Measured Approaches to IPv6 Address Anonymization and Identity Association

CARIS2 Workshop – Cambridge, MA, 1 Mar 2019

David Plonka <plonka@akamai.com|dave@plonka.us> & Arthur Berger

“kIP: a Measured Approach to IPv6 Address Anonymization” (pre-print)
https://arxiv.org/abs/1707.03900

“In the IP of the Beholder: Strategies for Active Topology Discovery” (IMC 2018)
https://arxiv.org/abs/1805.11308

Premise: an intersection of Privacy and Security

IPv6 poses (at least) two challenges in facets of coordinated attack response:

1. **Sharing IP address-related info while respecting victim and even potential/candidate attacker's privacy.**
2. **Mitigating abuse by dropping or rate-limiting only traffic associated with an attackers' (or victims') identities.**

Meeting these challenges depends on knowledge - or on assumptions - about IP address identities, typically in the form of a public, globally-routed IP address prefix – the *Identity Associations* (or IAs) – of the victimized or attacking parties.

What is a best practice for anonymization of these identities?

Can the identity association be reliably determined, remotely?

IP Address Anonymization and Identity Association

Today we'll consider:

- **Truncation and/or aggregation-based anonymization**
  i.e., for sharing network identifiers for attack response or, generally, in traffic data, *e.g.,* correlating with network topology, routing, service providers, and geographic locations.

- **Nascent IPv6 topology discovery results and implications for determining associated identify**
  i.e., for sharing topology information for attack response *e.g.,* anonymization and identity association involving router addresses.

IP Address Anonymization and Identity Association

Consider these questions:

- *How can passive and active Internet measurements inform decisions about address anonymization and identity association?*

IP Address Anonymization and Identity Association

Consider these questions:

- How can passive and active Internet measurements inform decisions about address anonymization and identity association?

- *Is there reason to believe that any one IP prefix length would perform satisfactorily for either?*

IP Address Anonymization and Identity Association

We consider these questions:

- How can passive and active Internet measurements inform decisions about address anonymization and identity association?
- Is there reason to believe that any one IP prefix length would perform satisfactorily for either?
- *In the face of attack, when, where, and how should IP addresses be de-aggregated or coalesced to effectively associate them with victims or attackers?*

Background: IPv4 Address Anonymization by aggregation

```
10.0.42.24 1    10.0.42.31 1
10.0.42.30 1    10.0.42.10 1
10.0.42.25 1    10.0.42.22 1
10.0.42.6  1    10.0.42.16 1
10.0.42.17 1    10.0.42.4  1
10.0.42.17 1    10.0.42.21 1
10.0.42.9  1    10.0.42.8  1
10.0.42.19 1    10.0.42.20 1
10.0.42.29 1    10.0.42.3  1
10.0.42.26 1    10.0.42.14 1
10.0.42.11 1    10.0.42.1  1
10.0.42.27 1    10.0.42.15 1
10.0.42.13 1
10.0.42.7  1
10.0.42.0  1
10.0.42.12 1
10.0.42.28 1
10.0.42.2  1
10.0.42.23 1
10.0.42.5  1
```

Background: IPv4 Address Anonymization by aggregation to a fixed length

```
10.0.42.24 1    10.0.42.31 1
10.0.42.30 1    10.0.42.10 1
10.0.42.25 1    10.0.42.22 1
10.0.42.6  1    10.0.42.16 1
10.0.42.17 1    10.0.42.4  1
10.0.42.17 1    10.0.42.21 1
10.0.42.9  1    10.0.42.8  1
10.0.42.19 1    10.0.42.20 1
10.0.42.29 1    10.0.42.3  1
10.0.42.26 1    10.0.42.14 1
10.0.42.11 1    10.0.42.1  1
10.0.42.27 1    10.0.42.15 1
10.0.42.13 1
10.0.42.7  1
10.0.42.0  1
10.0.42.12 1
10.0.42.28 1
10.0.42.2  1
10.0.42.23 1
10.0.42.5  1
```


10.0.42.0/27 32

IP Address Anonymization

- ***Truncation-based anonymization is ideal if, and only if, it can be guaranteed to improve privacy.***

  We propose *k*IP anonymization, *i.e.*,
  make an individual appear indistinguishable amongst a set of [*k*] individuals

  [https://en.wikipedia.org/wiki/K-anonymity,
  RFC 6973: “Privacy Considerations for Internet Protocols”]

*k*IP: a measurement-based approach…

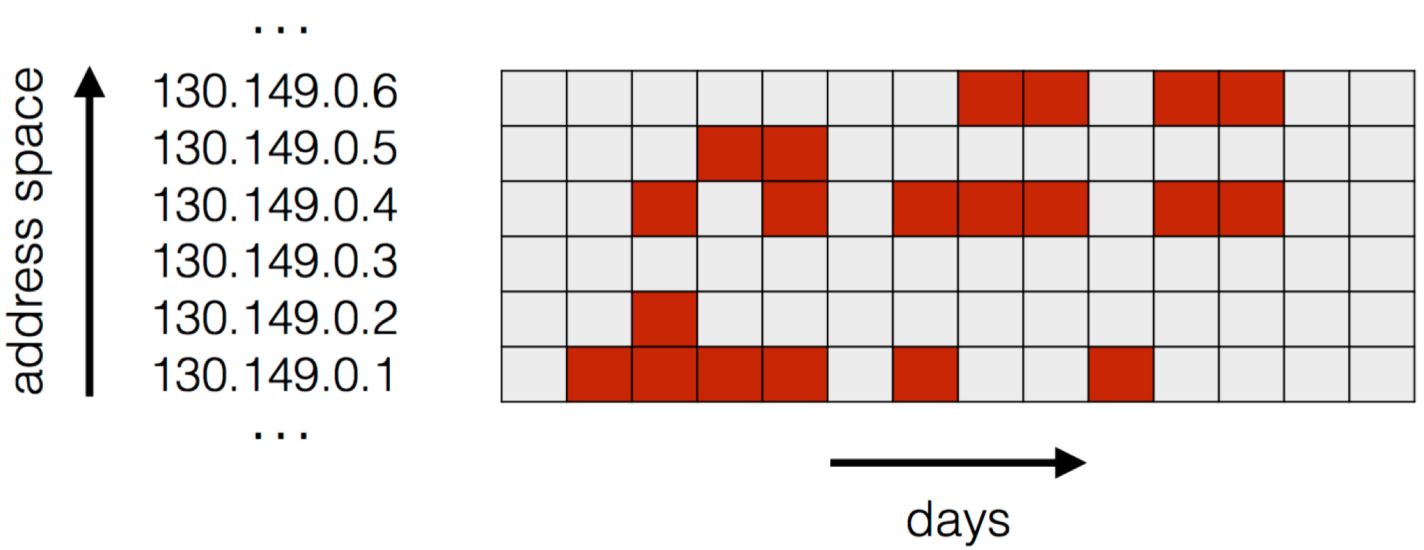
1. **Temporal & Spatial Address *Classification***
   See “kIP: a Measured Approach to IPv6 Address Anonymization”
   Slides/video: https://trac.ietf.org/trac/irtf/wiki/map

2. **Address *Activity Matrix* Analysis:**
   estimating a lower bound on simultaneously assigned addresses

3. **Anonymous *Aggregate* (Prefix) *Synthesis*:**
   then perform longest-prefix match to produce results

Step 2. Address Activity Matrix Analysis

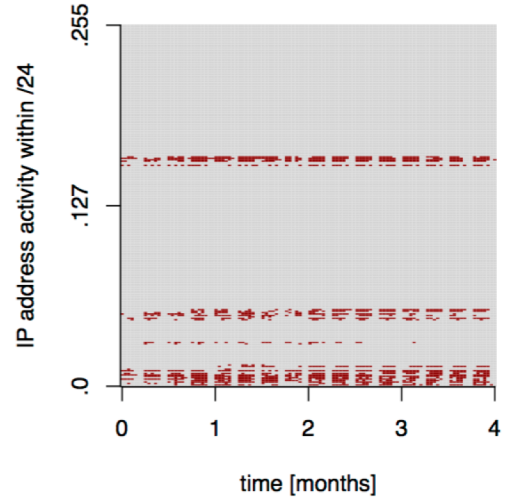
Related Work: IPv4 Address Activity Matrix introduced in “Beyond Counting …”, MAPRG Meeting July 2016



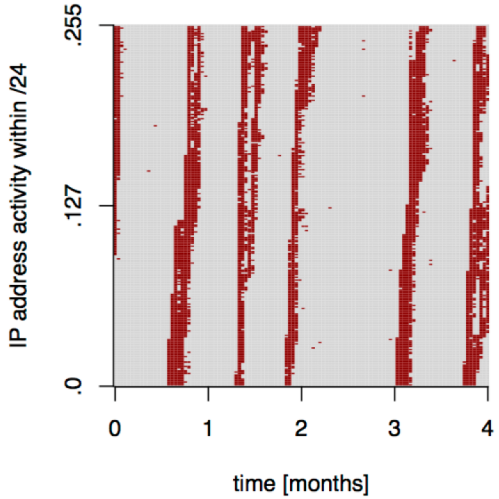
for each day on which an IP address was active (requested content), we draw a red dot.

Beyond Counting: New Perspectives on the Active IPv4 Address Space (Richter *et al.* IMC 2016): https://arxiv.org/abs/1606.00360

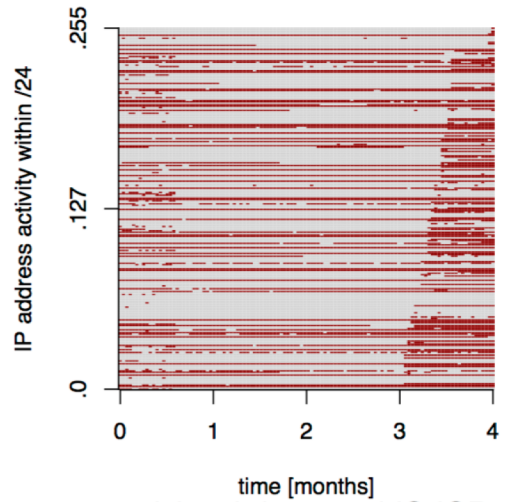
Related Work: IPv4 Address Activity Matrix



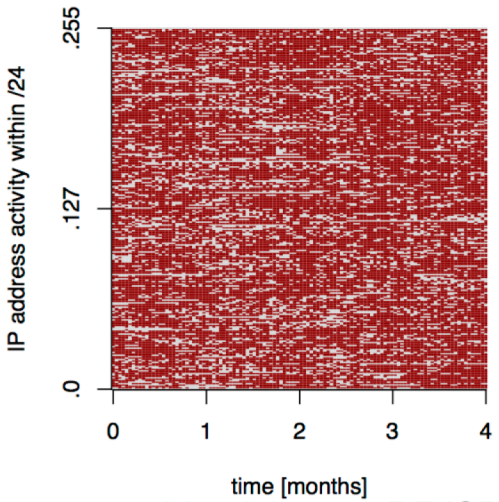
static block DE University



DHCP pool US University



residential users US ISP



residential users DE ISP

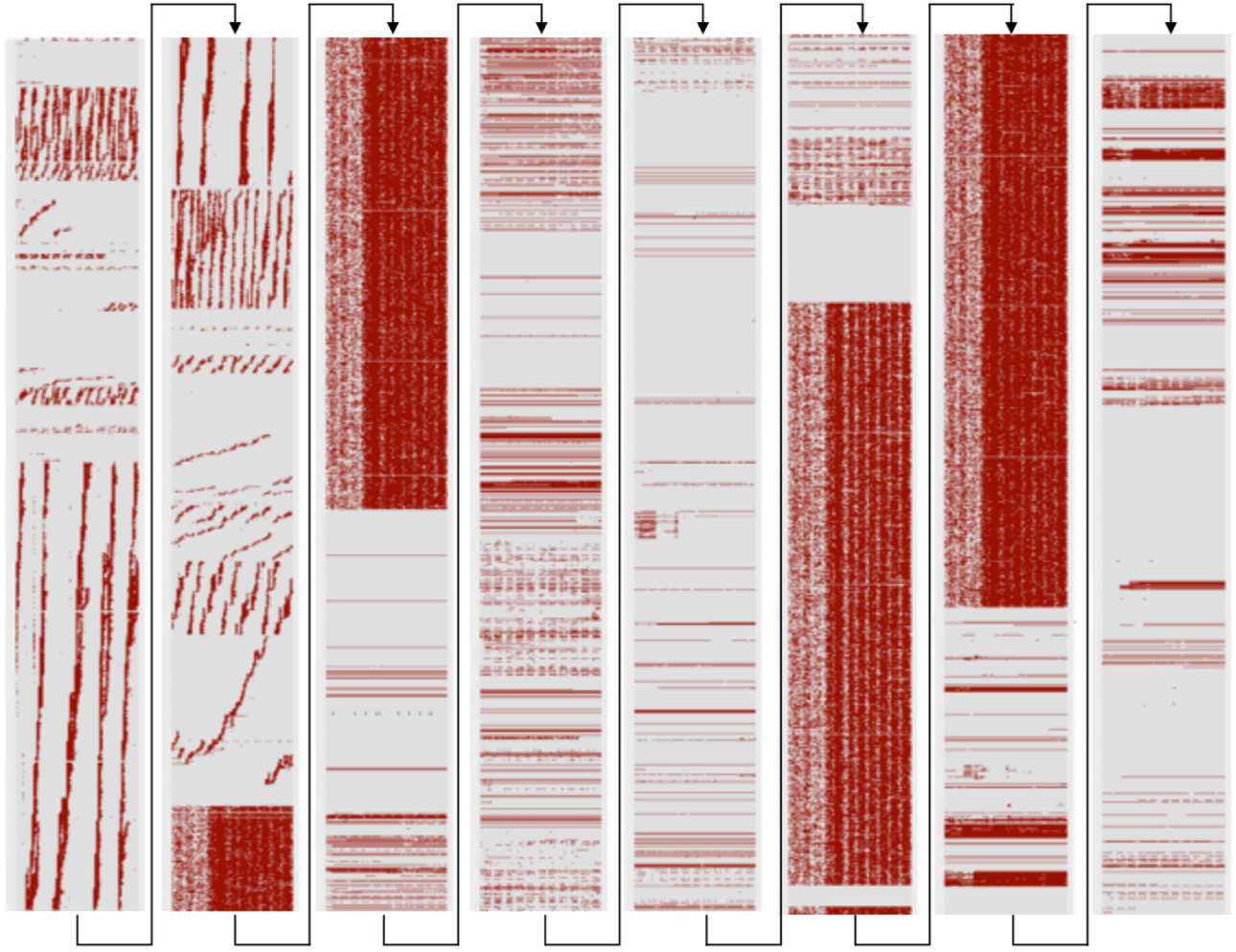
"in situ" activity:
address assignment practice
+
user behavior

(no visible modification of address assignment practice)

Beyond Counting: New Perspectives on the Active IPv4 Address Space (Richter *et al.* IMC 2016): https://arxiv.org/abs/1606.00360

Related Work: IPv4 Address Activity Matrix

20k adjacent IP addresses (in active /24s), University Network



Beyond Counting: New Perspectives on the Active IPv4 Address Space (Richter *et al.* IMC 2016): https://arxiv.org/abs/1606.00360

IPv6 Address Activity Matrix

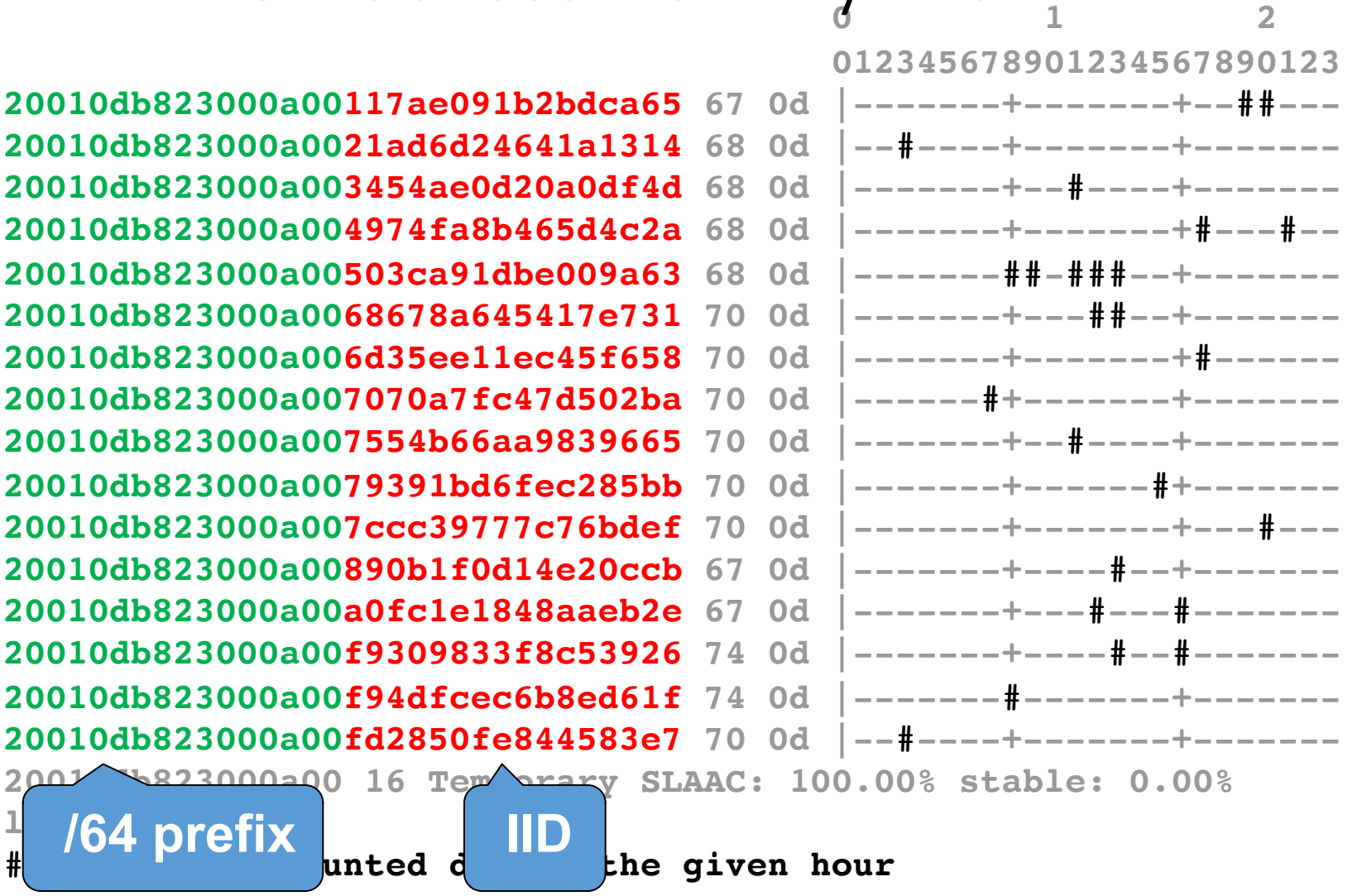
```
                                          0         1         2
                                          012345678901234567890123
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--##---
20010db823000a0021ad6d24641a1314 68 0d |--#----+-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--#----+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+#---#--
20010db823000a00503ca91dbe009a63 68 0d |-------##-###--+-------
20010db823000a0068678a645417e731 70 0d |-------+---##--+-------
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+#------
20010db823000a007070a7fc47d502ba 70 0d |------#+-------+-------
20010db823000a007554b66aa9839665 70 0d |-------+--#----+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------#+-------
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---#---
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----#--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+---#---#-------
20010db823000a00f9309833f8c53926 74 0d |-------+----#--#-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------#-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--#----+-------+-------
20010db823000a00 16 Temporary SLAAC: 100.00% stable: 0.00%
legend:
# = activity counted during the given hour
```

IPv6 Address Activity Matrix

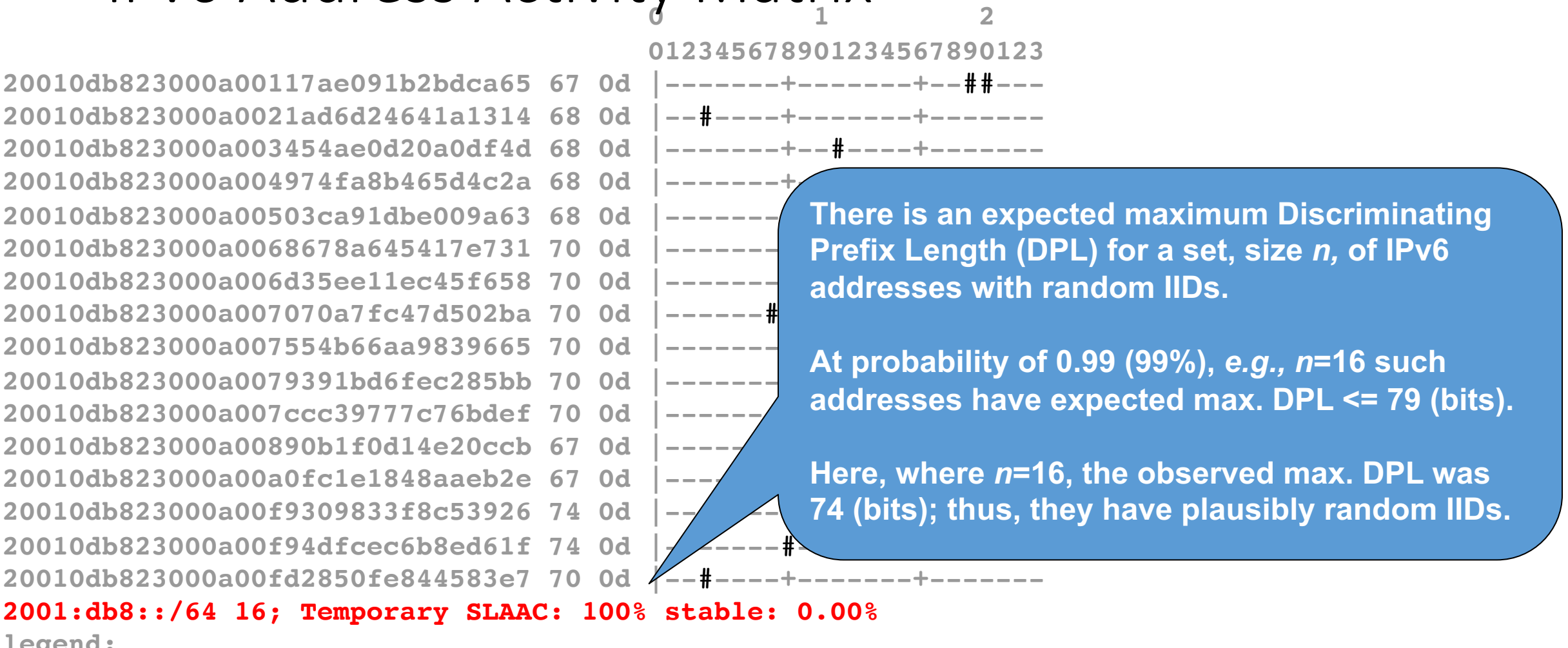
```
                                          0         1         2
                                          012345678901234567890123
20010db823000a00117ae091b2bdca65 67 0d  |-------+-------+--##---
20010db823000a0021ad6d24641a1314 68 0d  |--#----+-------+-------
20010db823000a003454ae0d20a0df4d 68 0d  |-------+--#----+-------
20010db823000a004974fa8b465d4c2a 68 0d  |-------+-------+#---#--
20010db823000a00503ca91dbe009a63 68 0d  |-------##-###--+-------
20010db823000a0068678a645417e731 70 0d  |-------+---##--+-------
20010db823000a006d35ee11ec45f658 70 0d  |-------+-------+#------
20010db823000a007070a7fc47d502ba 70 0d  |------#+-------+-------
20010db823000a007554b66aa9839665 70 0d  |-------+--#----+-------
20010db823000a0079391bd6fec285bb 70 0d  |-------+------#+-------
20010db823000a007ccc39777c76bdef 70 0d  |-------+-------+---#---
20010db823000a00890b1f0d14e20ccb 67 0d  |-------+----#--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d  |-------+---#---#-------
20010db823000a00f9309833f8c53926 74 0d  |-------+----#--#-------
20010db823000a00f94dfcec6b8ed61f 74 0d  |-------#-------+-------
20010db823000a00fd2850fe844583e7 70 0d  |--#----+-------+-------
2001:db823000a00 16 Temporary SLAAC: 100.00% stable: 0.00%
1
# = 1 address counted during the given hour
```

/64 prefix

IPv6 Address Activity Matrix



IPv6 Address Activity Matrix

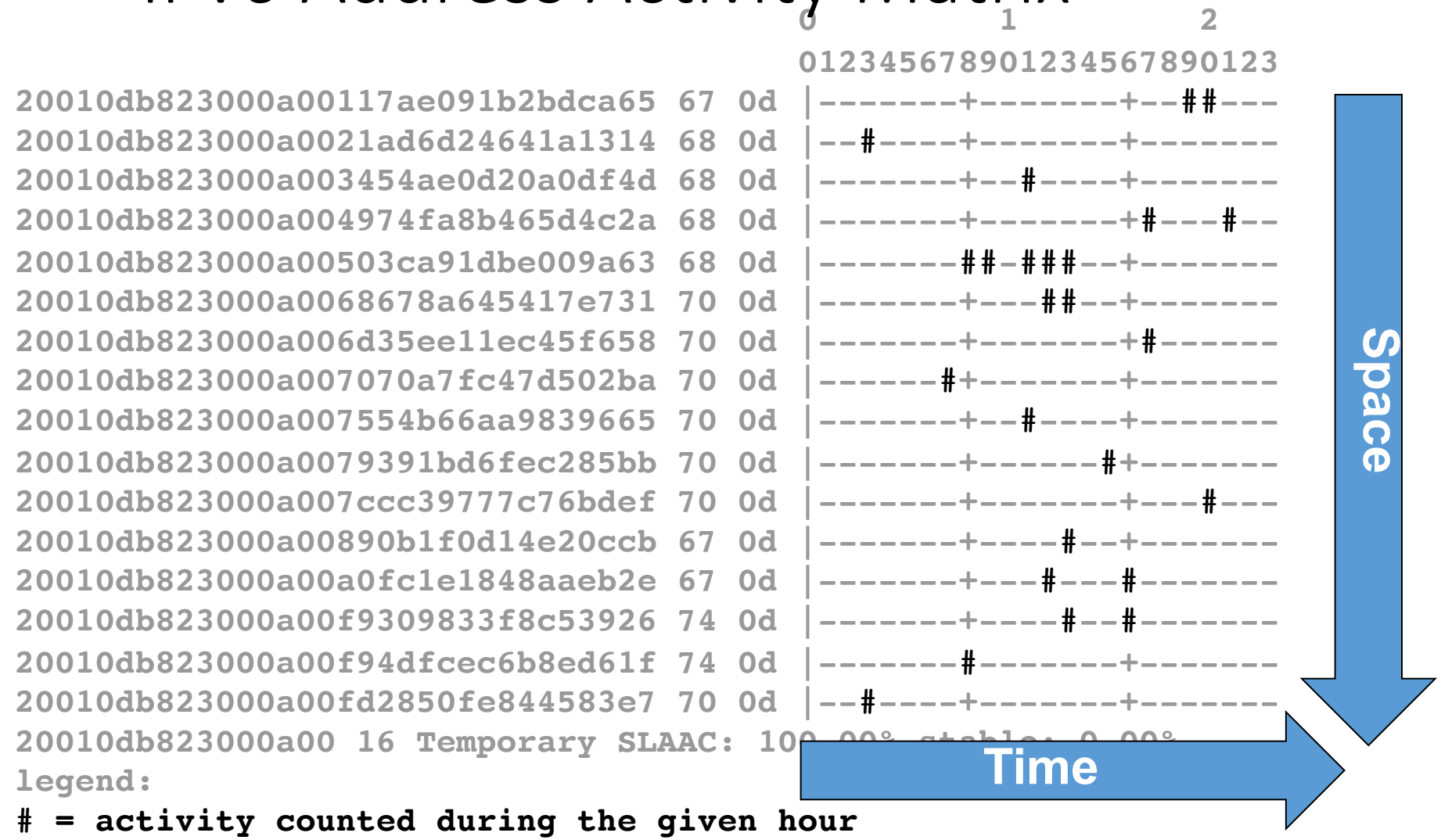


There is an expected maximum Discriminating Prefix Length (DPL) for a set, size *n,* of IPv6 addresses with random IIDs.

At probability of 0.99 (99%), *e.g., n*=16 such addresses have expected max. DPL <= 79 (bits).

Here, where *n*=16, the observed max. DPL was 74 (bits); thus, they have plausibly random IIDs.

IPv6 Address Activity Matrix



legend:

 # = activity counted during the given hour

IPv6 Address Activity Matrix

```
                                          0         1         2
                                          012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--#----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--#----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------#+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |-------##-###--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------#-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--#----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--#----+-------
20010db823000a0068678a645417e731 70 0d |-------+---##--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+---#---#-------
20010db823000a00f9309833f8c53926 74 0d |-------+----#--#-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----#--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------#+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+#---#--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+#------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--##---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---#---
20010db823000a00 16 Temporary SLAAC: 100.00% stable: 0.00%
legend:
# = activity counted during the given hour
```

IPv6 Address Activity Matrix

```
                                          0         1         2
                                          012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--#----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--#----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------#+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |-------#@@@@#--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------#-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--#----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--#----+-------
20010db823000a0068678a645417e731 70 0d |-------+---##--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+---#@@@#-------
20010db823000a00f9309833f8c53926 74 0d |-------+----#@@#-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----#--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------#+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+#@@@#--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+#------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--##---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---#---
20010db823000a00 16 Temporary SLAAC: 100.00% stable: 0.00%
legend:
# = activity counted during the given hour
@ = assignment of address inferred throughout the given hour
```

IPv6 Address Activity Matrix

```
                                          0         1         2
                                          012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--#----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--#----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------#+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |-------#@@@@#--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------#-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--#----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--#----+-------
20010db823000a0068678a645417e731 70 0d |-------+---##--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+---#@@@#-------
20010db823000a00f9309833f8c53926 74 0d |-------+----#@@#-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----#--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------#+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+#@@@#--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+#------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--##---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---#---
20010db823000a00 16 Temporary SLAAC: 100.00% stable: 0.00%
legend:
# = activity counted during the given hour
@ = assignment of address inferred throughout the given hour
```

IPv6 Address Activity Matrix

```
                                                0         1         2
                                                012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--X----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--X----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------X+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |------->@@@@<--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------X-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--X----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--X----+-------
20010db823000a0068678a645417e731 70 0d |-------+---><--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+--->@@@<-------
20010db823000a00f9309833f8c53926 74 0d |-------+----->@@<-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----X--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------X+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+>@@@<--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+X------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--><---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---X---
```

legend:

 # = activity counted during the given hour

@ = assignment of address inferred throughout the given hour

X = activity started and ended during the given hour (within this whole window, e.g., 1 day)

> = starting activity during the given hour (within this whole window, e.g., 1 day)

< = ending activity during the given hour (within this whole window, e.g., 1 day)

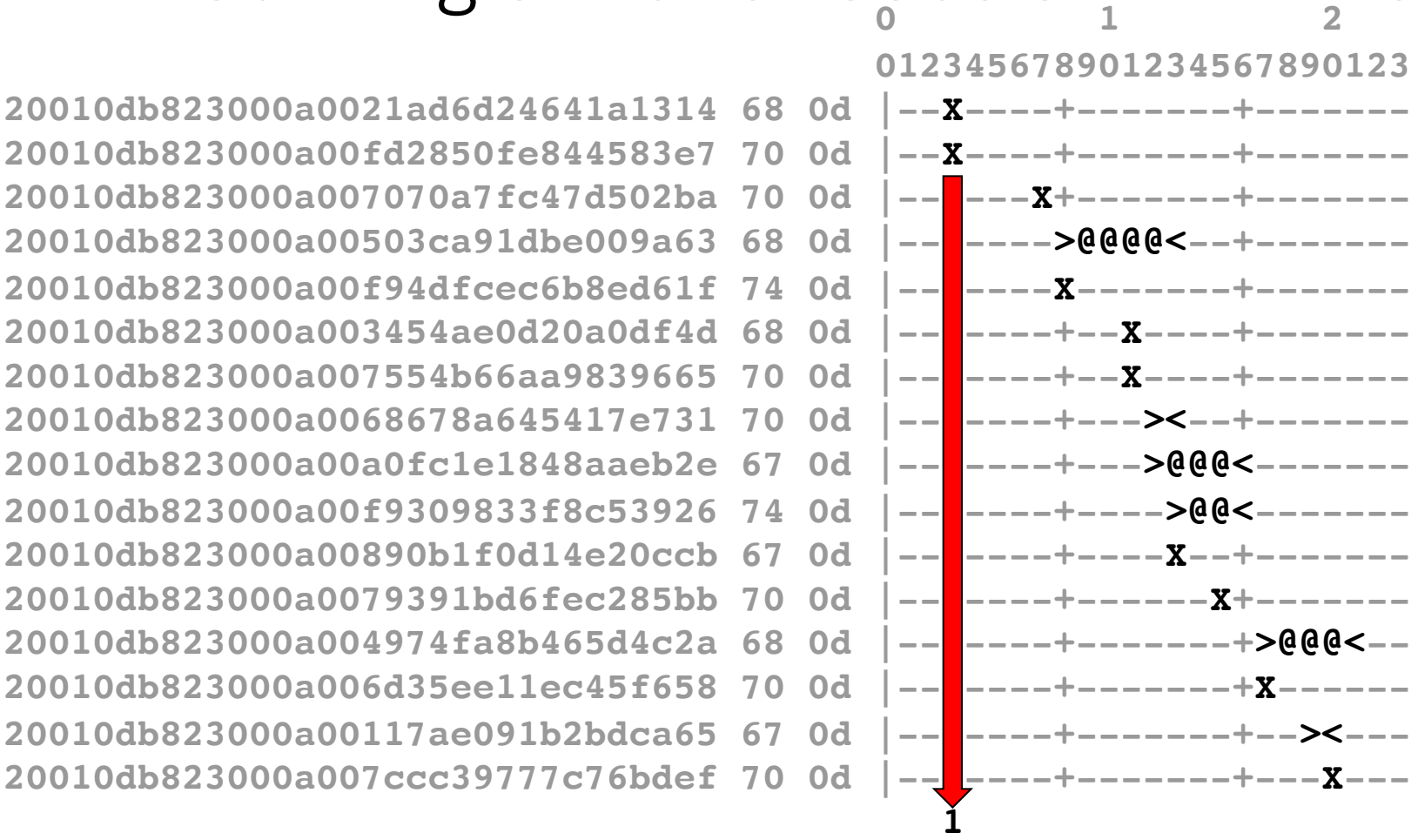
IPv6 Address Activity Matrix

```
                                          0         1         2
                                          012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--X----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--X----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------X+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |------->@@@@<--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------X-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--X----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--X----+-------
20010db823000a0068678a645417e731 70 0d |-------+---><--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+--->@@@<-------
20010db823000a00f9309833f8c53926 74 0d |-------+----->@@<-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----X--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------X+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+>@@@<--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+X------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--><---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---X---
```

legend:

 # = activity counted during the given hour
@ = assignment of address inferred throughout the given hour
X = activity started and ended during the given hour (within this whole window, e.g., 1 day)
> = starting activity during the given hour (within this whole window, e.g., 1 day)
< = ending activity during the given hour (within this whole window, e.g., 1 day)

Counting Simultaneous SLAAC IIDs



legend:

 # = activity counted during the given hour

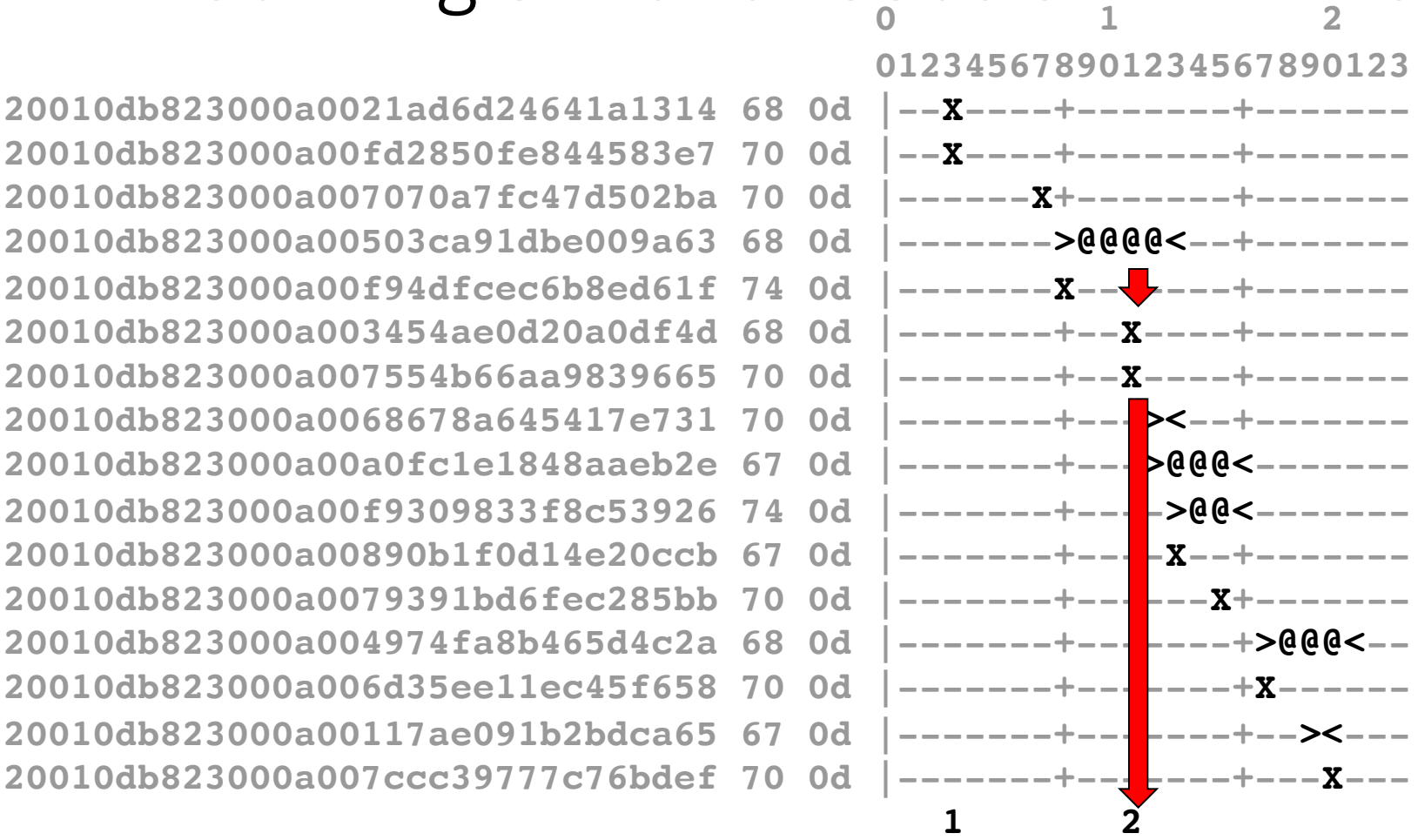
X = activity started and ended during the given hour (within this whole window, e.g., 1 day)

> = starting activity during the given hour (within this whole window, e.g., 1 day)

< = ending activity during the given hour (within this whole window, e.g., 1 day)

@ = assignment of address inferred throughout the given hour

Counting Simultaneous SLAAC IIDs



legend:

 # = activity counted during the given hour

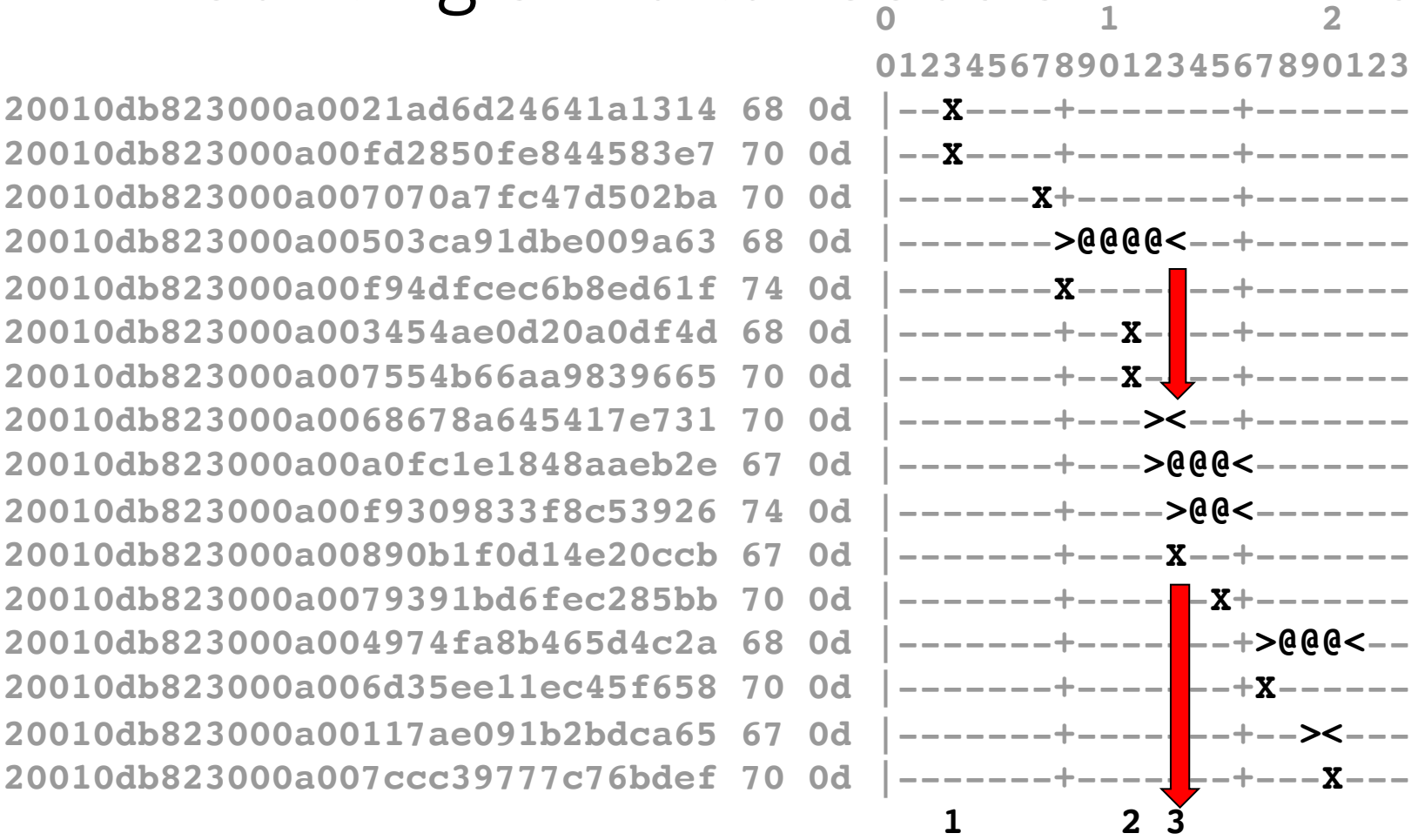
X = activity started and ended during the given hour (within this whole window, e.g., 1 day)

> = starting activity during the given hour (within this whole window, e.g., 1 day)

< = ending activity during the given hour (within this whole window, e.g., 1 day)

@ = assignment of address inferred throughout the given hour

Counting Simultaneous SLAAC IIDs



legend:

 # = activity counted during the given hour

X = activity started and ended during the given hour (within this whole window, e.g., 1 day)

> = starting activity during the given hour (within this whole window, e.g., 1 day)

< = ending activity during the given hour (within this whole window, e.g., 1 day)

@ = assignment of address inferred throughout the given hour

Counting Simultaneous SLAAC IIDs

```
                                          0         1         2
                                          012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--X----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--X----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------X+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |------->@@@@<--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------X-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--X----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--X----+-------
20010db823000a0068678a645417e731 70 0d |-------+---><--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+--->@@@<-------
20010db823000a00f9309833f8c53926 74 0d |-------+----->@@<-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----X--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------X+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+>@@@<--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+X------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--><---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---X---
                                          000100011112332321122100
```

legend:

 # = activity counted during the given hour

X = activity started and ended during the given hour (within this whole window, e.g., 1 day)

> = starting activity during the given hour (within this whole window, e.g., 1 day)

< = ending activity during the given hour (within this whole window, e.g., 1 day)

@ = assignment of address inferred throughout the given hour

Counting Simultaneous SLAAC IIDs

```
                                                0         1         2
                                                012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--X----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--X----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------X+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |------->@@@@<--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------X-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--X----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--X----+-------
20010db823000a0068678a645417e731 70 0d |-------+---><--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+--->@@@<-------
20010db823000a00f9309833f8c53926 74 0d |-------+----->@@<-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----X--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------X+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+>@@@<--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+X------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--><---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---X---
                                                000100011112332321122100 => 3 simultaneous IIDs, maximum
```

legend:

 # = activity counted during the given hour

X = activity started and ended during the given hour (within this whole window, e.g., 1 day)

> = starting activity during the given hour (within this whole window, e.g., 1 day)

< = ending activity during the given hour (within this whole window, e.g., 1 day)

@ = assignment of address inferred throughout the given hour

IPv6 Address Activity Matrix: Identity Assignment

```
                                          0         1         2
                                          012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d  |--X----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d  |--X----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d  |------X+-------+-------
20010db823000a00503ca91dbe009a63 68 0d  |-------@@@@<--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d  |-------X-------+-------
20010db823000a003454ae0d20a0df4d 68 0d  |-------+--X----+-------
20010db823000a007554b66aa9839665 70 0d  |-------+--X----+-------
20010db823000a0068678a645417e731 70 0d  |-------+---><--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d  |-------+---@@@<-------
20010db823000a00f9309833f8c53926 74 0d  |-------+----@@<-------
20010db823000a00890b1f0d14e20ccb 67 0d  |-------+----X--+-------
20010db823000a0079391bd6fec285bb 70 0d  |-------+------X+-------
20010db823000a004974fa8b465d4c2a 68 0d  |-------+-------+@@@<--
20010db823000a006d35ee11ec45f658 70 0d  |-------+-------+X------
20010db823000a00117ae091b2bdca65 67 0d  |-------+-------+--><---
20010db823000a007ccc39777c76bdef 70 0d  |-------+-------+---X---
                                          000100011112332321122100 => 3 simultaneous IIDs, maximum
2001:db8::/64 16; Temporary SLAAC: 100%--------!!!!!!!!-!!!!--?
                                          000000001111111110111100? => /64 assignment @ fenceposts
legend:
! = infer /64 prefix assigned at the "fencepost" moments between intervals
```

IPv6 Address Activity Matrix: Identity Assignment

```
                                        0         1         2
                                        012345678901234567890123
20010db823000a0021ad6d24641a1314 68 0d |--X----+-------+-------
20010db823000a00fd2850fe844583e7 70 0d |--X----+-------+-------
20010db823000a007070a7fc47d502ba 70 0d |------X+-------+-------
20010db823000a00503ca91dbe009a63 68 0d |------->@@@@<--+-------
20010db823000a00f94dfcec6b8ed61f 74 0d |-------X-------+-------
20010db823000a003454ae0d20a0df4d 68 0d |-------+--X----+-------
20010db823000a007554b66aa9839665 70 0d |-------+--X----+-------
20010db823000a0068678a645417e731 70 0d |-------+---><--+-------
20010db823000a00a0fc1e1848aaeb2e 67 0d |-------+--->@@@<-------
20010db823000a00f9309833f8c53926 74 0d |-------+---->@@<-------
20010db823000a00890b1f0d14e20ccb 67 0d |-------+----X--+-------
20010db823000a0079391bd6fec285bb 70 0d |-------+------X+-------
20010db823000a004974fa8b465d4c2a 68 0d |-------+-------+>@@@<--
20010db823000a006d35ee11ec45f658 70 0d |-------+-------+X------
20010db823000a00117ae091b2bdca65 67 0d |-------+-------+--><---
20010db823000a007ccc39777c76bdef 70 0d |-------+-------+---X---
                                        000100011112332321122100 => 3 simultaneous IIDs, maximum
2001:db8::/64 16; Temporary SLAAC: 100%--------!!!!!!!! !!!!--?
                                        000000001111111101111100? => /64 assignment @ fenceposts
legend:
! = infer /64 prefix assigned at the "fencepost" moments between intervals
```

Results: simultaneously-assigned addresses and prefixes

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	**Simultaneously-assigned /64 prefixes max. (median)**	**Simultaneously-assigned addresses max. (median)**	Active addresses (7 days)
Meeting Network	1	3	**3 (2)**	**309 (84)**	15.4K
EU ISP	163K	21.4M	**2.02M (1.52M)**	**3.80M (2.63M)**	125M
JP ISP	2.46M	2.46M	**1.21M (897K)**	**2.26M (1.54M)**	72.2M
US ISP	8.16K	2.42M	**1.81M (1.66M)**	**4.71M (3.82M)**	84.5M

Results: simultaneously-assigned addresses and prefixes

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	**Simultaneously-assigned /64 prefixes max. (median)**	**Simultaneously-assigned addresses max. (median)**	Active addresses (7 days)
Meeting Network	1	3	**3 (2)**	**309 (84)**	15.4K
EU ISP	163K	21.4M	2.02M (1.52M)	3.80M (2.63M)	125M
JP ISP	2.46M	2.46M	1.21M (897K)	2.26M (1.54M)	72.2M
US ISP	8.16K	2.42M	1.81M (1.66M)	4.71M (3.82M)	84.5M

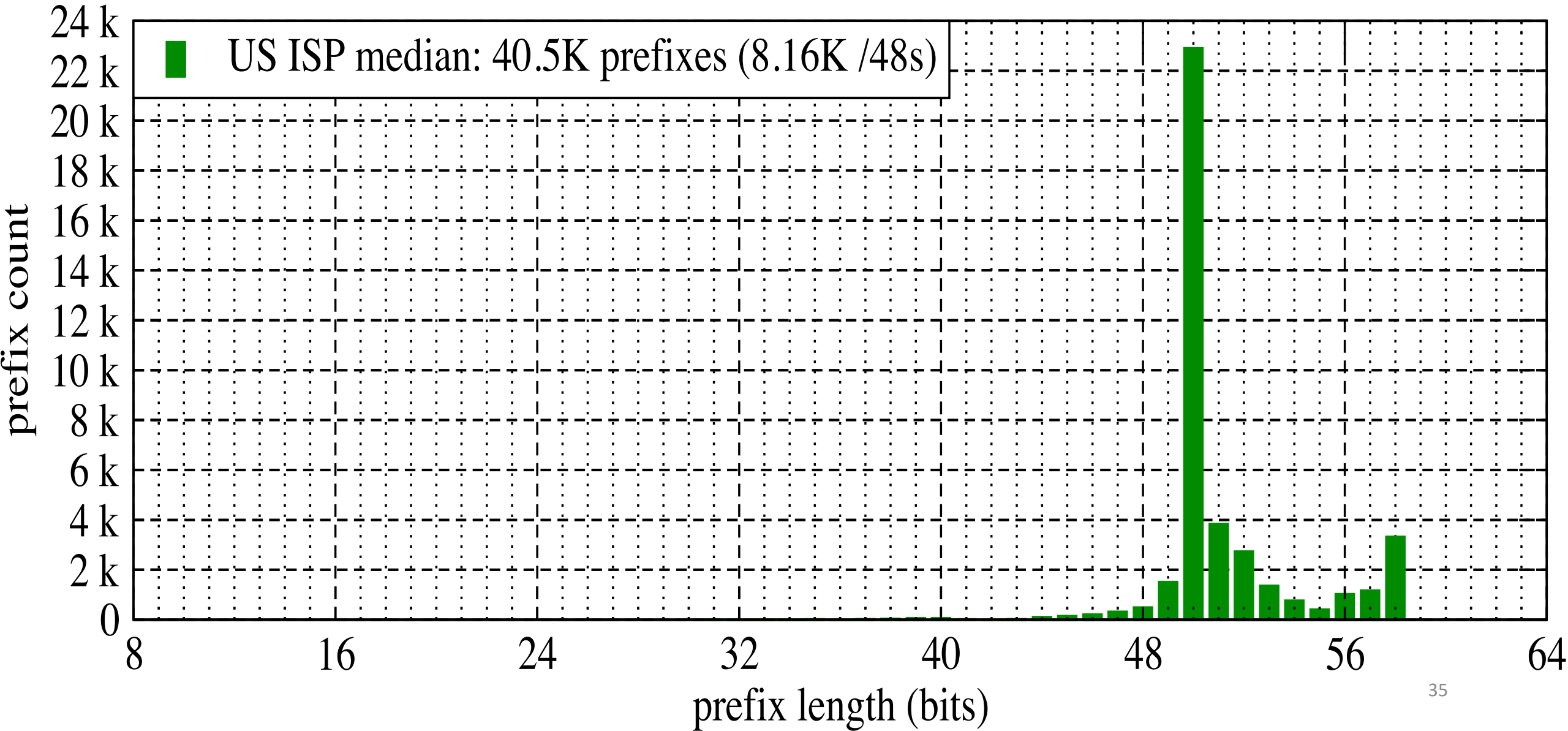
Results: simultaneously-assigned addresses and prefixes

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	**Simultaneously-assigned /64 prefixes max. (median)**	**Simultaneously-assigned addresses max. (median)**	Active addresses (7 days)
Meeting Network	1	3	3 (2)	309 (84)	15.4K
EU ISP	163K	21.4M	**2.02M (1.52M)**	**3.80M (2.63M)**	125M
JP ISP	2.46M	2.46M	**1.21M (897K)**	**2.26M (1.54M)**	72.2M
US ISP	8.16K	2.42M	**1.81M (1.66M)**	**4.71M (3.82M)**	84.5M

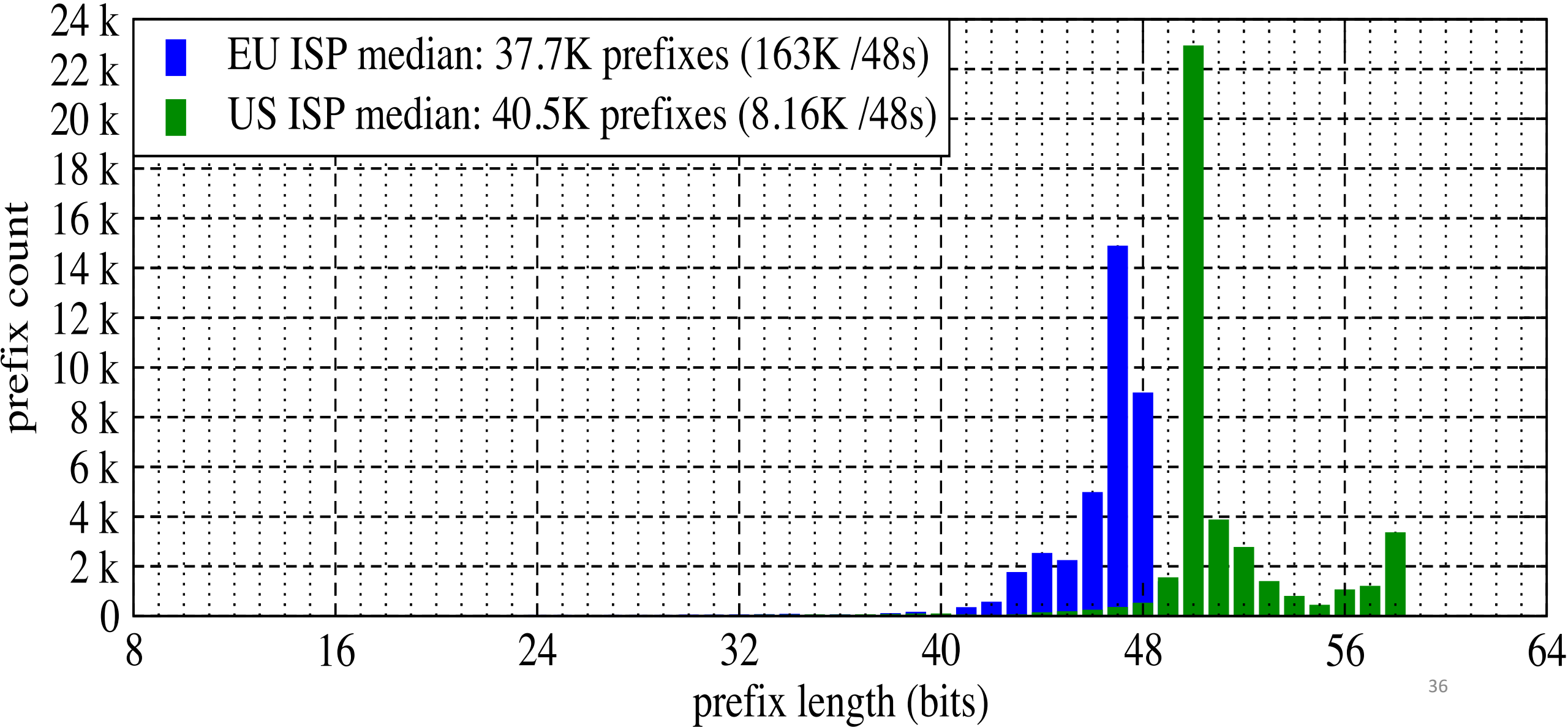
Results: simultaneously-assigned addresses and prefixes

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	**Simultaneously-assigned /64 prefixes max. (median)**	**Simultaneously-assigned addresses max. (median)**	Active addresses (7 days)
Meeting Network	1	3	3 (2)	309 (84)	15.4K
EU ISP	163K	21.4M	**2.02M (1.52M)**	**3.80M (2.63M)**	125M
JP ISP	2.46M	2.46M	**1.21M (897K)**	**2.26M (1.54M)**	72.2M
US ISP	8.16K	2.42M	**1.81M (1.66M)**	**4.71M (3.82M)**	84.5M

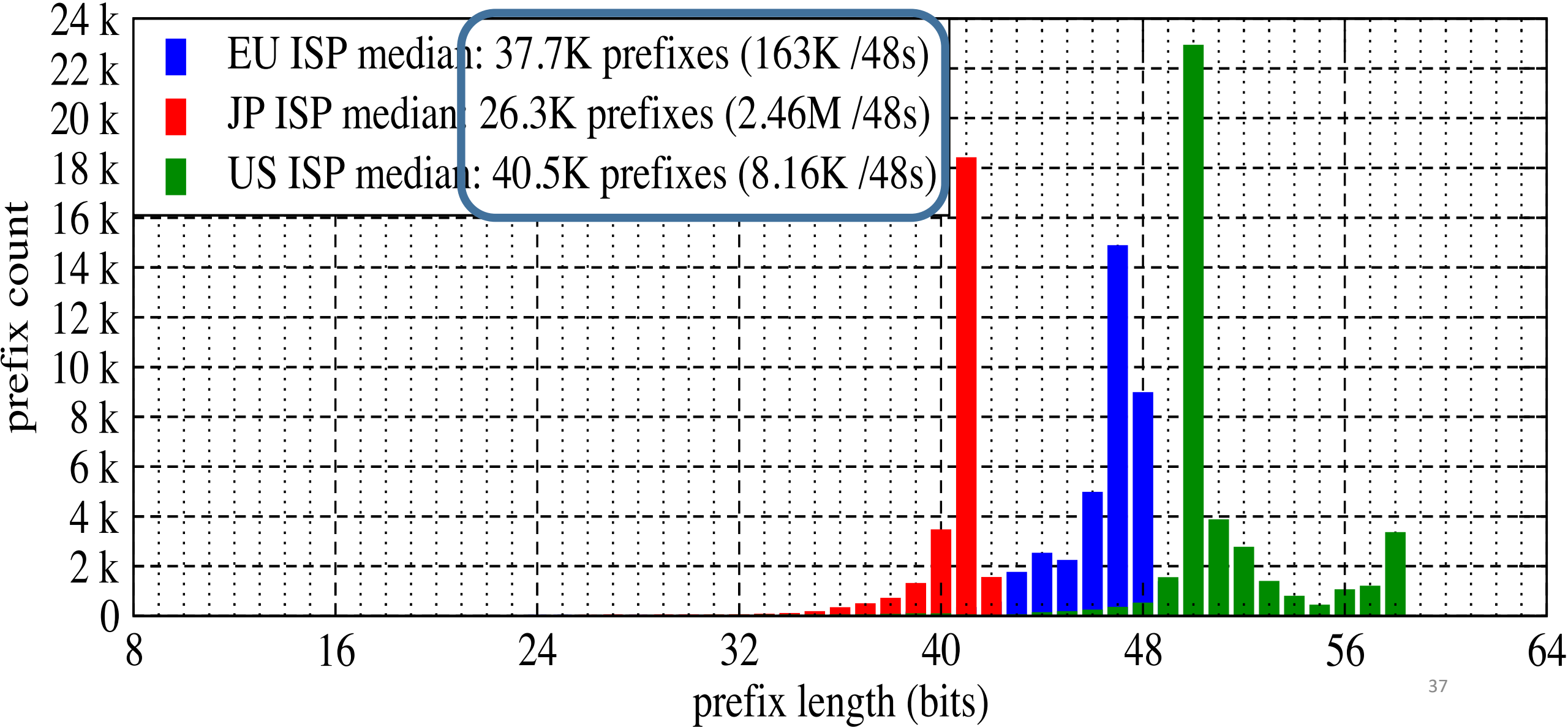
Histogram k=32 anonymous aggregate prefix lengths (w=7d, i=1h)



Histogram: k=32 anonymous aggregate prefix lengths (w=7d, i=1h)



Histogram: k=32 anonymous aggregate prefix lengths (w=7d, i=1h)



IPv6 & Privacy & IoT: Strategies for Active Topology Discovery

You know traceroute.
What is Yarrp?

“Yelling at Random Routers Progressively”

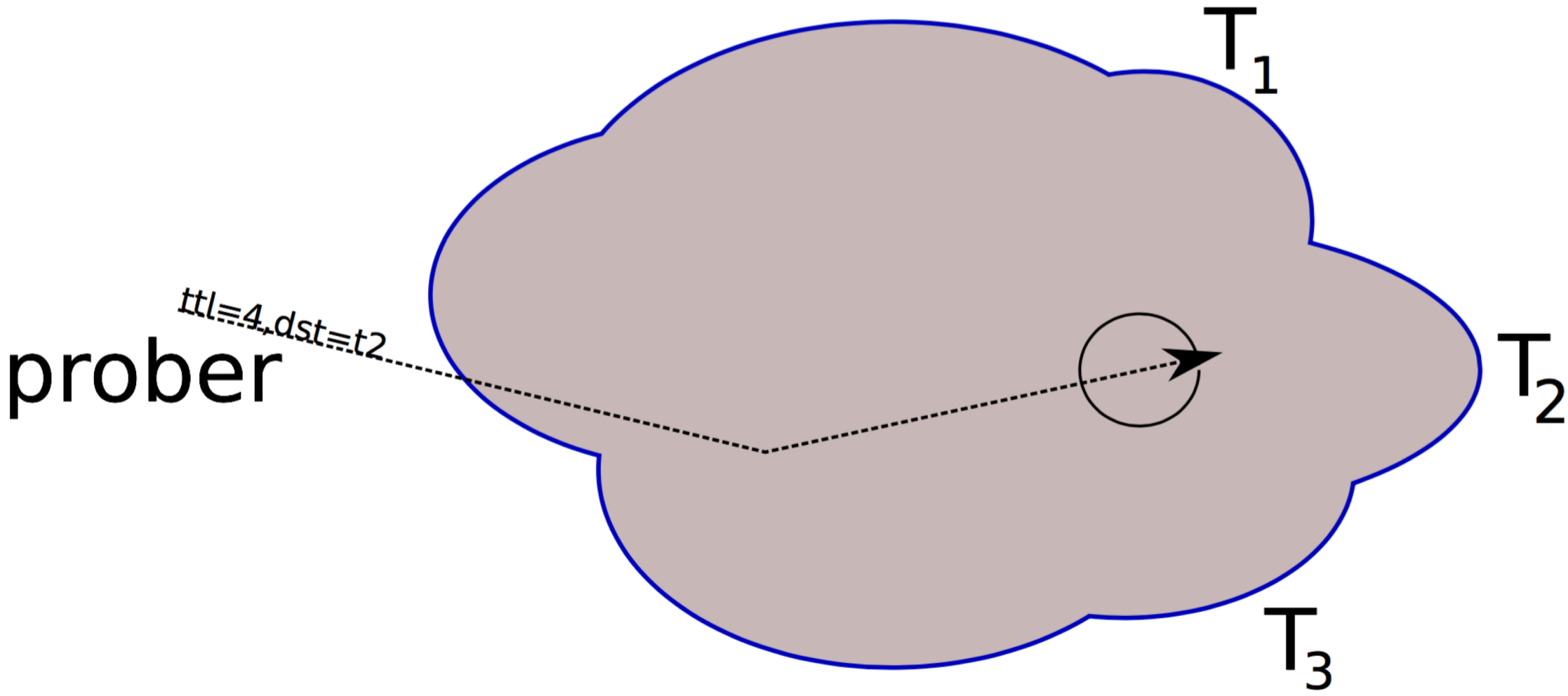


Image credit: R. Beverly, 2016.

In contrast, Yarrp iterates through randomly permuted $< \textit{Target}, \textit{TTL} >$ pairs

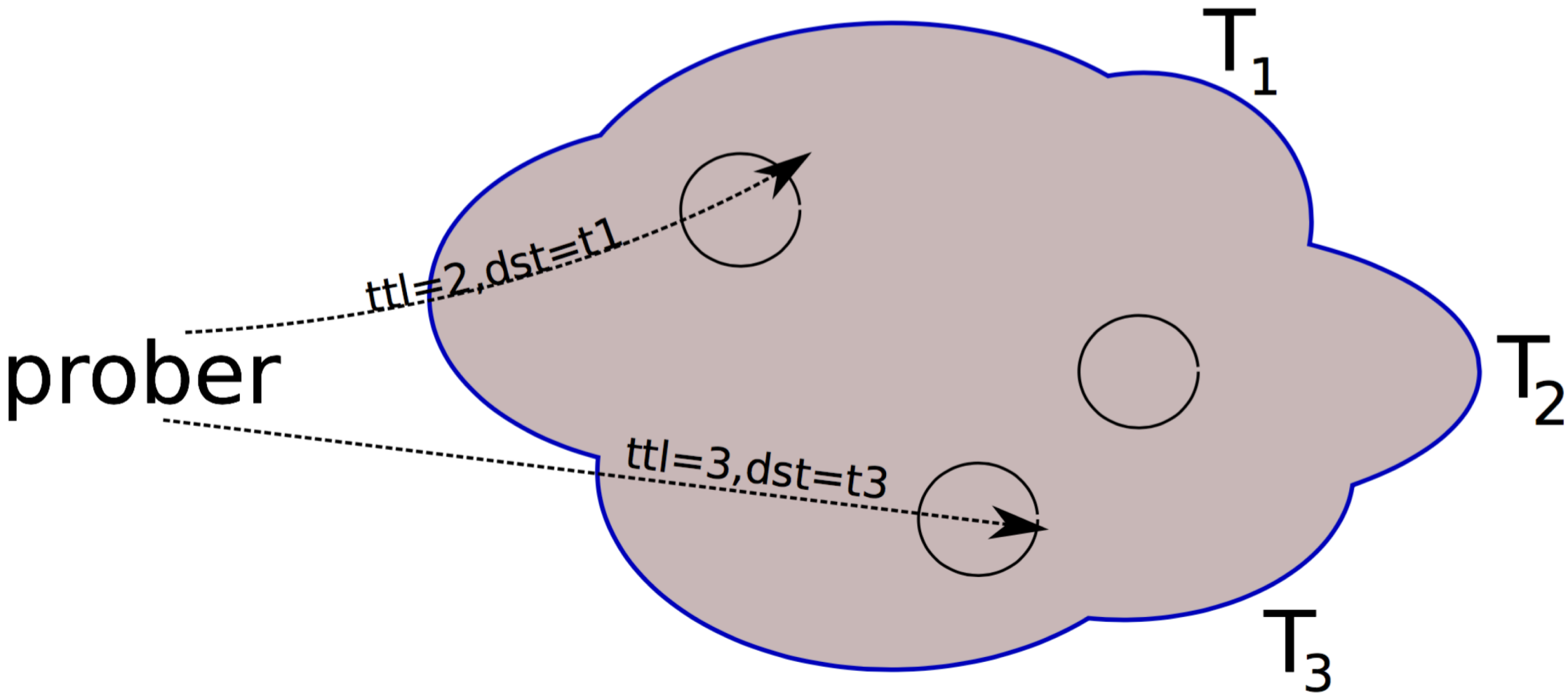


Image credit: R. Beverly, 2016.

In contrast, Yarrp iterates through randomly permuted $< \textit{Target}, \textit{TTL} >$ pairs

prober

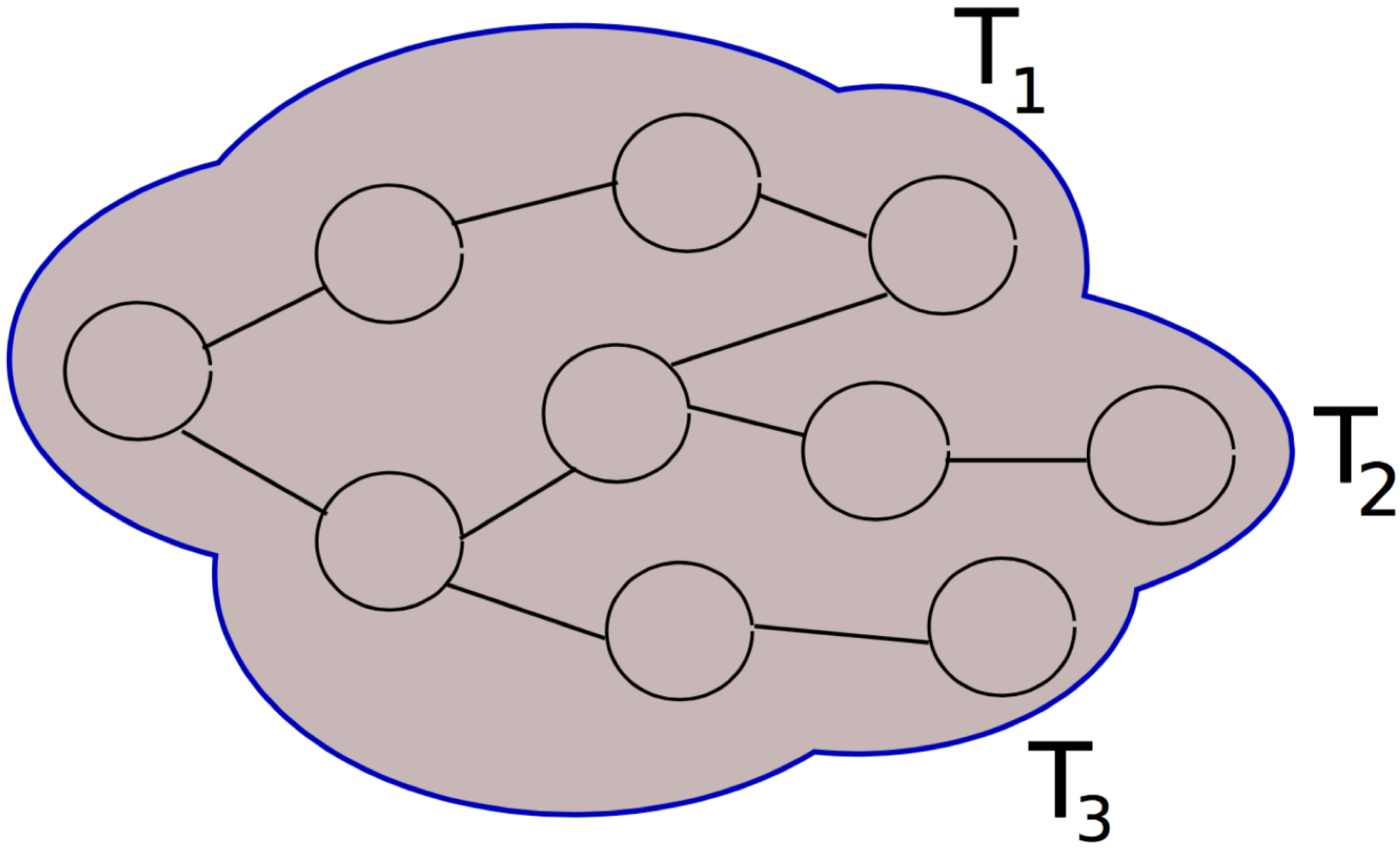


Image credit: R. Beverly, 2016.

Finally, stitch together topology. Requires state and computation, but decoupled (off-line after probing completes).

Rendezvous: How do we [best] choose targets for probes?

Seed address set (a.k.a. “hit list”) Characterization: Finding where the IPv6 action is

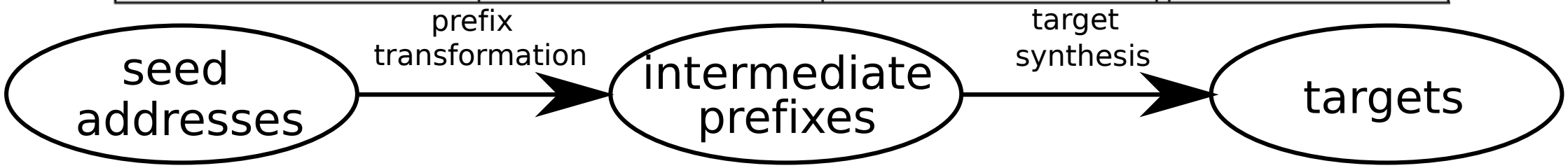
Name	Method	Date yyyy-mm-dd	# Addrs
CAIDA [7]	BGP-derived	2018-05-09	105.2k
DNSDB [41]	Passive DNS	2018-04-28	5.4M
Fiebig [16]	Reverse DNS	2018-03-27	11.7M
FDNS [38]	Fwd. DNS	2018-04-27	24.8M
CDN Clients [36]	$k256$ Agg.	2018-03-03	N/A
	$k32$ Agg.	2018-03-03	N/A
6gen [34]	Generative	2018-02-13	4.9M
Combined	Join Sets	Varies	50.8M
TUM [18]	Collection	Varies	5.6M
Random	Random	2018-05-23	26.5M

Seed address set (a.k.a. “hit list”) Characterization: Finding where the IPv6 action is

Name	Method	Date yyyy-mm-dd	# Addrs
CAIDA [7]	BGP-derived	2018-05-09	105.2k
DNSDB [41]	Passive DNS	2018-04-28	5.4M
Fiebig [16]	Reverse DNS	2018-03-27	11.7M
FDNS [38]	Fwd. DNS	2018-04-27	24.8M
CDN Clients [36]	*k*256 Agg.	2018-03-03	N/A
	*k*32 Agg.	2018-03-03	N/A
6gen [34]	Generative	2018-02-13	4.9M
Combined	Join Sets	Varies	50.8M
TUM [18]	Collection	Varies	5.6M
Random	Random	2018-05-23	26.5M

Seed address set (a.k.a. “hit list”) Characterization: Finding where the IPv6 action is

Name	Method	Date yyyy-mm-dd	# Addrs
CAIDA [7]	BGP-derived	2018-05-09	105.2k
DNSDB [41]	Passive DNS	2018-04-28	5.4M
Fiebig [16]	Reverse DNS	2018-03-27	11.7M
FDNS [38]	Fwd. DNS	2018-04-27	24.8M
CDN Clients [36]	$k256$ Agg.	2018-03-03	N/A
	$k32$ Agg.	2018-03-03	N/A
6gen [34]	Generative	2018-02-13	4.9M
Combined	Join Sets	Varies	50.8M
TUM [18]	Collection	Varies	5.6M
Random	Random	2018-05-23	26.5M



Yield in yarrp campaigns

Table 7: Results of aggregate yarrp campaigns run from three vantages, reverse sorted by yield. Rtr Hop Addrs are sources of ICMPv6 Time-Exceeded messages.

Yarrp6 Campaign	Agg	Traces	Target Addrs	Rtr Hop Addrs	Excl Hop Addrs	Rtr BGP Pfxs	Excl Rtr BGP	Rtr ASNs	Excl Rtr ASNs	EUI-64 Hop Addrs
ALL	both	45.8M	12.6M	1.4M	0	9.9k	0	7.1k	0	651.4k
EU-NET	both	15.0M	12.2M	1.3M	136.0k	9.5k	236	6.9k	110	613.0k
US-EDU-1	both	15.4M	12.6M	1.3M	84.9k	9.4k	75	6.8k	31	602.7k
US-EDU-2	both	15.4M	12.6M	881.4k	20.7k	7.4k	148	5.5k	76	540.6k
cdn k32	z64	9.6M	3.2M	756.6k	91.6k	2.0k	31	1.2k	7	297.2k
tum	z64	6.2M	2.1M	582.4k	113.7k	7.9k	234	6.0k	131	311.2k
cdn k32	z48	1.6M	524.2k	203.7k	16.5k	1.8k	21	1.2k	7	79.8k
fdns	z64	2.2M	746.9k	185.2k	33.8k	6.2k	147	5.0k	80	15.4k
dnsdb	z64	698.6k	233.0k	154.0k	26.8k	7.8k	223	6.0k	132	10.1k
6gen	z64	13.4M	4.5M	126.4k	21.5k	7.1k	242	5.2k	115	24.8k

Yield in yarrp campaigns

Table 7: Results of aggregate yarrp campaigns run fr… …ield. Rtr Hop Addrs are sources of ICMPv6 Time-Exceede…

Yarrp6 Campaign	Agg	Traces	Target Addrs	Rtr Hop Addrs	Ex… H…	…64 …s
ALL	both	45.8M	12.6M	1.4M		…4k
EU-NET	both	15.0M	12.2M	1.3M	136.0k	…0k
US-EDU-1	both	15.4M	12.6M	1.3M	84.9k	…7k
US-EDU-2	both	15.4M	12.6M	881.4k	20.7k	…6k
cdn k32	z64	9.6M	3.2M	756.6k	91.6k	…2k
tum	z64	6.2M	2.1M	582.4k	113.7k	…2k
cdn k32	z48	1.6M	524.2k	203.7k	16.5k	…8k
fdns	z64	2.2M	746.9k	185.2k	33.8k	…4k
dnsdb	z64	698.6k	233.0k	154.0k	26.8k	…1k
6gen	z64	13.4M	4.5M	126.4k	21.5k	….8k

45.8 M traceroutes, from three vantages, to 12.6 M target addresses yields 1.4 M IPv6 router addresses.

15.0 M traceroutes, from one vantage to 12.2 M target addresses yields 1.3 M router addresses.

Yield in yarrp campaigns

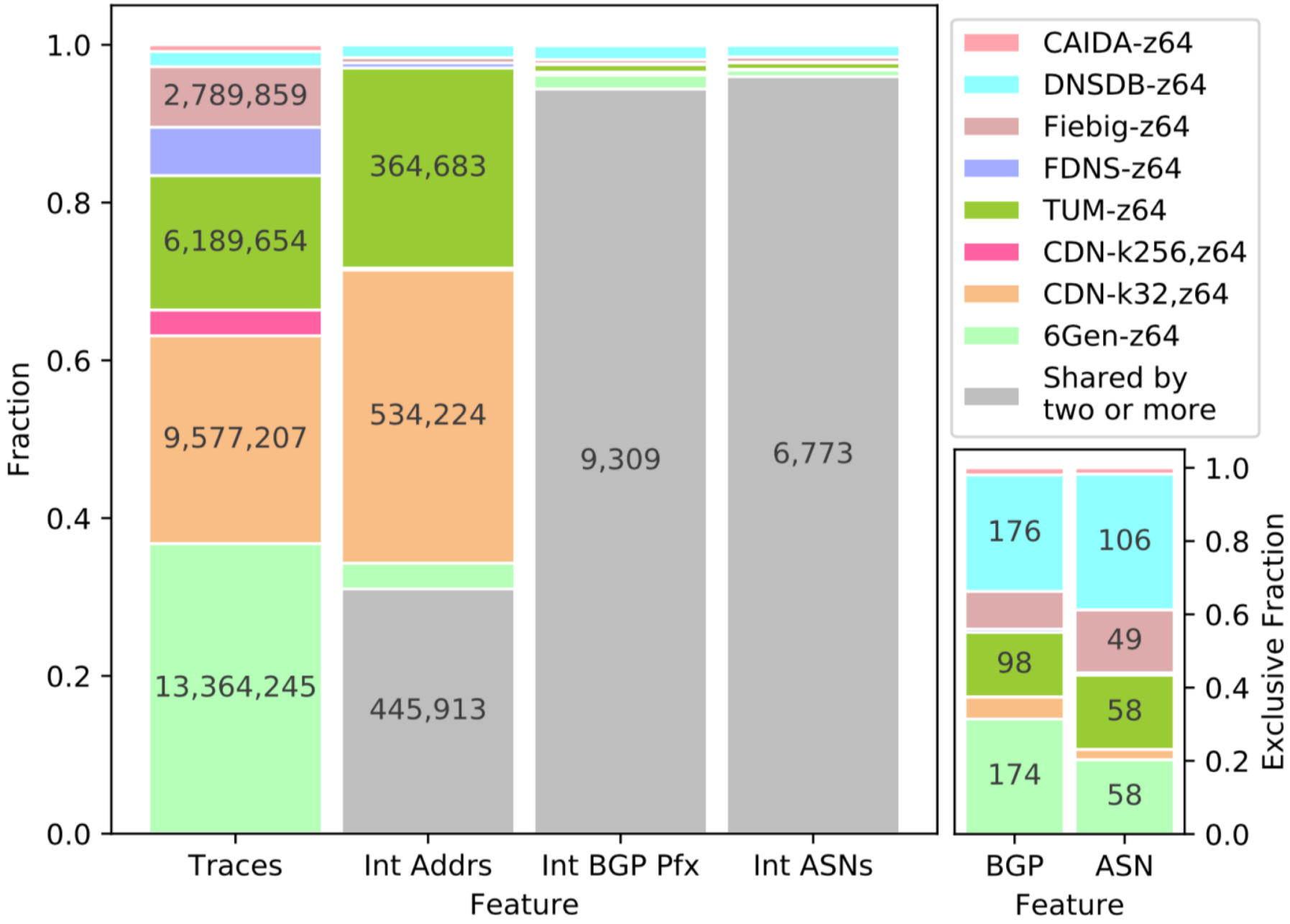
Very many EUI-64 router hop addresses, *i.e.*, having embedded MAC addresses!

59% of these MAC addresses were just two manufacturers: 99.9% of each of these are in just two ISP networks, *i.e.*, large sets of homogenous, IPv6-connected things: ostensibly CPE routers.

Table 7: Results of a… …everse sorted by yield. Rtr Hop Addrs are s…

Yarrp6 Campaign	Agg							tr NNs	Excl Rtr ASNs	EUI-64 Hop Addrs
ALL	both							.1k	0	651.4k
EU-NET	both								110	613.0k
US-EDU-1	both								31	602.7k
US-EDU-2	both								76	540.6k
cdn k32	z64									297.2k
tum	z64							.0k	131	311.2k
cdn k32	z48							1.2k	7	79.8k
fdns	z64							5.0k	80	15.4k
dnsdb	z64	698.6k	233.0k	154.0k	26.8k	7.8k	223	6.0k	132	10.1k
6gen	z64	13.4M	4.5M	126.4k	21.5k	7.1k	242	5.2k	115	24.8k

Yield in yarrp campaigns



Path Divergence Analysis

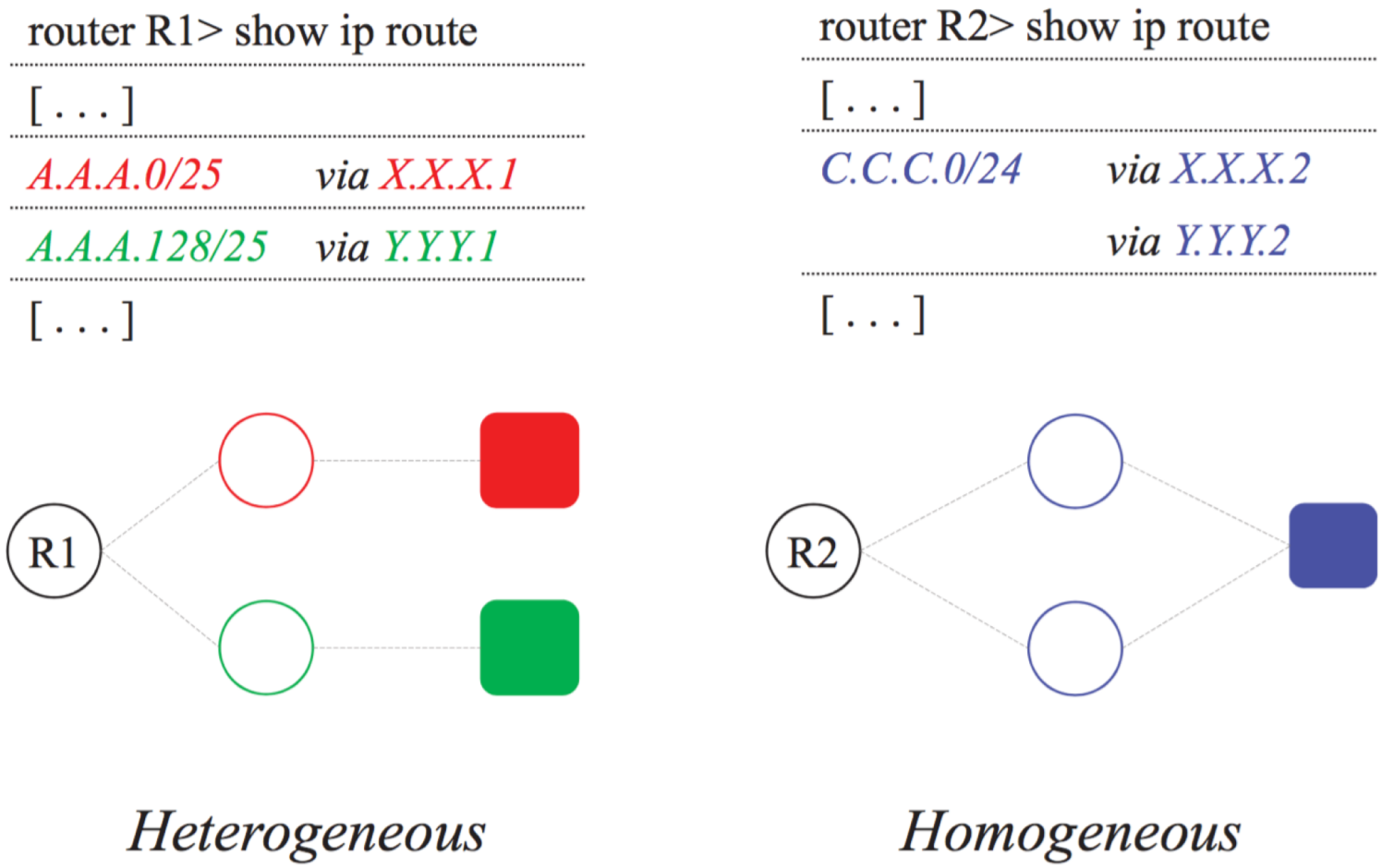
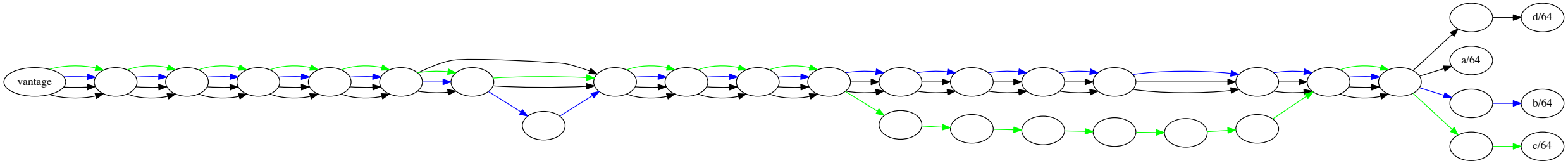


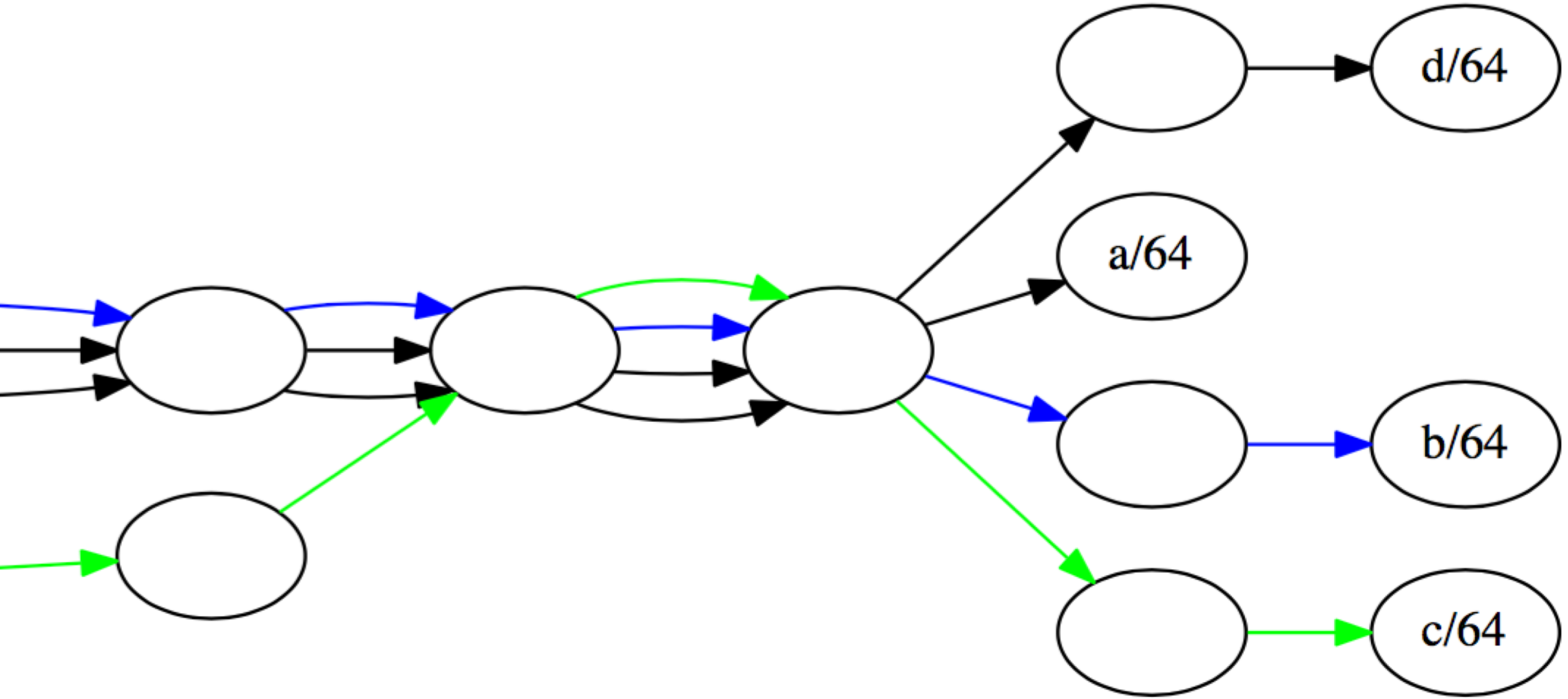
Figure 1: Different last-hop routers due to distinct route entries (left) and per-destination load-balancing (right).

Image credit: Lee & Spring, 2016.

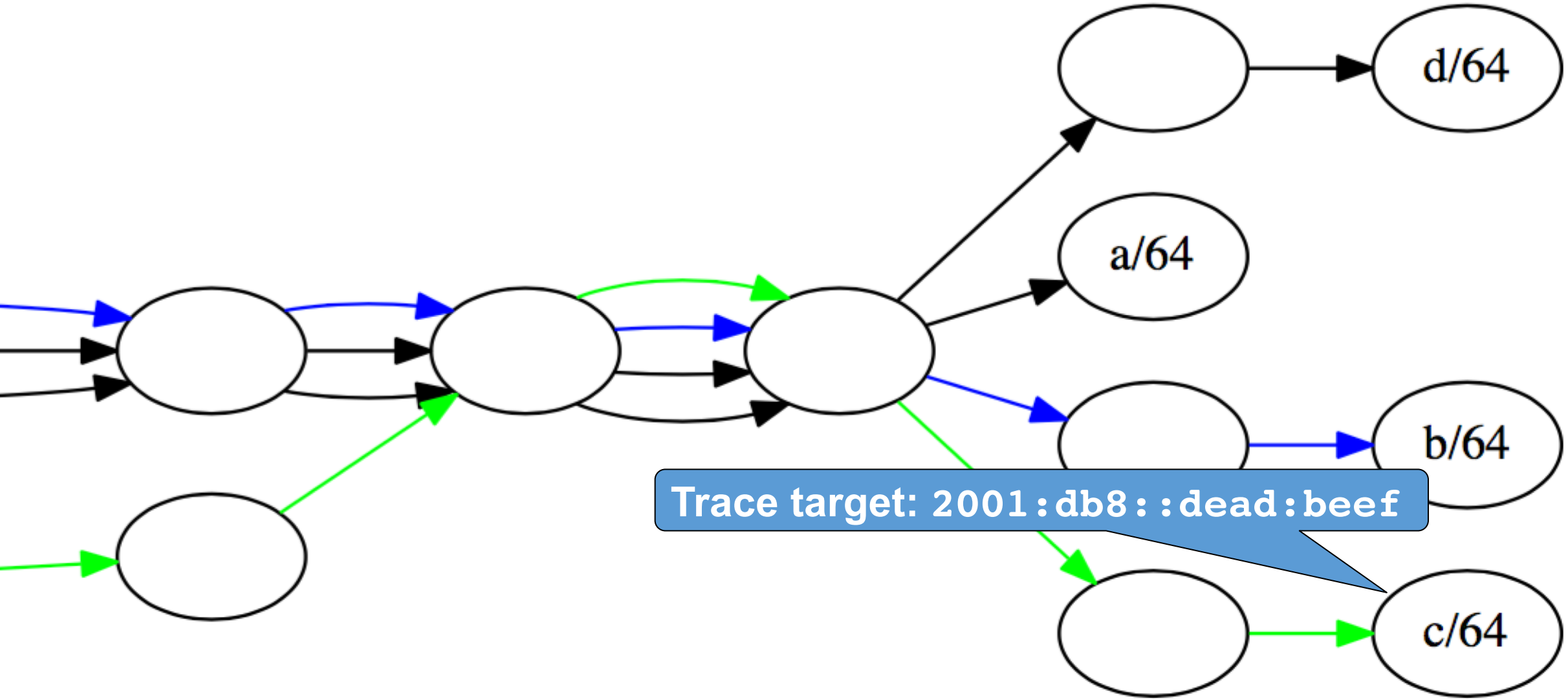
Path Divergence Analysis



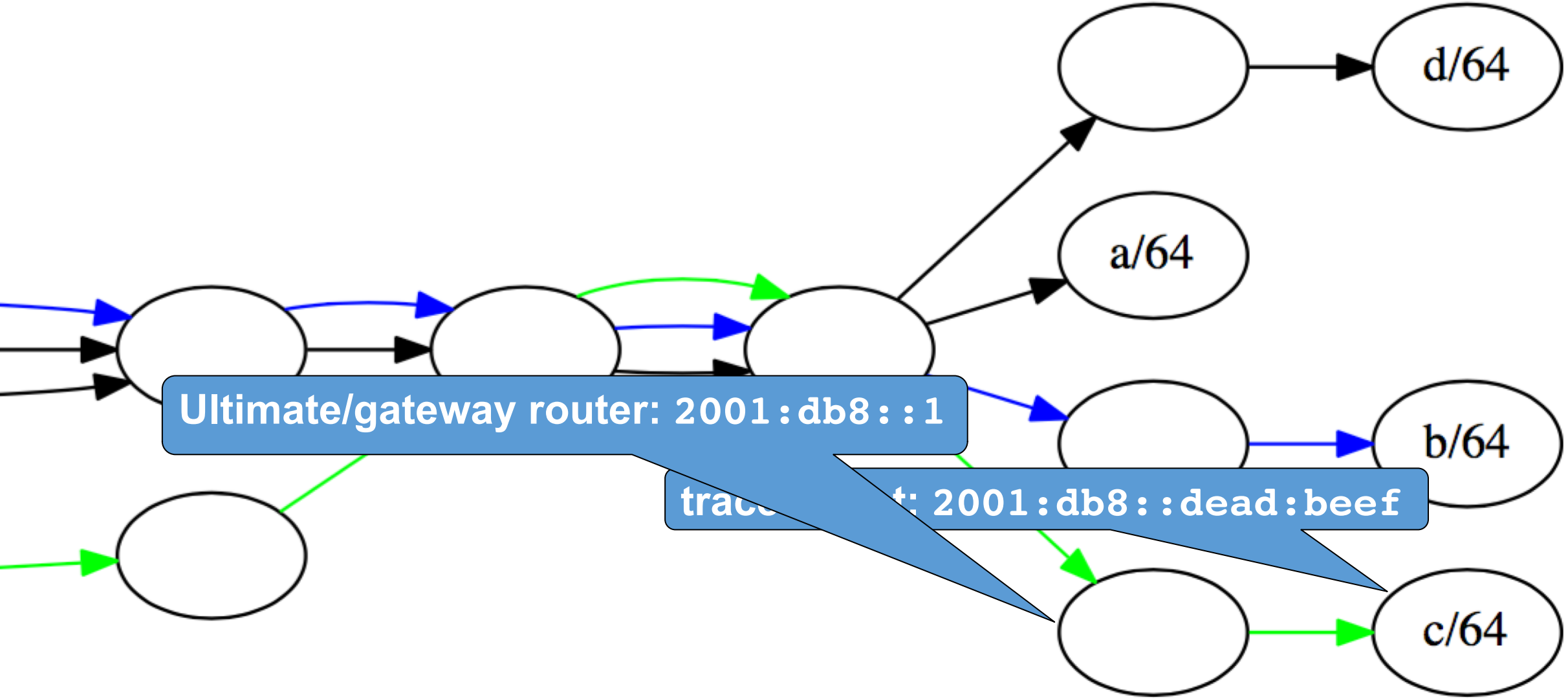
Path Divergence Analysis



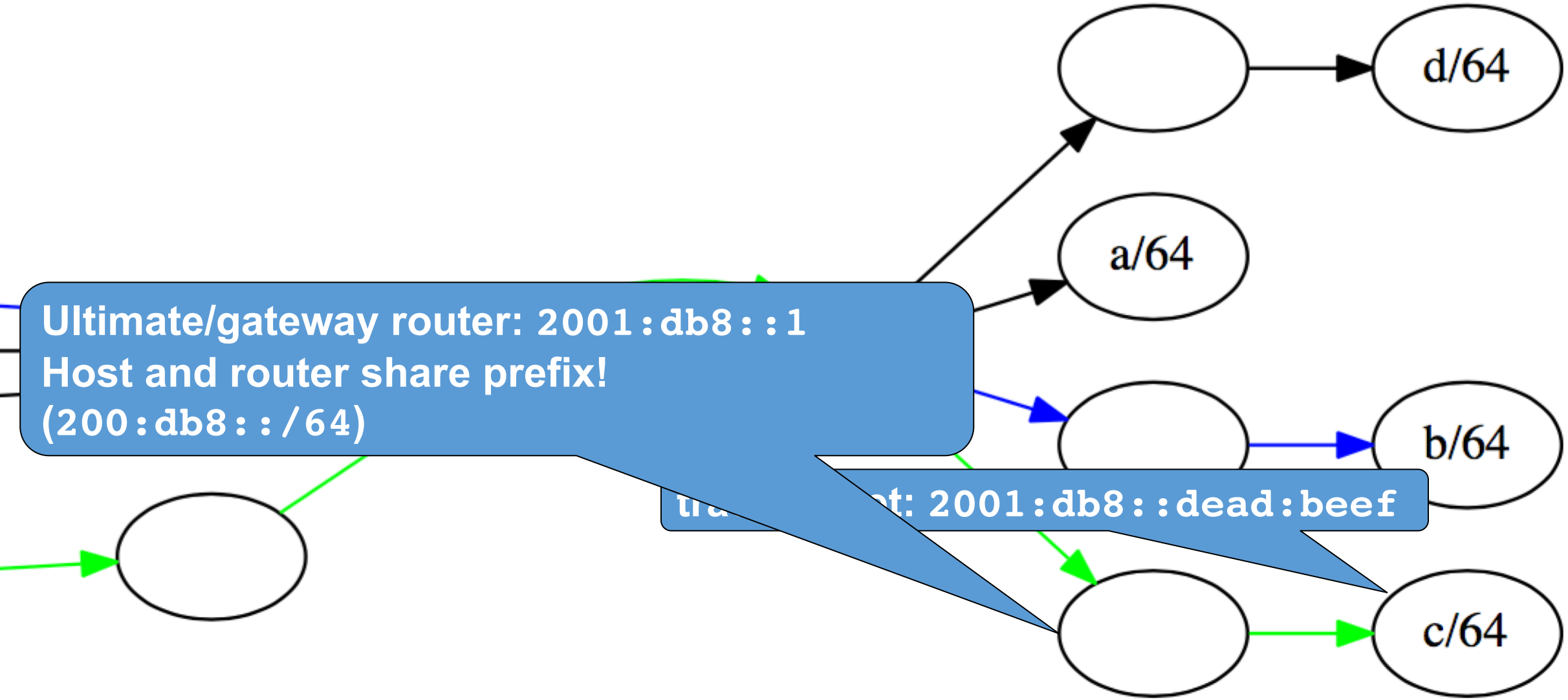
Identity Association “Hack”



Identity Association “Hack”



Identity Association “Hack”



What's left to do?

- *k*IP: Address activity changes over time, and in both v6 *and v4,* activity migrates across address space.

  A consequence of this is that anonymous aggregates have a (potentially short) lifetime, and must be continually reassessed, *e.g.,* in a sliding window of time, complicating time-series measurements based on aggregates.

- *k*IP: Perhaps surprisingly, its easier to "count" simultaneous hosts/subscribers having ephemeral privacy addresses than those having static, low numbered (*e.g.*, `::1`) or EUI-64 addresses.

  We need heuristics acceptable to the community.
  Perhaps what would be used with IPv4 will suffice.

What's left to do?

- *k*IP: Can we generate federated/coordinated anonymous aggregates from many observation points, *e.g.,* CDNs, in MAP and/or SMART Research Groups?

- What work is there to do regarding attacks on privacy?

Thanks! Questions?

David Plonka <plonka@akamai.com|dave@plonka.us>

Measurement and Analysis for Protocols Research Group (MAPRG)

For details, search for “maprg” in Google ☺
https://irtf.org/maprg

https://trac.ietf.org/trac/irtf/wiki/map