IPv6:
przestrzeń (nie) do ogarnięcia?

Paweł Foremski
IITiS PAN, pjf@iitis.pl
Farsight Security, pjf@fsi.io

Net::IP Meetup #10 @ OVH
Wrocław, 25.10.2018

IITiS

Agenda

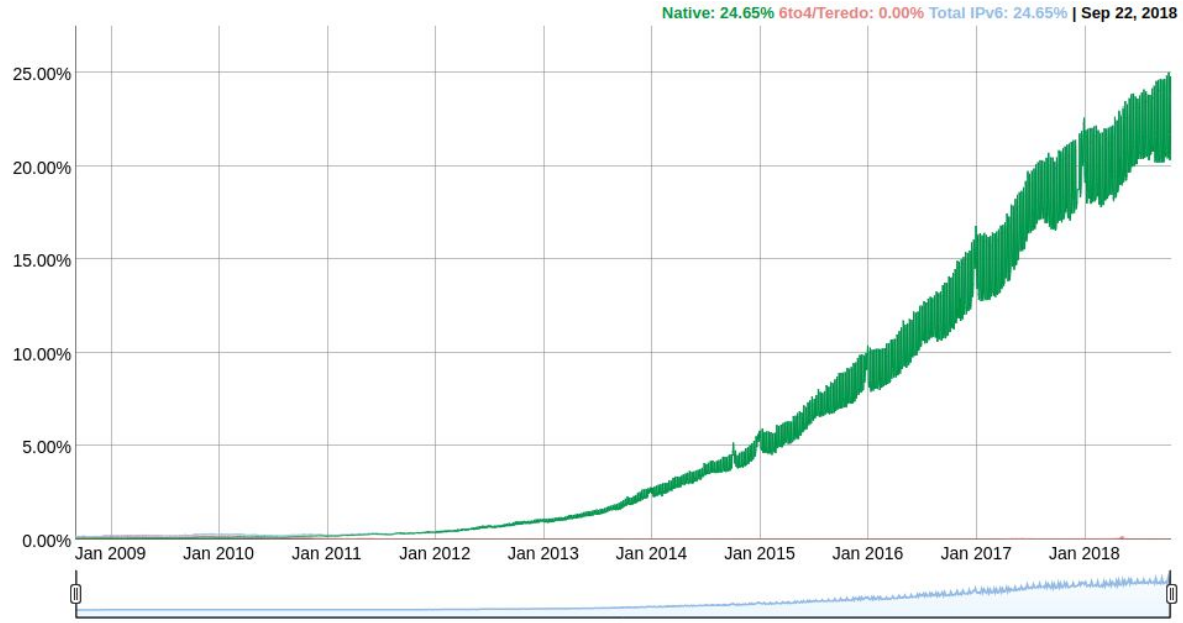
1. IPv6
   a. O RLY? Ktoś tego używa?
   b. Adresacja
   c. Funkcjonalność
   d. Bezpieczeństwo
   e. HE Tunnel Broker

2. Skanowanie IPv6
   a. Igła w stogu siana
   b. Źródła adresów
   c. DNSDB
   d. Skanowanie “hitlist”
   e. Entropy clustering
   f. Modele probabilistyczne
   g. Demo: Entropy/IP

Google

~25%

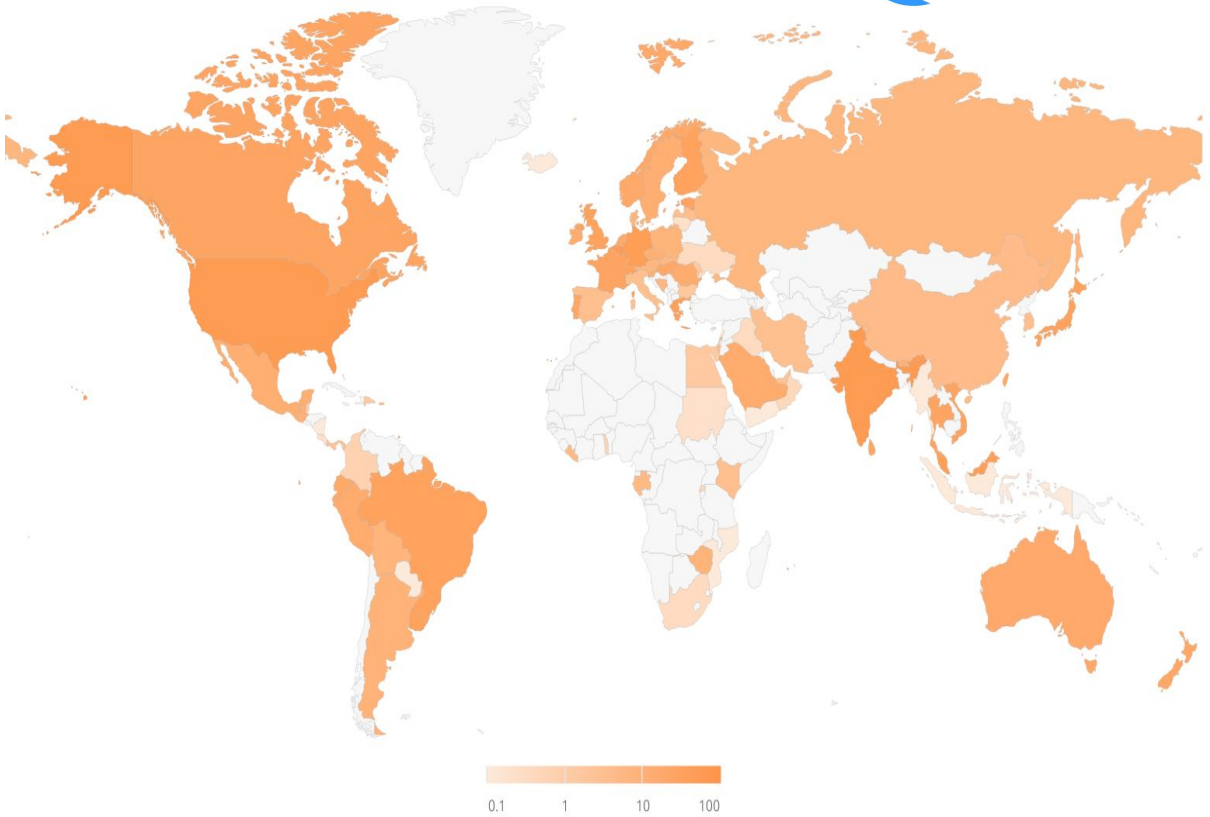
(Świat)



https://www.google.com/intl/en/ipv6/statistics.html

~8%

(Polska)



https://www.akamai.com/us/en/about/our-thinking/state-of-the-internet-report/state-of-the-internet-ipv6-adoption-visualization.jsp

Kilka dat z historii...

- ~1992: kończy się IPv4 :)
- 1993: CIDR w IPv4, NAT
- 1995: RFC1883 - IPv6 proposal
- 1998: RFC2460 - IPv6 draft
- (~20 lat “draftów”)
- 2011: koniec IPv4 w APNIC (Azja)
- 2012: ...RIPE (Europa)
- 2014: ...LACNIC (Ameryka Pd)
- 2015: ...ARIN (Ameryka Pn)
- 2017: RFC8200 - IPv6 standard

- 1996: IPv6 w Linuksie 2.1.8
- 2000: *BSD
- 2001: Cisco
- 2002: Windows
- 2003: MacOS
- 2004: DNS - .jp, .kr
- 2008: DNS - root zone
- 2013: IPv6 to 1% ruchu Google
- 2015: ...5%
- 2018: ...25%

Czemu tak długo? Network Address Translation

- Maskarada na cały LAN:

  192.168.1.0/24 <-> 91.200.172.1

- “Firewall za free”
  - Przekierowanie portów
  - P2P: STUN itp.
- Publiczny IP za $$$

- “Upchać jak najwięcej w /24”

- Na serwerze Virtual Hosting

- Infrastruktura NAT droga
- ...ogranicza rozwój protokołów
- Problemy P2P (gry, voip)
- Nie taki Internet wymyśliliśmy
- Content IPv6 w końcu dostępny
- (Podobno) IoT potrzebuje IPv6

Adresacja: adres 128-bitowy

2001:DB8:70:874::2

2001:0DB8:0070:0874::0002

2001:0DB8:0070:0874:0000:0000:0000:0002

2001:0DB8:0070:0874:0000:0000:0000:0002

Network Identifier (NID) | Interface Identifier (IID)

2001:0DB8:0070:0874:0000:0000:0000:0002

Routing Prefix | Subnet ID

Adresacja: sieci IPv6

2001:DB8:70:874::2/64

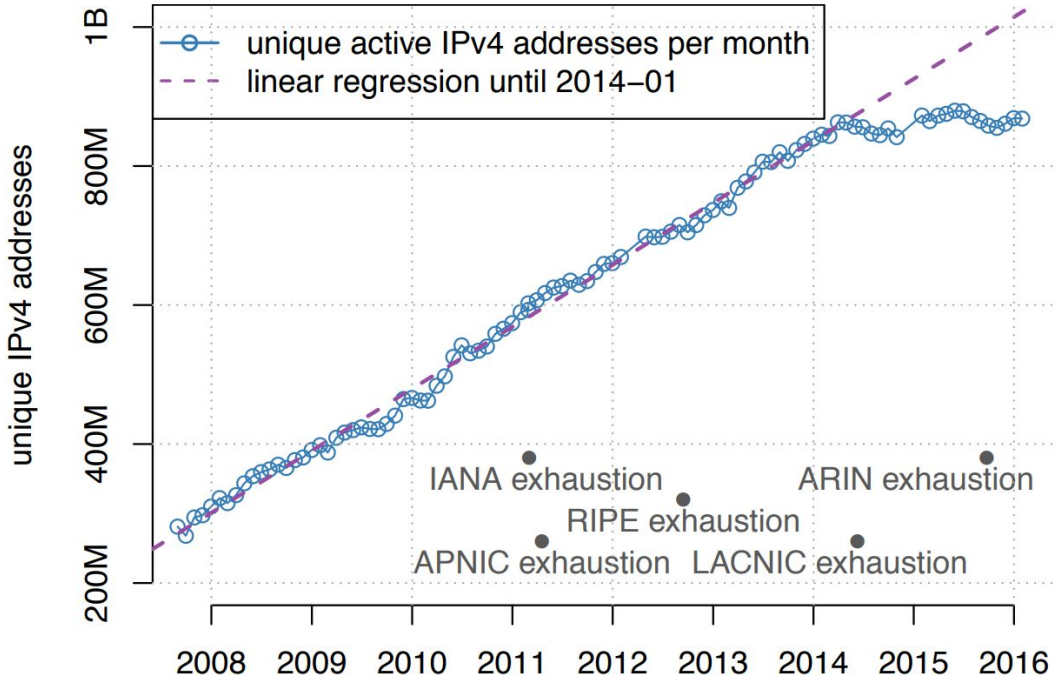
2001:0DB8:0070:0874:0000:0000:0000:0000
2001:0DB8:0070:0874:ffff:ffff:ffff:ffff

2000::/3 -> global unicast
2001:DB8::/32 -> doc prefix
::1/128 -> loopback
fc00::/7 -> LAN
fe80::/10 -> link-local

https://www.iana.org/assignments/ipv6-unicast-address-assignments/ipv6-unicast-address-assignments.xhtml

Adresacja: perspektywa vs. IPv4

- IPv4: 32 bity
- 2^{32} = 4.3 mld adresów
- Spora część “zmarnowana”
- Ile mają użytkownicy?



Beyond Counting: New Perspectives on the Active IPv4 Address Space
https://www.akamai.com/us/en/multimedia/documents/technical-publication/beyond-counting-new-perspectives-on-the-active-ipv4-address-space.pdf

Adresacja: perspektywa vs. IPv4

- IPv4: 32 bity
- 2^{32} = 4.3 mld adresów
- Spora część “zmarnowana”
- Ile mają użytkownicy?

- IPv6: 128 bitów
- 2^{128} = 340 282 366 920 938 463 463 374 607 431 768 211 456 ≈ 3.4×10^{38}
- Całe IPv6 to 2^{96} “internetów”
  - Aktualnie 2^{93}
- 1 komputer to 2^{32} “internetów”
  - IID to /64

Adresacja: metody przydzielania adresu (IID)

2001:DB8:70:874::	od ISP
2001:DB8:70:874::2	na sztywno
2001:DB8:70:874::af:10:2	struktura
2001:DB8:70:874::21e:c2ff:fec0:11db	SLAAC EUI-64
2001:DB8:70:874::3031:f3fd:bbdd:2c2a	SLAAC privacy

...

SLAAC = Stateless Address Autoconfiguration -> RFC docs

IPv6: funkcjonalność

!

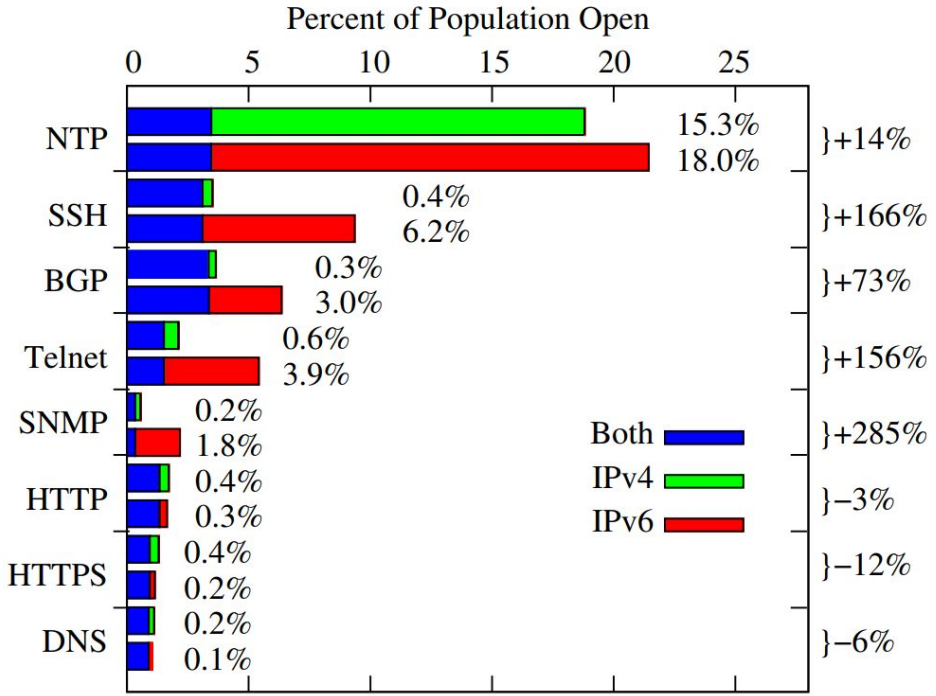
- Większa przestrzeń adresacji
- Łączność end-to-end (P2P)
- Łatwiejsza adresacja
  - autokonfiguracja
  - SLAAC
- Prostszy multicasting
- Szybszy routing
  - fragmentacja end-to-end
  - brak sumy kontrolnej (TTL)

?

- IPsec
- Jumbograms (pakiety 4GB)
- Brak broadcastu
- Mobilność
- Lepszy QoS

IPv6: bezpieczeństwo

- Extension Headers
  - Strzał w stopę :)
  - Omijanie ACL, DoS, itp.
- Skanowanie
  - Czy jest możliwe?
  - Brak NAT - smartfony, IoT, …
  - Wyciek adresów (CDN, P2P, …)
  - Proste schematy adresacji
- “Przecież mam iptables!”
  - *Dual*-stack
  - ip6tables



Don’t Forget to Lock the Back Door! A Characterization of IPv6 Network Security Policy
https://www.caida.org/publications/papers/2016/dont_forget_lock/dont_forget_lock.pdf

HE Tunnel Broker

Jak podłączyć się do IPv6?

tunnelbroker.net

```
modprobe sit

ip tunnel add he0 mode sit \
     remote 216.66.80.162 \
     local 192.168.X.X ttl 255

ip l set dev he0 up

ip addr add 2001:470:7X:XXX::2/64 dev he0

ip route add ::/0 dev he0

ping6 fb.com
```

Cz. 2: Skanowanie IPv6

Igła w stogu siana^{128} - teoria

- IPv4: 4 mld adresów
- 1 milion pakietów ICMP
- Prawdopodobieństwo 1 trafienia?

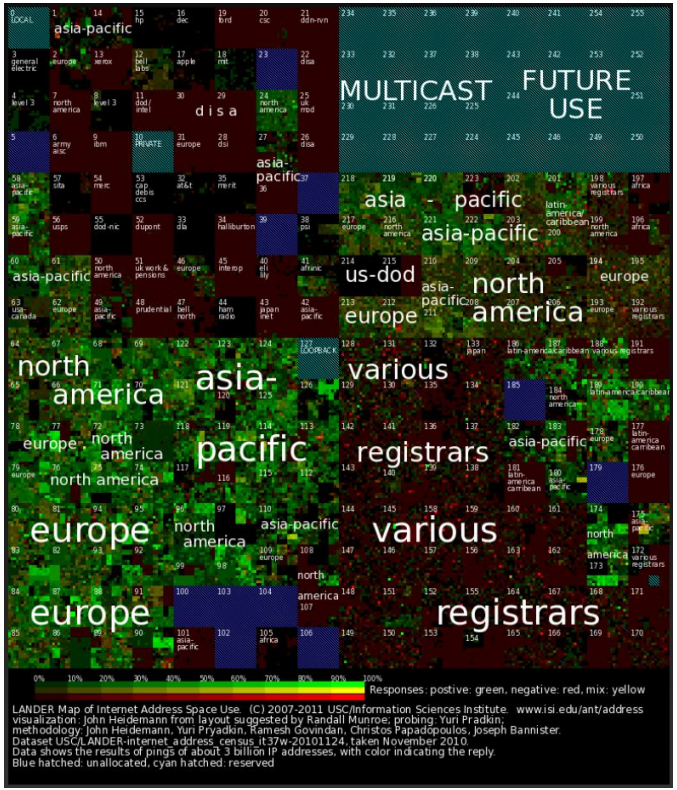
$$\frac{10^6}{2^{32}} \approx 0.023\%$$

- IPv6: 2^{128} adresów
- 1 milion pakietów ICMP
- Prawdopodobieństwo 1 trafienia

$$\frac{10^6}{2^{128}} \approx 2.9 \cdot 10^{-31}\%$$

Igła w stogu siana^{128} - praktyka

- Przeskanować cały Internet?!
- Brute-force powszechny w IPv4
  - ZMap github.com/zmap/zmap
  - 1GbE: <45 min
  - 10GbE: <5 min
- IPv6: nikt normalny nie próbuje
- Wzory 1mln/2^x na pewno OK?
  - Jest więcej niż 1 igła :)
  - IPv4: gęste upakowanie
  - IPv6: bardzo rzadka przestrzeń



https://ant.isi.edu/lander/index.html
https://datatracker.ietf.org/meeting/101/materials/slides-101-maprg-zesplot-an-attempt-to-visualise-ipv6-address-space-00

Igła w stogu siana^{128} - praktyka IPv6

- Skanowanie = informacja + schemat
  - Redukcja przestrzeni poszukiwań
- Proste schematy (RFC7721)
  - scan6 (apt install ipv6toolkit)
  - si6networks.com/tools/ipv6toolkit
  - SLAAC: MAC danego producenta
  - “Low-byte”: licznik na końcu
  - “Port-based”: np. …:53
  - “IPv4”: np. …:192:168:1:1
- Hitlisty
  - DNS (+ rDNS, AXFR, DNSSEC, …)
  - Traceroute
  - P2P (BitTorrent, Bitcoin, …)
  - Crowdsourcing
  - …NTP! (shodan.io)
- Metody probabilistyczne
  - Entropy/IP http://entropy-ip.com/2016-v6structures.pdf
  - 6Gen https://conferences.sigcomm.org/imc/2017/papers/imc17-final245.pdf

Papier #1: Clusters in the Expanse (**CitE**)

Clusters in the Expanse: Understanding and Unbiasing IPv6 Hitlists

Oliver Gasser
Technical University of Munich
gasser@net.in.tum.de

Quirin Scheitle
Technical University of Munich
scheitle@net.in.tum.de

Pawel Foremski
IITiS PAN
pjf@iitis.pl

Qasim Lone
Grenoble Alps University
qasim.lone@univ-grenoble-alpes.fr

Maciej Korczyński
Grenoble Alps University
maciej.korczynski@univ-grenoble-alpes.fr

Stephen D. Strowes
RIPE NCC
sdstrowes@gmail.com

Luuk Hendriks
University of Twente
luuk.hendriks@utwente.nl

Georg Carle
Technical University of Munich
carle@net.in.tum.de

ACM Internet Measurement Conference 2018
Boston, MA, USA

https://ipv6hitlist.github.io/

CitE: źródła danych

Table 2: Overview of hitlist sources, as of May 11, 2018.

Name	Public	Nature	IPs	new IPs	#ASes	#PFXes	Top AS1	Top AS2	Top AS3
DL: Domain Lists[1]	Yes	Servers	9.8 M	9.8 M	6.1 k	10.3 k	89.7 % ★	2.0 % ●	1.5 % ■
FDNS: Rapid7 FDNS	Yes	Servers	3.3 M	2.5 M	7.7 k	13.6 k	16.7 % ★	8.9 % ▲	6.7 % ✣
CT: Domains from CT logs[2]	Yes	Servers	18.5 M	16.2 M	5.3 k	8.7 k	92.3 % ★	1.6 % ✣	0.8 % ★
AXFR: AXFR&TLDR	Yes	Mixed	0.7 M	0.5 M	3.2 k	4.7 k	57.0 % ★	14.0 % ●	8.3 % ■
BIT: Bitnodes	Yes	Mixed	31 k	27 k	695	1.4 k	8.0 % ★	6.0 % ■	6.0 % ▲
RA: RIPE Atlas[3]	Yes	Routers	0.2 M	0.2 M	8.4 k	19.1 k	6.6 % ✣	3.5 % ★	3.1 % ✣
Scamper	–	Routers	26.0 M	25.9 M	6.3 k	9.8 k	38.9 % ★	23.8 % ●	12.0 % ■
Total			58.5 M	55.1 M	10.9 k	25.5 k	45.4 % ★	18.4 % ★	11.5 % ●

1: Zone Files, Toplists, Blacklists (partially with NDA); 2: Excluding DNS names already included in Domain Lists; 3: Traceroute and ipmap data

★Amazon, ●Host Europe, ■Cloudflare, ▲Linode, ✣DTAG, ★ProXad, ●Hetzner, ■Comcast, ▲Swisscom, ✣Google, ★Antel, ●Versatel, ■BIHNET

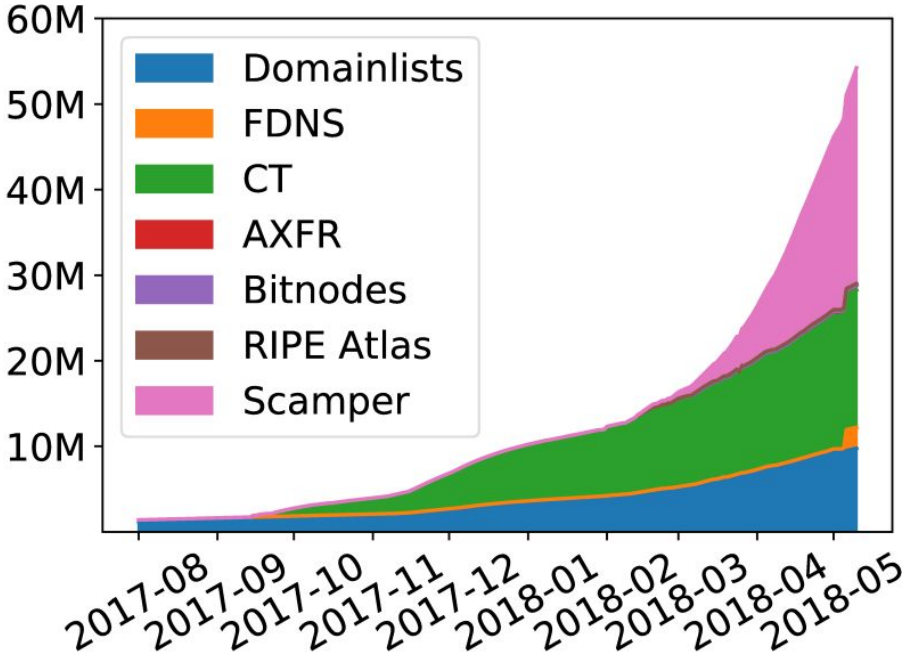
CitE: źródła danych (DNS + massdns)

- Strefy DNS, np.
  - https://zonedata.iis.se/
  - https://czds.icann.org/en
- Transfery strefy DNS przez AXFR, np.
  - https://github.com/tldr-pages/tldr
- Rapid7 Forward DNS
  - https://opendata.rapid7.com/sonar.fdns_v2/
- Google Certificate Transparency
  - https://www.certificate-transparency.org/
  - https://github.com/google/certificate-transparency-go/tree/master/client/ctclient
  - https://crt.sh/

CitE: źródła danych (nie-DNS)

- Bitcoin, np.
  - https://bitnodes.earn.com/nodes/?q=ipv6
- RIPE Atlas
  - https://atlas.ripe.net/
  - https://ftp.ripe.net/ripe/ipmap/
- Scamper
  - https://www.caida.org/tools/measurement/scamper/

CitE: źródła - akumulacja

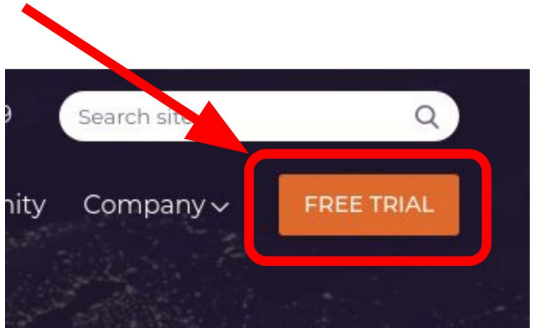


(a) Cumulative runup of IPv6 addresses.

DNSDB

"All your DNS
are belong to us"

1. www.farsightsecurity.com



2. Chrome -> DNSDB Scout
chrome.google.com/webstore/detail/farsight-dnsdb-scout/pkkfiklolnimjhgfokgicminkbfnfclb

3. GitHub -> dnsdbq
github.com/dnsdb/dnsdbq

CitE: skanowanie hitlist

- Zebrano: 55.1M IPv6
- Duplikaty (aliasy): 25.7M (47%)
  - Te same maszyny, inne adresy
- Zostaje:
  - 29.4M adresów
  - 10,900 ASów (sieci)
  - 24,600 prefiksów BGP
- Skanowanie (ZMapv6)
  - ICMP (ping)
  - TCP/80, TCP/443, UDP/443
  - UDP/53
- Powtarzane codziennie :)

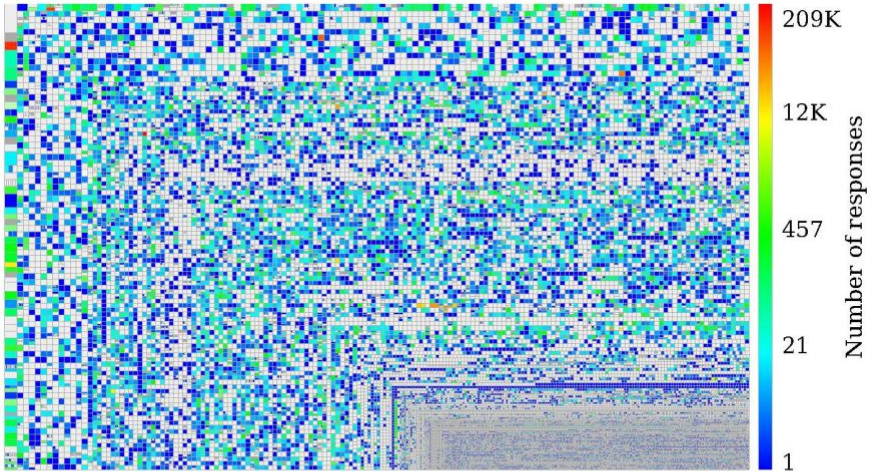


Figure 6: All 56 k BGP prefixes, colored based on the number of responses to ICMP Echo requests on May 11, 2018.

- 1.9M adresów (6.46%)
- 9970 ASów (91.5%)
- 21600 prefiksów BGP (87.8%)

Entropia ujawnia słabe sieci

```
2001:0db8:0010:0013:0000:0000:0000:07fe
2001:0db8:0010:0000:0000:0000:0000:0ed3
2001:0db8:0010:0003:0000:0000:0000:0fb5
2001:0db8:0020:d05f:882f:6082:f768:710d
2001:0db8:0010:0004:0000:0000:0000:04dc
2001:0db8:0010:0003:0000:0000:0000:03ce
2001:0db8:0010:0008:0000:0000:0000:0794
2001:0db8:0010:000a:0000:0000:0000:0923
2001:0db8:0010:0006:0000:0000:0000:003c
2001:0db8:0022:1014:aef6:60af:d029:63cd
2001:0db8:0010:0012:0000:0000:0000:0c7b
2001:0db8:0022:10c0:5100:ac7d:96f5:5851
(...)
```

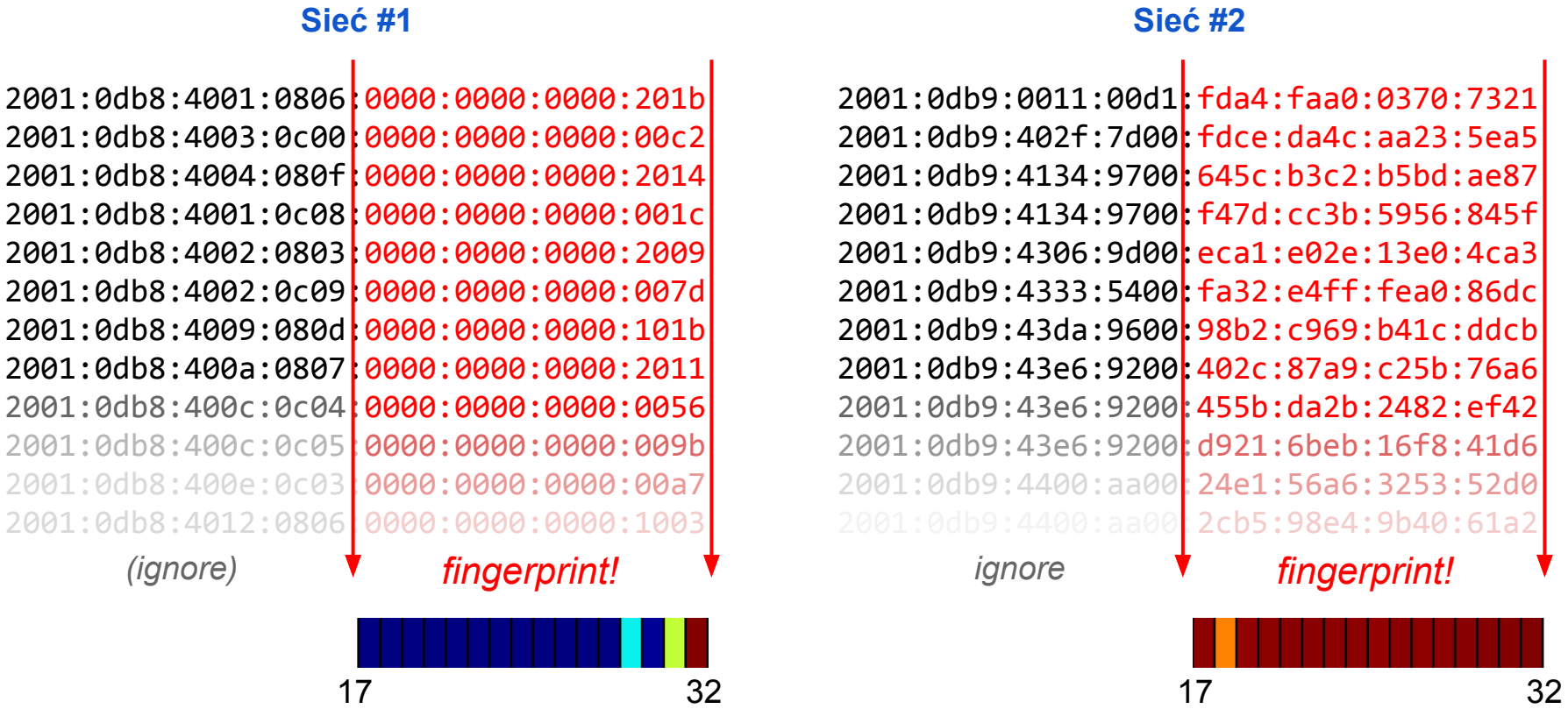
$$H(X) = -\sum_{i=1}^{k} P(x_i) \log P(x_i)$$

$H(X_{16}) = 3.8/4$

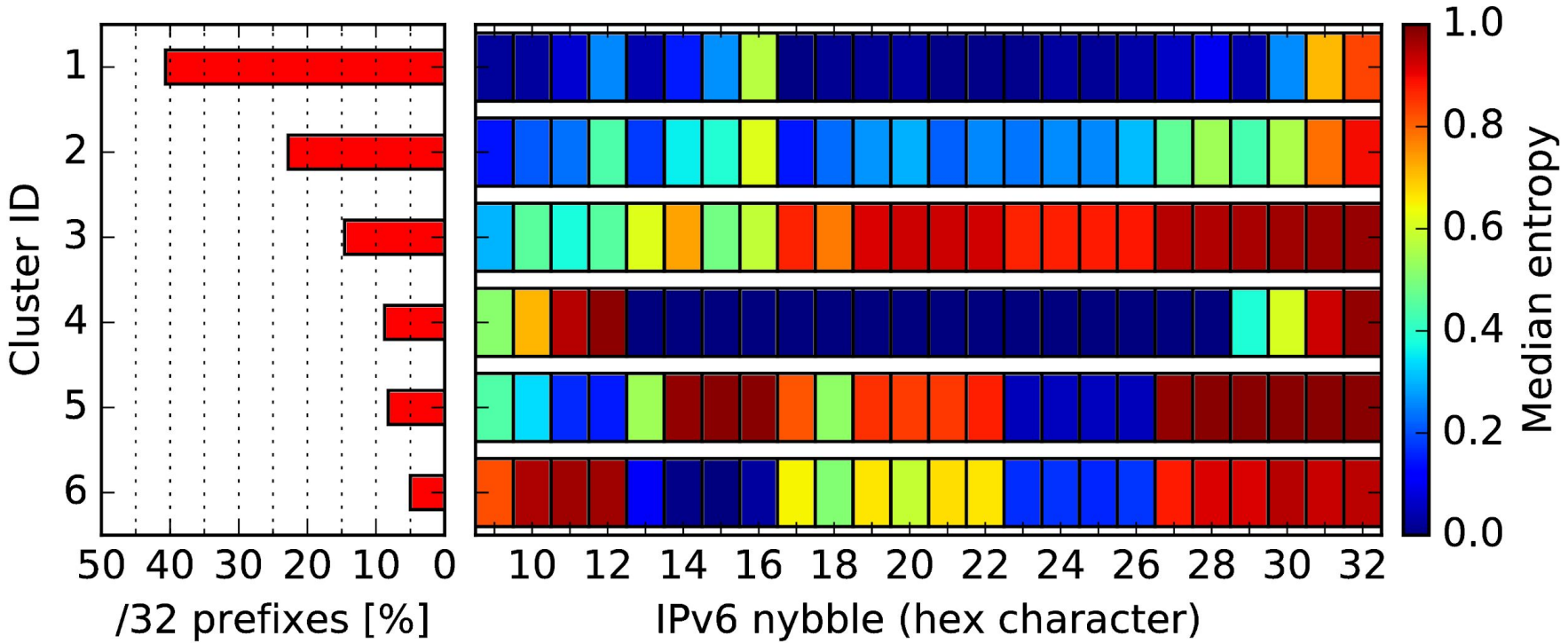
$H(X_{18}) = 2.2/4$

http://entropy-ip.com/

CitE: entropy clustering

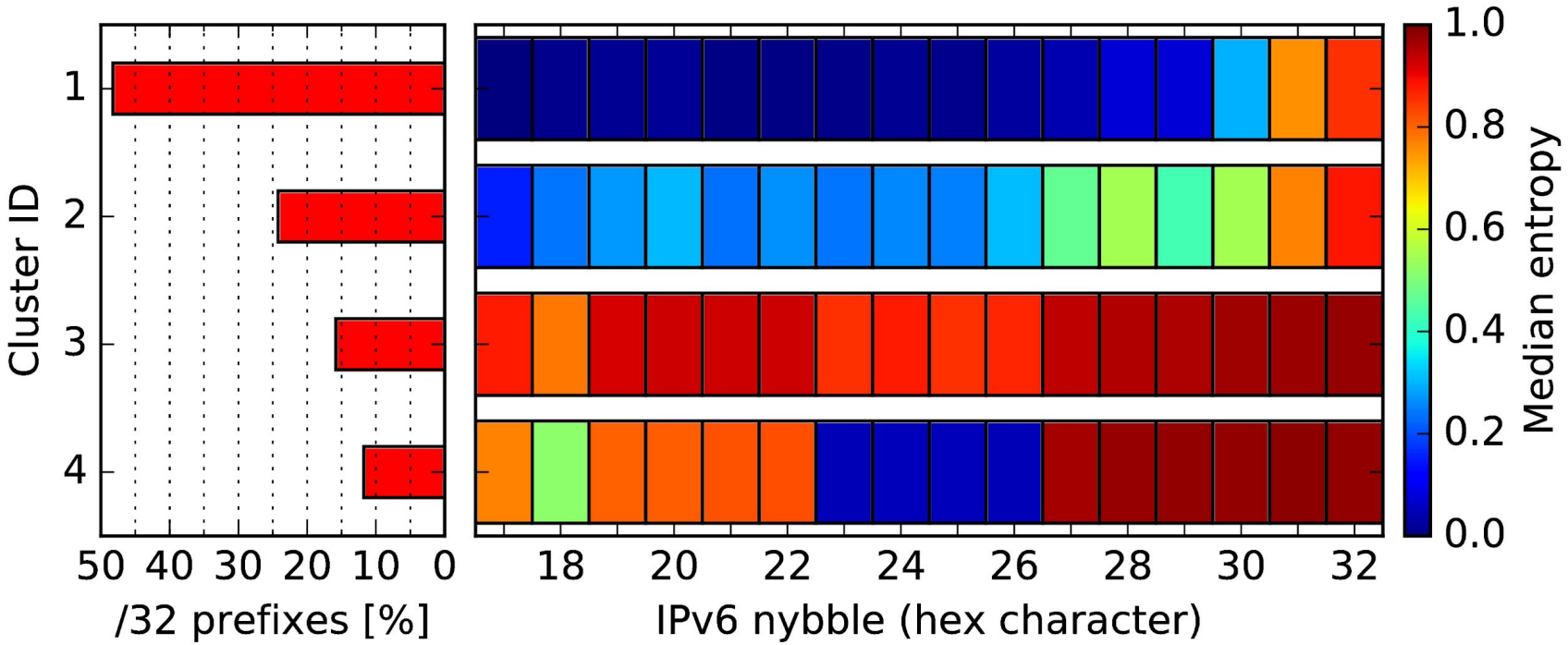


CitE: entropy clustering (prefiksy /32)



https://github.com/pforemski/entropy-clustering

CitE: entropy clustering (IID)



https://github.com/pforemski/entropy-clustering

Modele probabilistyczne danej sieci: Entropy/IP (**eIP**)

Entropy/IP: Uncovering Structure in IPv6 Addresses

Paweł Foremski
Akamai Technologies
IITiS PAN
pjf@iitis.pl

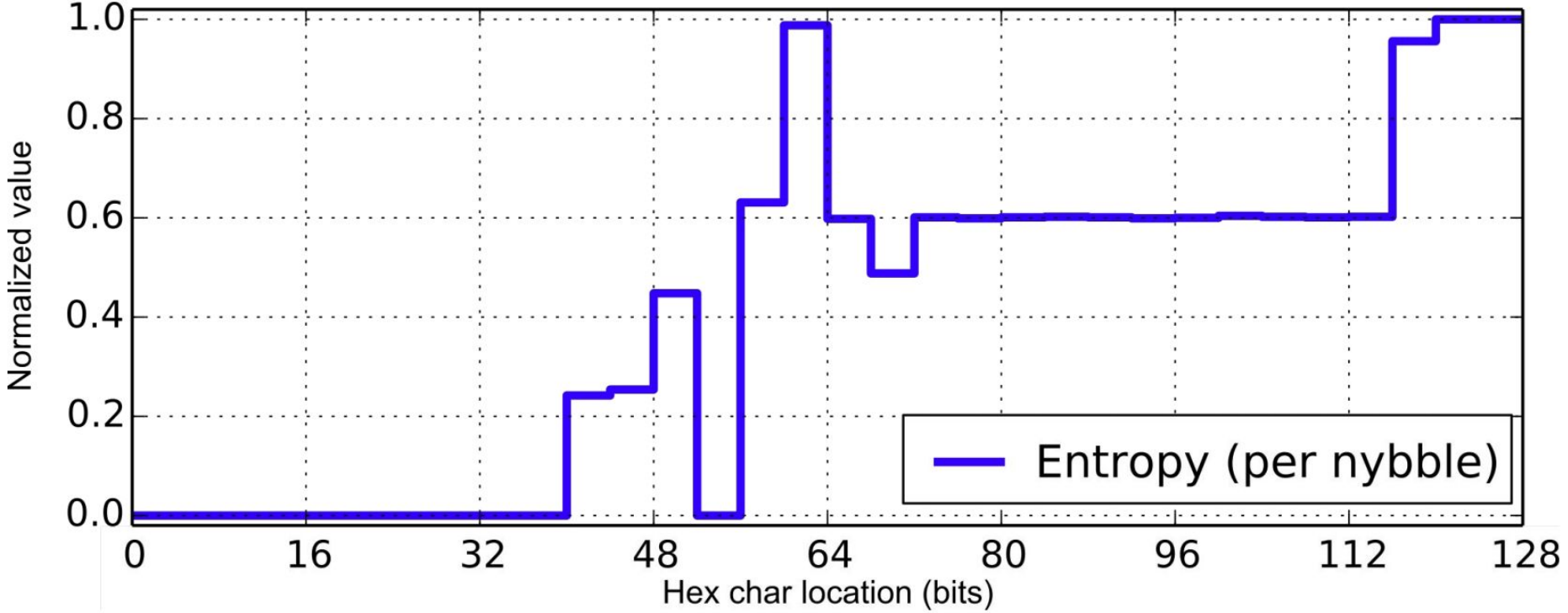
David Plonka
Akamai Technologies
plonka@akamai.com

Arthur Berger
Akamai Technologies
MIT CSAIL
arthur@akamai.com

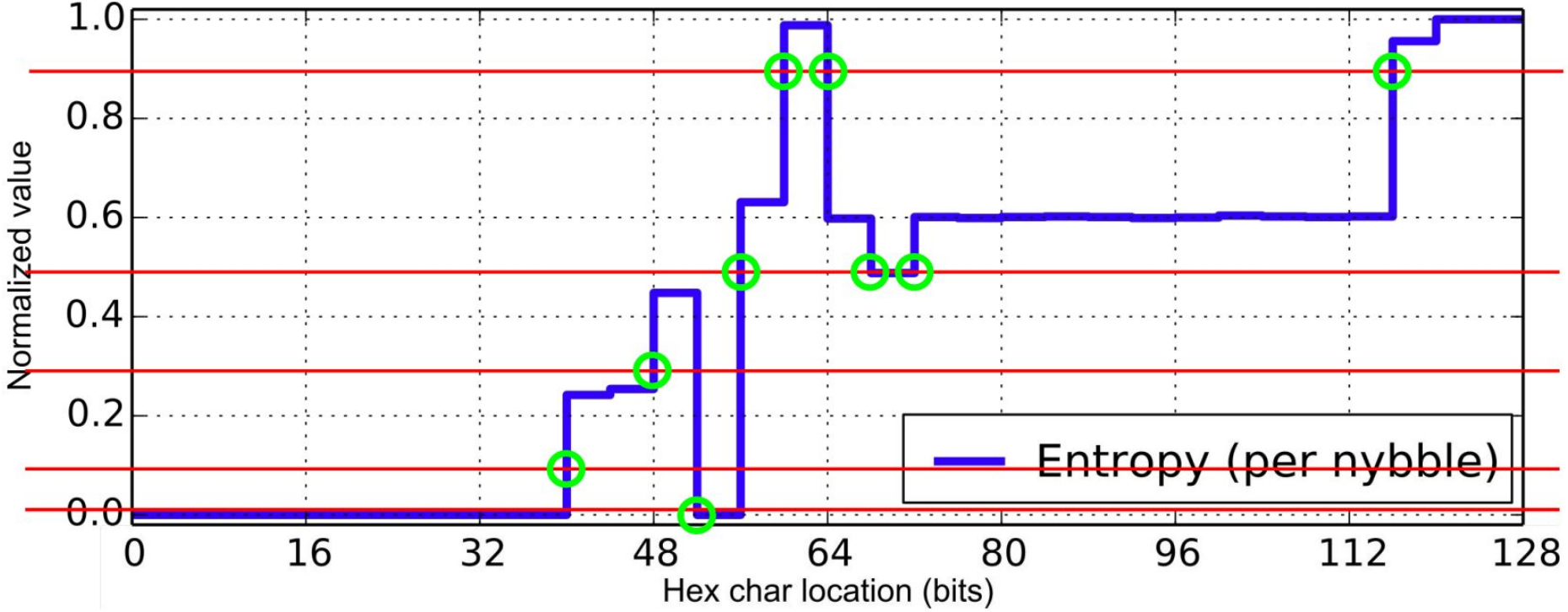
ACM Internet Measurement Conference 2016
Santa Monica, CA, USA

http://www.entropy-ip.com/

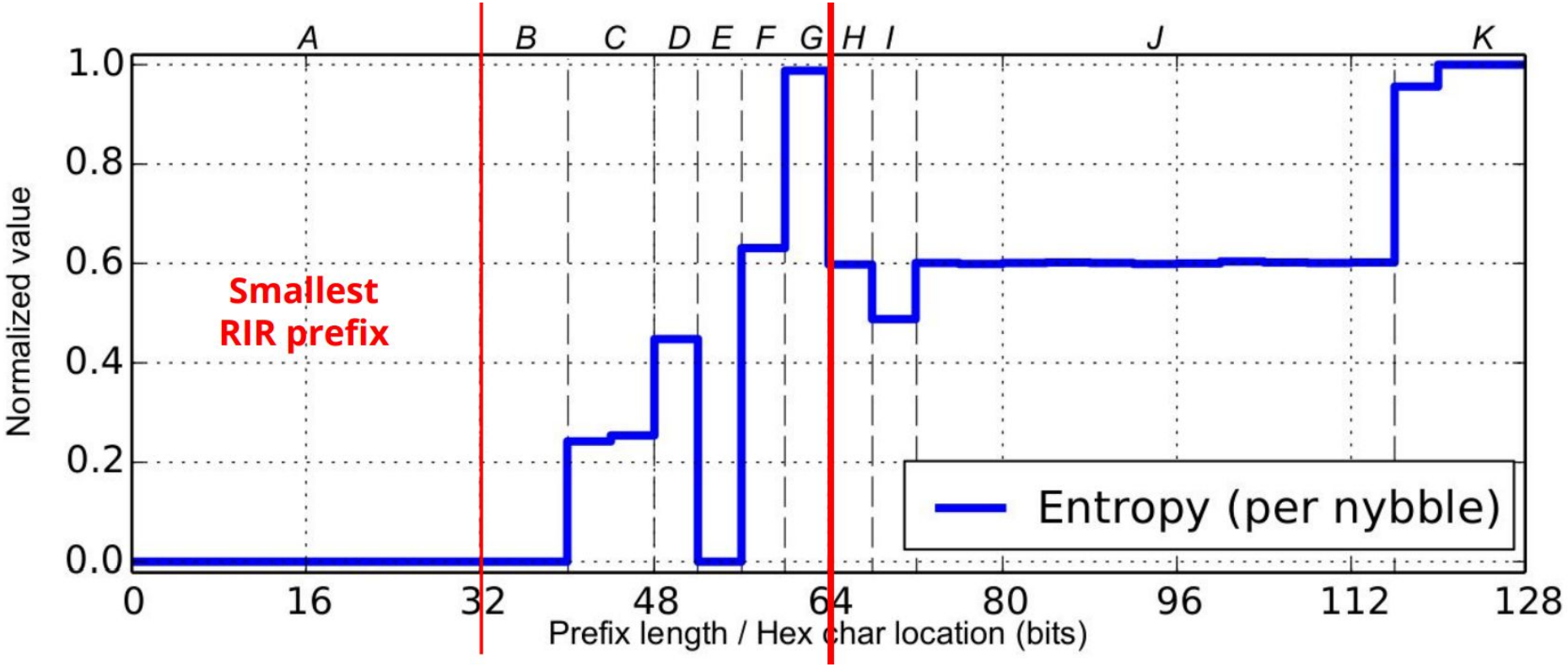
eIP: krok #1 - entropia



eIP: krok #2 - segmenty



eIP: krok #2 - segmenty



eIP: krok #3 - data mining w każdym segmencie

	Code	Value	Frequency
C (40-48)	C1	00	67.02%
	C2	01	11.13%
	C3	c2	0.67%
	C4	fe	0.41%
	C5	ff	0.41%
	C6	02-5b	11.94%
	C7	5c-fd	8.42%

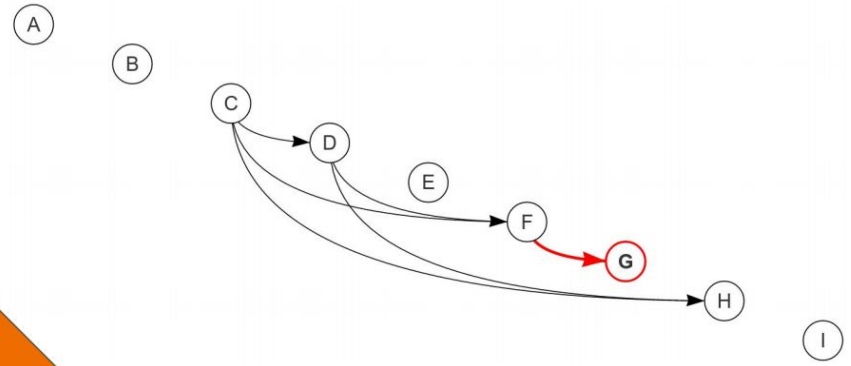
2001:0db8:0841:2500:0000:d9a0:5345:0012

2001:0db8:0841:2500:0000:d9a0:5345:0012

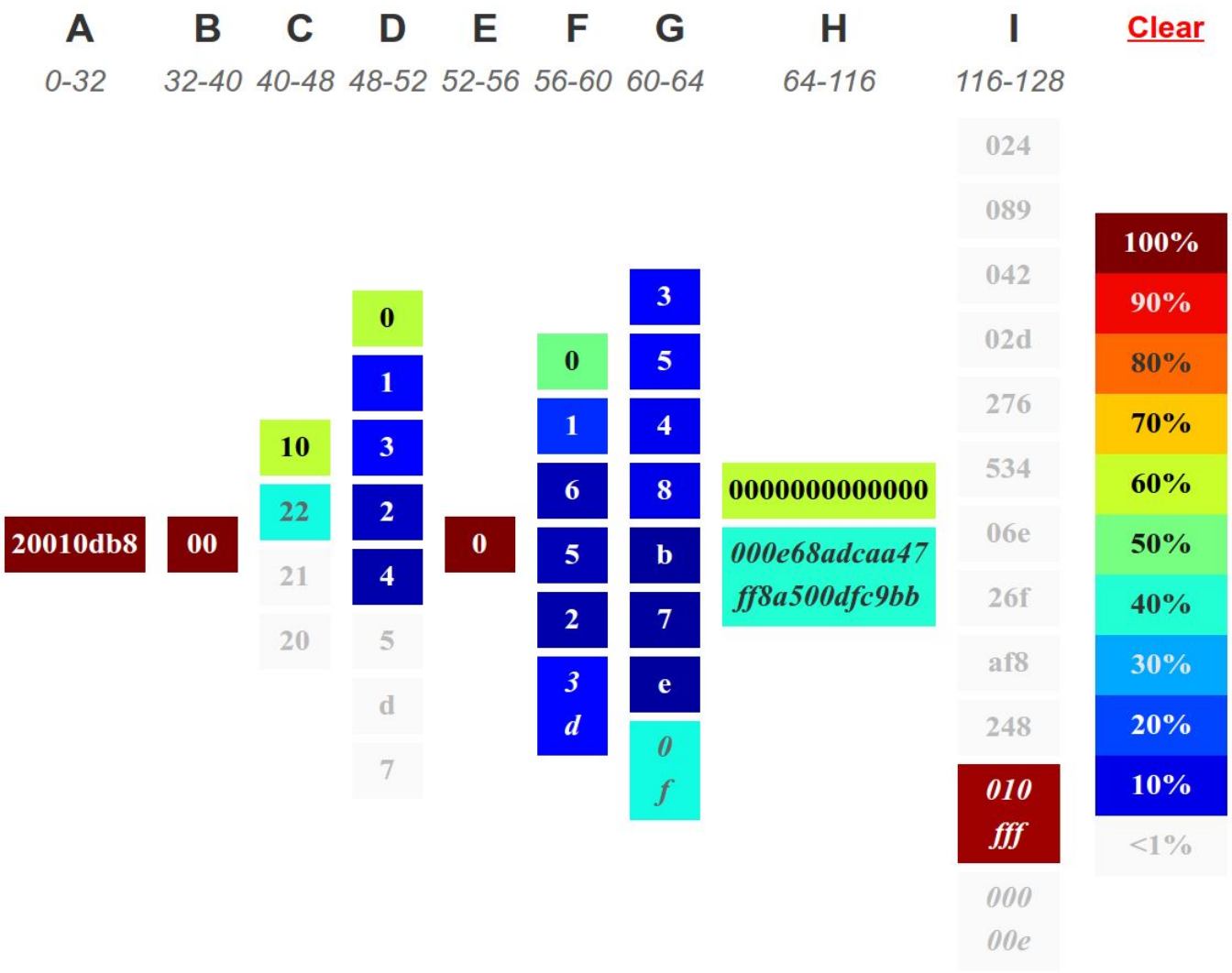
(A1, B2, C6, D4, E5, F1, G12, H1, I2, J3)

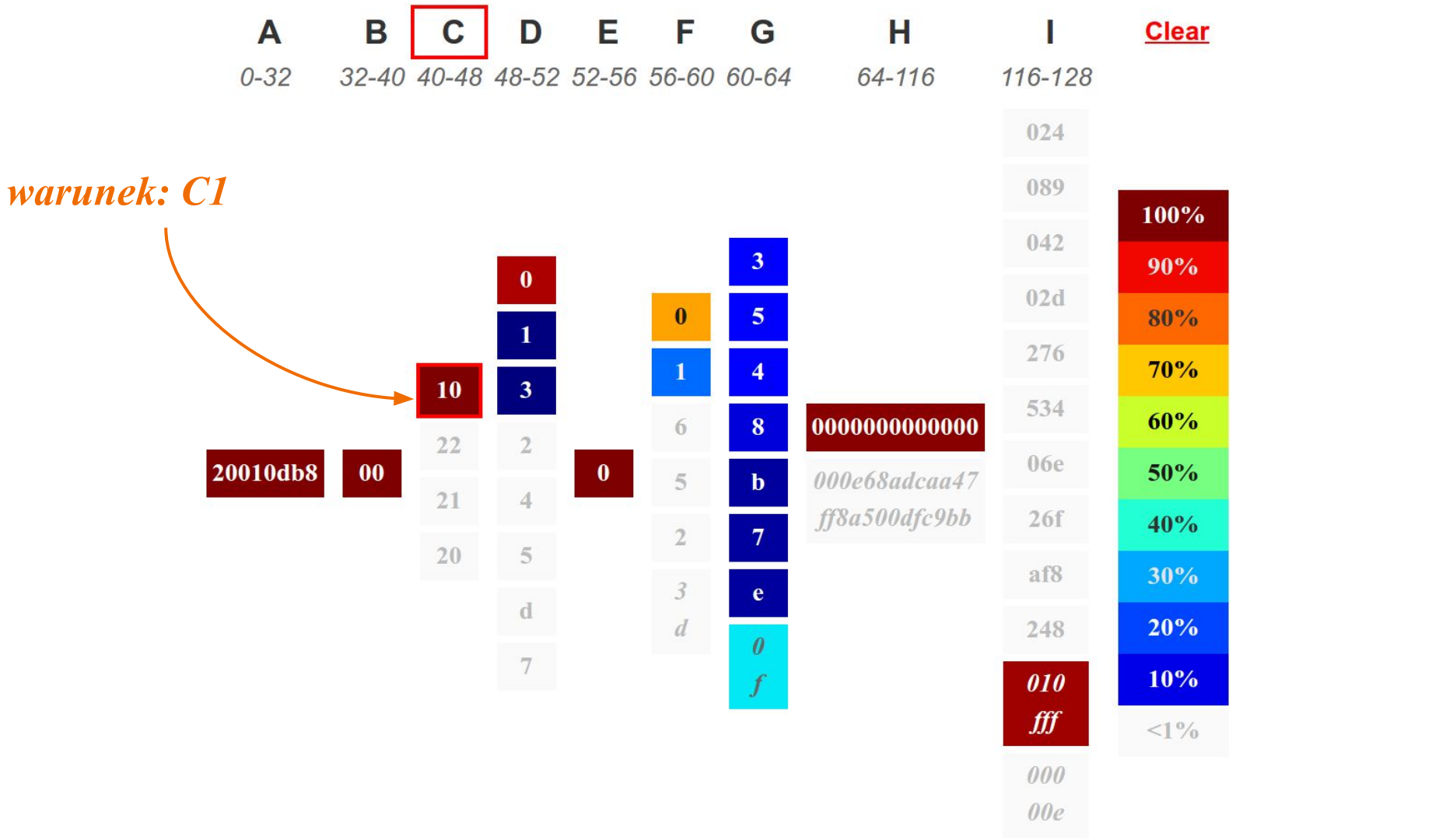
eIP: krok #4 - sieci Bayesa

(A1, B1, C1, D1, E1, F1, G3, H1, I11)
(A1, B1, C1, D1, E1, F1, G1, H1, I11)
(A1, B1, C2, D2, E1, F5, G4, H2, I11)
(A1, B1, C2, D3, E1, F3, G3, H2, I11)
(A1, B1, C1, D1, E1, F2, G3, H1, I11)
(A1, B1, C1, D1, E1, F2, G3, H1, I11)
(A1, B1, C1, D1, E1, F1, G1, H1, I11)
(A1, B1, C3, D1, E1, F4, G8, H2, I11)
(A1, B1, C1, D1, E1, F1, G1, H1, I11)
(A1, B1, C1, D1, E1, F1, G8, H1, I11)
(A1, B1, C1, D1, E1, F2, G1, H1, I11)
(A1, B1, C2, D4, E1, F6, G3, H2, I11)
(A1, B1, C3, D1, E1, F2, G3, H2, I11)
(A1, B1, C1, D1, E1, F1, G8, H1, I11)



	G:		
F:	G1	G2	G3
F1	13%	10%	10%
F2	18%	20%	20%
F3	13%	7%	9%
F4	16%	9%	10%


A
0-32
20010db8
B
32-40
00
C
40-48
10
22
21
20
D
48-52
0
1
3
2
4
5
d
7
E
52-56
0
F
56-60
0
1
6
5
2
3
d
G
60-64
3
5
4
8
b
7
e
0
f
H
64-116
0000000000000
000e68adcaa47
ff8a500dfc9bb
I
116-128
024
089
042
02d
276
534
06e
26f
af8
248
010
fff
000
00e
Clear
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
<1%


A
0-32
B
32-40
C
40-48
D
48-52
E
52-56
F
56-60
G
60-64
H
64-116
I
116-128
Clear
warunek: C1
20010db8
00
10
22
21
20
0
1
3
2
4
5
d
7
0
0
1
6
5
2
3
d
3
5
4
8
b
7
e
0
f
0000000000000
000e68adcaa47
ff8a500dfc9bb
024
089
042
02d
276
534
06e
26f
af8
248
010
fff
000
00e
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
<1%

Demo: Entropy/IP

Jak przeskanować prefix /32?

github.com/akamai/entropy-ip

```
git clone git@github.com:akamai/entropy-ip.git

cd entropy-ip/

# (instalacja + fping)

./ALL.sh ../dnsdb/isp.hexip ./out/

./c1-gen.sh -n 1000 ./out/cpd \
     | ./c2-decode.py --colons \
          /dev/stdin ./out/analysis \
     > targets.txt

cat targets.txt | fping6 -r 0 -a -s -e

# (ew. powtórz uczenie)
```

Podsumowanie

- IPv6 naprawdę ma znaczenie i rośnie w siłę
- IPv6 to nie “IPv4 ale więcej adresów”
- Stos IPv6 ma osobny firewall
  - Uważaj na wildcard bind :)
- Na rynku brakuje specjalistów od bezpieczeństwa IPv6
- IPv6 da się skanować

Skanowanie:

- Brute-force niemożliwy
- Proste schematy adresacji (scan6)
- Wszystko co publikujesz w DNS zostanie zeskanowane
- Jeśli łączysz się po IPv6 do P2P (gry, bitcoin, bittorrent) - j.w.
- Używaj pseudo-losowych IID

IPv6: przestrzeń do ogarnięcia

ipv6hitlist.github.io

entropy-ip.com

Dziękuję!

Paweł Foremski
pjf@iitis.pl, @pforemski

Net::IP Meetup #10 @ OVH
Wrocław, 25.10.2018

IITiS