Route flap damping: harmful?

2002.09.10
RIPE/Rodos

Randy Bush (IIJ)

Tim Griffin (AT&T Labs – Research)

Zhuoqing Morley Mao (UC Berkeley)

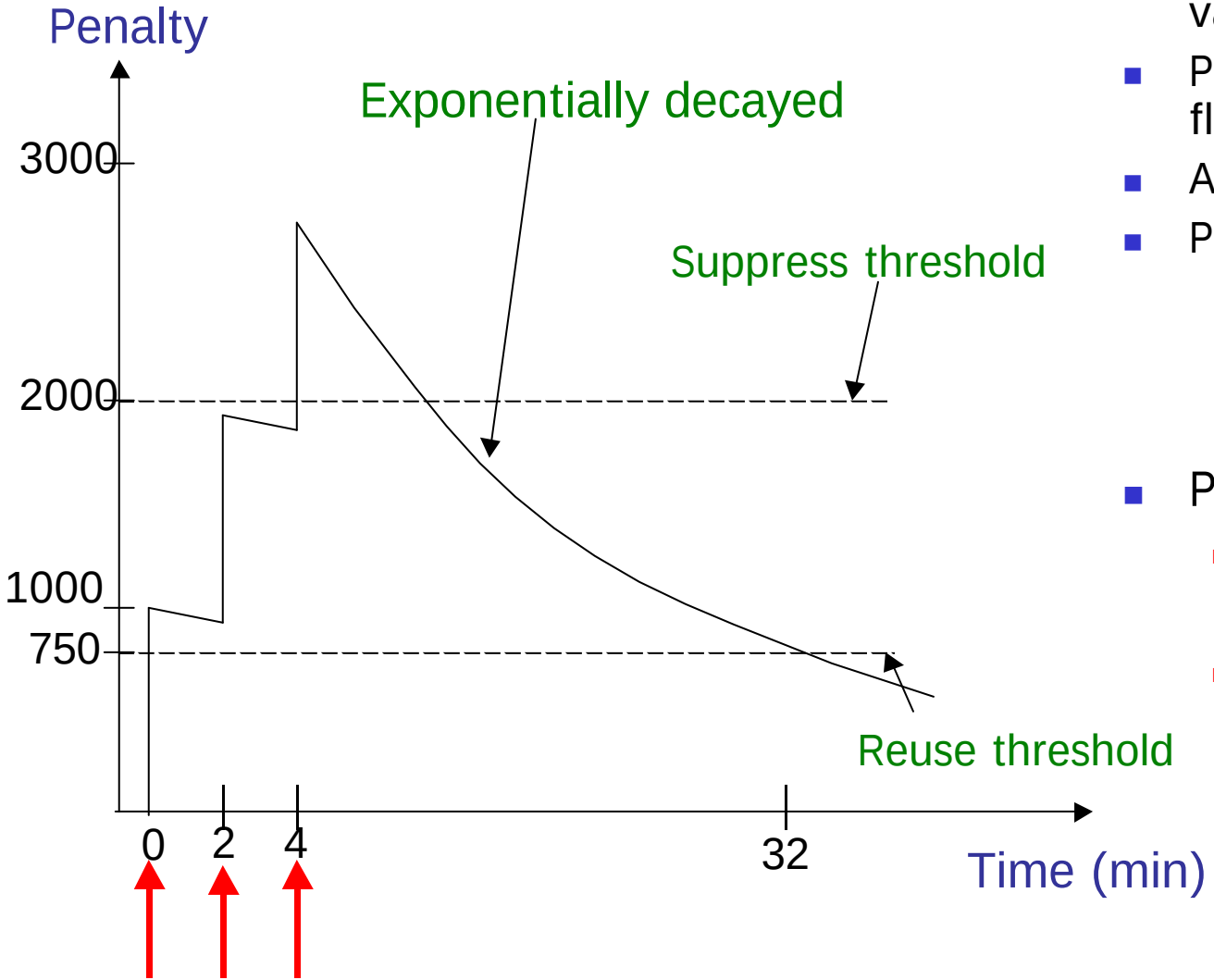
Route flap damping

- RFC2439/RIPE-229
  - Supported by all major router vendors
  - Believed to be widely deployed
  - Responsible for Internet stability?

- Goals:
  - Reduce router processing load due to instability
  - Do not sacrifice convergence times for well-behaved routes (!?)

How does route flap damping work?

(using default Cisco parameters)



- For each peer, per destination, keep a penalty value
- Penalty increases for each flap
- A flap is a route change
- Penalty decays exponentially

$$P(t') = P(t)e^{-l(t'-t)}$$

- Parameters:
  - Fixed:
    Penalty increment
  - Configurable:
    half-life, suppress-, reuse-threshold, max suppressed time

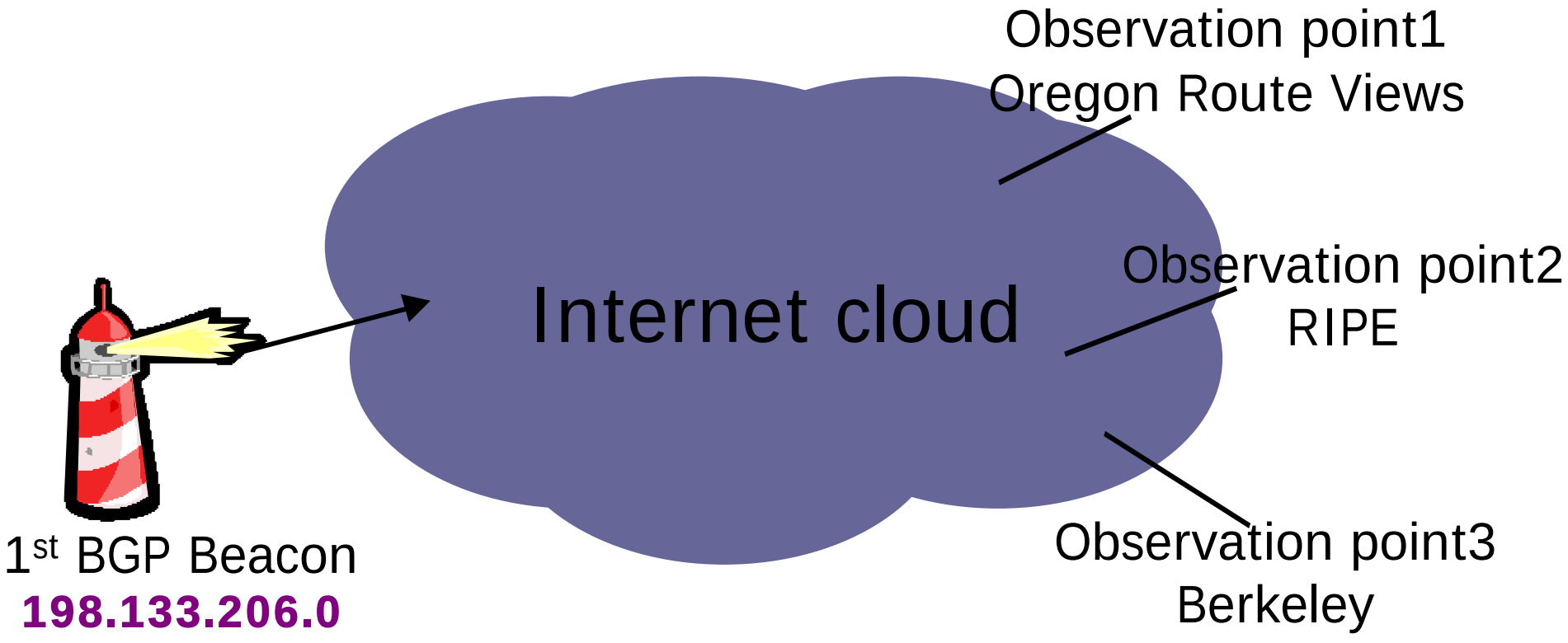
Router vendor default values

Parameter	Cisco	Juniper
Withdrawal penalty	1000	1000
Re-advertisement penalty	0	1000
Attributes change penalty	500	500
Suppress threshold	2000	3000
Half-life (min)	15	15
Reuse threshold	750	750
Max suppress time (min)	60	60

- Cisco
  - Three flaps can suppress route
- Juniper
  - Minimum four flaps to suppress route
- Example:
  - Three flaps with 2 min interval
  - Cisco: suppress on the third flap for more than 28 minutes

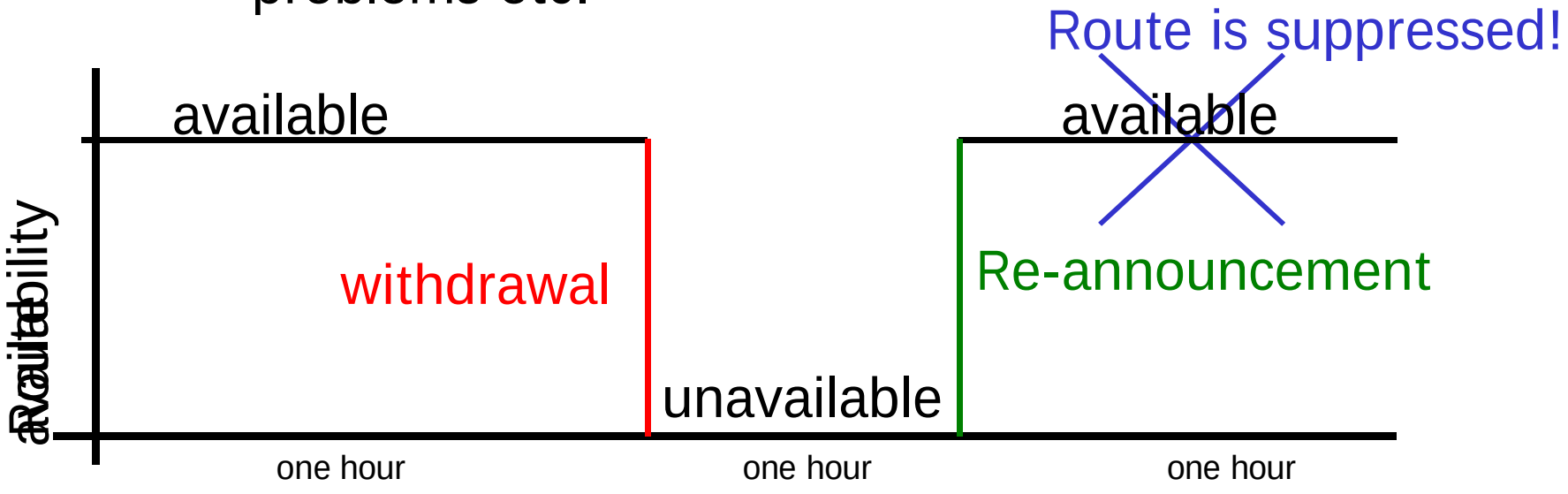
Verified using BGP Beacons

- BGP Beacon:
  - A prefix that is announced and withdrawn at **well-known** times



Transient instability

- Router reboot
  - Due to circuit or software upgrade, etc.
- A single link flap
  - Due to network congestion, link connectivity problems etc.



Cascaded withdrawals! (1)

- Peer: 212.47.190.1, AS=9177 from RIPE
- In response to WD-beacon at 18:00, Aug 10th.
- Using Cisco setting + RIPE229 recommendation

Time 8/10	A/W	ASPath	Penalty
18:00:15	A	9177 3320 1 2914 3130 3927	500
18:00:41	A	9177 6730 5400 2914 3130 3927	990
18:01:41	A	9177 3320 2914 3130 3927	1445
18:03:06	A	9177 3320 1239 2914 3130 3927	1853
18:03:35	W		2812
18:04:03	A	9177 6730 5400 2914 3130 3927	2752
18:04:31	W		3694

Above suppress threshold

Cascaded withdrawals! (2)

- Peer: 213.200.87.254, AS=3257 from RouteViews
- In response to WD-beacon at 01:00, Aug 20th.
- Using Cisco setting + RIPE229 recommendation
  - (Note: first 2 announcements differ in community attributes)

Time 8/20	A/W	ASPath	Penalty
01:00:16	A	3257 1299 2914 3130 3927	500
01:00:47	A	3257 1299 2914 3130 3927	988
01:00:50	W		1985
01:00:50	A	3257 1299 4200 2914 3130 3927	1985
01:01:13	A	3257 1299 701 2914 3130 3927	2451
01:02:05	W		3354

Above suppress threshold

Why does this happen?

- BGP is a path vector protocol
  - Explores alternate routes before withdrawal
  - Topology dependent
- Delay in messages due to variations in
  - MinRouteAdver timer values
  - Propagation delays
  - Router processing overhead
- Route flap damping parameter setting
  - Cisco/Juniper punishes virtually all route changes
  - Default setting and RIPE-229 recommendation are too aggressive

What to do?

- Redesign flap damping parameters
  - Less aggressive:
    - e.g., increase the suppress threshold
  - Need to understand impact on router load
  - Take into consideration of alternate routes
    - Less aggressive if no alternate routes, because reachability can be affected.

We need your help!

- To install more BGP beacons
  - To improve understanding of BGP dynamics
  - Need topology diversity
  - Undertand the function:
    - BGP Behavior = Function (Topology, Propagation delays, Routing Policies, ???)

- Contact zmao@research.att.com for more information

Reference

- RIPE-229
- RFC 2439
- C. Labovitz, A. Ahuja, A. Bose, F. Jahanian, “Delayed Internet Routing Convergence” Sigcomm 2000
- Tim Griffin, “What is the sound of one route flapping” talk slides, 2002
- Z. Mao, R. Govindan, G. Varghese, R. Katz “Route Flap Damping Exacerbates Internet Routing Convergence” Sigcomm 2002