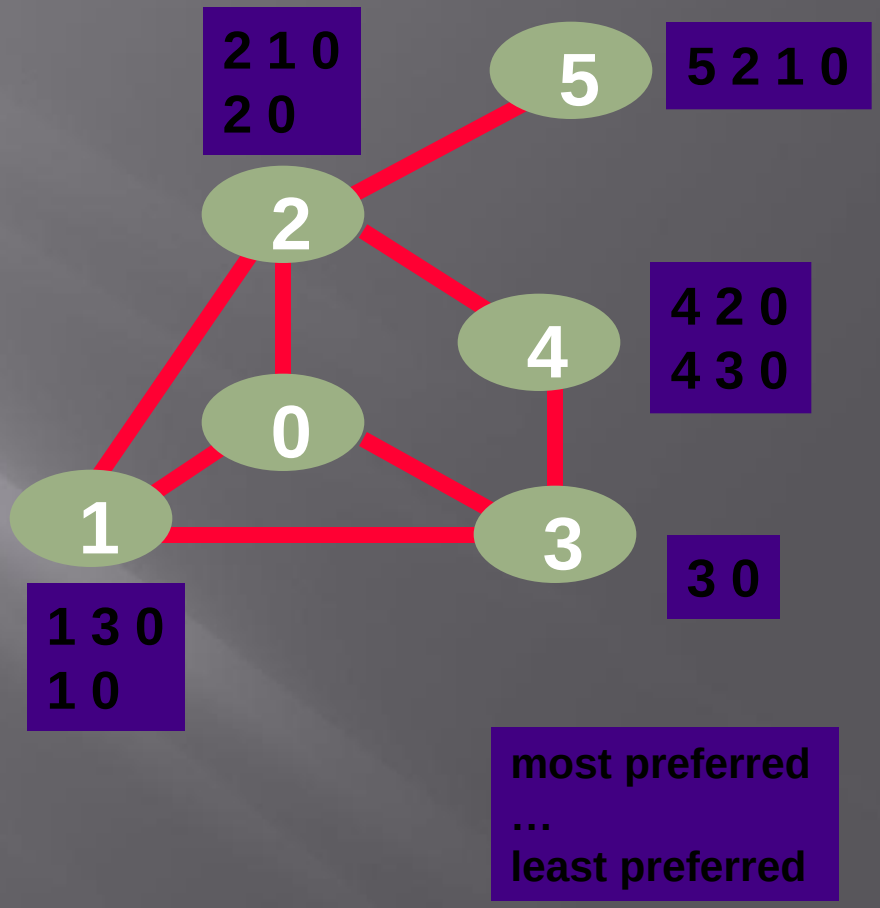
IP ROUTING

Bad News

- **BGP is not guaranteed to converge on a stable routing. Policy interactions could lead to "livelock" protocol oscillations.**
  See "Persistent Route Oscillations in Inter-domain Routing" by K. Varadhan, R. Govindan, and D. Estrin. ISI report, 1996
- **Corollary: BGP is not guaranteed to recover from network failures.**

An instance of the *Stable Paths Problem* (SPP)

- A graph of nodes and edges,
- Node 0, called *the origin*,
- For each non-zero node, a set or permitted paths to the origin. This set always contains the "null path".
- A ranking of permitted paths at each node. Null path is always least preferred. (Not shown in diagram)



1

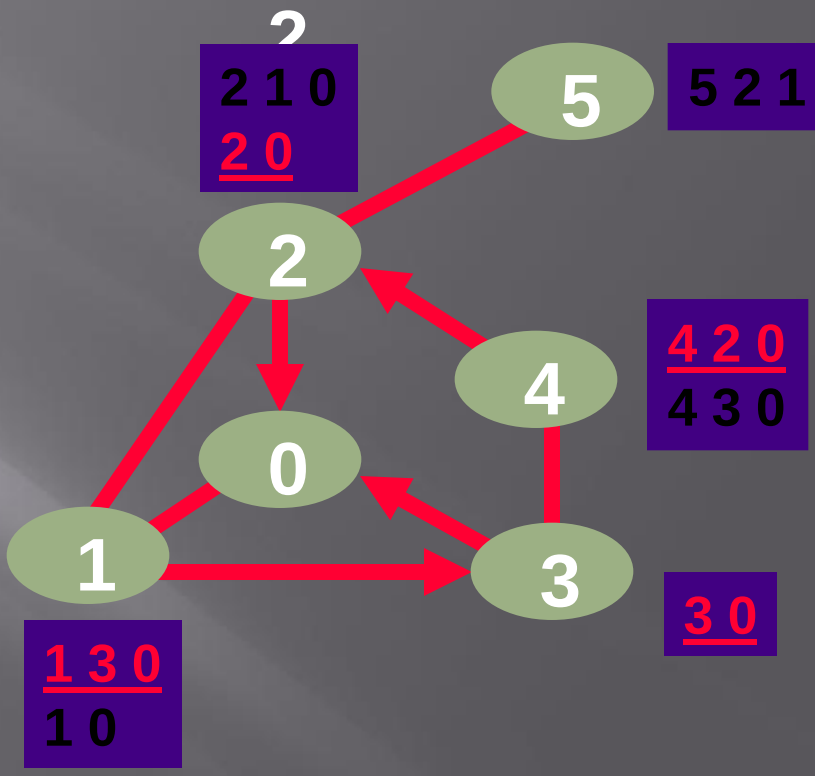
When modeling BGP : nodes represent BGP speaking routers, and 0 represents a node originating some address block

Yes, the translation gets messy!

A Solution to a Stable Paths Problem

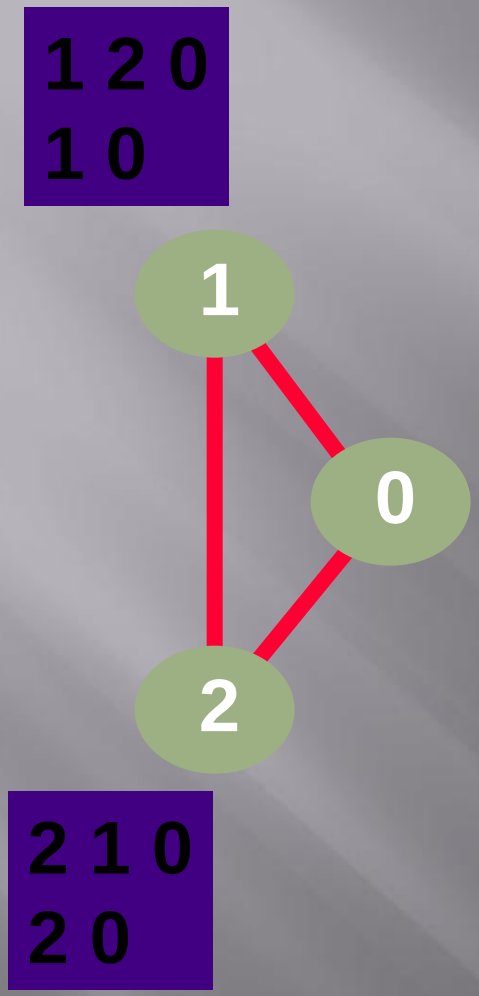
A *solution* is an assignment of permitted paths to each node such that

- node u's assigned path is either the null path or is a path uwP, where wP is assigned to node w and {u,w} is an edge in the graph,
- each node is assigned the highest ranked path among those consistent with the paths assigned to its neighbors.

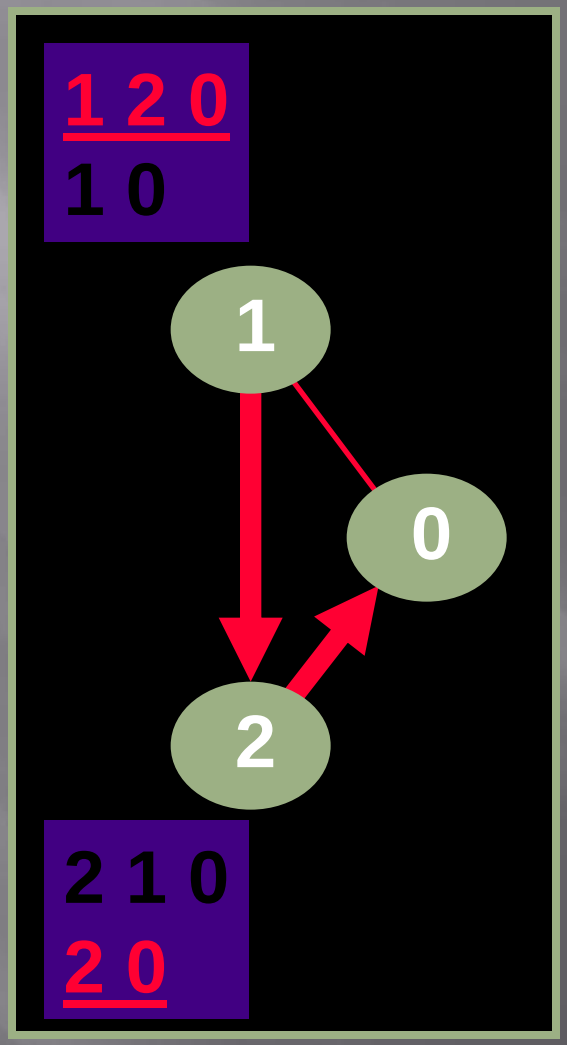


A Solution need not represent a shortest path tree, or a spanning tree.

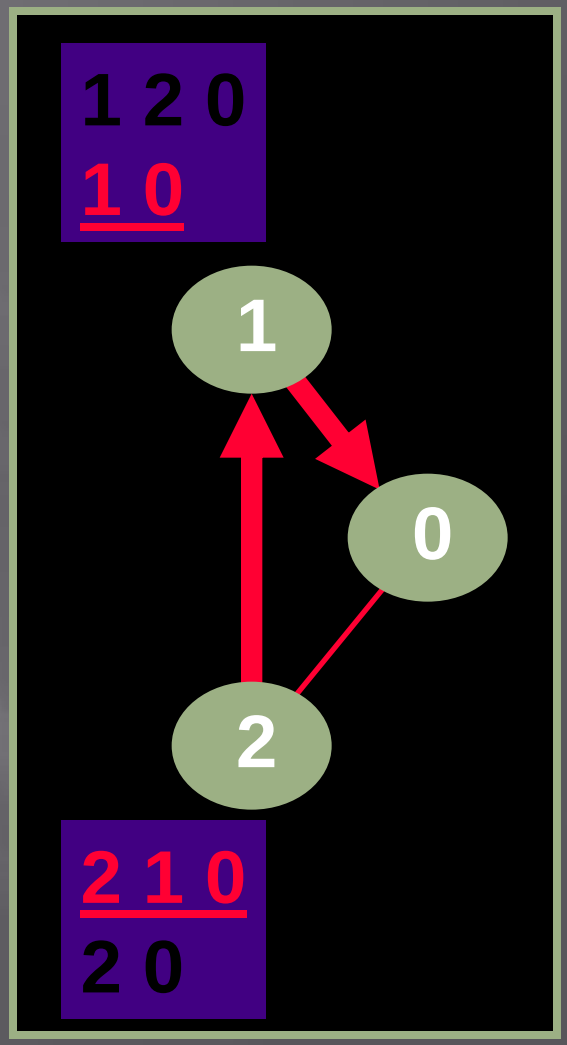
An SPP may have multiple solutions



DISAGREE

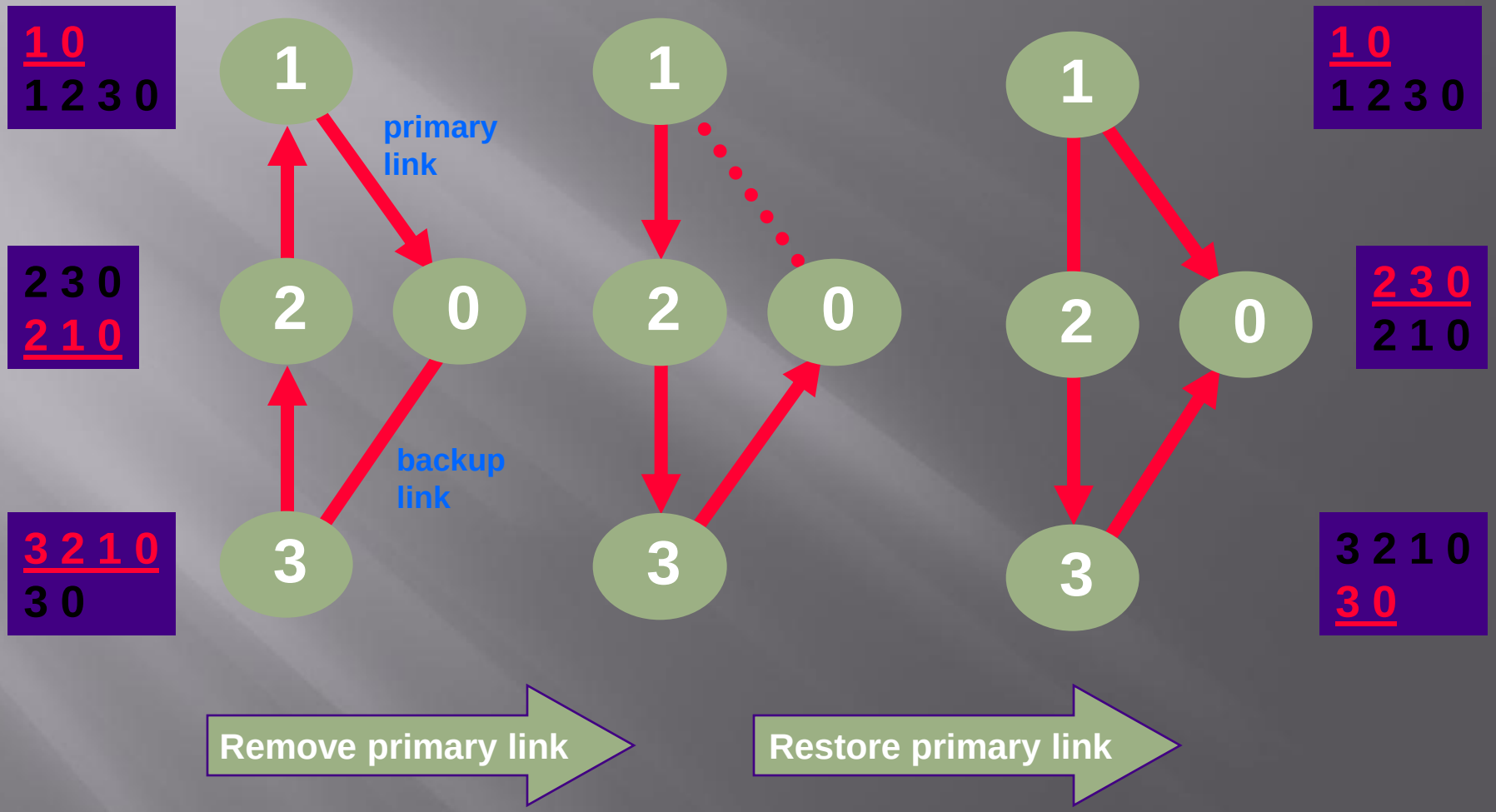


First solution

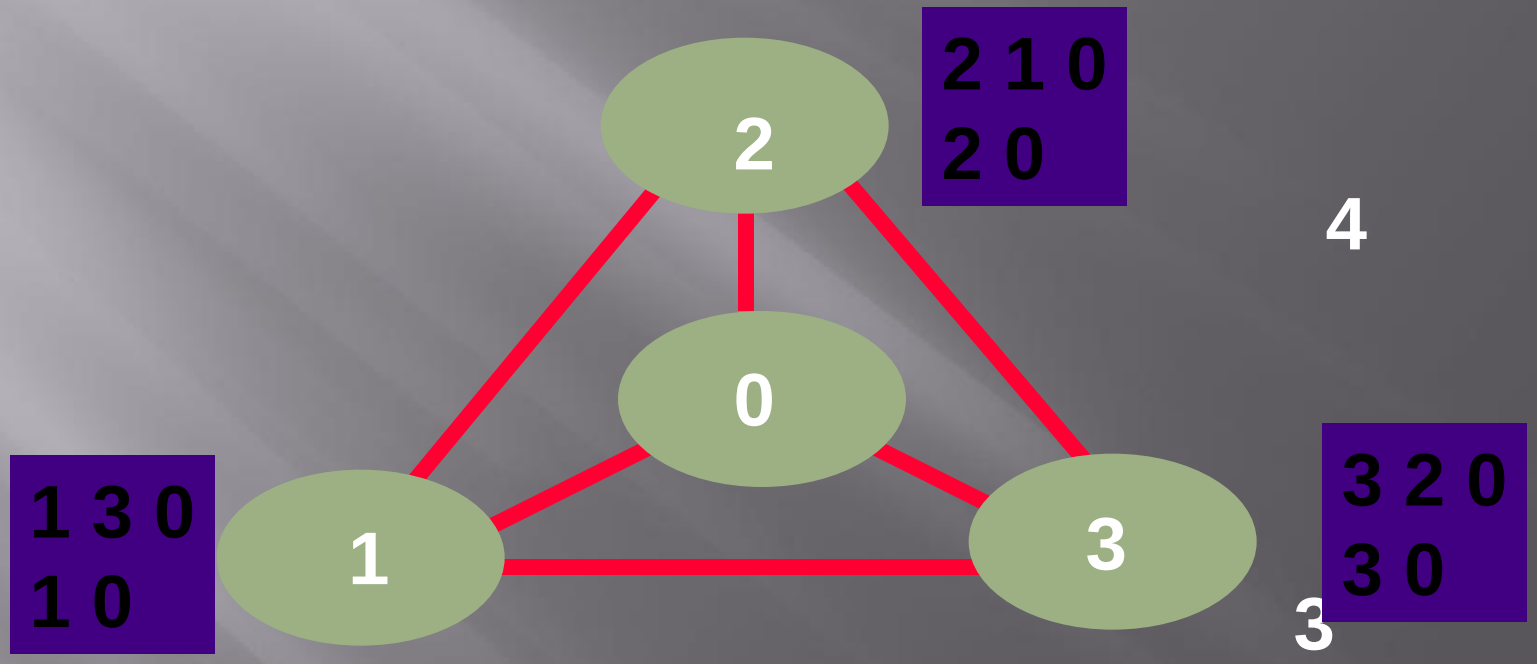


Second solution

Multiple solutions can result in "Route Triggering"



BAD GADGET : No Solution



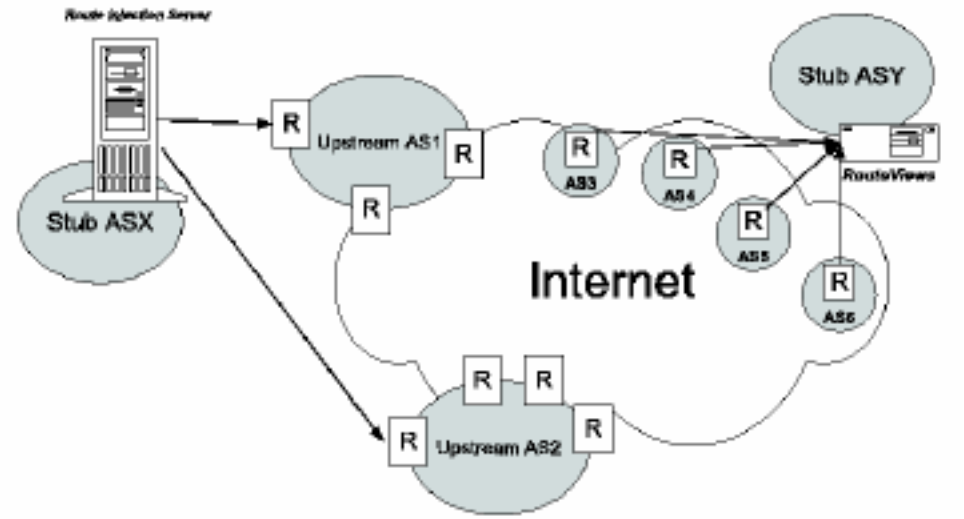
Persistent Route Oscillations in Inter-Domain Routing. Kannan Varadhan, Ramesh Govindan, and Deborah Estrin. Computer Networks, Jan. 2000

Labovitz 00 (slides by S. Savage)

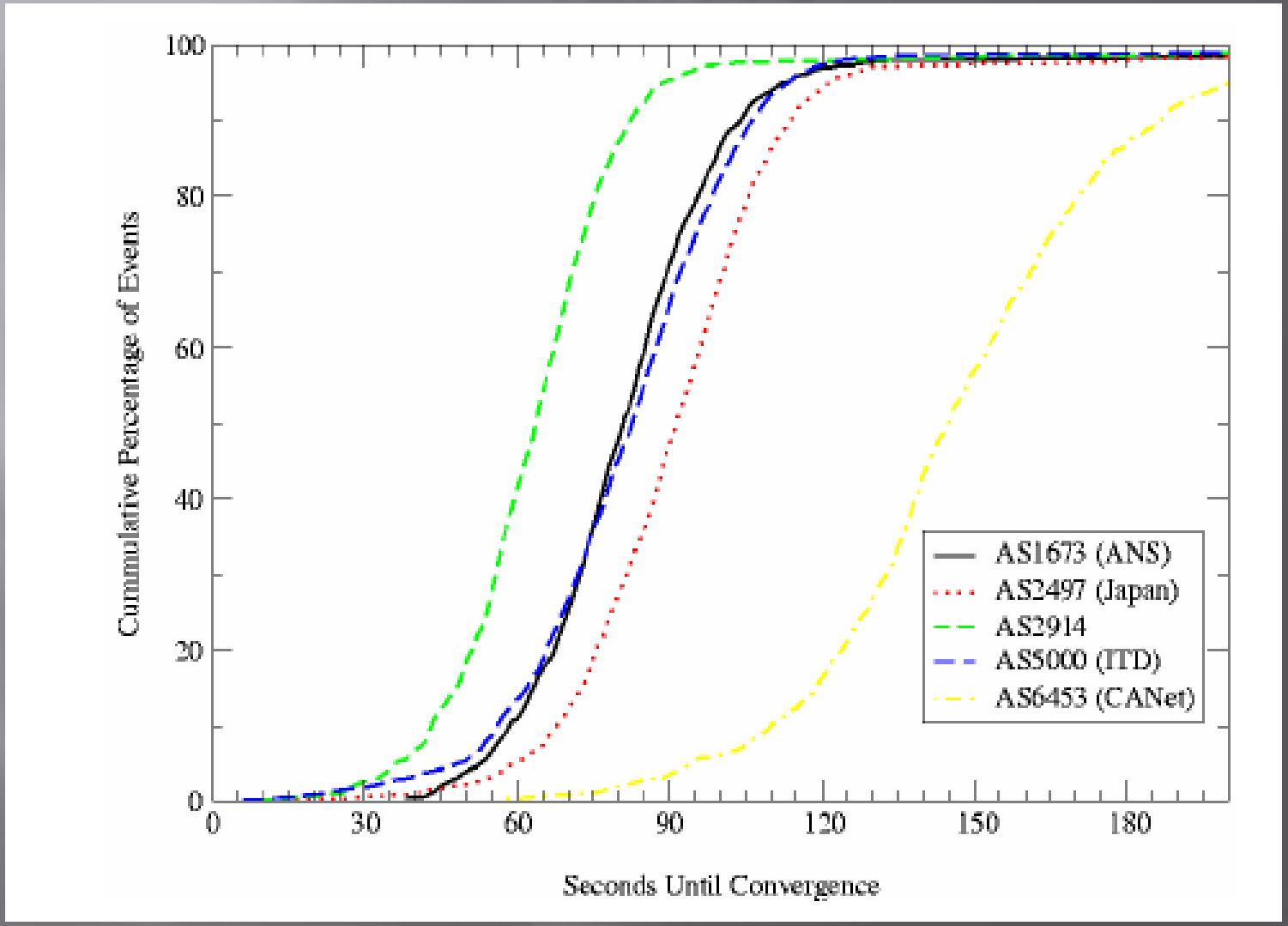
- When a node/link fails, how quickly can BGP recover?
  - Common wisdom
    - Very fast, a few milliseconds
    - Route withdrawals sent immediately
  - ASPATH loop detection eliminates problems with DV
  - Reality somewhat different…
- Meta-issue
  - Does BGP ever converge?
  - Griffin et al00 show that with unconstrained policies it doesn't have to and its NP-hard to tell if it does
  - However, Labovitz et al, deal with constrained policies that do converge

Active Route Measurement

- Inject routes into geographically and topologically points in the network
- Periodically fail and change these routes
- Time events using ICMP echos (ping), HTTP GET and time-synchronized monitoring machines



Time to converge after route failure



Time to Repair or Failover