

Desperately Seeking Default!

Geoff Huston
APNIC

In the Telephone Network

- Anyone could dial anyone else
- All telephones were equally reachable

anyone could d

l telephones v

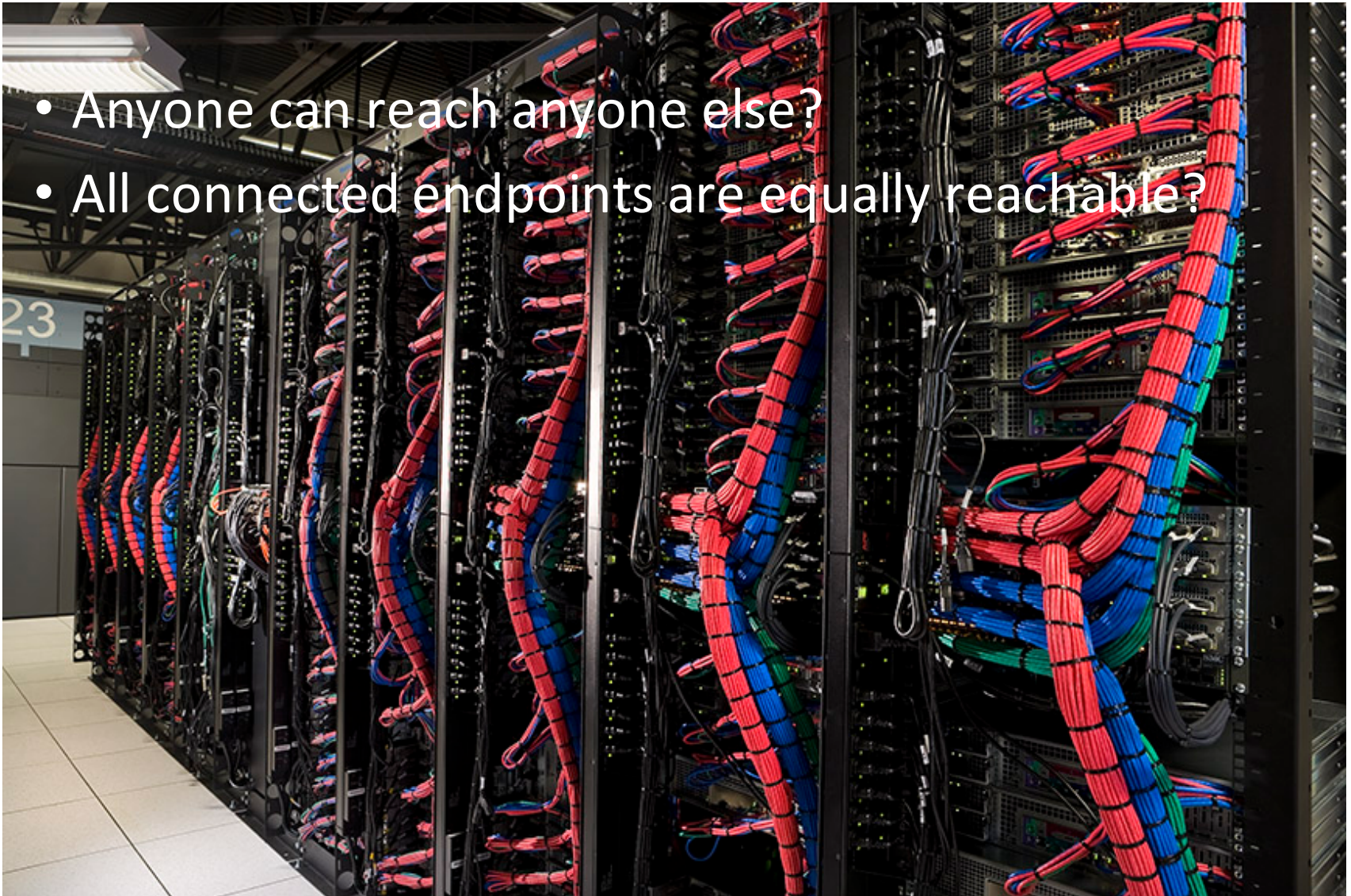
JAMES 1725 Bridgeway Dr	415 729-0095
HARRIMAN Barbara	415 339-8933
Jeff	415 488-0200
Jeff	415 488-0322
Jeff	415 488-9862
Nancy 121 Redwood Dr Wdacr	415 488-1218
S 111 Albert Av Sau	415 331-1857
Thomas R	415 488-9862
GT 10000	415 488-9862
Don 52 Corte Morada Kntld	415 888-9304
Don 52 Corte Morada Kntld	415 461-1310
Don 52 Corte Morada Kntld	415 461-1310
Don 52 Corte Morada Kntld	415 925-9045
Frank 28 Cambridge Dr S E	415 487-2141
Jane	415 488-2201
Jeff 140 Lagunitas Rd Laguntas	415 488-1271
Karen E	415 485-1771
Kenneth	415 269-2647
Laurel 7	
Laurie	
Margar	
Mark --	
Mary 42	
Robert	
Scott --	
Scott 27	
Stephen	
T 9	
Timoth	
Uta --	
Harrington	
4415 Para	
HARRIS B	
Alan &	
Andrew	
Anne 10	
Anne 10	
Anne 10	
Ariene 1	
B --	
Harris B	
HARRIS B	
Barbra	
Barry --	
Bernard	
Bernice	
Bourke	
Brent &	
C --	
C --	
C --	
C & B 3	
Carol J	
Charles	
Charles	
Charles	
Christin	
Cynthia	
D --	
D & G --	
Damas	
Daniel	
David 20	
David 8	
David 1	
David 28	
David G	
Don 213	

HARRIS TANYA U DDS	
Tiburon Bl & North Knoll Rd M Wly	415 388-7466
HARRIS Ted & Ada Beth 305 San Rafael Ave Bely	415 435-3228
Victor & Elizabeth	415 454-8613
Harris Victor Law Offices Of 1050 Northgate Dr S R	415 479-8000
HARRIS Wyman C 306 Bella Vista Av Bely	415 435-0245
Yvonne	415 256-9895
Yvonne	415 464-9445
HARRIS-KUNZ Gillian	
1115 Sir Francis Drake Bl Knitfld	415 455-9290
HARRISON A	415 454-4872
Anna	415 891-8931
Anna 135 Barbara Wy Tibirn	415 383-6948
Anthony Sr 105 S S R	415 457-1799
Anthony Sr 105 S S R	415 861-8300
Anthony Sr 105 S S R	415 384-0271
Anthony Sr 105 S S R	415 453-6334
Cory T 2 Crescent Rd C M	415 758-7027
David	415 456-3098
David	415 457-8771



internet In the ~~Telephone~~ Network

- Anyone can reach anyone else?
- All connected endpoints are equally reachable?



internet In the ~~Telephone Network~~

- Anyone can reach anyone else?
- ~~All connected endpoints are equally reachable?~~

NATs changed all that - so our current expectations are that if you stand up a public service on port 80 (or 443 for that matter) then everyone can reach you

internet In the ~~Telephone Network~~

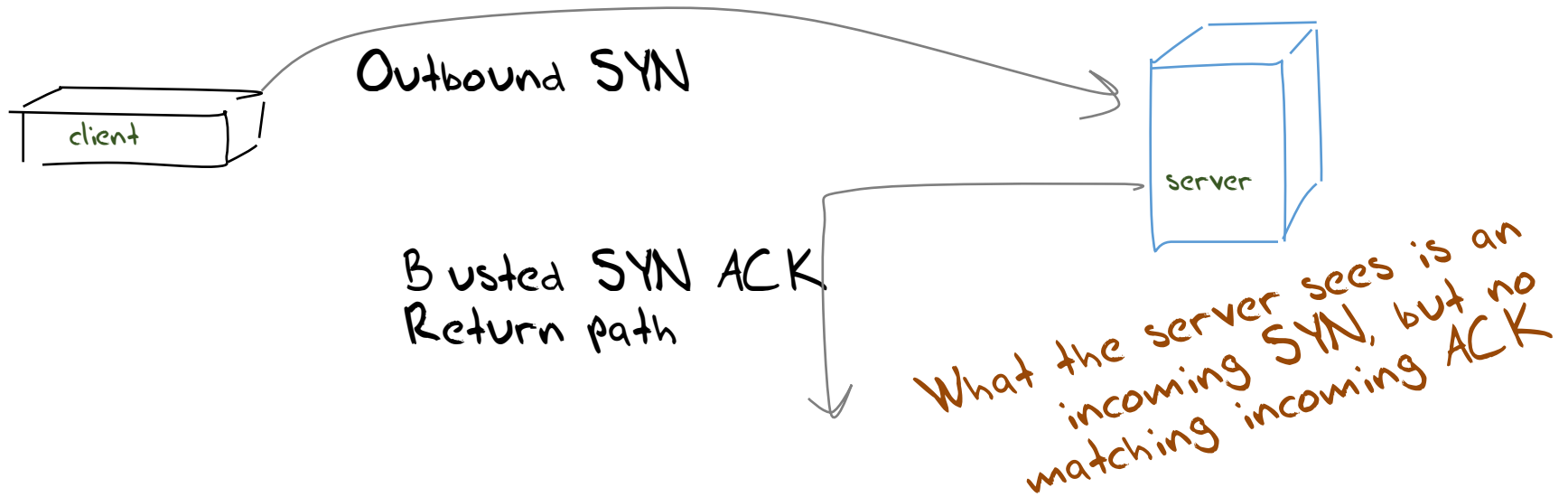
- Anyone can reach anyone else
- ~~All connected endpoints are~~ ~~shable~~

NATs changed all expectations.
public IP address
ma We might THINK this, but is it true
ALL the time?
... it 80 (or 443 for that
... you stand up a
... everyone can reach you

What do we see?

is everyone really connected to everyone else?

Connection Failure



Daily IPv6 Failures

RIP
Flash!



HTML5 + TLS +
Mobile Devices

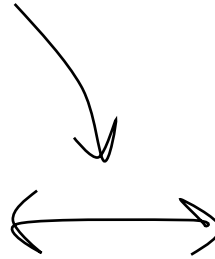
Daily IPv6 Failures

IPv6 Failures

- 1.8% failure for non-6to4 V6 is unacceptable!
- Why is this happening
 - Auto-tunnelling?
 - Lousy CPE firmware?
 - Strange firewall filters?
- But is all of this due to local configuration / equipment? What is the comparable view in IPv4?

IPv4 Connection Failure

Missing PCAP data



IPv4 Failures

- IPv4 failures are around 1 in 500
- And we are pretty sure its NOT:
 - Auto-tunnelling
 - Lousy CPE firmware
 - Strange firewall filters
- So what is the reason for this residual asymmetric failure rate?
- Is it asymmetric routing connectivity?

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

Route Views Routing Table

20 Years of Routing the Internet

This is a view pulled together from each of the routing peers of Route Views

2015/16, Route Views + RIS

This is a view pulled together from each of the routing peers of Route Views and RiS



Different peers see a slightly different Internet

- But is this just traffic engineering more specifics?
- Or do different peers see a different set of reachable addresses in the routing table?

Address Span
(Route Views + RIS data sets)

Address Span (Route Views + RIS data sets)

15 M Addresses



What does this mean?

- Each peer of RouteViews and RIS announces a span of addresses that appears to be a unique span.
- In total, these spans agree to within 15M addresses, but this means that there are at least 15M uniquely addressed endpoints that cannot be reached from all other endpoints.
- This variation is stable, so its not transient routing that is generating this – the reasons for this difference are structural

What about IPv6?

The Route Views view of IPv6

World IPv6 Day

IANA IPv4 Exhaustion



Number of IPv6 Routes in 2015/16



IPv6 Announced Address Span Variation (RV + RIS)

What is "default"?

- There is no definition of the route set that is "default"
- At best it's an informal quorum
 - So let's define this quorum by arbitrarily setting the quorum threshold at $2/3$
 - i.e. if $2/3$ of the peers of a route collector advertise a route then it is part of the default quorum.
- Individual peer networks will contain route sets that differ from this quorum by having both additional prefixes and holes.
 - Let's look at the variance from the quorum

A "Quorum" deviation view of IPv4

A magnified view

IPv6 "Quorum" Deviation

Zooming In

And Again

And Again!



Quorum Deviation for RVA IPv4 Peers (25th April 2016)

AS	RIB	SPAN	MISSING	EXTRA	NAME
AS7660	567,098	162.00	92,521,424	1,581,512	APAN-JP Asia Pacific Advanced Network - Japan, JP
AS1299	575,011	164.50	49,041,408	94,464	TELIANET , SE
AS31019	455,474	166.09	25,482,432	3,108,816	MEANIE Paulus M. Hoogsteder trading as Meanie, NL
AS23673	607,435	167.79	8,280,064	14,452,736	ONLINE-AS Cogetel Online, Cambodia, ISP, KH
AS3549	579,281	167.12	5,098,752	72,064	LVL3-3549 - Level 3 Communications, Inc., US
AS3356	577,651	167.12	5,031,936	1,536	LEVEL3 - Level 3 Communications, Inc., US
AS2497	587,831	167.44	4,726,016	4,982,272	IIJ Internet Initiative Japan Inc., JP
AS34224	582,228	167.36	4,071,936	3,120,896	NETERRA-AS Neterra Ltd., BG
AS1239	578,472	167.26	2,734,080	0	SPRINTLINK - Sprint, US
AS293	597,785	167.67	1,988,864	6,097,502	ESNET - ESnet, US
AS37100	585,091	167.50	1,966,848	3,241,984	SEACOM-AS, MU
AS20912	587,189	167.49	1,954,046	3,075,385	ASN-PANSERVICE Panservice, IT
AS6939	596,059	167.62	1,952,512	5,359,360	HURRICANE - Hurricane Electric, Inc., US
AS11686	584,696	167.39	1,651,456	1,086,464	ENA - Education Networks of America, US
AS3741	581,388	167.35	1,425,408	248,320	IS, ZA
AS47872	607,976	167.71	1,421,736	6,310,048	SOFIA-CONNECT-AS SOFIA CONNECT EOOD, BG
AS20771	606,620	167.71	1,418,752	6,313,728	CAUCASUS-CABLE-SYSTEM Caucasus Online Ltd., GE
AS58511	601,973	167.67	1,376,768	5,586,309	CONNECTIVITYIT-AU Connectivity IT Pty Ltd, AU
AS3130	577,821	167.34	1,338,880	5,120	RG-BIWA - RGnet, LLC, US
AS13030	579,067	167.54	1,129,216	3,130,881	INIT7 Init7 (Switzerland) Ltd., CH
AS8492	591,539	167.59	1,119,488	4,010,240	OBIT-AS OBIT Ltd., RU
AS3277	592,537	167.60	1,067,520	4,031,088	RUSNET-AS NPO RUSnet, RU
AS200130	581,449	167.55	650,240	2,863,616	DIGITALOCEAN-ASN-1 Digital Ocean, Inc., EU
AS202018	581,334	167.55	650,240	2,862,080	DIGITALOCEAN-ASN-3 Digital Ocean, Inc., NL
AS852	581,196	167.40	589,568	280,320	ASN852 - TELUS Communications Inc., CA
AS3303	581,407	167.41	560,640	323,352	SWISSCOM Swisscom (Switzerland) Ltd, CH
AS2152	581,886	167.46	444,416	1,070,336	CSUNET-NW - California State University, Office of the Chancellor, US
AS53364	580,271	167.41	316,416	93,440	AS-PRE2POST-2 - ZEROFAIL, US
AS3257	580,274	167.41	315,648	93,440	GTT-BACKBONE Tinet Spa, DE
AS2914	579,671	167.40	308,736	2,560	NTT-COMMUNICATIONS-2914 - NTT America, Inc., US
AS7018	579,063	167.41	293,888	54,272	ATT-INTERNET4 - AT&T Services, Inc., US
AS6762	581,338	167.41	249,856	7,936	SEABONE-NET TELECOM ITALIA SPARKLE S.p.A., IT
AS22652	584,493	167.43	203,264	397,312	FIBRENOIRE-INTERNET - Fibrenoire Inc., CA
AS1221	582,020	167.42	199,936	217,984	ASN-TELSTRA Telstra Pty Ltd, AU
AS286	580,287	167.45	93,952	521,104	KPN KPN B.V., NL
AS40191	581,427	167.43	67,007	201,732	AS-PRE2POST-1 - ZEROFAIL, CA
AS3561	579,635	167.42	53,248	2,816	CENTURYLINK-LEGACY-SAVVIS - Savvis, US
AS1668	579,942	167.43	37,376	118,528	AOL-ATDN - AOL Transit Data Network, US

Quorum Deviation for RIS IPv4 Peers (25th April 2016)

AS	RIB	SPAN	MISSING	EXTRA	NAME
AS37989	239,180	165.16	38,465,536	442,368	DCS1-02-AS-AP DCS1 Pte Ltd, at DC02, SG
AS3549	542,285	167.11	5,340,416	100,480	LVLT-3549 - Level 3 Communications, Inc., US
AS4777	553,629	167.48	5,263,360	6,312,004	APNIC-NSPIX2-AS Asia Pacific Network Information Centre, JP
AS50300	545,929	167.33	4,324,864	2,865,664	CUSTDC CustodianDC Limited, GB
AS50763	357,933	167.44	2,508,032	2,746,880	MCKAYCOM MCKAYCOM LTD, GB
AS2914	543,207	167.33	1,572,608	64,768	NTT-COMMUNICATIONS-2914 - NTT America, Inc., US
AS4608	557,106	167.67	1,251,840	5,418,368	APNIC-SERVICES Asia Pacific Network Information Centre, AU
AS30132	555,823	167.67	1,204,736	5,389,440	ISC-F-AS - Internet Systems Consortium, Inc., US
AS8758	547,071	167.56	1,023,744	3,357,880	IWAY Iway AG, CH
AS1103	552,261	167.6	1,000,192	3,990,793	SURFNET-NL SURFnet, The Netherlands, NL
AS13030	541,779	167.54	990,720	2,991,360	INIT7 Init7 (Switzerland) Ltd., CH
AS1836	546,526	167.56	888,832	3,171,328	GREEN green.ch AG Autonomous System, CH
AS8455	544,289	167.62	856,576	4,184,320	ATOM86-AS Schuberg Philis B.V. trading as atom86, NL
AS8283	548,631	167.63	791,040	4,368,936	COLOCLUE-AS Netwerkvereniging Coloclue, Amsterdam, Netherlands, NL
AS50304	546,357	167.56	698,880	2,999,992	BLIX Blix Solutions AS, NO
AS57381	546,362	167.56	698,112	2,999,736	FNUTT PE Fnutt Consulting Daniel Husand, NO
AS56730	547,896	167.56	675,840	2,947,584	WIREHIVE-AS Wirehive Limited, GB
AS12859	544,975	167.56	631,040	2,875,136	NL-BIT BIT BV, NL
AS8468	544,890	167.55	625,408	2,812,928	ENTANET ENTANET International Limited, GB
AS15435	544,370	167.56	617,472	2,878,272	KABELFOON CAIW Diensten B.V., NL
AS29611	544,190	167.52	606,208	2,187,776	ELITE-AS Elite Limited, GB
AS29636	544,589	167.53	567,040	2,363,136	CATALYST2-AS catalyst2 Services Limited, IE
AS25160	545,579	167.52	538,624	2,190,336	VORBOSS_AS Vorboss Limited, GB
AS29608	544,205	167.4	422,656	131,840	WAN2MANY-AS Absolight, FR
AS6453	542,804	167.4	416,256	68,608	AS6453 - TATA COMMUNICATIONS (AMERICA) INC, US
AS7018	542,580	167.4	412,416	115,456	ATT-INTERNET4 - AT&T Services, Inc., US
AS3257	543,349	167.41	375,296	145,920	GTT-BACKBONE Tinet Spa, DE

Quorum Deviation: IPv6, RVA

AS	RIB	SPAN (/32s)	MISSING (/48)	EXTRA (/48)	NAME
AS30071	25,172	73,510	985,303,395	196,610	OCCAID - TowardEX Technologies International, Inc., US
AS33437	27,843	88,591	18,522,729	21,799,422	HOTNIC - Hotnic LLC, US
AS6939	27,826	88,763	18,522,729	21,799,422	HURRICANE - Hurricane Electric, Inc., US
AS209	28,868	88,289	17,709,956	32,232,516	CENTURYLINK-US-LEGACY-QWEST - Qwest Communications Company, LLC, US
AS1239	27,439	88,433	16,542,293	0	SPRINTLINK - Sprint, US
AS701	27,620	88,446	7,138,817	65,536	UUNET - MCI Communications Services, Inc. d/b/a Verizon Business,
AS7018	27,809	88,449	6,491,217	229,387	ATT-INTERNET4 - AT&T Services, Inc., US
AS2914	27,781	88,450	6,095,925	65,537	NTT-COMMUNICATIONS-2914 - NTT America, Inc., US
AS53364	28,037	88,451	5,964,357	256	AS-PRE2POST-2 - ZEROFAIL, US
AS3257	28,038	88,469	5,898,820	256	GTT-BACKBONE Tinet Spa, DE
AS40191	28,625	88,539	4,849,970	131,329	AS-PRE2POST-1 - ZEROFAIL, CA
AS22652	28,979	89,126	2,097,156	1,966,356	FIBRENOIRE-INTERNET - Fibrenoire Inc., CA
AS2497	29,064	88,871	1,080,077	39,402,456	IJ Internet Initiative Japan Inc., JP
AS34224	29,965	88,603	851,984	22,455,041	NETERRA-AS Neterra Ltd., BG
AS37100	27,914	88,869	786,444	4,850,192	SEACOM-AS, MU
AS57463	28,569	88,873	327,694	21,799,422	NETIX NetIX Communications Ltd., BG
AS47872	29,152	88,889	327,692	22,062,852	SOFIA-CONNECT-AS SOFIA CONNECT EOOD, BG
AS3277	29,784	88,887	262,269	23,057,167	RUSNET-AS "NPO RUSnet", RU
AS31019	29,877	88,872	262,157	22,913,800	MEANIE Paulus M. Hoogsteder trading as Meanie, NL
AS393406	28,571	88,871	262,157	21,930,752	DIGITALOCEAN-ASN-NY3 - Digital Ocean, Inc., US
AS200130	28,568	88,871	262,157	21,864,961	DIGITALOCEAN-ASN-1 Digital Ocean, Inc., EU
AS202018	28,565	88,877	262,157	21,864,961	DIGITALOCEAN-ASN-3 Digital Ocean, Inc., NL
AS20912	28,683	88,874	262,157	22,259,368	ASN-PANSERVICE Panservice, IT
AS3741	28,866	88,871	262,157	22,061,825	IS, ZA
AS62567	28,570	88,551	262,157	21,864,961	DIGITALOCEAN-ASN-NY2 - Digital Ocean, Inc., US
AS13030	28,530	88,551	131,097	786,440	INIT7 Init7 (Switzerland) Ltd., CH

Quorum Deviation: IPv6, RIS

AS	RIB	SPAN (/32s)	MISSING (/48)	EXTRA (/48)	NAME
AS50300	13,173	22,789	4,312,799,213	3,899,410	CUSTDC CustodianDC Limited, GB
AS6939	27,015	88,596	18,309,740	22,127,121	HURRICANE - Hurricane Electric, Inc., US
AS6453	26,798	88,389	9,941,631	147,455	AS6453 - TATA COMMUNICATIONS (AMERICA) INC, US
AS25160	26,931	88,412	8,520,663	245,760	VORBOSS_AS Vorboss Limited, GB
AS7018	27,036	88,440	6,819,172	376,844	ATT-INTERNET4 - AT&T Services, Inc., US
AS2914	27,024	88,450	5,964,817	212,993	NTT-COMMUNICATIONS-2914 - NTT America, Inc., US
AS29608	27,463	88,477	5,440,109	1,462,540	WAN2MANY-AS Absolight, FR
AS29636	27,609	88,479	4,390,914	507,907	CATALYST2-AS catalyst2 Services Limited, IE
AS1273	27,120	88,506	4,259,864	2,130,090	CW Cable and Wireless Worldwide plc, GB
AS8758	27,294	88,886	826,754	23,662,663	IWAY Iway AG, CH
AS42708	27,092	88,559	720,910	2,081,050	PORTLANE Portlane AB, SE
AS8455	27,230	88,545	528,391	999,429	ATOM86-AS Schuberg Philis B.V. trading as atom86, NL
AS56730	28,120	88,891	245,773	23,405,075	WIREHIVE-AS Wirehive Limited, GB
AS50763	27,521	88,875	180,498	22,290,962	MCKAYCOM MCKAYCOM LTD, GB
AS12859	27,845	88,875	180,241	22,258,454	NL-BIT BIT BV, NL
AS8283	28,793	88,889	114,705	23,142,932	COLOCLUE-AS Netwerkvereniging Coloclue, Amsterdam, Netherlands, NL
AS15435	28,494	88,876	114,702	22,290,962	KABELFOON CAIW Diensten B.V., NL
AS31122	28,108	88,883	114,702	22,716,946	DIGIWEB-AS Digiweb Ltd., IE
AS8928	29,037	88,881	114,702	22,619,195	INTERROUTE Interoute Communications Limited, GB
AS57821	27,764	88,876	114,701	22,290,962	NONATTACHED-AS Mr Dan Luedtke, DE
AS6881	27,841	88,886	114,701	22,946,323	NIXCZ NIX.CZ z.s.p.o., CZ
AS57381	28,464	88,881	114,701	22,618,906	FNU TT PE Fnutt Consulting Daniel Husand, NO
AS50304	28,518	88,879	114,701	22,487,831	BLIX Blix Solutions AS, NO
AS1103	28,473	88,602	65,886	4,247,864	SURFNET-NL SURFnet, The Netherlands, NL
AS13030	27,739	88,552	261	933,892	INIT7 Init7 (Switzerland) Ltd., CH
AS9002	27,148	88,601	12	4,145,440	RETN-AS RETN Limited, UA
AS1836	27,107	88,579	2	2,670,604	GREEN green.ch AG Autonomous System, CH

It's structural, not temporal

- There is a visible stability to deviation from the quorum route set
 - The variation from the quorum is long – term stable, and does not rapidly self-correct
- We appear to assume that all Tier 1 providers, and their Tier 2, 3, ... resellers offer the same reachability set as each other – i.e. “default” is consistent everywhere
- But this is not necessarily the case all the time for every address in the routing system
- “Default” appears to vary by provider and by location
 - E.g.: 25 April, 1600 UTC:
AS2914: RouteViews 2,808,560,896 addresses
RIS: 2,807,358,208

So What?

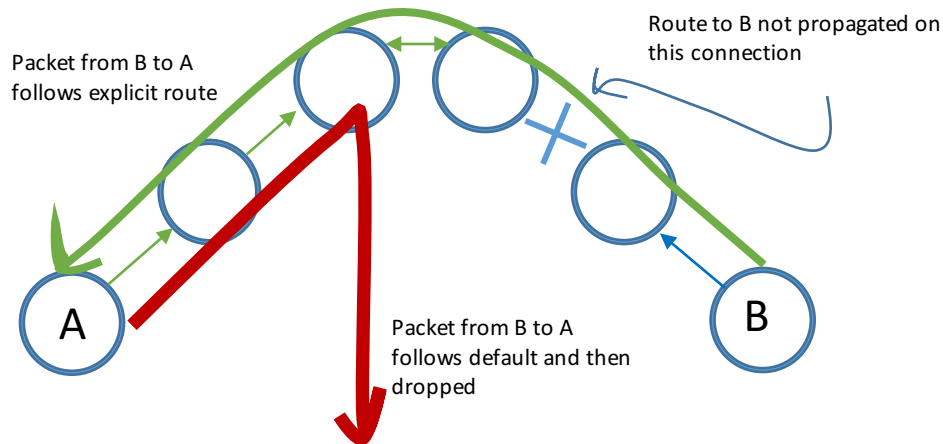
Surely all this is patched up by the widespread use of default in addition to specific routes? (*)

* Internet Optometry: Assessing the Broken Glasses in Internet Reachability”,
R. Bush, O. Maennel, M. Roughan, S Uhlig, ACM SIGCOMM IMC, 2009

So What?

Surely all this is patched up by the widespread use of default in addition to specific routes?

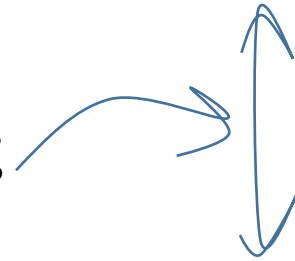
- Well, not really
- Default points along upstream transits
- It does not patch downstreams



To Recap: What is causing
this?

To Recap: What is causing this?

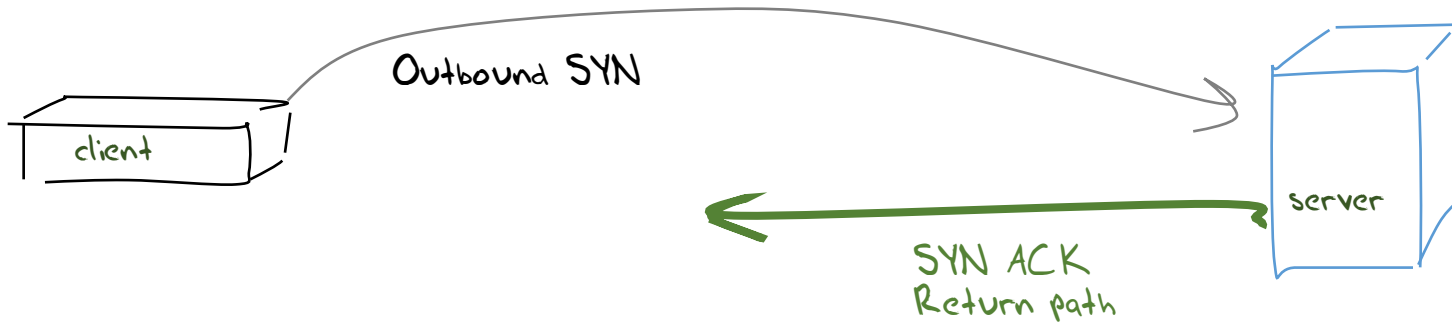
Could part of this be due to connectivity failure?



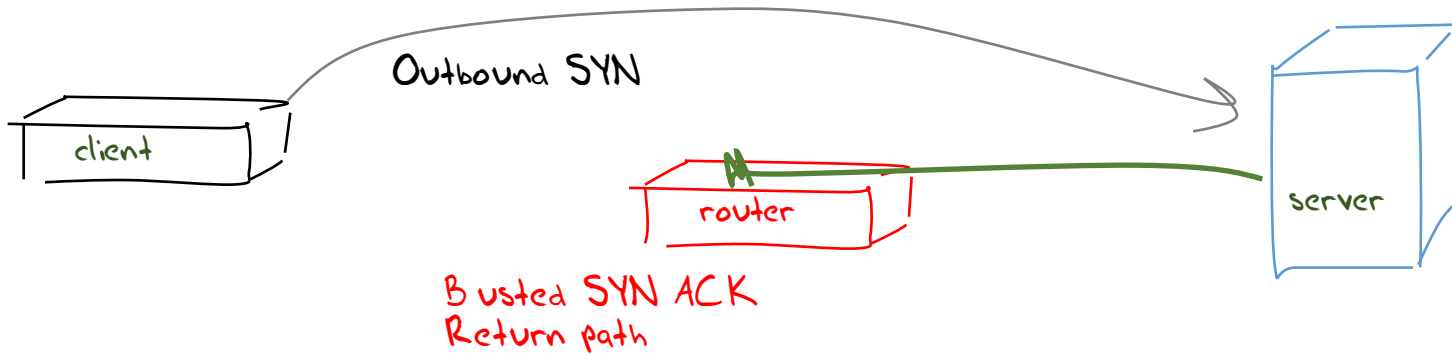
Connection Failure



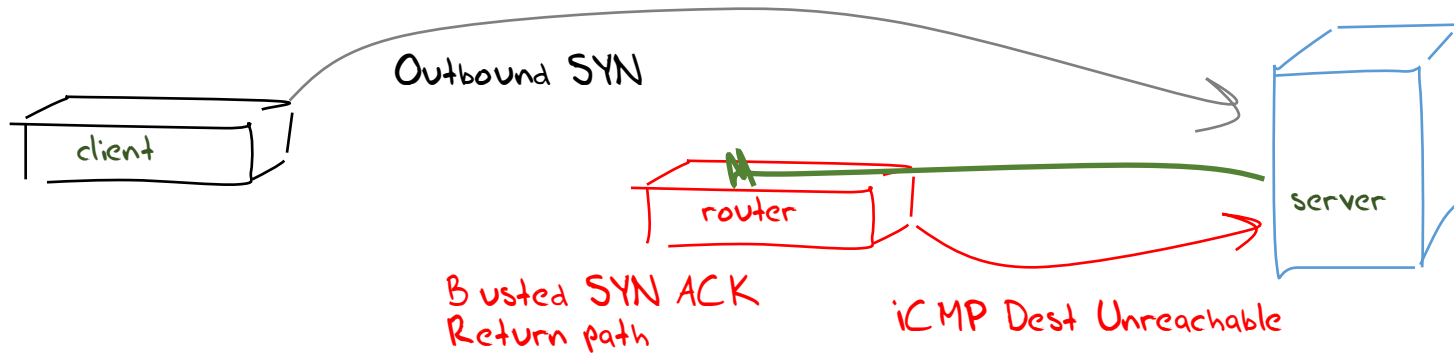
Connection Failure



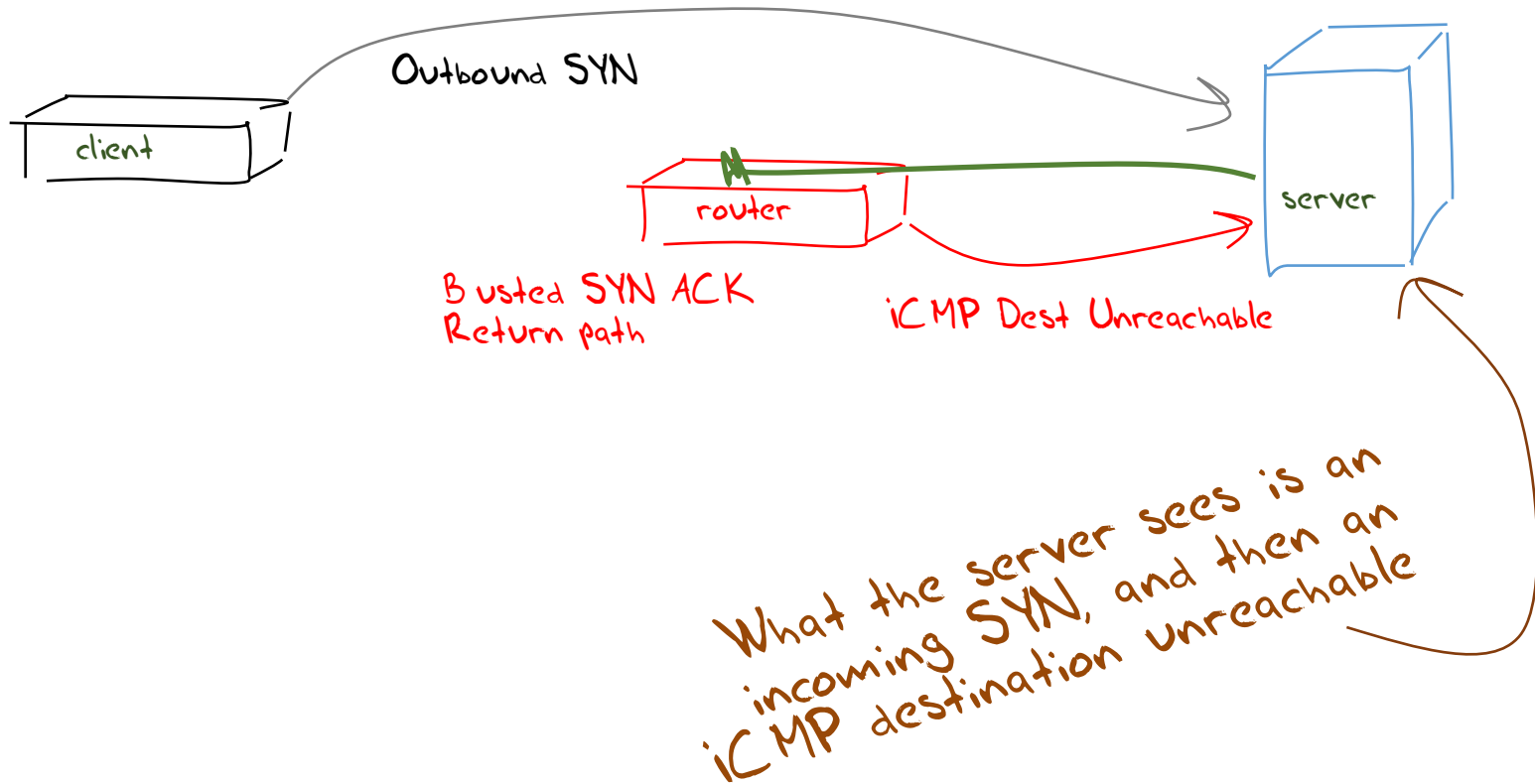
Connection Failure



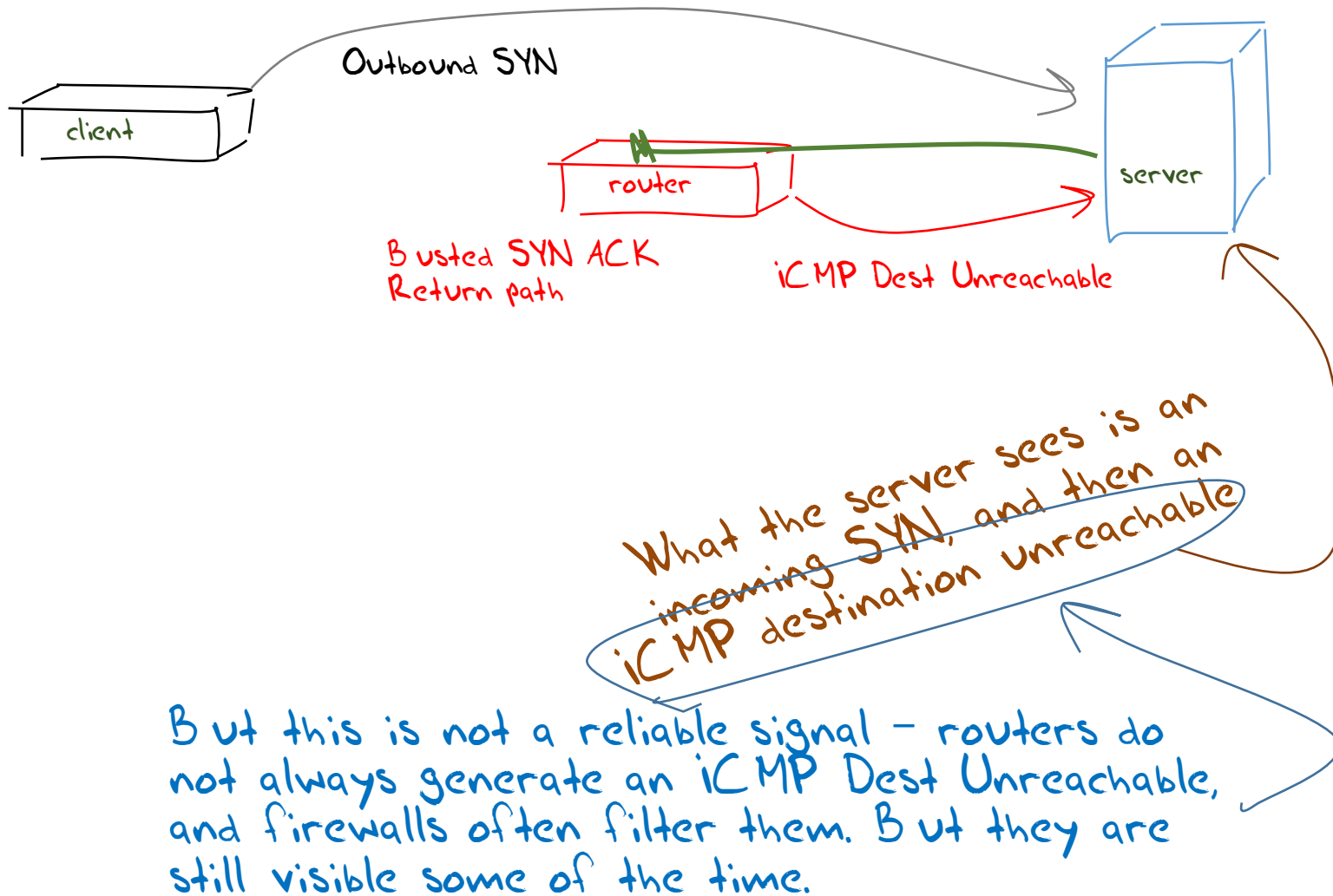
Connection Failure



Connection Failure



Connection Failure



And we see this ... here's an example

14:16:05.999497 IP (tos 0x0, ttl 55, id 31005, offset 0, flags [none], proto ICMP (1), length 80)

84.41.108.74 > 139.162.146.97: ICMP host 46.163.63.47 unreachable, length 60

↑
Outer packet is an ICMP Packet with a "destination unreachable" code sent to the server from the router at address 84.41.108.74

IP (tos 0x0, ttl 57, id 0, offset 0, flags [DF], proto TCP (6), length 52)

139.162.146.97.443 > 46.163.63.xx.52087: Flags [S.], cksum 0x5130 (correct), seq 3917125220, ack 685287936, win 29200, options [mss 1460,nop,nop,sackOK,nop,wscale 7], length 0



Payload packet is a SYN+ACK packet

Of course it's not the only reason why connections fail in IPv4

NATs are also erratic:

```
14:16:37.959604 IP (tos 0x38, ttl 55, id 40091, offset 0, flags [none], proto ICMP (1), length 80)
197.114.50.xx > 139.162.146.97: ICMP host 197.114.50.20 unreachable, length 60
    IP (tos 0x20, ttl 49, id 0, offset 0, flags [DF], proto TCP (6), length 52)
    139.162.146.97.443 > 197.114.50.xx.55648: Flags [S.], cksum 0xeb1e (correct), seq 3454637726, ack
3599361621, win 29200, options [mss 1400,nop,nop,sackOK,nop,wscale 7], length 0
```

Why is the destination telling us that it is unreachable?

Of course it's not the only reason why connections fail in IPv4

NATs are also erratic:

```
14:16:37.959604 IP (tos 0x38, ttl 55, id 40091, offset 0, flags [none], proto ICMP (1), length 80)
197.114.50.xx > 139.162.146.97: ICMP host 197.114.50.20 unreachable, length 60

IP (tos 0x20, ttl 49, id 0, offset 0, flags [DF], proto TCP (6), length 52)
139.162.146.97.443 > 197.114.50.xx.55648: Flags [S.], cksum 0xeb1e (correct), seq 3454637726, ack
3599361621, win 29200, options [mss 1400,nop,nop,sackOK,nop,wscale 7], length 0
```

Why is the destination telling us that it is unreachable?

it's probably a CGN that has lost its binding state while the TCP session was being established!

Connectivity appears to have a 1 in 500 failure rate in IPv4

- Some of it is based on asymmetric views of connectivity from the routing system
- Some of it is based on anomalous NAT behaviours
- And some of it ... we just can't tell!

internet

In the ~~Telephone~~ Network

Almost

Almost

most of the time

- Anyone can reach anyone else
- ~~All connected endpoints are equally reachable~~

Thanks!