

Quantifying BGP Path Exploration in the Internet

Rafit Izhak-Ratzin, Beichuan Zhang, Dan Pei,
Daniel Massey, Lixia Zhang

July 2005

1

What is known/unknown about BGP Path Exploration/Slow Convergence

- It exists
 - Controlled experiments (i.e., beacons)
 - Small number of prefixes
 - Artificially injected routing events
 - Simulations, theoretical analysis
- To *what extent* does it exist in the operational Internet?
 - How often does it occur?
 - How bad is it?
 - Do all prefixes suffer to the same degree?
 - Do all AS observe the same?

2

What We Have Done

- An initial attempt to *quantify* path exploration and slow convergence in the Internet
 - Randomly picked 3 weeks from this year
 - Future work: take more samples
 - Collected BGP updates from 28 RouteViews Oregon monitor
 - Future work: add data from other RouteViews collectors, and from RIPE
 - Measured the routing changes of *all* prefixes in the global routing table to identify slow convergence

3

What we did with the data

- Group updates into events
- Classify events based on the AS paths before and after the event
- Calculate update count and duration of different types of events
- Examine the locations of origin AS and observation point on slow convergence

4

Event Identification

A monitor = a RouteView's peer

Given a sequence of BGP updates from one monitor regarding one prefix, $U_1, U_2, \dots, U_i, \dots$

- $\{U_i, \dots, U_{i+n}\}$ belong to the same event if the inter-arrival time < 2 min
 - The same scheme used by others previously
 - tried with 4 min, not much difference in results
- This scheme cannot separate out events that overlap in time

5

Event Classification

Examining the AS paths before and after each event

- **Same Path:** Same AS path throughout the event
- **Path Disturbance:** Same AS path before and after the event, but different AS path during the event.
 - Overlapping events, e.g. due to a transient failure
- **Path Change:** Different AS paths before & after the event

6

Further sorting of path change events

- Tup: a new route
- Tdown: the route is withdrawn
- Tshort: moved to a "better" path
- Tlong: the new path is "longer"
- Tequal: cannot tell

7

How to rank the paths

- Don't know the routing policy of the monitors
- Used the following heuristics for path ranking:
From each monitor's view: to reach a prefix
 - A path's usage time: most used path is most preferred
 - No-valley routing policy
 - Path length

8

Path Trend of An Event

- Compare consecutive updates in the event, count the number of rank increase (R_+) and rank decrease (R_-)
- Event = increase if $R_+ > R_-$
- Event = decrease if $R_+ < R_-$
- Event = equal otherwise

9

Using Beacon Prefixes to Verify the path ranking scheme

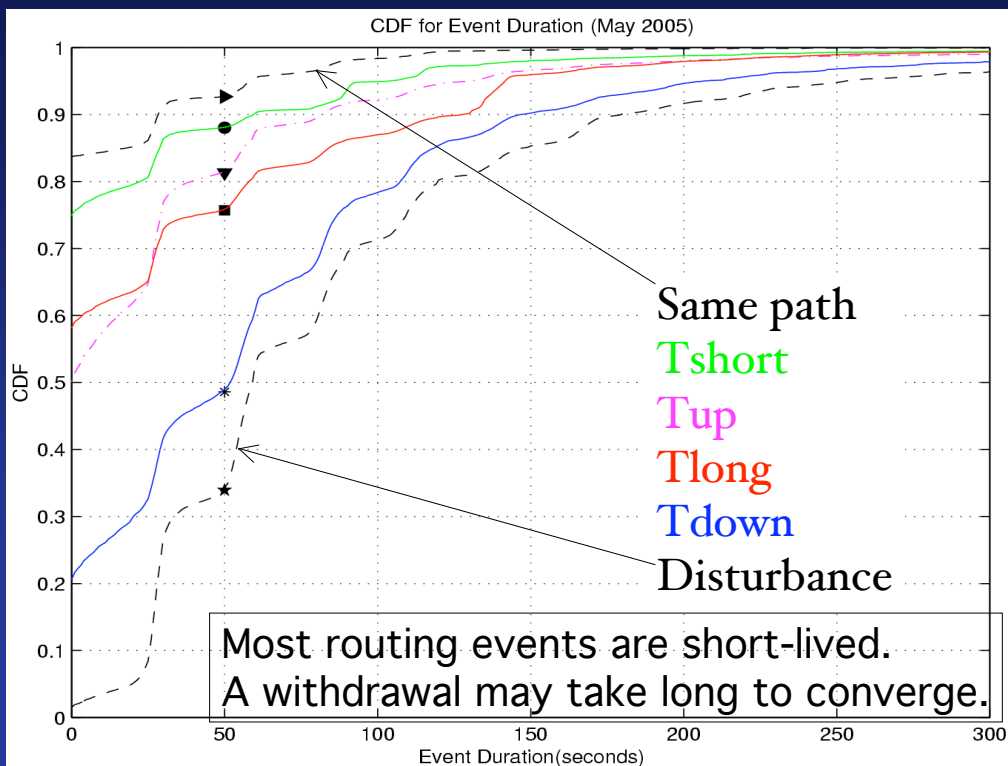
	increase	equal	decrease
Tdown	97.67%	2.27%	0.06%
Tup	0.71%	2.72%	96.57%

10

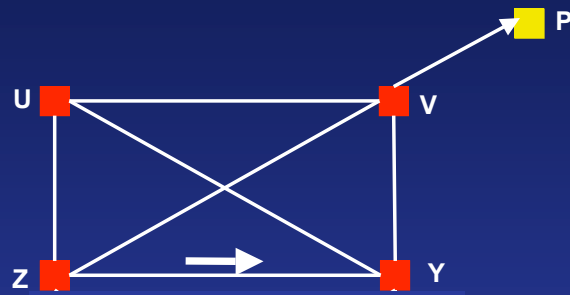
Events in One Week

Class	#Events	% of total events	#Updates	% of total Updates	U/E
Path change	2792581	42.67%	7578801	44.51%	2.7

Event Duration



Policy and Tier

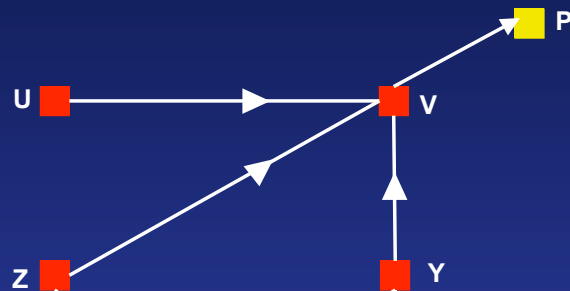


provider → customer

peer — peer

13

Policy and Tier

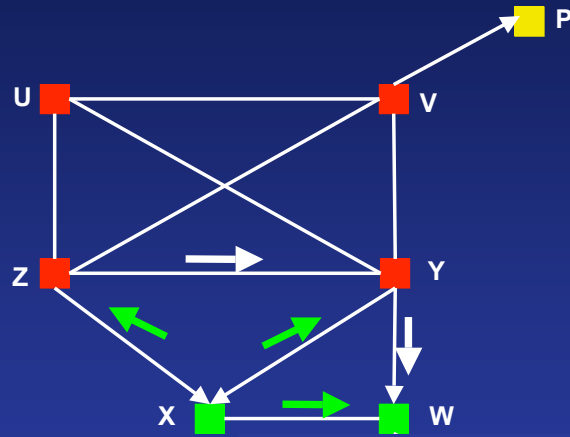


provider → customer

peer — peer

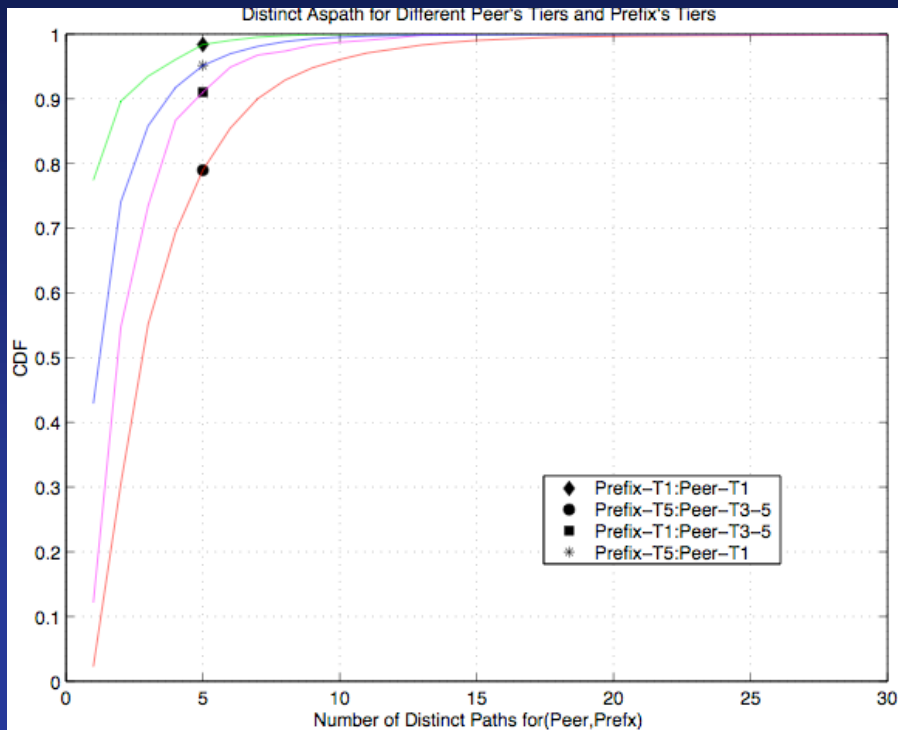
14

Policy and Tier

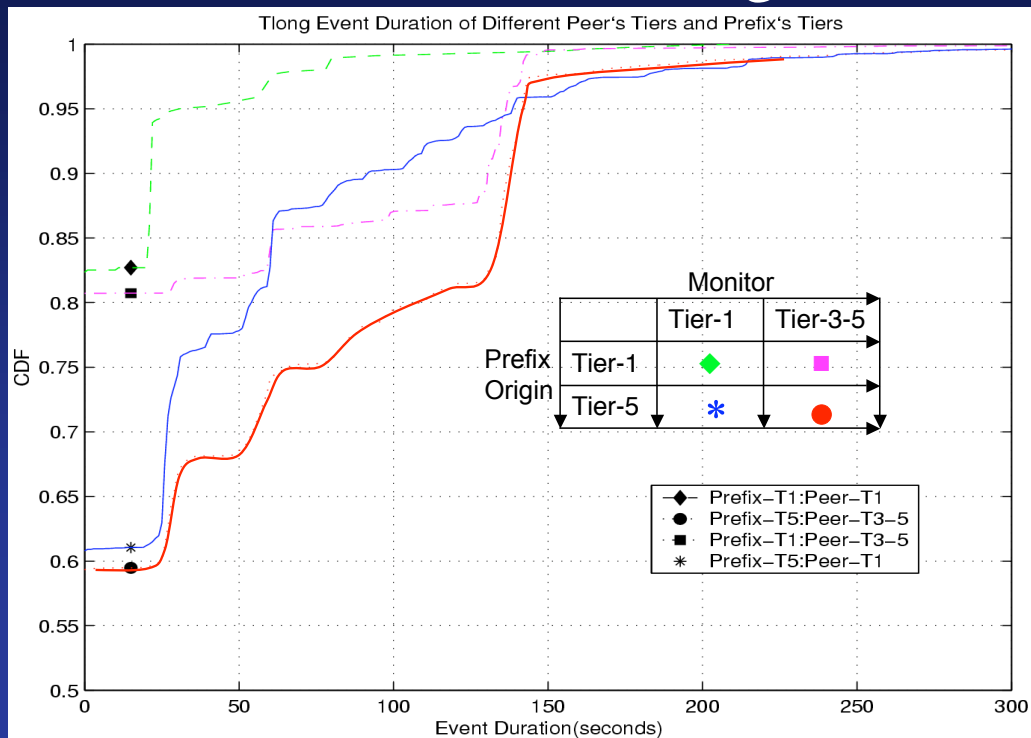


provider → customer
 peer — peer

#distinct paths observed over one week



Duration of Tlong



Summary:

How bad is BGP slow convergence?

Recall we asked 4 questions earlier

1. How often does it occur?
2. How bad is it?
3. Do all prefixes suffer to the same degree?
4. Do all ASes have the same observation?

Summary

(preliminary results)

1. How often does it occur?
 - Occur in most Tdown events
 - In less than half of Tlong events; much less in others
2. How bad is it?
 - Very long tail; >90-95% converged in < 2.5 min
3. Prefixes originated from tier-1 ASes suffer less than those from lower tiers
4. tier-1 ASes observe less slow convergence than lower tier ASes
 - False flat dampening more likely to happen at lower tiers

19

Comments & Questions?

Send to lixia@cs.ucla.edu

20