Toward Lowering the Load on Root DNS Servers

Duane Wessels
The Measurement Factory, and
CAIDA
wessels@measurement-factory.com

NANOG 26
October 2002

The Roots

- Thirteen "starting points" for finding anything in the DNS.
- [a-m].root-servers.net
- Each handles $\sim$ 100,000,000 queries per day.
  - See `http://www.caida.org/~kkeys/dns/`
- Records in the root zone have large TTLs (`.` 6 days, `com.` 2 days, `in-addr.arpa.` 1 day).
- Why so many queries then?

Data For This Talk

- 24 hours of tcpdump from f.root-servers.net.
- 04/Oct/2002, 00:00:00 – 23:59:59 UTC.
- 152,744,325 total queries.
- 382,708 source IP addresses.
- 220 busiest sources generate 50% of the queries.

Query Types

QTYPE	Count	Percent
A?	84,710,847	55.5
PTR?	30,462,666	19.9
AAAA?	7,213,988	4.7
MX?	7,019,561	4.6
A6?	6,900,619	4.5
SOA?	6,403,621	4.2
ANY?	4,786,327	3.1
NS?	2,636,004	1.7
SRV?	1,819,762	1.2
CNAME?	662,553	0.4
other	128,377	<0.1

Recursion Desired?

- 3,389,462 queries (2.22%) have recursion desired bit set.
- 23,945 source IP addresses (6.26%) send queries with recursion desired.
- Presumably from (stupid) stub resolvers, rather than name servers.

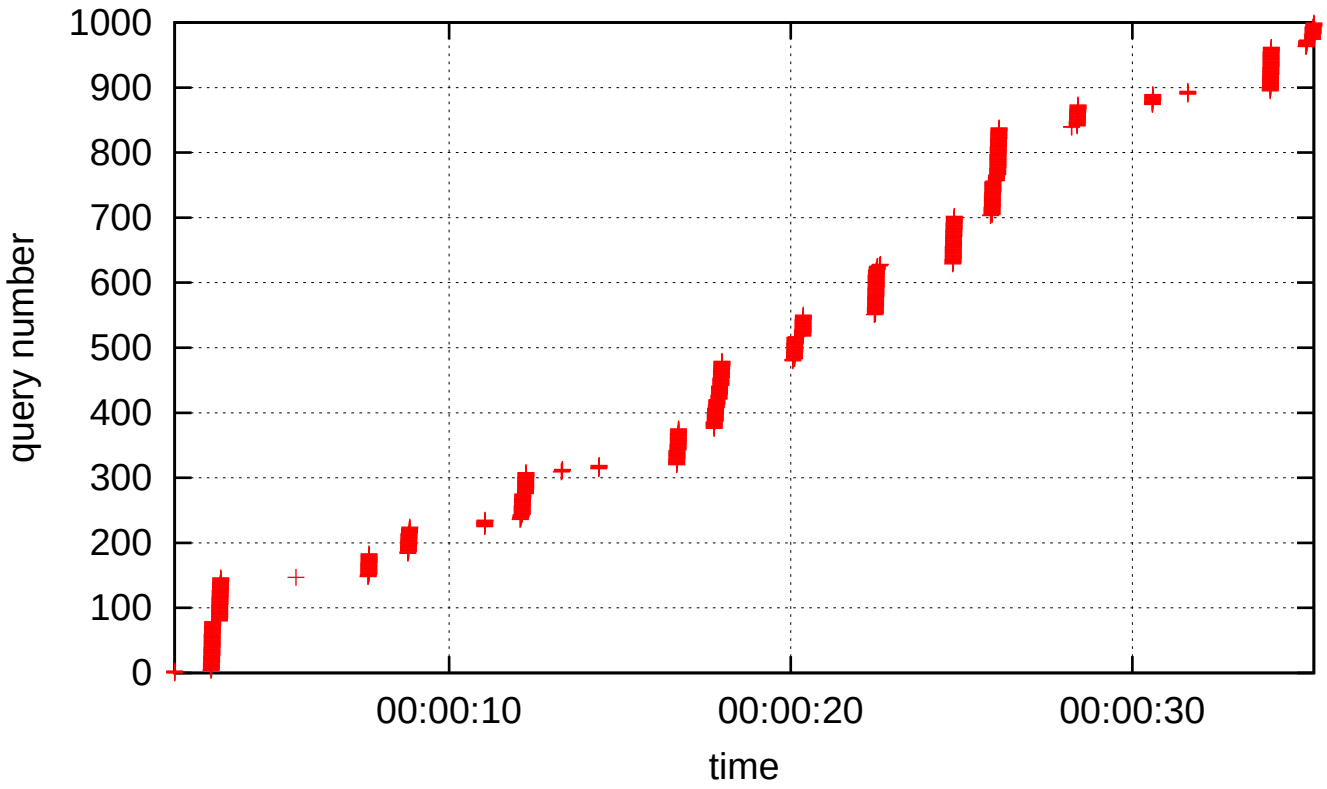
Let's Look at
Some Busy Clients

#1: Somewhere in `.mil`

- Route Views says: `Network not in table`
- 3,052,825 queries (2.00% of all queries this day)
- 2,331,857 queries (76.4%) are:

```
00:00:01.961516 160.30.209.71.1069 > f.53:  118 ANY? BURRBXR1.
00:00:01.961525 160.30.209.71.1069 > f.53:  8318 ANY? BURRBXR1.
00:00:01.961533 160.30.209.71.1069 > f.53:  6272 ANY? BURRBXR1.
00:00:01.961593 160.30.209.71.1069 > f.53:  8331 ANY? BURRBXR1.
00:00:03.027409 160.30.209.71.1069 > f.53:  10592 ANY? BURRBXR1.
```

#1: .mil: #Queries vs. Time



- Bursty on small time scale.

#2: A Name Registration Company

- 2,465,092 queries
- 10% are for a pair of names within the company's own domain.
- 58.5% are for [a-m].root-servers.net.
- 33.8% A?
  33.1% A6?
  33.1% AAAA?

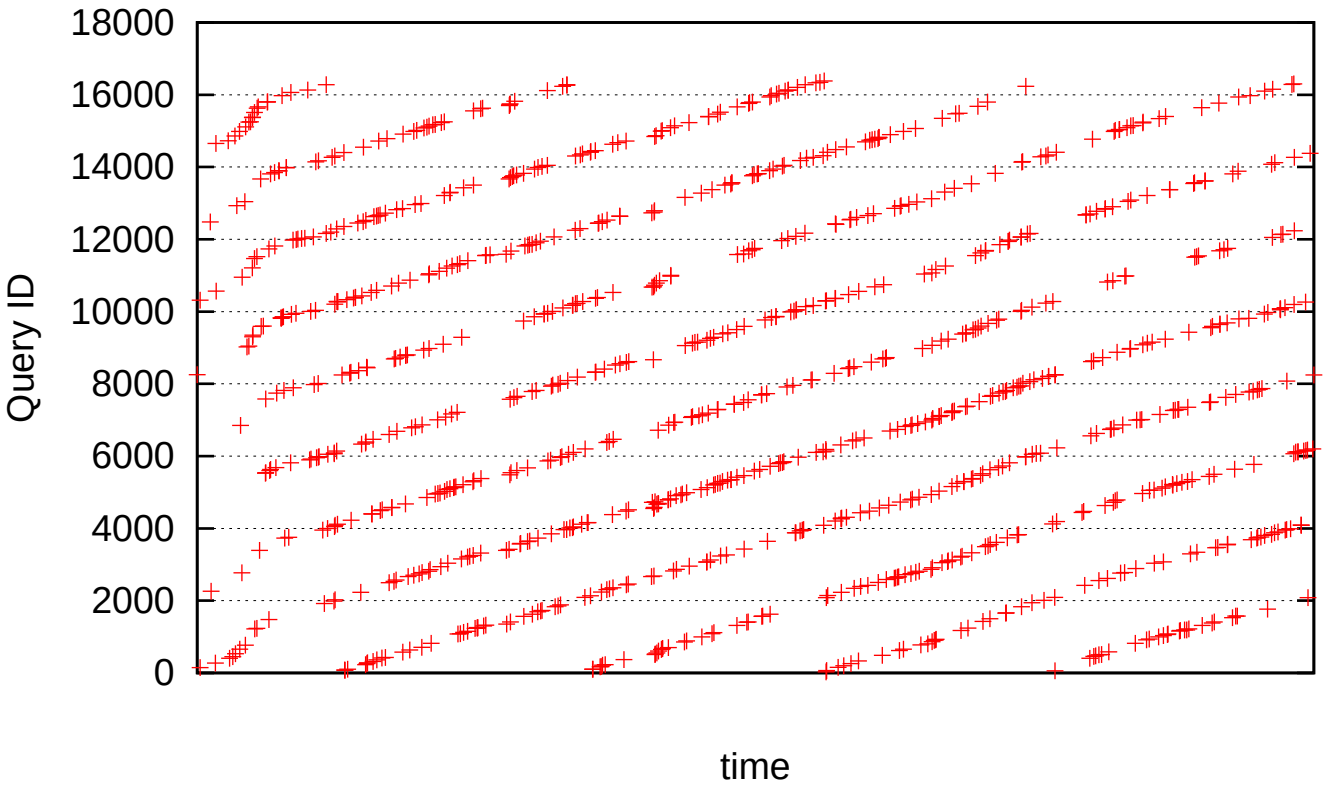
More on the Name Registration Company

- 30 source addresses in the trace resolve to this company's hosts.
- 23,300,020 total queries, 15.3% of queries this day.
- 14,761,664 (63.4%) for [a-m].root-servers.net.
- 3,390,206 (14.6%) for hosts in the company's own domain.

#3: Customer of a DSL Provider

- 2,138,697 queries.
- 2,137,830 (99.96%) are PTR queries.
- 1,894,315 (88.6%) are for 13.30.7.19.in-addr.arpa.
- Exhibits "QID $<$ 16,384 syndrome."

#6: Somewhere in C&W Land



- Anybody know what software this is?

#22: Equipment Manufacturer

- 760,106 (100%) queries for `localhost.`
- Each query repeated 8–12 times per second.
- Also exhibits QID $<$ 16,384 syndrome.

```
00:00:53.933610 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.933941 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.947688 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.947986 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.954369 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.954686 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.988930 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.989352 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.989602 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:53.990605 152.67.20.25.1111 > 192.5.5.241.53:  11648 A? localhost. (27)
00:00:55.308157 152.67.20.25.1111 > 192.5.5.241.53:  7567 A? localhost. (27)
00:00:55.308479 152.67.20.25.1111 > 192.5.5.241.53:  7567 A? localhost. (27)
00:00:55.410761 152.67.20.25.1111 > 192.5.5.241.53:  7567 A? localhost. (27)
00:00:55.411082 152.67.20.25.1111 > 192.5.5.241.53:  7567 A? localhost. (27)
```

Quite a Pair

- Two pairs of sources send exactly the same queries from different networks:

```
00:00:00.394937 108.67.19.5.53 > f.53:  23391 PTR? 175.94.9.64.in-addr.arpa.
00:00:00.910620 108.67.19.5.53 > f.53:  23410 PTR? 122.46.67.200.in-addr.arpa.
00:00:01.184493 108.67.19.5.53 > f.53:  23426 PTR? 175.94.9.64.in-addr.arpa.
00:00:02.348900 108.67.19.5.53 > f.53:  23456 A? 66.192.125.116.
00:00:03.437992 108.67.19.5.53 > f.53:  23471 A? 66.250.80.183.
00:00:04.361163 108.67.19.5.53 > f.53:  23484 A? 66.121.246.94.
00:00:04.882863 108.67.19.5.53 > f.53:  23487 A? 159.110.14.192.
00:00:05.103204 108.67.19.5.53 > f.53:  23507 SOA? 50.168.192.in-addr.arpa.
00:00:05.541302 108.67.19.5.53 > f.53:  23522 A? 222.95.109.41.
00:00:06.437053 108.67.19.5.53 > f.53:  23531 A? 66.250.80.183.

00:00:00.394928 135.220.118.5.53 > f.53:  23391 PTR? 175.94.9.64.in-addr.arpa.
00:00:00.910612 135.220.118.5.53 > f.53:  23410 PTR? 122.46.67.200.in-addr.arpa.
00:00:01.184485 135.220.118.5.53 > f.53:  23426 PTR? 175.94.9.64.in-addr.arpa.
00:00:02.348868 135.220.118.5.53 > f.53:  23456 A? 66.192.125.116.
00:00:03.437984 135.220.118.5.53 > f.53:  23471 A? 66.250.80.183.
00:00:04.361154 135.220.118.5.53 > f.53:  23484 A? 66.121.246.94.
00:00:04.882840 135.220.118.5.53 > f.53:  23487 A? 159.110.14.192.
00:00:05.103196 135.220.118.5.53 > f.53:  23507 SOA? 50.168.192.in-addr.arpa.
00:00:05.541294 135.220.118.5.53 > f.53:  23522 A? 222.95.109.41.
00:00:06.437035 135.220.118.5.53 > f.53:  23531 A? 66.250.80.183.
```

Classifying Queries

Unused Query Class

- IANA defines: IN, CH, HS, NONE, ANY.
- Others counted here as Unused Query Class
- For example:

```
04:00:30.076655 1.95.24.164.1025 > f.53:  40690 Type49275 (Class 45036)? localhost.

07:26:32.174479 171.250.57.109.53 > f.53:  31266 Type0 (Class 0)? ems.att.com
07:27:18.580880 171.250.57.109.53 > f.53:  4153 Type116 (Class 97)? ems
07:27:52.990604 171.250.57.109.53 > f.53:  36226 Type116 (Class 97)? ems
07:27:52.993551 171.250.57.109.53 > f.53:  17643 Type116 (Class 97)? ems
07:27:55.222638 171.250.57.109.53 > f.53:  29052 Type0 (Class 0)? ems.att.com
07:27:58.570414 171.250.57.109.53 > f.53:  55558 Type0 (Class 0)? ems.att.com
```

A For A

- QNAME is already an IPv4 address
- For example:

```
06:33:49.375586 236.197.47.135.32772 > f.53:  55624 [1au] A? 200.157.40.217.
06:45:38.573855 236.197.47.135.32772 > f.53:  25927 [1au] A? 207.244.8.2.
10:24:05.185055 236.197.47.135.32772 > f.53:  62470 [1au] A? 209.205.98.9.
12:19:23.049895 236.197.47.135.32772 > f.53:  60175 [1au] A? 149.131.144.151.
```

- Evi says its a buggy Microsoft implementation, and its been fixed in service pack 2 for Win2k.

Unknown TLD

- TLD should be one of the gTLDs or ccTLDs.
- For example:

```
21:58:39.674682 236.197.47.135.32772 > f.53:  54650 [1au] SOA? _ldap._tcp.training-its.
21:59:43.591884 236.197.47.135.32772 > f.53:  4213 [1au] SRV? _ldap._tcp.dc._msdcs.BRZTUIC4.
22:00:09.110046 236.197.47.135.32772 > f.53:  54376 [1au] SOA? _kpasswd._tcp.BRZTUIC3.

00:04:50.045238 237.123.109.108.53 > f.53:  52509 A? M-5M-5M-<M--M-0M-|.qq.
00:04:50.046362 237.123.109.108.53 > f.53:  56760 A? M-?M-\M-1M-9M->M-nM-1M-3M-@M-0M-:M-N.aa.

00:00:07.043692 154.88.153.27.16987 > f.53:  32425 A? www.abrawicca.hpg.combr.
00:00:18.688592 154.88.153.27.23615 > f.53:  50951 SOA? _ldap._tcp.pdc._msdcs.MASP.local.
00:00:18.728244 154.88.153.27.23632 > f.53:  17665 SRV? _ldap._tcp.dc._msdcs.ba-zxvhkkckcmtx.
00:00:19.049214 154.88.153.27.23800 > f.53:  36967 [1au] SOA? servidor.sciesp.local.
00:00:19.760810 154.88.153.27.24152 > f.53:  42099 A? pop3.informatica.mygra.com.br.lan.
00:00:26.254022 154.88.153.27.28133 > f.53:  47793 SOA? SPO-ETESMP1-W2.smp.ete.
00:00:46.147211 154.88.153.27.39843 > f.53:  3068 [1au] A? 48.17.35.162.bcocacique.
```

Non-printable Characters in QNAME

- Query name must not contain non-printable characters.
- For example:

```
18:41:05.677480 217.75.180.184.53 > f.53:  16372 A? www.launchstats.com.biz^A|.busycorp.net.

10:13:16.614163 93.242.134.180.21146 > f.53:  3222 A? M-2M-NM-$@M-5oM-2M-<.tw.

00:00:25.762806 126.218.103.79.8394 > f.53:  11446 A? M-0M-fM-@M-NM-AM-vM-?M-*M-=M-G.hyudai-m
00:05:05.692834 126.218.103.79.8394 > f.53:  14794 A? userM-!M-Z.simmani.com.co.kr.
00:05:48.244663 126.218.103.79.8394 > f.53:  7764 A? M-@M-NM-CM-5M-AM-vM-AM-!.n-top.com.
```

RFC1918 Addresses in PTR Queries

- RFC1918 addresses must not "escape" into the public Internet.
- For example:

```
07:11:08.024562 45.183.188.225.299 > f.53:  7935 PTR? 211.4.19.172.in-addr.arpa.

00:00:03.166185 208.207.185.73.53 > f.53:  58600 PTR? 2.215.25.10.in-addr.arpa.
00:00:03.195603 208.207.185.73.53 > f.53:  64053 PTR? 4.48.102.10.in-addr.arpa.
00:00:04.162770 208.207.185.73.53 > f.53:  60567 PTR? 1.215.25.10.in-addr.arpa.
00:00:04.163108 208.207.185.73.53 > f.53:  34584 PTR? 4.48.102.10.in-addr.arpa.
00:00:04.164180 208.207.185.73.53 > f.53:  56324 PTR? 1.215.25.10.in-addr.arpa.
00:00:06.109734 208.207.185.73.53 > f.53:  54124 PTR? 4.48.102.10.in-addr.arpa.
00:00:07.118211 208.207.185.73.53 > f.53:  53138 PTR? 3.20.29.10.in-addr.arpa.
```

Repeat Queries

- When Query ID, QTYPE, QCLASS, and QNAME are all the same (for a single source IP address).
- No time constraints.
- For example:

```
04:47:58.787556 30.23.54.57.13419 > f.53:  1664 A? www.equinox.ie.
04:47:58.794944 30.23.54.57.13419 > f.53:  1664 A? www.equinox.ie.
04:48:00.788963 30.23.54.57.13419 > f.53:  3725 A? www.equinox.ie.
04:48:00.795968 30.23.54.57.13419 > f.53:  3725 A? www.equinox.ie.

15:47:52.288311 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:52.295315 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:52.848003 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:52.854951 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:53.328908 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:53.348655 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:53.368517 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:53.404849 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:53.736279 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
15:47:53.749166 186.55.87.69.22316 > f.53:  6482 PTR? 7.0.0.240.200.in-addr.arpa.
```

Repeat QNAMEs

- When QTYPE, QCLASS, and QNAME are all the same (for a single source IP address).
- No time constraints.
- For example:

```
00:10:41.190734 106.221.10.67.53 > f.53:  7322 A? irc.webgiro.se.
00:10:41.290840 106.221.10.67.53 > f.53:  1188 A? irc.webgiro.se.
00:10:45.185877 106.221.10.67.53 > f.53:  7339 A? irc.webgiro.se.
00:10:45.233263 106.221.10.67.53 > f.53:  3254 A? irc.webgiro.se.
00:12:58.308428 106.221.10.67.53 > f.53:  1209 A? irc.webgiro.se.
00:12:59.304985 106.221.10.67.53 > f.53:  11456 A? irc.webgiro.se.
00:13:00.305877 106.221.10.67.53 > f.53:  1227 A? irc.webgiro.se.
00:13:02.294644 106.221.10.67.53 > f.53:  7378 A? irc.webgiro.se.
00:13:02.355269 106.221.10.67.53 > f.53:  7385 A? irc.webgiro.se.
00:13:06.391755 106.221.10.67.53 > f.53:  11488 A? irc.webgiro.se.
00:13:06.436114 106.221.10.67.53 > f.53:  7407 A? irc.webgiro.se.
```

Referral Not Cached

- Tricky!
- Assume client should have received and cached a referral NS for its query to the root.
- Require 3 second gap.
- Root servers are authoritative for some zones, like `arpa`, `in-addr.arpa`, `edu`, `mil`, `gov`, `root-servers.net`.
- For example:

```
10:22:03.625834 203.128.145.55.398 > f.53:  13952 A? urca1979.dns2go.com.
10:22:05.875210 203.128.145.55.398 > f.53:  13998 A? silkenwings.dns2go.com.
10:22:45.128622 203.128.145.55.398 > f.53:  2336 A? rarror.dns2go.com.
10:22:48.611913 203.128.145.55.398 > f.53:  12685 A? spectrum.dns2go.com.
10:23:54.827486 203.128.145.55.398 > f.53:  3482 A? overhung.dns2go.com.
10:23:56.839771 203.128.145.55.398 > f.53:  3544 A? twilight-zerver.dns2go.com.
10:24:07.838068 203.128.145.55.398 > f.53:  15901 A? nuanda.dns2go.com.
10:25:14.815513 203.128.145.55.398 > f.53:  15296 A? pippobis.dns2go.com.
10:27:24.817265 203.128.145.55.398 > f.53:  15508 A? sperrow1.dns2go.com.
10:29:54.864385 203.128.145.55.398 > f.53:  13680 A? urca1979.dns2go.com.
```

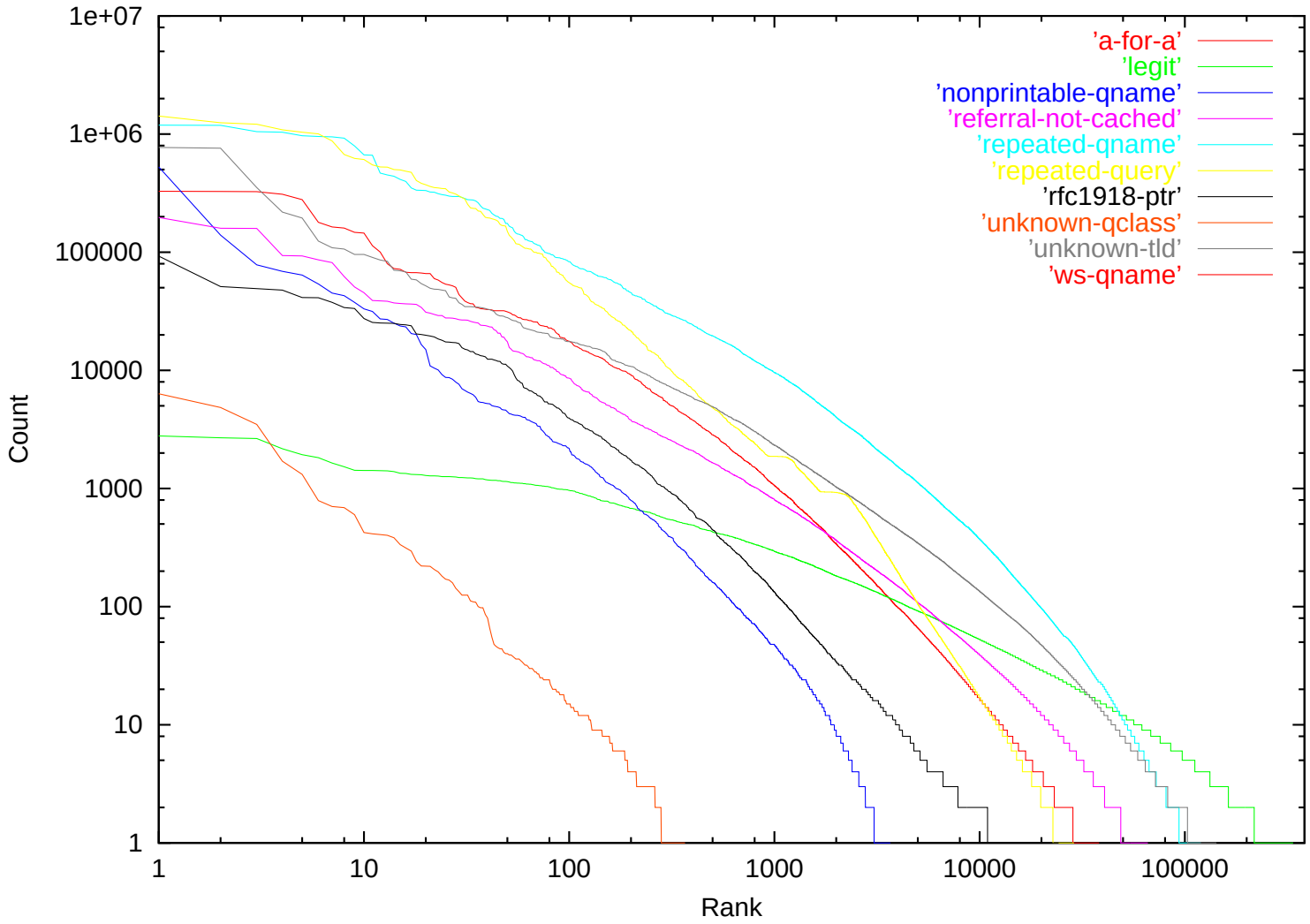
Legitimate Query

- Any query not categorized as one of the previous is assumed to be legitimate.

Query Classifications

Type	Count	Percent
Repeated QNAME	68,610,091	44.9
Repeat Query	38,838,688	25.4
Unknown TLD	19,165,840	12.5
A for A	10,739,857	7.03
Referral Not Cached	6,653,690	4.36
Legitimate	3,284,569	2.15
Nonprintable in QNAME	2,962,471	1.94
rfc1918 PTR	2,452,806	1.61
Unused Query Class	36,313	.024

Query Classifications

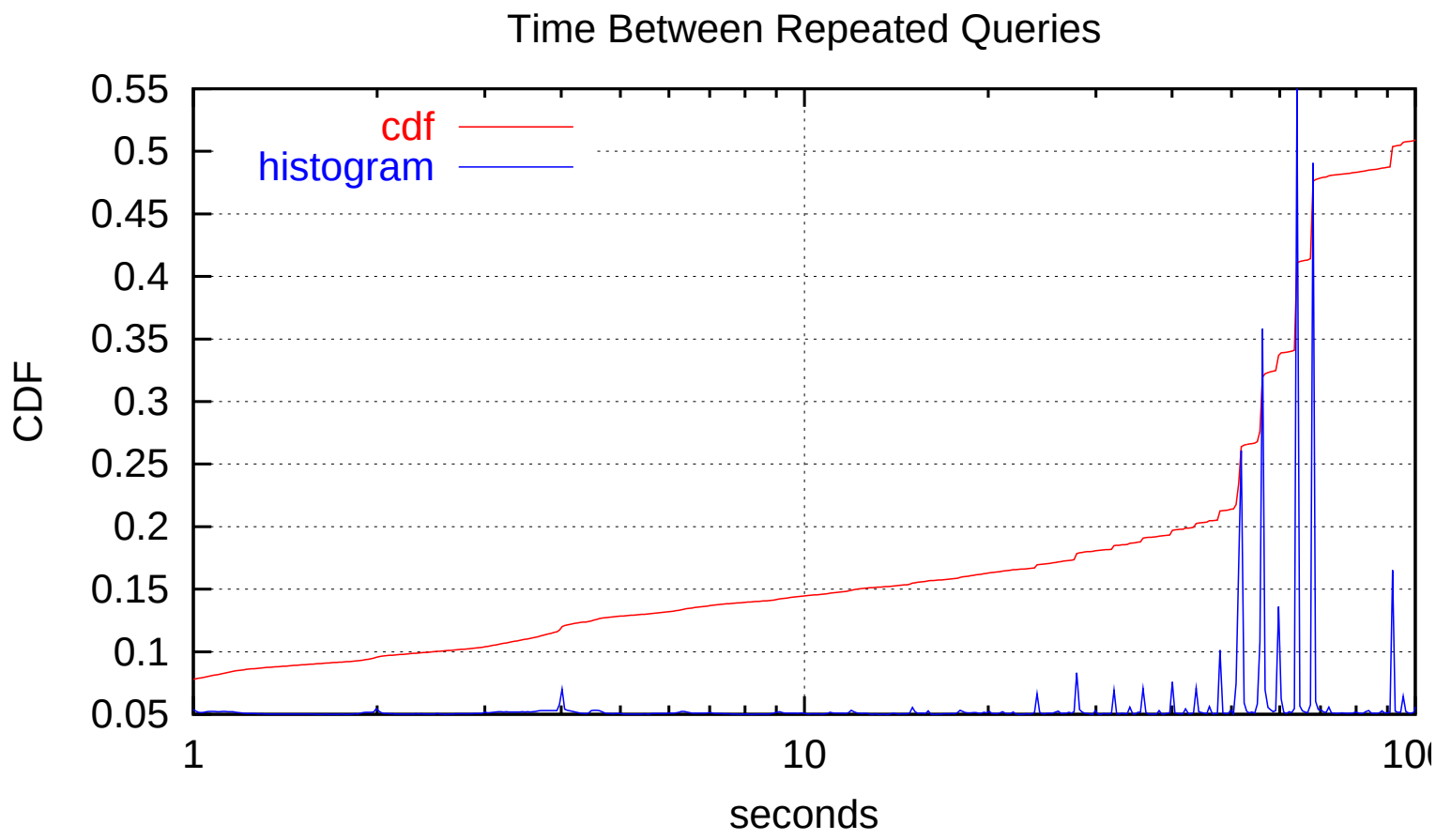


Compared to Earlier Study

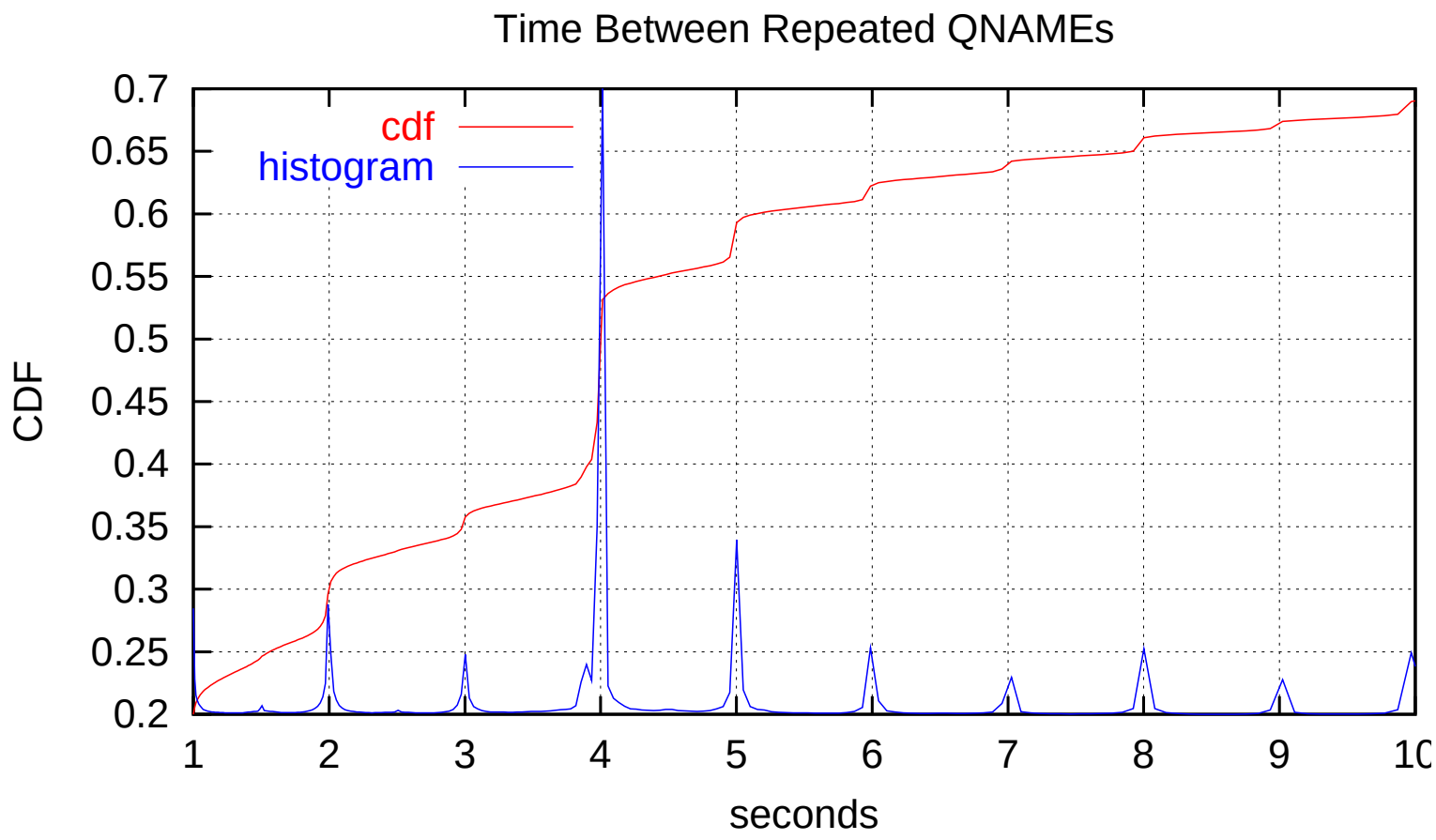
	Percentages	
Type	Jan 2001	Oct 2002
Repeated Queries	85	70.3
Unknown TLD	20	12.5
A for A	12–18	7.03
RFC1918 PTR	7	1.61

- The Jan 2001 data comes from *DNS Measurements at a Root Server*, by Nemeth, Brownlee, and Claffy, Globecom 2001.
- The longest trace from the 2001 study was 2 hours.
- 2001 study also had RFC1918 source addresses (0.7% of queries).

Repeat Query Interarrival Times



Repeat QNAME Interarrival Times



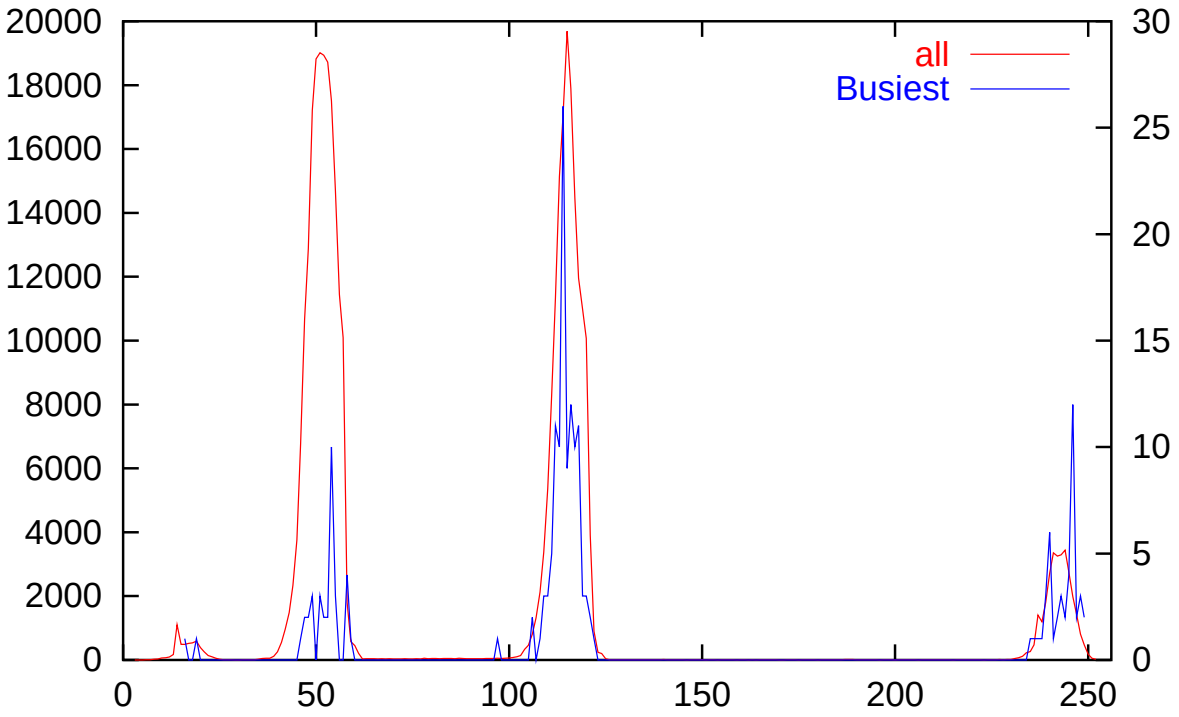
Popular QNAMEs (approx counts)

QNAME	Count	QNAME	Count
.	2,939,011	wpad	644,717
burrbxr1	2,331,857	auto.search.msn.com	637,558
localhost	1,950,123	25.0/26.38.96.12.in-addr.arpa	559,341
11.34.8.12.in-addr.arpa	1,894,427	www.opasoft.com	526,935
b.root-servers.net	1,617,809	in-addr.arpa	494,300
d.root-servers.net	1,602,463	mxmail.register.com	423,714
i.root-servers.net	1,598,830	7.too.co.il	422,627
a.root-servers.net	1,585,663	1.187.193.208.in-addr.arpa	372,789
f.root-servers.net	1,582,860	11.218.183.194.in-addr.arpa	370,533
c.root-servers.net	1,579,265	7.160.39.12.in-addr.arpa	367,106
h.root-servers.net	1,578,517	www.math.uwaterloo.ca	365,221
e.root-servers.net	1,577,431	tgp-gfn.trulyglobal.com	354,207
g.root-servers.net	1,573,570	21.9.128.in-addr.arpa	321,712
l.root-servers.net	1,556,753	6.240.34.158.in-addr.arpa	304,241
m.root-servers.net	1,548,165	130.128/26.227.218.63.in-addr.arpa	295,089
k.root-servers.net	1,534,249	32.0.54.34.in-addr.arpa	293,653
j.root-servers.net	1,526,195	d11-c5.data-hotel.net	273,526
philorch.com	1,186,696	ms51.hinet.net.tw	266,984
⟨*nonprintable*⟩.tw	987,504	ems.att.com	261,628
⟨*nonprintable*⟩.2ndpower.com	665,414	edu	244,426

Unknown TLDs

TLD	Rank	Count	Percent	#Sources
burrbxr1	6	2,331,857	1.53	1
local	7	2,001,210	1.31	19550
localhost	9	1,962,413	1.28	12298
wpad	16	651,230	.426	6869
eder003	33	218,210	.143	1
domain	41	162,682	.107	3716
lan	43	142,791	.093	1988
workgroup	47	121,547	.080	6079
5⟨C7⟩⟨D0⟩⟨B3⟩⟨E2⟩	50	110,880	.073	25
elvis	51	106,654	.070	47
admin	59	94,482	.062	740
ns1	68	71,556	.047	1485
msft	86	56,652	.037	1269
kornet	87	54,564	.036	2391
corp	92	51,776	.034	1742
loc	96	51,186	.034	1207
mailhost	97	50,698	.033	511
rcnet	109	48,196	.032	3
server	110	48,075	.031	4858

Some OS Fingerprinting



- Linux/BSD=64, Windows=128, Solaris=255
- Busiest corresponds to the busiest sources generating 50% of all queries, excluding 28 *Registration Company* addresses.

Some OS Fingerprinting

		Percent of Sources	
OS	TTLs	All	Busiest
BSD/Linux	35–64	49	17
Windows	100–128	40	58
Solaris/?	227–255	7.7	23

Some Thoughts

What's Causing This?

- Firewalls and packet filters.
- Incorrect name server configuration.
  - Use of RFC1918 address space, but no zone files.
- Aggressive retransmission.
  - Developers don't understand or follow the protocol.
- Other implementation bugs.

About Retransmission

- Originally hoped that most repeated queries could be eliminated by fixing retransmission bugs.
- However, looks like most repeat queries actually have acceptable delays.
- Number of queries repeated less than 2 seconds: 24,311,541 (15.9%).
- Increasing the delay wouldn't necessarily eliminate these repeats.

Possible Solutions

- When a name server doesn't get any answers from another server, it could:
  - rate-limit itself
  - complain to syslog
  - exit
- Name servers could refuse to forward "bad" queries
  - Unknown TLD
  - Non-ASCII characters in QNAME
  - Unused query types and classes

Possible Solutions, cont'd

- Install RFC1918 in-addr.arpa zones by default.
- Better QA testing.
- Contact the abusers.
  - Registries have contact information, but large organizations often do not reply.
- Post weekly root server abuse reports, ala CIDR Report.

You Too Can Play

dnstop

- A *libpcap* application, ala *tcpdump*
- Displays (via curses) various tables of DNS statistics
  - Source IP addresses
  - Destination IP addresses
  - Query types
  - Top level domains
  - Second level domains
  - more later...?
- Can replay a *tcpdump* "savefile."
- Get it at `http://dnstop.measurement-factory.com`

Thanks!

Thanks to the Root Server Operators Working With CAIDA

- Internet Software Consortium
- VeriSign Global Registry Services
- NASA Ames Research Center
- Autonomica
- Reseaux IP Europeens
- WIDE Project
-
-
-
-
-
-
-

Thanks to the Organizations Who Replied to Our Emails

- Flextronics International, Inc.
- That One Name Registration Company
-

Thanks to Funding Organizations

- WIDE (Jun Murai)

More Info

- CAIDA
  `http://www.caida.org/projects/dns-analysis/`
- dnstop
  `http://dnstop.measurement-factory.com/`
- Rob's DNS Data Page
  `http://www.cymru.com/DNS/`
- Root Server Technical Operations Assn
  `http://www.root-servers.org/`

The End