

# Exploring usable path MTU in the Internet

Ana Custura, Gorrry Fairhurst, Iain Learmonth

University of Aberdeen

June 2018

# What is PMTU

- ▶ Maximum Transmission Unit (MTU) = Largest size of packet that may be sent across a link without fragmentation
- ▶ Path MTU (PMTU) = Largest size of packets an endpoint can send on a specific Internet path
- ▶ PMTU Discovery (PMTUD) = network layer mechanism to dynamically determine the PMTU using ICMP
- ▶ MSS = Maximum Segment Size; TCP mechanism to discover remote link MTU

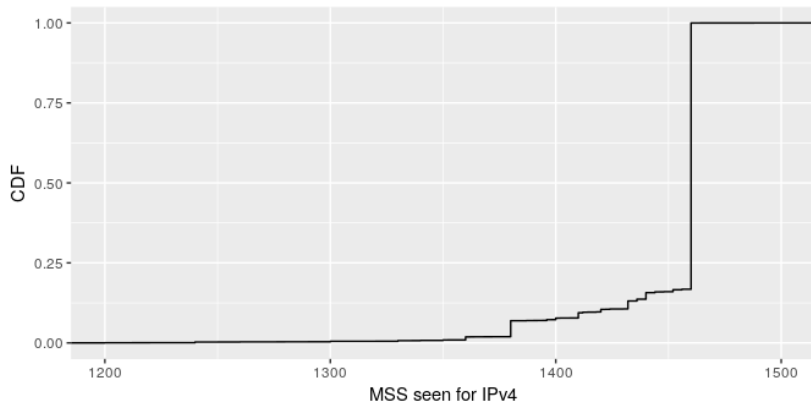
# Avoiding black holes

- ▶ Middleboxes can change TCP MSS option to avoid PMTUD failures: MSS Clamping
- ▶ PLPMTUD - slow adoption
- ▶ Other black-hole detection mechanisms: clearing the DF bit

# Experiments and datasets

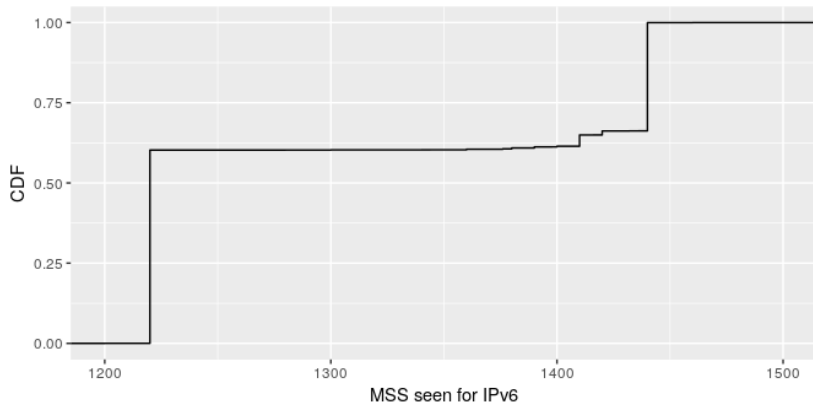
Purpose	Tool used	Dataset Name
Collect server MSS	PATHspider	A.1 "PATHspider"
Validate server MSS	Ping	A.2 "Ping"
Collect wireless/mobile client MSS	Pathtrace	B.1 "MONROE"
Collect wired edge client MSS	RIPE Atlas Traceroute	B.2 "RIPE"
Explore server PMTUD	Scamper	C.1 "Scamper"
Explore edge PMTUD	Netalyzer Traceroute	C.2 "Netalyzer"
Inspect ICMP quotations	Pathtrace	D "ICMP"

# Server advertised MSS - PATHspider results



**Figure:** Advertised MSS (in bytes) on TCP SYN/ACK server response seen at Janet academic network

# Server advertised MSS - PATHspider results



**Figure:** Advertised MSS (in bytes) on TCP SYN/ACK server response seen at Janet academic network

# Server advertised MSS - Ping results

- ▶ For 295,000 PATHspider targets we sent an ICMP probe to determine if they can be reached with a 1500 byte probe (Dataset A.1 "Ping")
- ▶ The control probe was the size of the advertised MSS
- ▶ Out of the subset which advertised an MSS lower than 1460 (34,920), 93% were reachable with a probe of 1500 bytes.

# Client advertised MSS - Mobile edge results

- ▶ Dataset B.1, “MONROE”, consists of traceroute-style measurements collected from the MONROE platform
- ▶ A total of 888 (21%) hops returned an MSS option
- ▶ MSS Clamping to allow for fragmentation inside the tunnel

**Table:** Inserted MSS options by mobile network,  $n = 10$  *paths*

Network	Inserted MSS option
Telenor Norway	1410 bytes
Telia Sweden	1400 bytes
Vodafone Italy	1400 bytes
Wind Italy	1420 bytes



# Client advertised MSS - Mobile edge results

- ▶ We sent 1500 byte UDP probes with the DF flag set on 10 paths; they reached our server
- ▶ Tested 16 mobile operators from over 40 vantage points within the MONROE platform (Dataset C.2)
- ▶ Both experiments consistently reported a PMTU of 1500 bytes

# Client advertised MSS - Wired edge results

- ▶ TCP traceroute from 3000 RIPE Atlas probes towards our server (Dataset B.2, "RIPE")
- ▶ 4.8% of the sent probes arrive at our server carrying an MSS option, some larger than standard ethernet
- ▶ 23% or 764 of the MSS values received by the probes differ from sent value of 1460

# MTU in the Internet

- ▶ 60k Cisco Umbrella domains - dataset C.1, "Scamper"

	1420 MTU	576 MTU	576 MTU Black-hole
PMTUD Too Small	7.45%	3.7%	0.95%
PMTUD Success	68.2%	63.9%	8.2%
PMTUD Failure	16.4%	19.5%	67.4%
No DF set	12.5%	12.3%	15.2%
Clear DF	2.7%	4.1%	NIL

	1280 MTU	1280 MTU Black-hole
PMTUD Too Small	59.6%	53.1%
PMTUD Success	95.5%	32%
PMTUD Failure	4.5%	67.9%

# ICMP quotation health

- ▶ Total of 125,212 ICMP and ICMPv6 replies (Dataset D, "ICMP"); 14257 unique addresses
- ▶ No evidence of quotations smaller than the minimum permitted size to match a probe is returned
- ▶ For 12 replies, the quoted transport payload did not match the probe that was sent

# Take aways and discussion

- ▶ 99% of MSS values seen client-side range between 1200 and 1460 bytes (both v4 and v6)
- ▶ For server MSS advertisements, the same figure is 98.95
- ▶ More than half of IPv6 hosts are configured to advertise a MSS corresponding to the minimum IPv6 MTU

# Take aways and discussion

- ▶ 95% tested IPv6 and 68% for IPv4 servers succeed in performing PMTUD
- ▶ 20% fail rate for IPv4, twice the amount reported in 2010 [?]
- ▶ Many servers artificially lower their MSS
- ▶ MSS clamping is common, in both mobile and wired edge
- ▶ Manufacturers and network admins do not trust PMTUD

# Take-aways and discussion

- ▶ 68% IPv6 and 76% IPv4 webserver failed PMTUD when local messages were blackholed
- ▶ Mechanisms for black-hole detection are available for TCP
- ▶ PLPMTUD can also provide black-hole detection for TCP

# Conclusion

- ▶ Less servers in the internet use PMTUD than previously thought
- ▶ Need to make PMTUD more robust for UDP



# Next Steps

- ▶ Present findings at TMA 2018
- ▶ Edge measurements via passive probe
- ▶ Journal paper