*k*IP: a Measured Approach to IPv6 Address Anonymization

MAPRG Meeting – Prague, July 20, 2017

David Plonka <plonka@akamai.com>

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

IP Address Anonymization

- **Today we'll only consider truncation and/or aggregation-based anonymization** *e.g.,* for correlating web analytics with network topology, routing, service providers, and geographic locations.

Background: IPv4 Address Anonymization by aggregation

```
10.0.42.24 1
10.0.42.30 1
10.0.42.25 1
10.0.42.6  1
10.0.42.17 1
10.0.42.17 1
10.0.42.9  1
10.0.42.19 1
10.0.42.29 1
10.0.42.26 1
10.0.42.11 1
10.0.42.27 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
10.0.42.31 1
10.0.42.10 1
10.0.42.22 1
10.0.42.16 1
10.0.42.4  1
10.0.42.21 1
10.0.42.8  1
10.0.42.20 1
10.0.42.3  1
10.0.42.14 1
10.0.42.1  1
10.0.42.15 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"]

Characteristics of the data sets

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	Active addresses (7 days)
Meeting Network	1	3	15.4K
EU ISP	**163K**	**21.4M**	**125M**
JP ISP	**2.46M**	**2.46M**	**72.2M**
US ISP	**8.16K**	**2.42M**	**84.5M**

Characteristics of the data sets: *no aggregation?*

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	Active addresses (7 days)
Meeting Network	1	3	15.4K
EU ISP	**163K**	**21.4M**	**125M**
JP ISP	**2.46M**	**2.46M**	**72.2M**
US ISP	**8.16K**	**2.42M**	**84.5M**

Characteristics of the data sets: *bias?*

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	Active addresses (7 days)
Meeting Network	1	3	15.4K
EU ISP	**163K**	**21.4M**	**125M**
JP ISP	**2.46M**	**2.46M**	**72.2M**
US ISP	**8.16K**	**2.42M**	**84.5M**

Characteristics of the data sets: *comparably sized?*

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	Active addresses (7 days)
Meeting Network	1	3	15.4K
EU ISP	**163K**	**21.4M**	**125M**
JP ISP	**2.46M**	**2.46M**	**72.2M**
US ISP	**8.16K**	**2.42M**	**84.5M**

Characteristics of the data sets: *comparably sized?*

Data set	Active /48 prefixes (7 days)	Active /64 prefixes (7 days)	Active addresses (7 days)
Meeting Network	1	3	15.4K
EU ISP	**163K**	**21.4M**	**125M**
JP ISP	**2.46M**	**2.46M**	**72.2M**
US ISP	**8.16K**	**2.42M**	**84.5M**

*k*IP: a measurement-based approach…

1. **Temporal & Spatial Address *Classification*:**
   “address dendrachonology”

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

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

Step 1. Classification: address dendrachronology

introduced in "IPv6 Prefix Intelligence," MAPRG Meeting, April 2016

Classification: Discarding [Personally Identifiable] Information

```
20010db8000e000000172cd5fa4bd6b1 75 0d
20010db8000e0000002ae748ea083efb 75 0d
20010db8000e0000005d58e18441347a 79 1d
20010db8000e0000005f1dd3864f2d03 79 0d
20010db8000e000000872ce4d7e0d16c 76 0d
```

… (1594 more addresses) …

```
20010db8000e0000fdbefa6dce8d096c 80 1d
20010db8000e0000fdbf6e62e74a33a4 80 1d
20010db8000e0000fdd4f4f54264cc52 75 0d
20010db8000e0000fdf73310ae0043da 75 2d
20010db8000e0000feedfacedeadbabe 71 3d
```

Spatial Characteristic:
Discriminating Prefix Length (DPL)

Classification: Discarding [Personally Identifiable] Information

```
20010db8000e000000172cd5fa4bd6b1 75 0d
20010db8000e0000002ae748ea083efb 75 0d
20010db8000e0000005d58e18441347a 79 1d
20010db8000e0000005f1dd3864f2d03 79 0d
20010db8000e000000872ce4d7e0d16c 76 0d
```

… (1594 more addresses) ...

```
20010db8000e0000fdbefa6dce8d096c 80 1d
20010db8000e0000fdbf6e62e74a33a4 80 1d
20010db8000e0000fdd4f4f54264cc52 75 0d
20010db8000e0000fdf73310ae0043da 75 2d
20010db8000e0000feedfacedeadbabe 71 3d
```

Temporal Characteristic: Stable Days (SD)

Spatial Characteristic: Discriminating Prefix Length (DPL)

Classification: Discarding [Personally Identifiable] Information

```
20010db8000e000000172cd5fa4bd6b1 75 0d
20010db8000e0000002ae748ea083efb 75 0d
20010db8000e0000005d58e18441347a 79 1d
20010db8000e0000005f1dd3864f2d03 79 0d
20010db8000e000000872ce4d7e0d16c 76 0d
```

… (1594 more addresses) …

```
20010db8000e0000fdbefa6dce8d096c 80 1d
20010db8000e0000fdbf6e62e74a33a4 80 1d
20010db8000e0000fdd4f4f54264cc52 75 0d
20010db8000e0000fdf73310ae0043da 75 2d
20010db8000e0000feedfacedeadbabe 71 3d
```

Stateless Classification:
(from F. Gont's IPv6 Toolkit)

```
$ addr6 –a 20010db8000e0000feedfacedeadbabe
unicast=global=global=randomized=unspecified
```

Classification: Discarding [Personally Identifiable] Information

```
20010db8000e000000172cd5fa4bd6b1 75 0d
20010db8000e0000002ae748ea083efb 75 0d
20010db8000e0000005d58e18441347a 79 1d
20010db8000e0000005f1dd3864f2d03 79 0d
20010db8000e000000872ce4d7e0d16c 76 0d
```

… (1594 more addresses) ...

```
20010db8000e0000fdbefa6dce8d096c 80 1d
20010db8000e0000fdbf6e62e74a33a4 80 1d
20010db8000e0000fdd4f4f54264cc52 75 0d
20010db8000e0000fdf73310ae0043da 75 2d
20010db8000e0000feedfacedeadbabe 71 3d
```

Truncate here?

Classification: Discarding [Personally Identifiable] Information

```
20010db8000e000000172cd5fa4bd6b1 75 0d
20010db8000e0000002ae748ea083efb 75 0d
20010db8000e0000005d58e18441347a 79 1d
20010db8000e0000005f1dd3864f2d03 79 0d
20010db8000e000000872ce4d7e0d16c 76 0d
```

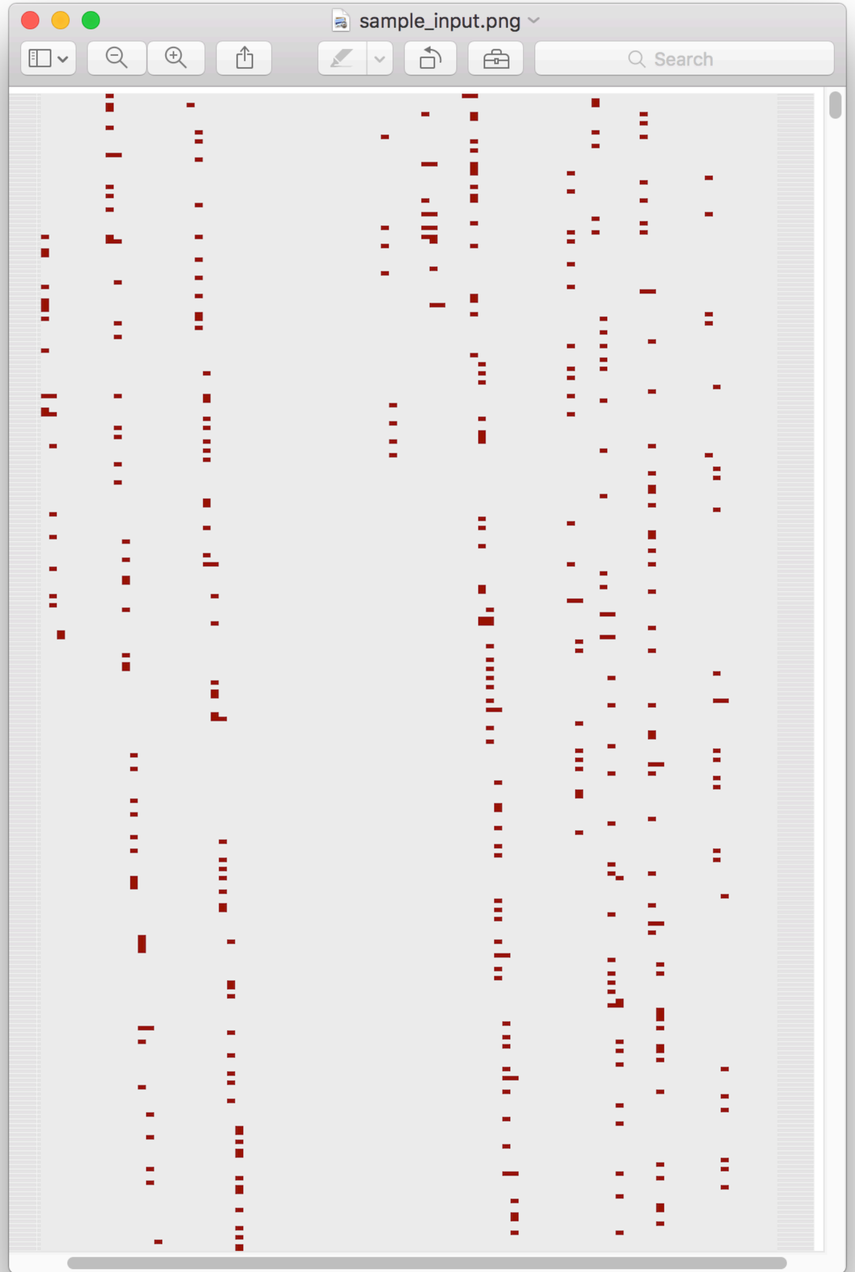
… (1594 more addresses) …

```
20010db8000e0000fdbefa6dce8d096c 80 1d
20010db8000e0000fdbf6e62e74a33a4 80 1d
20010db8000e0000fdd4f4f54264cc52 75 0d
20010db8000e0000fdf73310ae0043da 75 2d
20010db8000e0000feedfacedeadbabe 71 3d
```

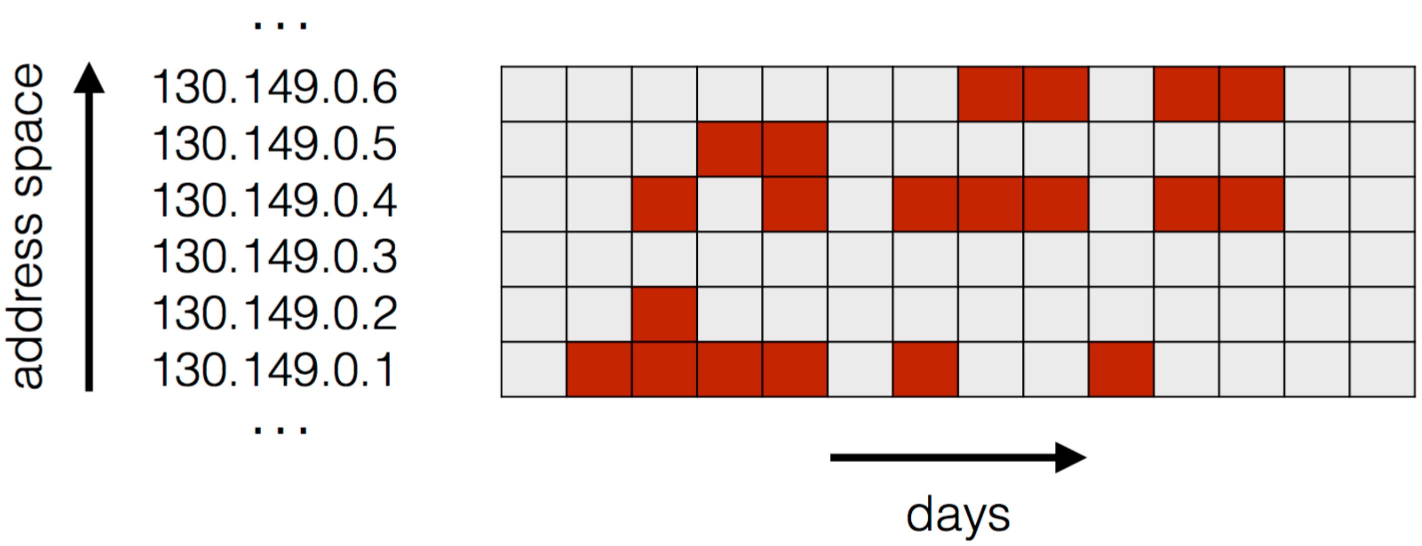
Truncate here?

Or here?

Step 2. Address Activity Matrix Analysis



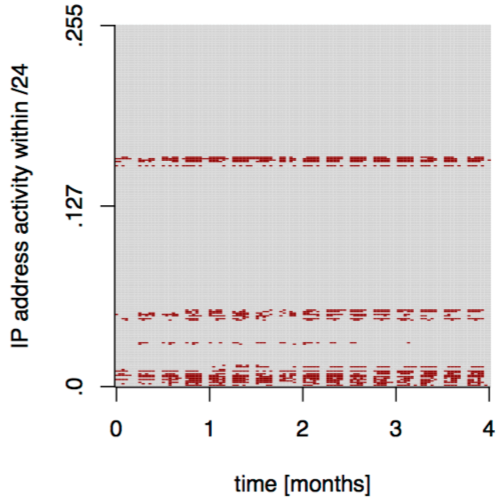
Related Work: IPv4 Address Activity Matrix introduced in “Beyond Counting …”, MAPRG Meeting July 2016



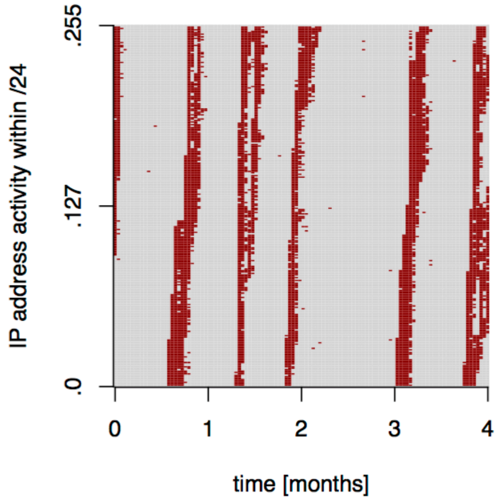
for each day on which an IP address was active (requested content), we draw a red dot.

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

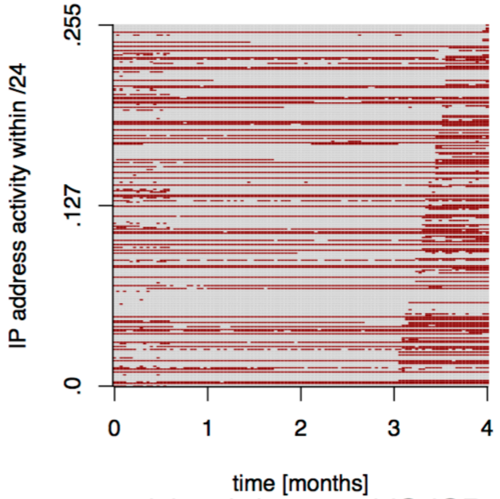
Related Work: IPv4 Address Activity Matrix



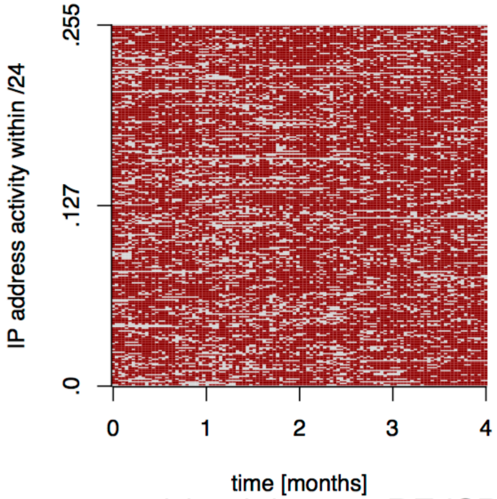
static block DE University



DHCP pool US University



residential users US ISP



residential users DE ISP

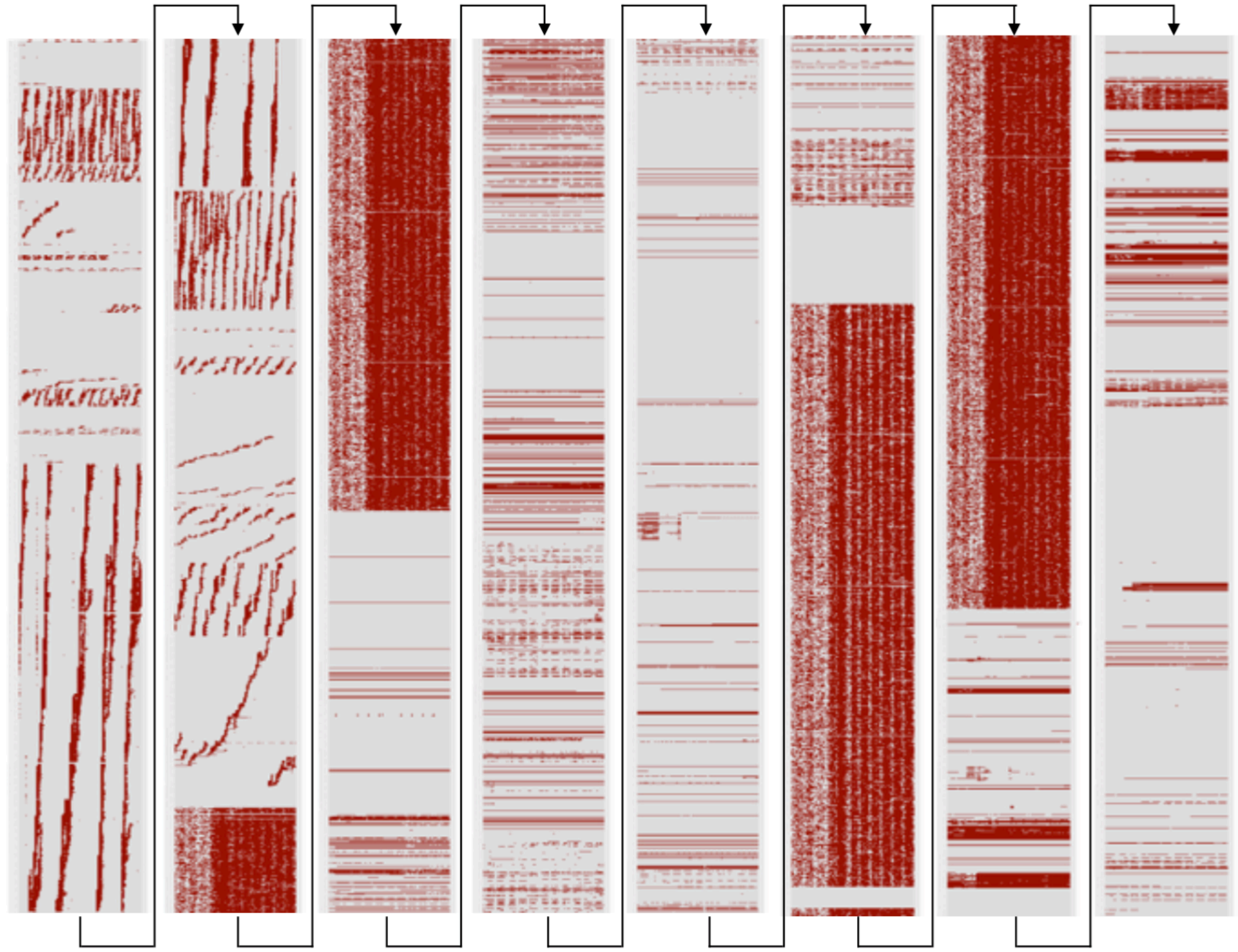
“in situ” activity:
address assignment practice
+
user behavior

(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

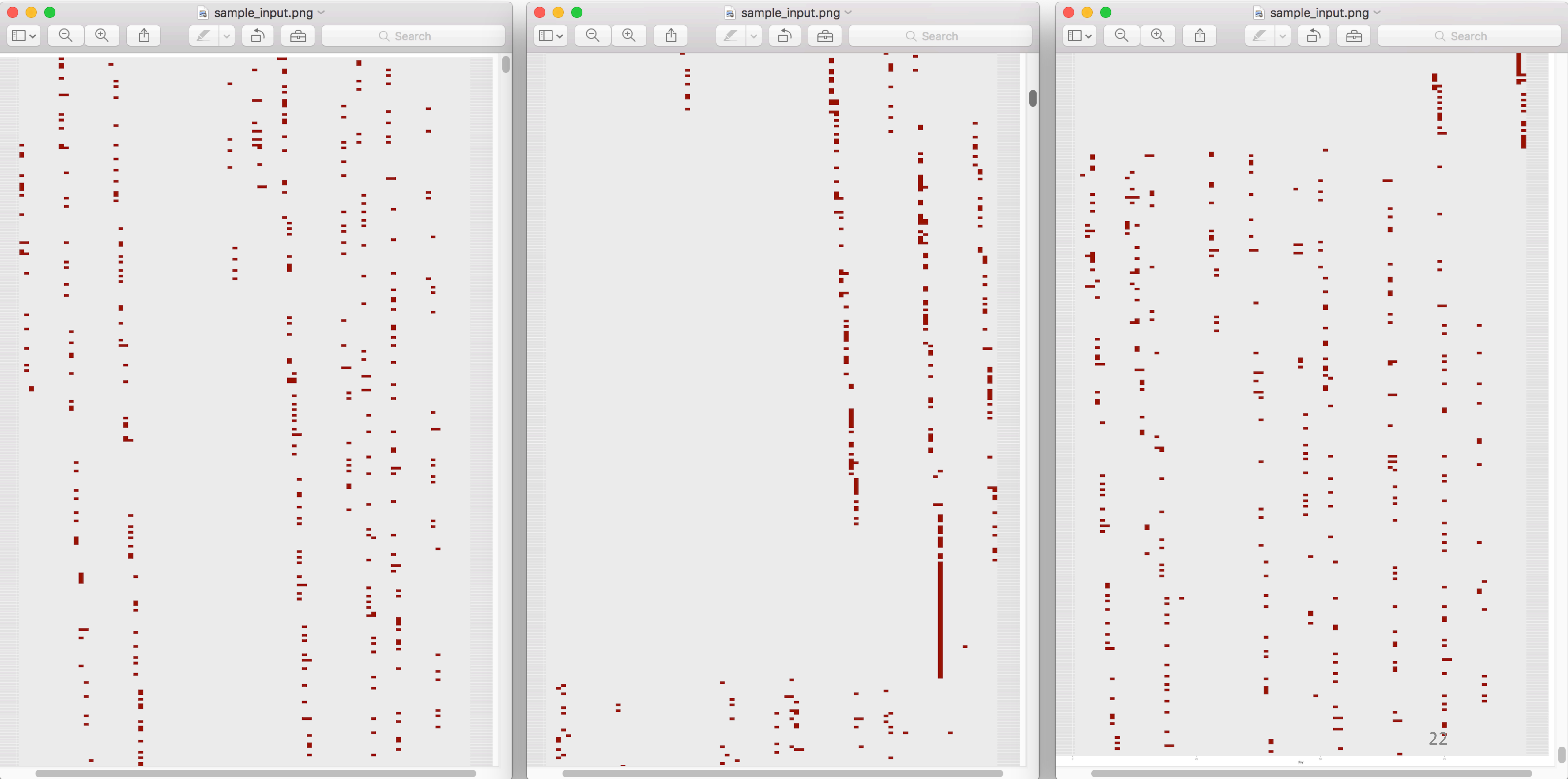
Related Work: IPv4 Address Activity Matrix

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



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 |--#----+-------+-------
200   b823000a00 16 Temporary SLAAC: 100.00% stable: 0.00%
1
#               unted during the given hour
```

/64 prefix

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 |--#----+-------+-------
```

/64 prefix

IID

```
2001 b823000a00 16 Te    rary SLAAC: 100.00% stable: 0.00%
1
#              unted d       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:db8::/64 16; Temporary SLAAC: 100% stable: 0.00%
legend:
# = activity counted during the given hour
```

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

```
                                       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
```

Space

Time

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
```

Time

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

```
                                          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---
                                             1
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---
                                        1       2
```

```
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---
                                          1       2 3

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
```

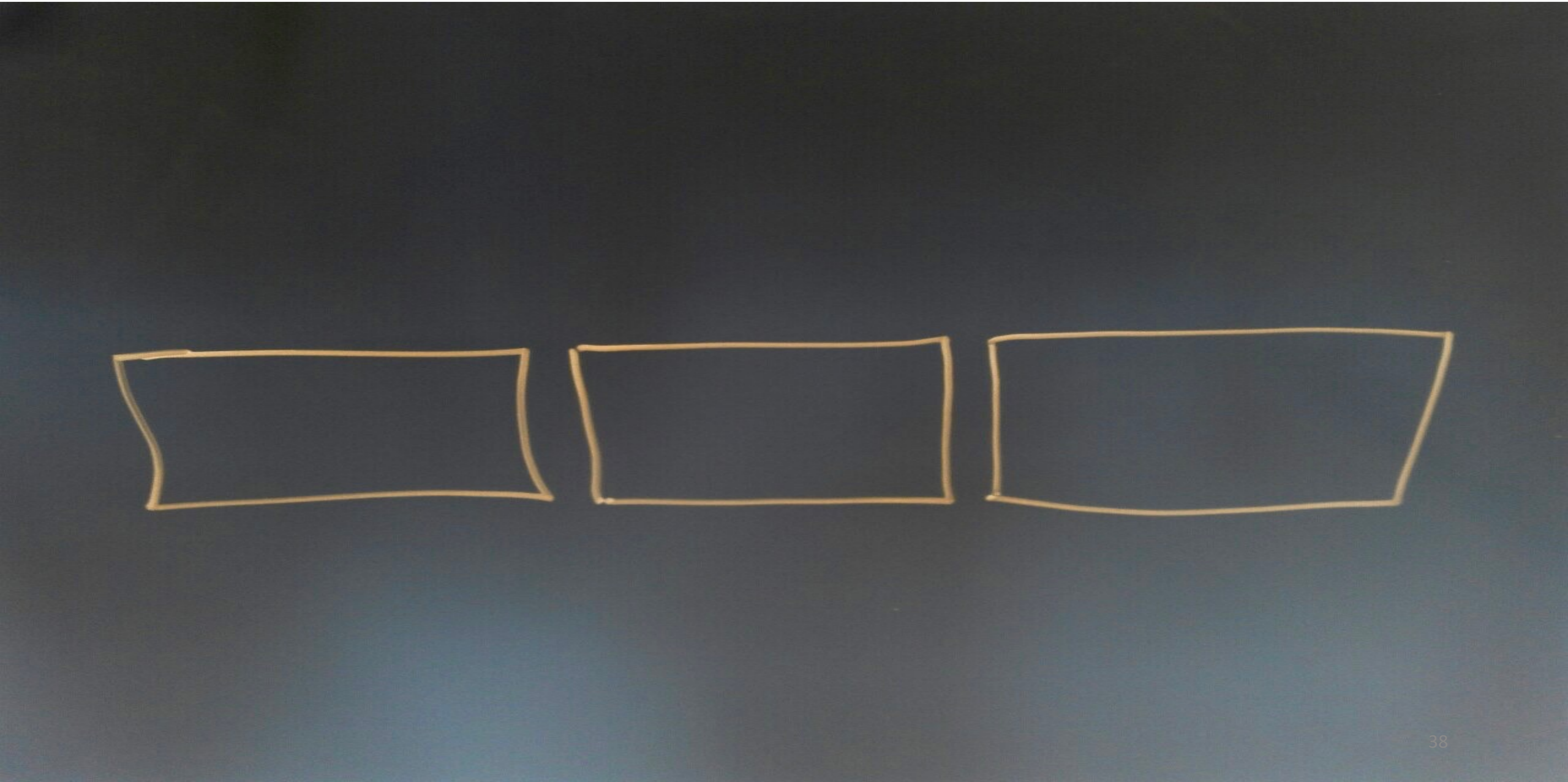
Counting Simultaneous SLAAC IIDs

```
                                        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%--------!!!!!!!!-!!!!--?
                                          000000001111111101111100? => /64 assignment @ fenceposts
legend:
! = infer /64 prefix assigned at the "fencepost" moments between intervals
```

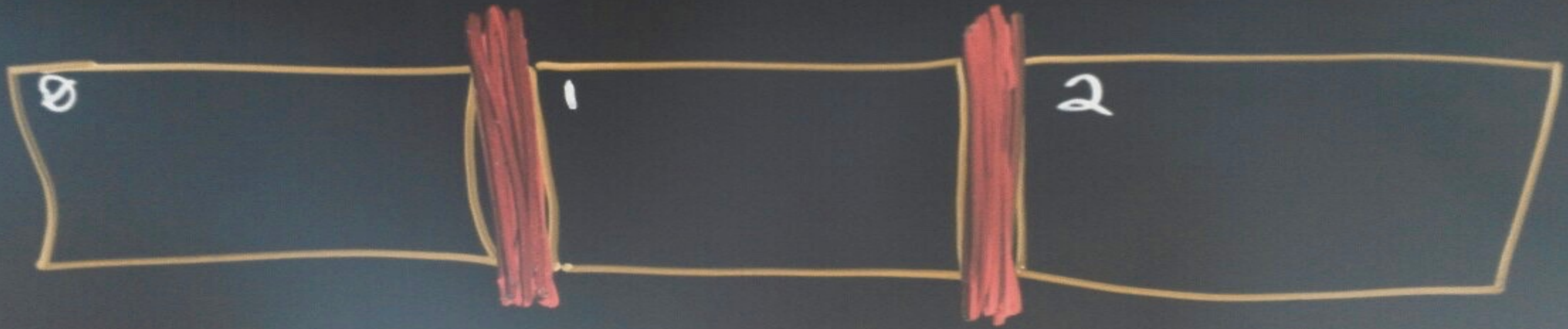
3. Synthesizing Anonymous Aggregates

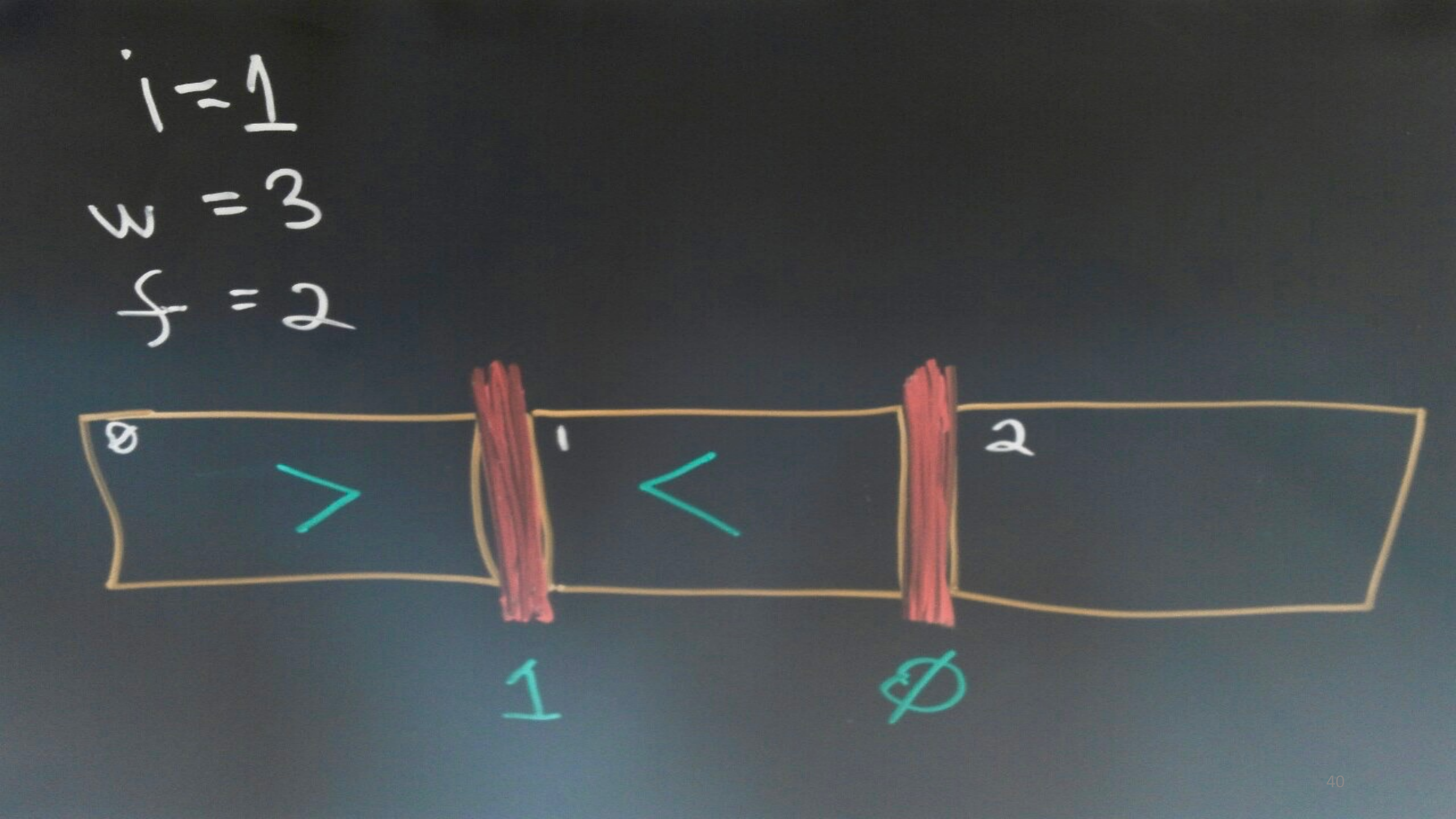


$i = 1$

$w = 3$

$f = 2$

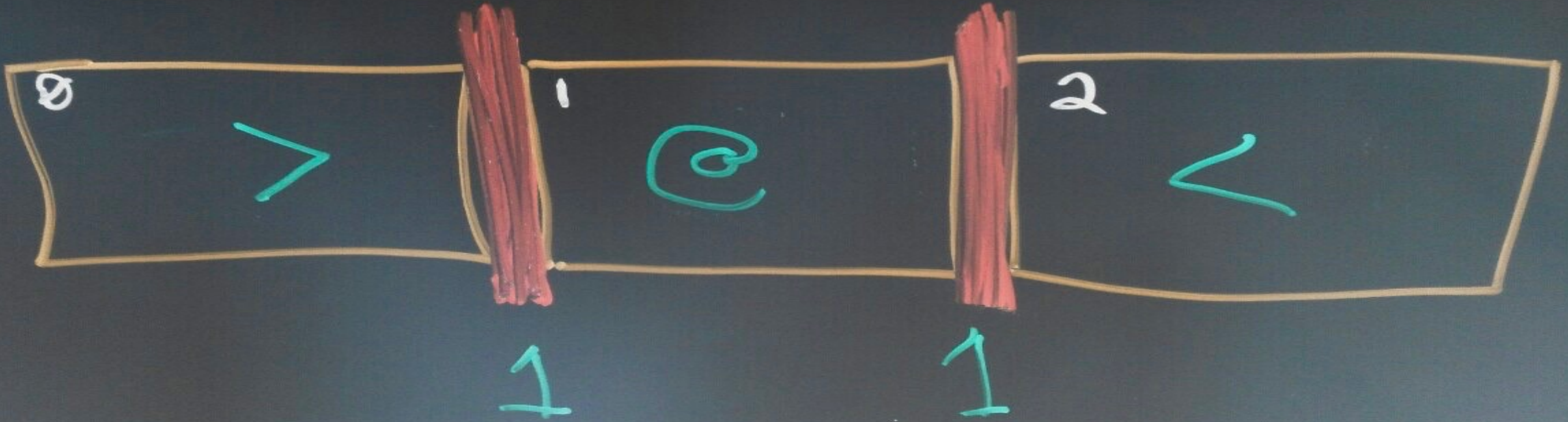



i = 1
w = 3
f = 2
0
1
2
>
<
1
Ø

$i = 1$

$w = 3$

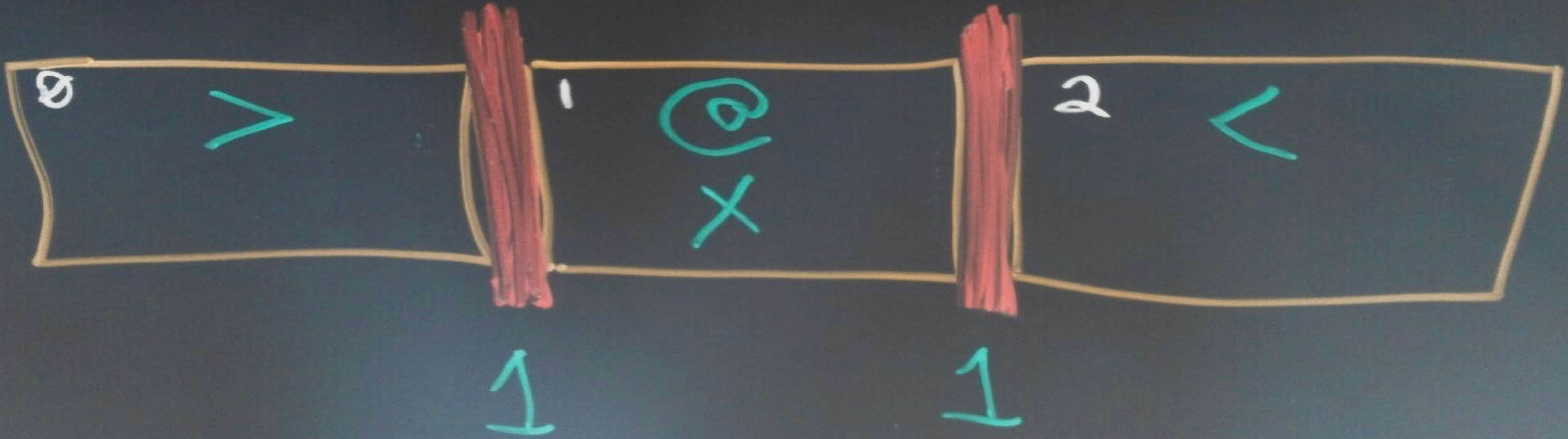
$f = 2$

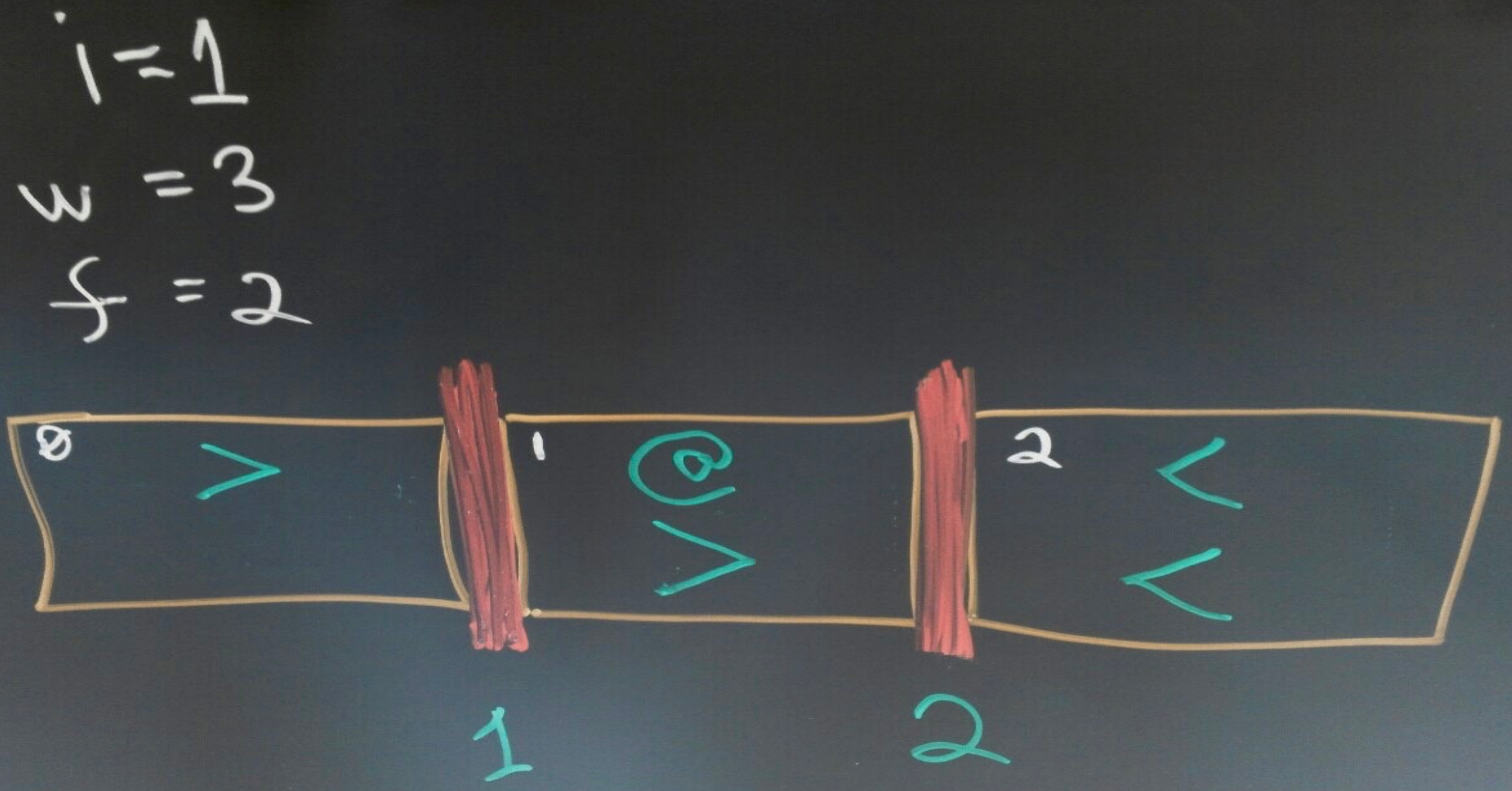


$i = 1$

$w = 3$

$f = 2$

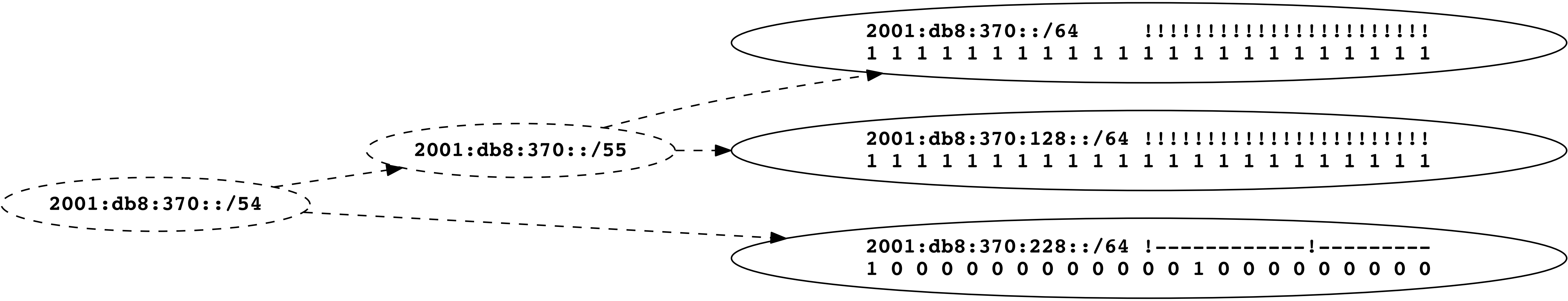



i=1
w=3
f=2
0
>
1
@
>
2
<
<
1
2

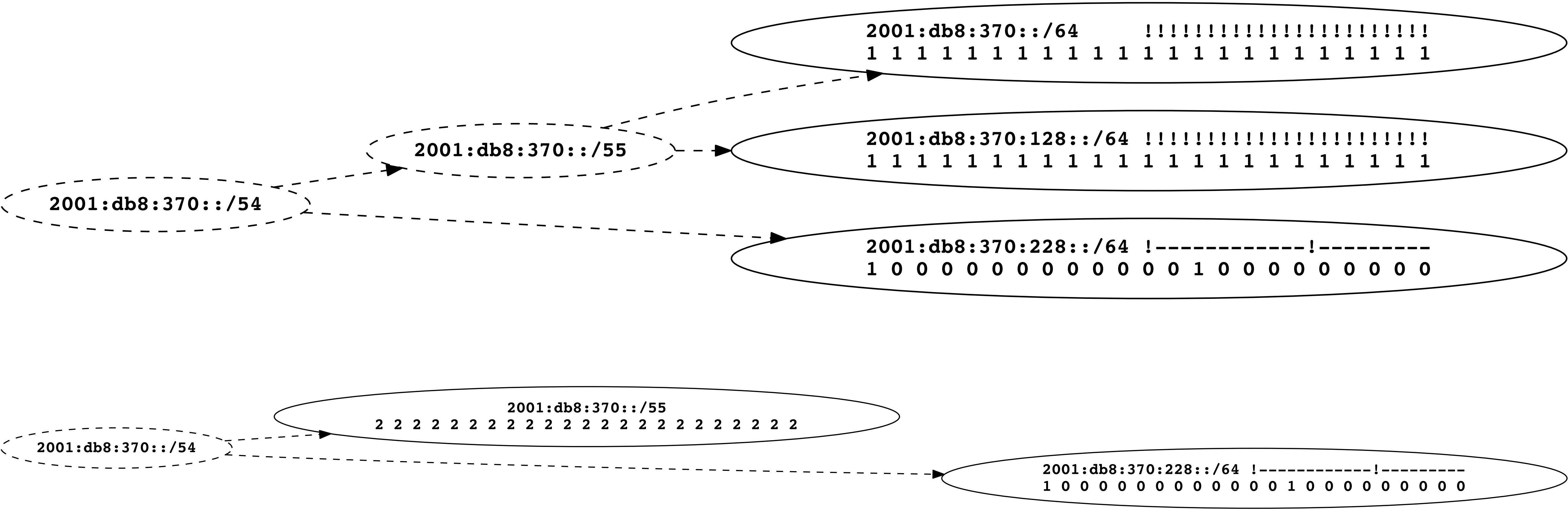
IPv6 Address Activity Matrix: Identity Assignment

```
                                         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%---     !!!!!!!! !!!! -?
                                         00000000111111110111100? => /64 assignment @ fenceposts
legend:
! = infer /64 prefix assigned at the "fencepost" moments between intervals
```

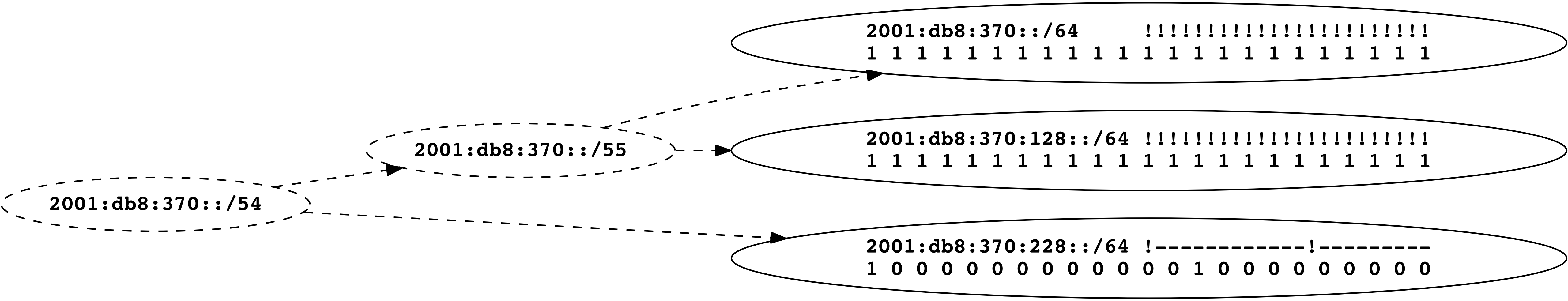
Step 3. Synthesizing Anonymous Aggregates
Example: *k=2* aggregates (w=1d, i=1h)

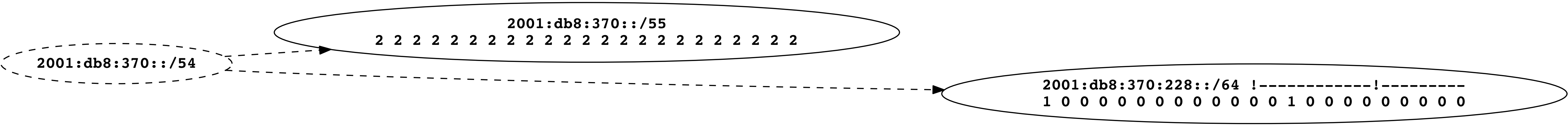


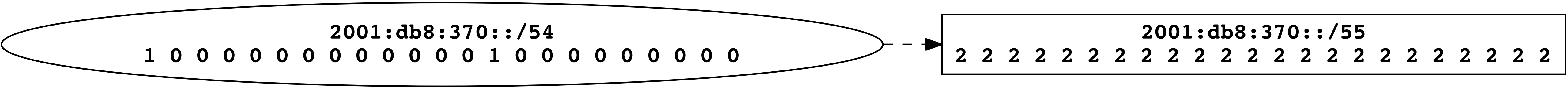
Example: Synthesizing anonymous *k=2* aggregates (w=1d, i=1h)



Example: Synthesizing anonymous *k=2* aggregates (w=1d, i=1h)







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

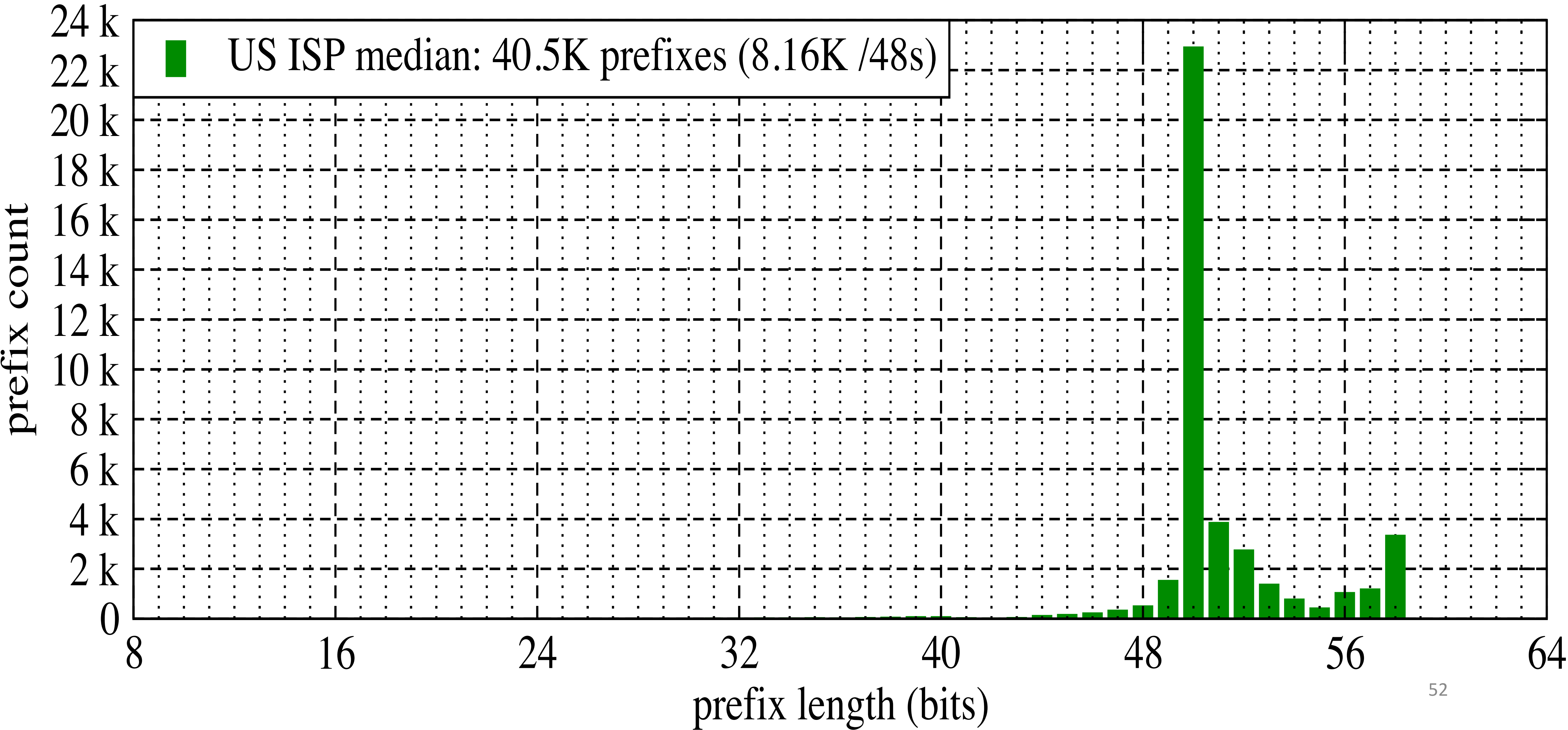
Results: simultaneously-assigned addresses and prefixes

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

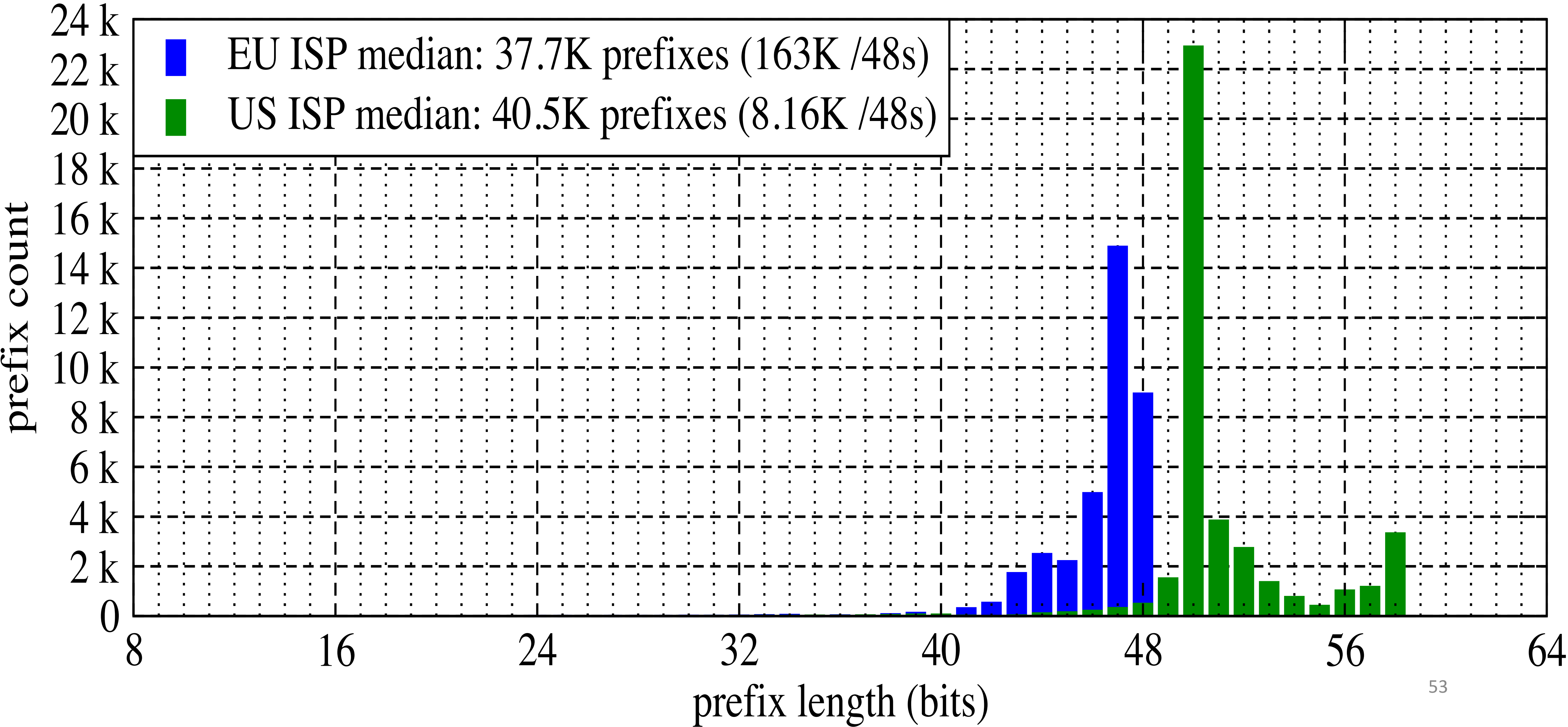
Results: simultaneously-assigned addresses and prefixes

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

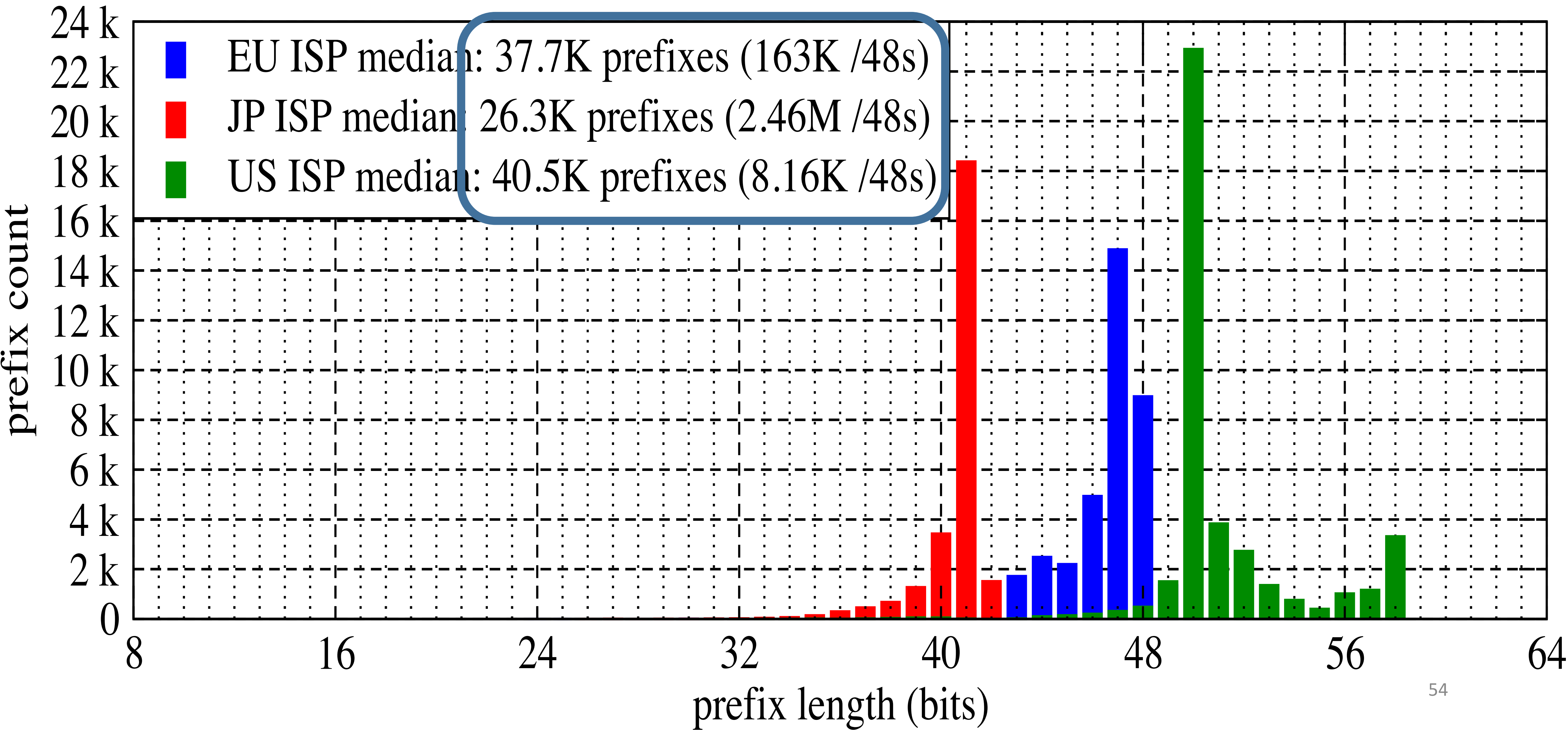
Histogram $\boxed{k=32}$ anonymous aggregate prefix lengths (w=7d, i=1h)



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



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



*k*IP: a Measured Approach to IPv6 Address Anonymization

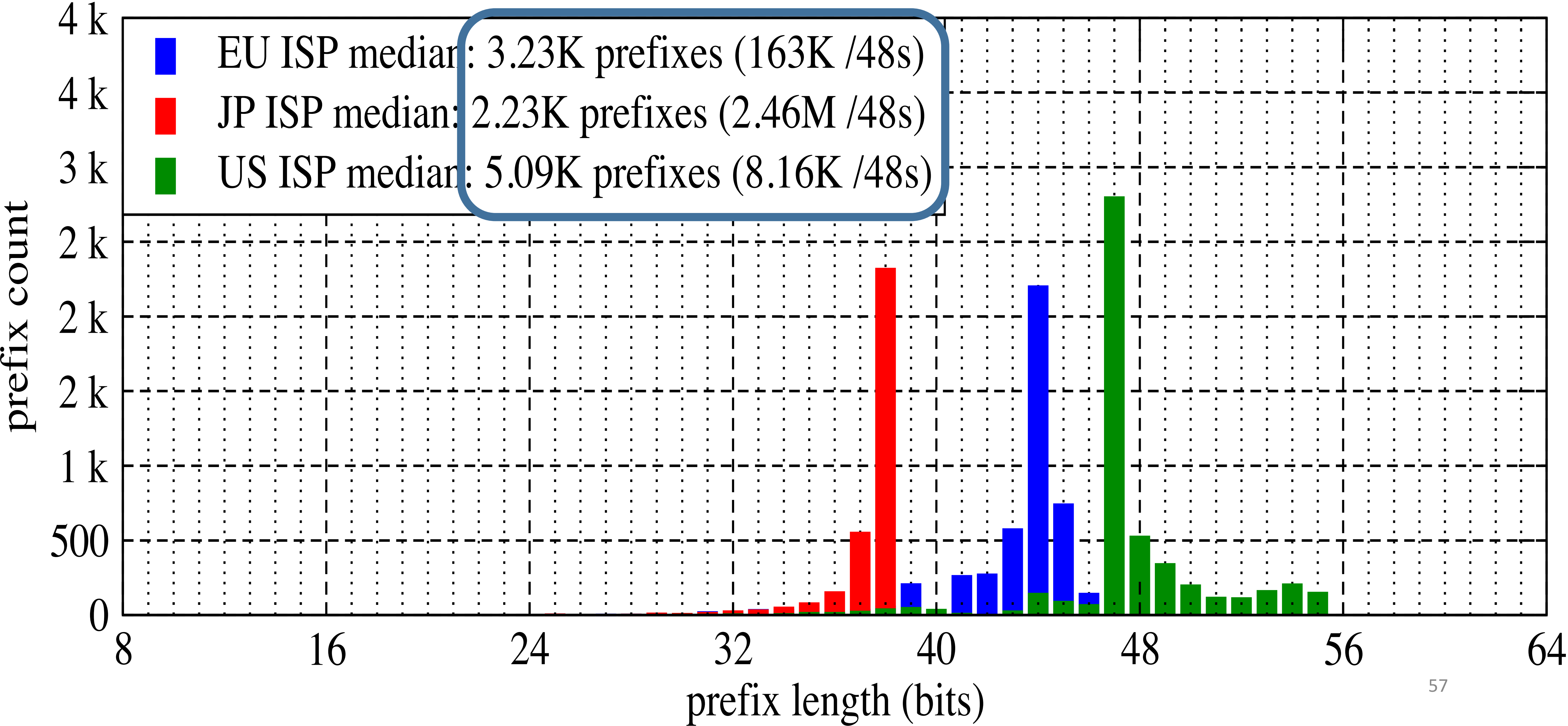
MAPRG Meeting – Prague, July 20, 2017

David Plonka <plonka@akamai.com>

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

The following are supplementary slides

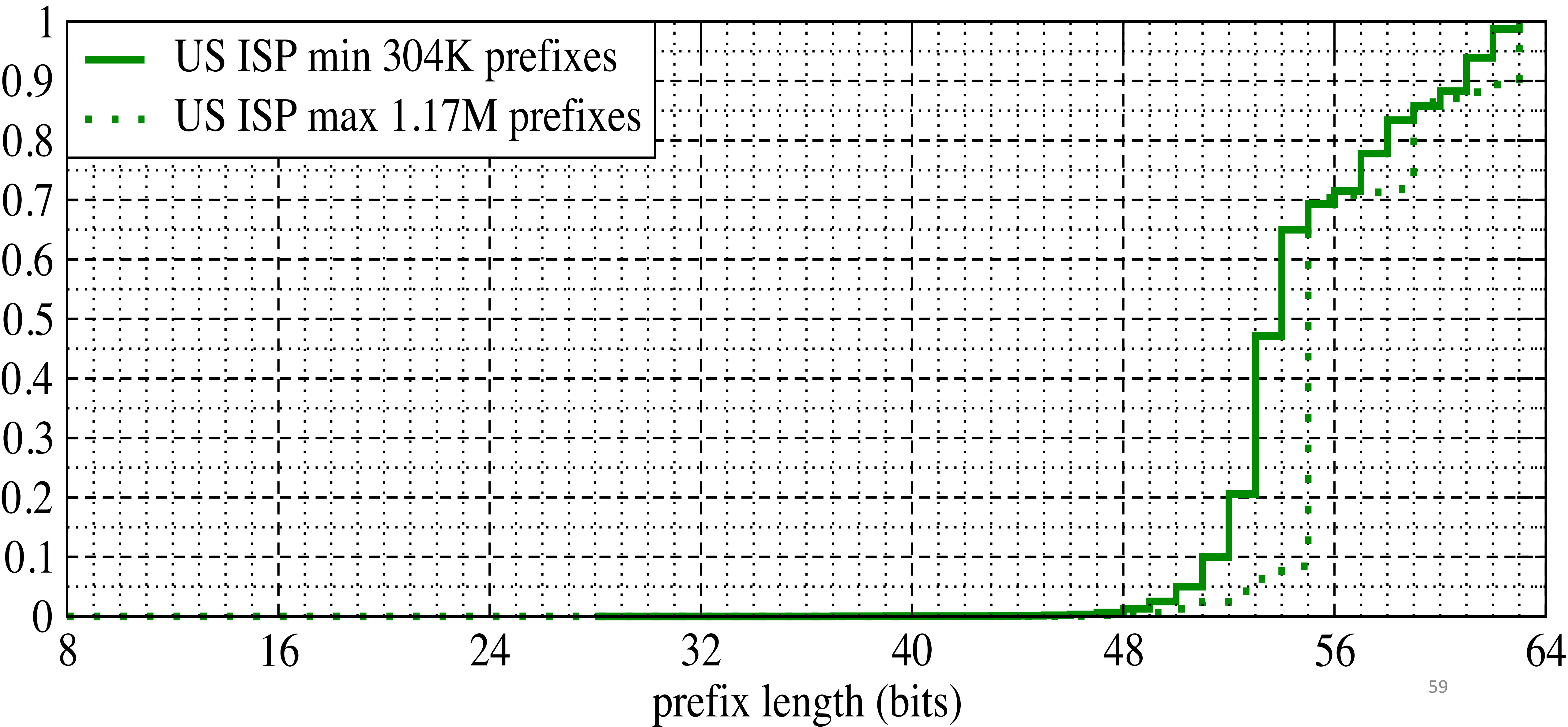
Histogram k=256 anonymous aggregate prefix lengths (w=7d, i=1h)



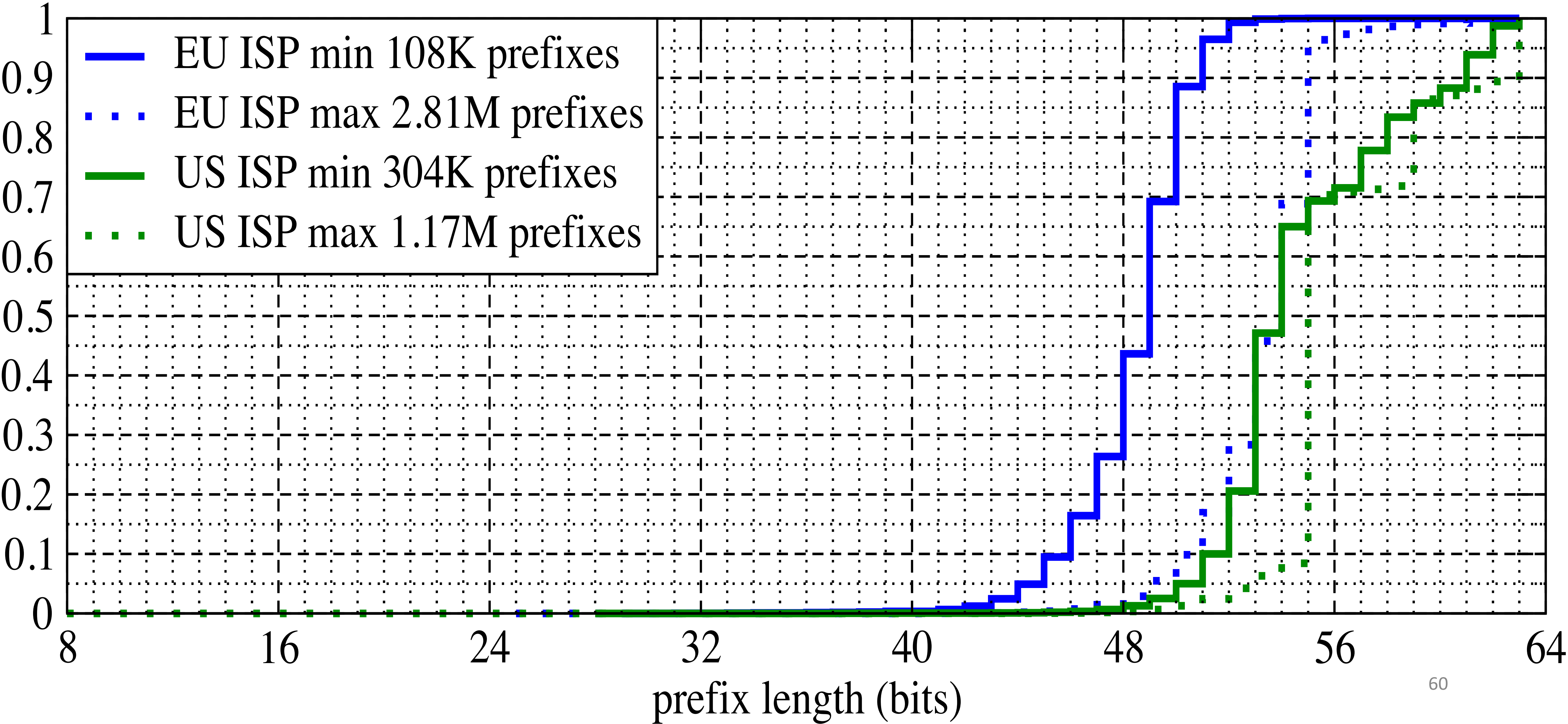
What are other applications of address activity matrix analysis and identifying simultaneously-assigned addresses?

Can we find the prefix length of an ISP's Identity Assignments (e.g., from DHCPv6 IA requests)?

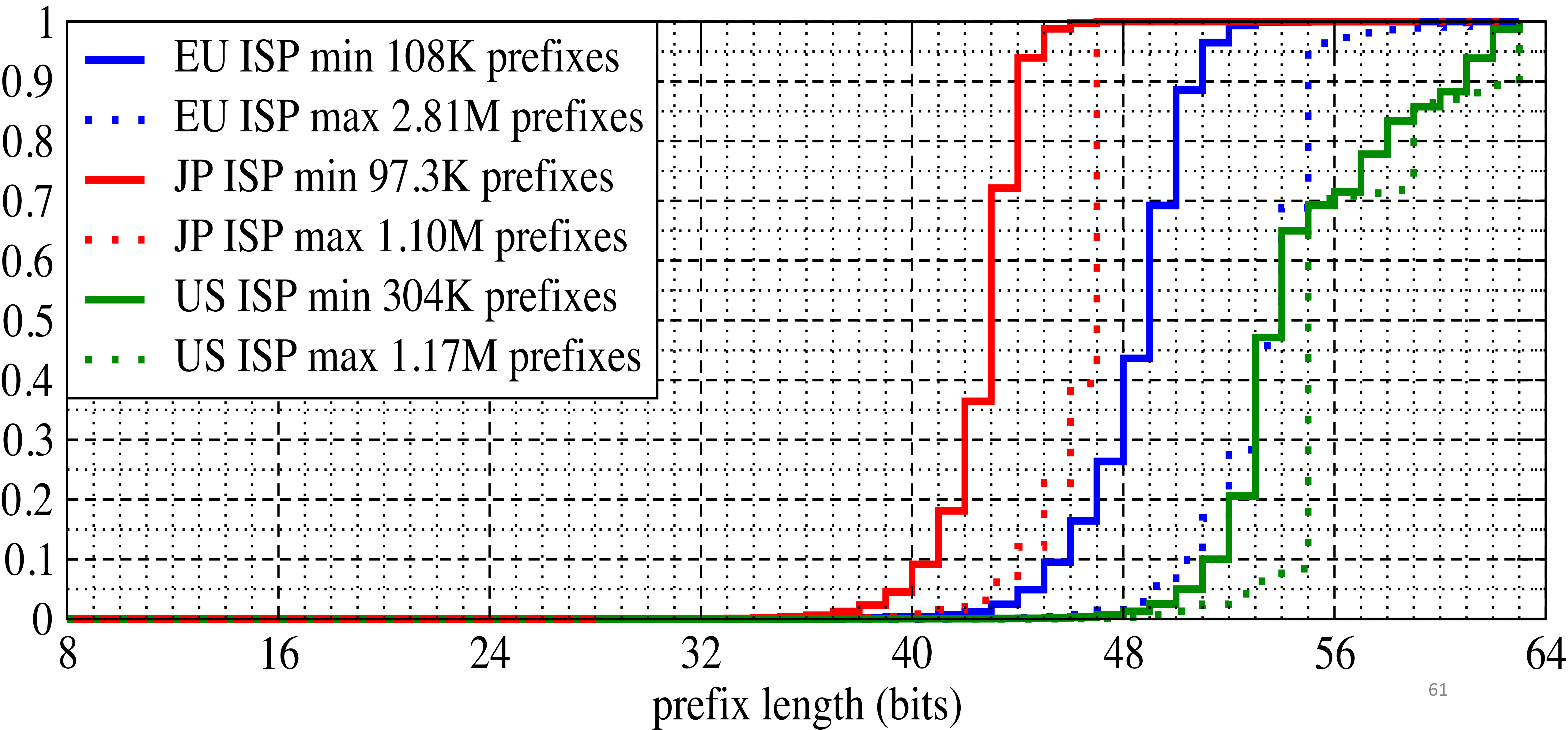
CDF $\boxed{k=2}$ anonymous aggregate prefix lengths (w=7d, i=1h)



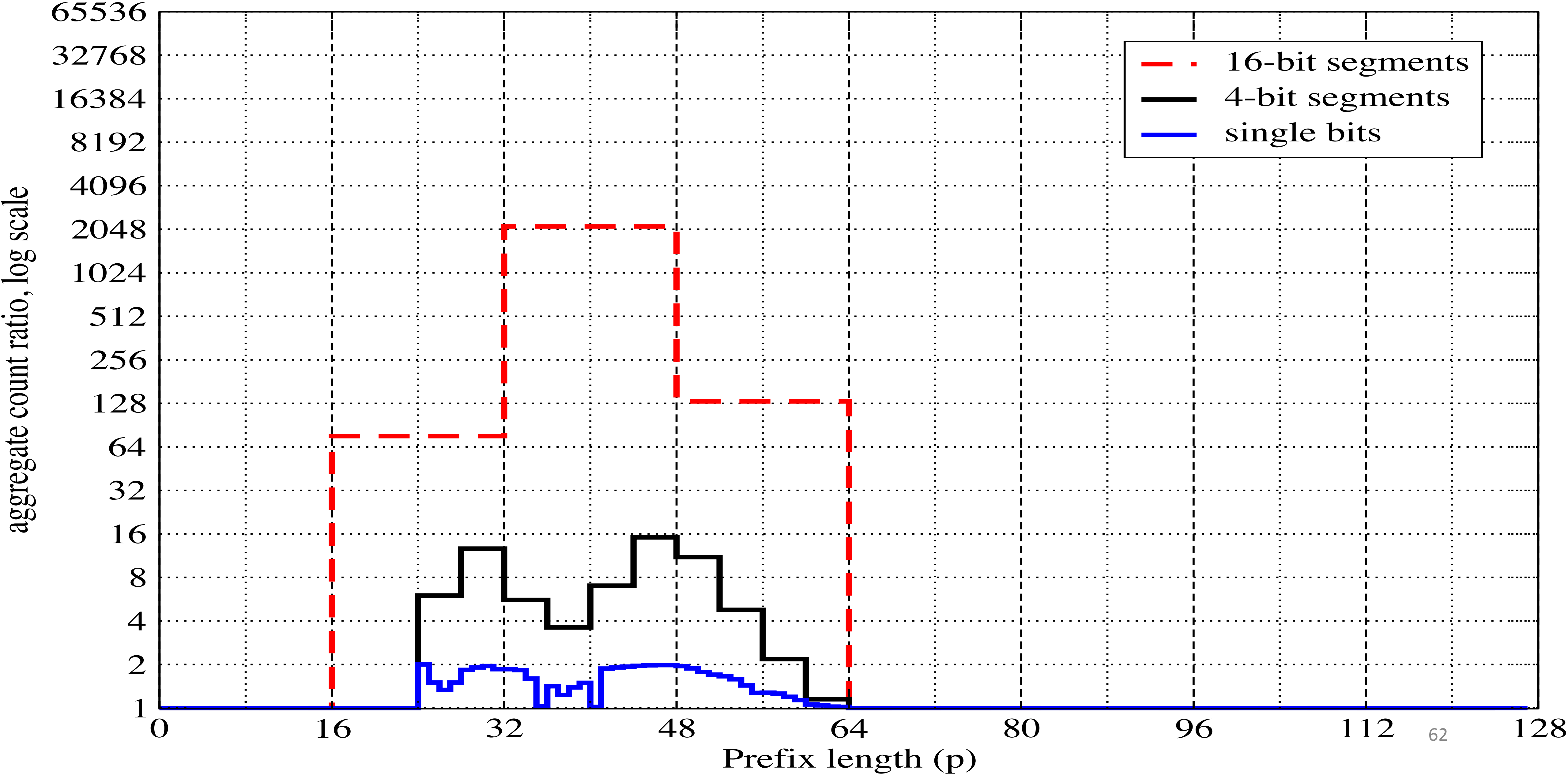
CDF: k=2 anonymous aggregate prefix lengths (w=7d, i=1h)



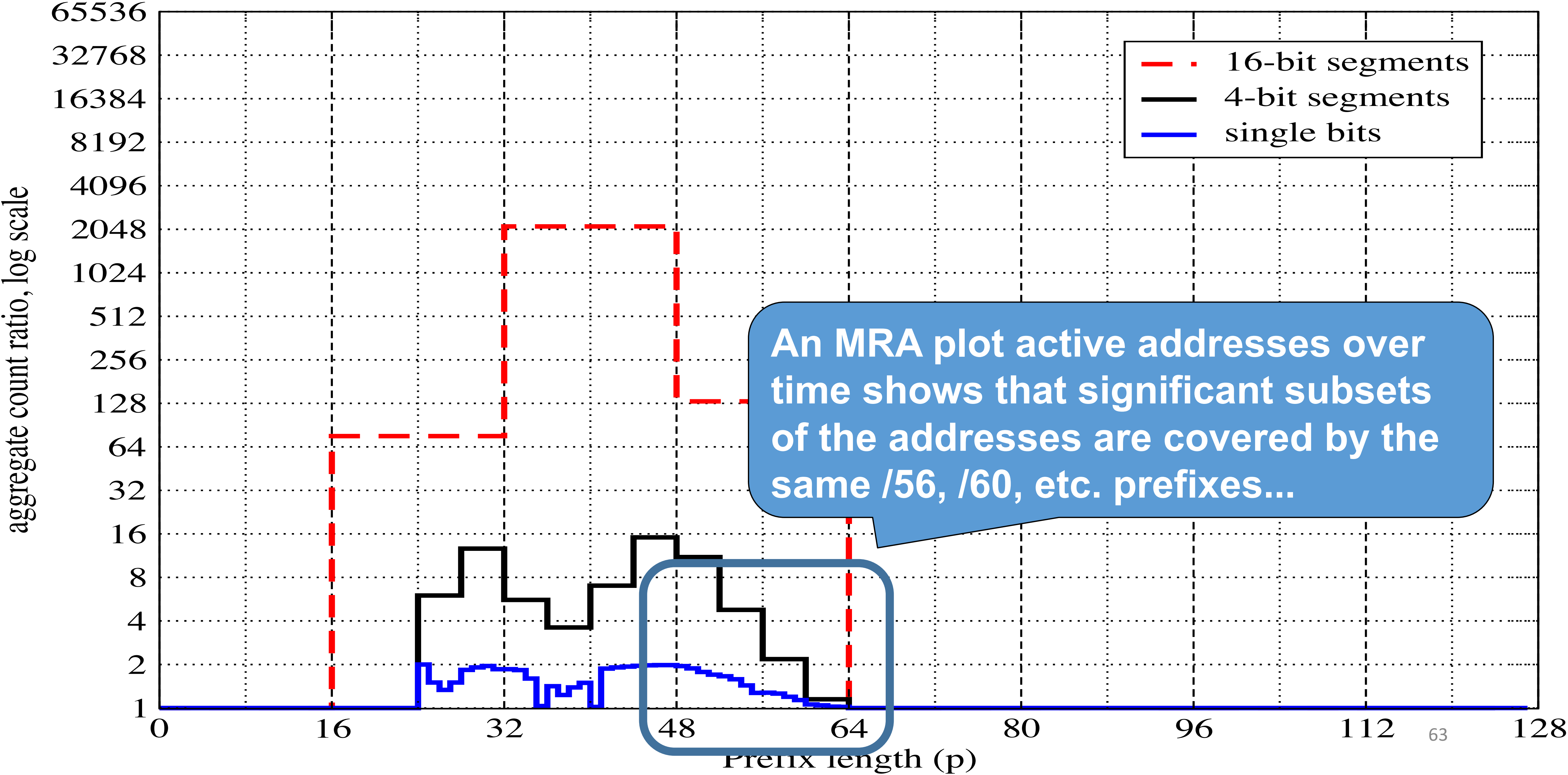
CDF: k=2 anonymous aggregate prefix lengths (w=7d, i=1h)



MRA Plot: EU ISP, 21.5M active addrs, 7 days

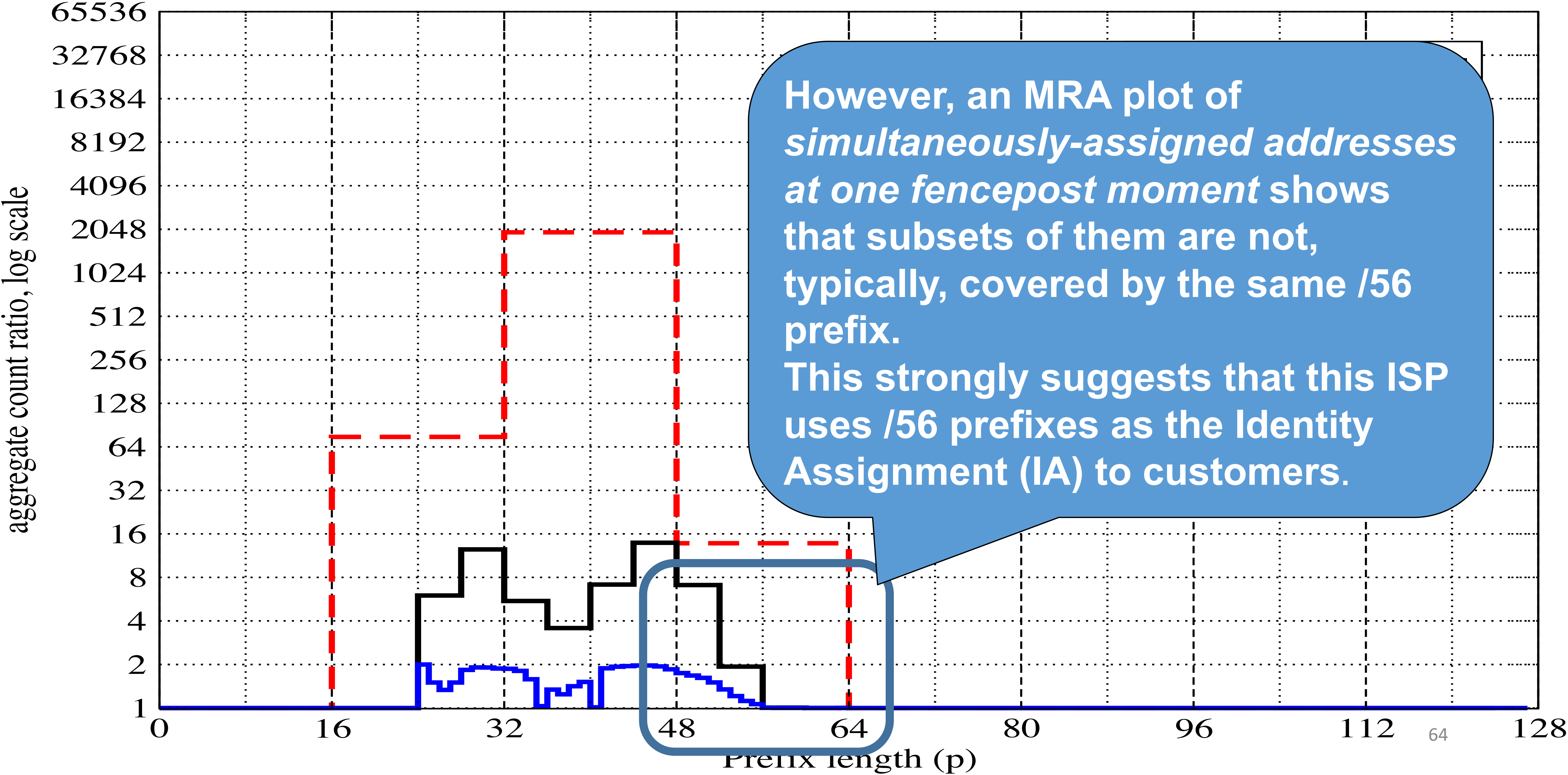


MRA Plot: EU ISP, 21.5M active addrs, 7 days



An MRA plot active addresses over time shows that significant subsets of the addresses are covered by the same /56, /60, etc. prefixes...

MRA Plot: EU ISP, 2.02M simultaneously assigned addrs, max.



However, an MRA plot of *simultaneously-assigned addresses at one fencepost moment* shows that subsets of them are not, typically, covered by the same /56 prefix.
This strongly suggests that this ISP uses /56 prefixes as the Identity Assignment (IA) to customers.