

Towards an Accurate AS-level Traceroute Tool

ACM SIGCOMM 2003
Karlsruhe Germany

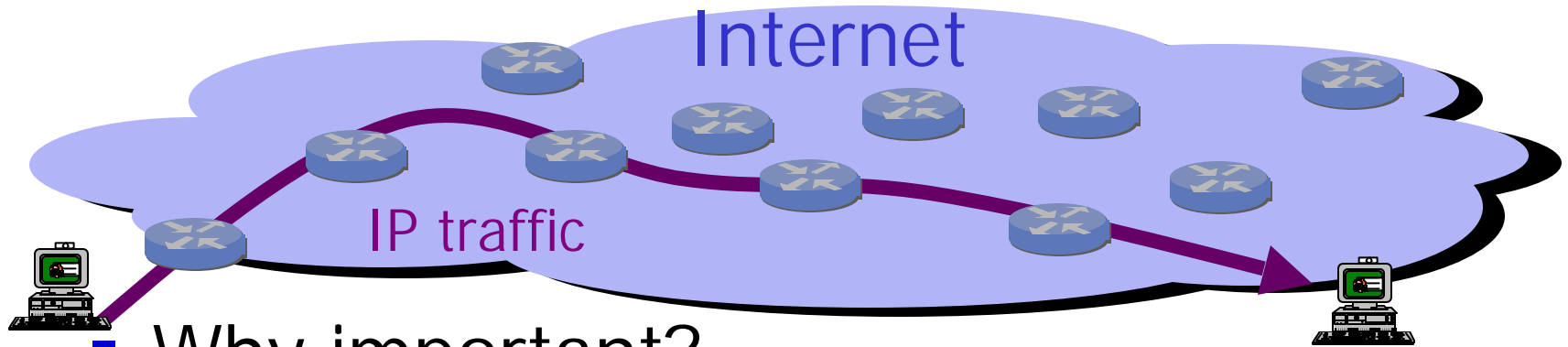
Z. Morley Mao^{*}, Jennifer Rexford^x,
Jia Wang^x, Randy Katz^{*}

^{*}University of California at Berkeley

^xAT&T Labs--Research

Motivation

- What is the **forwarding path**?
 - The path packets traverse through the Internet.



- Why important?
 - Characterize end-to-end network paths
 - Discover Internet topology
 - Detect routing anomalies

Traceroute gives IP-level forwarding path

Traceroute output: (hop number, IP address, DNS name)

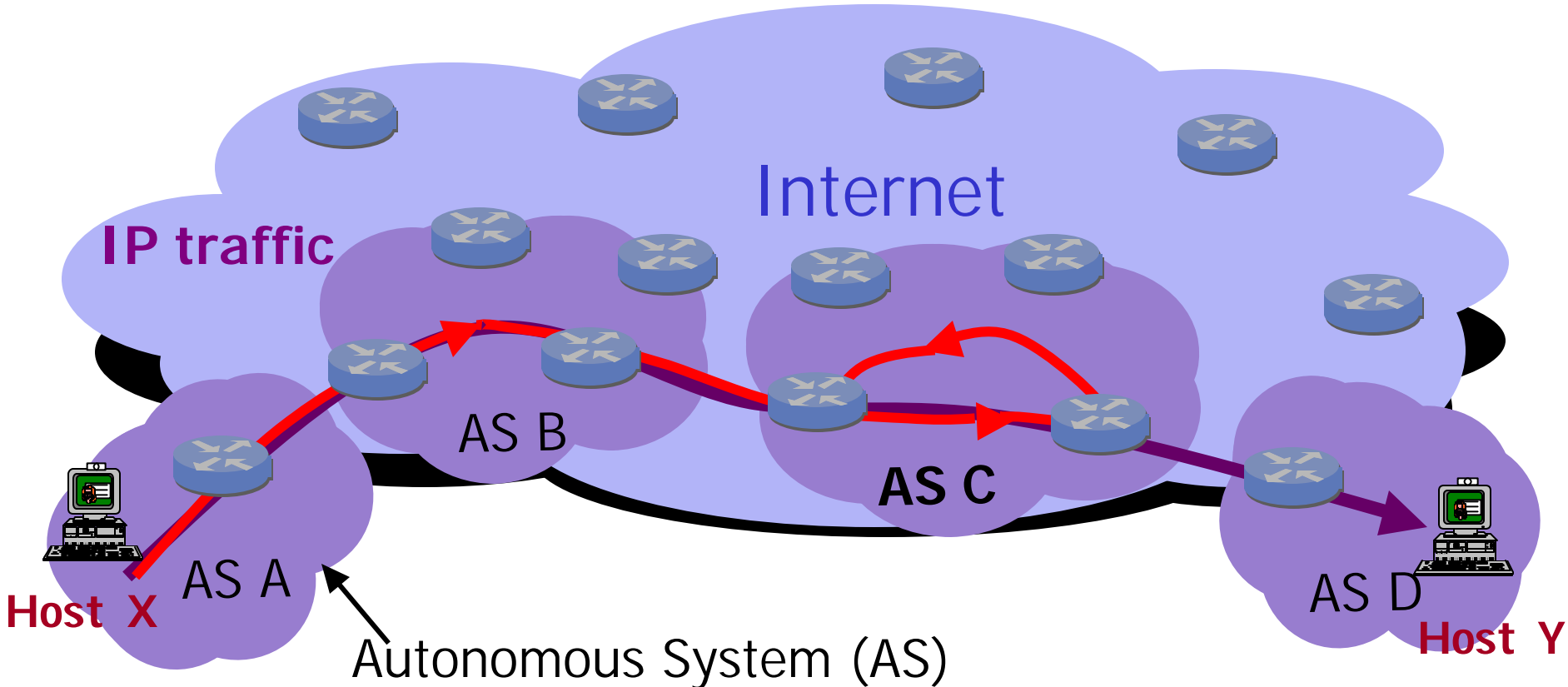
1	169.229.62.1	inr-daedalus-0.CS.Berkeley.EDU
2	169.229.59.225	soda-cr-1-1-soda-br-6-2
3	128.32.255.169	vlan242.inr-202-doecev.Berkeley.EDU
4	128.32.0.249	gigE6-0-0.inr-666-doecev.Berkeley.EDU
5	128.32.0.66	qsv-juniper--ucb-gw.calren2.net
6	209.247.159.109	POS1-0.hsipaccess1.SanJose1.Level3.net
7	*	?
8	64.159.1.46	?
9	209.247.9.170	pos8-0.hsa2.Atlanta2.Level3.net
10	66.185.138.33	pop2-atm-P0-2.atdn.net
11	*	?
12	66.185.136.17	pop1-atl-P4-0.atdn.net
13	64.236.16.52	www4.cnn.com

Traceroute from
Berkeley to
www.cnn.com
(64.236.16.52)

Why is AS-level path useful?

Example use:

Locating routing loops to find responsible networks:
Need **AS-level** forwarding path!

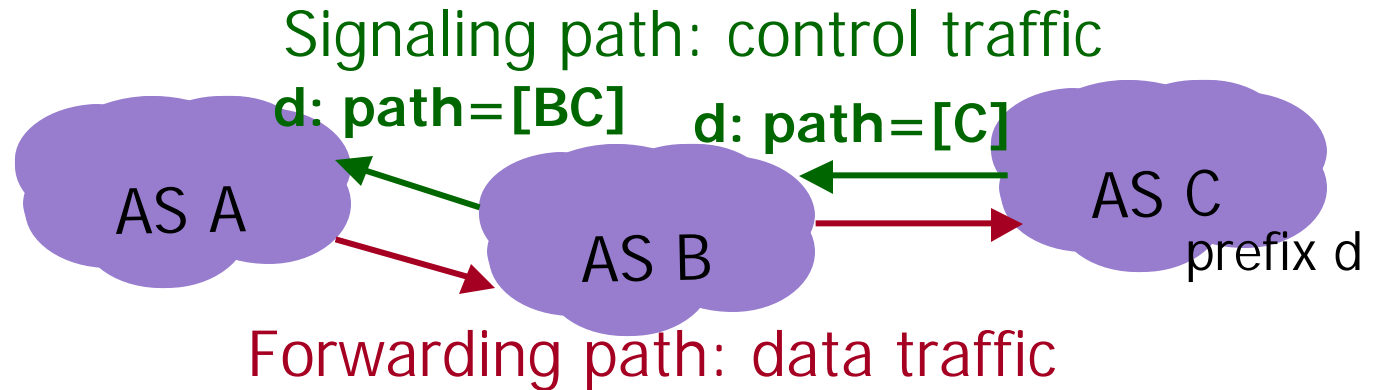


BGP path is not the answer.

Interdomain Routing using Border Gateway Protocol (BGP)

A's local BGP table:

Prefix	BGP Path
d	[A B C]
...	...



- Requires timely access to BGP data
- Signaling path **may differ from** forwarding path:
 - Routing anomalies: e.g., deflections, loops [Griffin2002]
 - Route aggregation and filtering
 - BGP misconfigurations: e.g., incorrect AS prepending

Our approach to obtain AS-level path

Traceroute output: (hop number, IP)

1	169.229.62.1	AS25	Berkeley
2	169.229.59.225	AS25	
3	128.32.255.169	AS25	
4	128.32.0.249	AS25	
5	128.32.0.66	AS11423	Calren
6	209.247.159.109	AS3356	Level3
7	*	AS3356	
8	64.159.1.46	AS3356	
9	209.247.9.170	AS3356	
10	66.185.138.33	AS1668	GNN
11	*	AS1668	
12	66.185.136.17	AS1668	
13	64.236.16.52	AS5662	CNN

1. Start with traceroute IP paths
2. Translate IPs to ASes

Need **accurate** IP-to-AS mappings (for network equipment).

Strawman approaches to get IP-to-AS mappings

- Routing address registry, e.g., whois.radb.net
 - Incomplete and out-of-date
 - Due to acquisitions, mergers, break-ups of institutions
 - Used by NANOG traceroute, prtraceroute
- Origin AS in BGP paths, e.g., RouteViews
 - Multiple origin AS (MOAS)
 - Misconfiguration, multi-homing, Internet eXchange Points
 - Equipment addresses not advertised globally
 - Addresses announced by someone else
 - Supernet: shared, provider-announced

Assumptions

- BGP data:
 - BGP paths and forwarding paths **mostly** match.
- Equipment IP-to-AS mappings:
 - Mappings from BGP tables are **mostly** correct.
 - Change slowly.
- Based on observations, analysis, and survey
 - E.g., **70%** of BGP paths and traceroute paths match

Solution: combine BGP and traceroute data to find a better answer!

Our approach to obtain IP-to-AS mappings

Initial mappings from origin AS of a **large** set of BGP tables

Traceroute paths from multiple locations

(Ignoring unstable paths)

For each location:

Local BGP paths

Traceroute AS paths

For each location:

- Compare

- Look for known causes of mismatches (e.g., IXP, sibling ASes)

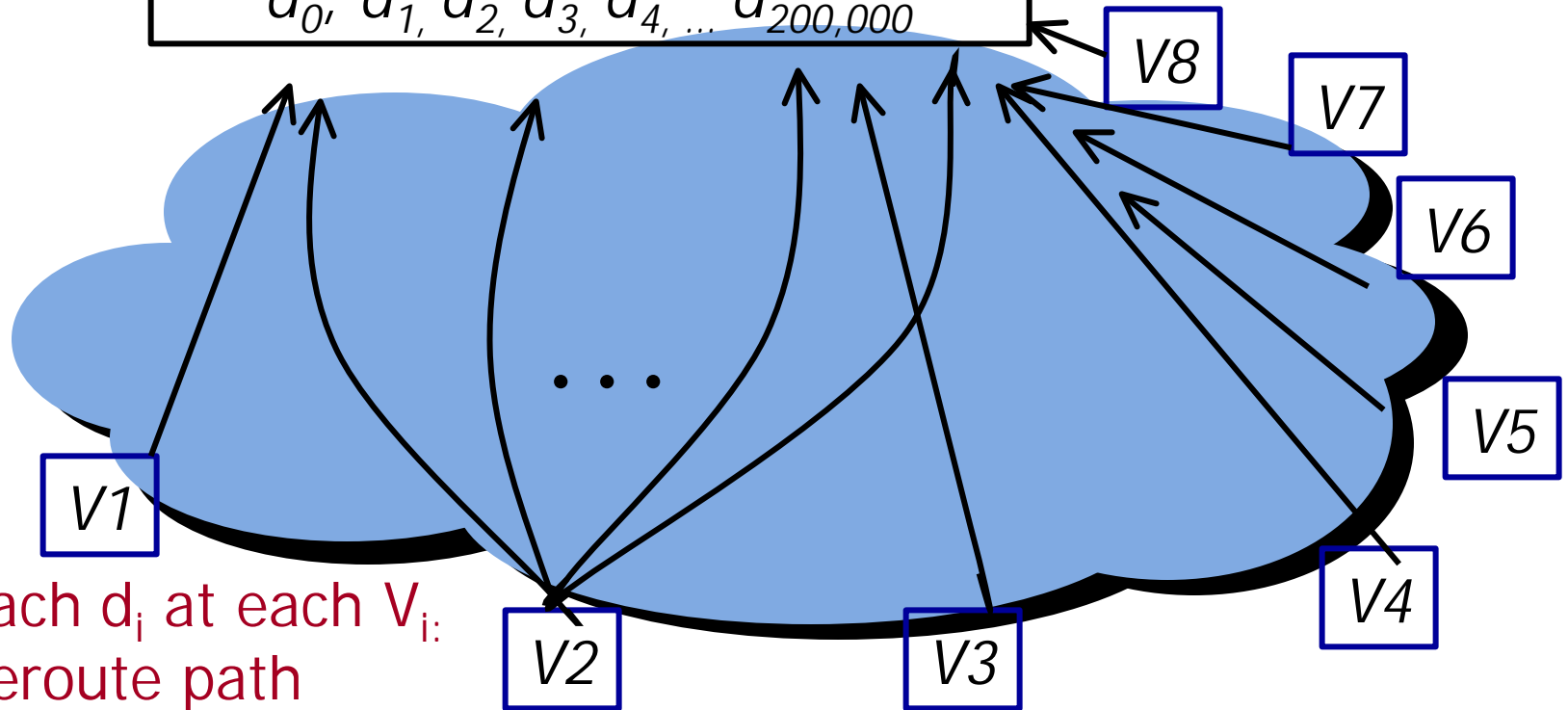
- Edit IP-to-AS mappings

(a single change explaining a large number of mismatches)

Combine **all** locations:

Experiment methodology

200,000 destinations:
 $d_0, d_1, d_2, d_3, d_4, \dots, d_{200,000}$



For each d_i at each V_i :
-Traceroute path
-BGP path

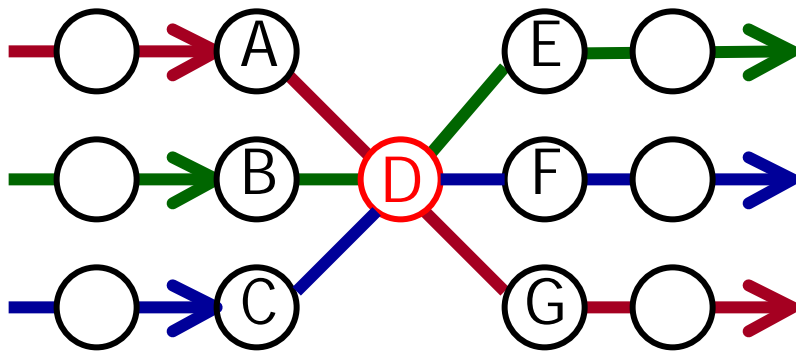
Combine data from multiple vantage points
to modify IP-to-AS mappings.

Why BGP and traceroute paths differ?

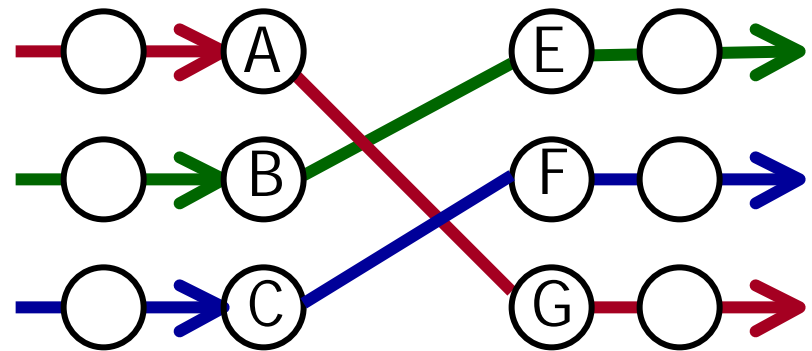
- “Inaccurate” mappings: (corrected)
 - Internet exchange points
 - “Sibling” ASes owned by the same institution
 - Unannounced infrastructure addresses
- Traceroute problems:
 - Forwarding path changing during traceroute
 - Interface numbering at AS boundaries
 - ICMP response refers to *outgoing* interface
- Legitimate mismatches: (interesting to study)
 - Route aggregation and filtering
 - Routing anomalies, e.g., deflections

Extra AS due to IXPs

- Internet eXchange Points (IXP) identification
 - E.g., Mae-East, Mae-West, PAIX
 - Large number of fan-in and fan-out ASes
 - Non transit AS, small address block, likely MOAS



Traceroute AS path

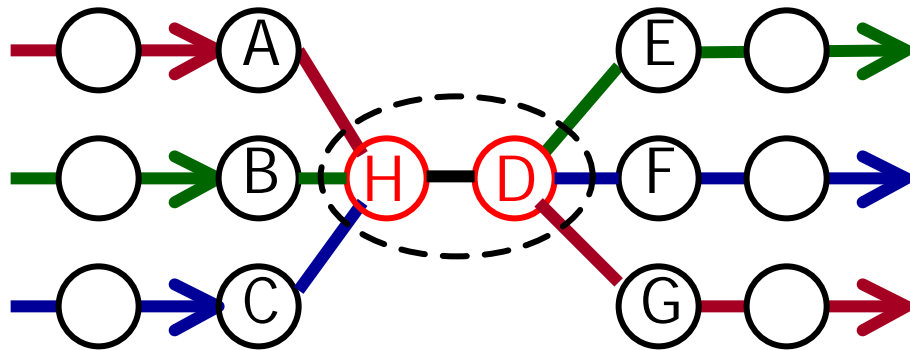


BGP AS path

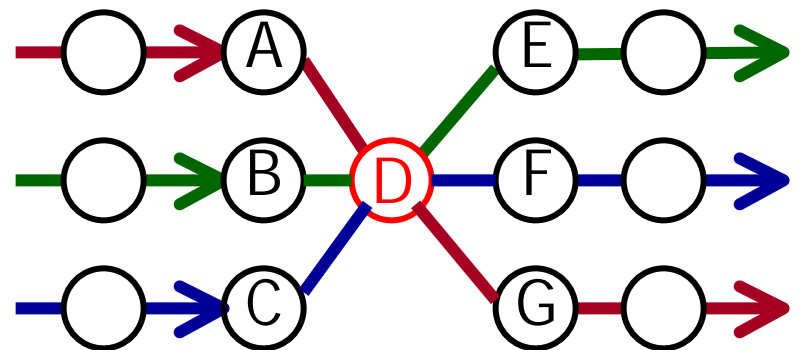
Physical topology and BGP session graph do not always match!

Extra AS due to sibling ASes

- Sibling: organizations with multiple ASes:
 - Sprint (AS1239, AS1791)
 - Mergers, acquisitions
- Identification: Large fan-in and fan-out for the "sibling AS pair"



Traceroute AS path



BGP AS path

Measurement set up

- Eight vantage points
 - Upstream providers: US-centric tier-1 ISPs
- Sweep all routable IP address space
 - About 200,000 IP addresses, 160,000 prefixes, 15,000 destination ASes

Organization	Location	Upstream provider
AT&T Research	NJ, US	UUNET, AT&T
UC Berkeley	CA, US	Qwest, Level3, Internet 2
PSG home network	WA, US	Sprint, Verio
Many thanks to people who let us collect data!		
ArosNet	UT, US	UUNET
Nortel	ON, Canada	AT&T Canada
Vineyard.NET	MA, US	UUNET, Sprint, Level3
Peak Web Hosting	CA, US	Level 3, Global Crossing, Teleglobe

Preprocessing BGP paths

- Discard prefixes with BGP paths containing
 - Routing changes based on BGP updates
 - Private AS numbers
 - Empty AS Paths (local destinations)
 - AS loops from misconfiguration
 - AS SET instead of AS sequence
- Less than 1% prefixes affected

Preprocessing traceroute paths

- Resolving incomplete traceroute paths
 - Unresolved hops within a single AS map to that AS
 - Unmapped hops between ASes
 - Try match to neighboring AS using DNS, Whois
 - Trim unresponsive (*) hops at the end
 - Compare with the beginning of local BGP paths
 - MOAS at the end of paths
 - Assume multi-homing without BGP
- Validation using AT&T router configurations
 - More than 98% cases validated

Vantage point: UC Berkeley

	Initial Mappings			Heuristics		
	Whois	Combined BGP tables	Resolving incompletes	IXP	Sibling ASes	Unannounced address space
Match	44.7%	73.2%	78.0%	84.4%	85.9%	90.6%
Mismatch	29.4%	8.3%	9.0%	8.7%	7.8%	3.5%
Ratio	1.5	8.8	9.0	9.7	11.0	26.0

- Overall modification to mappings:
 - 10% IP-to-AS mappings modified
 - 25 IXPs identified
 - 28 pairs of sibling ASes found
 - 1150 of the /24 prefixes shared

Validations – IXP heuristic

- 25 inferences: 19 confirmed
- Whois/DNS data confirm 18 of 25 inferences
 - AS5459 -- “London Internet Exchange”
 - 198.32.176.0/24:
part of “Exchange Point Blocks”
DNS name: sfba-unicast1-net.eng.paix.net
- Known list from *pch.net* confirm 16 of 25
- Missing 13 known IXPs due to
 - Limited number of measurement locations
 - Mostly tier-1 US-centric providers

Validations - Sibling heuristic

- 28 inferences: all confirmed
- Whois for organization names (15 cases)
 - E.g., AS1299 and AS8233 are TeliaNet
- MOAS origin ASes for several address blocks (13 cases)
 - E.g., 148.231.0.0/16 has MOAS:
AS5677 and AS7132
(Pacific Bell Internet Services and SBC Internet Services)

Conclusion

- Proposed techniques to improve infrastructure IP to AS mappings
 - Match/mismatch ratio improvement: 8-12 to 25-35
 - Reduction of incomplete paths: 18-22% to 6-7%
- Dependence on operational realities:
 - Most BGP routes are relatively stable
 - Few private ASes, AS_SETs
 - Public, routable infrastructure addresses
 - Routers respond with ICMP replies

Ongoing work

- Tool construction and usage:
 - IP-to-AS mapping is available at http://www.research.att.com/~jiawang/as_traceroute
 - Combining with router-level graphs
 - Automatically downloading the most up-to-date mappings
- Systematic optimization:
 - Dynamic-programming and iterative improvement
 - 95% match ratio
 - Write up available at Astrace Web page
- Continuous and scalable data collection
 - Efficient and robust probing techniques
 - Need more diverse vantage points (PlanetLab?)

Towards an Accurate AS-level Traceroute Tool

Tool information available at:

http://www.research.att.com/~jiawang/as_traceroute

Z. Morley Mao^{*}, Jennifer Rexford^x,
Jia Wang^x, Randy Katz^{*}

^{*}University of California at Berkeley

^xAT&T Labs--Research