BGP policy interactions can be difficult to debug

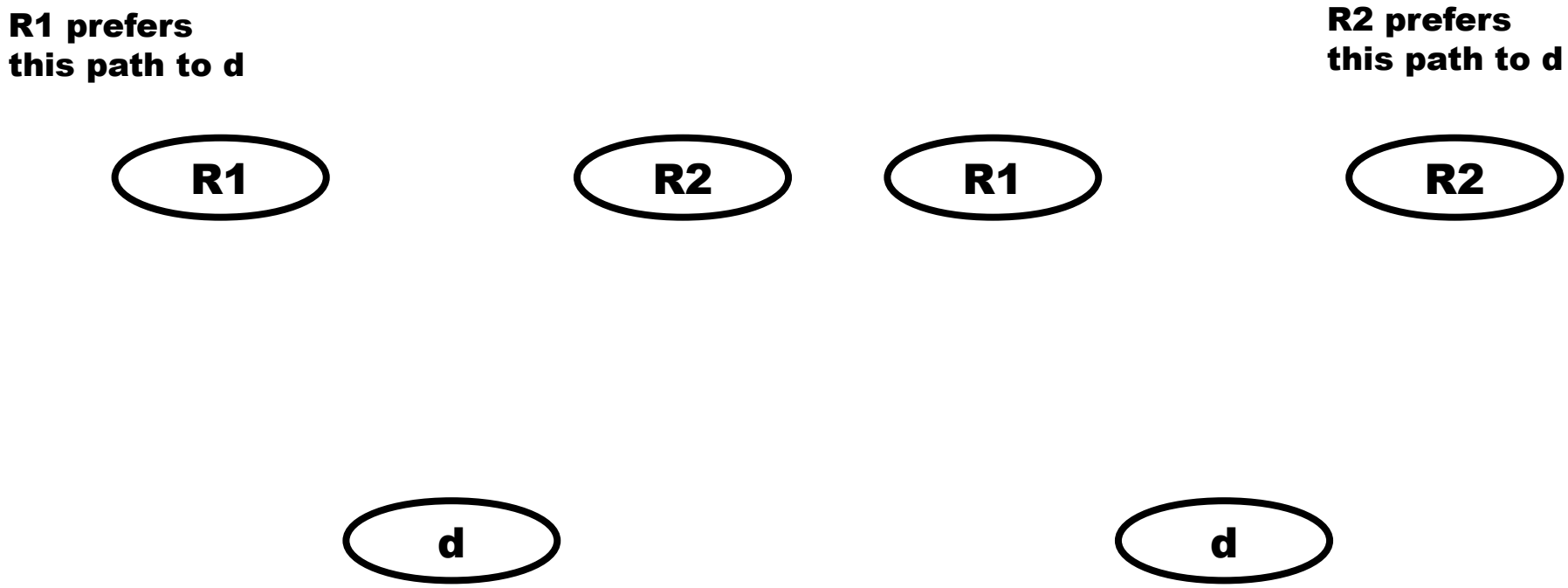
(BGP Wedgies, RFC 4264)

Complexity can arise in unexpected ways

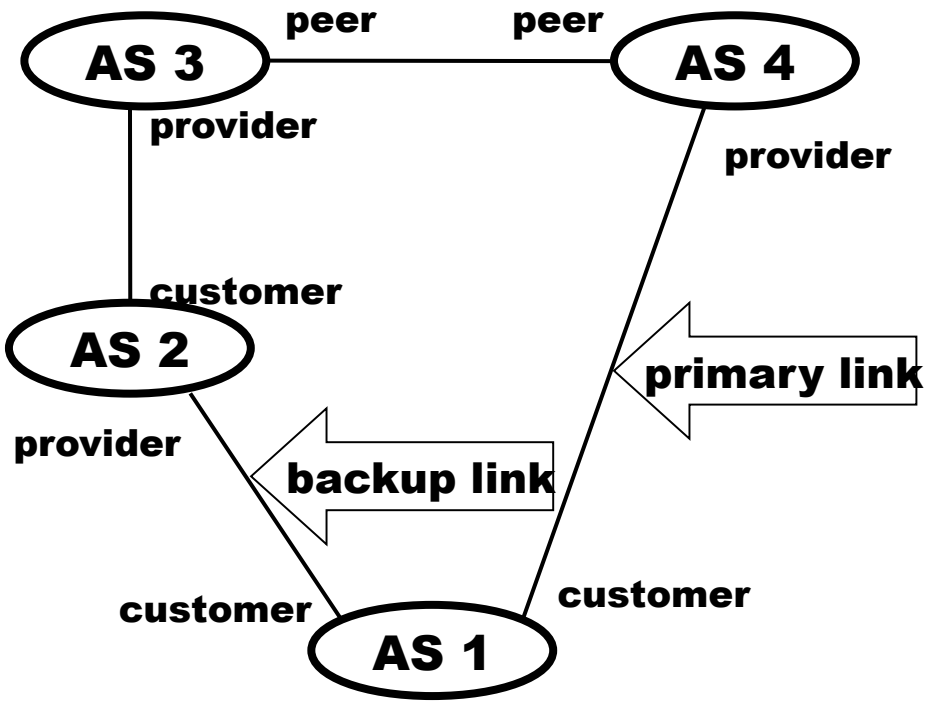
Timothy G. Griffin
University of Cambridge
tgg22@cam.ac.uk

IETF 83 --- Paris
Network Complexity Research Group (ncrg)
March 26 2012

Multiple Stable States

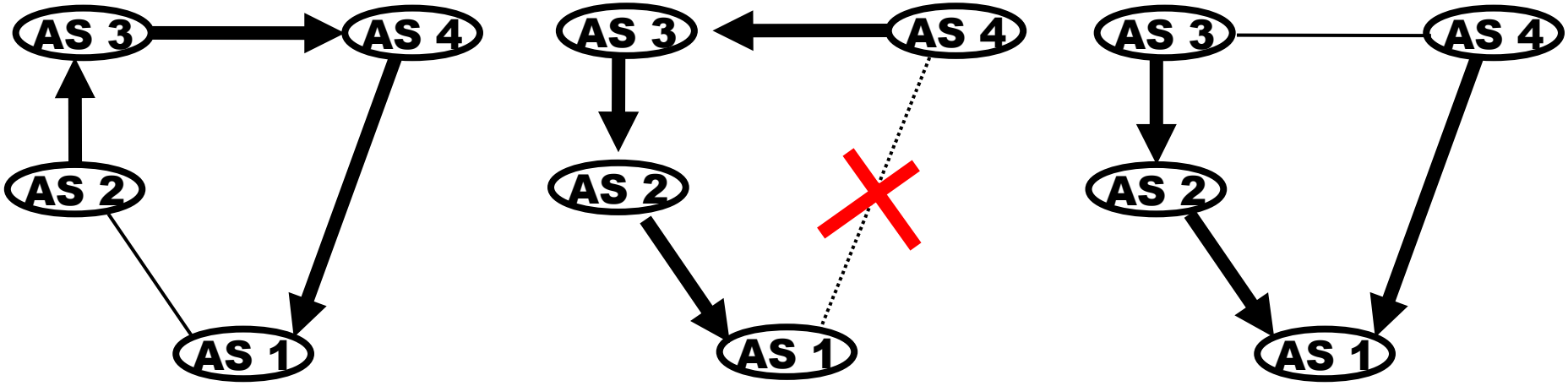


¾ Wedgie Example



- AS 1 implements backup link by sending AS 2 a "depref me" community.
- AS 2 implements this community so that the resulting local pref is below that of routes from it's upstream provider (AS 3 routes)

Getting wedged...



Happy, happy, joy, joy

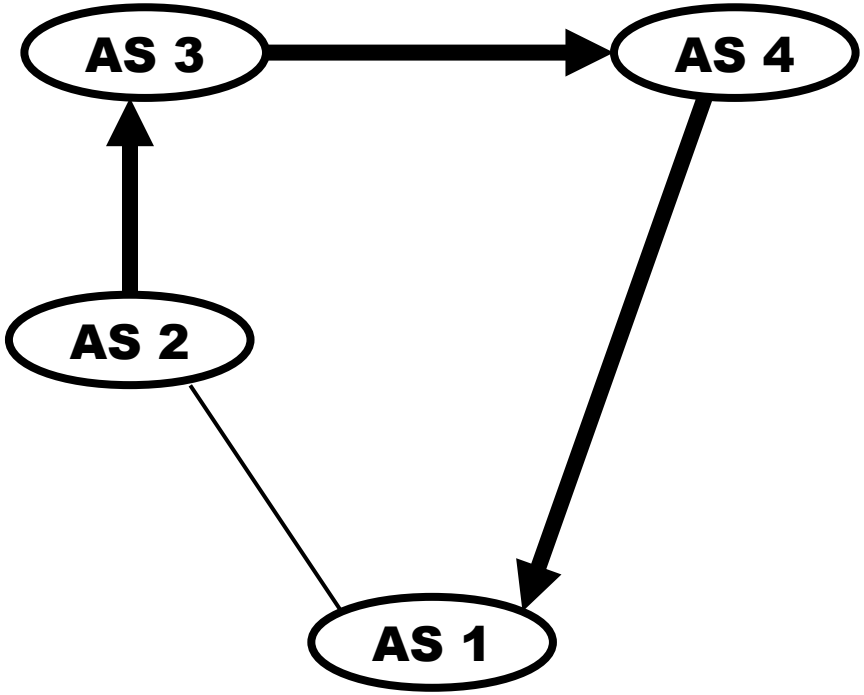
Backups are good!

OH NO, I'M WEDGED!

Primary fails

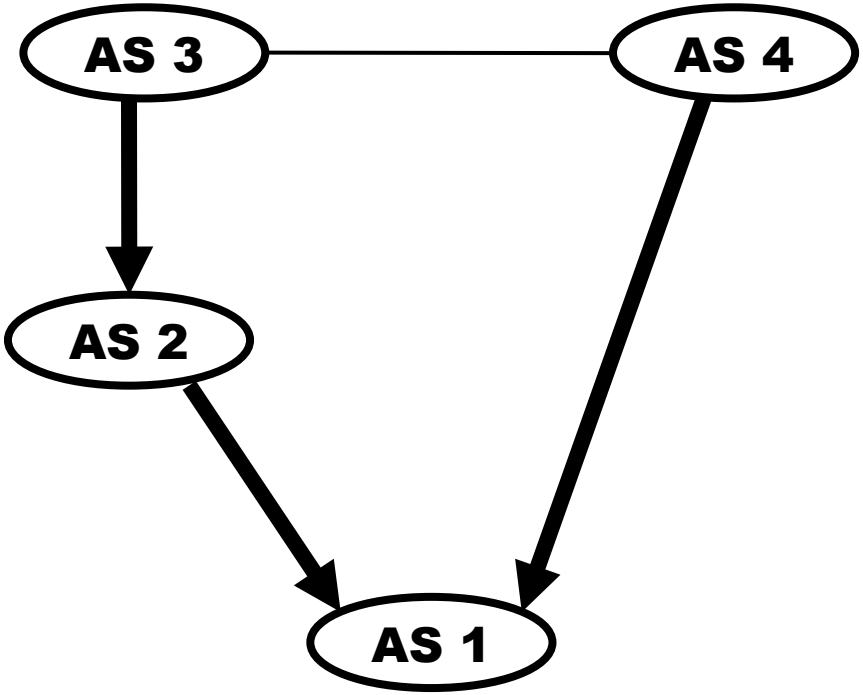
Primary comes back up!

And the Routings are…



Intended Routing

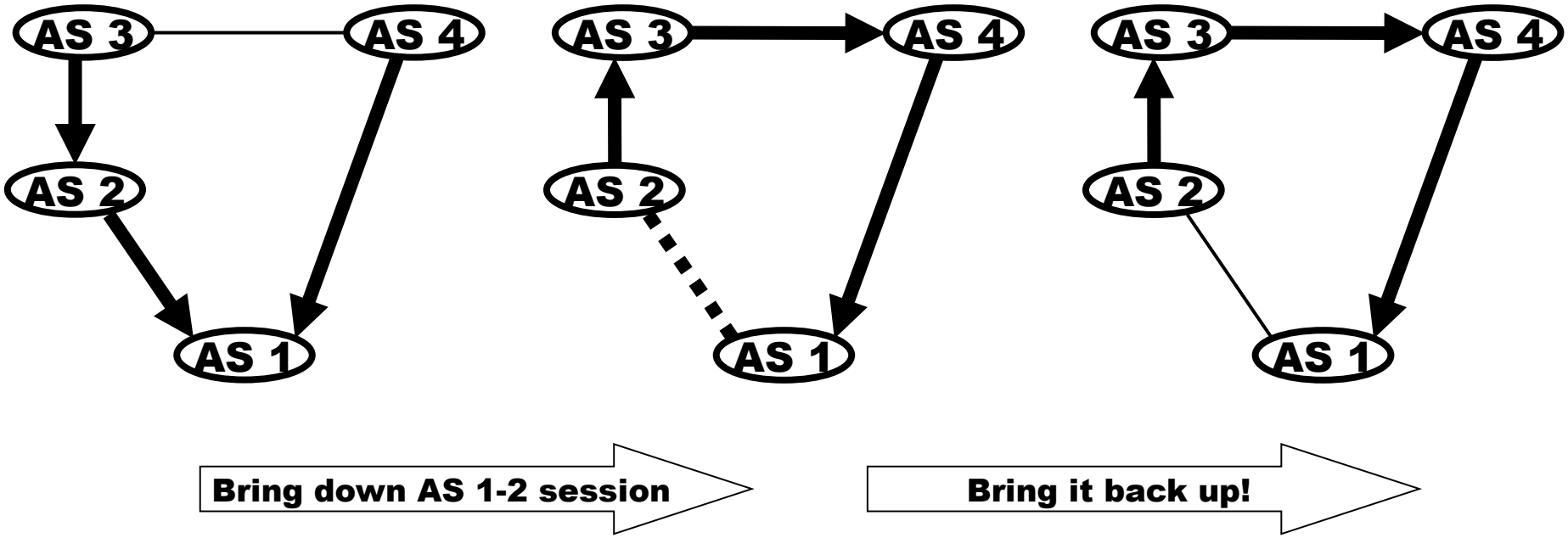
Note: this would be the ONLY routing if AS2 translated its "depref me" community to a "depref me" community of AS 3



Unintended Routing

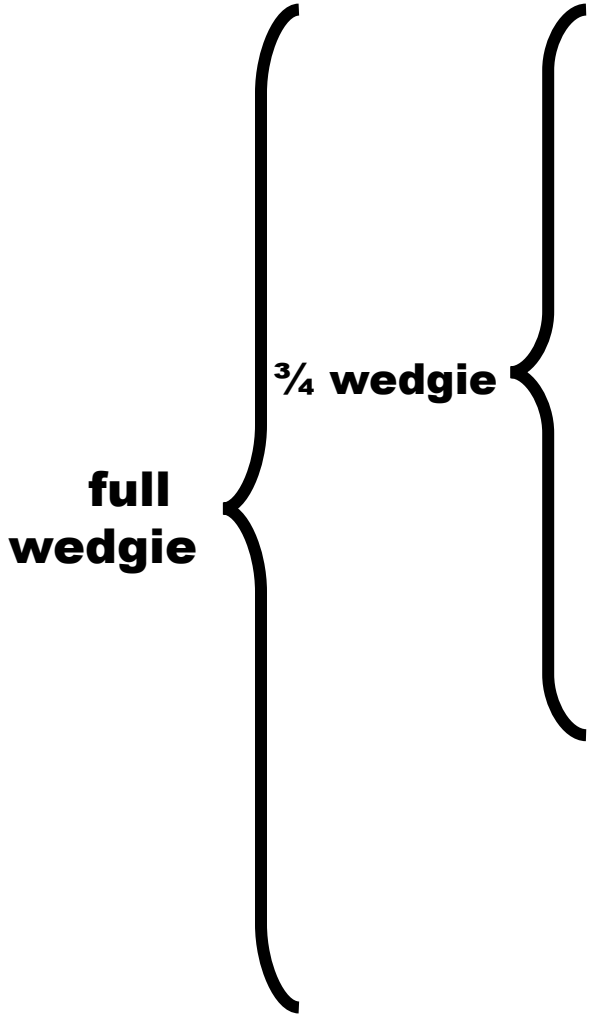
Note: This is easy to reach from the intended routing just by "bouncing" the BGP session on the primary link.

Recovery



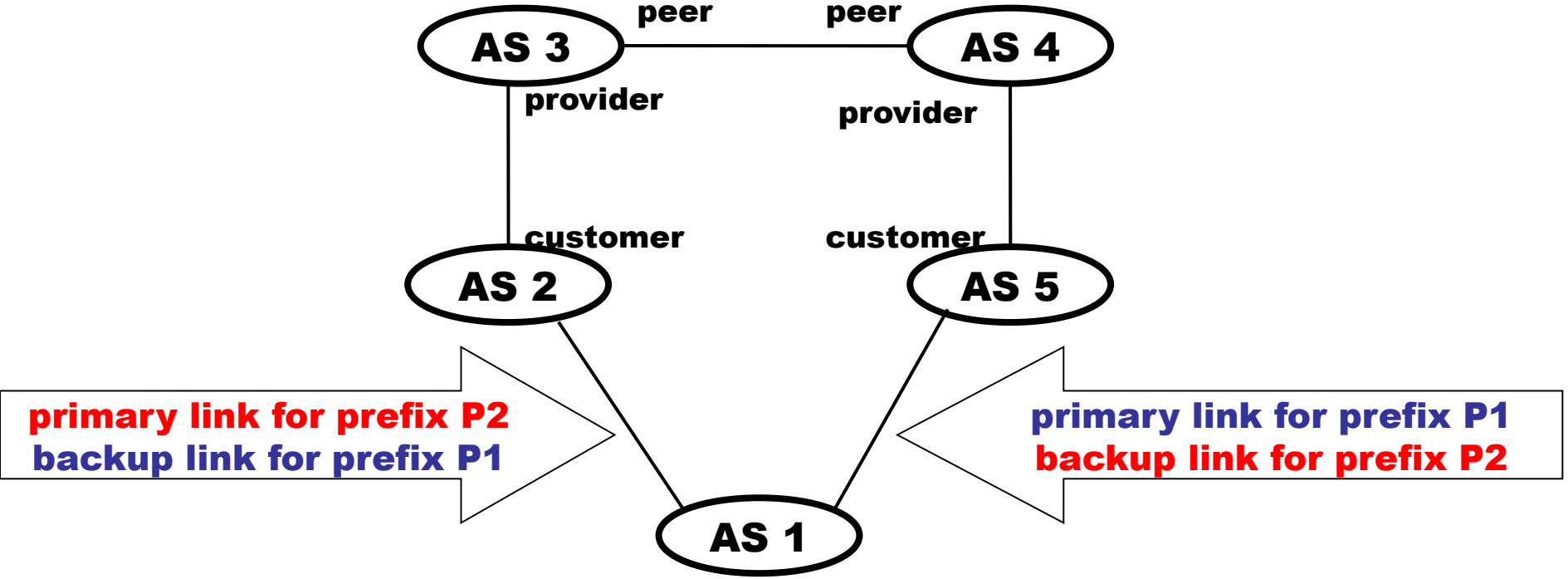
- **Requires manual intervention**
- **Can be done in AS 1 or AS 2**

What is a BGP Wedgie?



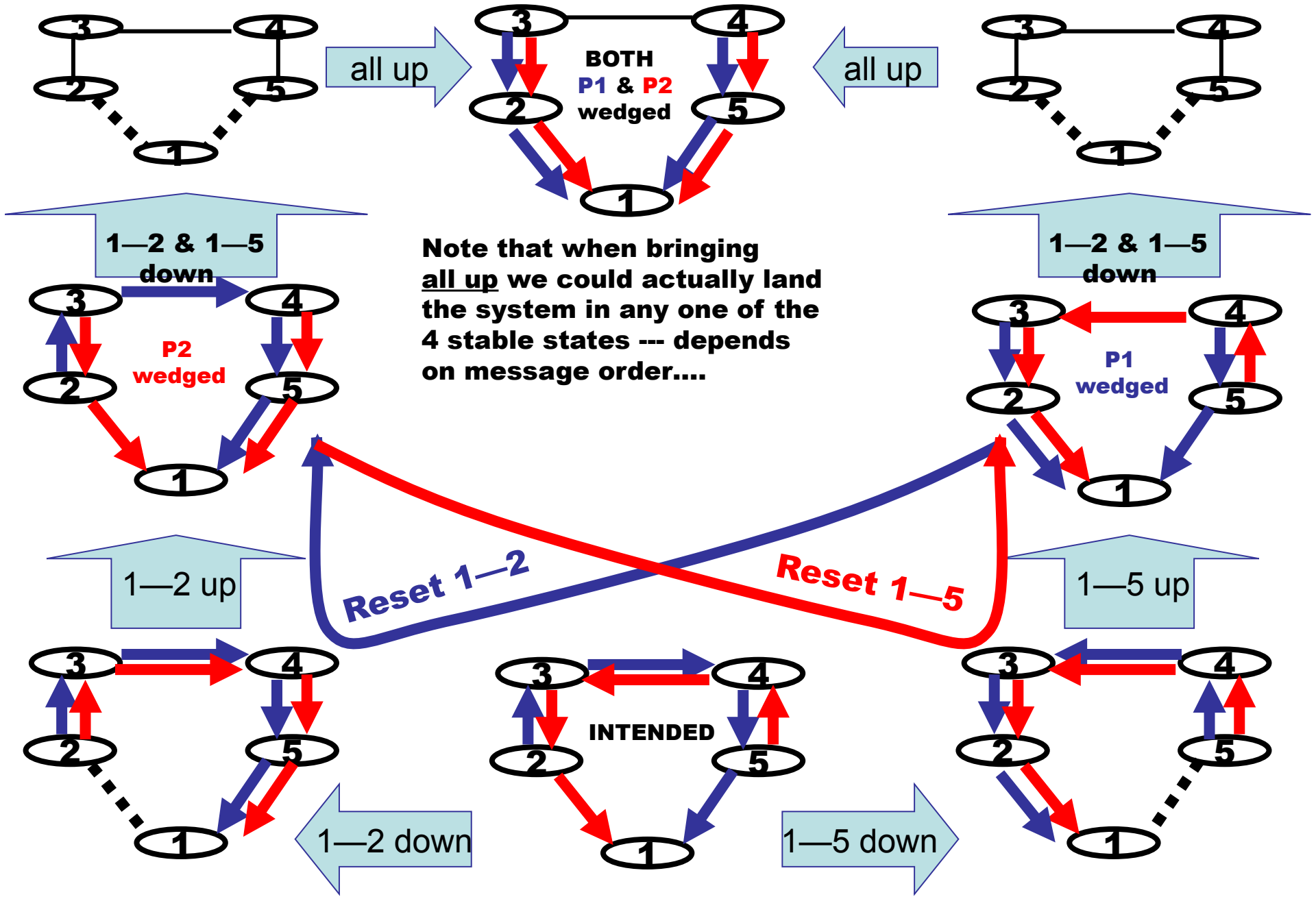
- BGP policies make sense locally
- Interaction of local policies allows multiple stable routings
- Some routings are consistent with intended policies, and some are not
  - If an unintended routing is installed (BGP is "wedged"), then manual intervention is needed to change to an intended routing
- When an unintended routing is installed, no single group of network operators has enough knowledge to debug the problem

Load Balancing Example

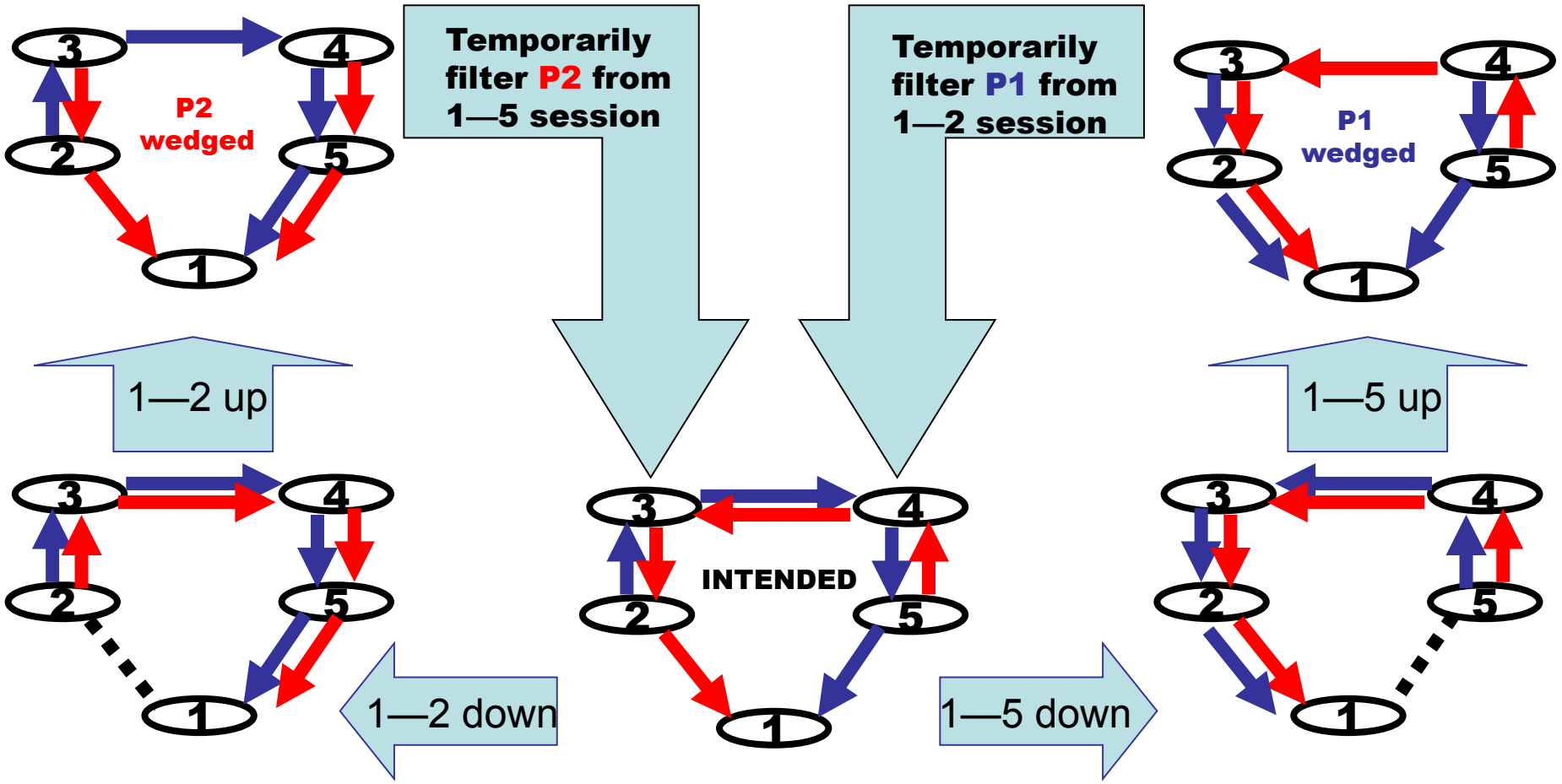


Simple session reset my not work!!

Can’t un-wedge with session resets!

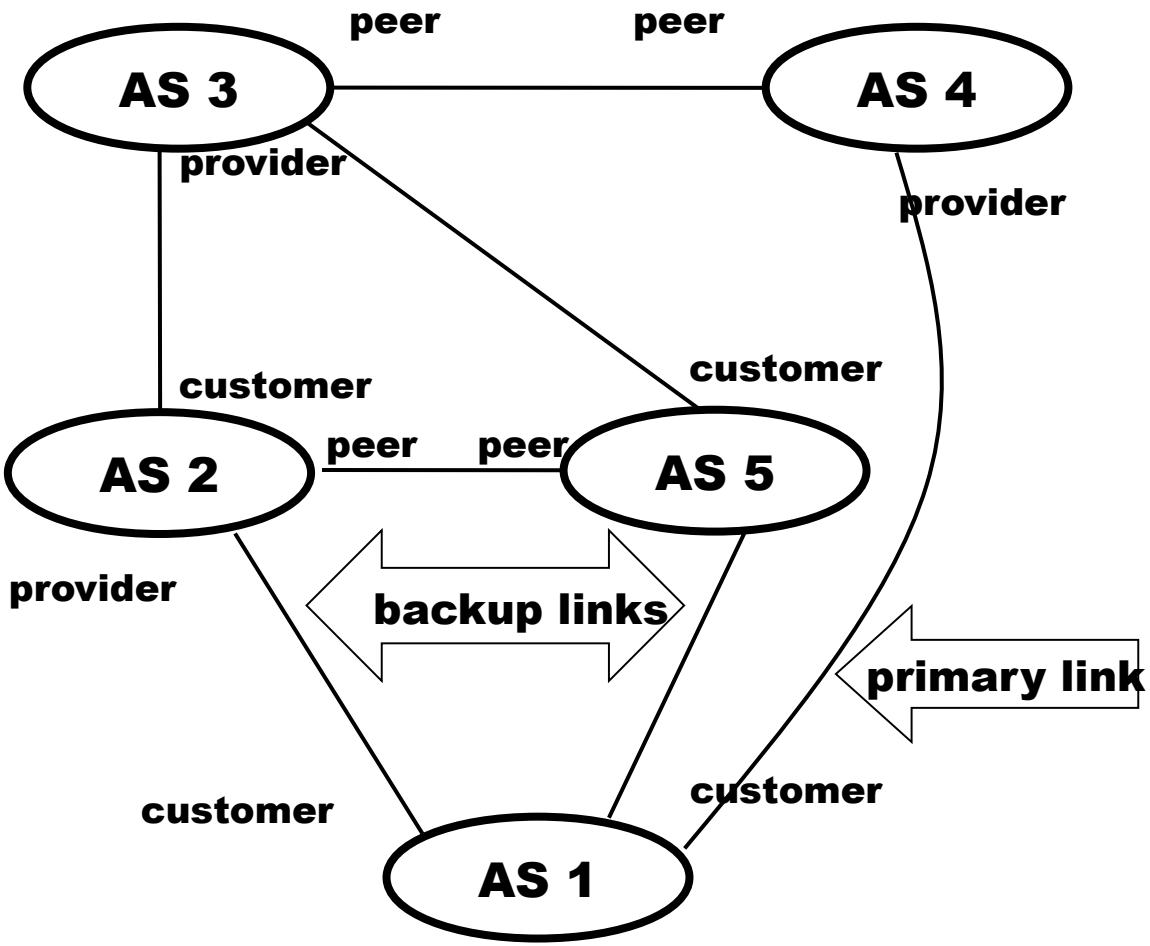


Recovery

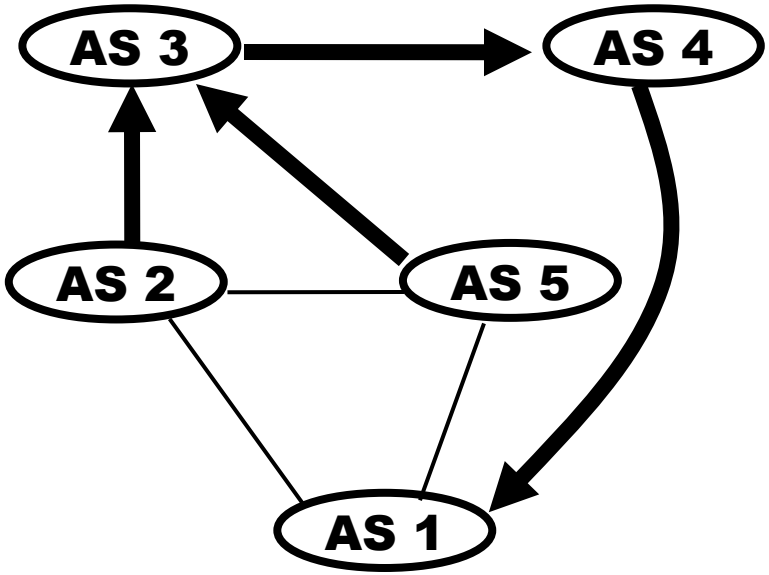


Who among us could figure this one out?
When 1—2 is in New York and 1—5 is in Tokyo?

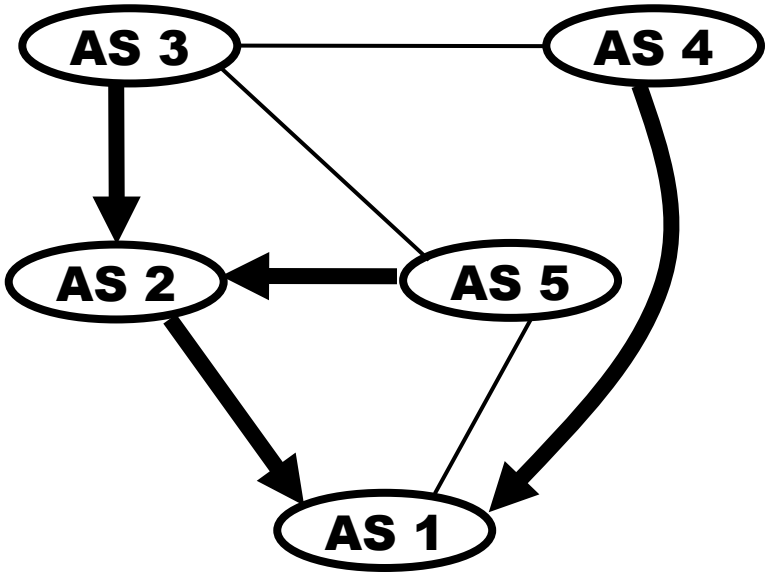
Full Wedgie Example



And the Routings are…

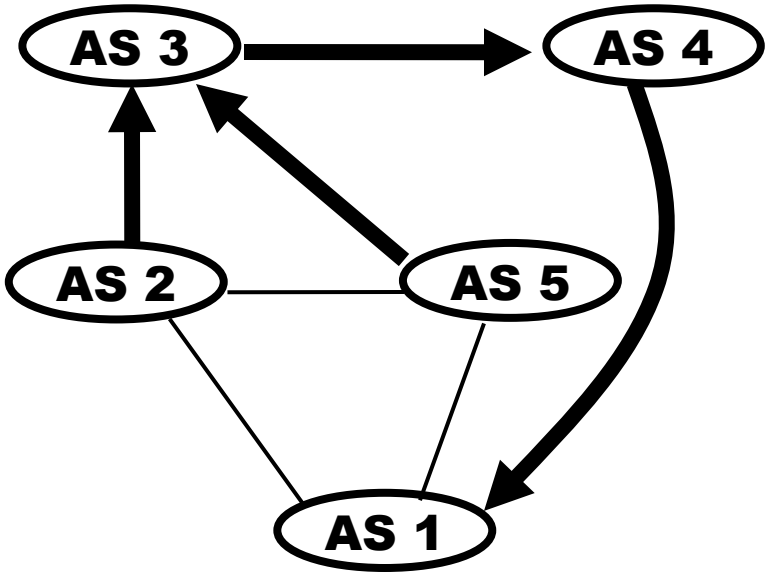


Intended Routing

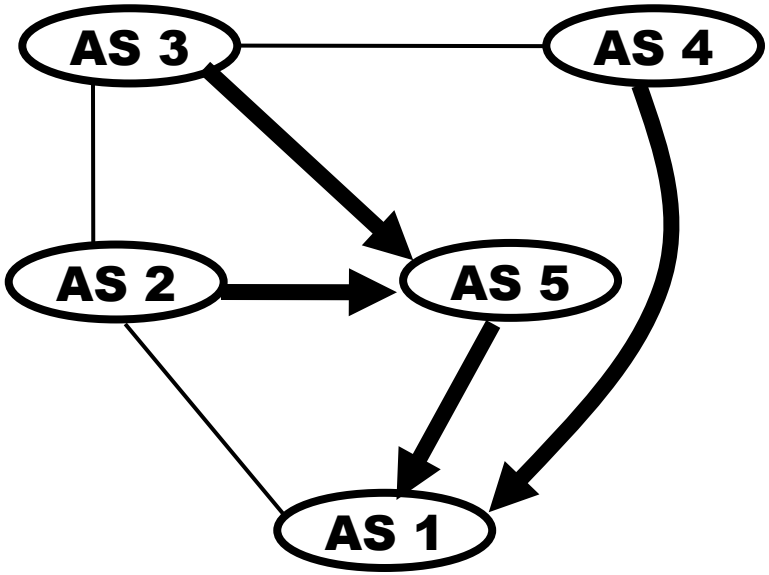


Unintended Routing 1

And the Routings are…

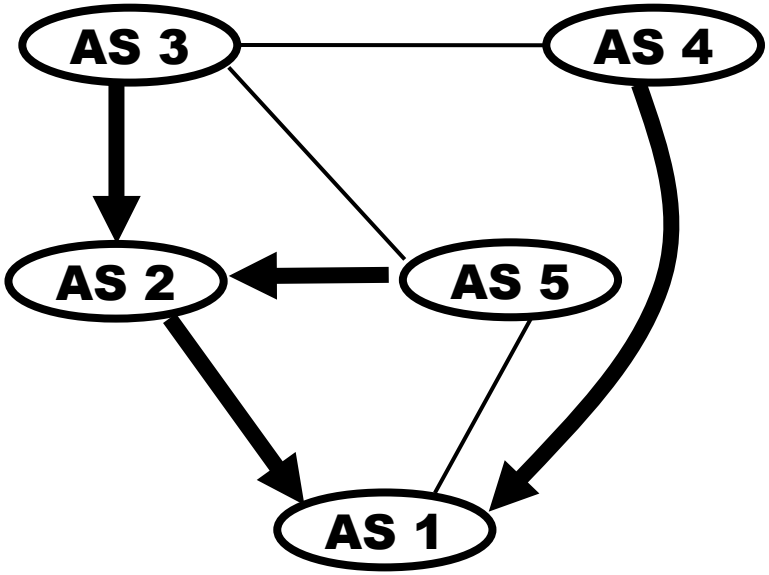


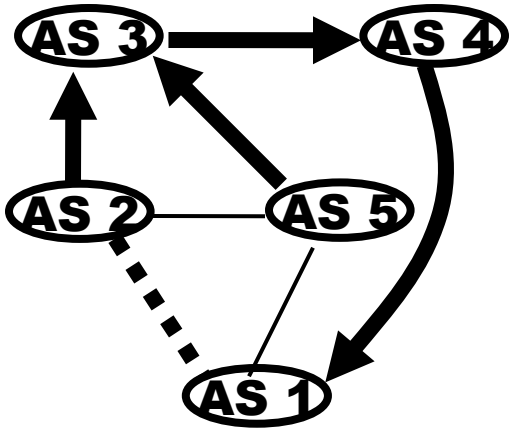
Intended Routing



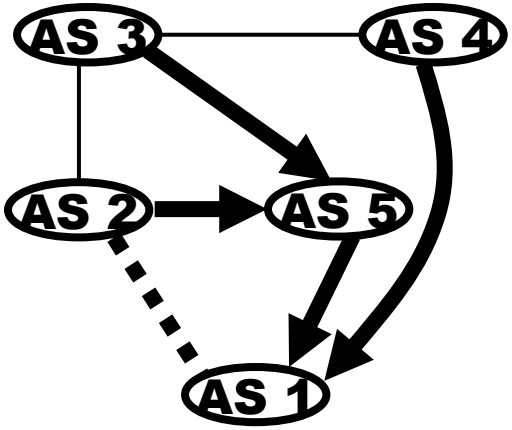
Unintended Routing 2

Resetting 1—2 may not help!!

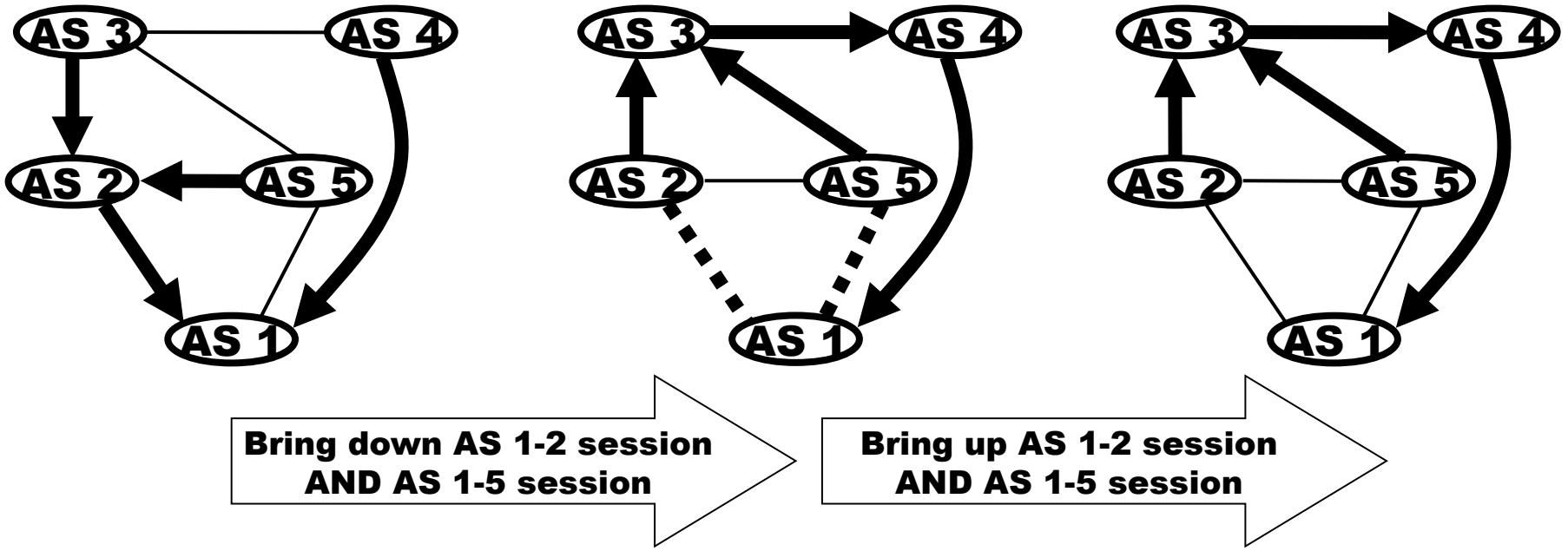




OR



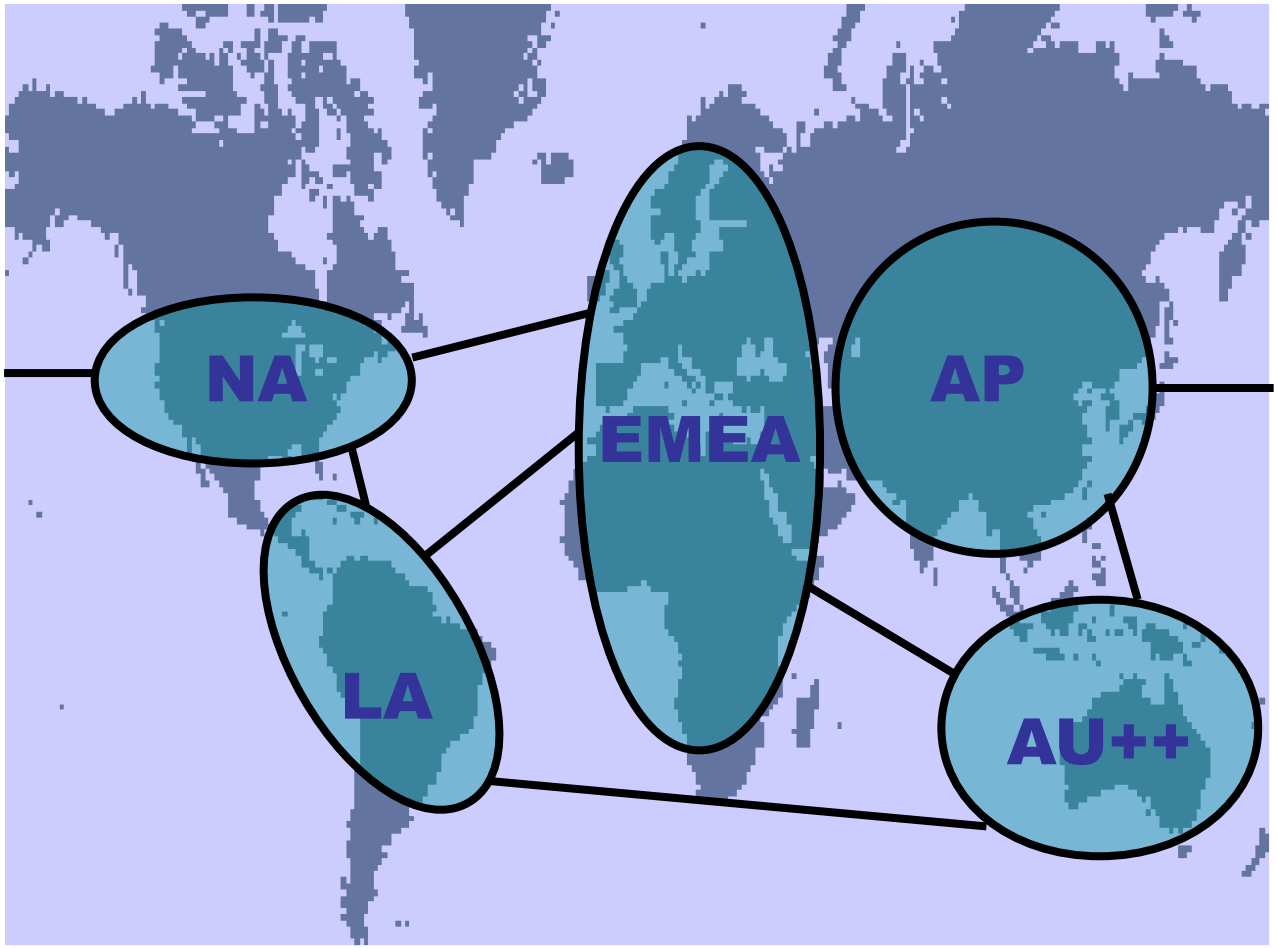
Guaranteed Recovery



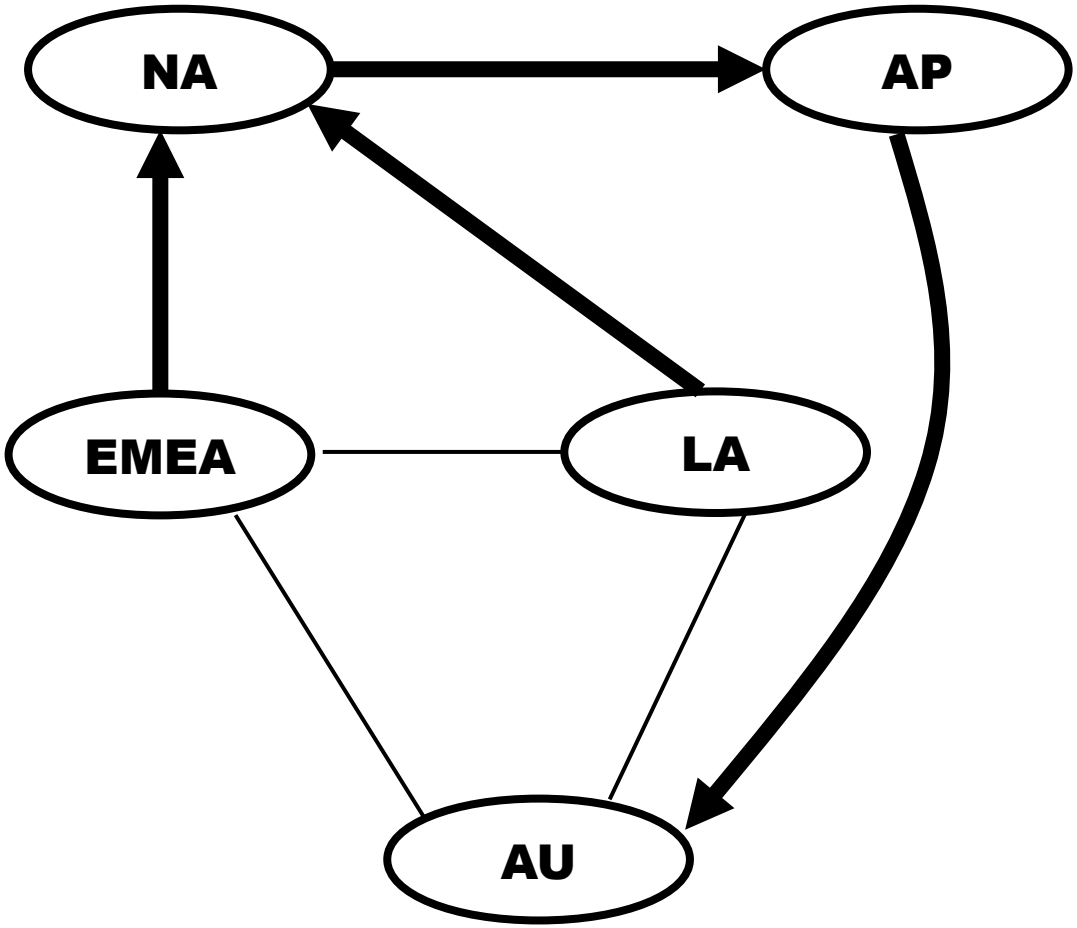
A lot of non-local knowledge is required to arrive at this recovery strategy!

Try to convince AS 5 that their session has be reset (or filtered) even though it is not associated with an active route!

Same problems can arise with “traffic engineering” across regional networks.



Look familiar?



Intended Routing for some prefixes in AU….

What is going on?

- There is no guarantee that a BGP configuration has a unique routing solution.
  - When multiple solutions exist, the (unpredictable) order of updates will determine which one is wins.
- Complex policies (weights, communities setting preferences, and so on) increase chances of routing anomalies.
  - … yet this is the current trend!

How to manage?

- Study the interactions in the wild
  - Several research groups working on this …
  - Could more edu/research nets follow example of Internet2 and publish configs?
- Guidelines for configuration?
  - This may be as simple as translate depref me communities in a consistent way
  - Or it may be more complicated, depending on what ISPs are actually doing …
- Unsolved research problem: autonomy vs global sanity