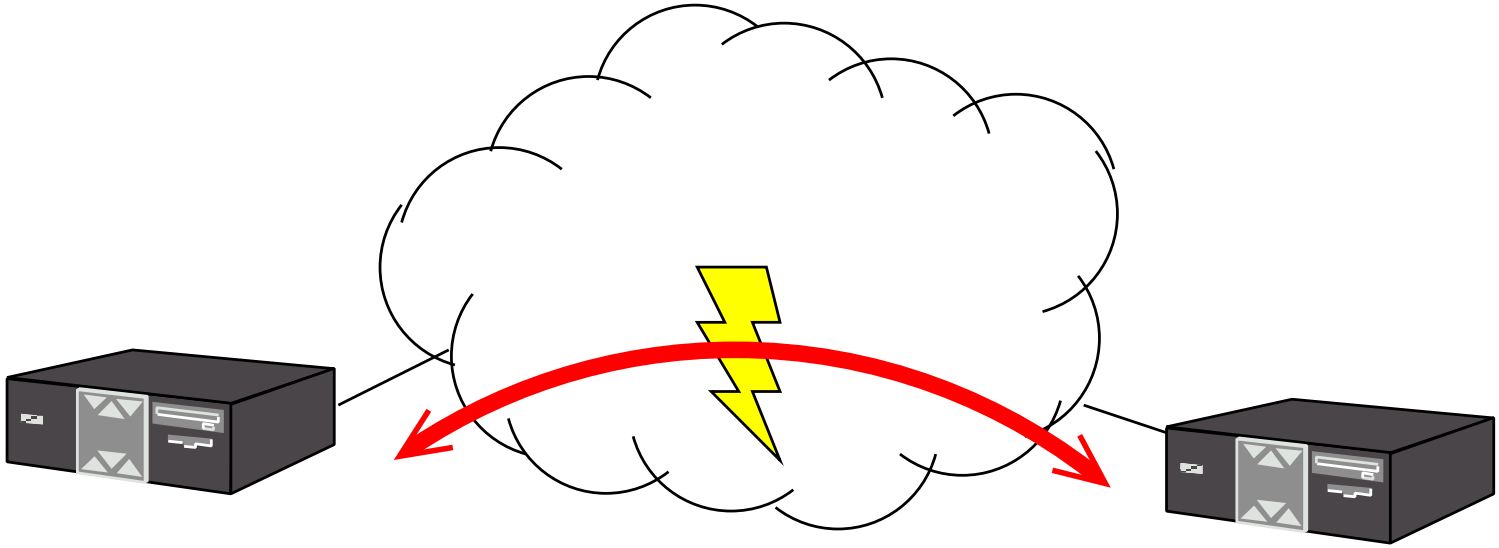
Measuring the Effects of Internet Path Faults on Reactive Routing

Nick Feamster, David Andersen,
Hari Balakrishnan, Frans Kaashoek

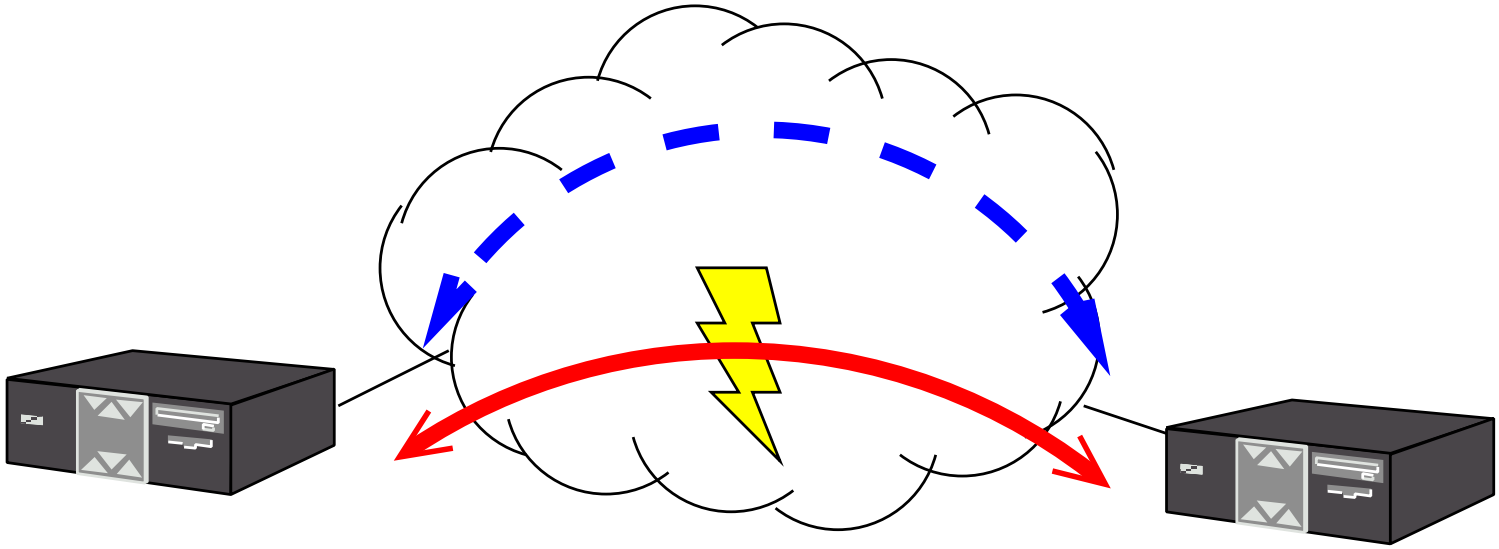
A Familiar Scenario

Internet paths are flaky!



Routing to the Rescue?

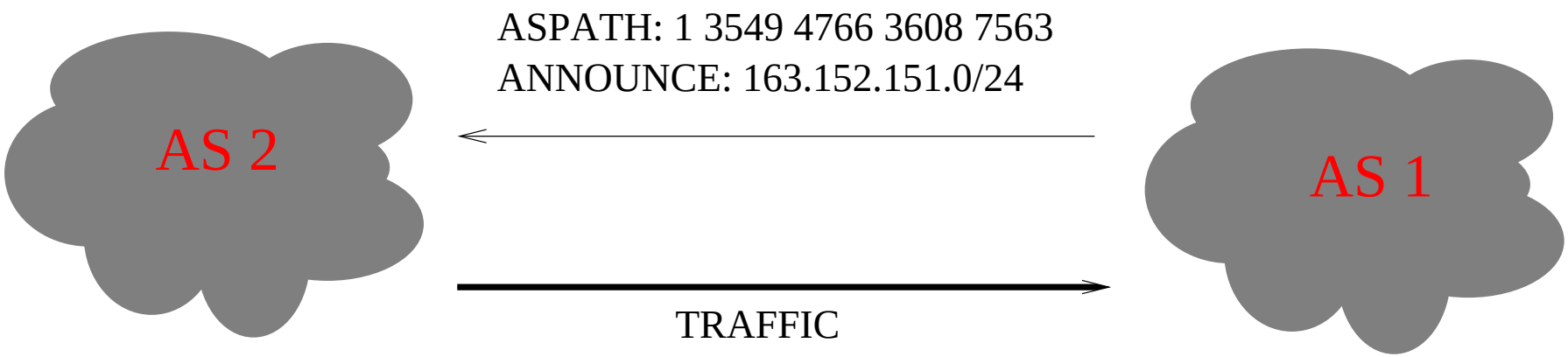
One option: wait for the routing protocol to react.



- "Border Gateway Protocol" (BGP).

BGP Crash Course

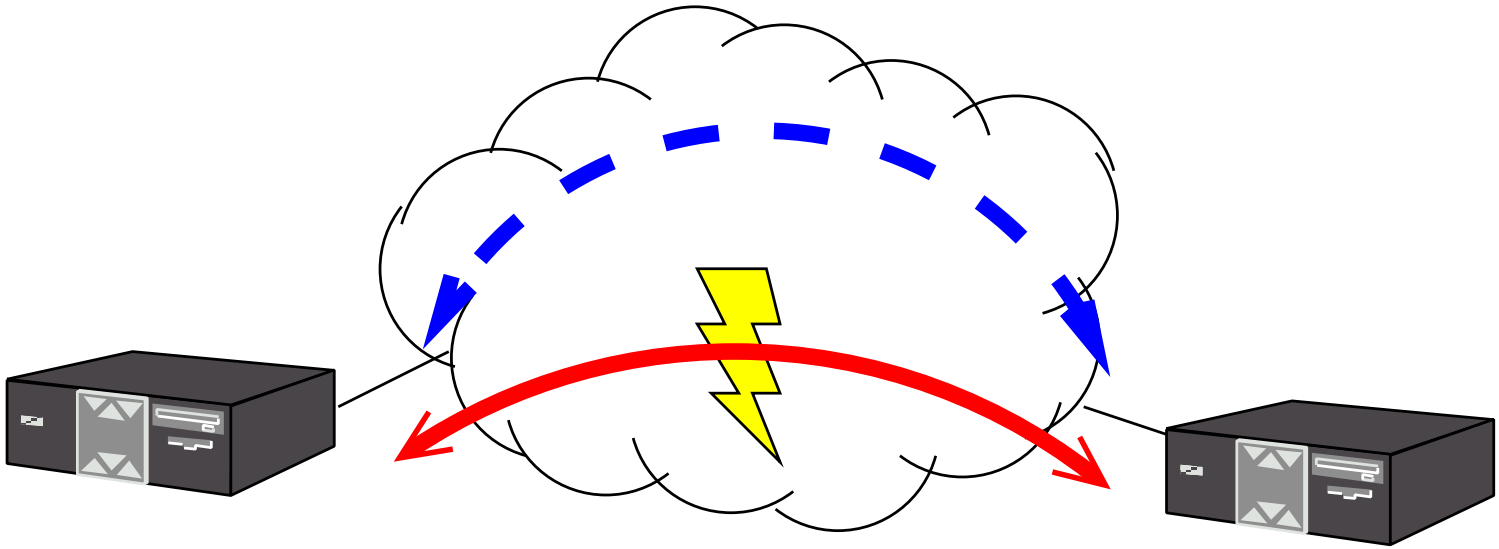
- Routing information is exchanged via updates between multiple autonomous systems (ASes).



- When a path fails, we expect to see changes in reachability reflected in BGP.
- *The routing protocol itself may cause failures.*

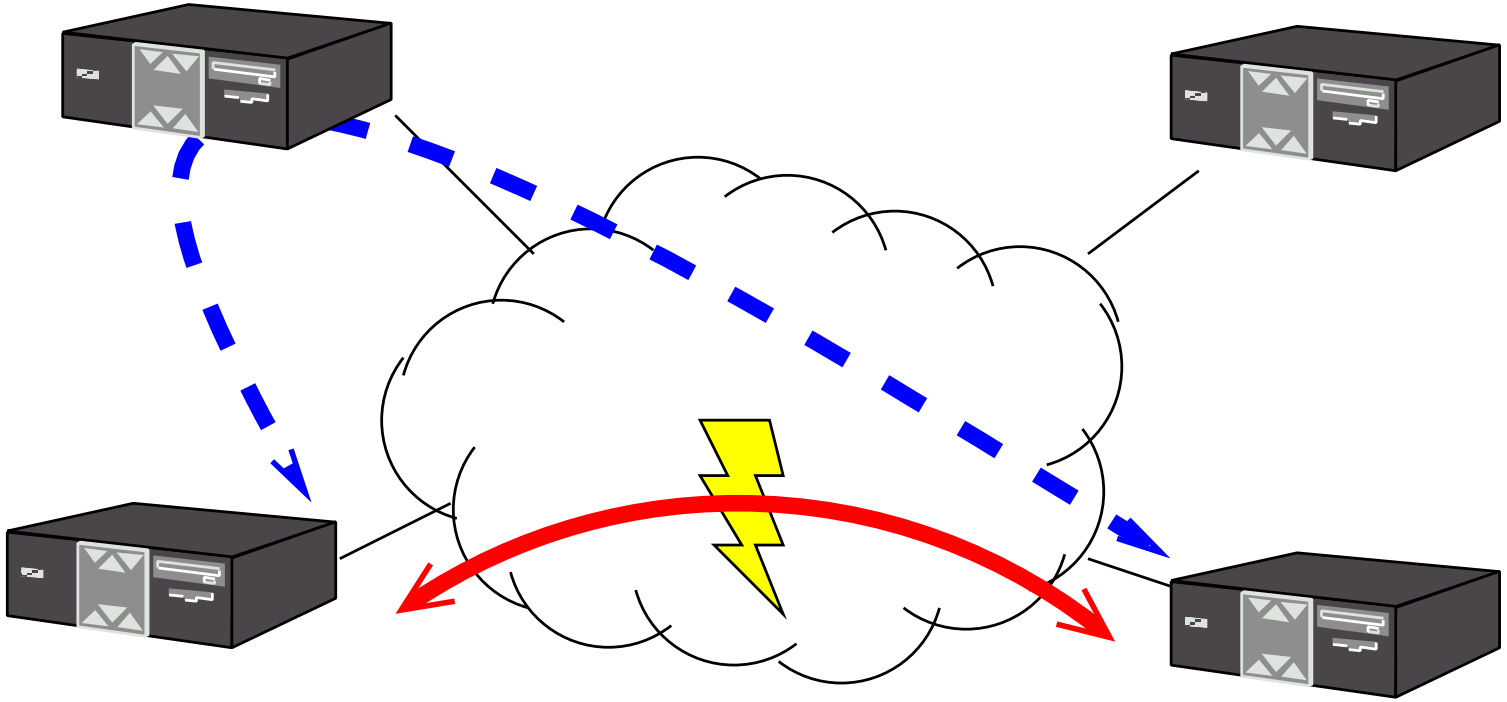
Routing to the Rescue?

BGP is too slow,
and it may not mask all failures.



Reactive Routing to the Rescue

Reactive routing can mask failures. How?
Use end-to-end measurements!



- End-to-end failure detection
- Actively search for better alternative paths

Why/when does Reactive Routing Work?

- **Location:** Where do failures *appear?*
- **Duration:** How long do failures last?
- **Correlation:** Do failures correlate with BGP instability?

A Foggy Telescope Lens

We can’t answer the following questions.
(For reactive routing, this is OK.)

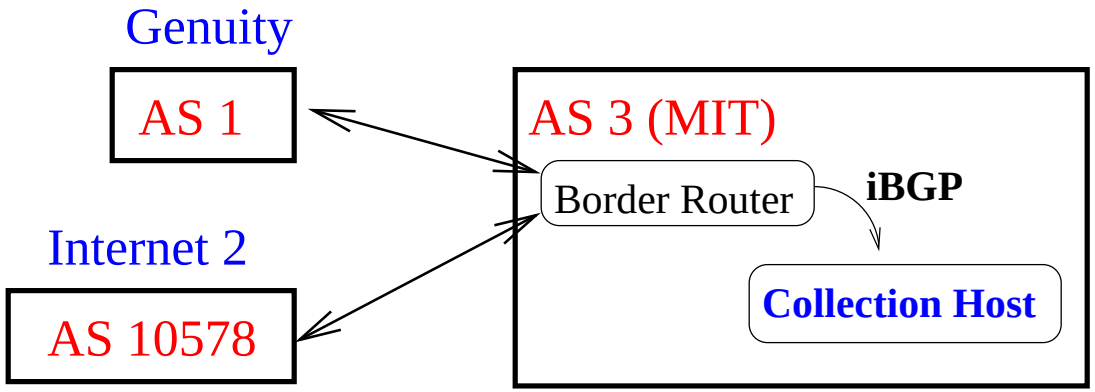
- Where do failures occur? (chasing a moving target)
- What causes these failures? (speculation)
- What about really short failures? (probing granularity)

Highlights

- Data Collection and Caveats
- Inputs and Analysis Heuristics
- Findings
  - *Location:* Some links experience many path failures, but many experience some failures.
  - *Duration:* 70% of failures last less than 5 minutes.
  - *Correlation:* BGP traffic often coincides with failures.
- Conclusions

Data Collection

- Topology: 31 widely distributed nodes
- Active Probes
  - Periodic pairwise probing; logging detects one-way loss.
  - Failure: 3 consecutive lost probes, >2 minutes
- Failure-triggered traceroutes
- BGP Data Feeds at 8 measurement hosts

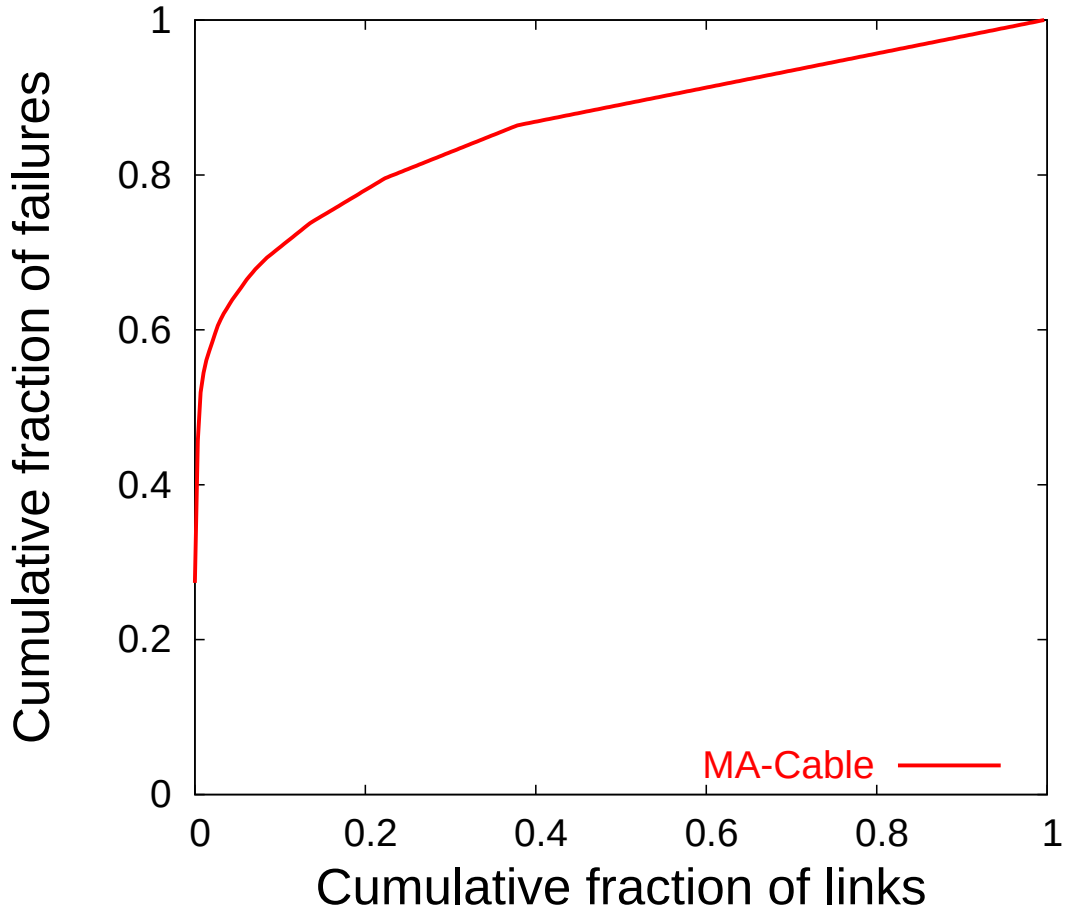


Analysis Heuristics and Inputs

- Analysis Heuristics
  - Resolving AS Boundaries
  - Estimating Network Topology
- Per (one-way) failure, we determine:
  - link (traceroute hop)
  - router
  - autonomous system
  - approximate distance from network edge
  - duration
  - consecutive number of probes lost
  - involves multihomed hosts?
- For some hosts, we have BGP updates, too.

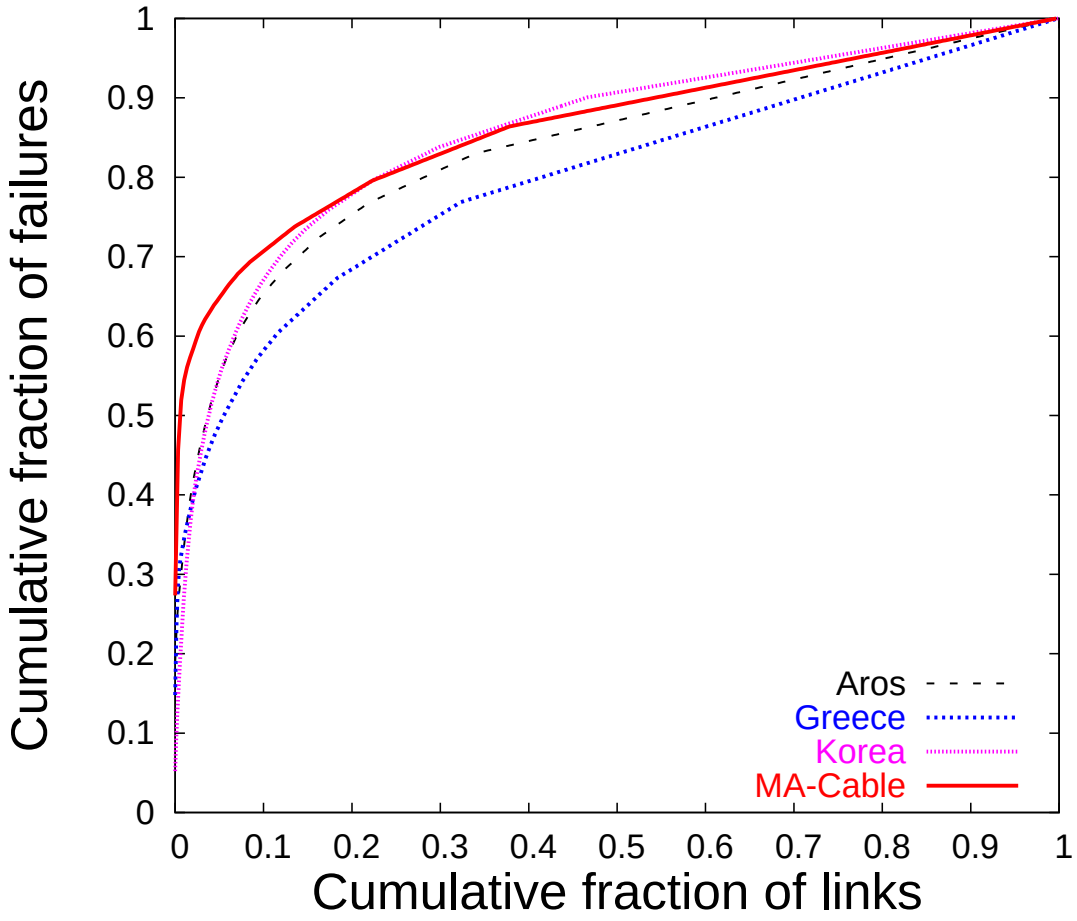
You can reroute, but you can’t hide.

Some links experience many path failures,
but many experience some failures.

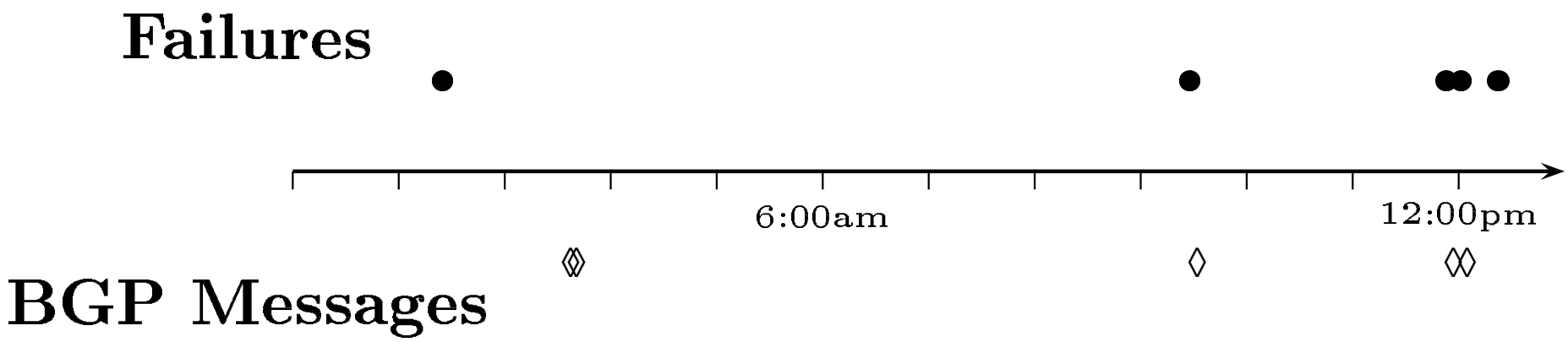


You can reroute, but you can’t hide.

Some links experience many path failures,
but many experience some failures.



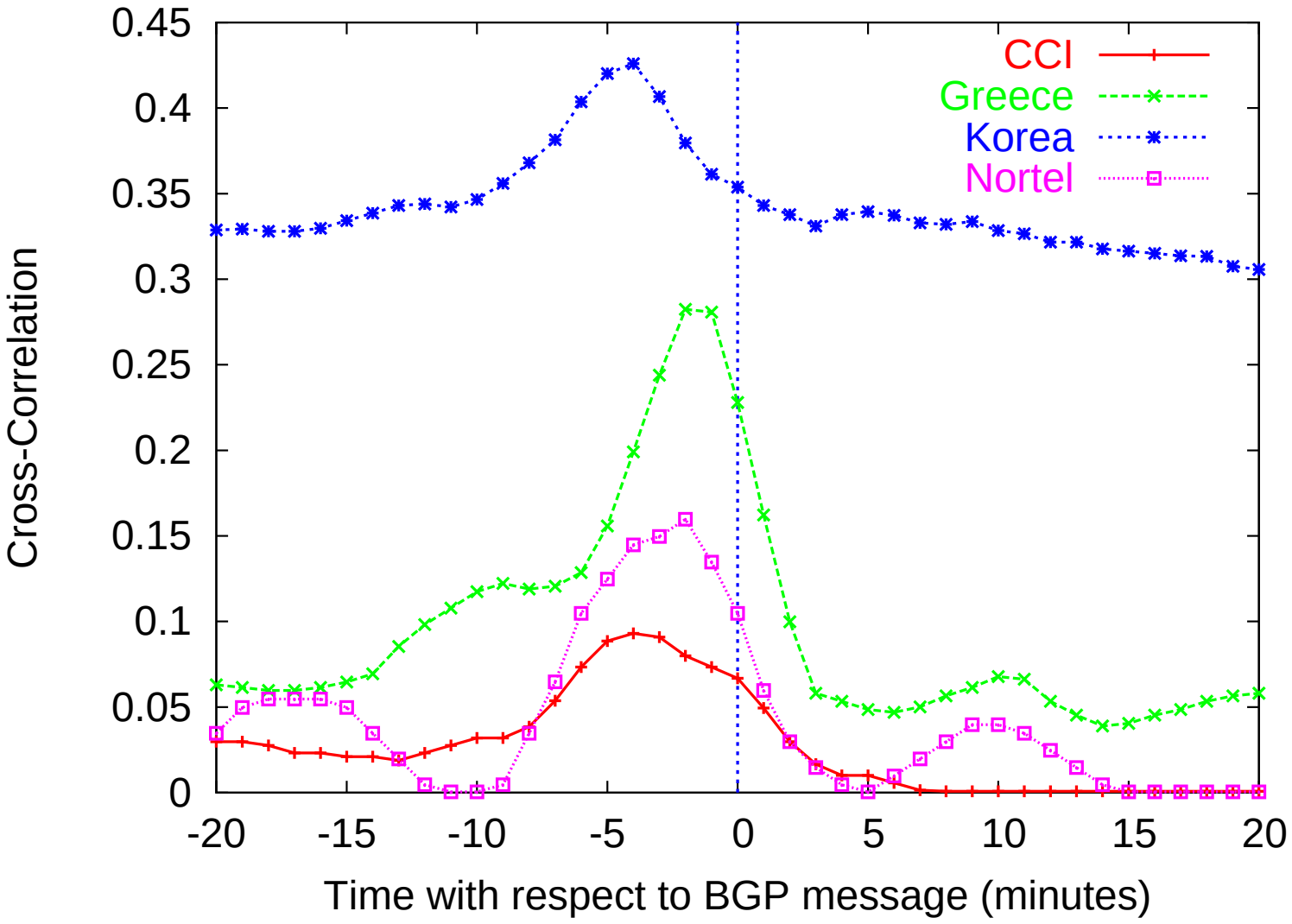
Relating Path Failures and BGP messages



- *Technique 1:* Cross-correlation of time-based signals
- *Technique 2:* Consider a failure and look for BGP (and vice versa)

Do failures correlate with routing instability?

Failures typically occur several minutes before BGP activity



Which failures correlate with instability?

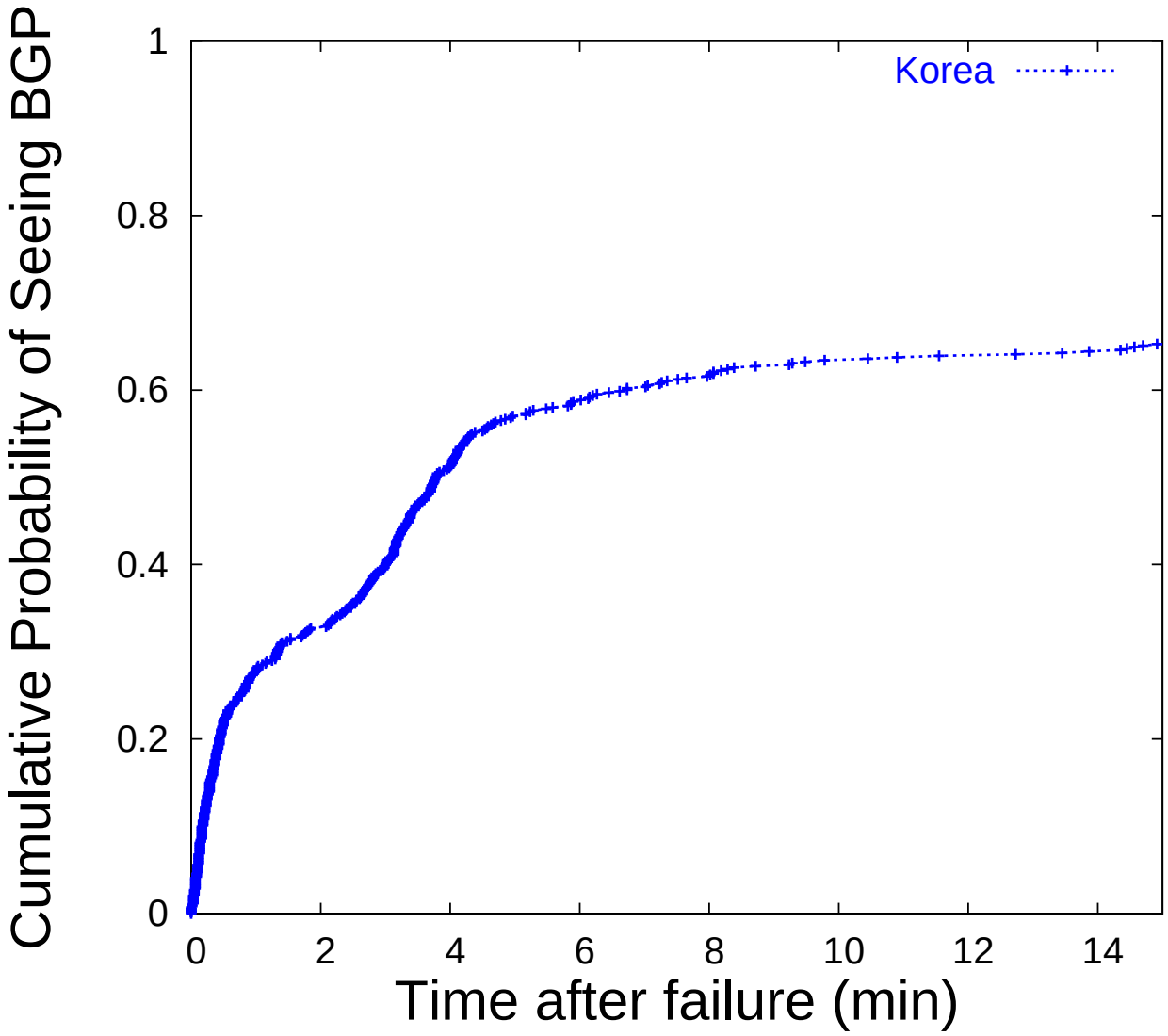
Failures that appear near end hosts are less likely to coincide with BGP instability.

- 60% of failures that appeared at least three hops from an end host coincided with at least one BGP message.

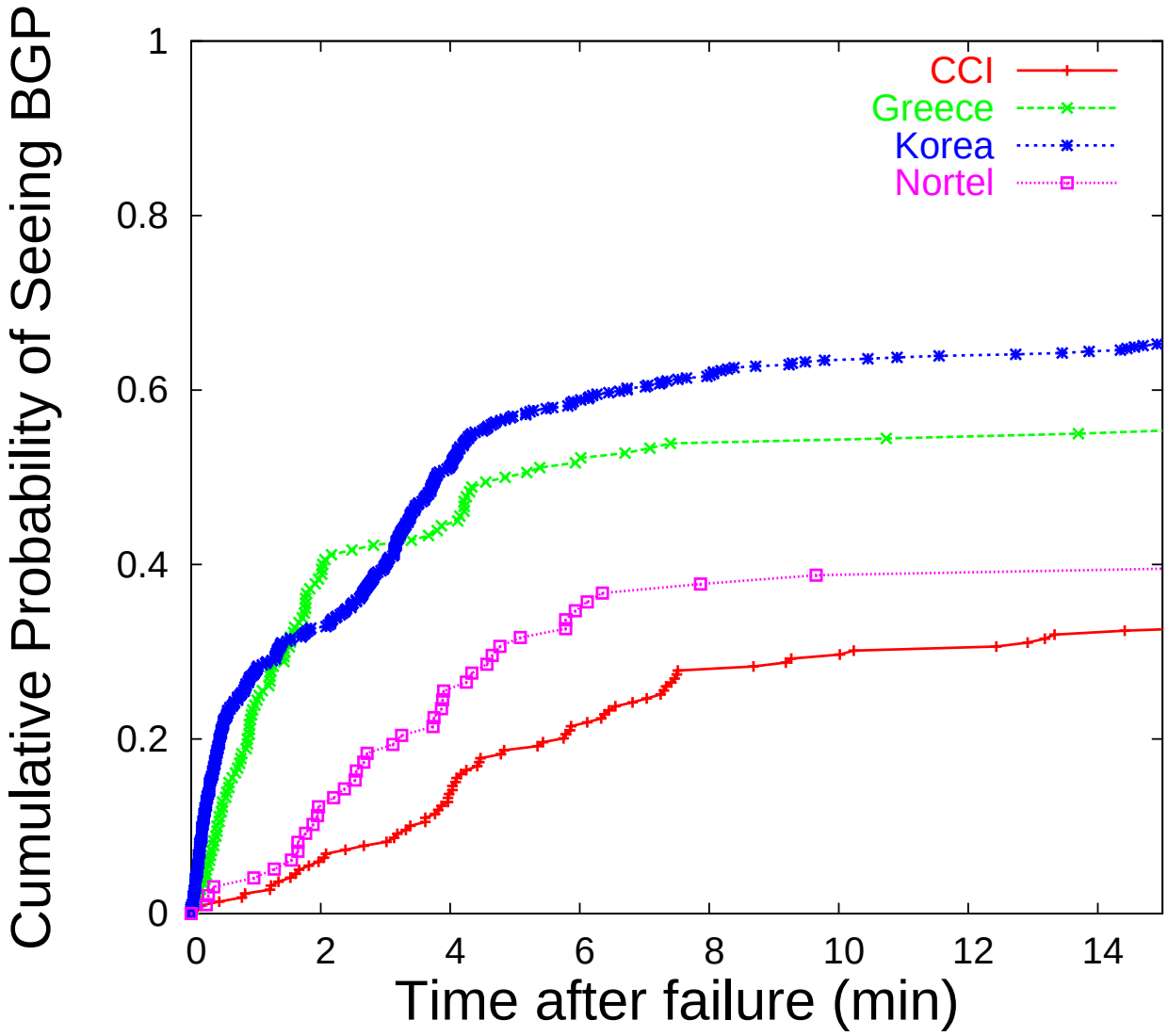
- 22% of failures within one hop of an end host coincided with at least one BGP message.

Just because an ISP is reachable doesn't mean its customers are reachable!

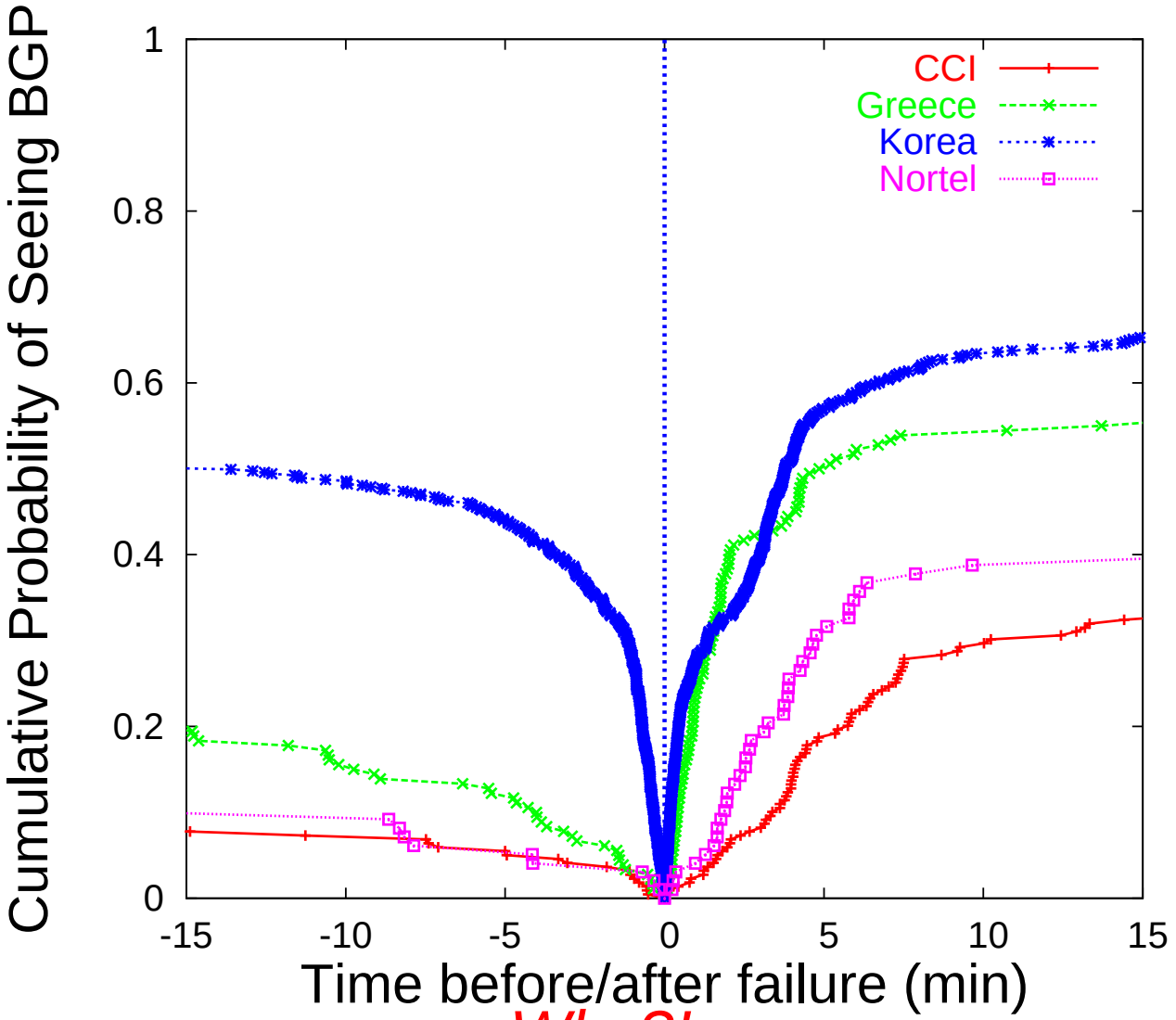
To put it another way...



To put it another way...



Surprise: BGP messages precede failures!



Why?!

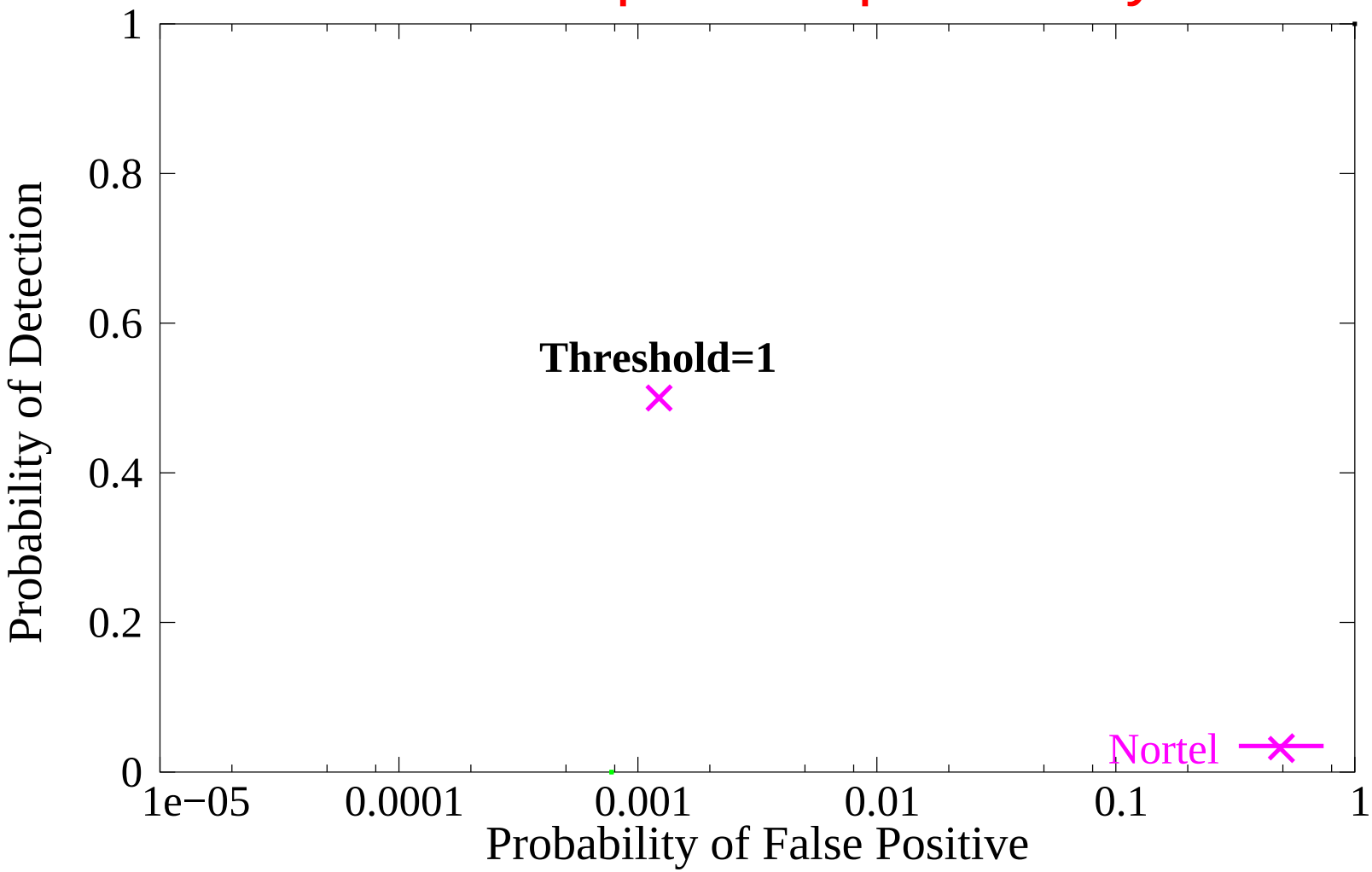
Route flap damping, maintenance, misconfiguration, etc.

Can BGP help predict failures?

- Observe BGP messages over a time window.
- If the number of BGP messages exceeds some threshold, predict failure.
- Varying this threshold affects the probabilities of detection and false positive.

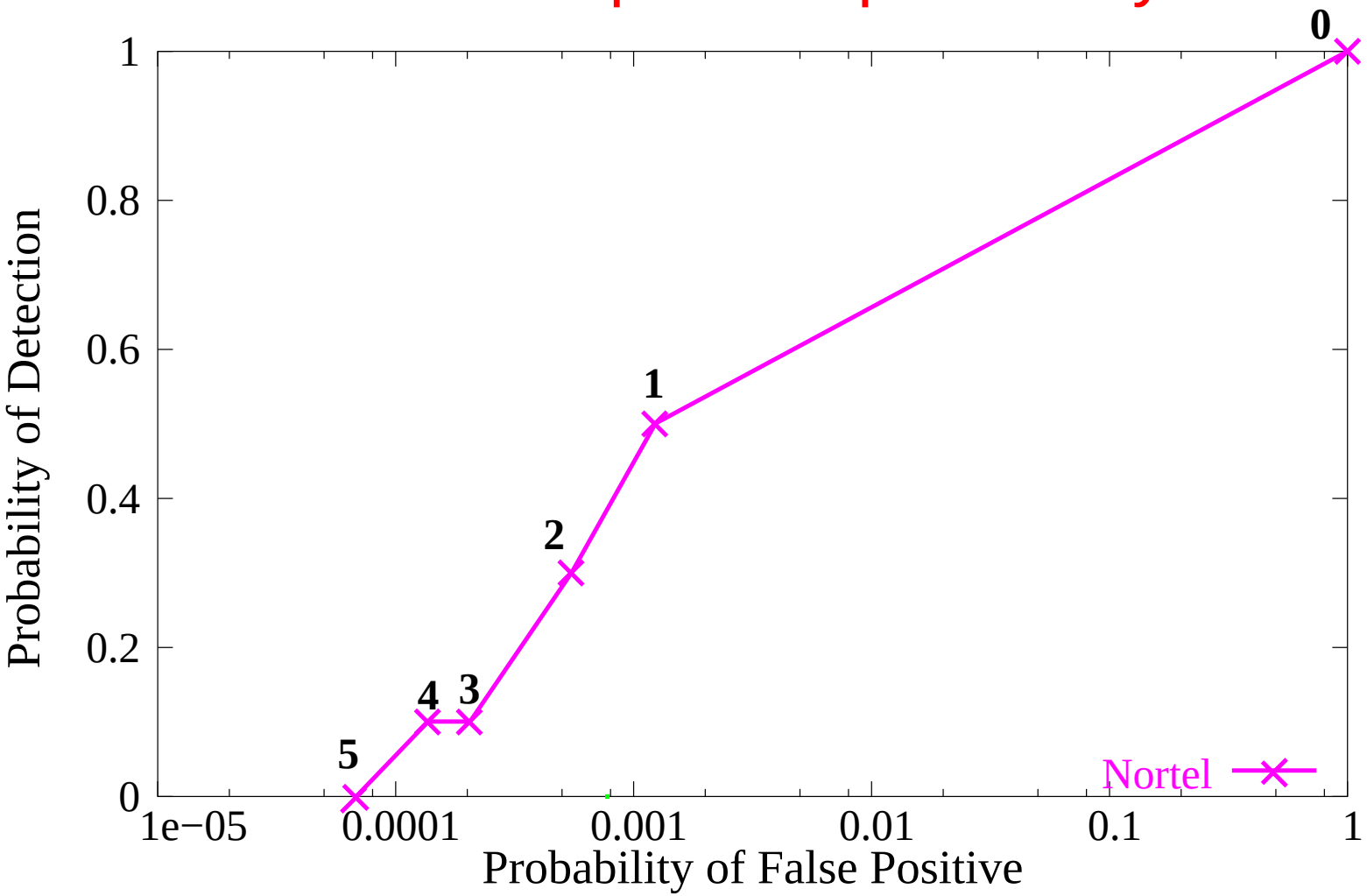
Can BGP help predict failures?

We want a high probability of detection with low false positive probability.



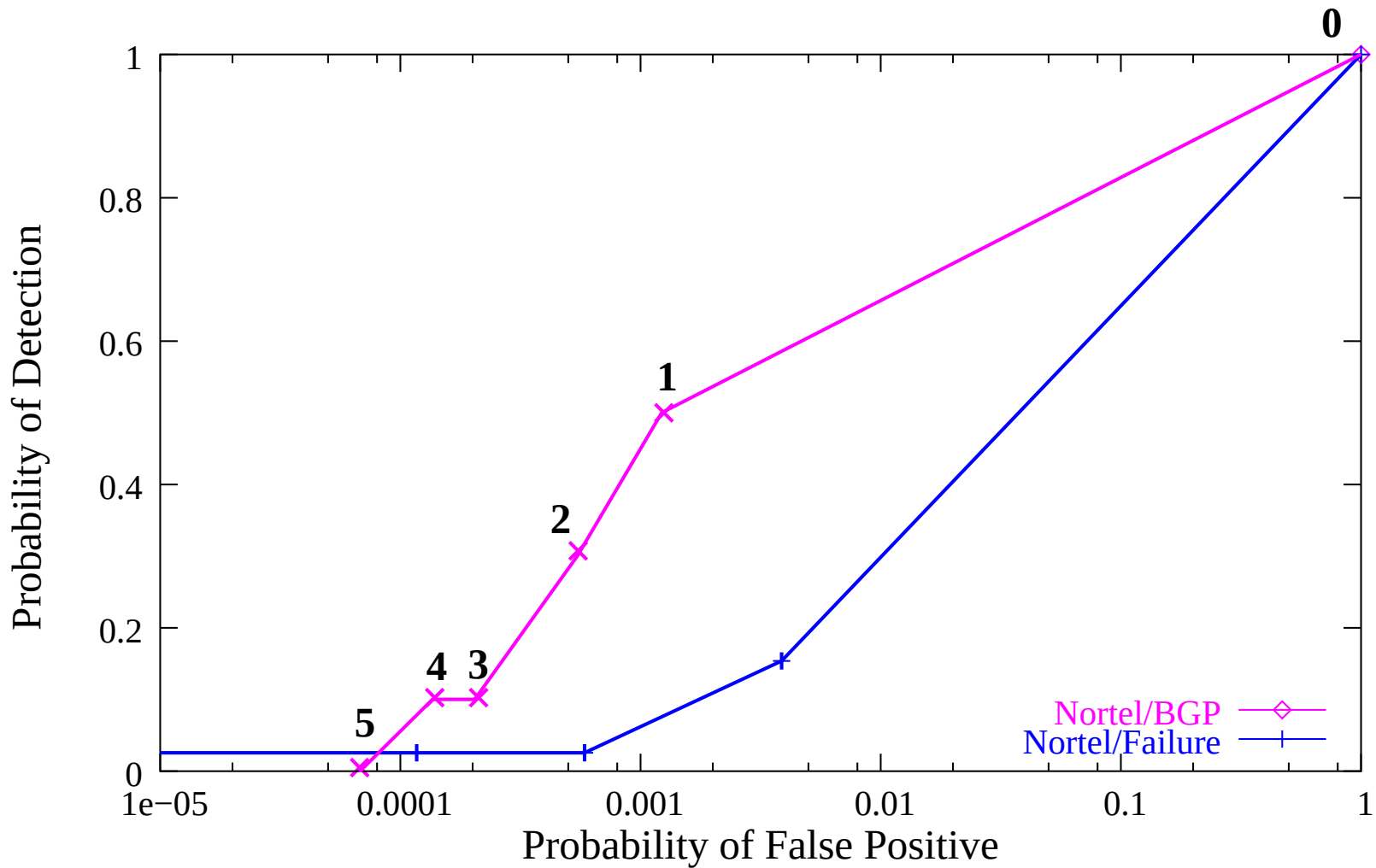
Can BGP help predict failures?

We want a high probability of detection with low false positive probability.



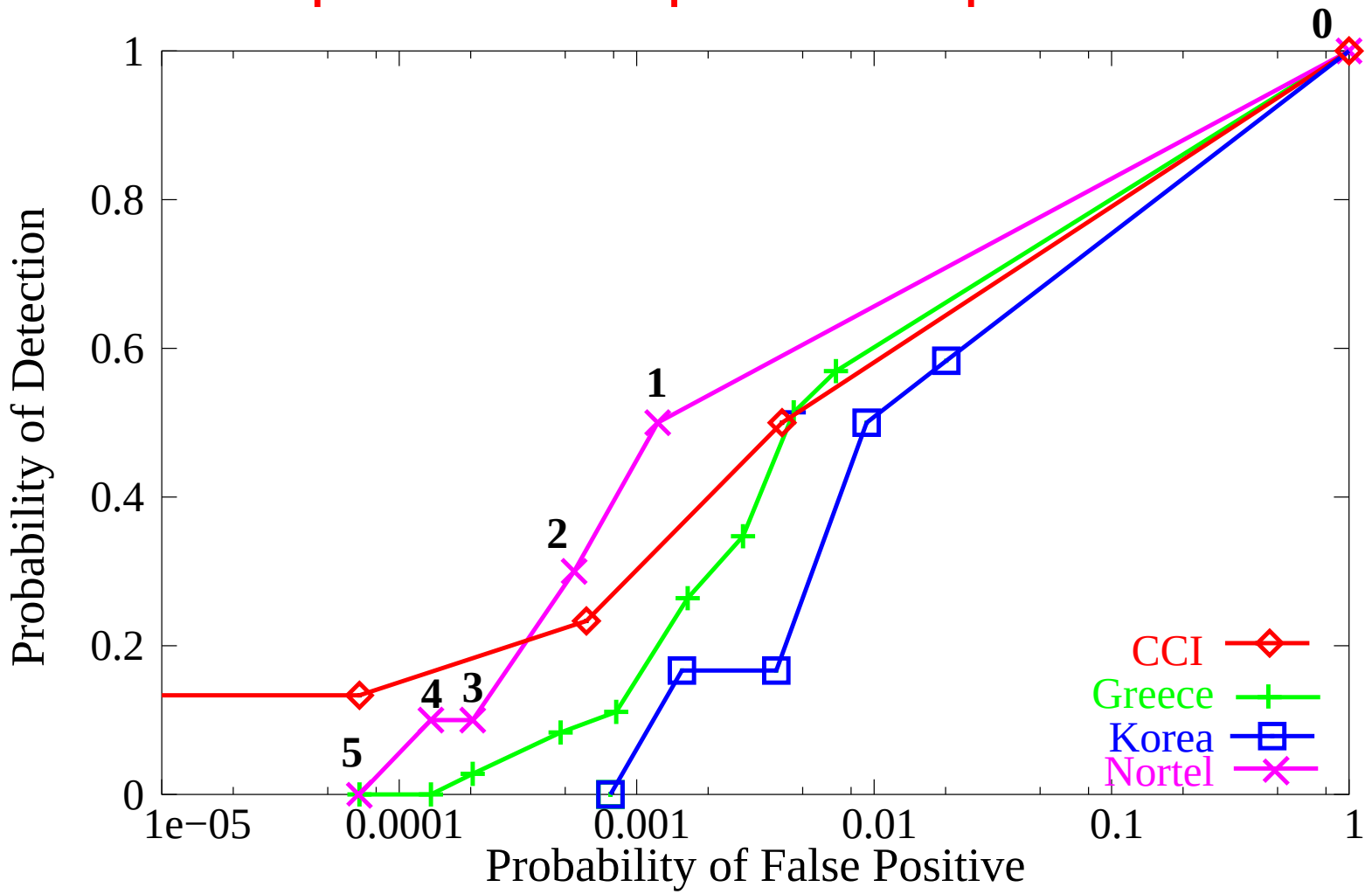
Can BGP help predict failures?

Using BGP to predict failures has a higher probability of detection for a given false positive rate than using failures.



Can BGP help predict failures?

Your mileage may vary:
Effectiveness of predictor depends on path characteristics.



Conclusion

- Analysis Heuristics
  - Resolving AS Boundaries
  - Estimating Network Topology
- Findings
  - *Location:* Some links experience many path failures, but many experience some failures.
  - *Duration:* 70% of failures last less than 5 minutes
  - *Correlation:* BGP is a better predictor of failures than failures themselves, but it doesn't catch everything.
- How can we improve reactive routing?
  - topologies, decision algorithms, etc.

Measurement Caveats

- Representativeness of data set

  Answer: Paths cross 9 of 11 of "large" ASes.

- Could reflect failure on the reverse path

  Solution: Correlate with failures from active probes.

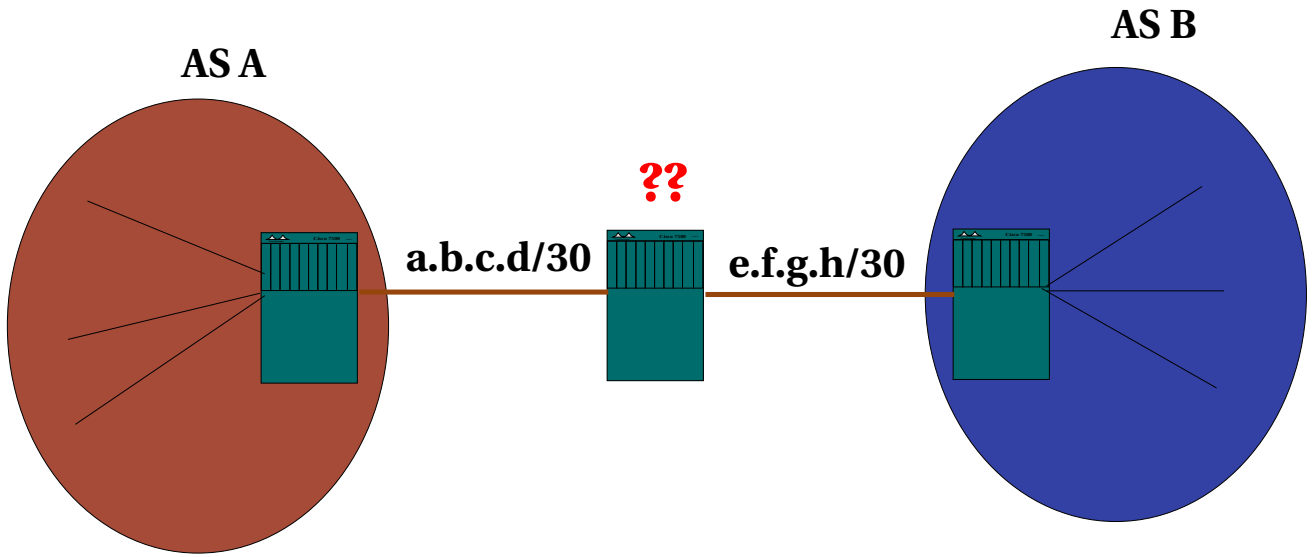
- Tells us about interfaces, not routers
  - Maybe even the outgoing interface of the return packet!

  Solution: Alias resolution techniques.

- Little information about AS boundaries

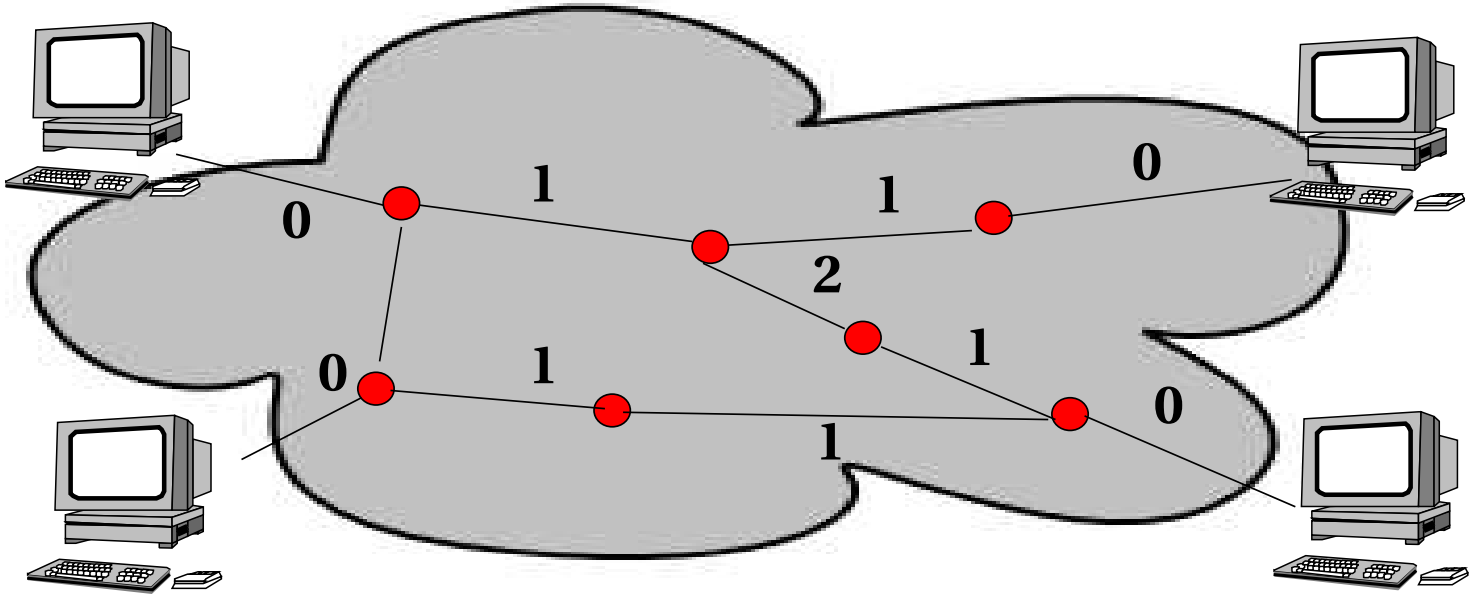
  Solution: Use knowledge about the topology...

Resolving AS Boundaries



- No (or few) extra traceroutes, please.
- If a router has interfaces from multiple ASes, see what other routers it's connected to.
- For each failure: router, AS, location in AS
- Most failures not on AS boundaries.

Estimating Network Topology



- How "deep" into the network is a particular link?
- New links introduce two interfaces.
  - At least one of these must connect to a host for which we've assigned a distance.
  - Assign minimum distance to end host.
- Assumes measurement hosts are at network edges

Reactive Routing

- Of the failures that reactive routing could mask, only half coincided with BGP messages.
- Multihoming: a near-necessary (but not sufficient) condition for always masking failures
  - 25 paths: RON always masked failures
    - All of these paths had at least one multihomed end
    - 18 were multihomed at both ends
  - 136 paths: RON never masked failures
    - 106 paths had at least one multihomed end
    - 35 were multihomed at both ends

Fleeting failures

Reactive routing must be fast;
multihoming alone isn't enough.

- 70% of failures last less than 5 minutes.
  - Reactive routing schemes should be fast.
- Multihomed hosts have the same failure duration.
  - Multihoming per se does not improve failover time (BGP is slow).