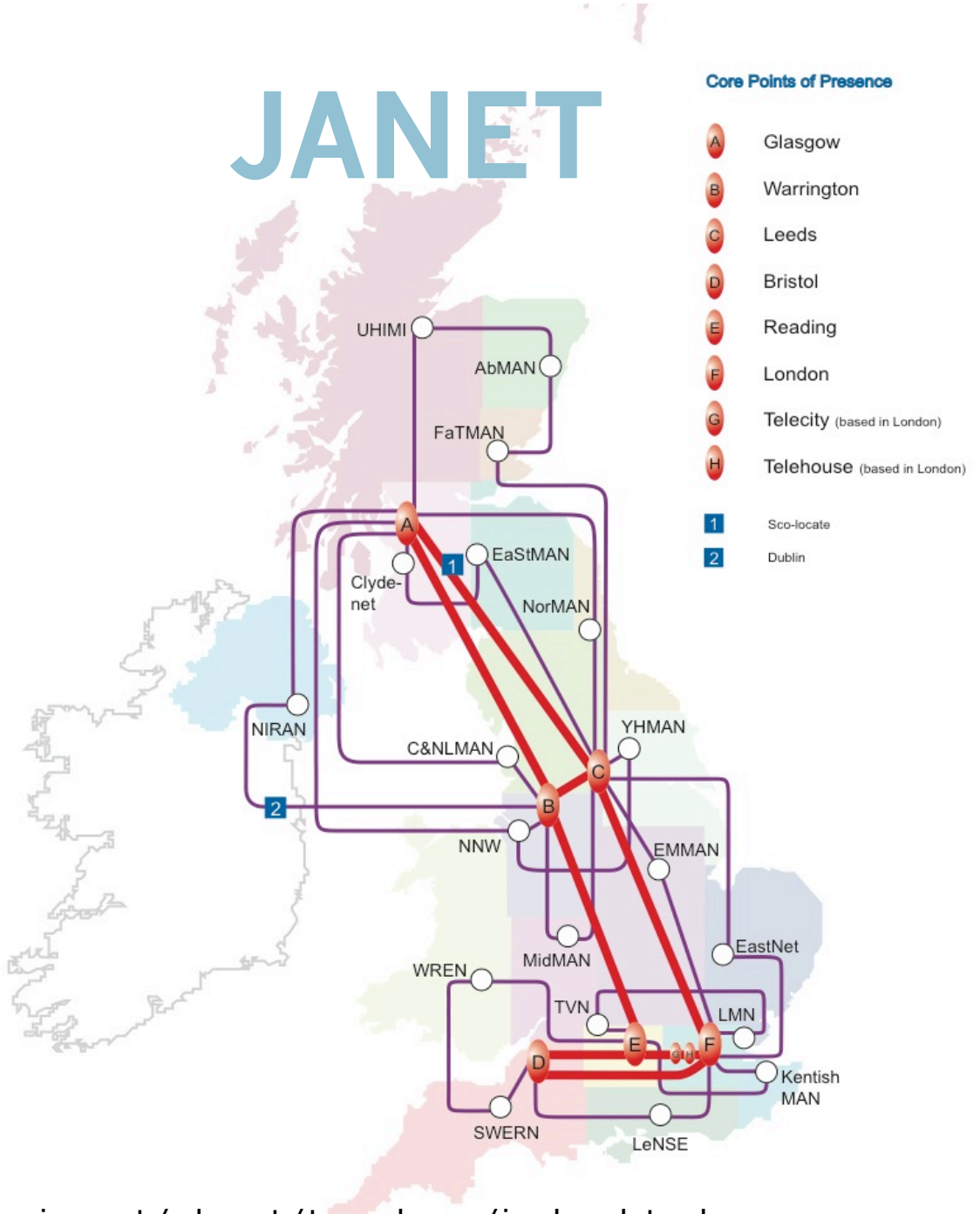
10 Lessons from 10 Years of Measuring and Modeling the Internet's Autonomous Systems

Olaf Maennel

this work would not have been possible without the help of friends and co-workers:

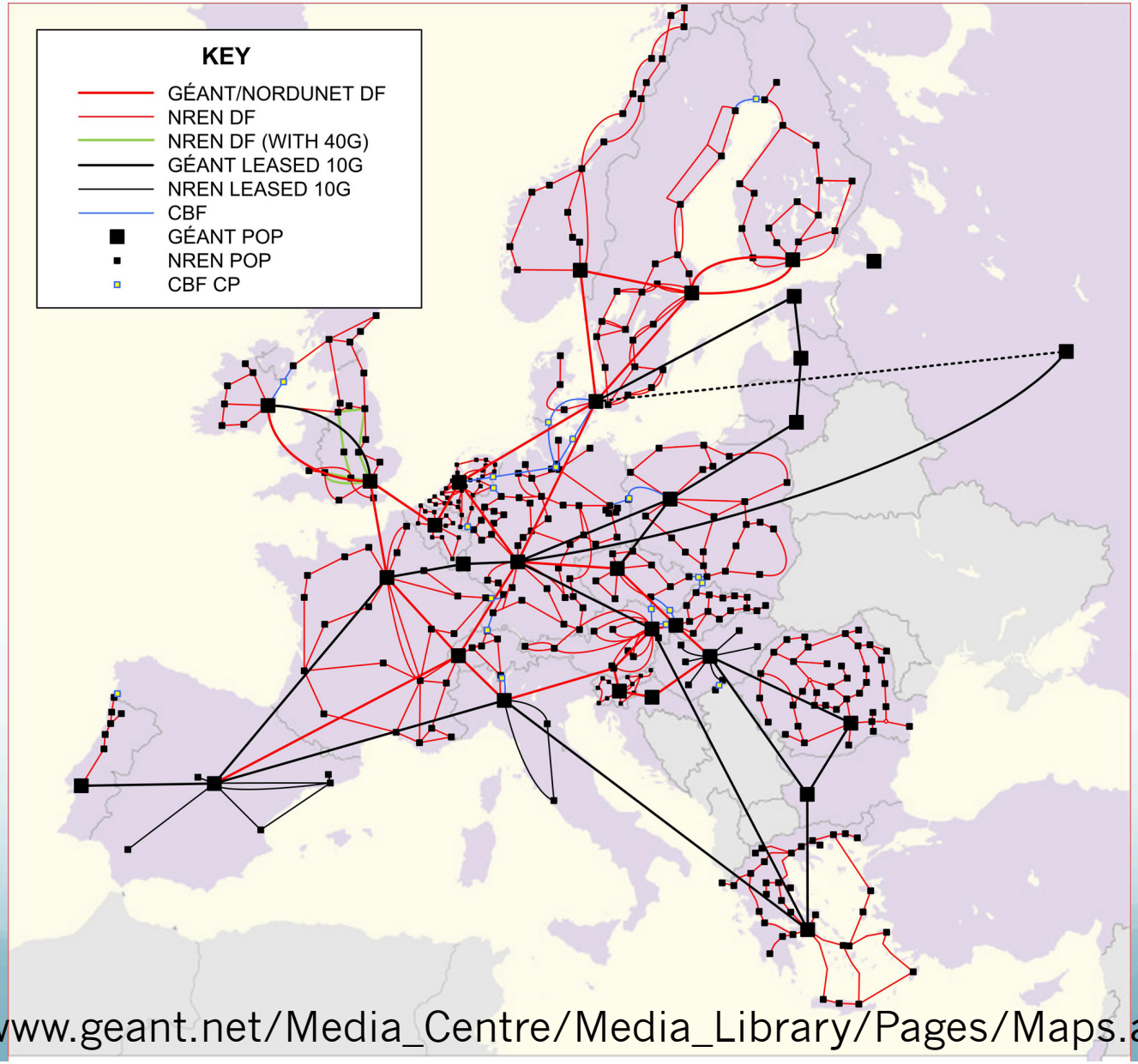
Randy Bush, Tim Griffin, Yoshinobu Matsuzaki, Matthew Roughan, Steve Uhlig, and Walter Willinger.

JANET



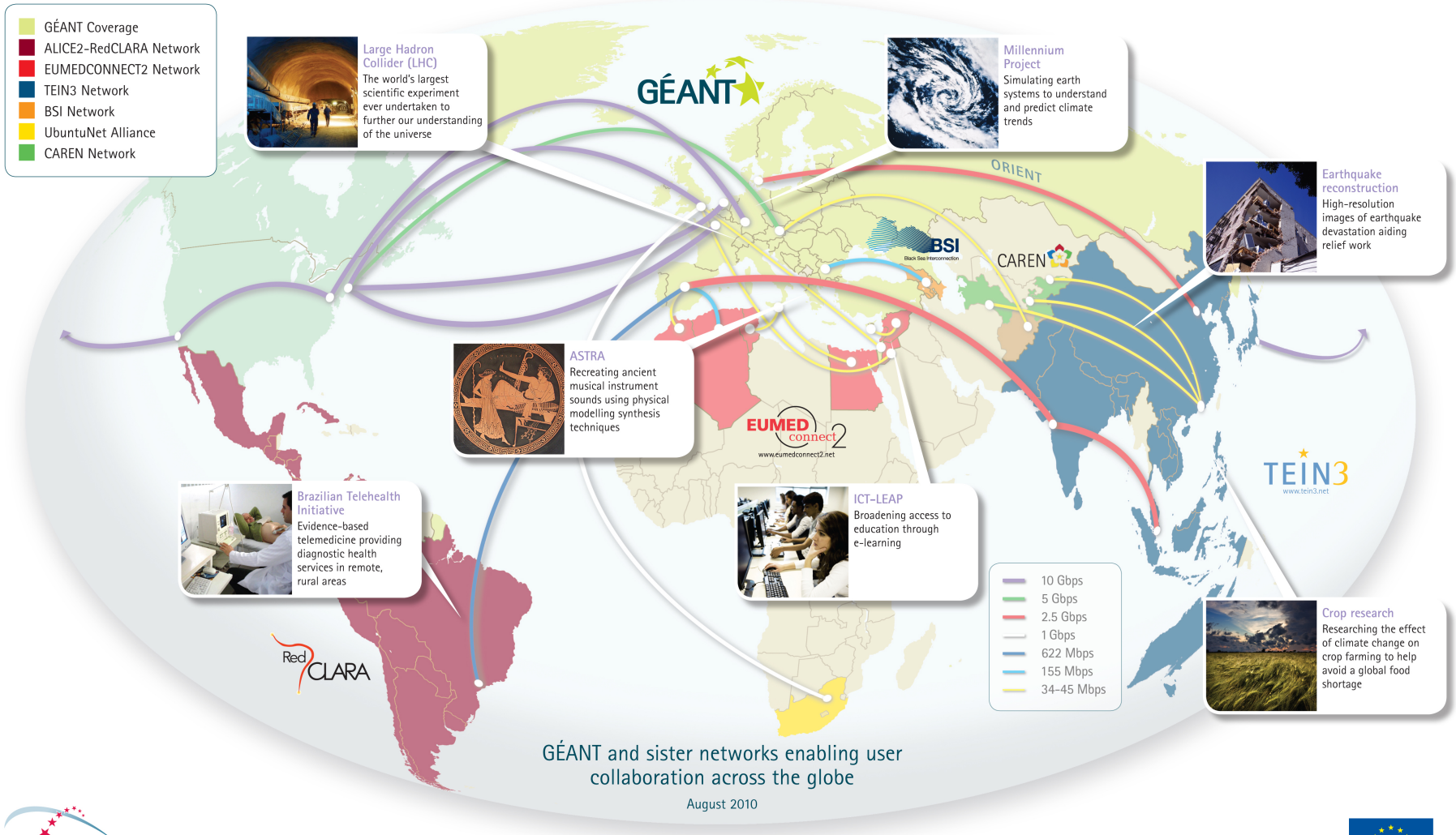
http://webarchive.ja.net/about/topology/index.html

GEANT



http://www.geant.net/Media_Centre/Media_Library/Pages/Maps.aspx

GÉANT At the Heart of Global Research Networking



GÉANT and sister networks enabling user collaboration across the globe

August 2010

DANTE
www.dante.net

connect • communicate • collaborate

This document has been produced with the financial assistance of the European Union. The contents of this document are the sole responsibility of DANTE and can under no circumstances be regarded as reflecting the position of the European Union.

European Commission
Information Society and Media

http://www.geant.net/Media_Centre/Media_Library/Pages/Maps.aspx

Network of Networks

- There are about 38,224 autonomous systems in the Internet today. (http://www.cidr-report.org/as2.0/)

- What we observe is often based on what protocols (e.g., BGP) show us.
- We need to understand the limitations of the observation to be able to understand the limitations of the inference.

The Goal of the Internet

- Different companies interconnect to build “the Internet”.
- They may be **competitors** in the business space…
- But they are **working together** to achieve one goal: *global reachability*.

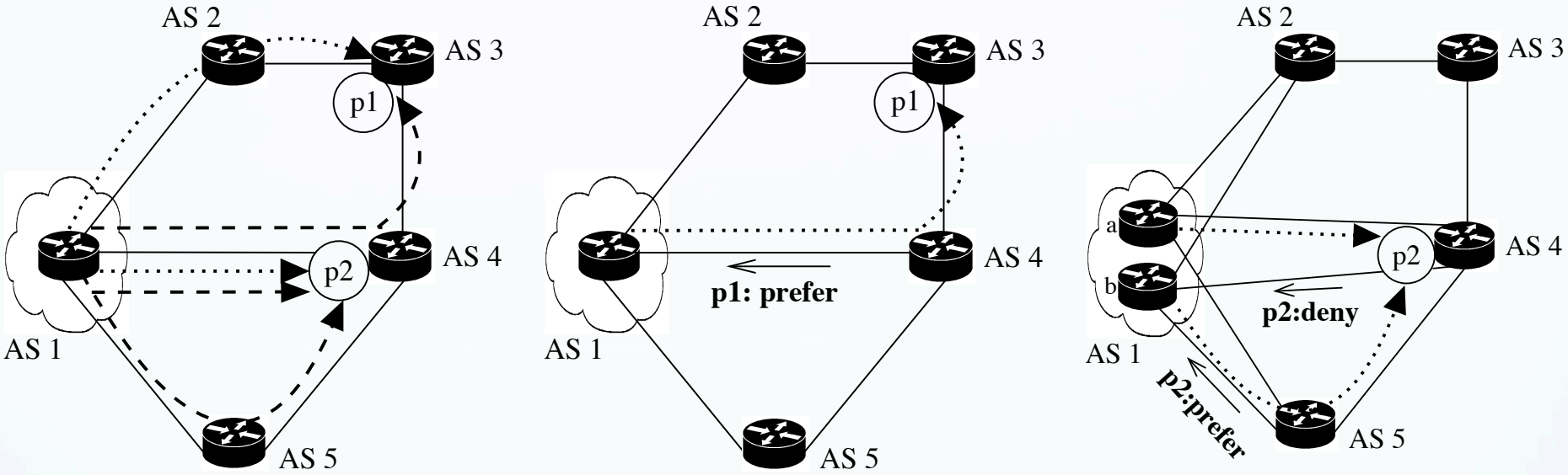
“Border Gateway Protocol” BGP

- BGP is the “**glue**” that keeps the Internet together. ;-)
- It’s a policy protocol, that allows companies to express their business needs, while exchanging at the same time reachability information.

BGP in a Nutshell

- BGP is a path-vector protocol
- Uses “attributes”, such as:
  - AS-path
  - Local-Preference
  - Communities
- Information hiding:
  - Scalability concerns
  - Policies for commercial policy reasons

Common approaches to modeling the Internet



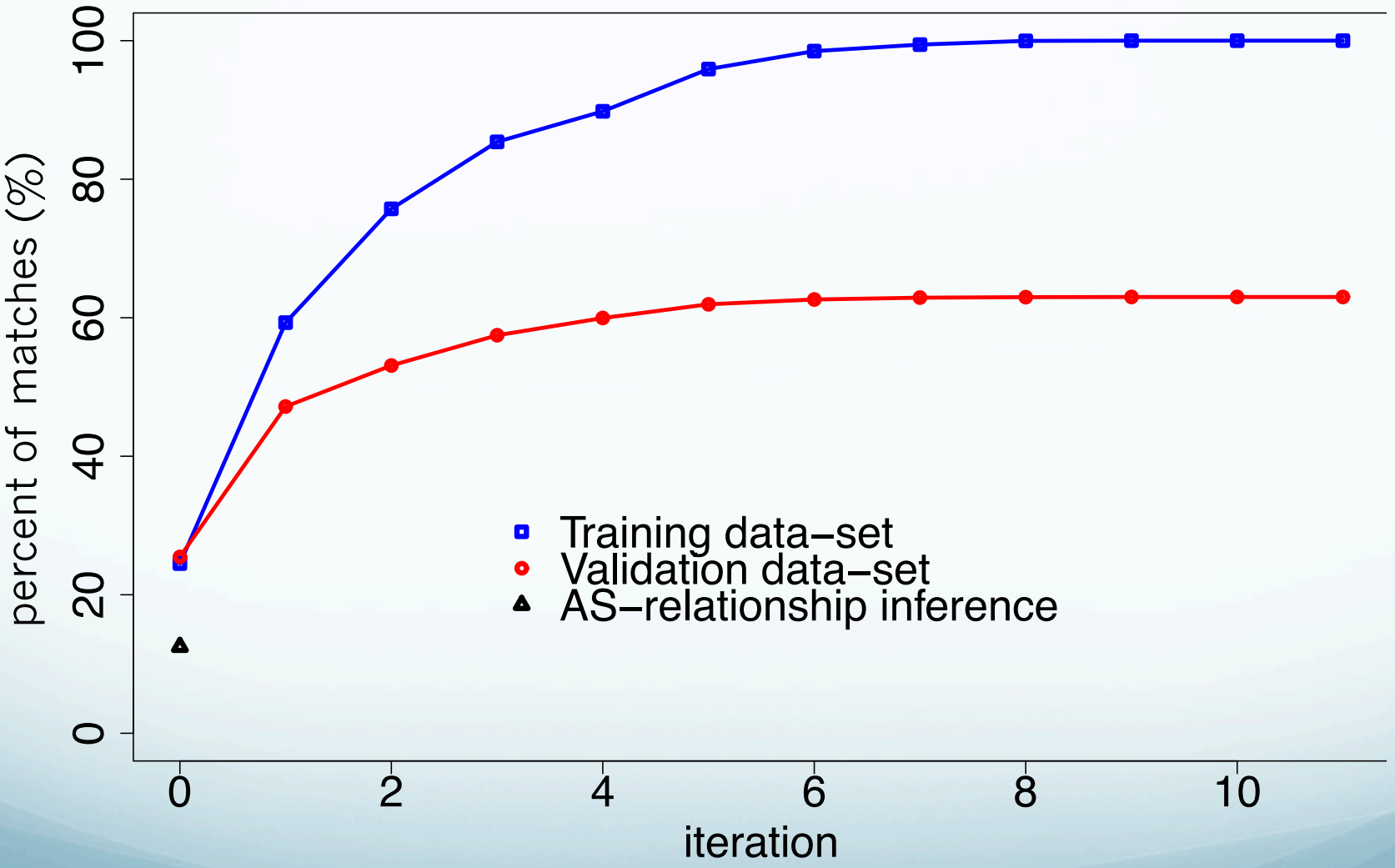
Methodology:

- **Take observations and tweak model until it fits observations…**

Data-sources may include:

- BGP
- traceroute
- IRR
- AS-relationship inference techniques
- topology-zoo.org
- …and if you are lucky: a few real configurations

(Typical?) Results



Building an AS-topology model that captures route diversity – W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, S. Uhlig. In ACM SIGCOMM, 2006.

I. Examples of limitations.

Our view is systematically biased!

Any model of the Internet needs to take the limitations of BGP into account.

Implications on Internet Measurement Research?

A Reason to Reject Future Papers Based on Data from RIS/RV?

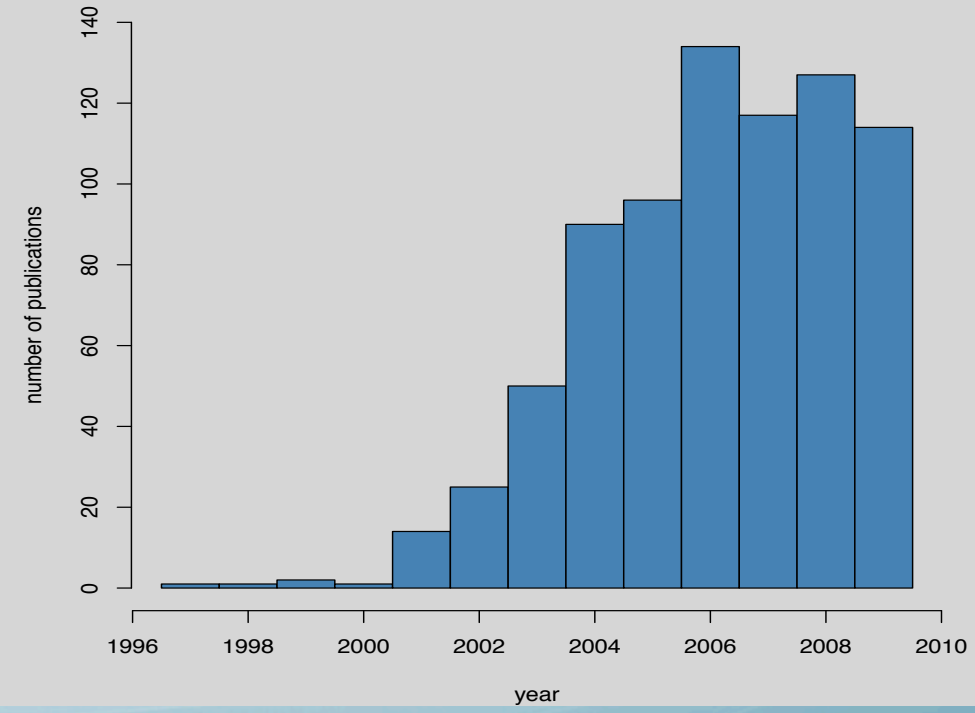
Internet Measurements – The State of the Art?

- What are the limitations of our measurements?
- Why are models of the Internet wrong?
- Why can't we "see" the bias?
- Are Internet Measurements still not understood?

- RIPE RIS/RouteViews were designed for operators
- Researchers discovered them – some without consideration of limitations

<u>Google Scholar search:</u>

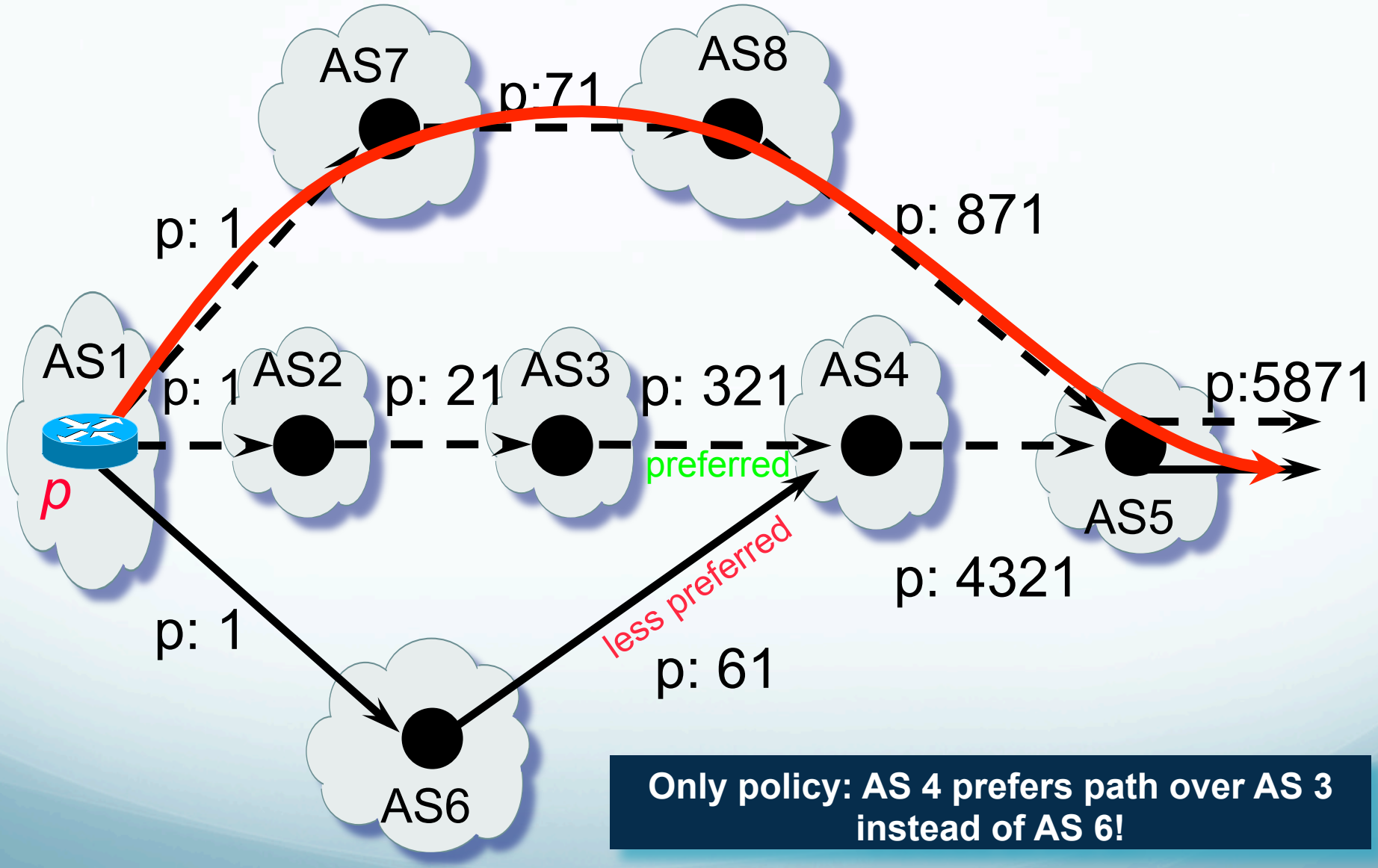
- papers mentioning term *'routeviews'*



Some Examples

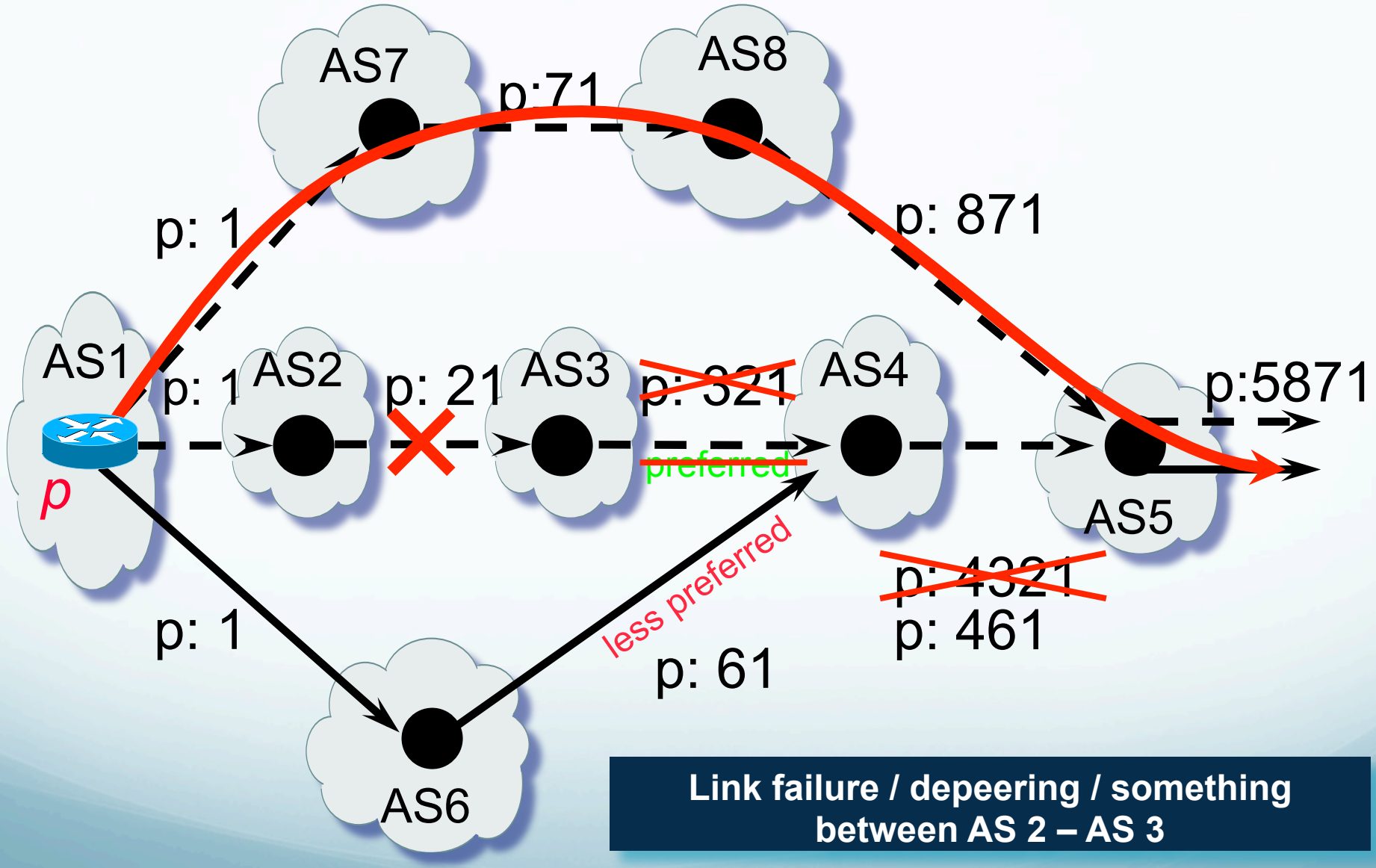
- Policy Interactions –
  BGP an “information-hiding” protocol
- Impact of default at the edge
- Routing vs. Forwarding and the story of modern complex policies…
- Tim Griffin’s “BGP Wedgies”

Policy Interactions – the "fun" of BGP research... ;-)

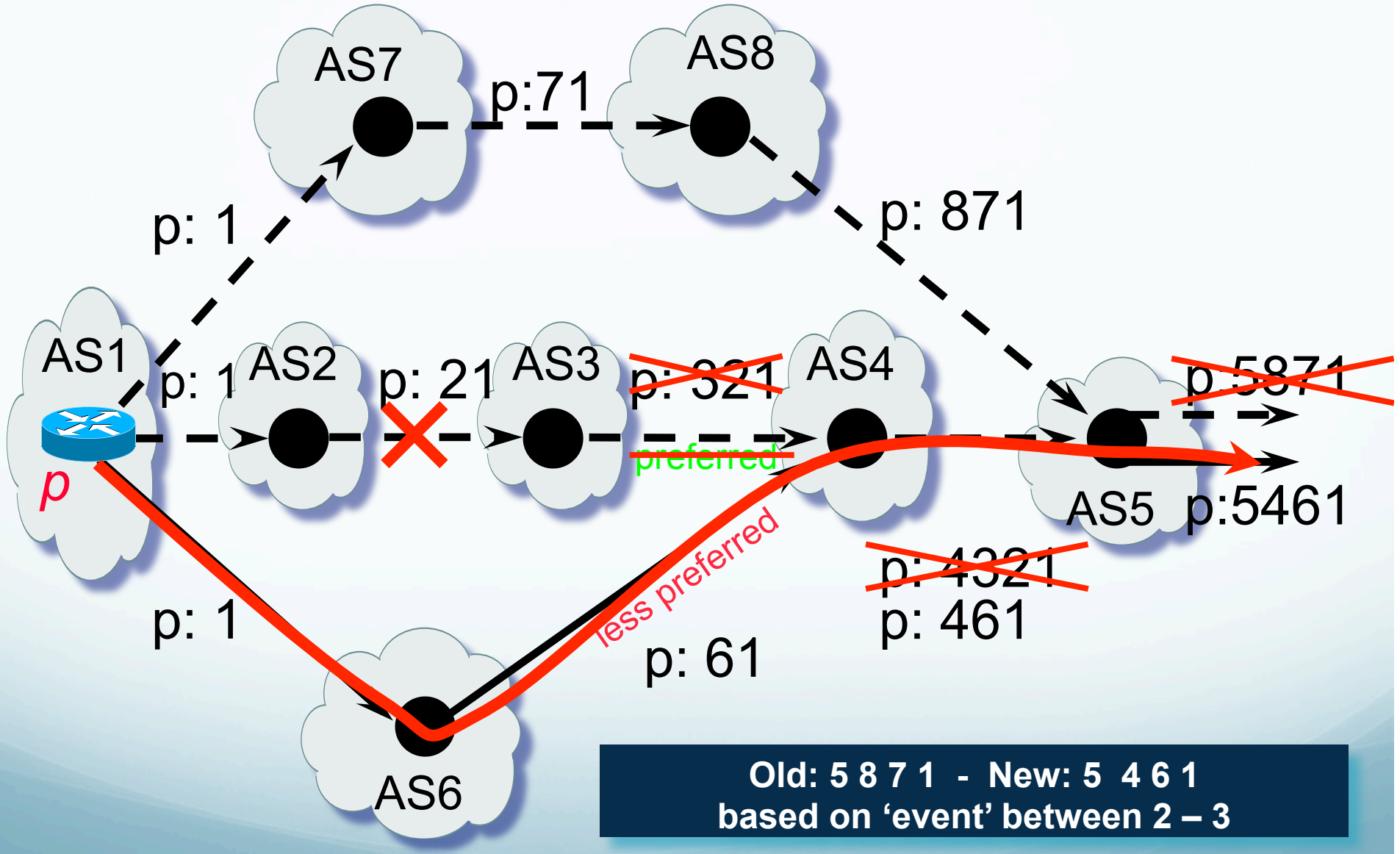


Only policy: AS 4 prefers path over AS 3 instead of AS 6!

Policy Interactions – the "fun" of BGP research... ;-)



Policy Interactions – the "fun" of BGP research... ;-)



Internet Measurements: Limitations in our Understanding

A Reason to Reject Future Papers Based on Data from RIS/RV?

Measurement Failures

- What is the real routing graph of the Internet?
- What is the AS topology of BGP routing?
- How biased is our methodology?
- How do we debug our network?
  - Are ping and traceroute the best we can do?

How to design controlled Internet topology experiments?

Experiment Setup

- Statement: “prefixes ≥ /25 are typically filtered in the Internet”
- Announced a /25 to NTT
- NTT passed it to customers

‘Traditional’ BGP Observations

- Public data: RIPE RIS and RV
- RV/RIS showed **15** ASes

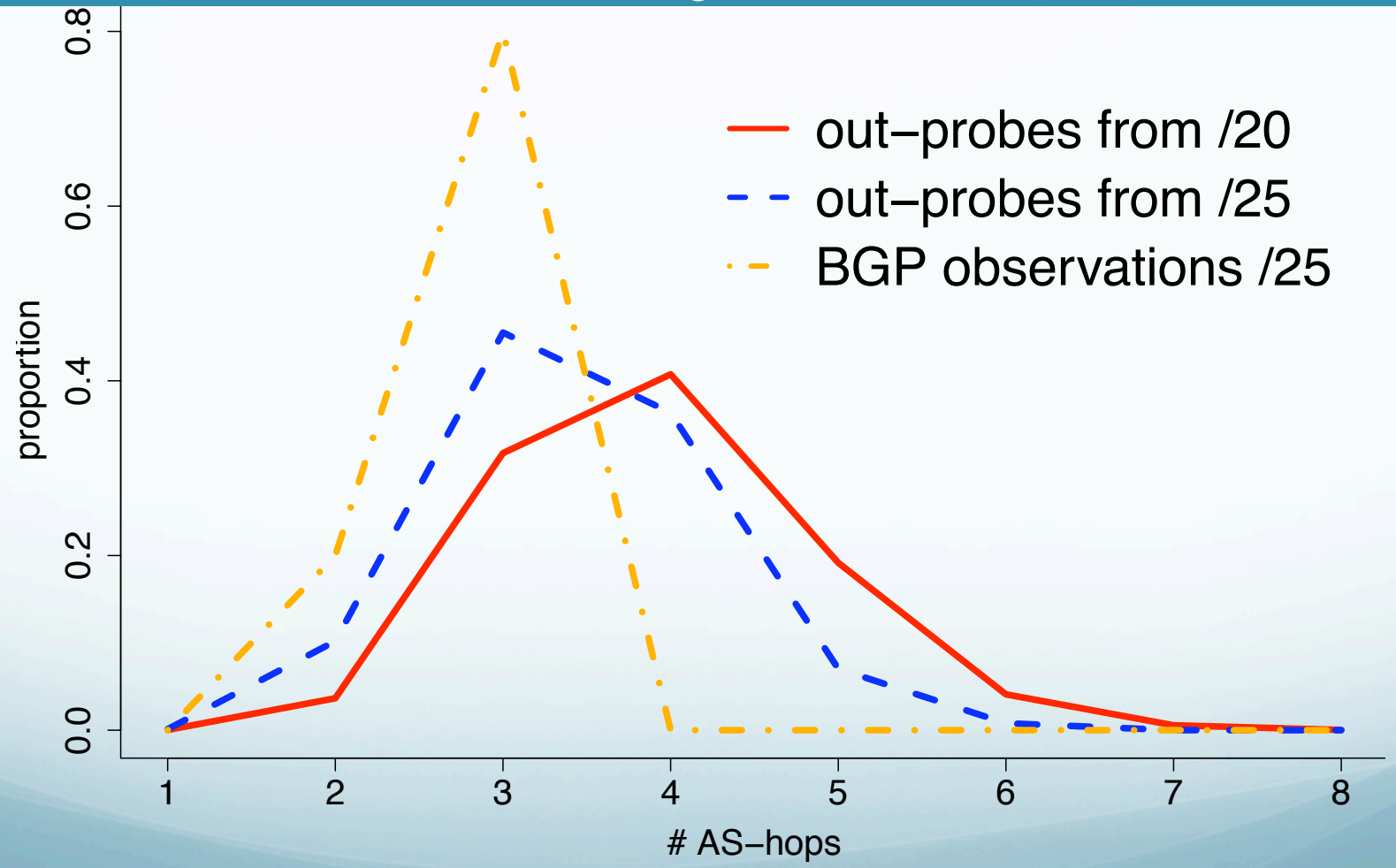
“Out-Probe” Technique

- We ping from /25 to “all” ASes
- **1024** ASes had connectivity!

How far does a /25 propagate?

Control Plane vs. Data Plane Measurements – Expecting a correlation?

Bias in Routing Measurements



How far does a /25 propagate?

Control Plane vs. Data Plane Measurements

Implications

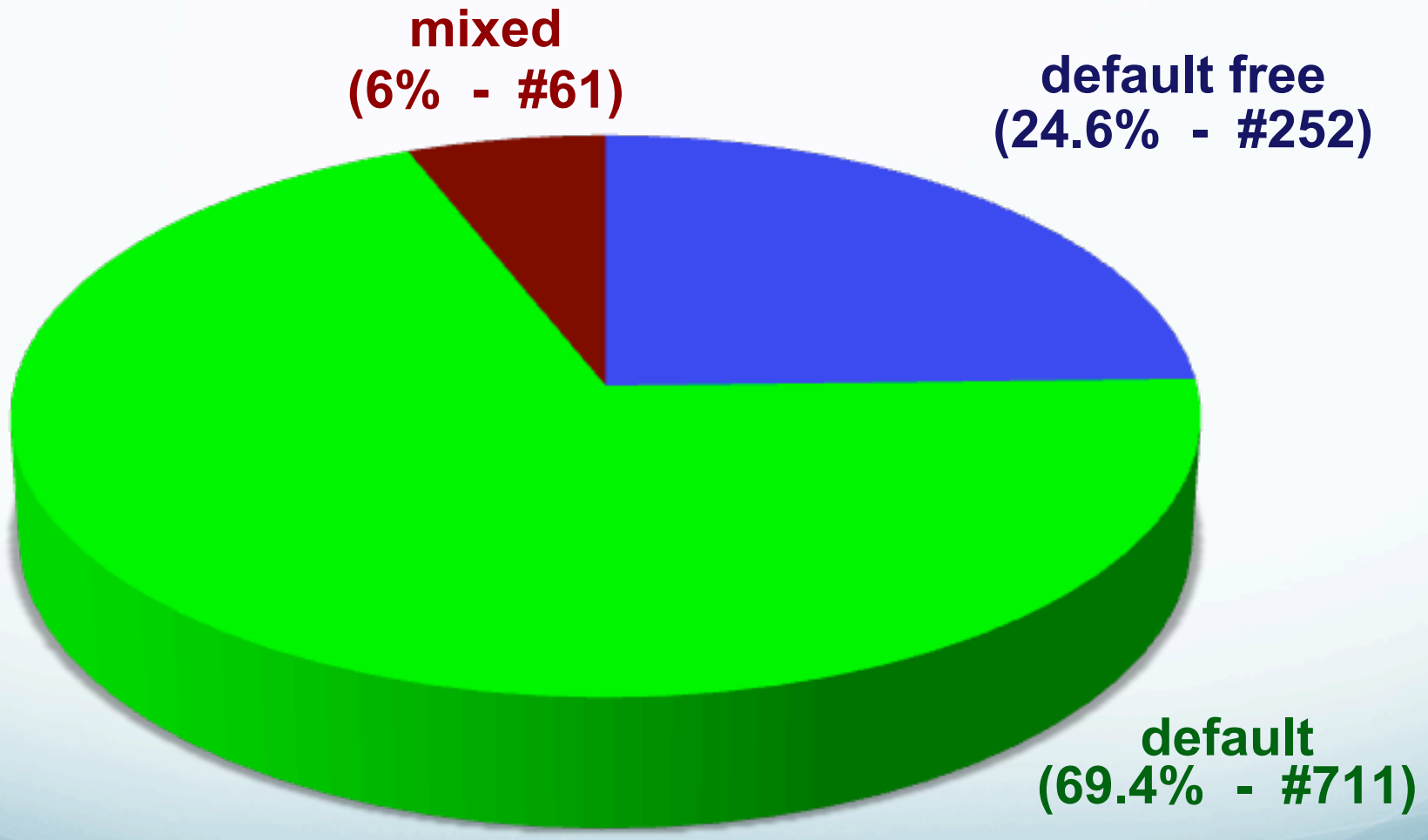
- Did they receive the BGP prefix and it just did not show in Route Views/RIS?
- Bias of Route Views or RIPE/RIS ?
- Did they have a “default-route” to someone who could reach us??

Follow-up questions:

- How much of this was due to default routes as opposed to poor BGP ‘visibility’?

Use of Default Routing in the /25-Experiment

Measurement Results for those 1024 reachable ASes



Use of Default Routing in the Internet

Measurement Results for ≈96% of transit ASes and ≈77% of the 'edge'.

	tested/total	default	default-free	mixed
stub	24,224/31,517	77.1%	19.3%	3.6%
small ISP	1,307/1,361	44.5%	42.2%	13.3%
large ISP	246/255	17.1%	60.6%	22.3%

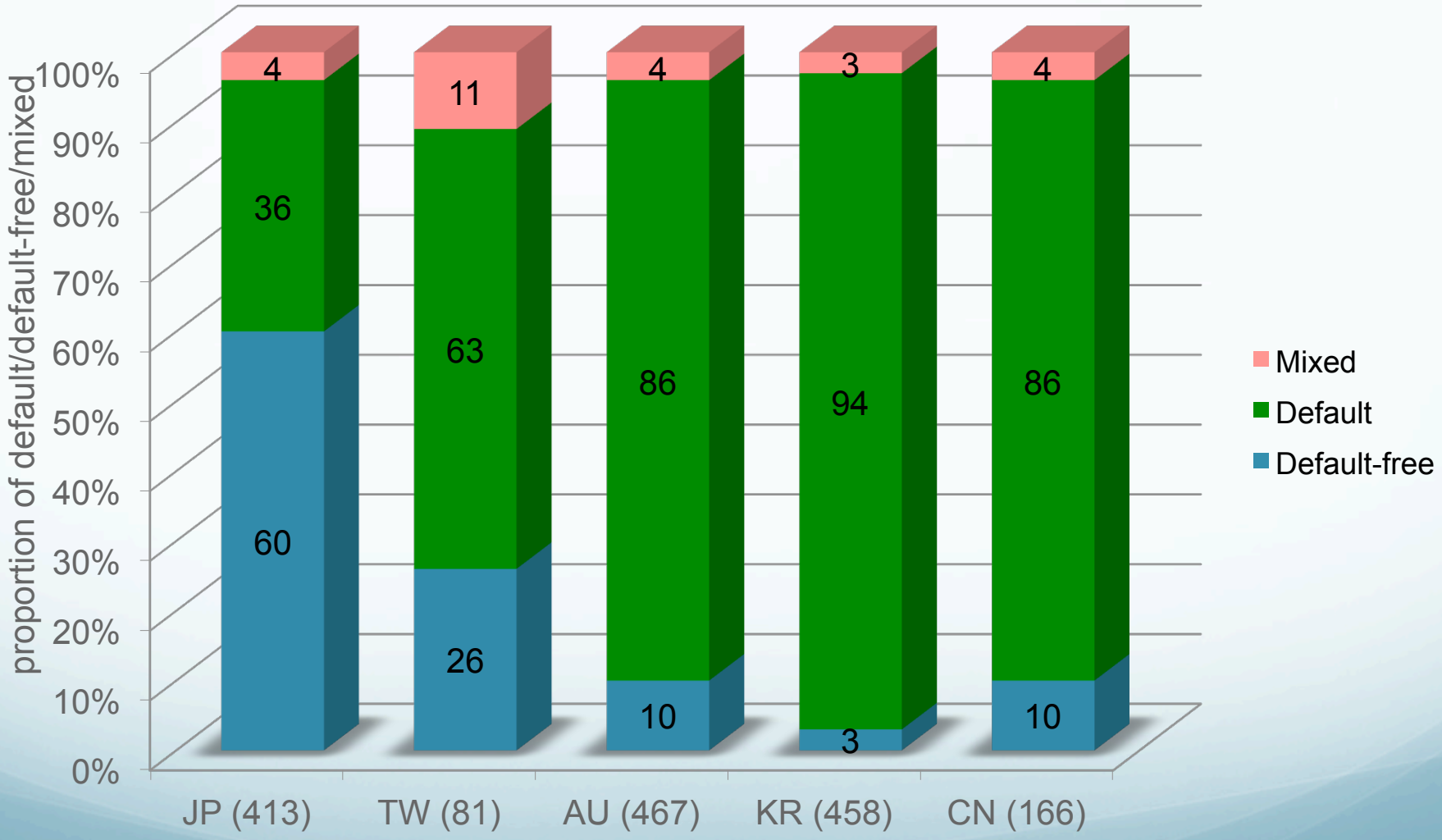
Validation from operator survey

191 operators answered,

- 158 (82.7%) said "correct",
- 12 (6.3%) "almost" correct (e.g., correctly measured, but network is more complex),
- 9 (4.7%) believed we are right (did not recheck),
- 7 (3.7%) we measured wrongly (e.g., AS address space from different provider),
- 5 (2.6%) believed we must be wrong.

Default in Different Regions

Different Countries Seems to Have Different Properties…

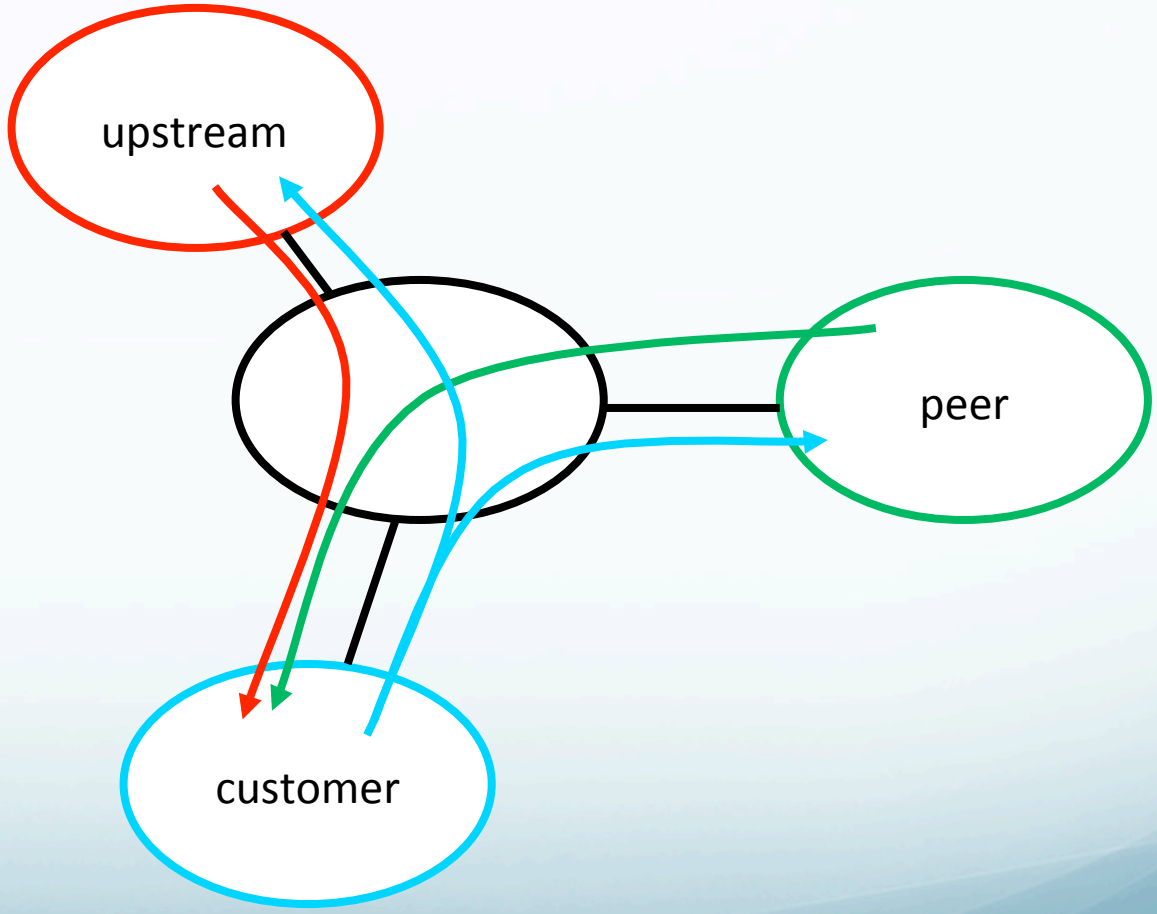


thanks to Tomoya Yoshida for this contribution!

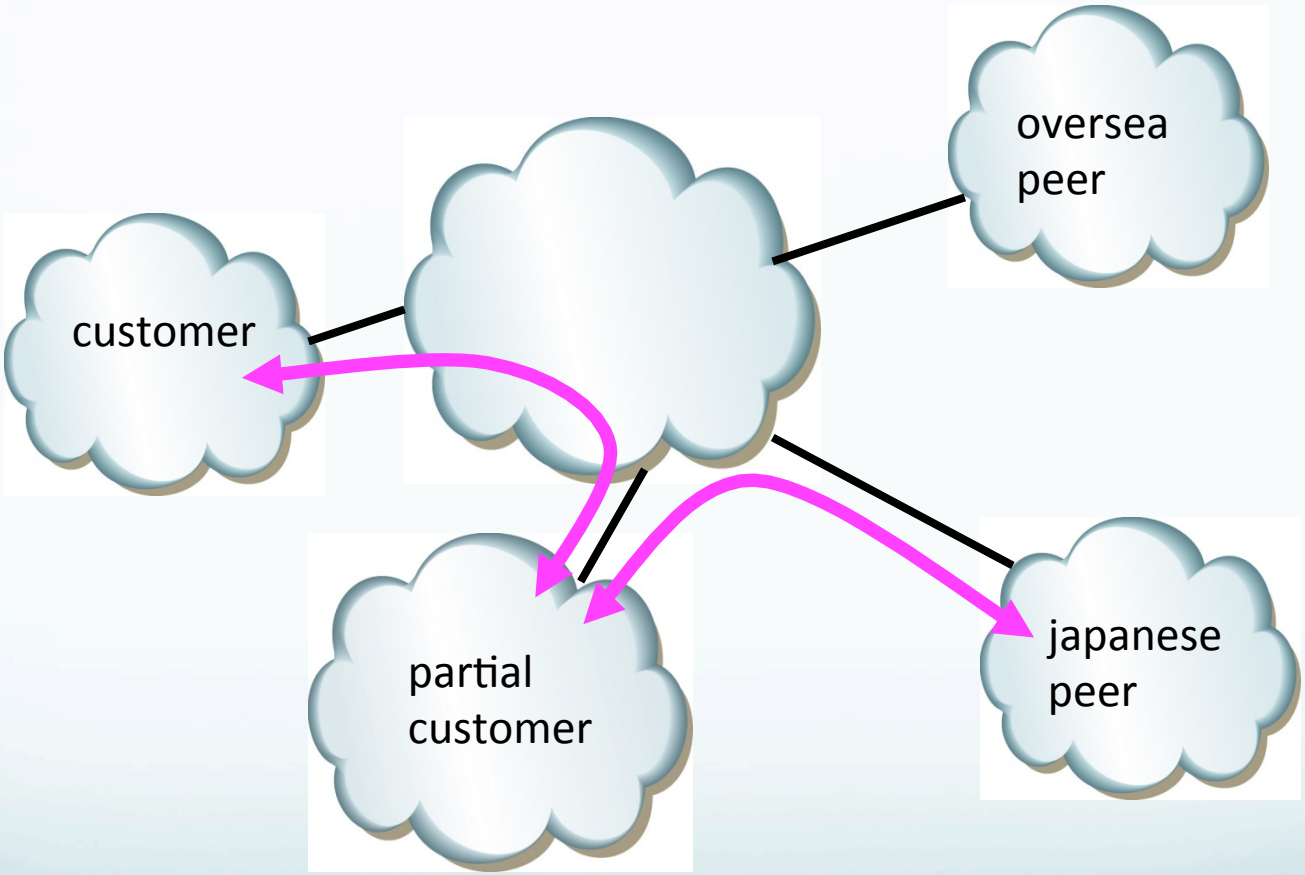
Routing vs. Forwarding

and the story of modern complex policies…

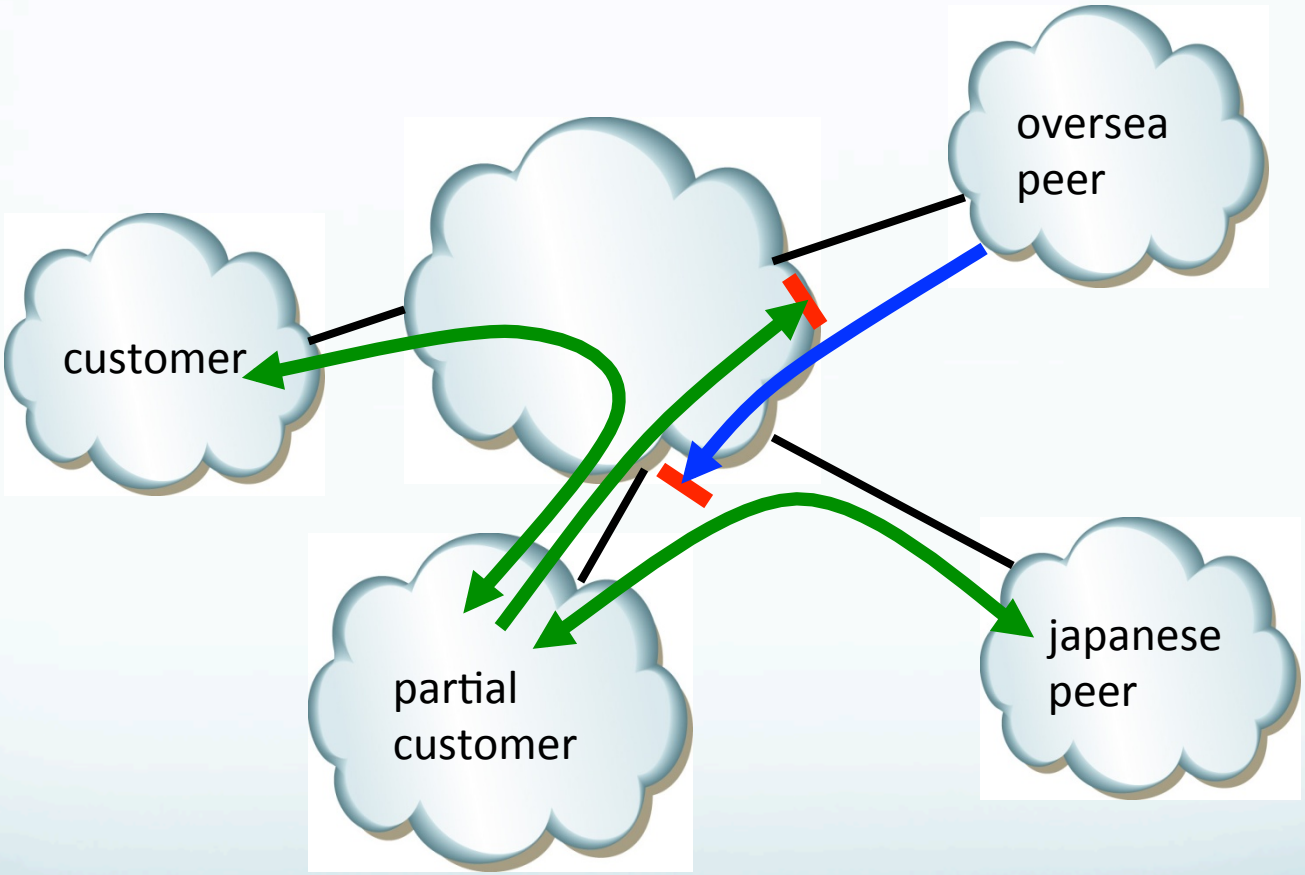
- customer
- peer
- upstream



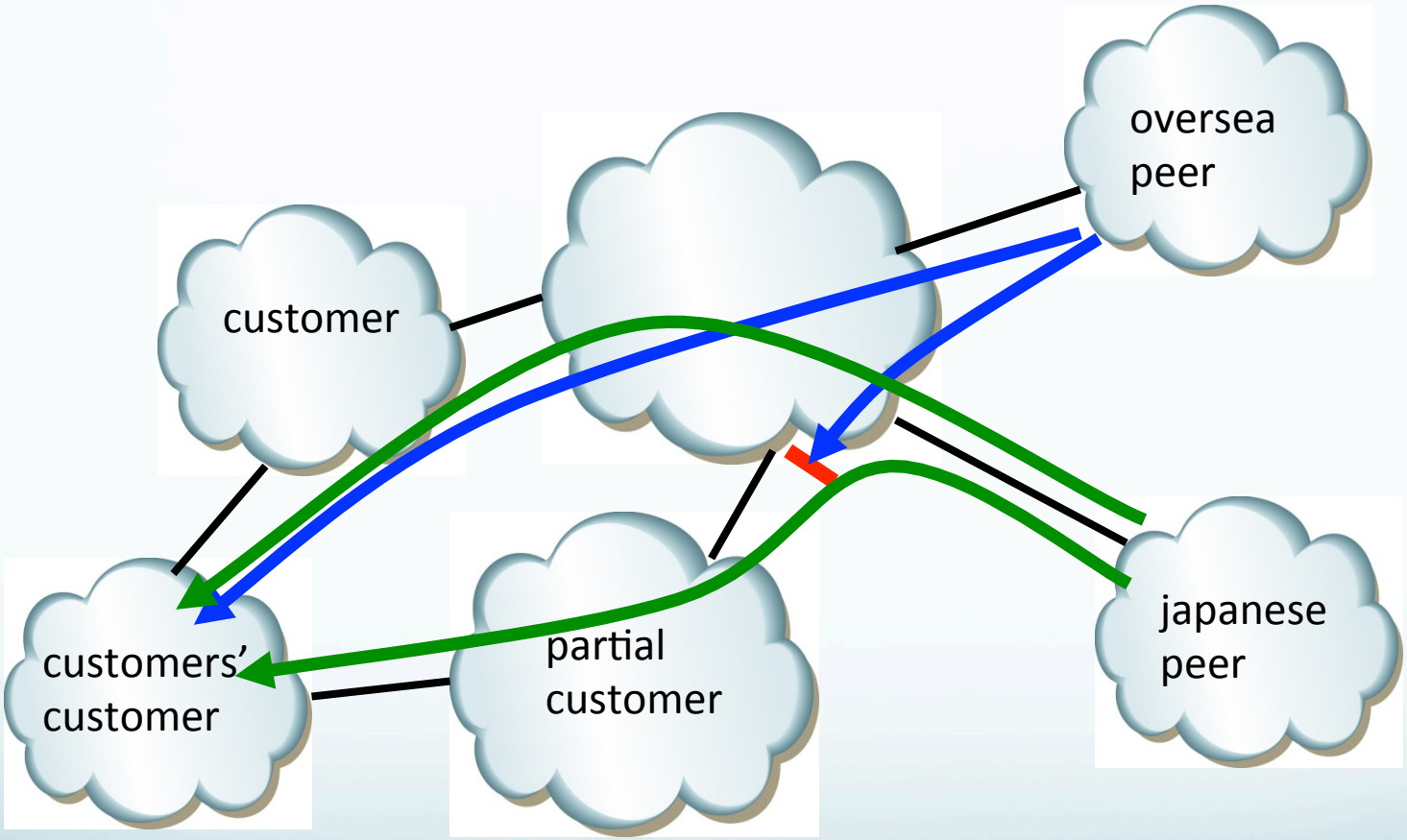
partial transit



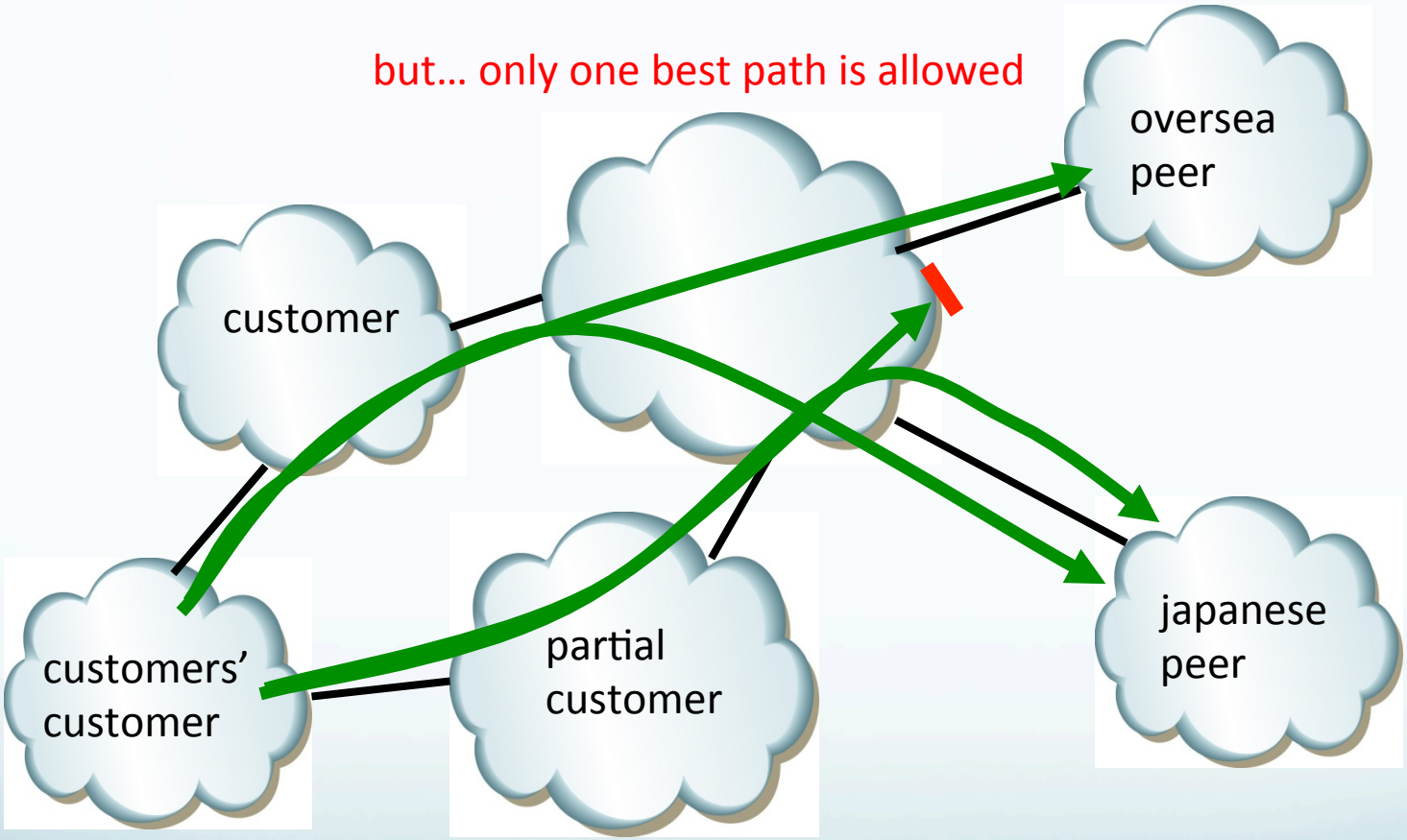
partial transit: desired propagation



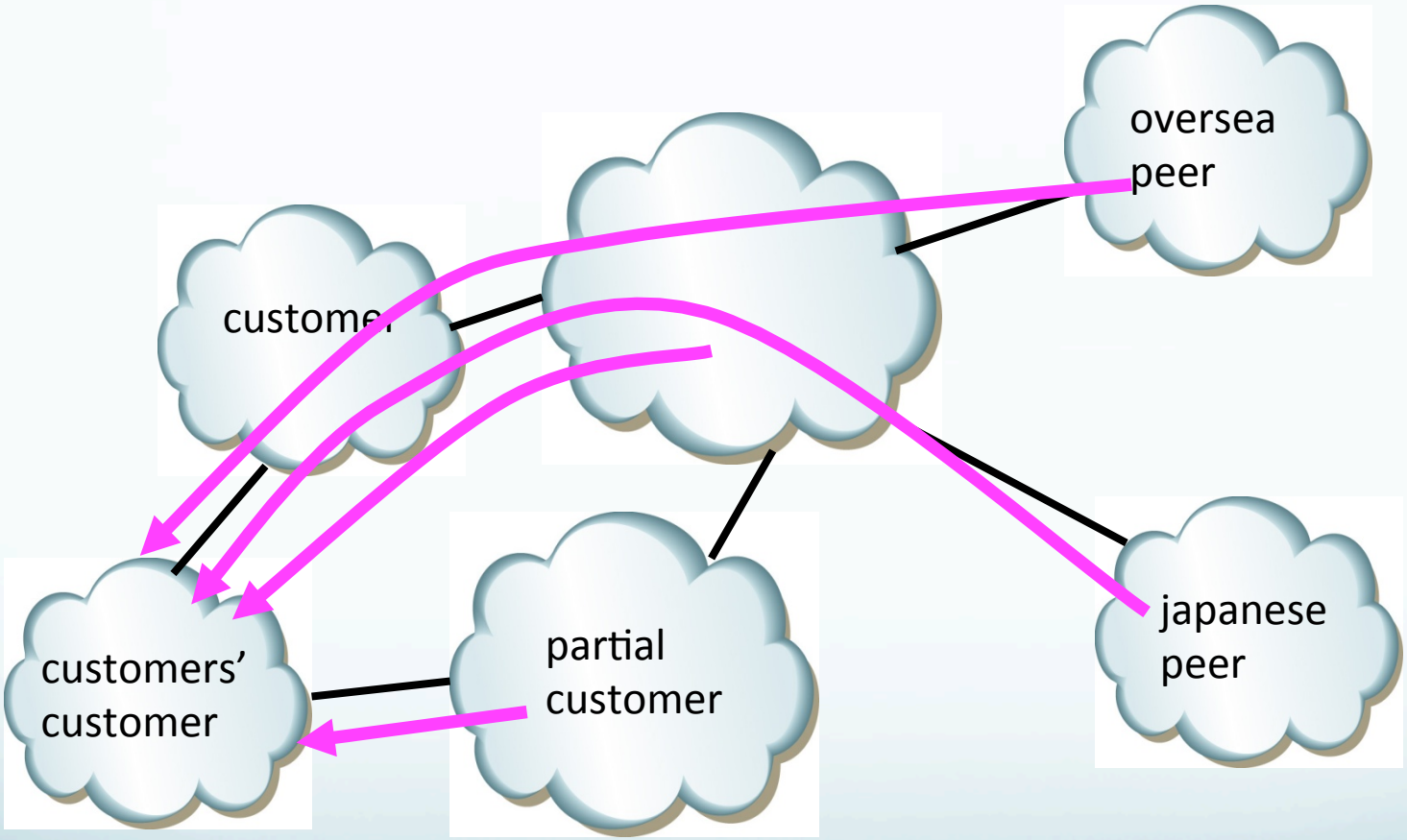
traffic: customers' customer



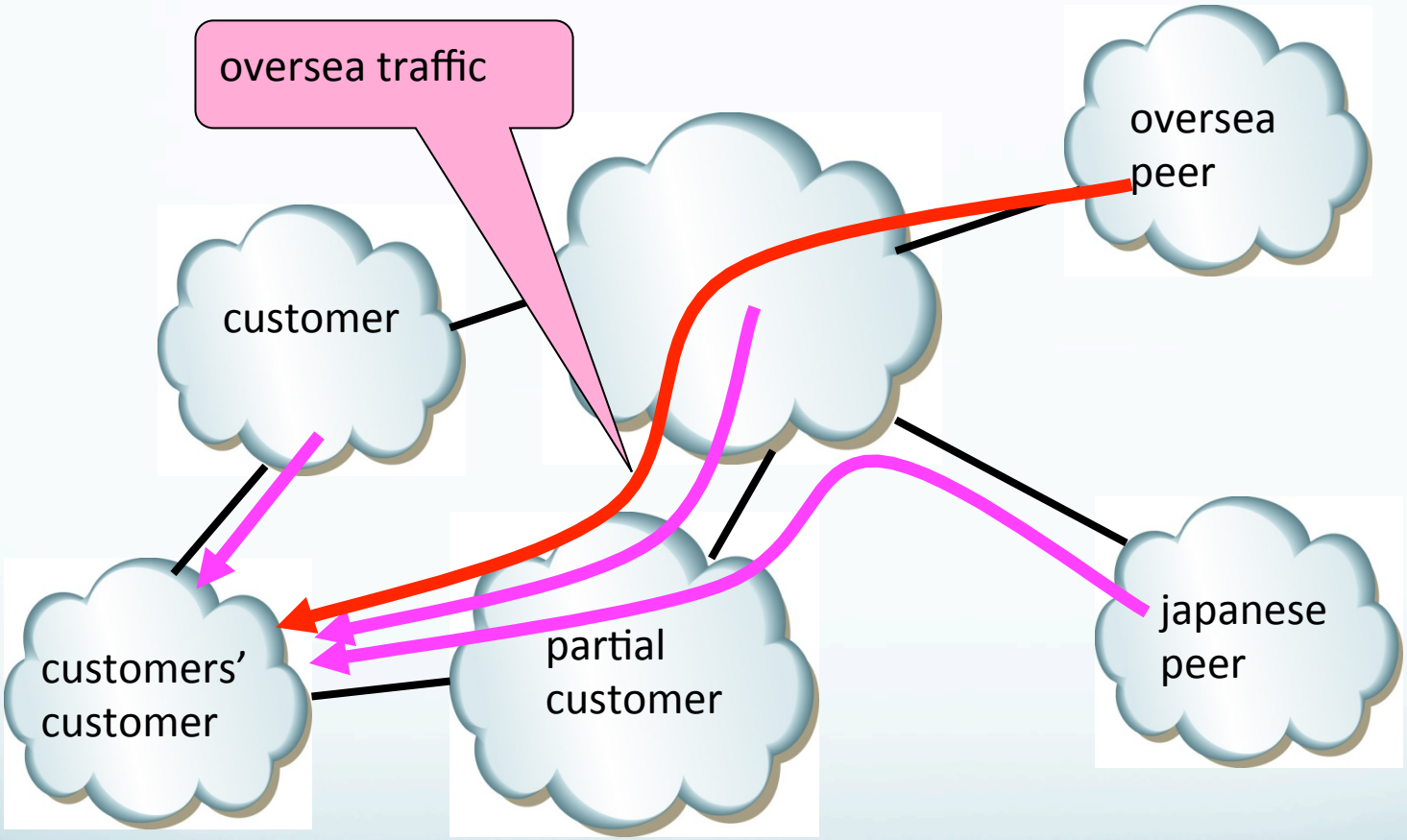
announcement from the customers' customer



expected: traffic



observed traffic

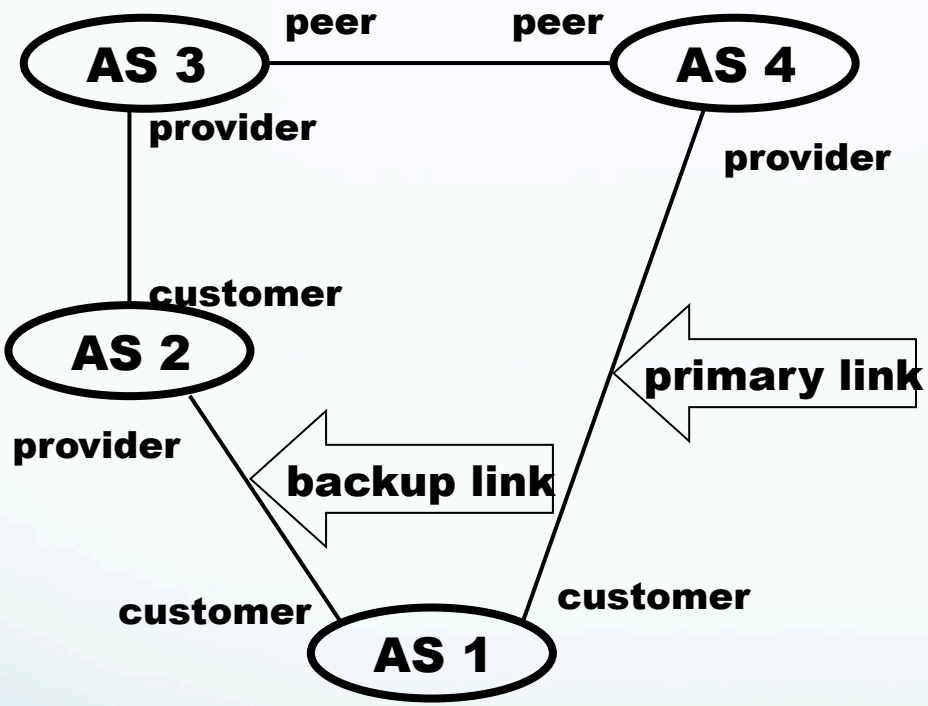


Conclusion

- Routing Research Is Fun! :)
  - Disagreement of control plane and data plane measurements. May explain counter-intuitive results in “Happy packets” (Bush)
  - What do we actually know about routing, ASes, policies??
  - What information is needed to debug the Internet, to identify problems?
- How to assure the robustness of the Internet??
  - Timothy G. Griffin’s “BGP wedgies” (RFC4264) are just another good example of poor “visibility”.
  - Architectural implications? How to design protocols and networks in the future?

Thanks!

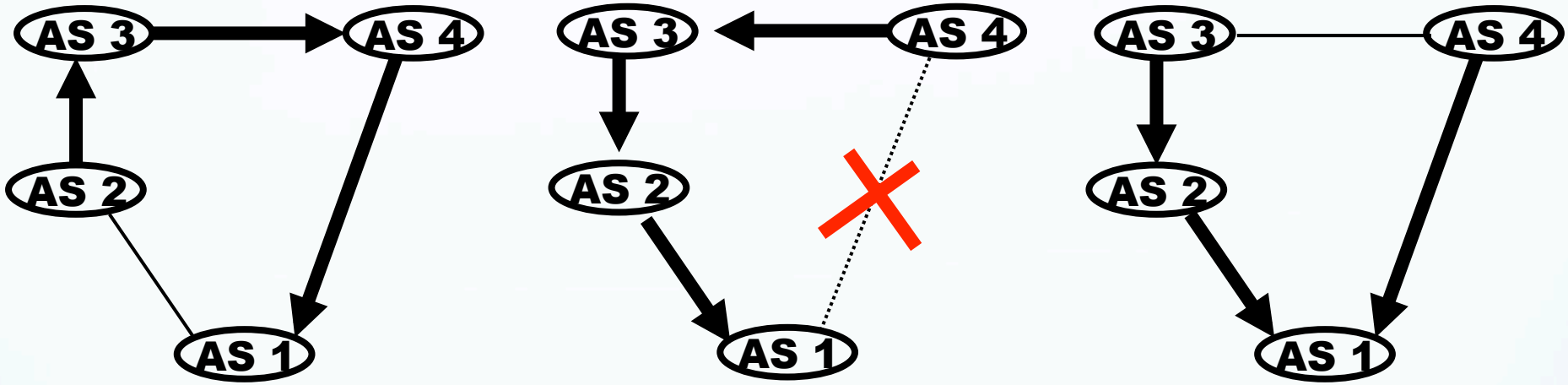
Wedgie Example



- 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)

Tim Griffin, "BGP Wedgies", RFC 4264.

Getting wedged…



Happy, happy, joy, joy

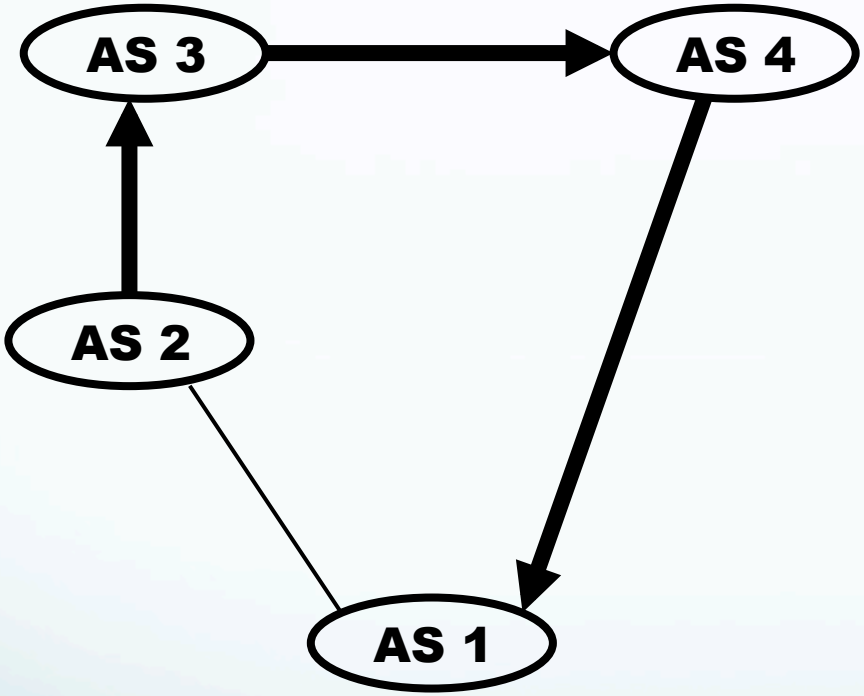
Backups are good!

OH NO, I'M WEDGED!

Primary fails

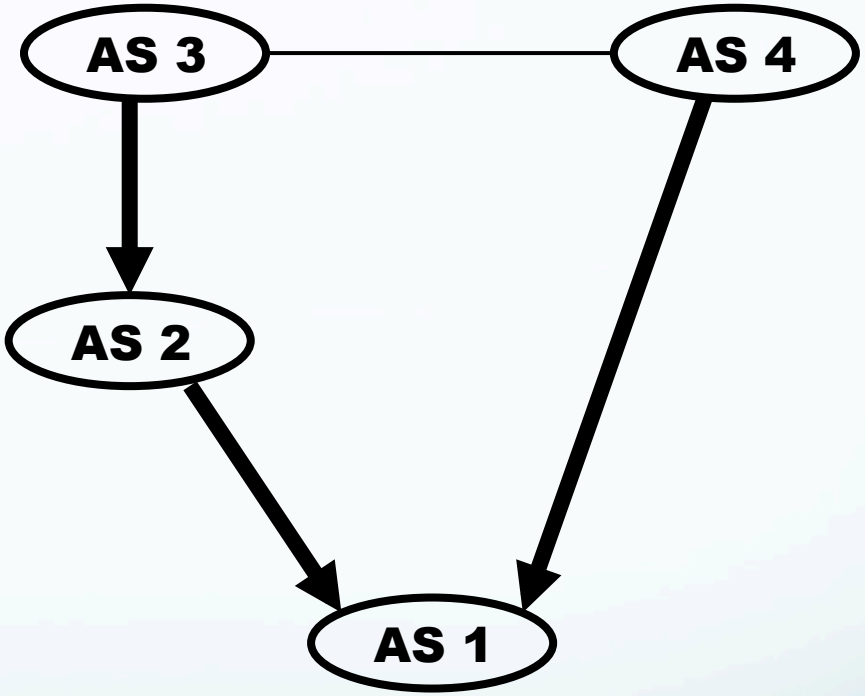
Primary comes back up!

And the Routings are…



Intended Routing

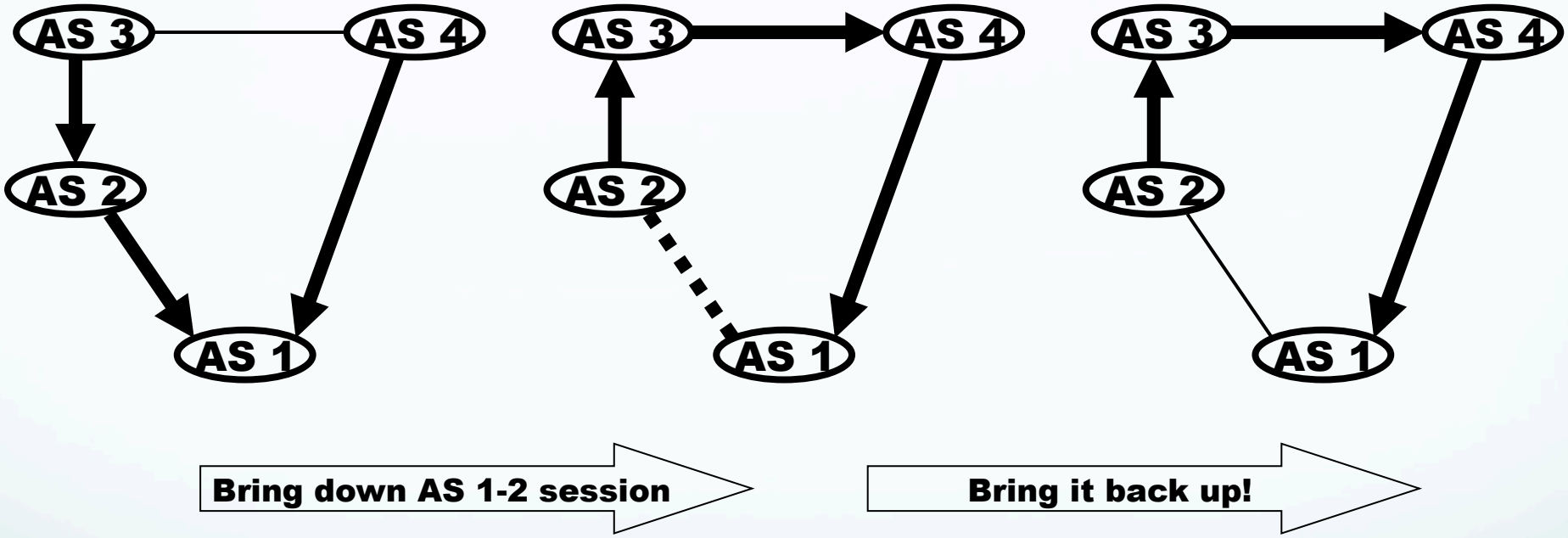
Note: this would be the ONLY routing if AS2 translated its "depref me" community to a "depref me" community of AS 3



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 the heck 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.
- There is no guarantee that a BGP configuration has any solution!
  - And checking configurations NP-Complete
  - Lab demonstrations of BGP configs never converging
- Complex policies (weights, communities setting preferences, and so on) increase chances of routing anomalies.
  - … yet this is the current trend!