



Internet Initiative Japan

# Happy Packets: Some Initial Results

APNIC - Nadi, Fiji 2004.09.01

Randy Bush <randy@psg.com>

Timothy G. Griffin <tim.griffin@intel.com>

Z. Morley Mao <zmao@eecs.umich.edu>

Eric Purpus <epurpus@cs.uoregon.edu>

Dan Stutsbach <dstutsba@cs.uoregon.edu>

<<http://psg.com/~randy/040901.apnic-happy.pdf>>

# Central Question

- What is the relationship between **control plane** instability and **data plane** instability?
- Related Questions:
  - Is the quantity of BGP updates good or bad?
  - Who wants to see zero BGP updates?

# Internet Weather

We frequently hear comments such as

- Internet routing is fragile, collapsing, ...,
- BGP is broken or is not working well,
- Day X was a bad routing day on the internet,
- Change X to protocol Y will improve routing,
- Etc.

And we often measure routing dynamics and say that some measurement is better or worse than another

# Internet [Routing] Instability

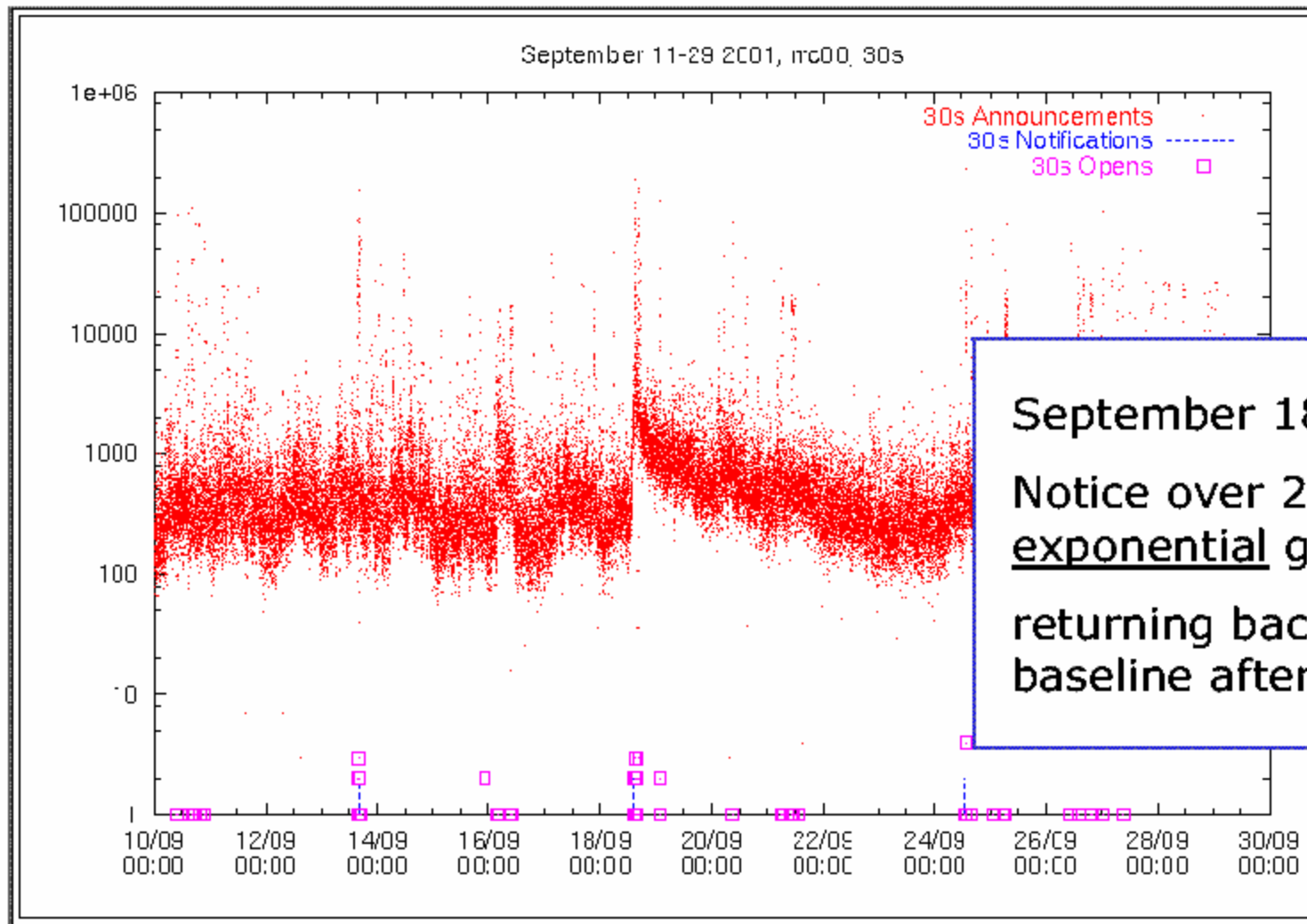
- We are told that a lot of BGP updates is equated with internet instability
- "There are too many BGP updates, so BGP must be broken."

# White Blood Cells

- Perhaps BGP announcements are like white blood cells
- Their presence may signal a problem
- But they are often part of the cure, not necessarily part of the problem

# A view on content of the same messages

Number of prefix announcements in 30 sec intervals



September 18:

Notice over 20-fold  
exponential growth

returning back to  
baseline after 4 days!

renesys

# Big Events

- The Renesys folk and others are looking at big events
- We are at looking at single announcements
- So our work does not contradict Renesys, but it does suggest we consider some of the assumptions (see Lan Wang et alia)
- And we are measuring data plane performance waiting for the next big event

# Routing Quality

- But what is *good* routing? How can we say one measurement shows routing is better than another unless we have **metrics** for routing quality?
- We often work on the assumption that number of prefixes, speed or completeness of convergence, etc. are measures of routing quality



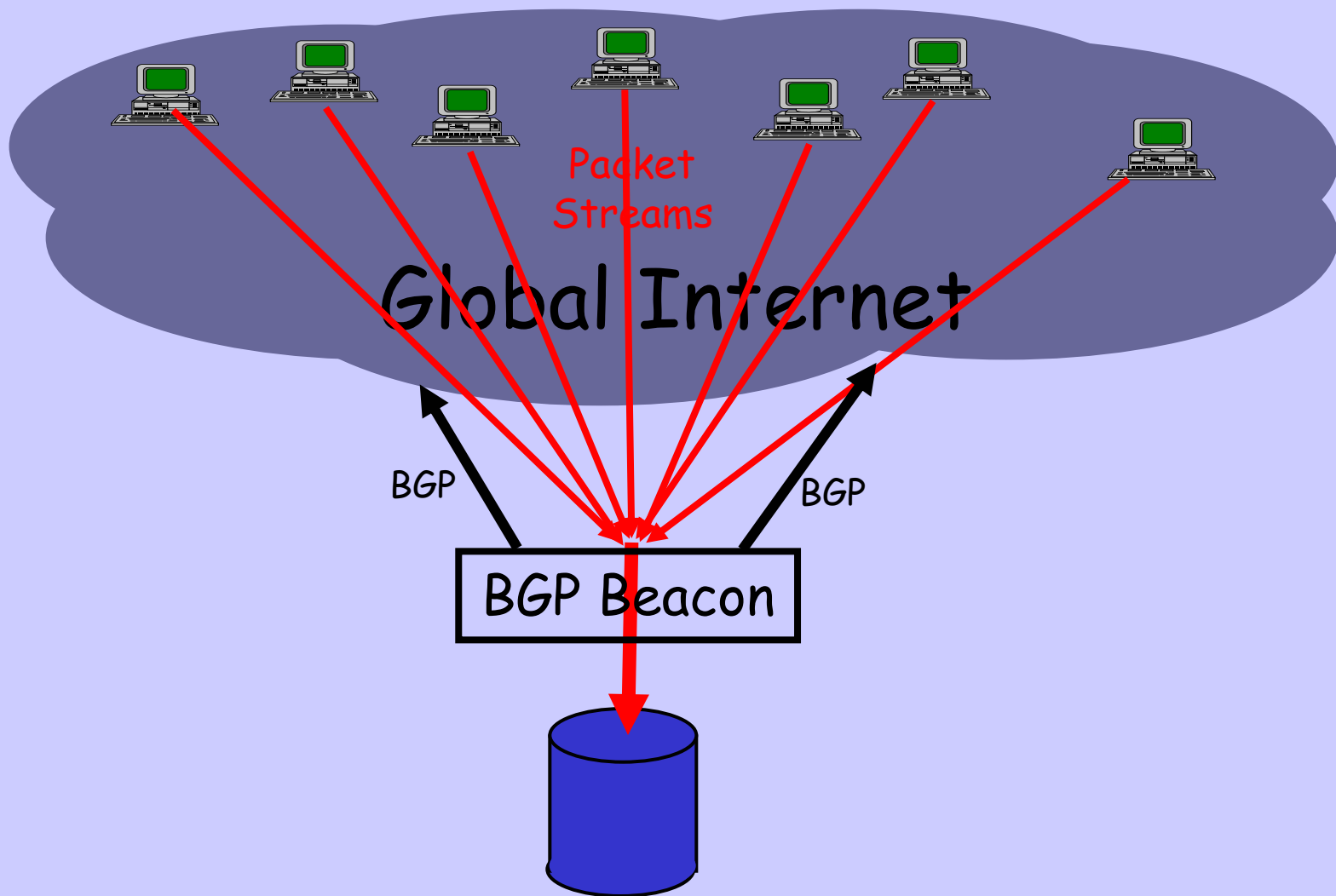
# Happy Packets

- The measure which counts is whether the users' packets reach their destination
- If the users' packets are happy, the routing system, and other components, are doing their job
- We call these *Happy Packets*
- There are well-known metrics for the data plane, Delay, Drop, Jitter, and Reordering
- So we set out to measure Control Plane quality by measuring the Data Plane

# Router Scaling

- While data plane performance is the goal, we can't have routers falling over processing chatty BGP
- But, as long as network BGP growth increases load on the routers below Moore's law, it is not clear we are in danger

# Experimental Setup



# BGP Beacon

BGP Beacon:

A prefix that is Announced and Withdrawn at well-known times



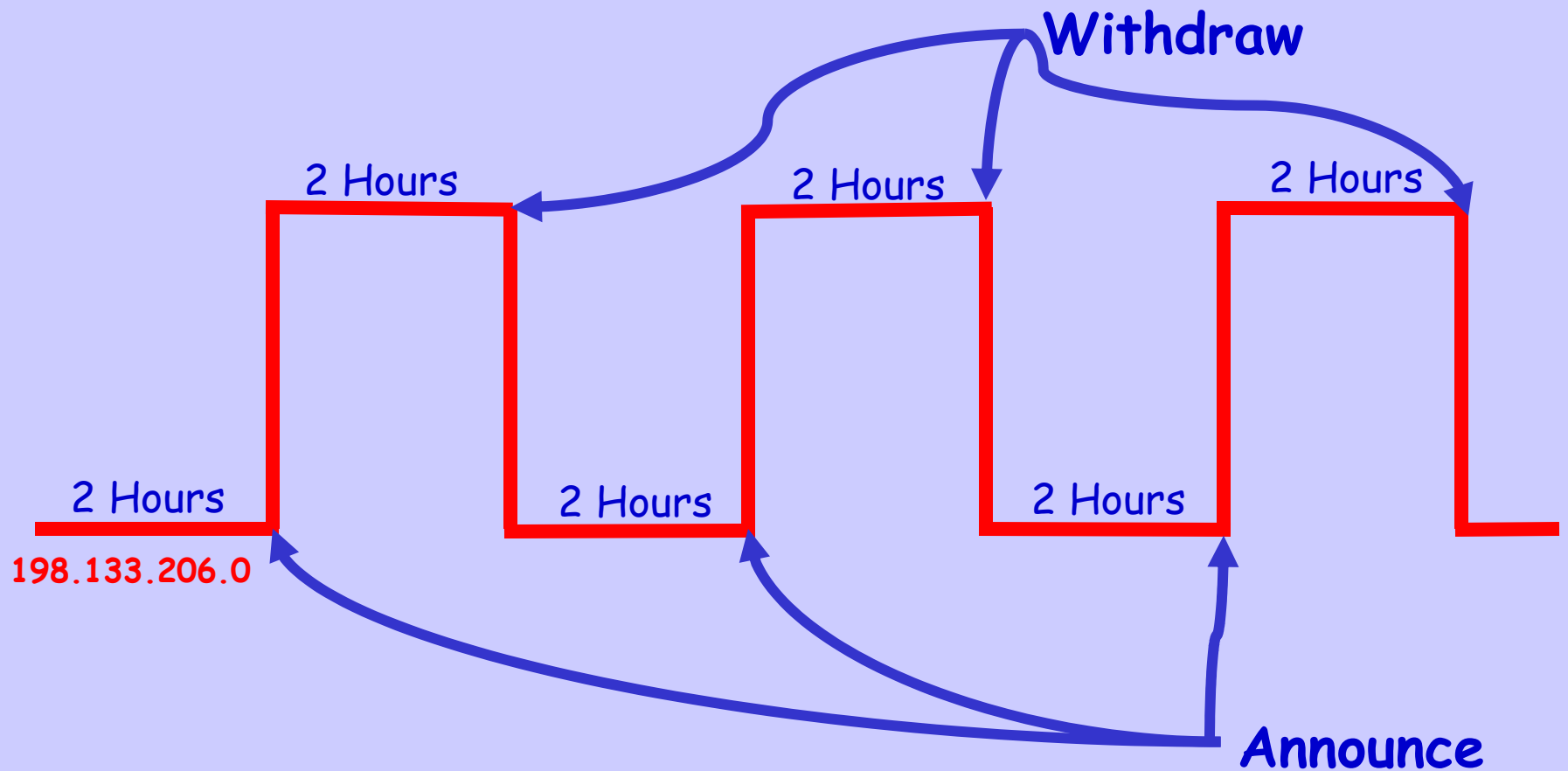
BGP Beacon

192.83.230.0

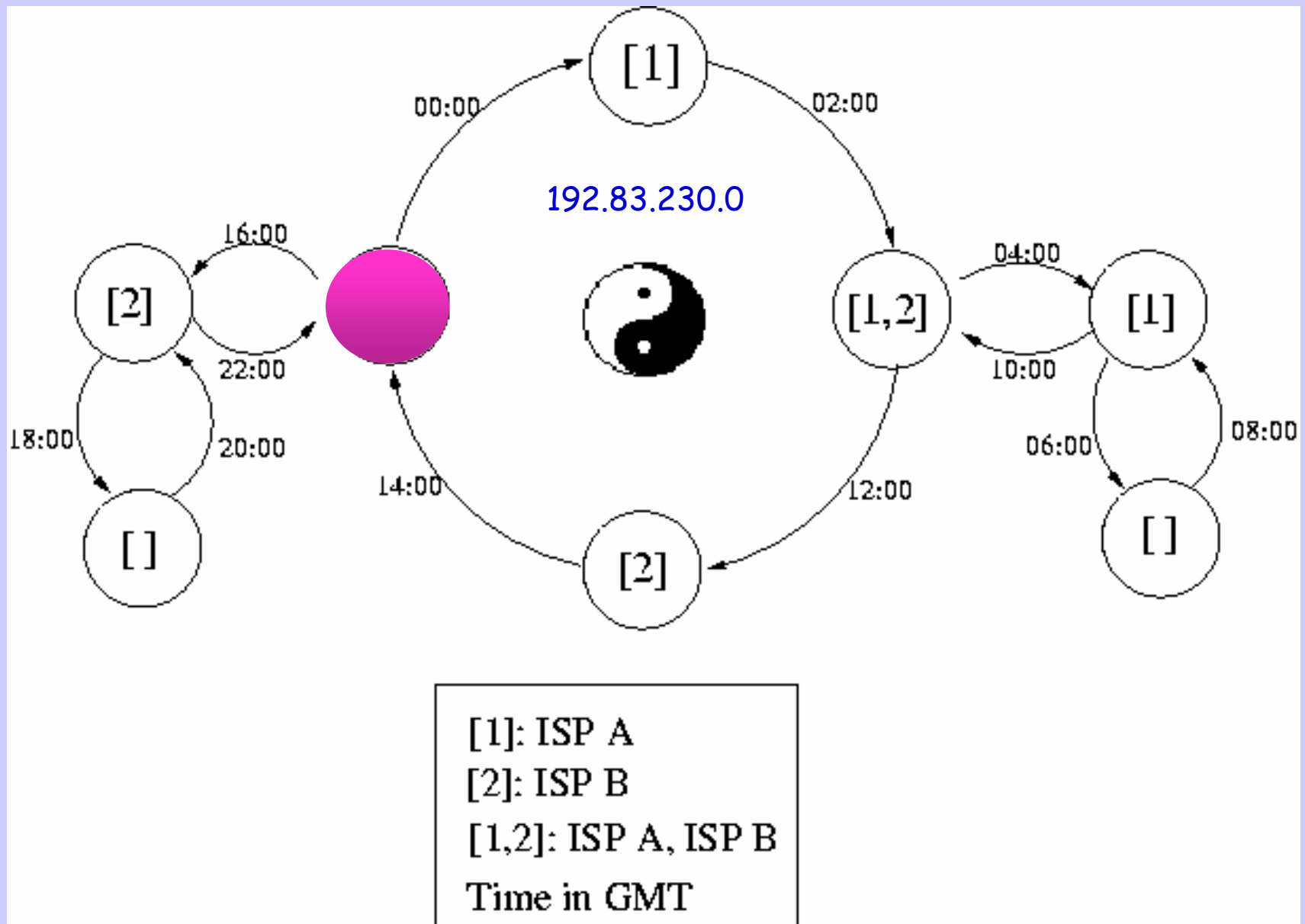
Global  
Internet

# BGP Beacons

## Announce & Withdraw



# Multi-Homed Beacon

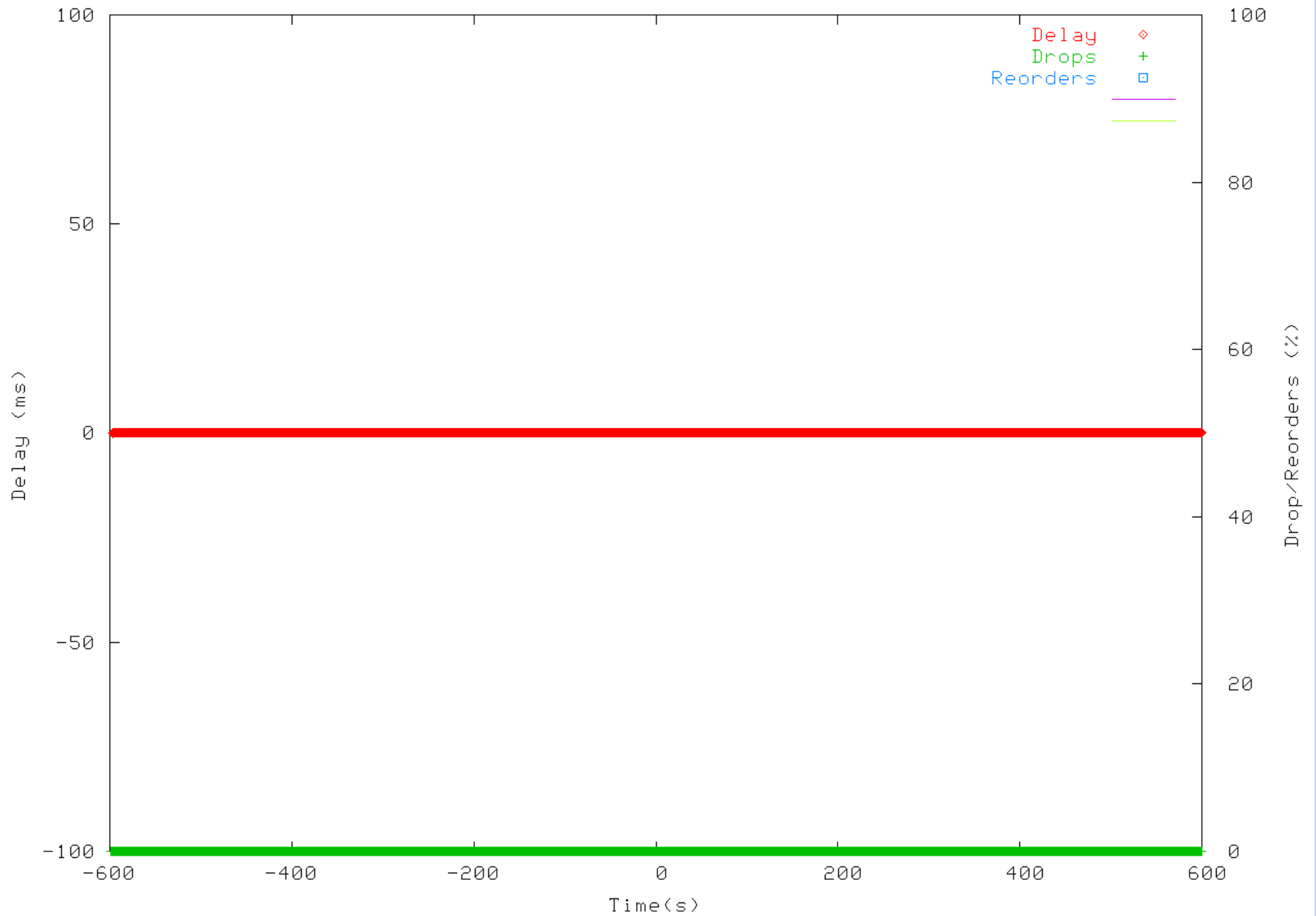


# Packet Stream Sources on PlanetLab (and RON)



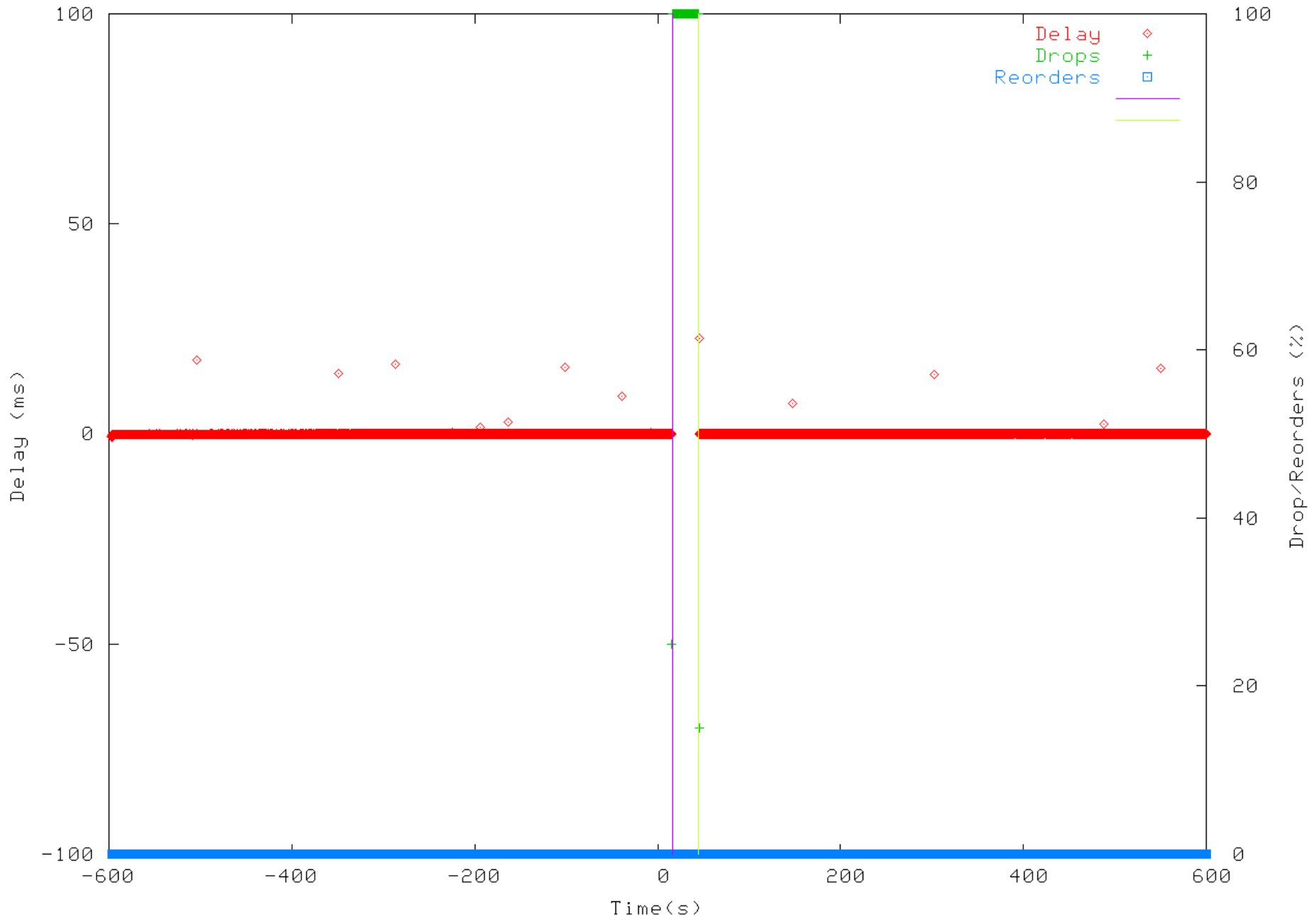
370 nodes at 155 sites  
Biased toward R&E Networks  
<<http://planet-lab.org>>

Transition: ispA,ispB -> ispB at 2003-11-6 12:00  
from probe: 128.95.219.192 (128.95.219.192)

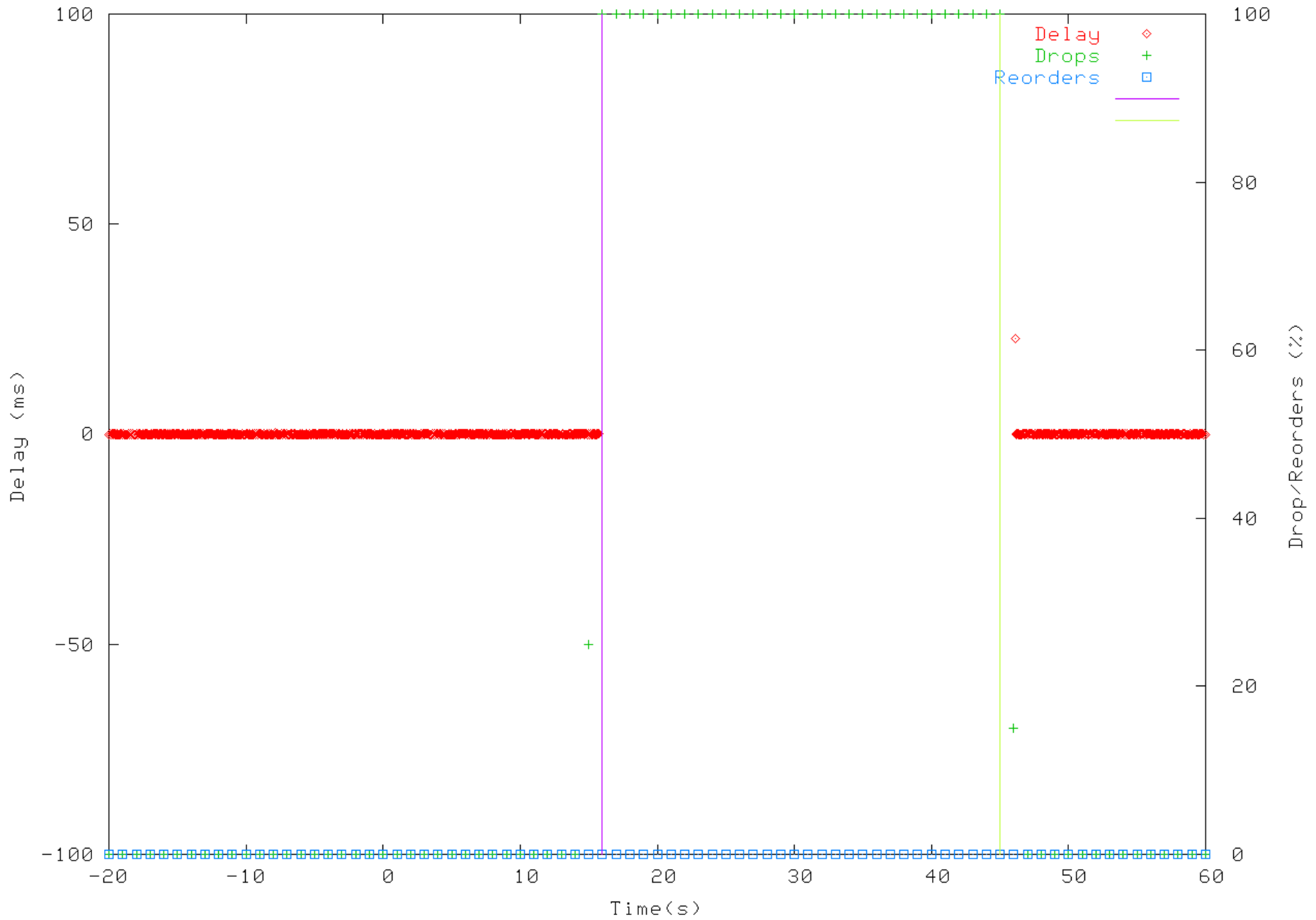




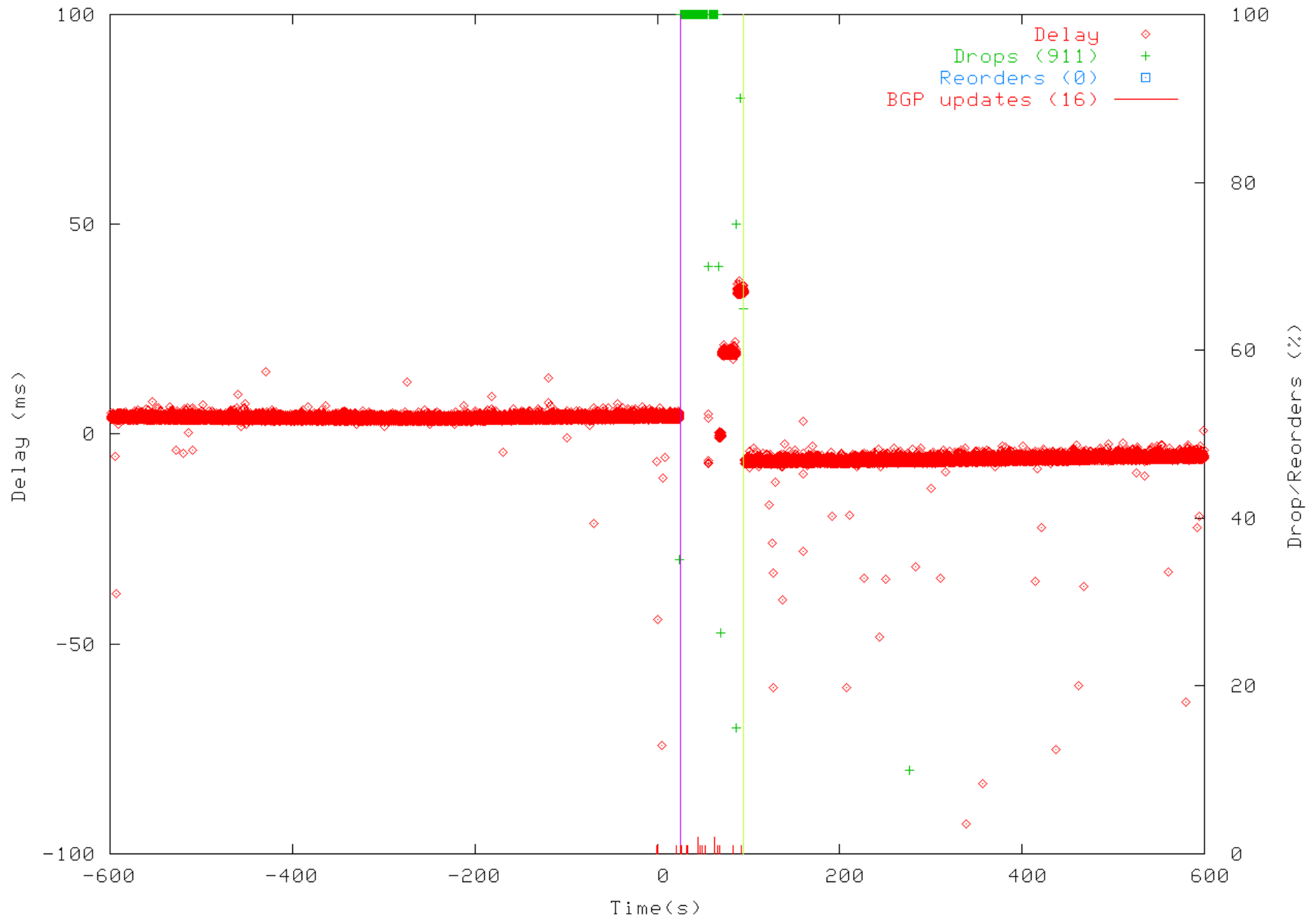
Transition: ispA,ispB -> ispA at 2003-11-6 4:00  
from probe: 128.95.219.192 (128.95.219.192)



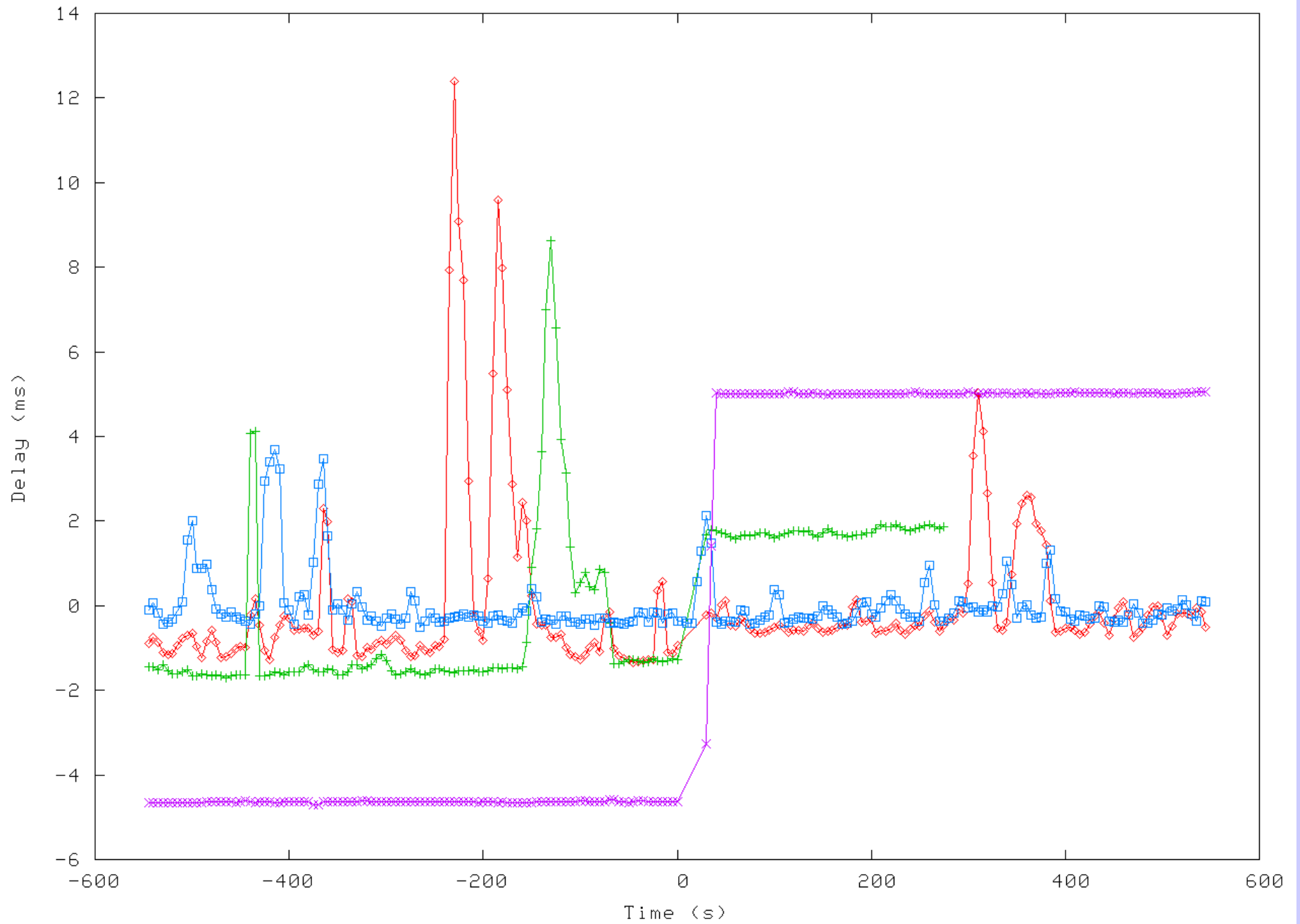
Transition: ispA,ispB -> ispA at 2003-11-6 4:00  
from probe: 128.95.219.192 (128.95.219.192)



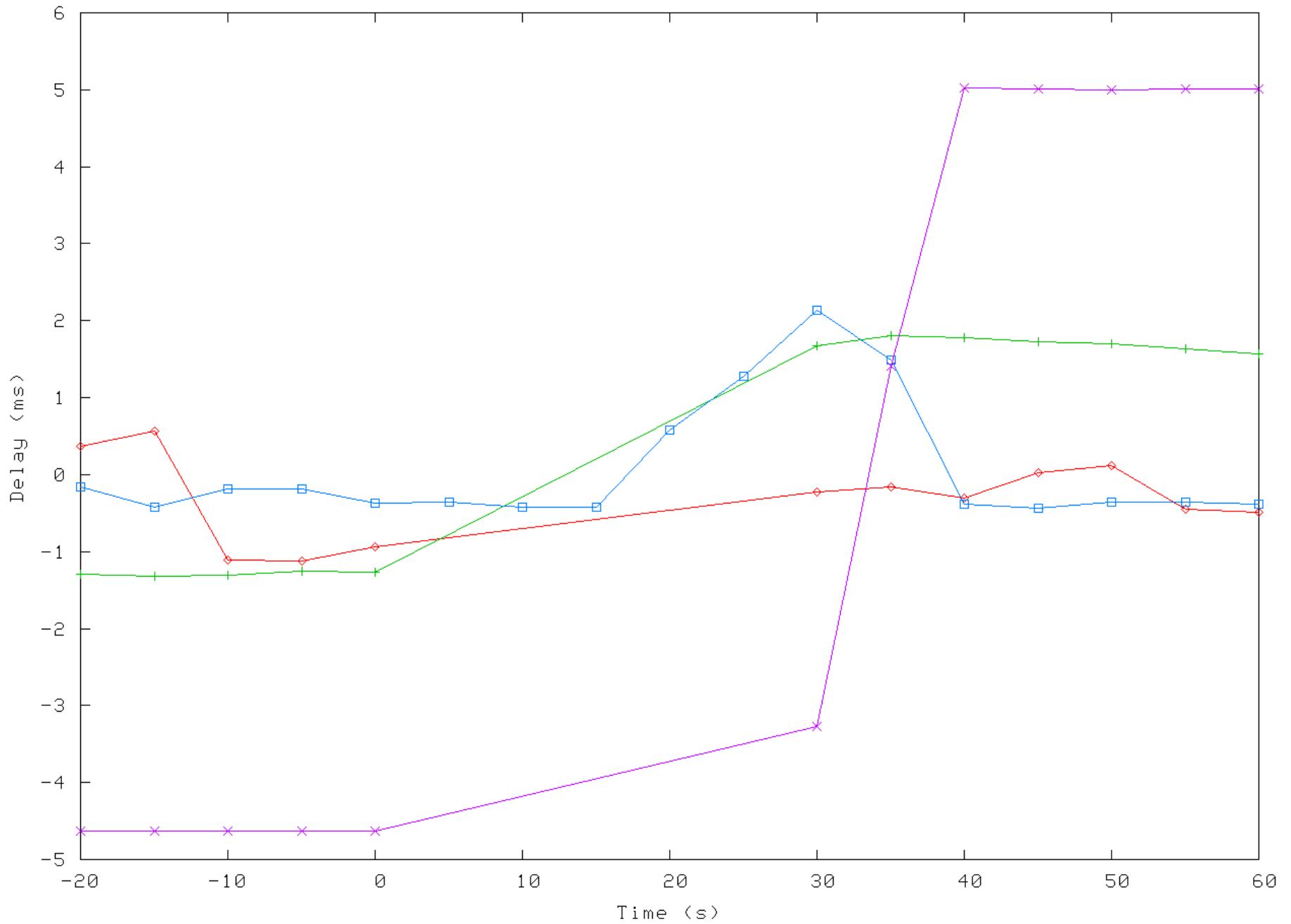
Transition: ispA,ispB -> ispA at 2003-11-2 4:00  
from probe: 212.192.241.155 (212.192.241.155)



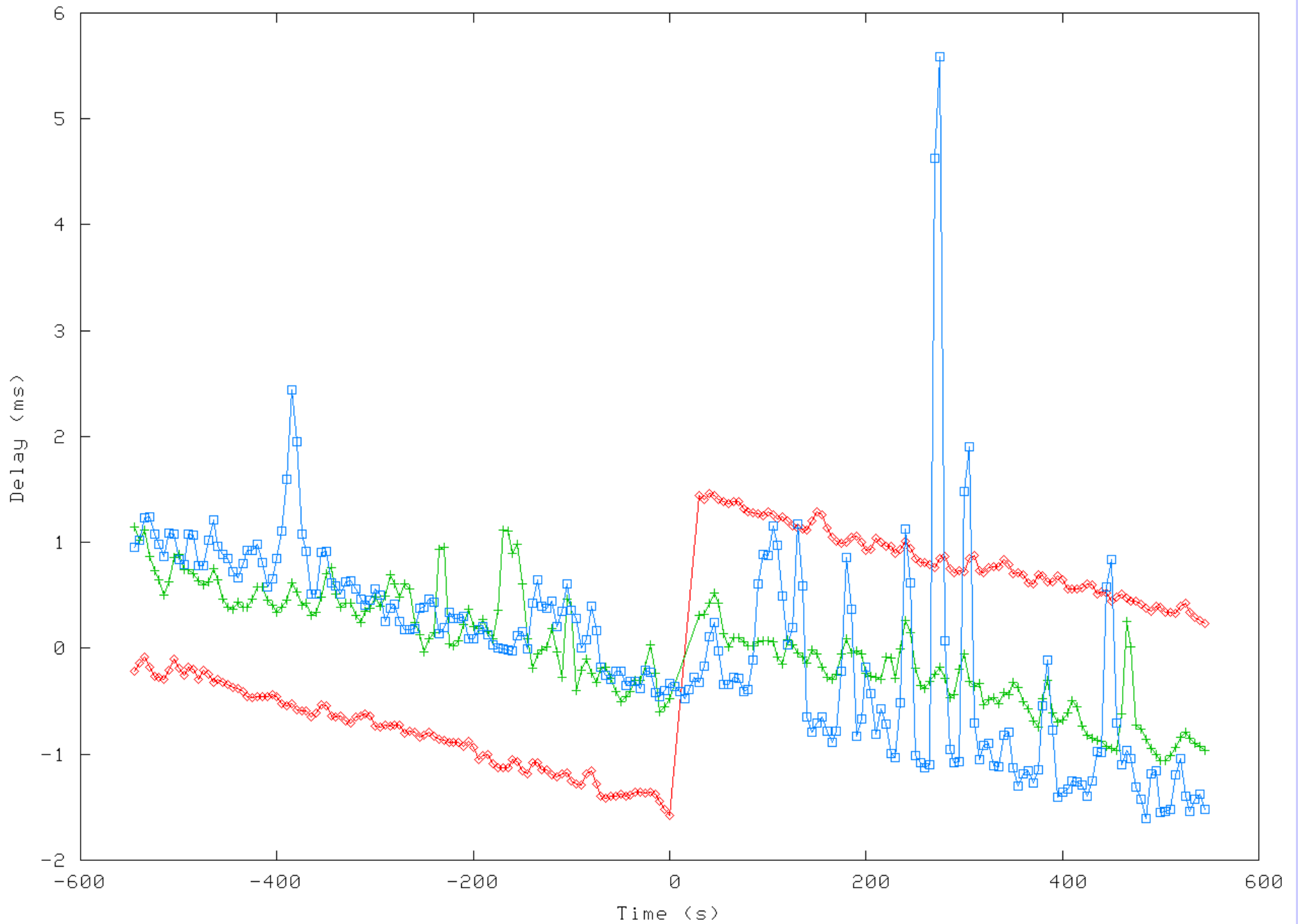
Transition: ispA,ispB -> ispA at 2003-12-20 04:00



Transition: ispA,ispB -> ispA at 2003-12-20 04:00



Transition: ispA,ispB -> ispA at 2004-01-01 04:00



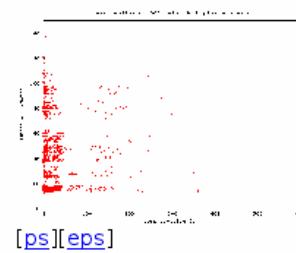
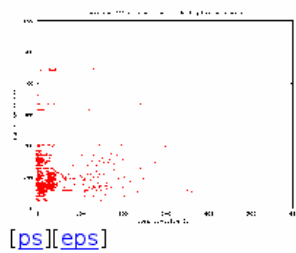
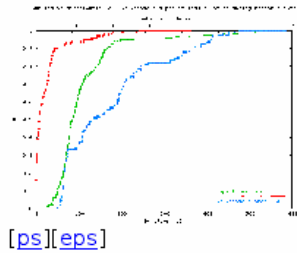
Event

CDF

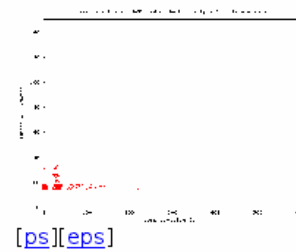
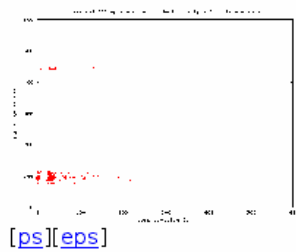
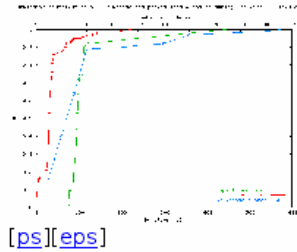
Loss duration vs BGP duration

Loss duration vs BGP count

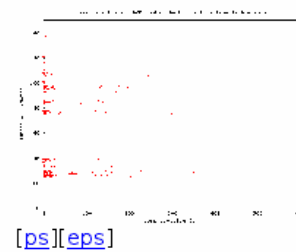
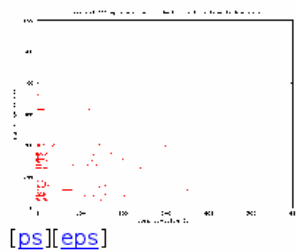
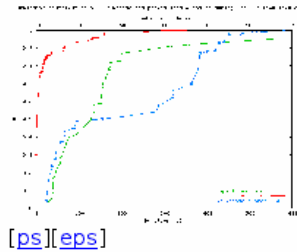
All



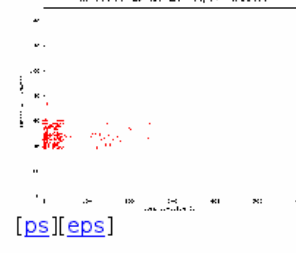
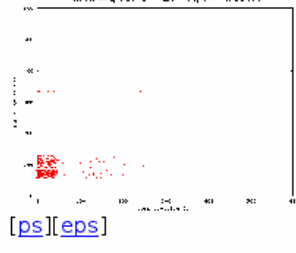
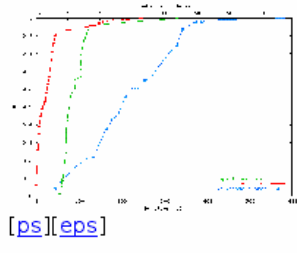
ispA,ispB -> ispA



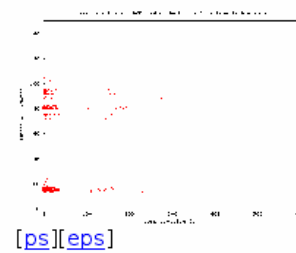
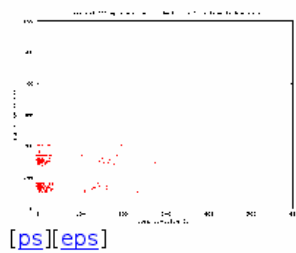
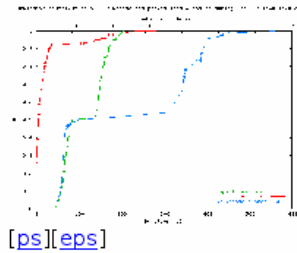
ispA -> ispA,ispB



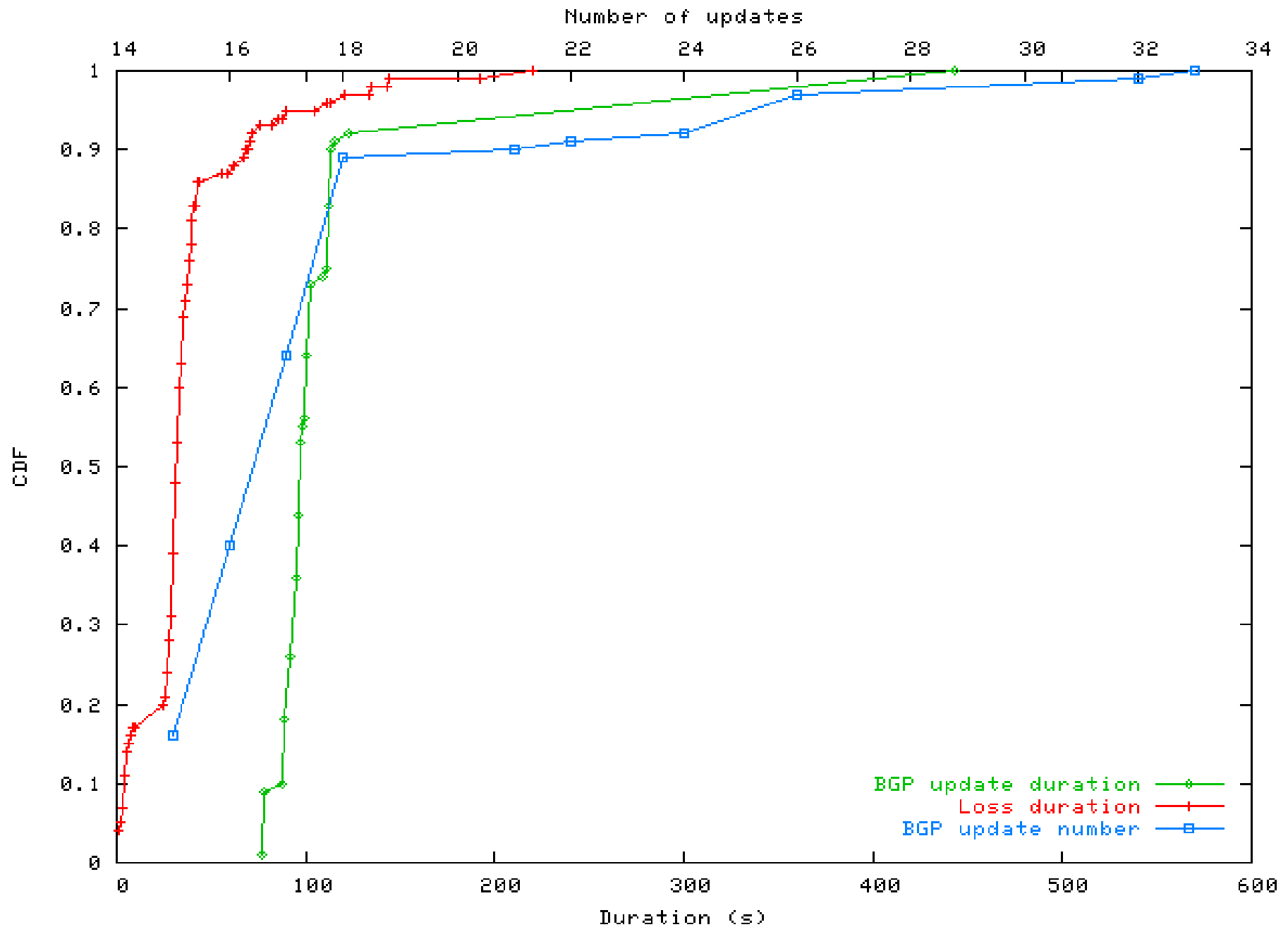
ispA,ispB -> ispB



ispB -> ispA,ispB

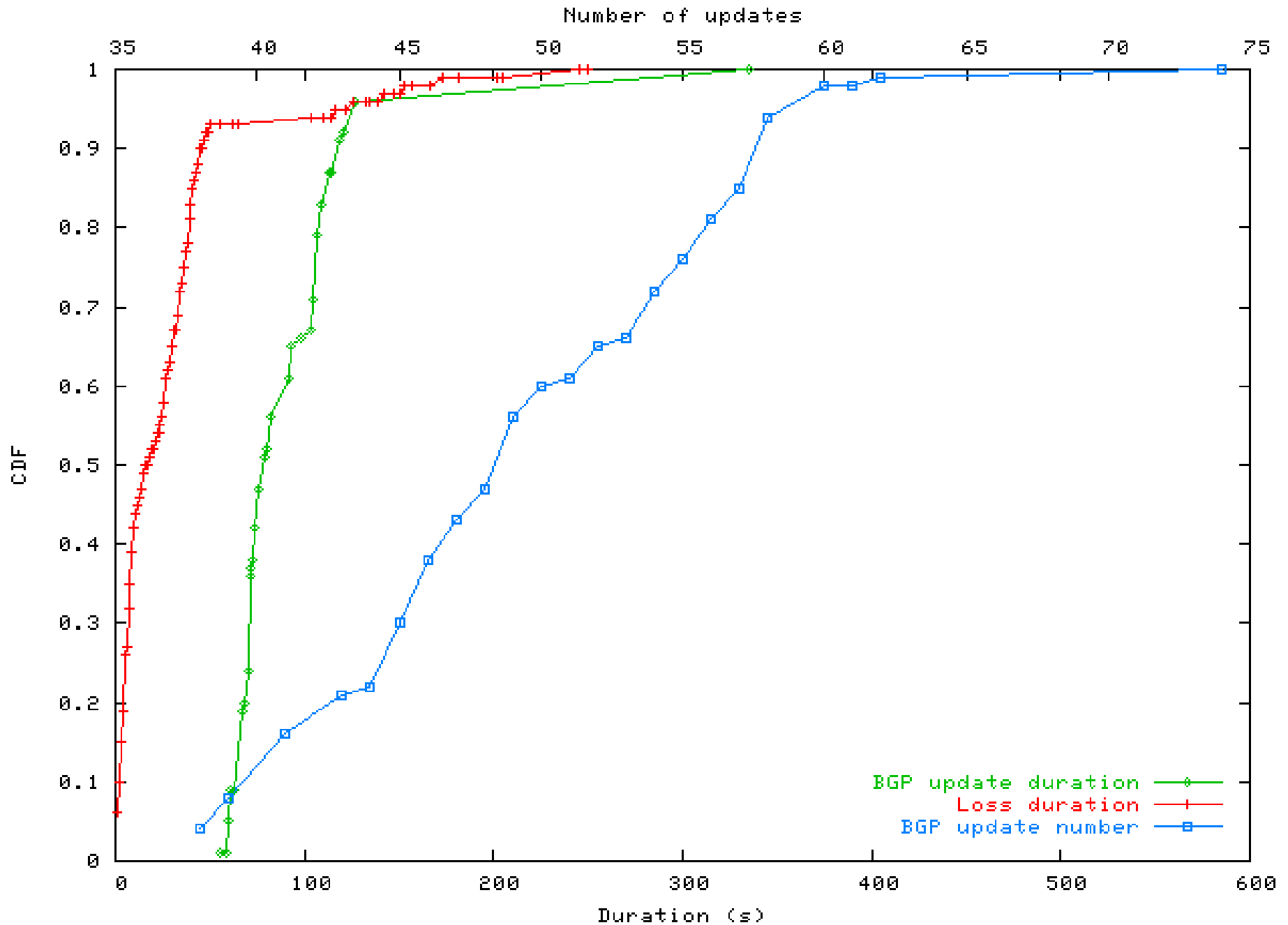


Cumulative distribution of BGP update and packet loss duration during ispA,ispB -> ispA ev

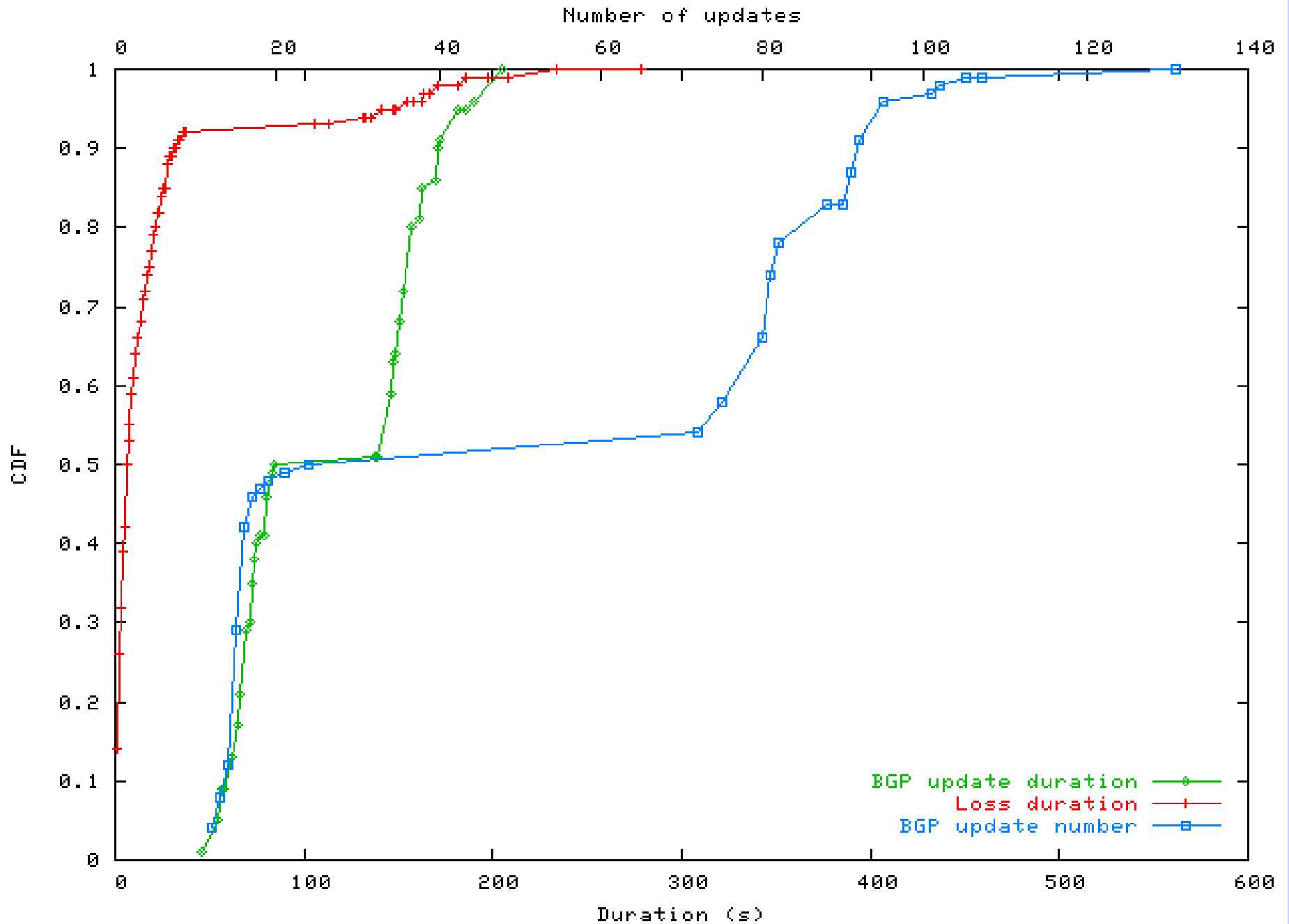




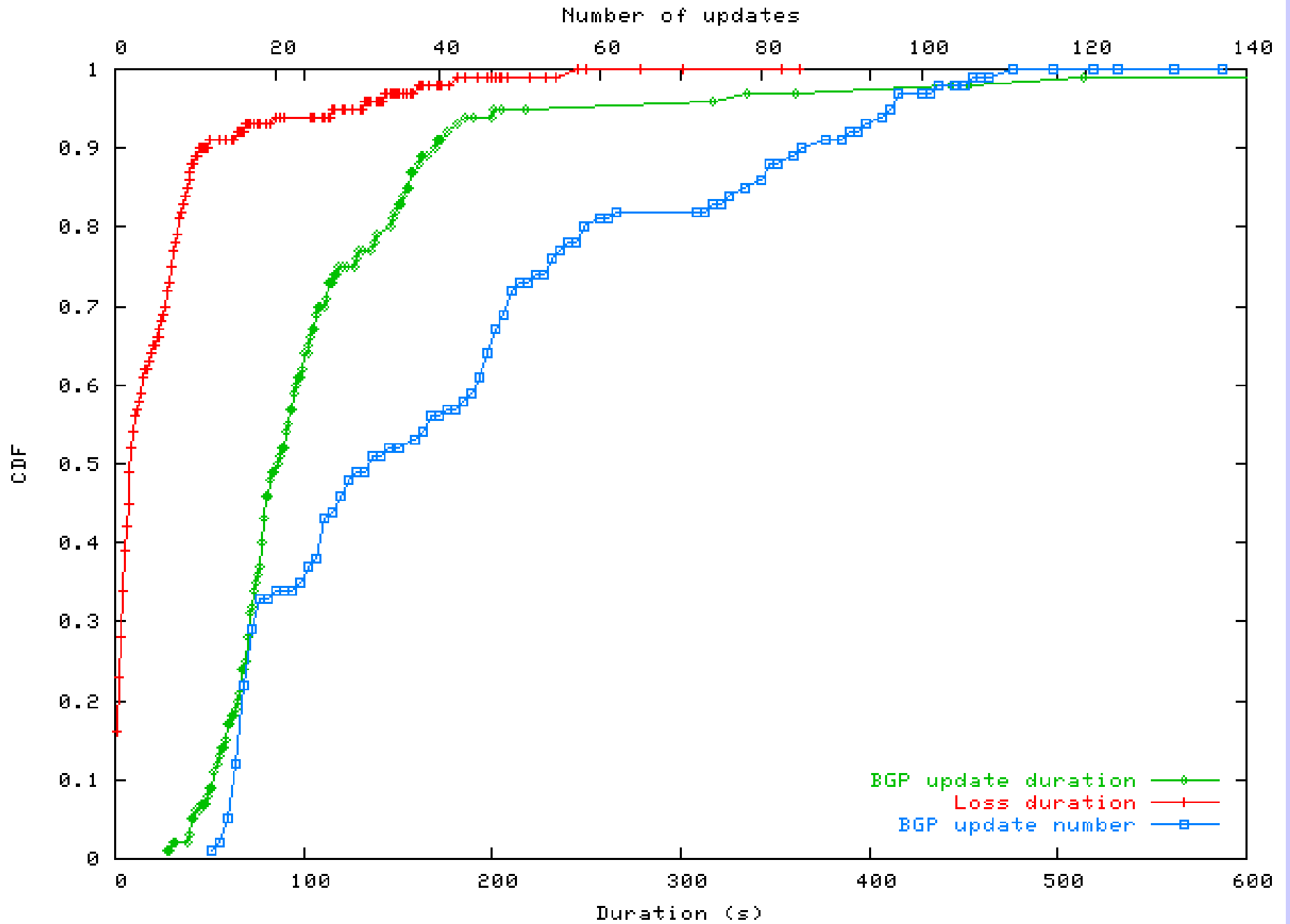
Cumulative distribution of BGP update and packet loss duration during ispA,ispB -> ispB ev



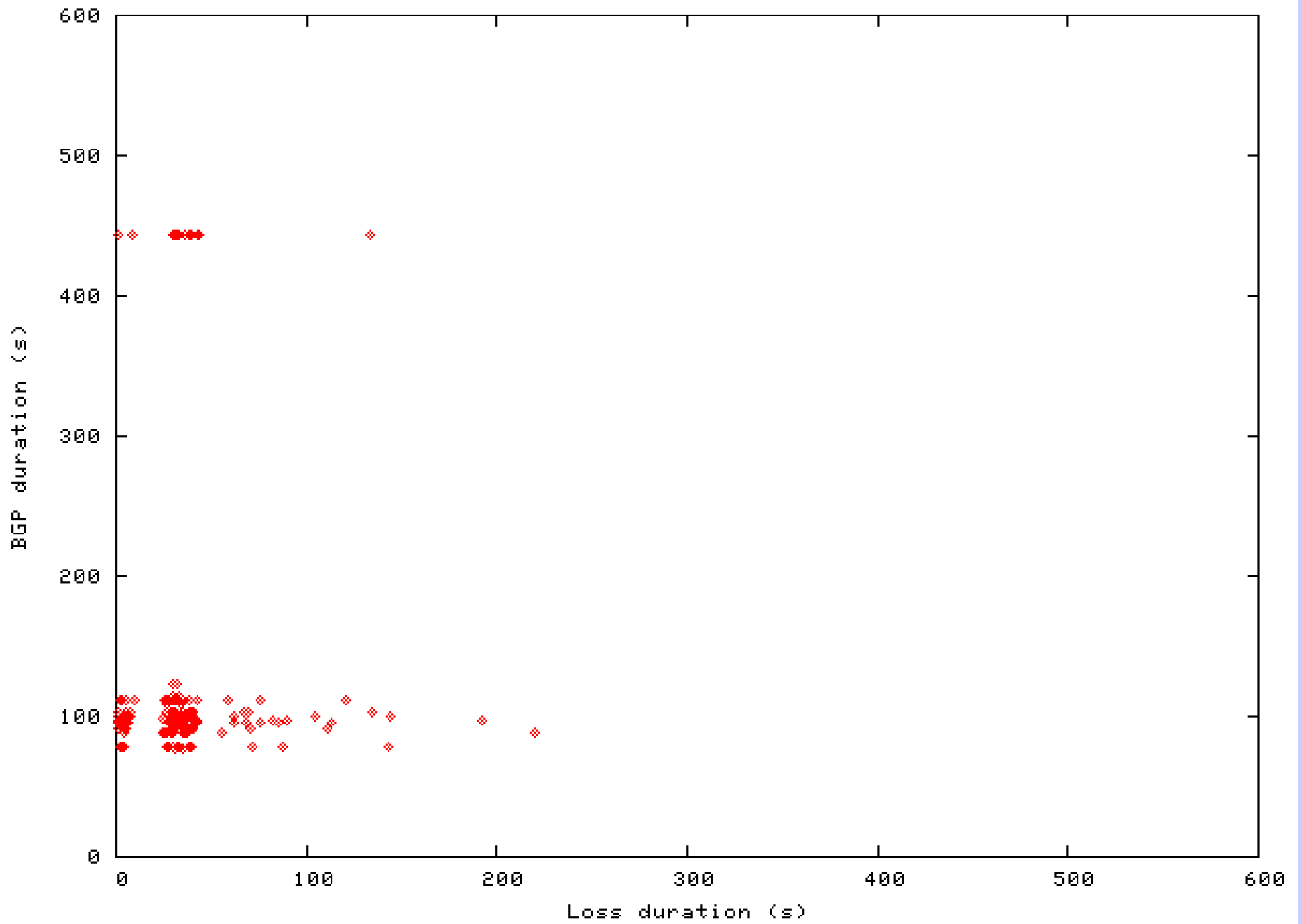
Cumulative distribution of BGP update and packet loss duration during ispB -> ispA, ispB ev



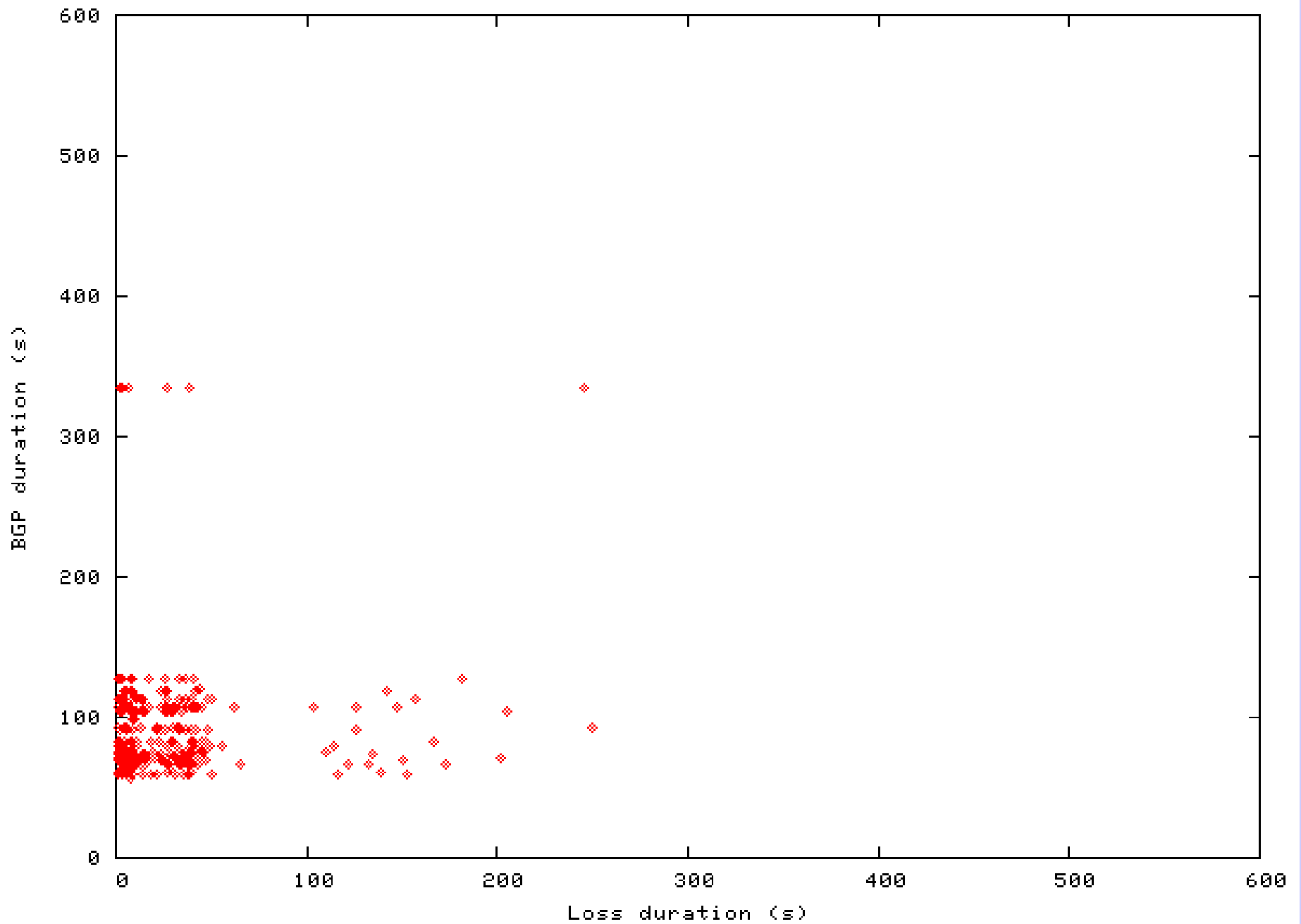
Cumulative distribution of BGP update and packet loss duration during beacon events



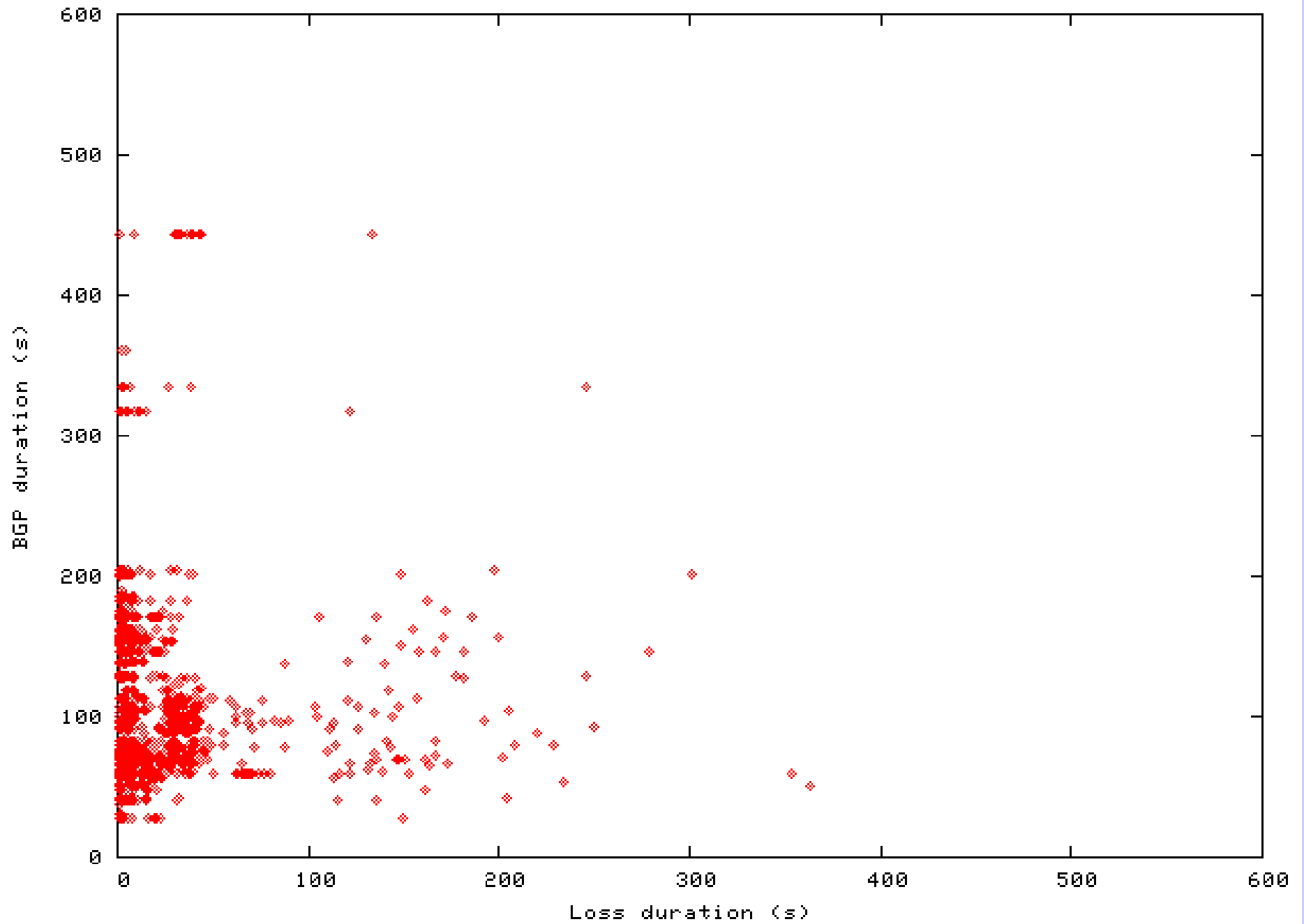
Loss and BGP update duration during ispA,ispB -> ispA events



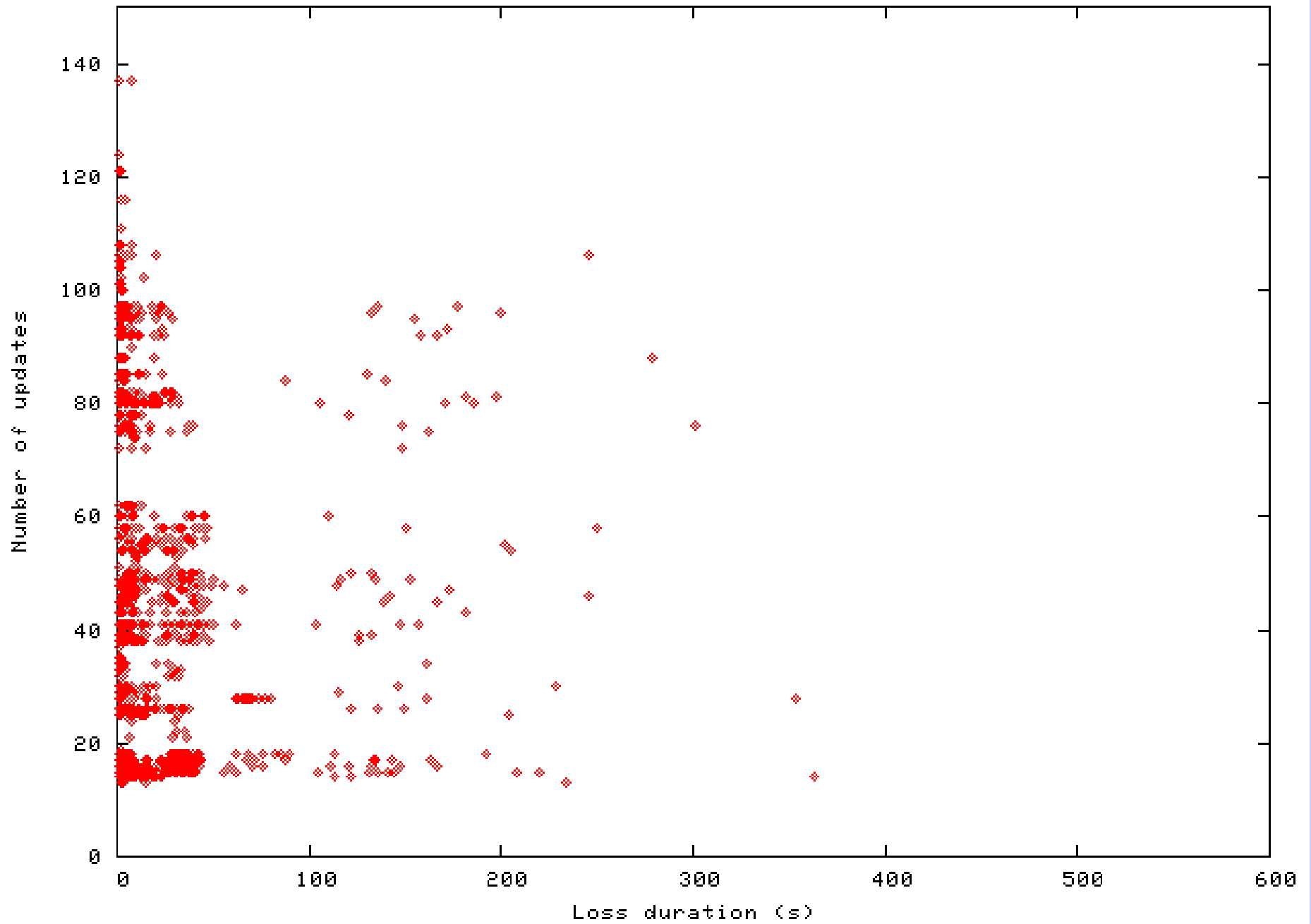
Loss and BGP update duration during ispA,ispB -> ispB events



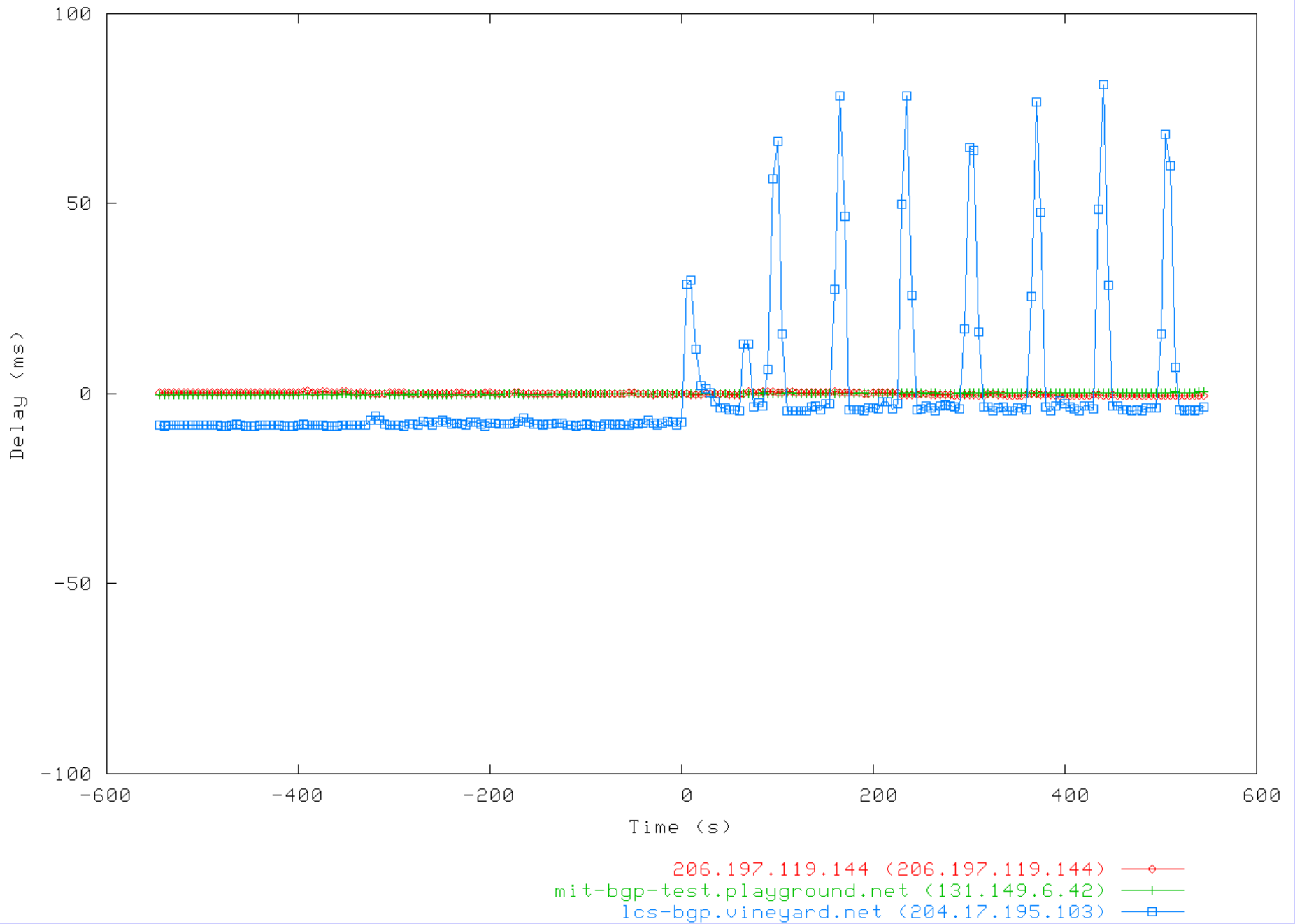
Loss and BGP update duration during beacon events



Loss duration and BGP number during beacon events

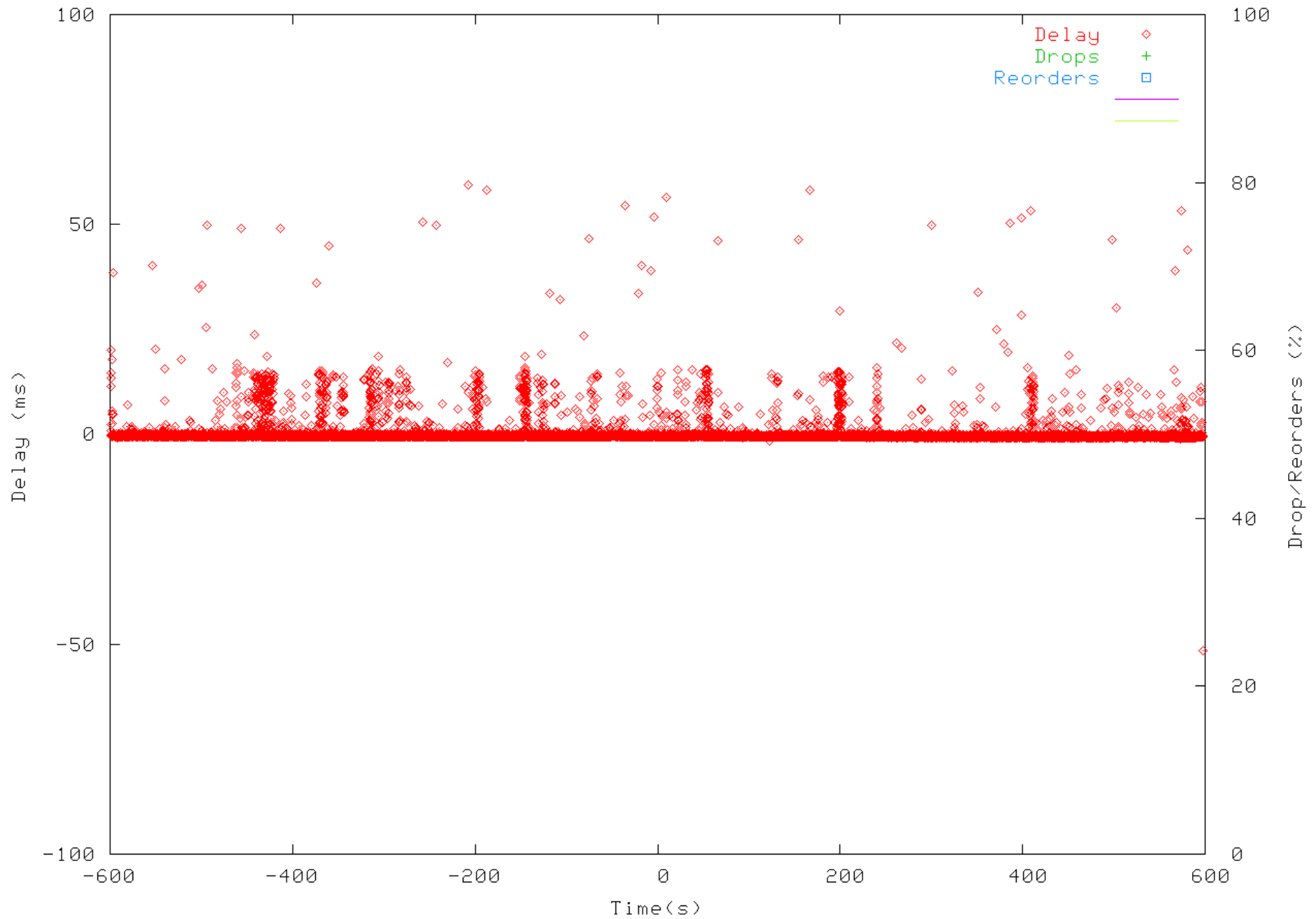


Transition: ispA,ispB -> ispB at 2004-01-03 12:00

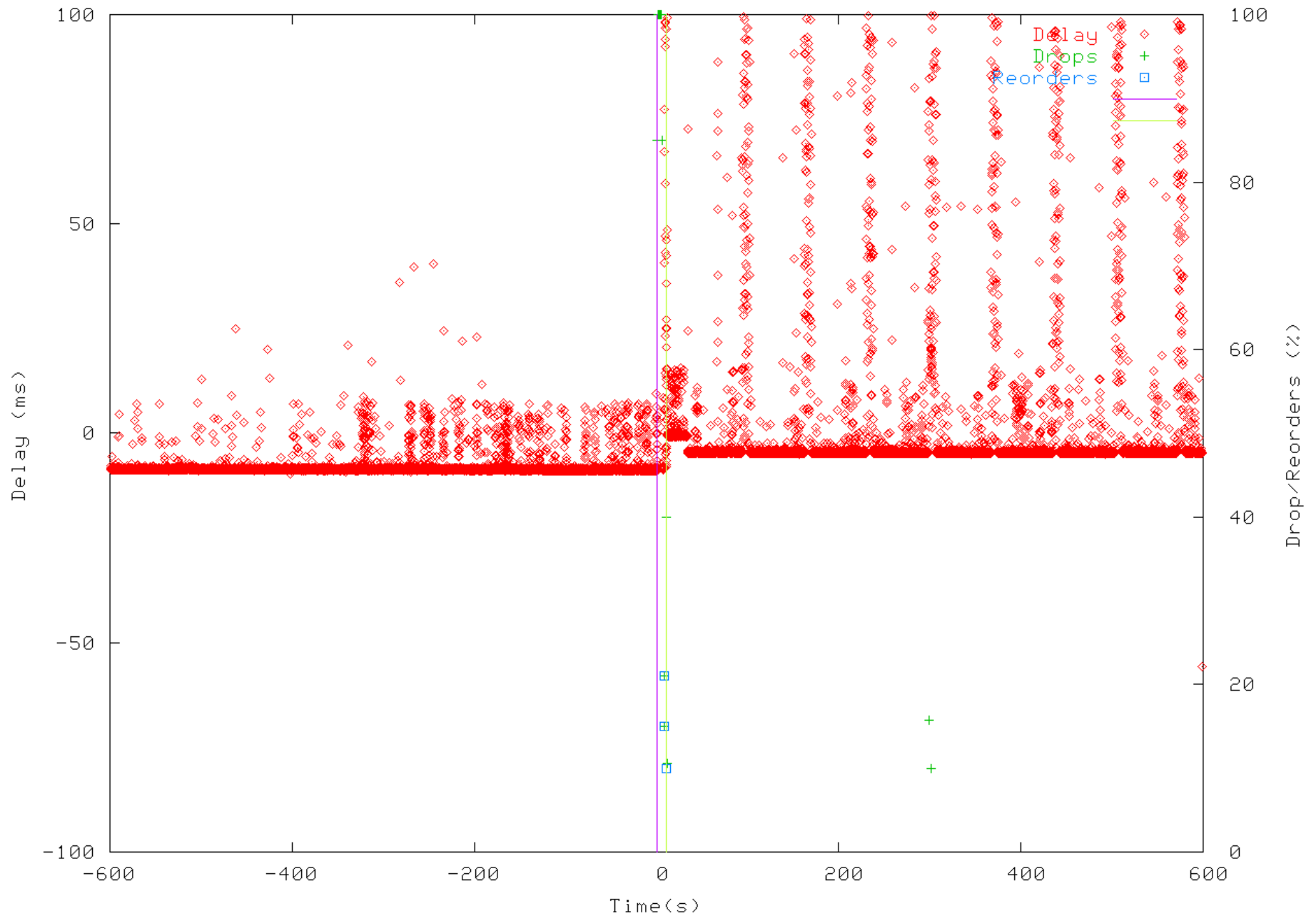




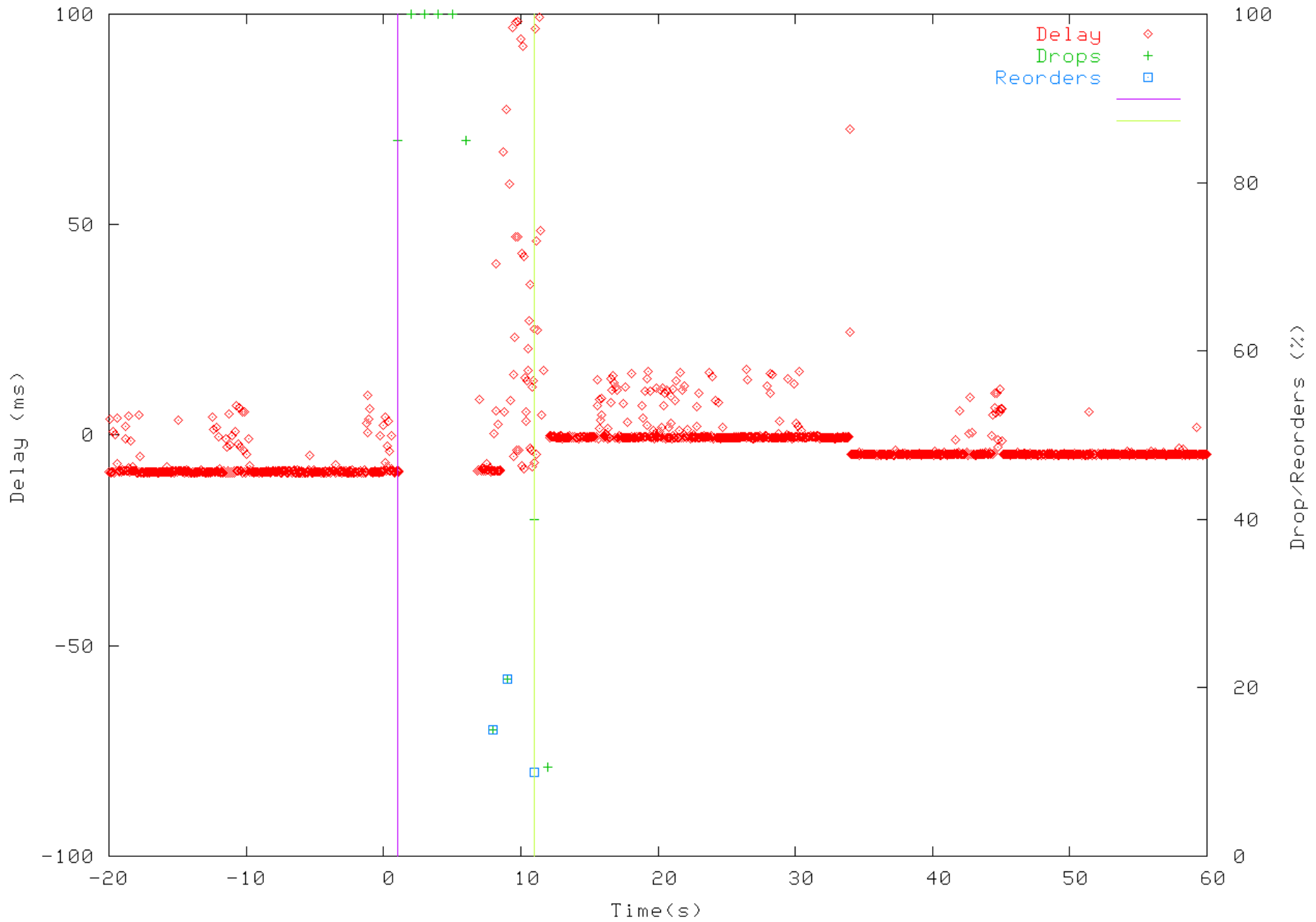
Transition: ispA,ispB -> ispA at 2004-1-3 4:00  
from probe: lcs-bgp.vineyard.net (204.17.195.103)



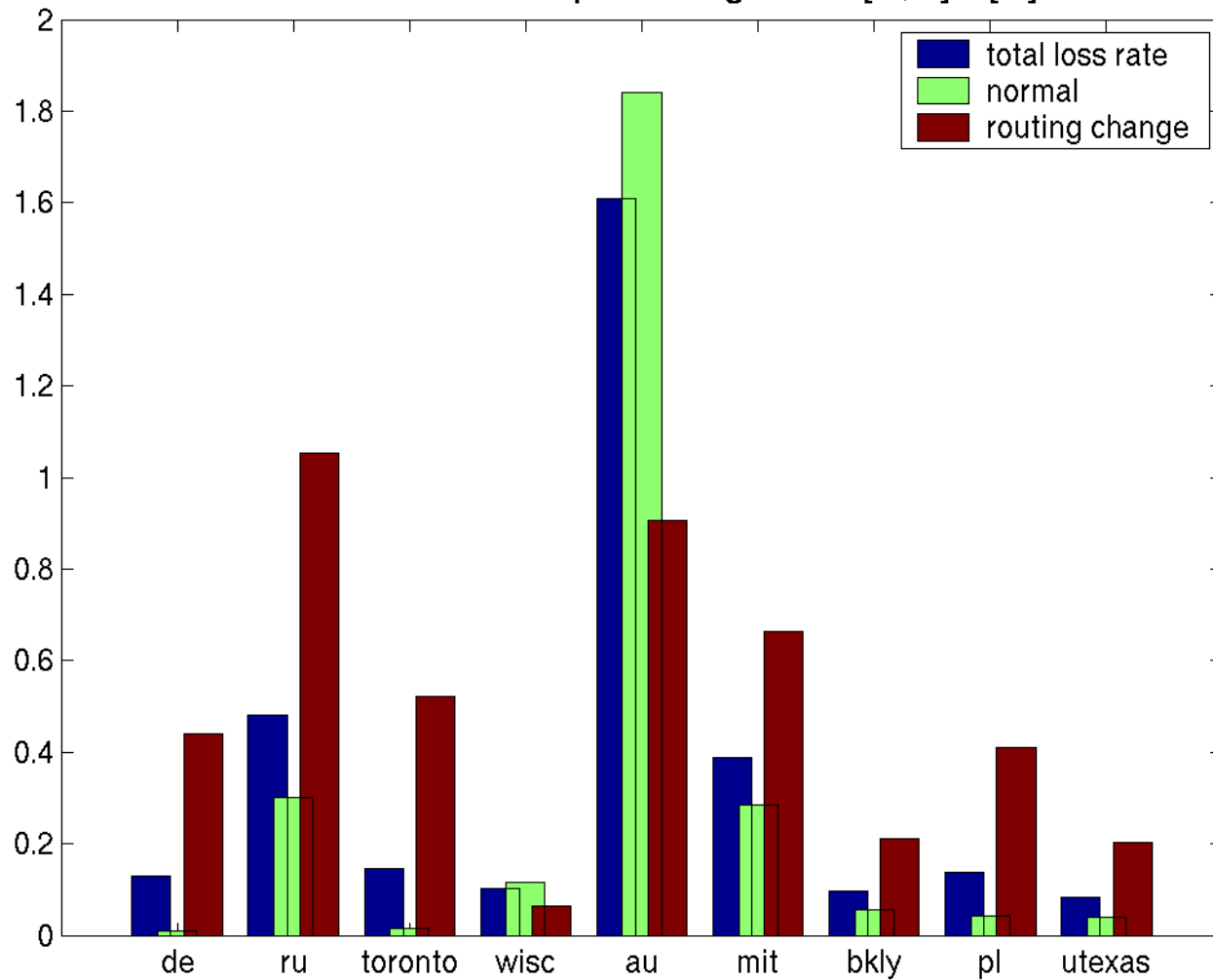
Transition: ispA,ispB -> ispB at 2004-1-3 12:00  
from probe: lcs-bgp.vineyard.net (204.17.195.103)

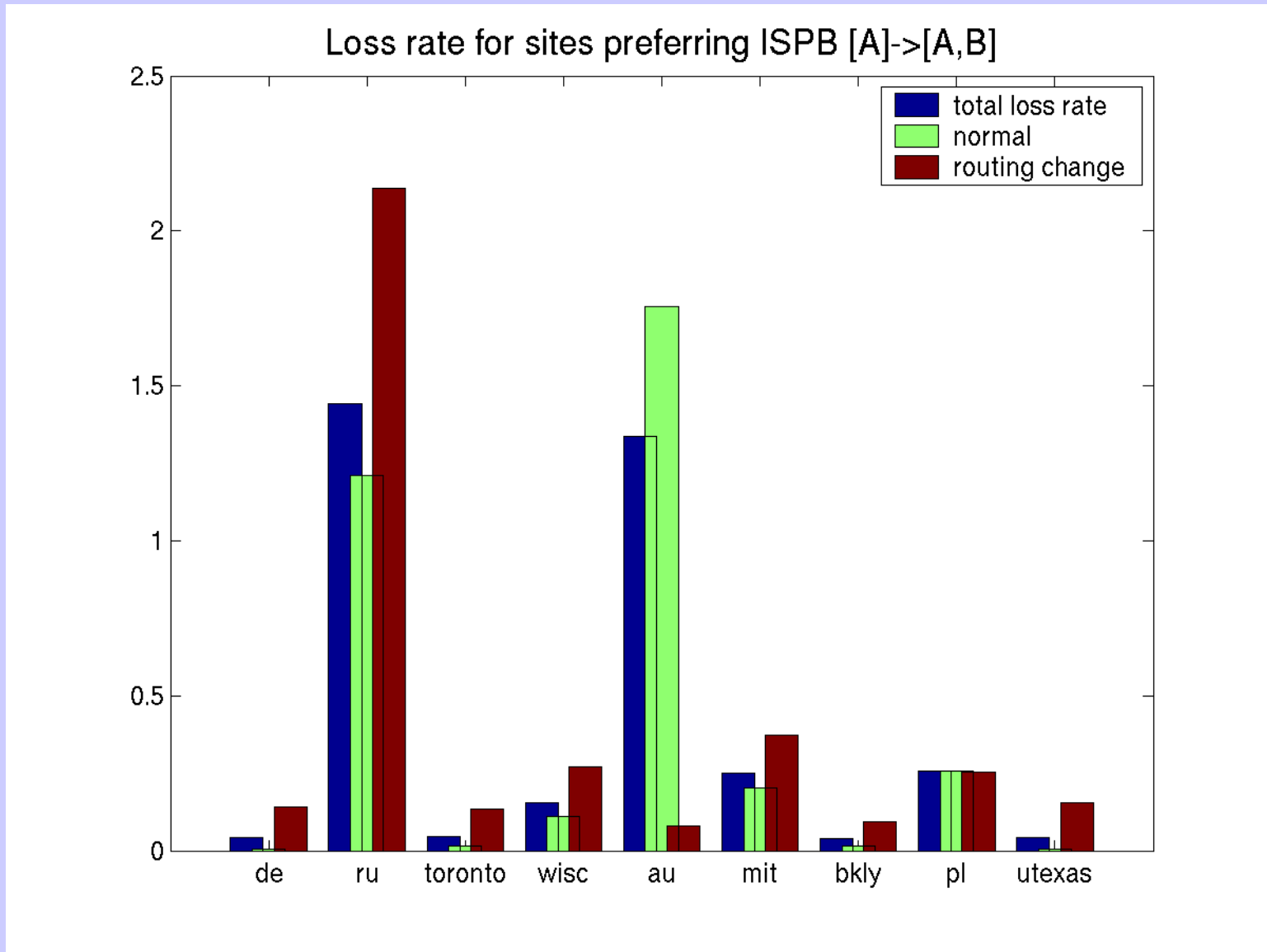


Transition: ispA,ispB -> ispB at 2004-1-3 12:00  
from probe: lcs-bgp.vineyard.net (204.17.195.103)

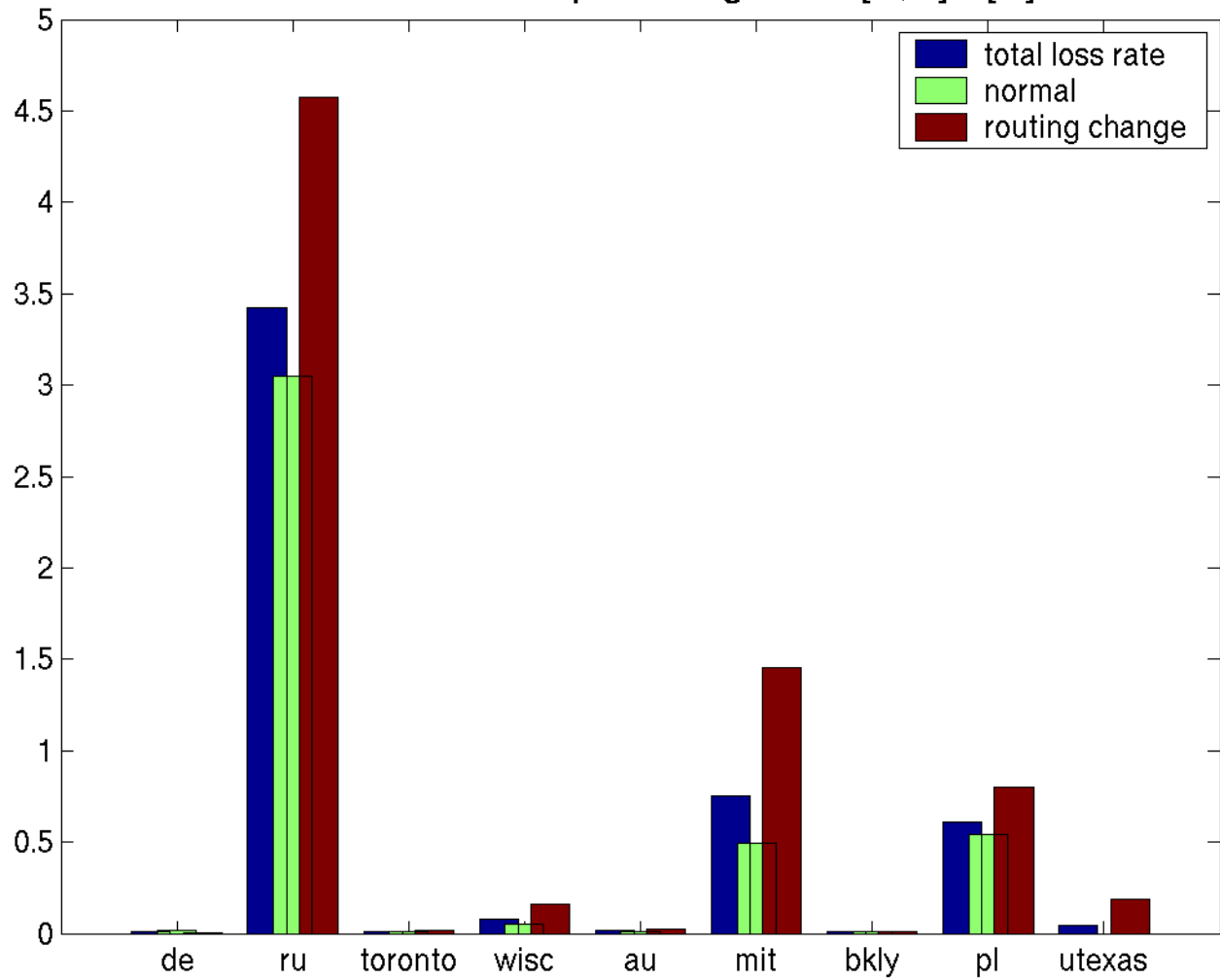


Loss rate for sites preferring ISPB [A,B]->[A]

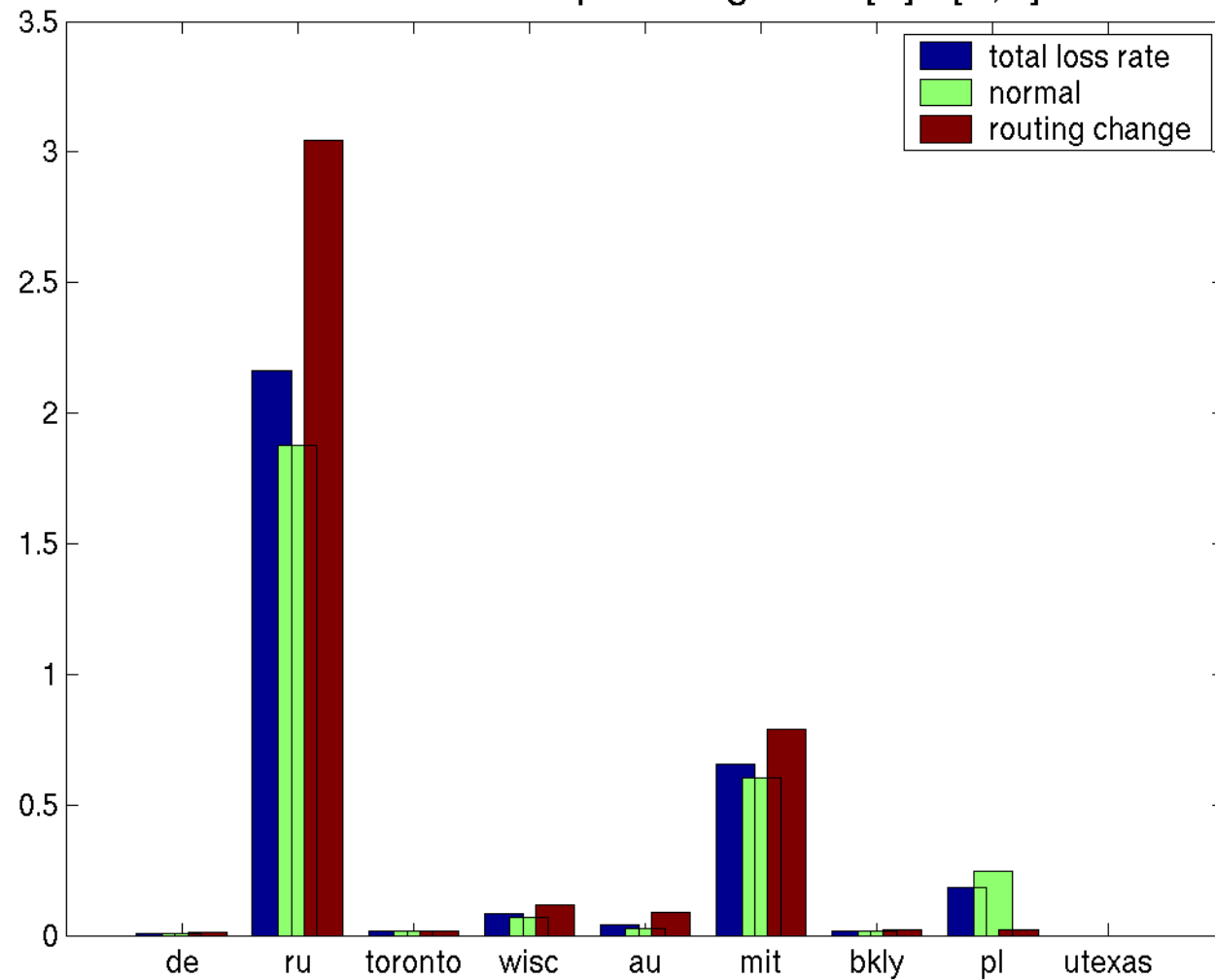




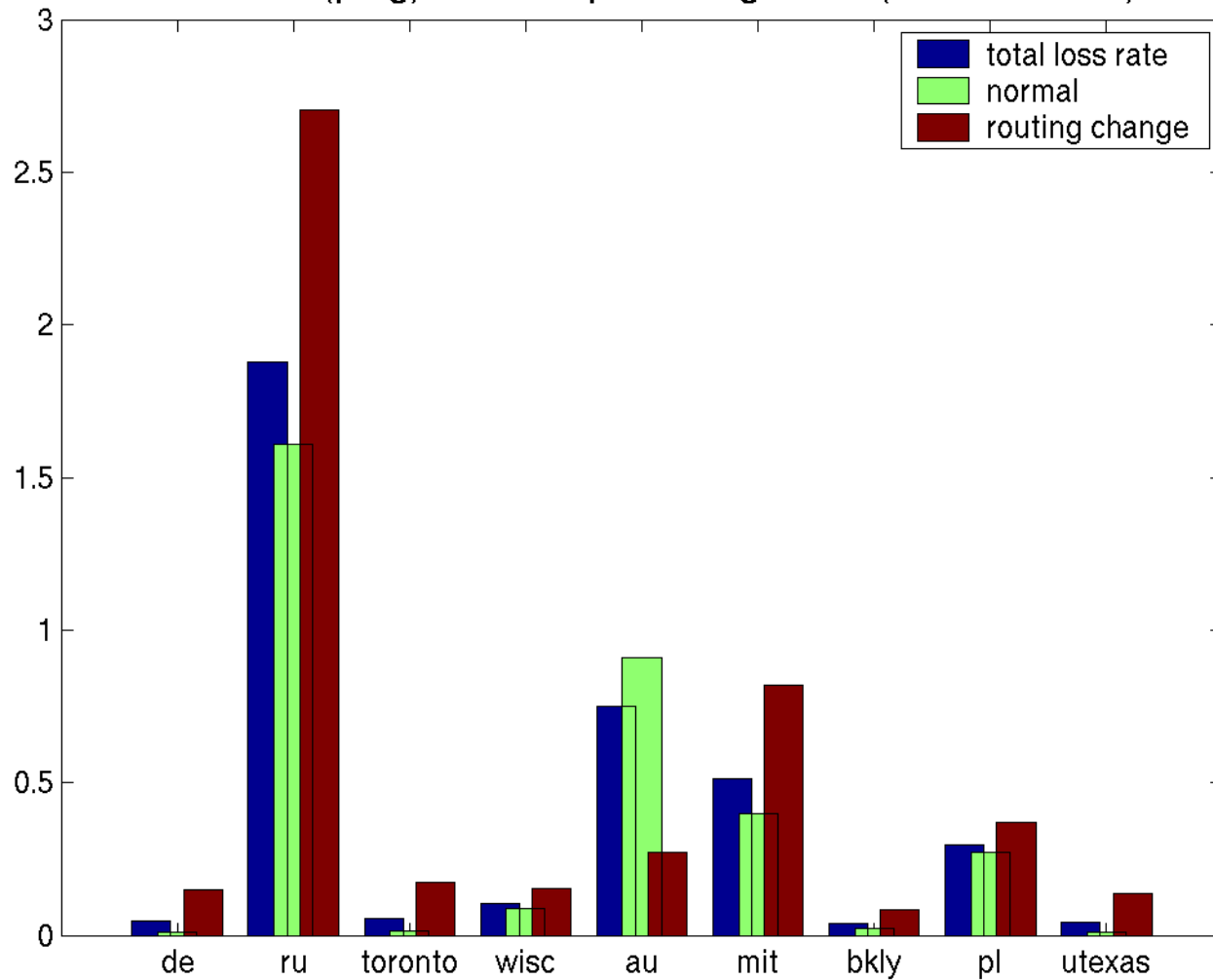
Loss rate for sites preferring ISPB [A,B]->[B]



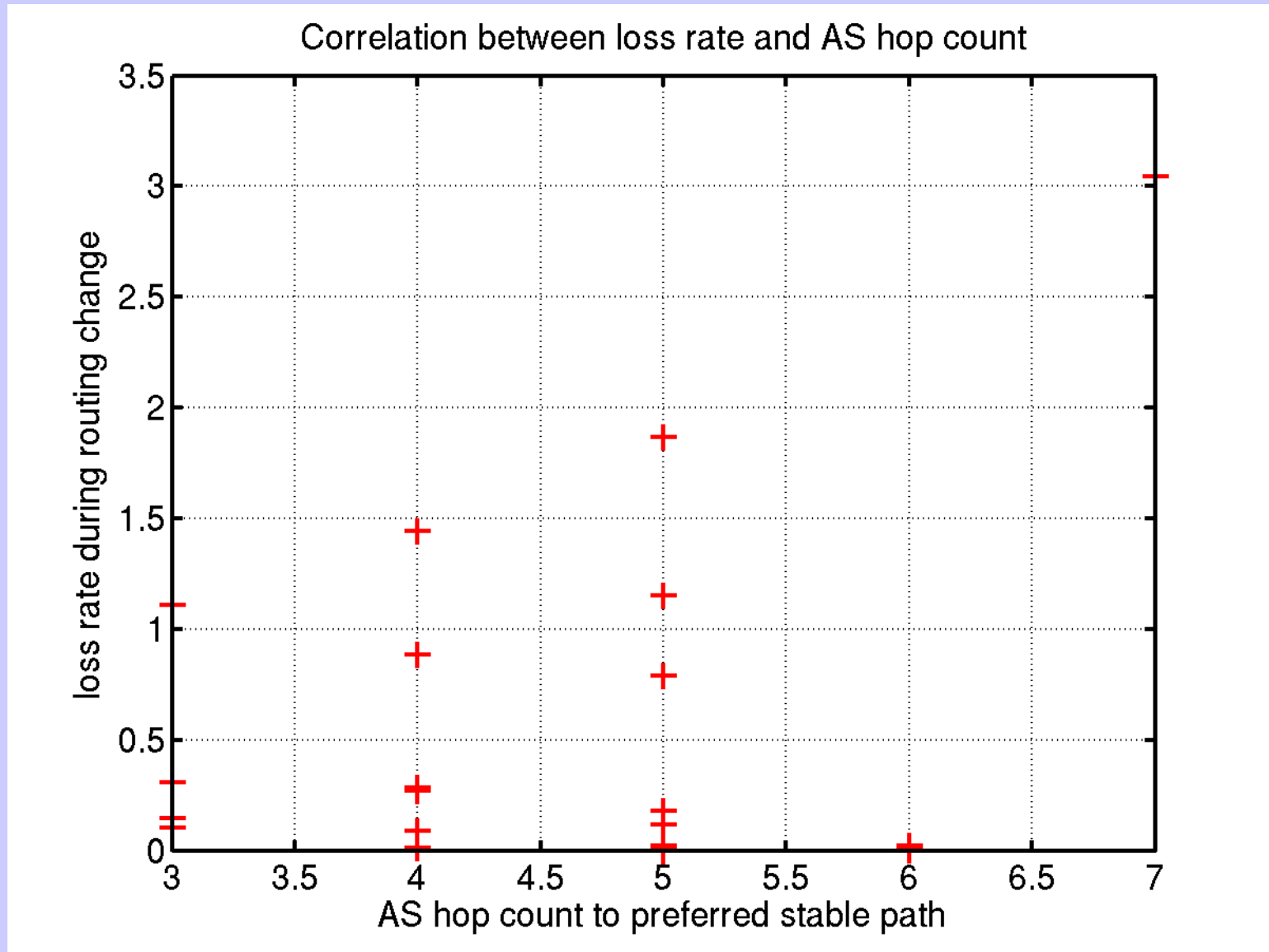
Loss rate for sites preferring ISPB [B]->[A,B]



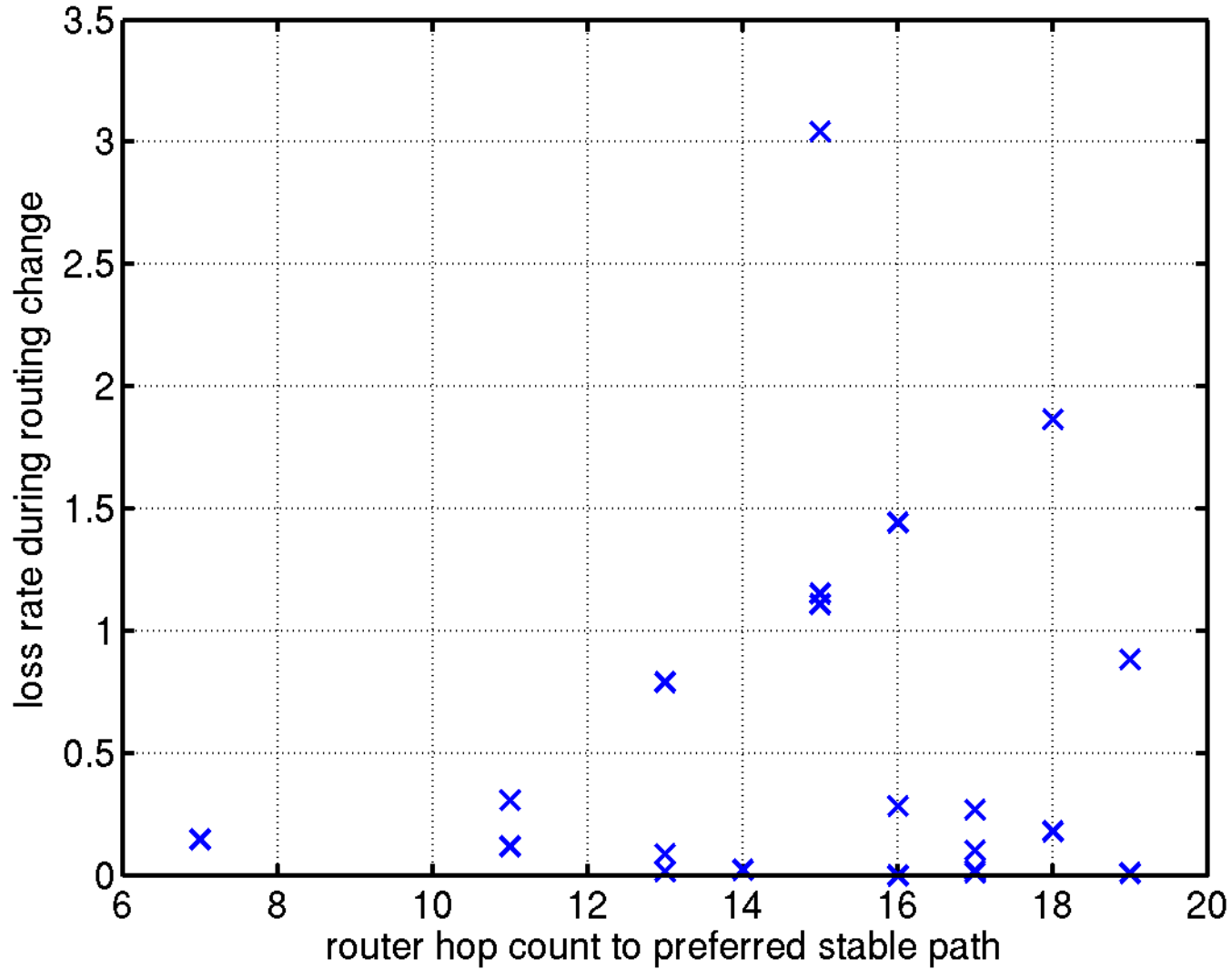
Loss rate (pctg) for sites preferring ISPB (all transitions)







Correlation between loss rate and router hop count



# This Seems to Say

- Distant sites experience more loss.
- There is a correlation between a site's routing preference and the type of transition: sites preferring ISP A have more loss rate during AB→B than AB→A, similarly more loss rate during B→AB than A→AB.
- The correlation between loss rate and AS or router hop count is quite weak. (we need more data points here).
- At some sites, the loss rate during 'normal' periods (i.e., no injected routing change) is higher than that during 'routing change' periods. Maybe those paths' inherent loss may be due to congestion.

# References

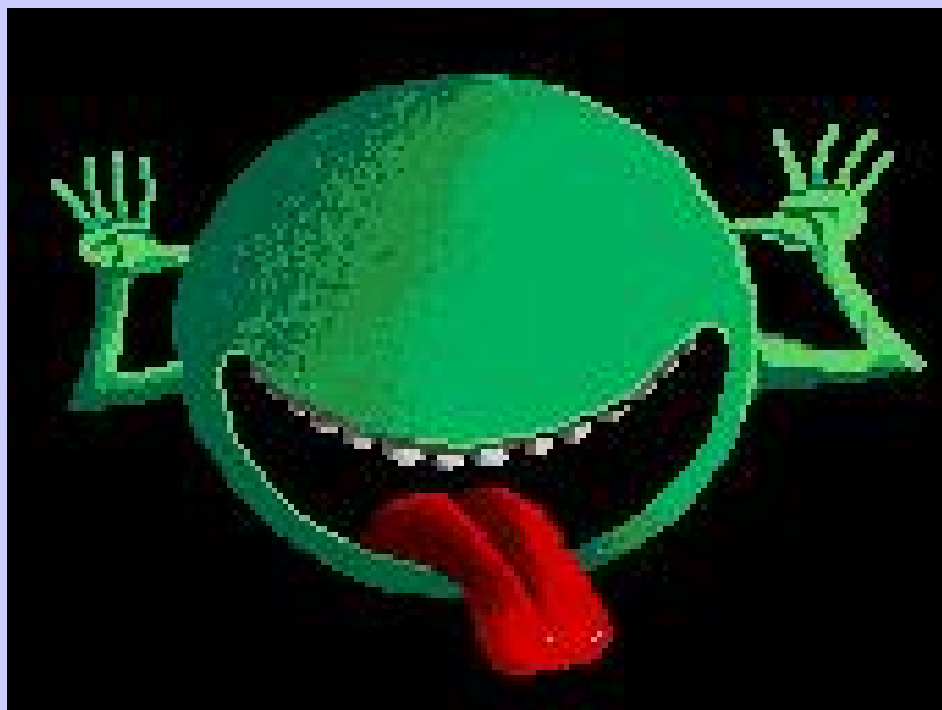
- Tim Griffin, "What is the sound of one route flapping" Dartmouth talk slides, June 2002
- "Global Routing Instabilities During Code Red II and Nimda Worm Propagation" Jim Cowie and Andy Ogielski, Renesys. NANOG 23, October 2001
- C. Labovitz, G. R. Malan, F. Jahanian, "Internet Routing Instability", TON 1998
- C. Labovitz, R. Malan, F. Jahanian, "Origins of Internet Routing Instability", Infocom 1999

## References (2)

- C. Labovitz, A. Ahuja, F. Jahanian, "Experimental Study of Internet Stability and Wide-Area Network Failures", FTCS 1999
- C. Labovitz, A. Ahuja, A. Bose, F. Jahanian, "Delayed Internet Routing Convergence" Sigcomm 2000
- C. Labovitz, A. Ahuja, R. Wattenhofer, S. Venkatachary, "The Impact of Internet Policy and Topology on Delayed Routing Convergence", Infocom 2001

# Thanks to Our Sponsors

- NSF via award ANI-0221435
- The University of Oregon (Dan & Eric)
- The University of Michigan (Morley)
- Internet Initiative Japan (Randy)
- Intel Corporation (Tim)
- Verio and Sprint (bandwidth)
- Juniper & Cisco (routers)



Don't Panic,  
Engineer Prudently