Temporal and Spatial Classification of Active IPv6 Addresses

David Plonka • Arthur Berger

Proc. of ACM IMC '15, Oct 2015.

Presented by Yibo Pi and Ryan Marcotte

IPv4 Address Exhaustion

IPv4 Address Report

This report generated at 01-Dec-2016 08:14 UTC.

IANA Unallocated Address Pool Exhaustion:
03-Feb-2011

Projected RIR Address Pool Exhaustion Dates:

RIR	Projected Exhaustion Date	Remaining Addresses in RIR Pool (/8s)
APNIC:	**19-Apr-2011** (actual)	0.4360
RIPE NCC:	**14-Sep-2012** (actual)	0.7997
LACNIC:	**10-Jun-2014** (actual)	0.0267
ARIN:	**24 Sep-2015** (actual)	
AFRINIC:	**27-Jun-2018**	1.3309

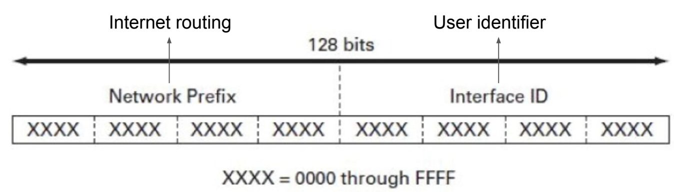
Regional Internet registry

African Network Information Center

Interesting Facts about IPv6

- You can use IPv6 to create 18,446,744,073,709,551,616 copies of the current Internet
- An IPv6 address has 128 bits
  - Forget about remembering a non-special IPv6 address

IPv6 Address Format



IPv4:
1. Static configuration
2. Dynamic Host Configuration Protocol (DHCP)

IPv6:
1. Static configuration
2. Dynamic Host Configuration Protocol (DHCP)
3. Stateless Address Autoconfiguration (SLAAC)

Stateless Address Autoconfiguration (SLAAC)

```
2001:db80:0000:0000:0000:0000:0000:0001  Gateway
0123:4567:89ab        48 bits            MAC
0123:45ff:fe67:89ab   48 + 16 bits = 64 bits   MAC after padding
2001:db80:0000:0000:0123:45ff:fe67:89ab  Final IP address
```

First 64 bits of gateway — Last 64 bits: padded MAC

SLAAC = network prefix + padded MAC address

Privacy extension to SLAAC

- MAC address is unique: users can be identified by their MAC addresses
- Privacy extension: randomly select IIDs
  - Question: two users select the same IID?
  - IPv6 Duplicate Address Detection (DAD) monitors if addresses duplicate
- Random IIDs deprecate typically after a day

Active IPv6 addresses

- Dataset: aggregated logs of servers (55000 of CDN's IPv6-capable servers)

Characteristic	Mar 17, 2014	Sep 17, 2014	Mar 17, 2015
Teredo addresses	1.98K (0.00%)	3.28K (0.00%)	20.1K (0.01%)
ISATAP addresses	90.2K (0.06%)	101K (0.04%)	133K (0.04%)
6to4 addresses	12.8M (7.97%)	12.5M (5.90%)	13.9M (4.19%)
Other addresses	**149M (92.0%)**	**199M (94.1%)**	**318M (95.8%)**
Other /64 prefixes	**61.4M**	**82.9M**	**121M**
ave. addrs per /64	2.41	2.40	2.63
EUI-64 addr (!6to4)	3.13M (1.94%)	3.66M (1.73%)	4.49M (1.35%)
EUI-64 IIDs (MACs)	2.85M	3.23M	3.81M

(a) Address characteristics per day

Characteristic	Mar 17-23, 2014	Sep 17-23, 2014	Mar 17-23, 2015
Teredo addresses	15.1K (0.00%)	24.5K (0.00%)	131K (0.01%)
ISATAP addresses	210K (0.02%)	238K (0.02%)	346K (0.02%)
6to4 addresses	64.9M (7.22%)	78.3M (6.34%)	64.2M (3.43%)
Other addresses	**833M (92.8%)**	**1.17B (94.9%)**	**1.80B (96.5%)**
Other /64 prefixes	**157M**	**207M**	**307M**
ave. addrs per /64	5.32	5.64	5.88
EUI-64 addr (!6to4)	8.88M (0.99%)	13.1M (1.06%)	16.2M (0.866%)
EUI-64 IIDs (MACs)	6.12M	8.16M	9.74M

(b) Address characteristics per week

Table 1: Active IPv6 WWW client address characteristics: March 2014 through March 2015.

Teredo: full connectivity for IPv6-capable hosts having no native connections
ISATAP: transmit IPv6 packets between dual-stack nodes
6to4: allows IPv6 packets to be transmitted over an IPv4 network w/o tunnels

Not native IPv6 (not all infrastructure supports IPv6)

Active IPv6 addresses

- Dataset: aggregated logs of servers (55000 of CDN's IPv6-capable servers)

Characteristic	Mar 17, 2014	Sep 17, 2014	Mar 17, 2015
Teredo addresses	1.98K (0.00%)	3.28K (0.00%)	20.1K (0.01%)
ISATAP addresses	90.2K (0.06%)	101K (0.04%)	133K (0.04%)
6to4 addresses	12.8M (7.97%)	12.5M (5.90%)	13.9M (4.19%)
Other addresses	**149M (92.0%)**	**199M (94.1%)**	**318M (95.8%)**
Other /64 prefixes	**61.4M**	**82.9M**	**121M**
ave. addrs per /64	2.41	2.40	2.63
EUI-64 addr (!6to4)	3.13M (1.94%)	3.66M (1.73%)	4.49M (1.35%)
EUI-64 IIDs (MACs)	2.85M	3.23M	3.81M

(a) Address characteristics per day

Characteristic	Mar 17-23, 2014	Sep 17-23, 2014	Mar 17-23, 2015
Teredo addresses	15.1K (0.00%)	24.5K (0.00%)	131K (0.01%)
ISATAP addresses	210K (0.02%)	238K (0.02%)	346K (0.02%)
6to4 addresses	64.9M (7.22%)	78.3M (6.34%)	64.2M (3.43%)
Other addresses	**833M (92.8%)**	**1.17B (94.9%)**	**1.80B (96.5%)**
Other /64 prefixes	**157M**	**207M**	**307M**
ave. addrs per /64	5.32	5.64	5.88
EUI-64 addr (!6to4)	8.88M (0.99%)	13.1M (1.06%)	16.2M (0.866%)
EUI-64 IIDs (MACs)	6.12M	8.16M	9.74M

(b) Address characteristics per week

Table 1: Active IPv6 WWW client address characteristics: March 2014 through March 2015.

Teredo: full connectivity for IPv6-capable hosts having no native connections
ISATAP: transmit IPv6 packets between dual-stack nodes
6to4: allows IPv6 packets to be transmitted over an IPv4 network w/o tunnels

Not native IPv6 (not all infrastructure supports IPv6)

Active IPv6 addresses

- Dataset: aggregated logs of servers (55000 of CDN's IPv6-capable servers)

Characteristic	Mar 17, 2014	Sep 17, 2014	Mar 17, 2015
Teredo addresses	1.98K (0.00%)	3.28K (0.00%)	20.1K (0.01%)
ISATAP addresses	90.2K (0.06%)	101K (0.04%)	133K (0.04%)
6to4 addresses	12.8M (7.97%)	12.5M (5.90%)	13.9M (4.19%)
Other addresses	**149M (92.0%)**	**199M (94.1%)**	**318M (95.8%)**
Other /64 prefixes	**61.4M**	**82.9M**	**121M**
ave. addrs per /64	2.41	2.40	2.63
EUI-64 addr (!6to4)	3.13M (1.94%)	3.66M (1.73%)	4.49M (1.35%)
EUI-64 IIDs (MACs)	2.85M	3.23M	3.81M

(a) Address characteristics per day

Characteristic	Mar 17-23, 2014	Sep 17-23, 2014	Mar 17-23, 2015
Teredo addresses	15.1K (0.00%)	24.5K (0.00%)	131K (0.01%)
ISATAP addresses	210K (0.02%)	238K (0.02%)	346K (0.02%)
6to4 addresses	64.9M (7.22%)	78.3M (6.34%)	64.2M (3.43%)
Other addresses	**833M (92.8%)**	**1.17B (94.9%)**	**1.80B (96.5%)**
Other /64 prefixes	**157M**	**207M**	**307M**
ave. addrs per /64	5.32	5.64	5.88
EUI-64 addr (!6to4)	8.88M (0.99%)	13.1M (1.06%)	16.2M (0.866%)
EUI-64 IIDs (MACs)	6.12M	8.16M	9.74M

(b) Address characteristics per week

Table 1: Active IPv6 WWW client address characteristics: March 2014 through March 2015.

EUI-64 IID: 64-bit interface ID (padded MAC address)

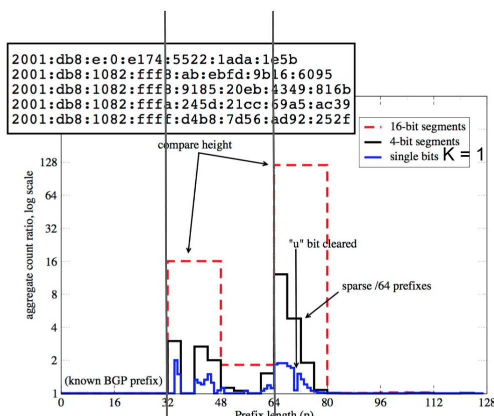
Classifying IPv6 Addresses

- Spatial classification
  - Multi-Resolution Aggregate (MRA) Count Ratio
  - Prefix Density
- MRA count ratio
  - Suppose we have N addresses
  - We need n_p prefixes of /p size to contain all N address
  - Count ratio $y_p = n_{p+k}/n_p$ (k = 1, 4, 16)

Example: k = 1
2001:db8:e:0:e174:5522:1ada:1
2001:db8:e:0:e174:5522:1ada:0
We need two /128 or one /127 or one /126 prefix to contain these two addresses
$y_{127} = n_{128}/n_{127} = 2$

$y_{126} = n_{127}/n_{126} = 1$

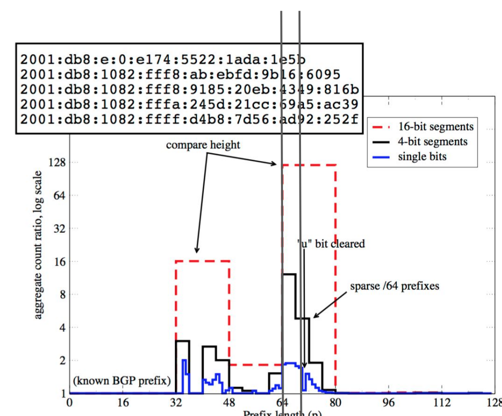
MRA plot



(a) US university

Subnetting

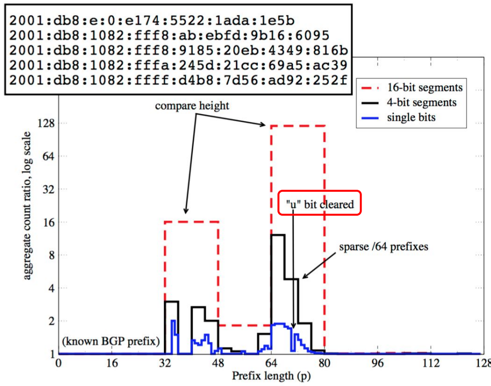
MRA plot



(a) US university

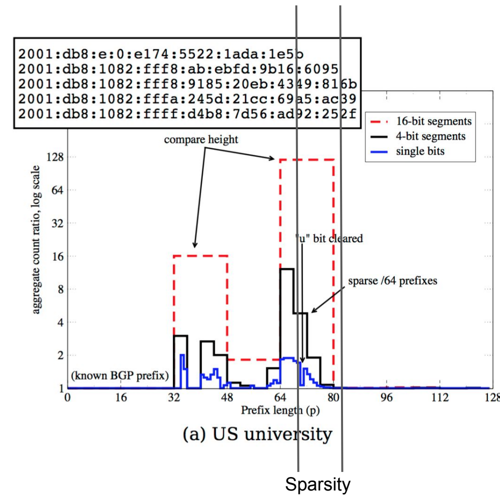
Privacy extension

MRA plot

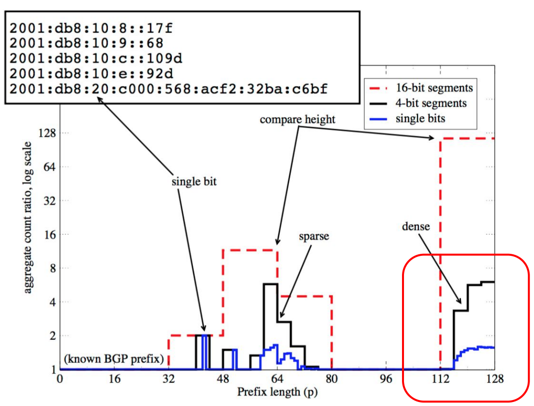


(a) US university

MRA plot



MRA plot



(b) JP telco

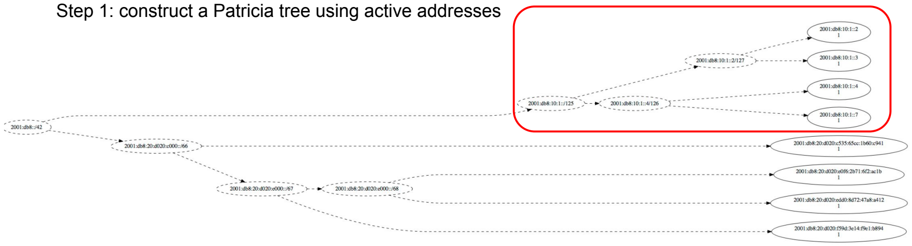
Figure 2: Sample MRA plots for active IPv6 client addresses (a) a university prefix with 7.22K addresses and (b) a telco prefix with 12.8K addresses.

Classifying IPv6 Addresses

- Spatial classification
  - Multi-Resolution Aggregate (MRA) Count Ratio
  - Prefix Density
- Prefix density
  - Two parameters
    - n: the number of addresses to be dense
    - p: prefix length
  - ‘n@p-dense’: the class of prefixes of length p that contain at least n addresses

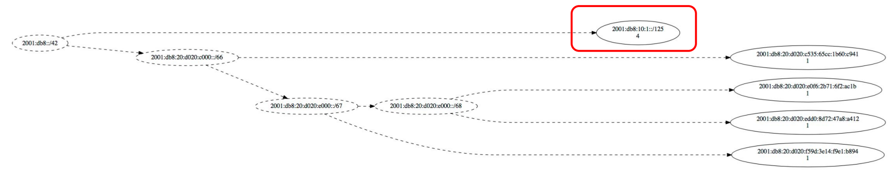
Computing dense prefixes

Step 1: construct a Patricia tree using active addresses



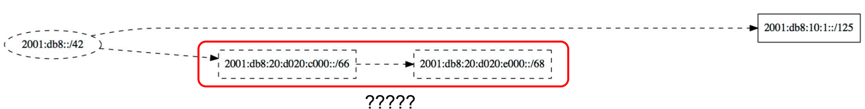
Computing dense prefixes

Step 2: aggregate nodes that satisfy the density requirement



Computing dense prefixes

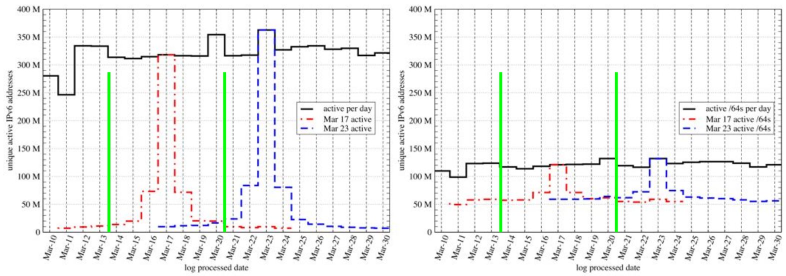
Step 3: output



Temporal Classification

- Goal
  - Determine address lifetime
  - Separate addresses that are persistent or stable from these not
- “6m-stable”: if x is observed at some time and also 6 months earlier
- “nd-stable”: there exists observations of activity on two days that are (n-1) days apart
- In daily stability analysis, a sliding 15-day window entered at the day is used

Address and Prefix Stability



(a) IPv6 address stability (b) /64 prefix stability

30.1M 3d-stable (9.44%) 109M 3d-stable (89.8%)

Summary of Address/Prefix Stability

addr class	Mar 17, 2014	Sep 17, 2014	Mar 17, 2015
3d-stable	13.7M (9.22%)	13.6M (6.84%)	30.1M (9.44%)
not 3d-stable	134M (90.8%)	185M (93.2%)	288M (90.6%)
6m-stable (-6m)		588K (.296%)	1.08M (.340%)
1y-stable (-1y)			328K (.103%)

(a) Stability of IPv6 addresses per day

/64 class	Mar 17, 2014	Sep 17, 2014	Mar 17, 2015
3d-stable	55.8M (91.0%)	74.6M (89.9%)	**109M (89.8%)**
not 3d-stable	5.53M (9.01%)	8.33M (10.1%)	12.3M (10.2%)
6m-stable (-6m)		23.4M (28.2%)	32.4M (26.7%)
1y-stable (-1y)			21.8M (18.0%)

(b) Stability of /64 prefixes per day

addr class	Mar 17-23, 2014	Sep 17-23, 2014	Mar 17-23, 2015
3d-stable	37.0M (4.44%)	34.0M (2.91%)	69.0M (3.82%)
not 3d-stable	796M (95.6%)	1.13B (97.1%)	1.74B (96.2%)
6m-stable (-6m)		3.25M (.280%)	3.66M (.202%)
1y-stable (-1y)			**1.81M (.100%)**

(c) Stability of IPv6 addr... per week

/64 class	Mar 17-23, 2014	Sep 17-23, 2014	Mar 17-23, 2015
3d-stable	131M (83.7%)	169M (81.8%)	**246M (80.3%)**
not 3d-stable	25.5M (16.3%)	37.7M (18.2%)	60.6M (19.7%)
6m-stable (-6m)		120M (58.1%)	**153M (49.9%)**
1y-stable (-1y)			**116M (37.8%)**

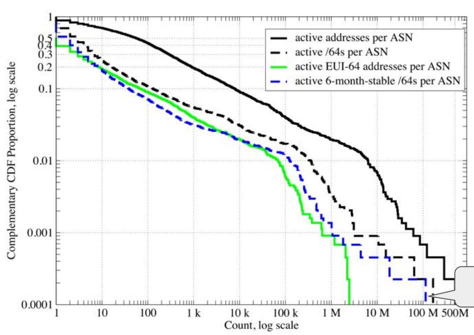
(d) Stability of /64 prefixes per week

Relatively few very long-lived client addresses

...WW client address and prefix counts, not 6to4 or Teredo

Many long-lived /64 prefixes

Address/Prefix Counts by ASN



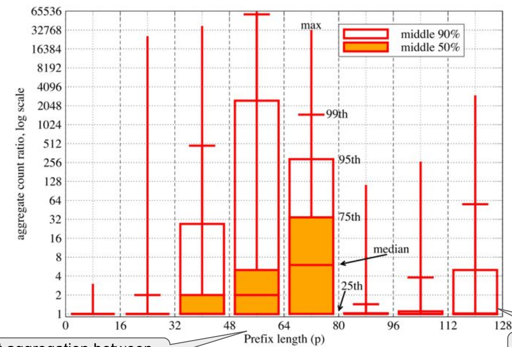
One ASN accounts for >100M /64s across 6 mo

Most long-lived /64s are in only a few networks!

Address/Prefix Stability Discussion

- Stable addresses make reliable targets for active probing
  - Randomly selected 3.18M 3d-stable addresses
  - Performed 52.6M traces to them from 20 source locations
  - Received response from 10.9% of "stable" targets compared to 0.3% of randomly selected "active" targets
- "Not 3d-stable" addresses
  - Vast majority are expected to be hosts using privacy-extension IID
  - Some EUI-64 addresses have same IID but different subnet prefix (e.g. device moved between networks)
- Future work: what temporal classes are most useful?
- Active/stable /64s as approximate lower-bound on subscribers or instances of IPv6-capable Internet connections?
  - 116M 1y-stable /64s observed March 2015

Distribution of Aggregation Ratios

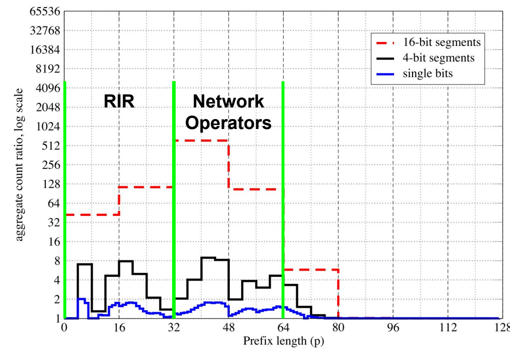


This data helps point in what address populations we should explore MRA plots

Most aggregation between bits 32 and 80

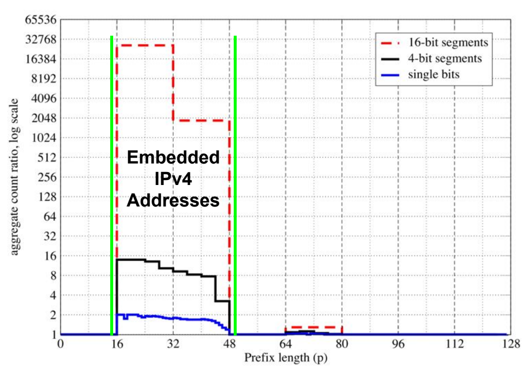
Significant aggregation in 112-128 bit segment

MRA plot - “30,000-foot view”

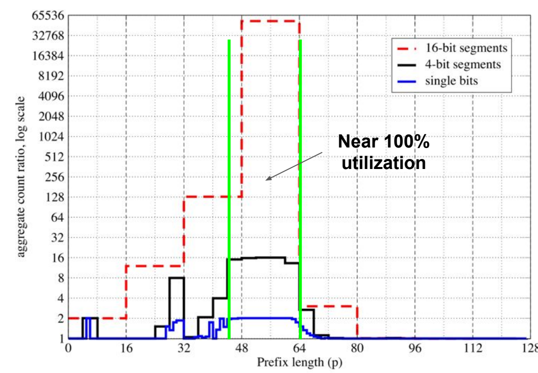


ALL active WWW client addresses observed in IPv6 unicast address space

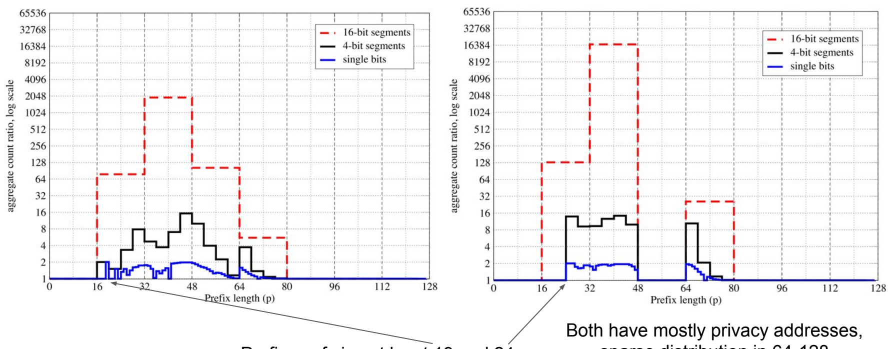
MRA plot - 6to4 Clients



MRA plot - U.S.-based Mobile Carrier



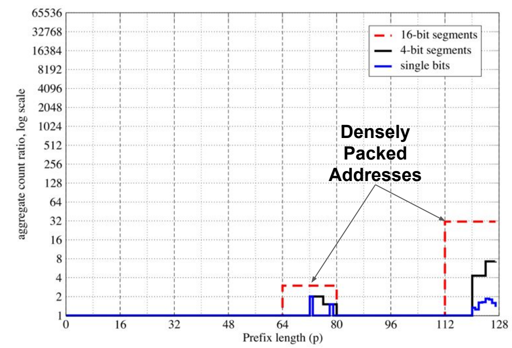
MRA plot - European ISP vs. Japanese ISP



Prefixes of size at least 19 and 24

Both have mostly privacy addresses, sparse distribution in 64-128

MRA plot - European University



Discussion

- Limitation
  - Experiments under fixed parameters
    - For example, single sliding window size or 3d-stable
  - Temporal classification does not provide enough insights
  - Temporal stability is not well-defined (may be restricted by the dataset they have)
- Extension
  - Study more attributes (geolocation, regions) of IPv6 addresses
  - Analyze the path performance (e.g., circuitous routing) of using IPv6 addresses
  - Study the performance of redirection on clients with IPv6 addresses in CDNs