What should networks do with IPv6 Extension Headers?

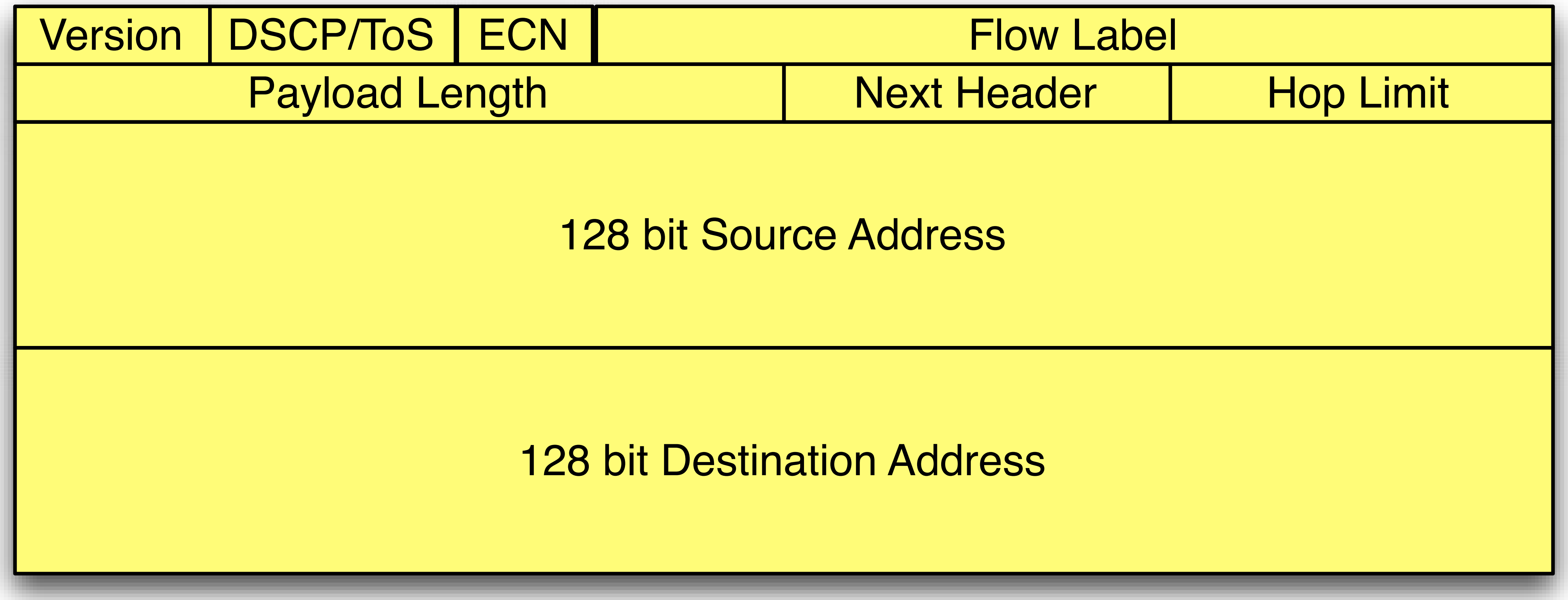
Ana Custura
Gorry Fairhurst

* University of Aberdeen
with funding from RIPE NCC

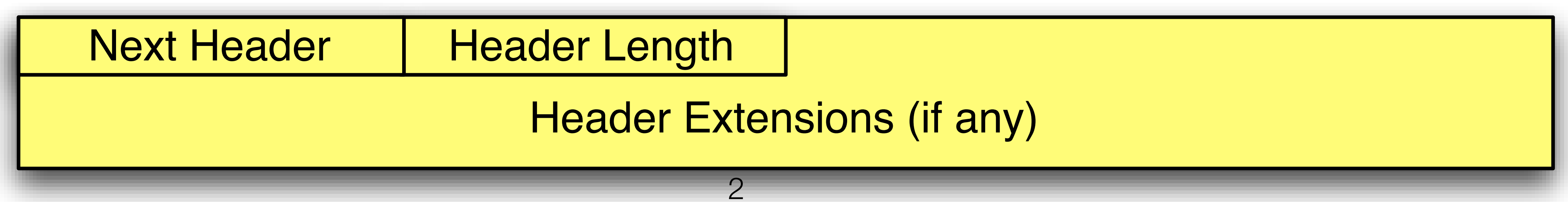
RIPE NCC

Electronics Research Group

IPv6



- IPv6 was standardised in the 1990's [RFC2460]
- Became Full Standard in 2017 [RFC 8200]



IPv6 promises

- Larger Address Space ✓
- More Efficient Forwarding/Routing ✓
- Improved IP Packet Fragmentation* ✓
- Multicast ?
- End-to-end Security (aka IPSEC) ?
- Extensibility ?

Fix to lack of IPv4 address space

Fix to lack of extension in IPv4

RIPE ‘86

IPv6 promises

- Larger Address Space ✓
- More Efficient Forwarding/Routing ✓
- Improved IP Packet Fragmentation* ✓
- Multicast ?
- End-to-end Security (aka IPSEC) ?
- Extensibility ?

Fix to lack of IPv4 address space

Fix to lack of extension in IPv4

After some refinements

RIPE ‘86

IPv6 promises

- Larger Address Space — Fix to lack of IPv4 address space ✓
- More Efficient Forwarding/Routing ✓
- Improved IP Packet Fragmentation* ✓ — **_*After some refinements_**
- Multicast ?
- End-to-end Security (aka IPSEC) ? — *Other ways have emerged, such as QUIC*
- Extensibility — Fix to lack of extension in IPv4 ?

RIPE '86

IPv6 promises

- Larger Address Space

Fix to lack of IPv4 address space

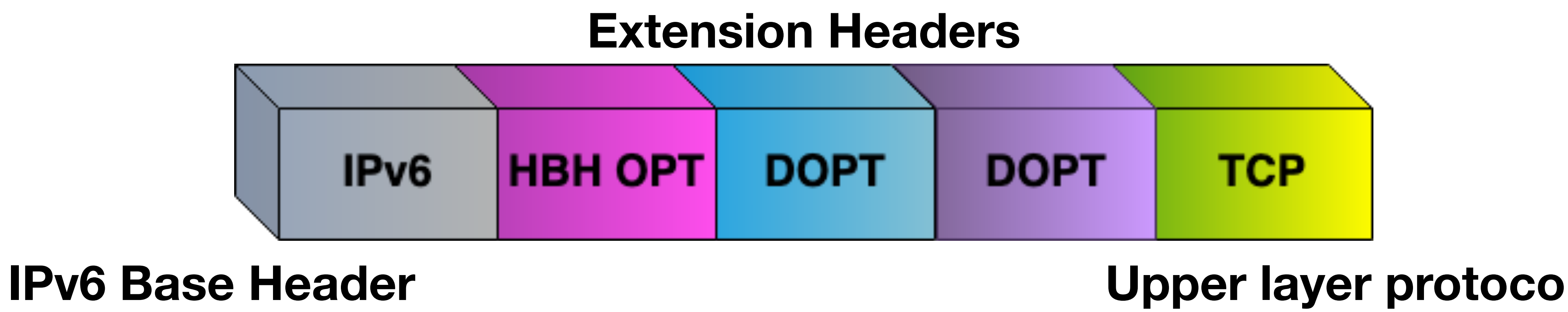
✓

- More Efficient Forwarding/Routing ✓
- Improved IP Packet Fragmentation* ✓ *After some refinements
- Multicast ?
- End-to-end Security (aka IPSEC) ? Other ways have emerged, such as QUIC
- Extensibility ? This talk!

Fix to lack of extension in IPv4

RIPE '86

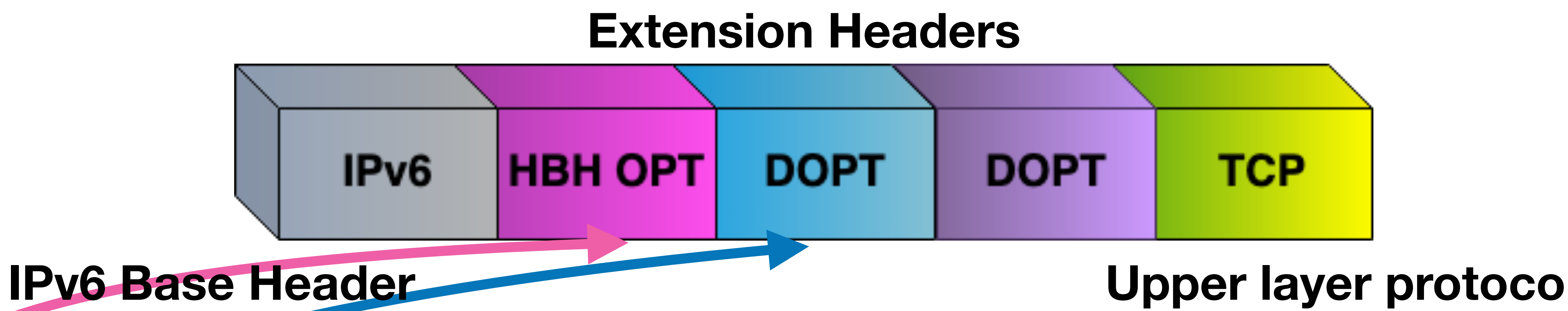
Extensibility - EH



Protocol Number	Description	References
0	IPv6 Hop-by-Hop Option	[RFC8200]
43	Routing Header for IPv6	[RFC8200] [RFC5095]
44	Fragment Header for IPv6	[RFC8200]
50	Encapsulating Security Payload	[RFC4303]
51	Authentication Header	[RFC4302]
60	Destination Options for IPv6	[RFC8200]
135	Mobility Header	[RFC6275]
139	Host Identity Protocol	[RFC7401]
140	Shim6 Protocol	[RFC5533]
253,254	Use for experimentation and testing	[RFC3692] [RFC4727]

https://www.iana.org/assignments/ipv6-parameters/ipv6-parameters.xhtml

Extensibility - EH



Protocol Number	Description	References
0	IPv6 Hop-by-Hop Option	C8200]
43	Routing Header for IPv6	00] [RFC5095]
44	Fragment Header for IPv6	[RFC8200]
50	Encapsulating Security Payload	[RFC4303]
51	Authentication Header	[RFC4302]
60	Destination Options for IPv6	[RFC8200]
135	Mobility Header	[RFC6275]
139	Host Identity Protocol	[RFC7401]
140	Shim6 Protocol	[RFC5533]
253,254	Use for experimentation and testing	[RFC3692] [RFC4727]

Some EHs carry 'Options'

https://www.iana.org/assignments/ipv6-parameters/ipv6-parameters.xhtml

EH concerns in RFC 9098 (2021)

- Slow-path processing of EHs
- Buggy implementations* -> DoS
- Complexity not bounded: can reduce router forwarding rate
- Large EH can exceed router parsing buffer

Some EHs had a rocky start

RIPE '86

* To this date, vulnerabilities still found: https://www.interruptlabs.co.uk/articles/linux-ipv6-route-of-death

EH concerns in RFC 9098 (2021)

- Slow-path processing of EHs
- Buggy implementations* -> DoS
- Complexity not bounded: can reduce router forwarding rate
- Large EH can exceed router parsing buffer

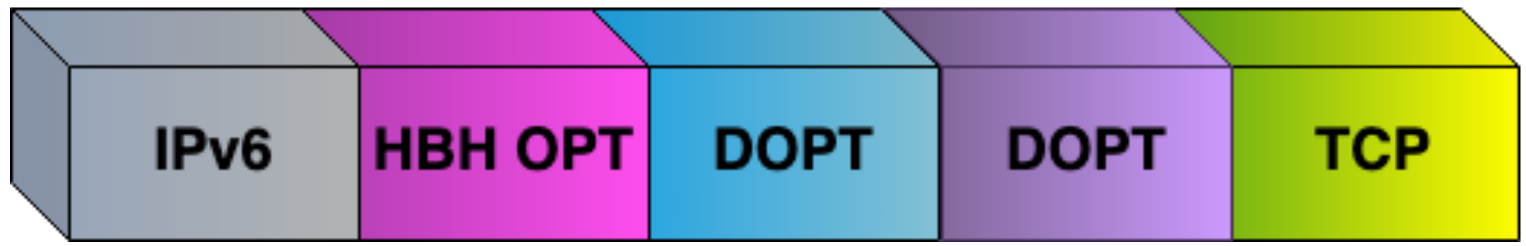
Some EHs had a rocky start

Measurements in RFC 7872 show many networks drop packets with EH

RIPE ‘86

* To this date, vulnerabilities still found: https://www.interruptlabs.co.uk/articles/linux-ipv6-route-of-death

Renewed Interest in EHs

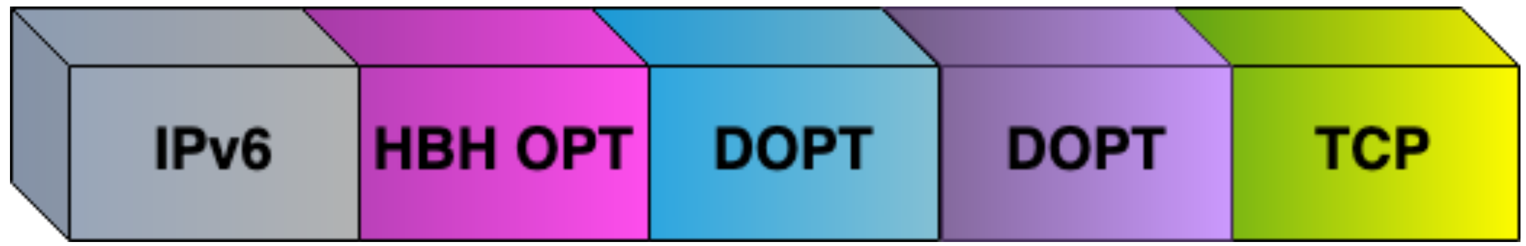


- IPv6 Segment Routing type (SRv6) [RFC8986]
- Service Management and Performance Measurement using PDM [RFC8250]
- In-situ Operations, Administration, and Maintenance [RFC9378]
- AltMark Measurement DO and HbH Options [RFC9343]
- minPMTU HBH Option [RFC9268]

ASICs are emerging that can process EHs at line speed!

RIPE ‘86

Renewed Interest in EHs



- IPv6 Segment Routing type (SRv6) [RFC8986]
- Service Management and Performance Measurement using PDM [RFC8250]
- In-situ Operations, Administration, and Maintenance [RFC9378]
- AltMark Measurement DO and HbH Options [RFC9343]
- minPMTU HBH Option [RFC9268]

Can Options be used more widely in the Internet?

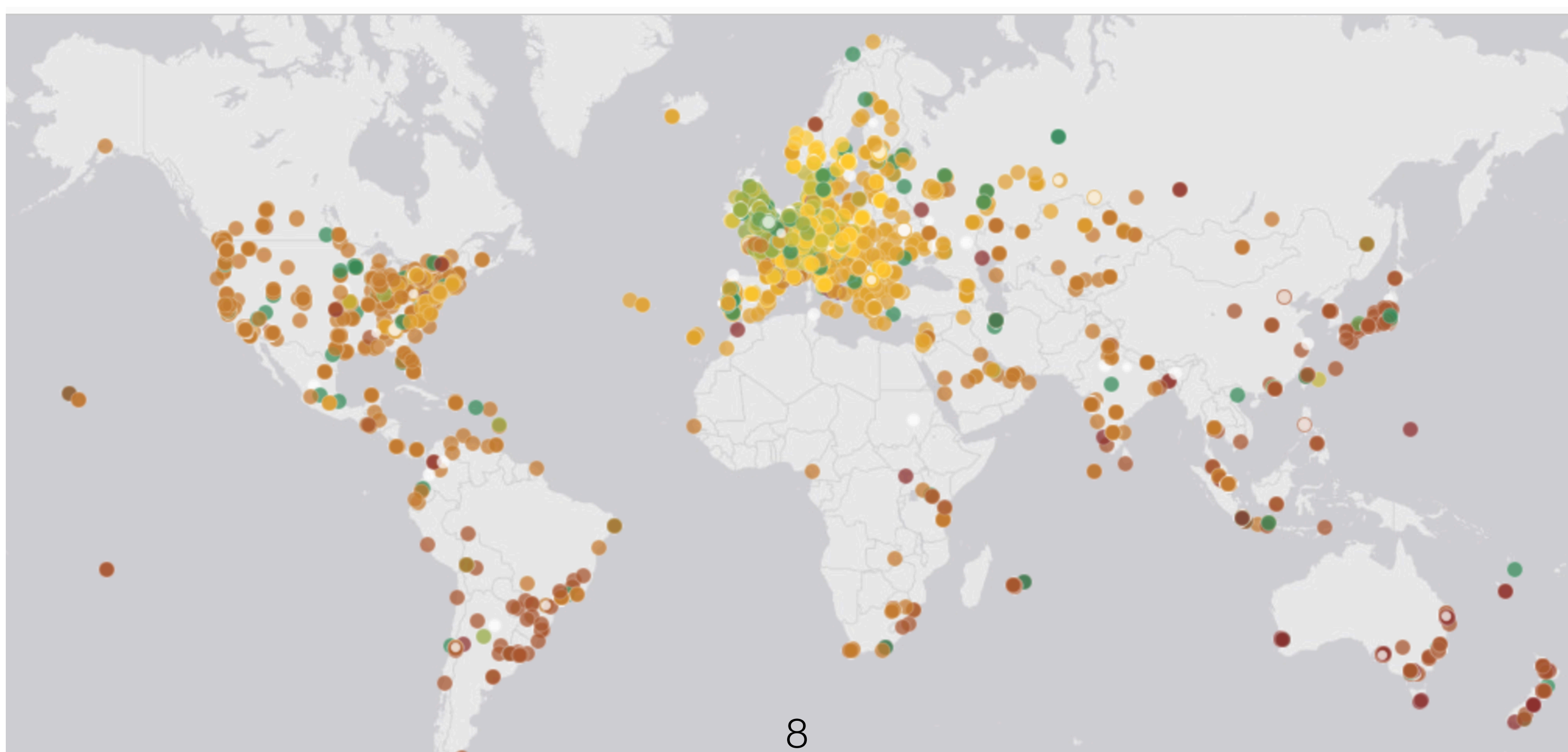
Existing Measurements

- Focus on Destination Options (DOPT) and Hop-by-Hop Options (HBHOPT) EHs
- Let's measure survival of packets with EH

	Destination Option EH	Hop-by-Hop Option EH
RFC 7872 (2016) [1] - server edge	80-90%	45-60%
My own (2018) data [2] - server edge	70-75%	15-20%
APNIC (2022) [3] - client edge	30-80%	0%
JAMES (2022) [4] - core	94-97%	8-9%

Experiment 1: Survival

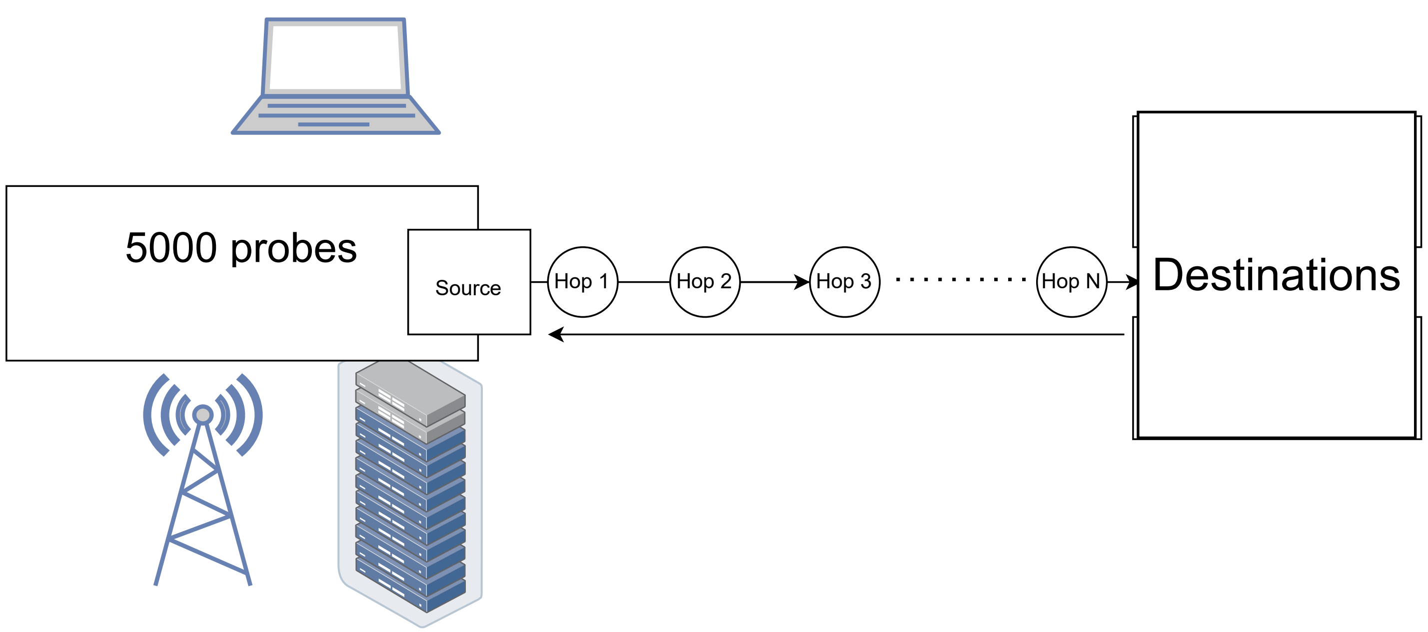
- ~5500 IPv6-enabled probes in RIPE, globally distributed
- Testing survival by sending packets to 7 targets (UK, US, Canada, Australia, Zambia, Kazakhstan, France)
  - {TCP, UDP} to port 443
    - {**DOPT**, **HBHOPT**} + control IPv6 packets
  - Survives if packet reaches destination AS



RIPE '86

Survival at a Glance DOPTs

- 8B PadN option
- High survival for **DOPTs**
- Difference between TCP and UDP



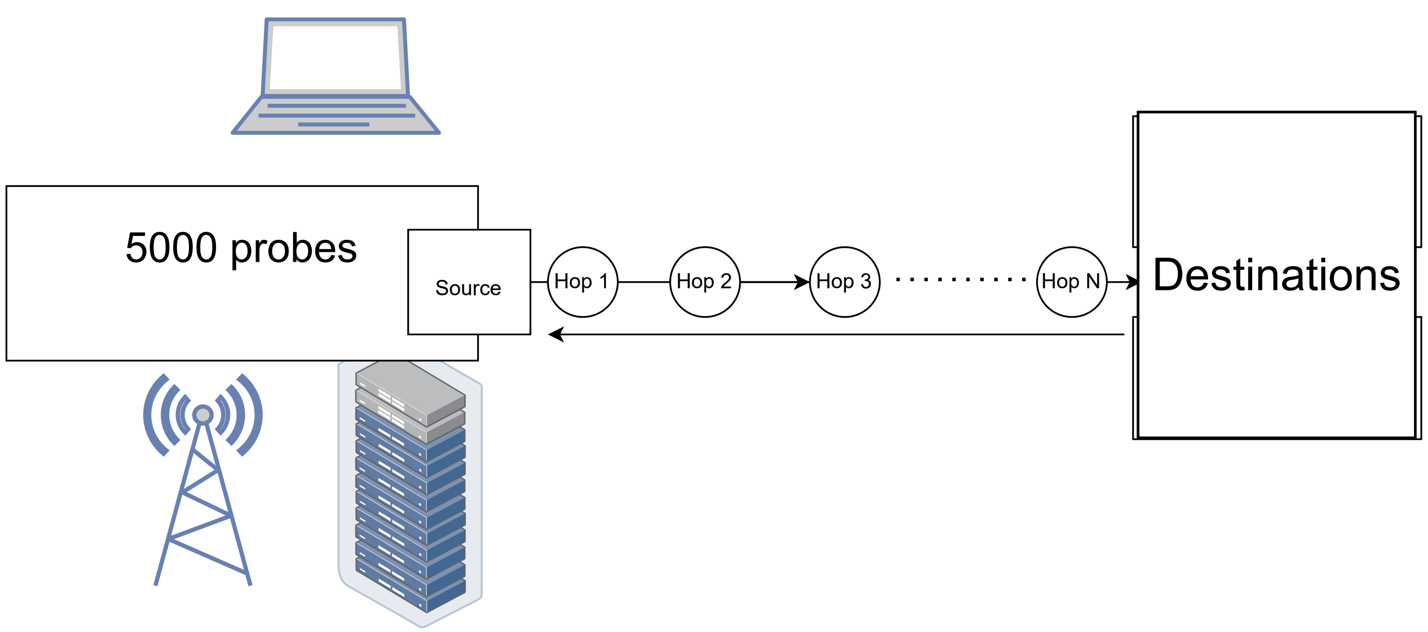
DOPT

~92%	UDP
~68%	TCP

RIPE ‘86

Survival at a Glance HBHOPTs

- 8B PadN option
- **HBHOPTs** survive some paths
- Difference between TCP and UDP



HBHOPT

~11%	UDP
~9%	TCP

Per-AS Survival (UK path)

DOPT

The **local AS** is responsible for most of the drops:

- 5% for UDP
- 25% for TCP

	1st AS	AS1>AS2	∞
DOPT UDP 8B	95.3%	93%	91.5%
DOPT TCP 8B	74.7%	70%	68.5%

HBHOPT

The **local AS** is responsible for most of the drops:

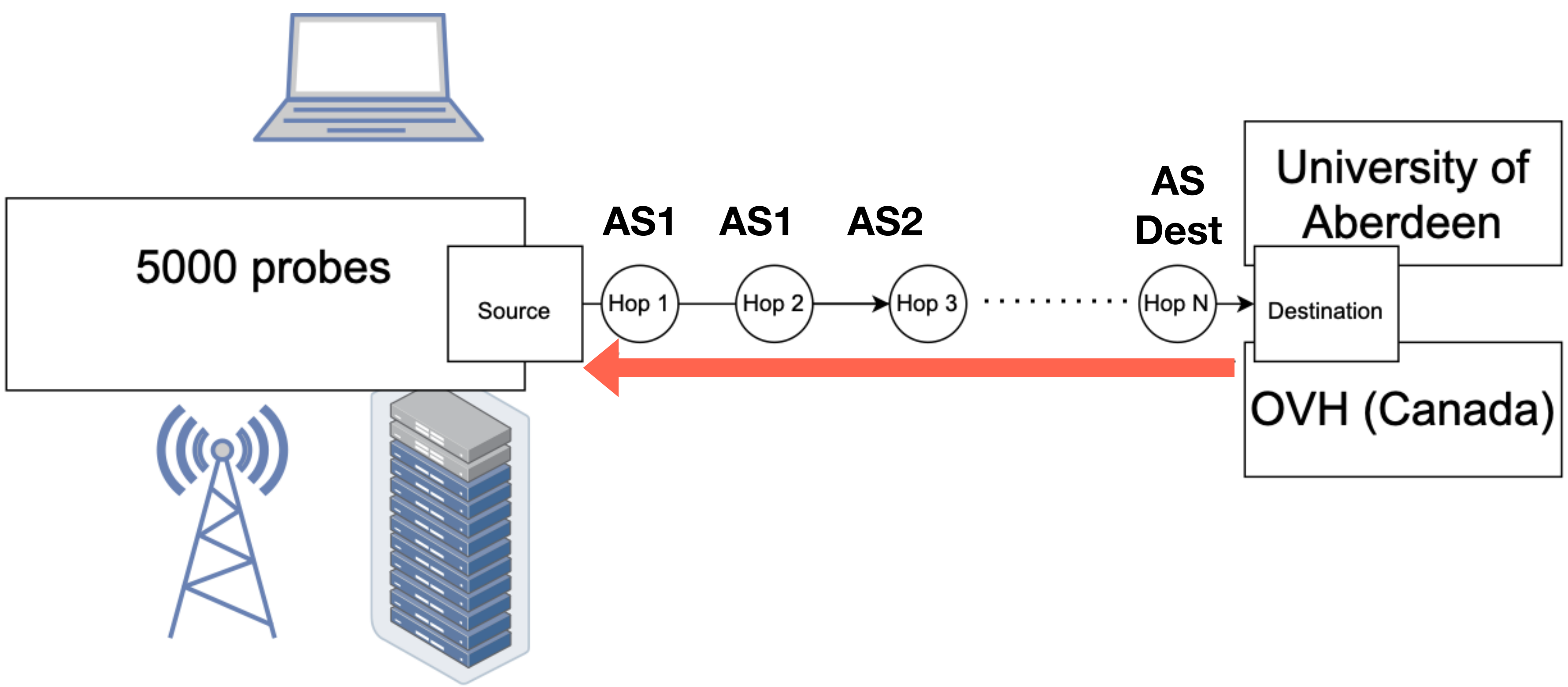
- 68% for UDP
- 74% for TCP

	1st AS	AS1>AS2	2nd AS	AS2>AS3	∞
HBHOPT UDP 8B	31.4%	20.1%	15%	12.2%	11.4%
HBHOPT TCP 8B	26.9%	16.3%	13.9%	9.7%	8.6%

Drops are considered to be within the AS if the next hop on a control measurement is also in that AS.
If the next hop would otherwise be in a different AS, then the drop is attributed to the AS boundary.

RIPE '86

What if packets would traverse the first AS?



- Most probes have public IPv6 addresses
- Reverse traceroute on paths where drops happen in first AS
- Same protocol/port
- Does the packet reach original AS?

RIPE '86

What if packets would traverse the first AS?

DOPTs

Reverse traceroute on paths with drops in first AS (n=271 paths for UDP):
95 - 97% make it back to the original AS.

	%predicted traversal
DOPT UDP (UK)	~96%
DOPT UDP (Canada)	~96%

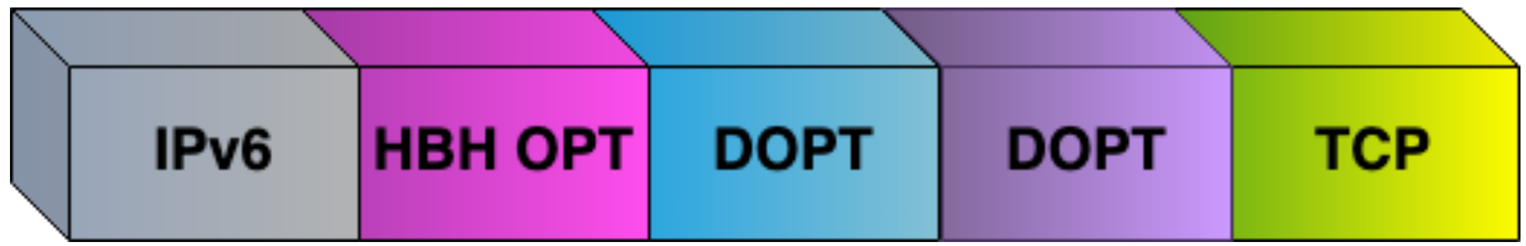
HBOPTs

Reverse traceroute on paths with drops in first AS (n=3150 paths for UDP):
10 - 17% make it back to the original AS.

	%predicted traversal	Notes
HBHOPT UDP (UK)	~17%	60% packets get dropped at LINX peering
HBHOPT UDP (Canada)	~25%	

Transit networks drop more packets with HBHOPTs

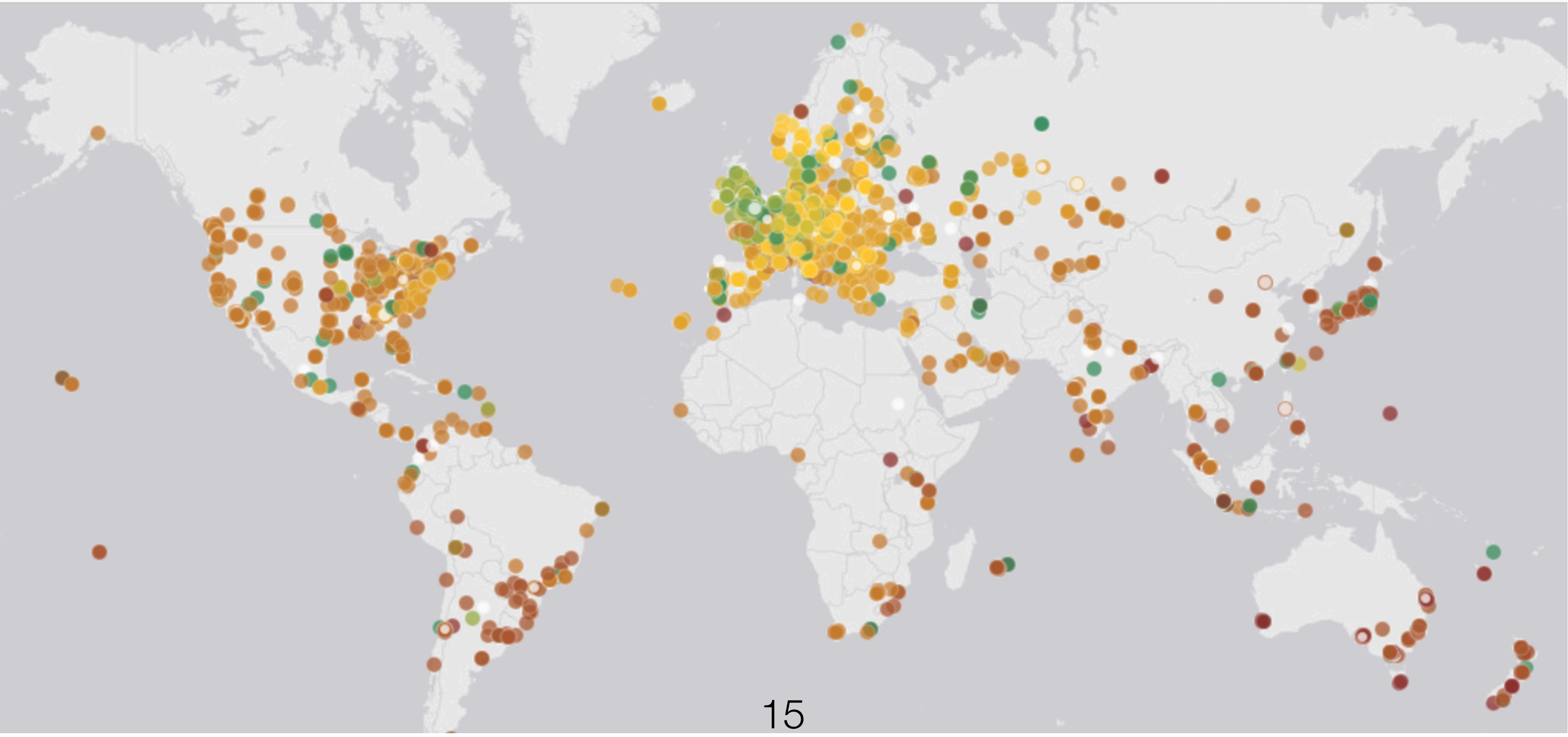
Why?



- Network/Firewall policy (e.g. Fastly)
- Different router designs
- Different devices (CPE, load balancers, firewalls, IDS) wanting access to upper layer protocols
- End-systems (NICs that do processing in hosts)
  - Is EH size a factor? Is full chain size a factor?

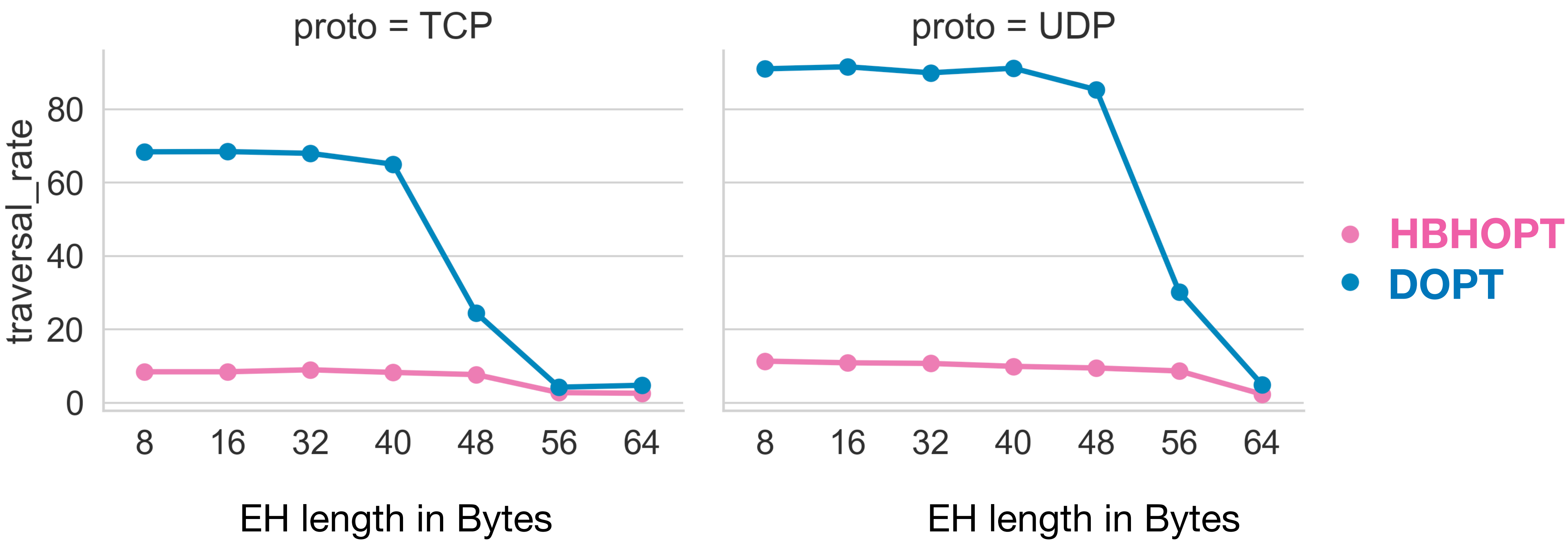
Experiment 2: Size

- {TCP, UDP} to port 443
  - {**DOPT**, **HBHOPT**} + control measurement
  - {8,16,32,40,48,56,64} B in size to one target
- Survival is successful if packet reaches destination AS



RIPE ‘86

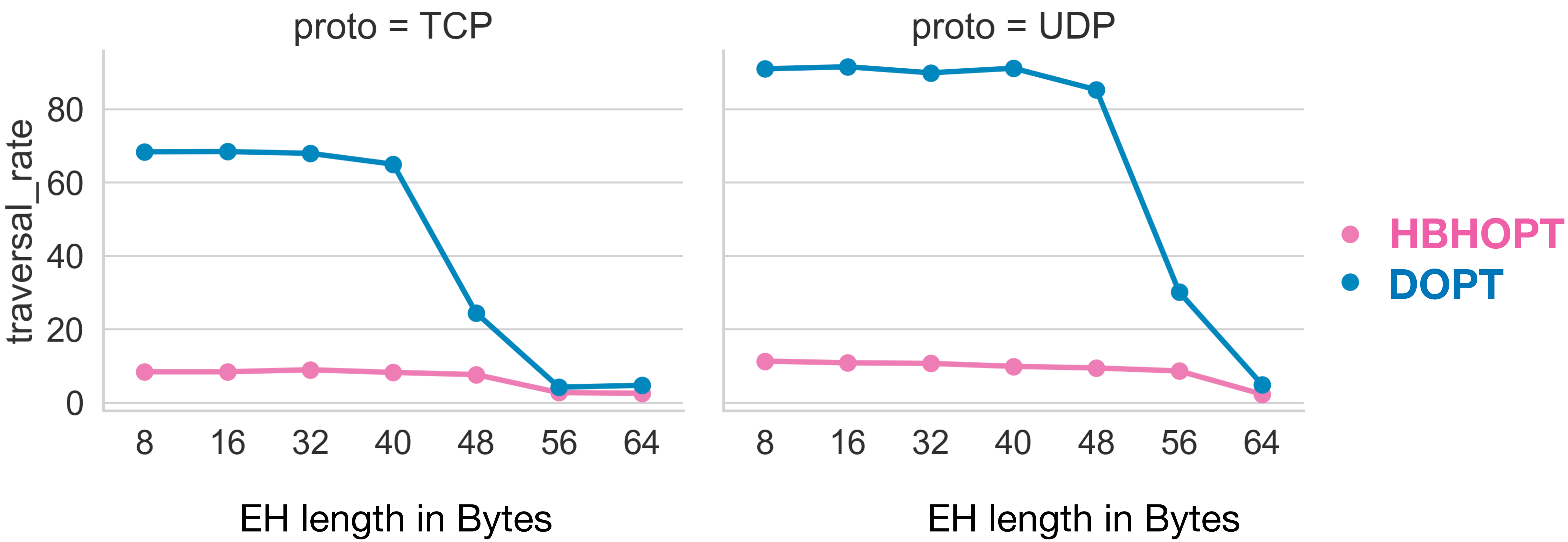
Traversal vs Size



- TCP sees the biggest drop in traversal at 48B: 48 + 20 = 68B (108B total)
- UDP sees the biggest drop at 56B: 56 + 8 = 64B (104B total)
- Is this due to EH size or IPv6 total chain size?
- 40B is the max for IPv4 options

RIPE '86

Traversal vs Size



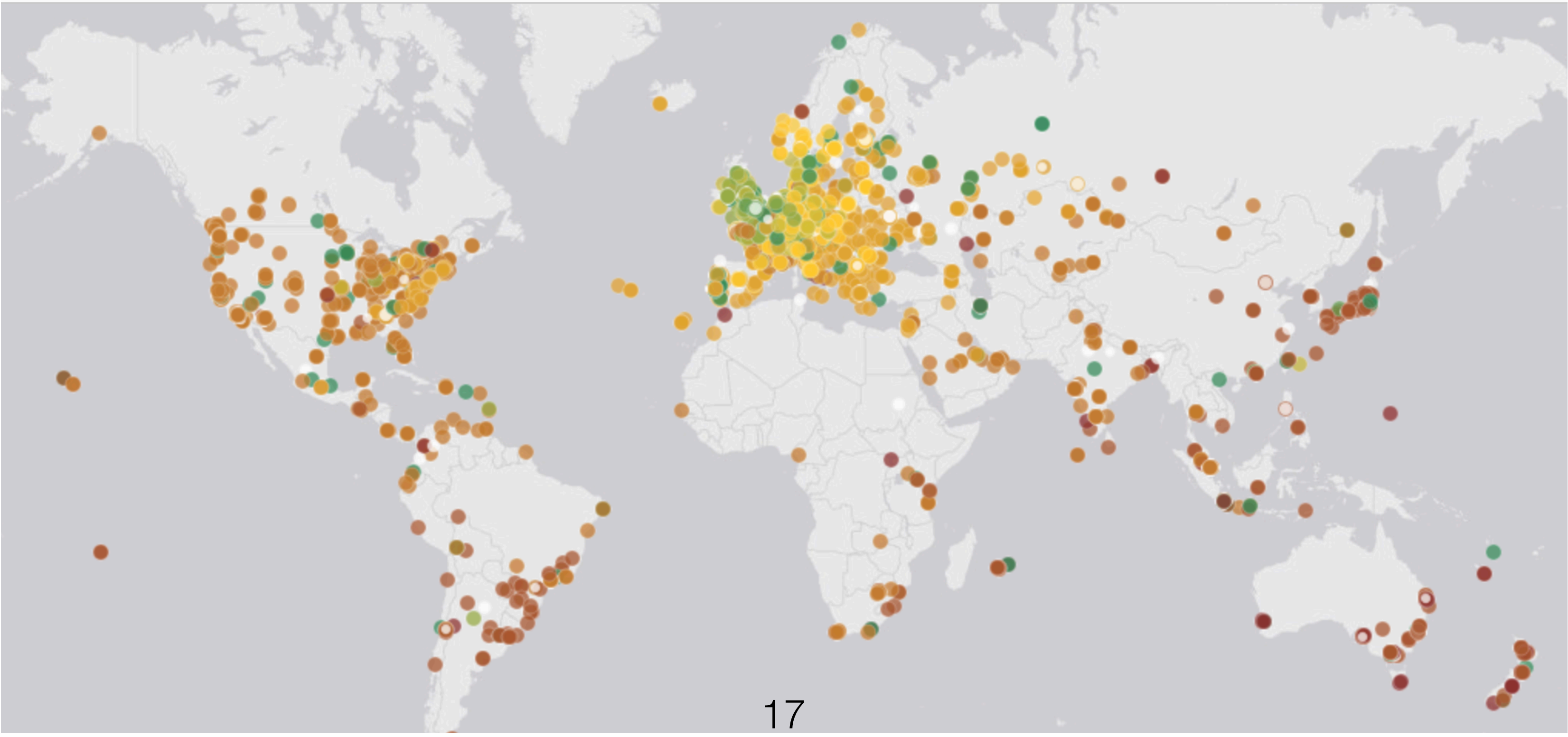
- TCP sees the biggest drop in traversal at 48B: 48 + 20 = 68B (108B total)
- UDP sees the biggest drop at 56B: 56 + 8 = 64B (104B total)
- Is this due to EH size or IPv6 total chain size?
- 40B is the max for IPv4 options

Where EHs can be used, 40B often works

RIPE '86

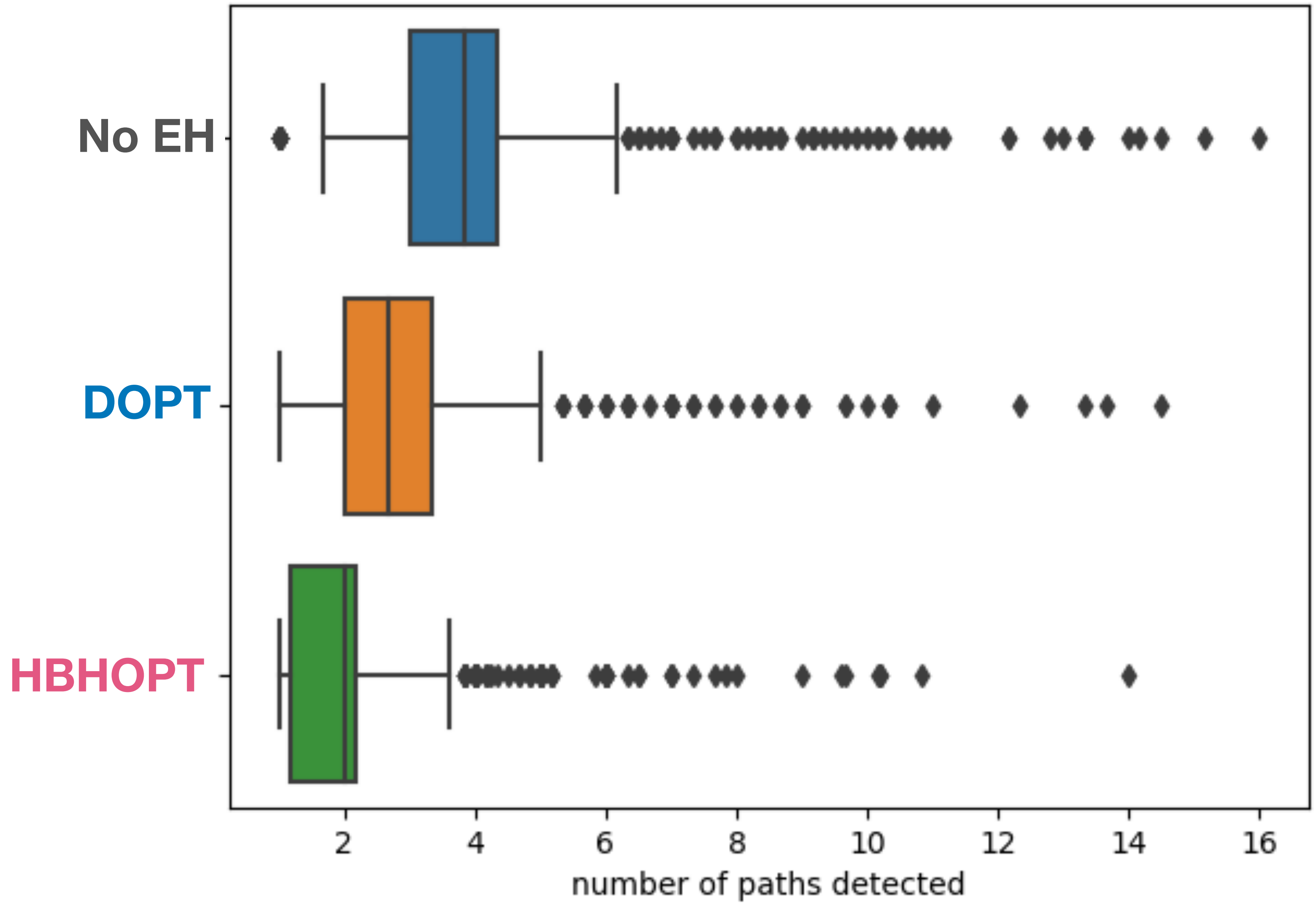
Experiment 3: ECMP

- ECMP uses header information for load-balancing
- UDP to port 443 from ~850 probes
  - {DOPT, HBHOPT} + control measurement
  - **We measure 16 Paris ID variations to the same target (Flow Label + source port combinations)**



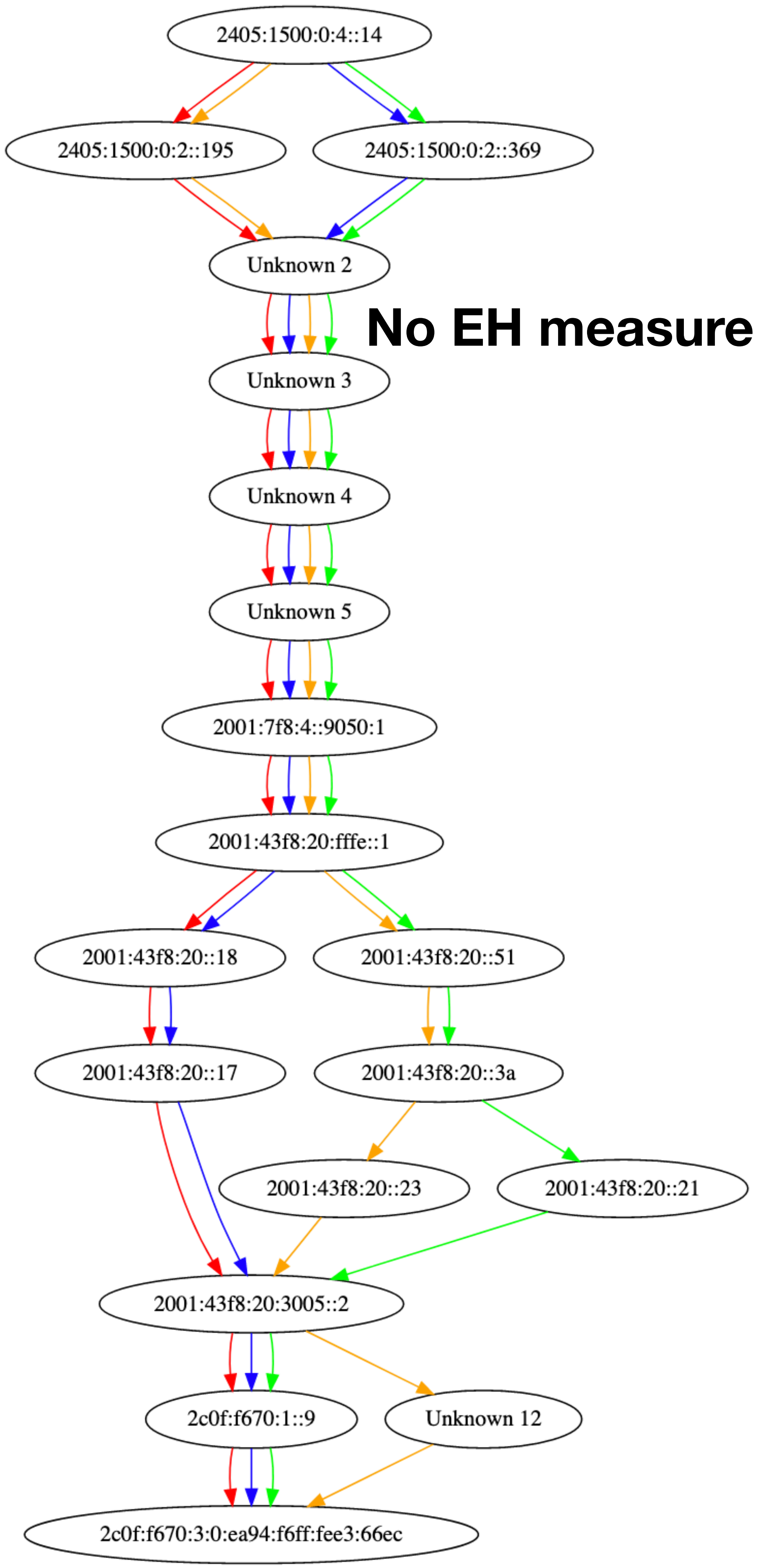
RIPE '86

Statistics: ECMP

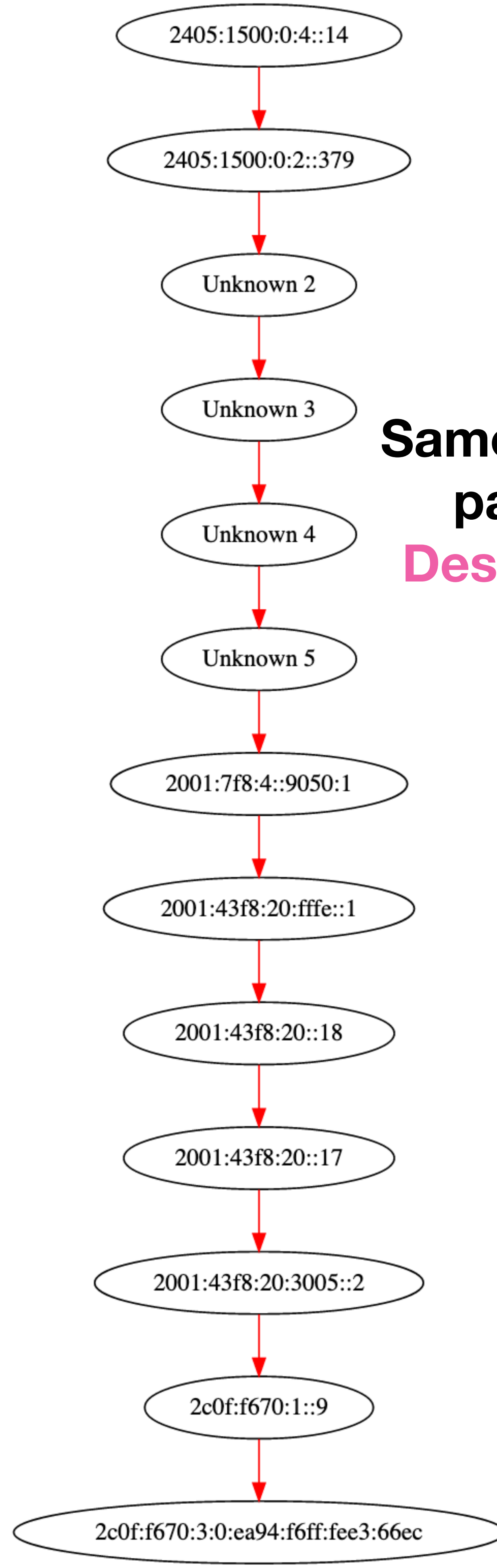


- Not all devices are equipped to handle flows that mix packets with and without EHs
  - Motivates the use of Flow Label for ECMP

RIPE '86

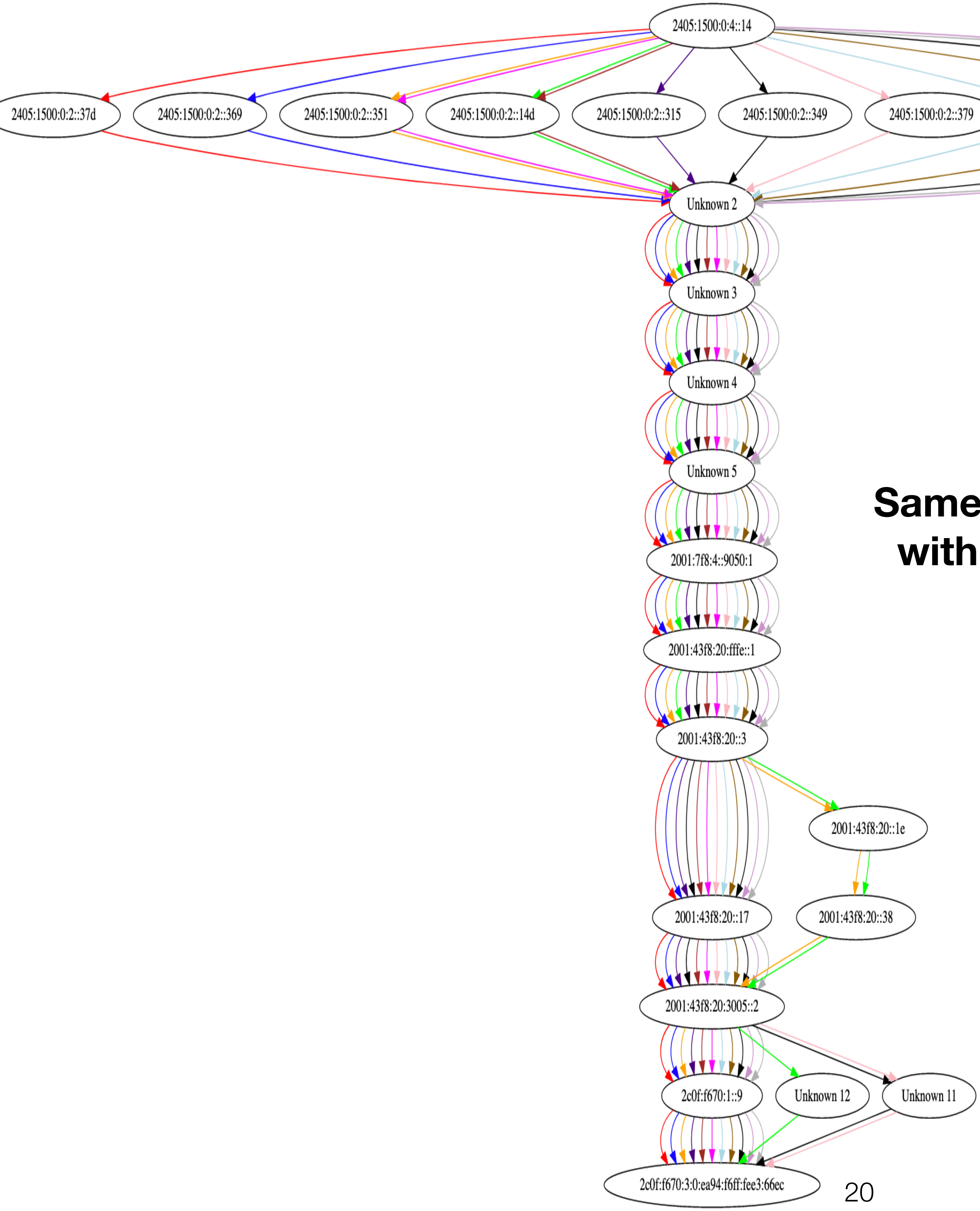


No EH measurement



Same source-destination pair measured with Destination Options EH Packets

RIPE '86



Same source-destination pair measured with Hop-by-Hop Options EH packets

RIPE '86

What should networks do with them?

- Firewall, firewall … **but only if you need to**!
  - IPv6 is being extended within domains
  - Unnecessary barriers bad for innovation
- More capable ASICs - > Forwarding + processing without impacting performance

3 new IETF drafts might help: draft-ietf-6man-eh-limits, draft-ietf-6man-hbh-processing, draft-ietf-v6ops-hbh

RIPE ‘86

What next?

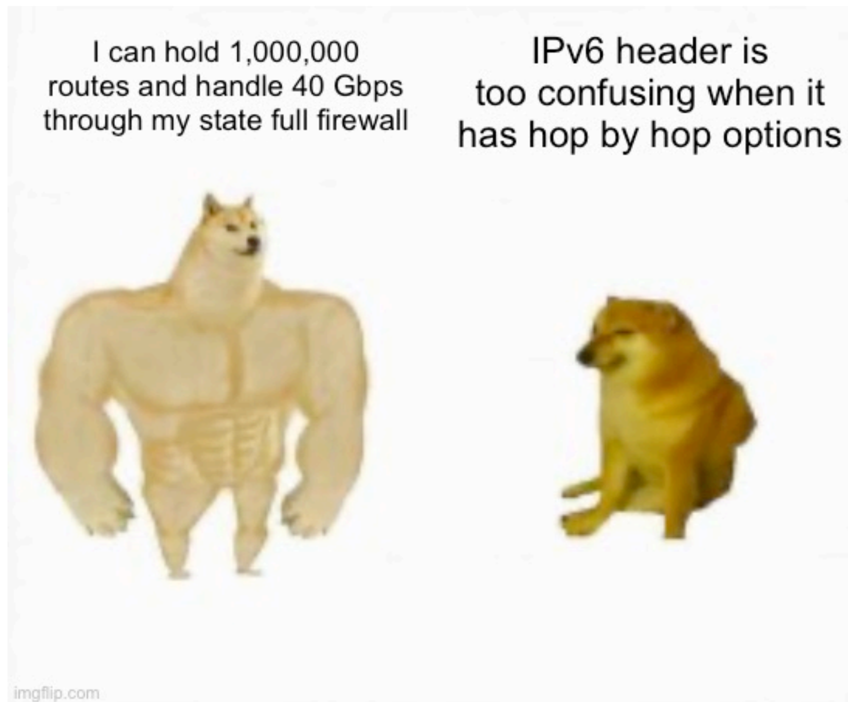
- Fragmentation got ‘fixed’ after trials and tribulations
- What about Options:

  …within a domain? It is low-risk, can be and IS done now

  …opportunistically in the Internet? DOPTs almost there

  What about in 5 years’ time?

References



- [1] https://www.rfc-editor.org/rfc/rfc7872
- [2] https://datatracker.ietf.org/meeting/108/materials/slides-108-6man-sessb-exploring-ipv6-extension-header-deployment-updates-2020-01
- [3] https://blog.apnic.net/2022/10/13/ipv6-extension-headers-revisited/
- [4] https://datatracker.ietf.org/doc/draft-vyncke-v6ops-james/