# Internet Measurement

# MIT 6.829 Computer Networks Fall 2018

Philipp Richter

richterp@csail.mit.edu

suggested reading:

M. Roughan et al. 10 Lessons from 10 Years of Measuring and Modeling the Internet's Autonomous Systems IEEE JSAC 29(9), 2011.

N. Spring et al. Measuring ISP Topologies with Rocketfuel IEEE/ACM Transactions on Networking 12(1), 2004.

Z. Durumeric et al. ZMap: Fast Internet-wide Scanning and Its Security Applications USENIX Security, 2013.

## **Internet Measurement**

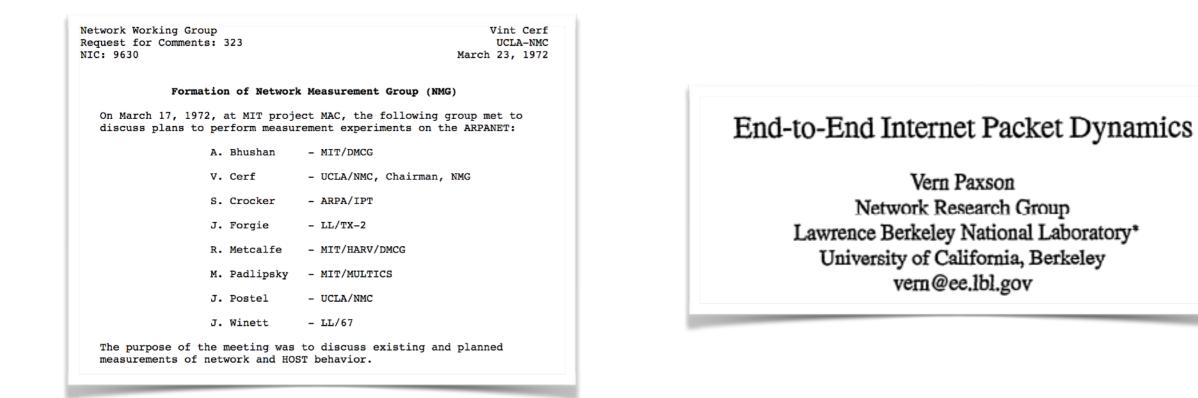
# "Reverse-Engineering the Internet"

"developing and applying techniques to empirically study properties (of interest) of the Internet"

# Motivation

- Network Debugging
- Performance
- Resilience
- Security
- Regulation and Policies
- Broader impact on society: state censorship, price and traffic discrimination, impact of social media, ...

# **Internet Measurements - The Origins**



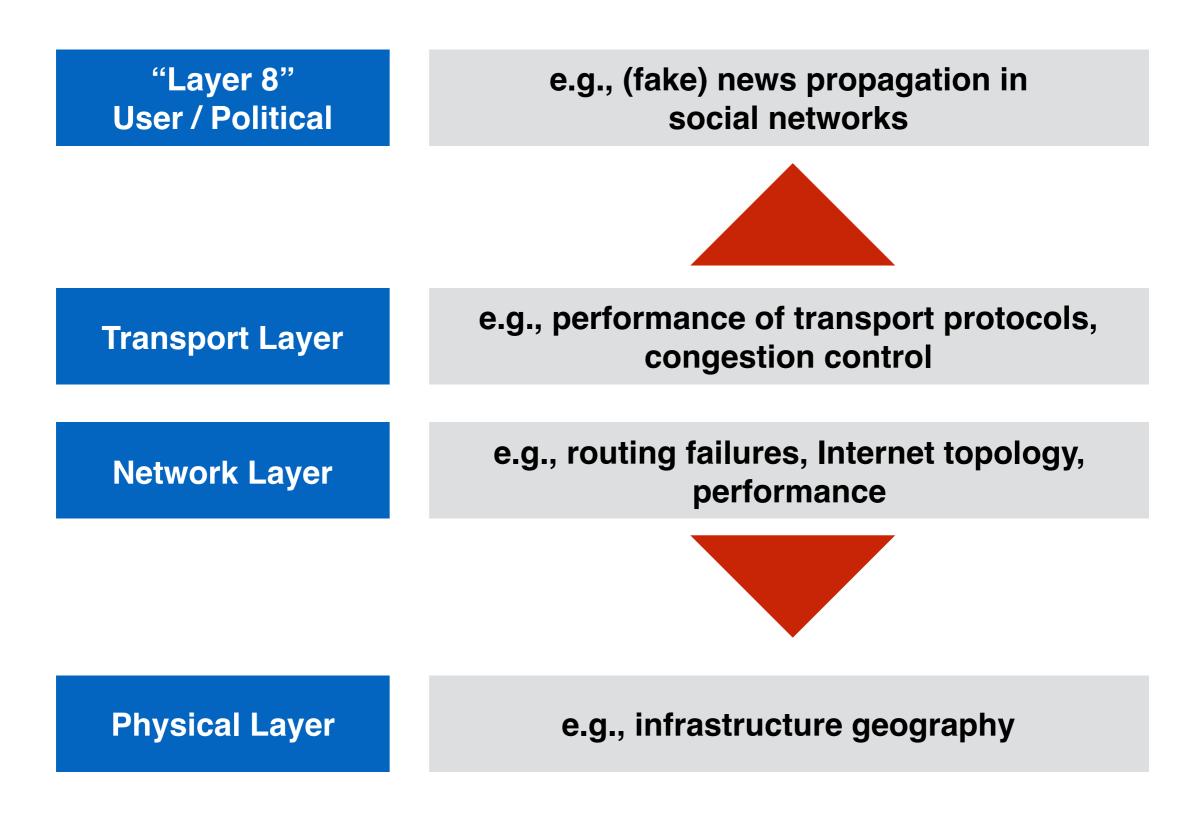
RFC323: IETF formed measurement group(s) as early as 1972 first major academic measurement studies (e.g., Paxson, SIGCOMM 1997)

2001: First ACM SIGCOMM Internet Measurement Workshop 2003: First ACM IMC (Internet Measurement Conference)

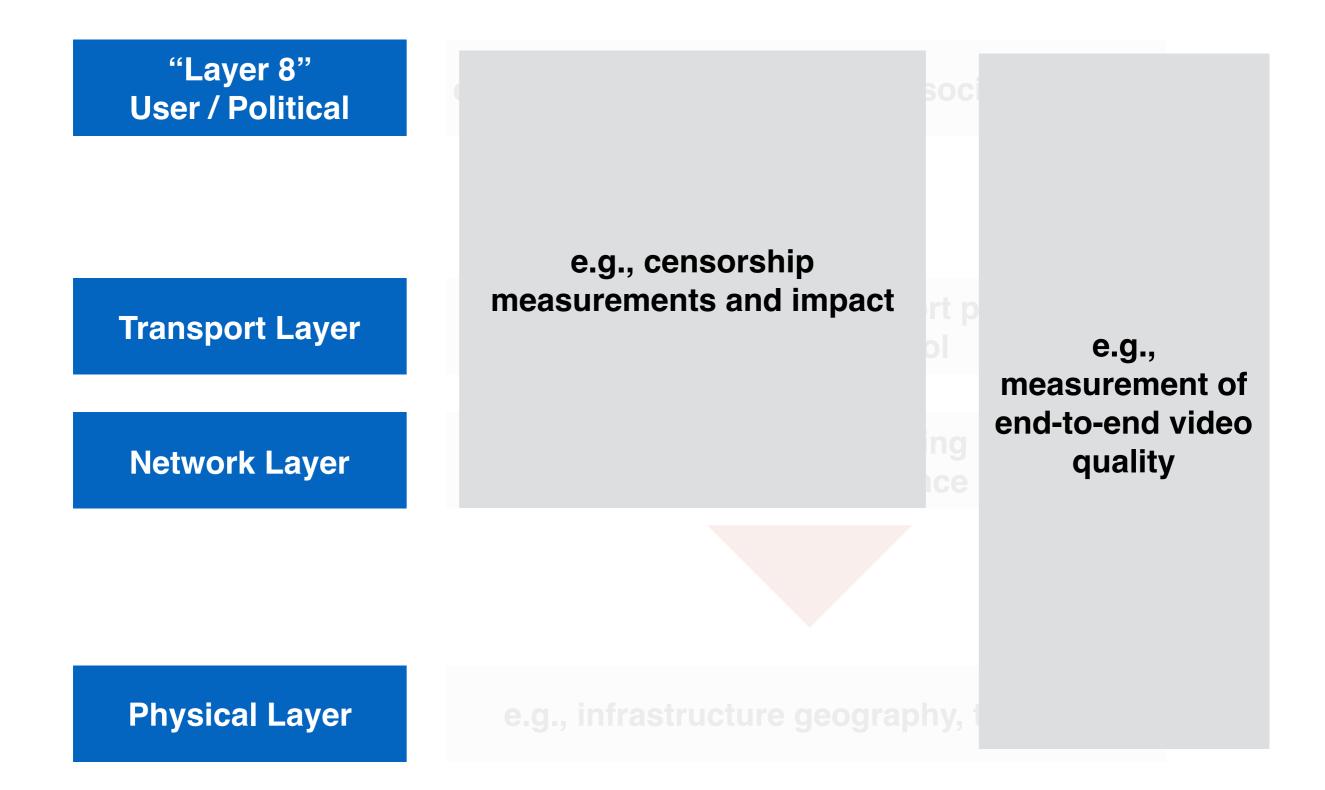
# Internet Measurements - "Classic" (yet highly relevant)

Transport Layer	e.g., performance of transport protocols, congestion control
Network Layer	e.g., routing failures, Internet topology, performance

# **Internet Measurements - A Broadening Field**



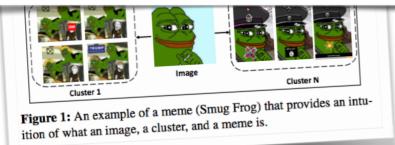
# **Internet Measurements - Cross-layer Measurements**



# **Internet Measurement - A Creative Field**

### On the Origins of Memes by Means of Fringe Web Communities

Savvas Zannettou<sup>\*</sup>, Tristan Caulfield<sup>‡</sup>, Jeremy Blackburn<sup>†</sup>, Emiliano De Cristofaro<sup>‡</sup>, Michael Sirivianos<sup>\*</sup>, Gianluca Stringhini<sup>‡</sup>, and Guillermo Suarez-Tangil<sup>‡+</sup>



### Email Typosquatting

Janos Szurdi Carnegie Mellon University jszurdi@andrew.cmu.edu Nicolas Christin Carnegie Mellon University nicolasc@andrew.cmu.edu

*.exampel.com.	300	MX	1	exampel.com.
exampel.com.	300	MX	1	exampel.com.
*.exampel.com.	300	Α	NA	1.1.1.1
exampel.com.	300	А	NA	1.1.1.1

#### If you are not paying for it, you are the product: How much do advertisers pay to reach you?

Panagiotis Papadopoulos FORTH-ICS, Greece panpap@ics.forth.gr

Pablo Rodriguez Rodriguez Telefonica Alpha, Spain pablo.rodriguezrodriguez@telefonica.com Nicolas Kourtellis Telefonica Research, Spain nicolas.kourtellis@telefonica.com

Nikolaos Laoutaris Data Transparency Lab, Spain nikos@datatransparencylab.org

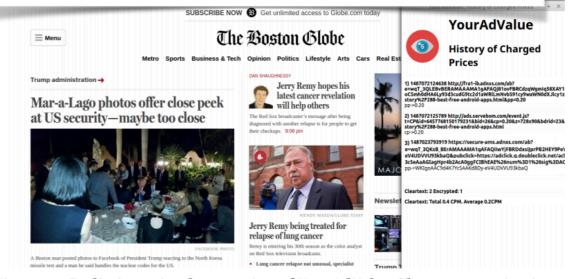


Figure 20: Preliminary implementation of YourAdValue Chrome extension in use.

### Measuring Price Discrimination and Steering on E-commerce Web Sites

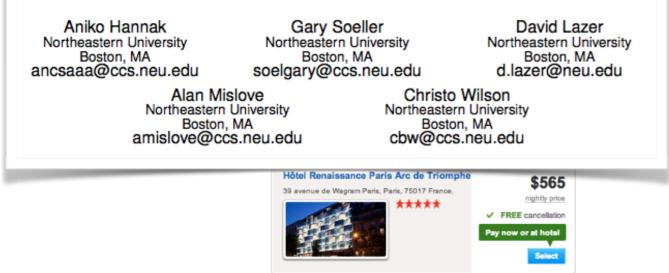
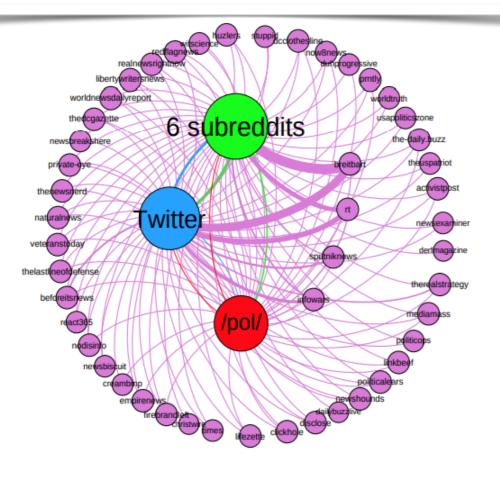


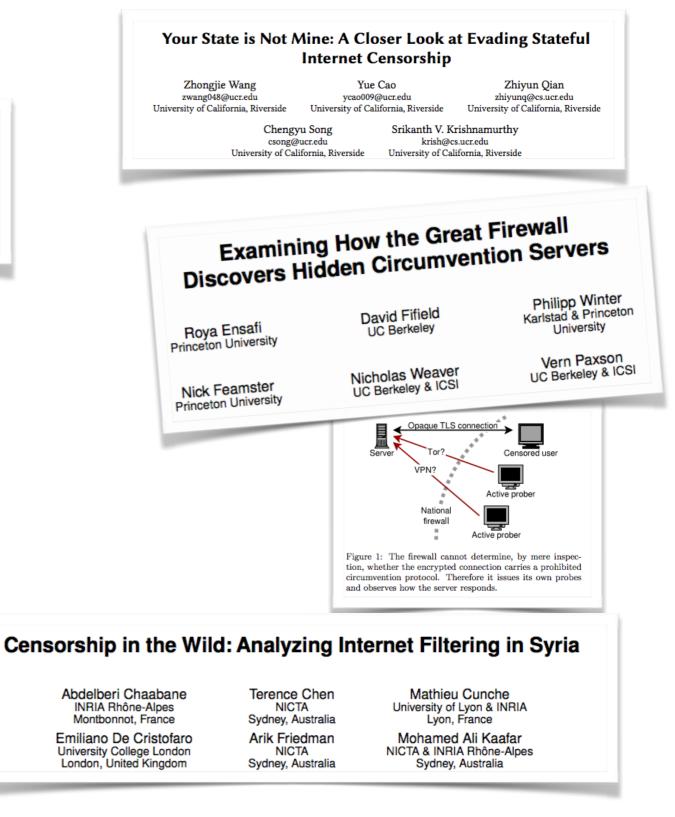
Figure 4: Example of price discrimination. The top result was served to the AMT user, while the bottom result was served to the comparison and control.

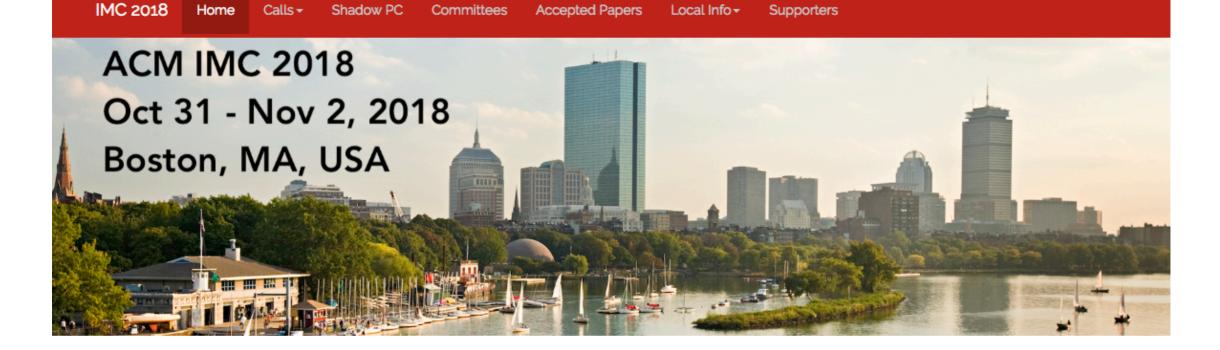
# **Internet Measurement - Broader Societal Impact**

### The Web Centipede: Understanding How Web Communities Influence Each Other Through the Lens of Mainstream and Alternative News Sources

Savvas Zannettou<sup>\*</sup>, Tristan Caulfield<sup>†</sup>, Emiliano De Cristofaro<sup>†</sup>, Nicolas Kourtellis<sup>‡</sup>, Ilias Leontiadis<sup>‡</sup>, Michael Sirivianos<sup>\*</sup>, Gianluca Stringhini<sup>†</sup>, and Jeremy Blackburn<sup>+</sup>







### ACM Internet Measurement Conference 2018

The 2018 Internet Measurement Conference (IMC) is a three-day event focusing on Internet measurement and analysis. The conference is sponsored by ACM SIGCOMM. IMC 2018 is the 18th in a series of highly successful Internet Measurement Workshops and Conferences.

The ACM IMC 2018 conference will be held in Boston, MA, USA on October 31 - November 2, 2018.

# come join us there!

# Internet Measurement - Fundamental Challenges (i)

# Internet: Not designed with measurability in mind

"current measurement practice often involves the exploitation of sideeffects and unintended features of the network, or, in other words, the **artful piling of hacks atop one another.** This state of affairs is a direct result of the relative paucity of diagnostic and measurement capabilities built into today's network stack."

M. Allman et al. "Principles for Measurability in Protocol Design" ACM CCR, 2017.

# Internet Measurement - Fundamental Challenges (ii)

- Lack of ground truth
- Lack of available data
- Heterogeneity of the network
   -> Generalizability of results
- Privacy concerns, Ethics

Internet Topology Measurement Topology (Oxford Dictionary):

"the way in which constituent parts are interrelated or arranged"

# model of the Internet:

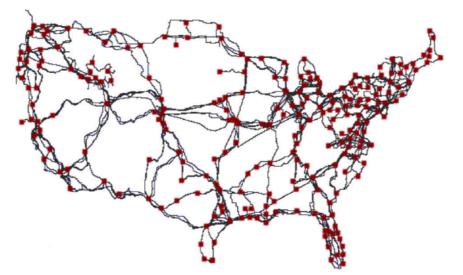
series of tubes? set of routers? nodes and vertices in a graph?

# why does it matter?

fundamental for systems design whatever testbed we have, is it realistic?

Trends in Interconnectivity

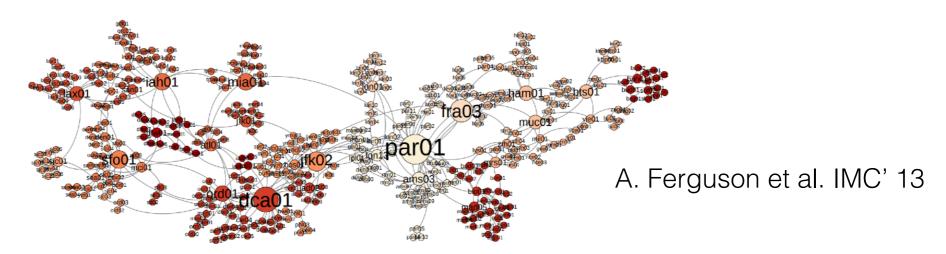
Internet resilience



### R. Durairajan et al. SIGCOMM '15

### Physical

Figure 1: Location of physical conduits for networks considered in the continental United States.



### Logical topology Router Level

Figure 6: Visualization of paths in Cogent's network based on data from the week of April 7, 2013; nodes represent routers, edges link routers sharing the same IP subnet, and nodes are scaled to represent *betweenness* – larger nodes have a greater number of paths passing through them. The layout is force-directed, with no geographical information.

bgp.he.net

Logical topology Autonomous Systems Level

# **AS-level Topology**

Within the Internet, an autonomous system (AS) is a collection of connected Internet Protocol (IP) routing prefixes under the control of one or more network operators on behalf of a single administrative entity or domain that presents a common, clearly defined routing policy to the Internet.

(Wikipedia)

# abstracts entire networks to be single nodes makes things (seemingly) easy!

# goal:

"find the ASes in the Internet and their BGP links"

(many follow-up questions possible)

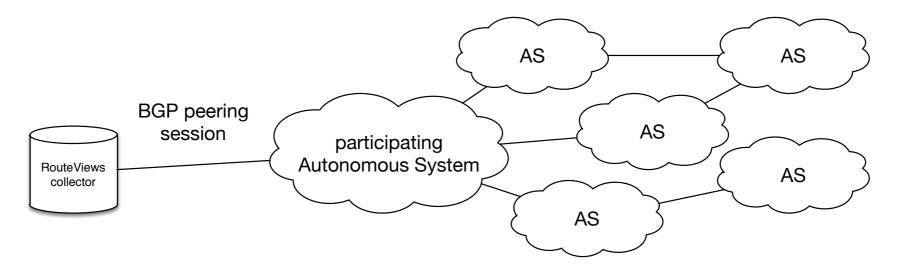
# Passive AS-level topology measurements: Tapping into the global routing system Publicly available data





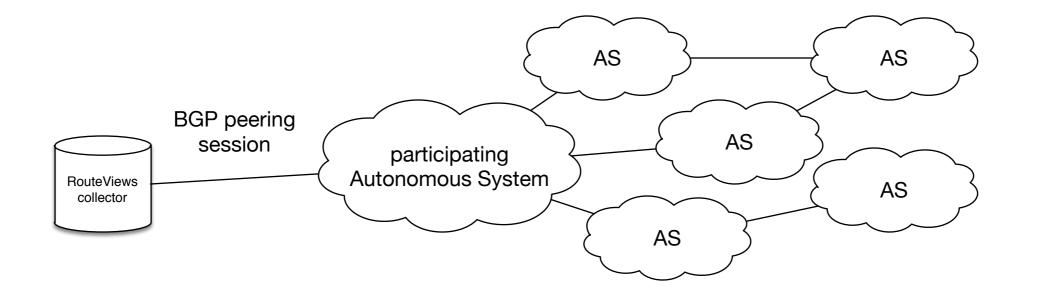


- 100+ route collectors, 1000+ peers ("participating" ASes)
- Collectors establish BGP session and collect messages
- But: they do not "peer" i.e., they do not exchange traffic



\* some ASes "participate" (provide direct feeds) unknowingly, if the route collector has BGP session(s) with IXP route servers. Further reading on IXP route servers: Richter et al., ACM IMC 2014

# Passive AS-level topology measurements: Tapping into the global routing system Publicly available data



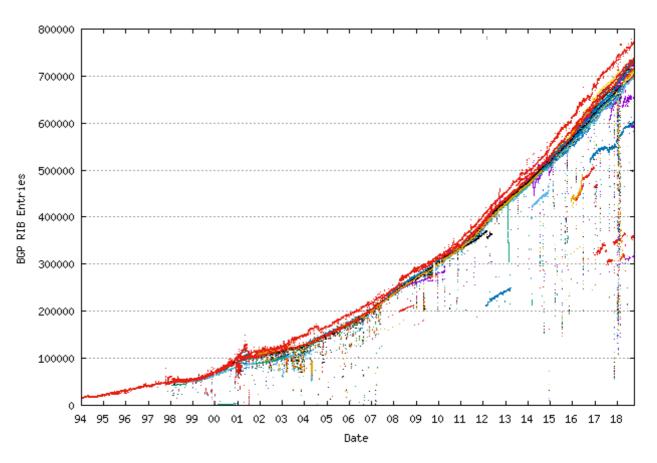
Route collector saves all BGP messages received from peers

- \* Route Announcements
- \* Route Withdrawals

# Statistics from a RouteView collector as of September 2018

# September '18: ~750K IPv4 prefixes originated from ~62K Autonomous Systems

"the global routing table"



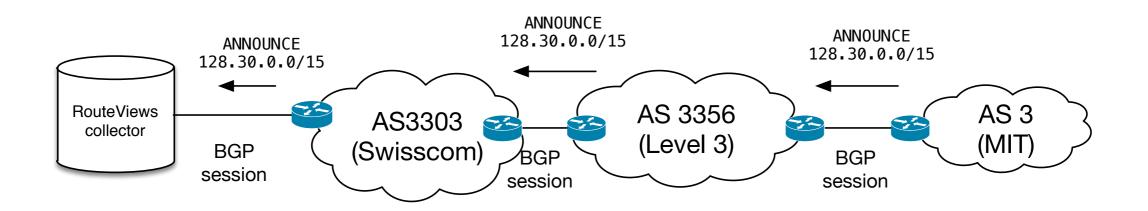
daily updated statistics: <u>http://bgp.potaroo.net/</u> live queries: <u>https://stat.ripe.net/widget/routing-status</u>

## prefix AS path

TABLE\_DUMP2|1536508822|B|217.192.89.50|3303|128.30.0.0/15|3303 3356 3|IGP| [...]

### prefix AS path

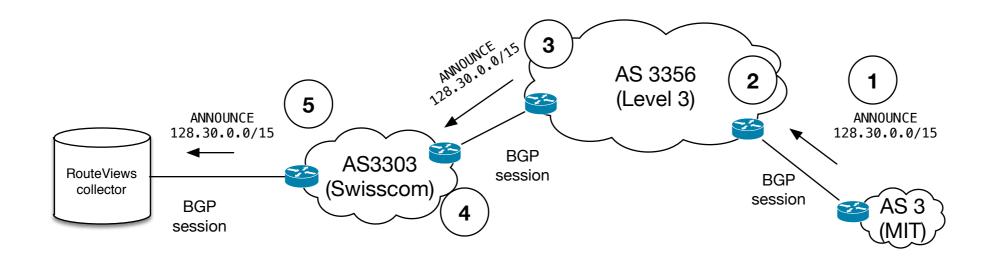
TABLE\_DUMP2|1536508822|B|217.192.89.50|3303|128.30.0.0/15|3303 3356 3|IGP| [...]



From this line, we derive:

-> AS3 is the origin of 128.30.0.0/15

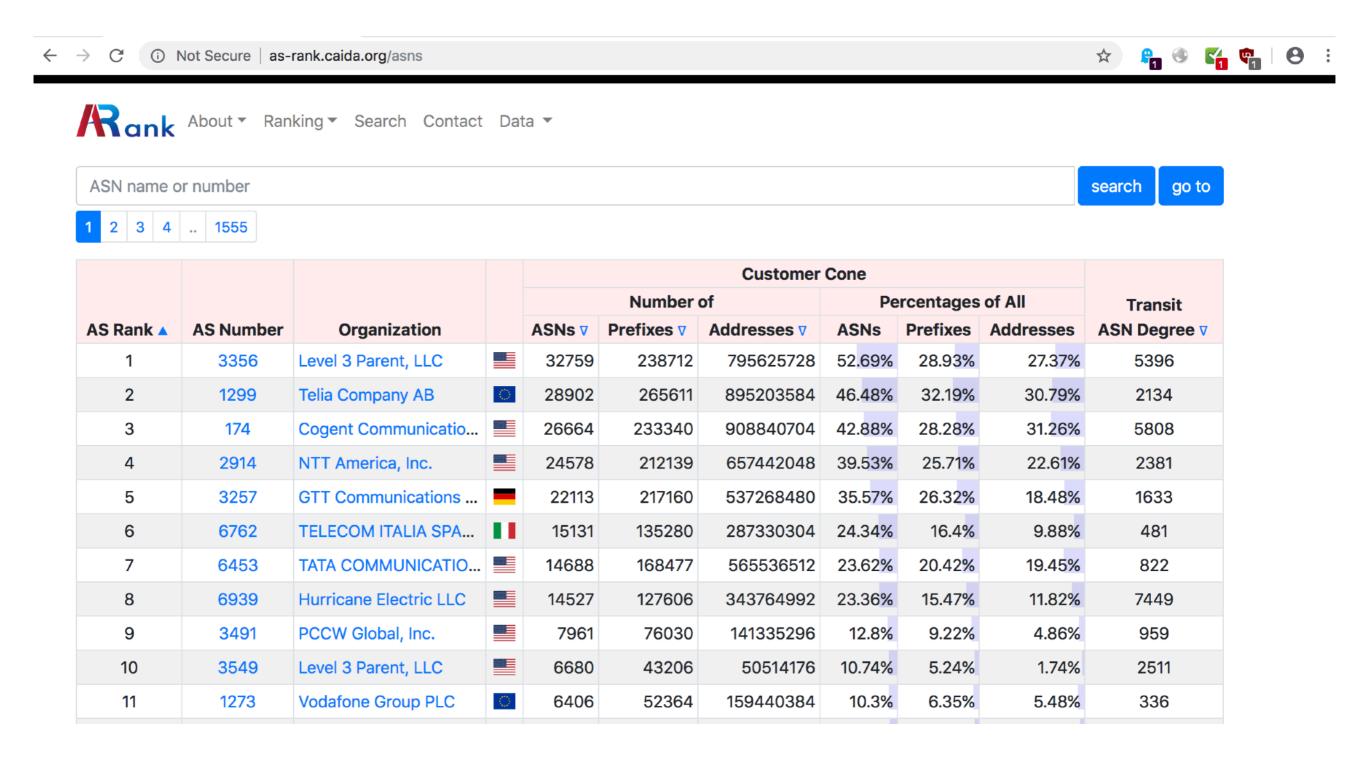
-> BGP peerings between: AS3303 <> AS3356 and AS3356 <> AS3



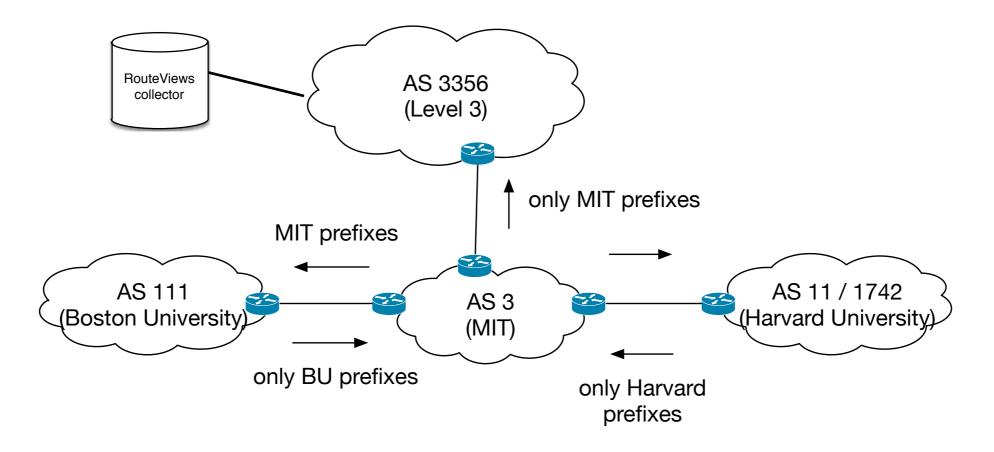
- 1. MIT announces its prefix to its **upstream**, *Level 3*
- 2. Level 3 decides to accept the announcement cause MIT is a customer.
- 3. Level 3 decides to propagate MIT prefix to its customers and peers
- 4. Swisscom receives Level 3 announcement and chooses it as best path
- 5. Swisscom propages to route collector.

# The AS path we see is the result of **policy routing**.

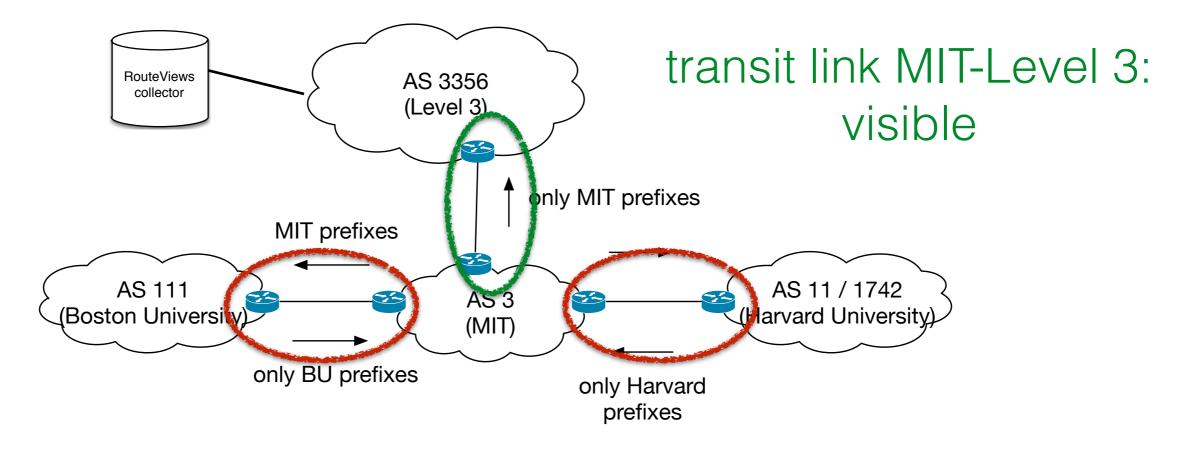
# **AS-Topology use case: Rank ISPs by Customer Cone**



further reading: Luckie et al. "AS Relationships, Customer Cones, and Validation" ACM IMC 2013.



\* this topology is made up, MIT and BU/Harvard to not peer directly, but via AS10578



\* this topology is made up, MIT and BU/Harvard to not peer directly, but via AS10578

peering links BU-MIT and MIT-Harvard invisible

# AS relationships derived from BGP data are (heavily?) biased towards Customer-Provider links.

Year/Methodology	Est. # of customer- provider links in the Internet	Est. number of peering links in the Internet
2008 (BGP)*	~60,000	~15,000

\* Dhamdhere et al., , ACM IMC 2008, IEEE/ACM Trans on Networking 2011
 \*\* Augustin et al., ACM IMC 2009
 \*\* K. Chen et al., ACM CoNEXT 2009
 \*\*\* Ager et al., SIGCOMM 2012

slide adapted from W. Willinger, "There is more to Internet measurement than meets the eye" @ KTH Stockholm

Year/Methodology	Est. # of customer- provider links in the Internet	Est. number of peering links in the Internet
2008 (BGP)*	~60,000	~15,000
2010 (BGP + traceroute)**	~90,000	~30,000

\* Dhamdhere et al., , ACM IMC 2008, IEEE/ACM Trans on Networking 2011
 \*\* Augustin et al., ACM IMC 2009
 \*\* K. Chen et al., ACM CoNEXT 2009
 \*\*\* Ager et al., SIGCOMM 2012

slide adapted from W. Willinger, "There is more to Internet measurement than meets the eye" @ KTH Stockholm

Year/Methodology	Est. # of customer- provider links in the Internet	Est. number of peering links in the Internet
2008 (BGP)*	~60,000	~15,000
2010 (BGP + traceroute)**	~90,000	~30,000
2012 (ground truth from a large IXP)***	~90,000	>200,000

Dhamdhere et al., , ACM IMC 2008, IEEE/ACM Trans on Networking 2011 \*\* Augustin et al., ACM IMC 2009 \*\* K. Chen et al., ACM CoNEXT 2009 \*

\*\*\* Ager et al., SIGCOMM 2012

slide adapted from W. Willinger, "There is more to Internet measurement than meets the eye" @ KTH Stockholm

Year/Methodology	Est. # of customer- provider links in the Internet	Est. number of peering links in the Internet
2008 (BGP)*	~60,000	~15,000
2010 (BGP + traceroute)**	~90,000	~30,000
2012 (ground truth from a large IXP)***	~90,000	>200,000

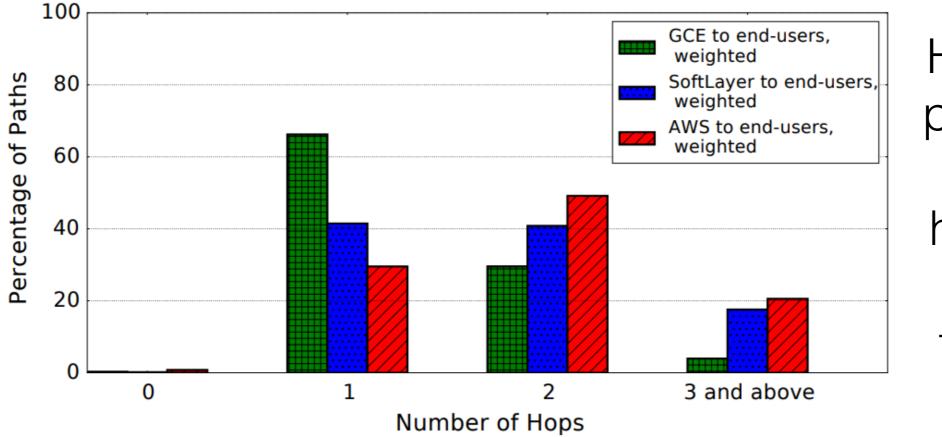
Topology much "flatter" than visible in BGP.

# Peering Links vs. Transit Links: Traffic?

majority of peering links, but majority of traffic still on transit?

# **Peering Links vs. Transit Links: Traffic?**

# majority of peering links, but majority of traffic still on transit?



Hypergiants peer directly with ASes home to the majority of their users.

Figure 4: Paths lengths from different cloud platforms to end-users.

Chiu et al., "Are We One Hop Away from a Better Internet?" ACM IMC 2015.

# **AS-level topology measurements: Recap**

- BGP data from RouteViews extremely useful
  - Studying Customer-Provider structure & economics
  - Studying BGP routing and routing anomalies
- But was never meant to be used for topology inference
- Hides most of peering links -> hides local connectivity
- Can easily lead to wrong conclusions
- "Know your data"

Is the data "fit" to answer your specific question?

# **Topology measurements: Active**

## Traceroute, introduced 1988 by Van Jacobson

Tue Dec 27 06:24:24 PST 1988

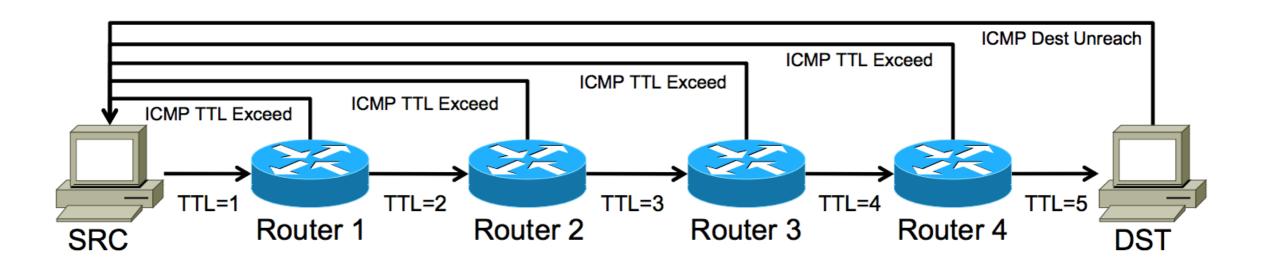
Traceroute is a system administrators utility to trace the route ip packets from the current system take in getting to some destination system. See the comments at the front of the program for a description of its use.

# (from traceroute.c, 1988)

```
* A more interesting example is:
*
      [yak 72]% traceroute allspice.lcs.mit.edu.
*
     traceroute to allspice.lcs.mit.edu (18.26.0.115), 30 hops max
*
         helios.ee.lbl.gov (128.3.112.1) 0 ms 0 ms 0 ms
*
       1
         lilac-dmc.Berkeley.EDU (128.32.216.1) 19 ms 19 ms
*
      2
                                                              19 ms
         lilac-dmc.Berkeley.EDU (128.32.216.1) 39 ms 19 ms
      3
                                                              19 ms
*
         ccngw-ner-cc.Berkeley.EDU (128.32.136.23) 19 ms 39 ms 39 ms
*
       4
         ccn-nerif22.Berkeley.EDU (128.32.168.22) 20 ms 39 ms 39 ms
*
      5
         128.32.197.4 (128.32.197.4) 59 ms 119 ms 39 ms
*
      6
         131.119.2.5 (131.119.2.5) 59 ms 59 ms
                                                 39 ms
*
      7
         129.140.70.13 (129.140.70.13) 80 ms 79 ms 99 ms
*
      8
         129.140.71.6 (129.140.71.6) 139 ms 139 ms 159 ms
      9
*
      10 129.140.81.7 (129.140.81.7) 199 ms 180 ms 300 ms
*
         129.140.72.17 (129.140.72.17) 300 ms 239 ms 239 ms
     11
*
         * * *
     12
*
         128.121.54.72 (128.121.54.72) 259 ms 499 ms 279 ms
     13
*
     14
         * * *
*
     15
         * * *
*
     16 * * *
*
     17
         * * *
*
      18
         ALLSPICE.LCS.MIT.EDU (18.26.0.115) 339 ms 279 ms 279 ms
*
*
* (I start to see why I'm having so much trouble with mail to
* MIT.)
```

# Traceroute

- 1. Launch a probe packet towards DST, with a TTL of 1
- 2. Every router hop decrements the IP TTL of the packet by 1
- 3. When the TTL hits 0, packet is dropped, router sends *ICMP TTL Exceeded* packet to SRC
- 4. SRC receives this ICMP message, displays as trace route "hop"
- 5. Repeat from step 1, with TTL incremented by 1 each time, until..
- 6. DST hop receives probe, returns ICMP Dest Unreachable
- 7. SRC stops the trace route upon receipt of ICMP Dest Unreachable



slide adapted from Richard Steenbergen, "A Practical Guide to (Correctly) Troubleshooting with Traceroute", NANOG 47

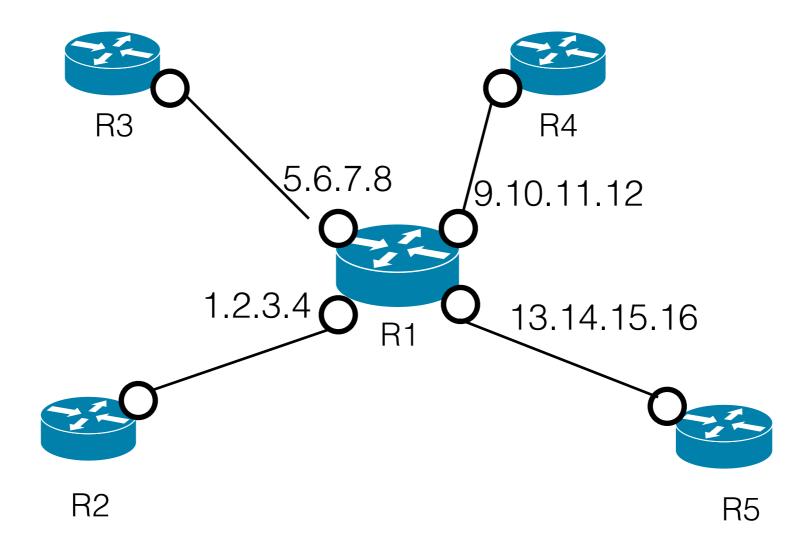
# **Traceroute Anomalies**

- Missing Hops
- Missing Destination
- Load Balancing
- No visibility into return path (asymmetric routing)
- Shows IP addresses = router aliases != routers

further reading on traceroute anomalies (not covered here):

Augustin et al., "Avoiding traceroute anomalies with Paris traceroute" ACM IMC 2006 Mao et al., "Towards an accurate AS-level traceroute tool" ACM SIGCOMM 2003 Luckie et al., "bdrmap: Inference of Borders Between IP Networks", ACM IMC 2016 Katz-Bassett et al., "Reverse Traceroute", NSDI 2010

#### **IP Address != Interface != Router**



traceroute via R2,R1,R4: R1 likely to show up with 1.2.3.4 traceroute via R5,R1,R4: R1 likely to show up with 13.14.15.16

routers typically (not always!) reply with the IP address of the **inbound** interface. (this violates RFC1812, but is common behavior).

Amini et al., "Issues with Inferring Internet Topological Attributes" Mao et al., "Towards an Accurate AS-Level Traceroute Tool"

# **Router Alias Resolution Example: Direct Probing**

```
Berkeley to MIT:
1 router1-vlan1.ICSI.Berkeley.EDU (192.150.186.1)
2 router12-ge0-0-0.ICSI.Berkeley.EDU (192.150.187.254)
3 ge-0-2-0.inr-667-sut.Berkeley.EDU (169.229.0.140)
...
MIT to Berkeley:
...
24 sut-mdc-ar1--xe-0-1-0.net.berkeley.edu (128.32.0.17)
25 router12-ge0-0-1.icsi.berkeley.edu (169.229.0.141)
26 router1-vlan5.icsi.berkeley.edu (192.150.187.249)
...
```

#### same router? send UDP probe to random high port:

<ul> <li>Wi-Fi: en0</li> <li>Wi-Fi: en0</li> <li>We send packets to each alia</li> <li>(different IP addresses)</li> </ul>										
No.	Time	Source	Destination	tocol Length	Info					
69	1537742992.833704	192.168.0.102	192.150.187.249 UD	P 47	64242→55022	Len=5				
71	1537742992.973948	192.150.187.249	192.168.0.102	MP 70	Destination	unreachable (Po	ort unreachable)			
103	1537743000.790409	192.168.0.102	192.150.186.1 UD	P 47	62171→55022	Len=5				
104	1537743000.884043	192.150.187.249	192.168.0.102 IC	MP 70	Destination	unreachable (Po	ort unreachable)			

router replies with one single IP address

# **Alias Resolution Example: Increasing IPID Field**

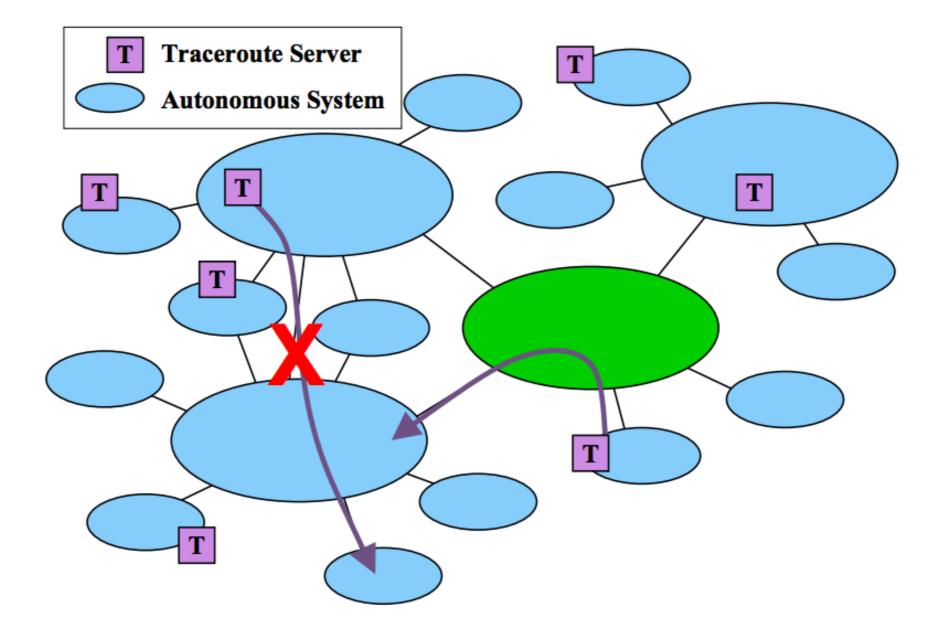
IP header has the IPID field. Original purpose: re-assemble fragmented IP packets.

#### Often implemented as counter:

•						Wi-Fi: en0				
		<u>a</u> 💿 🖿 🛅	🔀 🙆 🔍	🔶 🏓 🖉 🖣	<u>↓</u>	<b>.</b> ⊕ ⊖				
	ticmp or (udp and udp.port == 55022) Expression +									
No.		Time	Source	Destination	Protocol	Length Info		IPID		
Г	28	1537998445.314635	192.168.0.100	192.150.186.1	UDP	49 50874→55022 L	_en=7	0x=0ca (57546)		
L	29	1537998445.405180	192.150.187.249	192.168.0.100	ICMP	70 Destination u	<pre>inreachable (Port unreachable)</pre>	x7962 (31074) xe0ca (575		
	39	1537998449.187481	192.168.0.100	192.150.187.249	UDP	49 56132→55022 L	en=7	0x574b (22347)		
	40	1537998449.285354	192.150.187.249	192.168.0.100	ICMP	70 Destination u	nreachable (Port unreachable)	0x7967 (31079),0x574b (223…		
	42	1537998450.148436	192.168.0.100	192.150.186.1	UDP	49 52910→55022 L	en=7	0xeaf9 (60153)		
	43	1537998450.243982	192.150.187.249	192.168.0.100	ICMP	70 Destination u	ınreachable (Port unreachable)	0x7968 (31080),0x af9 (601…		
	44	1537998450.887651	192.168.0.100	192.150.187.249	UDP	49 53618→55022 L	en=7	0xbbaa (48042)		
	45	1537998450.987256	192.150.187.249	192.168.0.100	ICMP	70 Destination u	ınreachable (Port unreachable)	0x796b (31083),0x baa (480…		
	46	1537998451.564593	192.168.0.100	192.150.186.1	UDP	49 61181→55022 L	en=7	0xf1d9 (61913)		
	47	1537998451.657689	192.150.187.249	192.168.0.100	ICMP	70 Destination u	nreachable (Port unreachable)	0x796c (31084),0 f1d9 (619		
	50	1537998452.210720	192.168.0.100	192.150.187.249	UDP	49 60834→55022 L	en=7	🖗 xca1a (51738)		
	51	1537998452.315721	192.150.187.249	192.168.0.100	ICMP	70 Destination u	<pre>inreachable (Port unreachable)</pre>	0x796e (31086)/0xca1a (517…		

#### IPID field of ICMP replies of the router form a sequence

## **Traceroute for ISP Topology Inference**



### Traceroutes show single paths. How to effectively select target IP addresses?

Spring et al. "Measuring ISP Network Topologies with Rocketfuel", SIGCOMM 2002 slides

# Path Reductions

Want to choose unique paths – with new information.

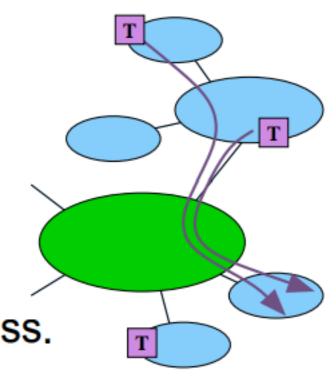
Skip repeated traces of the same path.

Expect the common case:

- Traceroute server has one ingress point
- Customer prefix has one egress point
- BGP peers have one early-exit per ingress.

If we're wrong, we might miss some paths.

New servers add paths or share load!



# **Reduction Effectiveness**

• Brute force:

All servers to all BGP prefixes, disaggregate ISP prefixes. 90-150 million traceroutes required

- BGP directed probes: All traceroutes identifiable from RouteViews.
   0.2-15 million traceroutes required
- Executed after path reduction: Traceroutes chosen by Rocketfuel.
   8-300 thousand traceroutes required

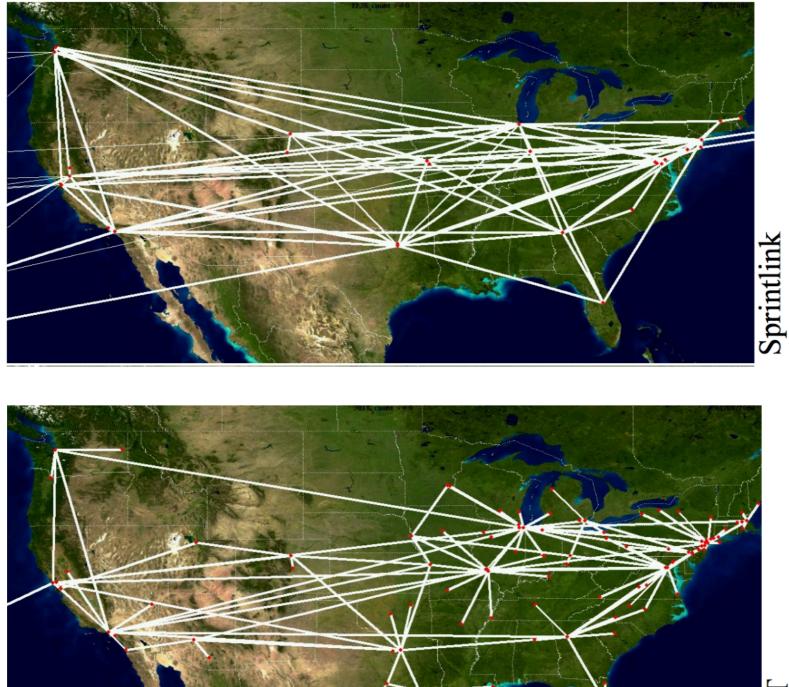
Directed probing and path reductions are effective at reducing the number of probes required to map an ISP

# **Traceroute for Large-Scale Topology Inference**

- Need sufficient number of vantage points
- Need a smart way to select target IPs
  - Brute-Force probing the whole space ineffective
- Need to deal with traceroute issues

Rocketfuel combines all these aspects together, leveraging BGP data to select target ranges, into a single system.

#### ISP Topologies inferred by Rocketfuel (back in 2002...)



AT&T

# Internet-Wide Scanning

# Scanning the entire IPv4 address space

entire IPv4 Space: 2 \*\* 32 addresses = 4.3B addresses

routable IPv4 space (excluding reserved ranges, multicast etc): ~3.7B addresses

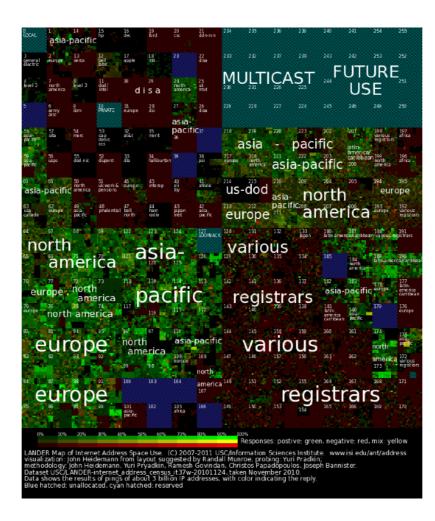
publicly routed IPv4 space: ~2.9B addresses (as of late 2017)

can we just scan (probe) every single routed IPv4 address?

further reading: Richter et al. "A Primer on IPv4 Scarcity" ACM CCR 2015

### Scanning the entire IPv4 address space

• First full scans of the IPv4 space took weeks to months



Heidemann et al., "Census and Survey of the Visible Internet" ACM IMC 2008

# **ZMap - Stateless Implementation**

Default case: We open a TCP socket, send a SYN packet wait for the destination to reply (or not to reply)

ZMap: Bypass the TCP/IP stack of the OS craft Ethernet frames directly, "fill up the pipe"

Encode destination IP address into probe packets, match responses on arrival.

TCP SRC port	TCP sequence number	
		-
TCP DST port	TCP ACK = SEQ + 1	

Adrian et al. "Zippier ZMap: Internet-Wide Scanning at 10 Gbps" WOOT 2014.

#### ZMap example: Track Heartbleed Vulnerability

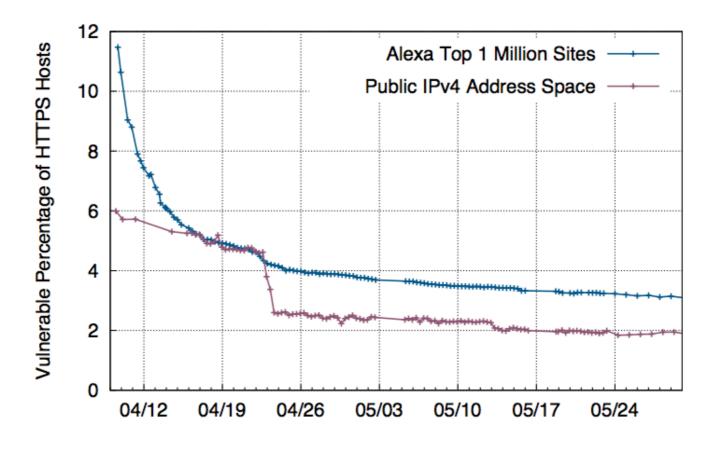




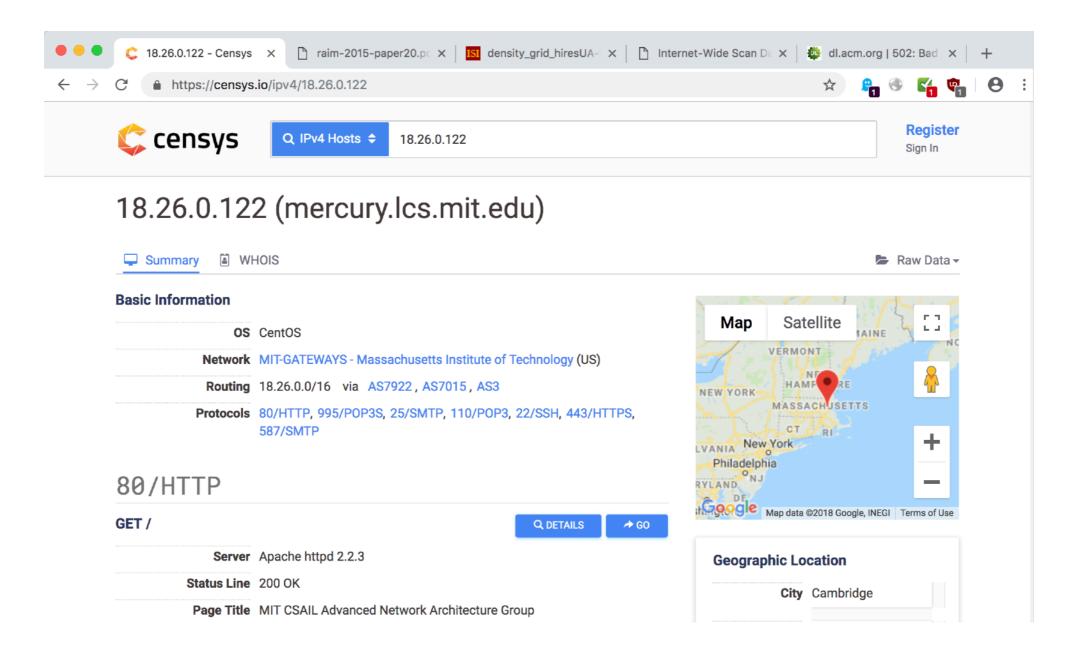
Figure 3: **HTTPS Patch Rate.** We track vulnerable web servers in the Alexa Top 1 Million and the public IPv4 address space. We track the latter by scanning independent 1% samples of the public IPv4 address space every 8 hours. Between April 9 and June 4, the vulnerable population of the Alexa Top 1 Million shrank from 11.5% to 3.1%, and for all HTTPS hosts from 6.0% to 1.9%.

Durumeric et al. "The Matter of Heartbleed" IMC 2014.

# ZMap Data Availability: <u>scans.io</u>

					x 🖁 🖉 😭
nternet-Wide Scan Data Repo	sitory				
110-pop3-starttls-full_ipv4	110	рорЗ	starttls	full ipv4	2018-09-23 00:50:46
143-imap-starttls-full_ipv4	143	imap	starttls	full ipv4	2018-09-23 23:18:48
1900-upnp-discovery-full_ipv4	1900	upnp	discovery	full ipv4	2018-09-24 02:36:51
1911-fox-device_id-full_ipv4	1911	fox	device id	full ipv4	2018-09-24 12:18:22
20000-dnp3-status-full_ipv4	20000	dnp3	status	full ipv4	2018-09-22 12:48:09
21-ftp-banner-full_ipv4	21	ftp	banner	full ipv4	2018-09-24 23:06:17
22-ssh-v2-full_ipv4	22	ssh	v2	full ipv4	2018-09-19 00:50:30
23-telnet-banner-full_ipv4	23	telnet	banner	full ipv4	2018-09-19 00:36:10
2323-telnet-banner-full_ipv4	2323	telnet	banner	full ipv4	2018-09-19 23:05:48
25-smtp-starttls-alexa_top1mil	25	smtp	starttls	alexa top1mil	2018-09-24 12:38:16
25-smtp-starttls-full_ipv4	25	smtp	starttls	full ipv4	2018-09-23 00:47:42
443-https-dhe-alexa_top1mil	443	https	dhe	alexa top1mil	2018-09-24 12:38:08
443-https-dhe-full_ipv4	443	https	dhe	full ipv4	2018-09-23 23:51:00
443-https-dhe_export-alexa_top1mil	443	https	dhe export	alexa top1mil	2018-09-24 11:09:59
443-https-dhe_export-full_ipv4	443	https	dhe export	full ipv4	2018-09-20 23:09:01
443-https-heartbleed-alexa_top1mil	443	https	heartbleed	alexa top1mil	2018-09-23 14:18:44

#### ZMap-driven search engine: censys.io



Durumeric et al. "The Matter of Heartbleed" IMC 2014.

#### **Interested in Internet Measurement Projects?**

richterp@csail.mit.edu