BENOCS

Stop guessing. Start knowing

Raw Network Analyser (RNA)

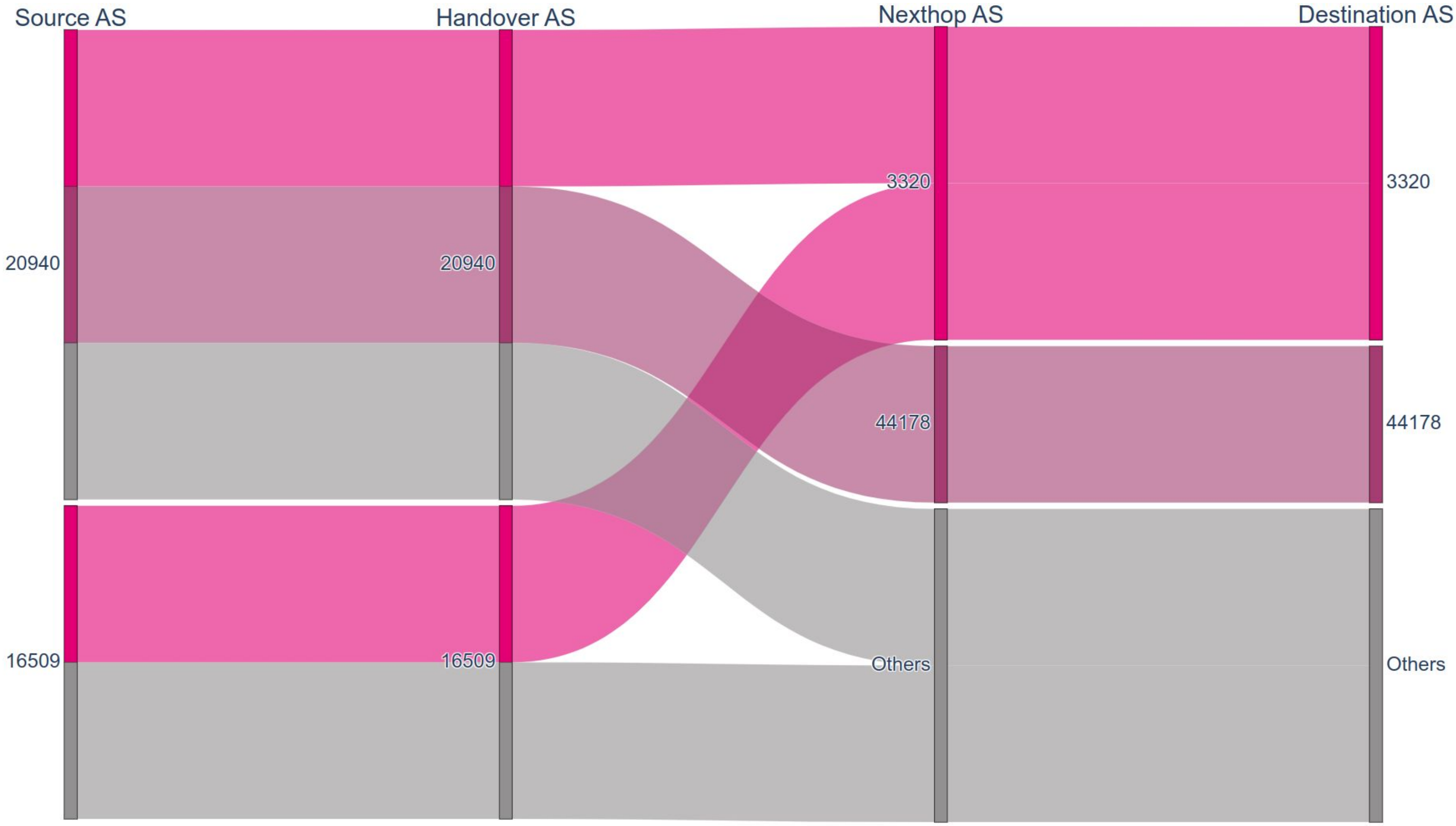
A DNS & Netflow Correlation System

Feb, 2023

Motivation

Correlations are widely used

- Examples:
  - Netflow:
    - IP addresses
    - Number of bytes/packets
  - BGP:
    - Autonomous systems (ASes)
- Multi-dimensional traffic information



https://www.benocs.com/analytics/

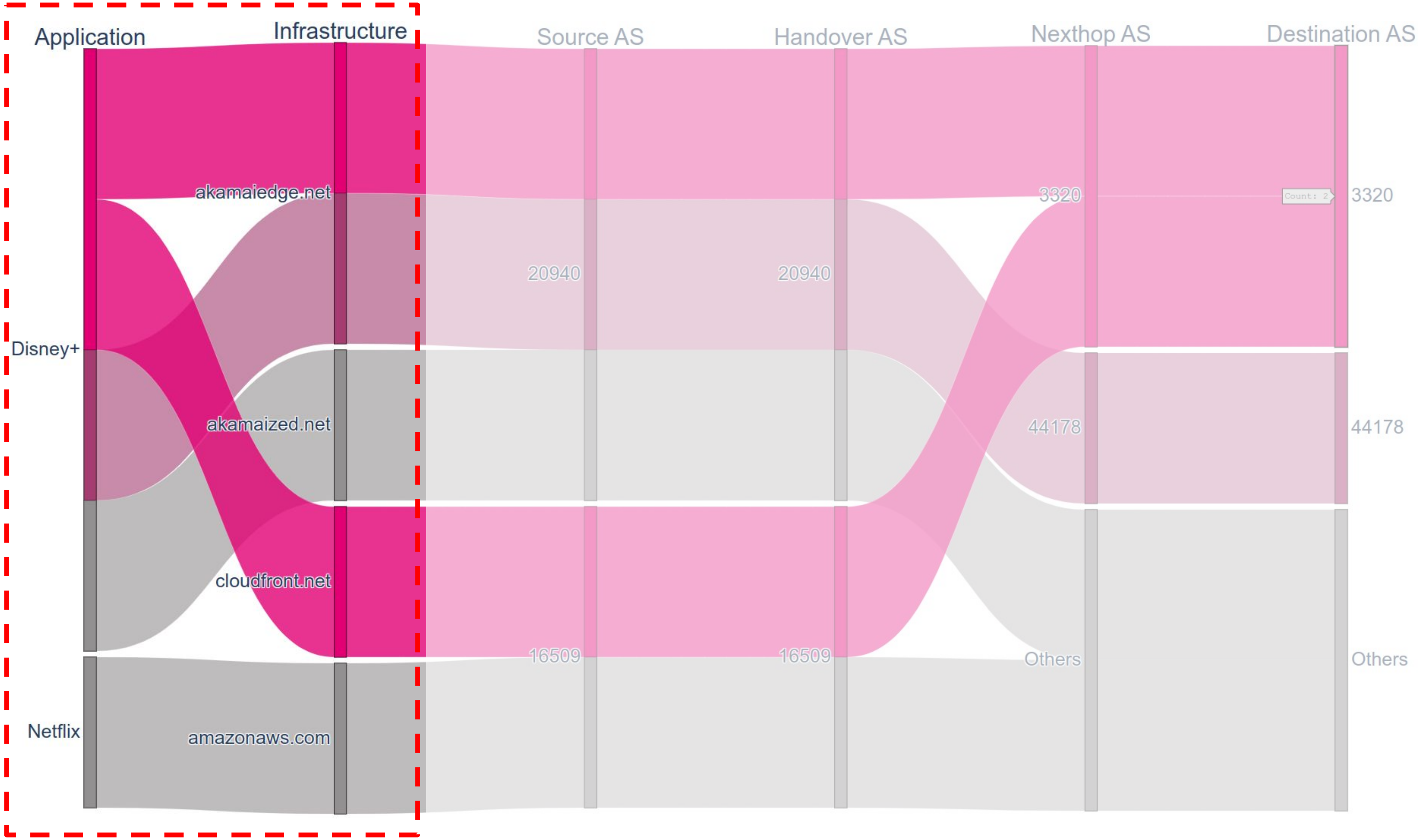
but …

- There is no information about infrastructures and applications behind those ASes or IP addresses
  - Disney+, DAZN, … (applications) use CDNs (infrastructures) for content distribution
- There is no answer for two questions:
  - What are those applications and how much traffic is coming from them?
  - What are those infrastructures and how much traffic is coming from them?

DNS correlation

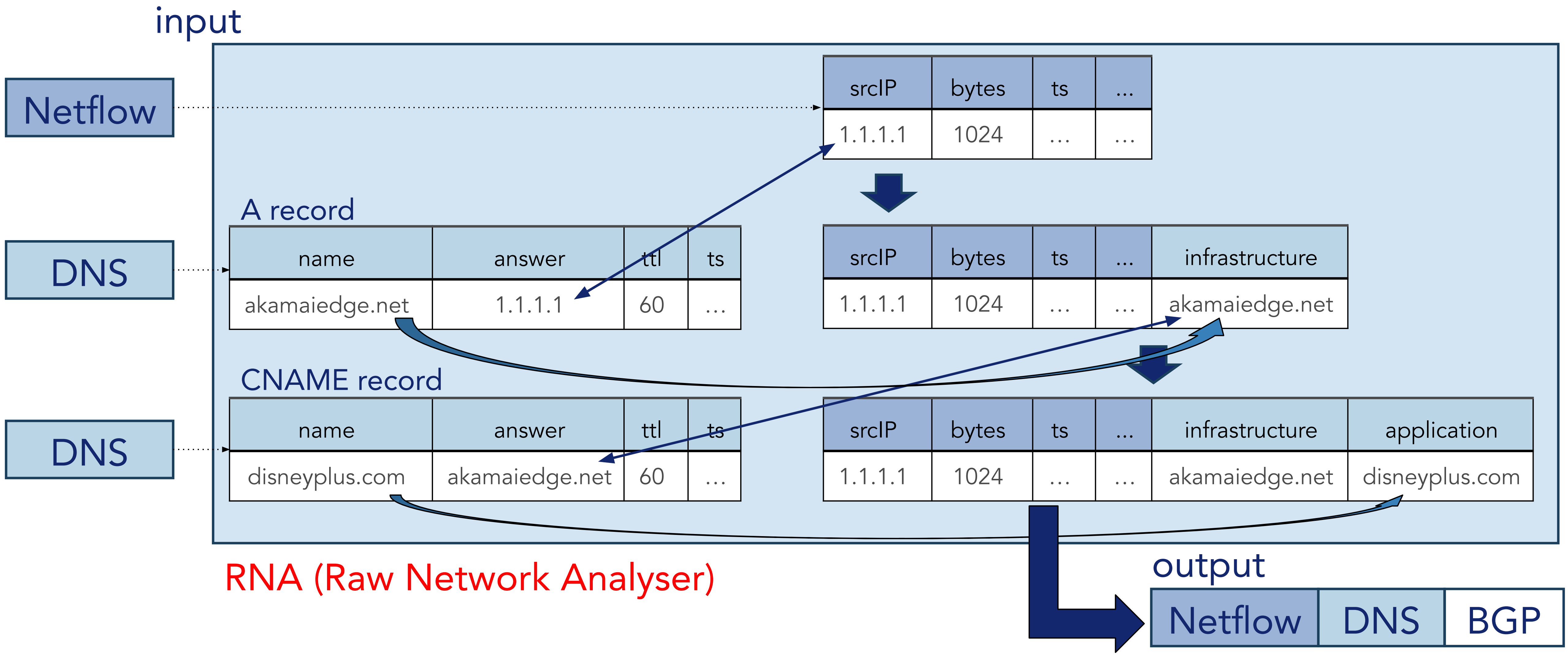
Annotate the traffic with the domain name(s) they came from so that TWO new dimensions are obtained:

- Infrastructure
- Application



HOW?

Netflow & DNS Correlation



input

Netflow

DNS

DNS

srcIP	bytes	ts	...
1.1.1.1	1024	…	…

A record

name	answer	ttl	ts
akamaiedge.net	1.1.1.1	60	…

srcIP	bytes	ts	...	infrastructure
1.1.1.1	1024	…	…	akamaiedge.net

CNAME record

name	answer	ttl	ts
disneyplus.com	akamaiedge.net	60	…

srcIP	bytes	ts	...	infrastructure	application
1.1.1.1	1024	…	…	akamaiedge.net	disneyplus.com

RNA (Raw Network Analyser)

output

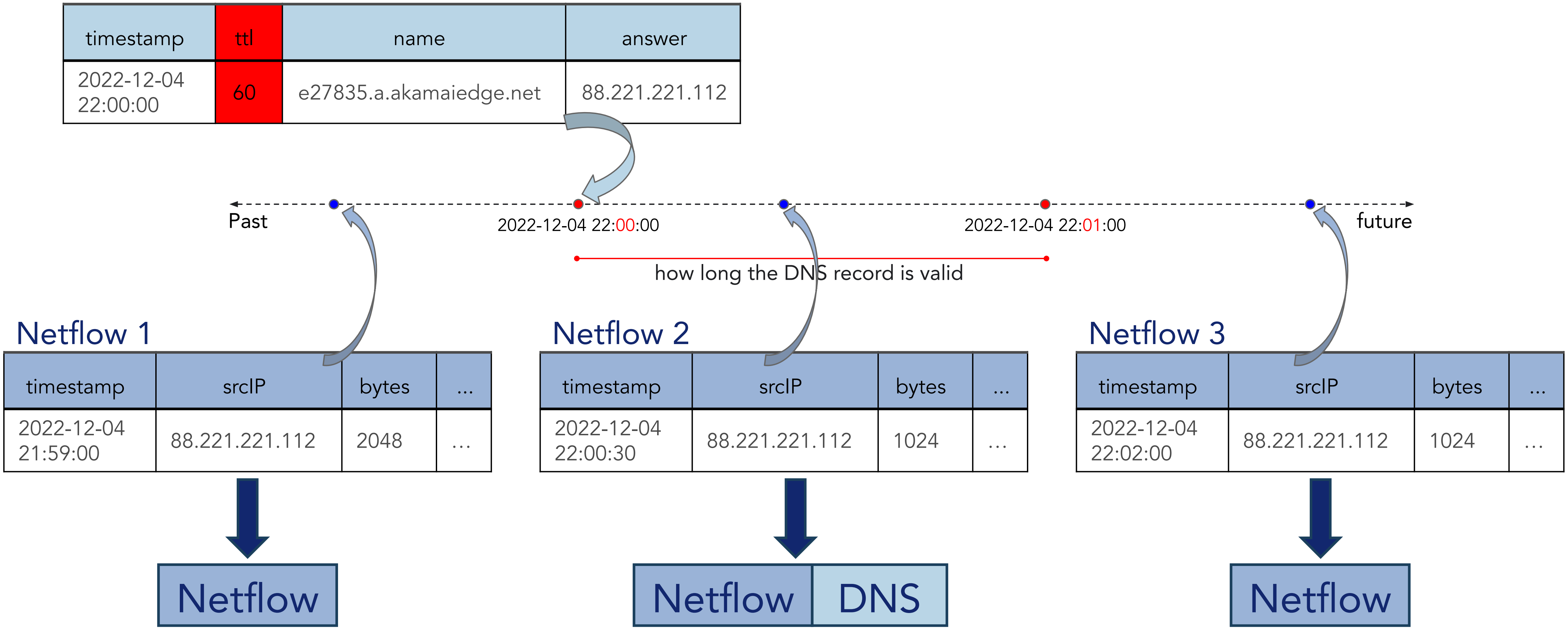
Netflow	DNS	BGP

Challenges

1. Expired DNS records

track to know if they are still valid or not

1. Expired DNS records



DNS record

timestamp	ttl	name	answer
2022-12-04 22:00:00	60	e27835.a.akamaiedge.net	88.221.221.112

Netflow 1

timestamp	srcIP	bytes	…
2022-12-04 21:59:00	88.221.221.112	2048	…

Netflow 2

timestamp	srcIP	bytes	…
2022-12-04 22:00:30	88.221.221.112	1024	…

Netflow 3

timestamp	srcIP	bytes	…
2022-12-04 22:02:00	88.221.221.112	1024	…

2. Domain name resolution

recursive lookup into DNS records

2. Domain name resolution

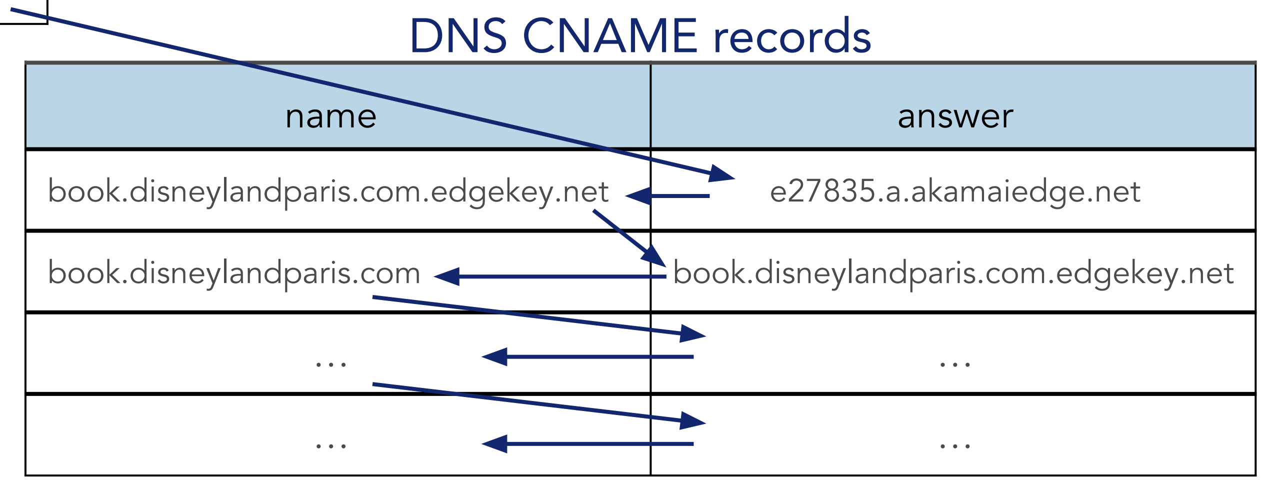
Netflow

srcIP	bytes	…
88.221.221.112	1024	…

DNS A/4A record

name	answer	…
e27835.a.akamaiedge.net	88.221.221.112	…

srcIP	bytes	…	Infrastructure
88.221.221.112	1024	…	e27835.a.akamaiedge.net



DNS CNAME records

name	answer
book.disneylandparis.com.edgekey.net	e27835.a.akamaiedge.net
book.disneylandparis.com	book.disneylandparis.com.edgekey.net
…	…
…	…

Recursive DNS Correlation

(sometimes) INFINITE LOOP

3. Infrastructure/Application classification

3. Infrastructure/Application classification

srcIP	bytes	ts	…	Infrastructure	Application
1.1.1.1	1024	…	…	e27835.a.akamaiedge.net	book.disneylandparis.com

Fully Qualified Domain Name (FQDNs)

Infrastructure
e27835.a.akamaiedge.net

- 2LD+TLD
- Top N infrastructures

Infrastructure
akamaiedge.net

Application
book.disneylandparis.com

- Top N applications

Application
disneylandparis.com

- Patter map

Application
Disney

dssott.com	Disney+
disney-plus.net	Disney+
pv-cdn.net	AmazonPrime
livedazn.secure.footprint.net	DAZN

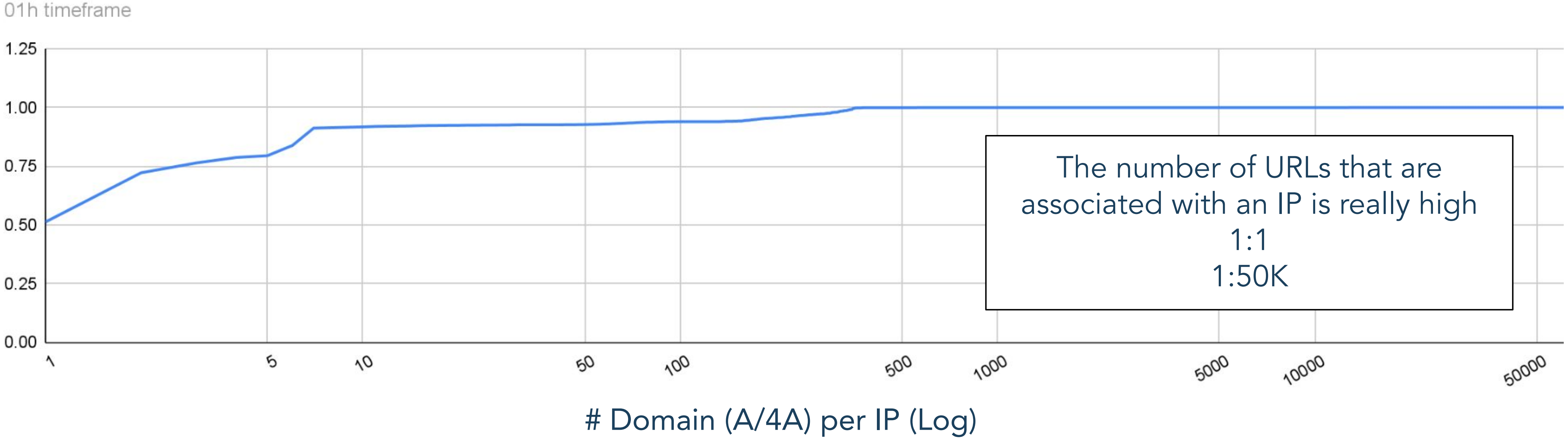
4. Ambiguity

4. Ambiguity

- Over the time the same IP address could be associated to different domain names

name	answer
cloudfront.net	1.1.1.1
amazonws.net	1.1.1.1
…	1.1.1.1

name	answer
akamaiedge.net	2.2.2.2
akamaized.net	2.2.2.2
…	2.2.2.2

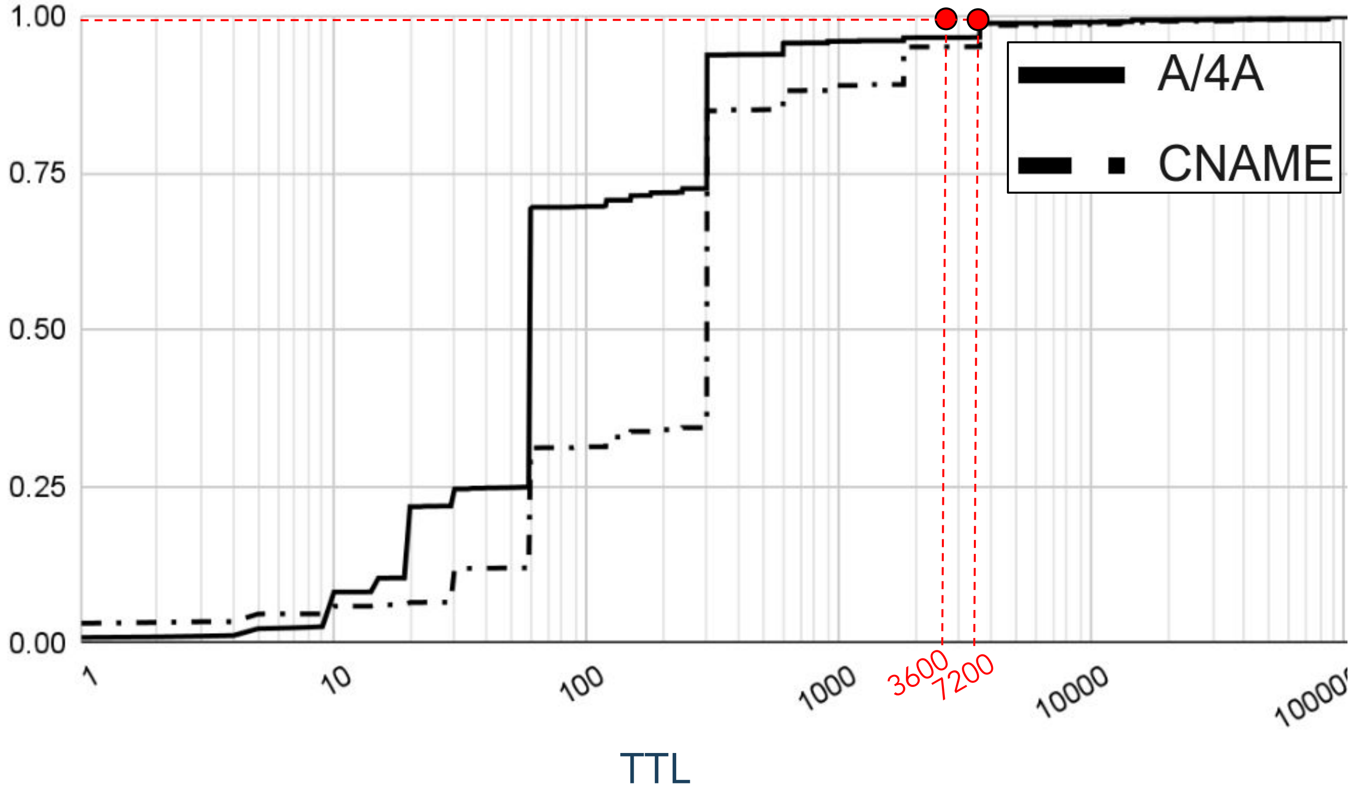


The number of URLs that are associated with an IP is really high
1:1
1:50K

RNA: Raw Network Analyser

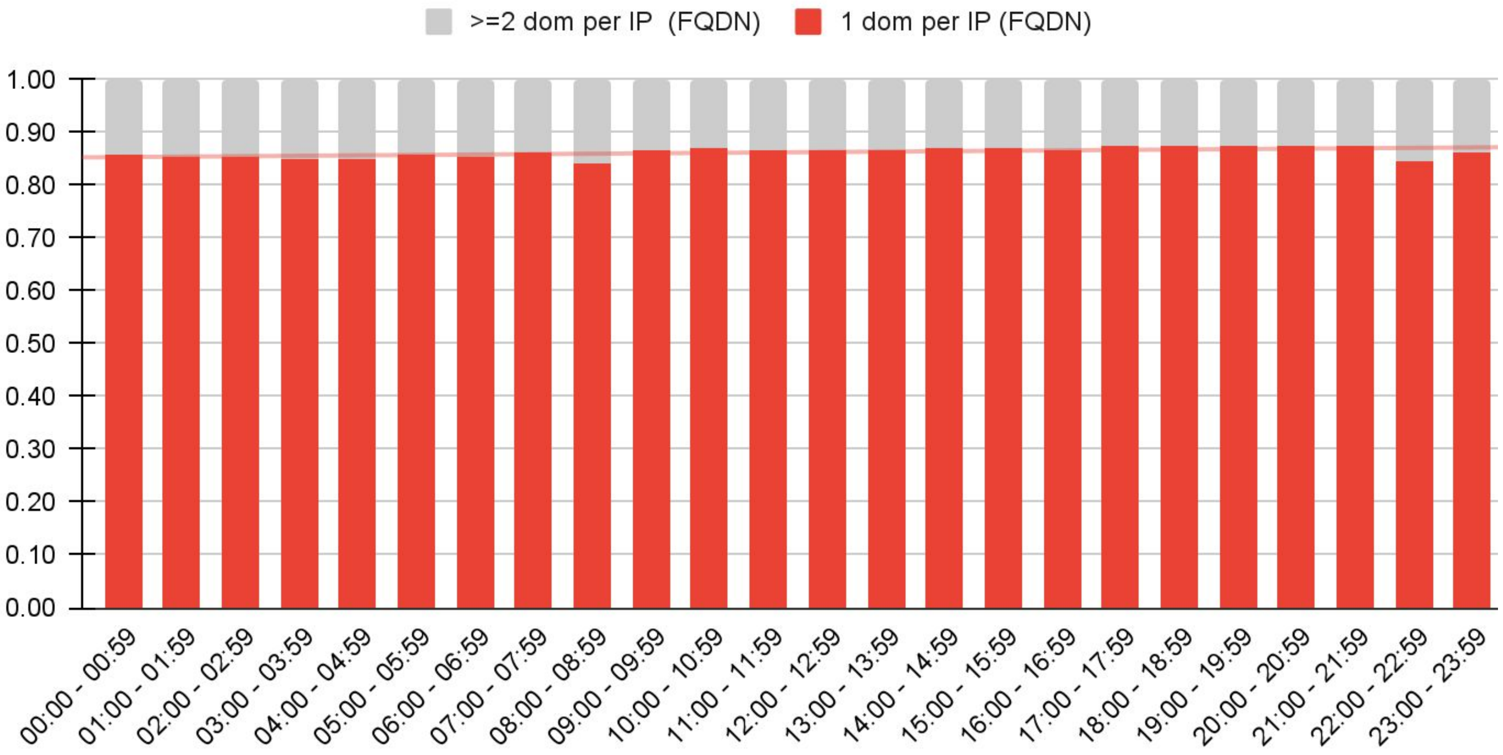
It has been developing a methodology that include the Analysis, Design, and Implementation of a real-time DNS-Netflow correlation system

TTL Distribution



- Memory constraints
  - What are the correct values for clear-up intervals?
- Analysis
  - A/4A
    - 3600 seconds
  - CNAME
    - 7200 seconds

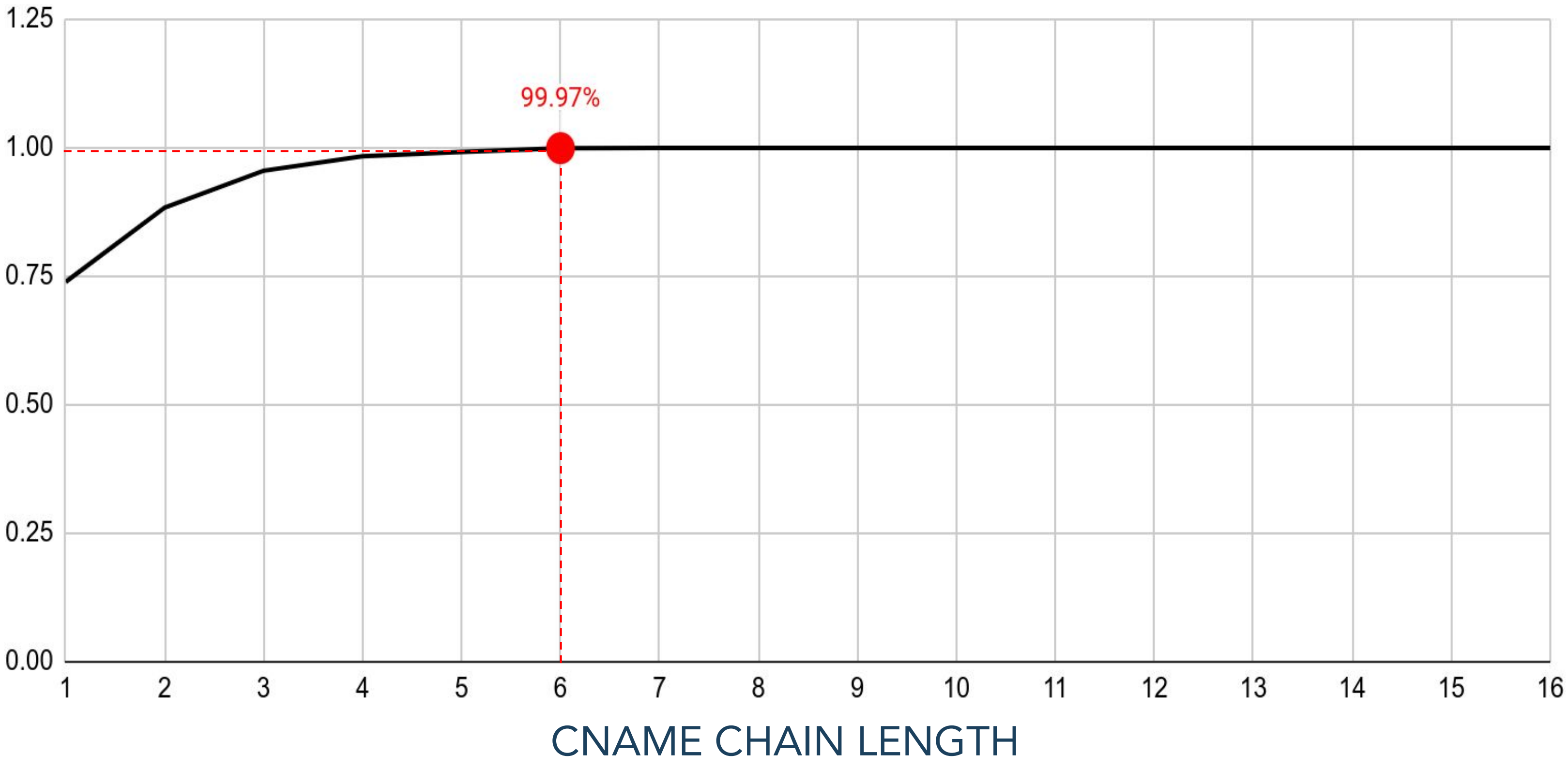
Domain names per IP Distribution



# DOMAINS PER IP

- Multiple domain names for one IP address
  - Affect the accuracy
  - Infrastructure level (A/4A Records)
- Analysis
  - Application level
    - CNAME starting the domain name resolution
    - Accuracy (FQDN):
      - ~86% of IP addresses are mapped to one single domain

Recursive CNAME Distribution



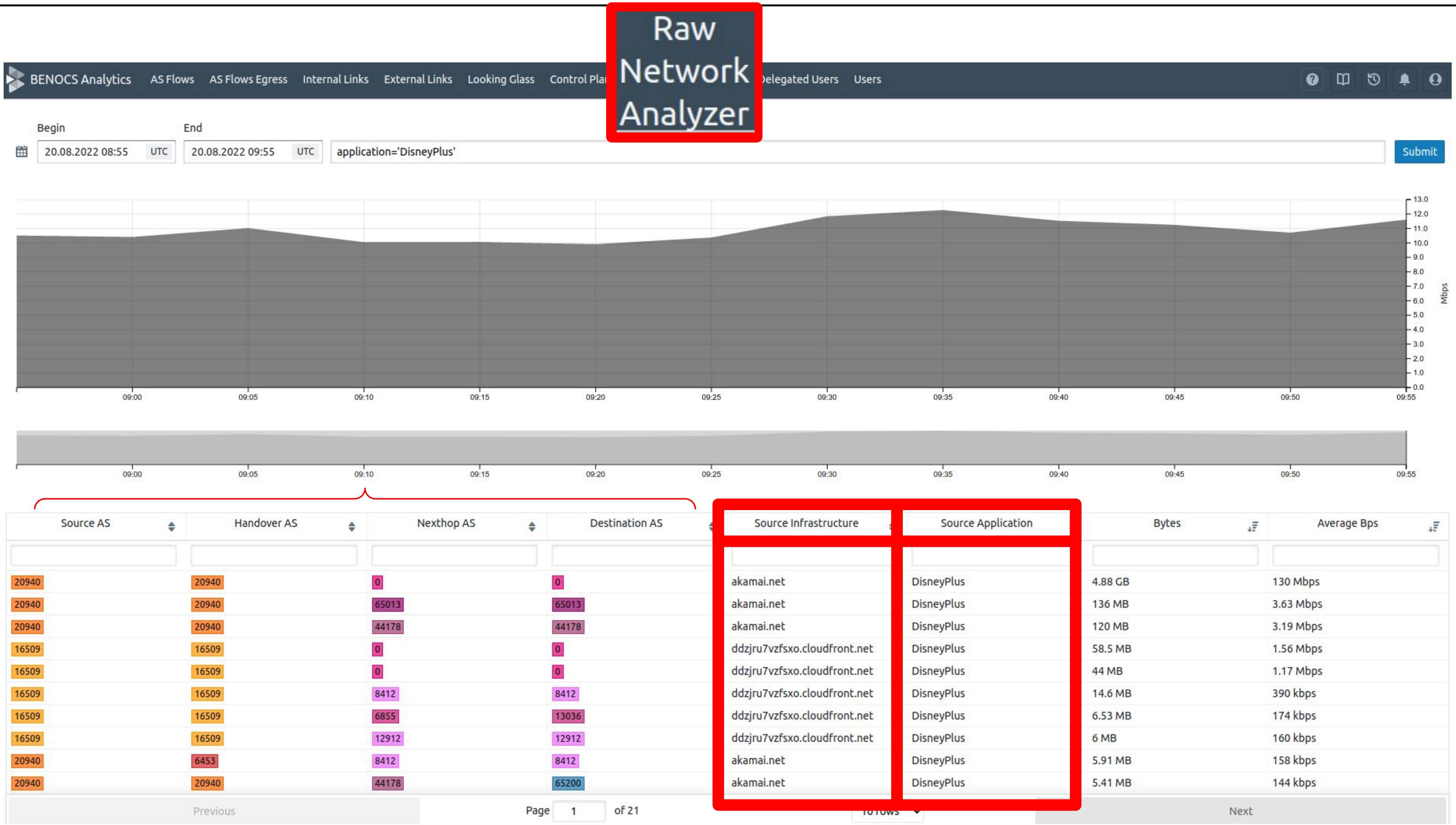
- Domain name resolution
  - Recursive lookup into the DNS records
  - What is a loop limit in this recursive searching process?
- Analysis
  - Loop limit of 6 is enough to complete the domain name resolution process

Public Suffix List & APP-URL Database

- Public Suffix List
  - Infrastructure classification:
    - 2LD + TLD
  - Source
    - Mozilla[1]: List of functional TLDs
- APP-URL Database
  - Application classification:
    - Associate specific domain names to the application they belong to
  - Source
    - Customers/Clients:
      - ~5500 URLs
    - External tools
      - ~199 APPs & ~3374 Urls
  - Proactive application classification
  - Only necessary DNS data

[1] https://publicsuffix.org/

RNA: Implementation



Raw Network Analyzer

Begin 20.08.2022 08:55 UTC

End 20.08.2022 09:55 UTC

application='DisneyPlus'

Source AS	Handover AS	Nexthop AS	Destination AS	Source Infrastructure	Source Application	Bytes	Average Bps
20940	20940	0	0	akamai.net	DisneyPlus	4.88 GB	130 Mbps
20940	20940	65013	65013	akamai.net	DisneyPlus	136 MB	3.63 Mbps
20940	20940	44178	44178	akamai.net	DisneyPlus	120 MB	3.19 Mbps
16509	16509	0	0	ddzjru7vzfsxo.cloudfront.net	DisneyPlus	58.5 MB	1.56 Mbps
16509	16509	0	0	ddzjru7vzfsxo.cloudfront.net	DisneyPlus	44 MB	1.17 Mbps
16509	16509	8412	8412	ddzjru7vzfsxo.cloudfront.net	DisneyPlus	14.6 MB	390 kbps
16509	16509	6855	13036	ddzjru7vzfsxo.cloudfront.net	DisneyPlus	6.53 MB	174 kbps
16509	16509	12912	12912	ddzjru7vzfsxo.cloudfront.net	DisneyPlus	6 MB	160 kbps
20940	6453	8412	8412	akamai.net	DisneyPlus	5.91 MB	158 kbps
20940	20940	44178	65200	akamai.net	DisneyPlus	5.41 MB	144 kbps

Page 1 of 21

BENOCS

Thanks!

(more) Comments/Questions?

✉ Danny Lachos
dlachos@benocs.com

BENOCS GmbH • Reuchlinstr. 10, 10553 Berlin
+49 30 577 000 4 – 17 •
www.benocs.com