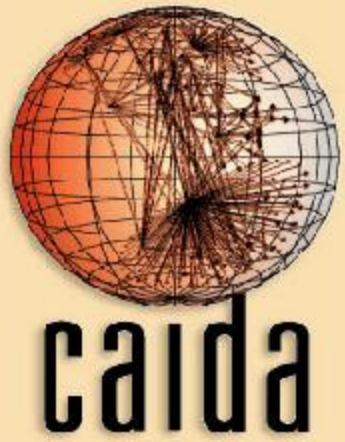
measurement and analyses activities relevant to NMS researchers



there're two possible outcomes:
if the result confirms the hypothesis,
then you've made a measurement.
if the result is contrary,
then you've made a discovery.
--fermi

ISMA workshop on routing & topology: oct.02

http://www.caida.org/outreach/isma/0210/ISMAagenda.xml

- DNS performance analysis by cc/TLD
- invariance of Internet RTT spectrum (surprising)
- evolution of bgp system at AS, prefix, and IP granularities
  - introduce semi-global prefix and related taxonomy
  - dispel myths of BGP growth and churn (rates and contributors)
- route dampening considered harmful
  - using BGP beacon (unused prefix announced and withdrawn at well-known times)
- lots about topology/policy/event inferences from BGP data
  - Patrignani, Gao, Schulzrinne, Nicol, Feamster
  - (Jean: Feamster's prefix clustering work)
- bgp convergence and scalability (anja&olaf)
- UDP/TCP performance during BGP update activity (avi)
- comparison of routeviews and ripe data for bgp analysis (tudor)
  - data sets seem congruent
- aggregated workload data behaves differently (globalcrossing)

joint NANOG/ARIN meeting -- oct 2002

http://www.nanog.org/

- route dampening considered harmful (again)
- damage in the DNS system (caida, more on this later)
- bgp behavior under stress (lixia et al)
  - http://www.cs.ucla.edu/~lanw/paper/imw02-bgp.ps
- bgp enhancements to prevent persistent route oscillations
  - http://www.nanog.org/mtg-0210/ppt/sue.pdf
- feds want to secure cyberspace (office of cyberspace security)
  - http://www.whitehouse.gov/pcipb/
  - `how to own the Internet' (again)
- ipv6 is really coming[here]
- complexity/robustness spiral (doyle-inspired) panel
  - http://www.maoz.com/~dmm/NANOG26/complexity_panel
  - considerable confusion over what complexity was being discussed
- measurement panel (not really)
- scriptroute: 'public measurement facility' proposal
  - www.scriptroute.org

joint NANOG/ARIN meeting -- oct 2002 (2)

bgp enhancements to prevent persistent route oscillations

- major factors for route oscillation:
  - dependency of IBGP updates
    - **sometimes circular dependency**
  - partial information by RR or confederation
    - **withdraws (over reduction) amplifies the issue**
  - partial order (due to MED) in route selection
- suggestions: modify route reflection spec, allow advertisement of multiple paths

issues:

- non-deterministic routing can be demonstrated within full mesh topology
- MEDs prevent reasoning about routing policies
  - though existing references encourage use of MEDs to influence inbound policies
- how much are we facing limitations of distance vector protocol and inherent limitations of same
- can't tell what tie-breakers cause route selection..

characterizing how traffic streams aggregate

dragonflies & tortoises

- method of measuring the size and lifetime of Internet streams
  - independent dimensions: mice..elephants; dragonflies..tortoises
  - two example sites
- most streams (about 45%) are dragonflies, lasting less than 2 seconds, 98% less than 15 minutes
  - tortoise: stream > 15 min
- significant number have lifetimes of hours to days, and can carry a high proportion (50% to 60%) of the total bytes on a given link
- distribution of stream sizes matters
  - any forwarding cache mechanisms in Internet routers must be able to cope with a high volume of short streams.
  - long-running (LR) streams can contribute a significant fraction of their packet and byte volumes -- shedding doubt on using traditional 'flat rate user bandwidth consumption' approaches to provisioning and engineering.

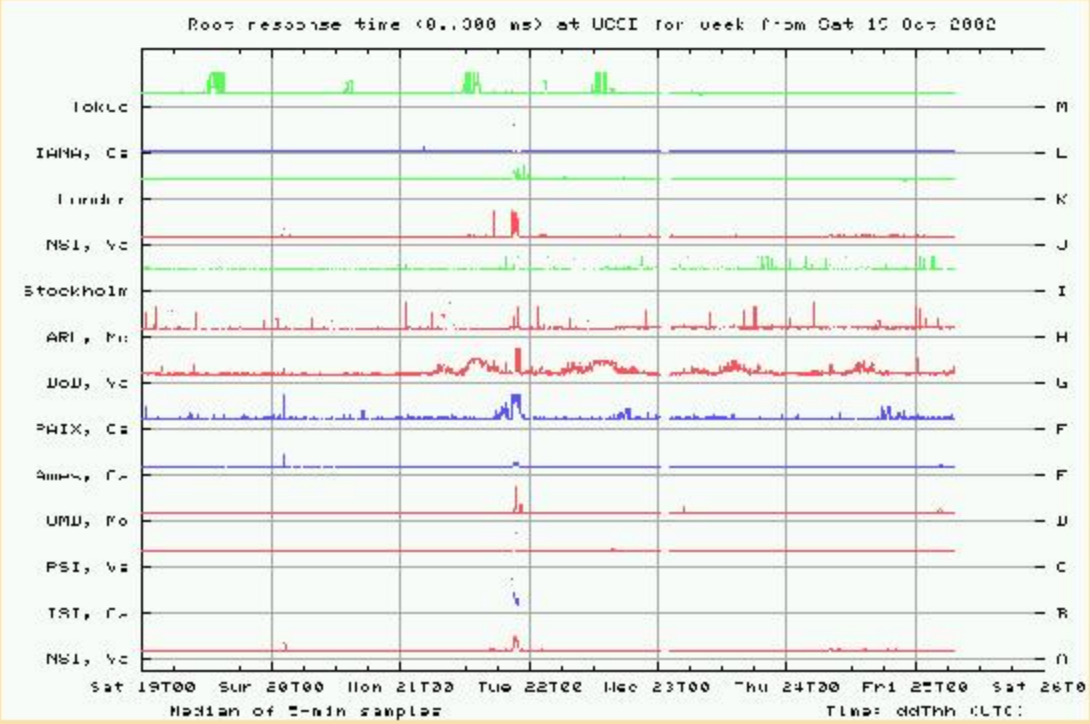
http://www.caida.org/outreach/papers/2002/Dragonflies/

(nevil@caida.org)

highlight: damage in the DNS system

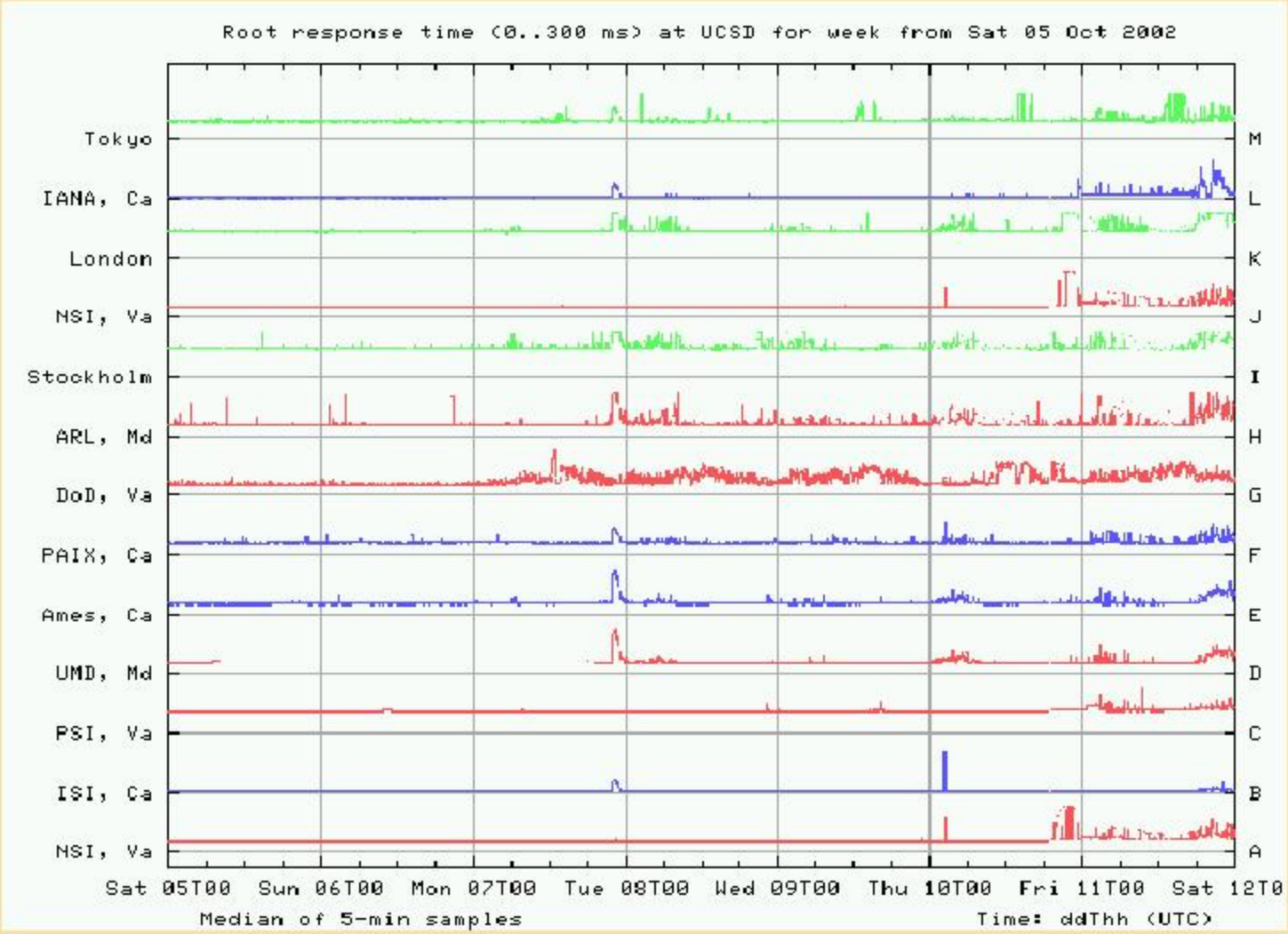
continuous monitoring of the DNS root servers performance

- root server attacks last month were 'practiced'
- as johnh said, it happens all the time



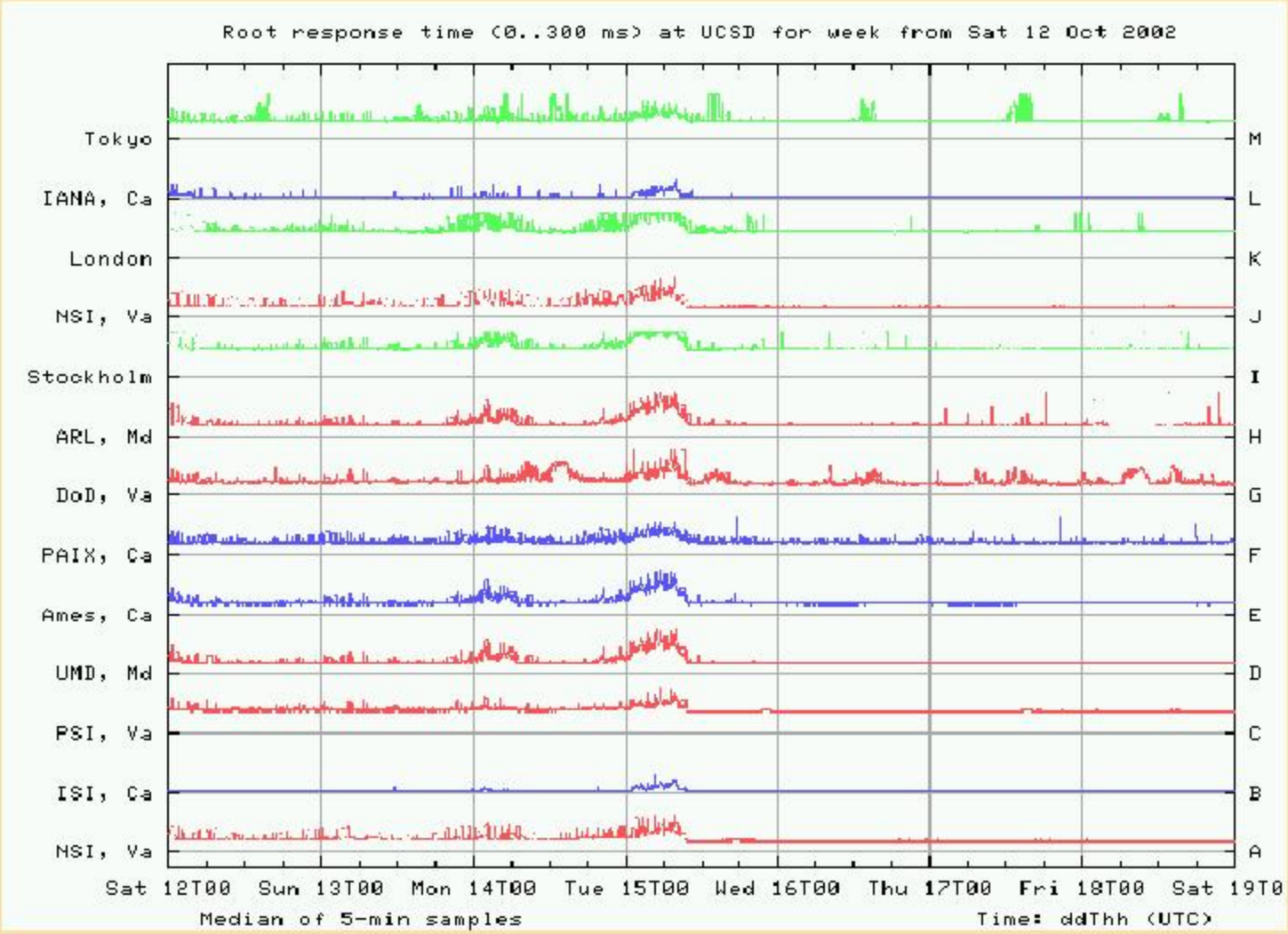
highlight: damage in the DNS system

the week of the oct 2002 attack (rtt to roots)



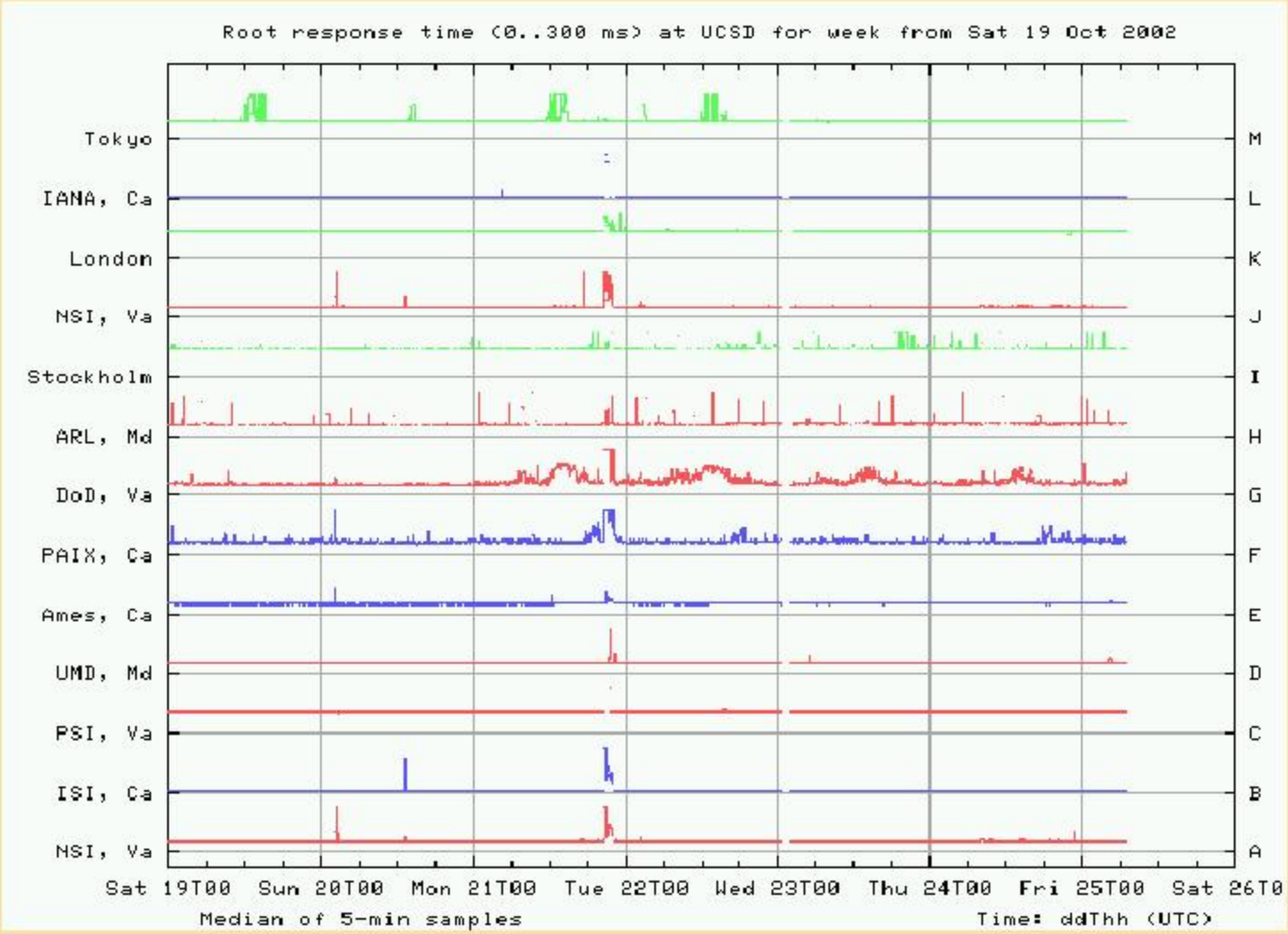
highlight: damage in the DNS system

2 weeks before (rtt to roots)



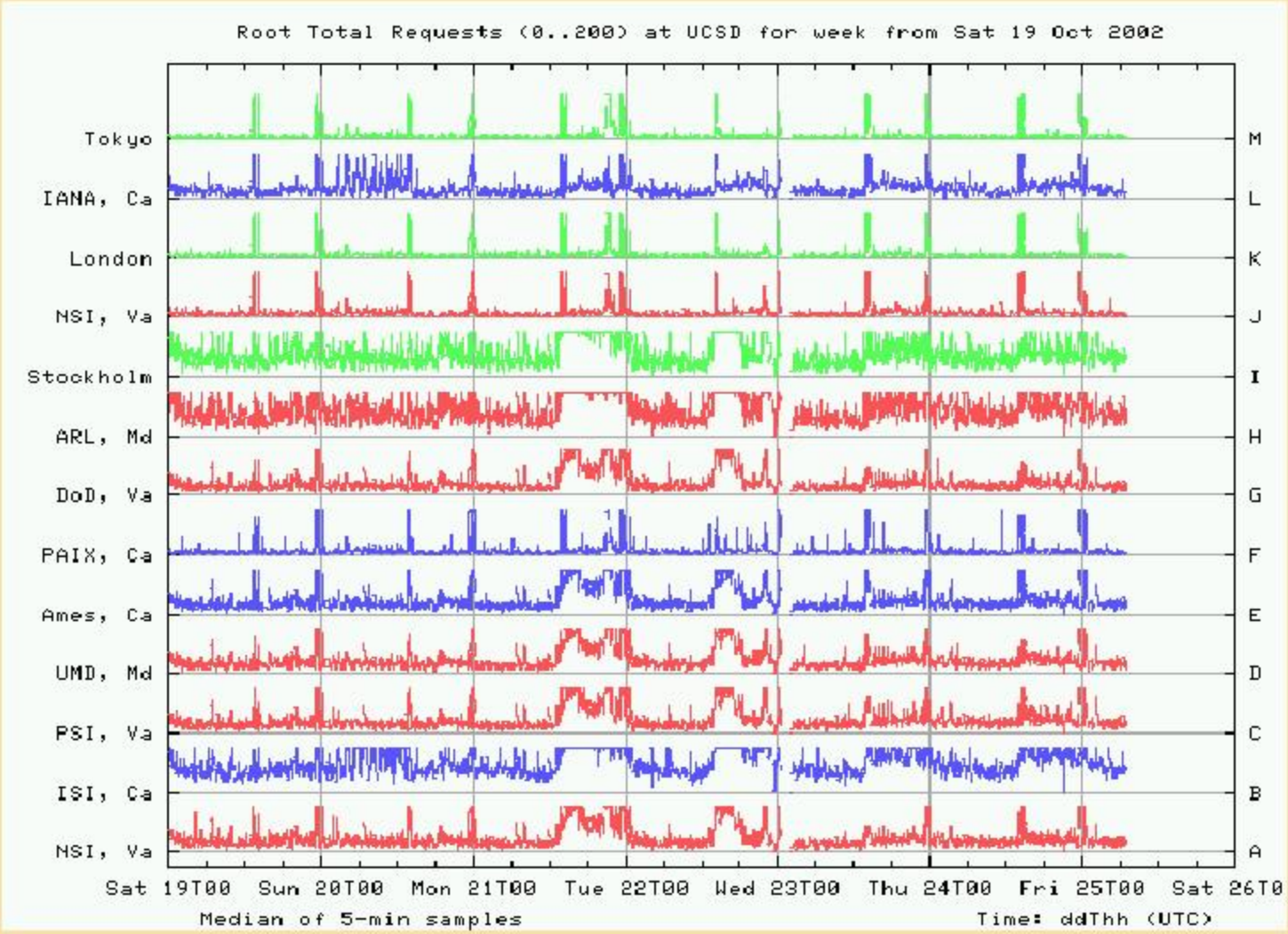
highlight: damage in the DNS system

again: the week of the attack (rtt to roots)



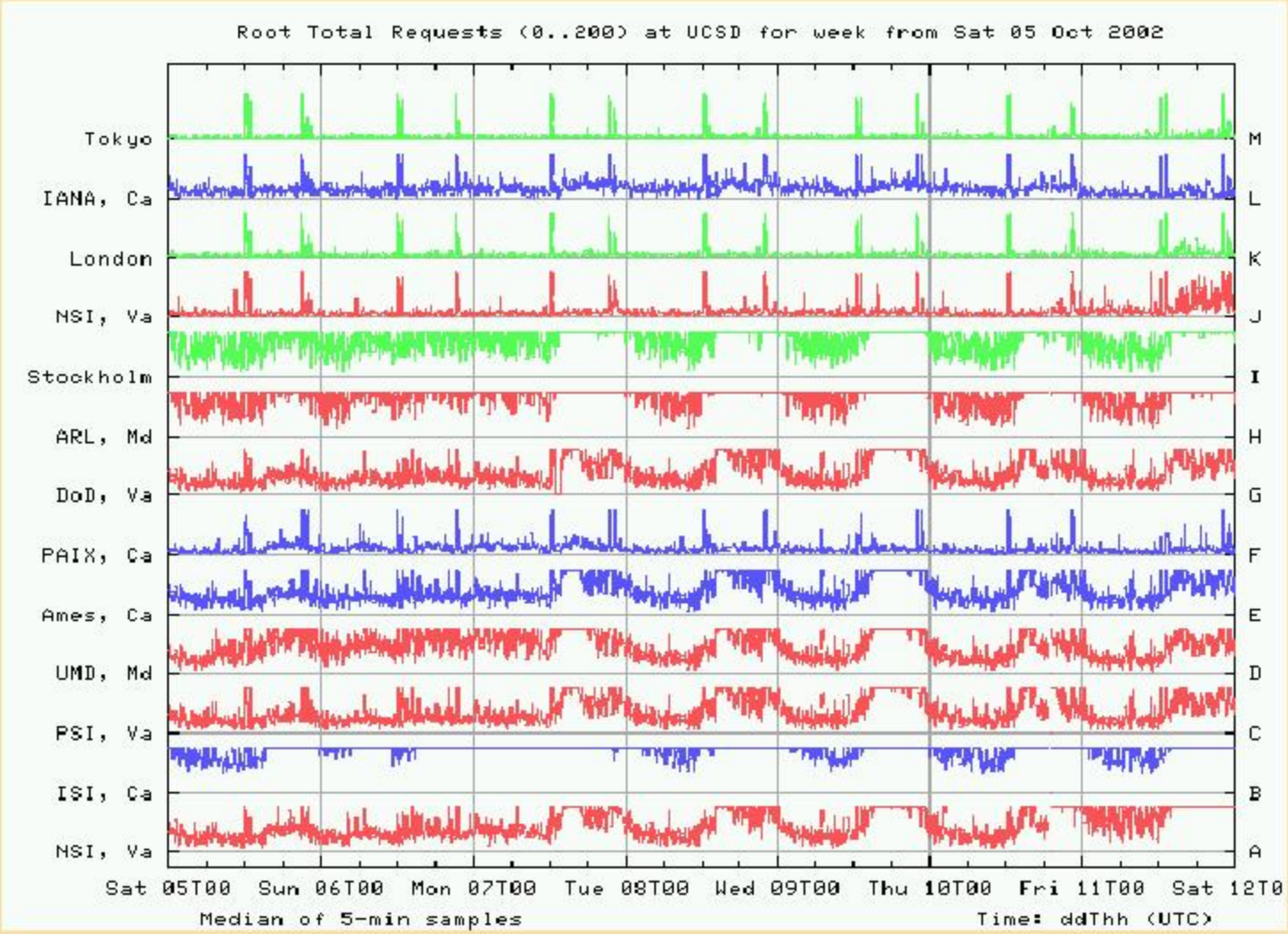
highlight: damage in the DNS system

the week of (# requests to roots)



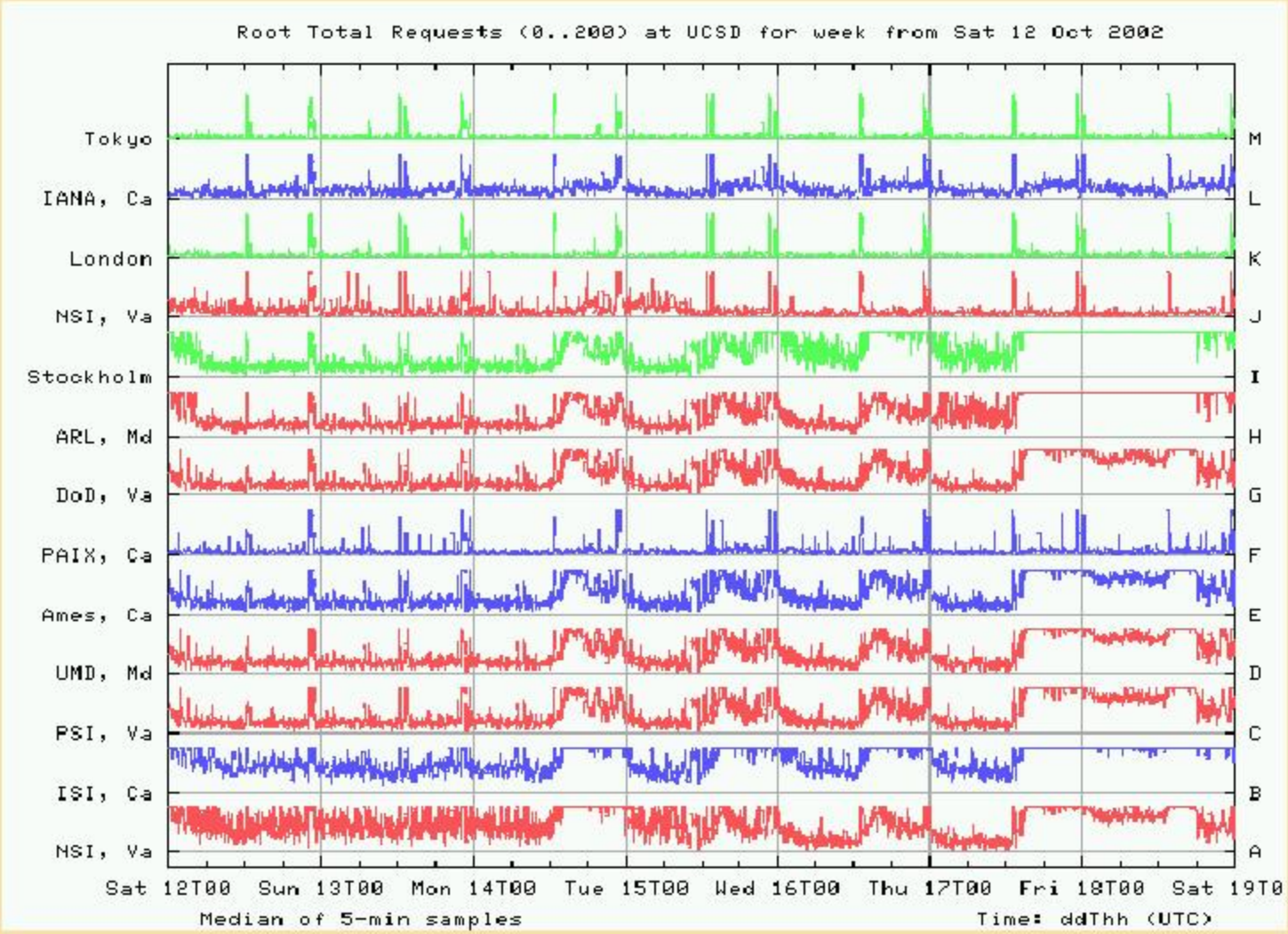
highlight: damage in the DNS system

2 weeks before (requests to roots)



highlight: damage in the DNS system

the week before (requests to roots)



- http://www.caida.org/projects/dns-analysis/

highlight: damage in the DNS system

if you think that's bad, you should see how bad baseline is

- analysis of 24 hours of traffic at f-root
  - http://www.nanog.org/mtg-0210/ppt/duane.pdf

only about 2% of the observed traffic to a root server appears to be "legitimate" queries!

part of larger CAIDA DNS research efforts (another talk)

- http://www.caida.org/projects/dns-analysis/
  - continuous monitoring of the DNS root servers performance
  - analysis of bogus queries and broken resolver configurations
  - investigation and modeling of BIND algorithm behavior
  - evaluation and optimization of root servers' placement
- tools for measuring (you can play too!)
  - dnstop: www.caida.org/tools/measurement/dnstop/
  - dnstat: www.caida.org/tools/measurement/dnstat/

highlight: damage in the DNS system

it gets worse... millions of illegitimate (spec-violating) updates per day to DNS root system

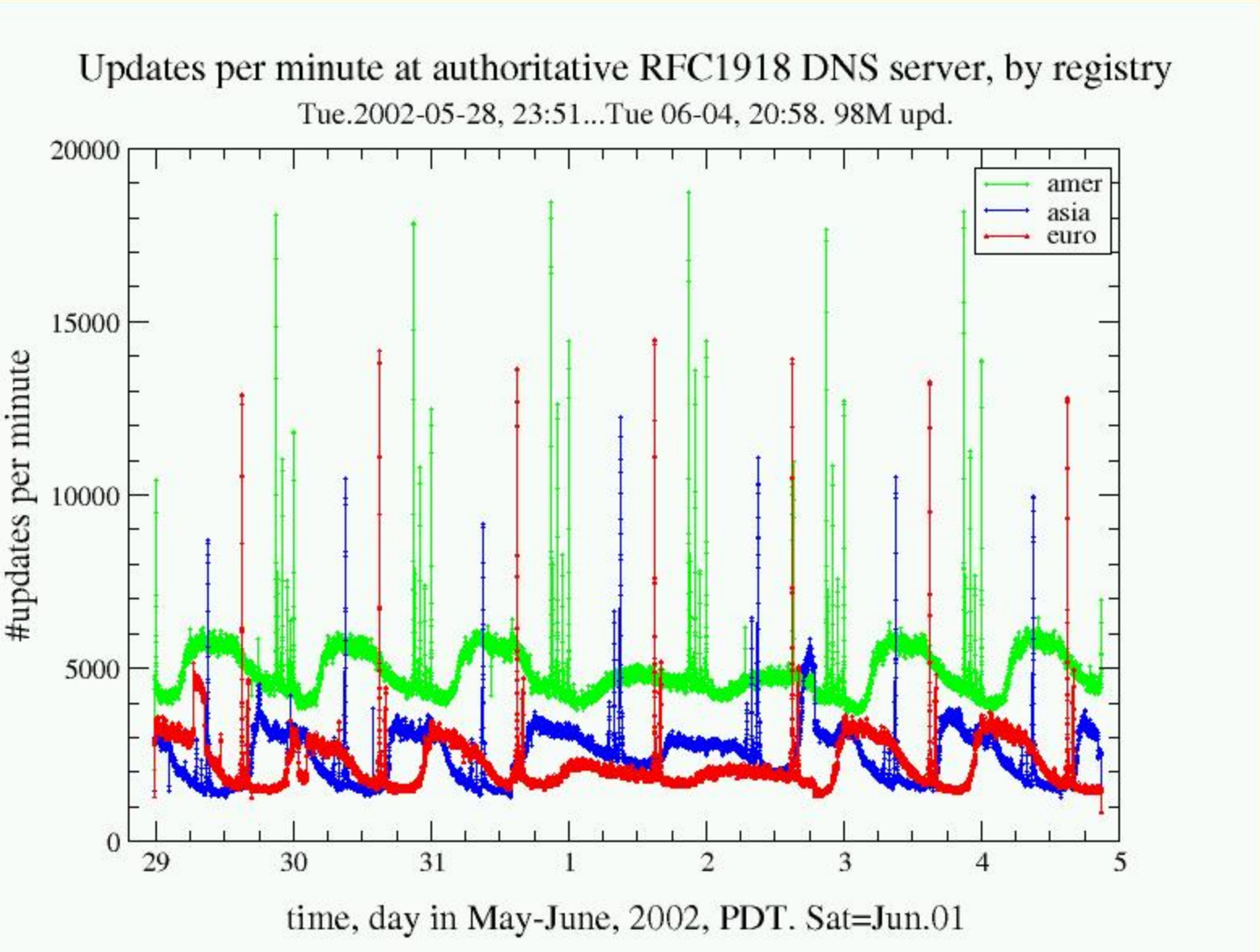
- spectroscopy analysis of RFC1918 updates
  - caida paper submitted to sigmetrics2003
- dynamic DHCP deployed since 1996
- RFC1918 updates coming from DHCP/nameservers
- should not leak outside local site
- millions a day getting to root name servers (whee)

--> AS112 project anycast to offload roots

- anycast servers dedicated to dealing with updates (tossing)
- experimenting at F-root first, expanding to others slowly
- logfiles allows us to analyze root causes (npi)

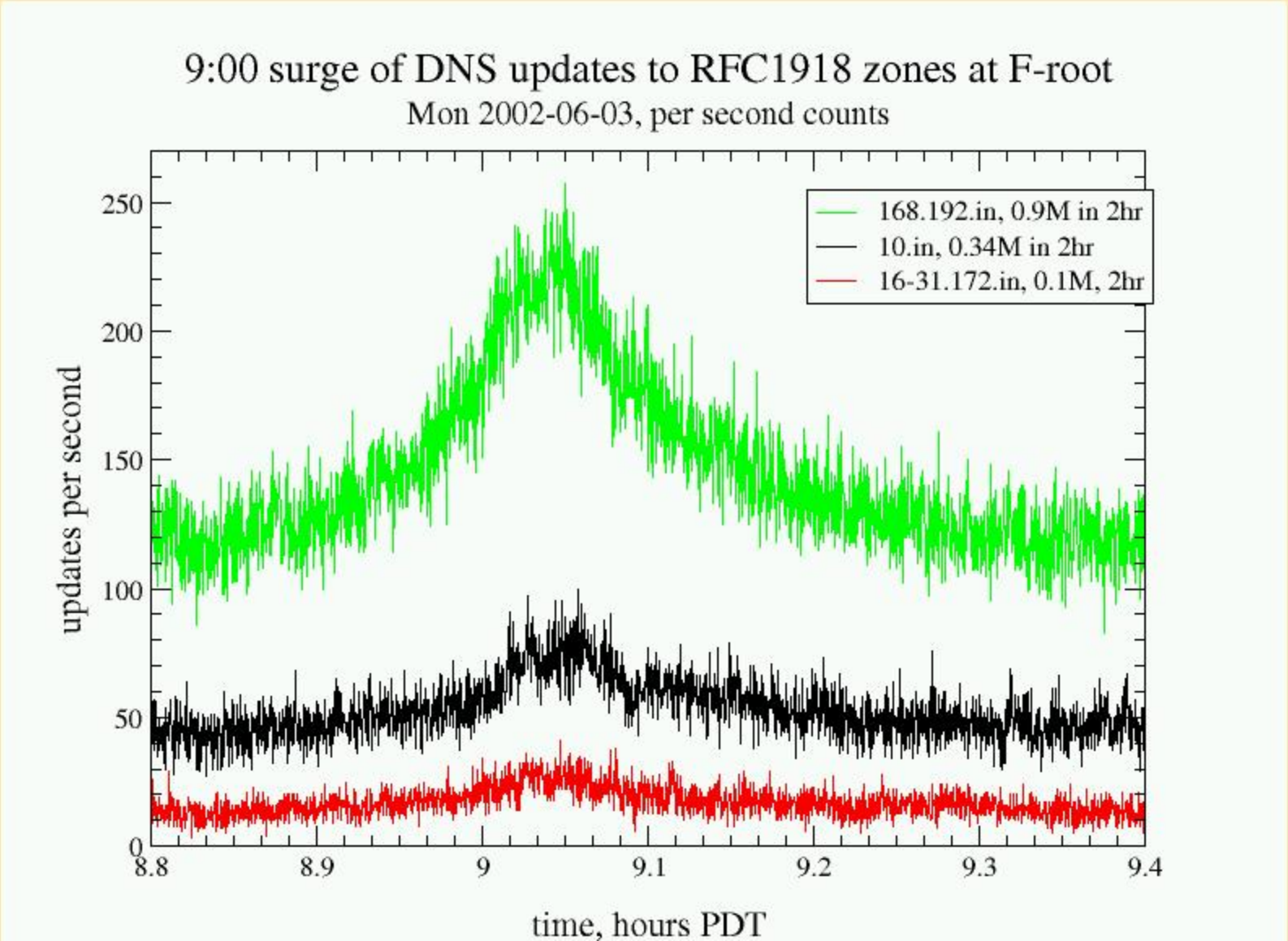
highlight: RFC1918 damage in the DNS system

a week of update counts



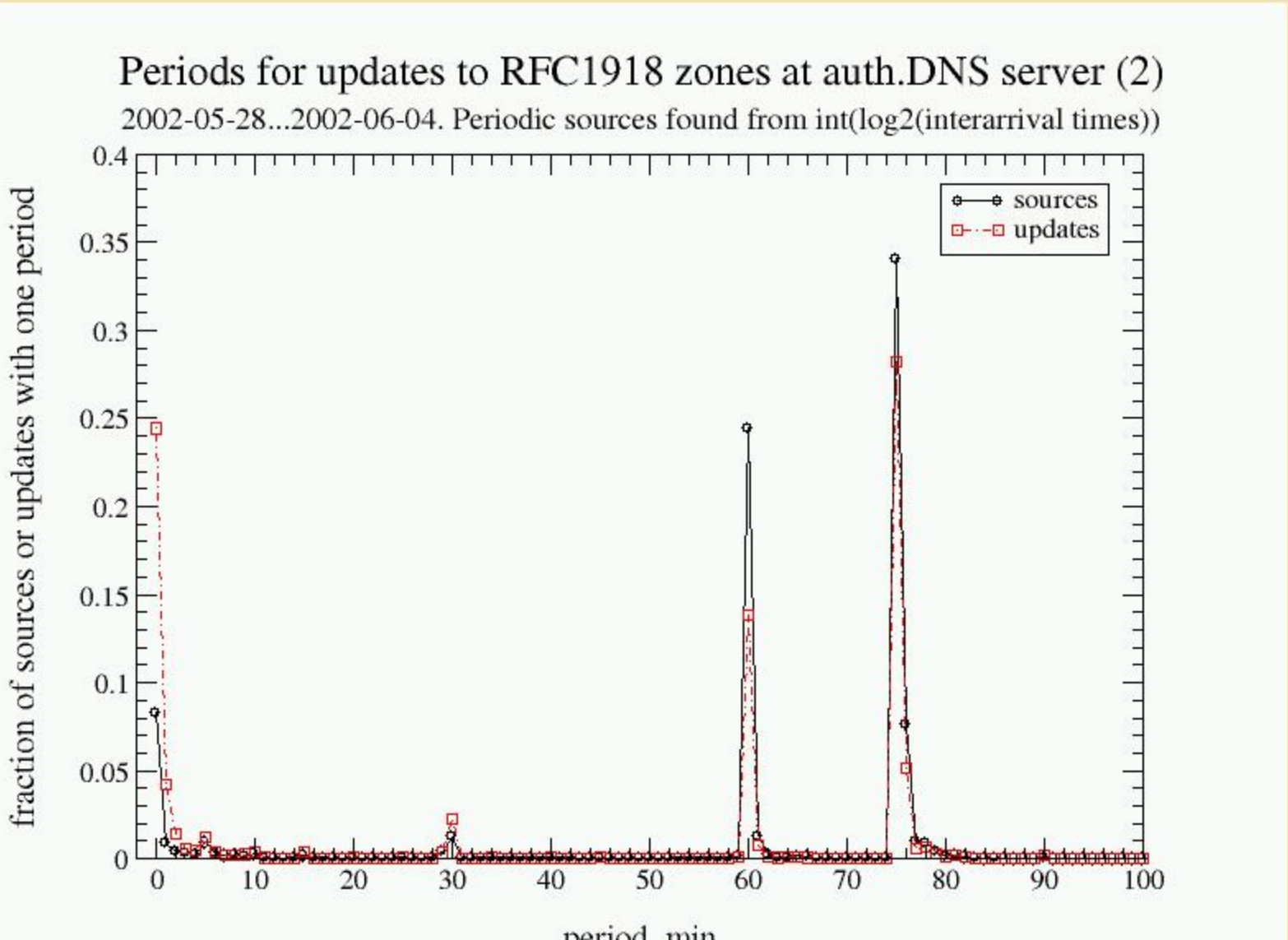
highlight: RFC1918 damage in the DNS system

a spike in detail



highlight: RFC1918 damage in the DNS system

periodicity



highlight: RFC1918 damage in the DNS system

who is trying to update the roots anyway?

- dsl, cablemodem, small population providers, developing countries
  - (hint: guess which operating system is pervasively pirated in china/hongkong)

Top 20 AS sources of RFC1918 updates

AS#	#Updates	%	Cumul.%	AS Name, Country
4134	7329178	7.51	7.51	CHINALINK, China
3352	6166266	6.32	13.84	Ibernet (TDE), Spain
7132	4559748	4.67	18.51	SW Bell, US
5673	3271669	3.35	21.86	Pac Bell, US
5676	2936073	3.01	24.87	Pac Bell, US
4813	2765227	2.83	27.71	China Telecom (Guandong)
4812	2644362	2.71	30.42	China Telecom (Shanghai)
852	2176242	2.23	32.65	Telus, Canada
6128	2083593	2.14	34.79	Cablevision, US
2828	1855065	1.90	36.69	XO, US
11427	1753091	1.80	38.49	Road Runner, US
7843	1504131	1.54	40.03	Adelphia, US
4760	1413921	1.45	41.48	Netvigator, Hong Kong
2914	1393102	1.43	42.90	Verio, US
1221	1378306	1.41	44.32	Telstra, AU
11509	1226816	1.26	45.58	Pajo, US
4436	1142608	1.17	46.75	SantaCruz Community I't, US
11426	1135058	1.16	47.91	Road Runner, US
10994	1129898	1.16	49.07	Time Warner, US
2548	1091393	1.12	50.19	Business Internet, US

highlight: RFC1918 damage in the DNS system

can we fingerprint the OSes causing this global behavior?

- updates are periodic
- periods are 60 minutes or 75 minutes
- 75 minute one is 5+10+60
  - try an update, get refused back
  - wait 5 minutes
  - try an update, get refused back
  - wait 10 minutes
  - try an update, get refused back
  - wait 60 minutes
  - repeat forever and ever

highlight: RFC1918 damage in the DNS system

can we really categorically blame microsoft?

- (hint: you bet your bootp)

- make sure all empirical data is consistent w hypothesis
- set up experimental setup of raw windows installations in our lab
- check documentation (<-- radical approach)

verified that vast majority derive from two OSes: Windows 2000 and Windows XP (empirically, experimentally, and based on documentation)

- majority of updates from sources that send them constantly. bulk of workload from contributions of medium size (not mice/elephants)
- most source IP addresses are those of home and small business users connected to the Internet via cable, DSL or phone-based ISPs, i.e., computers owned by individuals, not organizations. majority using software with default vendor settings.
- academic, corporate, backbone networks contribute little rfc1918 update traffic

highlight: RFC1918 damage in the DNS system

- combination of Microsoft software `features' and misconfigurations essentially causing a slowly paced massive distributed denial of service (DDOS) attack on the root name server system
  - everyone running windows is living in a glass house wrt DOS attacks, in more ways than one

- compelling example of why software and setups affecting stability of the Internet's infrastructure must be designed with more careful attention to potential effects of software engineering decisions, misimplementations, and misconfigurations on global systemic Internet stability

- current state of fielded desktop software poses substantial & increasing burden on, if not threat to, the robustness of the global Internet

not to be pessimistic, but

still the case: security, performance, configuration,
and fault management lack both effective solutions,
as well as an apparent lack of people able to state
a concise problem to be solved.

darpa nms researchers do have a valuable role to play
(but we have to get out of our labs)

measurement needs (1)

- workload
  - longer traces
    - tcp/ip headers, w options, other headers if possible
    - at least 24 hours
    - concurrent at several places
    - several across time for trend analysis
  - traffic matrices
    - same constraints as above
  - locations
    - disa
    - enterprise
    - university
    - backbone isp
    - peering point
    - cdn's (akamai)
- wireless
    - headers, location, signal strength
    - needs further requirements analysis

measurement needs (2)

- active measurements
  - [how to do] macroscopic monitoring of DNS, BGP system
  - correlated with passive (planned experiments)
- bandwidth estimation (not funded by NMS)
  - single source vs two end point
  - available bw vs capacity
- routing
  - BGP updates (requires more funding for equipment/maintenance)
  - IGP updates (requires contact/release from ISPs)
  - parameters from ISPs, e.g MRAI
  - implementation details from vendors

causes of Internet perf. problems/outages

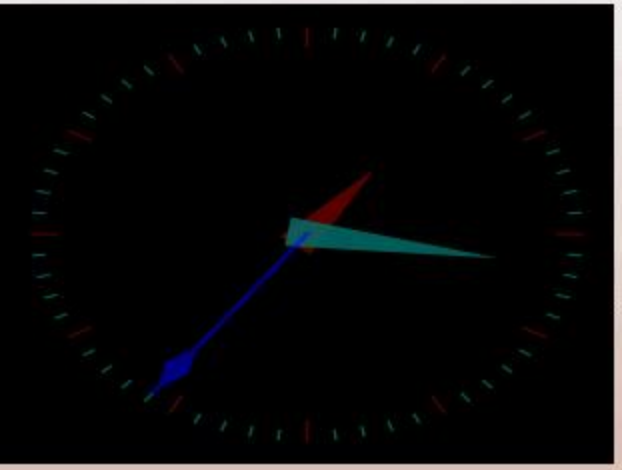
in roughly the order (courtesy sean donelan nanog post)

1. *network engineers (what's this command do?)*
2. *power failures (what's this switch do?)*
3. *cable cuts (backhoes, enough said)*
4. *hardware failures (what's that smell?)*
5. *congestion (more bandwidth! Captain, I'm giving you all she's got!)*
6. *attacks (malicious, you know who you are)*
7. *software bugs (your call is very important to us....)*

// ``I prefer the wicked rather than the foolish.
// the wicked sometimes rest." - alexandre dumas

problems that remain persistently insolvable should always be suspected as questions asked in the wrong way.

-- alan watts



www.caida.org/outreach/presentations/

kc claffy
UCSD/SDSC/CAIDA
kc@caida.org
www.caida.org