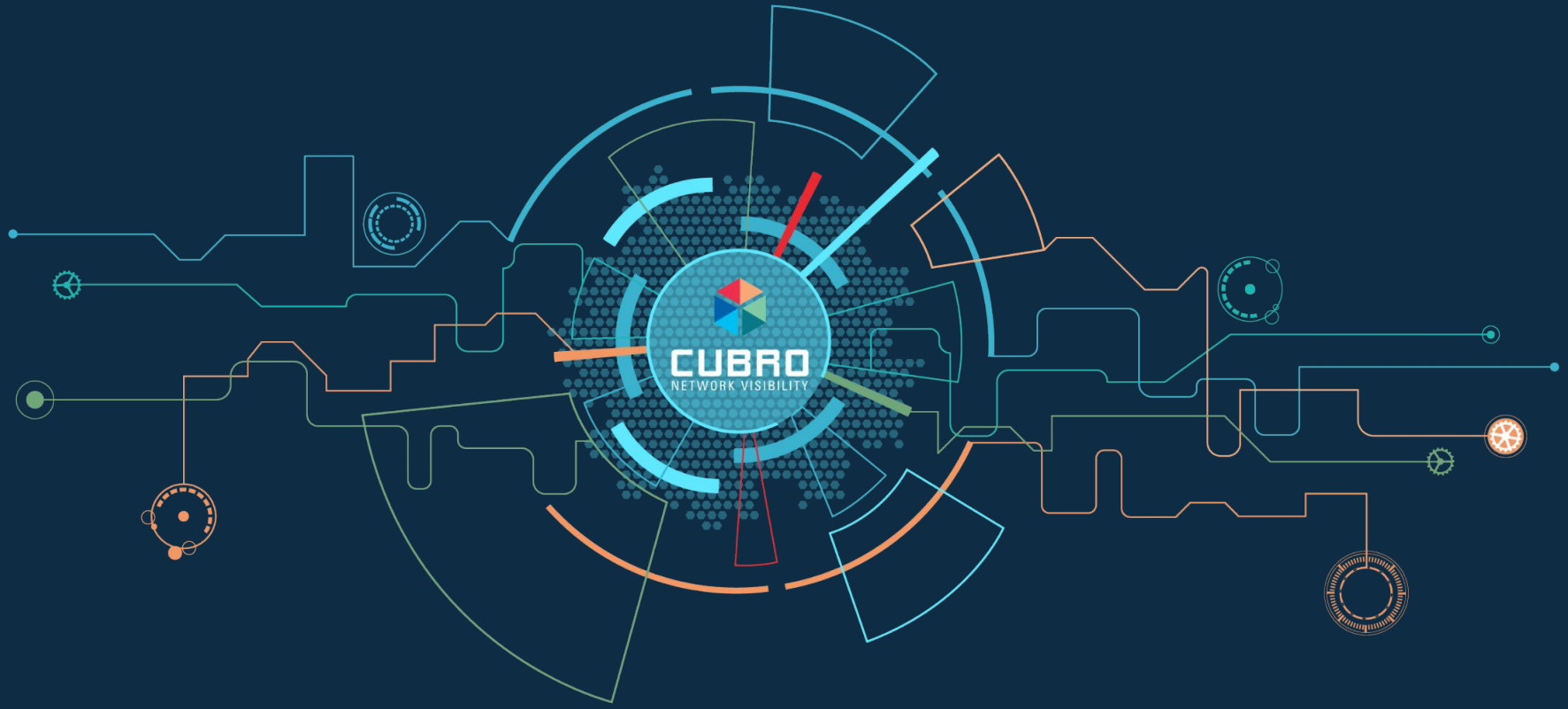


DPI Use Cases for Enterprise and Cubro Solution

Nov 2023

Introduction: DPI in Enterprise

Deep Packet Inspection (DPI) is a technology that enables the network owner to analyse internet traffic, through the network, in real-time and to differentiate them according to their payload.

DPI is often used for understanding the performance or behaviour of subscribers, which applications they use, how often etc. This helps operators to focus on improving service for the important applications. For instance, video streaming services like Netflix, YouTube, etc. consume a lot of bandwidth. DPI can be used to limit this.

Why DPI

Deep Packet Inspection (DPI) is used extensively by both enterprises and internet service providers for the following applications.

- Policy Definition and Enforcement
- Buffer Overflow Attack Detection
- Data Leak Prevention (DLP)
- Targeted Advertising
- Quality of Service (QoS)
- Tiered Services Offer
- Copyright Enforcement
- Net Neutrality Prevention
- Lawful Interception
- OTT application monitoring

(see our Video analytic approach)

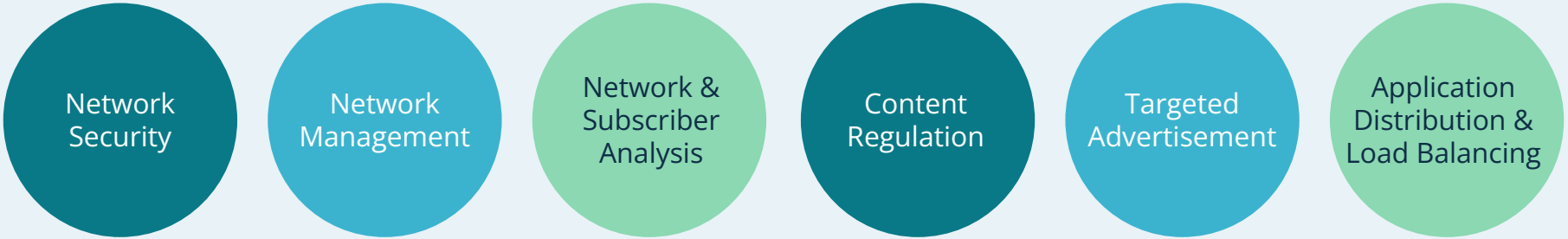
How is Omnia Metadata output used?

- Find, Identify, Classify, Reroute, and Block Packets with particular data/code payloads.
- Allocate available resources to smoothen traffic flow
- Ameliorate network performance and throughput
- Impose online privacy through sender-receiver identification
- Enable advanced network management, user service, internet data mining, internet censorship, and eavesdropping
- Ensure throttled data transfer, preventing P2P (Peer-to-Peer) misuse

Why is this needed?

Overview of DPI Applications

DPI facilitates analyzing and managing IP traffic and securing IP networks in real time by providing network visibility and real-time application awareness. Besides influencing bandwidth and traffic management decisions, DPI can provide insights into:



Use Case 1: Proactive Service Outage Alerts

Problem

Internet services are a major resource in any enterprise. Restoring a lost service is a significant responsibility of the IT department. The first step is always to identify precisely which service is affected, in which region, and at what time.

Solution

The Cubro solution offers real-time information about the service availability, like AWS, Microsoft, and other major services.

How does this help?

- The IT department is aware of this situation and can answer accordingly in real-time and save time on restoring the service. No need to generate support tickets at 3rd parties.
- Service outage monitoring to do better planning, and help with discussion in terms of SLA with the relevant provider.

Use Case 2: Misuse of Service/Resource

Problem

The value of IT resources and services is immense, but misusing them can lead to financial expenses and legal disputes.

Solution

The Cubro solution is capable of detecting this through unusual traffic/sessions/applications.

How does this help?

- To detect misuse of a resource by employees
- Prevent using illegal service because of bandwidth consumption (Video streaming)
- Preventing and blocking applications because of legal conflicts (P2P, Cryptomining, Darknet, etc)

Use Case 4: Security Detection

Problem

While security is crucial, it can be a costly investment due to the increasing demand for bandwidth, which drives up expenses.

Solution

The Cubro solution is capable of handling large amounts of data in an efficient way and produce detailed metadata of traffic/sessions/applications per device/user.

How does this help?

- Detect unusual network and subscriber behaviour
- Feed 3rd party tools with metadata
- Traffic filtering to offload and cost reduction on security tools

Measurement Metrics based on Cubro Metadata

User-based Metrics

- Monitor real-time total bandwidth usage of each individual user with one-second resolution.
- Calculate volume usage of each individual user for different time frames, such as Day/Week/Month/Year.
- Monitor real-time bandwidth usage per application of each individual user with one-second resolution.
- Calculate volume usage per application of each individual user for different time frames, such as Day/Week/Month/Year.
- Detect unwanted or forbidden applications per individual user.

Service Metrics

Service Metrics

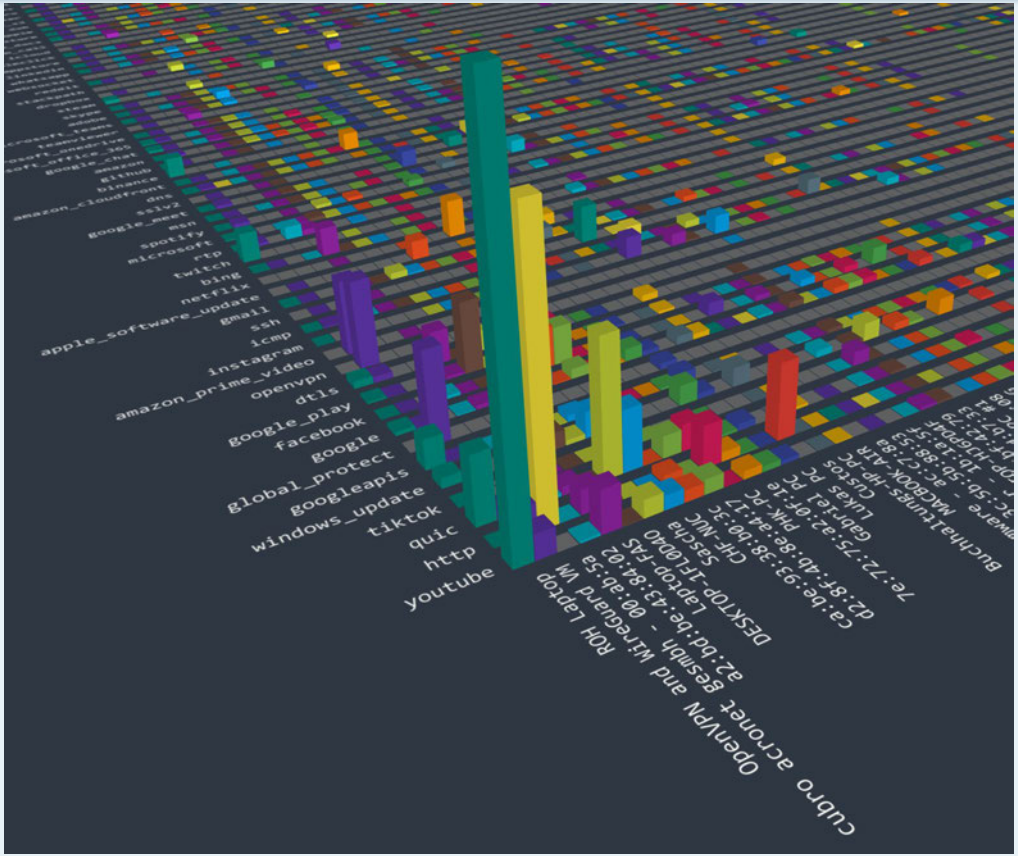
- Receive an alarm in the event of a service outage across the entire network, such as Netflix, Office364, and AWS services.
- Receive an alarm in the event of a service outage at specific sites, such as Netflix, Office364, and AWS services.
- Receive an alarm in the event of a service outage for a specific user or user group.
- Receive information when the service is restored.
- Monitor service usage statistics per individual service.
- Monitor service usage statistics per service groups, such as Social Media and Video Streaming.
- Monitor traffic distribution to Autonomous Systems (AS).
- Receive an alarm for unwanted services and apps. If necessary, inline blocking can be applied to the entire network or specific network sectors.

Network Element Metrics

Network Element Metrics

- Monitor near real-time bandwidth usage per cell with one-second resolution.
- Detect congestion near real-time per cell using burst detection.
- Monitor service distribution per cell for network planning purposes.
- Use machine learning to establish baselines and identify abnormal cell behaviour.

Graphical view out of the Cubro Metadata



Application vs. Endpoint

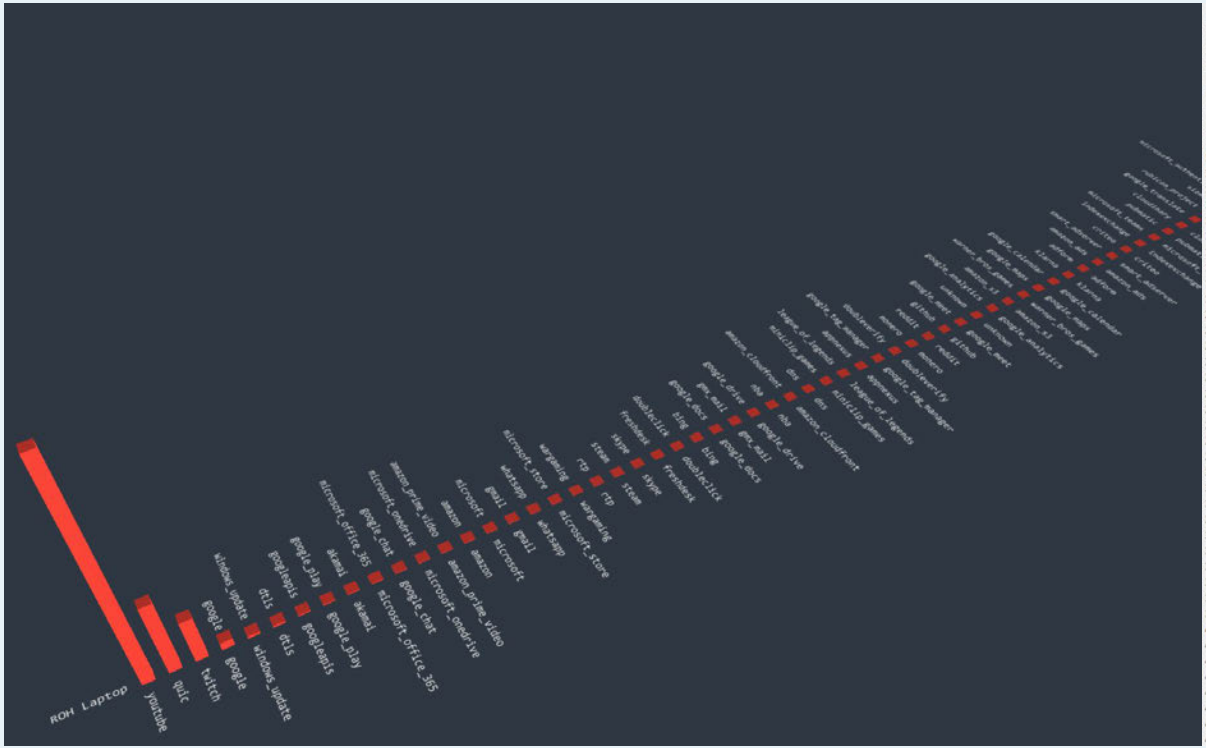
Bandwidth (one second resolution) per App
Volume per Application

Real-time total bandwidth usage of each individual subscriber in one second resolution

Volume calculation of each individual subscriber in any different time frame Day/Week/Month/year

The Cubro Metadata provides the opportunity to produce this kind of graphs for every individual subscriber.

Traffic Volume vs. application in any specific time frame.



Excel export

		ROH Laptop
2	youtube	23 296 311 820,00
3	windows_update	4 336 778 806,00
4	tls	3 832 502 635,00
5	quic	3 246 105 382,00
6	google	365 304 059,00
7	steam	214 476 206,00
8	twitch	135 345 653,00
9	dtls	82 897 466,00
10	microsoft	75 914 205,00
11	microsoft_onedrive	68 248 263,00
12	google_play	64 215 864,00
13	googleapis	63 583 620,00
14	microsoft_teams	48 722 924,00
15	gmail	44 883 040,00
16	google_chat	42 686 325,00
17	amazon	41 121 250,00
18	doubleclick	39 862 635,00
19	freshdesk	35 683 528,00
20	wargaming	35 002 858,00
21	riot_games	32 109 197,00
22	skype	25 141 635,00
23	whatsapp	23 534 390,00
24	microsoft_office_365	22 812 714,00
25	bing	20 108 655,00
26	gmx_mail	18 625 982,00
27	google_docs	17 275 120,00
28	microsoft_store	14 636 753,00
29	akamai	13 496 708,00
30	unknown	13 312 593,00
31	miniclip_games	12 912 396,00
32	google_drive	11 063 862,00
33	amazon_s3	10 684 587,00
34	appnexus	10 452 257,00
35	dns	9 967 262,00
36	microsoft_azure	8 330 188,00
37	google_tag_manager	7 432 996,00
38	twitter	6 997 179,00
39	github	6 641 741,00
40	doubleverify	6 619 527,00
41	google_meet	5 849 802,00
42	msn	5 342 842,00
43	amazon_cloudfront	5 331 667,00
44	monero	4 060 889,00
45	shopify	3 887 252,00
46	indexexchange	3 600 569,00
47	adform	3 374 687,00

DPI Applications

There are generally two different main applications for DPI

1. Analytics

A: In this application, the DPI engine can decode the full traffic and produce results in DB format for analytics purpose. This is only possible on CPU-based units like our OMNIA series. Since every packet has to be handled, it is a big effort in terms of CPU load and data output.

B: IPFIX with DPI enriched output. This is also a very common way of analyzing DPI data, but it is not very efficient and produces a lot of overhead. IPFIX on ISP level is very difficult because of the high amount of parallel sessions in the network. This often leads to issues on the installed probe like reaching memory limits.

2. Tagging/filtering/blocking

This application resonates with Cubro approach - remove an unwanted application type from the monitoring. **It is common to remove video streaming services.**

The same application is for blocking certain applications, or sending certain traffic to a special monitoring device. In this case, it is not needed to do a full decode because sampling gives a similar result but with much less effort.

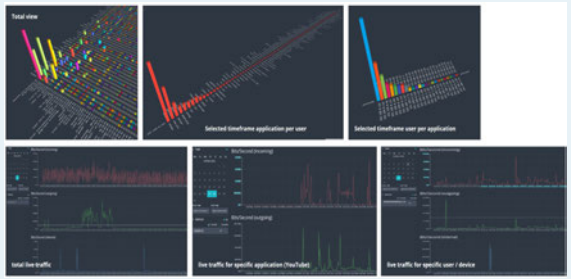
Cubro DPI Use Cases

DPI analytics in NPB

Advanced DPI analytics running on Cubro appliance.
Provide all kinds of measures.

- Performance per Application in select time frame
- Total volume per application in select time frame
- Volume user vs. Application in select time frame
- Total Volume per user in select time frame
- Export to Excel
- Geotraffic analytics
- Threat detection

and others



DPI metadata extraction

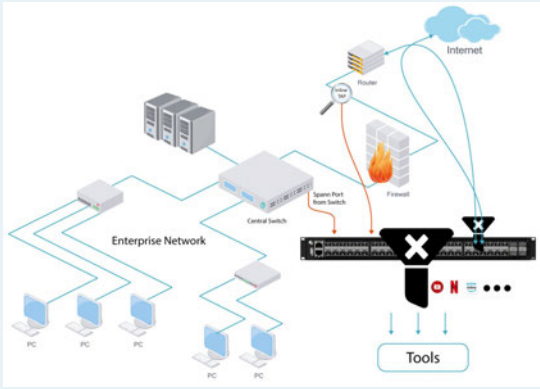
Metadata extraction based on the Cubro DPI engine, IPFIX / IPFIX & DPI / Cubro CDR.

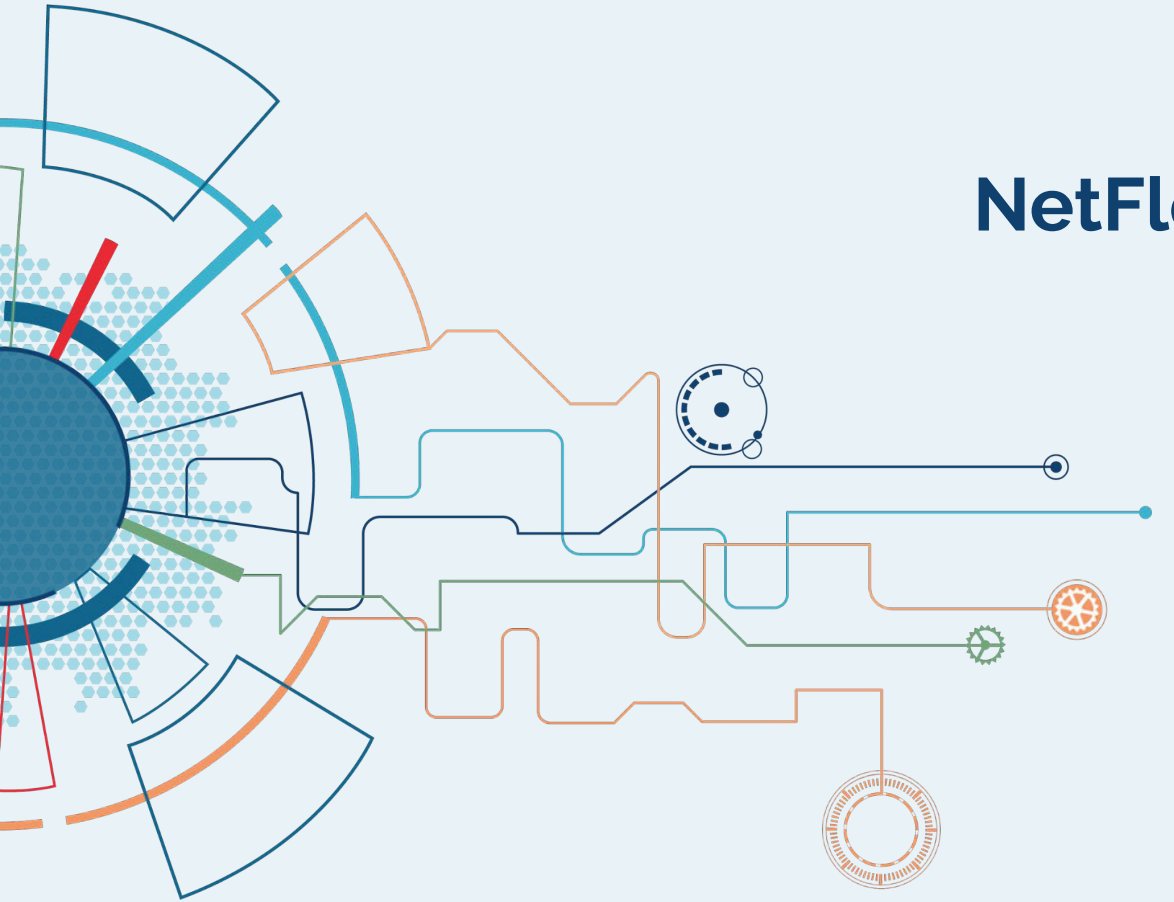
Cubro offers solutions from 100 Mbit/s up to multiple TBit/s input traffic.

Application blocking

Inline Application blocking from Gbit/s to TBit/s traffic. The solution also works based on the Cubro DPI.

This enables the application to block any application like Netflix, YouTube, WhatsApp, TikTok.



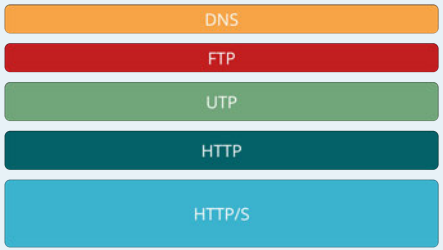


NetFlow DPI IPFIX

The Cubro DPI solution makes the difference

Deep packet inspection (DPI) and IP flow monitoring are frequently used network monitoring approaches. Although DPI provides application visibility, detailed examination of every packet is computationally intensive. IP flow monitoring achieves high performance by processing only packet headers, however, it fundamentally provides less detail about the traffic itself.

Application-aware flow monitoring is proposed to combine DPI accuracy and IP flow monitoring performance.



IPFIX

IPFIX with DPI

Time window based XDR
(Cubro XDR)

Time window-based Monitoring (1/2)

CUBRO NETWORK VISIBILITY

As described in the previous slides, there are several disadvantages to performing flow-based computation, mainly due to a lack of resources (CPU, memory, storage).

Aggregating metadata time window based solves most of these problems and allows a much higher computation throughput without losing any important metadata attributes.

The significant metadata cannot be extracted from single flows, as they mainly contain data from a technical perspective. The significant information is, for example, based on a device or user within the network. Which traffic is generated from a specific device, which services are used by a specific device and to what extent, how much traffic is generated by a user over time, which servers are involved, etc.

Time window-based Monitoring (2/2)

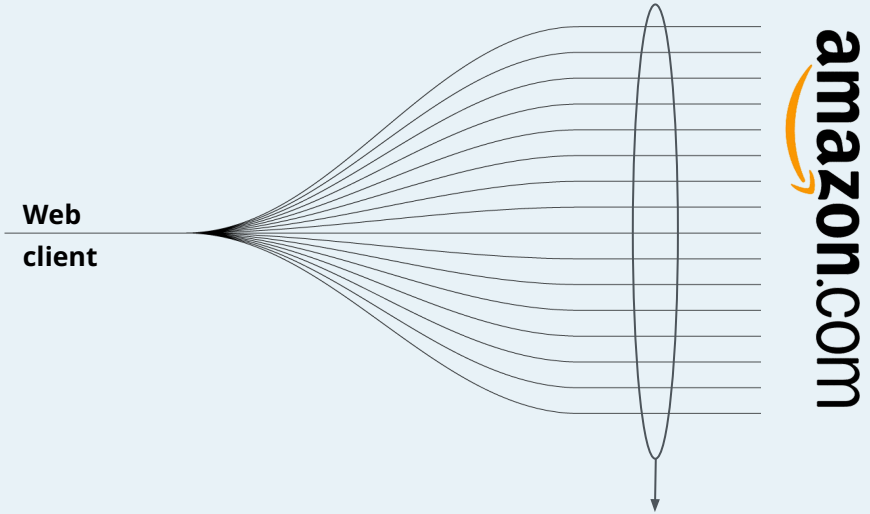
The time window-based approach focuses on this significant data.

The key is collecting Metadata for uploads, downloads, and internal traffic as well as DPI information from a device/user perspective over time. Extracting the essential information out of big data streams with the benefit of not wasting resources and storage for absolutely meaningless information (contained within single flows) is the important point.
The time window-based approach allows many essential views of the data:

- Service perspective: For every Service the amount of traffic, how often it is used over time, how much traffic is uploaded or downloaded, etc.
- Client/Device perspective: For every Client (user) the usage of the network by upload and download, which services and locations are used, how often they are used over time.

The difference (Single Flow to Session Flow aggregation)

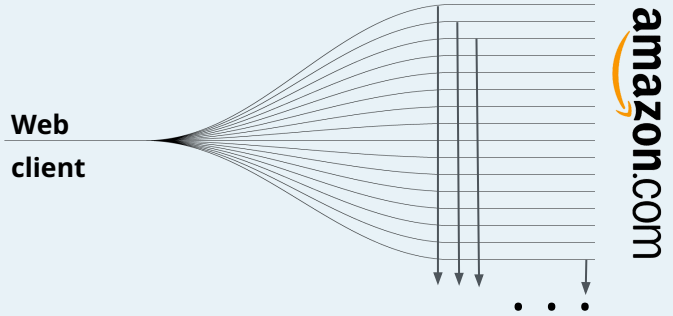
Time window based



For each client, there is a bucket for each application, if needed, and we collect/count all packets for a certain time window (configurable). When the window is closed, an XDR is produced/enriched and sent out. The advantage is, the traffic is reduced on most far points to avoid constraints on the workflow along to the database.

vs

Flow based



For each 5 tuple connection, one flow is produced. A lot of these flows cannot be detected, for instance, Amazon related because the external domain cannot be resolved.

After producing the flow, it is forwarded to the Flow Cache.
A Flow Cache can contain hundreds of thousands of entries, and in some cases, even millions of entries. This costs memory resources.

When the flows expire, they're exported off to the NetFlow Collector, which will constantly analyse and archive the flows for future reference.

The difference

Time window based

- We aren't concerned about sessions, (TCP handshake) perpetually open sessions are not an issue.
- Irregularly terminated/established sessions are also not an issue.
- The configurable time window offers the option to balance between performance constraints and granularity of the output.
- 0.1% - 0.5% of the input traffic is the size of the resulting metadata stream (configurable).

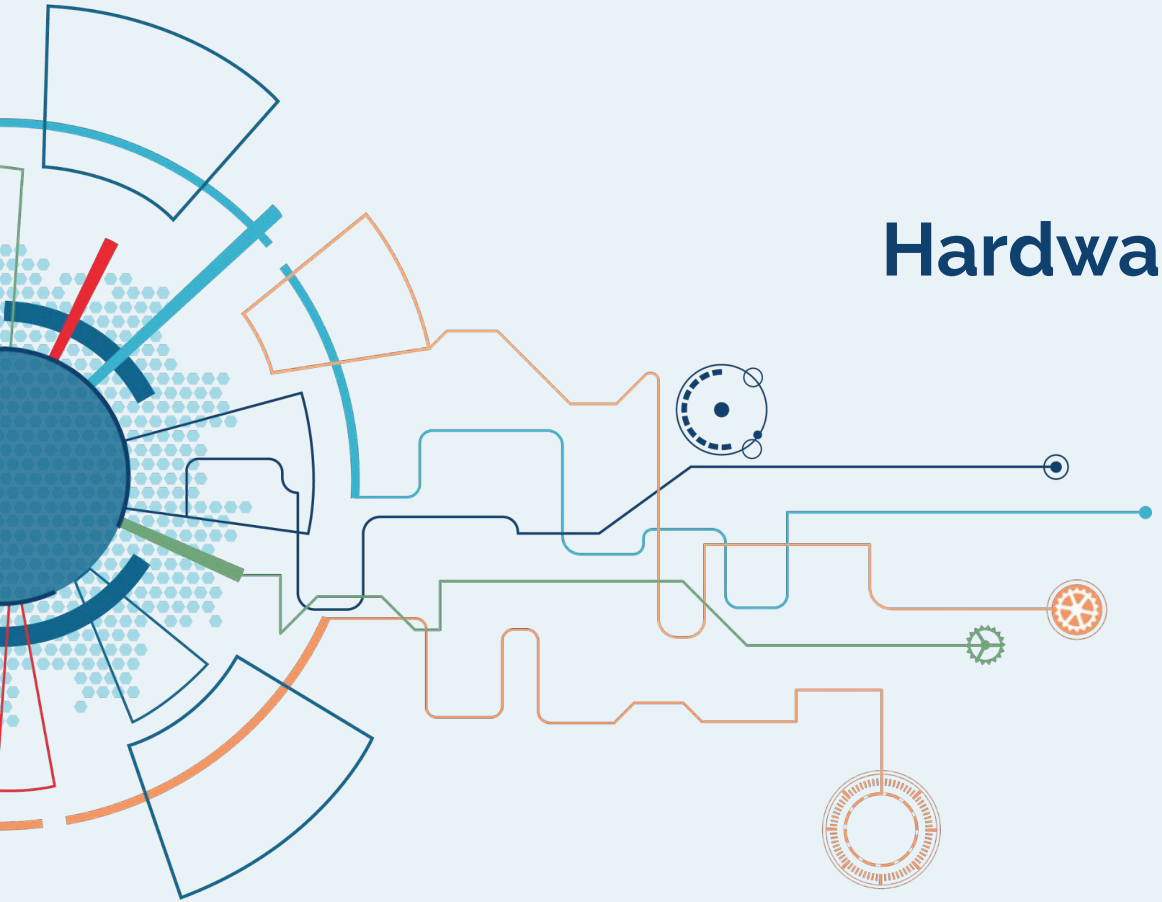
vs

Flow based

- The flow-based solution cares about sessions, (TCP handshake) perpetually open sessions are an issue. Typically, a flow probe has a limitation in terms of number of flows (FPS) not bandwidth.
- Irregularly terminated/established sessions are also an issue because an irregularly terminated session stays open until the timeout and consumes unnecessary resources.

  A session where the initial handshake is not seen for any reason will not be detected. For IoT especially, this could be an issue because such devices talk very rarely; it could be days until a session is detected again.
- 2% - 3% of the input traffic is the size of the resulting metadata stream.

For both solutions, there are pros and cons, but with the dramatic growth of traffic, a time window-based solution is much more efficient and saves Capex and Opex costs.

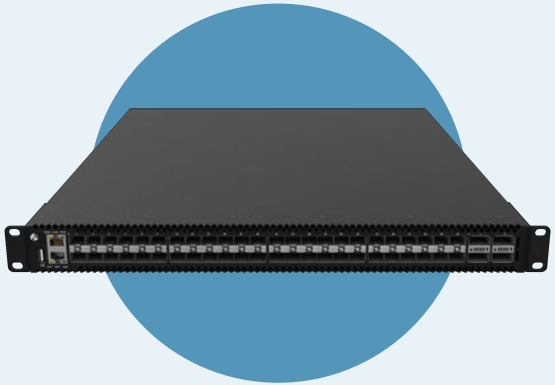


Hardware Appliance

CPU & Switch Omnia models overview

CUBRO NETWORK VISIBILITY

Omnia120



CPU only Omnia models overview

Omnia200

Omnia400

Omnic

Cubro Smart NIC

The implementation of high layer software on X86 server is becoming more and more complex. The performance of network service subsystem is a key factor to simplify application development and deployment.

Cubro i-NIC helps application software to offload network related processing from server CPU to a dedicated SoC, so that the application system can be accelerated with too much modification on existing software.

- **4 x 10/25 Gbit SFP+**
- **2 x 100 Gbit QSFP**
- **24 core ARM CPU**
- **64 GB Memory**
- **16 lane PCI connection (v4)**
- **Works also as a stand alone**

4 x 10/ 25 Gbit/s Interface

2 x 100 Gbit/s Interface

The passive DPI solution in detail

The super server holds three Cubro Smart NICs to perform 150 to 180 Gbit (or up to 500 Gbit with 1:4 sampling) DPI/ BGP correlation.

https://www.supermicro.com/en/products/system/Hyper/2U/SYS-220HE-FTNRD

3U small site solution up to 180 Gbit

Omnia120 as

- Load balancer
- Filter
- Signalling Probe

Super server with up to three Smart NICs for DPI analytics. Performance up to 150 to 180 Gbit/s.

User plane / Signalling correlation is done in the server CPU.

Technical Figures

Performance figures

(single use case estimations, can differ by traffic type and other external factors)

Product	IPFIX	IPFIX/DPI	Aggregated Cubro XDR with DPI (user plane)	Capture
Omnia120	20 Gbps on CPU 1	40 Gbps only one CPU 2	40 – 60 Gbps (on CPU 2)	Up to 6 TB / 8 - 10 Gbps performance
Omnia200	60 Gbps	60 Gbps	Up to 85 Gbps depending on the traffic situation	Up to 16 TB / 10 - 15 Gbps performance
Omnia400	60 Gbps (per CPU)	60 Gbps (per CPU)	160 Gbps (85 per CPU)	Up to 16 TB 10 - 15 Gbps performance
Omnic	40 Gbps	40 Gbps	Up to 85 Gbps depending on the traffic situation	No capture on NIC, but endless on the server via PCI
Omnic NG (road map Q1/23)	NA	NA	120+ Gbps depending on the traffic situation	No capture on NIC, but endless on the server via PCI

Raw packet storage calculation

Bandwidth in Gbit/s	1	3	6	12	24	Retention time in hours
0,1	0,08	0,25	0,5	1	2	
0,5	0,25	0,75	1,5	3	6	
1	0,46	1,375	2,75	5,5	11	
5	2,25	6,75	13,5	27	54	
10	4,50	13,5	27	54	108	
50	22,50	67,5	135	270	540	
100	45,00	135	270	540	1.080	in TB storage

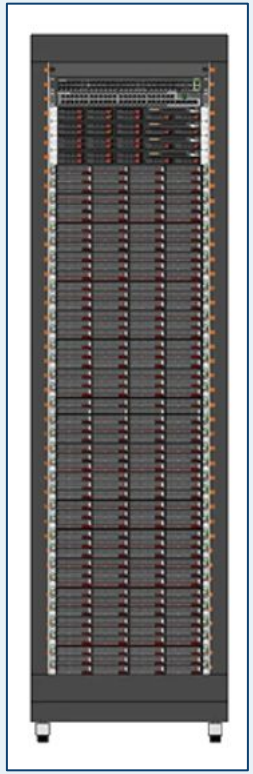
1024 TB = 1 PB Petabyte

Such a 42 Rack can support 3600 TB = 3,54 PB

432 x 8 TB HDD

2x 50A 208 3-Phase Metered PDU

Cost 600 - 890 k Euro depending on CPU RAM and HDD type



Bandwidth in Gbit/s	1	5	10	30	60	90	Retention time in days
0,1	2	6	11	33	65	98	
0,5	6	27	54	162	324	468	
1	11	54	108	324	648	972	
5	54	270	540	1.620	3.240	4.860	
10	108	540	1.080	3.240	6.480	9.720	
50	540	2.700	5.400	16.200	32.400	48.600	
100	1.080	5.400	10.800	32.400	64.800	97.200	in TB storage

The huge difference in Volume to store - 2,5 PB to 29 PB

Estimated IPFIX Metadata retention time

Bandwidth in Gbit/s	1	5	10	30	60	90	Retention time in days
0,1	0,060	0,180	0,330	0,990	1,950	2,940	
0,5	0,180	0,810	1,620	4,860	9,720	14,040	
1	0,330	1,62	3,240	9,720	19,440	29,160	
5	1,62	8,10	16,20	48,60	97,20	145,80	
10	3,24	16,20	32,40	97,20	194,40	291,60	
50	16,20	81	162	486	972	1.458	
100	32,40	162	324	972	1.944	2.916	
500	162	810	1.620	4.860	9.720	14.580	
1000	324	1.620	3.240	9.720	19.440	29.160	in TB storage

Estimated Cubro Metadata retention time

Bandwidth in Gbit/s	1	5	10	30	60	90	Retention time in days
0,1	0,00	0,014	0,029	0,086	0,173	0,259	
0,5	0,01	0,072	0,144	0,432	0,864	1,296	
1	0,03	0,14	0,29	0,86	1,73	2,59	
5	0,29	1,44	2,88	8,64	17,28	25,92	
10	0,72	3,60	7,20	21,60	43,20	64,80	
50	1,44	7,20	14,40	43,20	86,40	129,60	
100	2,88	14,40	28,80	86,40	172,80	259,20	
500	14,40	72	144	432	864	1.296	
1000	28,80	144	288	864	1.728	2.592	in TB storage

Cubro aggregated XDR volume calculation

UP Input	CDR/sec	CDR/H	Volume JSON/H Mbyte	Volume Binary/H Mbyte	Volume JSON/H Gbyte	Volume Binary/H Gbyte
25 Mbit	1,70	6.200	25	3		
100 Mbit	6,80	24.800	100	12		
250 Mbit	17	62.000	250	30		
500 Mbit	34	124.000	500	60		
1 Gbit	68	248.000	1.000	120	1	0,12
10 Gbit	680	2.480.000	10.000	1.200	10	1
25 Gbit	1.700	6.200.000	25.000	3.000	25	3
50 Gbit	3.400	12.400.000	50.000	6.000	50	6
100 Gbit	6.800	24.800.000	100.000	12.000	100	12
500 Gbit	34.000	124.000.000	500.000	60.000	500	60
1 Tbit	68.000	248.000.000	1.000.000	120.000	1.000	120

User plane Data (This data is based on the real output from a European Mobile SP)

Understand Internet traffic usage

	Static Web Page	Dynamic Web Page	Text e-mails	Attachment e-mails	Music streaming	Video streaming	IP voice call	IP Video call	Radio streaming	Application Data
15 MByte	151 pages	45 pages	15 k e-mails	30 e-mails	3 songs	2 min	27 min	3 min	44 min	2 Apps
500 MByte	5000 pages	1500 pages	500 k e-mails	1000 e-mails	100 songs	60 min	900 min	120 min	480 min	80 Apps

We have operations in all time zones.
Reach us at: support@cubro.com