

How The Pursuit of Truth Led Me To Selling Viagra®

Vern Paxson

*EECS Department, University of California
International Computer Science Institute
Lawrence Berkeley National Laboratory
Berkeley, California USA*

August 13, 2009



Outline:

- This is a broad, *retrospective* talk about network security **Data**
- Specifically, 2 decades' worth of Internet measurement:
 - What the data tells us about the lay of the land
 - ... what's changed
 - ... and what in fact doesn't change ("invariants")
- A personal (ivory tower research) view:
 - From general network characterization \Rightarrow manual attacks \Rightarrow worms \Rightarrow bots \Rightarrow spam
 - Why all this leads to selling Viagra



First, some acknowledgments:

- **ICSI**: Mark Allman, Christian Kreibich, Robin Sommer, Nicholas Weaver
- **LBL**: Craig Leres, Jim Rothfuss, Dwayne Ramsey, Brian Tierney, et al
- **UC San Diego**: Stefan Savage, Chris Kanich, Kirill Levchenko, Brandon Enright, Geoff Voelker



Part I

Pursuit of Truth +

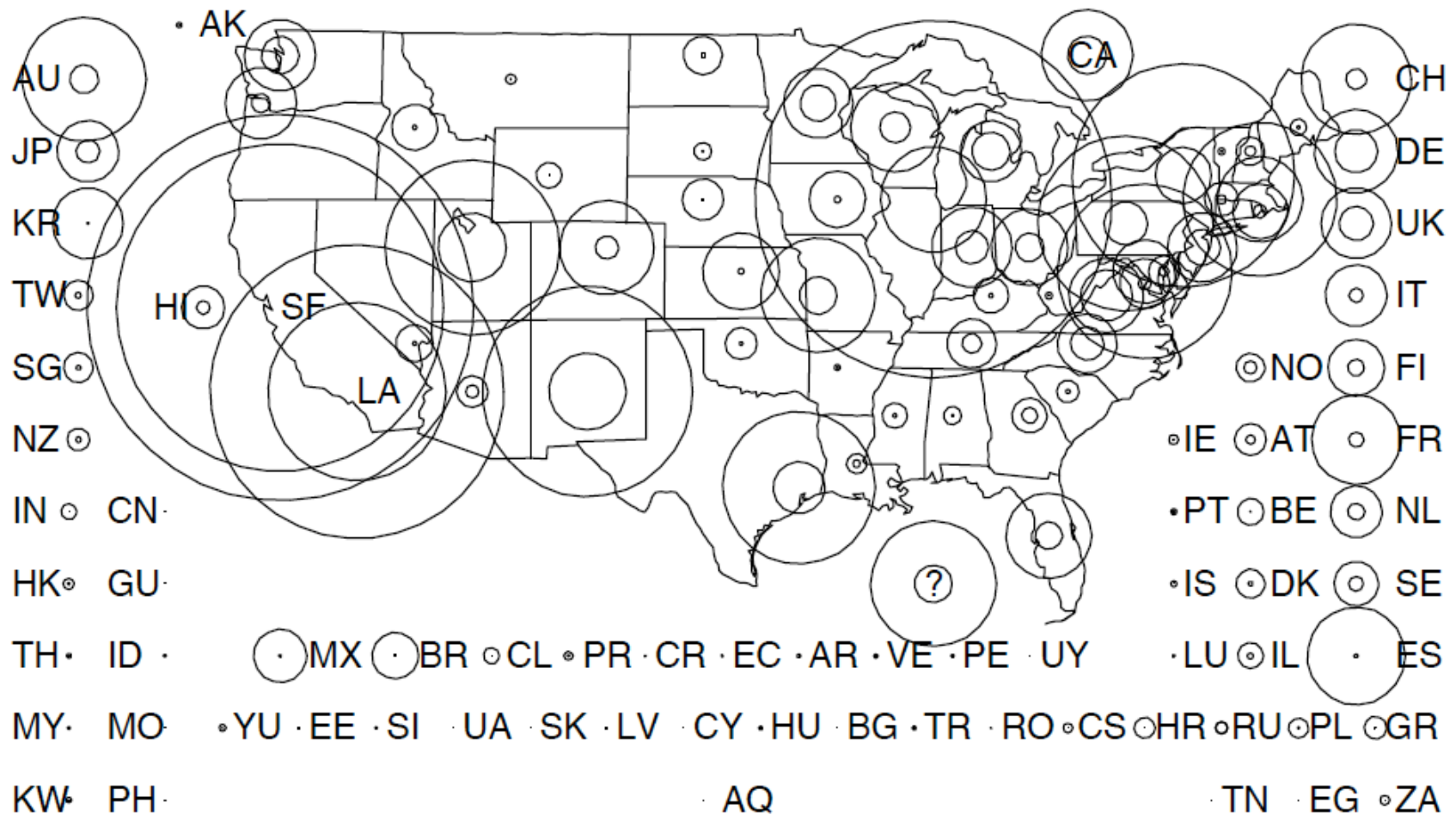
Phobia of Being Fooled =

Thirst for **Data**



Three Invariants: Growth, Explosive Onset, & Diversity

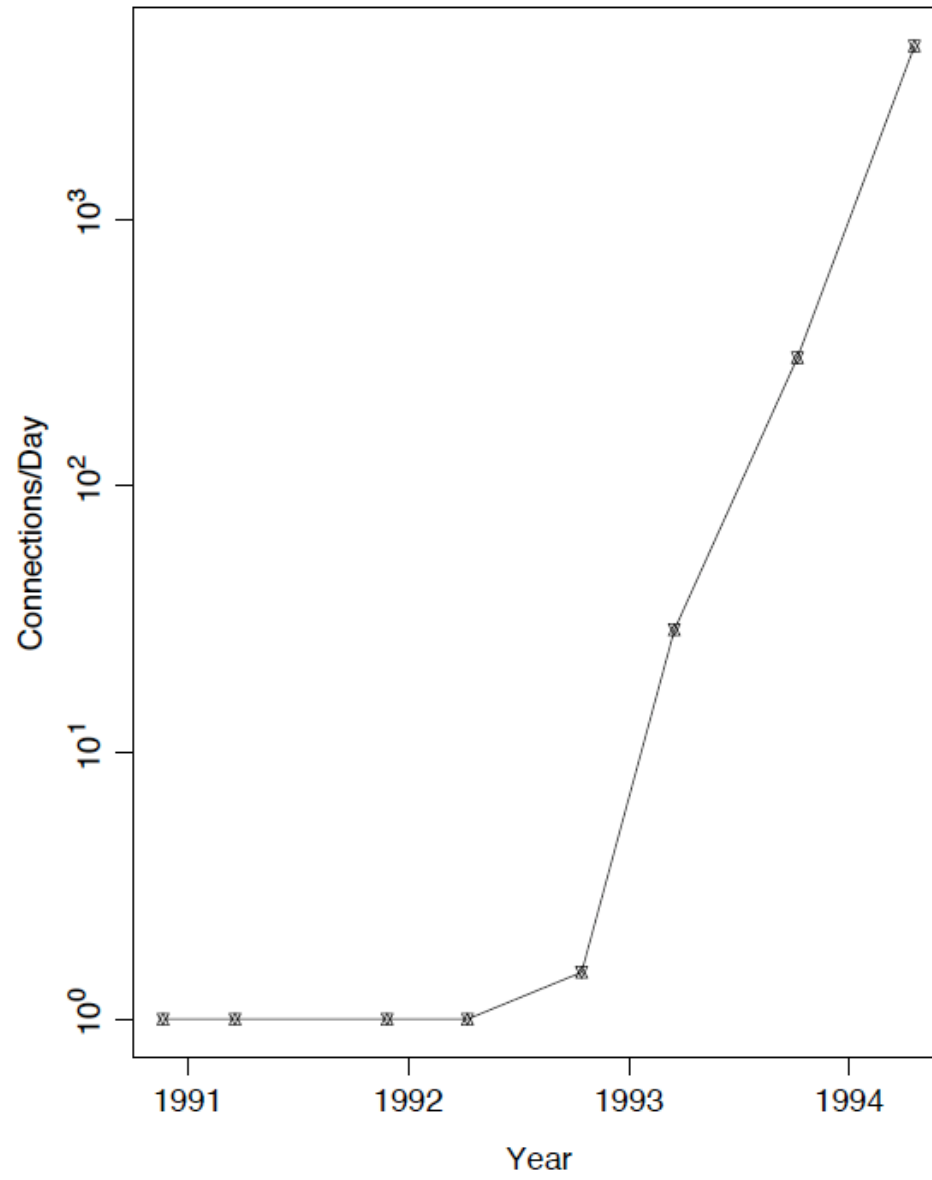
- Sep 1988: I apply to grad school
 - 56,000 Internet hosts (3.3 MB/day)
- Sep 1990: I enroll in grad “special topics” course on networking & start measuring traffic at LBL
 - 313,000 Internet hosts (9.5 MB/day)
- Oct 21 1991: I join Prof. Ferrari’s *Tenet* group
 - 617,000 Internet hosts (17.5 MB/day)
- May 11, 1994: My paper *Growth Trends in Wide Area TCP Connections* accepted for publication
 - \approx 3,000,000 Internet hosts (130 MB/day)



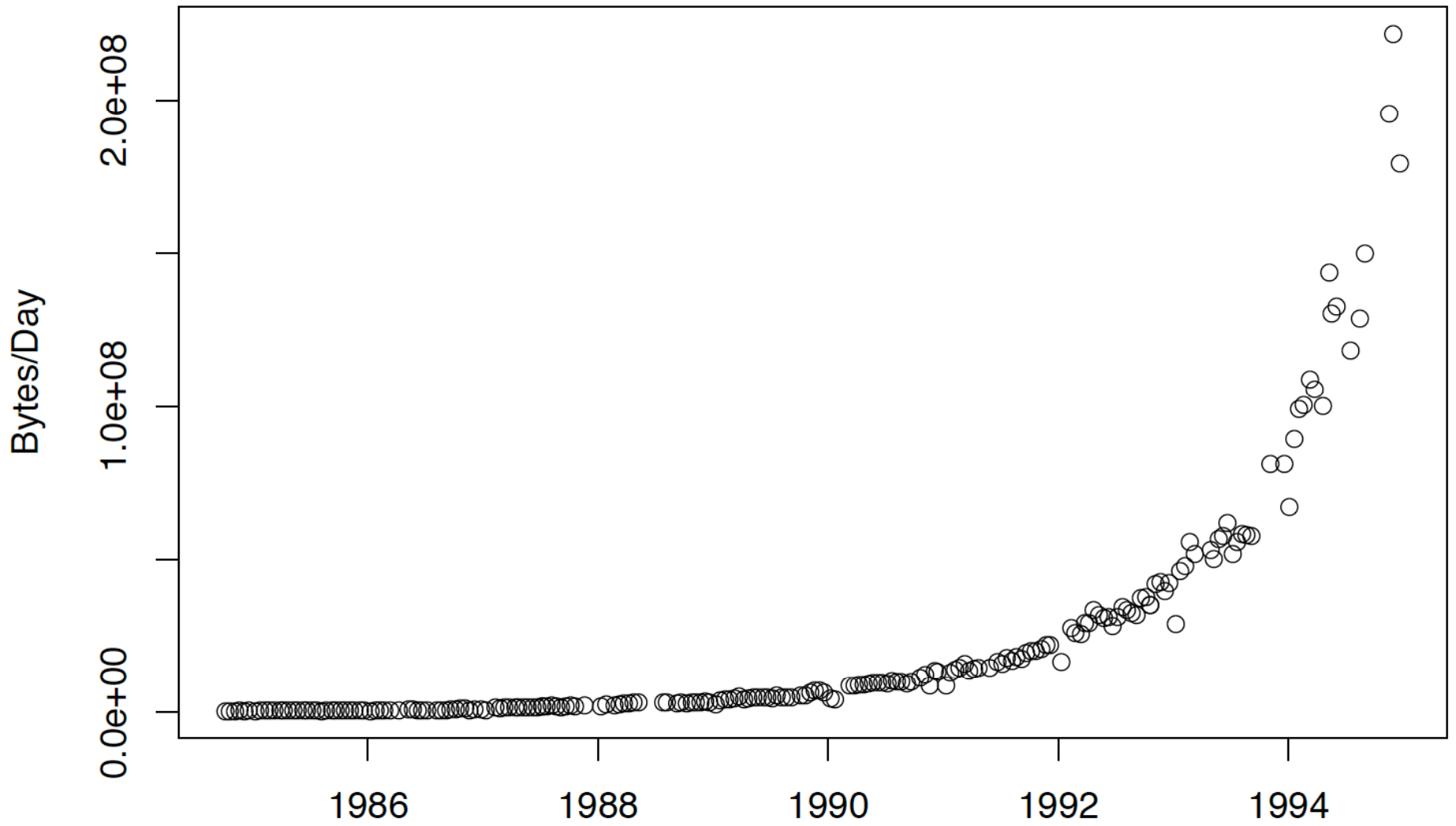
“Our data suggests a very recent explosion in commercial use of the Internet ...”

“... relatively new information-retrieval protocols such as Gopher and World-Wide Web exhibited explosive growth”

Growth of LBL's WWW Traffic



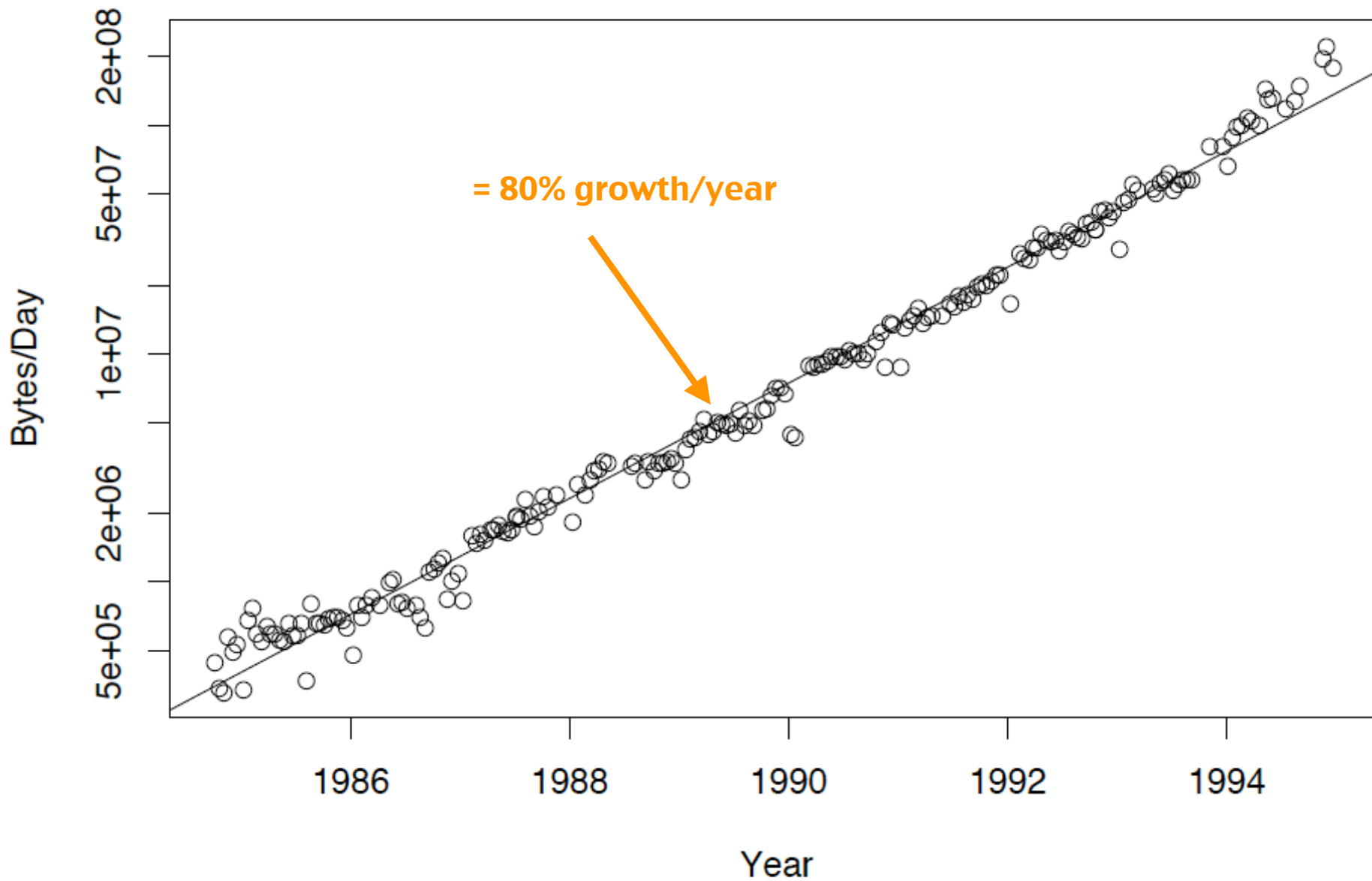
USENET Bulletin Board Traffic Volume



Year

Data courtesy of Rick Adams

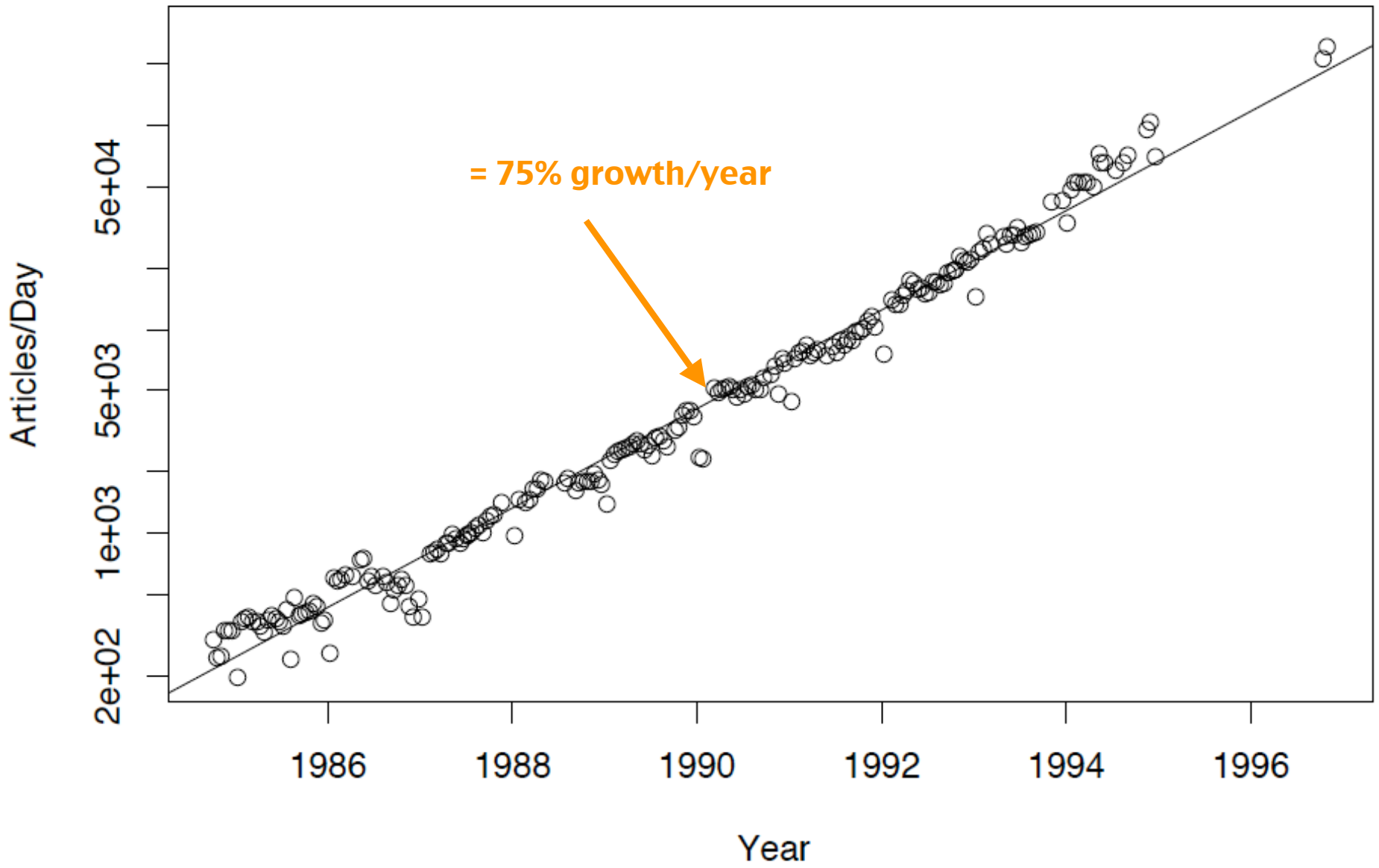
USENET Bulletin Board Traffic Volume



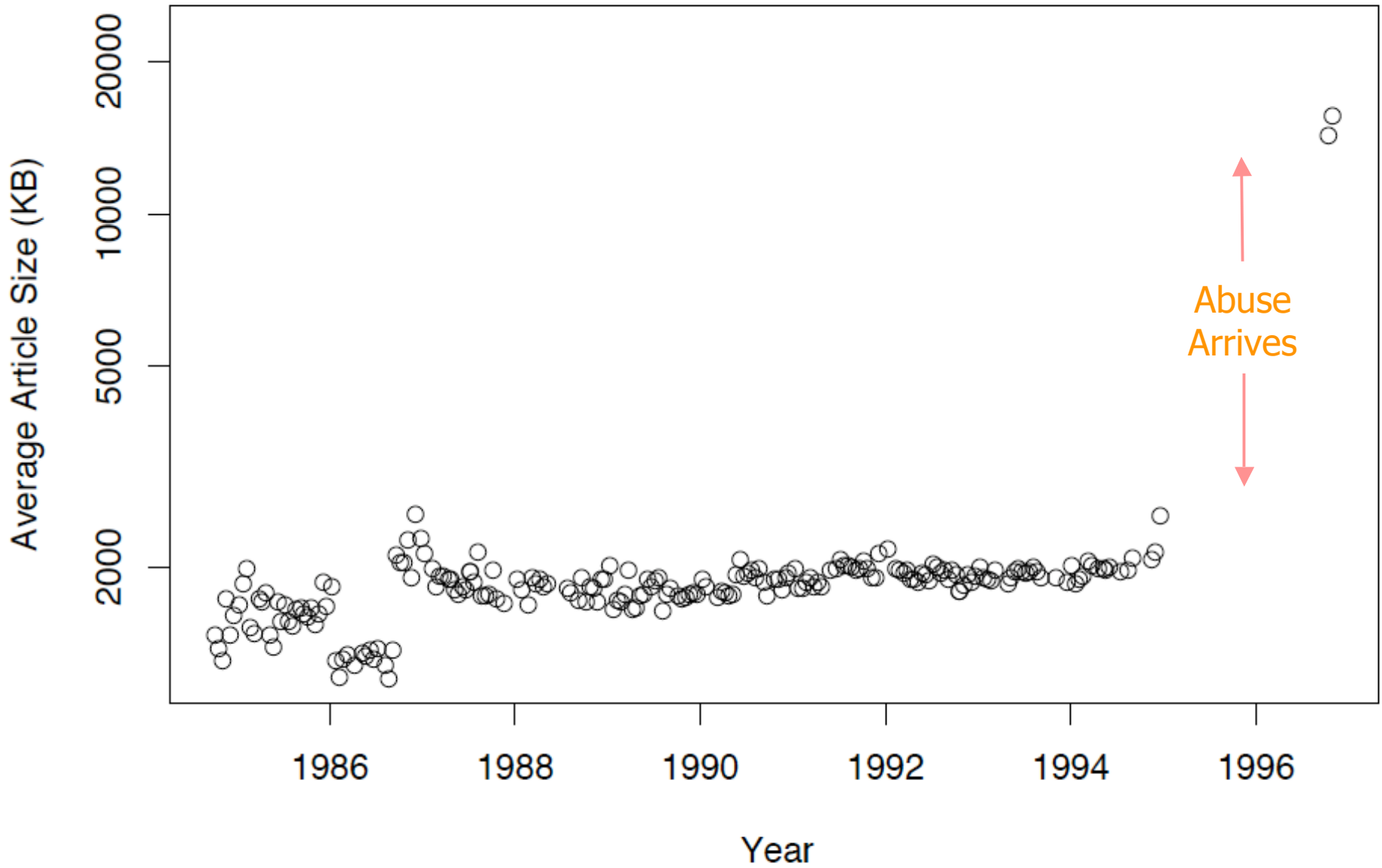
USENET Bulletin Board Traffic Volume



USENET Bulletin Board Traffic Volume



USENET Bulletin Board Traffic Volume



Mid-1990s: Internet Abuse Starts Becoming a Concern

- Observation: operators increasingly ask whether my network data sheds light on security incidents
 - Hmm, what about doing such measurement purposefully for security monitoring?
- Armed with equipment donation from DEC, the **Bro intrusion detection system** starts operating 24x7 in 1996
 - Inspects LBL border traffic in real-time
 - Who-talks-to-whom, what service, how much data
 - And, increasingly: what are the **semantics** of the conversations

Detecting Attackers, 1990s-style

- Inspect access to sensitive objects:
 - Hosts, usernames (“lp”, “r00t”), filenames (“/etc/passwd”), services (“mountd”, Windows file sharing)
- Look for specific forms of protocol abuse
 - E.g., FTP “site exec”, excessively long “finger” requests
- Check for telling behavior
 - Local host starts running an IRC chat server
 - Outbound requests to `www.uberhax0r.net`, `anticode.com`
 - Login sessions containing: `unset HISTFILE`; “eggdrop”;
`printf(“overflowing”;` “smurf.c by TFreak”; “u_char
`sparc_shellcode[] =`; “Coded by James Seter”
- Attackers exploit systems via interactive login sessions
 - Motivated by bragging rights / vandalism
 - Frequent community reuse of tools
 - Employment of “bots” for automating IRC management
- But what about “serious” attackers rather than weenies?

Real-World Security: *Threat Model*

- 1990s academic computer security research heavily influenced by cryptography's standard of mathematical assessment of security strength
 - Prove security properties ...
 - ... given a model of a powerful adversary
- In practice, goal is risk management, not bulletproof protection.
 - Much of the effort concerns “raising the bar” and *trading off resources*
- **Threat model**: what you are defending against
 - This can differ from what an academic might expect
 - Consider the Department of Energy ...

Network Security Research Grounded in Operational Use

- Ties with LBL operational deployment have been **research gold**
 - *Transformative* compared to working in small, self-contained environment like a lab
- Along with *threat model* (policy) realities, **scale** completely alters the problem landscape:
 - Performance - current target: analyze >> 100K pps
 - Research on: clustering; FPGA front end; multicore architecture
 - Diversity - you see the darnedest (benign) behavior & “crud”
 - Greatly complicates **anomaly detection** & detecting **evasion**



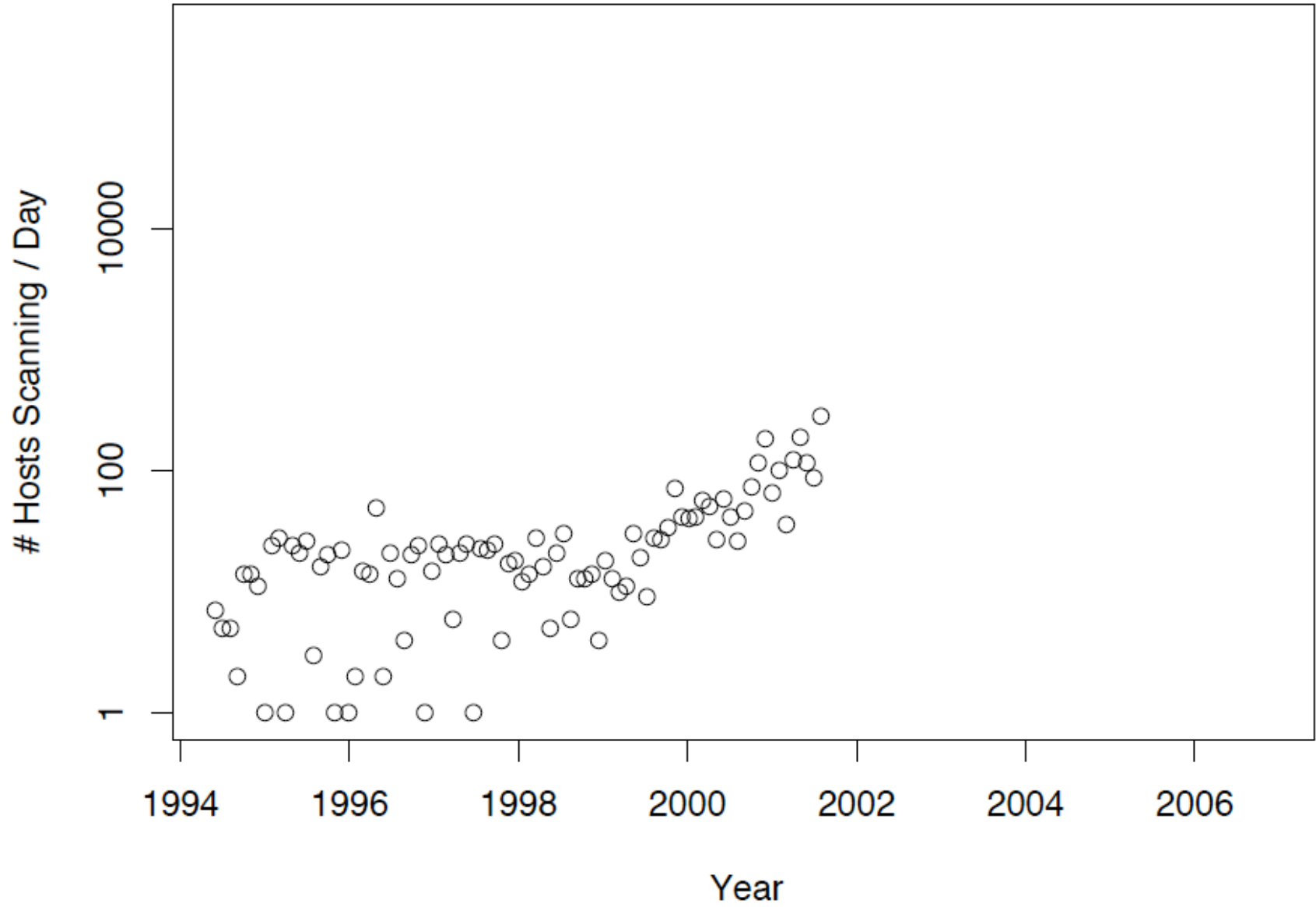
1 day of “crud” seen at ICSI (155K times)

active-connection-reuse	DNS-label-len-gt-pkt	HTTP-chunked-multipart	possible-split-routing
bad-Ident-reply	DNS-label-too-long	HTTP-version-mismatch	SYN-after-close
bad-RPC	DNS-RR-length-mismatch	illegal-%-at-end-of-URI	SYN-after-reset
bad-SYN-ack	DNS-RR-unknown-type	inappropriate-FIN	SYN-inside-connection
bad-TCP-header-len	DNS-truncated-answer	IRC-invalid-line	SYN-seq-jump
base64-illegal-encoding	DNS-len-lt-hdr-len	line-terminated-with-single-CR	truncated-NTP
connection-originator-SYN-ack	DNS-truncated-RR-rdlength	malformed-SSH-identification	unescaped-%-in-URI
data-after-reset	double-%-in-URI	no-login-prompt	unescaped-special-URI-char
data-before-established	excess-RPC	NUL-in-line	unmatched-HTTP-reply
too-many-DNS-queries	FIN-advanced-last-seq	POP3-server-sending-client-commands	window-recision

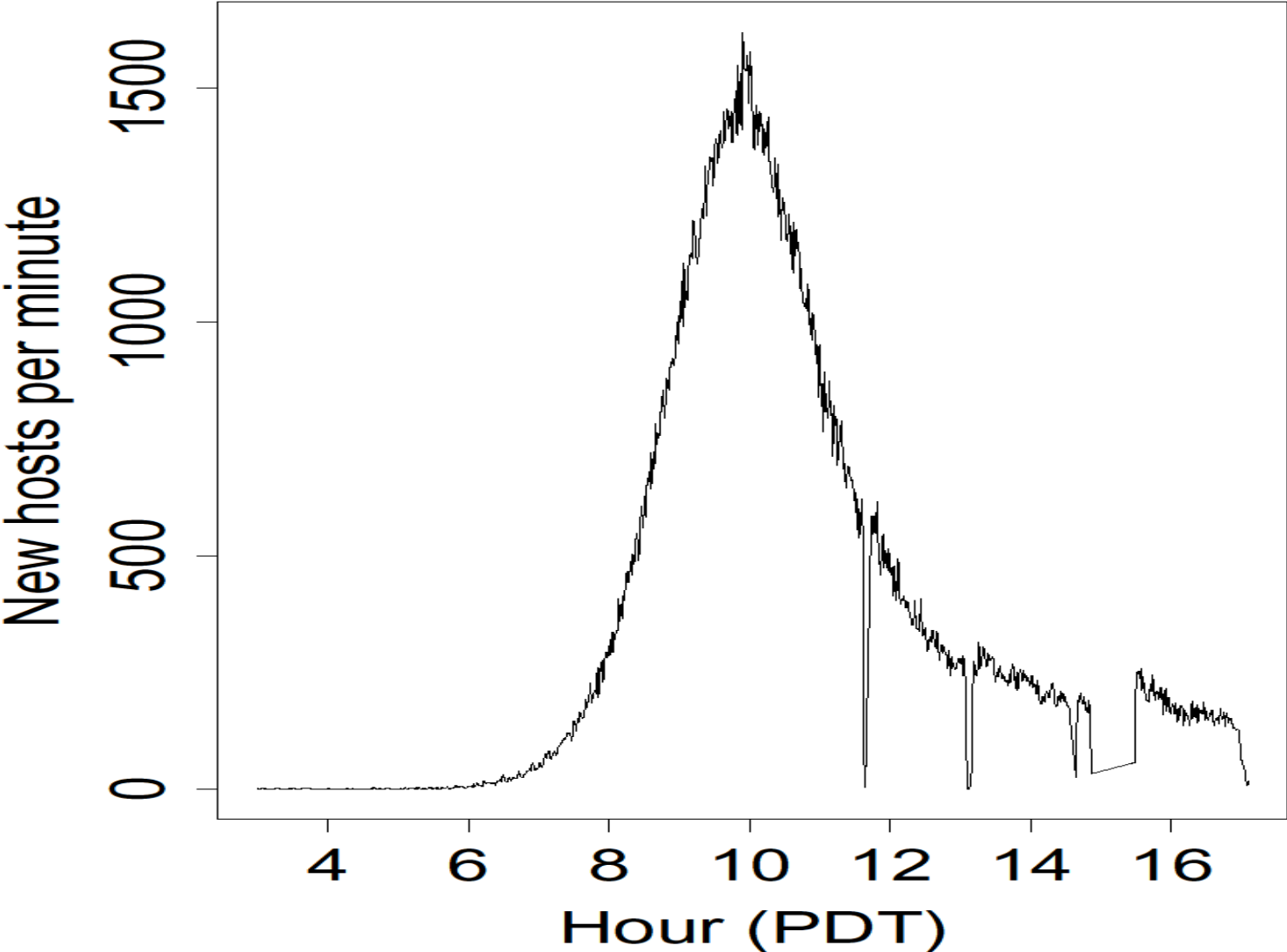
Network Security Research Grounded in Operational Use

- Ties with LBL operational deployment have been **research gold**
 - *Transformative* compared to working in small, self-contained environment like a lab
- Along with *threat model* (policy) realities, **scale** completely alters the problem landscape:
 - Performance - current target: analyze >> 100K pps
 - Research on: clustering; FPGA front end; multicore architecture
 - Diversity - you see the darnedest (benign) behavior & “crud”
 - Greatly complicates **anomaly detection** & detecting **evasion**
 - **Base Rate Fallacy** - detector w/ 10^{-6} error rate might not work!
- Another operational reality: intrusion **prevention**
 - Bro enabled to **automatically block** LBL traffic
 - (Very high standard for accuracy!)
 - #1 gain: dropping *scanners*

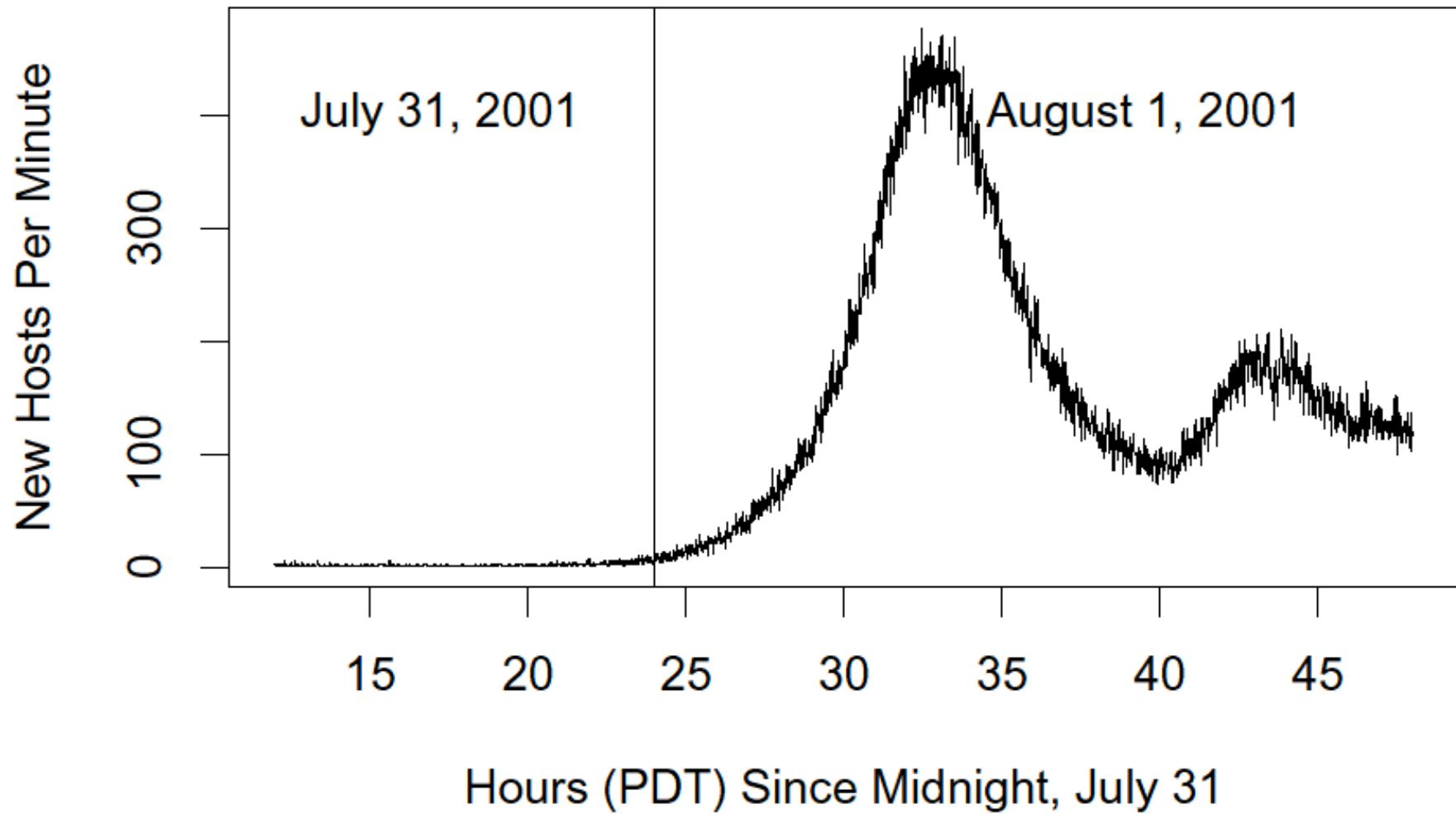
Scan Activity Seen At LBL



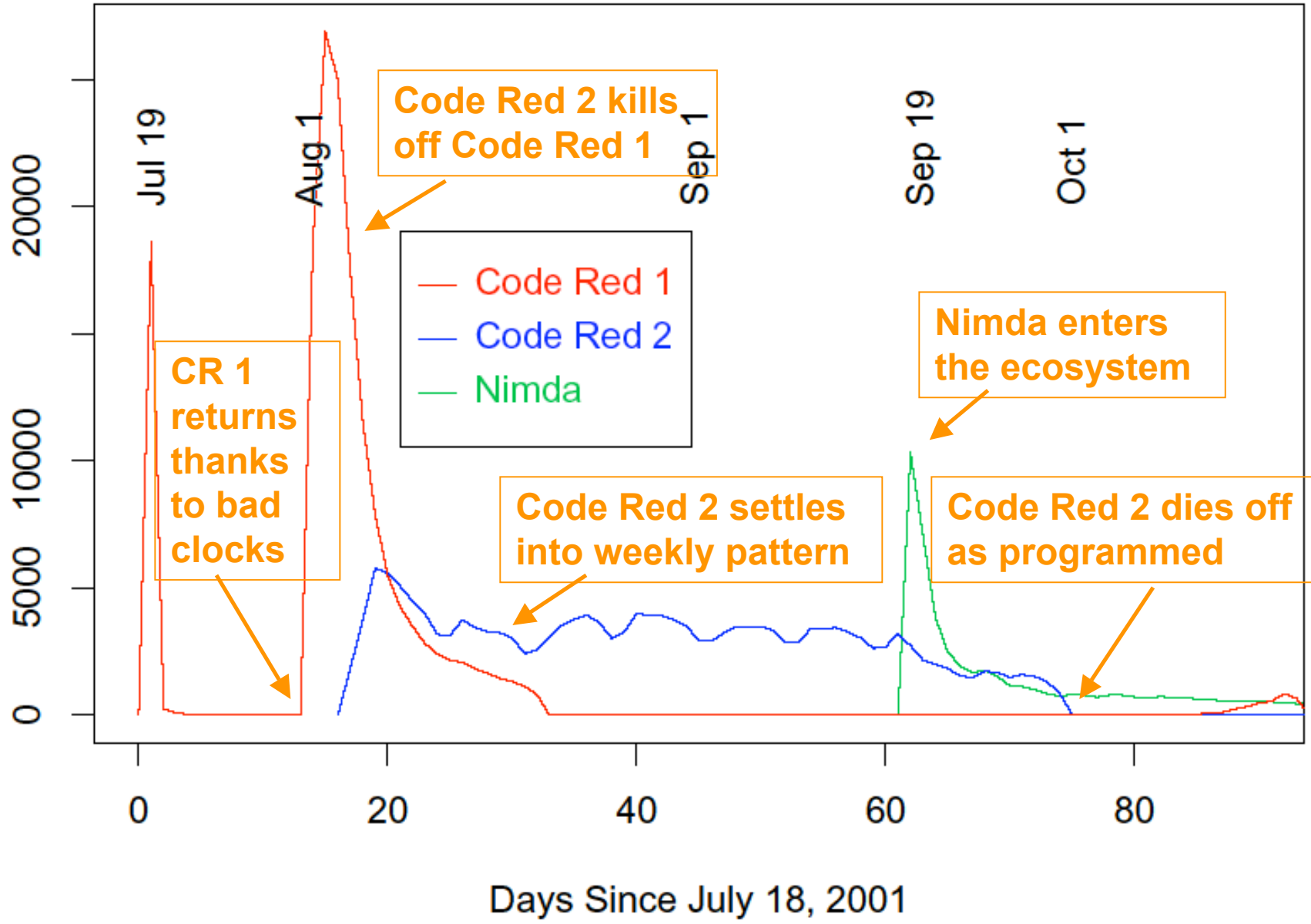
Growth of Code Red Worm



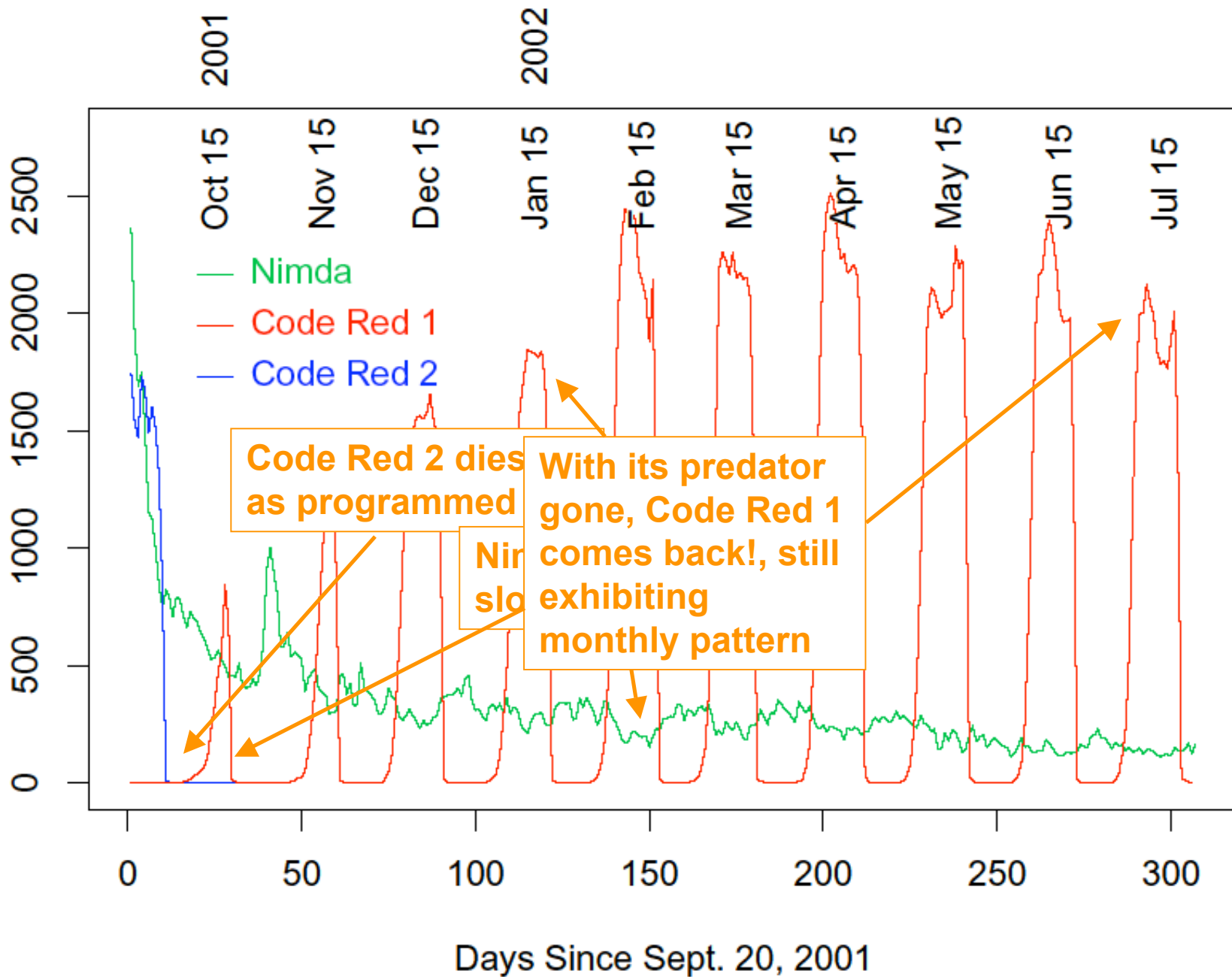
Return of Code Red Worm



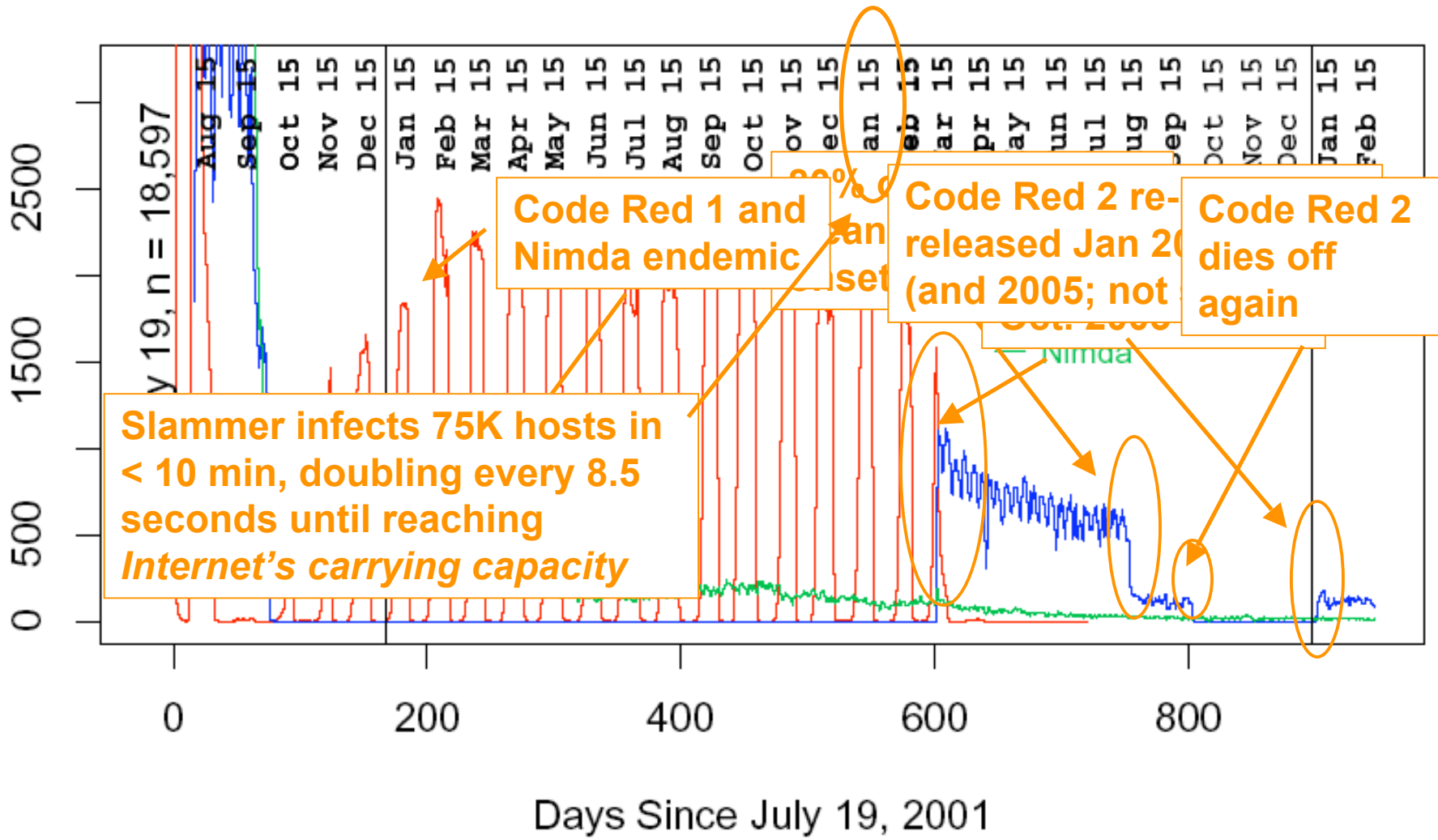
Distinct Remote Hosts Attacking LBNL



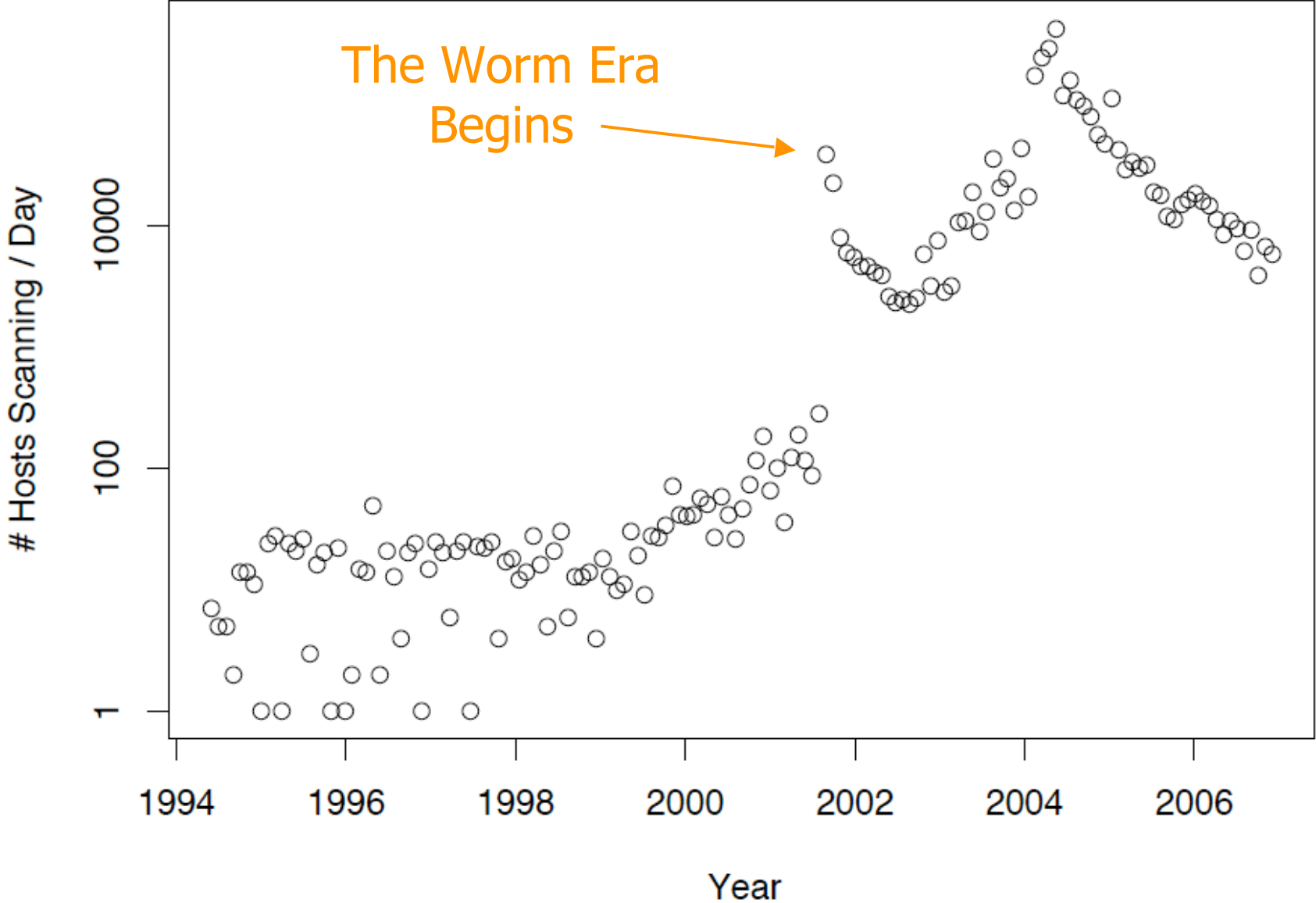
Distinct Remote Hosts Attacking LBNL



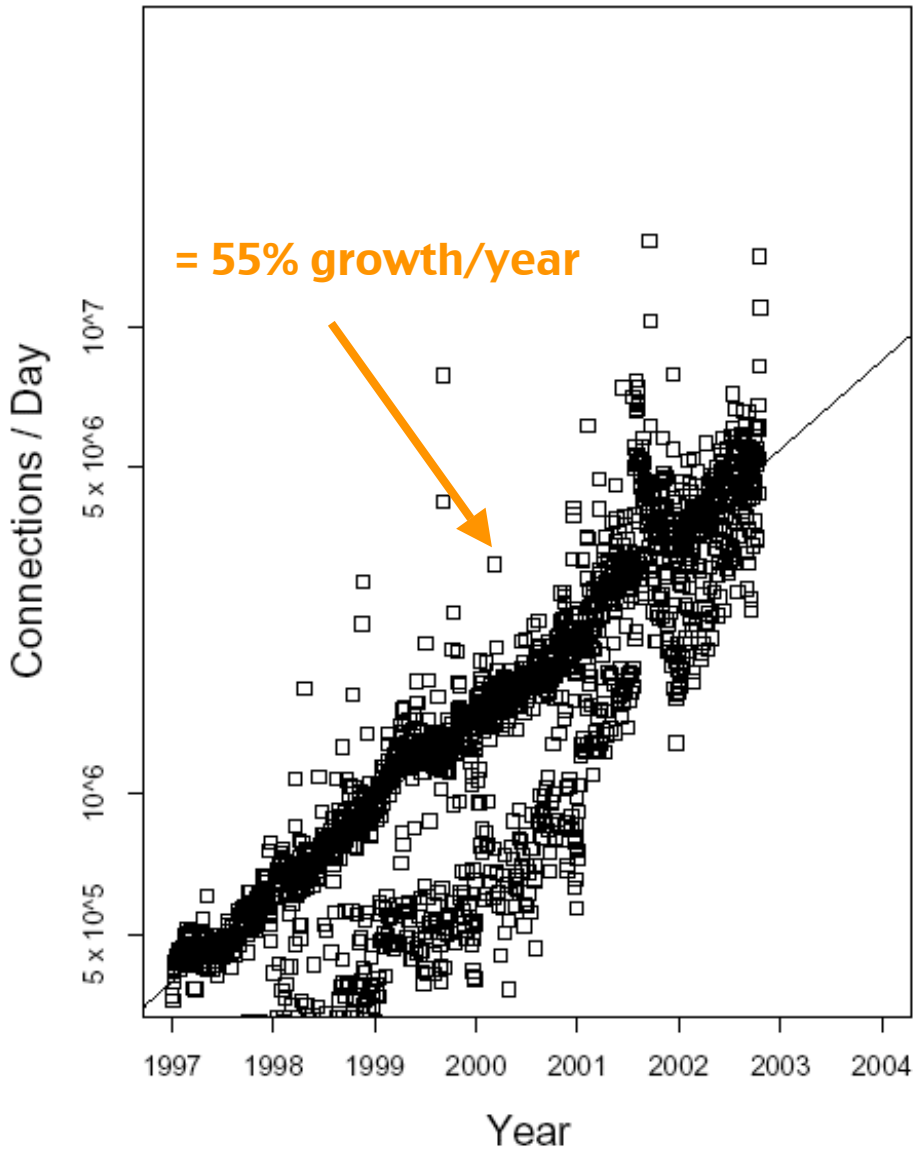
Distinct Remote Hosts Attacking LBNL



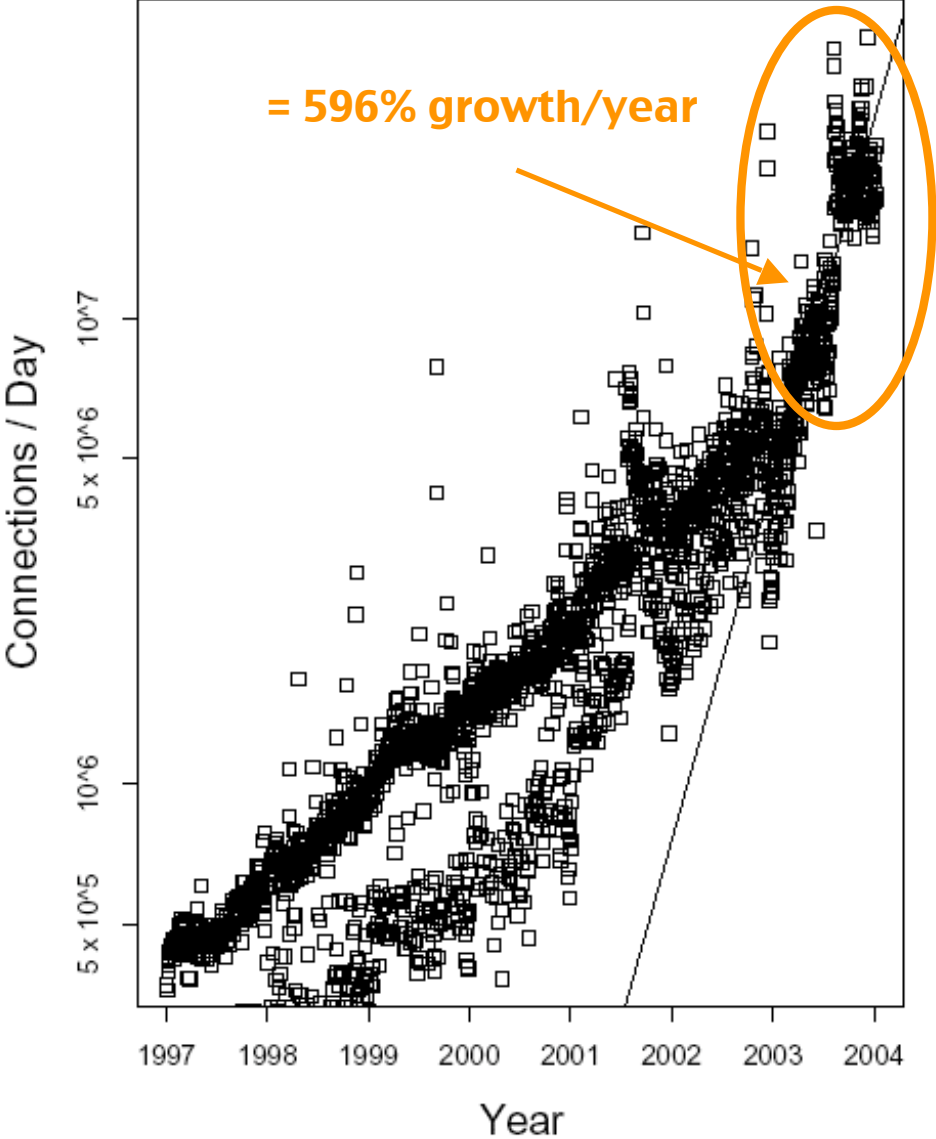
Scan Activity Seen At LBL



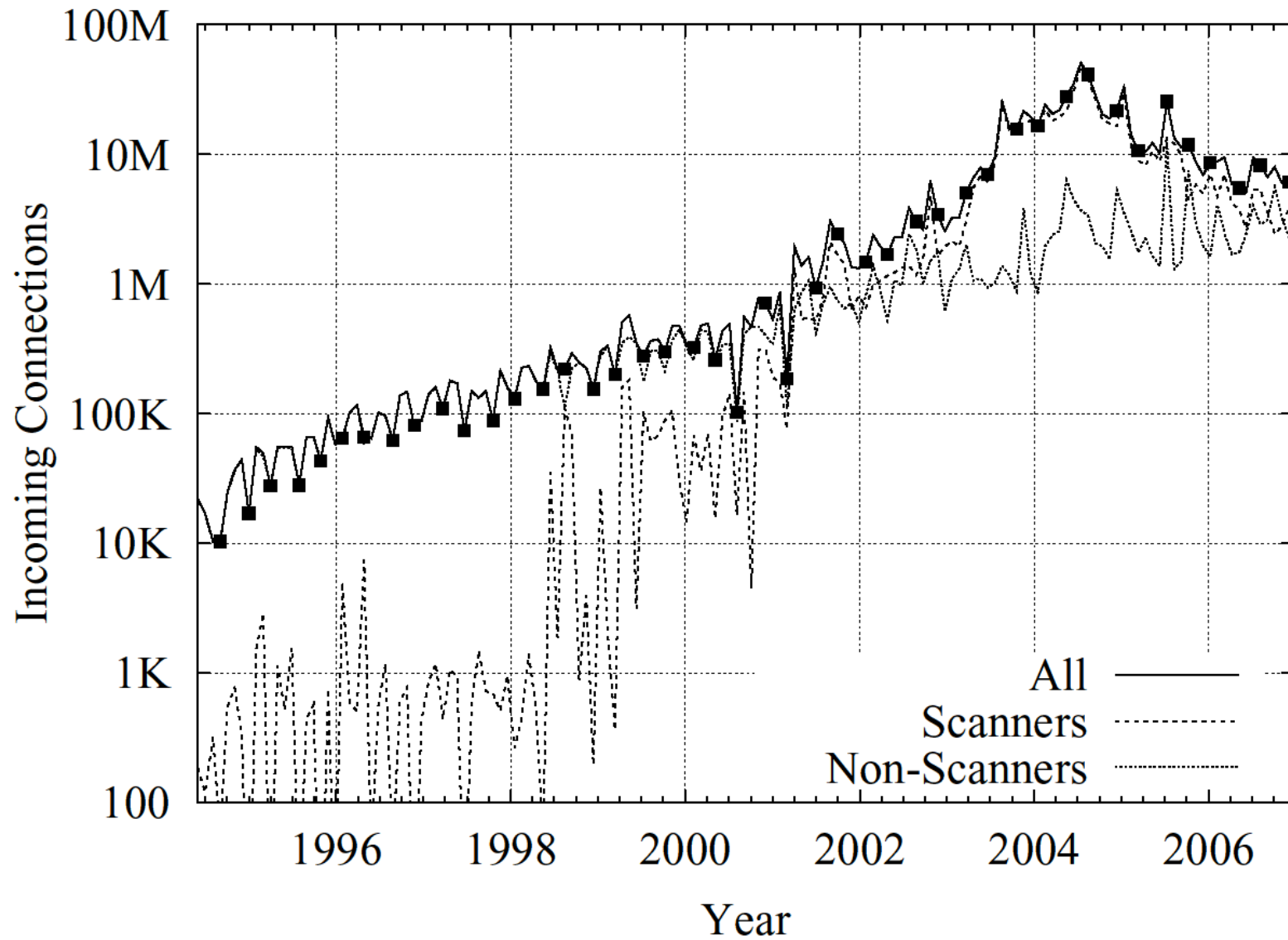
LBNL Traffic Volume, 1997-2004



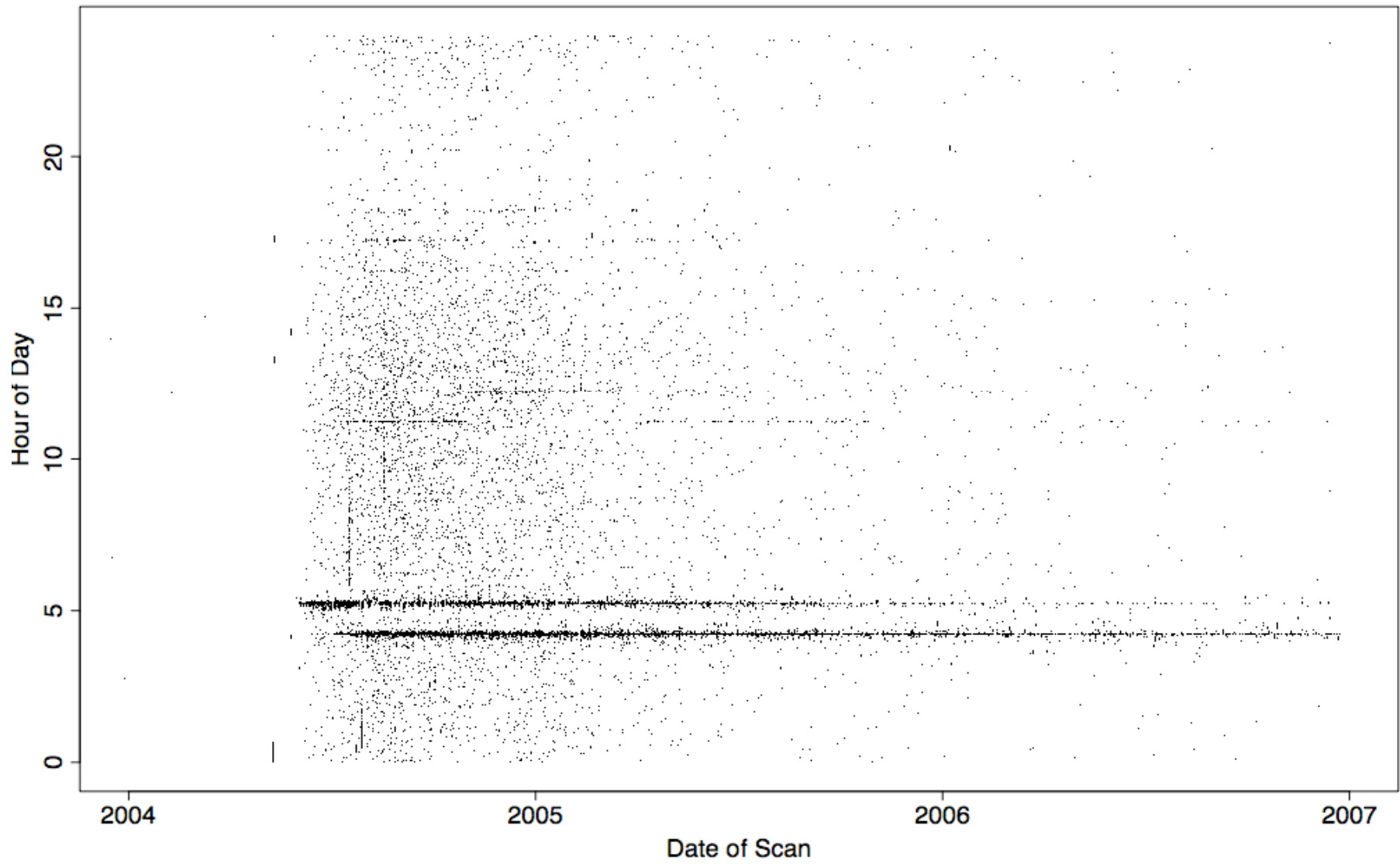
LBNL Traffic Volume, 1997-2004



Scanning Activity Seen @ LBL

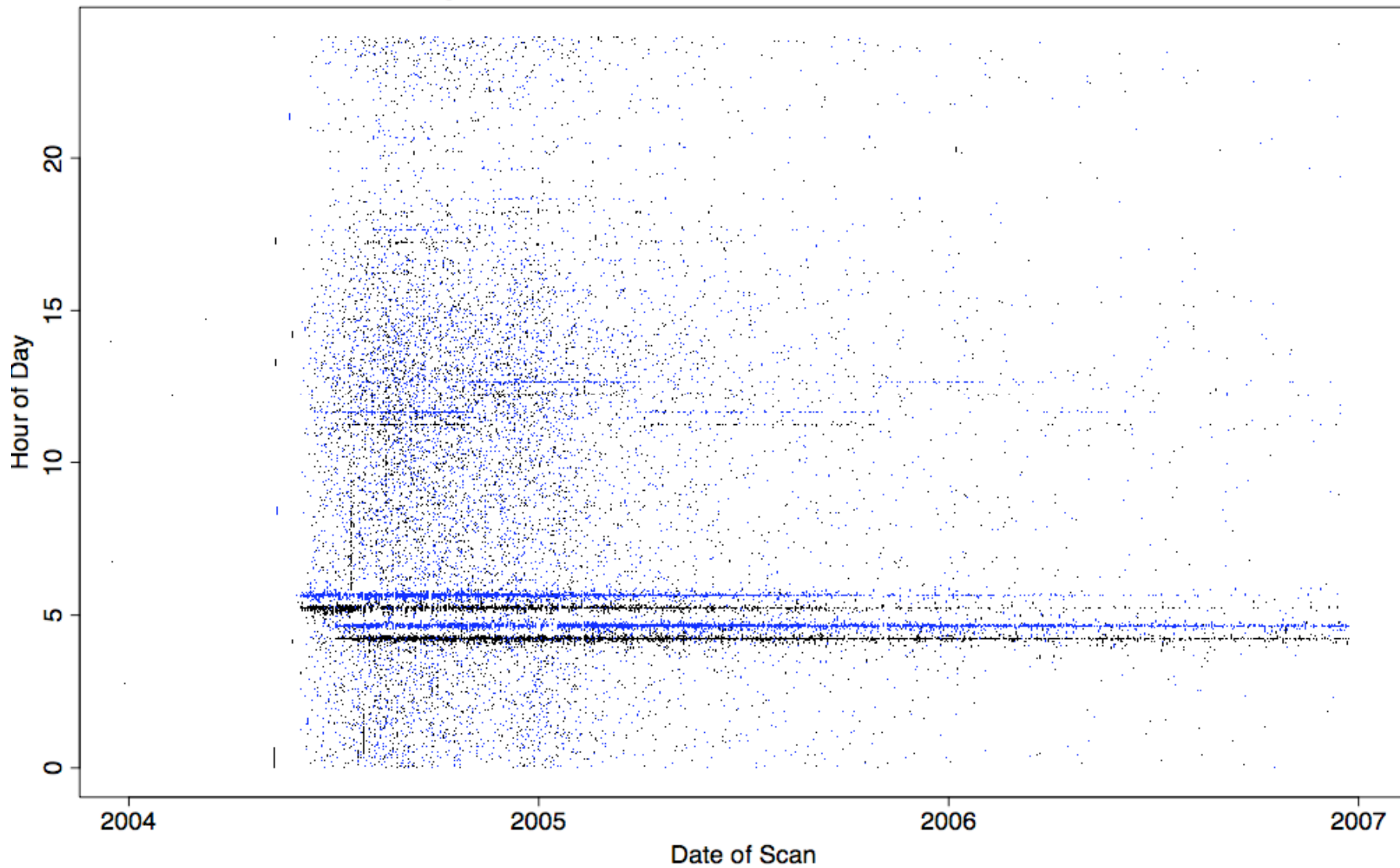


Daily Patterns Seen in 1023/TCP Scans



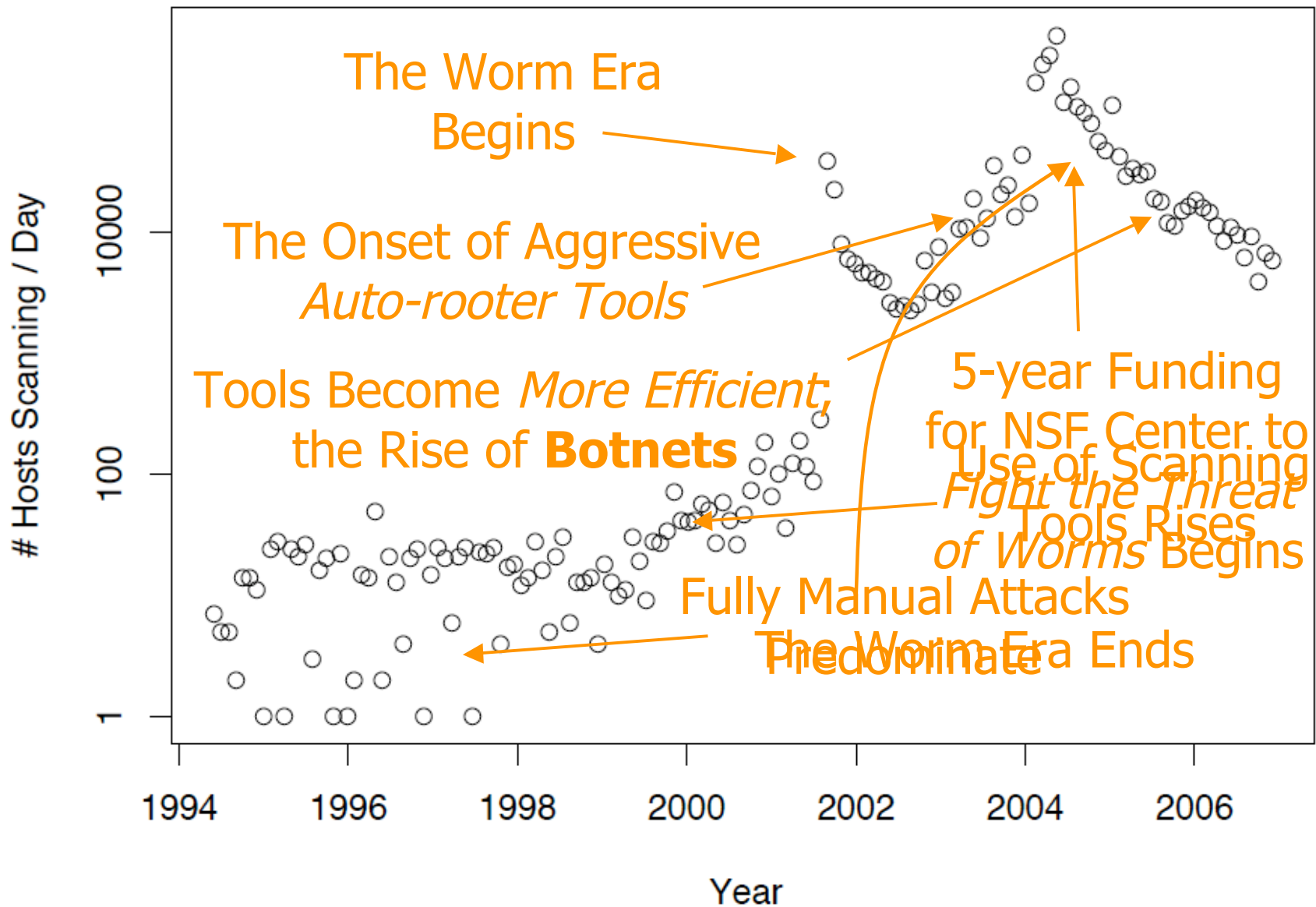
/16 at LBL, sampled 1-in-1K

Daily Patterns Seen in 1023/TCP Scans



/16 at LBL, sampled 1-in-1K
2nd /16, sampled 1-in-1K

Scan Activity Seen At LBL



Part II

Selling Viagra[®]



Know Your Enemy

- A sophisticated underground economy has emerged to **profit** from Internet subversion



My Documents

ProAgent V2.0 Public Edition

Send Menu

- Send Passwords
- Send CD-Keys
- Send KeyLog
- Send System Information
- Send Address Book
- Send URL History
- Send Processes Log

Options

- Give a fake error message
- Melt server on install
- Disable AntiVirus Programs
- Clear Windows XP Restore Points
- Protection for removing Local Server

Server Icon

You can choose any icon for server



Choose Icon

Bind with File

Bind with File

You can bind server with any files you want

Select File To Bind

Notification

Your e-mail address which you will to receive information from ProAgent.

E-Mail: bomberman@yahoo.com

Test

Decryptor

Remove Server

About

Buy Undetectable

Help

Create Server

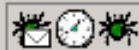
ProAgent - Professional Agent Copyright © 2005 SIS-Team



Recycle Bin



ProAgent



9:56 AM

ProAgent v2.1



- ProAgent Spy Software is one of the most powerful monitoring and surveillance applications available today.
- It is an ultimate solution for monitoring spouses, children, employees, or anyone else!
- ProAgent records all typed keystrokes, all active window texts, all visited web sites, usernames, passwords and more and sends e-mail reports to your e-mail address that you specified when creating the server, completely hidden!
- ProAgent can work in all kind of networks, it doesn't matter if the PC is behind a firewall or behind a router or in a LAN, ProAgent works in all of these conditions without any problems.

[Click here to purchase ProAgent v2.1 Special Edition...](#)

[Click here to download ProAgent v2.1 Public Edition](#)

SIS - Products

Purchase Program

Customer Support Department



Commercial Programs

Freeware Programs

Custom Special Programs

New Generation Software Solutions...

New Products

SIS-IExploiter v2.0

ProAgent v2.1



AntiDote v1.2

SIS-Downloader

Virtual Keyboard



Список доступных акков

Сервис по продаже аккаунтов аукциона eBay.

Добрые юзеры аукциона eBay предлагают вашему вниманию свои аккаунты.
Постоянным клиентам и тем, кто берет более 5 акков, различные бонусы и скидки.
Все аккаунты с доступом к мылу холдера.

Вы сами выбираете акк (несколько акков) из списка. Говорите мне. Оплачиваете и получаете.
Все акки предварительно проверяются перед продажей, в случае, если что-то не работает - 100% замена.

Актив/не актив смотрите сами по юзер ид. По активности не сортирую, так как это для каждого субъективно.

Также в продаже бывают акки PayPal. Цены рыночные. Постоянно не продаю.

Оплата по WM.

Перед покупкой следует обязательно ознакомиться с FAQ.

По работе с товаром не консультирую.

Работа через гарант сервис приветствуется.

Мои цены:

seller/баер акк до 10 фидов = 5\$

seller/баер акк 10-25 фидов = 10\$

seller/баер акк 25-50 фидов = 15\$

seller/баер акк более 50 фидов = 25\$



[Home page](#) > [Current bids](#)

Sign In

Username

Password

[Sign In](#)

New user? [Sign up here](#)

News

[PRESS RELEASE](#) 03/07/2007
Finally a Marketplace Site for Security Research

Current bids		MarketPlace history			
4 items found, displaying all items. Page 1					
Code	Time to live	Title	System	Offer type	Bid
ZD-00000007	9d 13h 26m	Local Linux kernel memory leak	Linux	Bidding	600€ 1 bid(s)
ZD-00000005	9d 13h 26m	Yahoo! Messenger 8.1 remote buffer overflow	Windows XP	Bidding	2,000€ 0 bid(s)
ZD-00000004	9d 13h 26m	Squirrelmail GPG Plugin Command Execution	Web application	Bidding Buy now at	600€ 1 bid(s) 1,750€
ZD-00000008	10d 13h 26m	MKPortal SQL injection	Web application	Bidding Buy now at	500€ 0 bid(s) 800€

Current bids		MarketPlace history			
4 items found, displaying all items. Page 1					
Code	Time to live	Title	System	Offer type	Bid
ZD-00000007	9d 13h 26m	Local Linux kernel memory leak	Linux	Bidding	600€ 1 bid(s)
ZD-00000005	9d 13h 26m	Yahoo! Messenger 8.1 remote buffer overflow	Windows XP	Bidding	2,000€ 0 bid(s)
ZD-00000004	9d 13h 26m	Squirrelmail GPG Plugin Command Execution	Web application	Bidding Buy now at	600€ 1 bid(s) 1,750€
ZD-00000008	10d 13h 26m	MKPortal SQL injection	Web application	Bidding Buy now at	500€ 0 bid(s) 800€

allBots Inc.

Social Networking Bots

GOOD News!!! We have something more for you! Yes, we have just integrated CAPTCHA Bypasser* in all of our bots.

Winsock (Multi-threaded) Bots

Become an **Affiliate** and **Start Earning Now**

Click here for 30+ MySpace Bots

Accounts Creator

(You Just Need To Type In The CAPTCHAs To Create Accounts)

Social Networks

MySpace Accounts Creator with Picture Uploader, Profile & Layout Manager		\$180.95	\$140.00
MySpace Accounts Creator with Picture Uploader, Profile & Layout Manager (Winsock)		\$360.95	\$320.00
YouTube Accounts Creator		\$120.95	\$95.00
Friendster Accounts Creator		\$120.95	\$95.00
Hi5 Accounts Creator		\$120.95	\$95.00
TopWorld Accounts Creator			

Friend Adders, Message Senders, Comment Posters & Others

(All Bots Work In A Conventional Manner, They Gather Friend IDs/Names And Send Friend Requests, Messages, Comments Automatically)

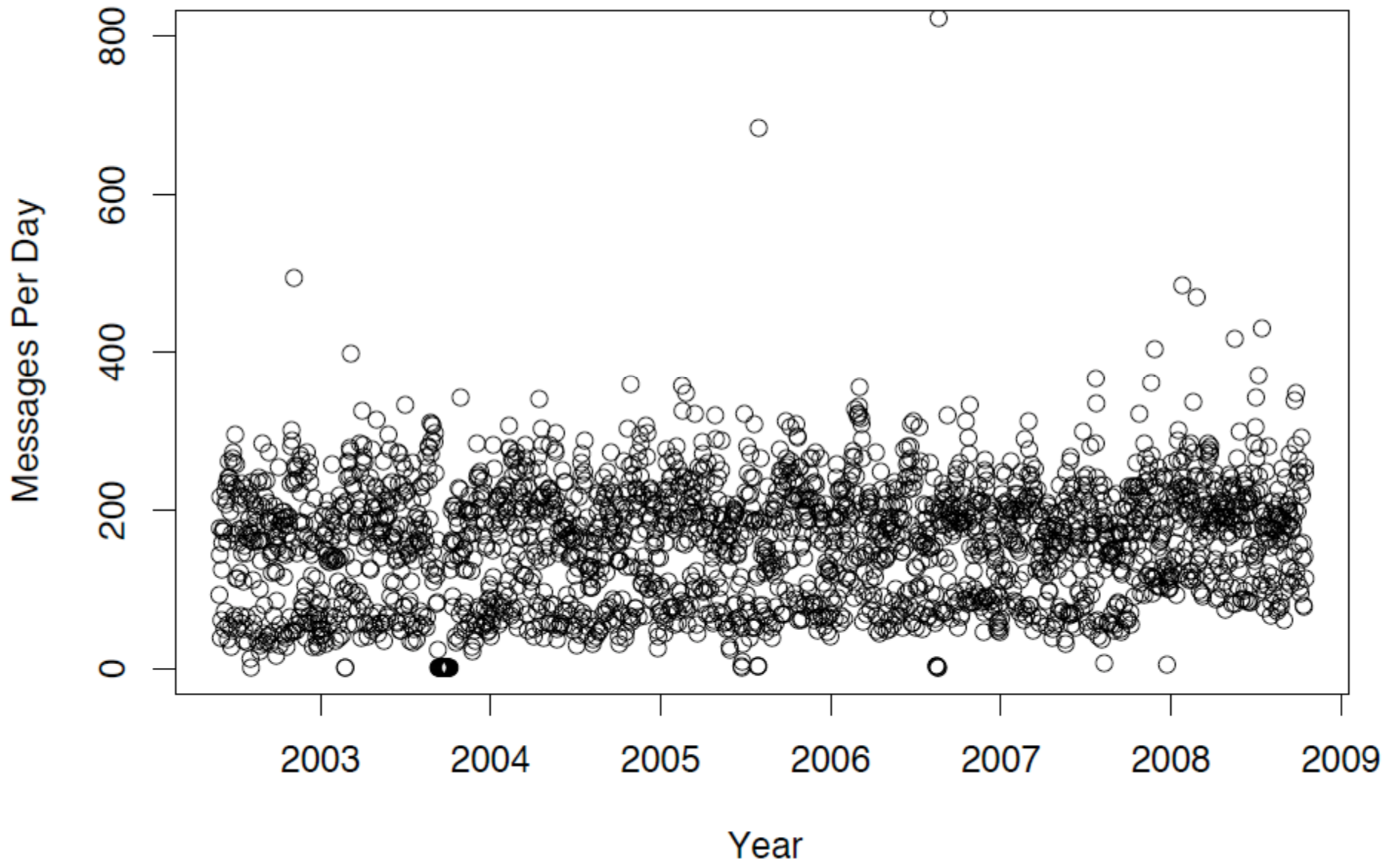
****Chaining Feature**** Is Available On All Bots for All Networks Except Facebook



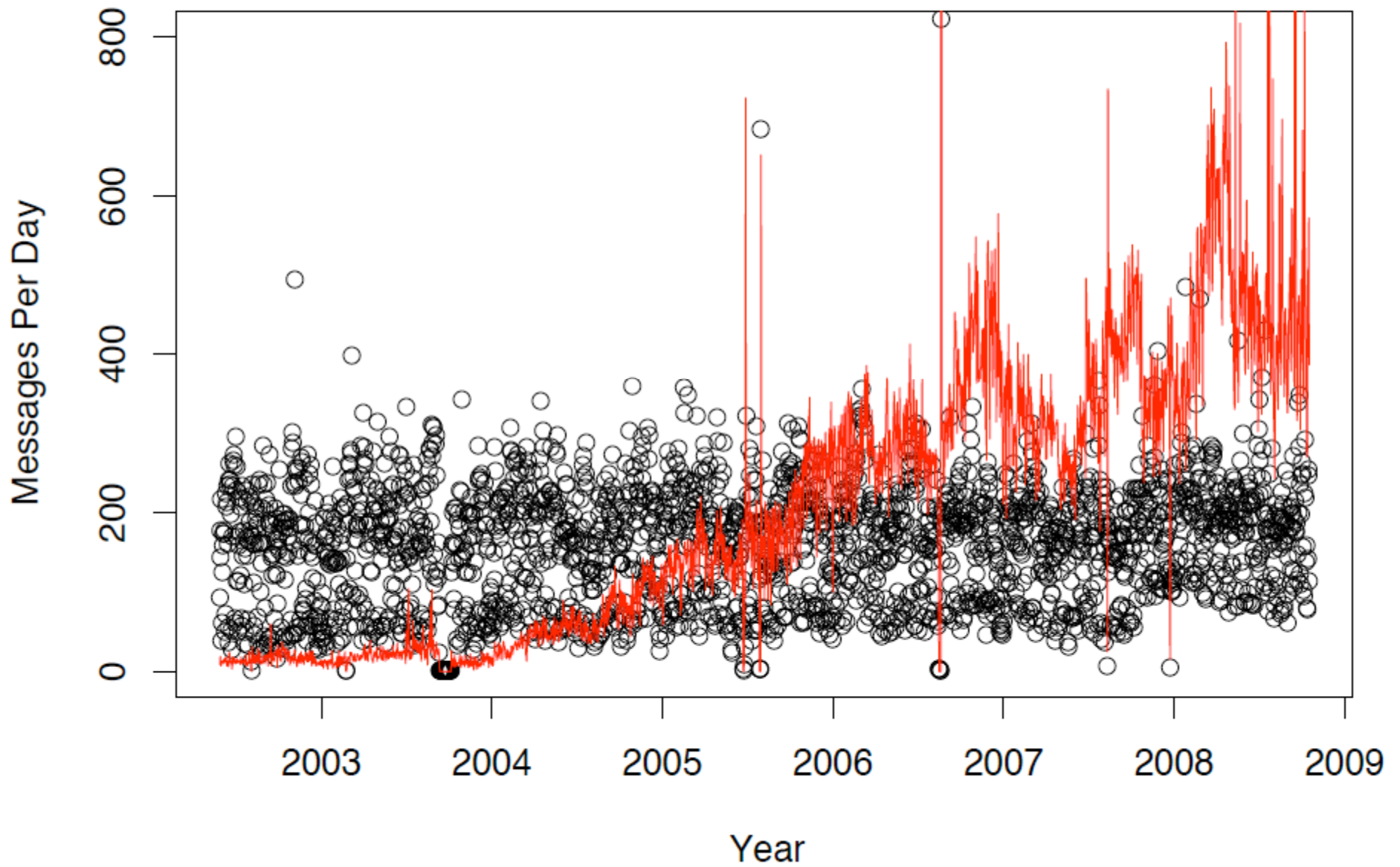
Know Your Enemy

- A sophisticated underground economy has emerged to profit from Internet subversion
- Empowered by virtually endless supply of **bots**
 - Internet systems under complete attacker control
- Dirt-cheap access to bots fuels *monetization* via relentless torrents of **spam**

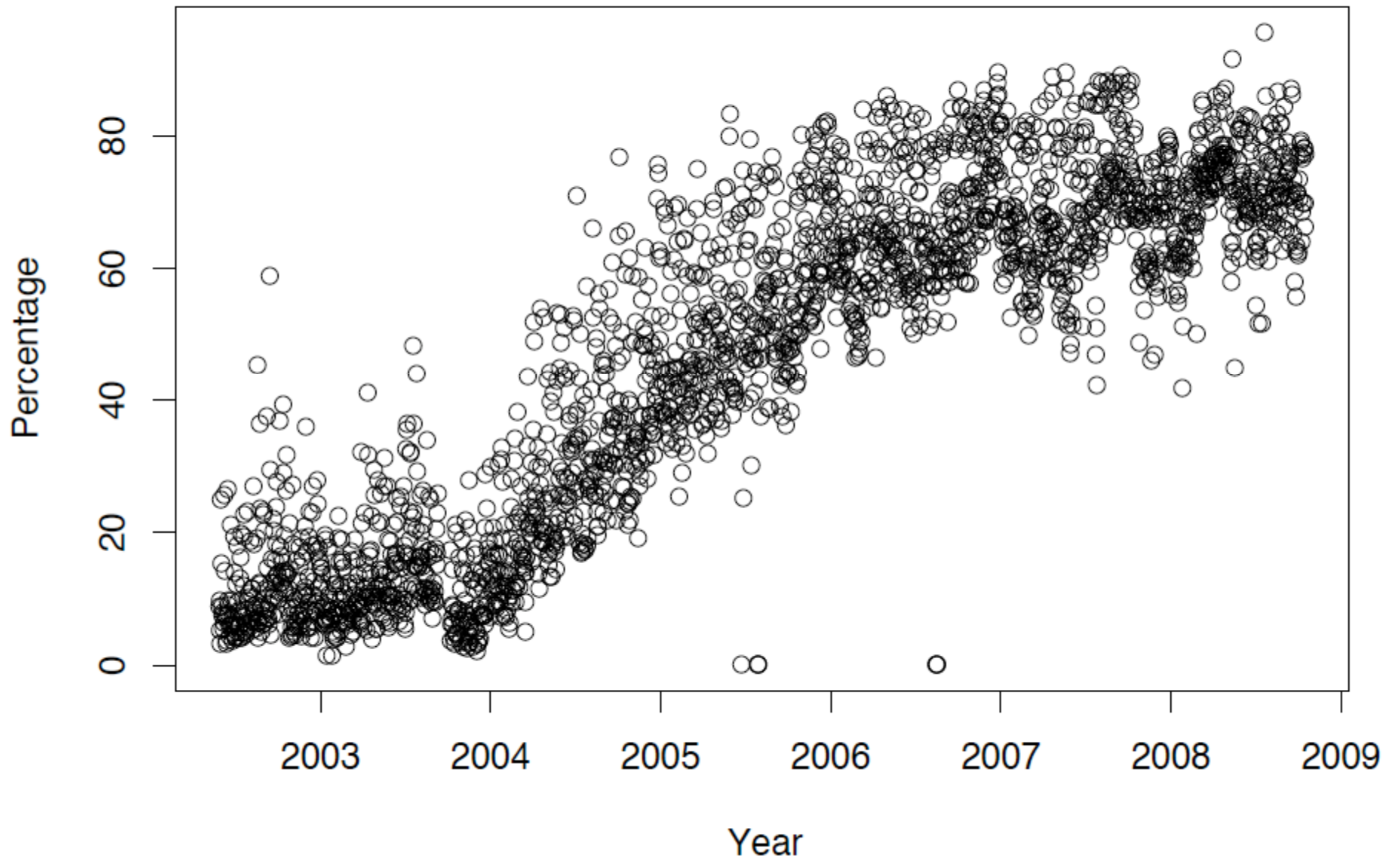
Mark Allman's Non-Spam Mail



Mark Allman's Non-Spam + Spam Mail



Fraction of Mark's Mail That is Spam



Know Your Enemy

- A sophisticated underground economy has emerged to profit from Internet subversion
- Empowered by virtually endless supply of “bots”
 - Internet systems under complete attacker control
- Dirt-cheap access to bots fuels *monetization* via relentless torrents of **spam**

- ***Just how profitable is all of this?***



Are Bots & Spam the New Black Gold?

Storm worm 'making millions a day'

Compromised machines sending out highly profitable spam, says IBM security strategist

Clive Akass, Personal Computer World 11 Feb 2008

The people behind the Storm worm are making millions of pounds a day by using it to generate revenue, according to IBM's principal web security strategist.

Joshua Corman, of IBM Internet Security Systems, said that in the past it had been assumed that web security attacks were essential ego driven.



How can we **measure** this?
Seemingly only knowable by
the spammers themselves.

- Spam finance elements:

- Retail-cost-to-send vs. Profit-per-response
- Key missing element: spams-needed-per-response, i.e., *conversion rate*

Welcome to Storm!



Would you like to be one of our newest bots?
Just read your postcard!

(Or even easier: just wait 5 seconds!)

September 6th, 2007

Storm Worm botnet could be world's most powerful supercomputer

Posted by Ryan Naraine @ 8:41 am

Categories: [Botnets](#), [Browsers](#), [Data theft](#), [Exploit code](#), [Firefox.....](#)

Tags: [Operation](#), [Supercomputer](#), [Malware](#), [Worm](#), [Ryan Naraine](#)



150 TalkBacks

ADD YOUR OPINION



SHARE



PRINT



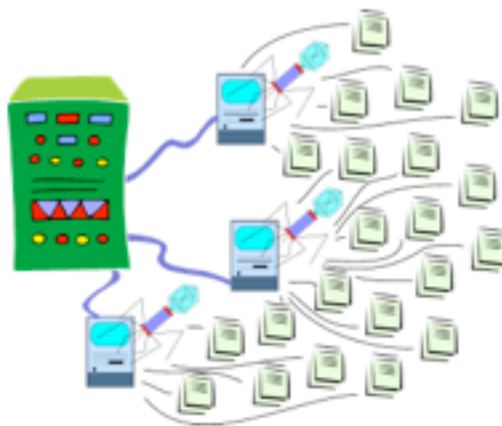
E-MAIL



WORTHWHILE?

+97

115 VOTES

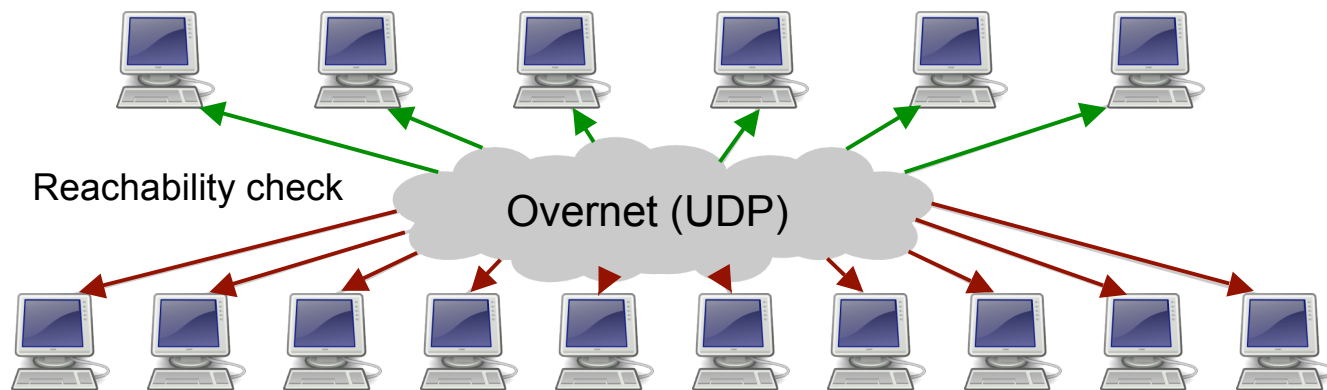


Nearly nine months after it was first discovered, the [Storm Worm](#) Trojan continues to surge, building what experts believe could be the world's most powerful supercomputer.

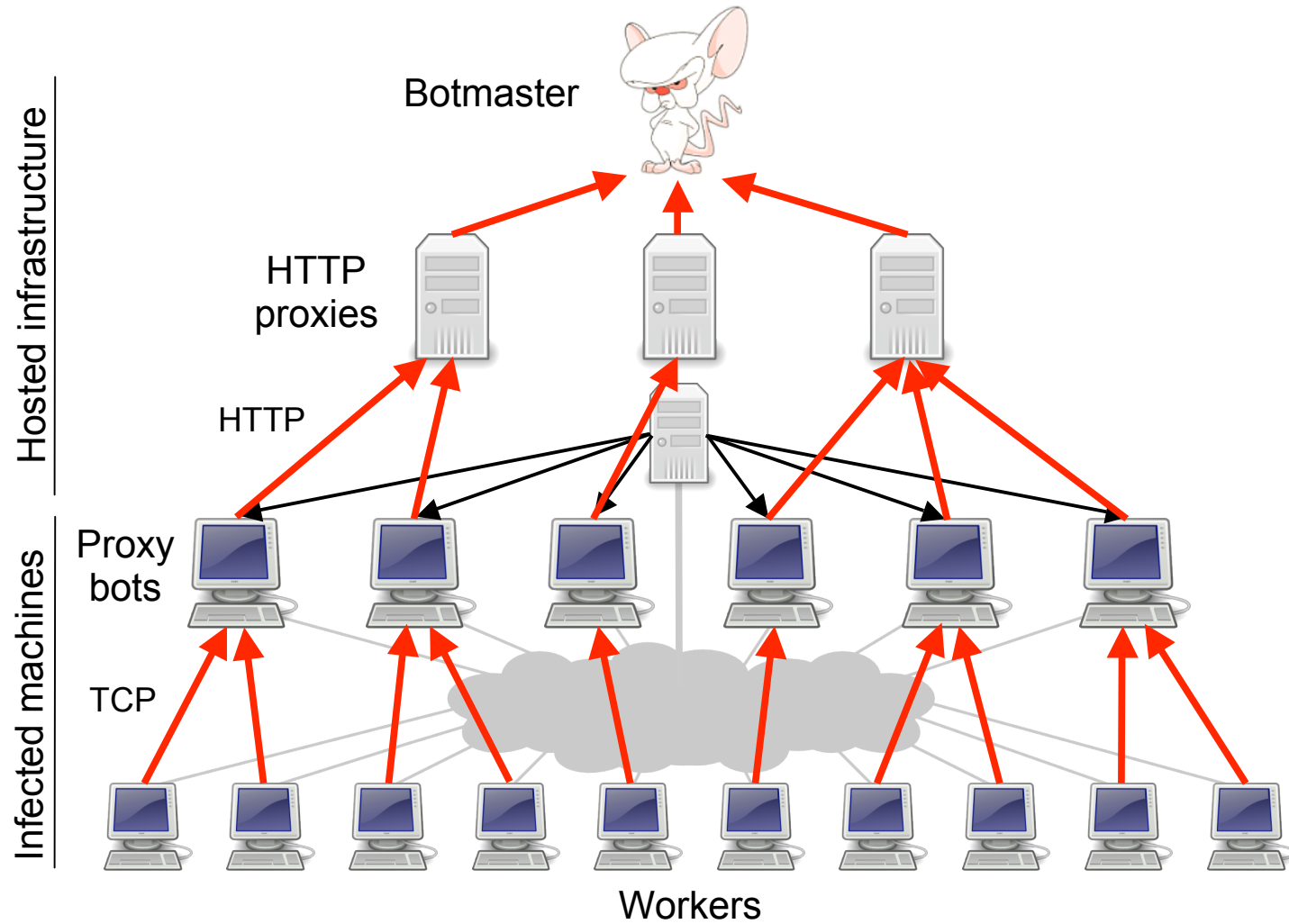
The Trojan, which uses a myriad of social engineering lures to trick Windows users into downloading malware, has successfully seeded a massive botnet — between one million and 10 million CPUs — producing computing power

to rival the world's top 10 supercomputers

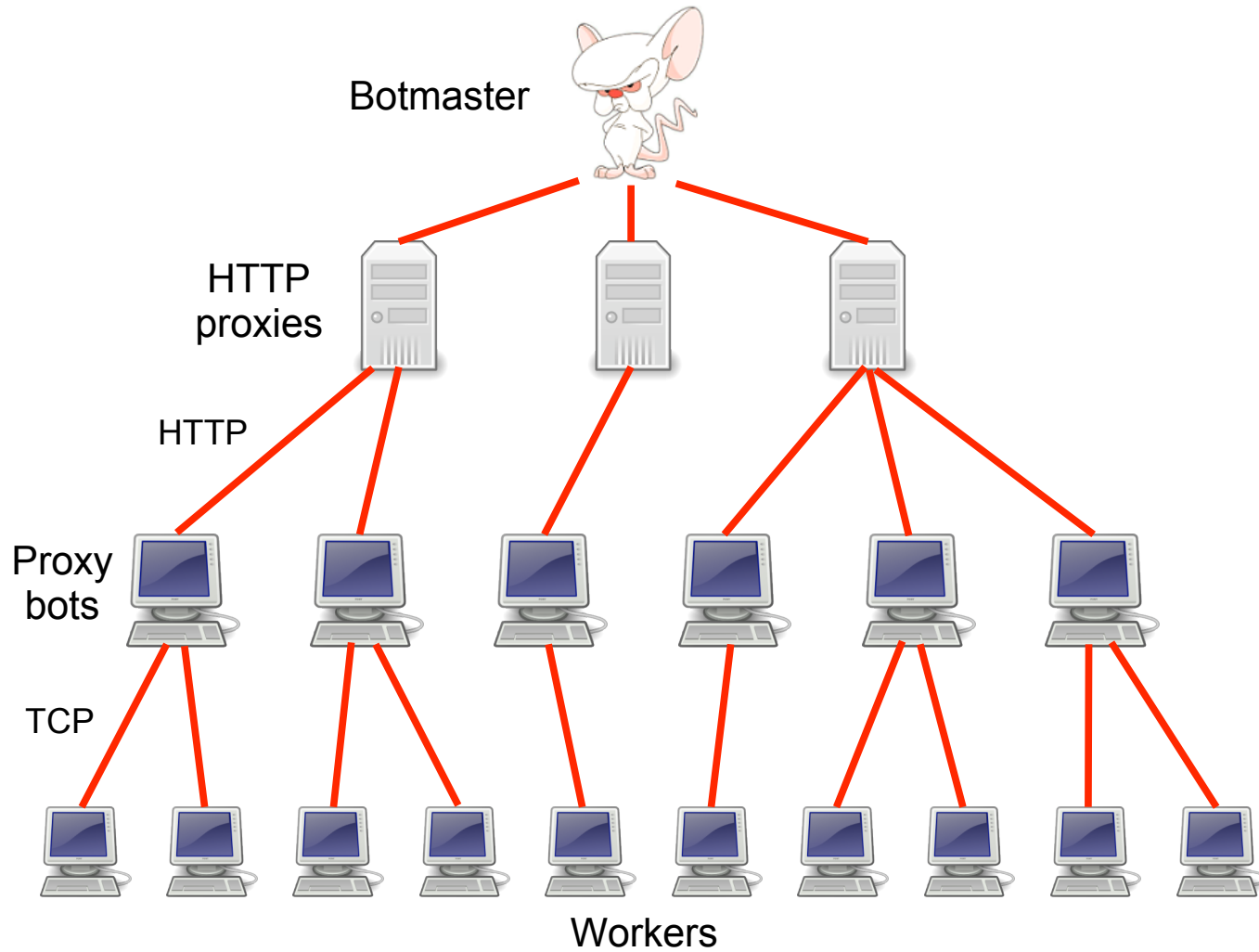
The Storm botnet



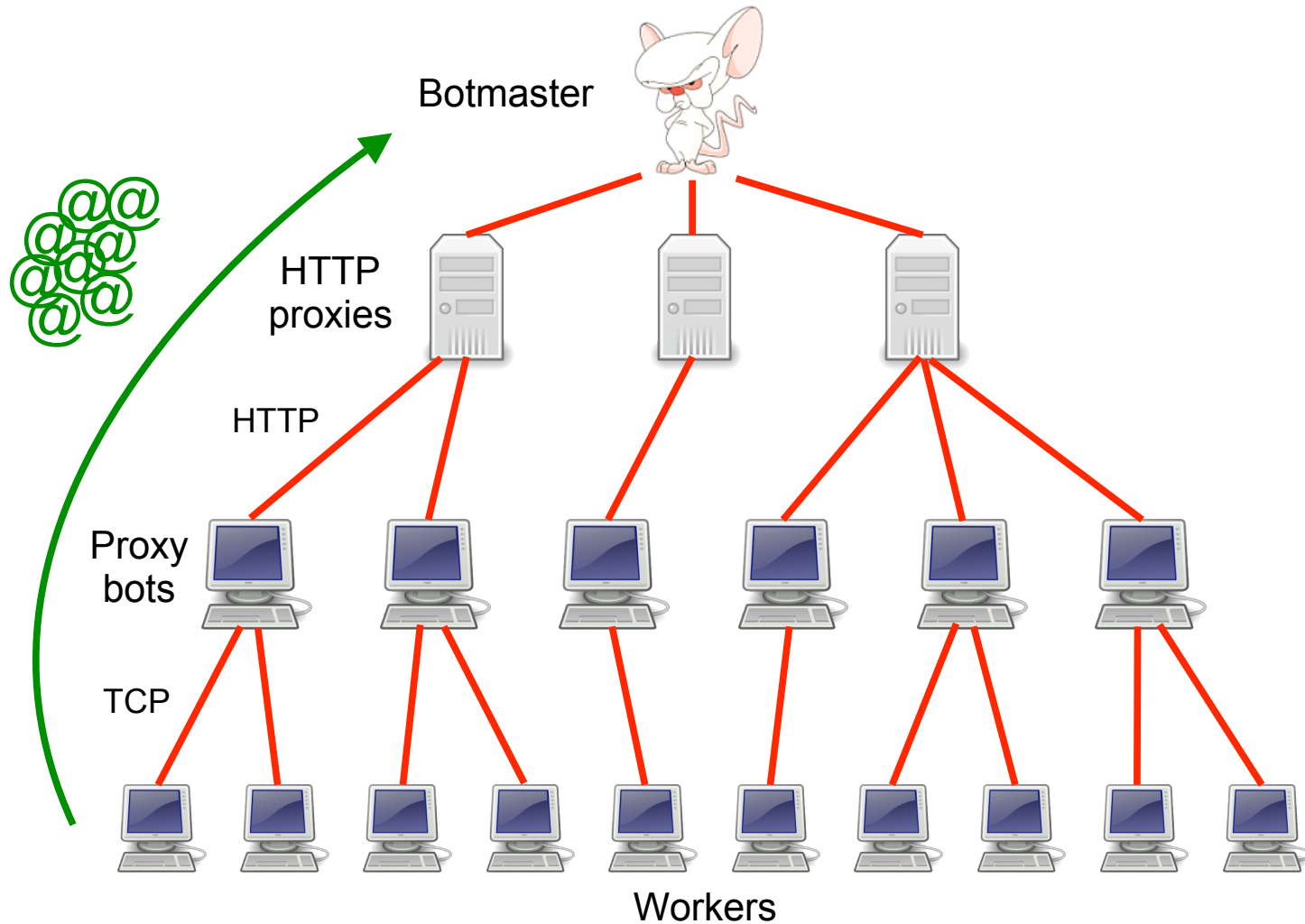
The Storm botnet



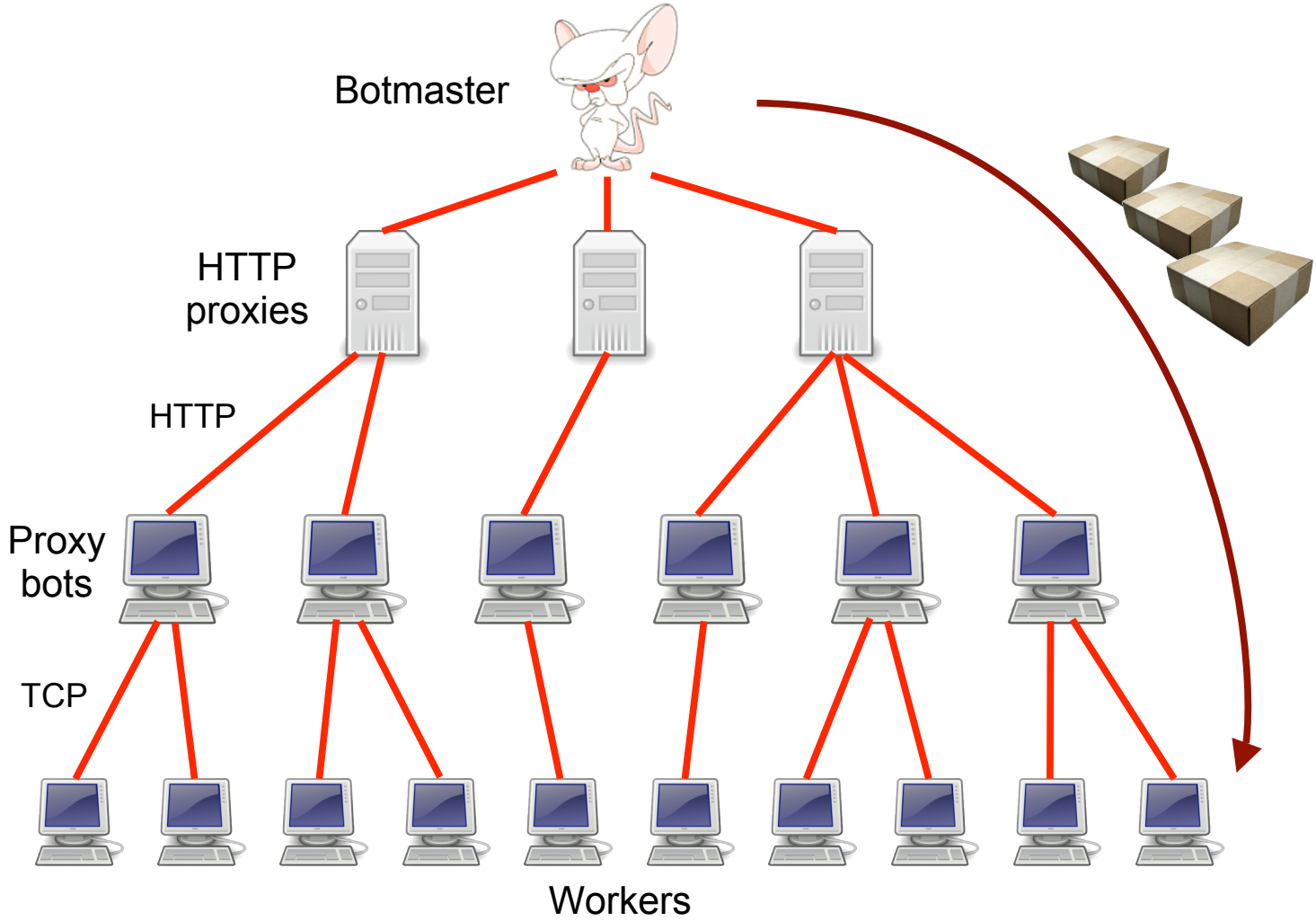
Spam campaign mechanics



Campaign mechanics: harvest



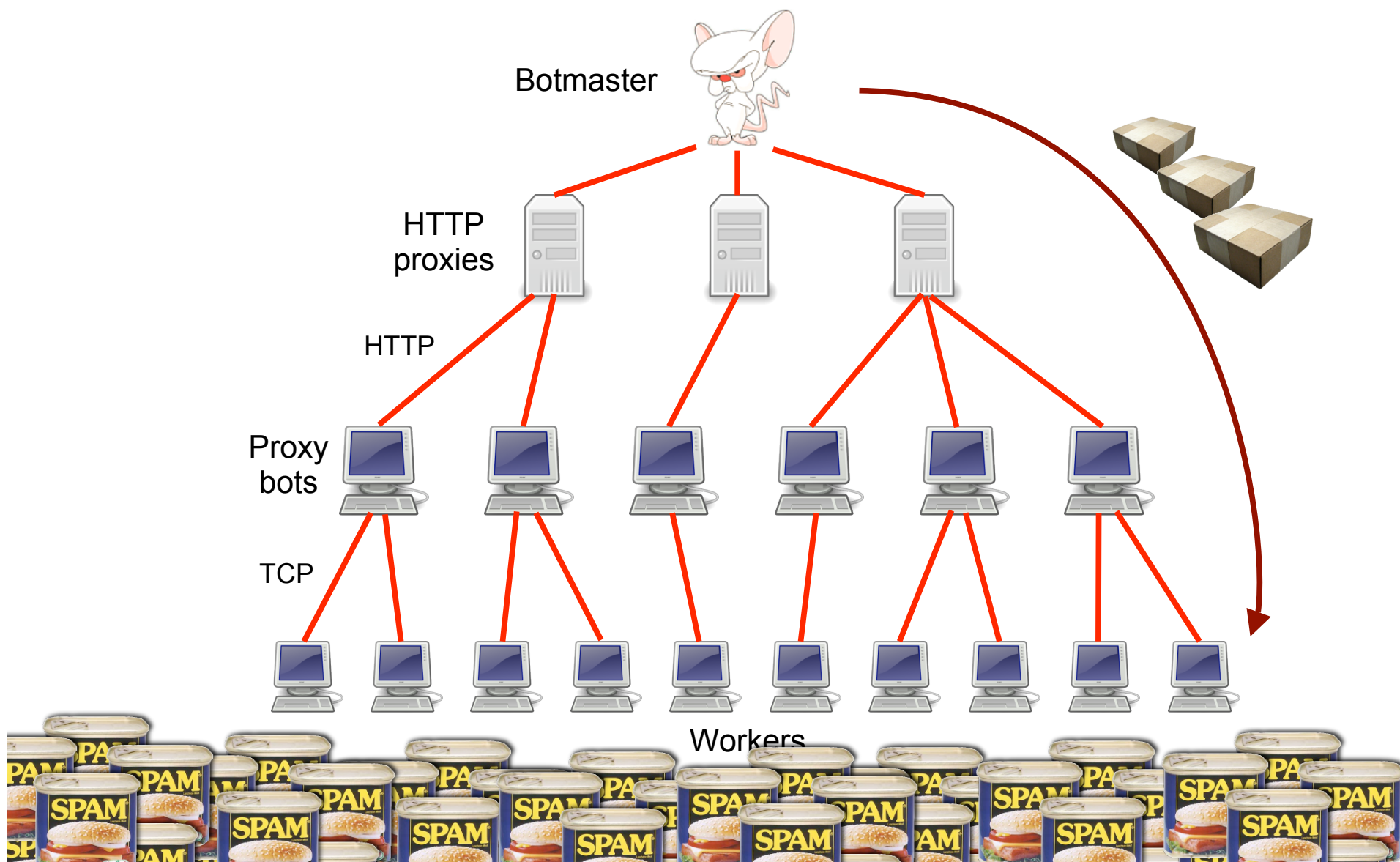
Campaign mechanics: spamming



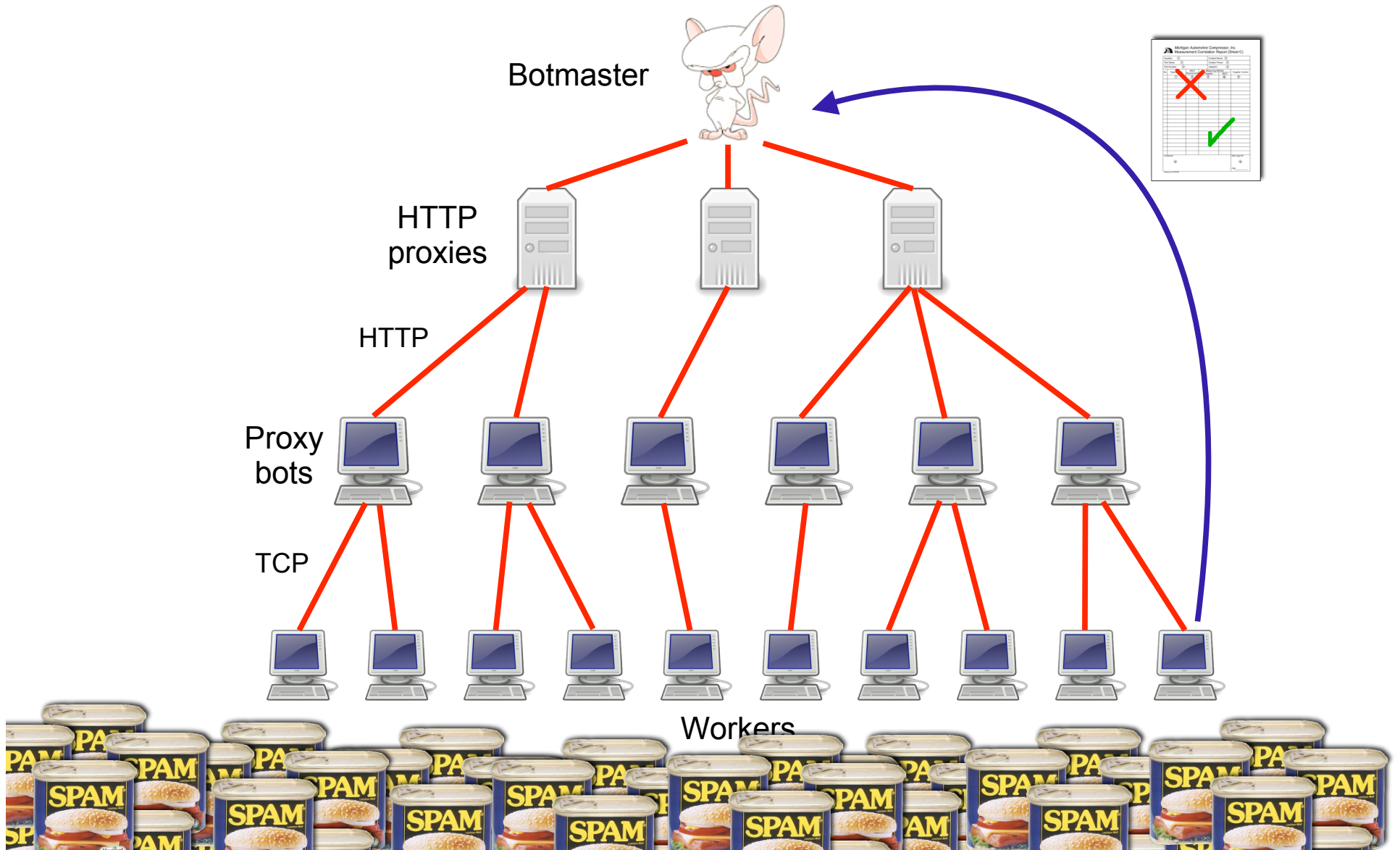
MACRO	SEEN LIVE	FUNCTIONALITY
(O)	✓	Spam target email address.
(A)	✓	FQDN of sending bot, as reported to the bot as part of the preceding C&C exchange.
(B)		Creates content-boundary strings for multi-part messages.
(Cnum)	✓	Labels a field's resulting content, so it can be used elsewhere through (V); see below.
(D)	✓	Date and time, formatted per RFC 2822.
(E)		ROT-3—encodes the target email address.
(Fstring)	✓	Random value from the dictionary named <i>string</i> . ²
(Gstring)	✓	Line-wrap <i>string</i> into 72 characters per line.
(Hstring)		Defines hidden text snippets with substitutions, for use in HTML- and plain-text parts.
(I)	✓	Random number between 1 and 255, used to generate fake IP addresses.
(Jstring)		Produces quoted-printable “=20” linewrapping.
(K)		IP address of SMTP client.
(M)	✓	6-character string compatible with Exim's message identifiers (keyed on time).
(N)		16-bit prefix of SMTP client's IP address.
(Ostring:num)	✓	Randomized message identifier element compatible with Microsoft SMTPSVC.
(Pnum ₁ [-num ₂]:string)	✓	Random string of num ₁ (up to num ₂ , if provided) characters taken from <i>string</i> .
(Qstring)		Quoted-printable “=” linewrapping.
(Rnum ₁ -num ₂)	✓	Random number between num ₁ and num ₂ . Note, special-cased when used with (D).
(Ustring)		Randomized percent-encoding of <i>string</i> .
(Vnum)	✓	Inserts the value of the field identified by (Cnum).
(W)		Time and date as plain numbers, e.g. “20080225190434”.
(X)		Previously selected member of the “names” dictionary.
(Ynum)	✓	8-character alphanumeric string, compatible with Sendmail message identifiers.
(Z)	✓	Another Sendmail-compatible generator for message identifiers.

Table 2: Storm's spam-generation templating language.

Campaign mechanics: spamming



Campaign mechanics: reporting



Welcome to Storm! What can we sell you?

The screenshot displays the Canadian Pharmacy website interface. At the top, there is a navigation menu with links for Home, Bestsellers, All products, FAQ, and Contact us. A currency selector shows \$, €, and £, along with a 'Pharma Bonus' icon. A shopping cart icon indicates 'Your cart: \$0.00 (0 items)' with a 'Proceed to Checkout' button.

The main banner features the Canadian Pharmacy logo and the tagline '#1 Internet Online Drugstore', accompanied by an image of two medical professionals. Below the banner, a 'Products list' section highlights three featured products:

- Viagra + Cialis:** 10 x Viagra 100 mg and 10 x Cialis 20 mg, priced at 69⁹⁹\$.
- Growth Pack:** 1 bottle x 60caps Growth Pills and 1 tube x 2oz Growth Oil, priced at 179⁹⁵\$.
- Viagra:** 120 pills 100 mg and +4 Free pills, priced at 225⁶¹\$.

Each product card includes an 'ORDER NOW' button. To the left of the product list is a sidebar with a 'Bestsellers' section and a list of categories: Male Enhancement, Men's Health, SALES - 20% OFF, Female Enhancement, Weight Loss, Gums New!, Body-Building, and Hypnotherapy.

Below the product list is a search bar with a 'Search by name' dropdown menu (A-Z) and a search input field. Underneath the search bar is a 'Today's Bestsellers' section featuring three products:

- Viagra:** Our price \$1.21.
- Cialis:** Our price \$2.18.
- Viagra Professional:** Our price \$3.73.

Each product in the 'Today's Bestsellers' section has a 'More info' link and an 'Add to cart' button.

Anatomy of a modern Pharma spam campaign

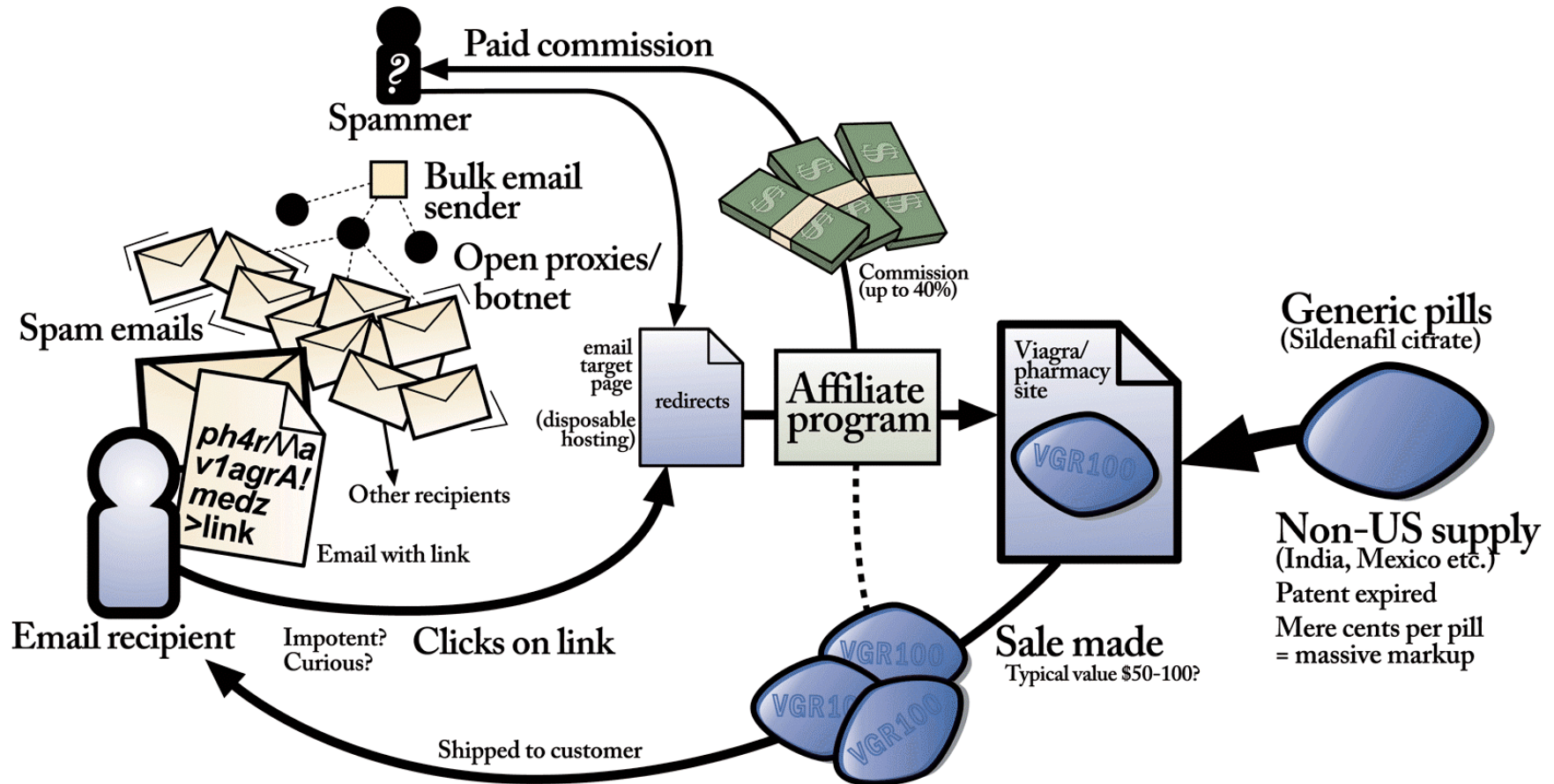


Diagram by Stuart Brown
modernlifeisrubbish.co.uk

These folks seem trustworthy ...



... how about these?





Bot master

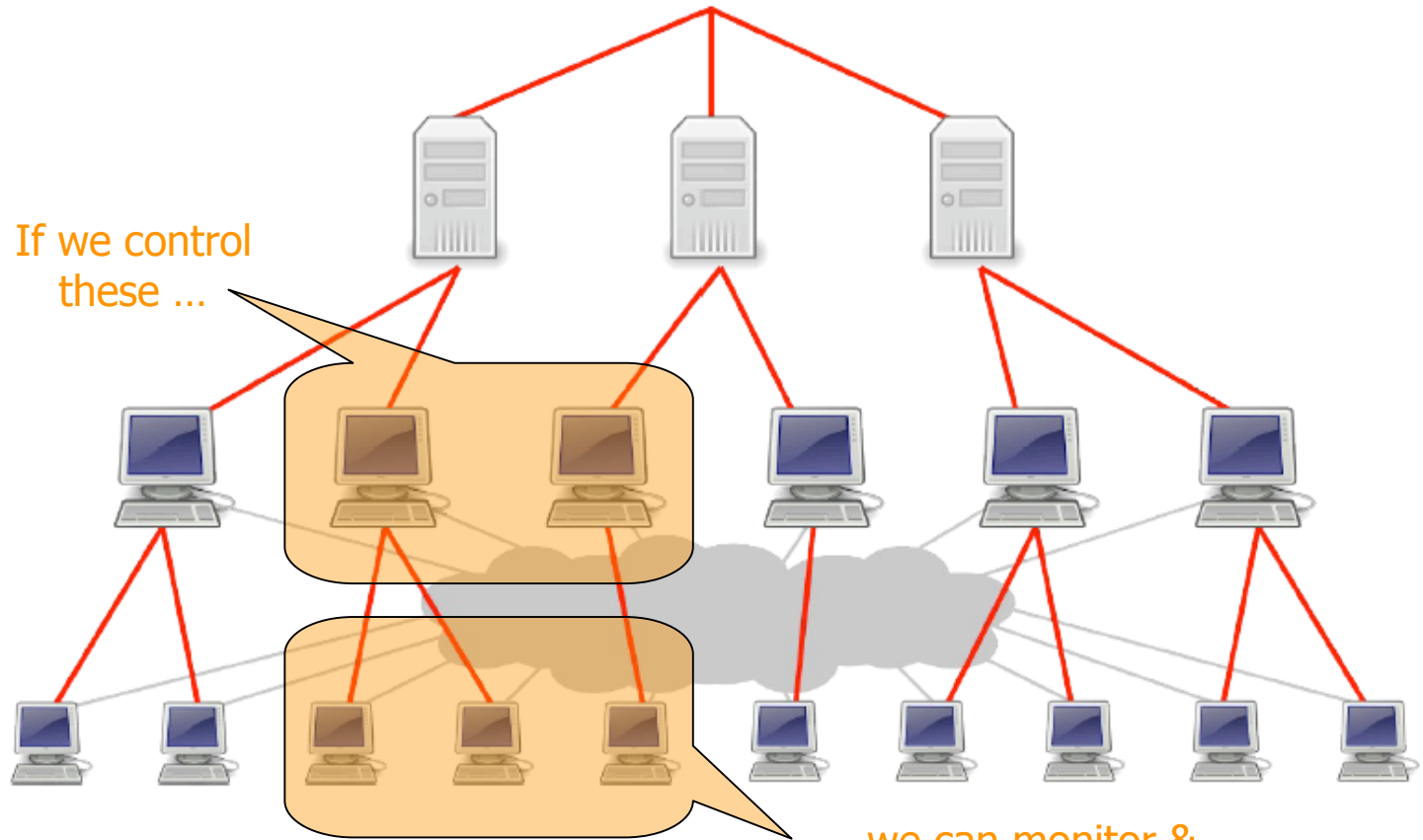
HTTP proxies

If we control these ...

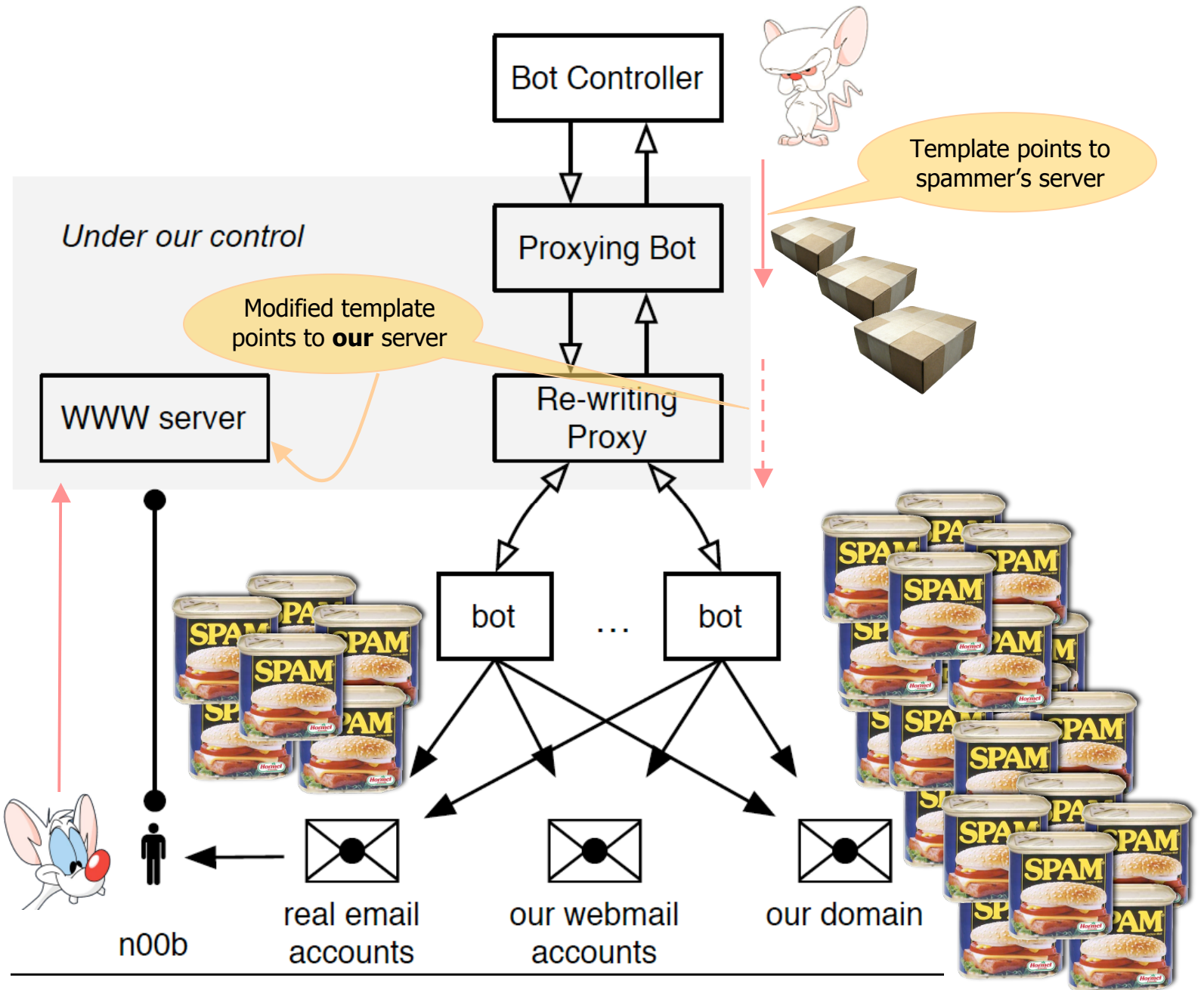
Proxy bots

Overnet

Worker bots



... we can monitor & **influence** these

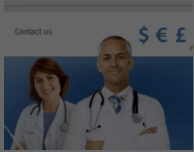


Spam conversion experiment

- Experimented with Storm March 21 – April 15, 2008
- Instrumented roughly 1.5% of Storm's total output

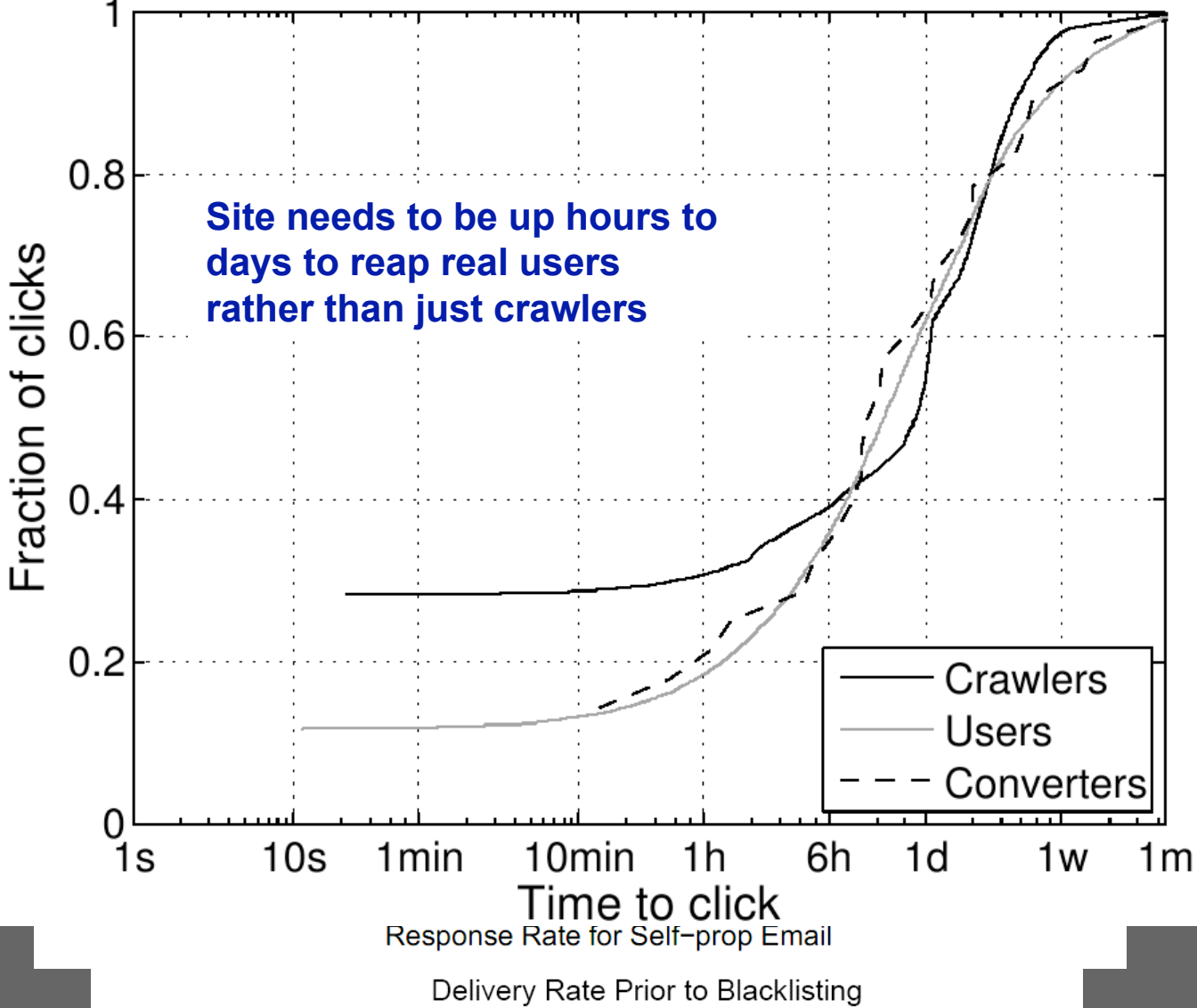
	Pharmacy Campaign	E-card Campaigns	
		Postcard	April Fool
Worker bots	31,348	17,639	3,678
Emails	347,590,389	83,665,479	38,651,124
Duration	19 days	7 days	3 days

Spa



load will start in 5 seconds.
load does not start, [click here](#)

targeted



sions

0008%)

037%)

056%)



Corresponding Revenue

- 28 purchases in 26 days, average “sale” ~\$100
 - Total: \$2,731.88, \$140/day
- **But:** we interposed on only ~1.5% of workers:
 - \$9,500/day (8,500 new bots per day)
 - \$3.5M/year
 - Though if selling Viagra via *Glavmed affiliation*, cut is **40%**
- Storm: service provider or integrated operation?
 - Retail price of spam ~\$80 per million
 - Pharmacy spam would have cost 10x the profit!
 - Strongly suggests Storm operates as an integrated operation rather than a reseller



Reflections on the Journey

- Network security research has seen enormous change over the last 15 years, from:
 - Not a field ...
 - ... to fending off ardent amateurs
 - ... to global worm epidemics
 - ... to botnets employed for spam campaigns that fuel an emergent **underground economy**
- The first of these was pretty tenable (and fun!)
- The second was daunting but the field made some surprising advances
 - Though **cyberwarfare** remains a major latent threat
- The third is even more daunting ...
 - ... deeply worrisome because it's fueled by criminals out to make **money** - **hastening the pace of adversary innovation**



Reflections on the Process

- *Measuring* is easy
- Measuring in a *meaningful* and *sound* way is hard ...
 - A lot of un-fun grunt work dealing with messiness & error
- But: only convincingly way to unearth **Truth**
- And sometimes you get surprised:
 - Pervasive diversity & exponential growth
 - Unanticipated threats & non-threats
 - Strikingly rapid changes in the landscape
- Security as a field is all about trading off resources vs. perceived risks
 - ⇒ Deep fundamental need for well-grounded empirical data
- In today's threat environment, biggest defense payoffs can come from understanding (= measuring) and then **undermining** attacker profit ...
 - ... rather than securing systems pointwise.