Searching for Evidence of Unallocated Address Space Usage in DITL 2008 Data

Duane Wessels
The Measurement Factory/CAIDA

OARC
June 4, 2008

Problem Statement

- From Leo Vegoda'a NANOG 42 presentation: *Analysis of PTR Queries for IPv4 Addresses Reserved for Future Allocation*
- All currently unallocated unicast space will eventually be allocated.
- Some networks and services already "secretly" use this space.
- Should IANA assign new /8's from the least secretly used space?

Leo's NANOG 42 Talk

- Counted PTR queries arriving at L.root-servers.net
- Top 10 Unallocated /8's:

Rank	/8
1.	2
2.	176
3.	1
4.	27
5.	107
6.	100
7.	23
8.	5
9.	46
10.	111

This Study

- Looking at DITL 2008 DNS Traces
- 48 hours
- Roots (8), Old-Roots (2), TLDs (5), RIRs (2), AS112's (6)
- 41 Unallocated /8's as of March 2008

What Do We Look For?

- Query from unused space

```
22:01:25.667048 IP 100.100.100.252.1047 > 192.5.5.241.53: 16873 A? yahoo.com. (27)|
```

- in-addr.arpa queries for unused space

```
22:00:21.327915 IP xxx.xx.x.xx.59822 > 192.5.5.241.53: 64 PTR? 43.88.184.100.in-addr.arpa. (44)
18:01:09.795632 IP xxx.xx.xxx.xx.44782 > 128.63.2.53.53: 8658 SOA? 2.in-addr.arpa. (32)
```

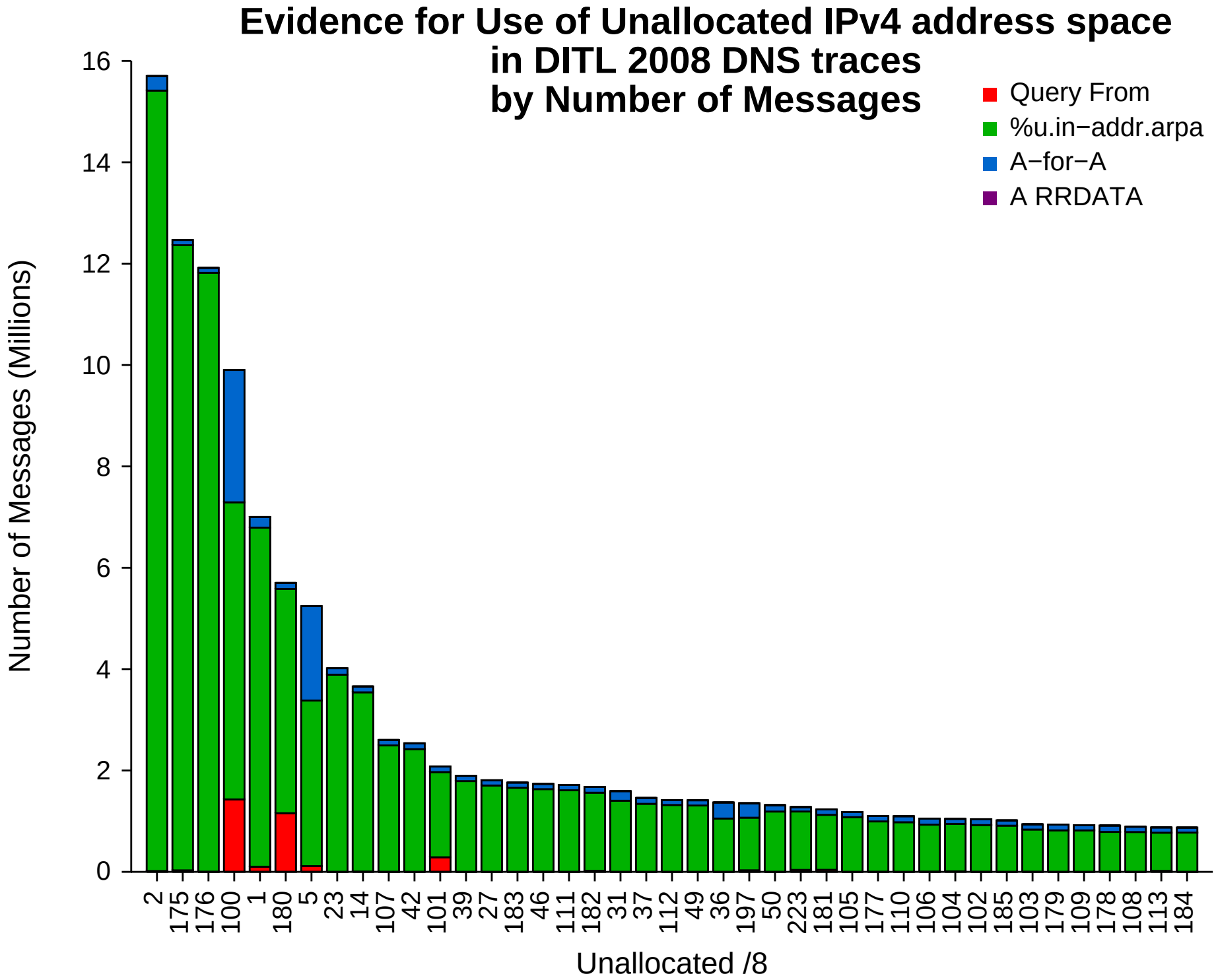
- A-for-A query

```
22:00:00.209799 IP xxx.xxx.xxx.xxx.40678 > 192.5.5.241.53: 5853% [1au] A? 100.100.131.192. (44)
```

- Addresses in A RRDATA (RRs in answer and additional sections)

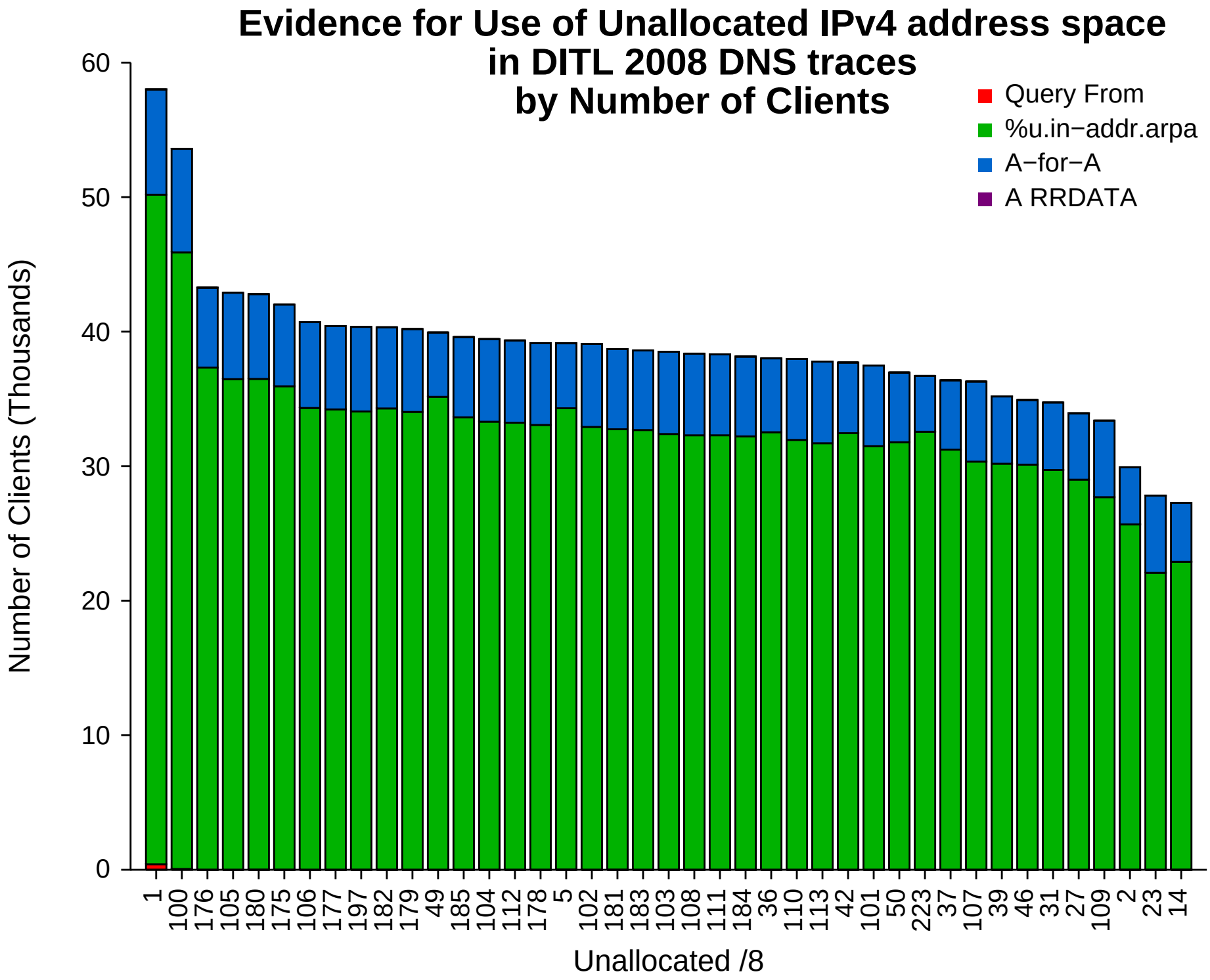
```
02:16:14.561809 IP xxx.xx.xxx.xxx.53 > 204.152.184.76.53: 34274*- q: A? www.microsoftliveupdates.com.
1/0/0 www.microsoftliveupdates.com. A 1.19.245.1 (62)
```

  - Very rare.
  - Most DITL 2008 DNS traces do not include replies.
  - Oops, note that 204.152.184.76 is *f.6to4-servers.net* which sits next to F-root's PAO1 node.


Evidence for Use of Unallocated IPv4 address space
in DITL 2008 DNS traces
by Number of Messages
Query From
%u.in−addr.arpa
A−for−A
A RRDATA
Number of Messages (Millions)
0
2
4
6
8
10
12
14
16
2
175
176
100
1
180
5
23
14
107
42
101
39
27
183
46
111
182
31
37
112
49
36
197
50
223
181
105
177
110
106
104
102
185
103
179
109
178
108
113
184
Unallocated /8

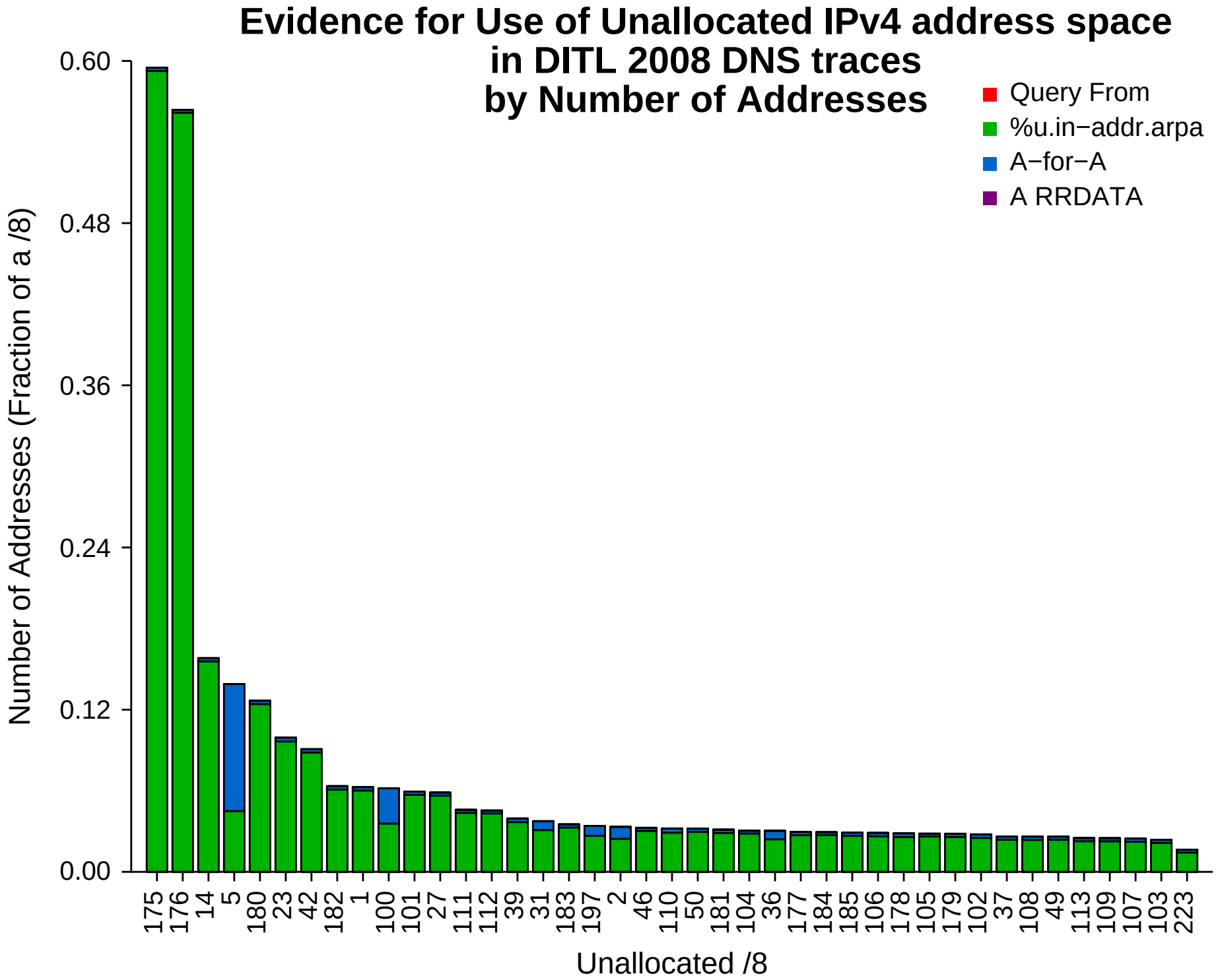
Number of Messages

- 2.0.0.0/8 is most popular, matching Leo's NANOG 42 data.
- This analysis could be skewed by "broken" DNS clients with very high query rates.
- So let's count the number of clients interested in unallocated space...


Evidence for Use of Unallocated IPv4 address space
in DITL 2008 DNS traces
by Number of Clients
Query From
%u.in−addr.arpa
A−for−A
A RRDATA
Number of Clients (Thousands)
0
10
20
30
40
50
60
1
100
176
105
180
175
106
177
197
182
179
49
185
104
112
178
5
102
181
183
103
108
111
184
36
110
113
42
101
50
223
37
107
39
46
31
27
109
2
23
14
Unallocated /8

Number of Clients

- Looks much different!
- 1.0.0.0/8 and 100.0.0.0/8 very popular in this analysis.
- But it doesn't tell us anything about *how many* addresses within those blocks might be in use.
- So let's count the number of unique addresses referenced from unallocated space...

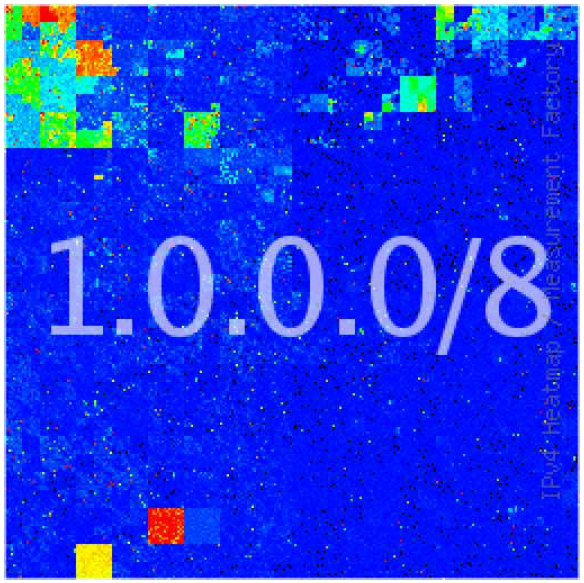

Evidence for Use of Unallocated IPv4 address space
in DITL 2008 DNS traces
by Number of Addresses
Query From
%u.in−addr.arpa
A−for−A
A RRDATA
Number of Addresses (Fraction of a /8)
0.60
0.48
0.36
0.24
0.12
0.00
175 176 14 5 180 23 42 182 1 100 101 27 111 112 39 31 183 197 2 46 110 50 181 104 36 177 184 185 106 178 105 179 102 37 108 49 113 109 107 103 223
Unallocated /8

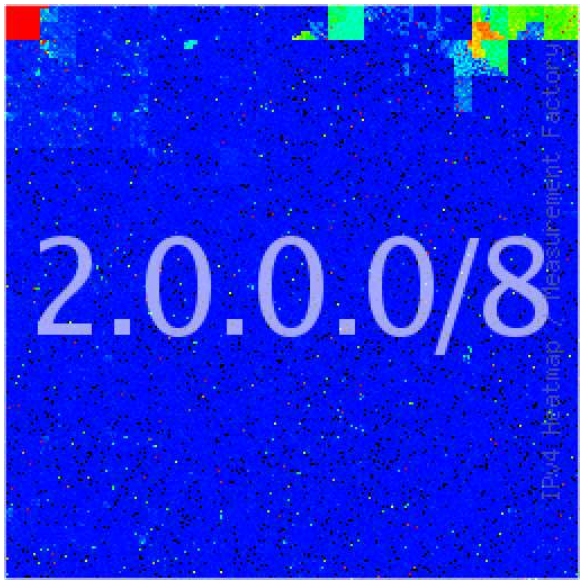
Number of Addresses

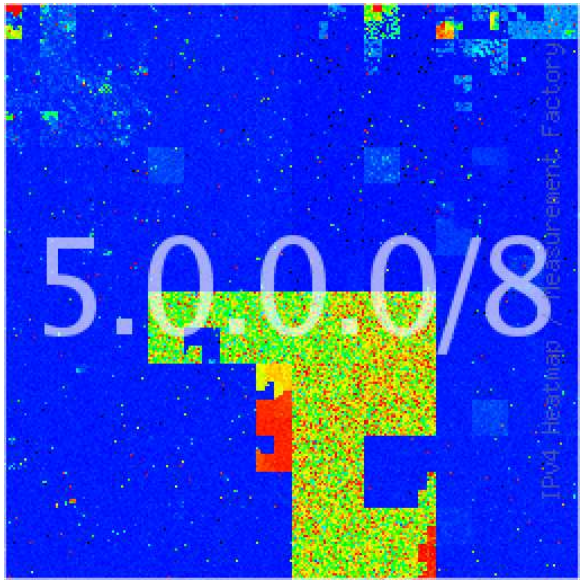
- Why are 175/8 and 176/8 so much higher?
- 14/8 – former Public Data Network block

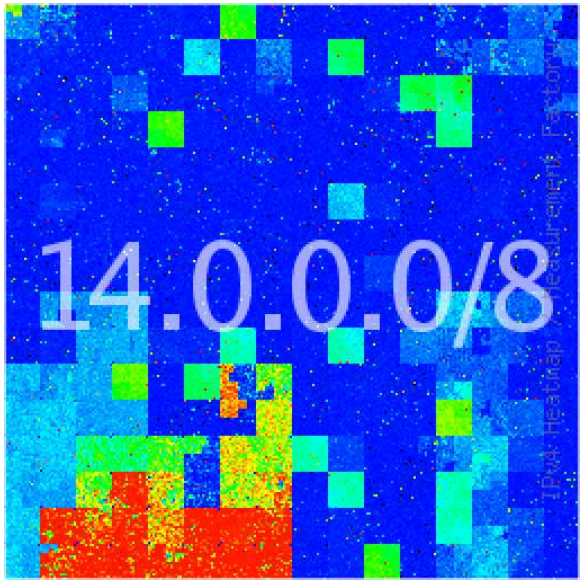
Hilbert Heatmaps

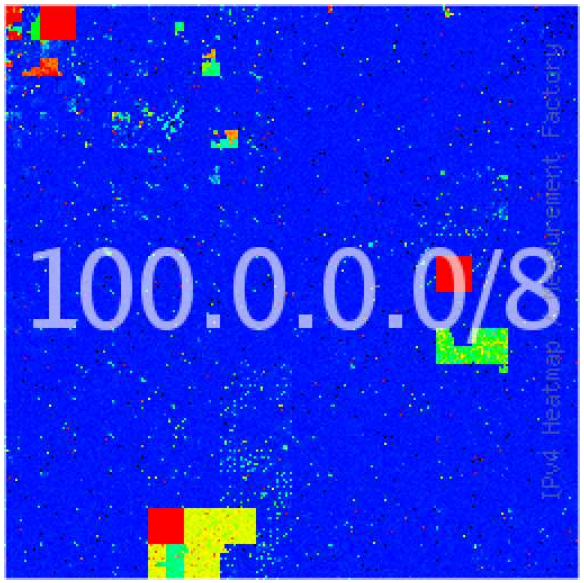
- Based on the "Number of Addresses" data
- Each pixel represents a /24
- Color represents number of addresses in the 24 where evidence of use was found.
  - Blue is low utilization ("background radiation")
  - Red is high utilization

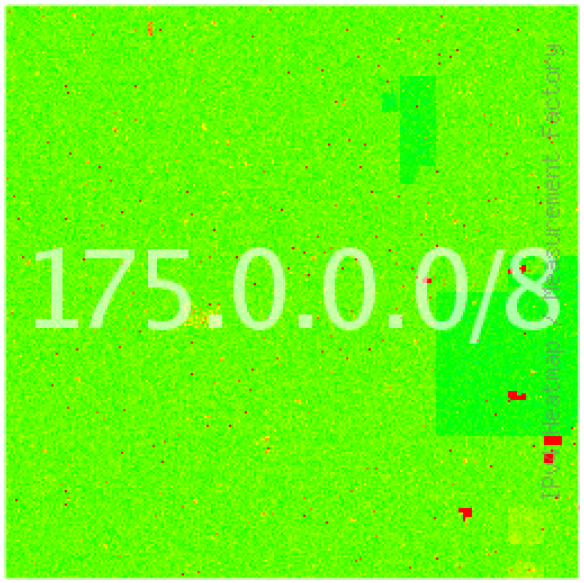

1.0.0.0/8
IPv4 Heatmap / Measurement Factory

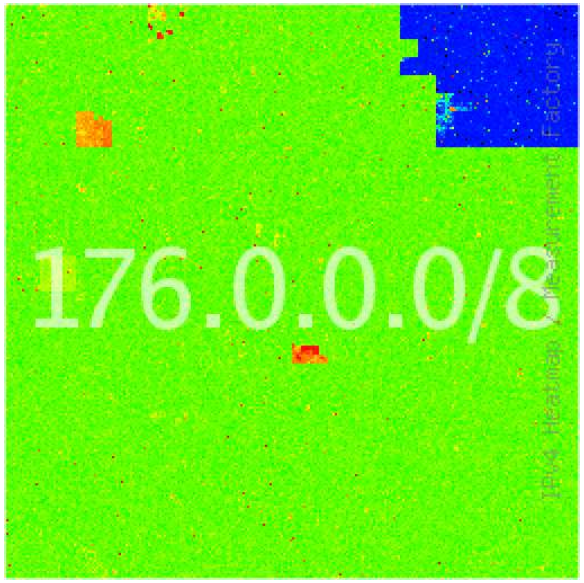

2.0.0.0/8
IPv4 Heatmap / Measurement Factory

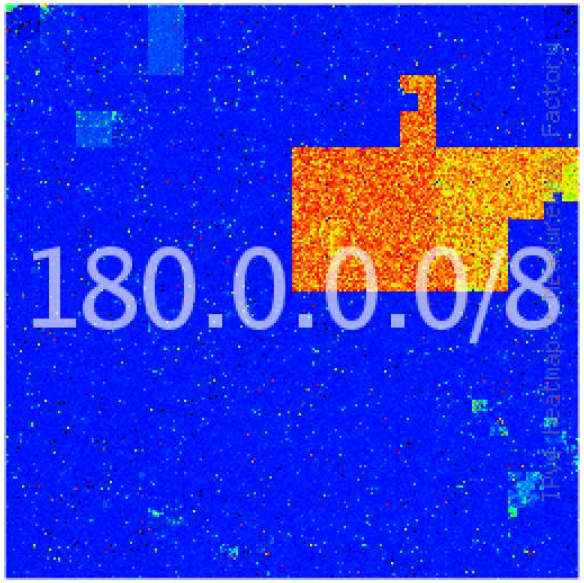

5.0.0.0/8
IPv4 Heatmap / Measurement Factory


14.0.0.0/8
IPv4 Heatmap / Measurement Factory



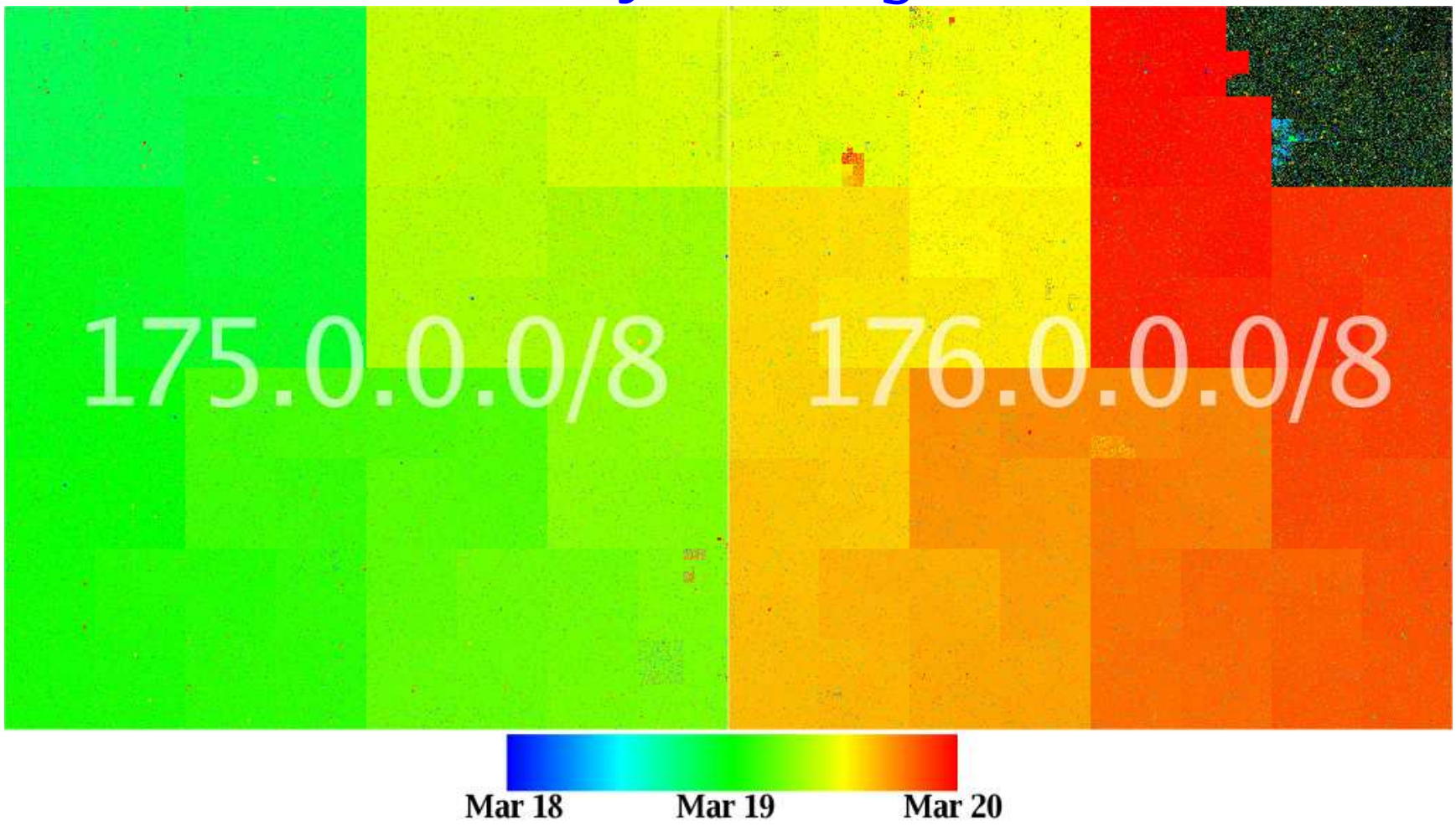






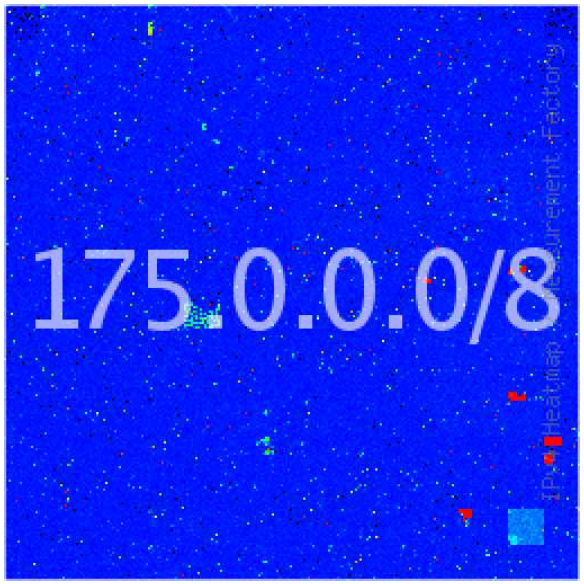
Note: green $\approx 60\% \approx \frac{8}{13}$.

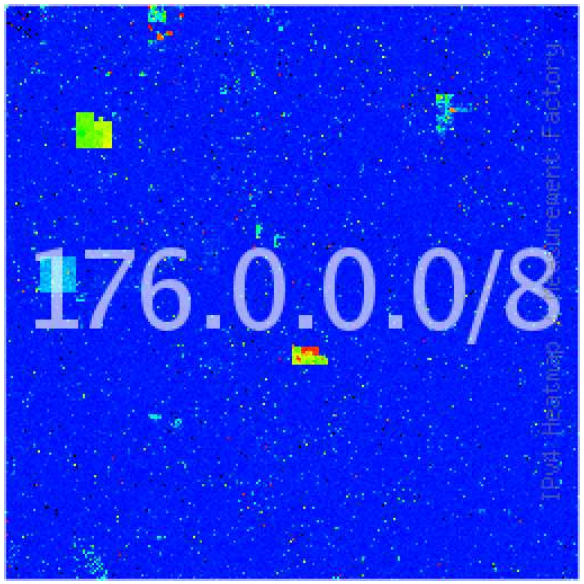
Colored by Message Time



- Scanning from a pair of nameservers on adjacent addresses in 205.209.x.x.

175 and 176 with Scanner Queries Removed





Executive Summary

- Tradeoffs from different ways of counting:
  - By number of messages: easy to count, but easily skewed by a handful of misbehaving sources.
  - By number of clients: eliminates biases from small number of busy sources, but doesn't measure extent of space used.
  - By number of addresses: easily biased by brute-force scanners.
- Most evidence comes from in-addr.arpa queries seen at root nameservers.
  - Queries from unused space are a stronger, but less-common indicator (if we assume they are not spoofed).
- Netblocks 1, 100, 175, 176, and 180 are in the top 10 for all three counting techniques.
- Netblocks 2, 14, and 23 are in the top 10 when counting by messages and by addresses, but are the bottom 3 when counting by clients.

The End