RFC1918 updates on servers near M and F roots

Andre Broido, work in progress

C A I D A

CAIDA / SDSC / UCSD

http://www.caida.org

CAIDA–WIDE Workshop

ISI, 2005-03-12

Previous projects

- IPv4 list (Young, Brad)
- Routing table growth (Evi Nemeth)
- BGP atoms (Patrick Verkaik)
- P2P traffic (Thomas Karagiannis)
- Spectroscopy
  - DSL/cable identification (Ryan King)
  - Remote device fingeprinting (Yoshi Kohno)
  - Router ICMP generation delays (Young)
  - OS fingerprinting by DNS updates (Evi)

Plan

Background
Routing changes
Microsoft sources
Conclusion

Two main questions

Is anycast stable against routing changes?

Are Microsoft boxes the largest update source?

History

- 1996: RFC1918 reserves address blocks 10/8, 172.16/12, 192.168/16 for private use
  People start using them for NATs
- 1997: RFC - dynamic DNS updates
- 2000: root servers see sharp increase in PTR updates for private addresses

Evi starts looking into this and other problems, suspects Microsoft

Transaction

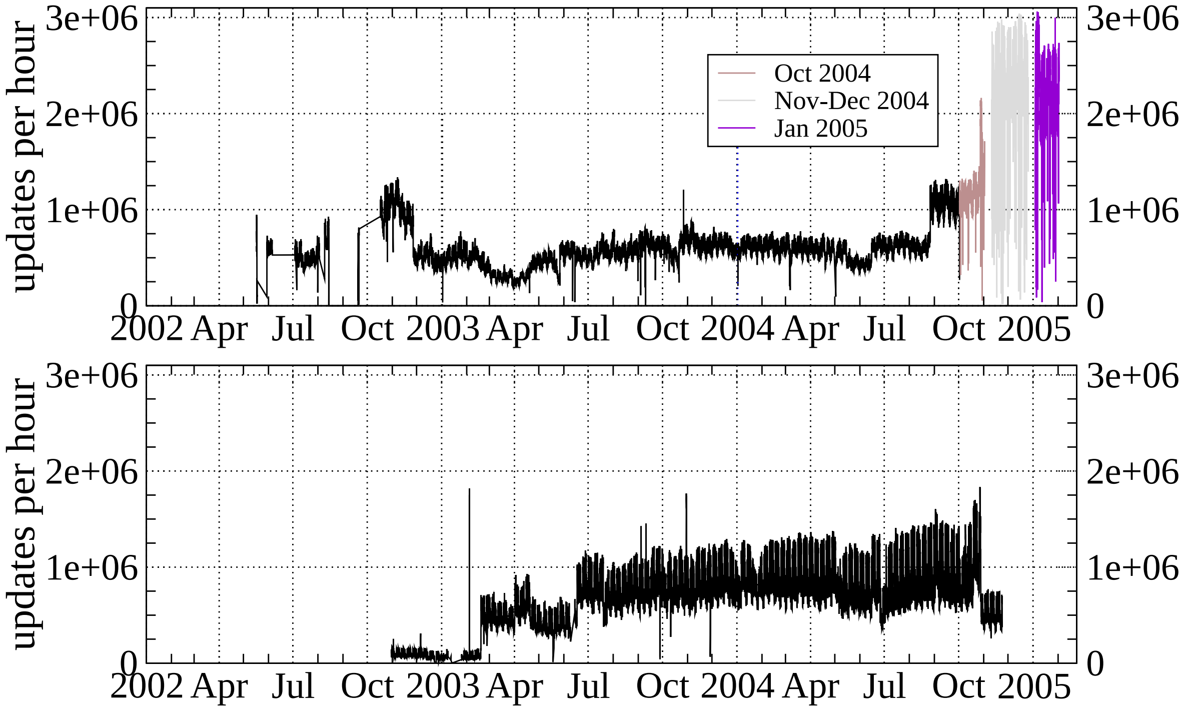
- A host with a globally routed IP address sends an update packet (UDP)
- PTR record (IP to name mapping) in the payload contains private IP address
- The server refuses
- The host tries the same update using TCP
- After a few attempts the host stops, waits for 5, 10 or 60 min, goes to step 1

An update fails in DNS layer; TCP/UDP are fine

Remedy: AS 112 project

- Vixie and other operators introduced three servers authoritative for rfc1918 space
- Two servers process queries, one – updates
- prisoner.iana.org (192.175.48.1) is anycasted
- In Jul.2004 12+ ASes provide this service
  - 40% Route Views peers see ISC
  - some peers see AS 7500 (WIDE)
- Our data consists of BIND logs from Palo Alto (hazel) and Osaka
- Courtesy Paul and Akira

The Routing Change Story



RFC1918 DNS updates, May 2002--Dec 2004

Top: hazel (ISC). Bottom: Server near M-root

Server at Osaka (below) has less traffic, but higher spikes

The changes are very abrupt, not long-term trends

Dynamics - Osaka as112 server

- Very bursty even on hourly scale
- The largest spike at 1 AM - Korea?
- Starts low in Oct 2002, under 100k/hour
- Jumps to 500k/hour in Feb 2003
- Jumps to 700k/hour in Jun 2003
- Grows slowly in 2003-2004
- Jumps up in mid-Oct 2004, about 1 M/hour
- Drops on Oct.27 to Feb 2003 level, 500k/hr

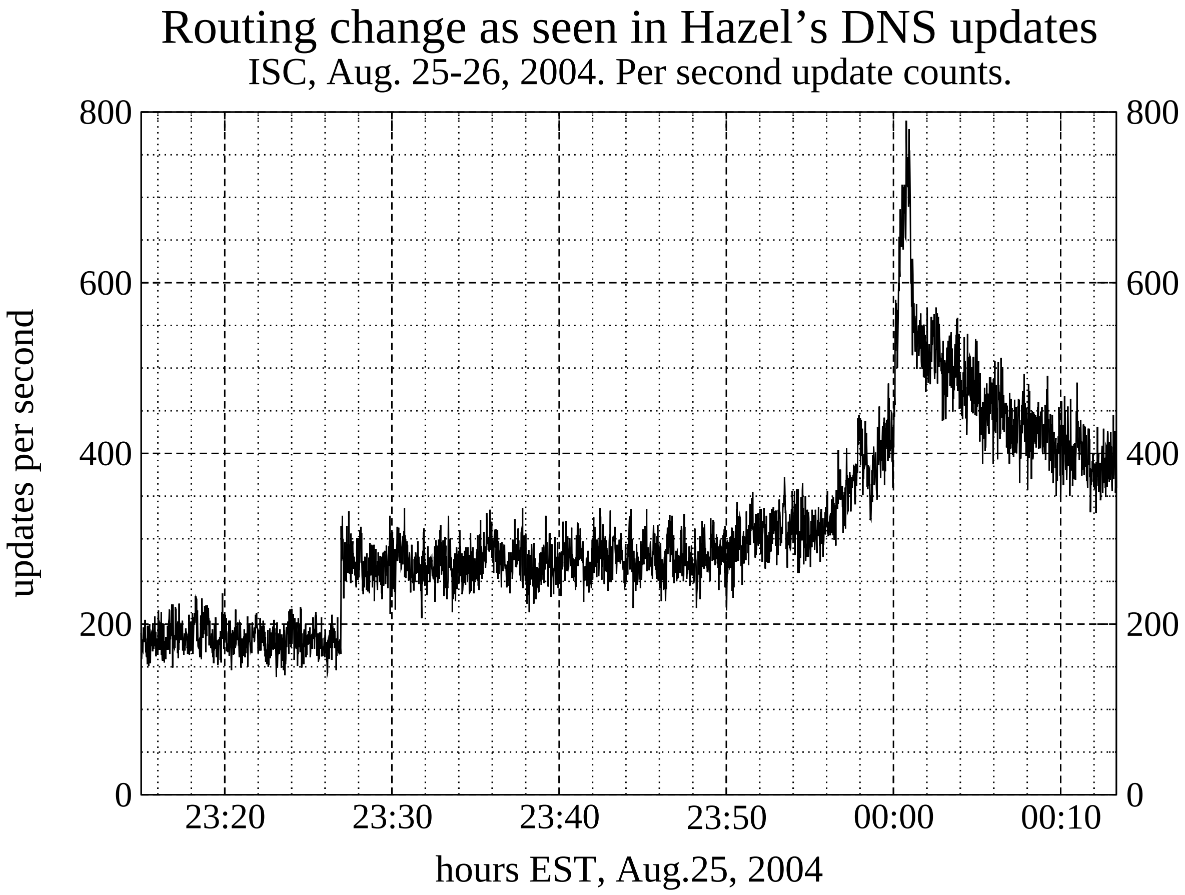
Are these jumps and drops
caused by routing changes?

Dynamics - Palo Alto

- Starts at 1M/hr in Oct 2002
- Drops to 500k/hr in Nov.2002
- Dips to 250k/hr and back in Jan-Jul 2003
- In 500k-700k/hr range, Jul 2003-Jul 2004
- Jumps up to 1 M/hr, Aug.25, 2004

The changes in update rates are very abrupt
Is it an artifact of hourly aggregation?

Palo Alto Aug.25, 2004 change



The change happens within one second
It is very likely we see a routing change

More evidence of routing change

- The weekly patern is qualitatively the same
- The update rate increased by 2/3
- The amplitude max/min increased by 2/3 too
- Everything scaled up - ”more of the same”

Routing table analysis

- Compare two sets of prefixes:
- 500K updates in 7 hours before the change
- And 500K in 4.3 hours after 03:00
  (we skipped midnight as a non-typical time)

Prefixes increase from 9k to 15k, by 62%
ASes increase from 1.7k to 3k, by 72%

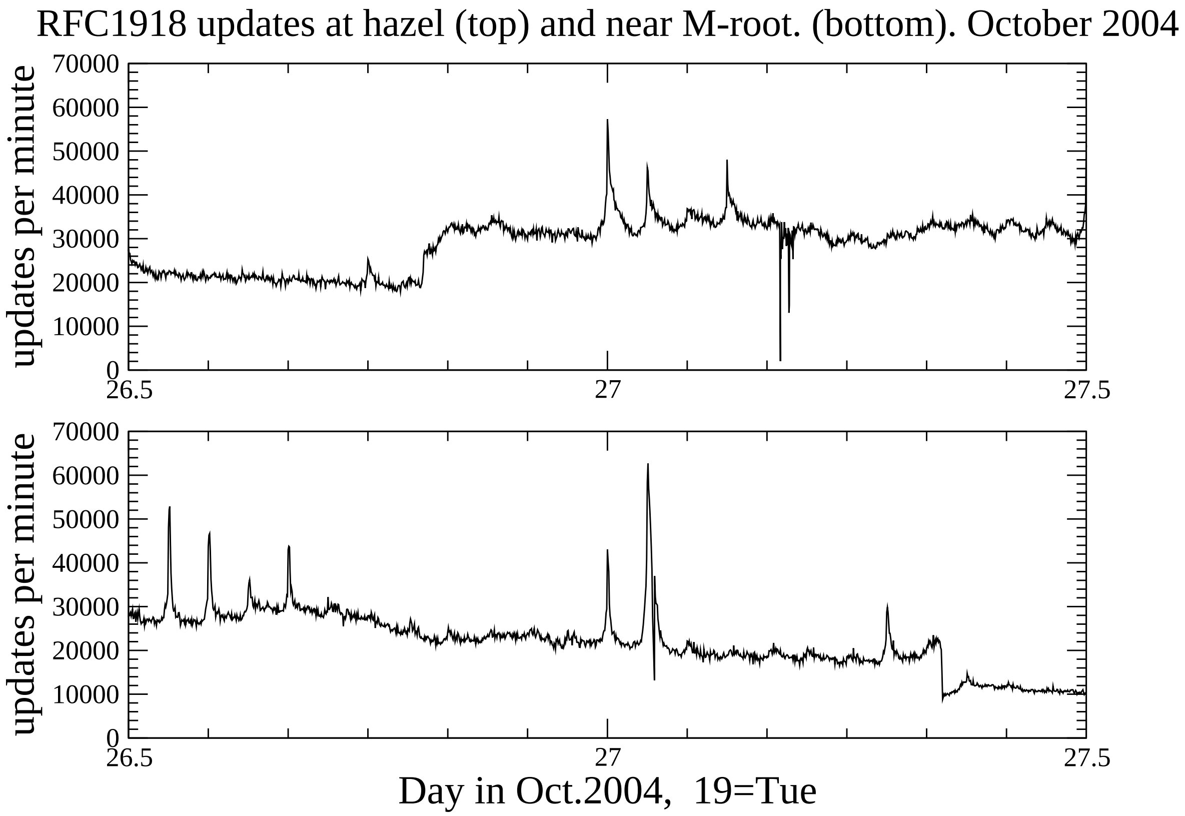
Rate, prefixes, AS counts changed proportionally

Representativeness - an aside

- Our data is contributed by:
  - 10% of all prefixes
  - 17% of all ASes

Taken with Osaka server, it represents even larger fraction of all networks

Load shift: Osaka to Palo Alto (hazel)



- Two load changes match in time
  - Palo Alto goes up (7pm EST Oct 26)
  - Osaka goes down (8am JST Oct.27)
- Magnitudes also comparable (170 upd/sec)

Conclusion – Part 1

- Route changes happen
- The load can suddenly move
- We observed almost 2-fold increase

Is our global anycast server system stable under these conditions?

The Microsoft Story

Highest update peaks

- Osaka as112 server:
  - 3889 in Apr 2004
  - 2584 in Sep 2004
- Palo Alto - Hazel
  - 3101 in Sep 2003
  - 2380 in Jan 2005
- One update = 30 packets
- 4k updates/sec = 120 kpps

Questions

- Who is doing updates?

- What happens if one server goes down?
- Can we have a domino effect?
- Why do we see stronger peaks at Osaka?

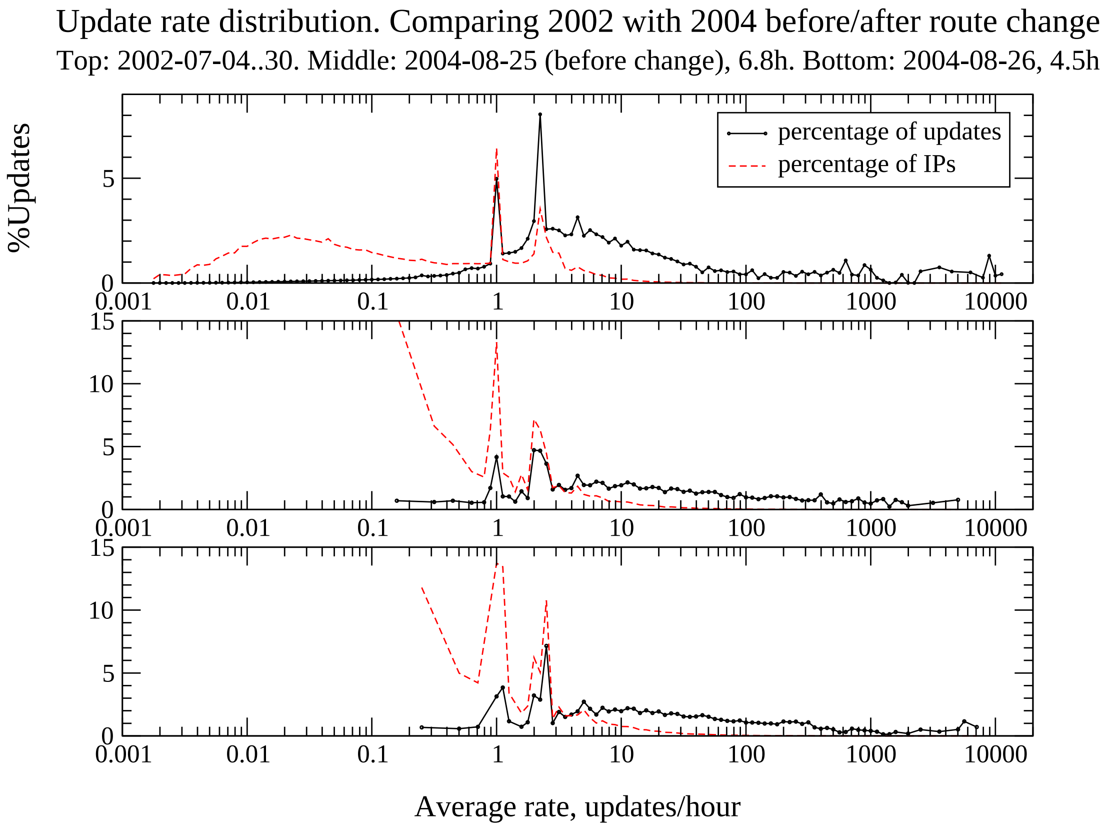
How should dynamic DNS updates
for RFC1918 addresses be done?

Update rates of individual hosts

- Our 2002 study: many boxes with
  - One update per hour
  - 3 updates per 75 min (2.4/hr)
- We find no qualitative changes

Many updates come from hosts with
1 or 2.4 updates/hour

Updates by host rate, Palo Alto



X axis: average rate, updates per hour
Y axis (black): percentage of updates
Y axis (red): percentage of IPs

Top: A histogram from 2002 paper
Middle: Aug.25, 2004 before route change
Bottom: Aug.26, 2004 after route change

TCP senders

- 2002 lab study of Microsoft boxes:
  - Always try Transact.Signature (secure upd.)
  - Done by TCP, three times in a row
  - Very few other boxes do TCP (see below)
- Duane ran tcpdump so we could check

I wish we did it in 2002

TCP senders - incoming packets

- TCP packets: 68.72% (1.7 M)
- UDP packets: 6.80% (0.17 M)
- TCP/UDP pkt: 75.52% (1.9 M)
- All incoming 100% (2.5 M)

TCP senders account for 3/4
of incoming packets at the server

Microsoft in the TCP payload

- ”gss.microsoft.com” in TCP DNS payload followed by domain name
- Sources with ”microsoft”: 56.5% (64k)
- Total #unique sources: 100% (114k)
- Sources saying ”microsoft” send 74.4% pkts

More than 1/2 sources
and about 3/4 packets are from MS boxes

Fingerprinting Microsoft boxes

- Passive OS fingerprinter p0f by Zalewsky
- Matches Syn packet with a list of signatures
- We have 70k IPs that sent a Syn
- p0f says 67k are Windows

p0f classifies 96% of TCP sources as Windows

microsoft is already in the payload
but p0f provides an independent confirmation

Conclusions

- Update rates are higher than in 2002
- Routing changes can potentially affect server system stability
- Windows machines are over 1/2 of all sources
- They send the majority (3/4) of packets
- The reason is their persistence:
- One UDP and 3 TCP attempts

Future work

- Fingerprinting individual boxes by event timing
- Potential clues:
  - The timer slop in the 5-10-60 min intervals, tends to be close for either interval
  - The offset in midnight update time
  - The drift of the midnight update time

(TCP timestamps are very rare
Usenix paper techniques may not work)

Acknowledgements:

- Paul Vixie
- Akira Kato
- Young Hyun
- Dan Andersen
- Duane Wessels
- Marina Fomenkov
- Brad Huffaker
- Evi Nemeth
- Dima Krioukov
- kc claffy