IP ROUTING

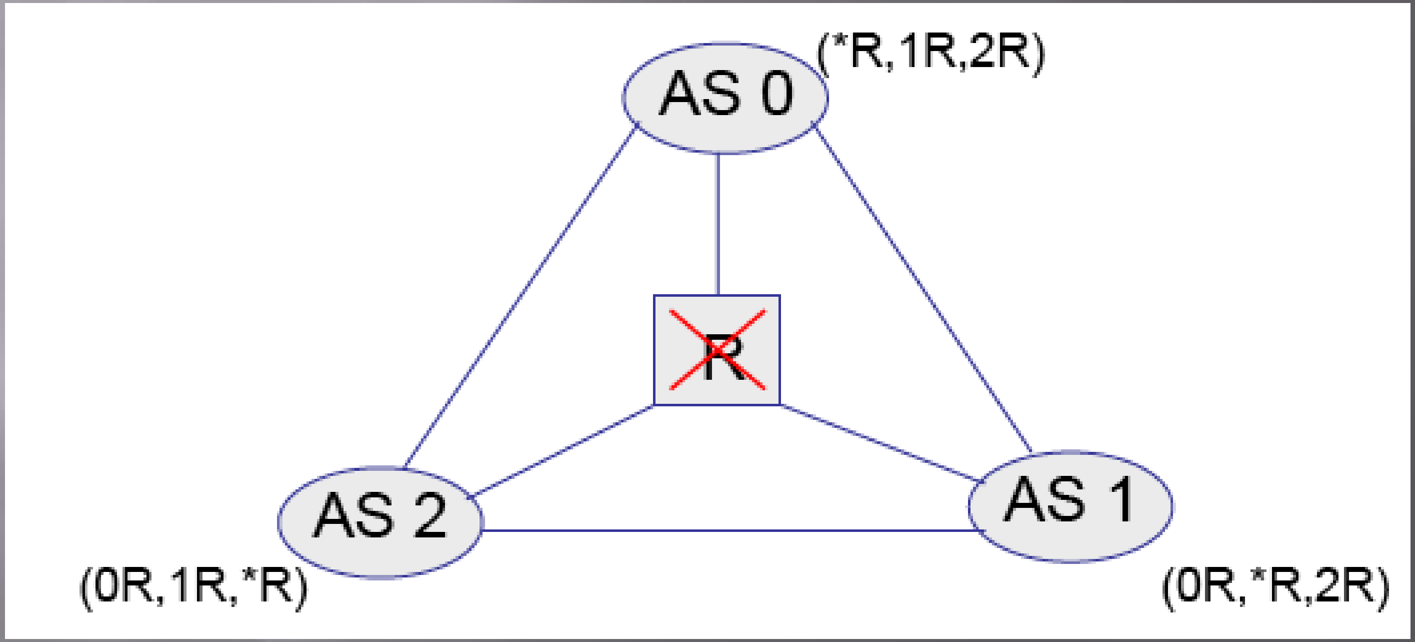
Observations

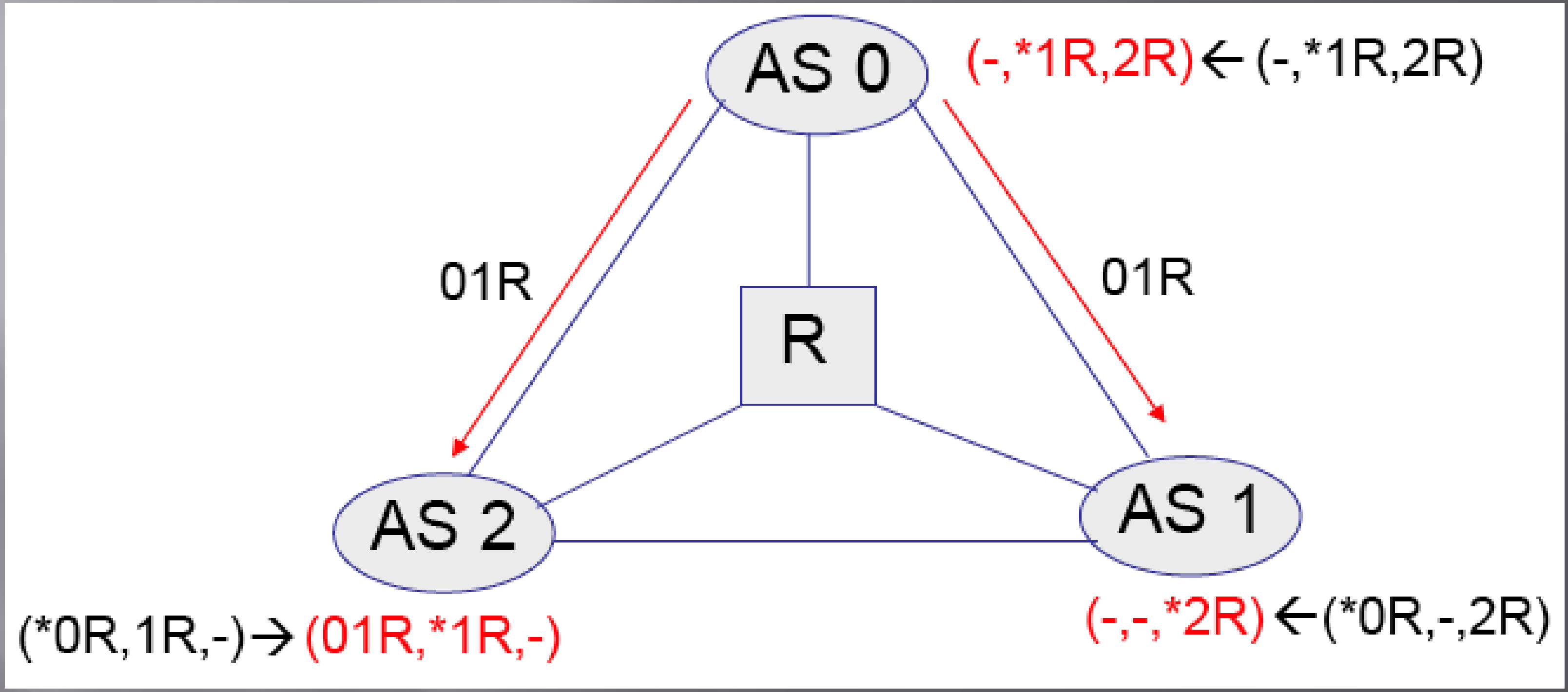
- Repairs (Tup time) exhibit similar convergence properties as long-short ASPath fail-over
- Failures (Tdown) and short-long fail-overs (e.g. primary to secondary paths) also similar
  - Slower than a repair (bad news doesn't travel fast, DV protocols)
  - 60% take > 2 minutes

Why "long" failovers?

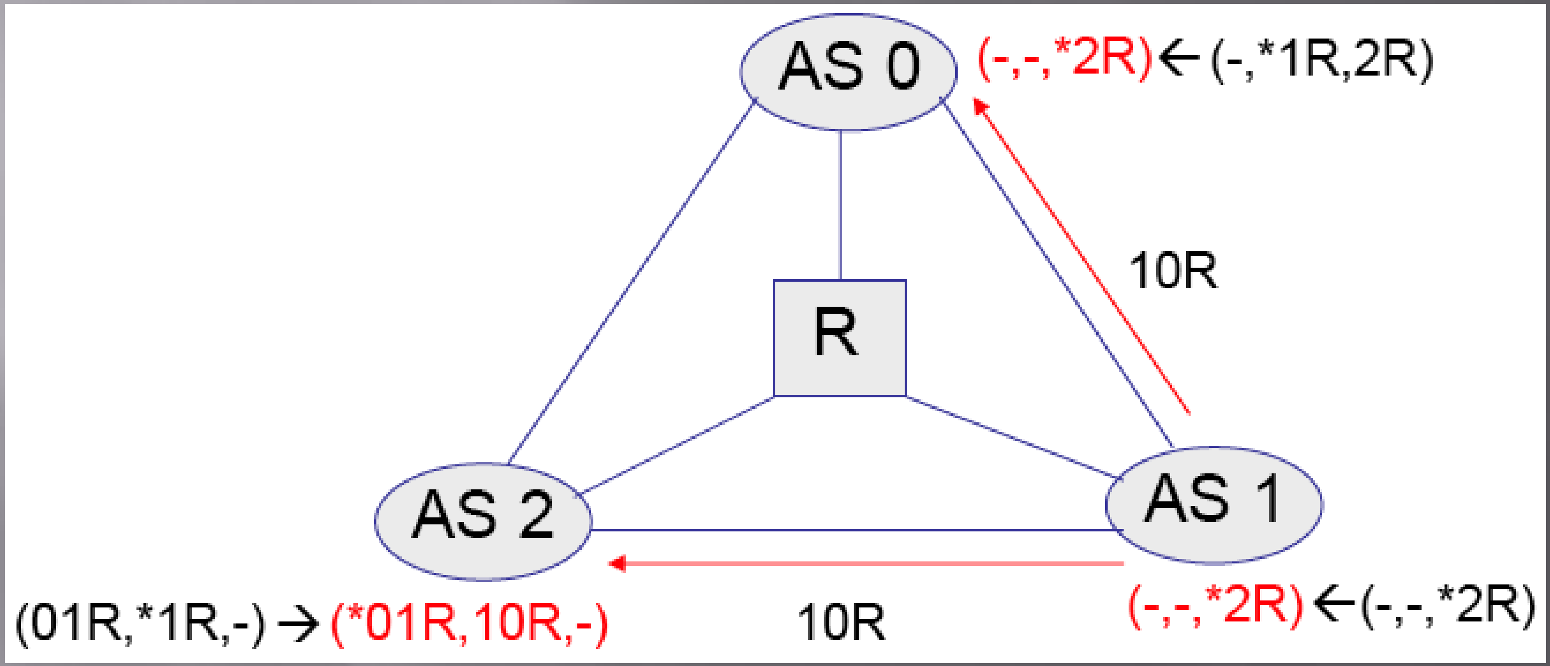
- **Route oscillation**
  - If policy just uses ASPATH length, then looks like a DV protocol; can still oscillate
  - Loop prevention doesn't prevent this
  - Can explore every possible path through network → (n-1)! Combinations
- **Timers**
  - 30MinAdver timer makes router wait between updates
  - Adds artificial "rounds" to propagation speed
- **Loop detection**
  - Waits to send routes that have loops in them (prevailing BGP implementation only does receiver based loop detection)
- **Typical Internet failover times**
  - New/shorter link → 60 seconds: simple replacement at nodes
  - Down link → 180 seconds: search of possible options
  - Longer link → 120 seconds: replacement or search based on length

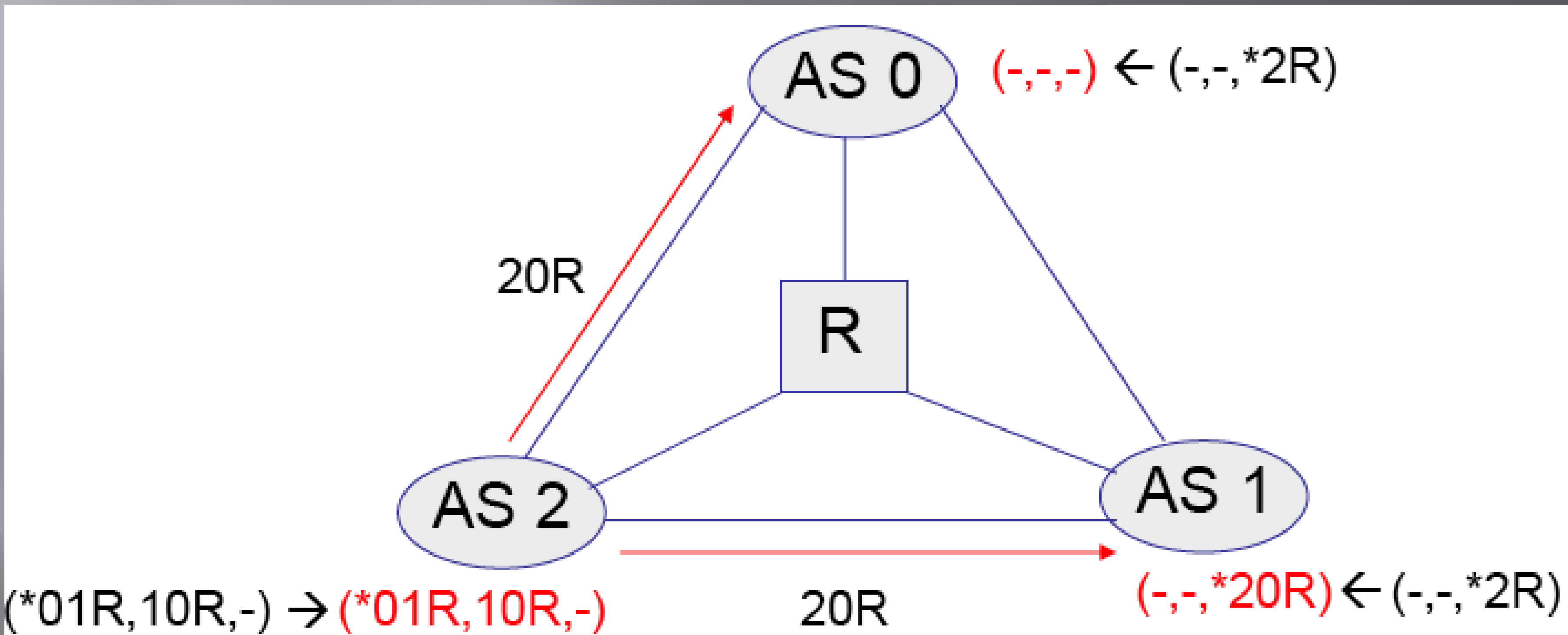
Example BGP Oscillation

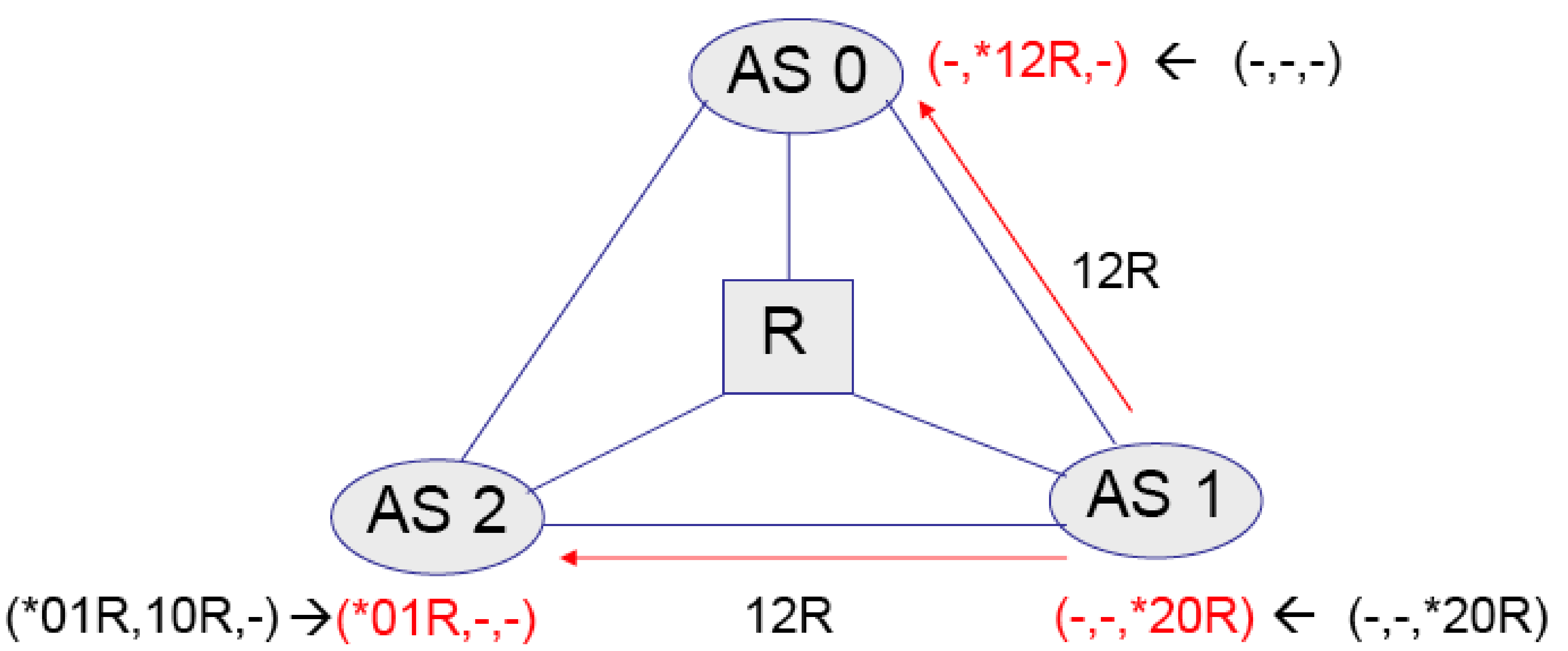




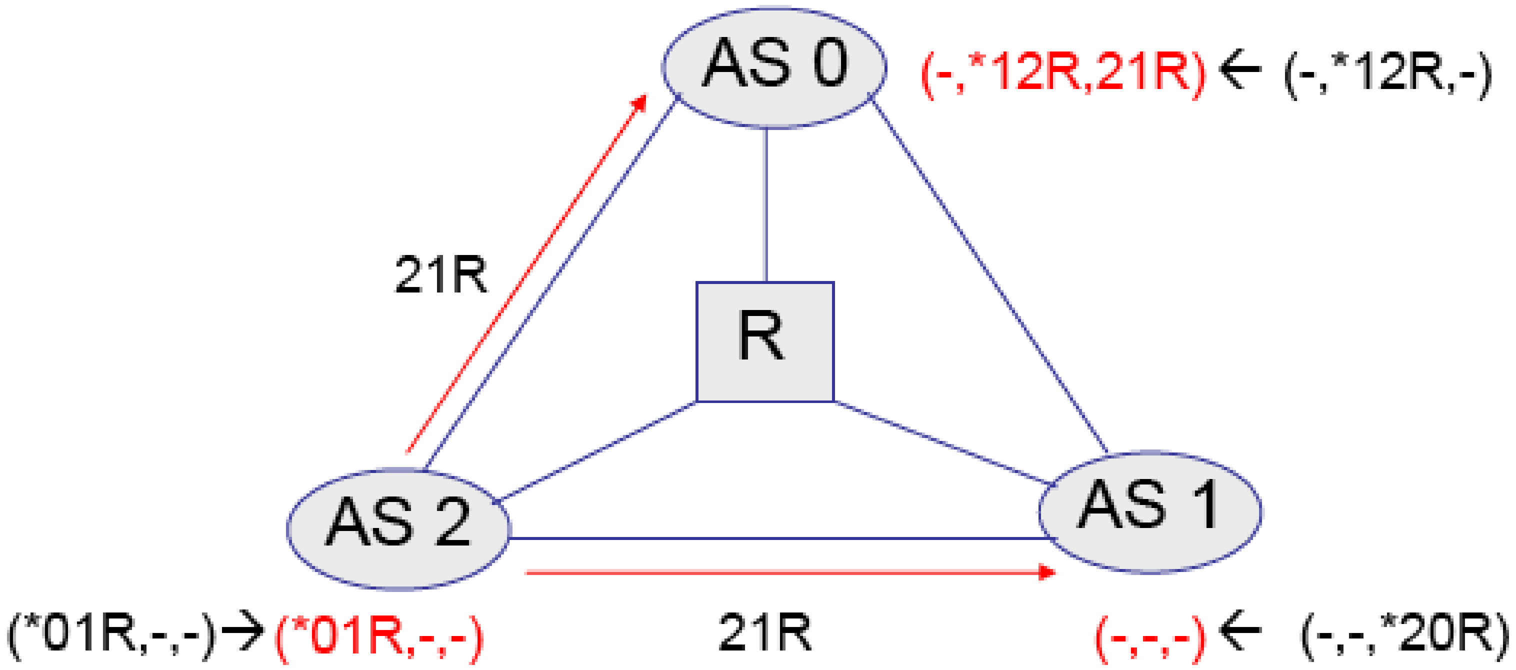
AS 0
(-,*1R,2R)←(-,*1R,2R)
01R
R
01R
AS 2
AS 1
(*0R,1R,-)→(01R,*1R,-)
(-,-,*2R)←(*0R,-,2R)


AS 0
(-,-,*2R)←(-,*1R,2R)
R
10R
AS 2
AS 1
(01R,*1R,-)→(*01R,10R,-)
10R
(-,-,*2R)←(-,-,*2R)


AS 0
(-,-,-) ← (-,-,*2R)
20R
R
AS 2
AS 1
(*01R,10R,-) → (*01R,10R,-)
20R
(-,-,*20R) ← (-,-,*2R)


AS 0
(-,*12R,-) ← (-,-,-)
12R
R
AS 2
AS 1
12R
(*01R,10R,-) →(*01R,-,-)
(-,-,*20R) ← (-,-,*20R)

Example Oscillation



BGP impact on flow

- Slides by Sharad Agarwal

Traffic Collection

- Packet-level data
  - Ingress, OC-12
  - TCP/IP header
  - GPS synched timestamp

- Analyzed 6 traces
  - 2 PoPs, 3 days, 6 different ASes