Clusters in the Expanse: De-Aliasing IPv6 Hitlists

Quirin Scheitle

July 19, 2018
IRTF Measurement and Analysis for Protocols Research Group (maprg)
IETF 102, Montreal

Chair of Network Architectures and Services
Department of Informatics
Technical University of Munich

Publications

This presentation is based on the following publications:

Fingerprinting Methodology:

> Large-Scale Classification of IPv6-IPv4 Siblings with Variable Clock Skew
> *Quirin Scheitle, Oliver Gasser, Minoo Rouhi, Georg Carle*
> Network Traffic Measurement and Analysis Conference (TMA), Dublin, Jun. 2017

Application to IPv6 Scanning:

> Clusters in the Expanse: Understanding and Unbiasing IPv6 Hitlists
> *Oliver Gasser, Quirin Scheitle, Pawel Foremski, Qasim Lone, Maciej Korczynski, Stephen D. Strowes, Luuk Hendriks, Georg Carle*
> arXiv:1806.01633, June 5th, 2018

Introduction

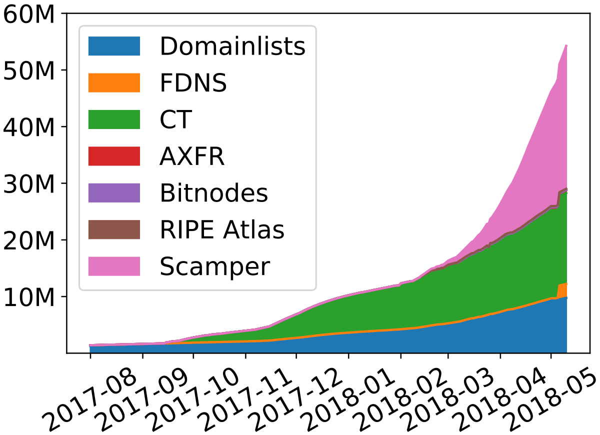
- Vast IPv6 space → Hitlists
- Approaches: Address Collection [1,2] & Generation [3,4]
- Biases towards some ASes and prefixes?

Single hosts can respond to entire IPv6 prefixes, which possibly adds vast clusters of responsive and valid IP addresses into a hit list.

Such IP addresses are called *aliases*, and prefixes containing aliased IP addresses can be called *aliased prefixes*.

[1] Oliver Gasser, Quirin Scheitle, Sebastian Gebhard, Georg Carle, "Scanning the IPv6 Internet: Towards a Comprehensive Hitlist, TMA'16
[2] Robert Beverly, Ramakrishnan Durairajan, David Plonka, Justin P. Rohrer, " In the IP of the Beholder: Strategies for Active IPv6 Topology Discovery, arXiv:1805.11308, 2018
[3] P. Foremski, D. Plonka, A. Berger , " In the IP of the Beholder: Entropy/IP: Uncovering Structure in IPv6 Addresses", IMC'16
[4] Austin Murdock, Frank Li, Paul Bramsen, Zakir Durumeric, Vern Paxson, " Target Generation for Internet-wide IPv6 Scanning", IMC'17

Brief Intro: Hitlist Sources and Growth over Time



- Many addresses from domainlists and CT
- Rapid increase of scamper addresses due to CPE routers

Multi-Level Aliased Prefix Detection

How to detect aliased prefixes?

State-of-the-art: Probe random (or fixed) addresses in prefixes suspected aliased [1,2]. Limitations:

- Requires only a subset of addresses to respond
- Typically conducted at a specific, fixed, prefix length
- Random address: targets may cluster as result of random process
- Fixed addresses (such as ...1111:1111:1111:) are predictable

Our approach:

- Send 16 well-spread probes, and require responses from **all** addresses
- Work at all levels of the prefix tree, and send probes on ICMP and TCP80

2001:0db8:0407:8000::/64
2001:0db8:0407:8000:0151:2900:77e9:03a8
2001:0db8:0407:8000:181c:4fcb:8ca8:7c64
2001:0db8:0407:8000:23d1:5e8e:3453:8268
⋮
2001:0db8:0407:8000:f693:2443:915e:1d2e

[1] T. Fiebig, K. Borgolte, S. Hao, C. Kruegel, and G. Vigna. Something from Nothing (There): Collecting Global IPv6 Datasets from DNS. PAM'17
[2] A. Murdock, F. Li, P. Bramsen, Z. Durumeric, and V. Paxson. Target Generation for Internet-wide IPv6 Scanning. IMC'17

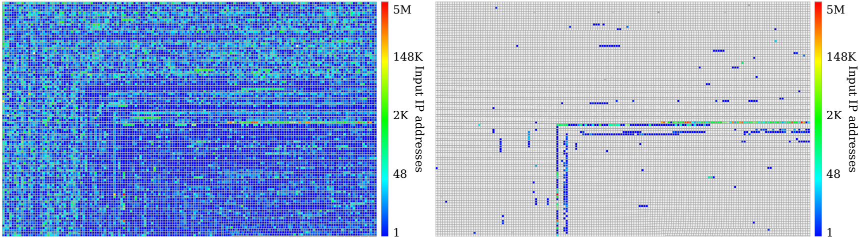
Multi-Level Aliased Prefix Detection – Details

- When do we suspect a prefix aliased?
  - $>$ 100 *IP addresses*
- How to cope with packet loss?
  - *Accept replies for either ICMP or TCP*
  - *Accept replies from past 3 days*
- Impact of Multi-Level Alias Detection?
  - We find several cases where subprefixes of aliased prefixes are **not** aliased
  - 2001:db8:/32 may be aliased
  - 2001:db8::/124 may be not
  - Build a multi-level binary tree and query it using longest-prefix matching

Filtering Aliased Prefixes

Result

- 55.1M raw IPv6 addresses in hit list
- 29.4M non-aliased IPv6 addresses (53.4%)



- Only few prefixes contain aliased prefixes
- But aliased prefixes contain about 47% of addresses in the hit list!

Plots created using zesplot, cf. Luuk Hendrik's maprg talk at IETF101.

Validation: Fingerprinting Aliased Prefixes

Can we validate our results, and learn more about the homogenity of aliased prefixes?
Recall the assumption: All IP addresses in an aliased prefix belong to the same host.

Deploying advanced fingerprinting, used earlier to detect IPv6-IPv4-aliases [1,2].

Features:

- iTTL (*Do all IP addresses in the prefix have the same iTTL value?*)
- TCP Options Fingerprint *(Do all IP addresses in the prefix offer the same TCP Options fingerprint?)*
- TCP Timestamp linearity (*Do remote TCP timestamps in a prefix behave linearly?*)

Scale: These fingerprinting features come at **no additional cost** on top of our liveness probing

Metrics can have confirming and falsifying confidences (e.g., iTTL)

Validation: Fingerprinting Aliased Prefixes

We fingerprinted 20.7k `/64` prefixes considered aliased.

Result confidence heavily dependent on test:

- same iTTL value: small confirmative confidence, large disapproving confidence
- Timestamp linearity: strong confirmative confidence, no negative indication at all (some OSes do not use linear timestamping).

Test	Σ Incs.	Σ Cons.
iTTL	6	20 686
Optionstext	110	20 581
WScale	215	19 515
MSS	1175	19 513
WSize	**1186**	**19 506**
Timestamps	n/a[1]	**13 202**

[1] A failed timestamping test does not indicate an inconsistent, but an indecisive prefix

Few subnets are inconsistent, and a majority is strongly consistent (linear timestamps), indicating that prefixes determined aliased are indeed bound to one host.

Conclusion

- IPv6 hitlists can contain large clusters of *aliased prefixes*
- Rigorous, multi-level aliased prefix detection provides accurate and confident detection, including proper outlier handling
  - (such as non-aliased subprefixes in aliased prefixes)
- Fingerprinting of aliased prefixes can can increase decision confidence
- Paper and Plots:
  `https://ipv6hitlist.github.io/`

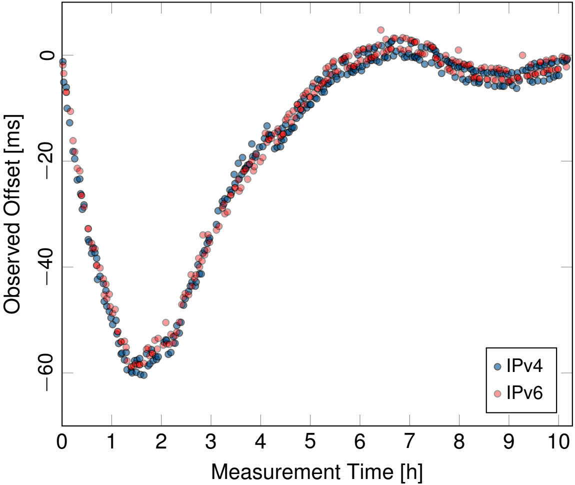
Other topics I am happy to discuss (cf. ANRW):

- HTTPS/TLS security scanning
- Web PKI topics, e.g., CAA DNS records
- TLS Client Certificates
- Internet Toplists

Backup

Backup

Example of Remote TCP Timestamping for Alias Detection



Quirin Scheitle, Oliver Gasser, Minoo Rouhi, Georg Carle, “Large-Scale Classification of IPv6-IPv4 Siblings with Variable Clock Skew", TMA'17