Wide-Area Internet Measurement at MIT: Data Collection and Analysis

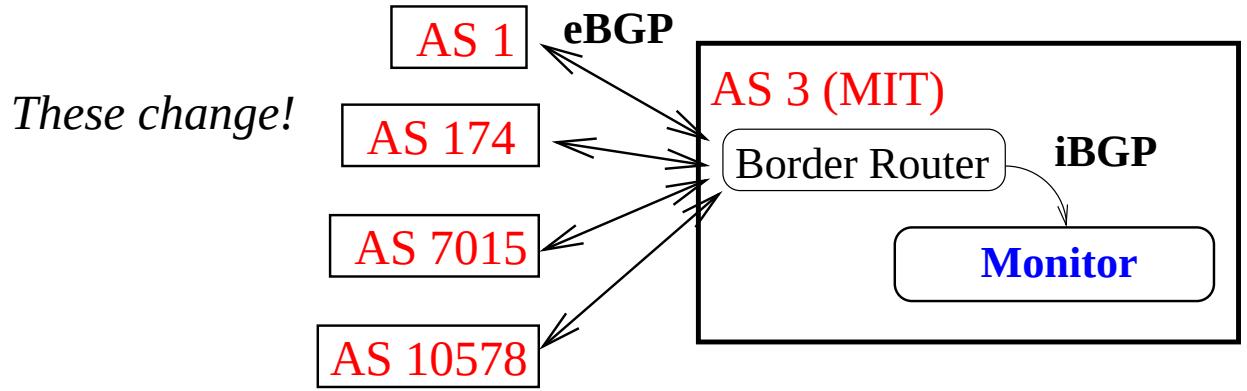
Nick Feamster, Dave Andersen, Hari Balakrishnan

M.I.T. Laboratory for Computer Science

{feamster,dga,hari}@lcs.mit.edu

Collection: Infrastructure and Data

- Topology: 31 widely distributed nodes (RON testbed)
  - Stratum 1 NTP servers, CDMA time sync
- Active Probes
  - Periodic pairwise probes; local logging for 1-way loss and delay.
  - Failure: 3 consecutive lost probes, >2 minutes
- Failure-triggered traceroutes
- Daily pairwise traceroutes over testbed topology
- iBGP Feeds at 8 measurement hosts (Zebra)



Data pushed to centralized measurement box.

General Issues with Data

- Changes in connectivity
  - IP renumbering sometimes breaks BGP sessions
  - Upstream providers change
- Home-brew tools (sometimes buggy...keep raw files!)
- Management
  - Continuous collection vs. archival (snapshots take space)
  - MySQL Table Corruption, Disk failures, etc.
  - Collection machine downtime (power outages, moves, etc.)
  - Complaints (pre-emption: DNS TXT record, mailing Nanog, etc.)
- Collection subtleties
  - Keeping track of downtimes, session resets, etc.
  - hosts are not firewalled
  - Some hosts located in "core" (e.g., GBLX hosts)
  - iBGP sessions to border router on the same LAN

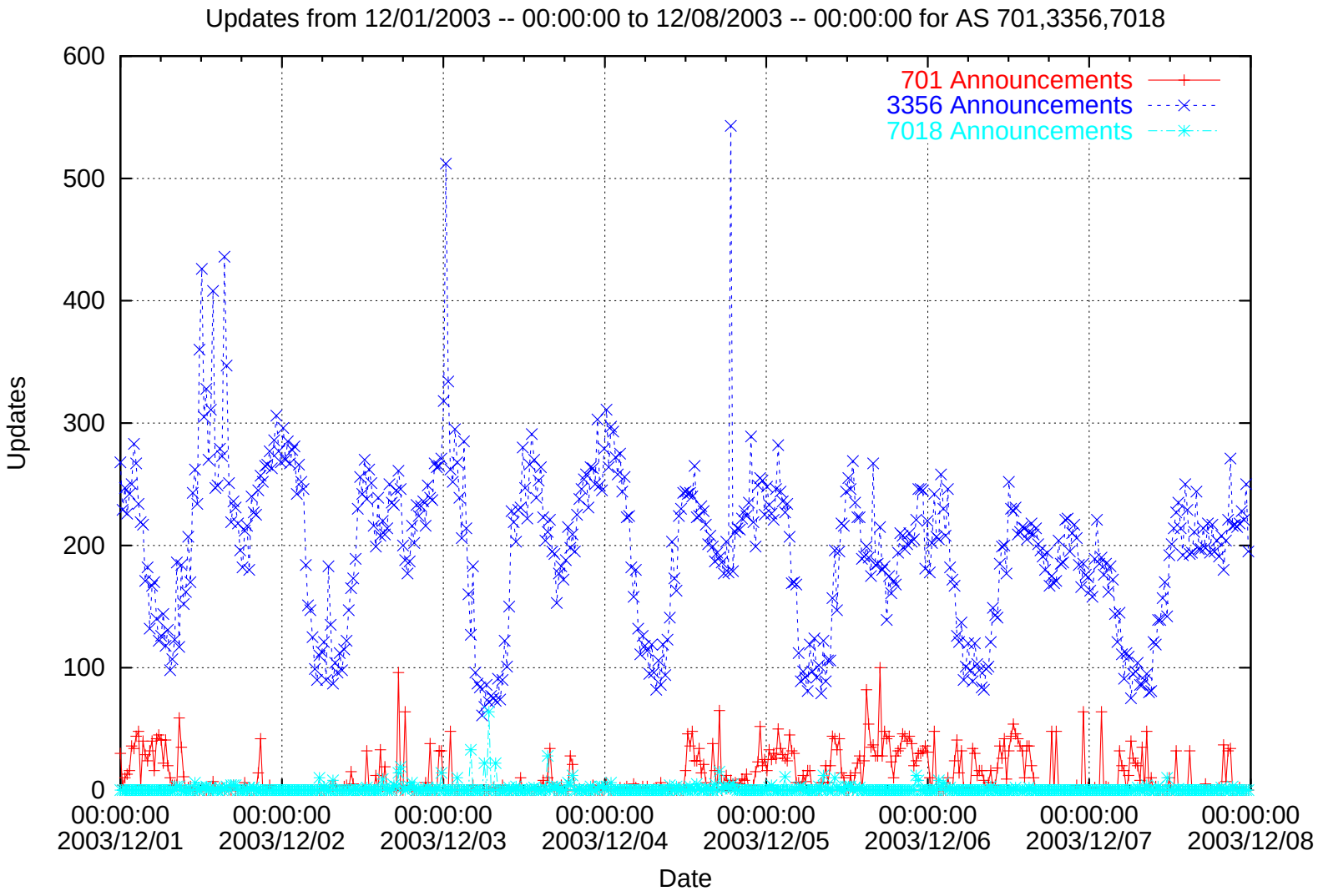
BGP Monitor Overview

http://bgp.lcs.mit.edu/

Table	Start time (EDT)	End time (EDT)	Entries
MIT (AS 3)	Thu Jun 28 16:23:21 2001	Thu Jun 3 11:12:41 2004	209590609
PSG (AS 3130)	Wed May 8 17:25:50 2002	Thu Jun 3 11:01:50 2004	133856658
NYC GBLX (AS 3549)	Mon Jan 27 08:58:13 2003	Thu Jun 3 11:05:12 2004	206068529
London GBLX (AS 3549)	Fri Feb 28 16:40:46 2003	Thu Jun 3 11:09:17 2004	245168742
Aros (AS 6521)	Mon Sep 2 16:58:44 2002	Thu Jun 3 11:02:57 2004	76544186
Nortel (AS 11085)	Mon Aug 19 11:01:09 2002	Thu Jun 3 11:05:38 2004	77446798
VNI (AS 10781)	Mon Aug 12 15:03:25 2002	Thu Jun 3 11:02:18 2004	151731464
PWH (AS 22208)	Fri Jul 12 14:14:00 2002	Thu Jun 3 11:08:06 2004	415609922

- General BGP update summaries by:
  - Time period
  - Origin AS, AS Path
  - Prefix (exact, all subnets, etc.)
- Graph and List Outputs
- Useful for diagnosis in practice
  - www.merit.edu/mail.archives/nanog/2002-11/msg00230.html

Diurnal BGP Update Activity from Level3

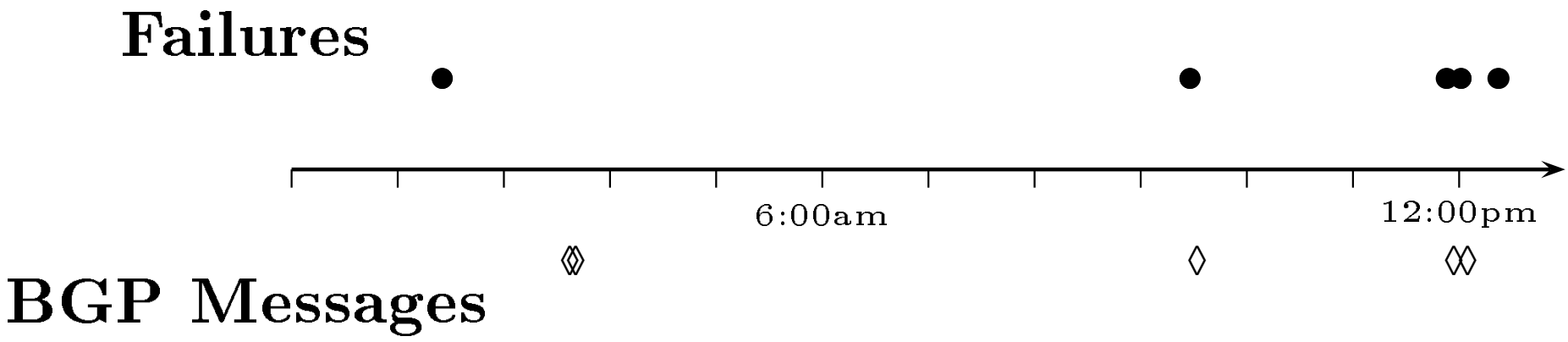


Project 1: Failure Characterization Study

"Measuring the Effects of Internet Path Faults on Reactive Routing"
N. Feamster, D. Andersen, H. Balakrishnan, M.F. Kaashoek
In *Proc. SIGMETRICS 2003*

- **Location:** Where do failures *appear?*
- **Duration:** How long do failures last?
- **Correlation:** Do failures correlate with BGP instability?

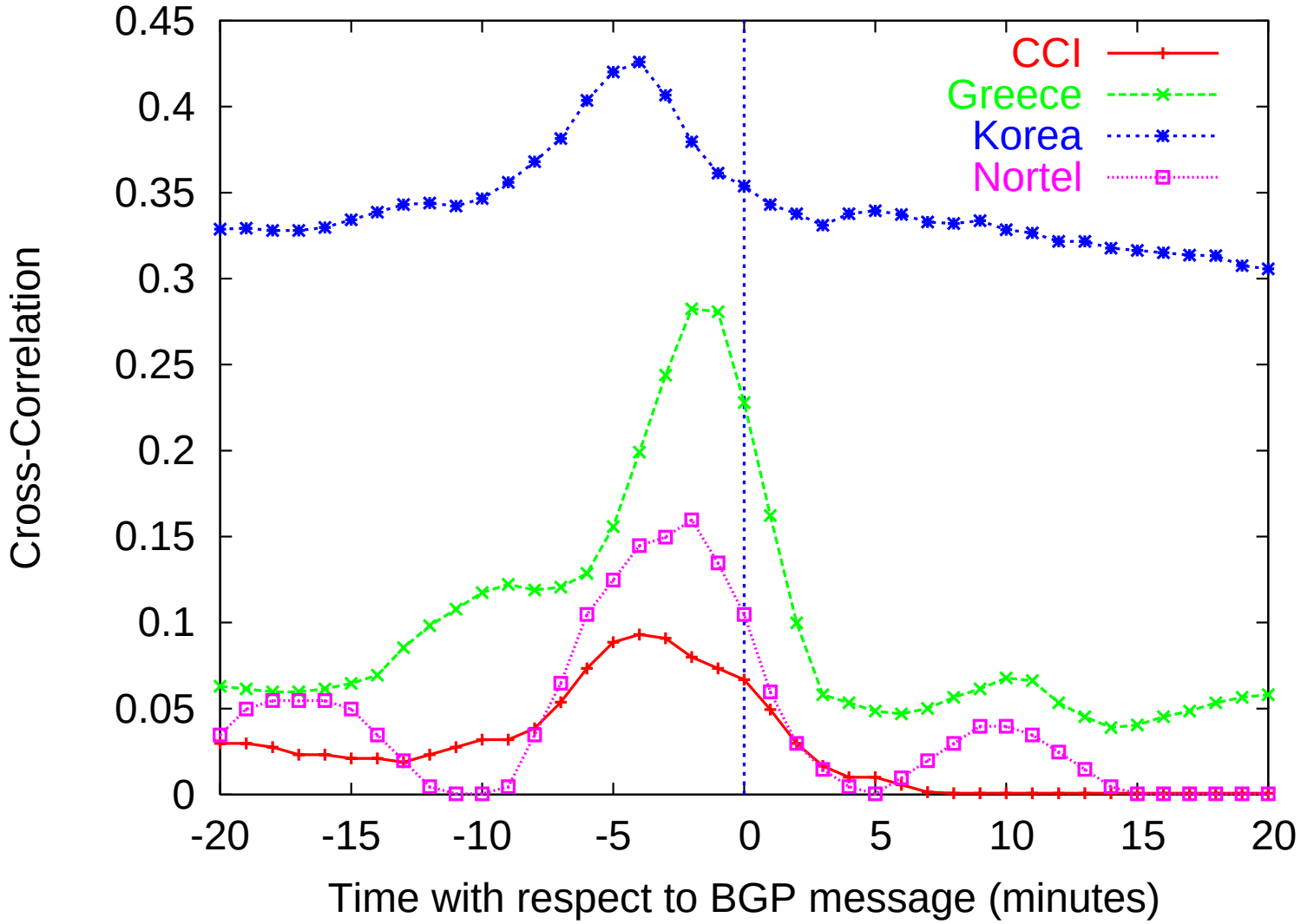
Relating Path Failures and BGP messages



- *Technique 1:* Cross-correlation of time-based signals
- *Technique 2:* Consider a failure and look for BGP (and vice versa)

Do failures correlate with routing instability?

Failures typically occur several minutes before BGP activity.



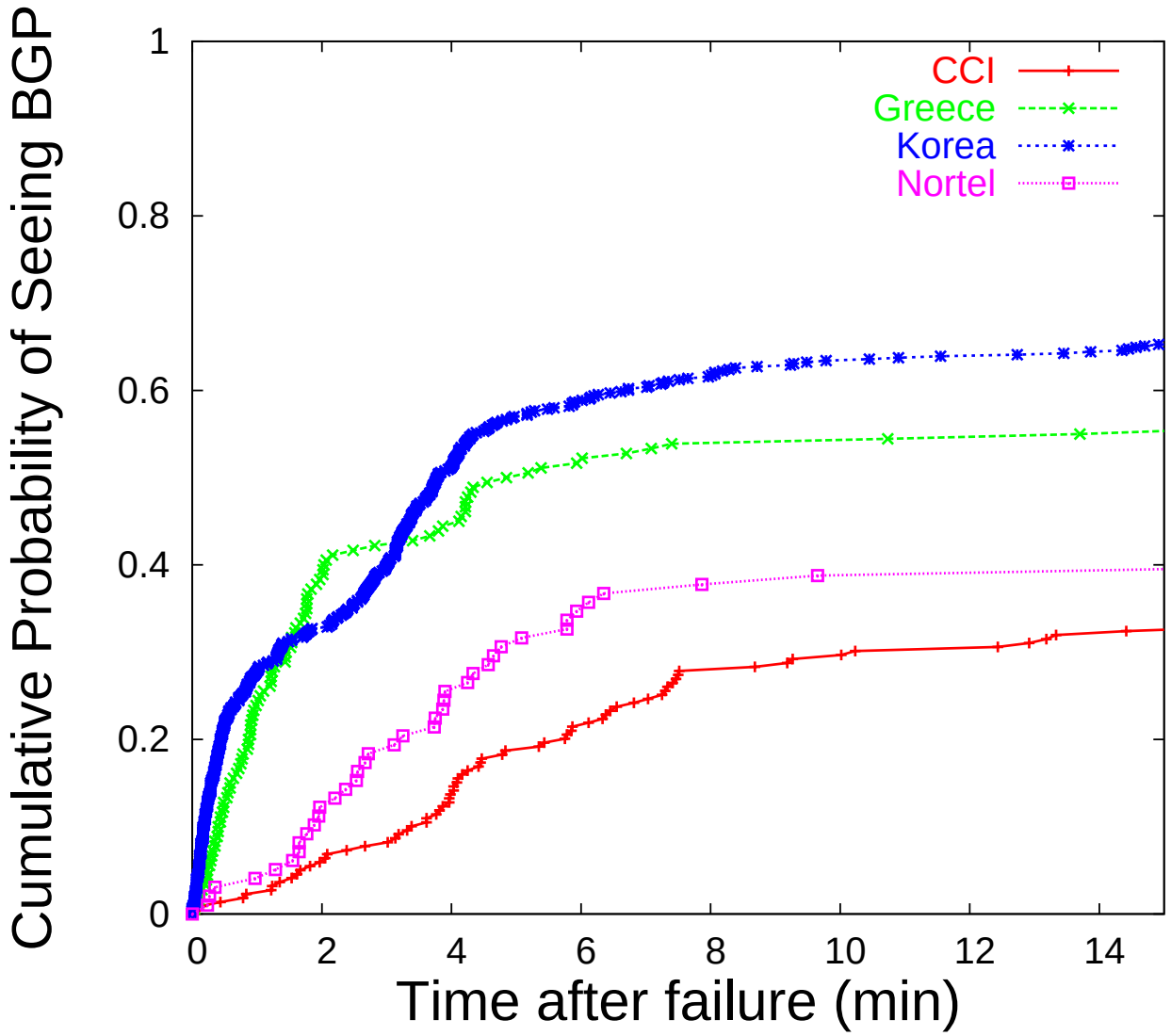
Which failures correlate with instability?

Failures that appear near end hosts are less likely to coincide with BGP instability.

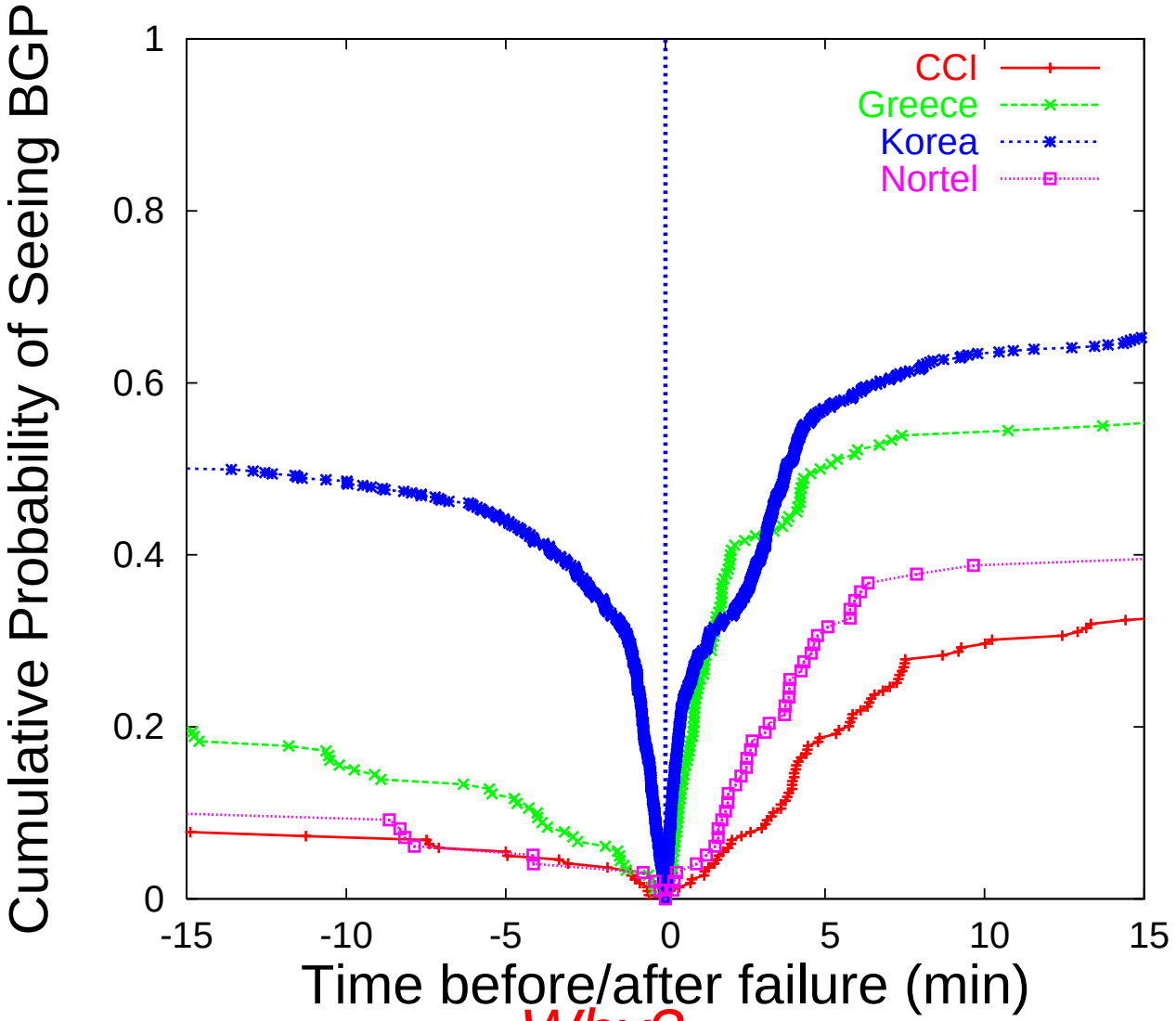
- 60% of failures that appeared at least three hops from an end host coincided with at least one BGP message.
- 22% of failures within one hop of an end host coincided with at least one BGP message.

Just because an ISP is reachable doesn't mean its customers are reachable!

To put it another way...



Surprise: BGP messages precede failures!



Why?

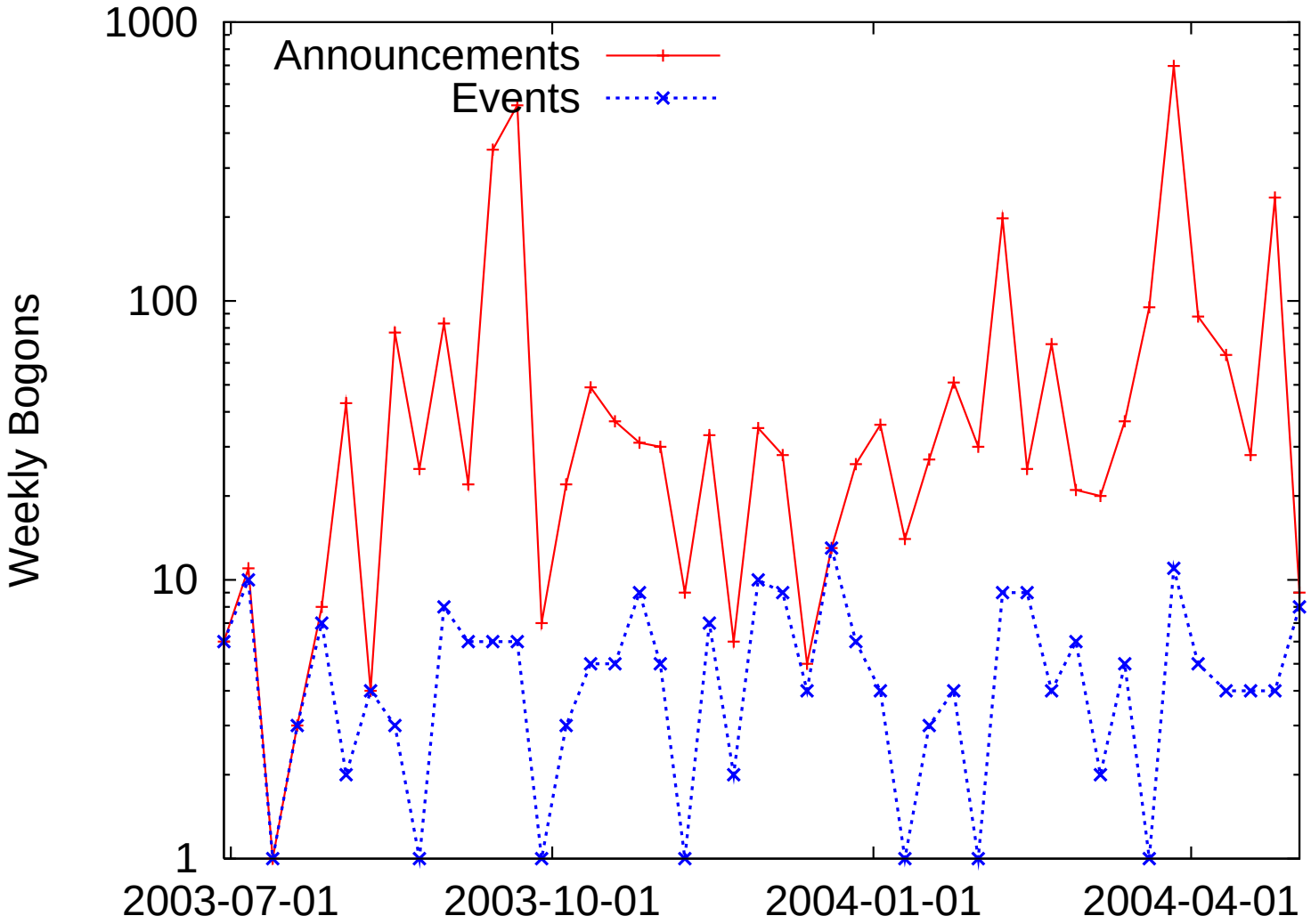
Route flap damping, maintenance, misconfiguration, etc.

Summary

- *Location*
  - Some links experience many path failures, but many experience some failures.
  - Failures appear more often inside ASes than between them.
- *Duration*
  - 90% of failures last less than 15 minutes
  - 70% of failures last less than 5 minutes
- *Correlation*
  - BGP messages coincide with only half of the failures that reactive routing could potentially avoid.
  - When BGP messages and failures coincide, BGP messages most often follow failures by 4 minutes.
  - BGP sometimes precedes failures.

Project 2: Invalid Prefix Advertisement Study

BGP route advertisements from July 2003 to May 2004.
http://bgp.lcs.mit.edu/bogons.cgi

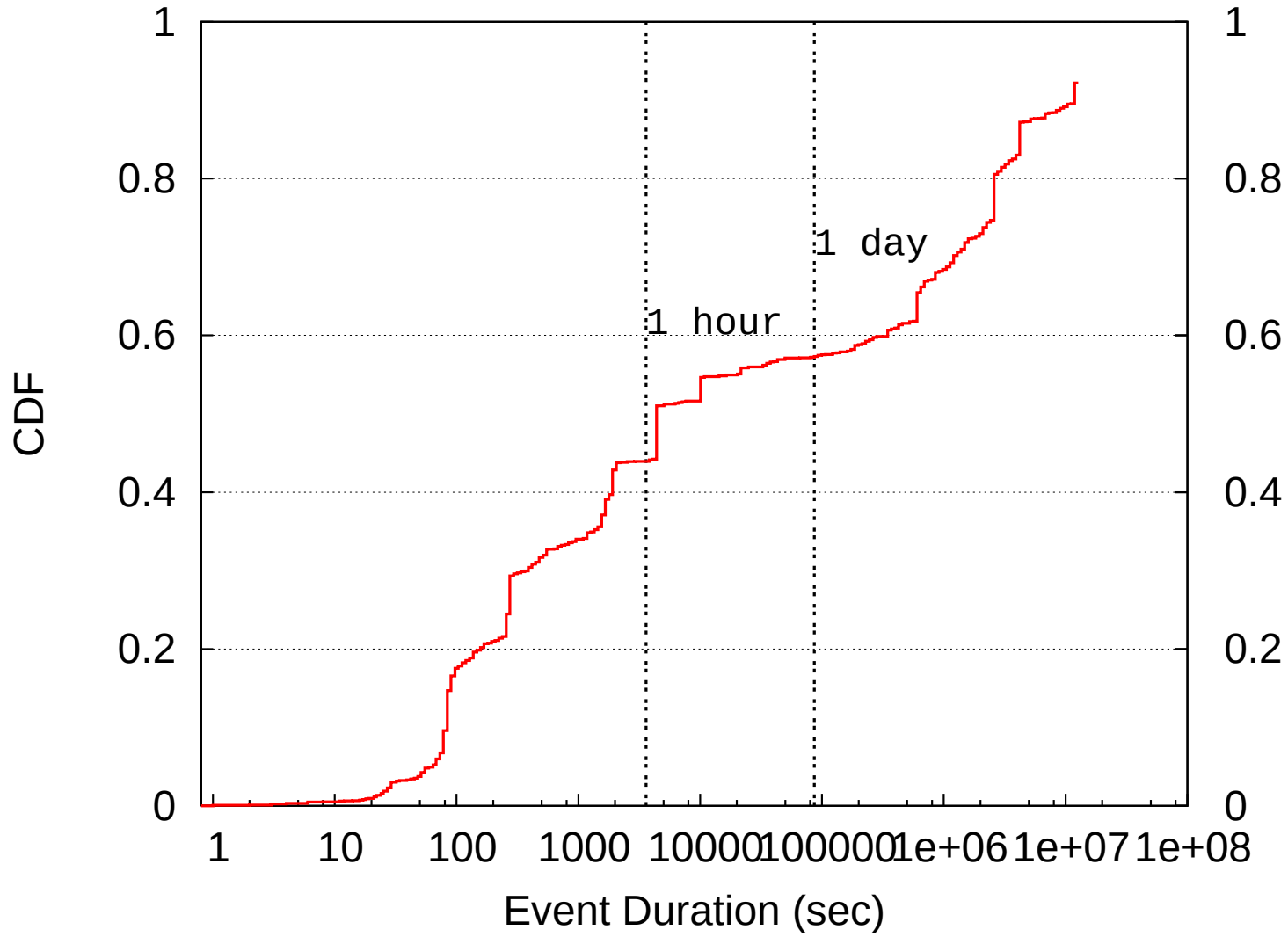


What Type of Prefixes Are Leaked?

Bogon Space	*Announce Ev.*	*Operational Ev.*	*Monitors*	*Origin ASes*
172.16.0.0/12	2652	90	1	4
0.0.0.0/7	239	80	8	69
192.0.2.0/24	7	23	2	3
10.0.0.0/8	26	14	2	3
96.0.0.0/3	90	11	7	7
189.0.0.0/8	34	10	6	2
169.254.0.0/16	5	4	3	4

- Many route leaks from private address space.
  - Large number of offending origin ASes
  - Many 0.0.0.0/7 widely visible
  - 0.0.0.0/8 often filtered, but not 0.0.0.0/7
- Simple, static filters could make a big difference.

How Long Do These Routes Persist?



Half of bogus route events persist for longer than an hour.