Deep Packet Inspection
深度封包檢測

國防大學理工學院 資訊工程學系
國立陽明交通大學 系統工程與科技學士學位學程
羅嘉寧

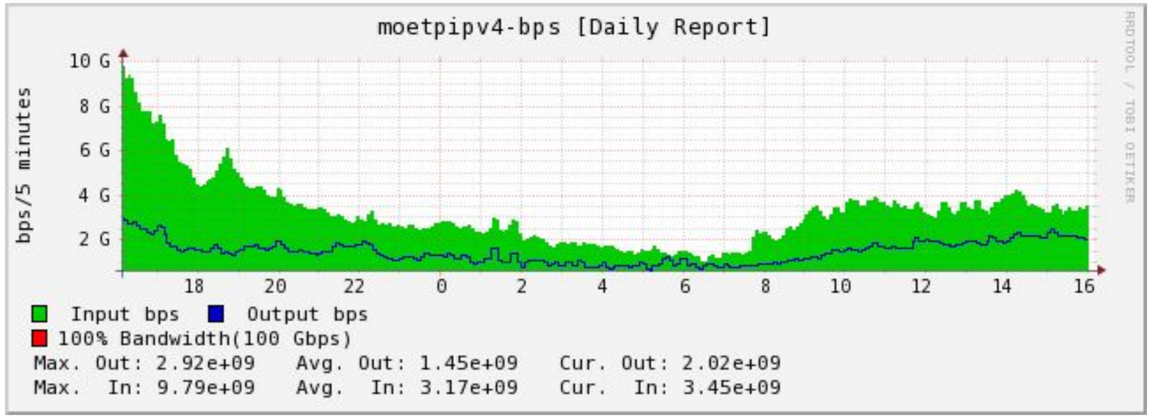
Background

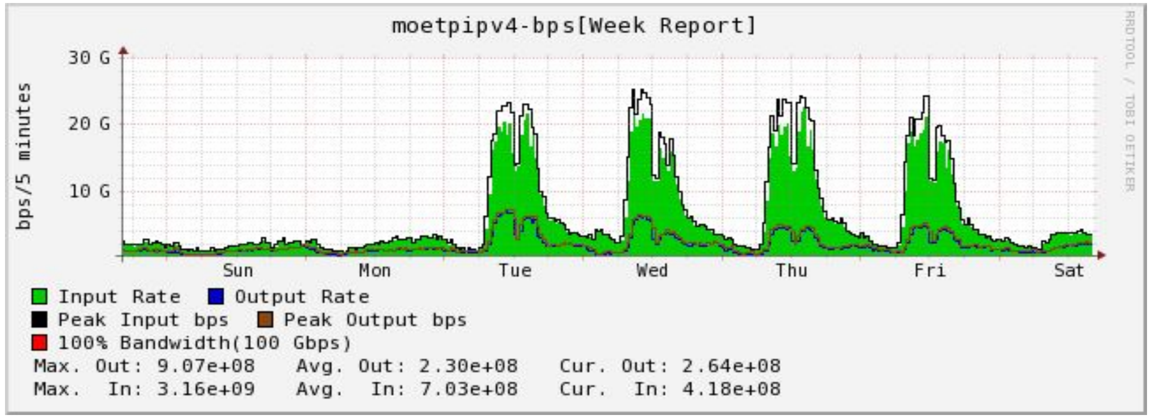
- Flow Analysis
  - How can I know the “Detail and Real-Time” situation of my network?
  - How can I do “Load Balance” job among the Outgoing Links?
  - How can I know the destination sites every “IP Block” belong to my network go to and the percentage of “Application” included on those flow?
  - How can I know “Who” connect to my network from outside? Is it attacking?
  - How I can estimate the “Growth” of my network?

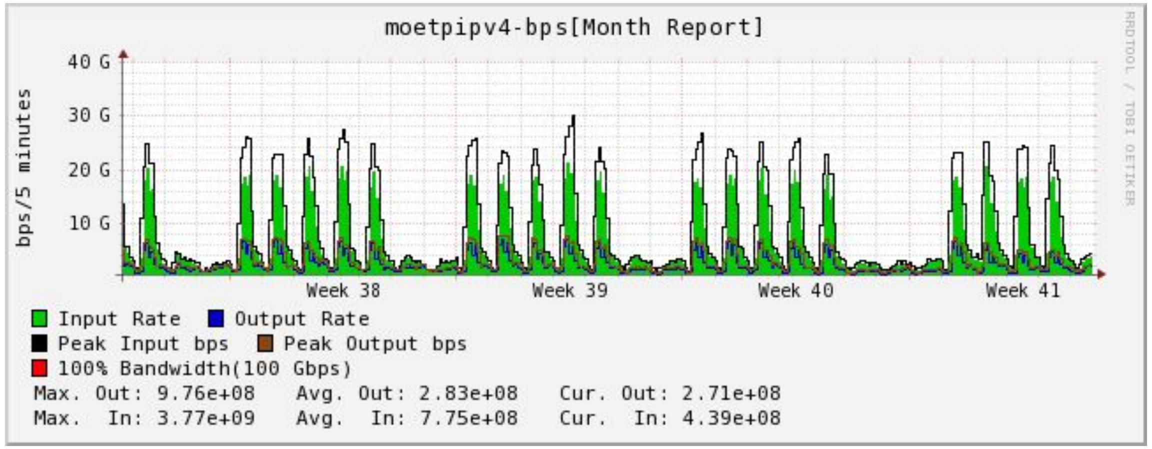
Background

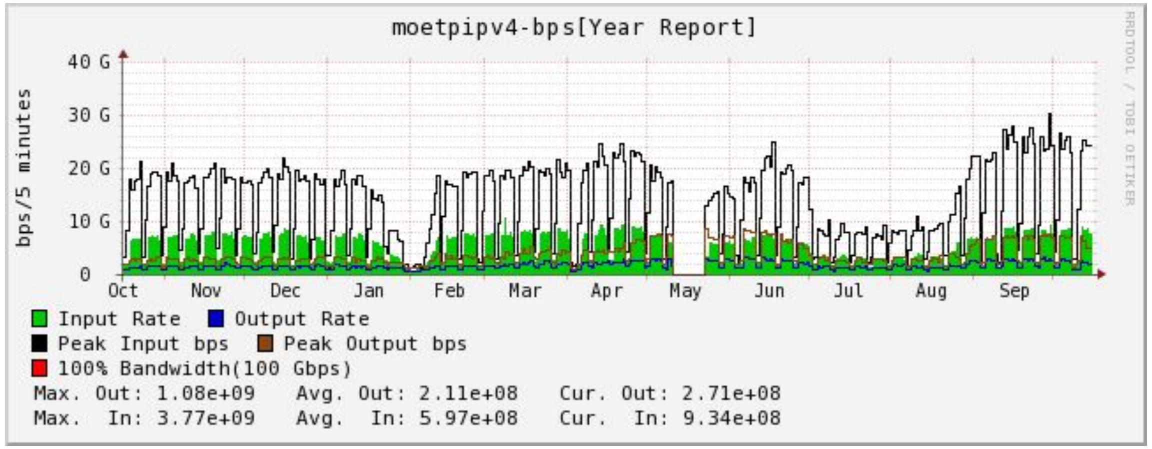
- Flow Analysis Techniques
  - Multi Router Traffic Grapher (MRTG)
  - Remote Network Monitoring (RMON)
    - enables various network monitors and console systems to exchange network-monitoring data
  - Cisco Netflow
  - Wireshark

Background – NCCU.edu.tw MRTG





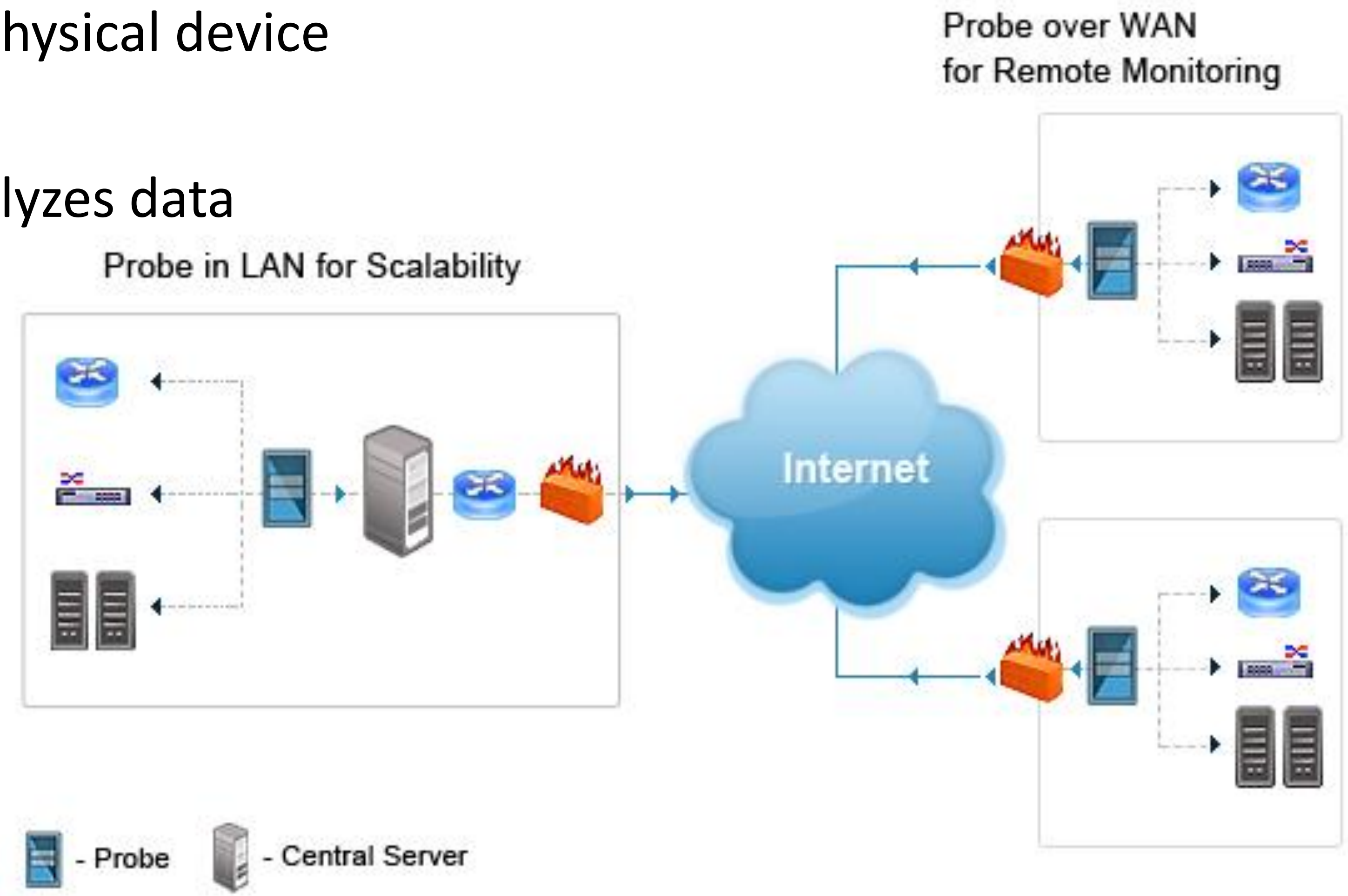




Data from: https://nms.moe.edu.tw/index.php/network-traffic

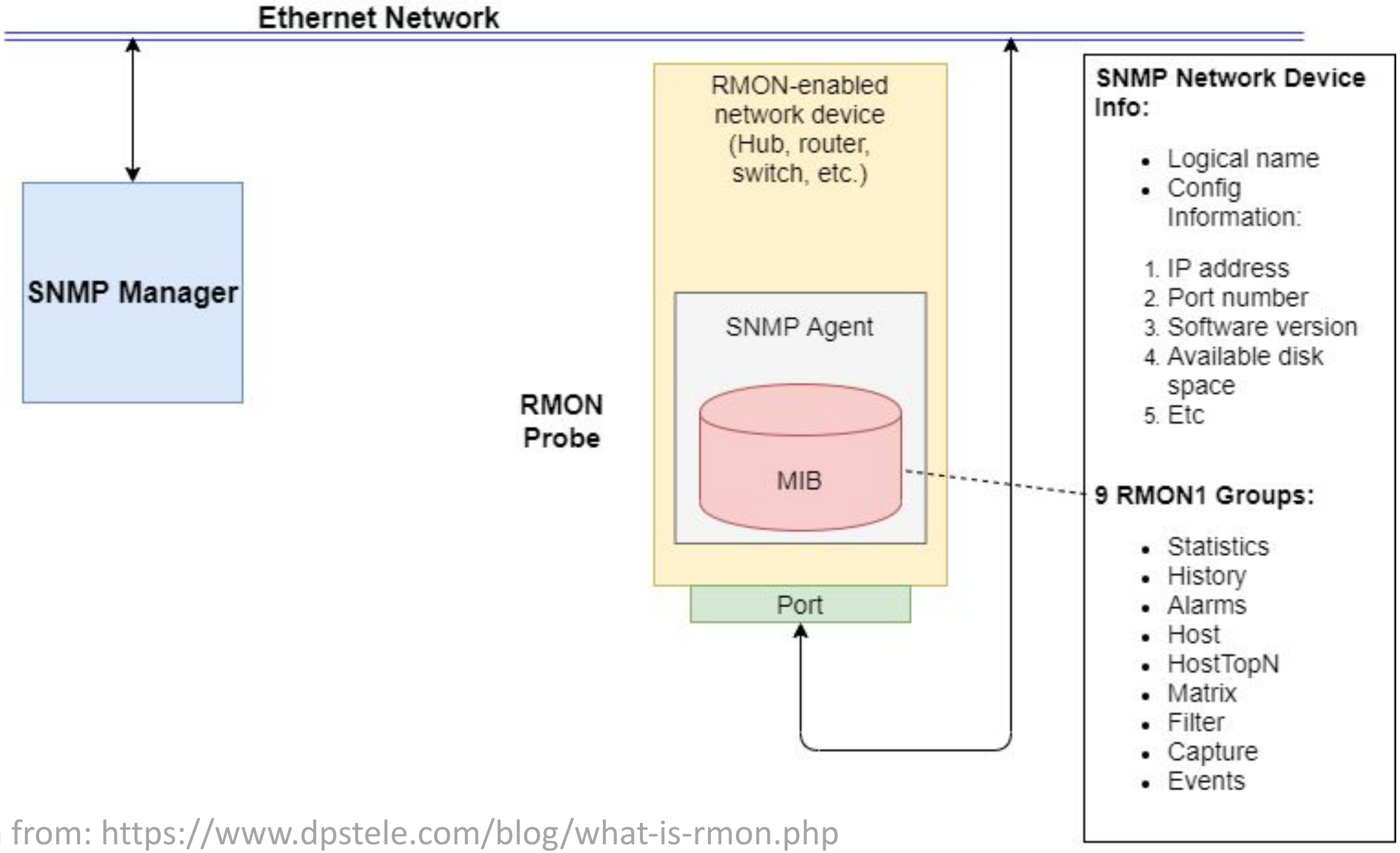
RMON: Remote Network Monitoring

- RMON Probe
  - Data gatherer - a physical device
- Data analyzer
  - Processor that analyzes data



Data from: https://www.manageengine.com/network-monitoring/remote-network-monitoring.html

Network with RMONs



SNMP Network Device Info:

- Logical name
- Config Information:

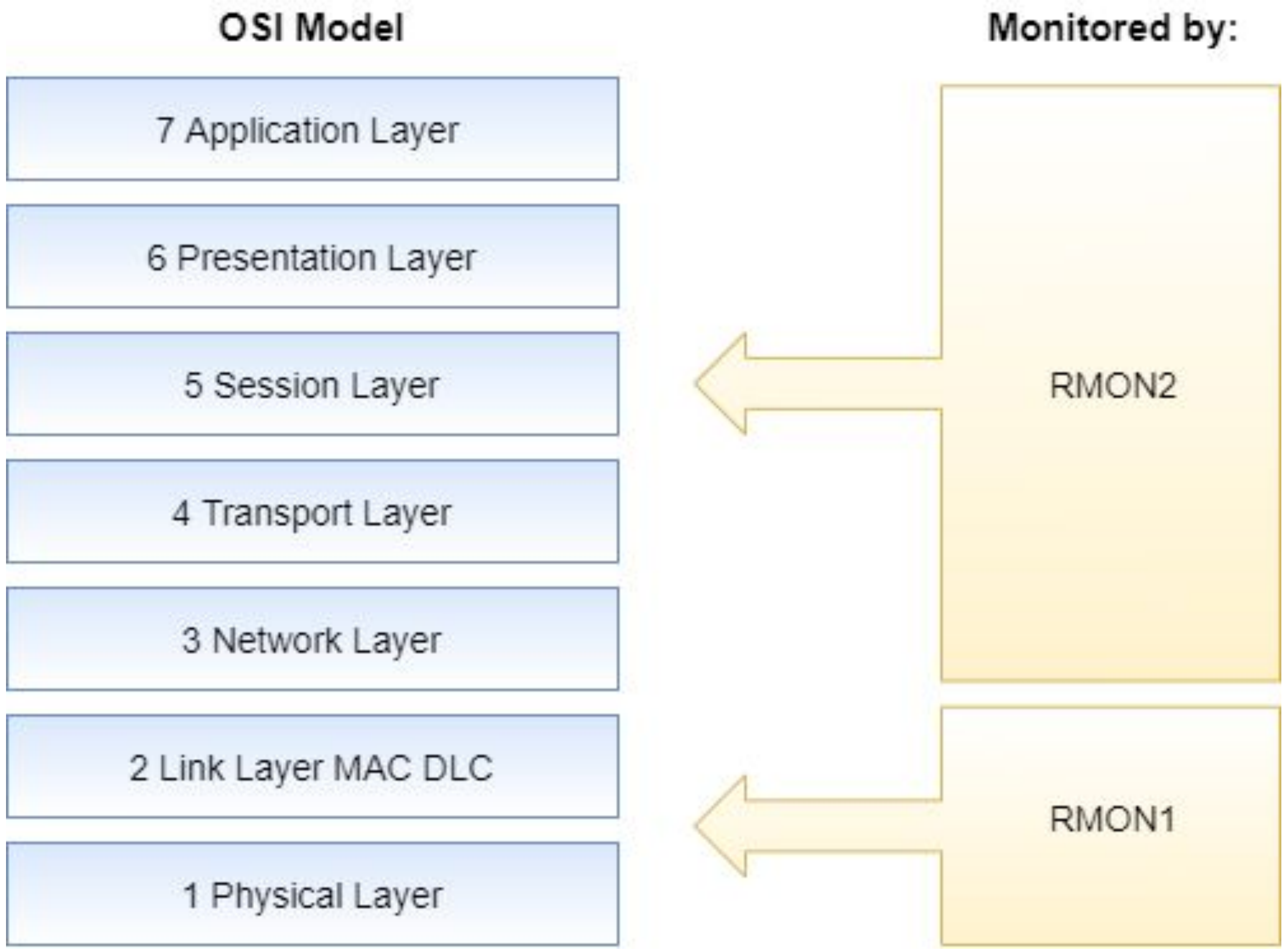
1. IP address
2. Port number
3. Software version
4. Available disk space
5. Etc

9 RMON1 Groups:

- Statistics
- History
- Alarms
- Host
- HostTopN
- Matrix
- Filter
- Capture
- Events

Data from: https://www.dpstele.com/blog/what-is-rmon.php

RMON1 vs RMON2



OSI Model

7 Application Layer

6 Presentation Layer

5 Session Layer

4 Transport Layer

3 Network Layer

2 Link Layer MAC DLC

1 Physical Layer

Monitored by:

RMON2

RMON1

Data from: https://www.dpstele.com/blog/what-is-rmon.php

Remote Network Monitoring Goals

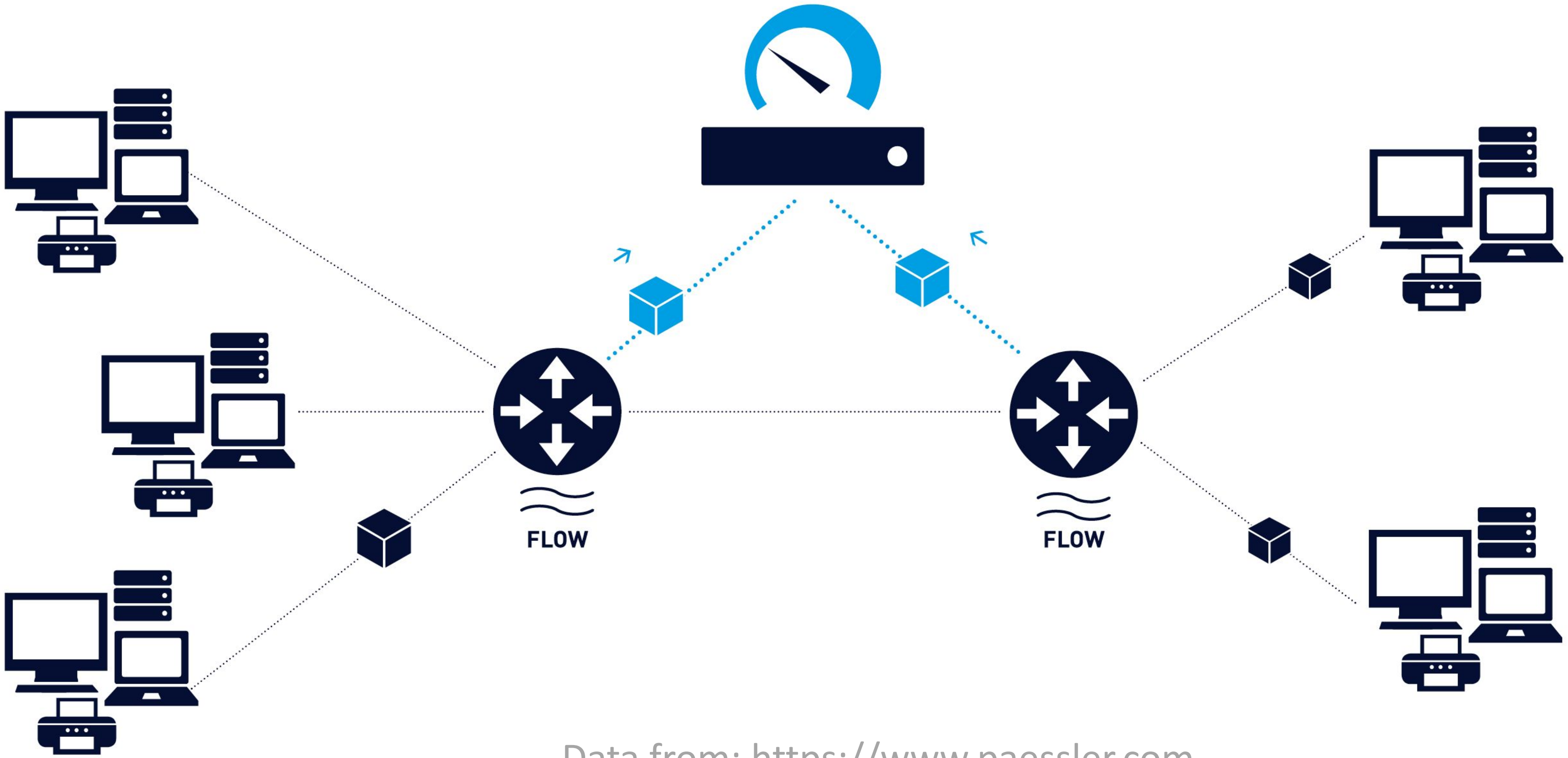
- Offline Operation
  - Perform diagnostics and to collect statistics continuously, even when communication with the management station may not be possible or efficient.
- Proactive Monitoring
  - Continuously run diagnostics and log network performance.
- Problem Detection and Reporting
  - Given conditions, the probe continuously to check for them.
  - If there any condition occurs, notify the manager.
- Value Added Data
  - Who generate the most traffic or errors, …
- Multiple Managers

Remote Network Monitoring Benefits

- Monitors and analyzes locally and relays data;
  Less load on the network
- Needs no direct visibility by NMS;
  More reliable information
- Permits monitoring on a more frequent basis and hence faster fault diagnosis
- Increases productivity for administrators

Netflow by Cisco

Netflow: collect IP network traffic as it enters an interface

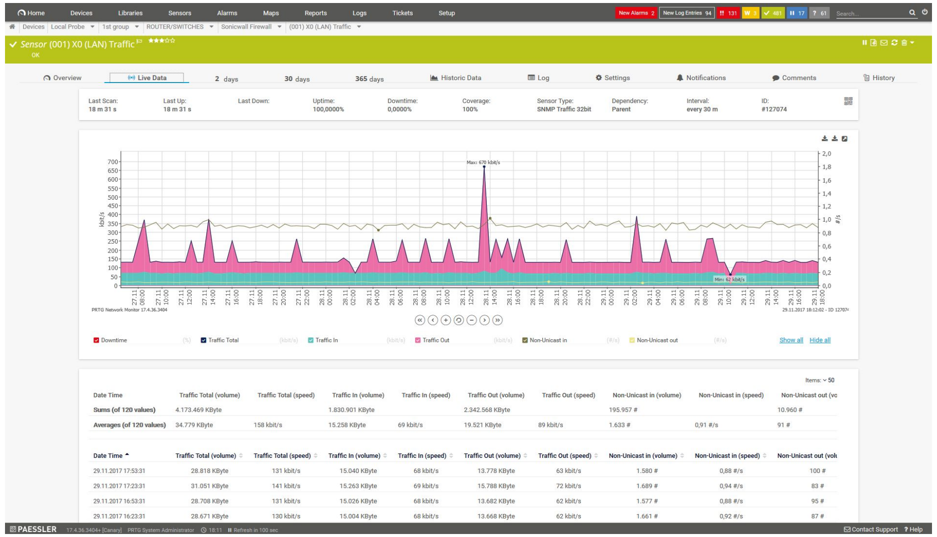


DATA ACQUISITION USING FLOW

Data from: https://www.paessler.com

Netflow collected information

- The following information can be obtained from Netflow packets
  - Source and Destination addresses
  - Input and Output interface numbers
  - Source and Destination port numbers
  - Layer 4 protocol
  - Number of packets in the flow
  - Total Bytes in the flow
  - Time stamp in the flow
  - Source and Destination autonomous system (AS) number
  - TCP_Flag and Type of Service (ToS)



✓ *Sensor* (001) X0 (LAN) Traffic ★★★☆☆

OK

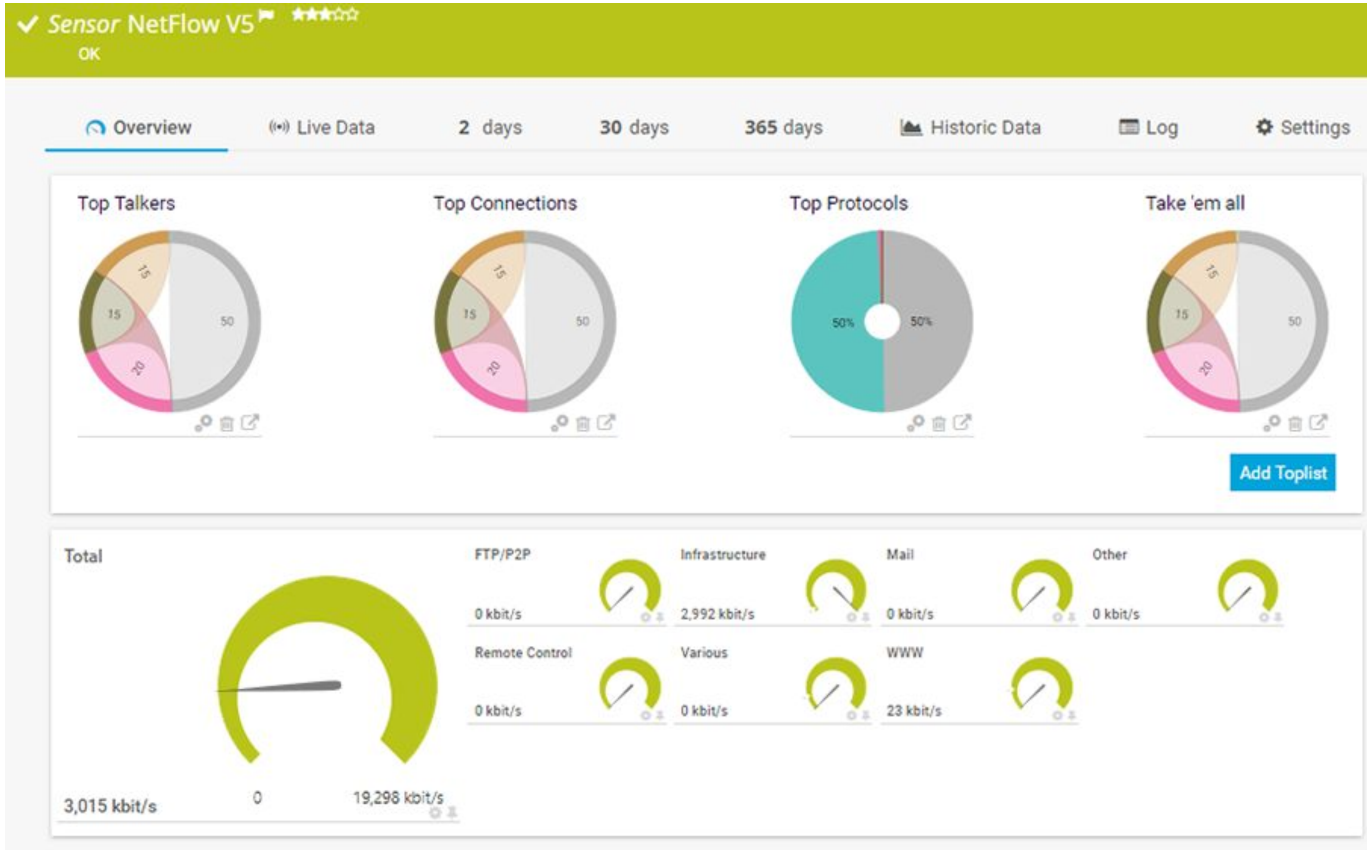
Last Scan:	Last Up:	Last Down:	Uptime:	Downtime:	Coverage:	Sensor Type:	Dependency:	Interval:	ID:
18 m 31 s	18 m 31 s		100,0000%	0,0000%	100%	SNMP Traffic 32bit	Parent	every 30 m	#127074

☑ Downtime (%) ☑ Traffic Total (kbit/s) ☑ Traffic In (kbit/s) ☑ Traffic Out (kbit/s) ☑ Non-Unicast in (#/s) ☑ Non-Unicast out (#/s)

Items: 50

Date Time	Traffic Total (volume)	Traffic Total (speed)	Traffic In (volume)	Traffic In (speed)	Traffic Out (volume)	Traffic Out (speed)	Non-Unicast in (volume)	Non-Unicast in (speed)	Non-Unicast out (vo
Sums (of 120 values)	4.173.469 KByte		1.830.901 KByte		2.342.568 KByte		195.957 #		10.960 #
Averages (of 120 values)	34.779 KByte	158 kbit/s	15.258 KByte	69 kbit/s	19.521 KByte	89 kbit/s	1.633 #	0,91 #/s	91 #

Date Time	Traffic Total (volume)	Traffic Total (speed)	Traffic In (volume)	Traffic In (speed)	Traffic Out (volume)	Traffic Out (speed)	Non-Unicast in (volume)	Non-Unicast in (speed)	Non-Unicast out (volu
29.11.2017 17:53:31	28.818 KByte	131 kbit/s	15.040 KByte	68 kbit/s	13.778 KByte	63 kbit/s	1.580 #	0,88 #/s	100 #
29.11.2017 17:23:31	31.051 KByte	141 kbit/s	15.263 KByte	69 kbit/s	15.788 KByte	72 kbit/s	1.689 #	0,94 #/s	83 #
29.11.2017 16:53:31	28.708 KByte	131 kbit/s	15.026 KByte	68 kbit/s	13.682 KByte	62 kbit/s	1.577 #	0,88 #/s	95 #
29.11.2017 16:23:31	28.671 KByte	130 kbit/s	15.004 KByte	68 kbit/s	13.668 KByte	62 kbit/s	1.661 #	0,92 #/s	87 #



✓ *Sensor* NetFlow V5 ★★★☆☆

OK

Top Talkers

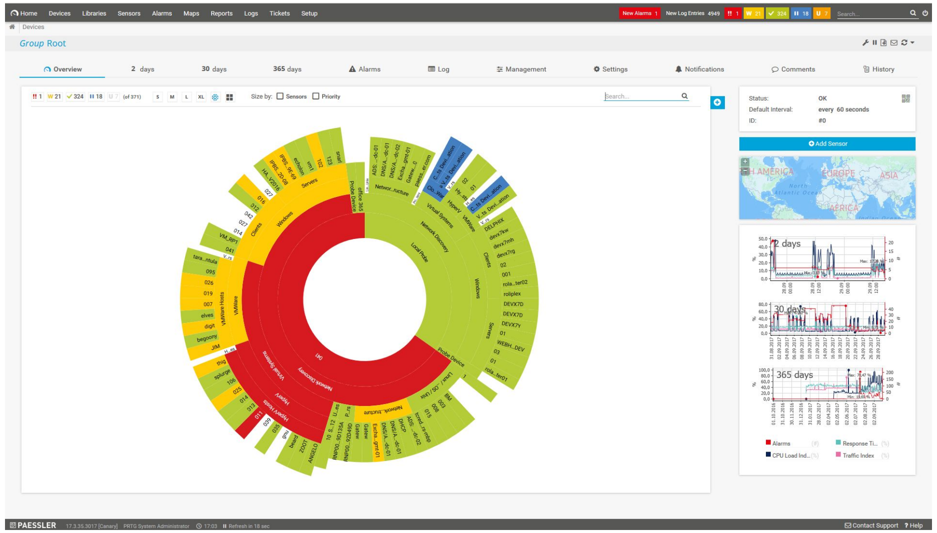
Top Connections

Top Protocols

Take 'em all

Total 3,015 kbit/s (0 – 19,298 kbit/s)

FTP/P2P	0 kbit/s
Infrastructure	2,992 kbit/s
Mail	0 kbit/s
Other	0 kbit/s
Remote Control	0 kbit/s
Various	0 kbit/s
WWW	23 kbit/s

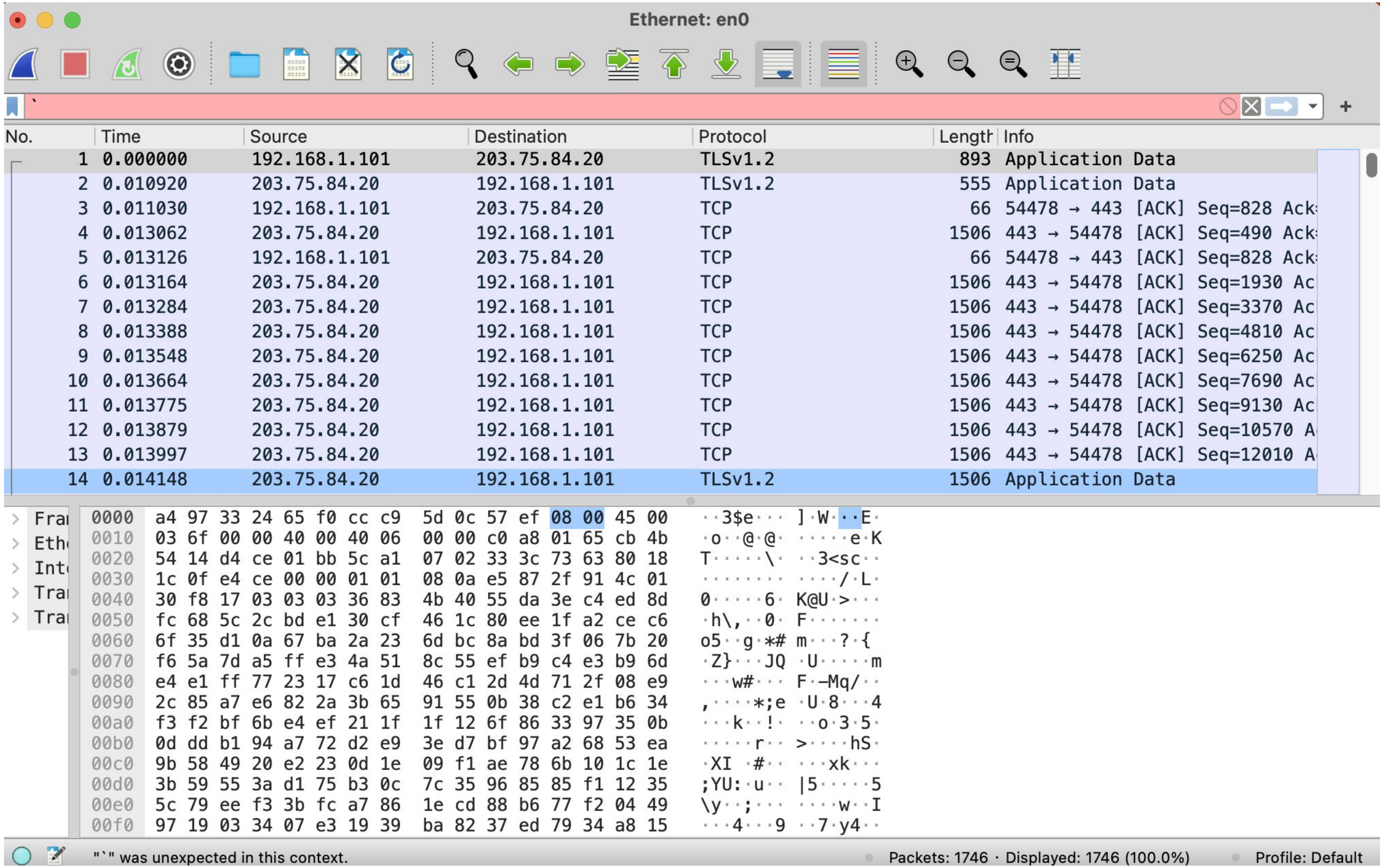


Data from: https://www.paessler.com

Wireshark

- Packet analyser / traffic sniffer
- Open-source
- Cross-platform
- Fancy GUI
- https://www.wireshark.org/

Wireshark



```
No.  Time      Source          Destination     Protocol  Length Info
   1 0.000000  192.168.1.101   203.75.84.20    TLSv1.2     893 Application Data
   2 0.010920  203.75.84.20    192.168.1.101   TLSv1.2     555 Application Data
   3 0.011030  192.168.1.101   203.75.84.20    TCP          66 54478 → 443 [ACK] Seq=828 Ack
   4 0.013062  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=490 Ack
   5 0.013126  192.168.1.101   203.75.84.20    TCP          66 54478 → 443 [ACK] Seq=828 Ack
   6 0.013164  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=1930 Ac
   7 0.013284  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=3370 Ac
   8 0.013388  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=4810 Ac
   9 0.013548  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=6250 Ac
  10 0.013664  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=7690 Ac
  11 0.013775  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=9130 Ac
  12 0.013879  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=10570 A
  13 0.013997  203.75.84.20    192.168.1.101   TCP        1506 443 → 54478 [ACK] Seq=12010 A
  14 0.014148  203.75.84.20    192.168.1.101   TLSv1.2    1506 Application Data
```

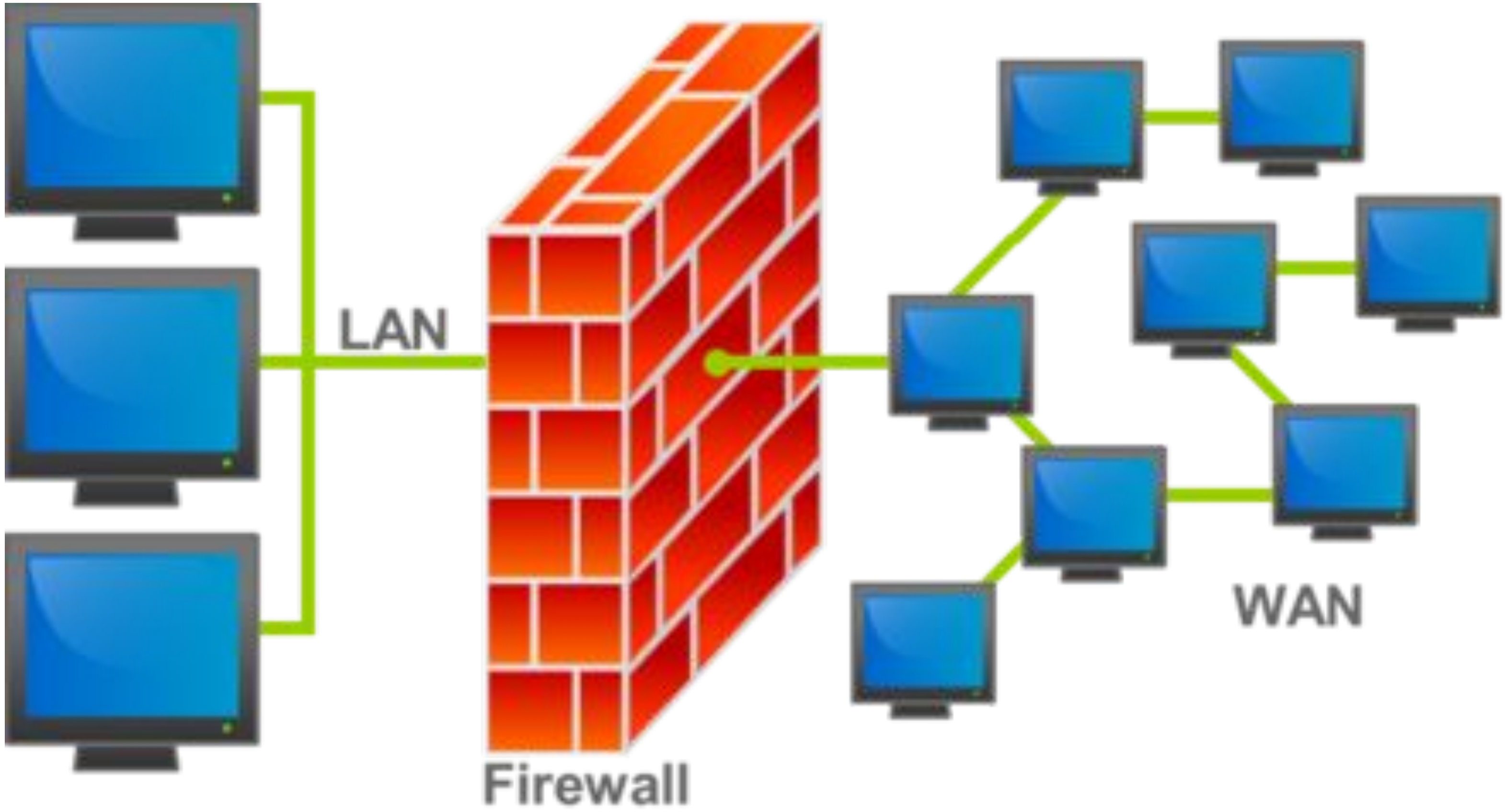
```
> Fra   0000  a4 97 33 24 65 f0 cc c9  5d 0c 57 ef 08 00 45 00   ··3$e···  ]·W··E·
> Eth   0010  03 6f 00 00 40 00 40 06  00 00 c0 a8 01 65 cb 4b   ·o··@·@·  ·····e·K
> Int   0020  54 14 d4 ce 01 bb 5c a1  07 02 33 3c 73 63 80 18   T·····\·  ··3<sc··
> Tra   0030  1c 0f e4 ce 00 00 01 01  08 0a e5 87 2f 91 4c 01   ········  ····/·L·
> Tra   0040  30 f8 17 03 03 03 36 83  4b 40 55 da 3e c4 ed 8d   0·····6·  K@U·>···
        0050  fc 68 5c 2c bd e1 30 cf  46 1c 80 ee 1f a2 ce c6   ·h\,··0·  F·······
        0060  6f 35 d1 0a 67 ba 2a 23  6d bc 8a bd 3f 06 7b 20   o5··g·*#  m···?·{
        0070  f6 5a 7d a5 ff e3 4a 51  8c 55 ef b9 c4 e3 b9 6d   ·Z}···JQ  ·U·····m
        0080  e4 e1 ff 77 23 17 c6 1d  46 c1 2d 4d 71 2f 08 e9   ···w#···  F·-Mq/··
        0090  2c 85 a7 e6 82 2a 3b 65  91 55 0b 38 c2 e1 b6 34   ,····*;e  ·U·8···4
        00a0  f3 f2 bf 6b e4 ef 21 1f  1f 12 6f 86 33 97 35 0b   ···k··!·  ··o·3·5·
        00b0  0d dd b1 94 a7 72 d2 e9  3e d7 bf 97 a2 68 53 ea   ·····r··  >····hS·
        00c0  9b 58 49 20 e2 23 0d 1e  09 f1 ae 78 6b 10 1c 1e   ·XI ·#··  ···xk···
        00d0  3b 59 55 3a d1 75 b3 0c  7c 35 96 85 85 f1 12 35   ;YU:·u··  |5·····5
        00e0  5c 79 ee f3 3b fc a7 86  1e cd 88 b6 77 f2 04 49   \y··;···  ····w··I
        00f0  97 19 03 34 07 e3 19 39  ba 82 37 ed 79 34 a8 15   ···4···9  ··7·y4··
```

"`" was unexpected in this context.   Packets: 1746 · Displayed: 1746 (100.0%)   Profile: Default

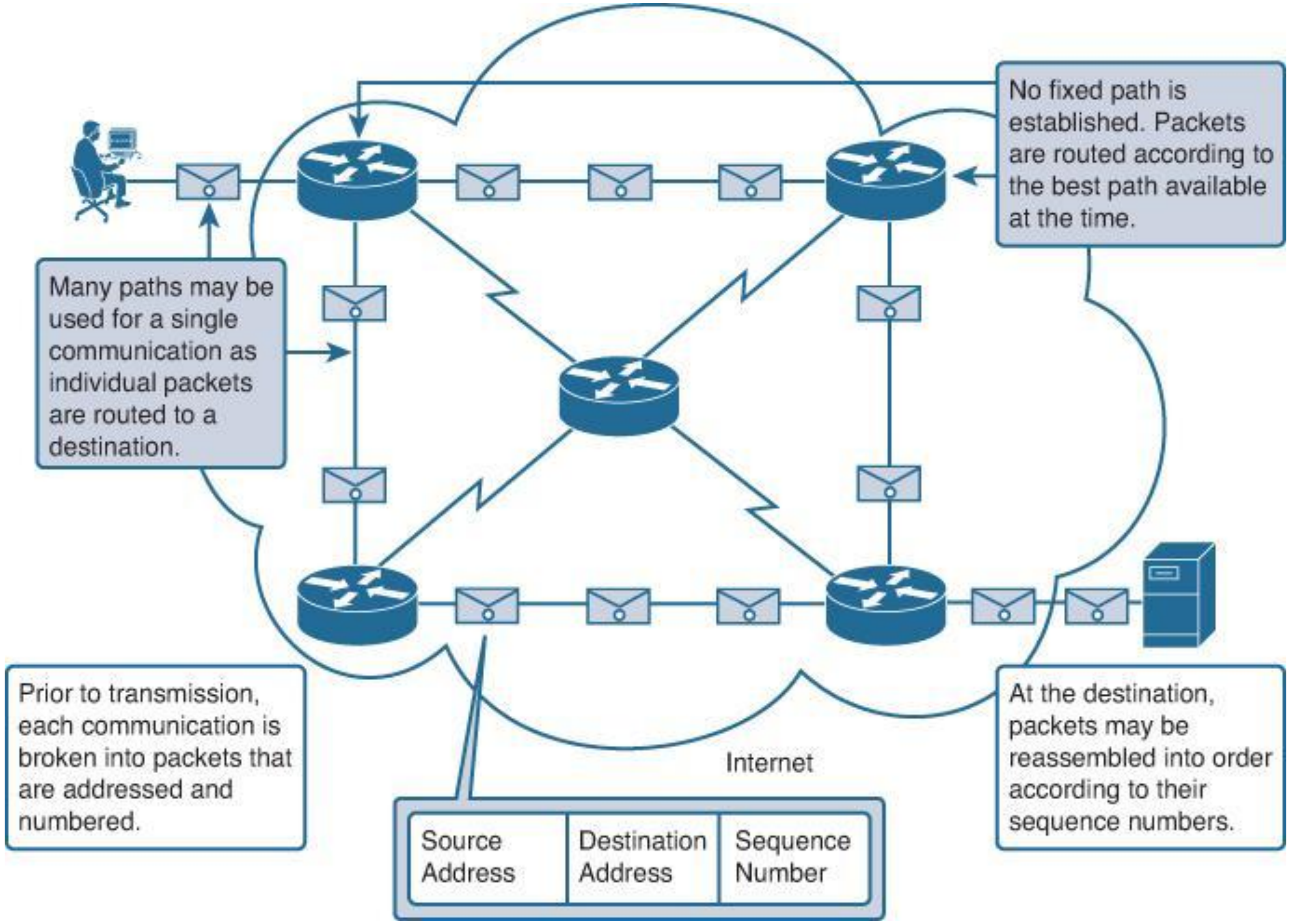
2022/10/17

Stateful Packet Inspection (SPI)

Firewall



Packet-Switched Network



No fixed path is established. Packets are routed according to the best path available at the time.

Many paths may be used for a single communication as individual packets are routed to a destination.

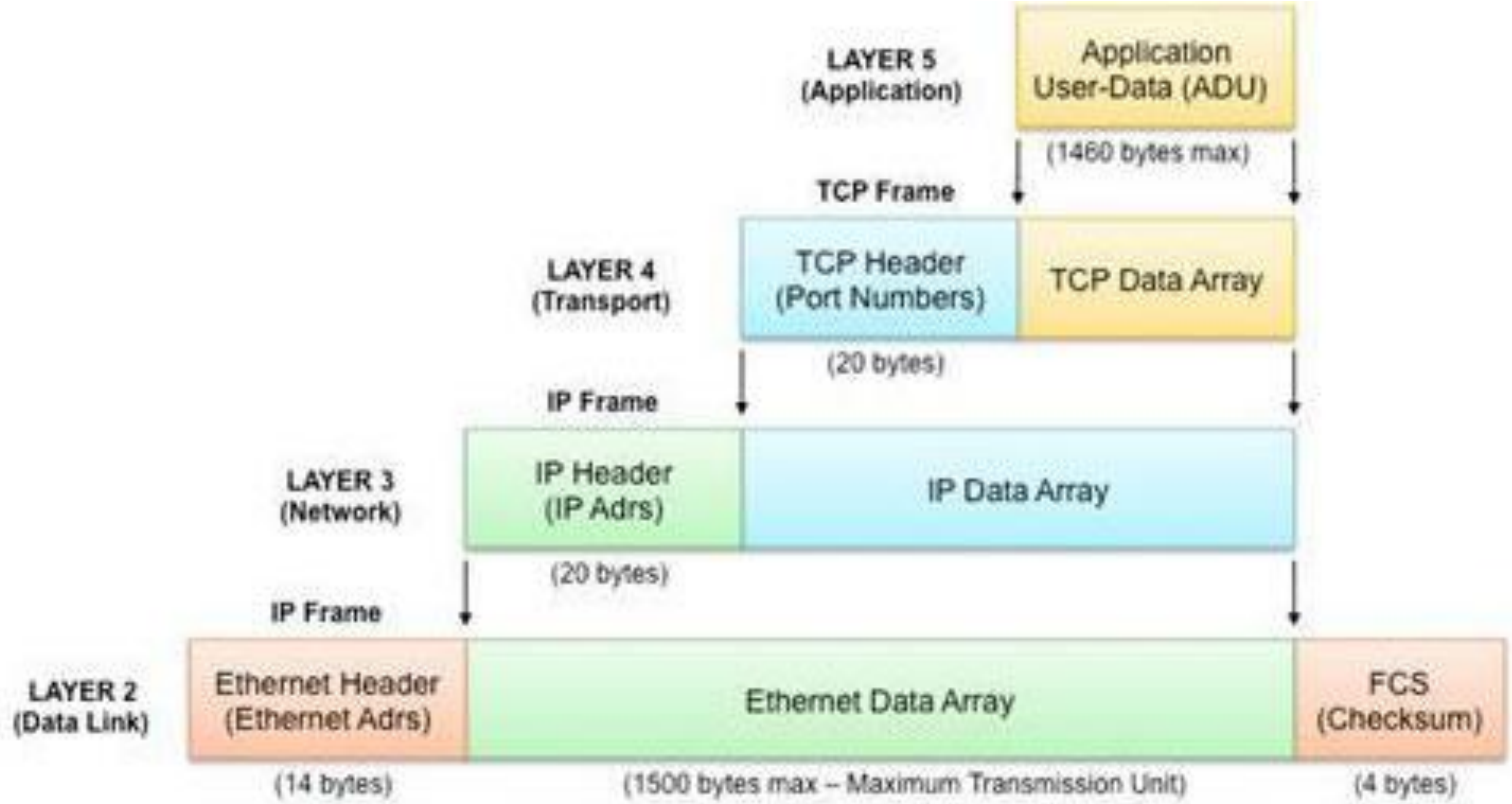
Prior to transmission, each communication is broken into packets that are addressed and numbered.

Source Address	Destination Address	Sequence Number

At the destination, packets may be reassembled into order according to their sequence numbers.

Data from: https://www.ciscopress.com/articles/article.asp?p=2158215&seqNum=5

TCP/IP Packets



LAYER 5 (Application)

Application User-Data (ADU)

(1460 bytes max)

TCP Frame

LAYER 4 (Transport)

TCP Header (Port Numbers)

TCP Data Array

(20 bytes)

IP Frame

LAYER 3 (Network)

IP Header (IP Adrs)

IP Data Array

(20 bytes)

IP Frame

LAYER 2 (Data Link)

Ethernet Header (Ethernet Adrs)

Ethernet Data Array

FCS (Checksum)

(14 bytes)

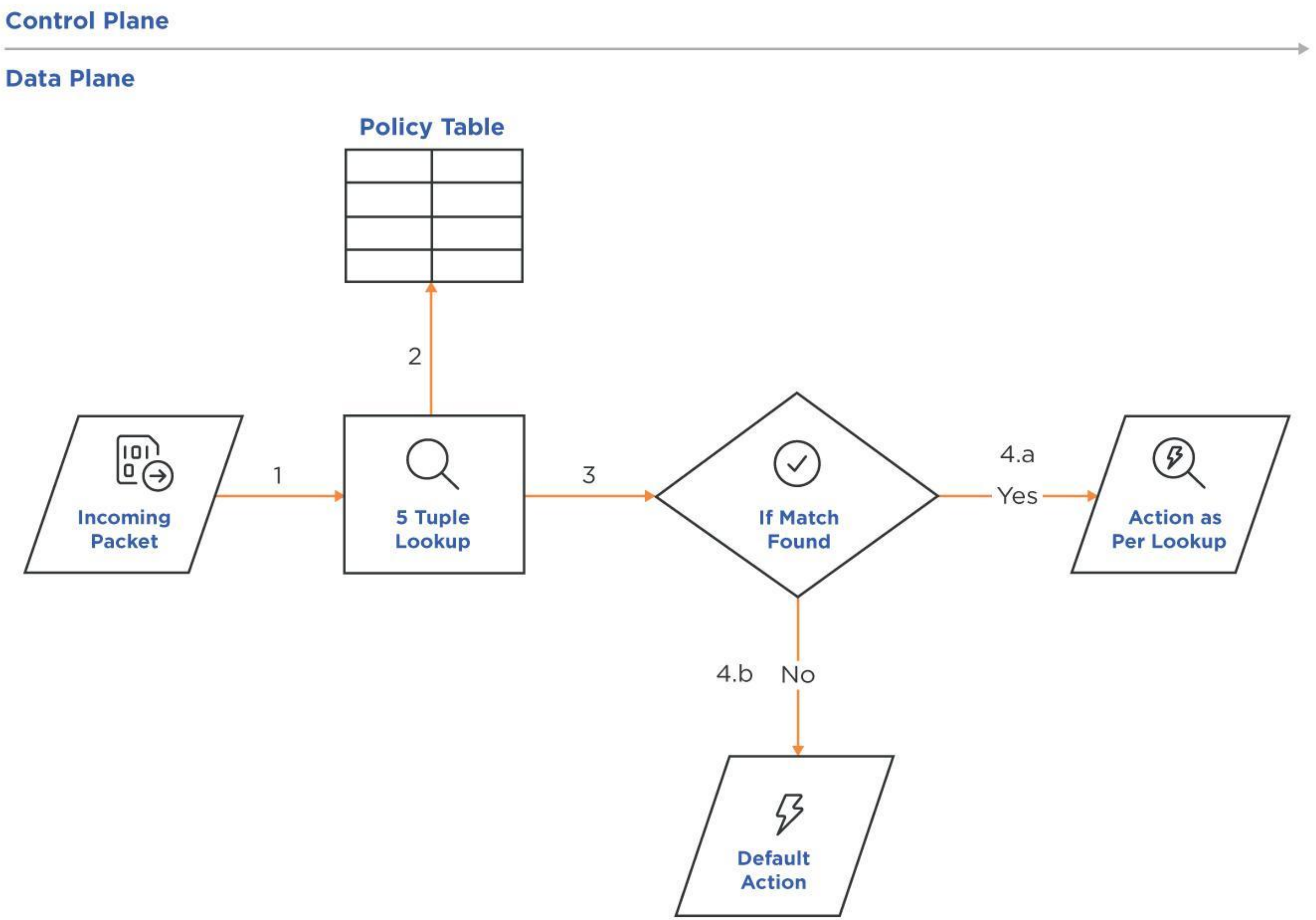
(1500 bytes max – Maximum Transmission Unit)

(4 bytes)

Connections

- Metadata
  - IP address and port of source and destination endpoints
  - Last packet received time for handling idle connections
  - Packet length
  - Layer 4 TCP sequence numbers and flags
  - Layer 3 data related to fragmentation and reassembly to identify session for the fragmented packet, etc.

Stateless Packet Inspection Firewall



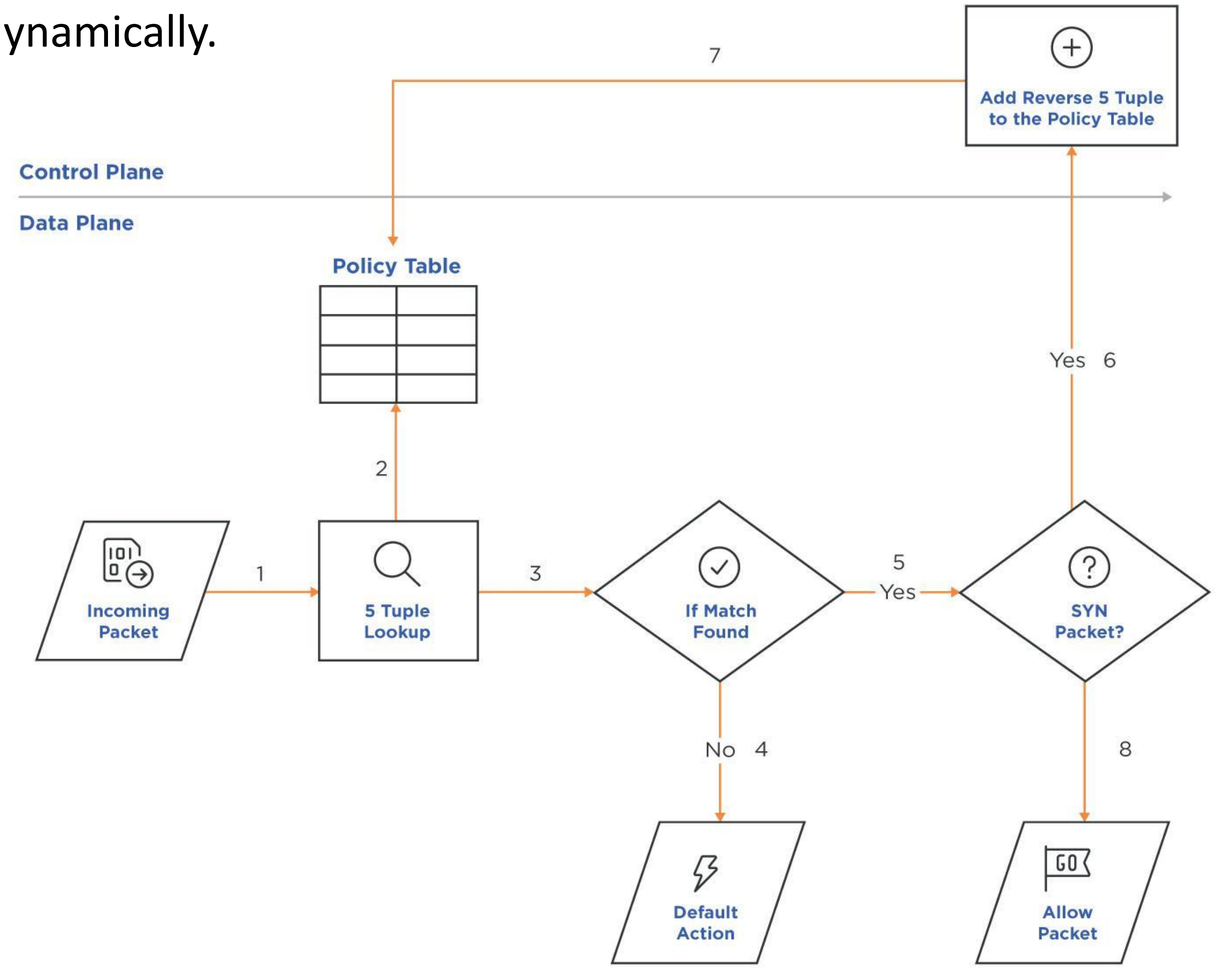
Data from: https://www.illumio.com/blog/firewall-stateful-inspection

Pros & Cons of Stateless Packet Inspection

- Pros
  - Less resource intensive
  - static packet data and policy table
    - the amount of CPU and memory resources required to do the lookup is low.
  - no increased latency
    - Additional processing adds no-to-minimal overhead latency
- Cons
  - Limited filtering
    - using low fidelity data from the firewall which provides limited filtering capability.
  - ACL configuration:
    - Hard to configure and manage large ACLs.

Reflexive firewall

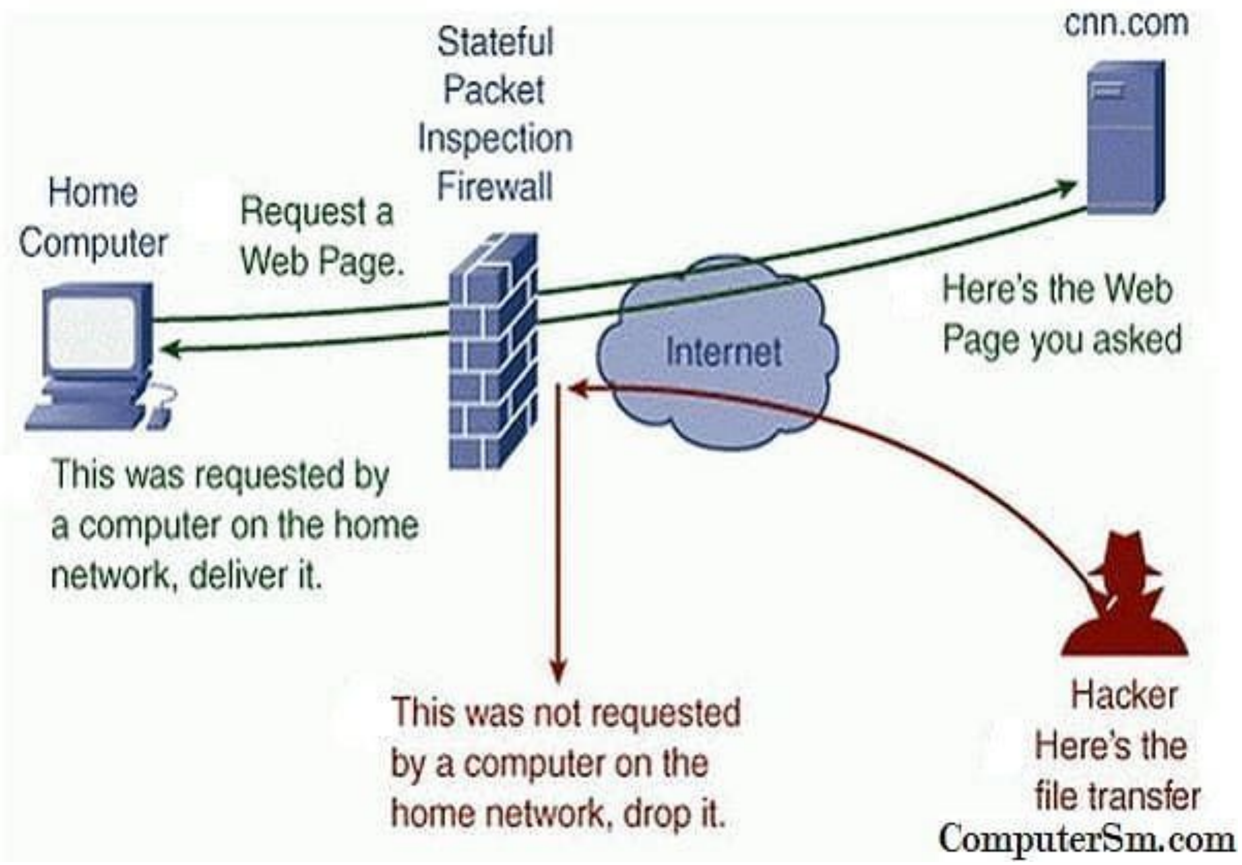
- whitelist return traffic dynamically.



Data from: https://www.illumio.com/blog/firewall-stateful-inspection

Stateful Packet Inspection firewall

- protects devices by checking incoming packets against existing connections.

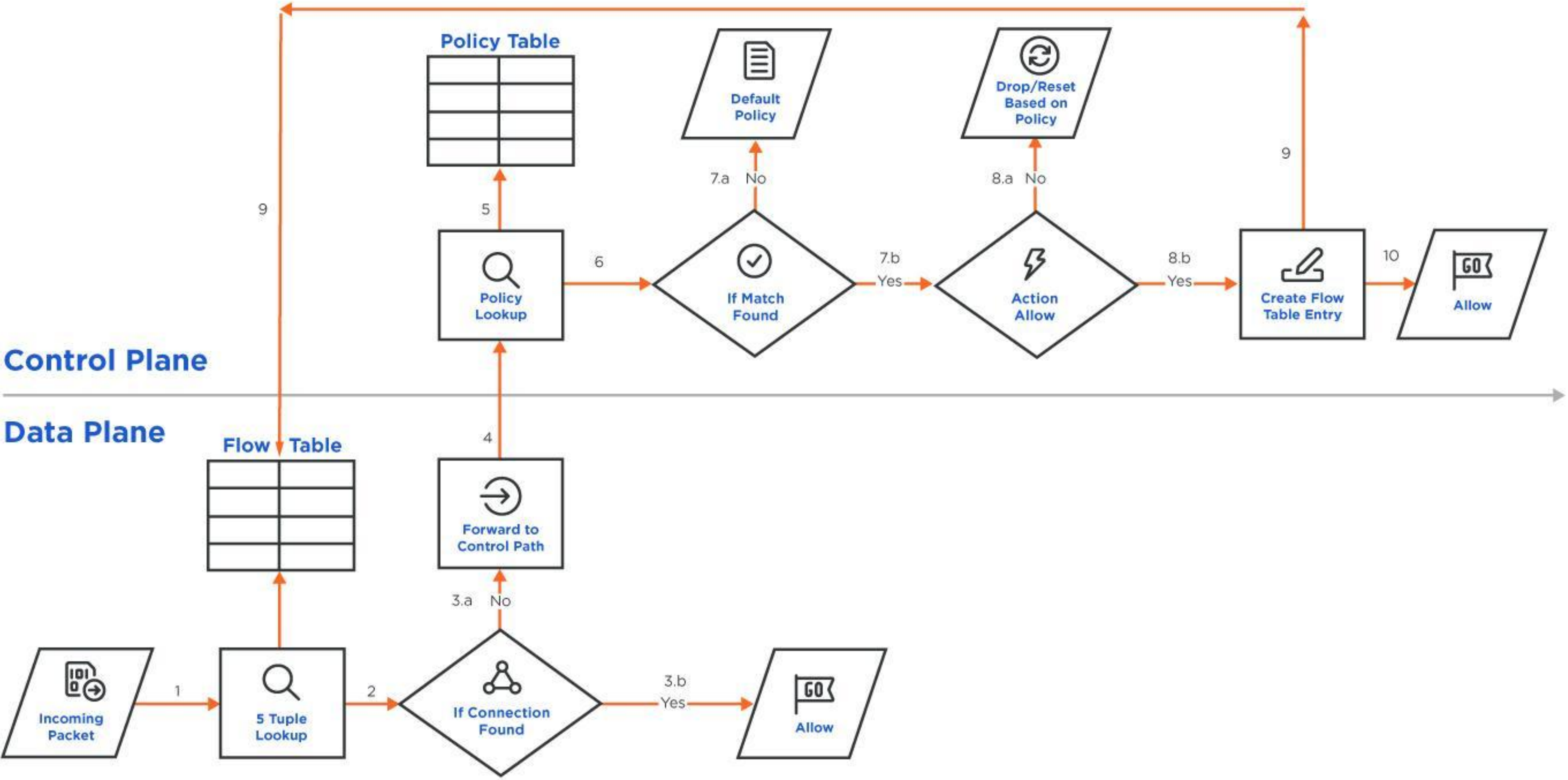


Data from: https://www.cyberangels.org/what-does-spi-stand-for-in-cyber-security/

Stateful Packet Inspection Firewall

- 5-tuple lookup
  - source IP, source port, destination IP, destination port, protocol) in a flow table to find a match
- Fast path / data plane processing
  - layer 3 IP sanitation check to avoid fragmentation & reassembly based attack
  - layer 4 check to prevent attacks like spoofing, DOS, etc.
  - layer 7: Application Layer Gateways (ALGs)
- Slow path / control panel processing:
  - new connections
  - needs additional policy checks
- Policy lookup:
  - using the STATE + CONTEXT of the connection.
  - ALLOW, DENY or RESET.

Stateful Packet Inspection firewall



Data from: https://www.illumio.com/blog/firewall-stateful-inspection

Pros and Cons of Stateful Packet Inspection

- Pros
  - Higher protection.
  - More advanced.
  - Configuring capability of network flow.
  - Complex protocols like FTP, P2P protocols, etc.
- Cons
  - Processing power
    - do additional checks to provide more security.

Deep Packet Inspection

What’s Deep Packet Inspection?

Stateful Packet Inspection

Stateful packet inspection looks at the header and footer of a packet.

Deep Packet Inspection

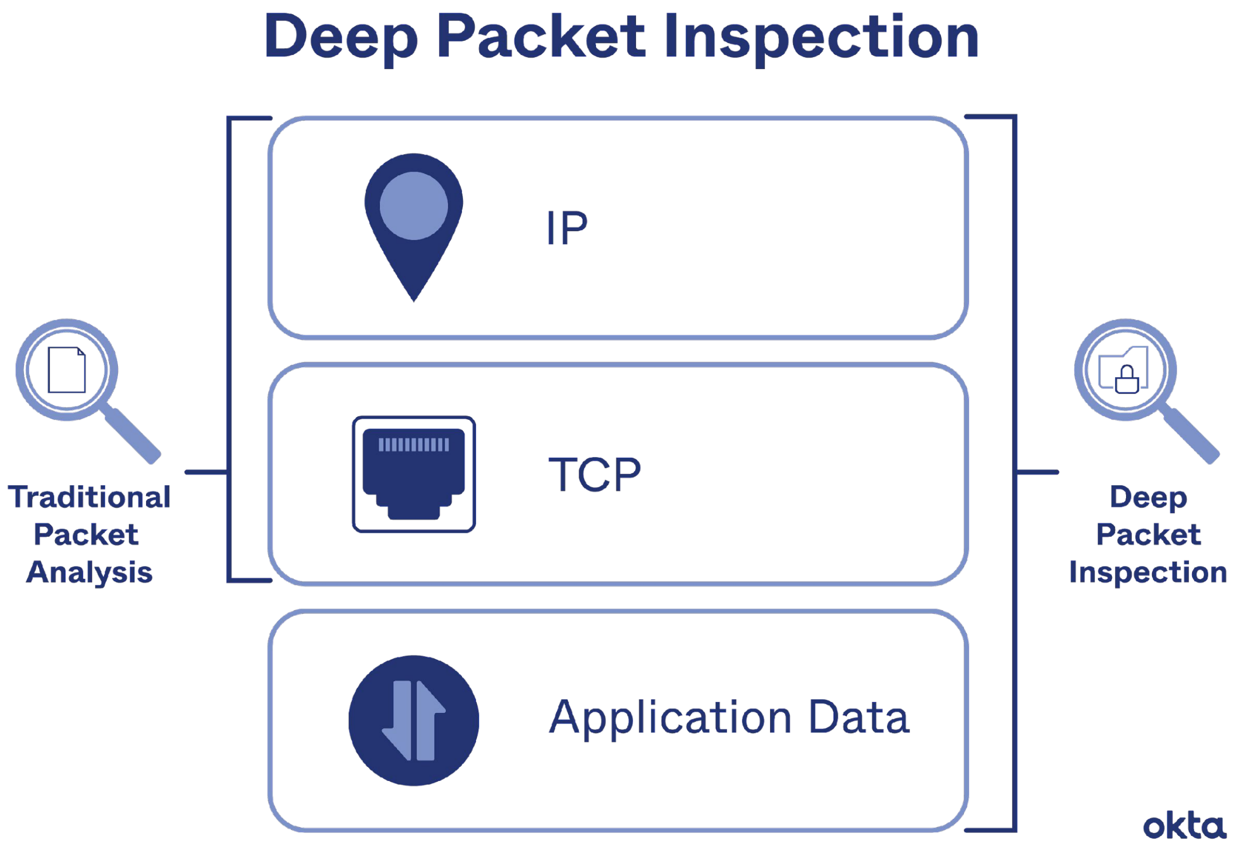
Deep packet inspection examines the data part of a packet.

Data from: https://devopedia.org/deep-packet-inspection

Deep packet inspection versus conventional packet filtering

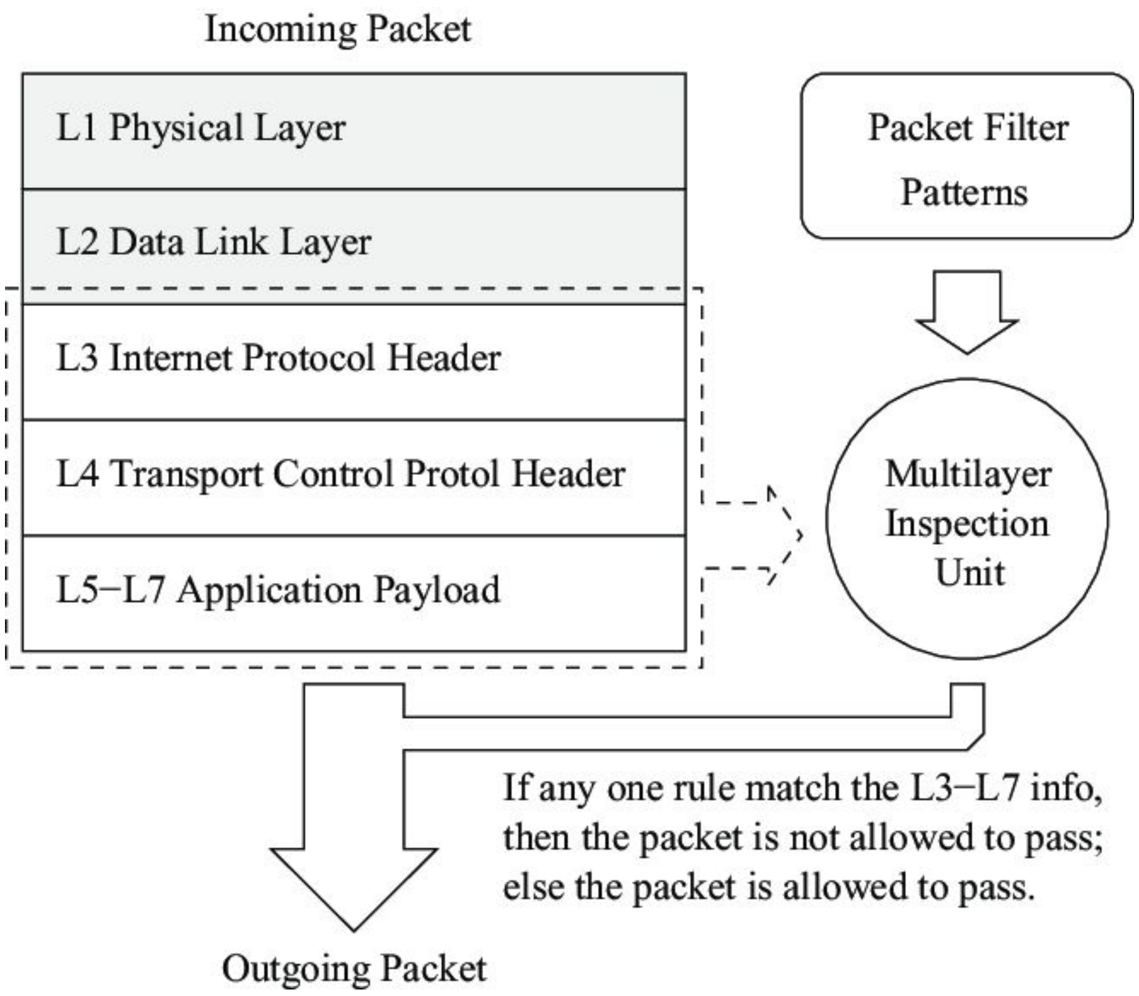
- Conventional packet filtering
    - only reads the header information of each packet
    - similar to reading the title of a book, without awareness or evaluation of the content inside the cover
    - Firewalls had very little processing power, and it was not enough to handle large volumes of packets
- Deep packet inspection
    - picking up a book, cracking it open, and reading it from cover to cover.

What’s Deep Packet Inspection?



Data from: https://www.okta.com/identity-101/deep-packet-inspection/

Incoming Packet filtering in DPI



If any one rule match the L3–L7 info,
then the packet is not allowed to pass;
else the packet is allowed to pass.

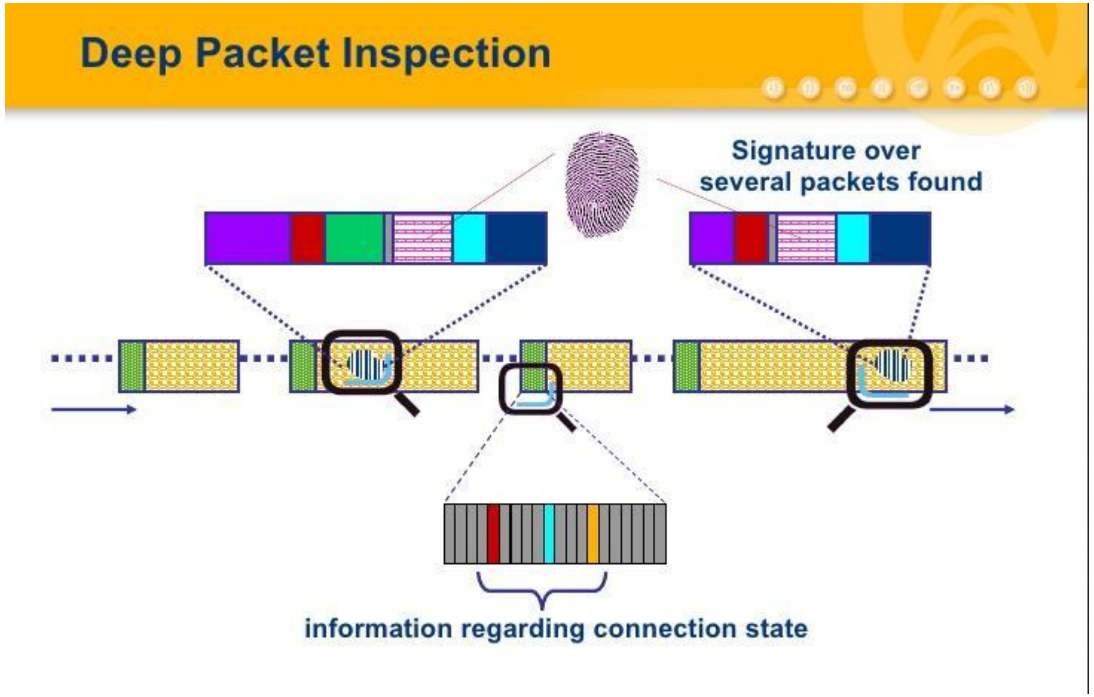
Data from: https://devopedia.org/images/article/262/6575.1587059062.jpg

How does DPI works?

- Flow tracking
  - 5-tuple identifier (SRC-IP, DEST-IP, SRC-PORT, DEST-PORT, PROTOCOL).
- Pattern matching
  - matching applications/protocols to their most common/standard ports.
    - ex. BitTorrent uses the TCP ports 6881-6889 by default.
  - some applications can ride on standard ports fooling detection systems.
    - Skype used TCP port 80/443 when its normal ports are blocked.

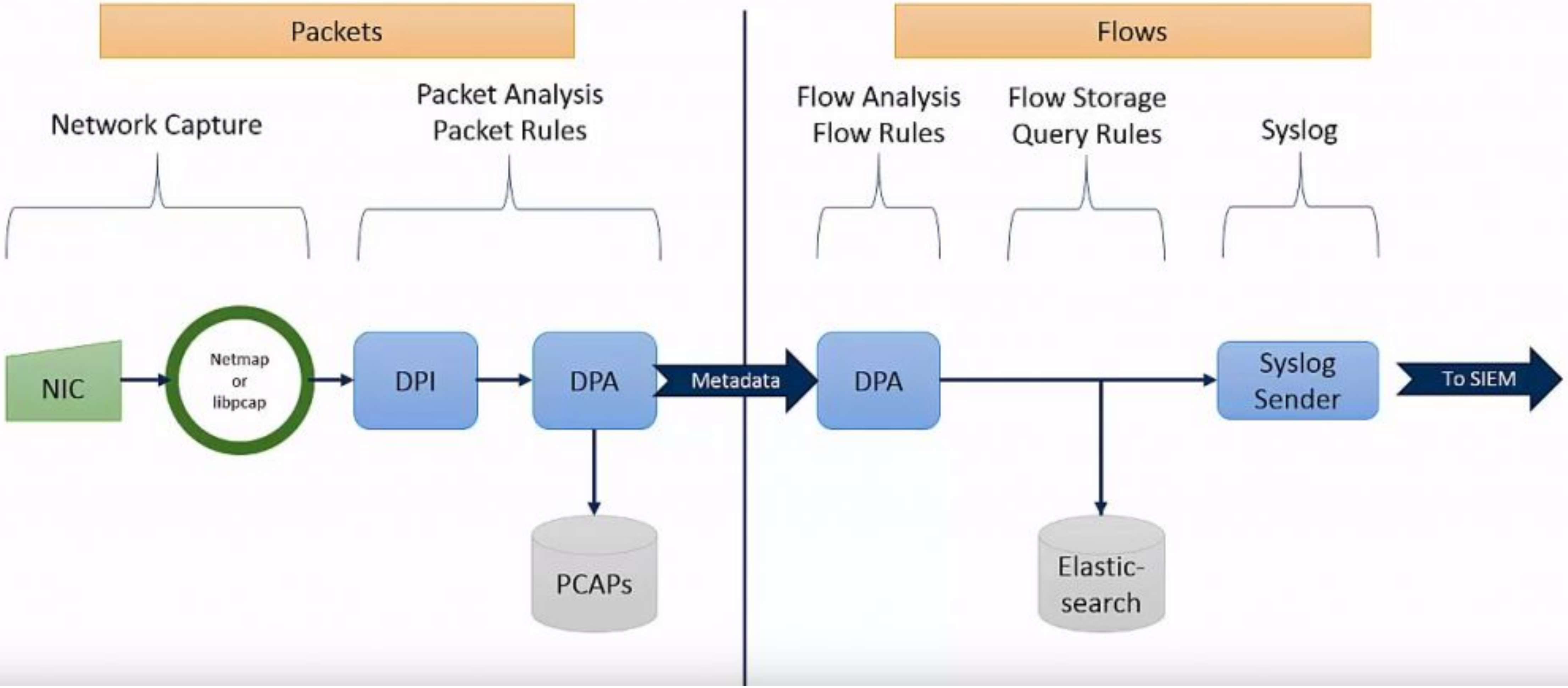
How does DPI works?

- **Signature Matching**
  - still some strings or patterns that may be recognizable in such applications.
    - old Skype begins with “80 46 01 03 …”.
  - applications are constantly updated new signatures.


Deep Packet Inspection
Signature over several packets found
information regarding connection state

Data from: https://devopedia.org/deep-packet-inspection

Life of Packets and Flows

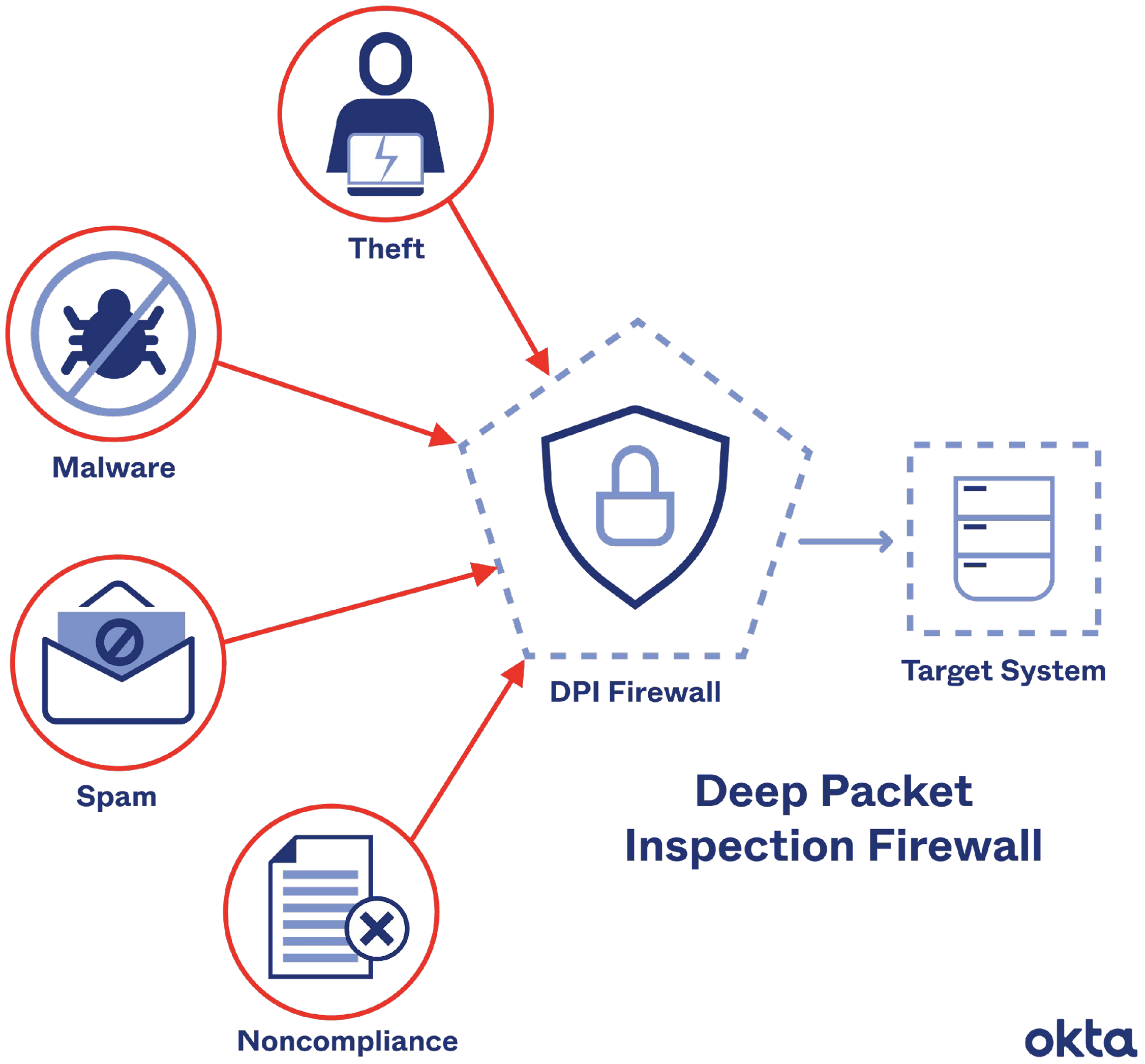


How does DPI works?

- **Heuristic and Behavior Analysis**
  - measuring packet sizes, flow rate per application.
    - Voice over IP (VoIP) starts with session initiation and then many small-sized UDP packets.
  - newer forms of detection are being developed especially those that rely on Machine Learning (ML) and Artificial Intelligence (AI).
- a mixture of these techniques can improve detection and increase accuracy
- Deny by default
  - restricting traffic to only what is necessary.
- System defaults
  - present DPI network rules.

Why DPI?

- Network and Endpoint Security
  - identify malicious traffic
  - prevent attacks caused by viruses, worms, ransomware, and so on.
  - similar to how antivirus programs work on end devices.
- Data Loss Prevention (DLP)
  - prevent sensitive information from leaving a company's network.



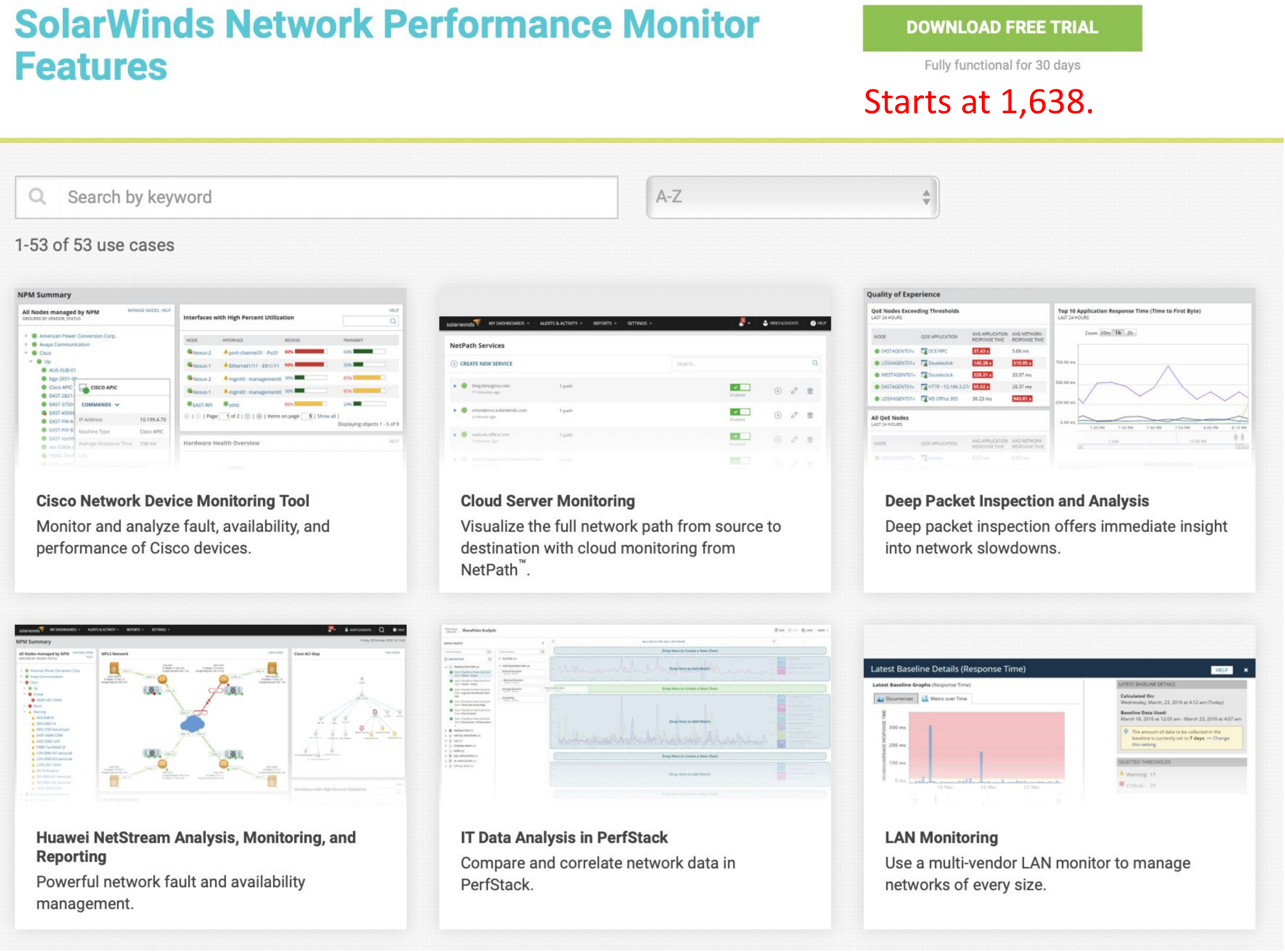
Data from: https://www.okta.com/identity-101/deep-packet-inspection/

Possible misuses of DPI

- QoS/Traffic Shaping
  - ISPs are able to “snoop” into the contents of the traffic flowing
  - ISP may perform traffic shaping
    - limit user’s download rate with large files.
- Behavioral targeting (BT)
  - harvesting user information anonymously (supposedly)
  - create ads that are targeted to the individual.

What are common uses and applications of deep packet inspection?

- SolarWinds Network Performance Monitor
- Paessler Packet Sniffing with PRTG
- oPManager
- nDPI
- Netifyd
- AppNeta
- NetFort LANGuardian



SolarWinds Network Performance Monitor Features

DOWNLOAD FREE TRIAL

Fully functional for 30 days

Starts at 1,638.

Search by keyword

A-Z

1-53 of 53 use cases

Cisco Network Device Monitoring Tool

Monitor and analyze fault, availability, and performance of Cisco devices.

Cloud Server Monitoring

Visualize the full network path from source to destination with cloud monitoring from NetPath™.

Deep Packet Inspection and Analysis

Deep packet inspection offers immediate insight into network slowdowns.

Huawei NetStream Analysis, Monitoring, and Reporting

Powerful network fault and availability management.

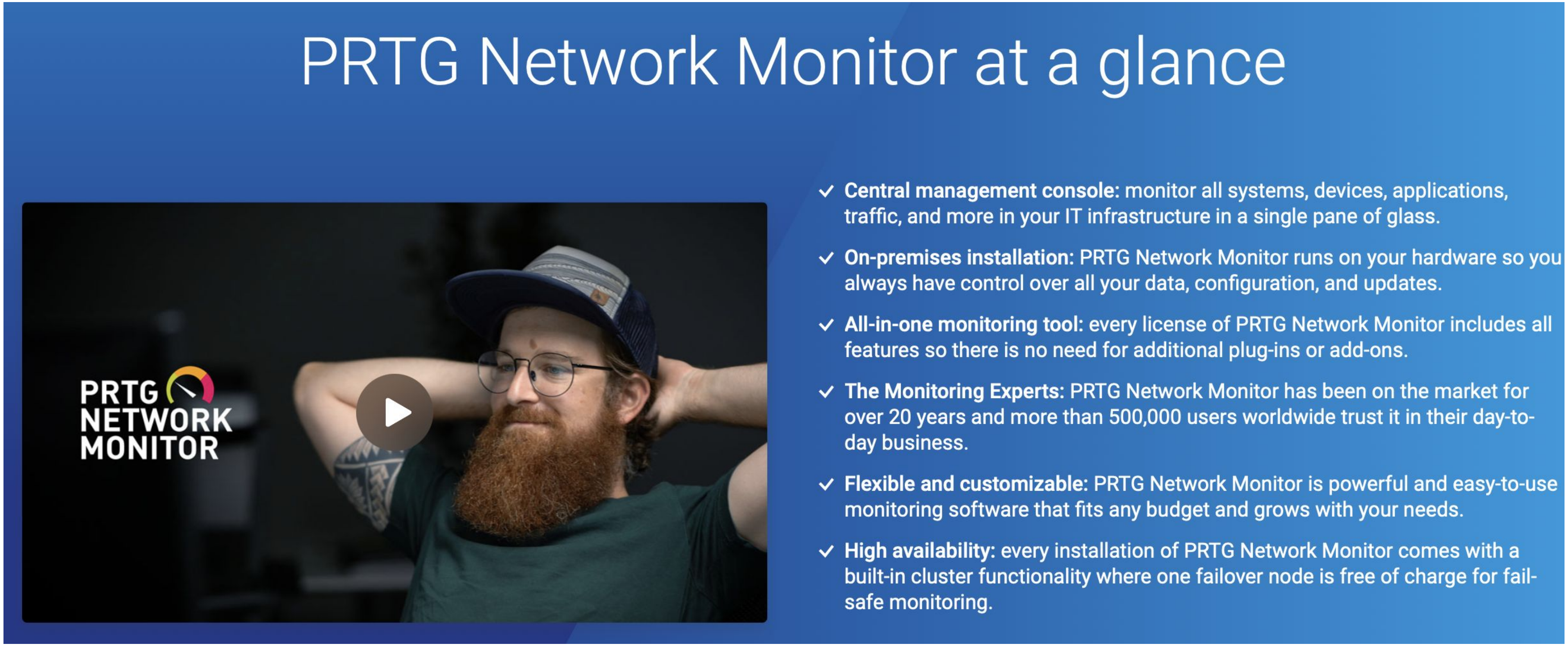
IT Data Analysis in PerfStack

Compare and correlate network data in PerfStack.

LAN Monitoring

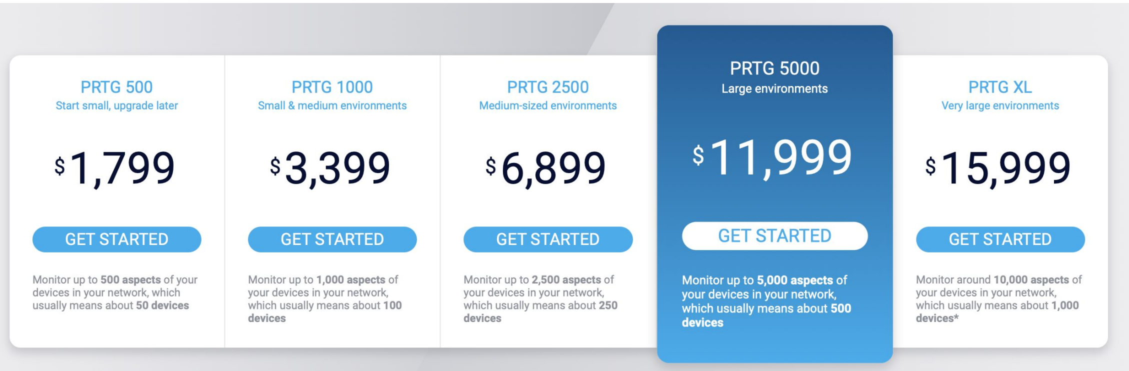
Use a multi-vendor LAN monitor to manage networks of every size.

Data from: https://www.solarwinds.com/network-performance-monitor



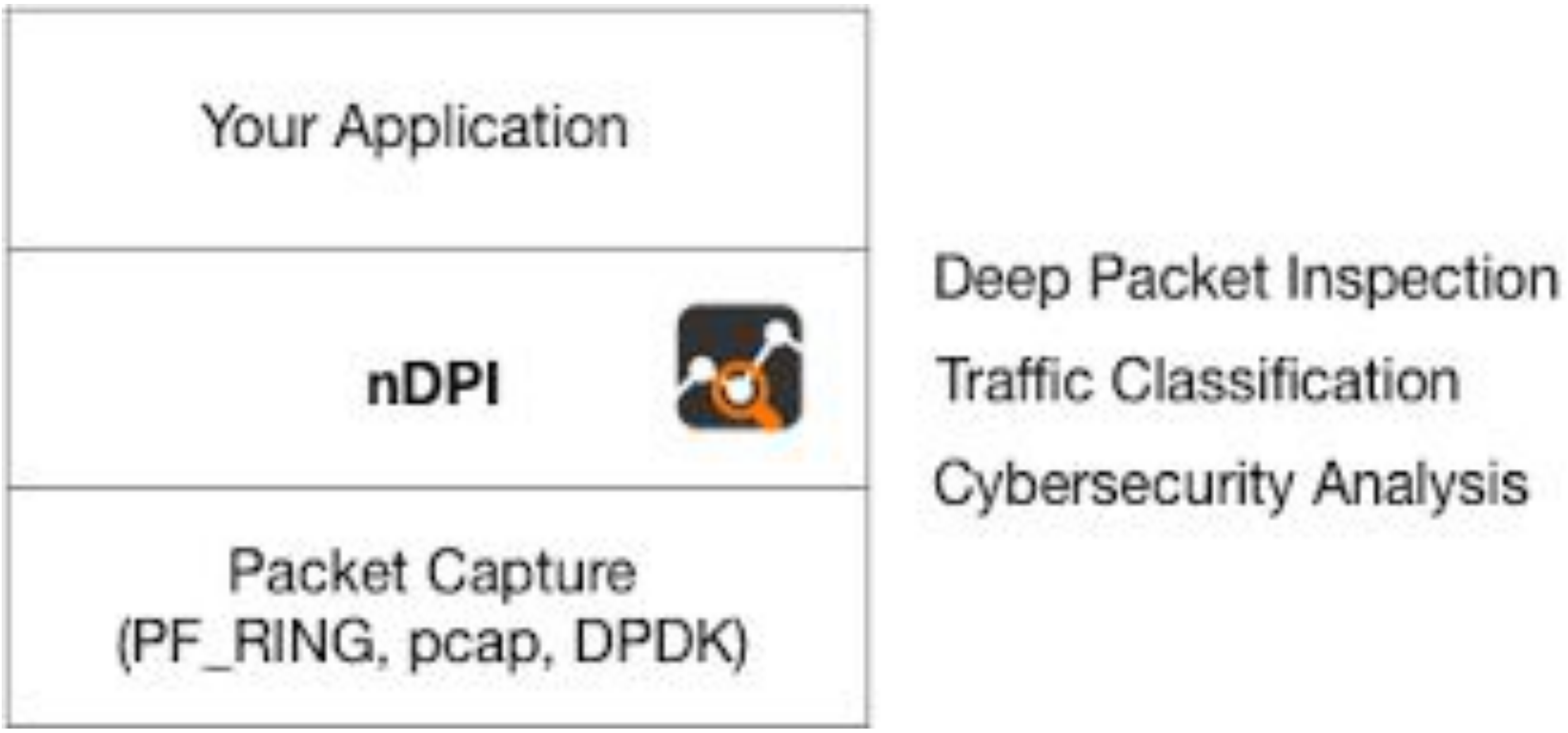
PRTG Network Monitor at a glance

- ✓ **Central management console:** monitor all systems, devices, applications, traffic, and more in your IT infrastructure in a single pane of glass.
- ✓ **On-premises installation:** PRTG Network Monitor runs on your hardware so you always have control over all your data, configuration, and updates.
- ✓ **All-in-one monitoring tool:** every license of PRTG Network Monitor includes all features so there is no need for additional plug-ins or add-ons.
- ✓ **The Monitoring Experts:** PRTG Network Monitor has been on the market for over 20 years and more than 500,000 users worldwide trust it in their day-to-day business.
- ✓ **Flexible and customizable:** PRTG Network Monitor is powerful and easy-to-use monitoring software that fits any budget and grows with your needs.
- ✓ **High availability:** every installation of PRTG Network Monitor comes with a built-in cluster functionality where one failover node is free of charge for fail-safe monitoring.



PRTG 500	PRTG 1000	PRTG 2500	PRTG 5000	PRTG XL
Start small, upgrade later	Small & medium environments	Medium-sized environments	Large environments	Very large environments
$1,799	$3,399	$6,899	$11,999	$15,999
GET STARTED	GET STARTED	GET STARTED	GET STARTED	GET STARTED
Monitor up to **500 aspects** of your devices in your network, which usually means about **50 devices**	Monitor up to **1,000 aspects** of your devices in your network, which usually means about **100 devices**	Monitor up to **2,500 aspects** of your devices in your network, which usually means about **250 devices**	Monitor up to **5,000 aspects** of your devices in your network, which usually means about **500 devices**	Monitor around **10,000 aspects** of your devices in your network, which usually means about **1,000 devices***

nDPI – https://github.com/ntop/nDPI



Your Application

nDPI

Packet Capture
(PF_RING, pcap, DPDK)

Deep Packet Inspection

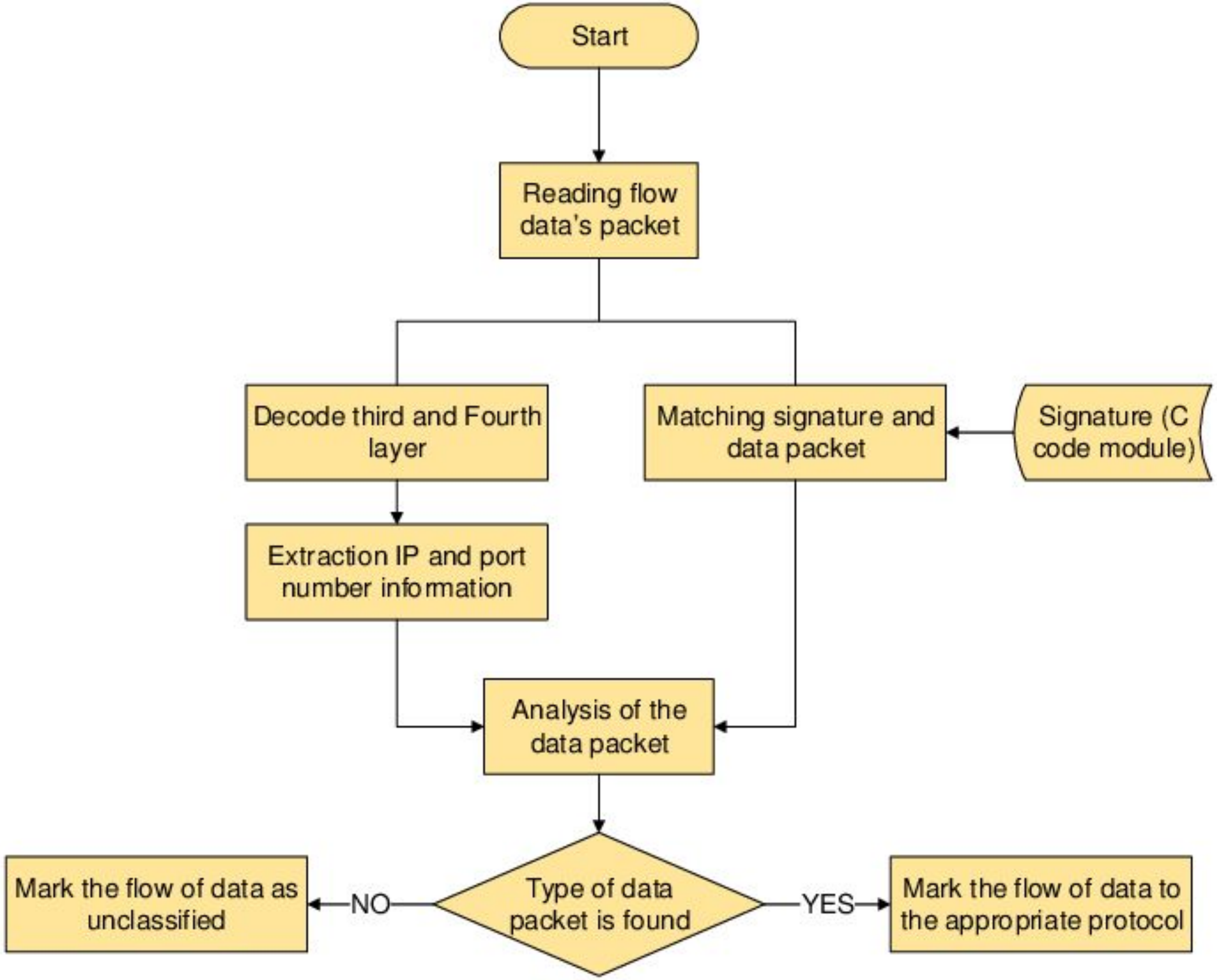
Traffic Classification

Cybersecurity Analysis

nDPI

- Data Forecasting and Anomaly Detection
  - Single, Double, Triple (Holt-Winters) Exponential Smoothing
  - RSI (Relative Strength Index)
  - Data Binning, Clustering, and Similarity Evaluation
- Network Data Analysis
  - Jitter
  - Entropy
  - GeoIP
  - Data Ratio (also known as PCR)
  - Rolling Average, Standard Deviation, Variance (all implemented as streamed versions)
- IP Address Retrieval
  - Radix (Patricia) Tree (trie)
- Cardinality Estimation
  - HyperLogLog
- (Sub-)String Searching
  - Aho-Corasick

nDPI


Start

Reading flow data's packet

Decode third and Fourth layer

Extraction IP and port number information

Matching signature and data packet

Signature (C code module)

Analysis of the data packet

Type of data packet is found

NO

YES

Mark the flow of data as unclassified

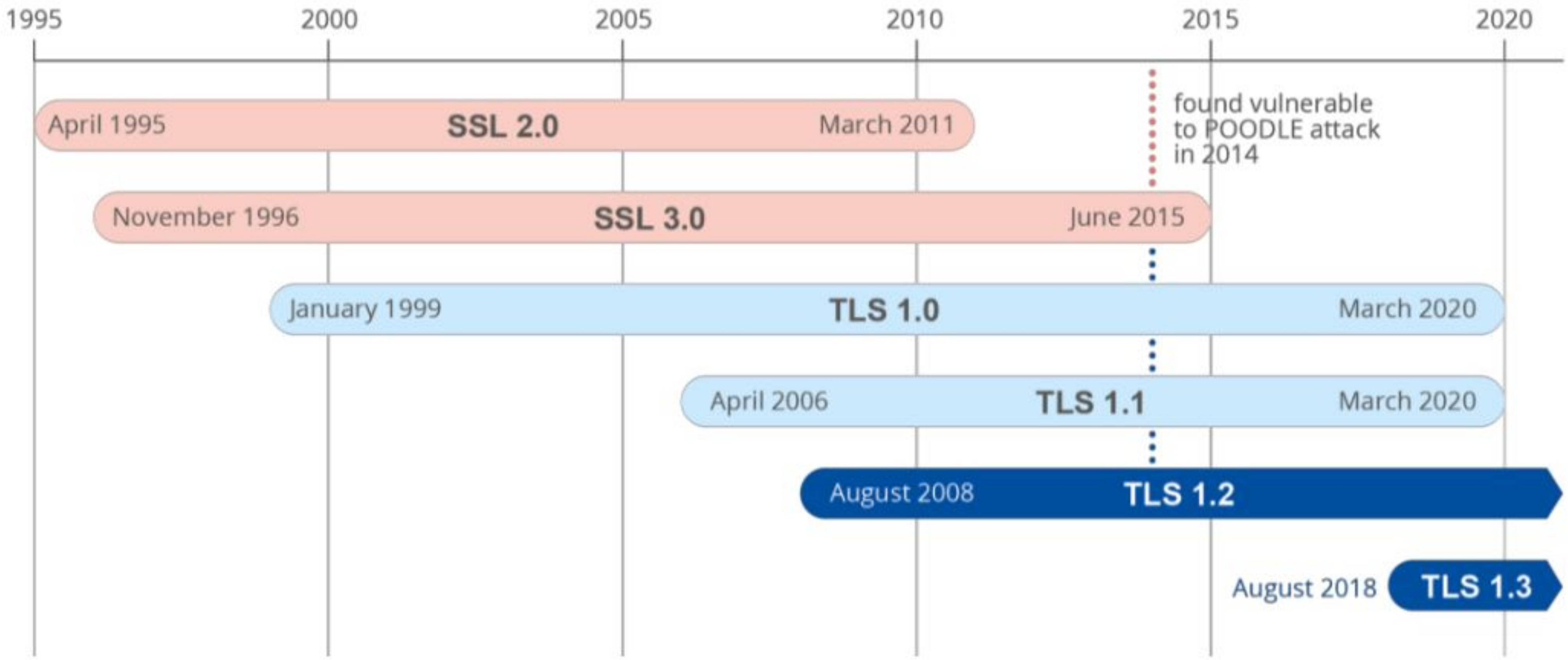
Mark the flow of data to the appropriate protocol

http://dx.doi.org/10.21108/IJOICT.2016.22.77

Encrypted Traffic Analysis

Data from: ENISA Report – Encrypted Traffic Analysis

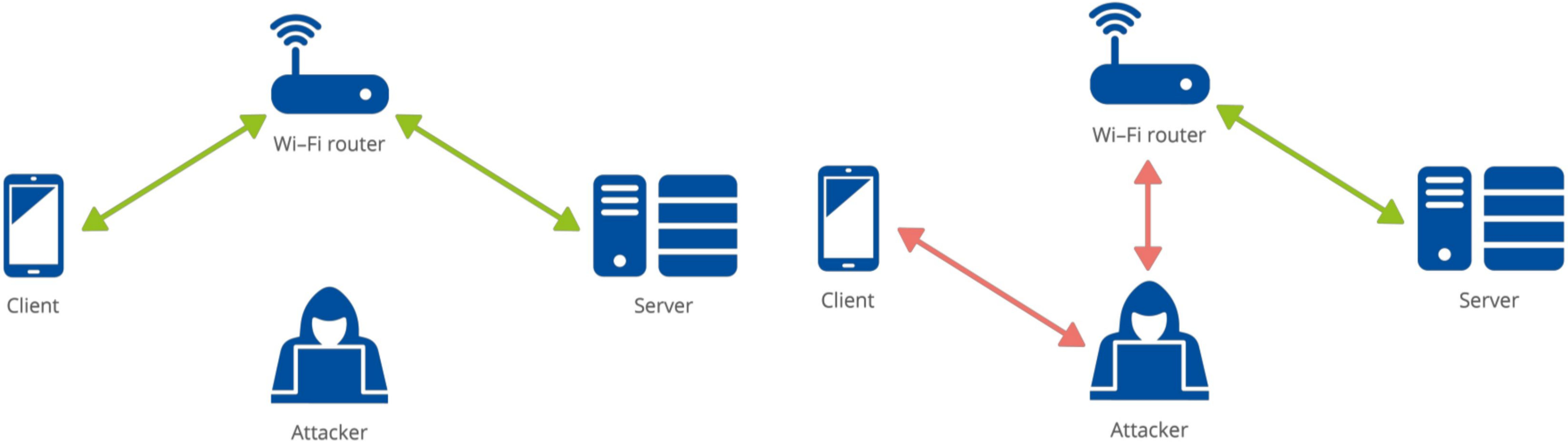
Timeline of SSL/TLS Protocols



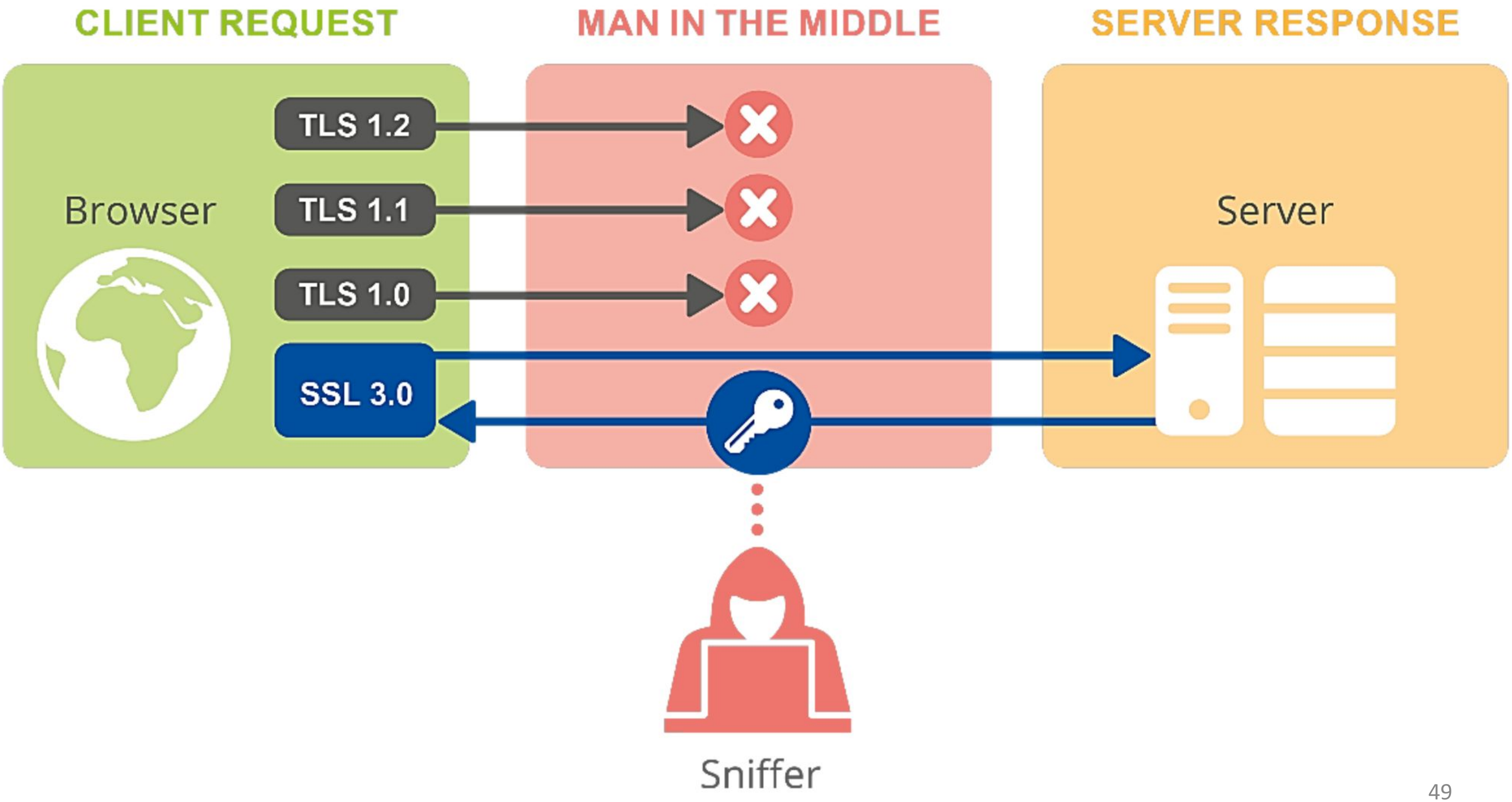
Attack TLS

- Lack of Certificate Validation
  - Self-signed certificate
  - Expired certificate
- Man-in-the-middle attack on a TLS connection
  - HTTP redirects
- Weak ciphers and Deprecated Protocols

Man-in-the-Middle Attack

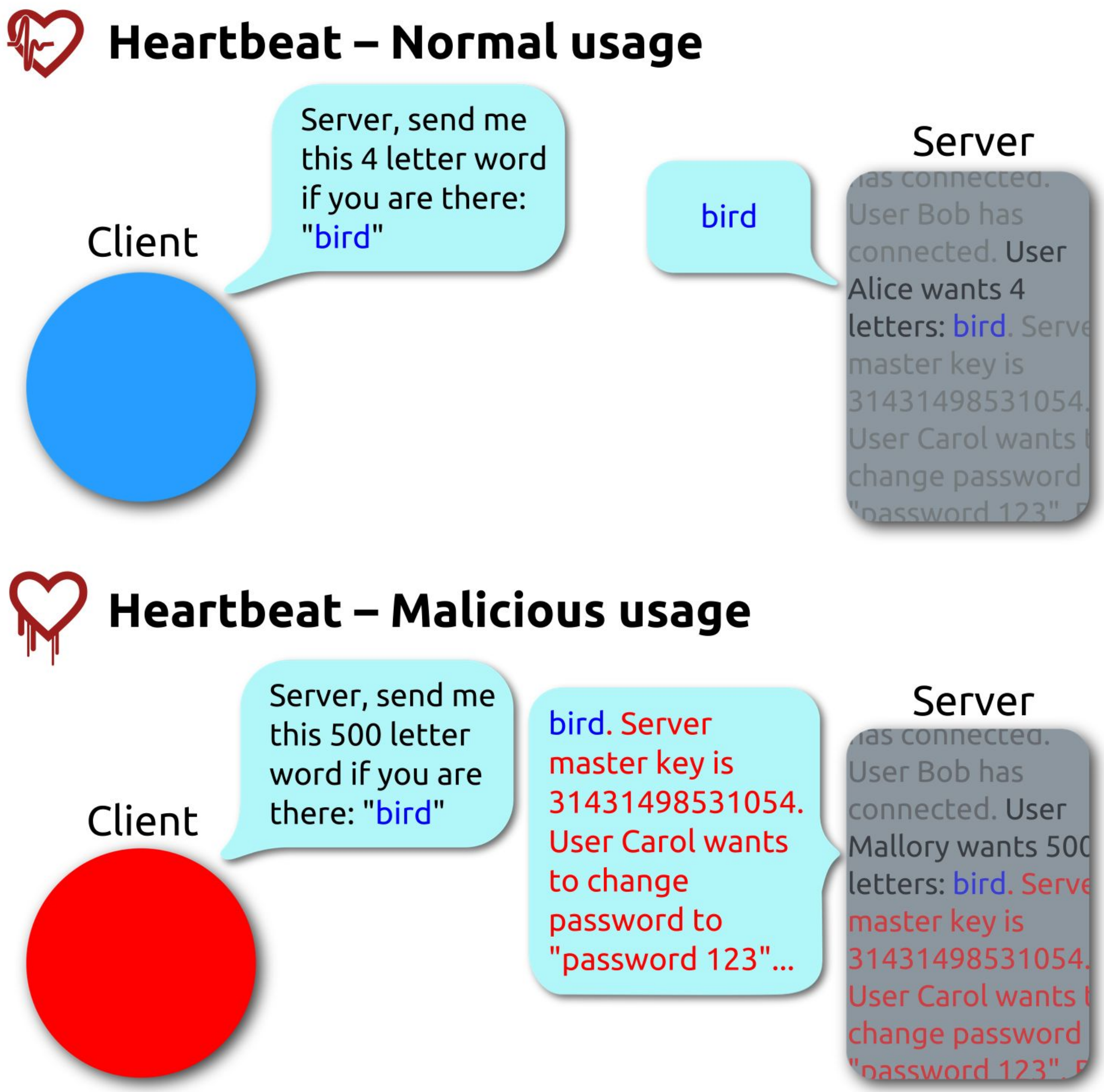


Protocol Downgrade



HeartBleed

- a serious bug in the OpenSSL library
- allows an attacker to decrypt the content that is encrypted using TLS.



Heartbeat – Normal usage

Client

Server, send me this 4 letter word if you are there: "bird"

bird

Server

has connected. User Bob has connected. User Alice wants 4 letters: bird. Serve master key is 31431498531054. User Carol wants t change password "password 123". P

Heartbeat – Malicious usage

Client

Server, send me this 500 letter word if you are there: "bird"

bird. Server master key is 31431498531054. User Carol wants to change password to "password 123"...

Server

has connected. User Bob has connected. User Mallory wants 500 letters: bird. Serve master key is 31431498531054. User Carol wants t change password "password 123". P

Encrypted Traffic Analysis properties

1. Goals.
   - Traffic Clustering, Application Type and Protocol Classification, Anomaly Detection or File Identification.
2. Information extraction.
   - observing behavioral properties (e.g. the round trip time, number of packets sent)
   - observing the encrypted payload itself
   - observing additional information such as protocol handshakes (e.g. TLS handshake)
3. Information processing.
   - Basic
     - by using heuristics, profiles or simple statistical means
   - Complex
     - data-driven / machine-learning

Feature Extraction

INTER-ARRIVAL-TIME

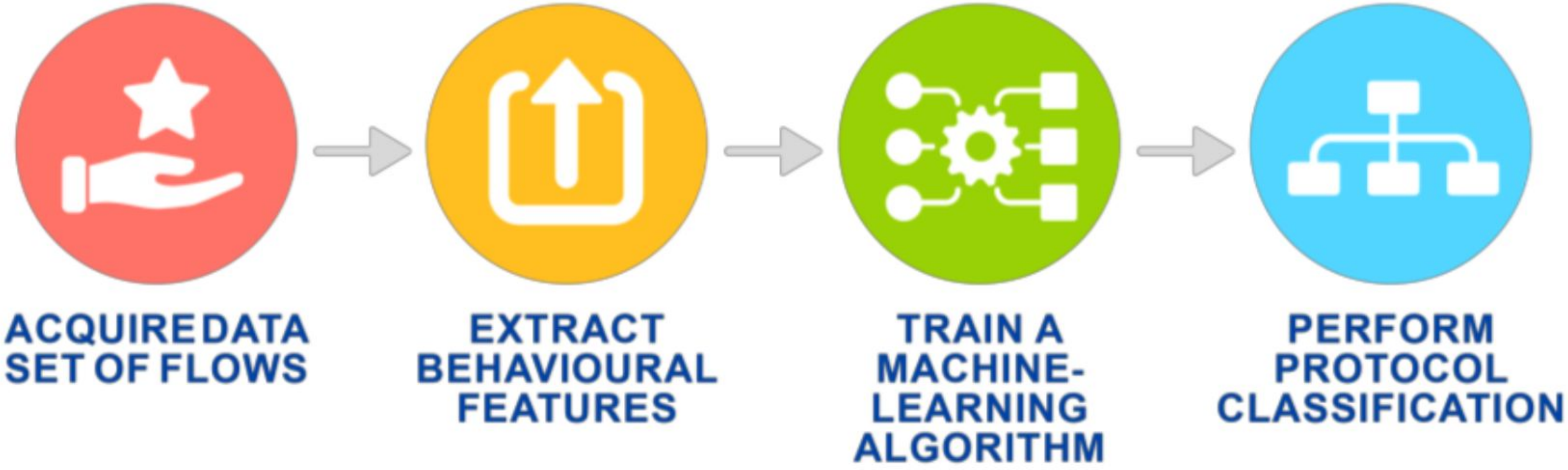
PACKET-LENGTH

NUMBER OF ACK PACKETSOBSERVED

NUMBER OF RETRANSMISSIONS

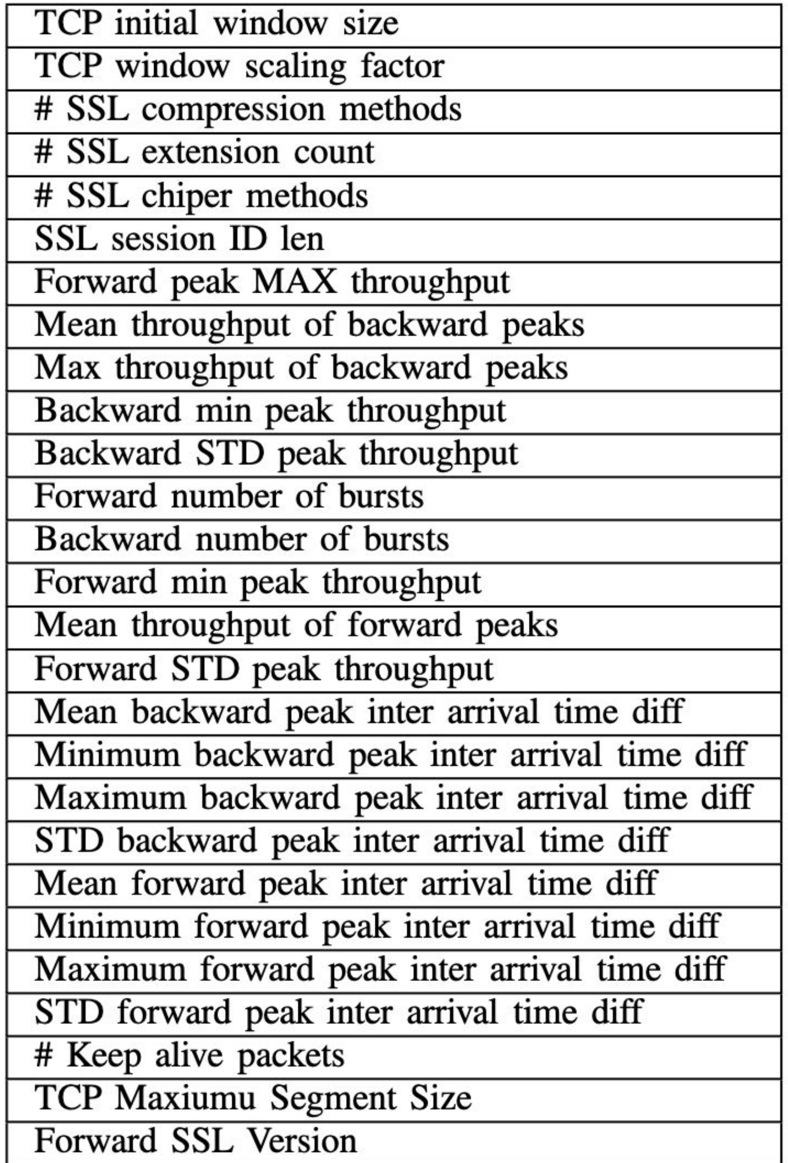
ROUND TRIP TIME

Use CASE: Application Protocol Classification


ACQUIRE DATA SET OF FLOWS
EXTRACT BEHAVIOURAL FEATURES
TRAIN A MACHINE-LEARNING ALGORITHM
PERFORM PROTOCOL CLASSIFICATION

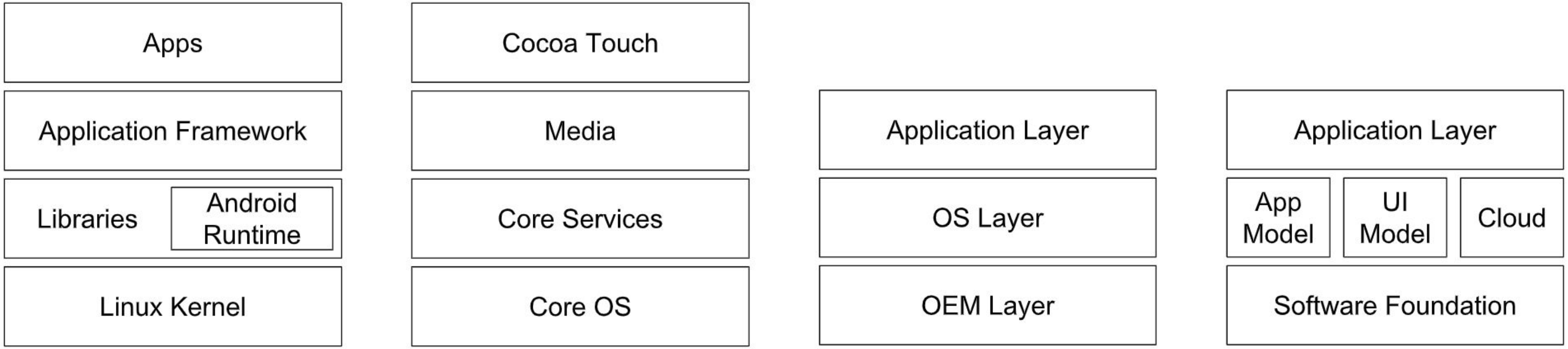
USE CASE: User Information Identification

- Detect OS/Browser/Application
  - https://arxiv.org/vc/arxiv/papers/1603/1603.04865v1.pdf



TCP initial window size
TCP window scaling factor
SSL compression methods
SSL extension count
SSL chiper methods
SSL session ID len
Forward peak MAX throughput
Mean throughput of backward peaks
Max throughput of backward peaks
Backward min peak throughput
Backward STD peak throughput
Forward number of bursts
Backward number of bursts
Forward min peak throughput
Mean throughput of forward peaks
Forward STD peak throughput
Mean backward peak inter arrival time diff
Minimum backward peak inter arrival time diff
Maximum backward peak inter arrival time diff
STD backward peak inter arrival time diff
Mean forward peak inter arrival time diff
Minimum forward peak inter arrival time diff
Maximum forward peak inter arrival time diff
STD forward peak inter arrival time diff
Keep alive packets
TCP Maxiumu Segment Size
Forward SSL Version

System Architecture of mobile operating system



Apps

Application Framework

Libraries | Android Runtime

Linux Kernel

(a) Android.

Cocoa Touch

Media

Core Services

Core OS

(b) iOS.

Application Layer

OS Layer

OEM Layer

(c) Windows Mobile.

Application Layer

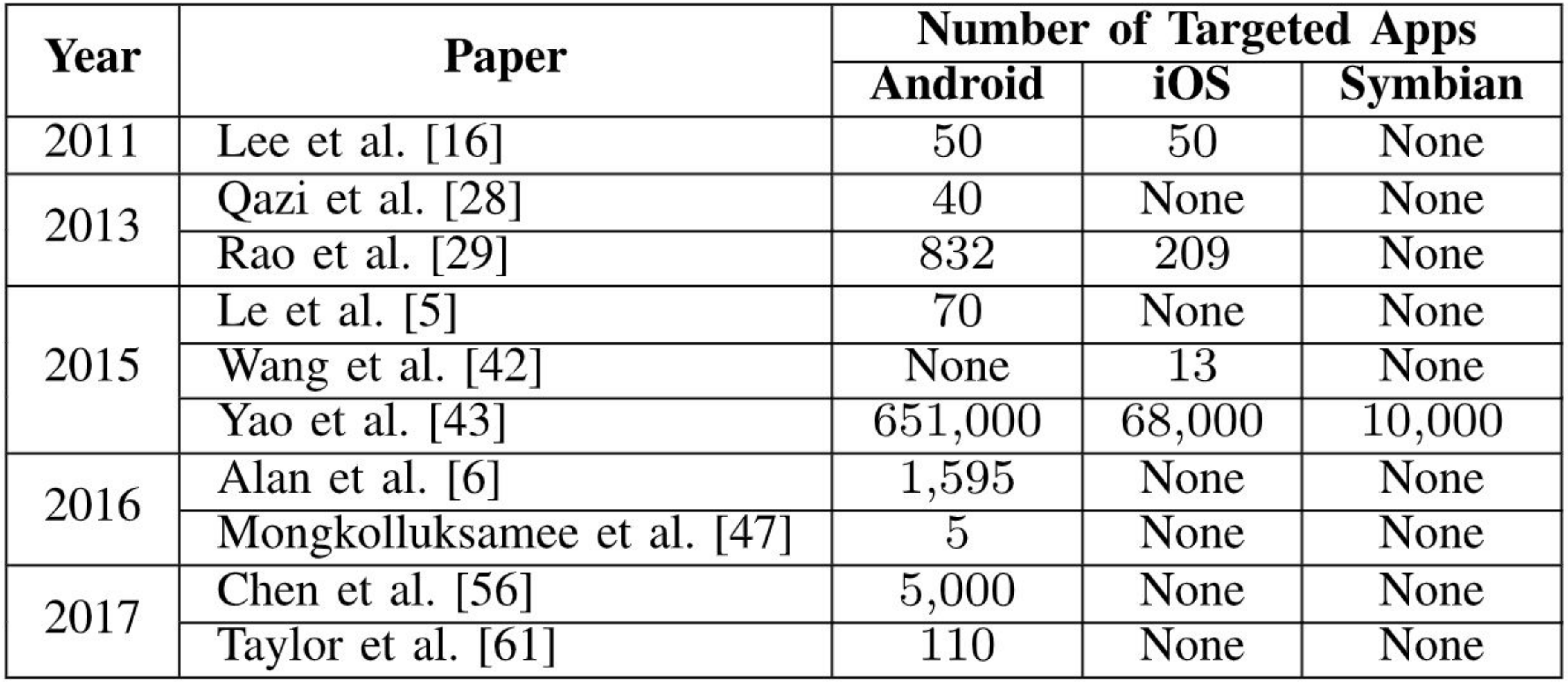
App Model | UI Model | Cloud

Software Foundation

(d) Windows Phone.

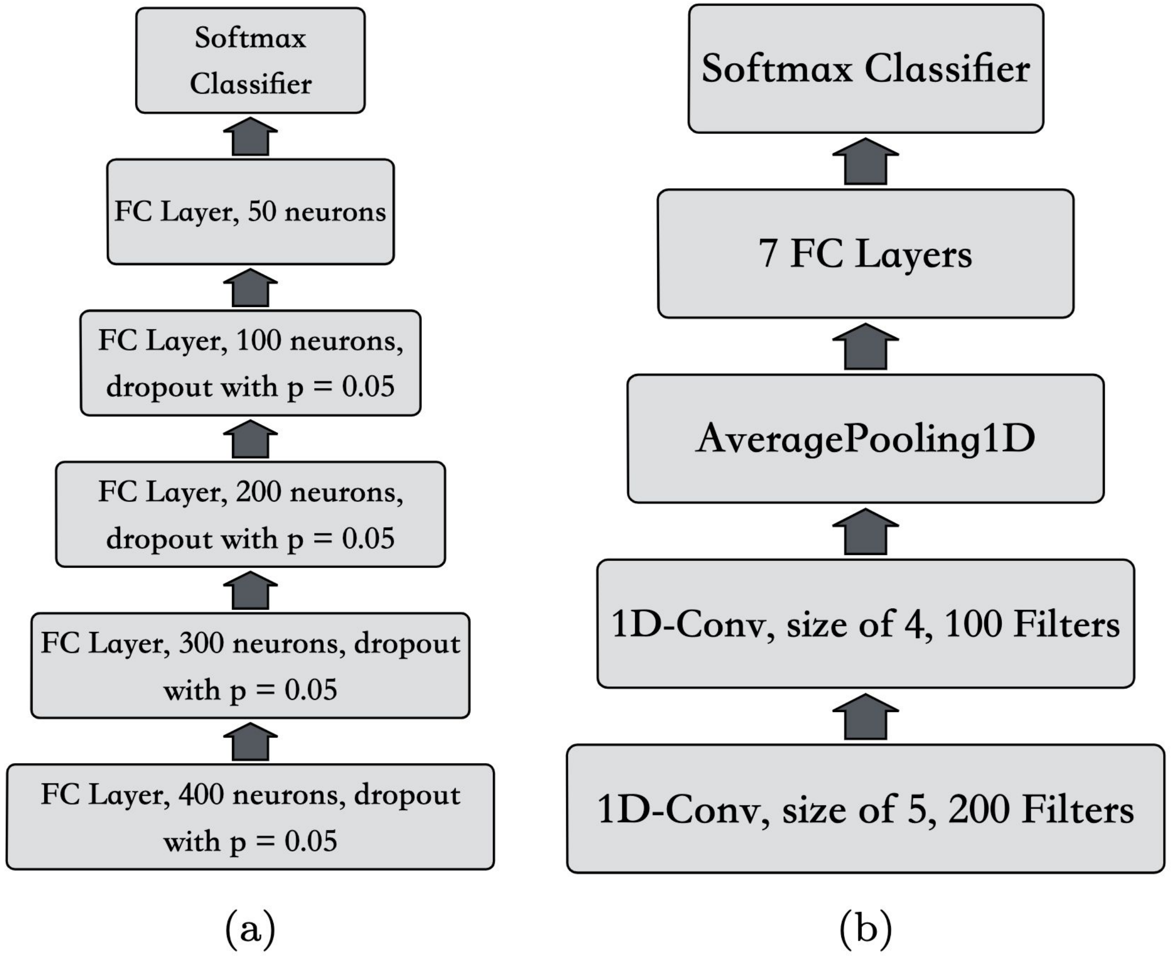
From: The Dark Side(-Channel) of Mobile Devices:A Survey on Network Traffic Analysis, IEEE COMMUNICATIONS SURVEYS & TUTORIALS, VOL. 20, NO. 4, FOURTH QUARTER 2018

App Identification



Year	Paper	Number of Targeted Apps		
		Android	iOS	Symbian
2011	Lee et al. [16]	50	50	None
2013	Qazi et al. [28]	40	None	None
	Rao et al. [29]	832	209	None
2015	Le et al. [5]	70	None	None
	Wang et al. [42]	None	13	None
	Yao et al. [43]	651,000	68,000	10,000
2016	Alan et al. [6]	1,595	None	None
	Mongkolluksamee et al. [47]	5	None	None
2017	Chen et al. [56]	5,000	None	None
	Taylor et al. [61]	110	None	None

Deep Learning for encrypted traffic classification



Softmax Classifier

FC Layer, 50 neurons

FC Layer, 100 neurons, dropout with p = 0.05

FC Layer, 200 neurons, dropout with p = 0.05

FC Layer, 300 neurons, dropout with p = 0.05

FC Layer, 400 neurons, dropout with p = 0.05

(a)

Softmax Classifier

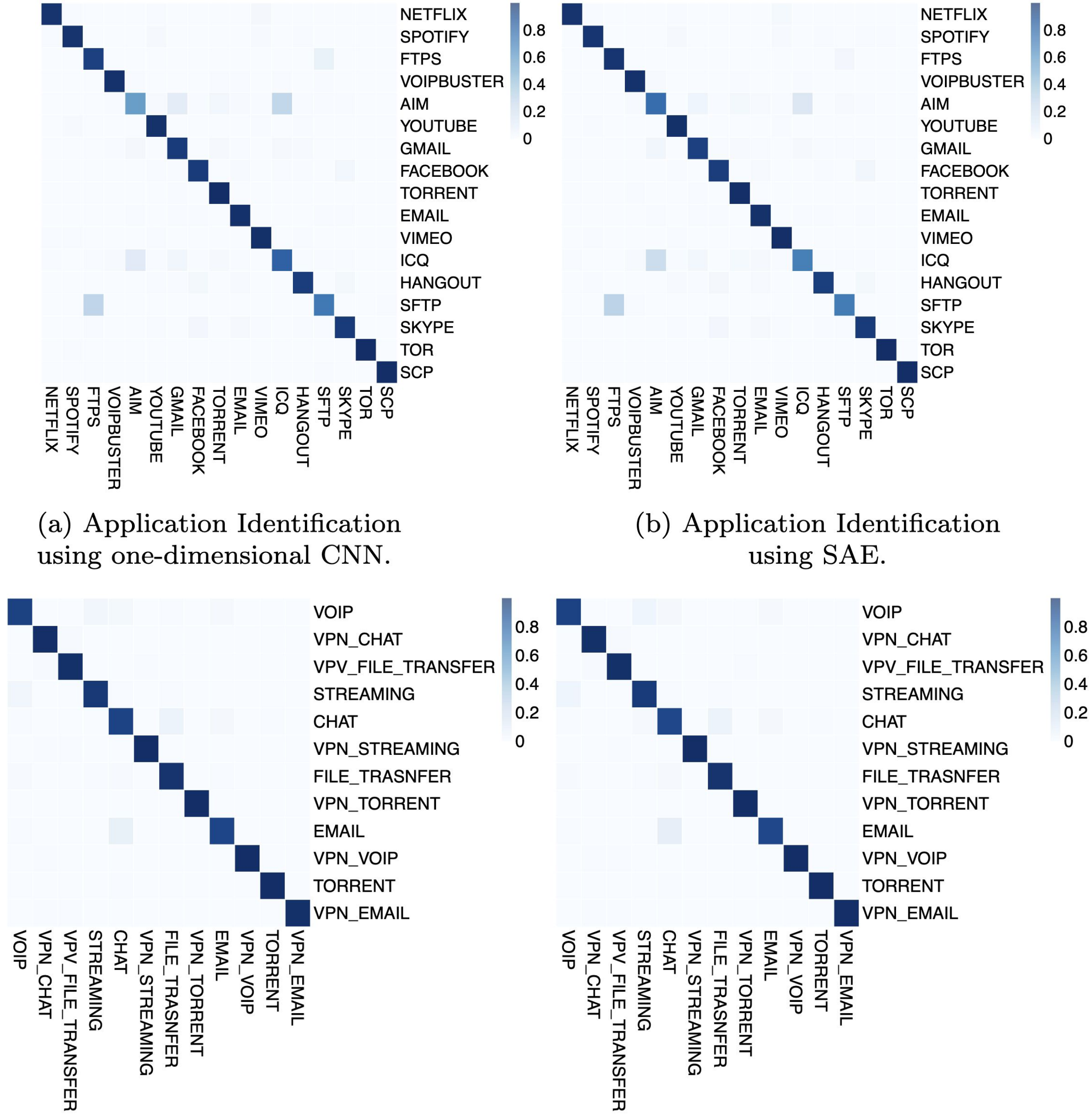
7 FC Layers

AveragePooling1D

1D-Conv, size of 4, 100 Filters

1D-Conv, size of 5, 200 Filters

(b)



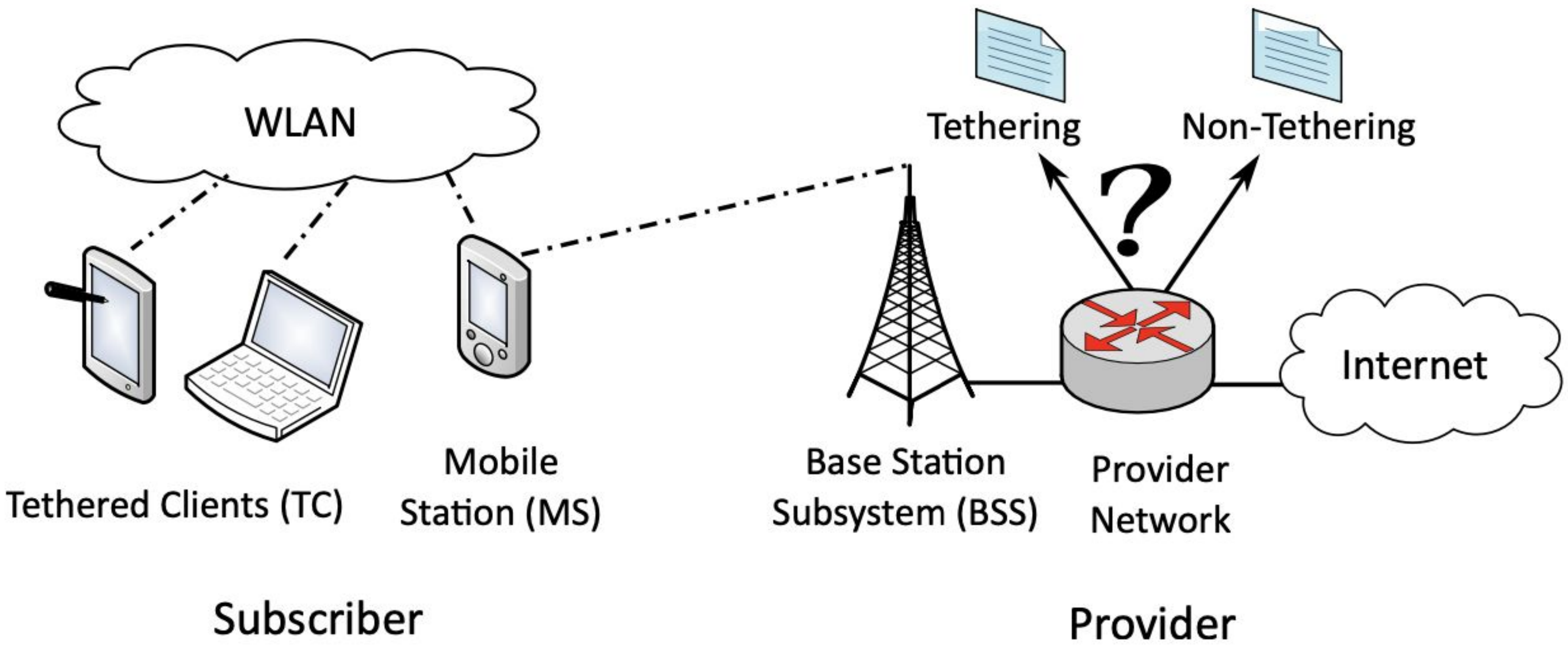
(a) Application Identification using one-dimensional CNN.

(b) Application Identification using SAE.

(c) Traffic Characterization using one-dimensional CNN.

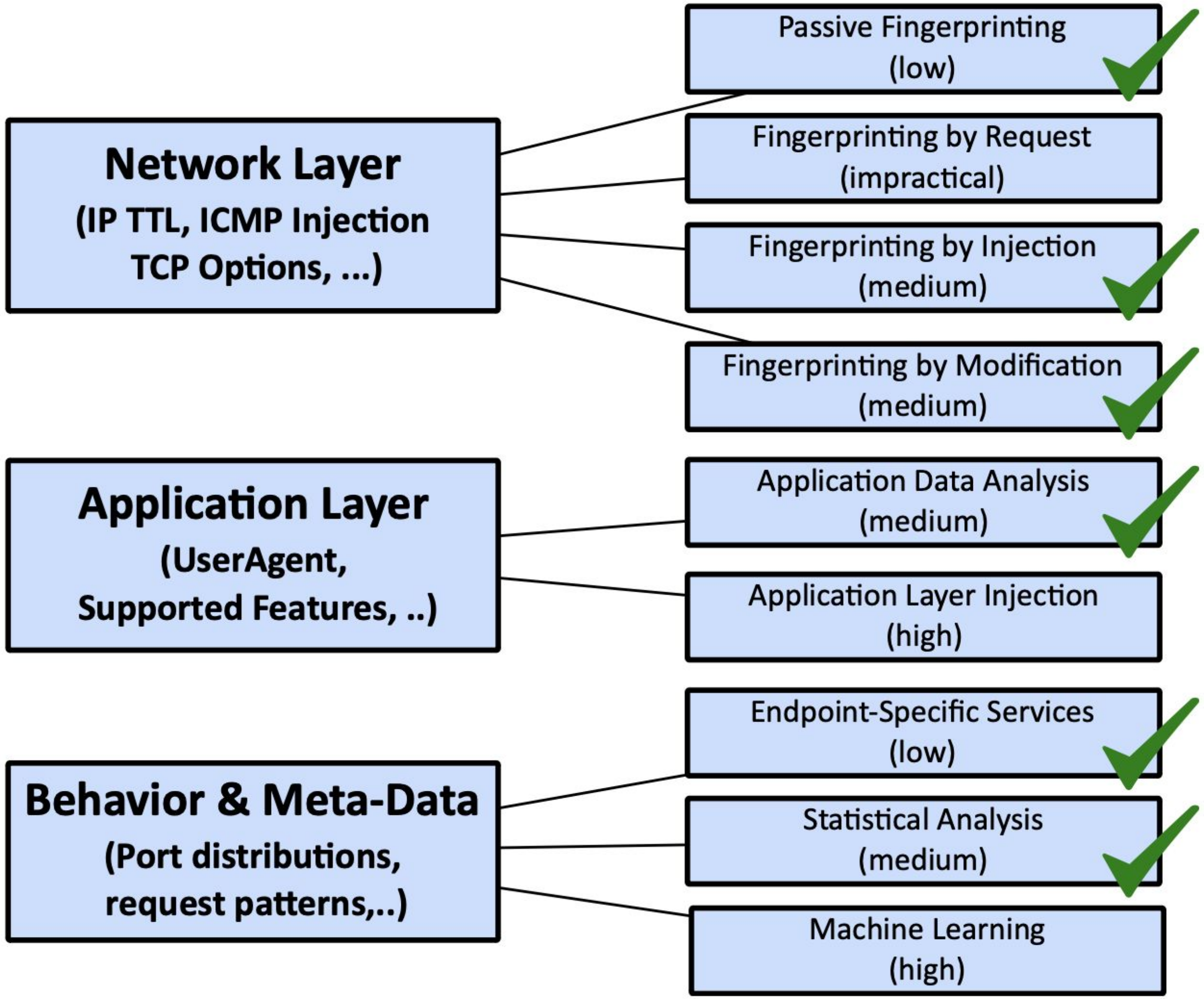
(d) Traffic Characterization using SAE.

USE Case: Mobile Tethering detection



https://www.researchgate.net/publication/230708469

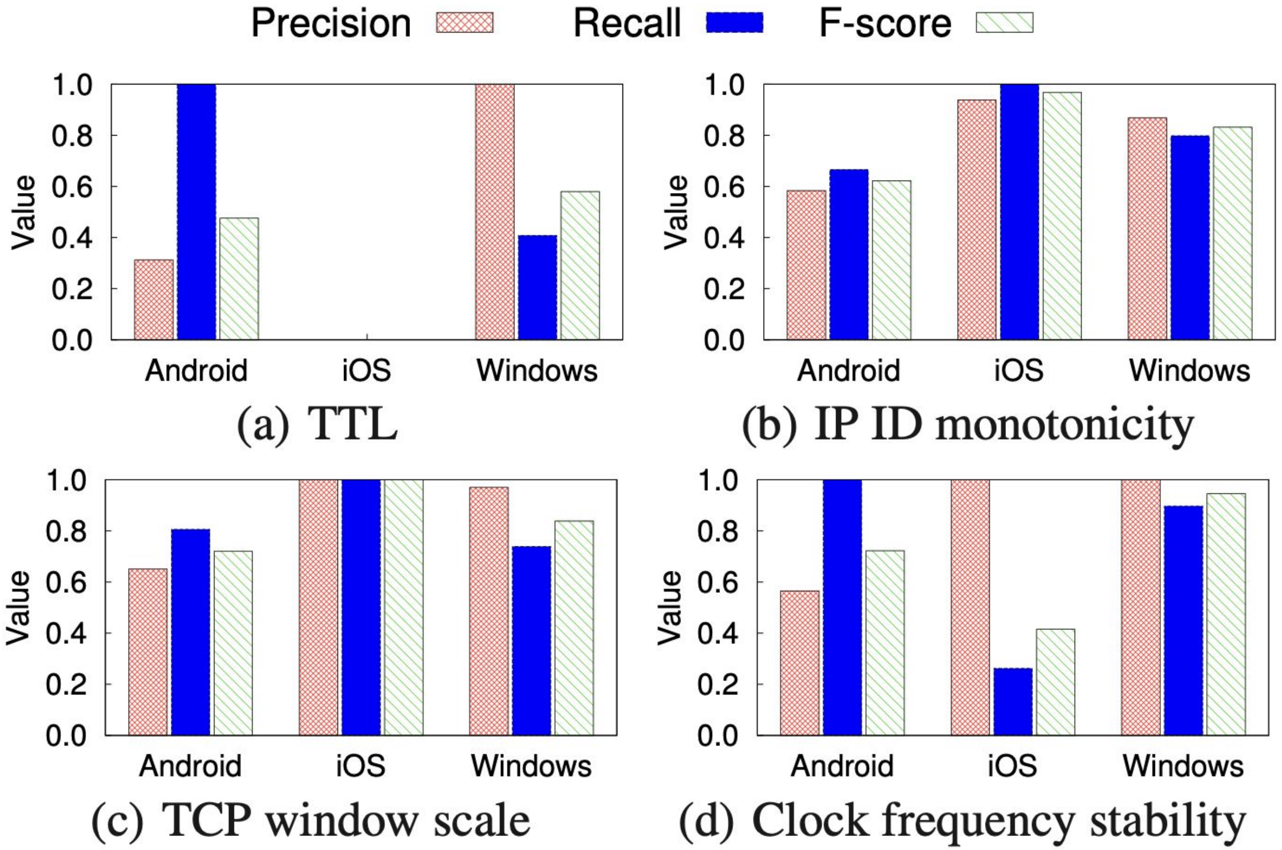
Classification of tethering detection mechanisms



USE CASE: Detect Mobile Tethering

- TCP/IP Stack Fingerprinting
    - Initial Packet Size
    - Initial TTL
    - IP ID
    - TCP Window Size
    - TCP Timestamp
- NAT Detection
- Destination IP/URL
    - Captive Portal Detection
        - when they first connect to a wifi network, they try to connect to a known web server across the internet, and checking to see if they get the response that they're expecting.
    - Windows Update Server

https://doi.org/10.1145/2663716.2663745



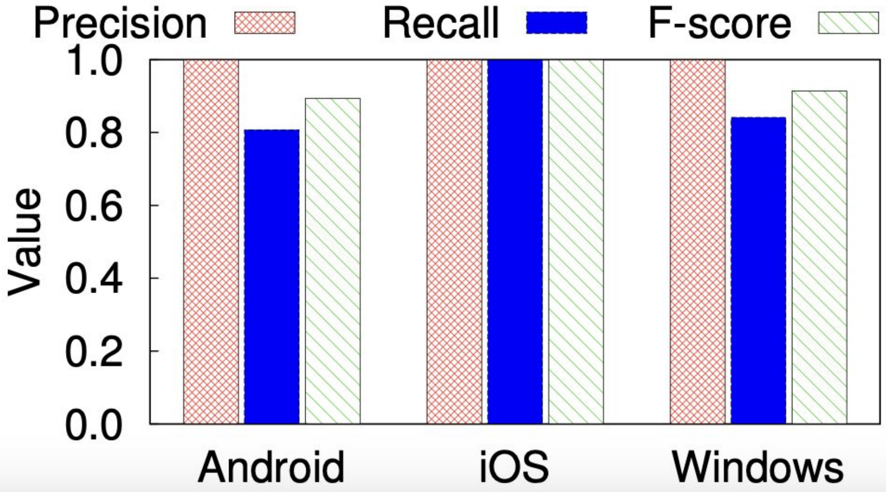
(a) TTL

(b) IP ID monotonicity

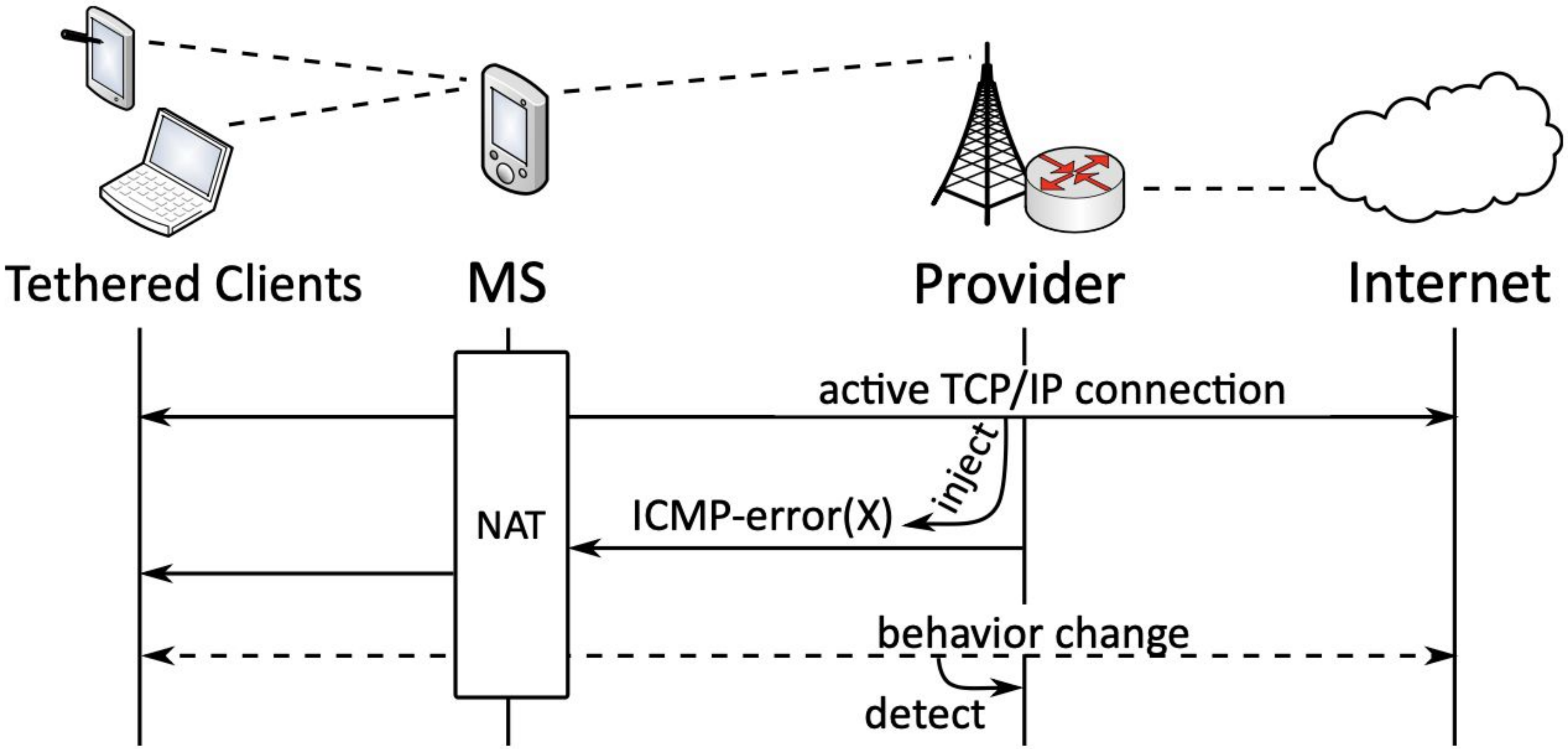
(c) TCP window scale

(d) Clock frequency stability

Figure 8: Accuracy of detecting OSes via individual features.



NAT Detection



USE CASE: Advertising monitoring/tracking

- What ad content the user monitors

Conclusion

- Deep Packet Inspection
  - Great network monitoring tool?
  - User’s privacy?