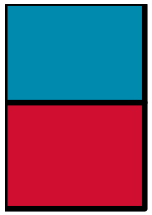
Revisions to RIPE210: Coordinated Route Flap Damping Parameters

Philip Smith

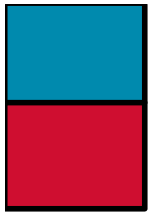
pfs@cisco.com

RIPE39, Bologna, May 2001



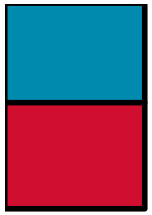
Route Flap Damping

- What is it?
- Why revise RIPE-210?
- Changes
- Summary



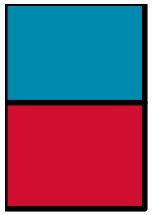
Route Flap Damping

- **Route flap**
  - **Going up and down of path or change in attribute**
    - **BGP UPDATE followed by WITHDRAW = 1 flap**
    - **eBGP neighbour going down/up is NOT a flap**
  - **Ripples through the entire Internet**
  - **Wastes CPU**
- **Damping aims to reduce scope of route flap propagation**



Route Flap Damping

- **Requirements**
  - **Fast convergence for normal route changes**
  - **History predicts future behaviour**
  - **Suppress oscillating routes**
  - **Advertise stable routes**
- **Described in RFC2439**



Operation

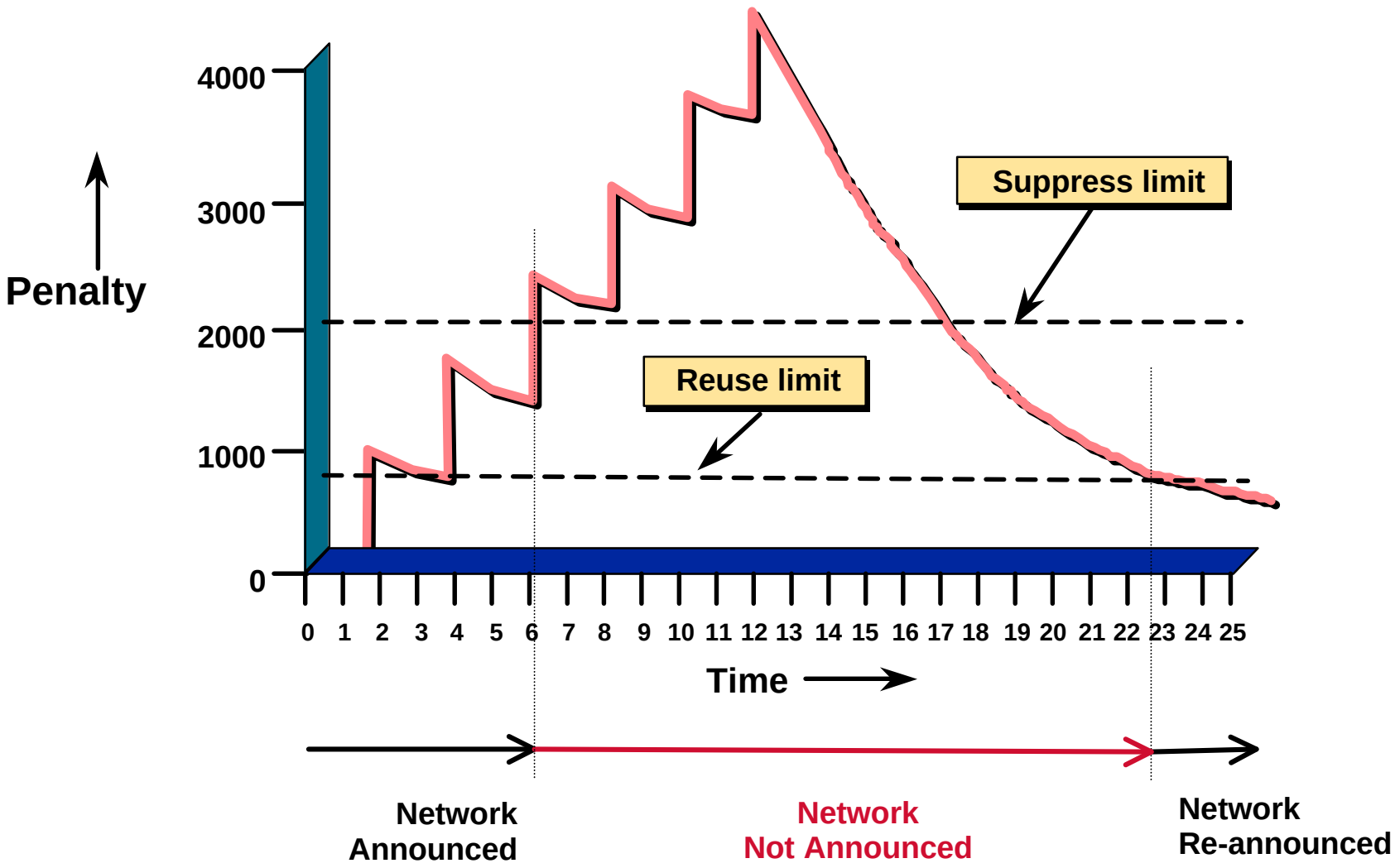
- **Add penalty for each flap**
- **Exponentially decay penalty**

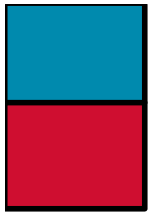
  half life determines decay rate
- **Penalty above suppress-limit**

  do not advertise route to BGP peers
- **Penalty decayed below reuse-limit**

  re-advertise route to BGP peers

Operation





Operation

- Only applied to inbound announcements from eBGP peers
- Alternate paths still usable
- In Cisco IOS controlled by:

  Penalty of 1000 per flap

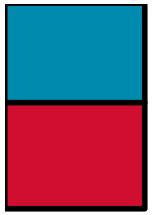
  (penalty of 500 for attribute change)

  Half-life (default 15 minutes)

  reuse-limit (default 750)

  suppress-limit (default 2000)

  maximum suppress time (default 60 minutes)



Operation

- **BGP WITHDRAW message received**

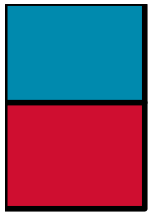
  penalty on prefix increased by 1000

  prefix is marked as having flap history

- **BGP UPDATE message received**

  if penalty ³ suppress-limit, prefix is not announced to any BGP peers and is marked as suppressed

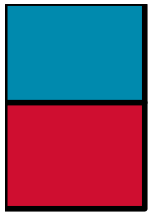
- **If prefix carries on flapping after being suppressed, penalty is incremented and decayed as normal**



Operation

- Once prefix is stable, it will be suppressed according to the decay rate given by the half life time
- Penalty value is decayed

  Decay rate is same whether prefix is or is not in the BGP table
- Once penalty reaches reuse-limit, prefix is re-advertised
- Once penalty is less than half reuse-limit, penalty is reset to zero (Cisco IOS)



Operation

- **Example – Cisco IOS default**

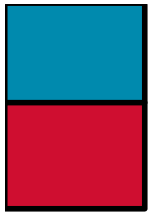
  bgp dampening 15 750 2000 60

  half-life of 15 minutes

  reuse-limit of 750 and suppress time of 60 minutes means maximum possible penalty of 12000

  once prefix stops flapping, penalty is decayed to 750 - this will take maximum of 60 minutes

  once penalty reaches 375, it is reset to zero and all damping history is removed



Operation

- Care required when setting parameters
- Penalty must be less than reuse-limit at the maximum suppress time
- Maximum suppress time and half life must allow penalty to be larger than suppress limit
- Decay rate pre-calculated when flap damping is configured

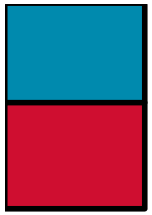
  numbers must be feasible, IOS does not check (yet)

Maths!

- **Maximum value of penalty is**

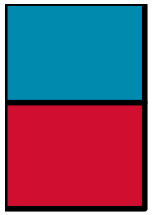
$$\text{max-penalty} = \text{reuse-limit} \times 2^{\left\langle \frac{\text{max-suppress-time}}{\text{half-life}} \right\rangle}$$

- **Always make sure that suppress-limit is LESS than max-penalty otherwise there will be no route flap damping**



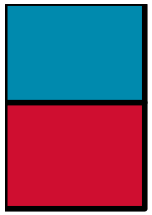
Route Flap Damping

- What is it?
- Why revise RIPE-210?
- Changes
- Summary



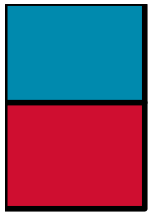
Why revise RIPE 210?

- Parameters for /24 damping are virtually unfeasible
- “Golden Networks” are changing
- Remaining open issues are long solved
- Restructuring and updating



Route Flap Damping

- What is it?
- Why revise RIPE-210?
- Revisions
- Summary



/24 parameters

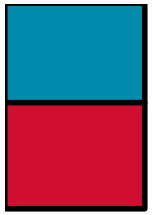
- **Original**

  bgp dampening 30 750 3000 60

  reuse-limit of 750 means maximum possible penalty is 3000

  ⇨no prefixes suppressed as penalty cannot exceed suppress-limit

  ⇨But damping is seen in real life – because if an update is received within 5 seconds of a withdraw it is possible to have the prefix suppressed. Rare!



/24 parameters

- **Revision**

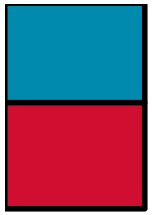
  bgp dampening 30 820 3000 60

  reuse-limit of 820 means maximum possible penalty is 3280

  ⇨Suppress limit is well below maximum possible penalty

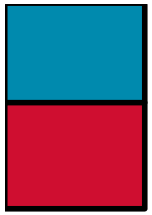
  ⇨Prefixes are suppressed

  ⇨Original design intentions achieved



“Golden Networks”

- **Networks which should not be damped when they flap**
  - **RIPE-178 listed the root nameservers**
  - **RIPE-210 claimed to list the revised root nameservers**
  - **RIPE-210 in fact listed a few root servers and a few gTLD servers – and the latter networks are frequently changing**

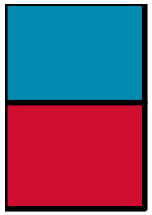


“Golden Networks”

- Revision redefines “Golden Networks” to be those which an operator does not want to be damped
- Revision lists possible networks as “Golden Networks” in an Appendix

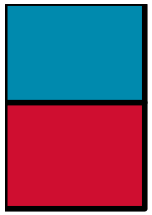
  The examples include the current list of root and gTLD servers

  Operators are encouraged to construct their own list as appropriate



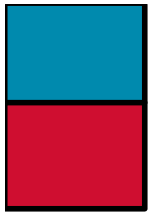
Open Issues in RIPE-210

- **Updates from a router arrive at a peer at different times through different paths**
  - **Line flap or circuit upgrade looks like multiple flaps**
  - **Considered a bug!**
- **Solution**
  - **Not a bug**
  - **Use IOS command “ip route …. permanent”**
  - **Static route always exists in routing table, even if interface is down**
  - **Therefore prefix/path is always in BGP ⇨ no flap**



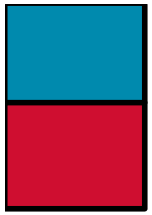
Restructuring

- "Golden Networks" moved to Appendix
- Configuration examples moved to Appendix
- Other stability features brought up to date
  - Route refresh is RFC2918 – proposed standard
  - Cisco IOS "soft reconfiguration" relegated to "recommended only if RFC2918 not supported by peer"
- Non-recommended flap damping parameter configuration discussed



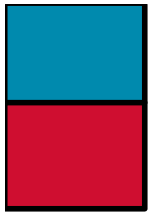
Updates

- **Appendix containing study of flap damping operation**
  - **Shows typical flaps for /24s, /22s & /23s, and £/21s**
  - **Intended to be helpful to port configurations to other vendor implementations**



Updates

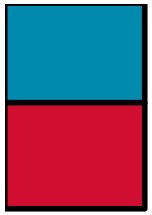
- **Appendix dedicated to configuration examples:**
  - **Received sample for Juniper's JunOS**
  - **Request to operators & vendors for other configuration examples:**
    - **GateD, Foundry, Redback,…?**



Updates

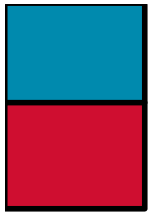
- First draft posted to routing-wg, NANOG, APOPS and AfNOG
- Several comments received

  Second draft will be posted soon



Route Flap Damping

- What is it?
- Why revise RIPE-210?
- Revisions
- Summary



Summary

- Errors fixed
- Restructured
- New document hopefully resembles best current practices

- Questions/Comments?