, and provide commentary on emerging trends in attacks, malicious code activity, phishing, and spam. Here are some highlights from the threat landscape of 2011:

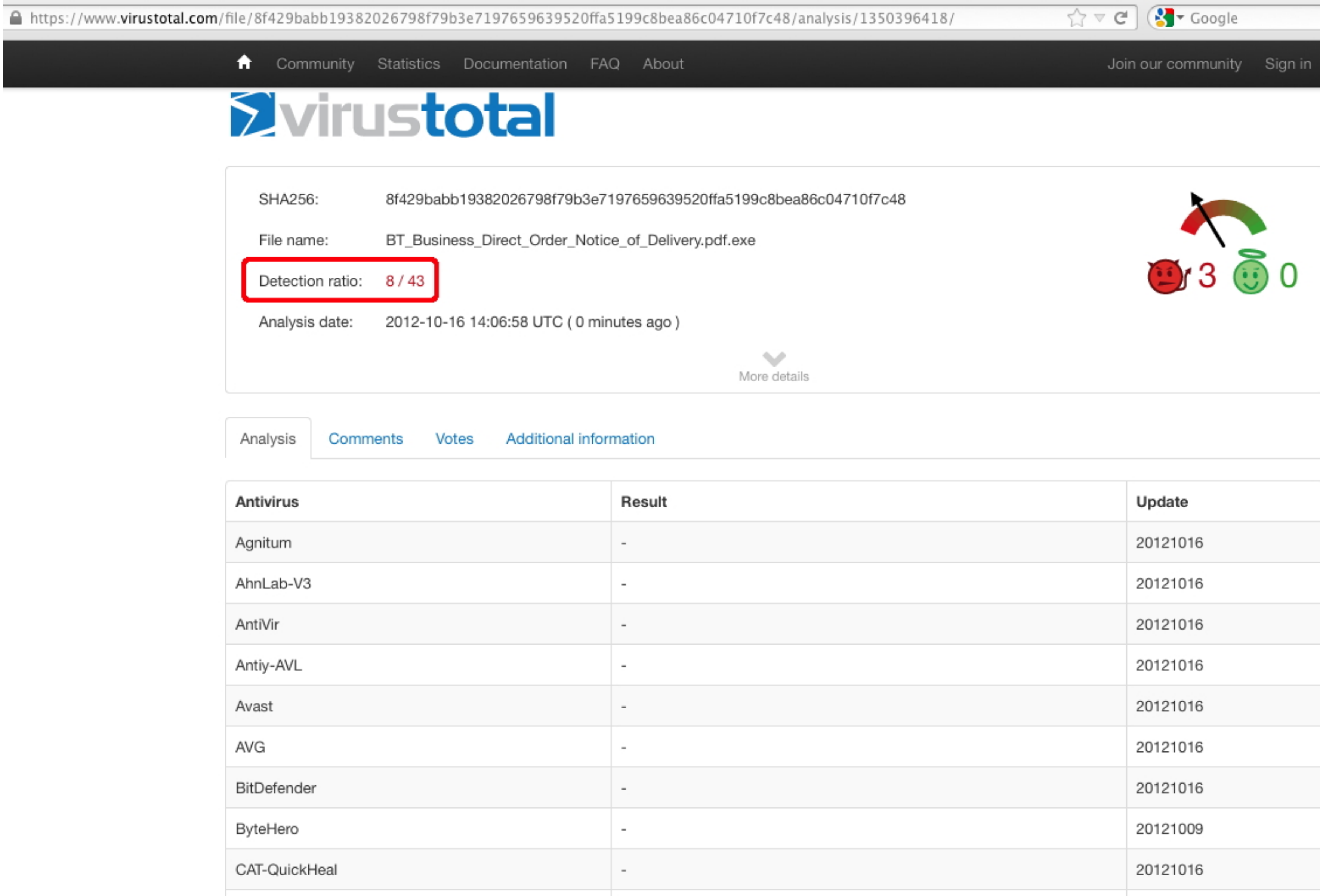
- Symantec blocked a total of over 5.5 billion malware attacks in 2011, an 81% increase over 2010.
- Web based attacks increased by 36% with over 4,500 new attacks each day.
- 403 million new variants of malware were created in 2011, a 41% increase of 2010.
- SPAM volumes dropped by 13% in 2011 over rates in 2010.
- 39% of malware attacks via email used a link to a web page.
- Mobile vulnerabilities continued to rise, with 315 discovered in 2011.

When You Mention "Cyber Security," Everyone Thinks "Malware"

- Computer viruses. Worms. Trojan horses. Rootkits. Spyware. There's a long list of terms that have been coined to refer to the various types of malicious software (malware) that's out there, but when you get right down to it, **malware is basically software you just don't want.**
- It can make your computer slow or unstable
- It can breach your privacy (and steal your money)
- It can mean that your computer is used to host stolen intellectual property (such as pirated software or ripped off movies or music), or child exploitation materials
- **It can also mean that your computer and network connectivity gets used to attack someone else, as part of a DDoS,** or to send spam.

"But *I* Use Antivirus Software! Surely *I'm* Safe, Aren't I?"

- I'm glad that virtually all of you are using antivirus software. Even though it isn't perfect, you definitely *should*.
- However, **you need to understand a brutal truth: signature-based antivirus software has largely failed to keep up with the pace at which malware is being tweaked and re-released**. The bad guys (and bad gals), are able to create modified versions of their malicious code -- versions that successfully slip past antivirus software -- faster than antivirus software companies can update A/V definitions.
- Thus, at any given point in time, there will be at least some malware in circulation that will slip past fully-updated signature-based A/V software and potentially infect your Windows PC. For example...



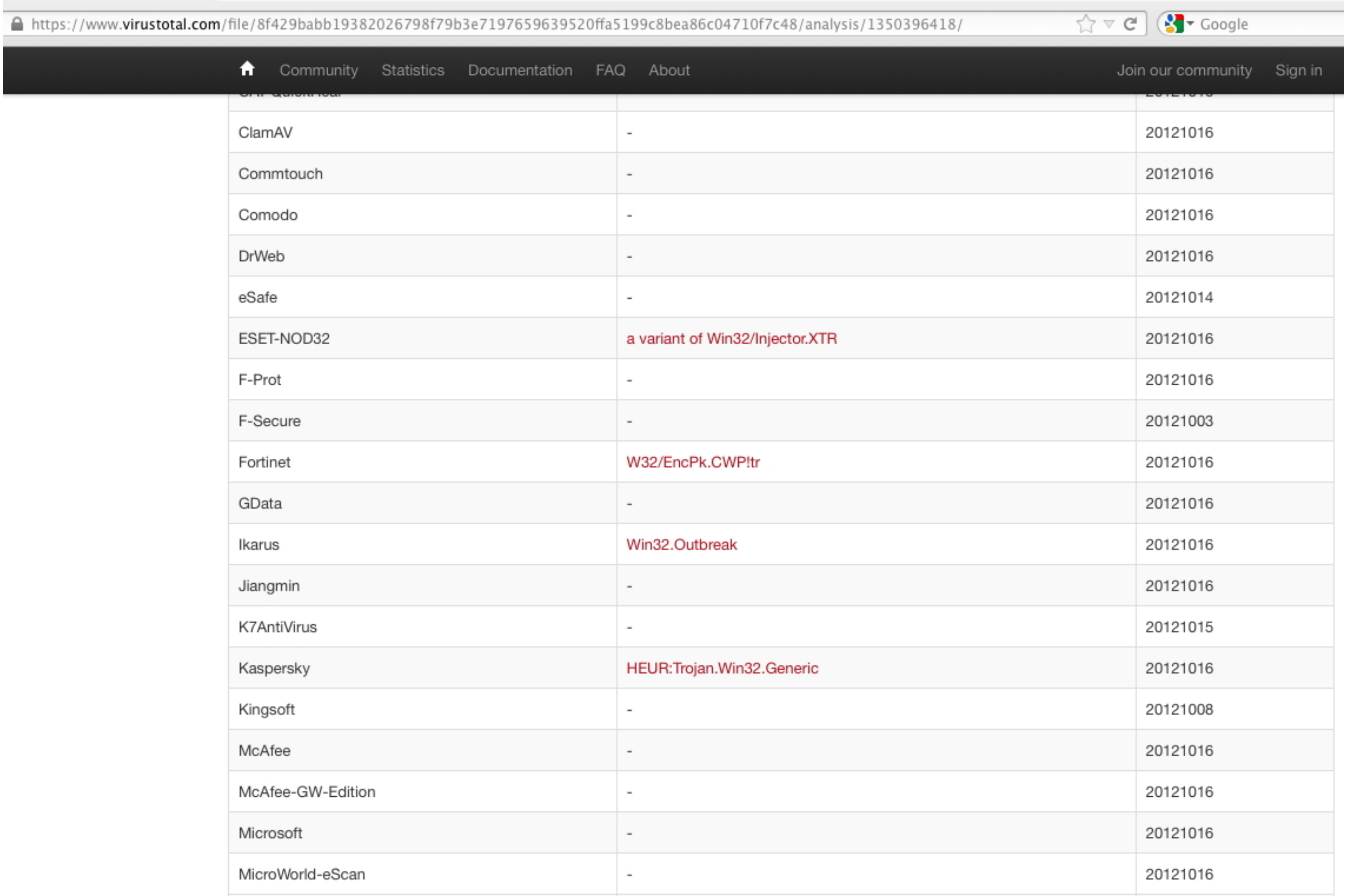
SHA256: 8f429babb19382026798f79b3e7197659639520ffa5199c8bea86c04710f7c48

File name: BT_Business_Direct_Order_Notice_of_Delivery.pdf.exe

Detection ratio: 8 / 43

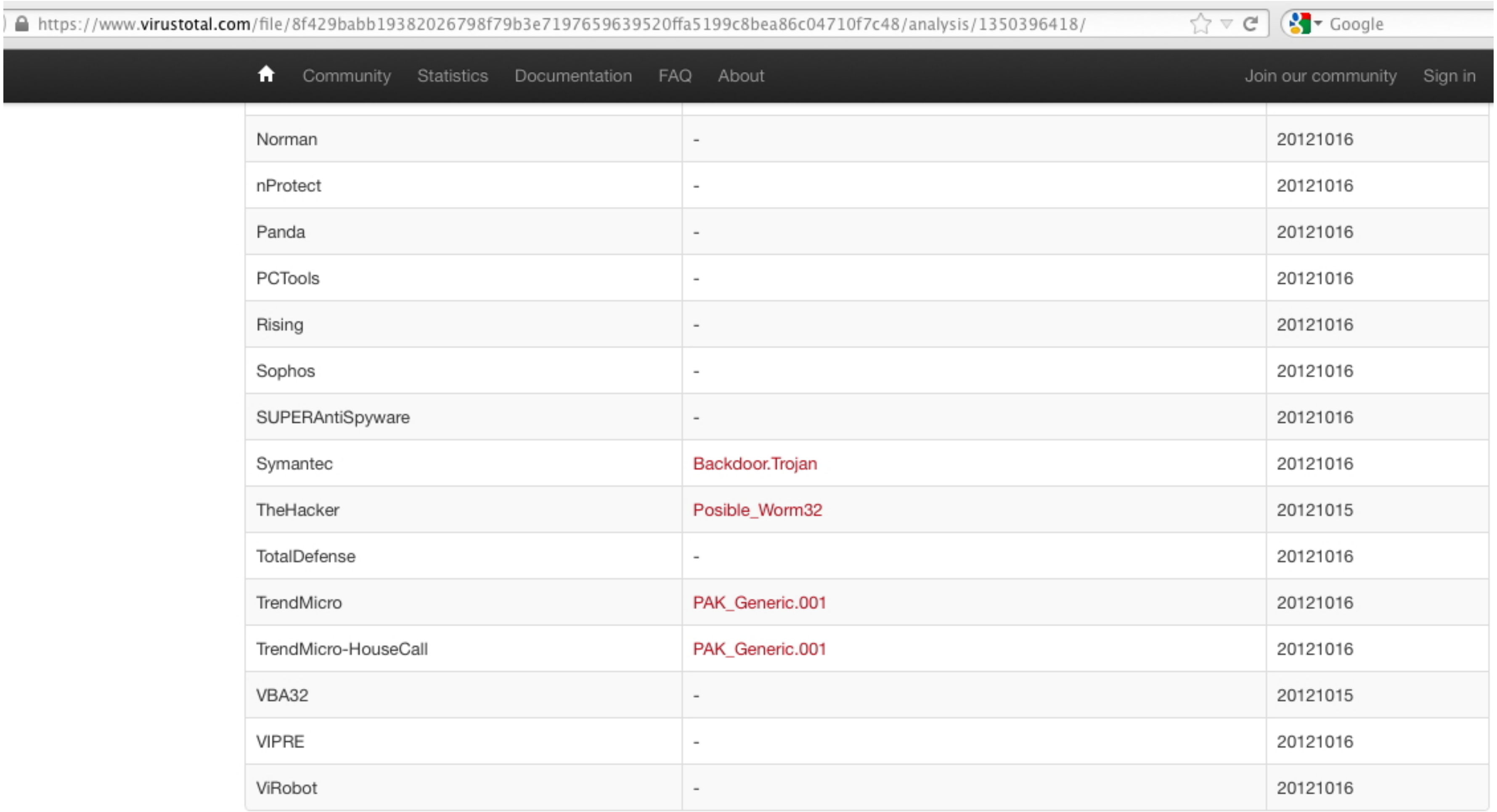
Analysis date: 2012-10-16 14:06:58 UTC (0 minutes ago)

Antivirus	Result	Update
Agnitum	-	20121016
AhnLab-V3	-	20121016
AntiVir	-	20121016
Antiy-AVL	-	20121016
Avast	-	20121016
AVG	-	20121016
BitDefender	-	20121016
ByteHero	-	20121009
CAT-QuickHeal	-	20121016



https://www.virustotal.com/file/8f429babb19382026798f79b3e7197659639520ffa5199c8bea86c04710f7c48/analysis/1350396418/

ClamAV	-	20121016
Commtouch	-	20121016
Comodo	-	20121016
DrWeb	-	20121016
eSafe	-	20121014
ESET-NOD32	a variant of Win32/Injector.XTR	20121016
F-Prot	-	20121016
F-Secure	-	20121003
Fortinet	W32/EncPk.CWP!tr	20121016
GData	-	20121016
Ikarus	Win32.Outbreak	20121016
Jiangmin	-	20121016
K7AntiVirus	-	20121015
Kaspersky	HEUR:Trojan.Win32.Generic	20121016
Kingsoft	-	20121008
McAfee	-	20121016
McAfee-GW-Edition	-	20121016
Microsoft	-	20121016
MicroWorld-eScan	-	20121016



https://www.virustotal.com/file/8f429babb19382026798f79b3e7197659639520ffa5199c8bea86c04710f7c48/analysis/1350396418/

Norman	-	20121016
nProtect	-	20121016
Panda	-	20121016
PCTools	-	20121016
Rising	-	20121016
Sophos	-	20121016
SUPERAntiSpyware	-	20121016
Symantec	Backdoor.Trojan	20121016
TheHacker	Posible_Worm32	20121015
TotalDefense	-	20121016
TrendMicro	PAK_Generic.001	20121016
TrendMicro-HouseCall	PAK_Generic.001	20121016
VBA32	-	20121015
VIPRE	-	20121016
ViRobot	-	20121016

You May Not Even *Know* You're Infected

- Like an asymptomatic carrier of a human communicable disease, you may not know that your system is infected.
- We've already mentioned that A/V software, even though it can help, still misses a lot of malware. Often, when that happens, **we can still find infected machines because of their behavioral outputs**: they source attack traffic or emit spam or they do other bad things over the network, and that misbehavior gets noticed and reported. But not all infected machines immediately "make noise." Some malware infected hosts may "lay doggo," at least for a while.
- More fundamentally, recognize that **you're asking a lot if you expect your potentially infected computer to "self-diagnose" and "heal" itself.** It is, after all, the electronic equivalent of a sick person, and shouldn't be expected to engage in effective introspection and self-remediation.

Many Of You *Have* Been Surreptitiously Infected, Only To *Also* Be Quietly Cleaned Up

- Many of you may laugh and feel somewhat self-satisfied and smug at the thought that while <u>others</u> may have been compromised by malware, <u>you</u> have not. Your antivirus software has never reported a problem, your system has never acted strangely, and your Internet service provider has never kicked you offline for being infected.
- Before you get too smug, you should understand that in many cases, you, too, <u>may</u> actually have been infected without your knowledge, it's just that you may <u>also</u> have been silently disinfected, perhaps by the Microsoft Malicious Software Removal Tool (MSRT), a malware removal tool that silently runs each month when your PC applies its monthly package of updates from Microsoft.
- There are a lot of "silent saves" happening out there…

All Malware Is Bad, But Many Centrally Coordinated Infected Systems Are Worse

- Anytime even one system gets infected, that's terrible.
- But what really worries me is when cyber criminals succeed in **compromising systems in bulk**, getting the ability to centrally command tens or hundreds of thousands of compromised zombie systems, often known as bots.
- When attack traffic gets routed via bots, an attack that is actually coming from a single attacker appears to be coming from thousands (or tens of thousands of sources), including many that are located overseas.
- "Working back upstream" to determine who's routing their traffic through those bots requires a well-instrumented network at a cooperative ISP, and of course, in some cases, bot traffic may be chained through multiple layers of bots (lather, rinse, and repeat).

Remediating Botted Hosts

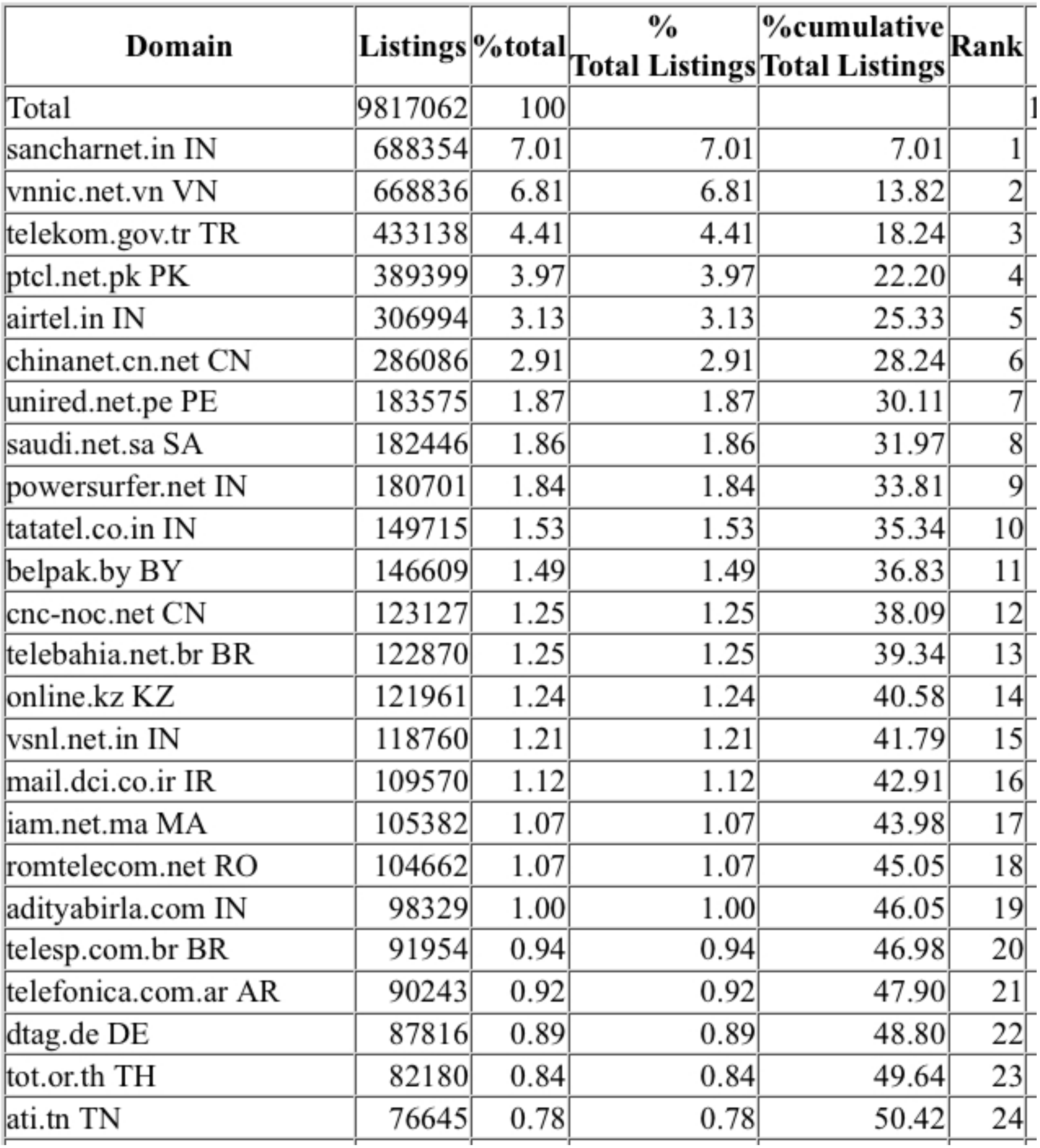
- Once you've identified a host that's been botted, wouldn't it be terrific if you could fix it, and make it whole again?
- But **who's going to clean up botted hosts**, and how?
- The **system owner** may not know how to clean things up, and they may not be motivated to pay someone to help
- **ISPs** can't afford to fix every customer's PC one-on-one
- How about the **operating system vendor**? Maybe they can help, but many compromises aren't due to operating system issues anymore -- should it be their responsibility to clean up all third party app issues, too?
- Who's going to clean up all the botted hosts **<u>overseas</u>** that are sending spam DoS'ing us, and otherwise acting bad?
- Maybe the government could help, if the problem's bad enough... but **how bad is the bot problem?**

Botnet Metrics

- That is, are one in three systems botted? One in ten? One in a hundred? One in a thousand? We just don't know.
- And where/when are we talking about? Infection rates in India will not be the same as infection rates in Indiana, and infection rates tomorrow may be dramatically different than infection rates today.
- **Our current botnet metrics are really, really disappointingly poor.** We know far more about tomorrow's weather or who may win an obscure congressional contest than the number of botted hosts here in Michigan today.
- Personally, I find this lack of "cyber epidemiology" rather surprising given the extent to which we all collectively view malware as a major cyber threat and our pretensions toward treating cybersecurity research as "science."

One Imperfect Bot Indicator: Spam Levels

- While we've been talking about bots as a way of performing DDoS attacks, bots *are* also routinely used for other purposes, such as sending spam or conducting click fraud.
- Thus, one way of indirectly getting some sense of the lower bound of the number of botted hosts is by looking at the number of IP addresses that have been seen to be doing something bad, like emitting spam.
- The best available metrics for botted hosts emitting spam are those that are made available by the CBL, see http://cbl.abuseat.org/statistics.html
- Two dozen ISP domains account for half of all known bots currently sending spam...



Domain	Listings	%total	% Total Listings	%cumulative Total Listings	Rank	
Total	9817062	100				1
sancharnet.in IN	688354	7.01	7.01	7.01	1	
vnnic.net.vn VN	668836	6.81	6.81	13.82	2	
telekom.gov.tr TR	433138	4.41	4.41	18.24	3	
ptcl.net.pk PK	389399	3.97	3.97	22.20	4	
airtel.in IN	306994	3.13	3.13	25.33	5	
chinanet.cn.net CN	286086	2.91	2.91	28.24	6	
unired.net.pe PE	183575	1.87	1.87	30.11	7	
saudi.net.sa SA	182446	1.86	1.86	31.97	8	
powersurfer.net IN	180701	1.84	1.84	33.81	9	
tatatel.co.in IN	149715	1.53	1.53	35.34	10	
belpak.by BY	146609	1.49	1.49	36.83	11	
cnc-noc.net CN	123127	1.25	1.25	38.09	12	
telebahia.net.br BR	122870	1.25	1.25	39.34	13	
online.kz KZ	121961	1.24	1.24	40.58	14	
vsnl.net.in IN	118760	1.21	1.21	41.79	15	
mail.dci.co.ir IR	109570	1.12	1.12	42.91	16	
iam.net.ma MA	105382	1.07	1.07	43.98	17	
romtelecom.net RO	104662	1.07	1.07	45.05	18	
adityabirla.com IN	98329	1.00	1.00	46.05	19	
telesp.com.br BR	91954	0.94	0.94	46.98	20	
telefonica.com.ar AR	90243	0.92	0.92	47.90	21	
dtag.de DE	87816	0.89	0.89	48.80	22	
tot.or.th TH	82180	0.84	0.84	49.64	23	
ati.tn TN	76645	0.78	0.78	50.42	24	

There Are Many Things That Make Even Those Measurements Less Than Perfect

- Blocking direct-to-port 25 SMTP traffic? If so, your botted customers can no longer directly send spam. Thus, they magically disappear from the CBL (even though they ARE still botted and could still participate in DDoS attacks)
- One IP might be in front of multiple NAT'd customers, resulting in the number of infected customers getting (potentially significantly) understated
- Or one infected customer might use multiple dynamic addresses, thereby resulting in a single infected system appearing to be multiple botted users (when it isn't)
- Or what if a system has *multiple* malware infections?
- And once an infected system gets cleaned up, how do we know that we can delist the associated IP address?

Malware Research Needs

- We need to have more security sites **contributing malware samples**
- We need more people working on malware analysis (including **automating analysis processes in scalable ways**)
- We need **better instrumented networks**, so we can look back up stream and see (and take down!) botnet C-and-C's that are steering all these botnets, and so we can better describe/characterize malware and botnets (just as we need metrics for denial of service attack traffic)
- And we need to understand **who will work on getting bots cleaned up, both here in the US and overseas**

IV. Spoofed Traffic

October 2012: "Internet2 Computer and Network Security Expectations (Version 1.2)" http://pages.uoregon.edu/joe/expect/page1.pdf at item 8.

TCP and UDP Traffic

- There are basically two types of network application traffic: TCP and UDP.

- TCP traffic is associated with relatively persistent connections (such as ssh sessions, web traffic, email, etc.), and has a variety of characteristics which are desirable from a network application programmer's point of view, including retransmission of lost packets, congestion control, etc.

- UDP traffic, on the other hand, is designed for "send-it-and-forget-it" apps where you don't want to/can't afford to maintain state or you don't want a lot of connection setup overhead. DNS, NFS, and IP video traffic all normally run as UDP.

The Spoofability of UDP Connections

- Unlike a fully established TCP connection (which only gets established after a bidirectional handshake is negotiated and which is therefore robust to spoofing attempts), UDP traffic can be created with virtually **any** apparent source address -- including IP addresses which have no relationship to the traffic's actual origin.

- Network traffic that's intentionally created with a bogus source address is usually said to be "spoofed."

- If allowed to reach the global Internet, spoofed traffic is generally indistinguishable from legitimate traffic.

Why Would Anyone Bother to Spoof Traffic?

- If you don't spend time "thinking like an attacker," you might not immediately "get" why an attacker would be interested in spoofing his attack traffic. The answer is actually quite simple: the attacker wants the systems he's using as part of his attack to stay online and unblocked as long as possible.

- Spoofing the source of the attack traffic...

  -- hinders backtracking/identification/cleanup of the system that's sourcing the traffic; and
  -- makes it harder for the attack victim to filter the attack traffic (the spoofed source addresses may be constantly changed by the attacker, and thus doesn't provide a stable "filterable characteristic").

"So Why Not Just Block All UDP Traffic?"

- Given that UDP can be easily spoofed by the bad guys/bad gals, sometimes you'll hear folks naively propose simply blocking all inbound or outbound UDP traffic (or at least heavily rate limiting all UDP traffic).

- Unfortunately, because some pretty basic services (like DNS) requires support for UDP, blocking (or heavily rate limiting) all inbound or outbound UDP traffic is generally **not** a good idea. :-;

- Warts and all, you have no choice but to learn to to live with UDP traffic. :-;

"Well, Can We Block SOME UDP Traffic?"

- For once, the answer is positive: yes, you can block some UDP traffic.

- For example, if you're the University of Oregon and your school has been assigned the IP address range 128.223.0.0-128.223.255.255 there's no reason for systems on your network to be sourcing packets that pretend to be from some other IP address range. We'd filter that spoofed traffic before it leaves our campus.

- This is a pretty basic sanity check, but you'd be surprised how many sites don't bother with even this trivial sort of filter.

Subnet-Level Filtering

- While it is great to prevent spoofing at the university-wide level, that sort of border router anti-spoofing filter does not prevent a miscreant from forging an IP address taken from one of your subnets for use on another of your subnets.

- *Cue subnet-level anti-spoofing filters….*
  You KNOW that hosts on each subnet should ONLY be originating packets with IP addresses legitimately assigned to that subnet, so at the uplink from each subnet, drop/block outbound packets that appear to be "from" any other IP address – another very basic sanity check.

BCP38/RFC2827

- Let me be clear that ingress filtering of traffic with spoofed IP addresses is not new and is not my idea – it is Best Current Practice (BCP) 38/RFC 2827, written by Ferguson and Senie in May 2000.

- Unfortunately, despite being roughly a dozen years old now, **many** sites still do **NOT** do BCP38 filtering -- perhaps as many as 14-23% Internet wide, depending on how you measure things. (See http://spoofer.csail.mit.edu/summary.php)

- **Does YOUR network do BCP38 filtering? If not, it should!**

"So Why Doesn't <u>Everyone</u> Do BCP38 Filtering?"

- "Too hard given the complexity of my network"

- Asymmetric costs/benefits: filtering my network protects you (which is nice), but filtering that traffic "costs" me w/o any tangible/economic "benefits." So what are these horrible "costs?"
  -- engineer time to configure and maintain the filters (one time/negligible for most .edu networks)
  -- overhead on the routers (but if that overhead is material enough to be a "show stopper," you should be upgrading anyway)

- "Too busy" (or other excuses)

"What's It To <u>You</u> Anyhow, Bub? Butt Out…"

- Some may question why others should care what they do with their networks – your network, your rules, right? Well, generally yes.

- However in this case, remember that if you're NOT doing BCP38 filtering, your network may be getting used to generate spoofed attack traffic that's pretending to be "from" someone else's network, and that's the point at which what you do (or don't do) potentially affects a lot of other people including the attack target itself, the entity whose IP addresses are being spoofed, etc.

"So How Should I Be Doing This Filtering?"

- Only you can make the final decision about the best approach for your network, but you may want to see BCP84/RFC3704, March 2004.
- I would note, however, that strict mode unicast reverse path forwarding ("strict uRPF") is **not** a good idea for the multihomed environment typical of I2 universities due to route asymmetry.
- I would also urge you to review (April 19, 2006) draft-savola-bcp84-urpf-experiences-00.txt
- Quoting RFC3704 "Ingress Access Lists require typically manual maintenance, but are the most bulletproof when done properly..."

An Example of A Spoofed Traffic Attack

- Most DNS traffic is UDP (and thus easily spoofable)
- Many recursive resolvers are "open" (aka (ab)usable by anyone), see "The Million Plus Open Resolver Challenge," http://www.team-cymru.org/Services/Resolvers/
- Small DNS queries may result in large answers: this makes DNS recursive resolvers into terrific "traffic amplifiers"
- If an attacker sends a large number of DNS queries to open recursive resolvers, and spoofs the source address to look as if those queries were all from some source he or she wants to attack, the open recursive resolvers will flood the attack target with "answers" to the faked queries (even though the target of the attack never actually made those queries)
- A ray of hope: http://www.redbarn.org/dns/ratelimits

Research Opportunities

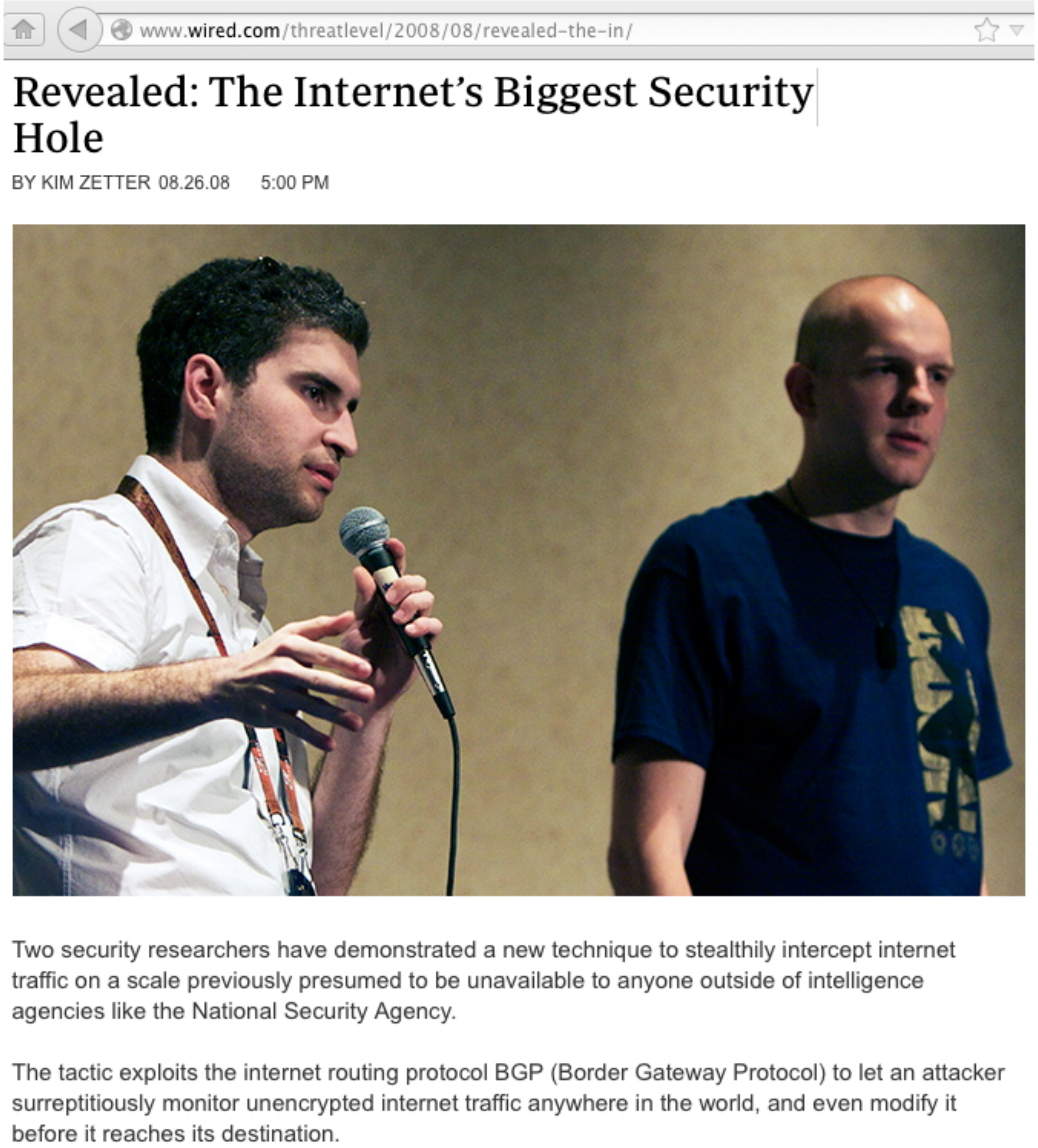
- We need people to continue documenting networks that are NOT filtering spoofed traffic.

- We need more research into any potential technical issues that may deter operators from filtering spoofed traffic.

- Given that we'll never convince everyone to filter spoofed traffic, we need solutions for containing that traffic. Heck, even continent-wide sanity checks would be better than nothing.

V. Route Injection and BGP Security

December 2006: "Route Injection and the Backtrackability of Cybermisbehavior," http://pages.uoregon.edu/joe/fall2006mm/fall2006mm.pdf



Revealed: The Internet's Biggest Security Hole

BY KIM ZETTER 08.26.08 5:00 PM

Two security researchers have demonstrated a new technique to stealthily intercept internet traffic on a scale previously presumed to be unavailable to anyone outside of intelligence agencies like the National Security Agency.

The tactic exploits the internet routing protocol BGP (Border Gateway Protocol) to let an attacker surreptitiously monitor unencrypted internet traffic anywhere in the world, and even modify it before it reaches its destination.

A Conceptual Intro to Wide Area Routing

- Customers have chunks of IP addresses called netblocks.
- ISPs on the Internet advertise routes for customer netblocks via a network protocol called "BGP."
- Traffic passing through a router will check that router's routing table for a "best match" to figure out where it should go next. That is, literally, which interface on the router should it go out to get (closer) to its destination?
- While various rules are used to determine what's a "best match," one of the most important rule says, "The **most specific route** that covers a given IP address 'wins,' if there's both a generic route, and more specific ones." This rule allows route tweaks to handle special case traffic.
- Unfortunately, this also creates a way for a malicious site to manipulate other sites' traffic -- they simply arrange to announce a more specific route for a prefix of interest.

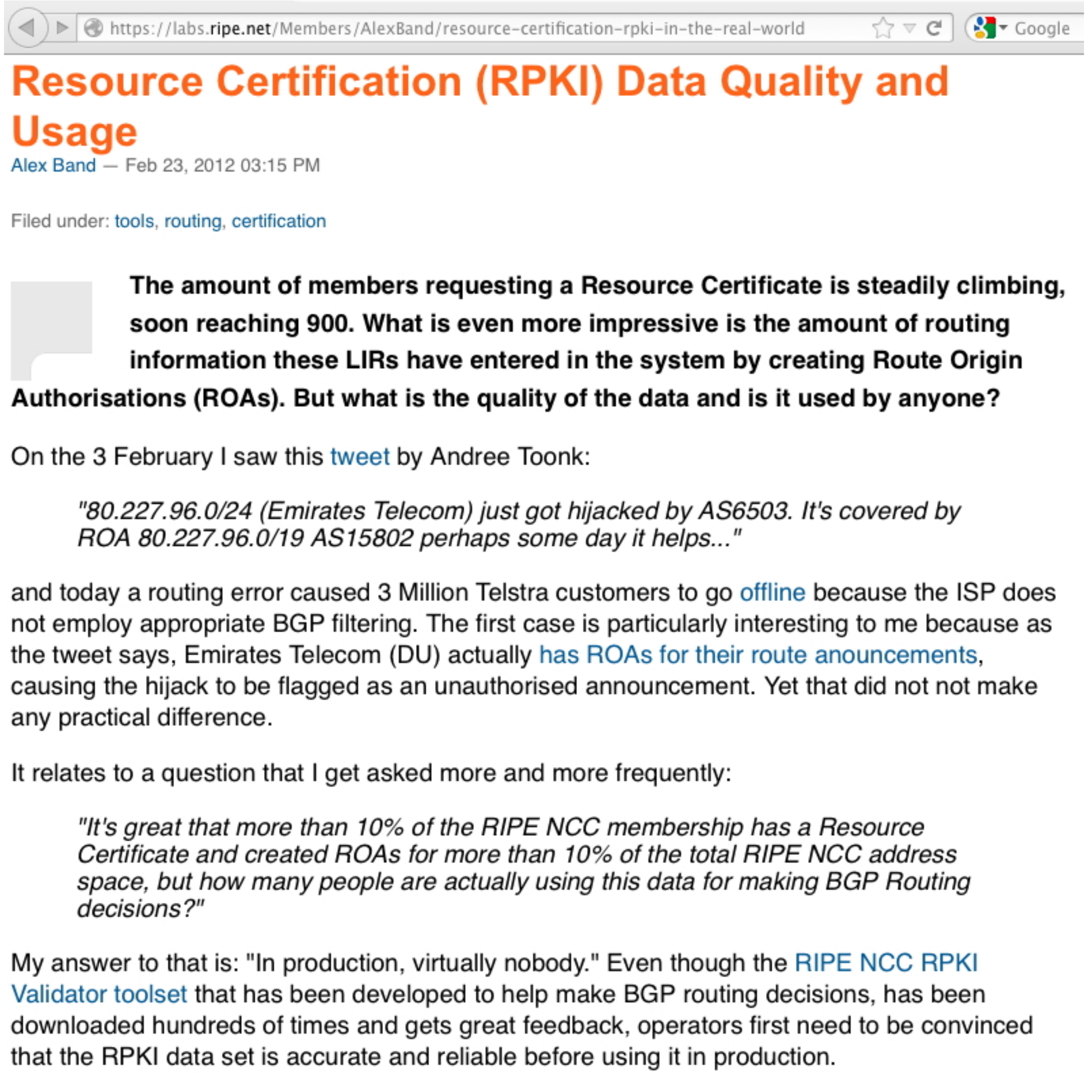
Intentionally Misrouted Traffic Can Result In Bad Things(tm) Happening

- I could "borrow" part of one of your netblocks that you're not currently using, and use it to send spam anonymously.
- If I inject a more specific route, I can convince others to send your traffic to me. I can then pretend that I'm delivering that traffic, while actually discarding it. This can make for a terrific DoS attack.
- And if I can get you to send your *unencrypted* traffic in my direction, I can **eavesdrop** on that traffic (or even surreptitiously modify it), and then, when I'm done playing with it, I can send it back out onto the network, back on its way to its actual destination. Unless you're paying **very** close attention, you'll never even know that that traffic has been taken on an unauthorized "detour" over the network. (This is sometimes called a "BGP shunt.")

Deterring Unauthorized Route Announcements

- Normally, customers need to provide an ISP with a **letter of authorization** (LOA) authorizing them to route a given prefix. Unfortunately, with Photoshop, anyone can create a fake LOA, and some ISPs don't even bother asking for one.
- In other cases, providers might require customers to register any prefixes they wanted to announce in a **routing registry ("RR")**, but many times what's in the RR and what's actually showing up on the network quickly get out of sync. **Poor RR accuracy over time has reduced the ability of ISPs to operationally filter based on RR data.**
- More recently, ARIN, RIPE, & the other regional registries have begun allowing users to get cryptographically signed "ROAs" or "route origin authorizations" with **RPKI.** For the first time, a technical mechanism exists for demonstrating authorization to originate a given netblock. But...

The Problem? So Far, Few Are Using RPKI


https://labs.ripe.net/Members/AlexBand/resource-certification-rpki-in-the-real-world
Google

Resource Certification (RPKI) Data Quality and Usage

Alex Band — Feb 23, 2012 03:15 PM

Filed under: tools, routing, certification

The amount of members requesting a Resource Certificate is steadily climbing, soon reaching 900. What is even more impressive is the amount of routing information these LIRs have entered in the system by creating Route Origin Authorisations (ROAs). But what is the quality of the data and is it used by anyone?

On the 3 February I saw this tweet by Andree Toonk:

> *"80.227.96.0/24 (Emirates Telecom) just got hijacked by AS6503. It's covered by ROA 80.227.96.0/19 AS15802 perhaps some day it helps..."*

and today a routing error caused 3 Million Telstra customers to go offline because the ISP does not employ appropriate BGP filtering. The first case is particularly interesting to me because as the tweet says, Emirates Telecom (DU) actually has ROAs for their route anouncements, causing the hijack to be flagged as an unauthorised announcement. Yet that did not not make any practical difference.

It relates to a question that I get asked more and more frequently:

> *"It's great that more than 10% of the RIPE NCC membership has a Resource Certificate and created ROAs for more than 10% of the total RIPE NCC address space, but how many people are actually using this data for making BGP Routing decisions?"*

My answer to that is: "In production, virtually nobody." Even though the RIPE NCC RPKI Validator toolset that has been developed to help make BGP routing decisions, has been downloaded hundreds of times and gets great feedback, operators first need to be convinced that the RPKI data set is accurate and reliable before using it in production.

What To Do In the Meantime?

- We need more eyes paying attention to global routing data, working to identify hijacked netblocks more-or-less in "real time."
- Dave Meyer's UofO Routeviews project, see http://www.routeviews.org/ , collects routing data from contributors all around the world, making that data freely available for researchers who might like to work with it.
- There's a lot of strange things happening in the BGP routing world out there, and to quote Jacqueline Kelly's *The Evolution of Calpurnia Tate,*

  > "It's amazing what you can see when you just sit quietly and look."

- **We need more researchers looking at BGP routing data!**

VI. Authentication

November 2009: "Passwords," http://pages.uoregon.edu/joe/passwords/passwords.pdf

February 2012: "Client Cert Deployment Models and Hardware Tokens/Smart Cards," http://pages.uoregon.edu/joe/client-cert-models/jt-louisiana.pdf

May 2012: "Client Certificates: A Security Professionals 2012 Pre-Conference Seminar," http://pages.uoregon.edu/joe/secprof2012/sec-prof-2012-client-certs.pdf

Passwords Are Ubiquitous

- It is hard to think of an online service that doesn't use traditional passwords of one sort or another…

  -- workstation, server, network device and mobile device logins including wireless auth, VPNs, etc.
  -- networked applications such as your email (and instant messaging, and calendaring, and…)
  -- many web sites (such as Amazon, eBay, Facebook, etc.)
  -- campus course management systems (e.g., Blackboard)
  -- campus administrative systems (with FERPA data?)
  -- online financial accounts (with GLB data?)
  -- medical and insurance-related sites (with HIPAA data?)
  -- etc., etc., etc.

- We truly use passwords everywhere.

Passwords Failures Can Have Major Impacts

- If one or more of your passwords gets compromised:
  -- confidential materials may be accessed or disclosed (resulting in you being sued/fired/arrested)
  -- critical files may be surreptitiously modified or deleted, (including potentially irreplaceable data)
  -- you may be denied access to your own resources (e.g., if the bad guys decide to "lock you out")
  -- your personal or institutional reputation may be damaged (for example if spam is sent from your account, your college may end up being blocklisted)
  -- miscreants may take your money or even co-opt your identity
- I think passwords play a **critical** security role, so if we're going to rely on them, then they'd BETTER be trustworthy. Unfortunately, as you know, they're not.

A Few Password Problems

- People will pick weak passwords.
- People will reuse the same ones across multiple sites.
- Everyone has way too many of the dang things, so they forget them (unless they write them down)
- Forgotten passwords often get reset via insecure mechanisms.
- People will willingly disclose the ones they've picked (e.g., phishing)
- Passwords will end up getting sniffed over the wire, or by malware running on a system.
- Passwords, once picked, are unlikely to get changed
- Passwords are a huge PITA to administer
- Passwords are no longer good enough (by definition!) for some high risk activities

(Maybe) We Can Fix Some of These Authentication-Related Issues

- For example, we could reduce or eliminate the need for individual usernames and passwords on many sites if only more sites used federated authentication (as offered through InCommon) [obDislcaimer: I work with InCommon]

- A list of Identity Providers and Service Providers can be found at https://incommon.org/federation/info/all-orgs.html Bravo to all listed there, bravo! You're my heroes!

- However, all too many other sites still don't support federated auth. Do we understand why sites are reluctant to become Identity Providers or Relying Parties? If not, maybe that's an area we should be researching...

Or How About Getting Beyond Using Plain Old Passwords?

- Many of you may be familiar with multifactor authentication, where something you know, such as a password, gets used along with something you have (such as a hardware cryptographic) token, or something you are (e.g., biometric methods, such as fingerprint scanners).
- **Multifactor authentication can make many authentication-related issues (such as phishing) virtually disappear**, and multifactor auth really is a BCP these days for accounts with access to sensitive data.
- We have cost effective methods for doing multifactor auth, including both open source solutions and commercially supported solutions (such as campus site licenses for Duo Security, see http://www.incommon.org/duo/)…
  and yet **people keep on using plain old passwords.**

Call Me A Dreamer: How About Getting All The Way to LOA-4?

- In case just deploying some sort of multifactor authentication isn't a big enough challenge, we could even talk about deploying client certificates using PKI hard tokens or smart cards in an effort to eventually get all the way to the highest level of assurance, NIST 800-63 LOA-4.
- We have a proof by example that this is possible – the Federal government has deployed millions of CAC/PIV cards for their employees and contractors.
- So why aren't we seeing similarly strong credentials used in higher education? Are there no use cases? Is it just more of a pain than it's worth? Is it too expensive? Are there usability issues? Privacy concerns? Is it the required identity proofing work?

VII. Conclusion

The Preceding Is Not a Comprehensive List

- We've mentioned many issues, we really haven't even scratched the surface. For example, we haven't covered:
  -- Securing DNS and DNSSEC
  -- Using IPv6 Securely
  -- Work That Needs to Be Done on OpenFlow/SDN
  -- Security in the Cloud
  -- Mobile Security
  -- Security and Privacy
  -- Physical IT Security (plus Disaster Recovery)
  -- Control System Security (so-called "SCADA" Security)
  -- and the list goes on...

- **We really NEED more people working on cyber security issues!** This is a tremendous potential area for academic researchers, and there's a LOT of funding available.

A Closing Quote About Nature That Could Have Been Written About the Internet

"If nothing else, school teaches that there is an answer to every question; only in the real world do young people discover that many aspects of life are uncertain, mysterious, and even unknowable. [...] The more you watch, the more mysterious the natural world becomes, and the more you realize how little you know. Along with its beauty, you may also come to experience its fecundity, its wastefulness, aggressiveness, ruthlessness, parasitism, and its violence. These qualities are not well-conveyed in textbooks.

"Perhaps the single most important less to be learned by direct experience is that the natural world, with all its elements and interconnections, represents a complex system and therefore we cannot understand it and we cannot predict its behavior."

Michael Crichton, *Micro*, Aug 28th, 2008 (Crichton died on Nov 4th)