

Discovering Path MTU black holes in the Internet using RIPE Atlas

Maikel de Boer
Jeffrey Bosma

5 July 2012

Introduction

- Black holes
 - “A sphere of influence into which or from which communication or similar activity is precluded.” ~ Wiktionary.org
 - In layman’s terms: what goes in is forever lost
 - The Internet is full of black holes
 - Many possible causes
 - E.g., misconfiguration, bugs in software, etc.
 - We focus on Path MTU black holes

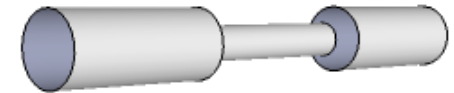
Research questions

Where on the Internet do Path MTU black holes occur?

Do Path MTU black holes occur more often in the IPv6-Internet compared to IPv4?

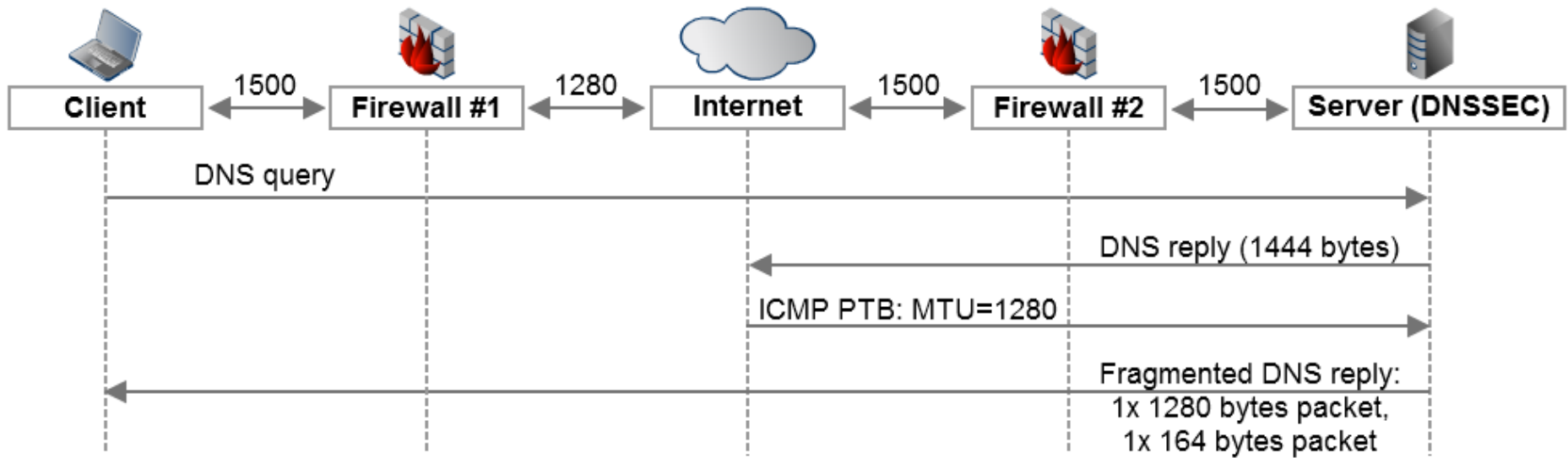
Theory

- The Internet: enormous collection of links
- Maximum Transmission Unit (MTUs) on network interface
 - Limits the amount of data in packets
 - Two-way limit: sending and receiving
- Path MTU (RFC1191)
 - Highest possible MTU for entire path
 - Determined by link with smallest MTU
- Internet Path MTU is commonly 1500 bytes
 - Not always the case
 - Requires Path MTU detection mechanism



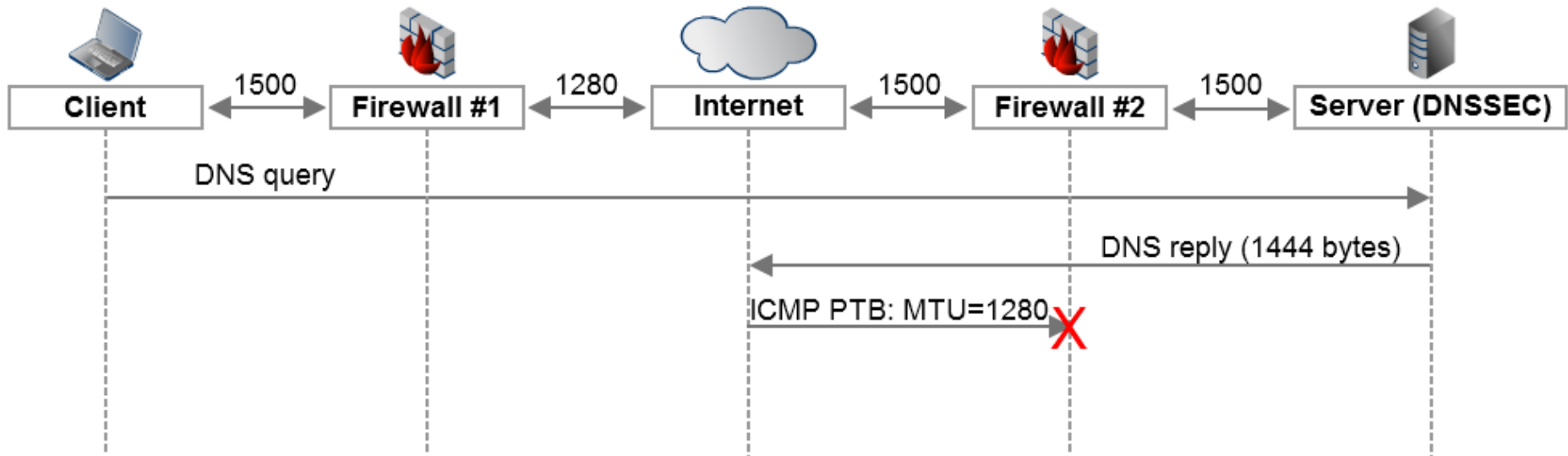
Theory

Path MTU Discovery (PMTUD)



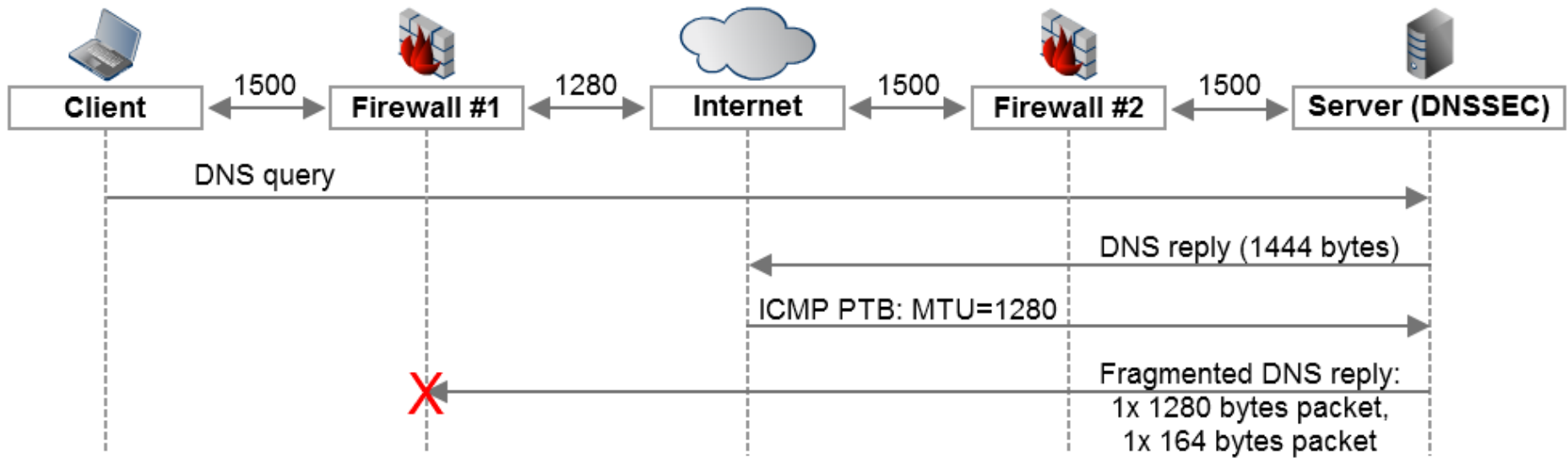
Theory

Problem #1: ICMP PTB filtering



Theory

Problem #2: fragment filtering



RIPE Atlas

- Internet measurement system
- Driven by probes
 - USB-powered embedded devices
- Default measurement functionality:
 - *ping*
 - *traceroute*
- Currently around 1700 probes up and running
 - Located primarily in the RIPE NCC service region
 - But also other regions around the globe



RIPE Atlas

Worldwide network of probes



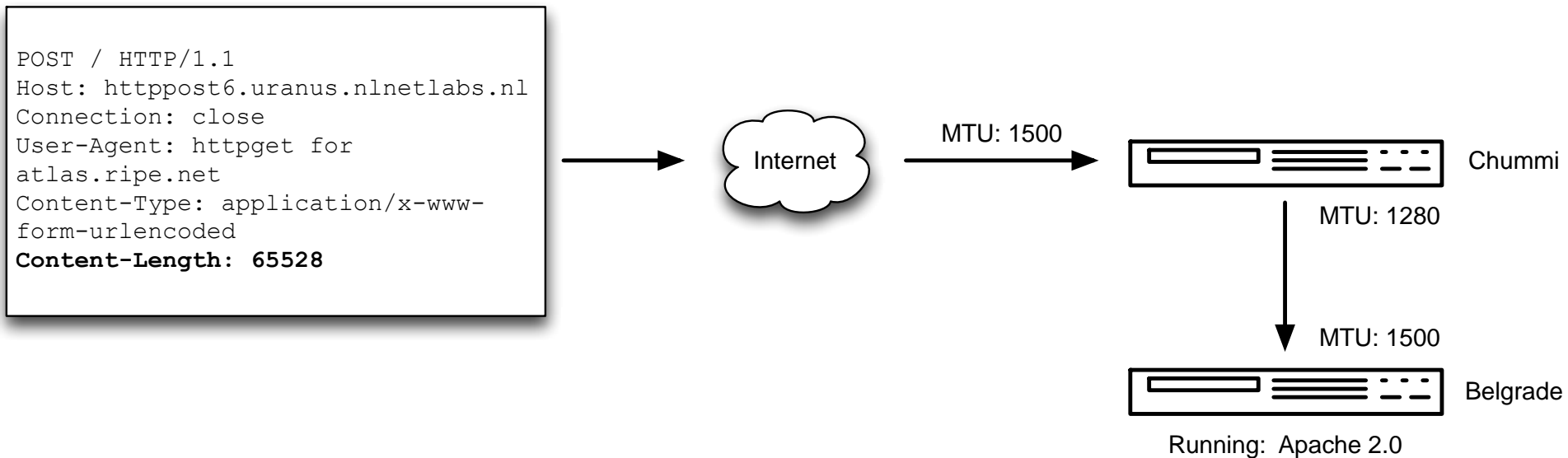
Research questions

Where on the Internet do Path MTU black holes occur?

Do Path MTU black holes occur more often in the IPv6-Internet compared to IPv4?

Experimental setup

ICMP PTB filtering



Experimental setup

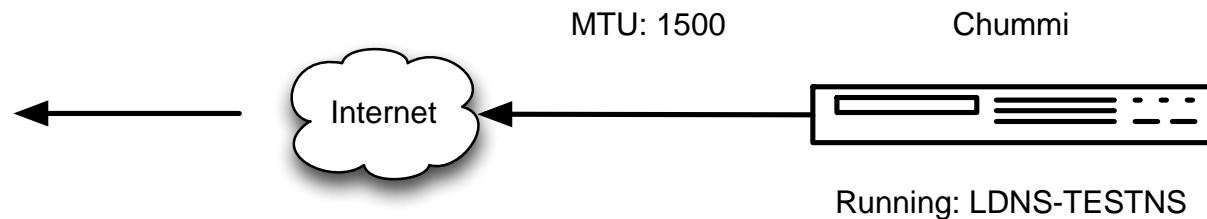
Fragment filtering

```
version.bind. 60 CH TXT
```

```
1,002,003,004,005,006,007,008,00  
9,010,011,012,013,014,015,016,01  
7,018,019,020,021,022,023,024,02  
5,026,  
33,334,335,336,337,338,339,340,3  
41,342,343,344,345,346
```

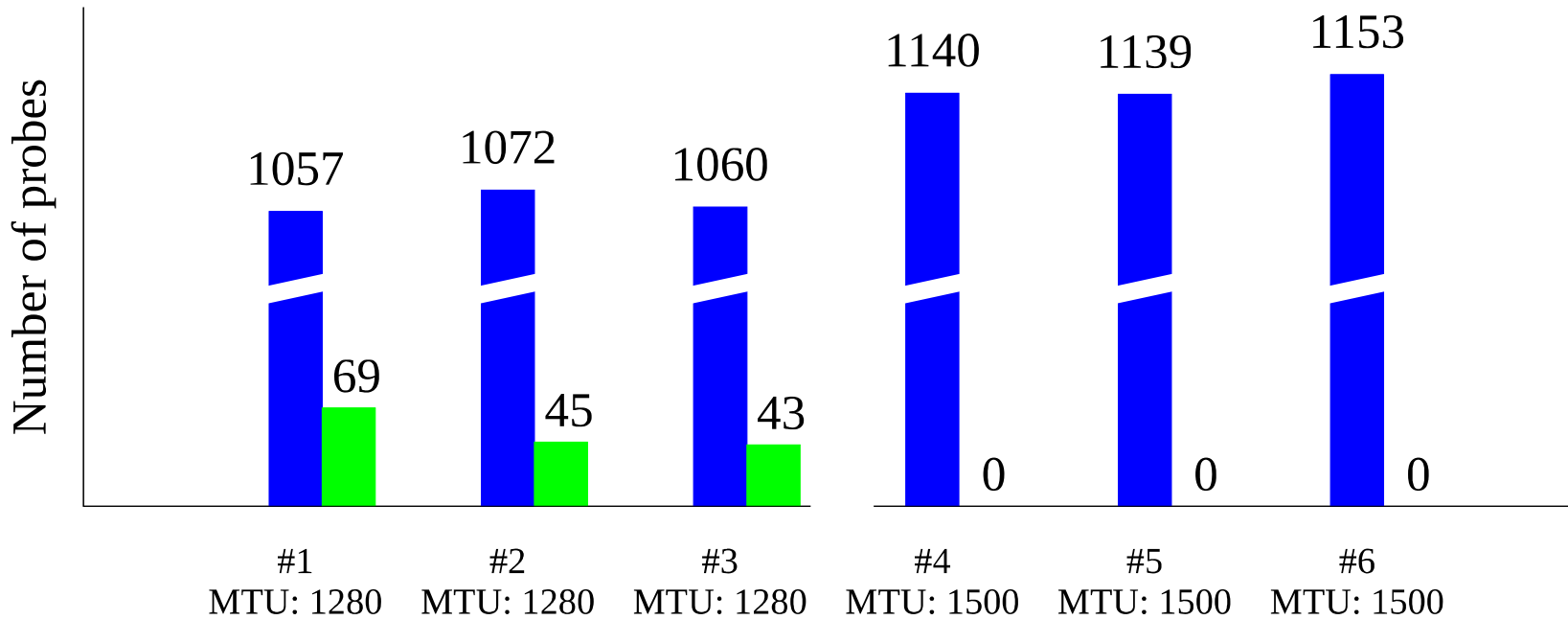
```
347,348,349,350,351,352,353,354,  
355,356,357,358,359,360,361,362,  
363,364,365,366,367,368,369,370,  
371,372,373,374,375,376,377,378,  
379,380,381,382,383
```

```
MSG SIZE snd: 1590
```



Results

ICMP PTB filtering IPv4

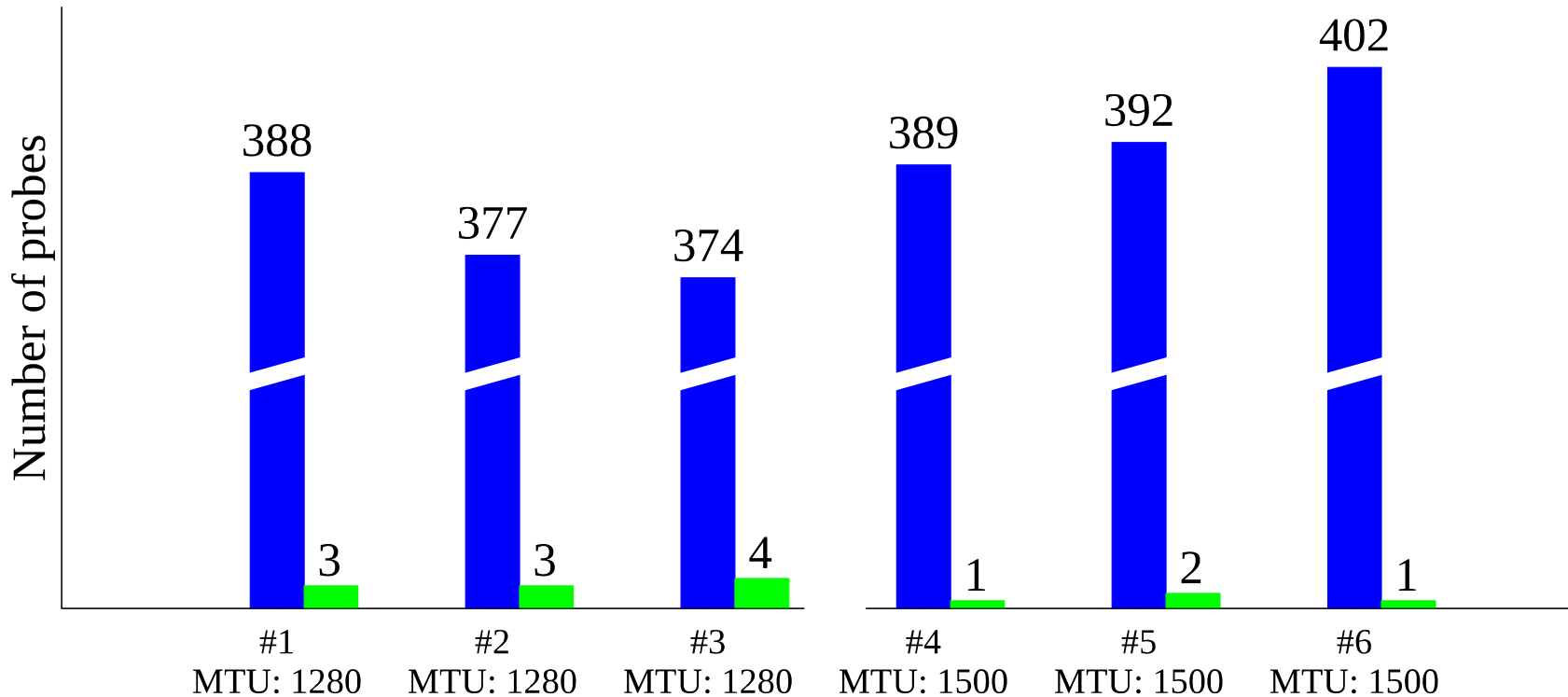


Experiment

- IPv4 ICMP OK
- IPv4 ICMP filtered

Results

ICMP PTB filtering IPv6

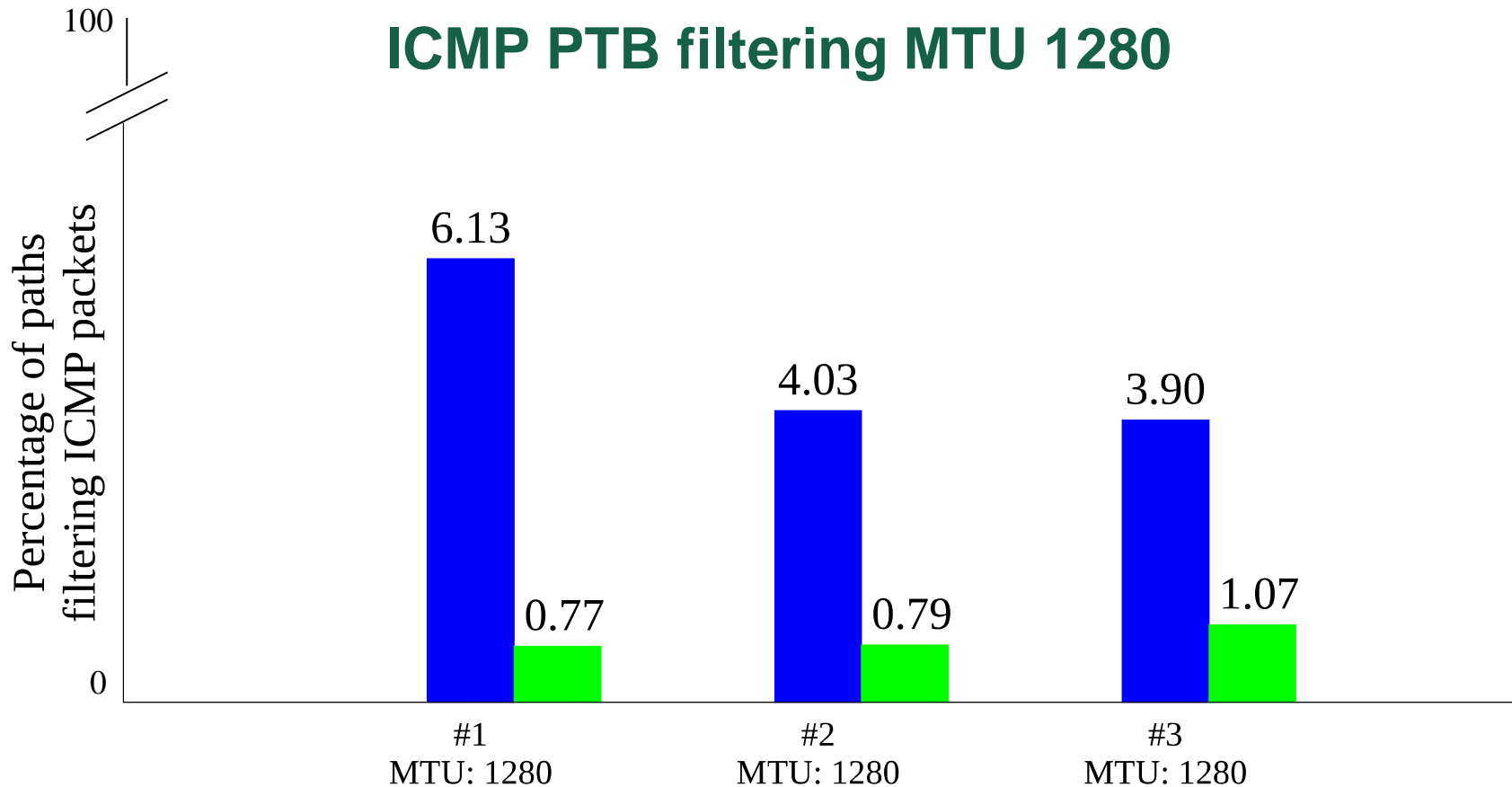


Experiment

■ IPv6 ICMP OK
■ IPv6 ICMP filtered

Results

ICMP PTB filtering MTU 1280

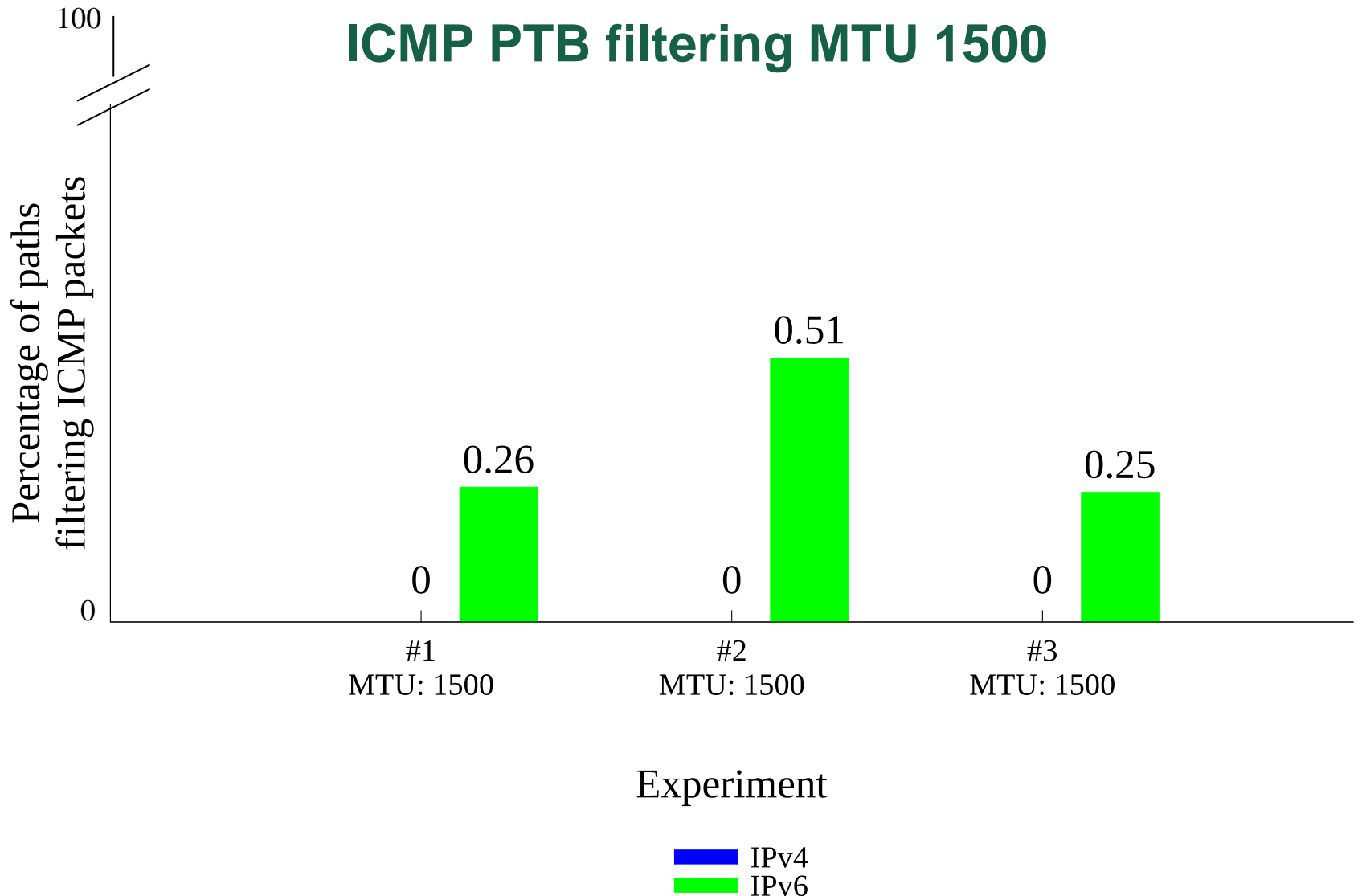


Experiment

■ IPv4
■ IPv6

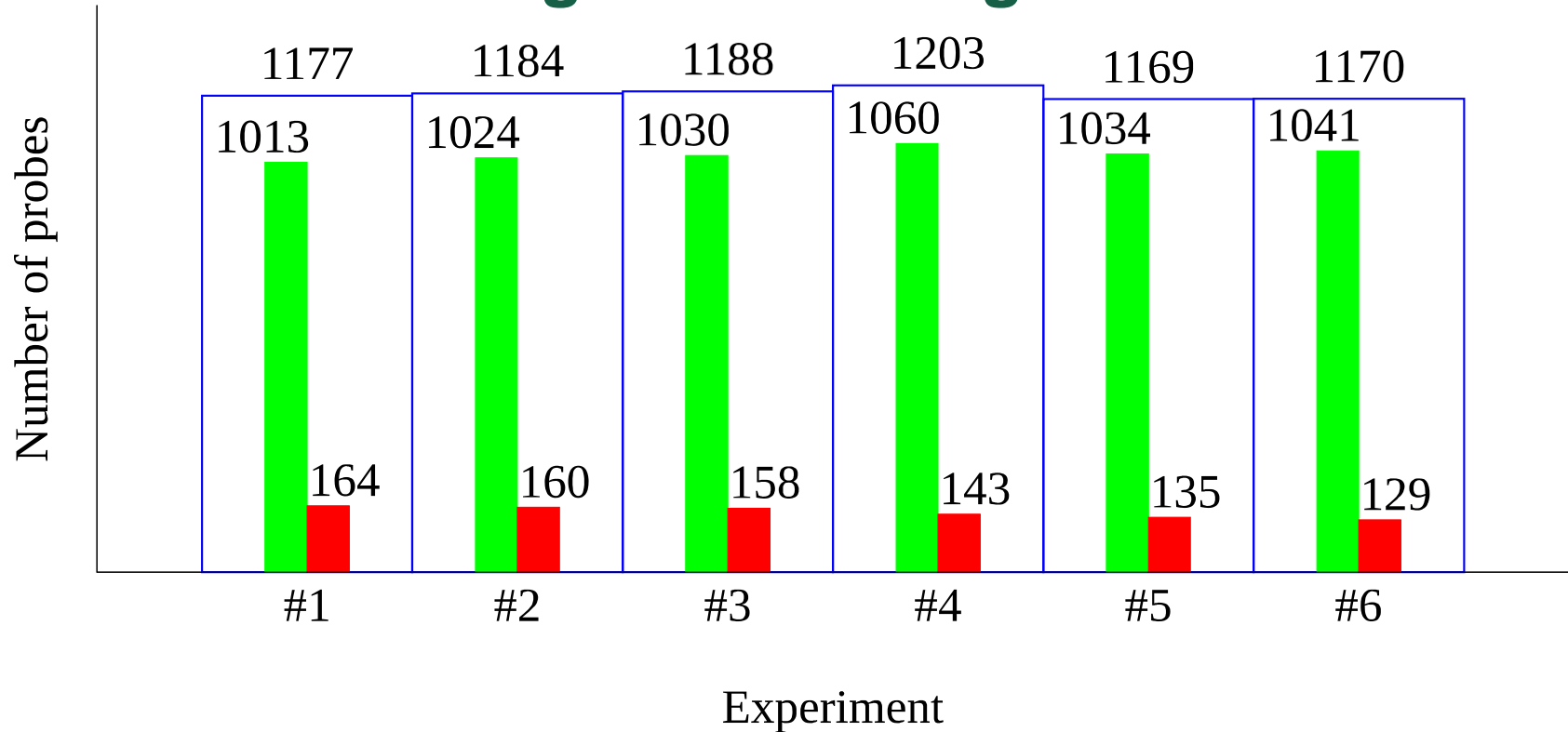
Results

ICMP PTB filtering MTU 1500



Results

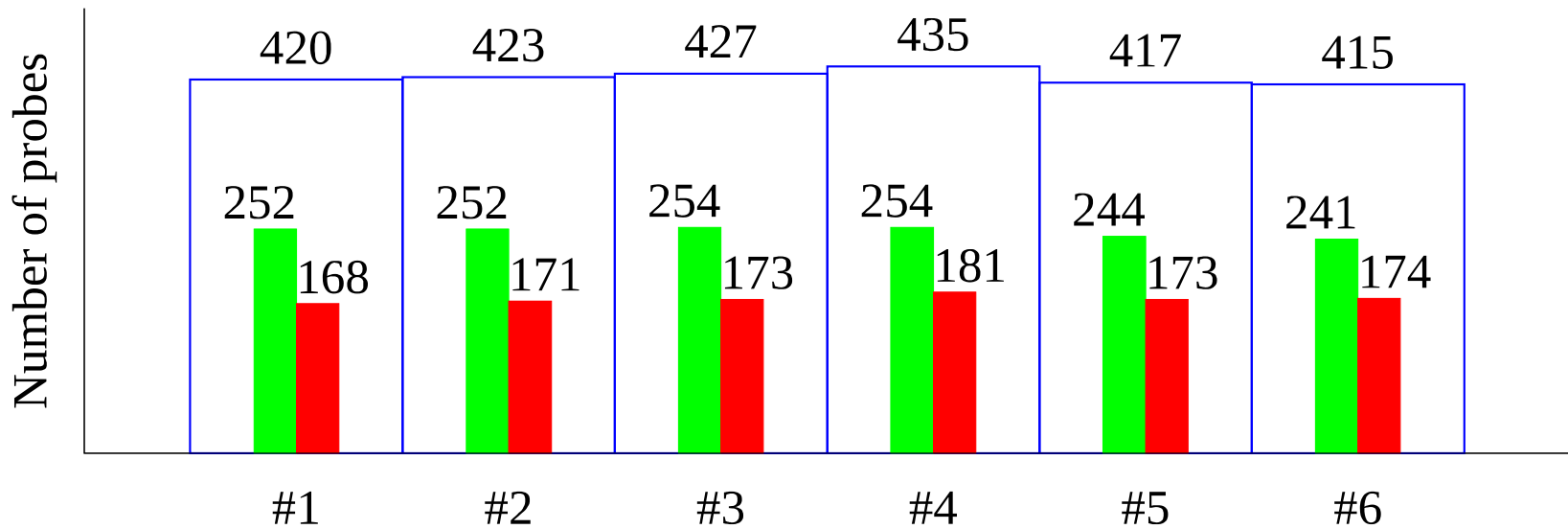
Fragment filtering IPv4



- Total
- Frag. OK
- Frag. filtered

Results

Fragment filtering IPv6



Experiment

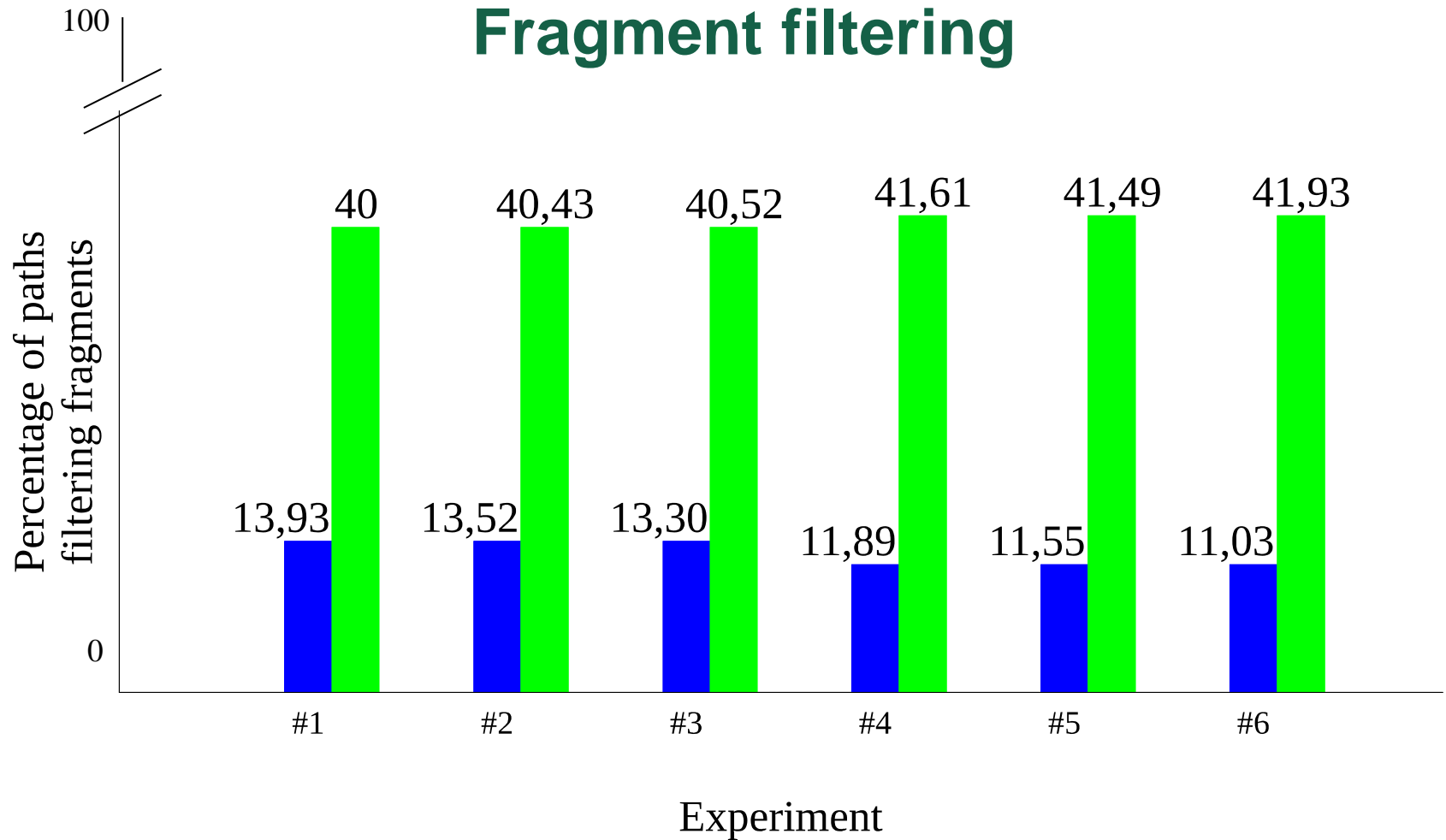
 Total

 Frags. OK

 Frags. filtered

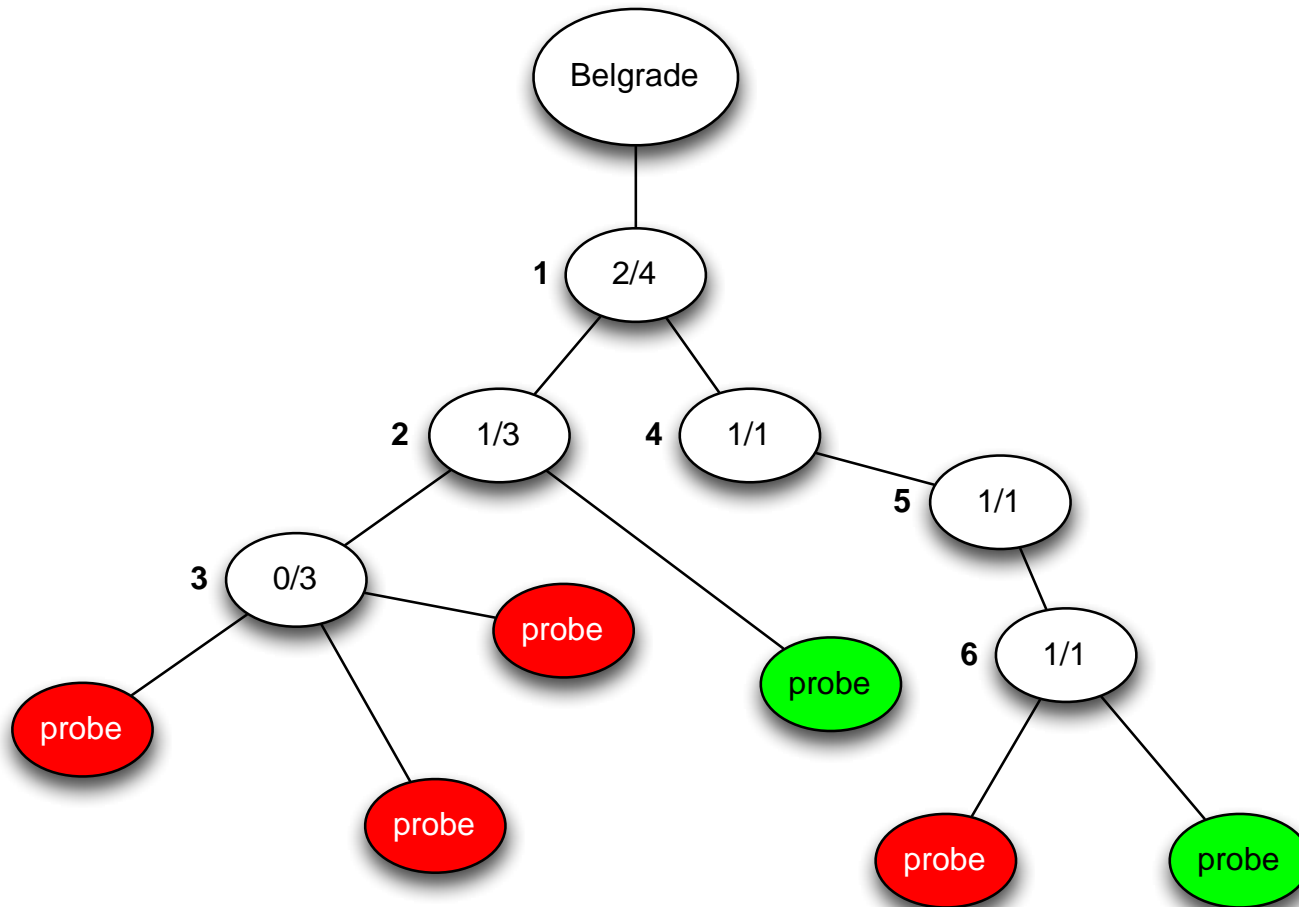
Results

Fragment filtering



■ IPv4
■ IPv6

Hop counting



Results

Where do IPv4 ICMP PTB messages get filtered?

Bad	Total	Error percentage	Ip
69	1126	6.1%	145.145.19.190
53	810	6.5%	145.145.80.65
16	311	5.1%	145.145.80.73
13	214	6.1%	77.67.72.109
7	199	3.5%	109.105.98.33
2	60	3.3%	62.40.124.157
...			
2	2	100.0%	203.50.6.78
2	2	100.0%	203.50.6.89
2	2	100.0%	61.10.0.118
2	2	100.0%	80.231.159.10
2	2	100.0%	84.116.238.49

Results

Where do IPv6 ICMP PTB messages get filtered?

Bad	Total	Error percentage	Ip
3	391	0.8%	2001:610:158:1916:145:100:99:17
2	292	0.7%	2001:610:e08:64::65
2	131	1.5%	2001:7f8:1::a500:6939:1
1	9	11.1%	2001:470:0:217::2
1	6	16.7%	2001:470:0:67::2
1	46	2.2%	2001:470:0:3f::1
...			
No routers with 100% failure rate			

Results

Where do IPv4 fragments get filtered?

Bad	Total	Error percentage	Ip
143	1203	11.9%	145.145.19.190
103	861	12.0%	145.145.80.65
40	337	11.9%	145.145.80.73
36	219	16.4%	77.67.72.109
23	226	10.2%	109.105.98.33
9	54	16.7%	62.40.124.157
...			
2	2	100.0%	212.188.29.138
2	2	100.0%	216.66.41.110
2	2	100.0%	46.19.96.235
2	2	100.0%	62.154.32.74
2	2	100.0%	80.241.177.86

Results

Where do IPv6 fragments get filtered?

Bad	Total	Error percentage	Ip
181	435	41.6%	2001:610:158:1916:145:100:99:17
138	322	42.9%	2001:610:e08:64::65
74	146	50.7%	2001:7f8:1::a500:6939:1
28	53	52.8%	2001:470:0:3f::1
27	91	29.7%	2001:610:e08:72::73
21	53	39.6%	2001:948:2:6::1
...			
6	6	100.0%	2001:610:f01:9012::14
4	4	100.0%	2001:16d8:aaaa:5::2
4	4	100.0%	2001:7f8:1::a503:9326:1
4	4	100.0%	2a01:348::10:0:1
4	4	100.0%	2a01:348::27:0:1

Conclusion

- ICMP PTB messages get dropped
 - More for IPv4 but nobody notices
 - But not that often (anymore)
- Fragments get dropped
 - More in IPv6
- Path MTU black holes
 - Occur on the edges of the Internet, not in the core

Recommendations

- Recommendations for Filtering ICMPv6 Messages in Firewalls – RFC4890
- Don't filter IPv4 ICMP type 3 code 4
- Packetization Layer Path MTU Discovery – RFC4821
- Don't filter fragments (problems for DNSSEC)
- Don't reduce MTU on interface
- No MSS clamping

Acknowledgements

- NLnet Labs
 - Benno Overeinder
 - Willem Toorop
- RIPE NCC
 - Philip Homburg
 - Andreas Strikos
 - Vesna Manojlovic
 - Emile Aben

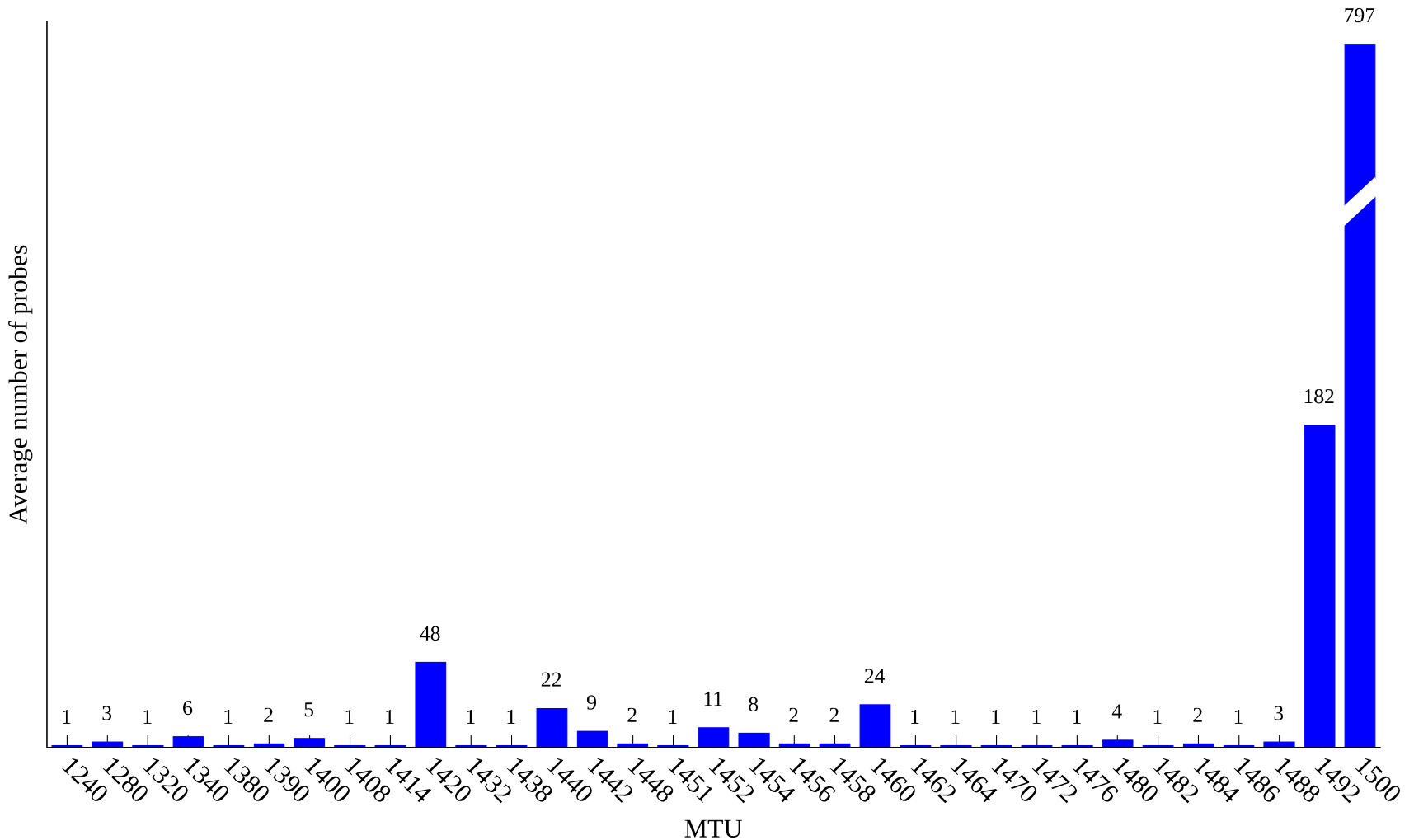
Questions?

Maikel de Boer – maikel.deboer@os3.nl

Jeffrey Bosma – jeffrey.bosma@os3.nl

Results

Path MTU determination IPv4



Results

Path MTU determination IPv6

