renesys

The Curious Incident of 7 Nov 2011

James Cowie
Andrew Hobgood

NANOG 54

Presentation Overview

- Remembering November 7th 2011
- Initial event review: what it *wasn't*
- Impairment analysis using traceroute
- Searching for the root cause in BGP
- *Lessons for next time*

A day in the life at Renesys

14:09 UTC, 7 November 2011 (09:09 EDT)

Once again, something “really bad” has happened to the Internet

- Anecdotal reports from #irc, twitter, @outages
- Widespread perception of unreachable destinations, dropped connections, BGP session resets, etc. with no clear common cause

A day in the life at Renesys

- 14:24 UTC, 7 November 2011 (T+00h15m)

```
<dtemkin> does renesys have a report
    ready for this yet? :)
```

A day in the life at Renesys

- 14:24 UTC, 7 November 2011 (T +00h15m)

```
<dtemkin> does renesys have a report
    ready for this yet? :)
```

- 13:56 UTC, 7 November 2011 (T -00h13m)

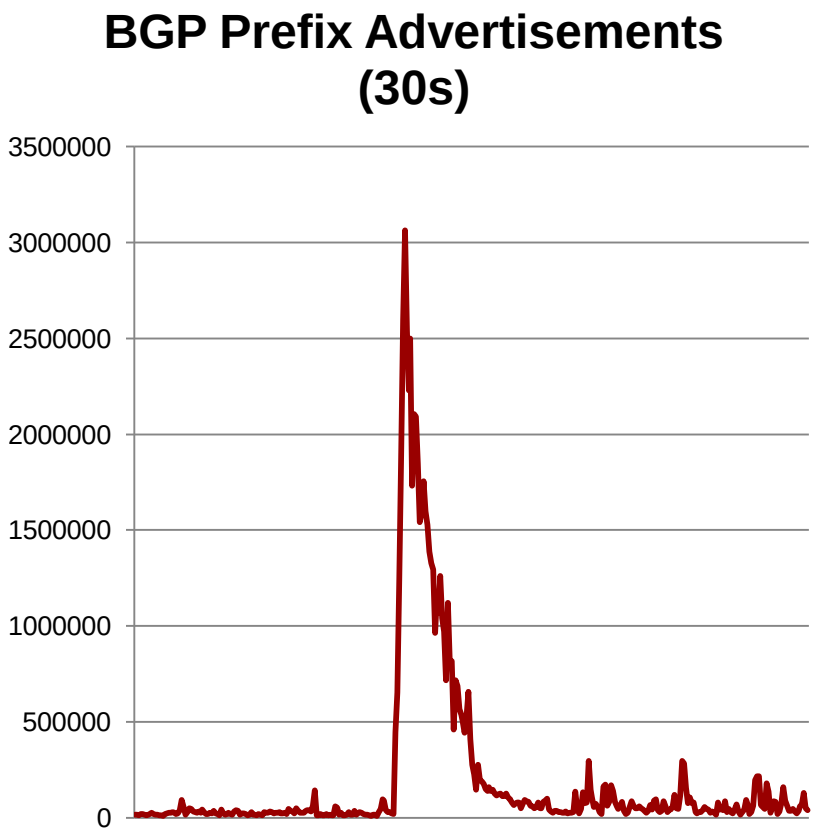
Renesys PreRouting Unit emerges from flotation tanks with initial report

“No, Seriously, what was that?”

- 14:09:30 UTC, 7 November 2011
- Initially, fingers point at Level(3)
  - Then at Tata
  - Gradual recognition that this was a global event affecting more than one or two providers
- Strong transatlantic impacts to traffic
- Dozens of countries visibly affected
- **7.4% of Renesys BGP peers reset**

First Symptom: Skyrocketing BGP Rates

- Classic signature of infrastructure instability
- Worms (Nimda 2001, Slammer 2003) used to cause storms like this
- More recently, tickling a router bug or two is the more likely scenario



Pathologies We Expected To Find, But Didn't

Weird ASPATH?

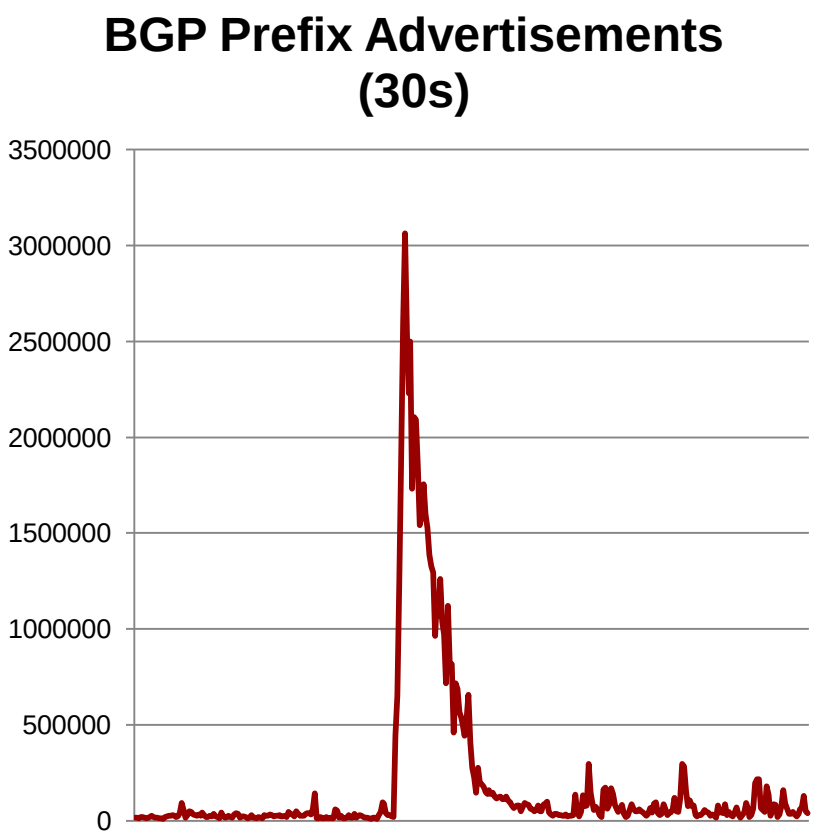
- Mikrotik range-checking bug causes runaway prepends, 255-length paths, tickles Cisco bug (Feb2009)
- Afrinic sends highly prepended paths for new 4-byte ASN, tickles Quagga buffer overflow problem (May2009)

Bad attribute?

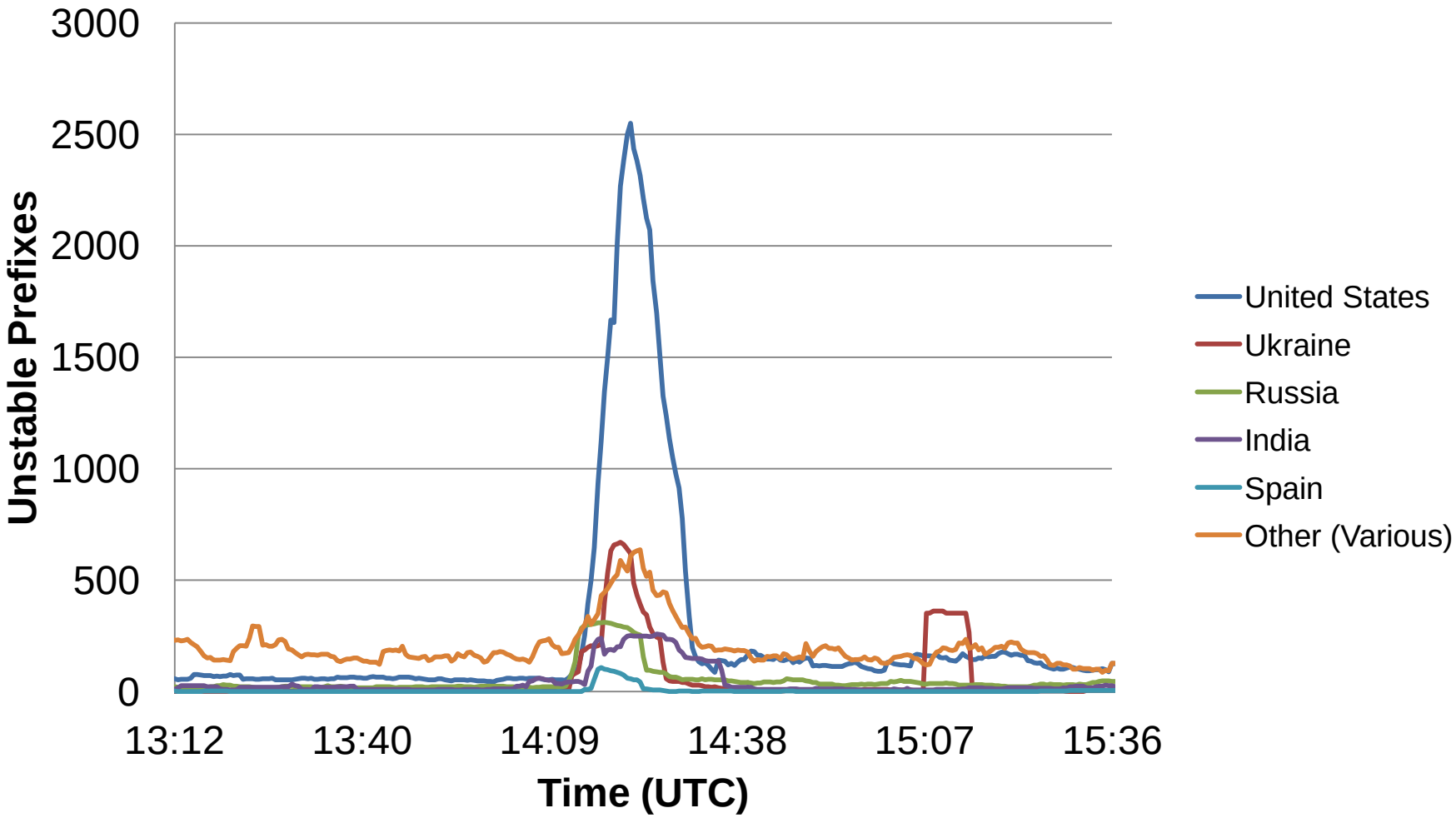
- Empty AS4_PATH attribute takes down Cisco IOS XR (Aug2009)
- RIPE RIS sends compliant-but-ugly BGP attribute to all peers at AMS-IX, another Cisco bug corrupts-and-forwards, recipients reset sessions (Aug2010)

No obvious trigger messages?

- Was this a new kind of cybernuke?™
- When one major vendor has a bug, and the others don't, they act like immune carriers
- Worldwide propagation to every vulnerable router follows within 30-60s
- BGP is a mess

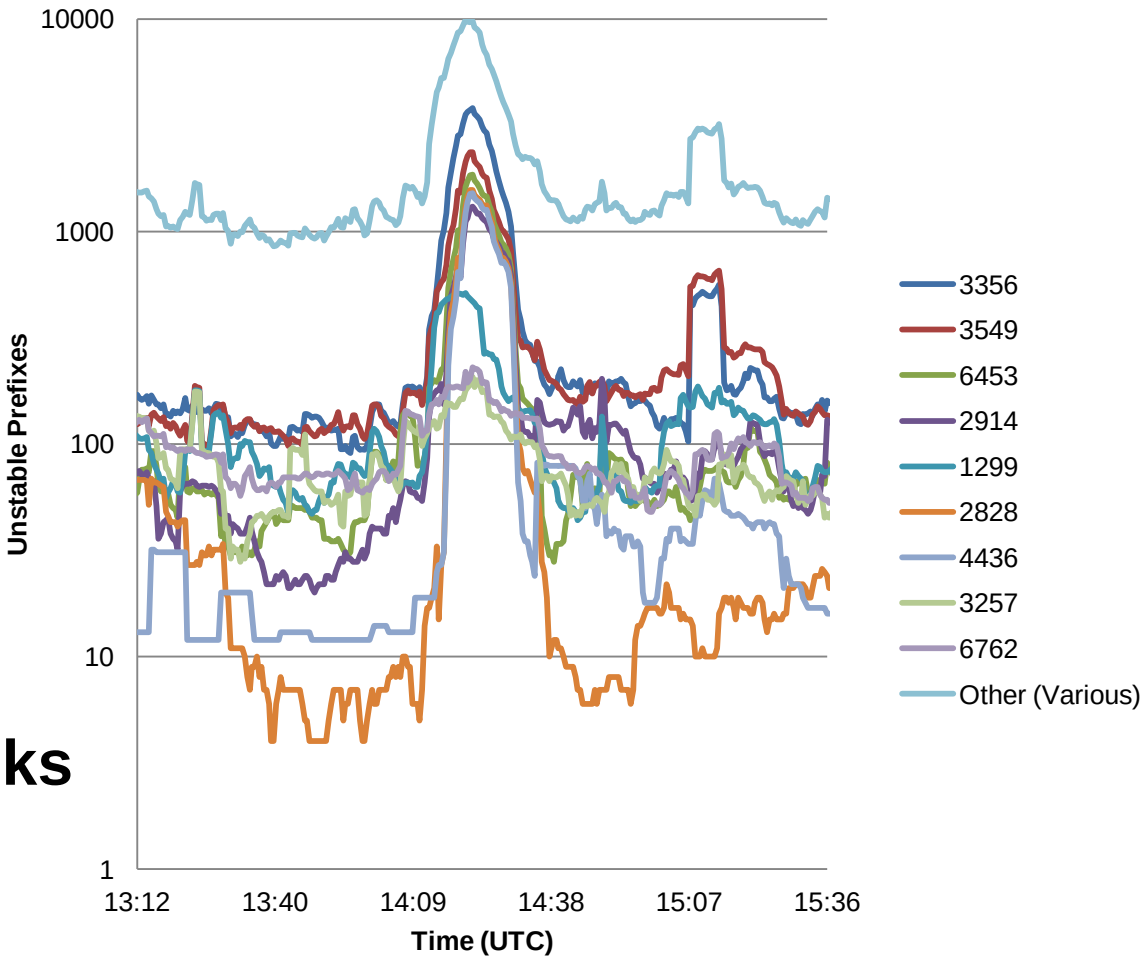


Newly unstable prefix counts, by country



Unstable prefix counts, by upstream

- **Broad-spectrum instability causes impacts in virtually every carrier's customer cone**
- **No single smoking gun in BGP, as one might expect: no single origin, no transit provider sticks out as the problem**



2. Inferring the scope of infrastructure impairment

Distributed Traceroute Dataset

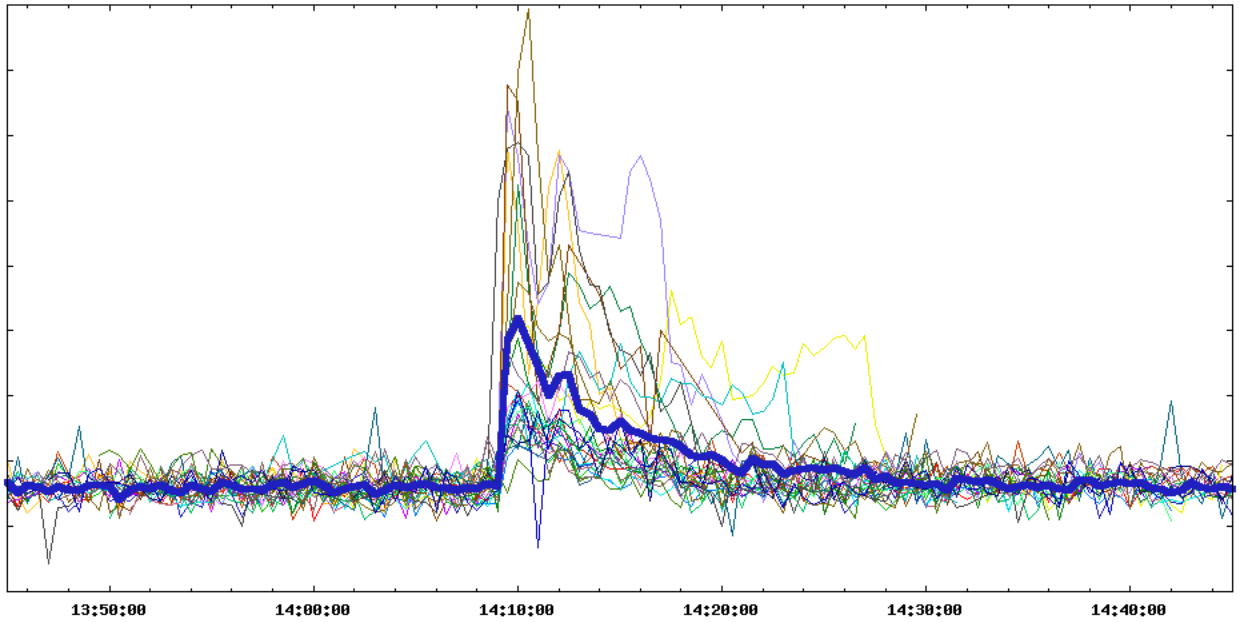
- Developed to complement BGP infrastructure analysis in “broad impact” events like this one
- Dozens of collection points worldwide
- Nearly 1 million IPs traced daily
- ~8 billion ICMP traces collected each year
- “Noncompletion rate” (did the destination fail to respond?) is a rough proxy for global end-to-end connectivity from a given site

Analysis Challenges

- Traceroutes can fail for many reasons
- Analysis limited to the traces that were “in flight” during the time of the event
- Problems close to collectors mask/simulate problems downstream (need diversity!)
- Event itself lasted only minutes, created widespread problems that were hard to distinguish

Traceroute Noncompletion, 7 Nov 2011

Simultaneous worldwide increase in rate of failure of end-to-end traceroutes



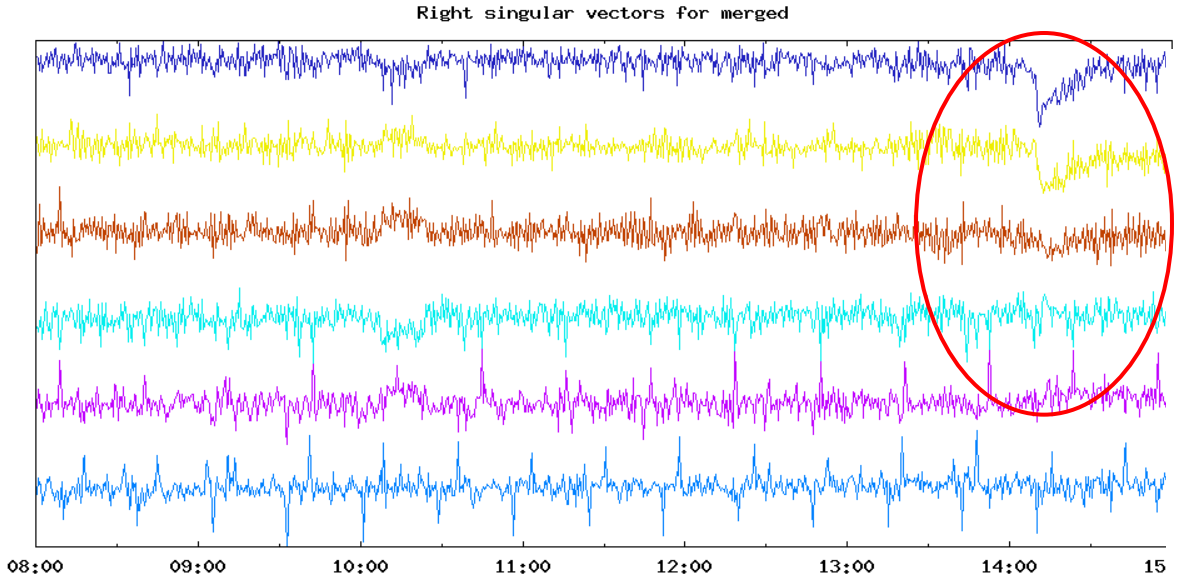
- Failure rates double worldwide *within 30s*
- Impairment decays, lasts 10-20 minutes

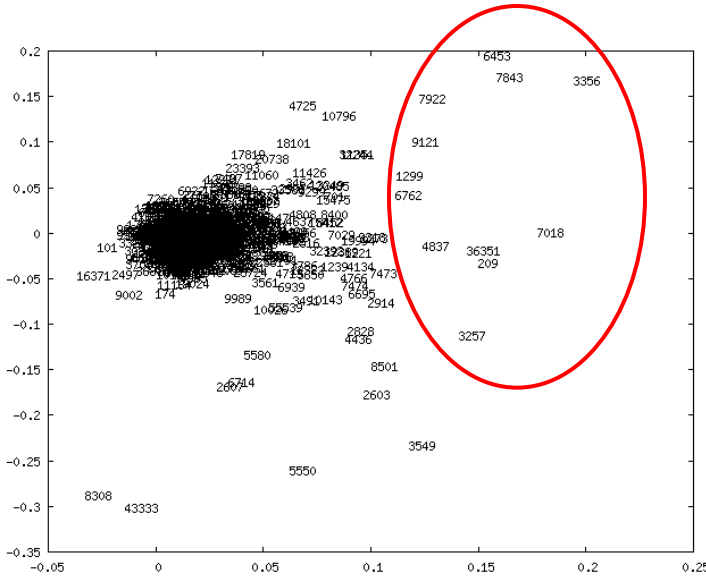
Seeking Evidence of Impairment

- Examined 34M traceroutes gathered during the first 15 hours of 7 November
- Identified router interface pairs (“edges”) that were reliably crossed by traces before and after the event, but not during the first 10 minutes
- 28,000 out of 2M edges are “highly traversed”
- Try to understand which providers participate in the “highly impaired, highly traversed” subset
- ***Warning: math***

Obvious impairment across key ASNs

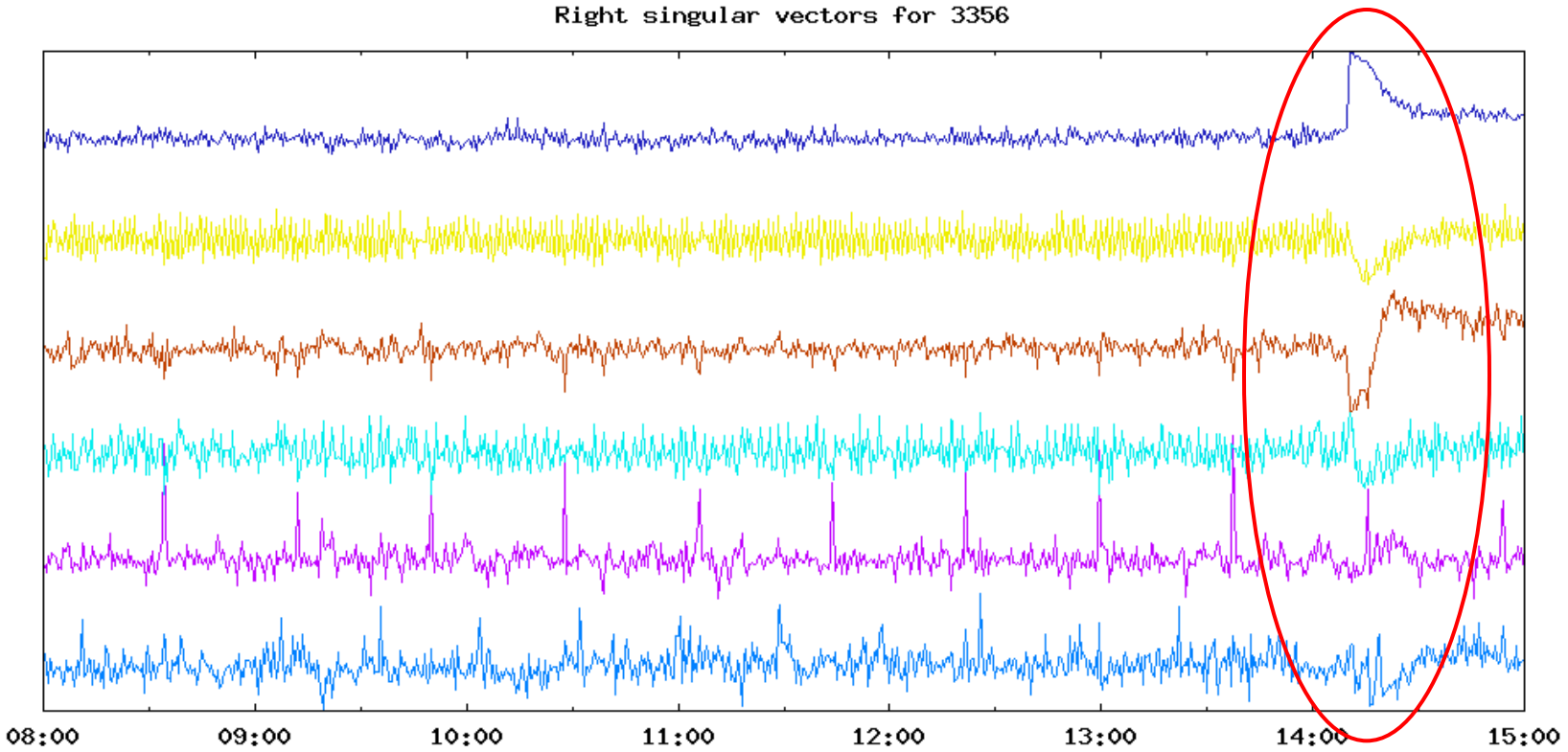
- Create normed matrix of 30s tracecounts per ASN
- Compute singular value decomposition
- 14:09 event visible as primary source of variance
- Confirms impacts on 6453, 3356, 7922, 7843, 7018





Level(3)'s Impairment

- Analyzed sample of 3,110 heavily traversed edges
- 650 edges correlated strongly with first 4 sing.vec.



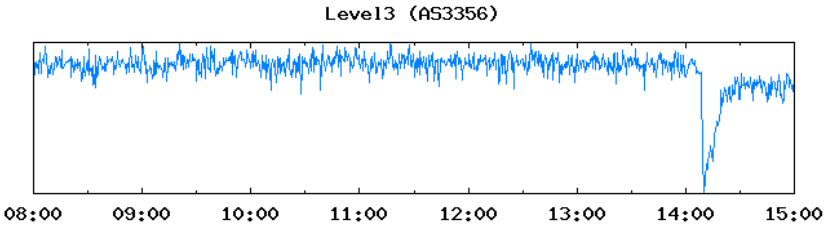
Level(3)'s Impairment

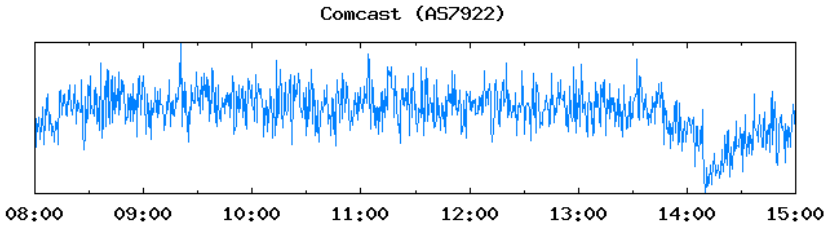
- 650 of 3,110 heavily-traversed traceroute edges were severely impacted in the first 10 minutes
  - 549 (84%) are internal Level(3) edges
  - The remainder were spread evenly among 46 peers and customers worldwide
- Heaviest impacts in Los Angeles (29%), San Jose (12%), Frankfurt (11%), London (10%)
  - LAX edges to Hutch, Hanaro, Chunghwa
  - SJC edges to KDDI, KT, Roadrunner
  - Frankfurt to Turk Telekom

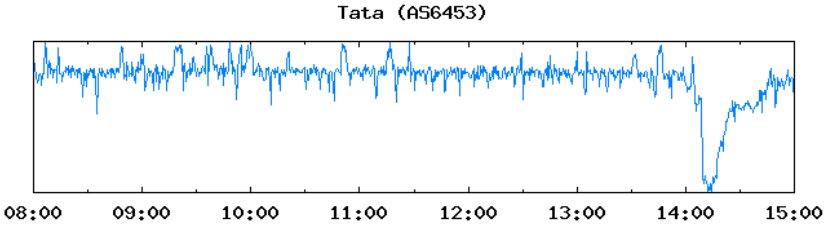
Tata's Impairment: similar but lower

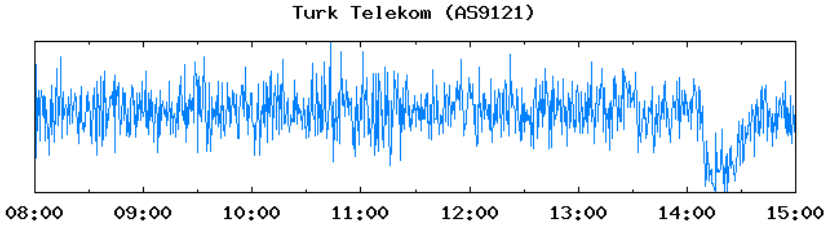
- 438 of 1,514 heavily-traversed traceroute edges were severely impacted in the first 10 minutes
  - 305 (70%) are internal Tata edges
  - The remainder were spread evenly among 40 peers and customers worldwide
- Heaviest impacts in New York (16%), Los Angeles (12%), London (12%)
  - But also Singapore, Hong Kong, Mumbai, Tokyo
  - NYC to Roadrunner
  - SJC to Comcast

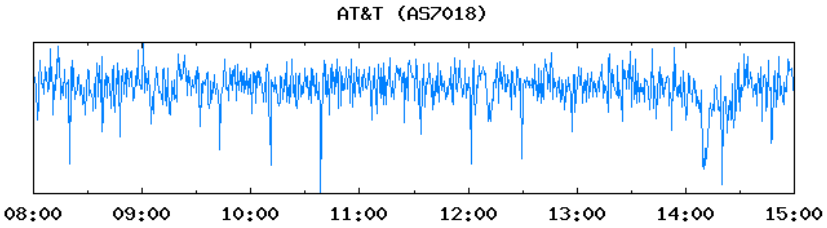
Traces with responding destination, 30s

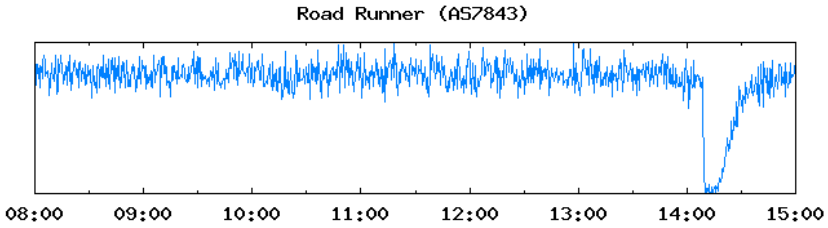




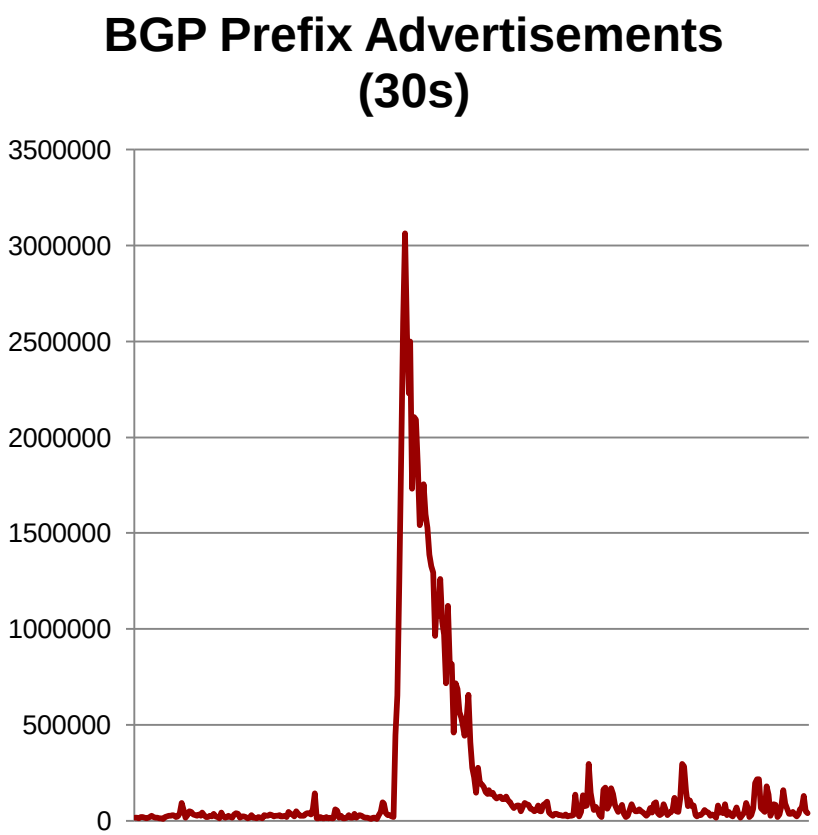




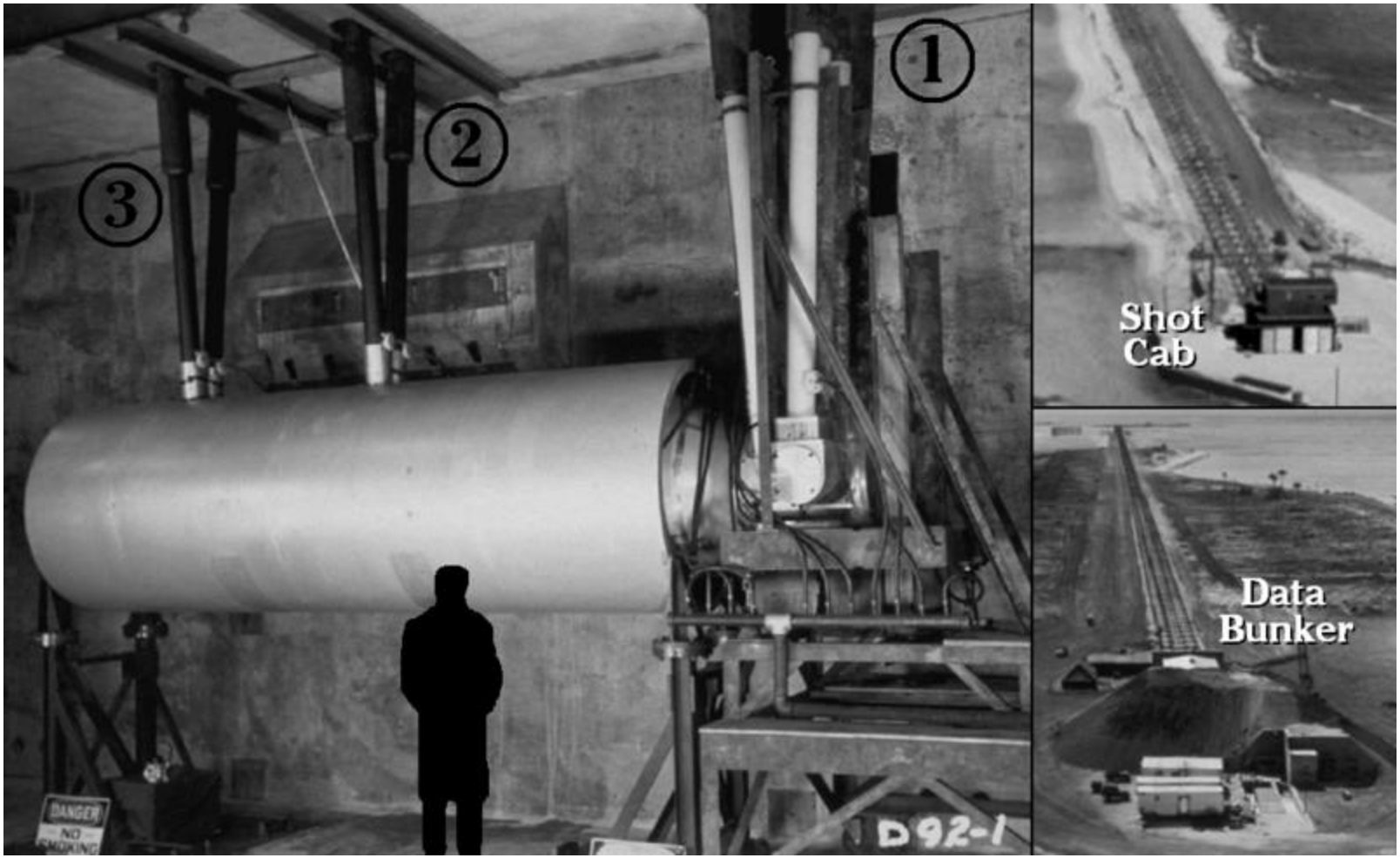




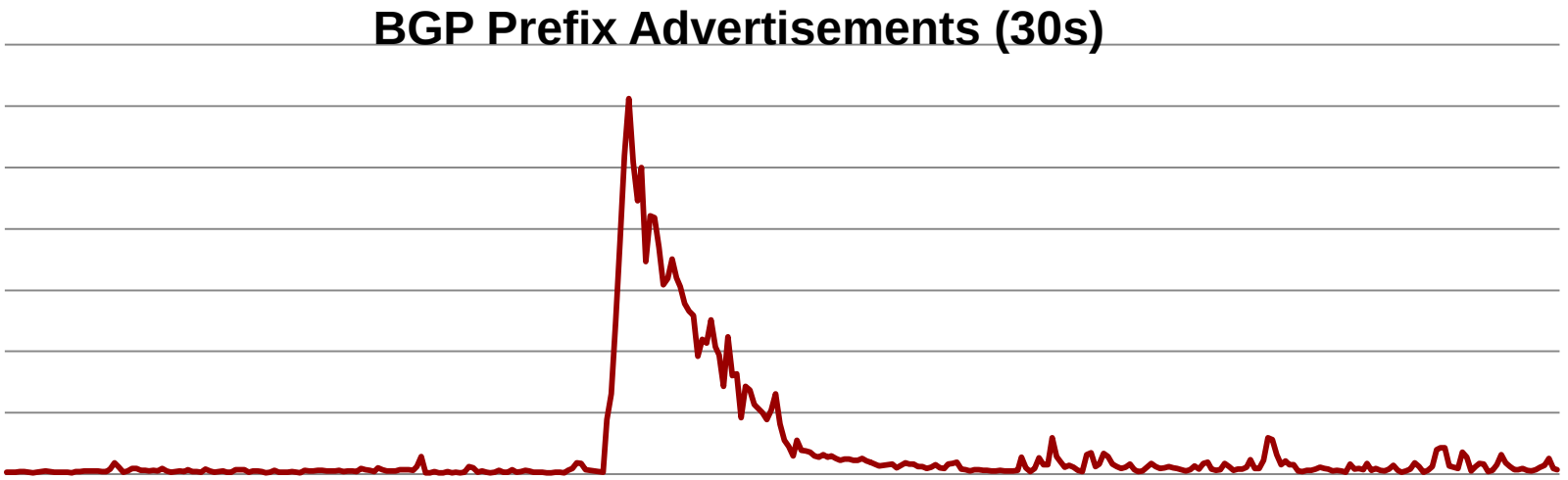
3. Searching for the Root Cause

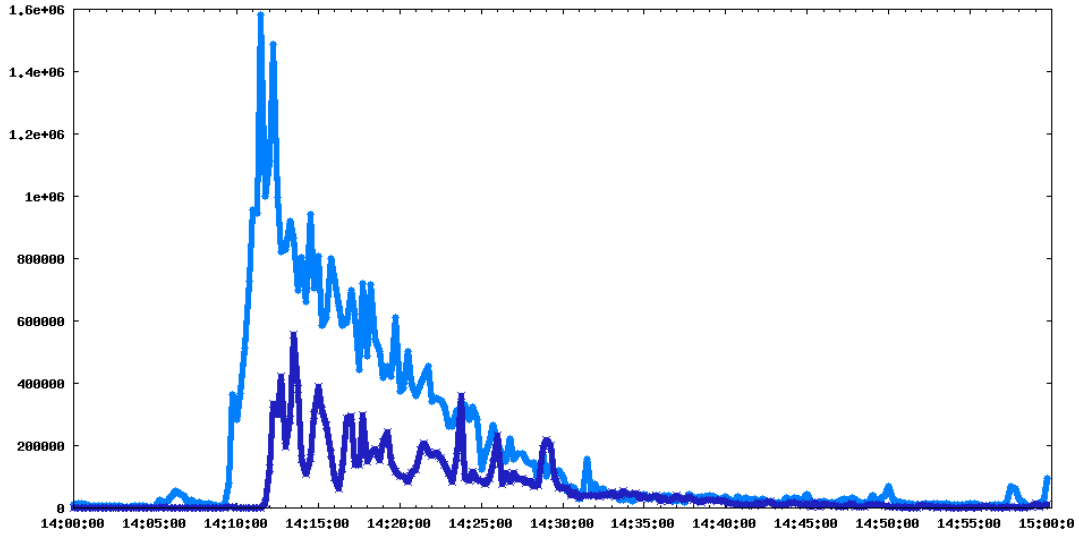


“Light Pipes” record the first moments



Don’t mistake session resets for signal

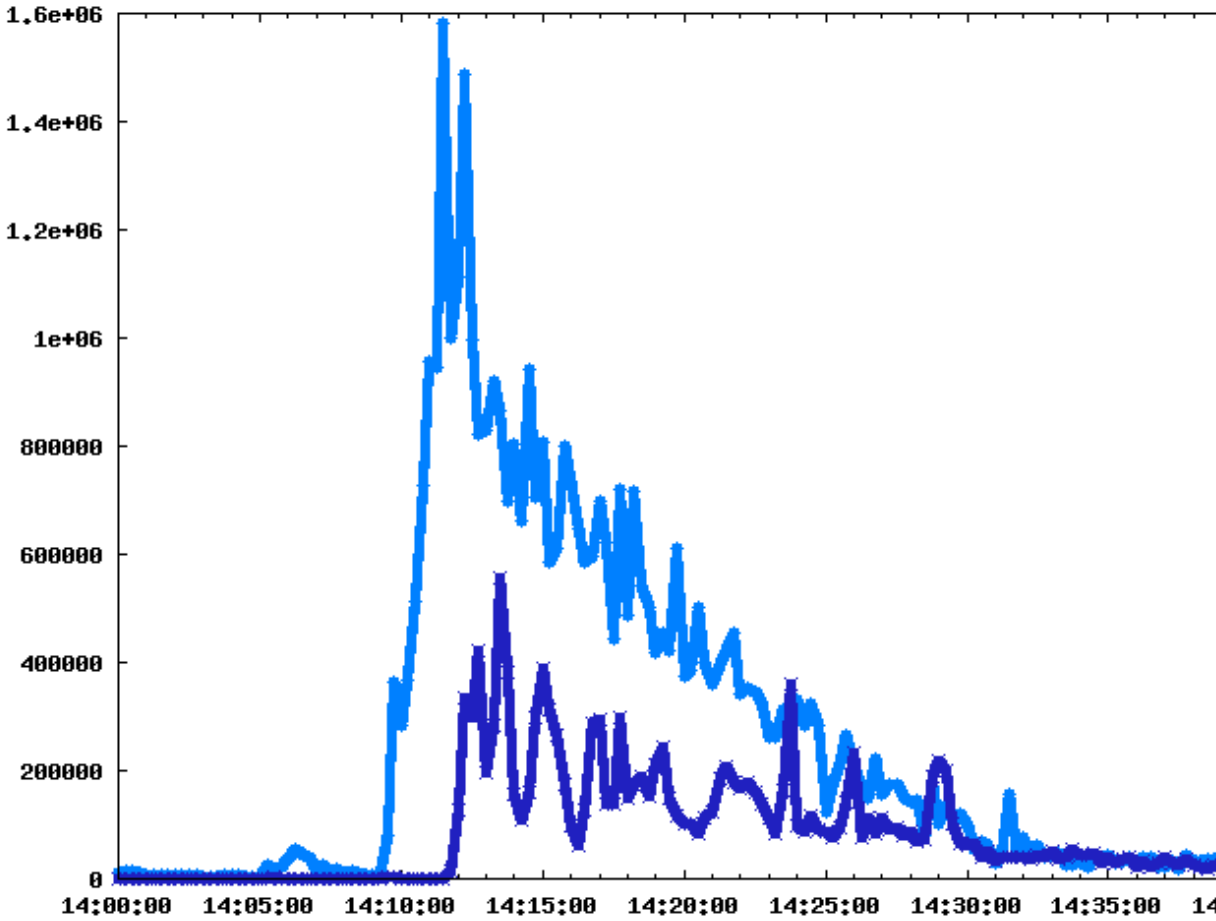




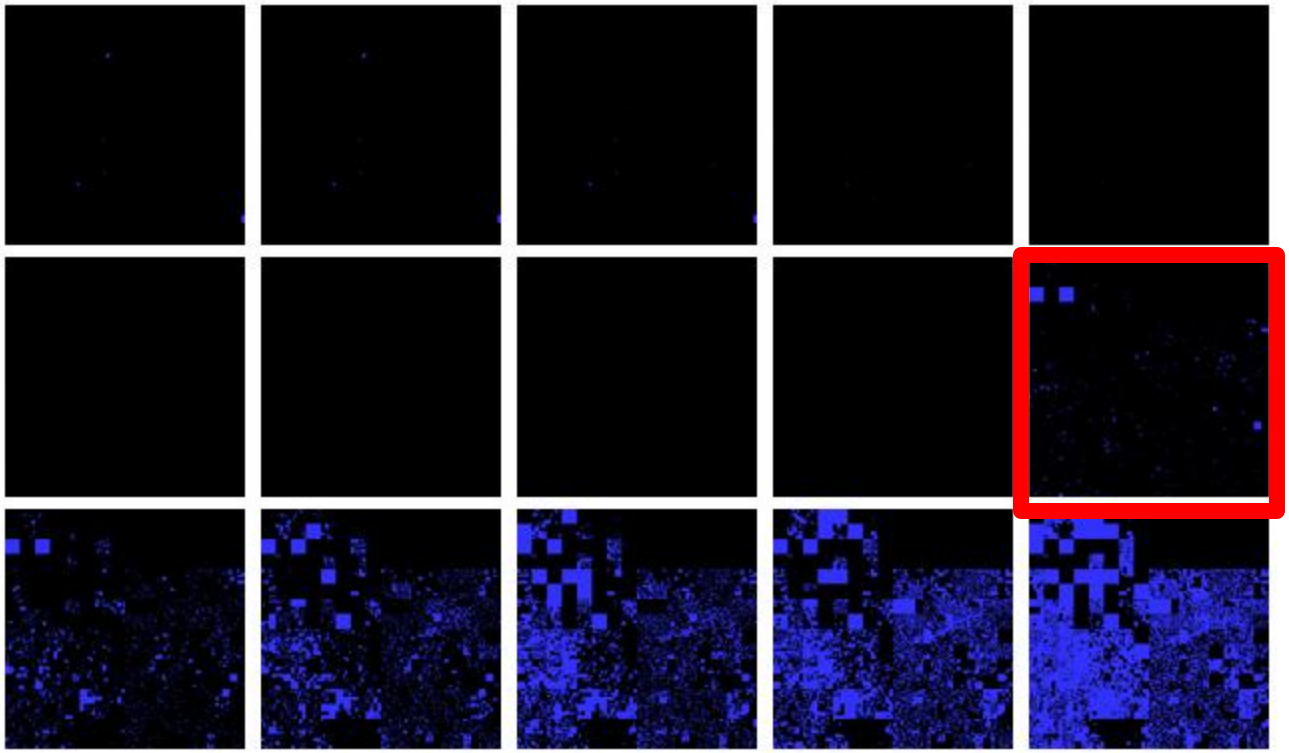
- **Dark blue**: session resets
- **Light blue**: *no resets*

Primary spike *not* simply session resets

- **'Echo instability'** from our peers' peers arrive as "first light" (14:09:30 UTC)
- Our peers start to reset as a **second wavefront** (14:11:22 UTC)



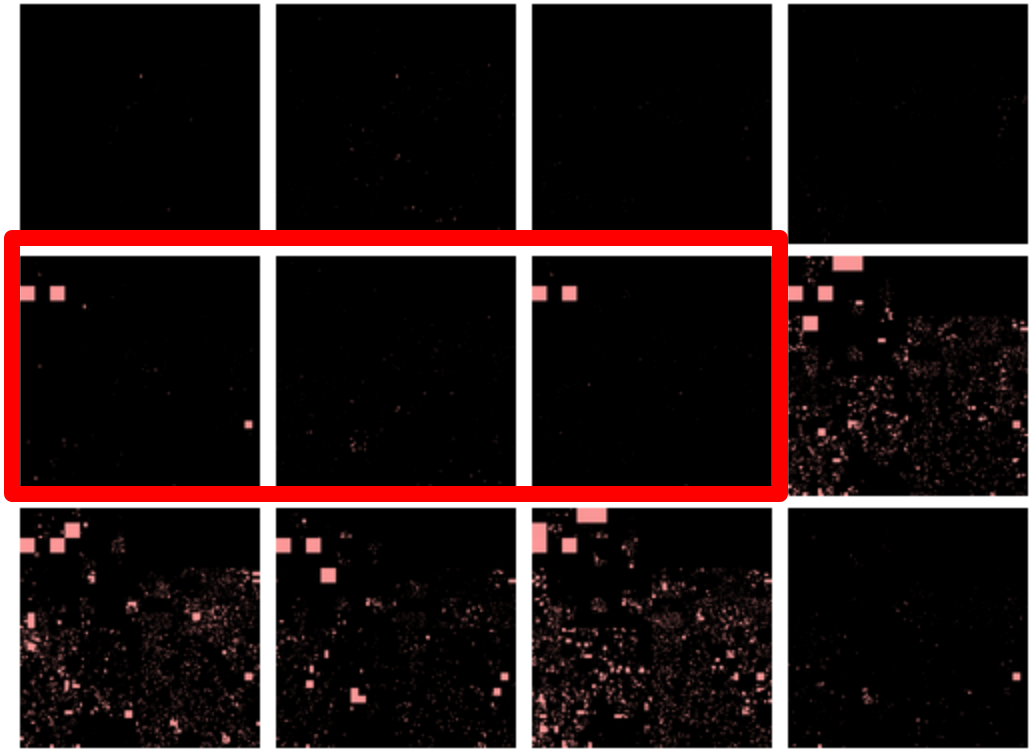
Prefixes in the 7-minute rising edge



- 30-second snapshots in IPv4 Hilbert-curve space
- 14:05:00 through 14:12:00

- **Note first light just after 14:09:30**

10x zoom; first 30 seconds

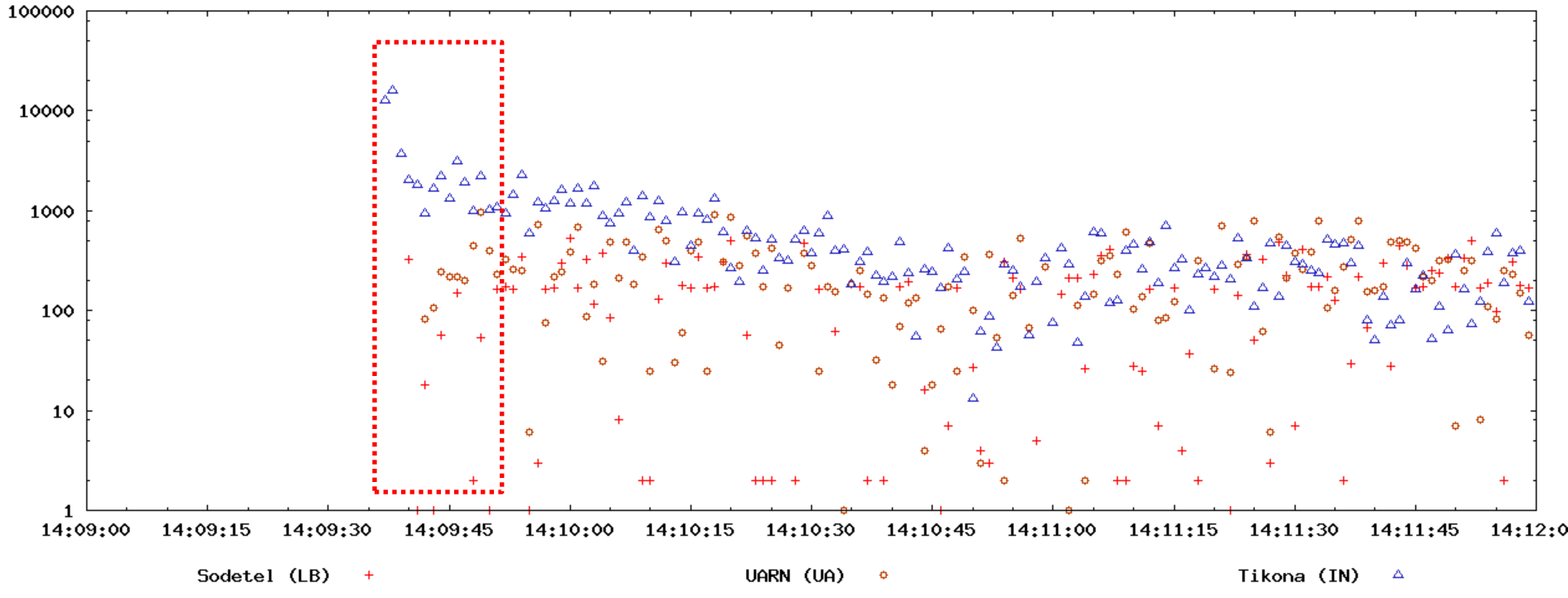


- 3-second snapshots in IPv4 Hilbert-curve space
- 14:09:25 through 14:09:58

- **Examine prefixes in 9s window: 14:09:37-46**

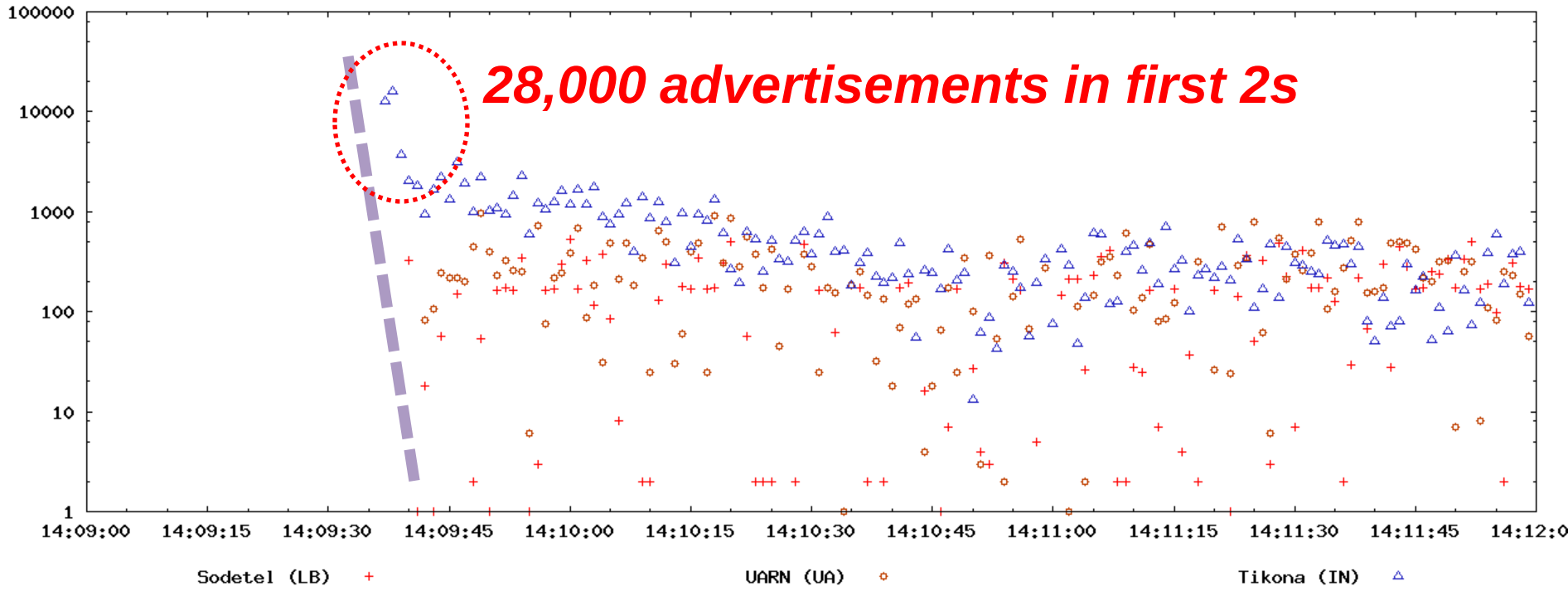
Trigger candidates (log(bgp adv/sec))

- **Sodetel** AS31126 (Beirut, Lebanon)
- **Tikona** Digital Networks AS45528 (Mumbai, India)
- **Ukrainian** Academic Res Net AS3255 (Kiev, UA)

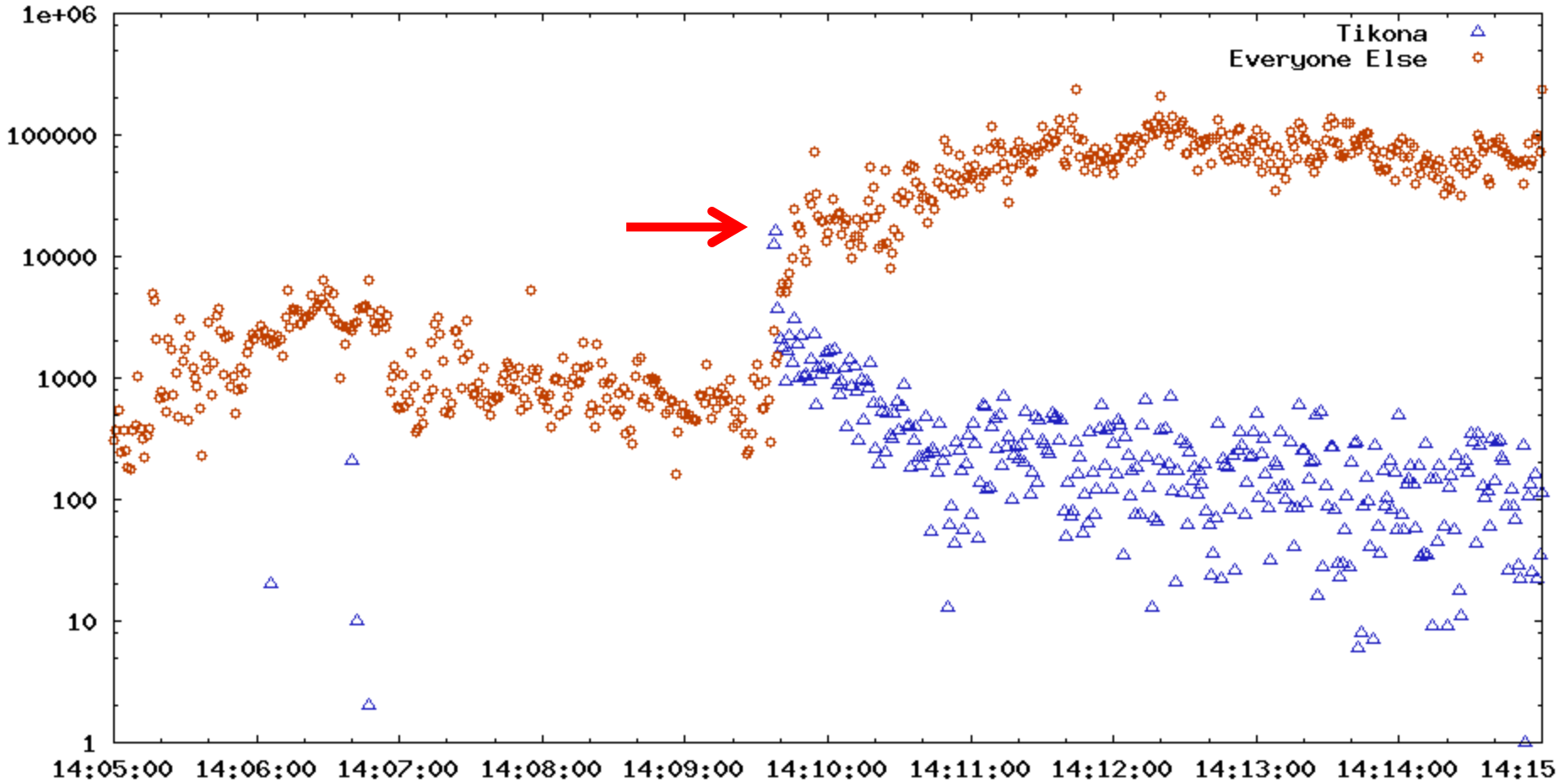


Trigger candidates (logscale, 1s adv)

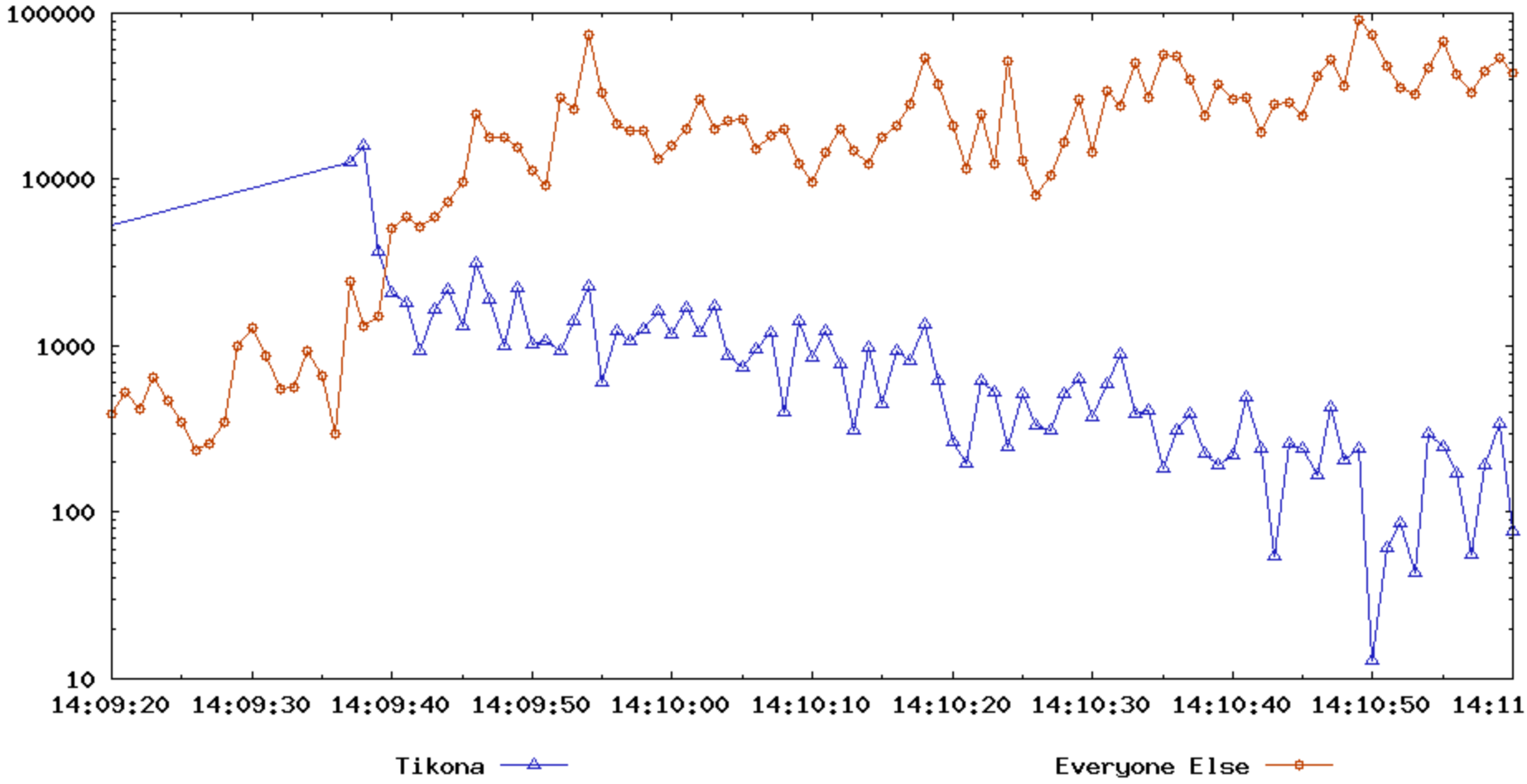
- **Sodetel** AS31126 (Beirut, Lebanon)
- **Tikona** Digital Networks AS45528 (Mumbai, India)
- **Ukrainian** Academic Res Net AS3255 (Kiev, UA)



Tikona Briefly Outshines the Rest of the Internet

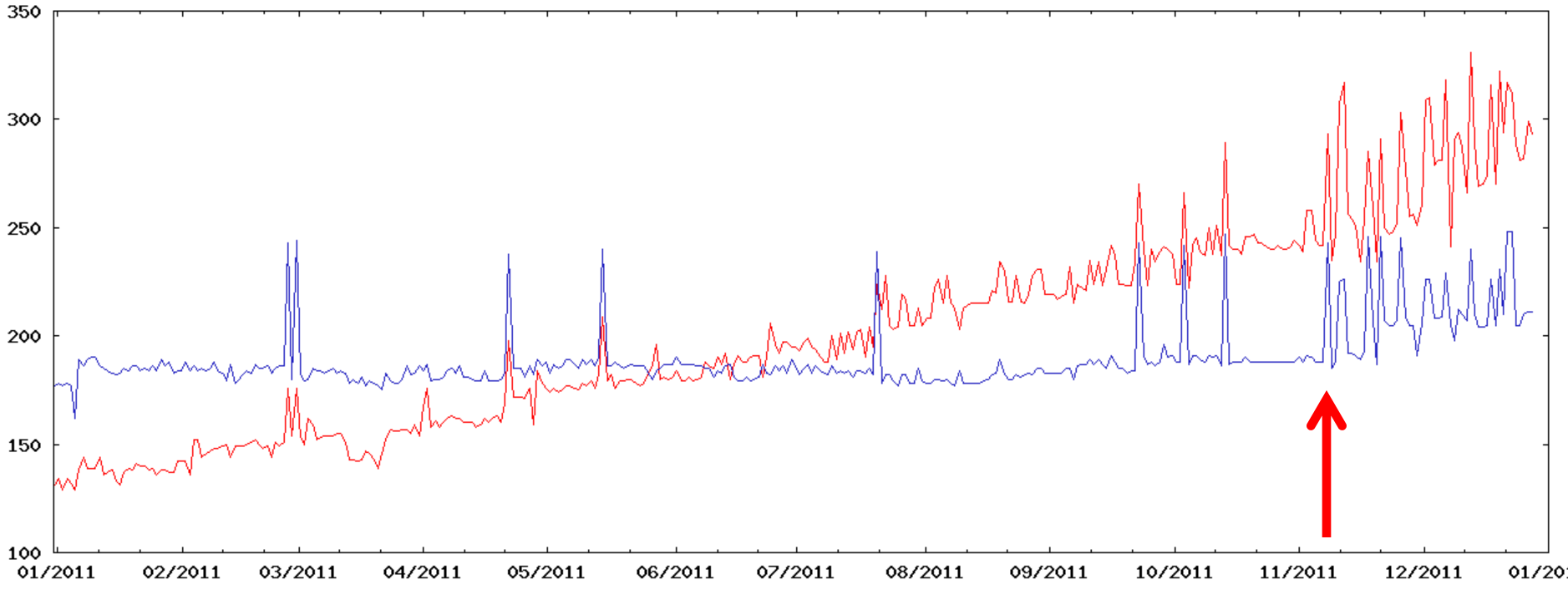


Tikona Briefly Outshines the Rest of the Internet



Tikona's increasing deaggregation history

- 113.193.0.0/16 (since Nov 2008)
- 1.22.0.0/15 (since Jul 2010)



JunOS PSN-2011-08-327 (8 August 2011)

"MPCs (Modular Port Concentrators) installed in an MX Series router may crash upon receipt of very specific and unlikely route prefix install/delete actions, such as a BGP routing update. ***The set of route prefix updates is non-deterministic and exceedingly unlikely to occur.*** [..] A complex sequence of preconditions is required to trigger this crash."

Hypothesis

Tikona's bursty deaggregation of their recently acquired space from 1/8 accidentally recreated the trigger conditions for PSN-2011-08-327.

These updates spread upstream via transit providers Tata, Vodafone India, and Bharti.

Ultimately it hit Level(3) and exploded.

Summary: Could Have Been Worse

- Vendor seems to have downplayed severity of this vulnerability, probability of rare-event loss
- Carriers assessed the odds, may have elected to postpone upgrades
- Law of large numbers is unforgiving
- Trigger appeared, failures manifested, upgrades ensued under less-than-ideal conditions
- *Did not recur on reboot (critically important)*
- Broad impact but limited duration (this time)

www.renesys.com