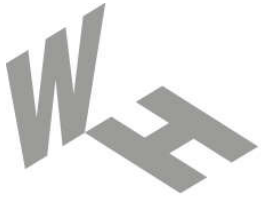


Westfälische Hochschule
Gelsenkirchen Bocholt Recklinghausen
University of Applied Sciences

Internet Analysis System → Part 1

Prof. Dr. (TU NN)
Norbert Pohlmann

Institute for Internet Security - if(is)
University of Applied Sciences Gelsenkirchen
http://www.internet-sicherheit.de

if(is) internet security.

Content

- Aim and outcomes of this lecture
- Idea of the Internet Analysis System
- Knowledge Base
- Outline of the Current State
- Detection of Attacks and Deflection
- Forecast of Patterns and Attacks
- Summary

Content

- **Aim and outcomes of this lecture**
- Idea of the Internet Analysis System
- Knowledge Base
- Outline of the Current State
- Detection of Attacks and Deflection
- Forecast of Patterns and Attacks
- Summary

Internet Analysis System (IAS)
→ Aims and outcomes of this lecture

Aims

- To introduce an (Internet) Early Warning System with a statistical approach
- To explore the structure of the Internet Analysis System
- To analyze the results of the Internet Analysis System
- To assess the value the Internet Analysis System

At the end of this lecture you will be able to:

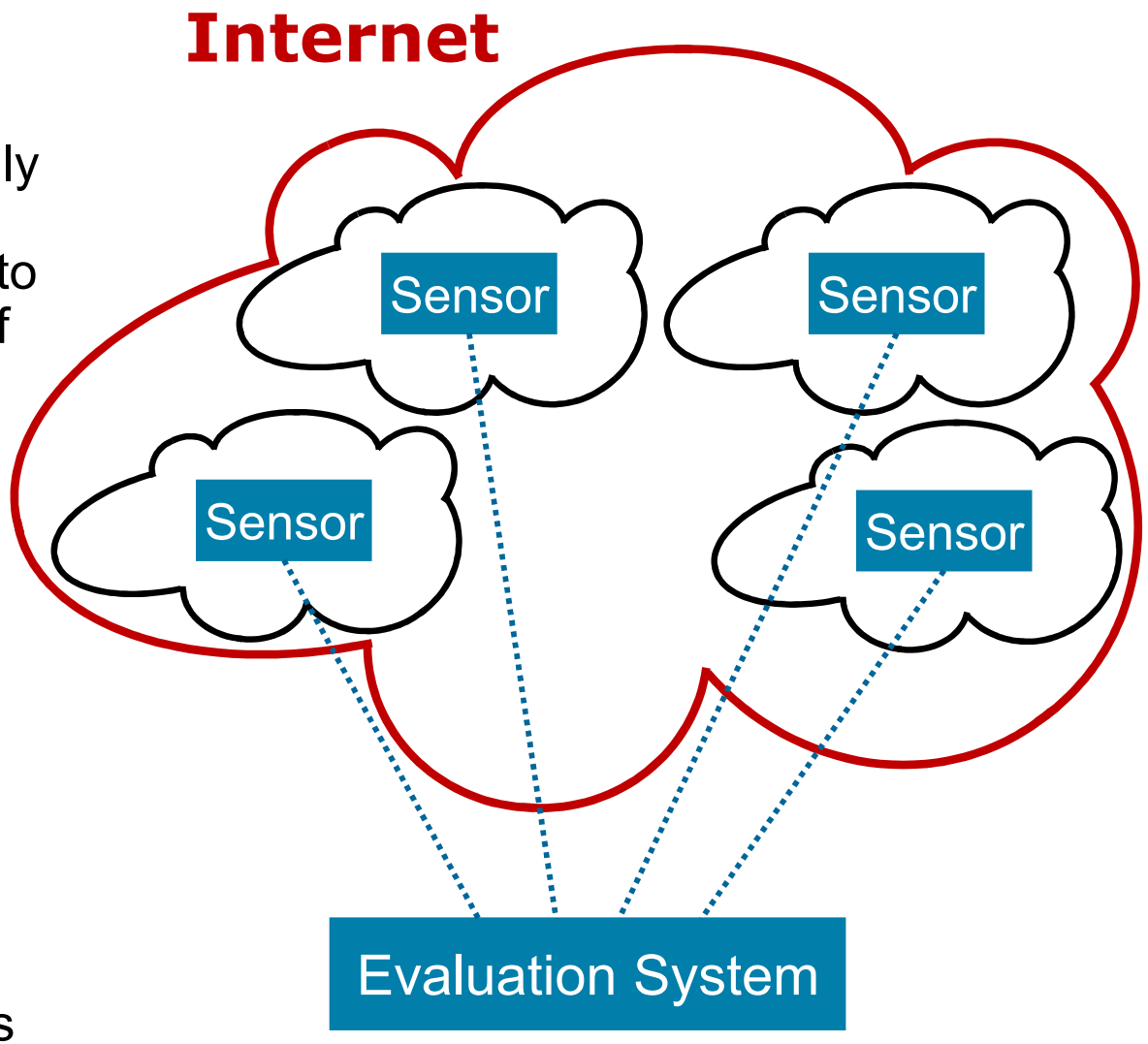
- Understand what is meant by the Internet Analysis System.
- Know something of the structure of the Internet Analysis System.
- Know what the results of the Internet Analysis System could be.
- Understand the capabilities and limitations of the Internet Analysis System.

Content

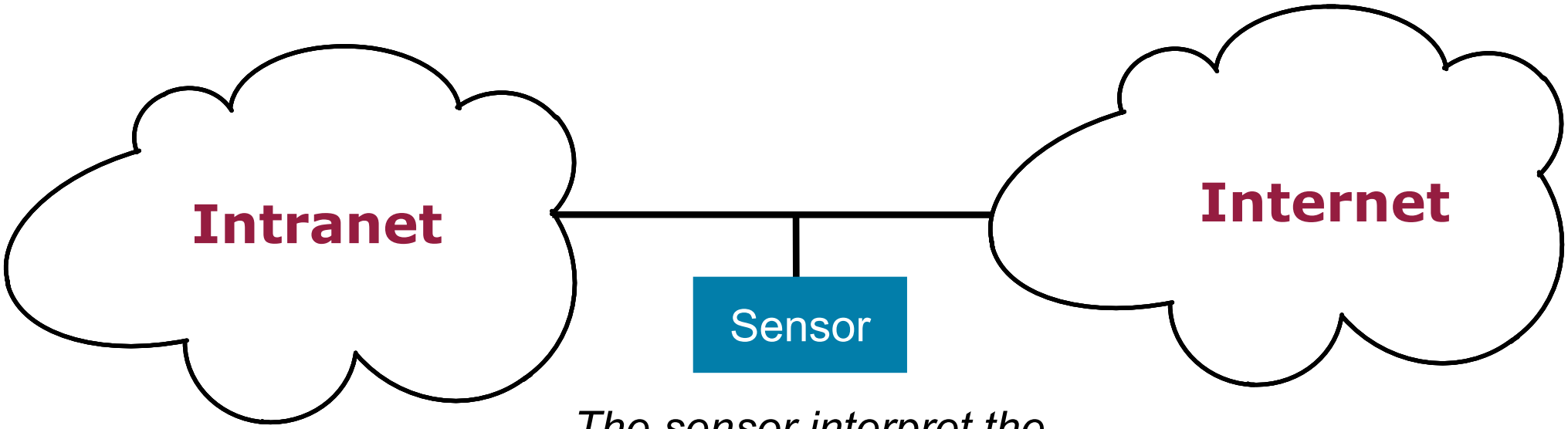
- Aim and outcomes of this lecture
- **Idea of the Internet Analysis System**
- Knowledge Base
- Outline of the Current State
- Detection of Attacks and Deflection
- Forecast of Patterns and Attacks
- Summary

Internet Analysis System (IAS)
→ Idea

- Observation of the critical infrastructure **„Internet".**
- **Sensor** are placed in thoughtfully selected spots of the **internet communication infrastructure** to gather the raw data, consisting of counted header information.
- Only header information is counted, which is **not considered as data privacy relevant**.
- The system gathers information over a **great period of time**!
- A centrally managed **Evaluation System** is used to analyse the raw data and to display the detailed results in an intuitive manner.

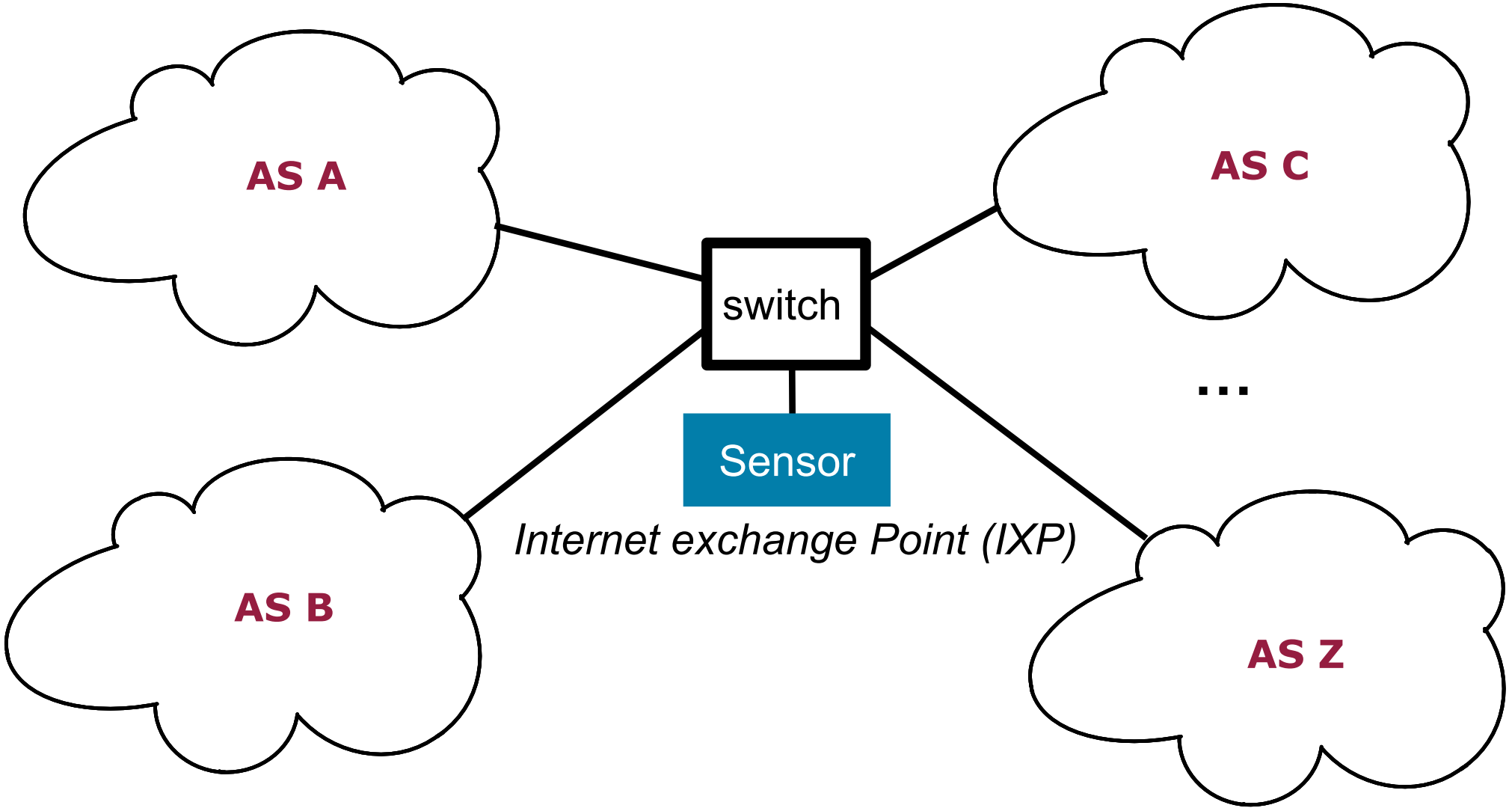


Position of the sensor
→ Between the Intranet and the Internet

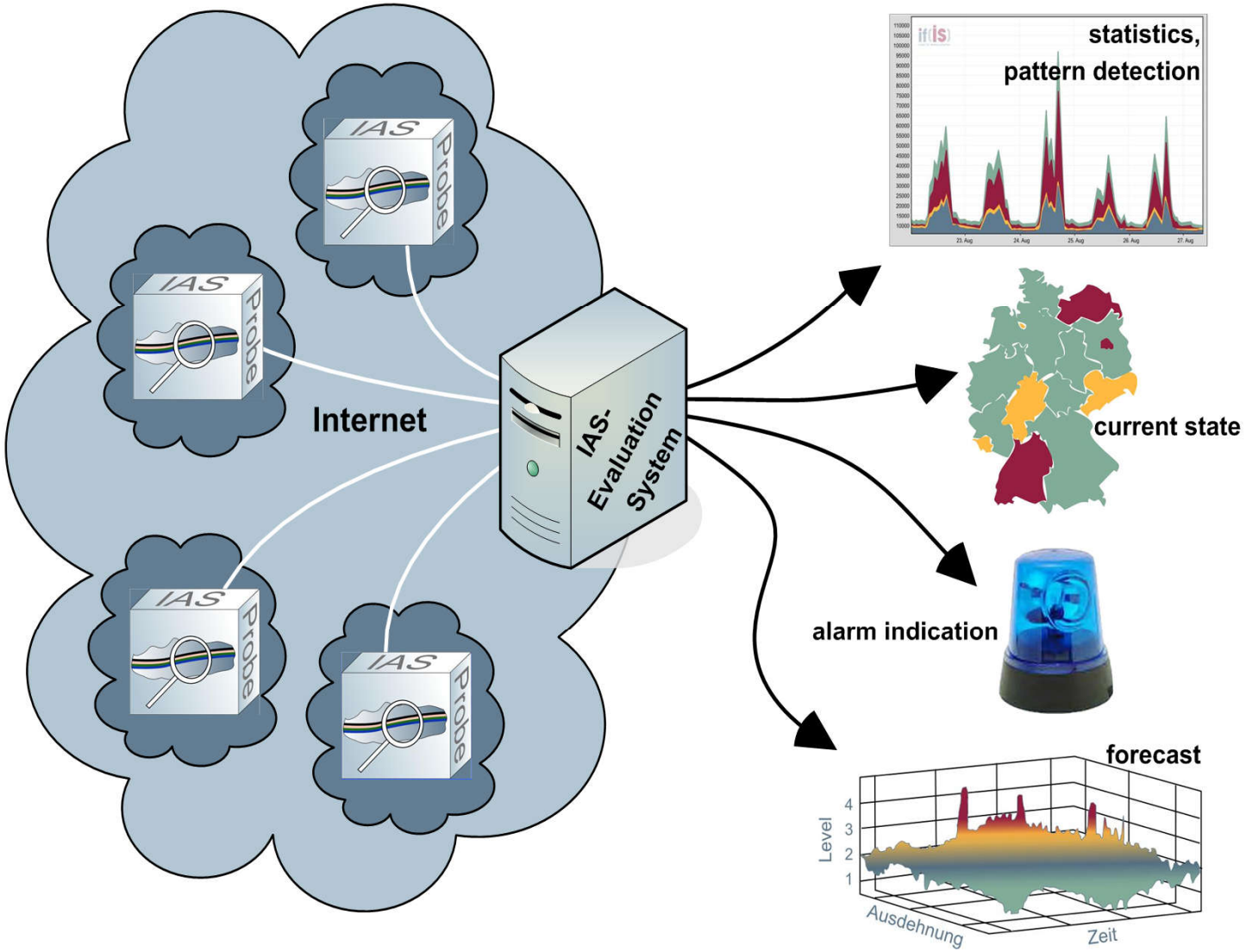


The sensor interpret the behavior in communication between an Intranet (e.g. corporate network) and the Internet

Position of the sensor
→ Between ASs (IXP)



Internet Analysis System
→ Targets



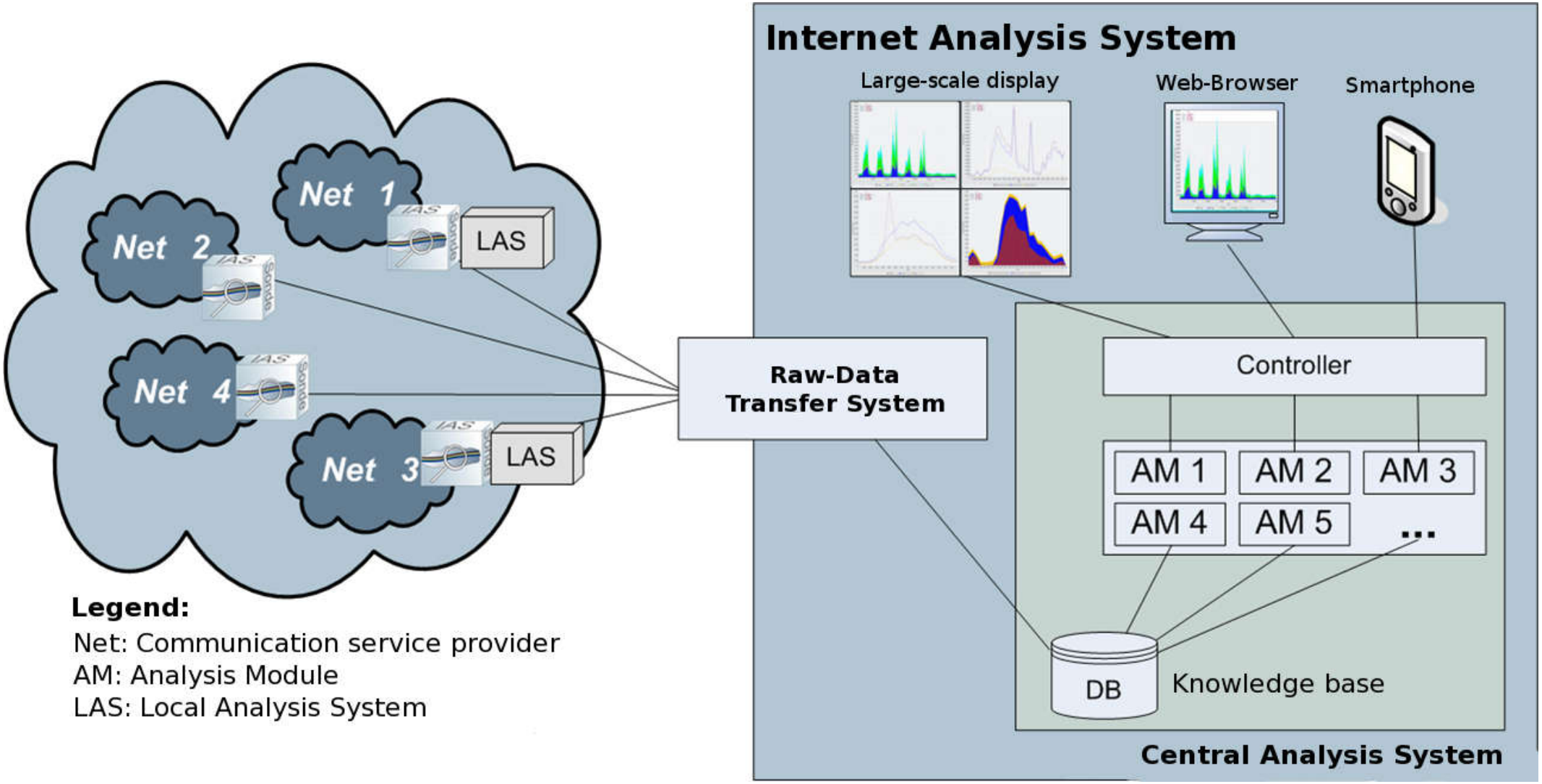
Description of profiles, patterns and coherences, creation of a knowledge base.

Outline of the current state of the internet.

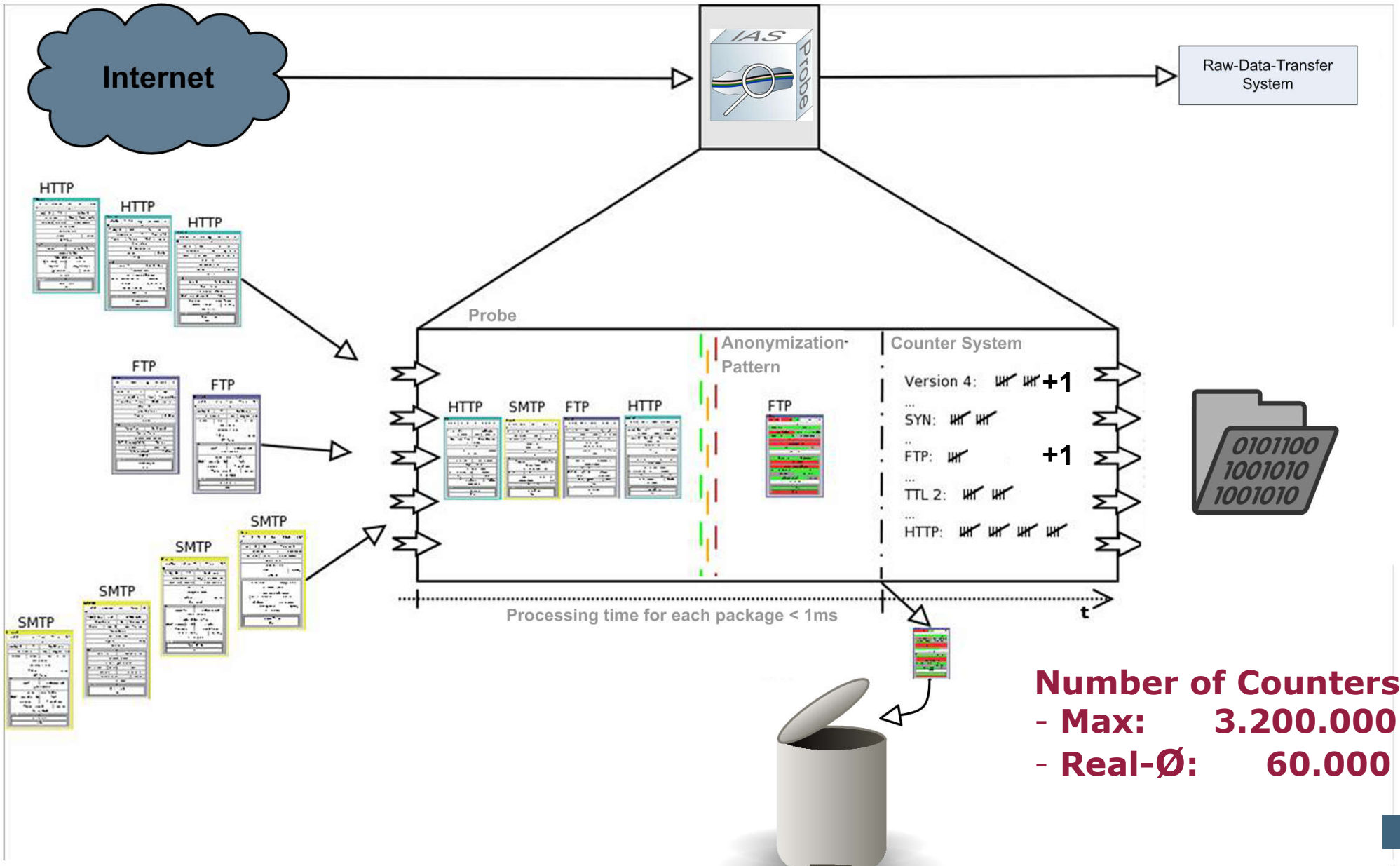
Detection of attacks and of deflections.

Forecast of patterns and attacks.

Implementation of the IAS
→ Overview



Internet Analysis System (IAS)
→ Counting of header information (1/3)



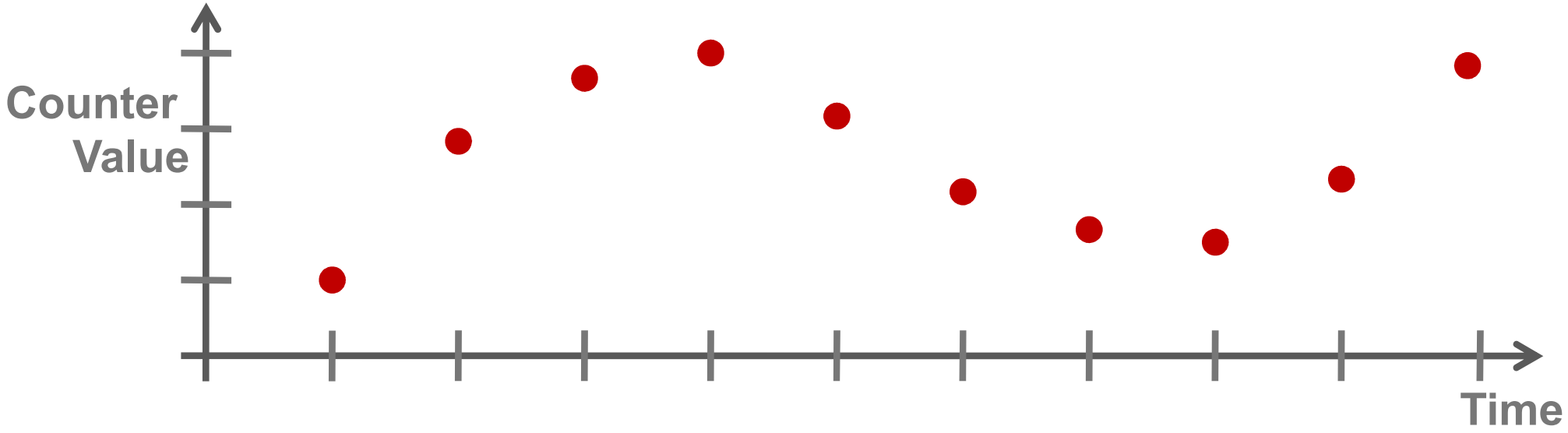
Number of Counters:
- **Max: 3.200.000**
- **Real-Ø: 60.000**

Internet Analysis System (IAS)
→ Counting of header information (2/3)

ID	Description		Count
131134	IP (Protocol Number 6)	:	18.854.151
131145	IP (Protocol Number 17)	:	1.123.149
327708	TCP (Flags: SYN)	:	334.435
327723	TCP (Flags: FIN/ACK)	:	480.697
327724	TCP (Flags: SYN/ACK)	:	275.779
545857	HTTP (Request Method POST)	:	2.026
545861	HTTP (Request Method GET)	:	293.616
545863	HTTP (Request Method HEAD)	:	18.992

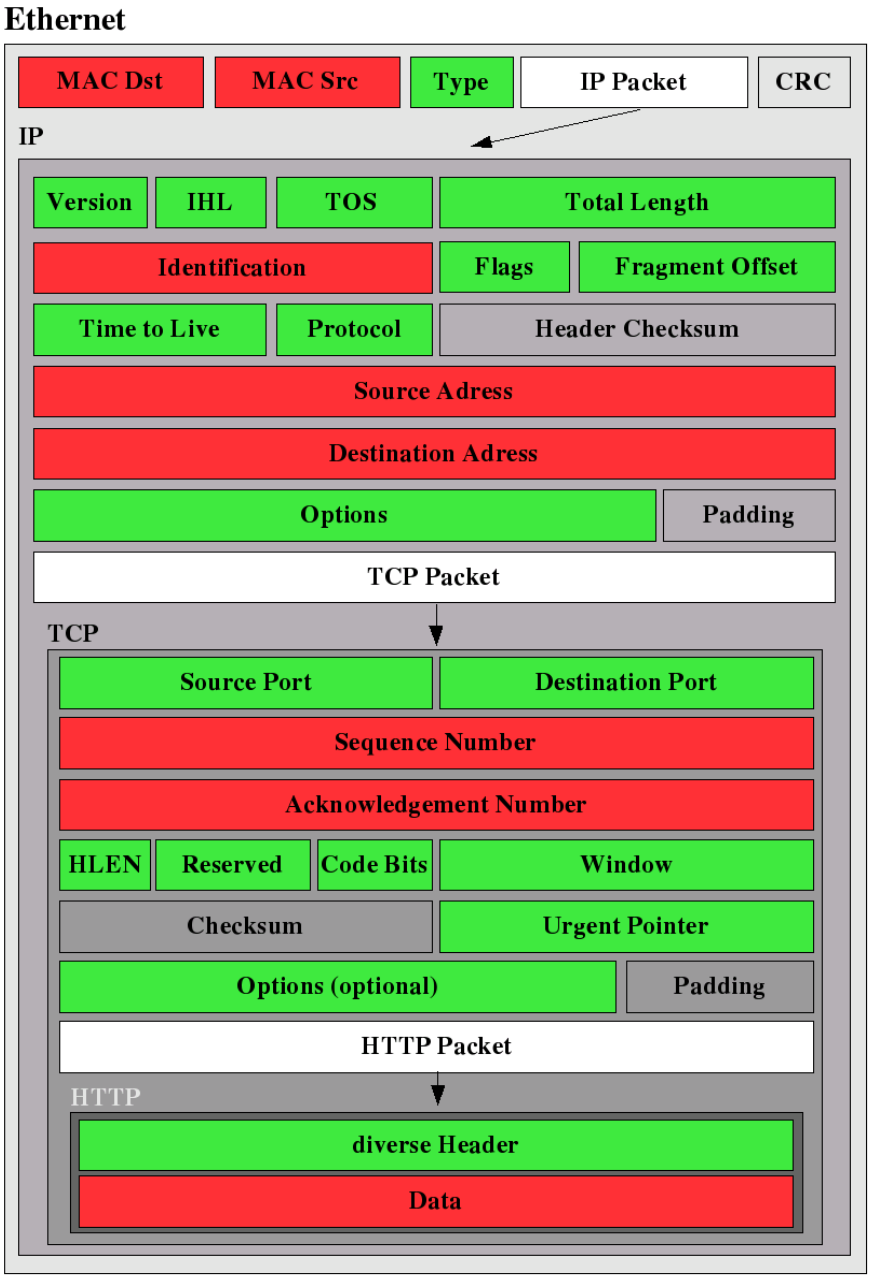
- On the right behind the colon character are the **counter values** for each parameter specified on the left.
- Each line stands for one counter.
- For example, line 2 indicates that 1,123,149 packets with the IP protocol number 17 (UDP) appeared in the prescribed time interval.
- All of this information is completely anonymous!

Internet Analysis System
→ Counting of header information (3/3)



Principle of raw data collection → Protocol stack (1/2)

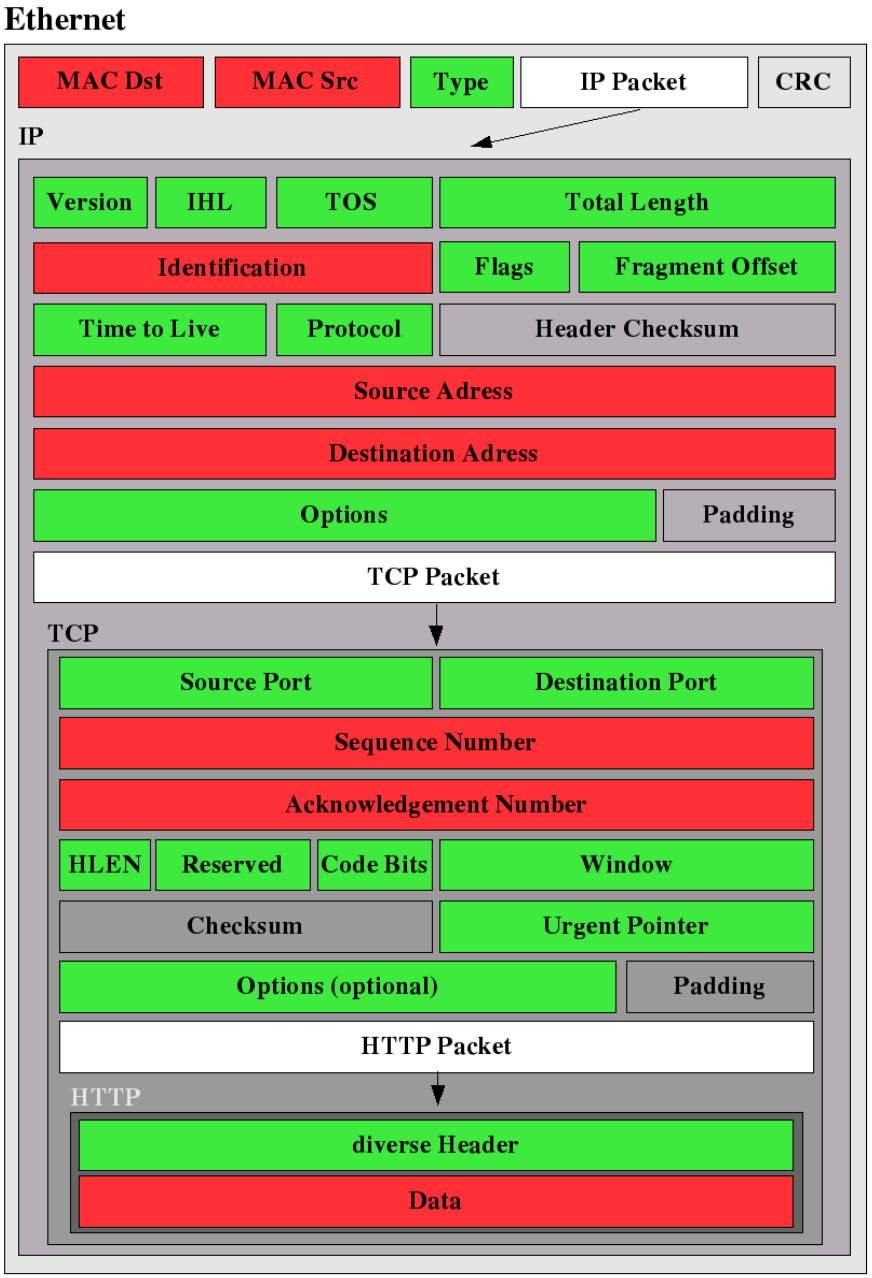
- **Ethernet**
  - Type: Type of the nested packets, in this case: 0x0800 (IP)
  - Checksum (CRC) irrelevant
- **Internet Protocol (IP)**
  - e.g.: Total Length of the packet
  - Protocol: Type of the nested Packet, in this case: 6 (TCP)
  - Source- and destination address privacy critical



Principle of raw data collection → Protocol stack (2/2)

- **Transmission Control Protocol (TCP)**
  - Port: end point of the connection
    - HTTP: 80 (WWW)
    - Others e.g.:
      SMTP (25), HTTPS (443)
  - Code Bits
    - Information about the connection establishment and shut down

- **Hypertext Transfer Protocol (HTTP)**
  - Header:
    - e.g.: User Agent:
      describes the user's browser
  - User data (DATA)
    e.g.: content of a web site



Internet Analysis System (IAS)
→ Counting of header information

- Description of the network traffic
  - Sequence of packets on the line

$$S = \langle P_1, P_2, \ldots, P_N \rangle$$

- A network packet (P) consists of

$$P = \langle H, PL \rangle$$

  - H:= Header := $\langle h_1, h_2, \ldots, h_k \rangle$
  - PL:= Payload := $\langle b_1, b_2, \ldots, b_l \rangle$ (the payload could be empty)
  - Header fields can belong to different protocols

- Each header field (h_i) can consist of number of values (w_j)

- For each of these values a counter is defined $z_i \in \aleph$ which indicates, how often a specific value of a header field has already occurred

$$h_i \in \{w_1, w_2, \ldots, w_l\}$$

(see examples)

Internet Early Warning System
→ Evaluation counter (1/5)

Protocol	Number	Protocol	Number
DNS	9.458	EDONKEY	53
EMULE	19	Ethernet II	6
FTP	103	HTTP	1.123
HTTPS	179	ICMP	318
IKEv2	10.764	IMAP	40
IMAPS	179	IP	9.089
IPCO	4	IPSEC-AH	513
IRC	499	ISAKMP	4.912
META	14	P2P	6
POP	1.015	POPS	179

Internet Early Warning System
→ Evaluation counter (2/5)

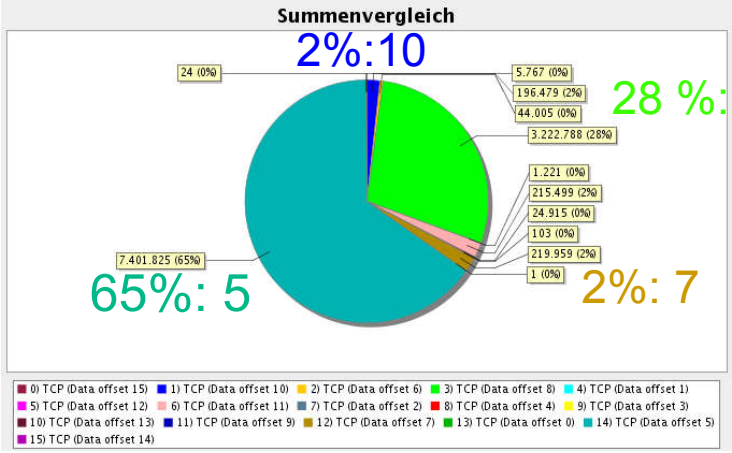
Protocol	Number	Protocol	Number
RTP	37	SIP	138
Skype	1	SMTP	1.624
SMTPS	179	TBURL	23.986
TCP	678.614	TFTP	17
UDP	131.590		

Internet Early Warning System
→ Evaluation counter (3/5)

TCP	**678.614**
reserved	1
High ports (P2P definition)	4
ecn	8
HLEN	16
TCP Flags (Code Bits)	66
Window (size)	255
options	522
Well-known P-D (1024) + TTL (8)	8.192 (1.024 * 8)
Well-known P-S/D (1024) + Flags combi (8)	16.384 (2 * 8.192)
Port (Source/Destination – S/D)	131.072 (2 * 2^{16})
Well-known P-S/D (1024) + total length (255)	522.240 (2 * 261.120)

Internet Early Warning System
→ Evaluation counter (4/5) – TCP - HLEN

- **TCP – HLEN** (16)
  - Specifies the size of the TCP header in 32-bit words. The minimum size header is 5 and the maximum is 15 words.
  - Data Offset 0
  - Data Offset 1
  - Data Offset 2
  - Data Offset 3
  - Data Offset 4 (standard **HLEN 5**)
  - Data Offset 5 (+ option)
  - Data Offset 6 (+ option **HLEN: 7**)
  - Data Offset 7 (+ option **HLEN: 8**)
  - …
  - Data Offset 15 (+ option)



HLEN: 7
MSS (4 Byte)
NOP (2 Byte)
SACK permitted (2 Byte)

HLEN: 8
NOP (2 Byte)
Timestamp (10 Byte)

HLEN: 10
NOP (2 Byte)
SACK (18 Byte)
oder
MSS (4 Byte)
SACK permitted (2 Byte)
NOP (1 Byte)
WScale (3 Byte)

Internet Early Warning System
→ Evaluation counter (5/5) – P2P

- **P2P Counter**
  - UDP Port 4672
  - **Source >= 1024 and Destination >= 1024 (« P2P »)**
    - If both the source- and the destination-port are greater than or equal to 1024, this is an approximate estimate of the client-to-client communication, as for that only ports in the upper part are elected.
  - **Source < 1024 and Destination < 1024 (« B2B »)**
    - If both the source- and the destination-port are lower than 1024, this is an approximate estimate of the server-to-server communication, as for that only ports in the lower part are elected.
  - **Source >= 1024 and Destination < 1024 (« P2B »)**
  - **Source < 1024 and Destination >= 1024 (« B2P »)**

Internet Early Warning System
→ Evaluation counter

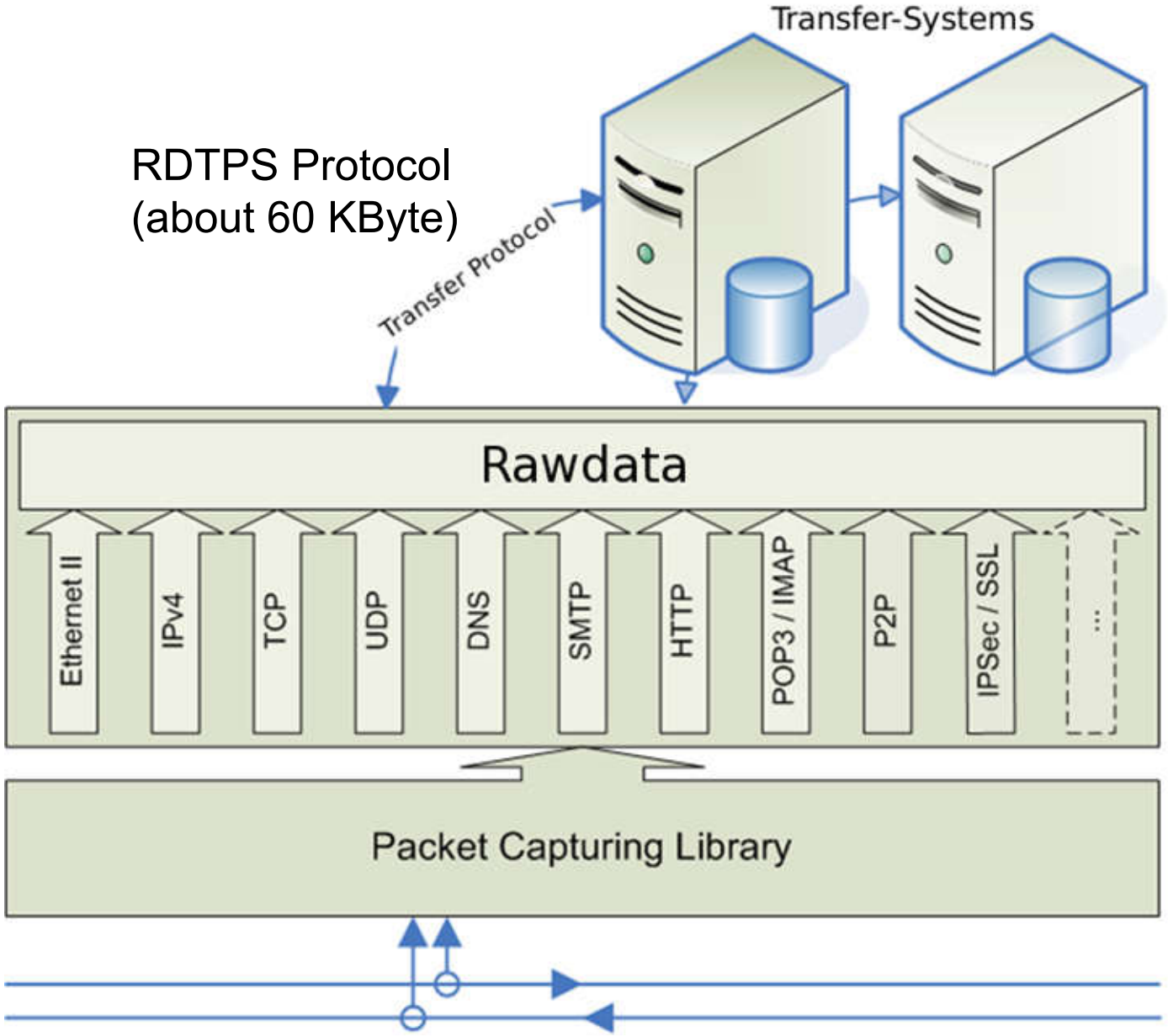
- Distribution of the counters

	UNI Santa Maria	Computer Science Department	Dt. Messe AG	Dr. Buelow & Masiak
Max	115.343	16.462	26.167	75.595
Min	45.889	7.474	4.343	26.120
average	72.795	11.264	12.263	48.376
Total count of packets (P2P)	555.555 (150.00)	36.111	61.361	154.400
Union	276.287	125.696	148.291	267.840

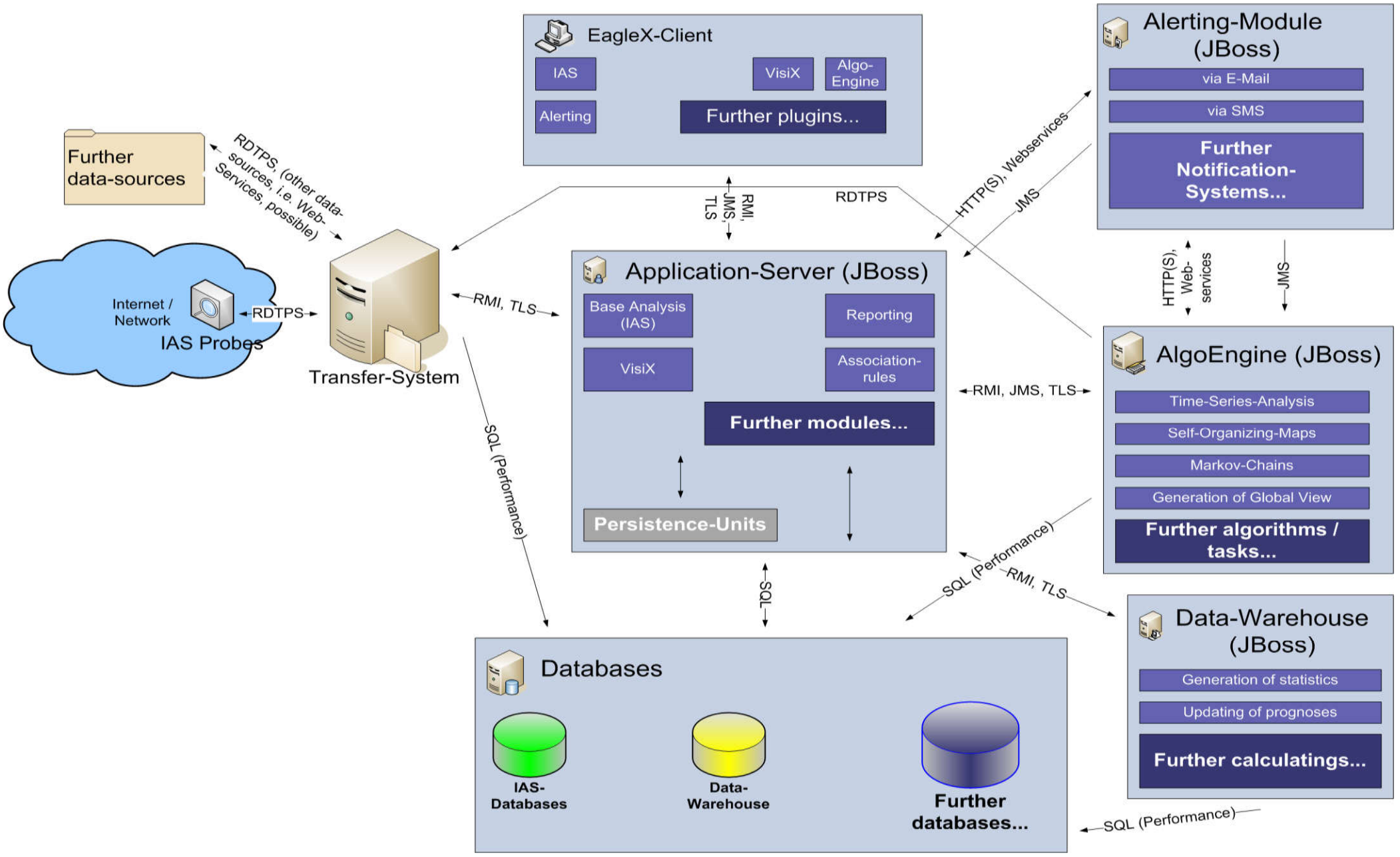
Principle of raw data collection
→ Why do we only use "tally sheets"?

- **Enhances performance**
  - No tracking of connections or sessions
  - Irrelevant information can be ignored
    - e.g.: Checksums
- **Protection of critical information**
  - Since the connections and sessions cannot be put together again
  - Since critical content is left out from further processing
    - IP/ MAC addresses
    - User data
  - **Anonymization by design**

IAS: Current State of Development
→ Sensor



IAS: Current State of Development
→ Architecture of the system



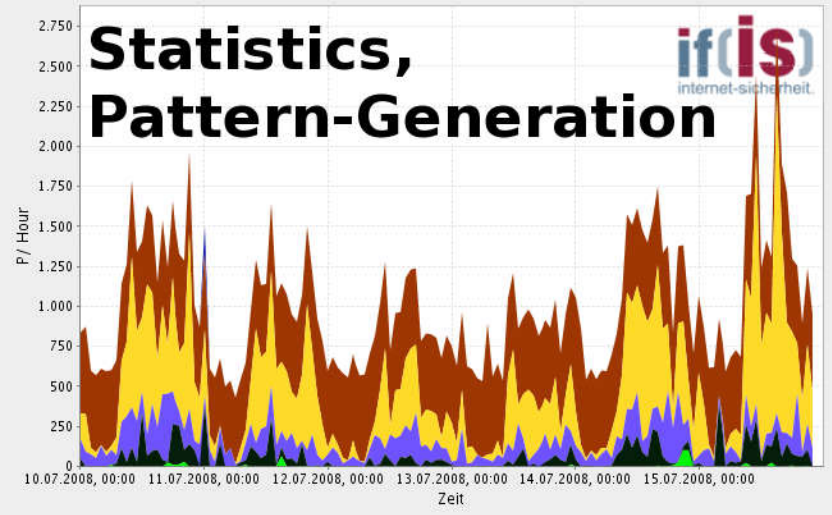
Content

- Aim and outcomes of this lecture
- Idea of the Internet Analysis System
- **Knowledge Base**
- Outline of the Current State
- Detection of Attacks and Deflection
- Forecast of Patterns and Attacks
- Summary

Internet Analysis System (IAS)
→ Defined targets

Target 1

- **Description of profiles, patterns and coherences**
- **Creation of a knowledge base.**



We want to create a knowledge base which we can use to understand the functioning of the internet from the “communication behaviour” point of view. The main task here is the support in analysing communication parameters – our row data - with the aim of identifying a pattern in the profiles, technological trends and correlations.

Internet Analysis System (IAS)
→ Target 1: Overview

- Counting of communication parameters by the sensor
- Transmitting of the counter readings (raw data) to the transfer system
- Long term storage in a database

Establishment of a knowledge base

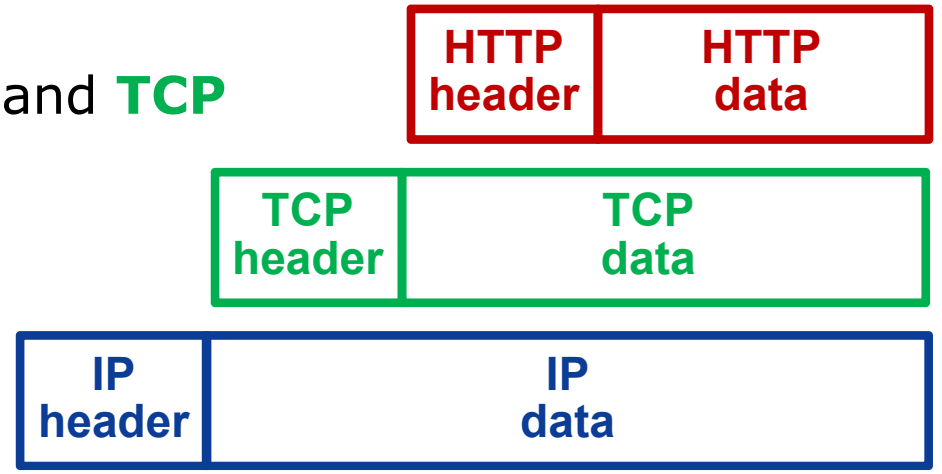
- Preservation of the raw data in a database
- Gaining on experience and collection of events / incidents

Description of profiles, patterns, technology trends and other coherences.

- Analyzing of the raw data with the expert tool
- (automated) generation of reports

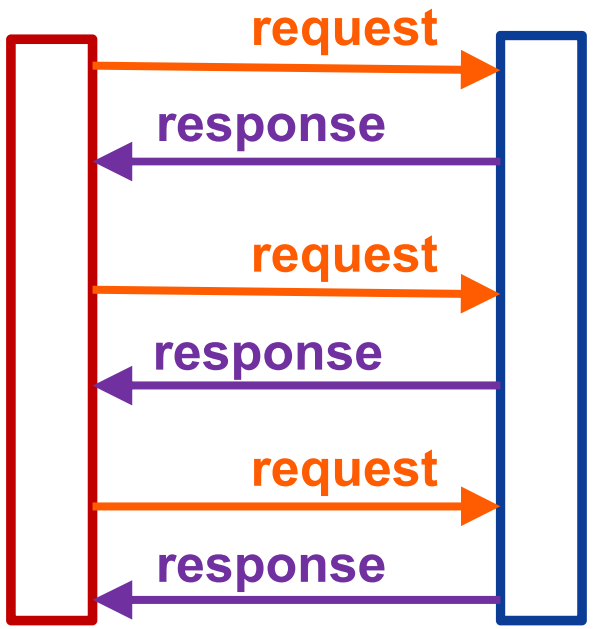
Knowledge Base - IAS
→ Coherences (1/4)

- **Coherences in architectural matters**
  - When **http** is detected, then also **IP** and **TCP**

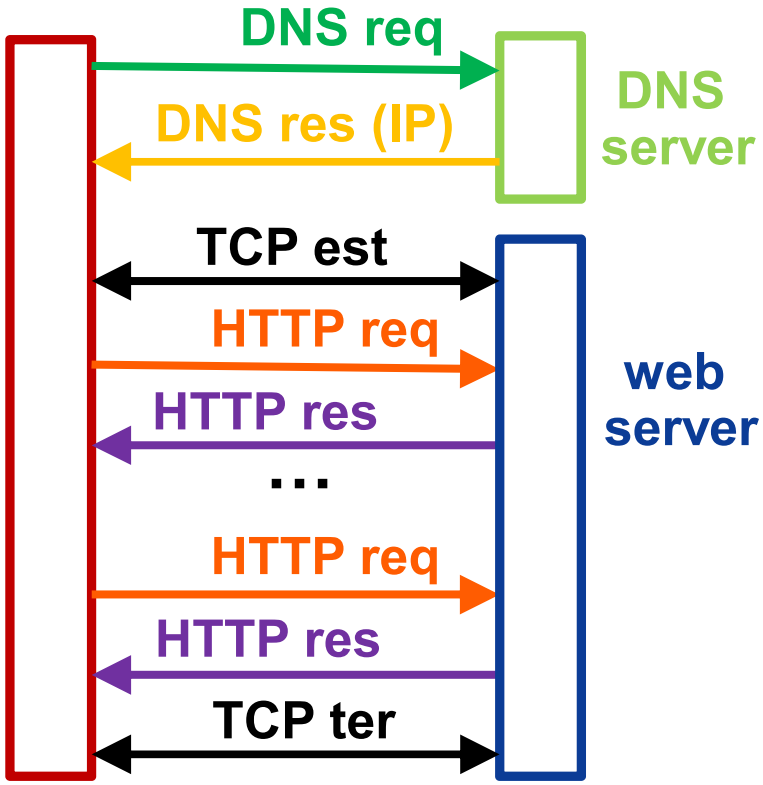


TCP/IP stack

- **Coherences in protocol matters**
  - When we detect a http **request**, then we should detect a http **response** as well



Knowledge Base - IAS
→ Coherences (2/4)

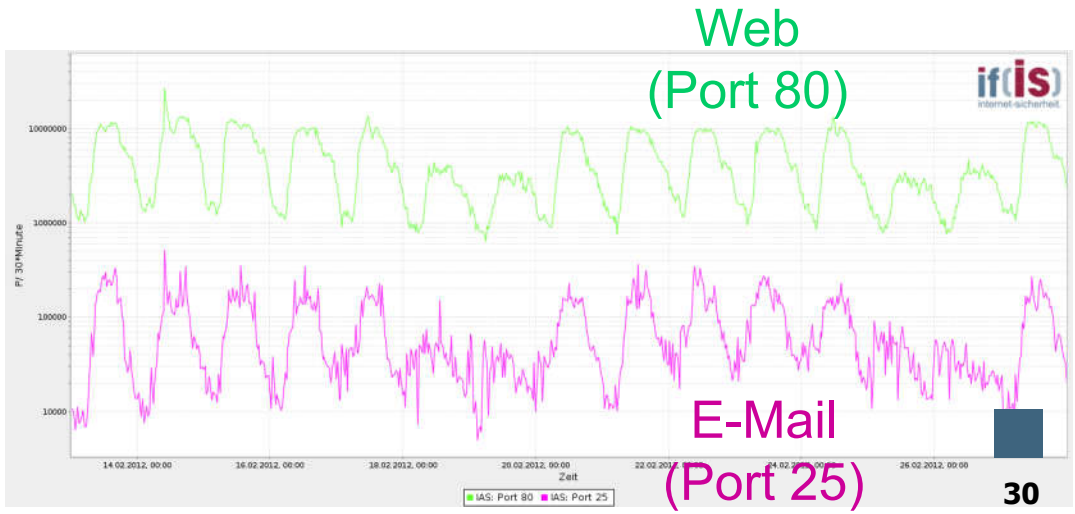


Coherences due to system matters

- When we detect **http traffic**, then in most cases we have also recorded **DNS traffic**

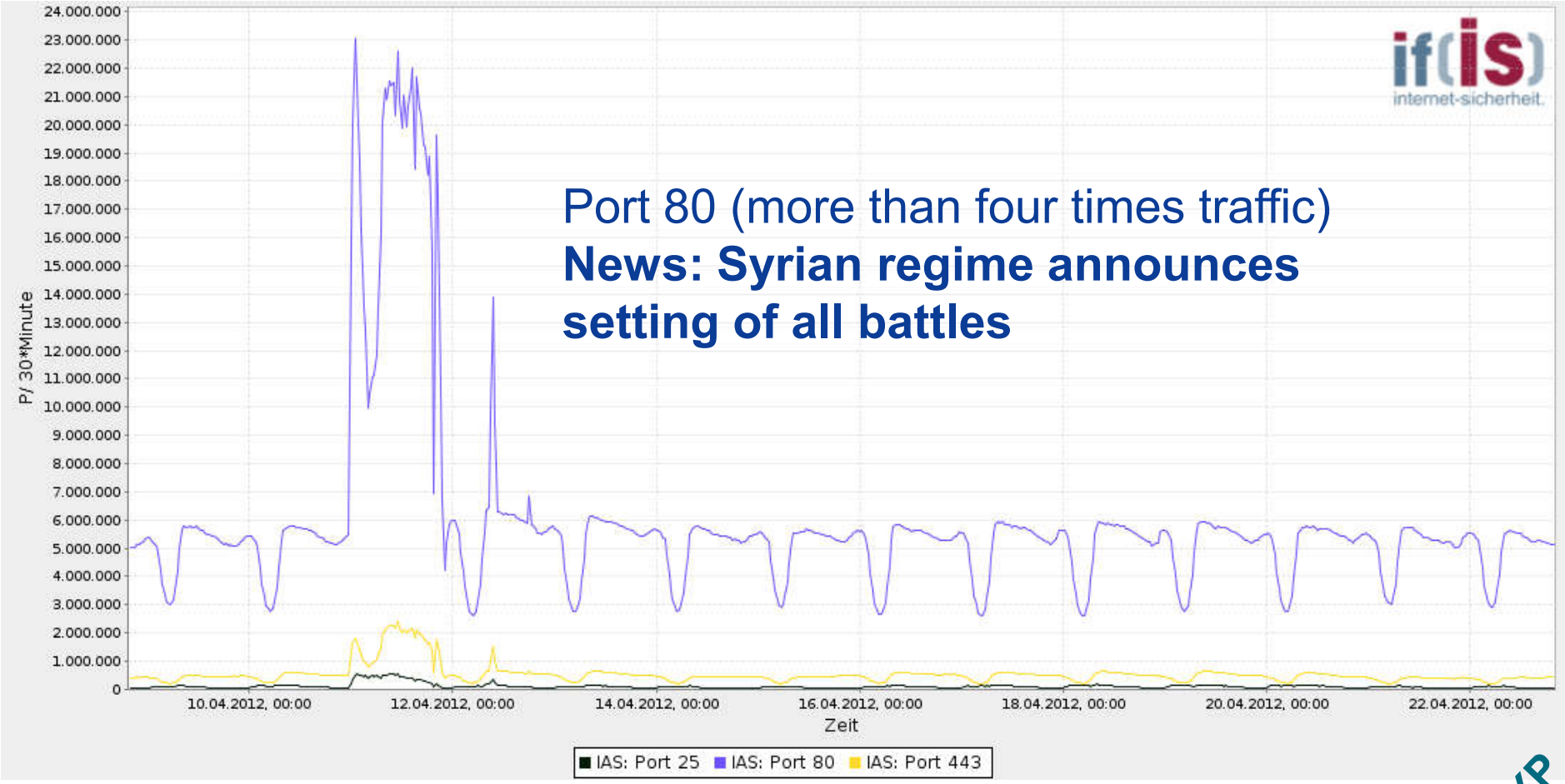
Coherences coming form behavior

- E.g. when we detect **http (port 80)**, then we also see **SMTP (port 25)**, this means, when we **surf** online we also **write e-mails**



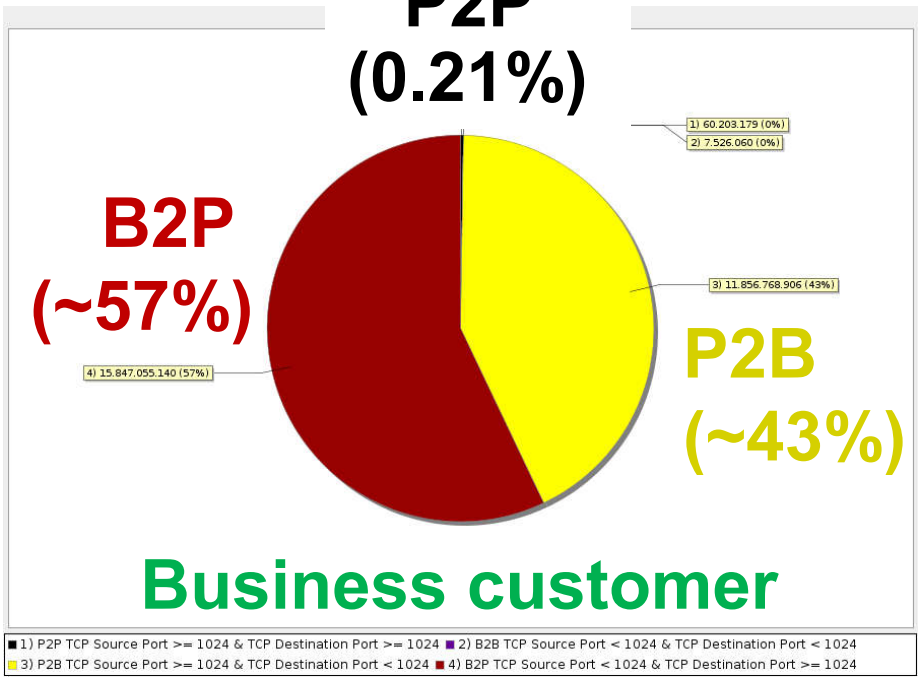
Knowledge Base - IAS
→ Coherences (3/4)

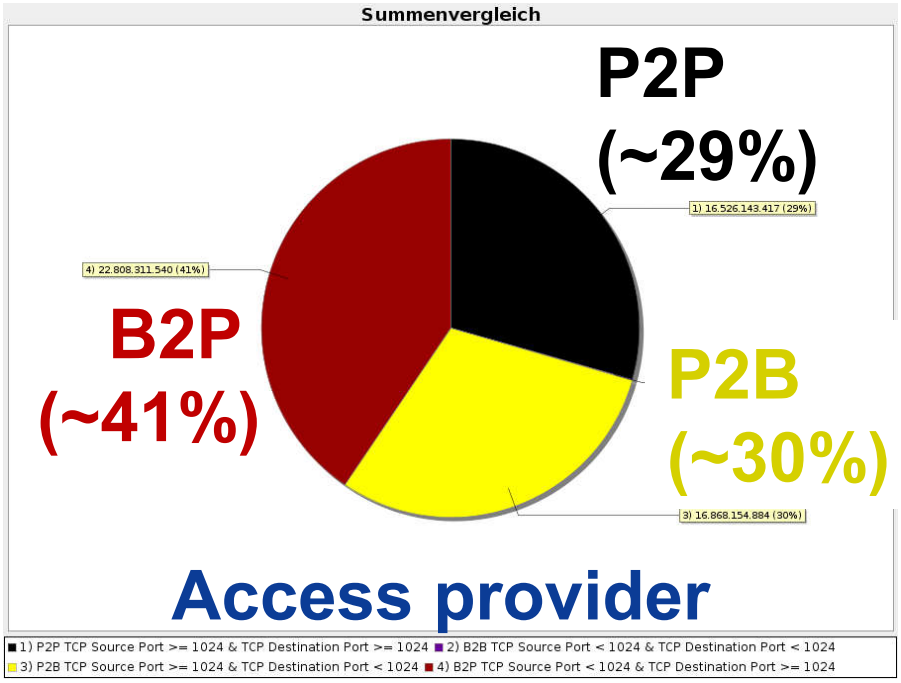
- **Coherences due to situations**
  - When a news with an **important impact** is broadcasted, e.g. an act of terror, then we can see a lot more **Internet traffic**



Knowledge Base - IAS
→ Coherences (4/4)

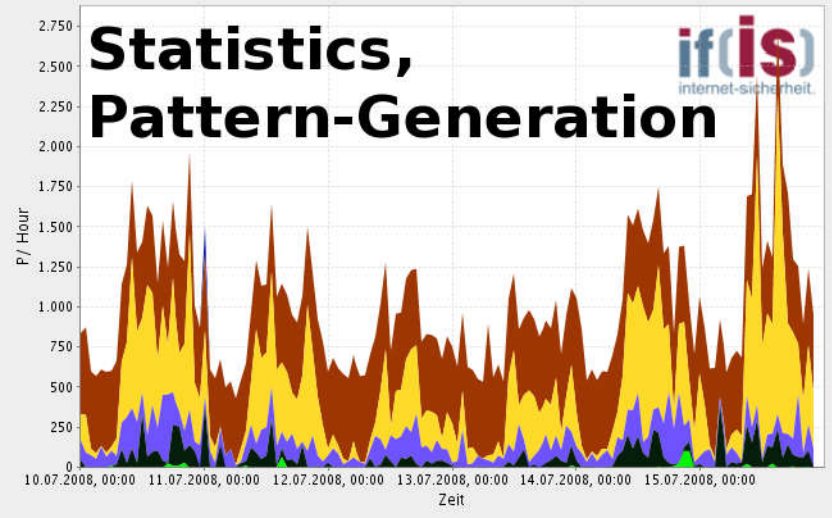
- **Coherences because of the location of the sensor and because of certain applications**
  - Access provider, content provider and business customer have very different coloured Internet traffic depending on the services and applications used
  - For instance we can detect a **lot more p2p traffic** in the network of a **Access provider** than in the network of a **Business Customer**





Internet Analysis System (IAS)
→ Results

- **Description of profiles, patterns and coherences**



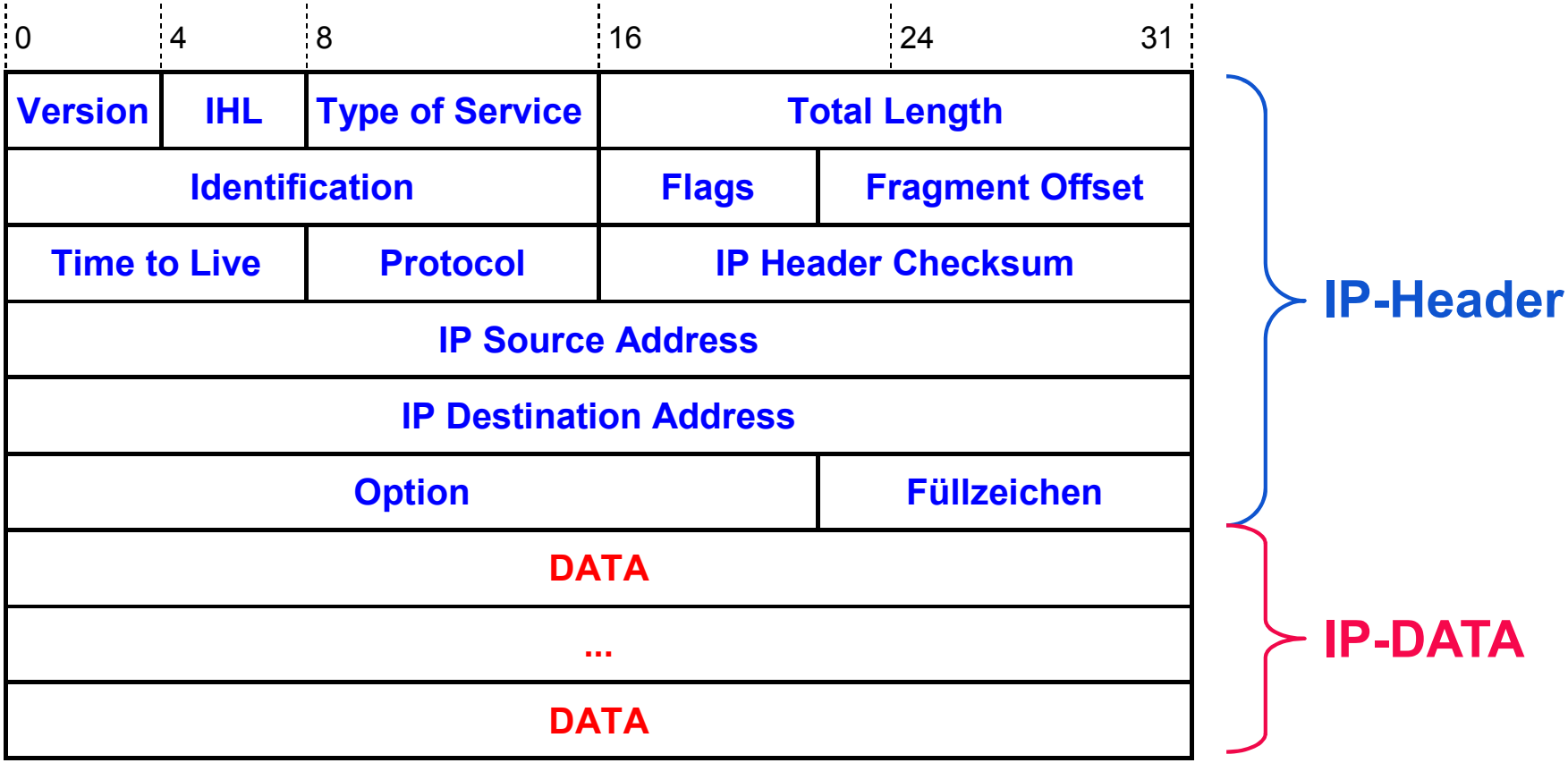
- **Creation of a knowledge base.**

sFlow
→ Important aspects

- The IAS is connected via sFlow
  - **Only every 16.384 packet is sent in form of a sFlow record**
    - **Of this package, the first 128 bytes transmitted as sFlow**

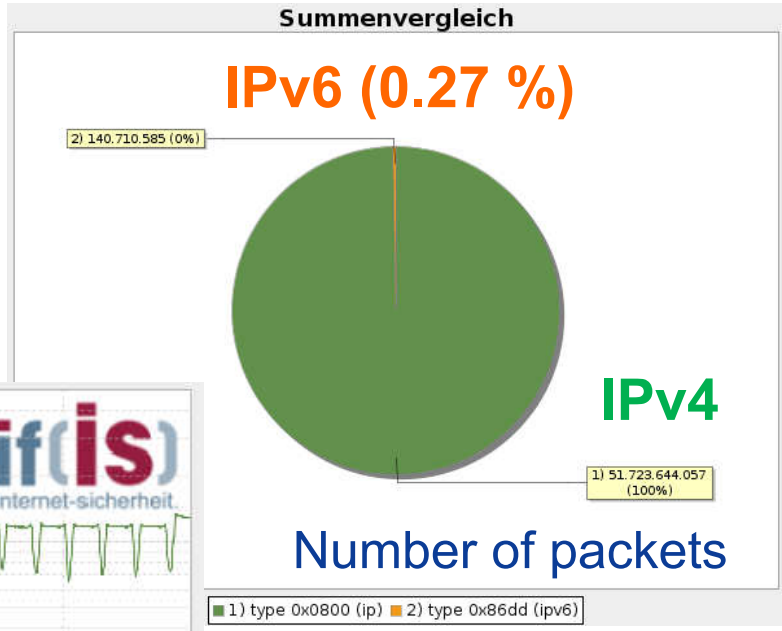
- Only the first 128 byte are available for the analysis.
  - 14 byte Ethernet Header + 4 byte (optional) VLAN
  - 20 byte IPv4 (without option) or 40 byte IPv6 Header (without extensions)
  - 20 byte TCP (without option) or 8 byte UDP Header
  - → **Best Case:** 128 byte – 42 byte = **86 bytes from layer 4**
  - → **Worst Case:** 128 byte – 78 byte = **50 bytes from layer 4**

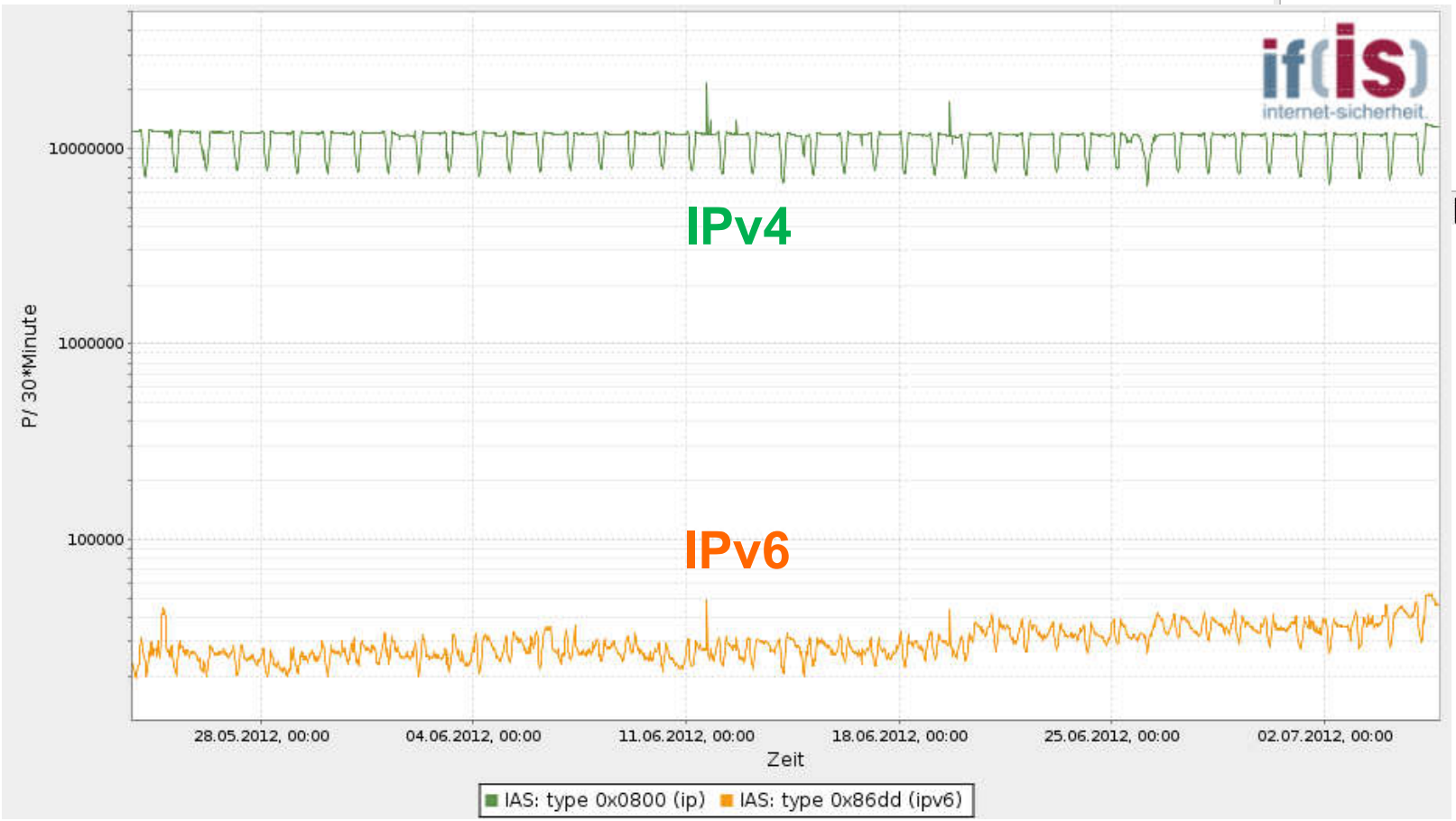
IP-Packet (Version 4)
→ Format



IP Header „Version “-field → IPv4 vs. IPv6

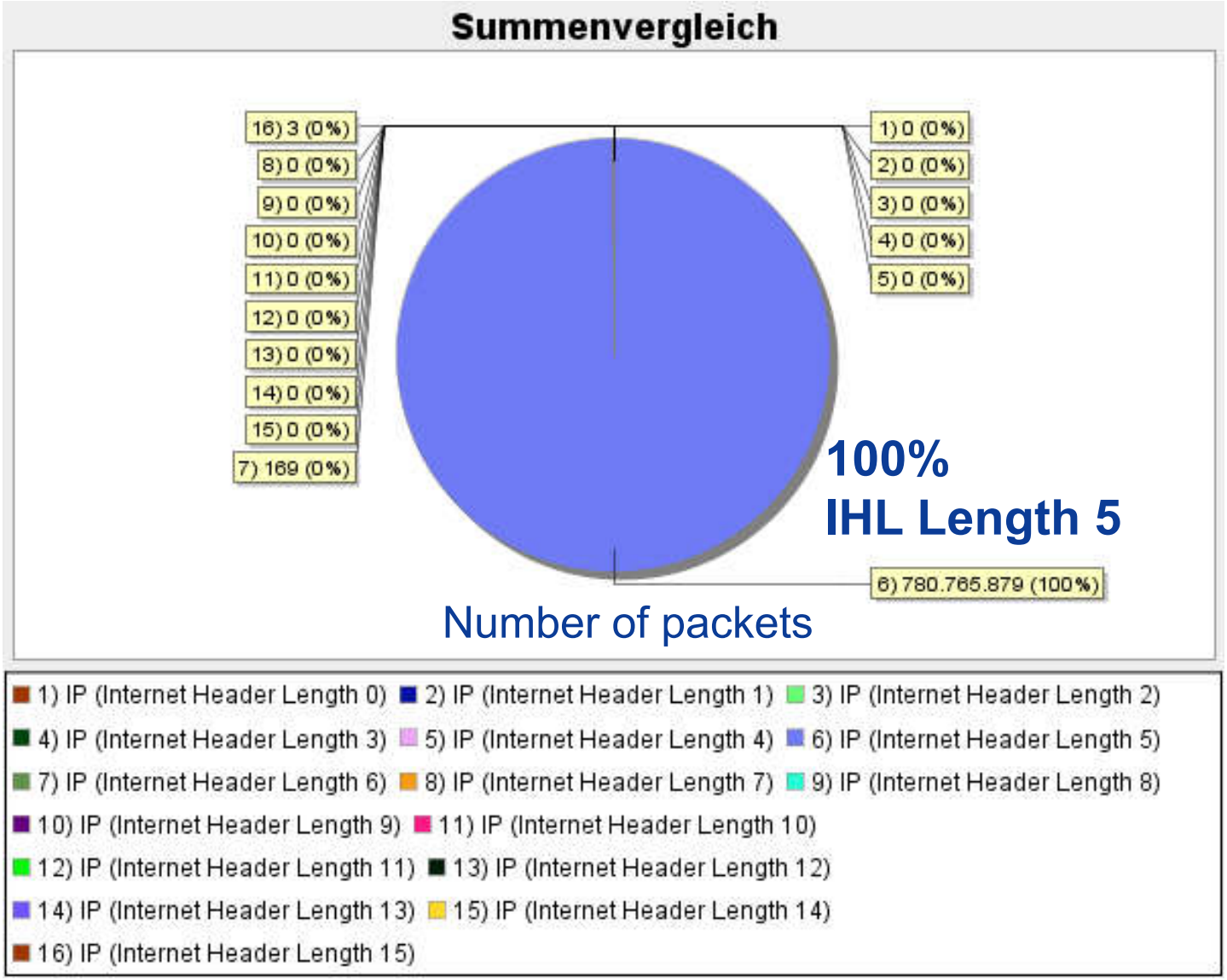
- Proportion between IPv4 and IPv6





IPv4 Header „IHL“-Feld

- How expected almost only a length of 5 for IHL (20 Byte)
- No “Option” via the Internet!



IPv4 Header „Time To Live"-field
→ Number of packets

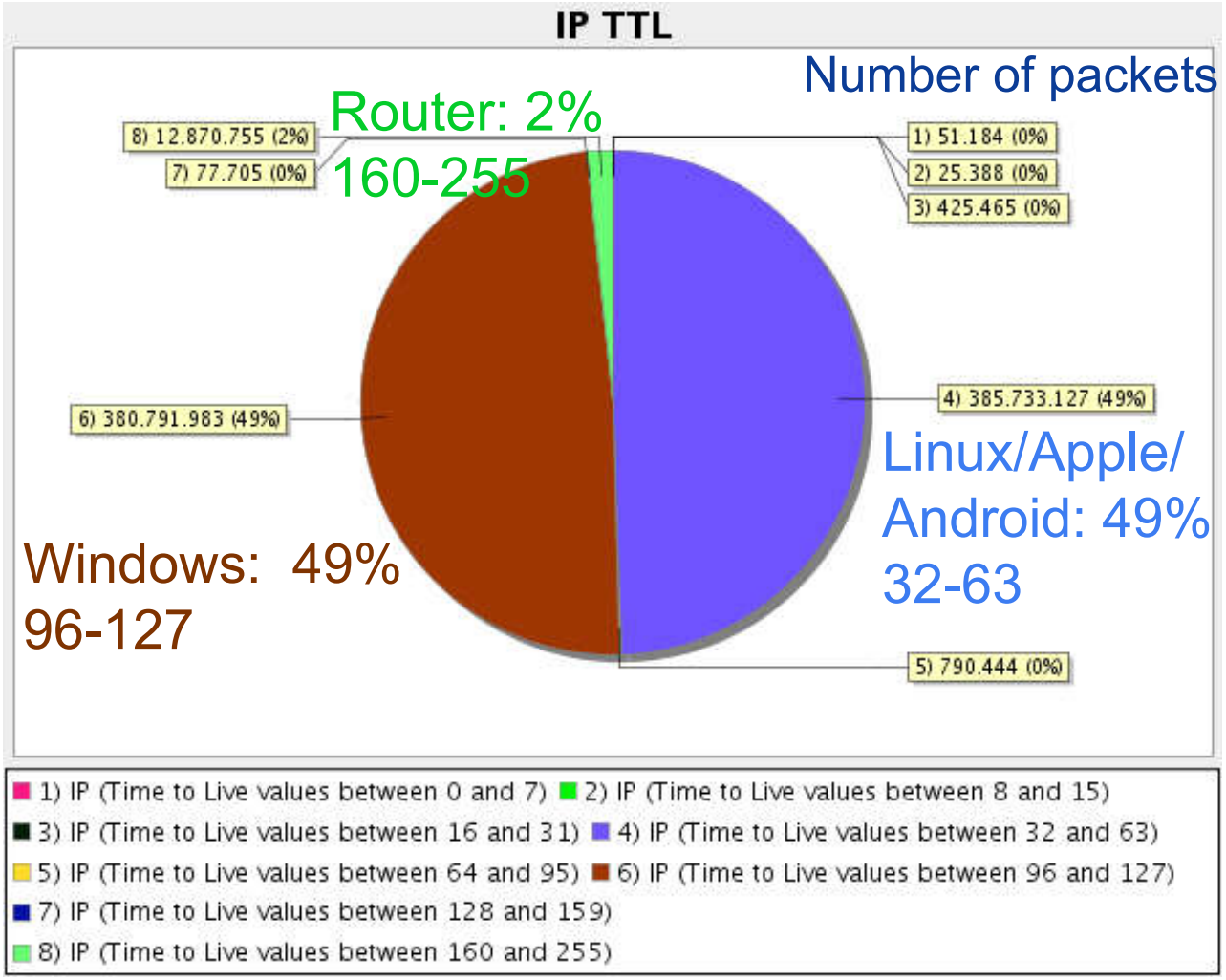
TTL allows conclusions on the used operating system.

Linux: TTL Default 64
Mac-/iOS: TTL Default 64
Android: TTL Default 64

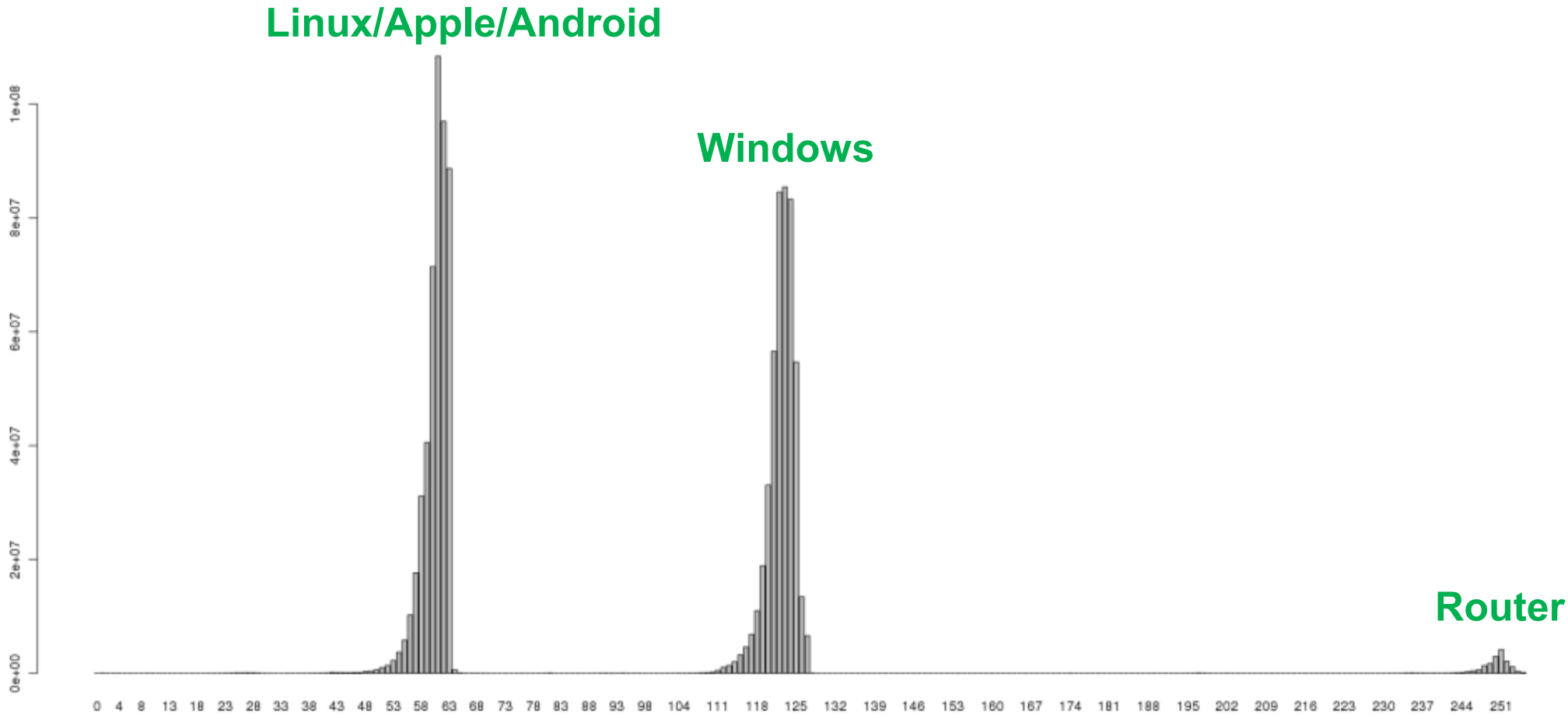
Windows : TTL Default 128

Router : TTL Default 254

Symbian: TTL Default 69



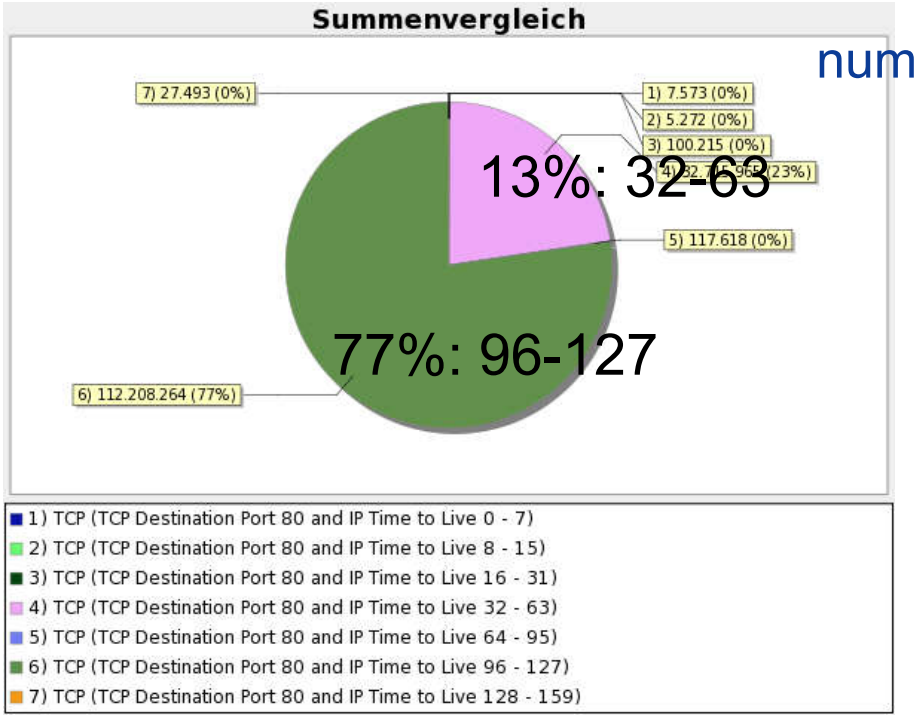
IPv4 Header „Time To Live"-field → Overview



- The distribution of TTL values is corresponding to the TTL characteristics of the operating systems

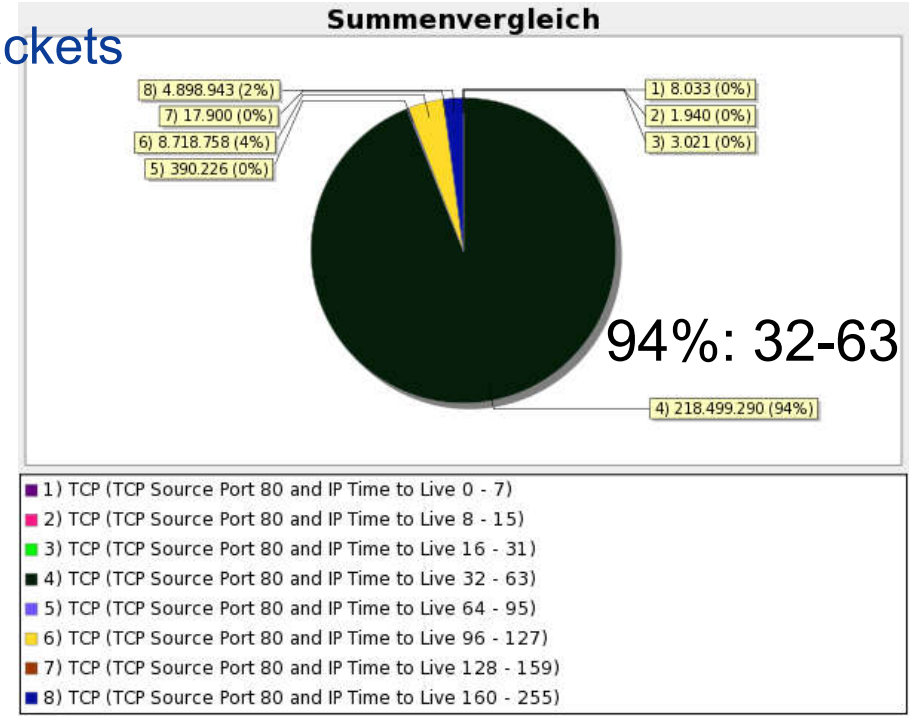
TTL in combination with HTTP
→ Port 80 (number of packets)

TCP Dst. Port 80 → Client



number of packets

TCP Src. Port 80 → Server



- **Client: 77 % Windows**
- **Client: 13 % Linux/Mac- & iOS/Android**

- **Server: 94 % Linux**
- **Server: 4 % Windows**

IPv4 Header „Time To Live"-field
→ Number of hops to the IXP

Linux **„Server"**

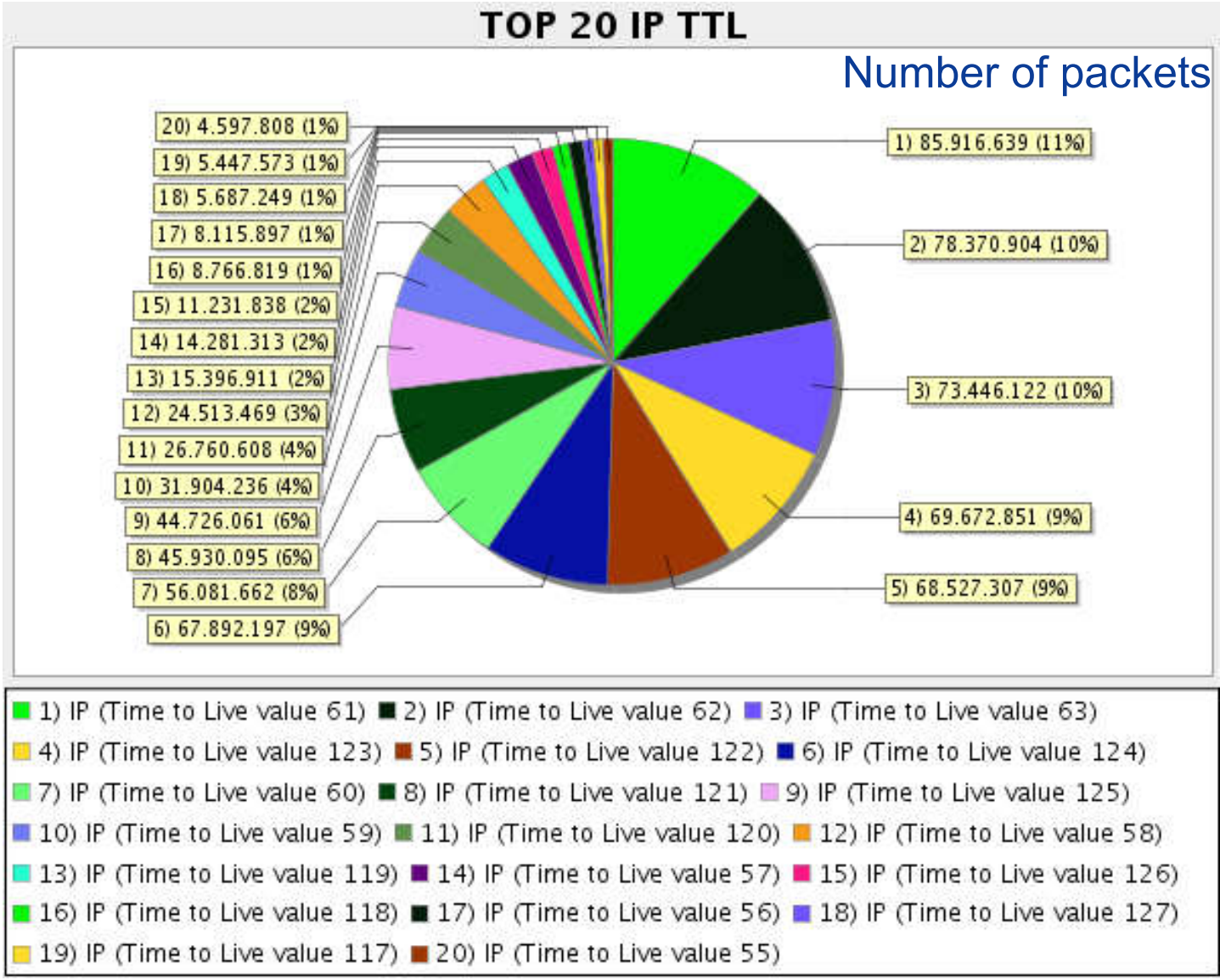
- 11% 3 hops
- 10 % 2 hops
- 10 % 1 hop
- 8 % 4 hops

distance to the IXP (average 2.5 hops)

Windows **„Client"**

- 9 % 5 hops
- 9 % 6 hops
- 9 % 4 hops
- 6 % 7 hops

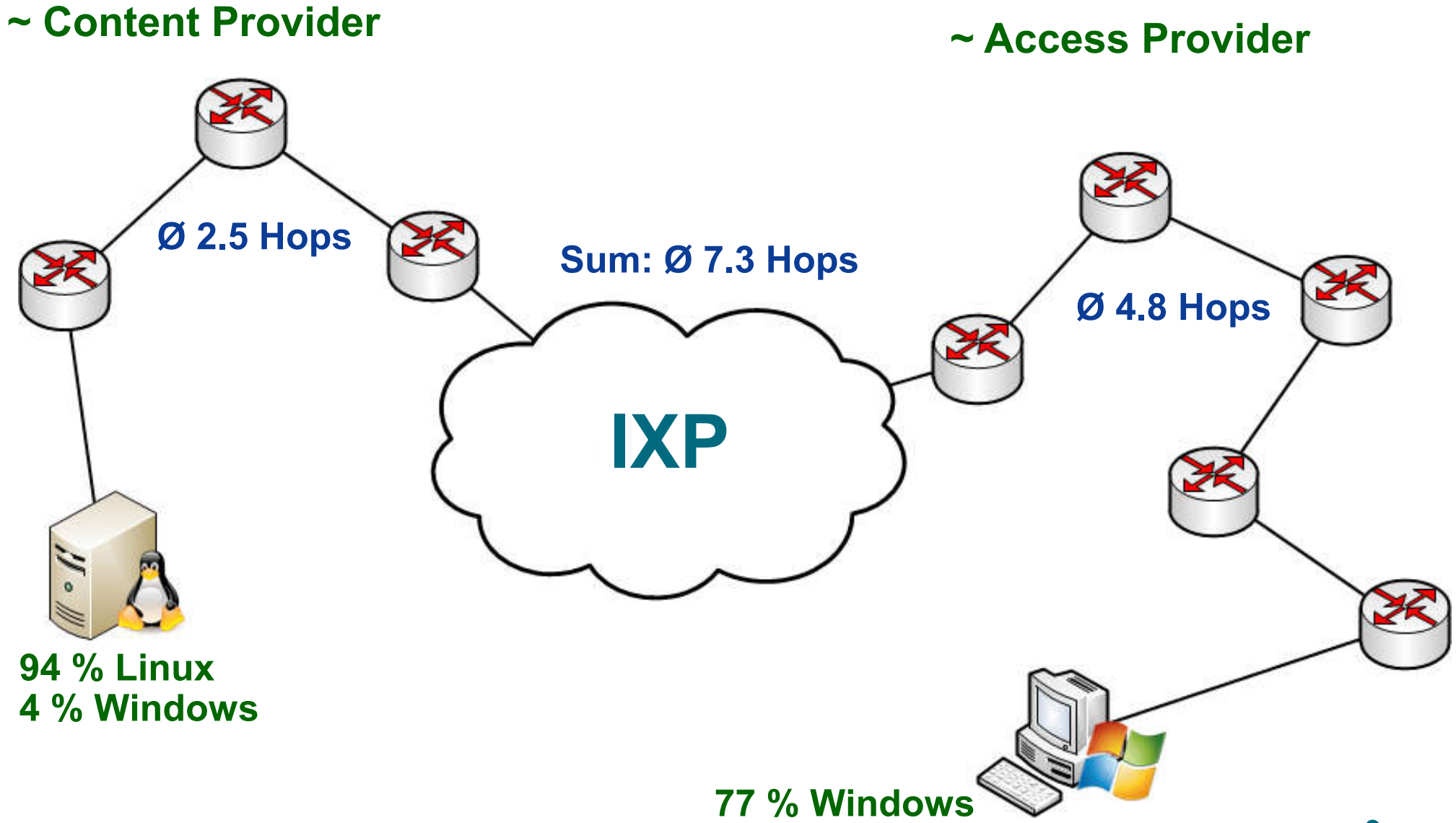
distance to the IXP (average 4.8 hops)



sum = ~ 7.3 hops

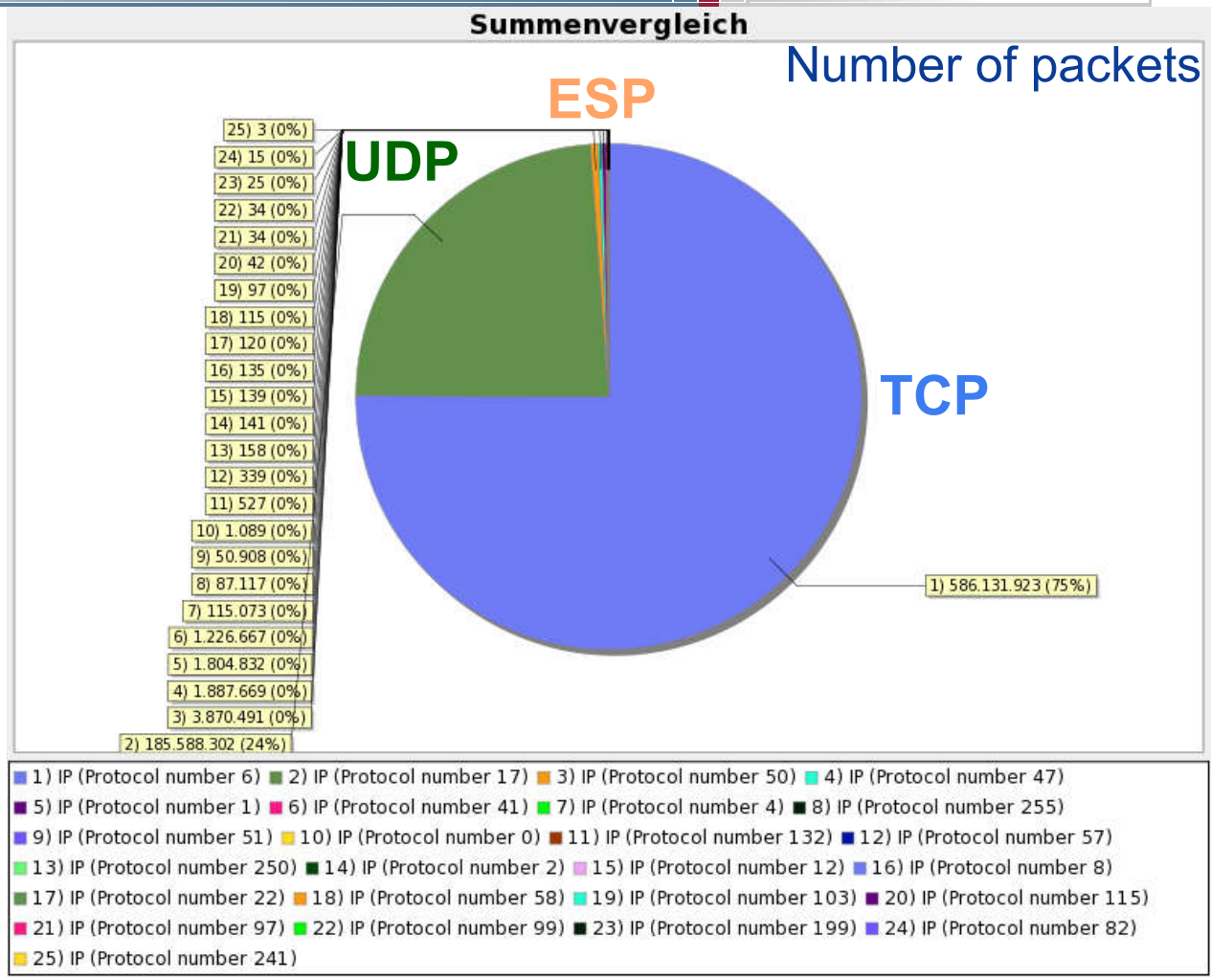
IXP

IPv4 Header „Time To Live"-field
→ Number of hops to the IXP



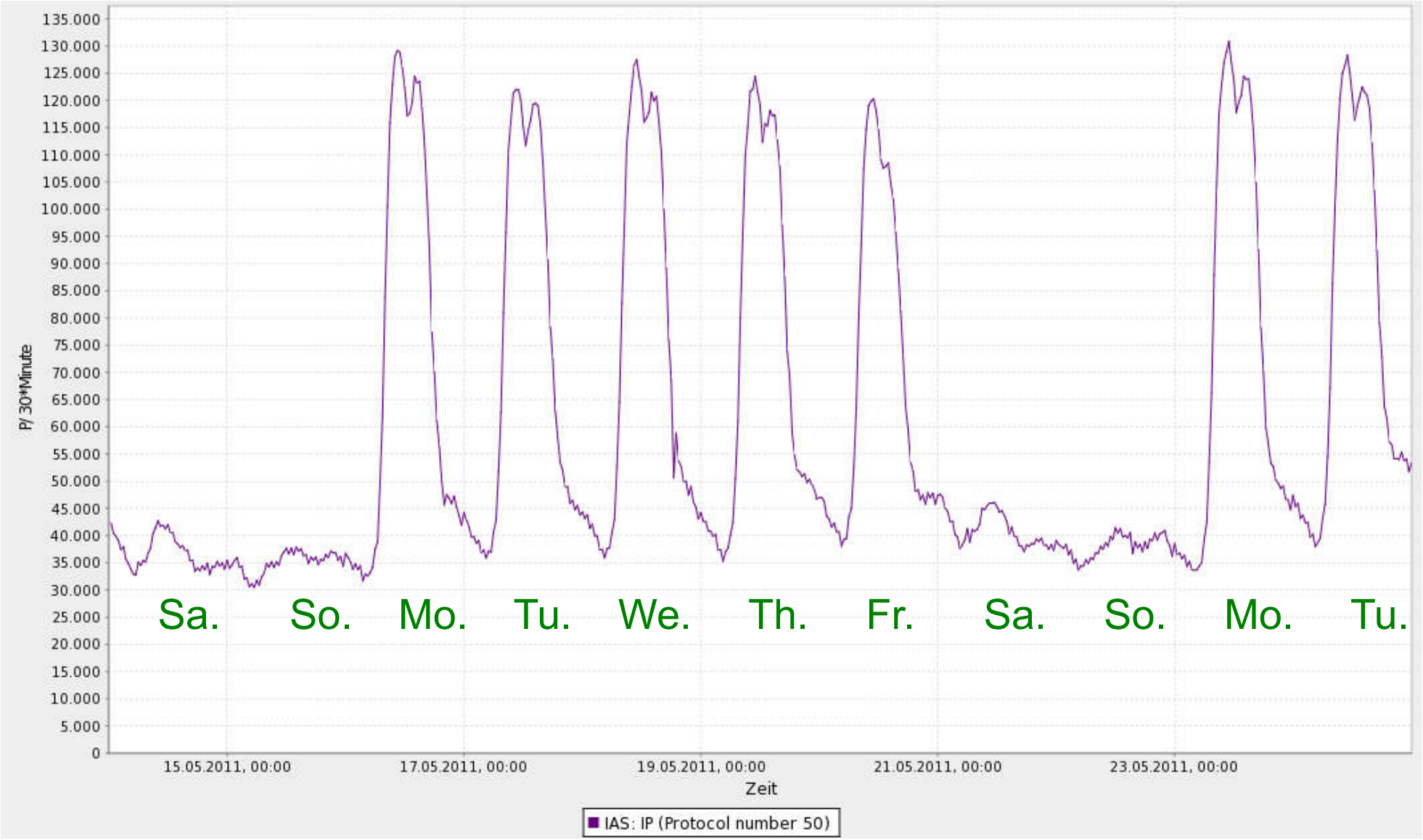
IPv4 Header „Protocol“-field → TOP25

- **75% TCP – (6)**
- **24% UDP – (17)**
- **<1% ESP (0,5%) – (50)**
- <1% GRE (0,24%) – (47)
- <1% ICMP (0,23%) – (1)
- <1% IPv6 Encapsulation - (41)
  (6over/to4 = 0,157%)
- <1% IPv4 Encapsulation
- <1% Reserved
- **<1% Authentication H. (0,007%)**
- <1% IPv6 Hop-by-Hop Option
- <1% SCTP
- <1% SKIP
- <1% Unassigned (199, 250, 241)
- <1% PUP
- <1% EGP
- <1% XNS-IDP
- <1% IPv6-ICMP
- <1% PIM



- <1% L2TMP
- <1% ETHERIP
- <1% any private encription scheme
- <1% SECURE-VMTP

„Protocol"-field 50 (IPSec → ESP) → User behavior

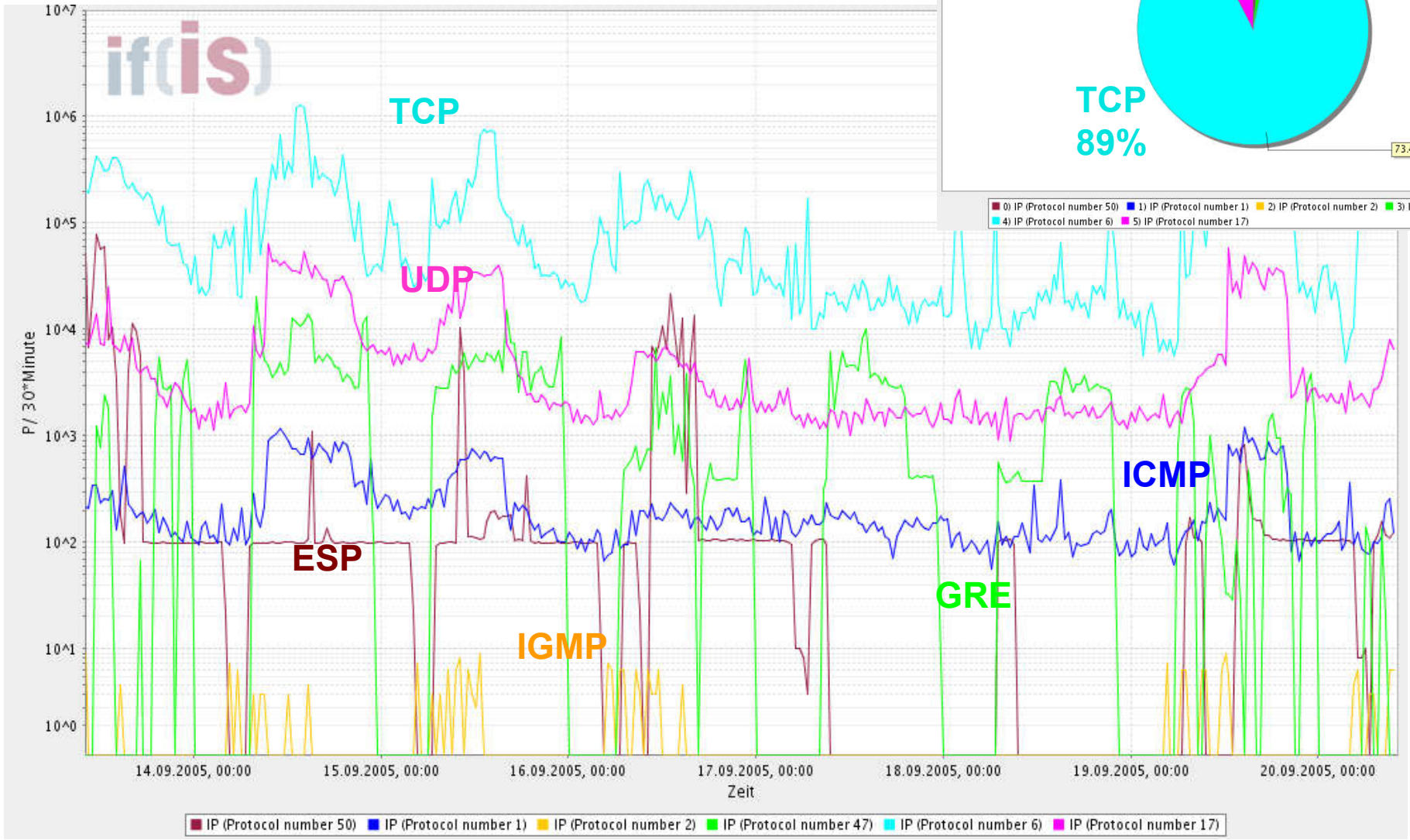


→ IPSec more common in the business environment

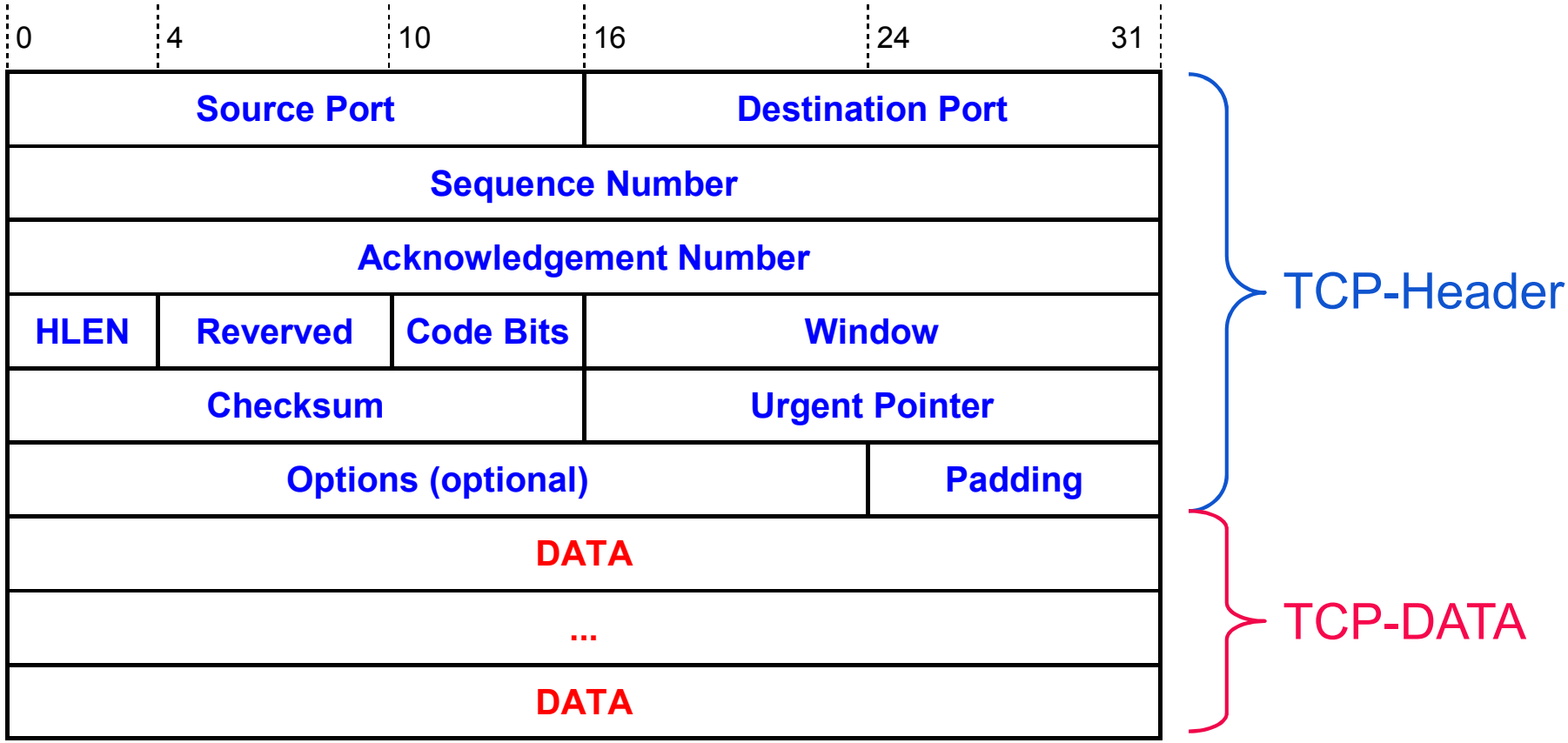
IPv4 Header „Protocol“-field
→ Result: Distribution transport protocol

- **Distribution of Transport Protocols (2005)**

  Profile shaping und trend development

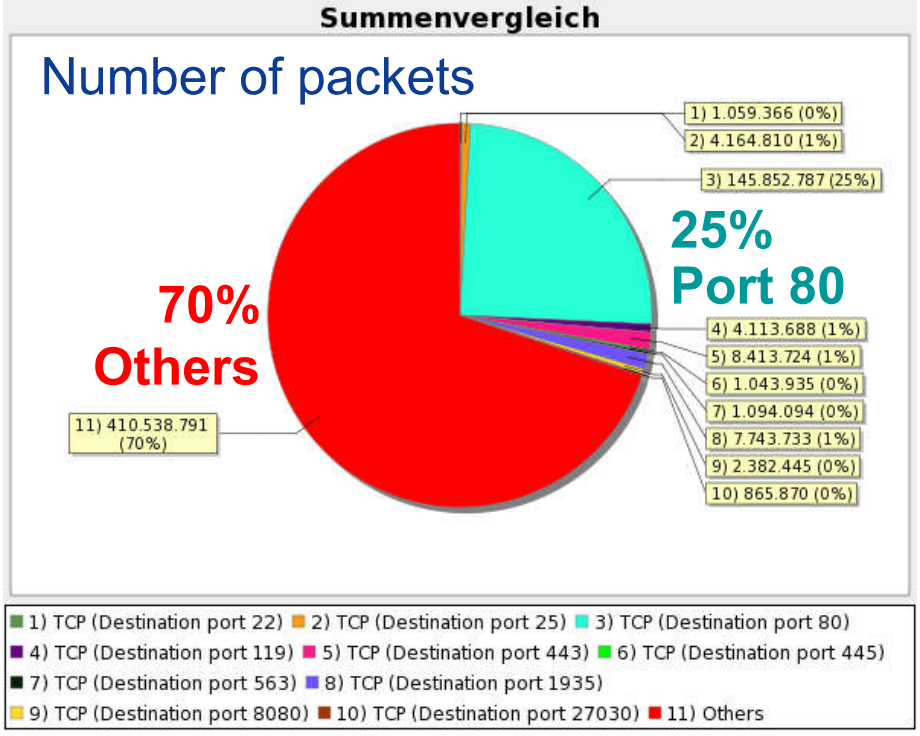


TCP - Transmission Control Protocol
→ Format



TCP Top10 Destination Ports (all)
(TCP ca. 75 % of Σ the IPv4 traffics)

1. 25% HTTP (80)
2. 1% HTTPS (443)
3. 1% RTMP (1935)
   (Flash Real Time Messaging Protocol)
4. 1% NNTP (119)
   (→ Usenet → Downloads)
5. 1% SMTP (25)
6. <1% HTTP (8080)
7. <1% NNTPS (563)
   (→ Usenet → Downloads)
8. <1% SSH (22)
9. <1% SMB (445)
   (bedenklich → Exploitversuche, direkt angeschlossene MS-Systeme?)
10. <1% Steam (27030) (→ Gaming)

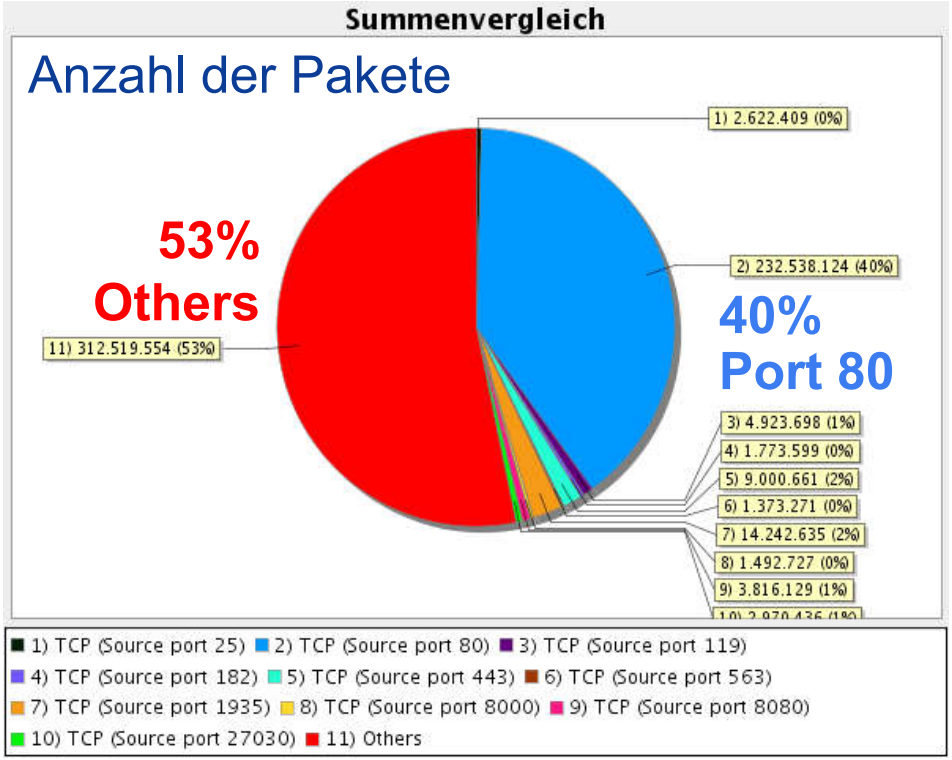


(Top10 = 30%)

<0.001% 445 SRC (SYN/ACK)
TCP connection (2.277)
IXP

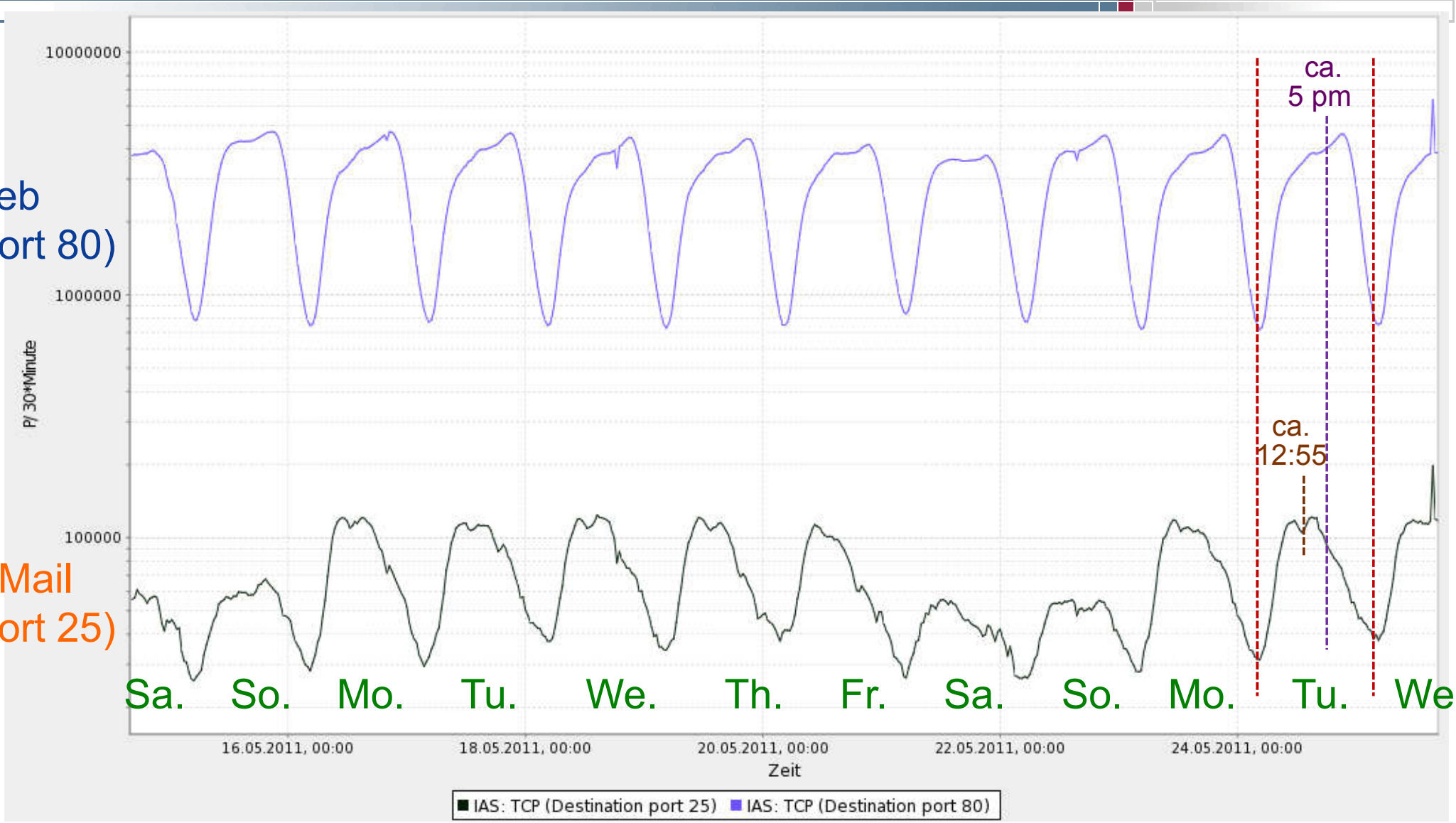
TCP Top10 Source Ports (all)
(TCP ca. 75 % of Σ the IPv4 traffics)

		DST/SRC
1.	40% HTTP (80)	*1 zu 1.6*
2.	2% RTMP (1935) (Flash Real Time Messaging Protocol)	*1 zu 1.8*
3.	2% HTTPS (443)	*1 zu 1.1*
4.	1% NNTP (119) (→ Usenet → Downloads)	*1 zu 1.2*
5.	1% HTTP (8080)	*1 zu 1.6*
6.	1% Steam (27030) (→ Gaming)	*1 zu 3.4*
7.	<1% SMTP (25)	*1.6 zu 1*
8.	<1% Unisys Audit SITP (182)	
9.	<1% HTTP (8000)	
10.	<1% NNTPS (563) (→ Usenet → Downloads)	*1 zu 1.3*



(Top10 = 47%)

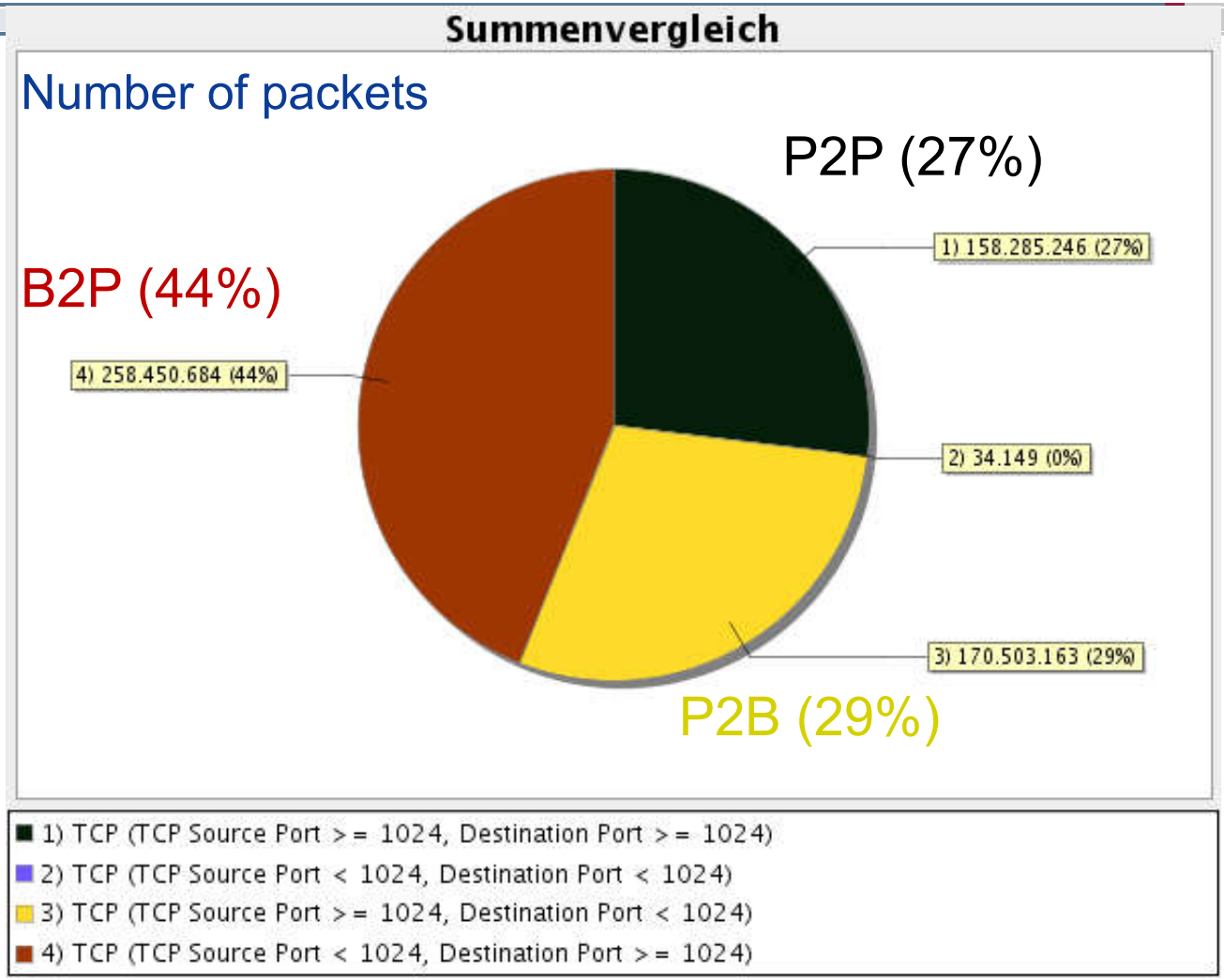
Port 25 (E-Mail) – Port 80 (Web)
→ User behavior



E-Mails take off for evening / much less on the weekend

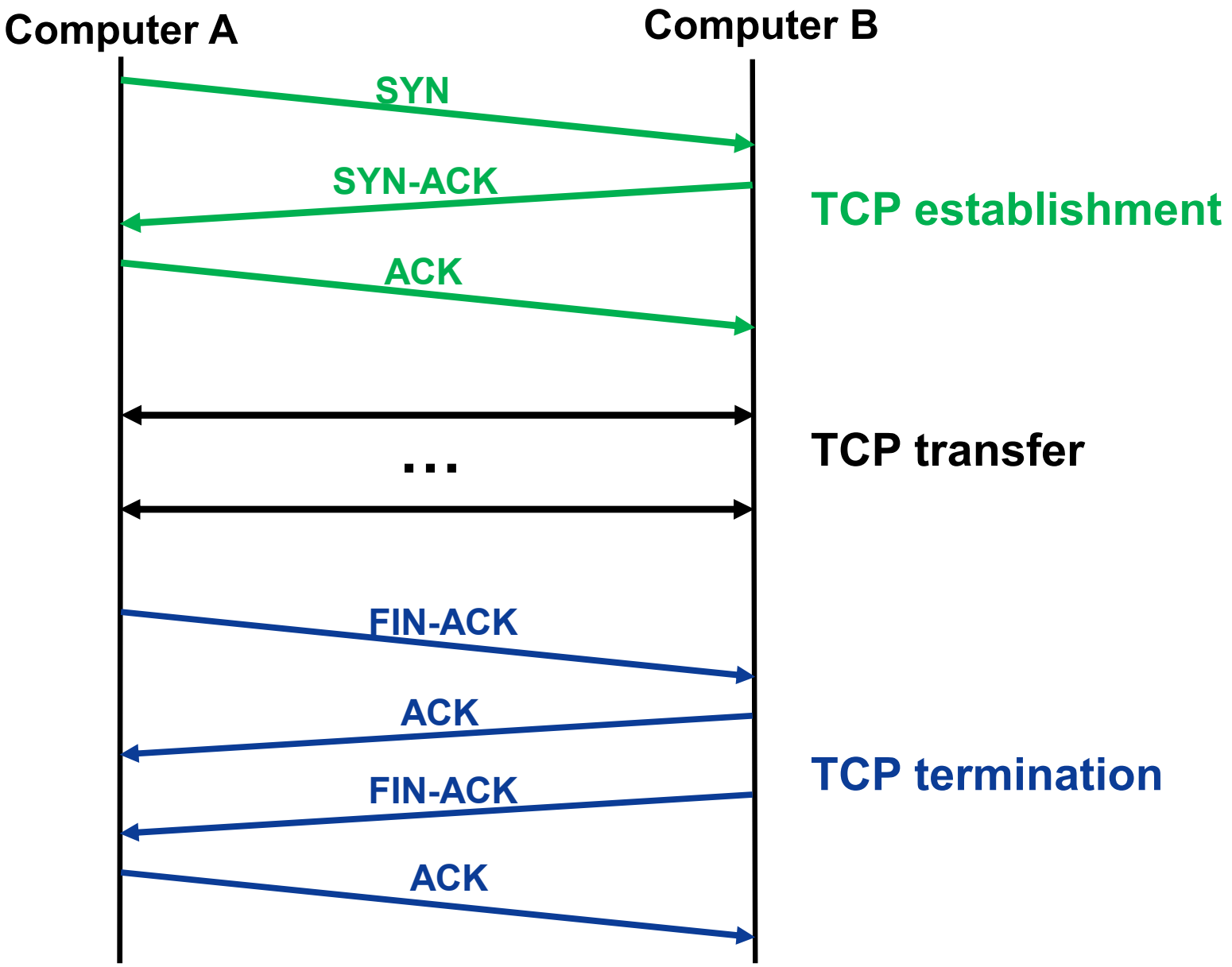
Web (2.0) rise to the evening / remain the same at the weekend

P2P Counter
→ Heuristic



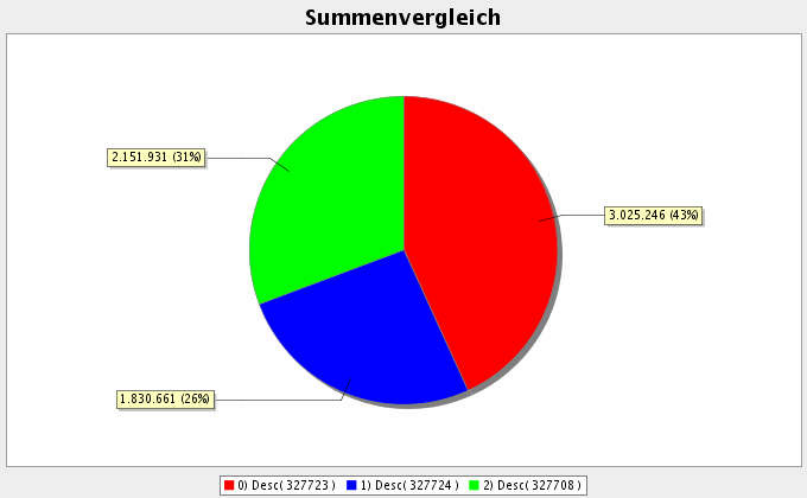
- Src >= 1024 and Dst >= 1024 (« P2P ») - client-to-client
- Src < 1024 and Dst < 1024 **(« B2B »)** - server-to-server
- Src >= 1024 and Dst < 1024 **(« P2B »)**
- Src < 1024 and Dst >= 1024 **(« B2P »)**

Analysis TCP
→ TCP communication: Overview

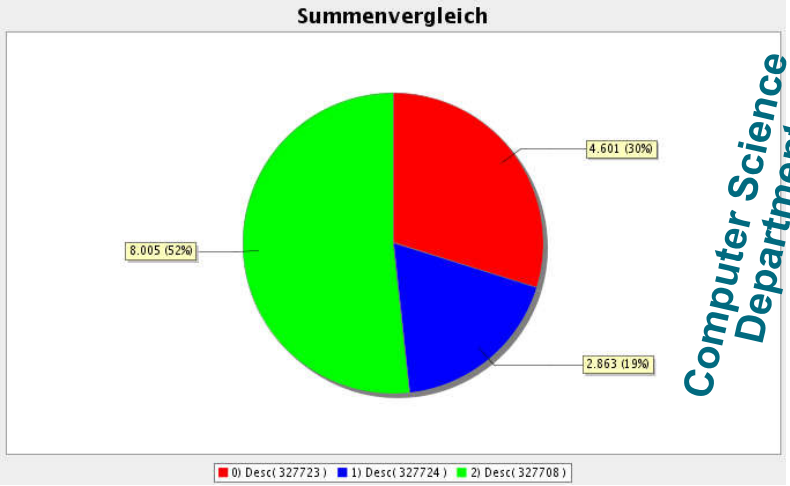


TCP Header „Code Bits“-field
→ Result: Expected distribution

- **SYN-Scan (Potential Attack)**
  - Comparison between different periods
    - Expected: SYN > SYN/ACK > 2xFIN/ACK (TCP teardown handshake)
  - Gap between expected spreading and spreading in case of an attack → Detection of attacks



SYN
(31% - 52%)
SYN/ACK
(26% - 19%)
FIN/ACK
(43% - 30%)

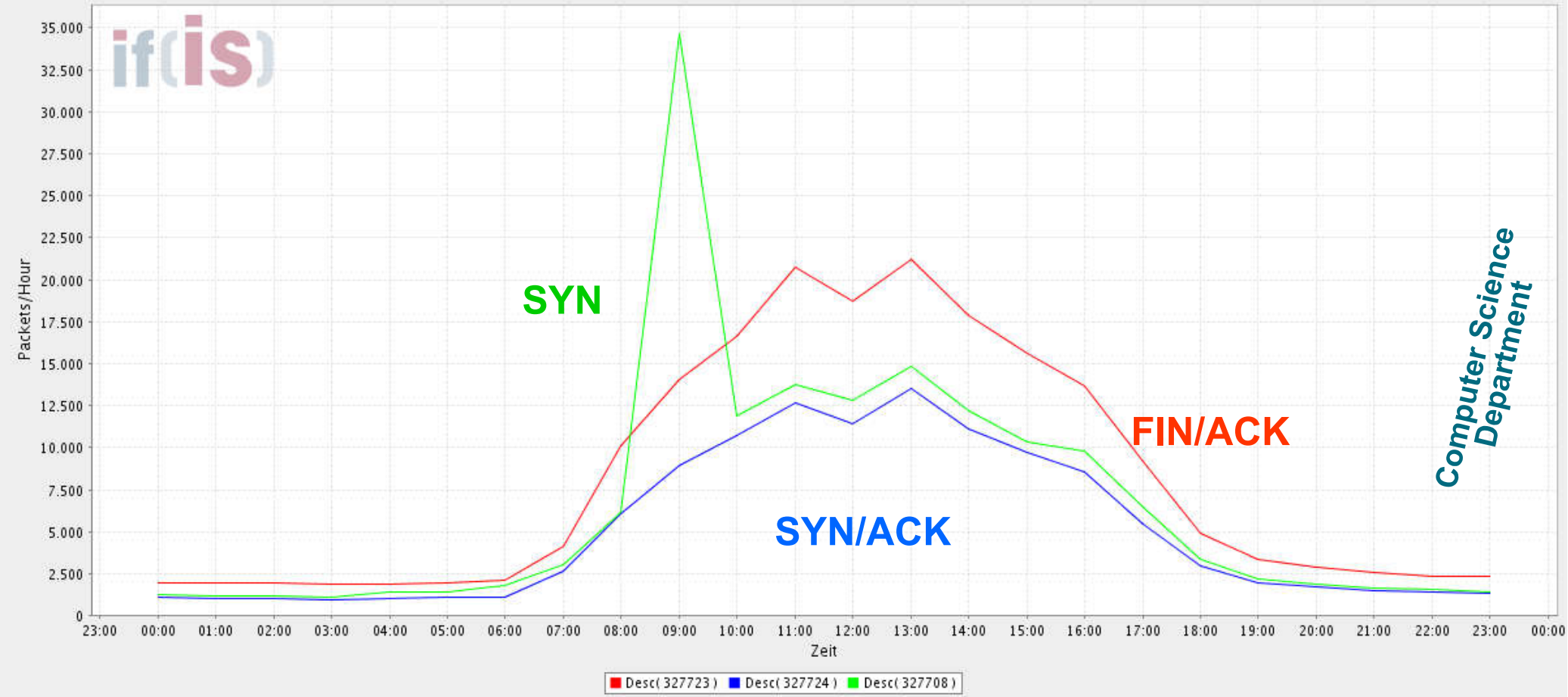


Expected Distribution

Unexpected Distribution

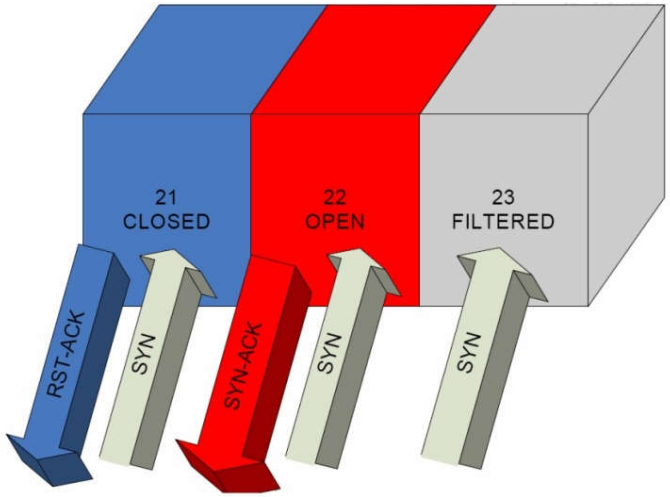
TCP Header „Code Bits“-field
→ Result: Detection of attacks

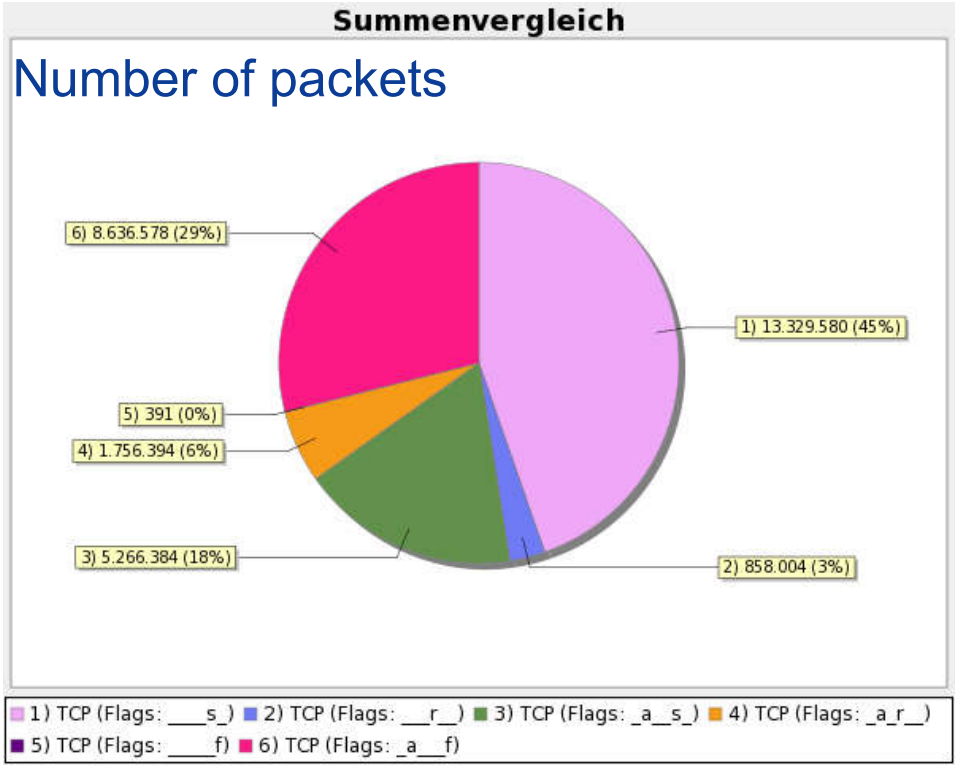
- **SYN-Scan (Potential Attack)**
  - Period of SYN scan can easily be detected



TCP-Header „Code Bits“-field
→ Establishment and termination

- 45% SYN → *45-18= 27%*
  (Scan (FILTERED), (CLOSED))
  7.997.748 packets - scan (1.02% der Σ IP)
- 29% FIN ACK (15% TCP termination)
- **18% SYN ACK (OPEN)**
  (→ 18% TCP establishment)
- 9% (ACK) RST (reset – termination),
  Scan (OPEN), (CLOSED))
- <1% FIN





Analysis TCP
→ TCP connection: calculation

- 586,131,923 TCP packets total
- **5,266,384 SYN-ACK packets total**
  → Indicator for succesful TCP connection
  - **4,326,855 (82,16%) well-known Ports**
    - 3,805,706 - Port 80 (HTTP)
    - 234,428 - Port 443 (HTTPS)
    - 170,622 - Port 25 (SMTP)
    - 30,524 - Port 110 (POP3)
    - 806 - Port 119 (NNTP)
    - 611 - Port 179 (BGP)

    7 times for TCP establishment and termination packets = 36,864,680
    + 15,085,982 (ACK-RST/+SYN)
- 534,181,261 TCP data transfer packets

- **101.43 TCP packets per TCP connection**

Analysis TCP
→ TCP-connection: Port 80 / Port 25

Port 80 (HTTP)

- 378,390,911 TCP packets (Port 80, DST+SRC)
- **3,805,706 SYN-ACK packets (Port 80)**
  7 times TCP establishment and termination packets = 26,639,942
  + 10,861,907 (ACK-RST/+SYN – 72% share)
- 340,889,061 TCP data transfer packets
- **89.57 TCP packets (Port 80) per TCP connection**

Port 25 (SMTP)

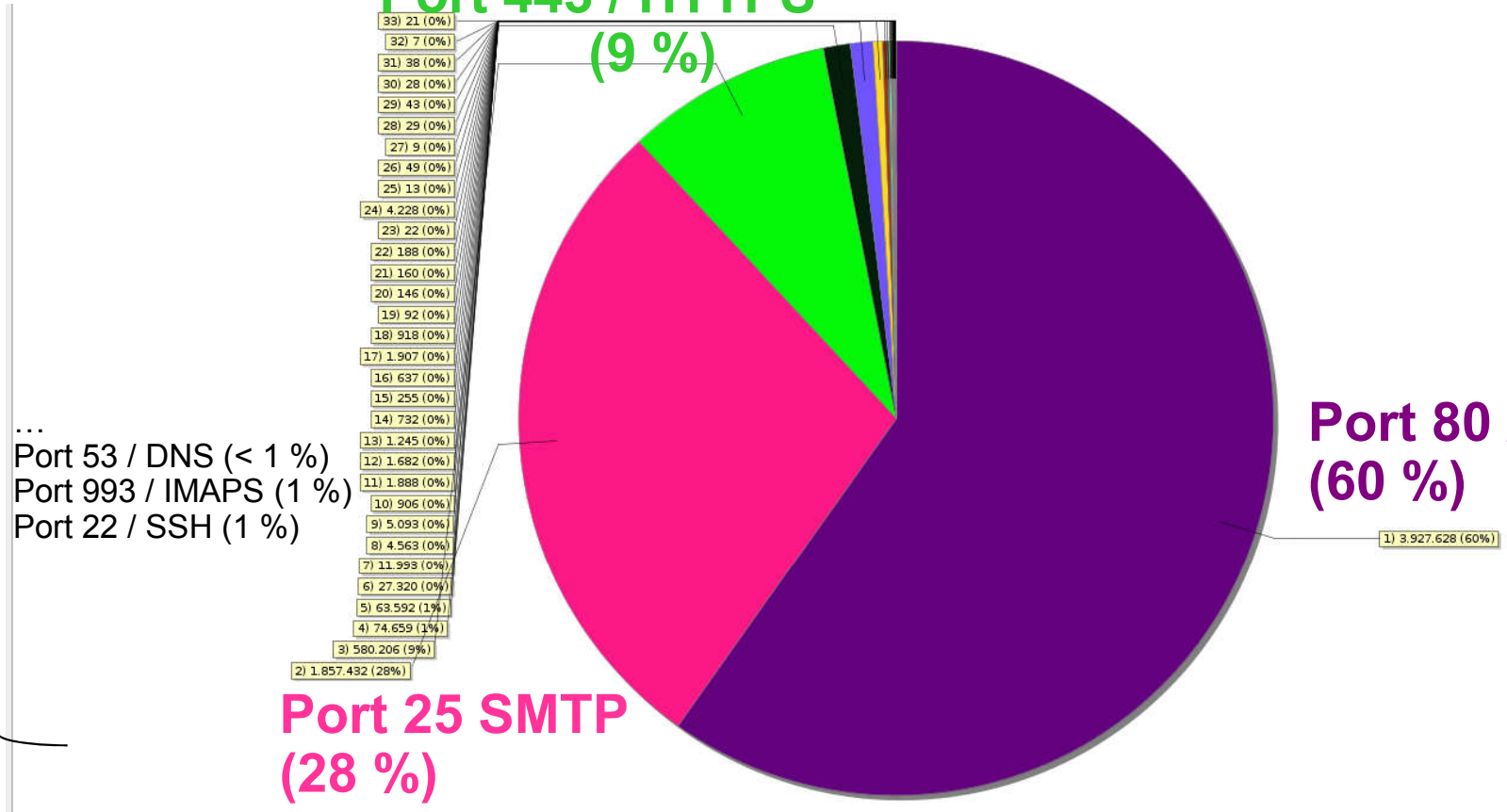
- 6,787,19 TCP packets (Port 25, DST+SRC)
- **170,622 SYN-ACK packets (Port 25)**
  7 times TCP establishment and termination packets = 1,194,354
  + 270,590 (ACK-RST/+SYN – 1.8% share)
- 5,322,275 TCP data transfer packets
- **31.19 TCP packets (Port 25) per TCP connection**

Note: The other applications are "streaming" more

TCP-Header „Code Bits“-field
→ TCP-connection: SYN ACK (TopX)

only
33 ports

…
Port 53 / DNS (< 1 %)
Port 993 / IMAPS (1 %)
Port 22 / SSH (1 %)



Computer Science Department

Data rate
→ Analysis and Overview

541,07 Gbyte/day (all IP-Packets, TCP + UDP and remaining) **= 100 %**

320,99 Gbyte/day (Port 80 - TCP) **= 59,33 % of total** (48% Σ Packets)
26,11 Gbyte/day (DST-Port 80)
293,15 Gbyte/day (SRC-Port 80)
→ DST/SRC = 0,089 (*DSL capable*)

10,25 GByte/day (Port 443 - TCP) **= 1,89 % of total** (2.2% Σ Packets)
2,64 Gbyte/day (DST-Port 443)
7,61 Gbyte/day (SRC-Port 443)
→ DST/SRC = 0,347

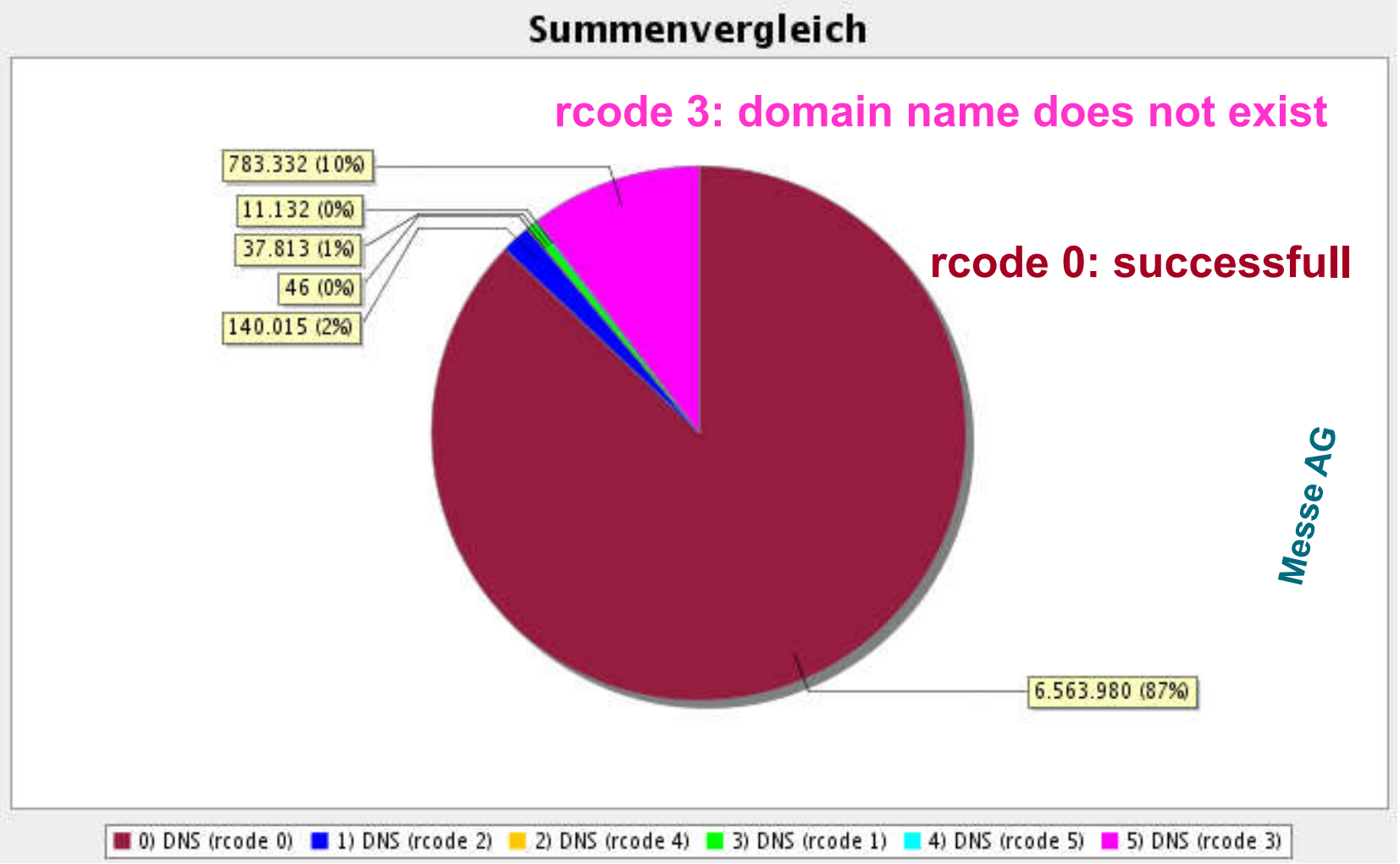
7,63 Gbyte/day (Port 119 - TCP) **= 1,41 % of total** (1,16% Σ Packets)
2,09 Gbyte/day (DST-Port 119)
5,55 Gbyte/day (SRC-Port 119)
→ DST/SRC = 0,377

3,83 Gbyte/day (Port 25 - TCP) **= 0,64 % of total** (0,86% Σ Packets)
3,468 Gbyte/day (DST-Port 25)
0,365 Gbyte/day (SRC-Port 25)
→ DST/SRC = 9,5

198,37 Gbyte/day (Rest TCP + UDP and remaining) **= ca. 36,66 %**

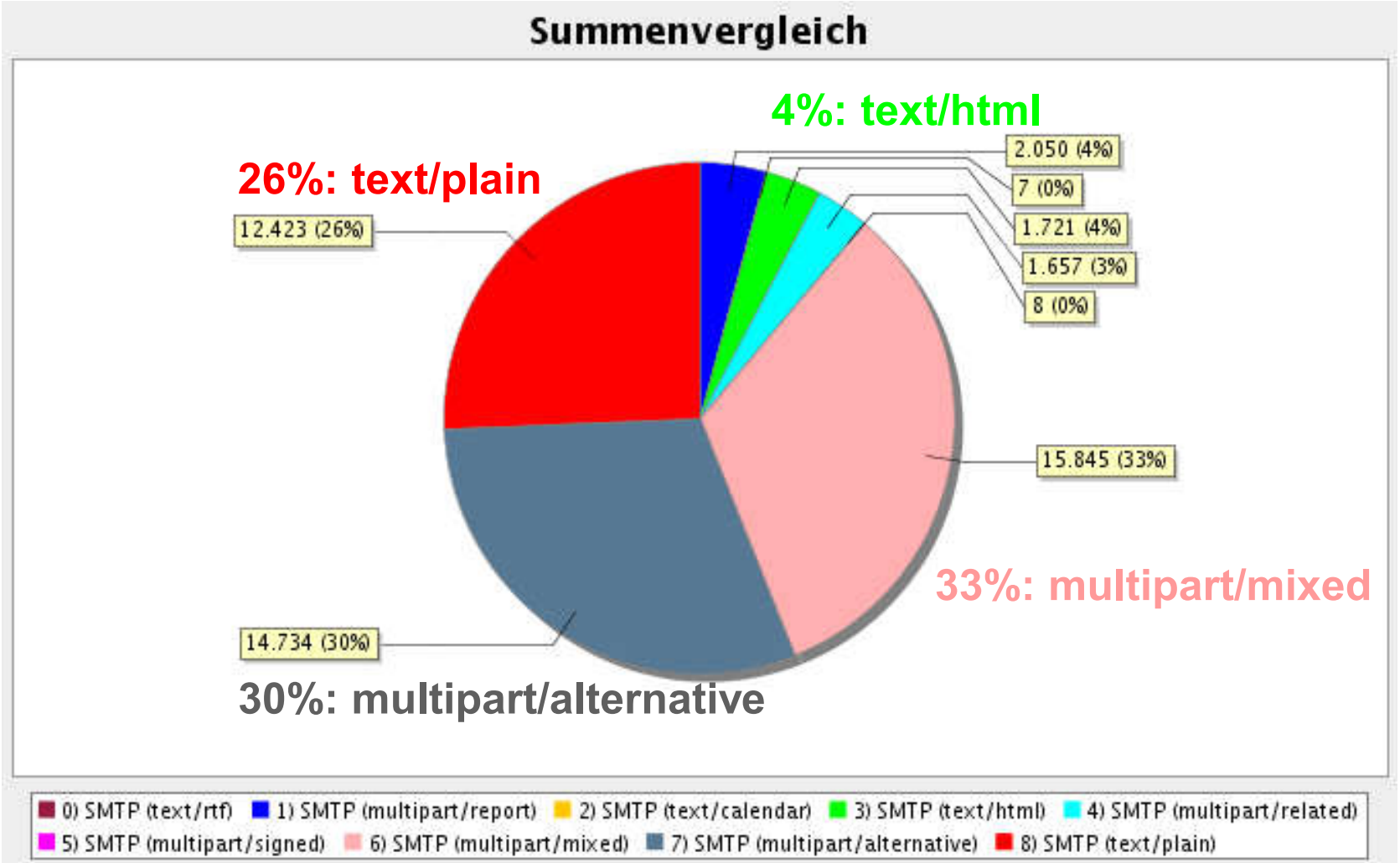
DNS Header
→ Result: Examples

- **DNS Server Return Codes**
  - Normal distribution: Everything Ok
  - About 10%: Domain name not found



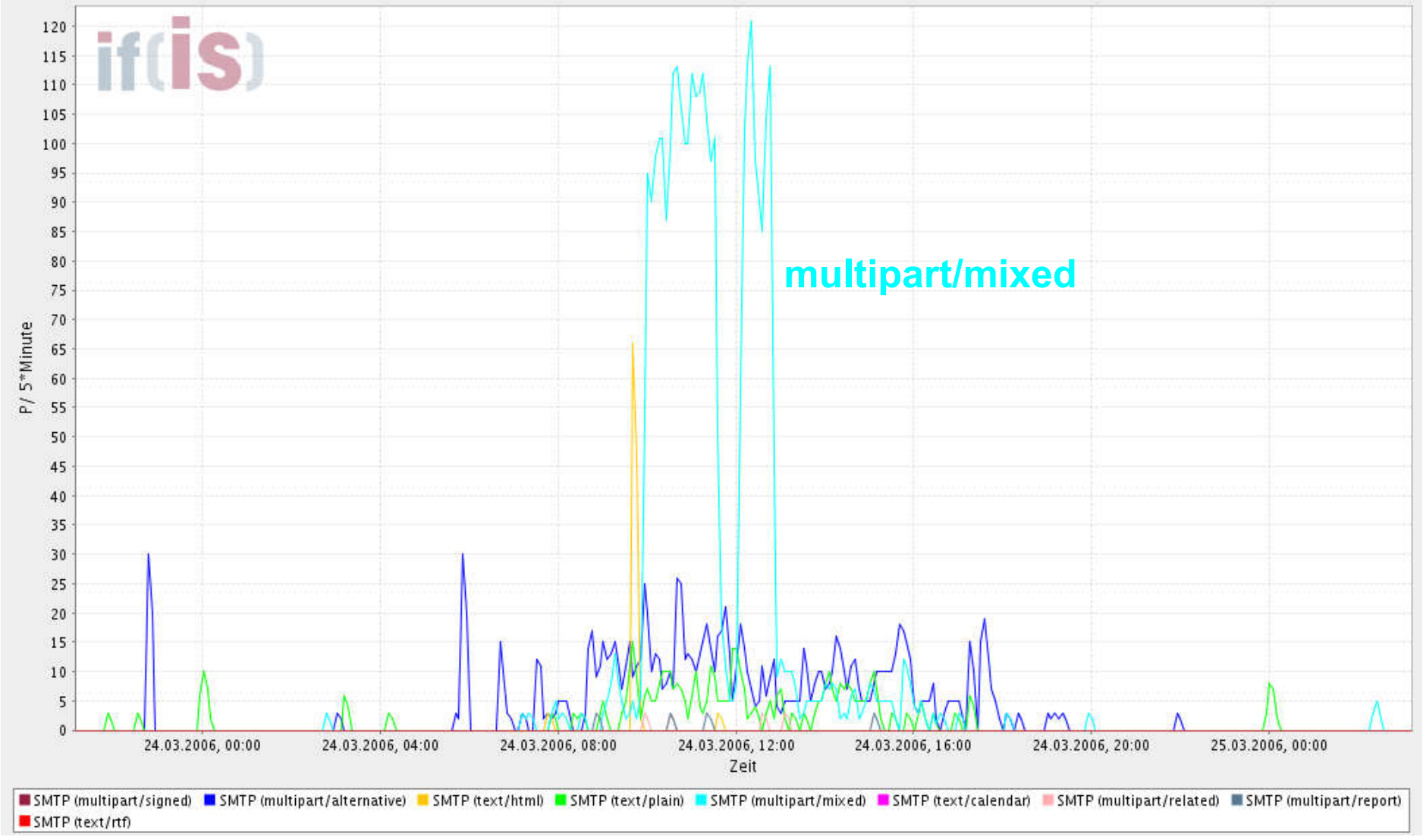
SMTP Header "MIME-Typ"
→ Result: Distribution SMTP Content Type

- **SMTP Content Type**
  - 60% "text" Mails
  - 33 % "attachments"



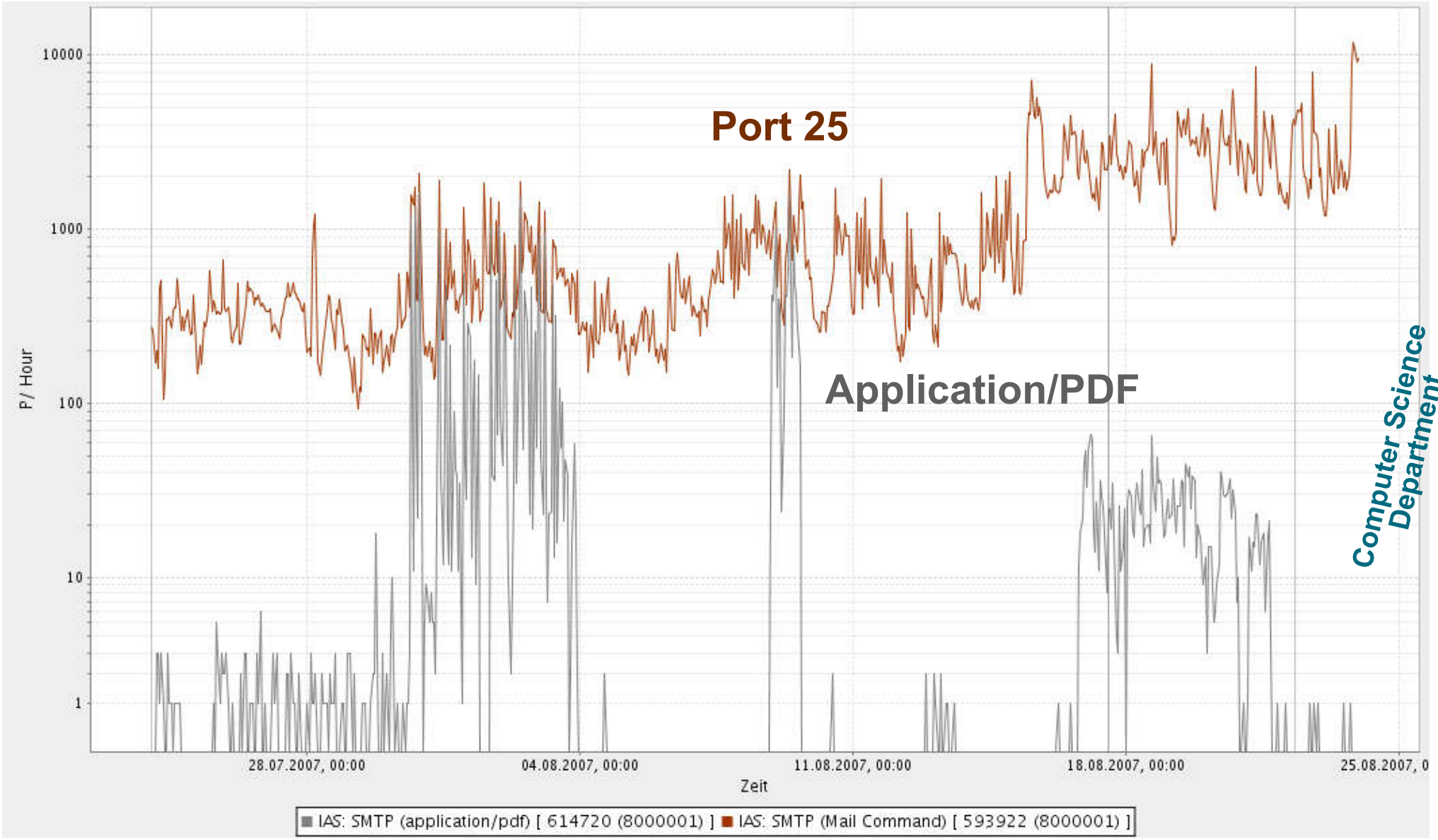
SMTP Header "MIME-Type"
→ Result: Detection of attacks (1/3)

- **SMTP Content Type**
  - Temporarily more e-mails with attachments -> Mail-(Worms/Virus)!



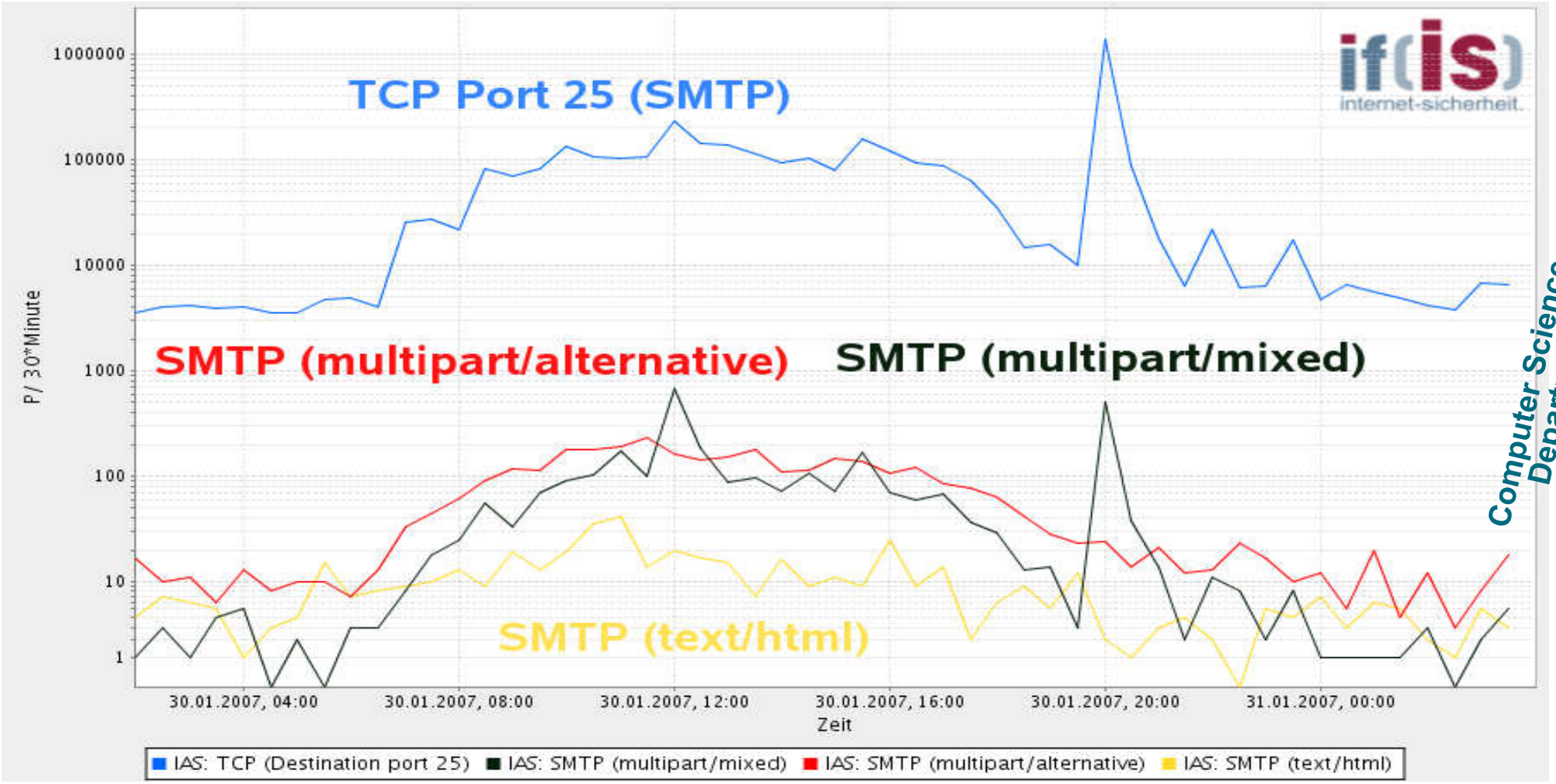
SMTP Header "MIME-Type"
→ Result: Detection of attacks (2/3)

- **PDF Spam Wave**



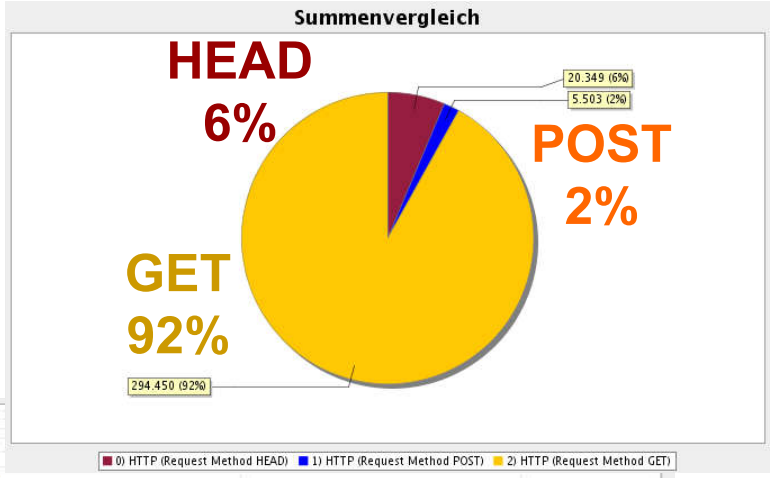
SMTP Header "MIME-Type"
→ Result: Detection of attacks (3/3)

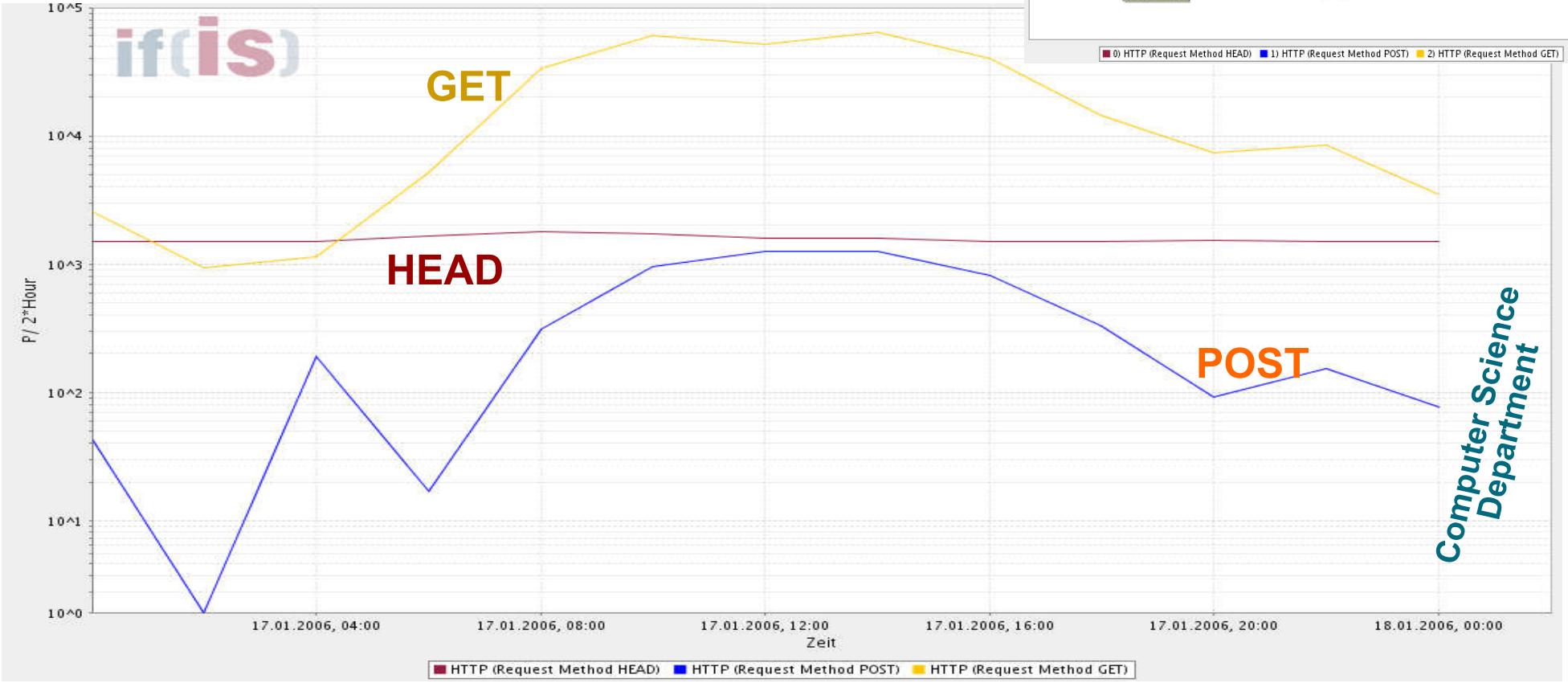
- **BKA worm (Sober.Z)**
  - The waves were transmitted in January 2007 concentrated at 3 pm and/or 8 pm.



HTTP Header "Methods"
→ Result: Distribution HTTP Methods

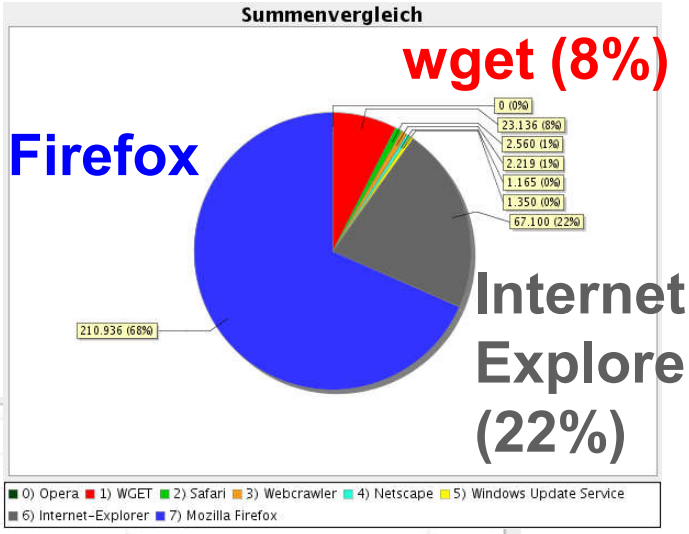
- **HTTP Methods**
  - Diurnal rhythm
    - HEAD used by automated processes
    - GET und POST usually used by human users

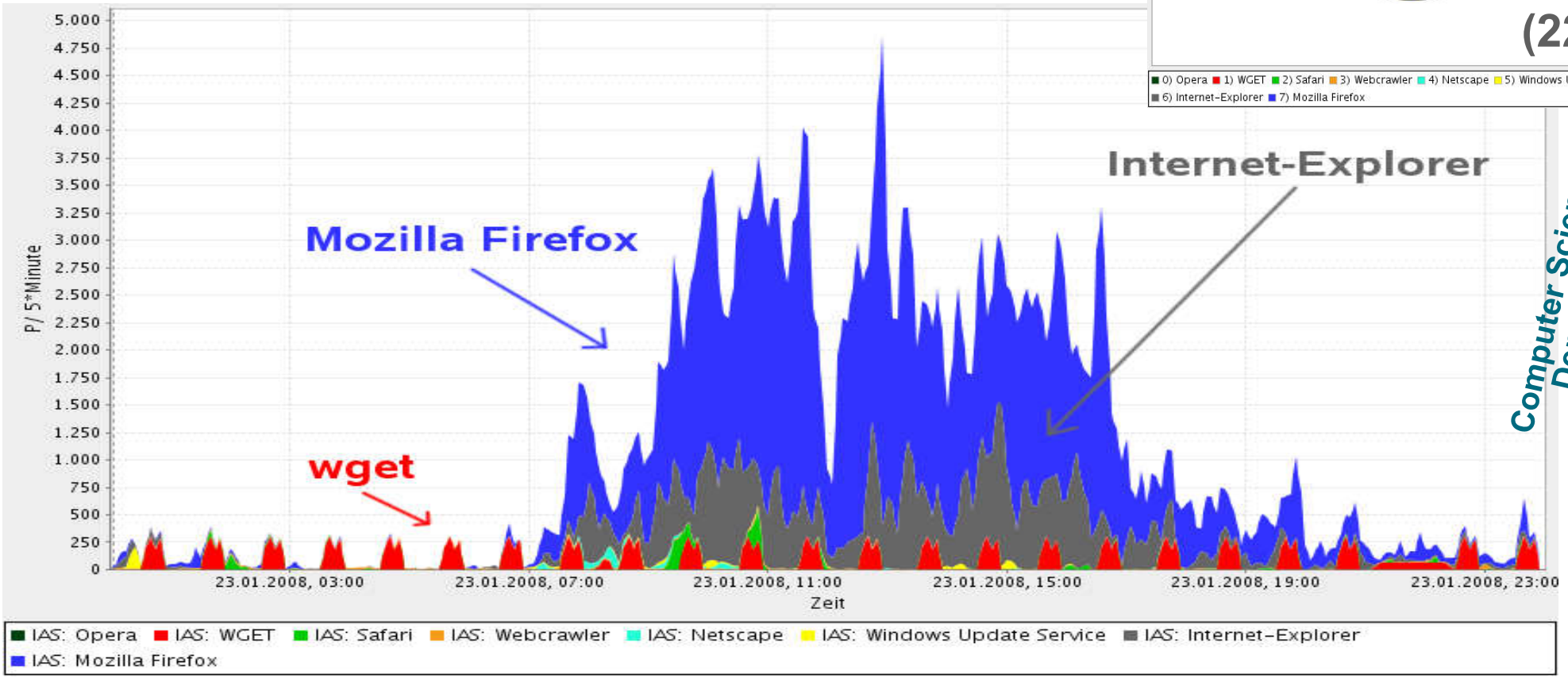




HTTP Header "M: User Agent"
→ Result: Technology trend (1/2)

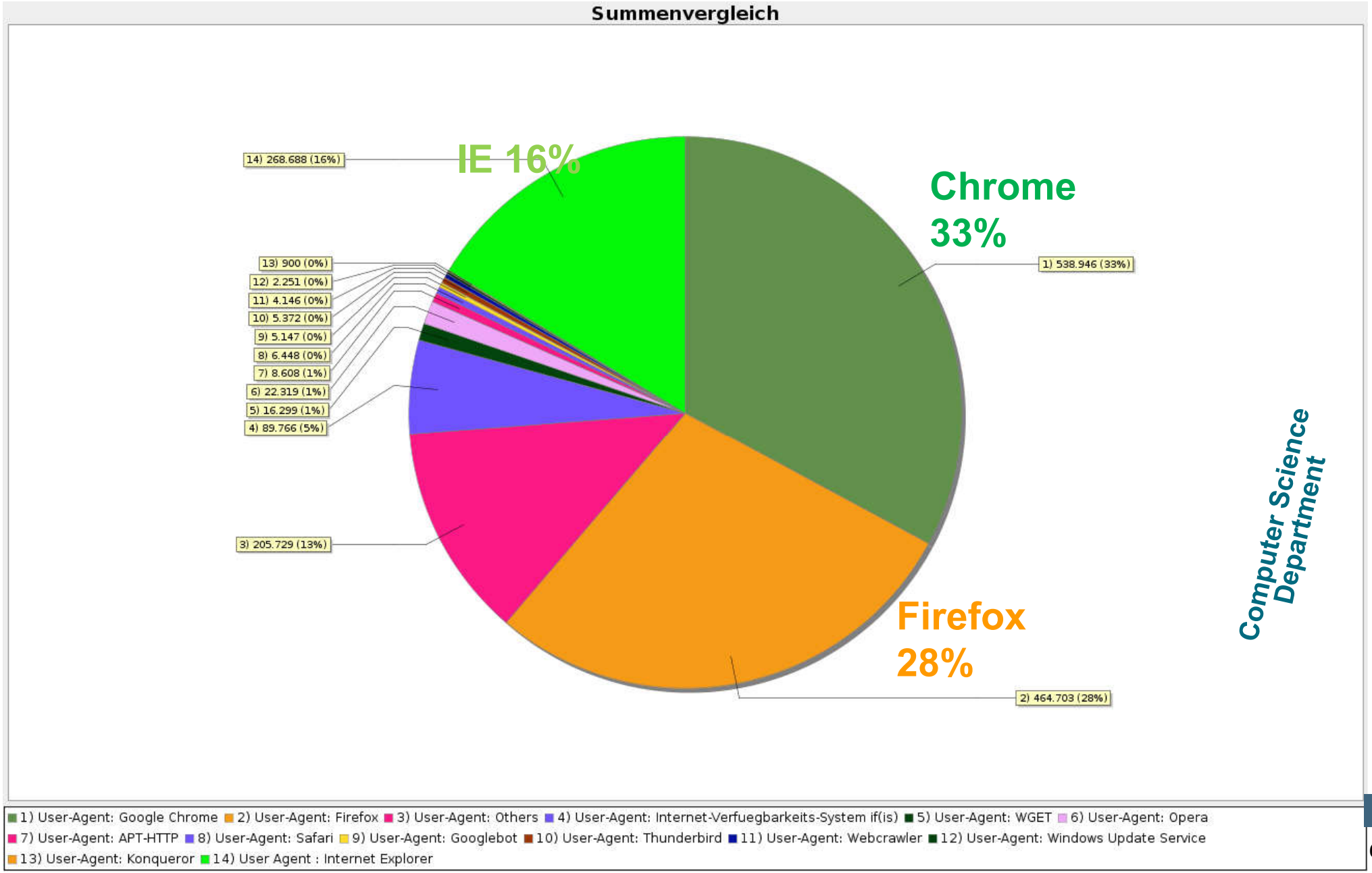
- **Distribution of browsers (2008)**
  - Diurnal profile
  - Differences between manual use (e.g. Internet Explorer und Firefox) and automated use (z.B. wget) are detectable.





Computer Science Department

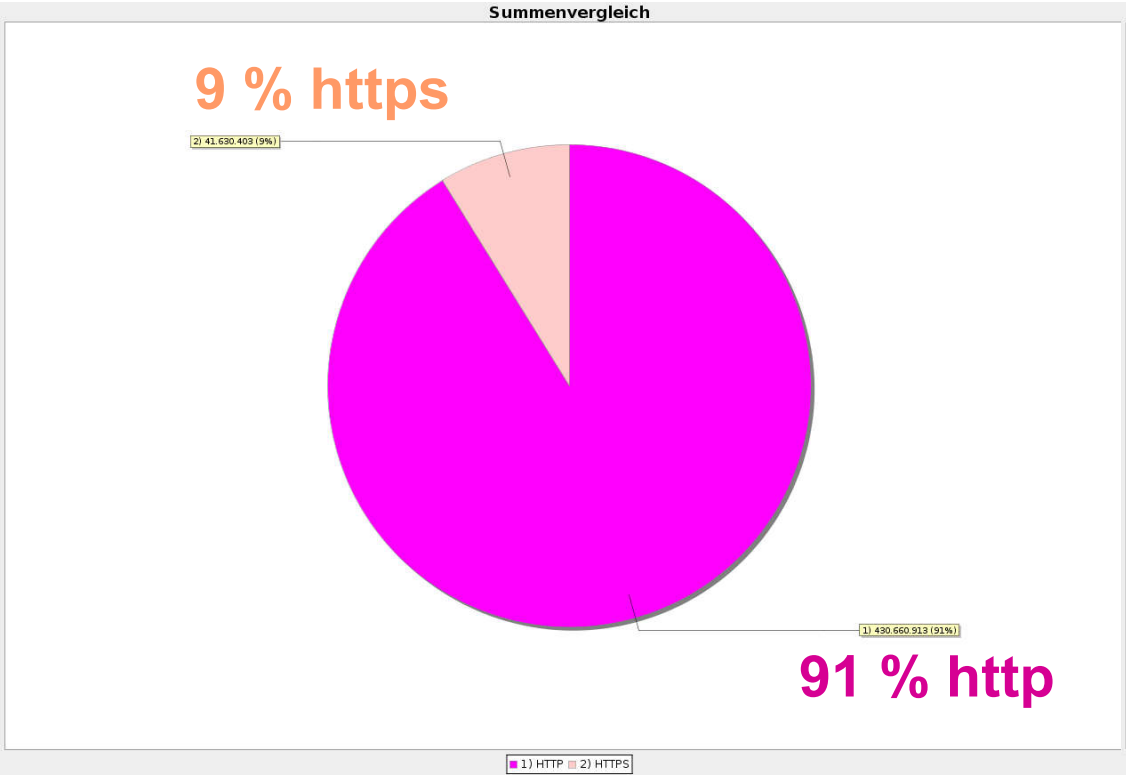
HTTP Header "M: User Agent"
→ Result: Technology trend (2/2)

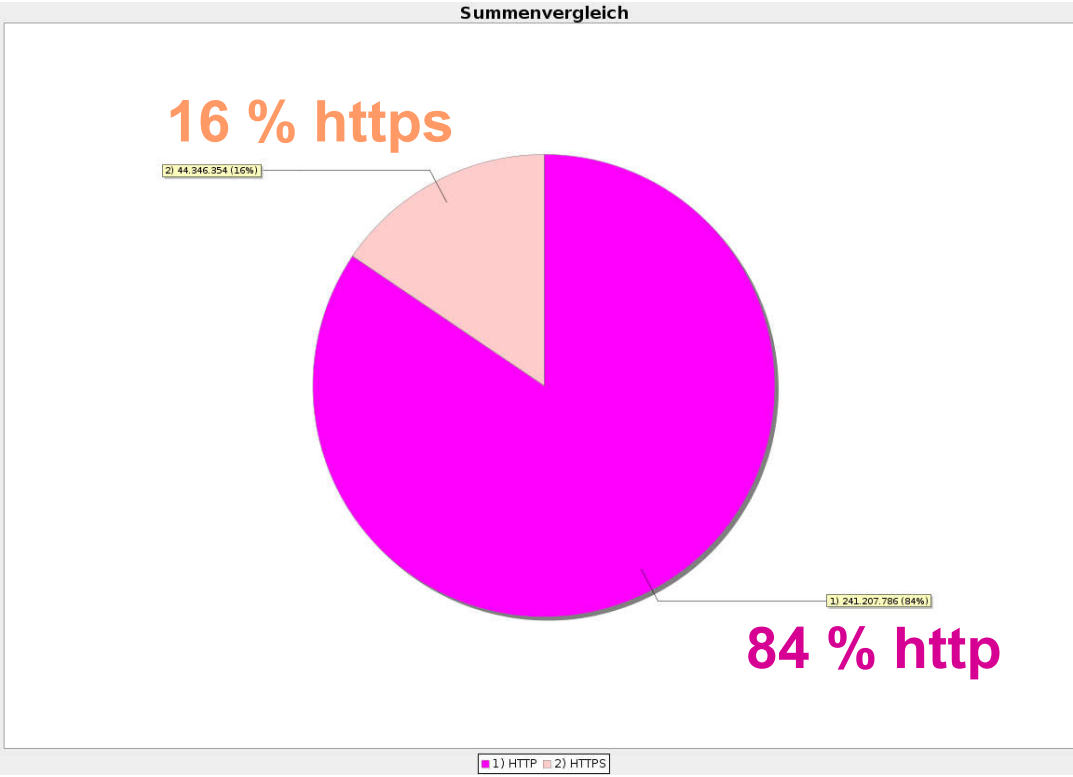


Computer Science Department

TCP-Header "Port Number"
→ Result: HTTP / HTTPS

- Distribution of encrypted HTTP-Session (IXP 10/90)





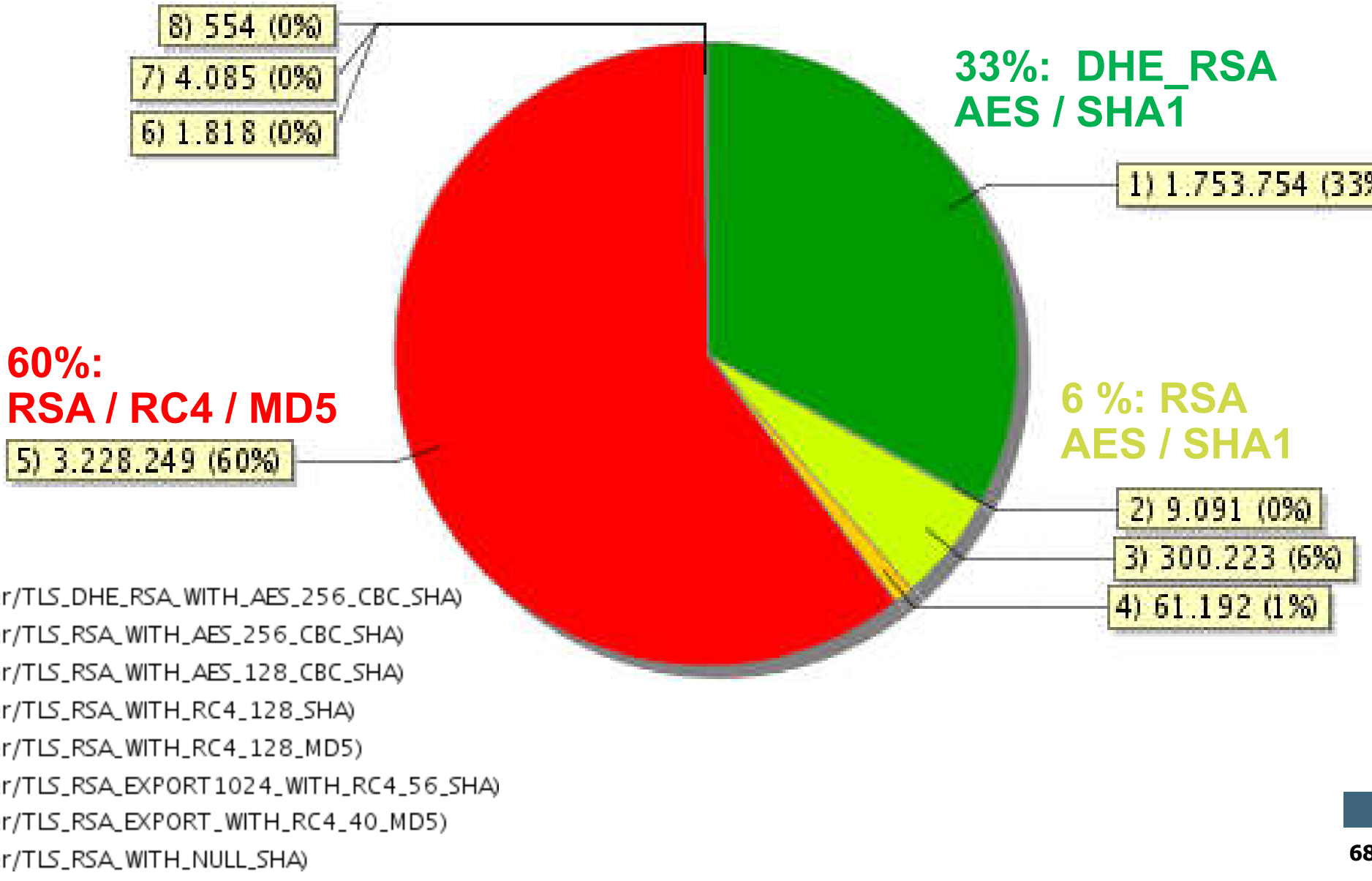
lecture-free time

lecture time

Computer Science Department

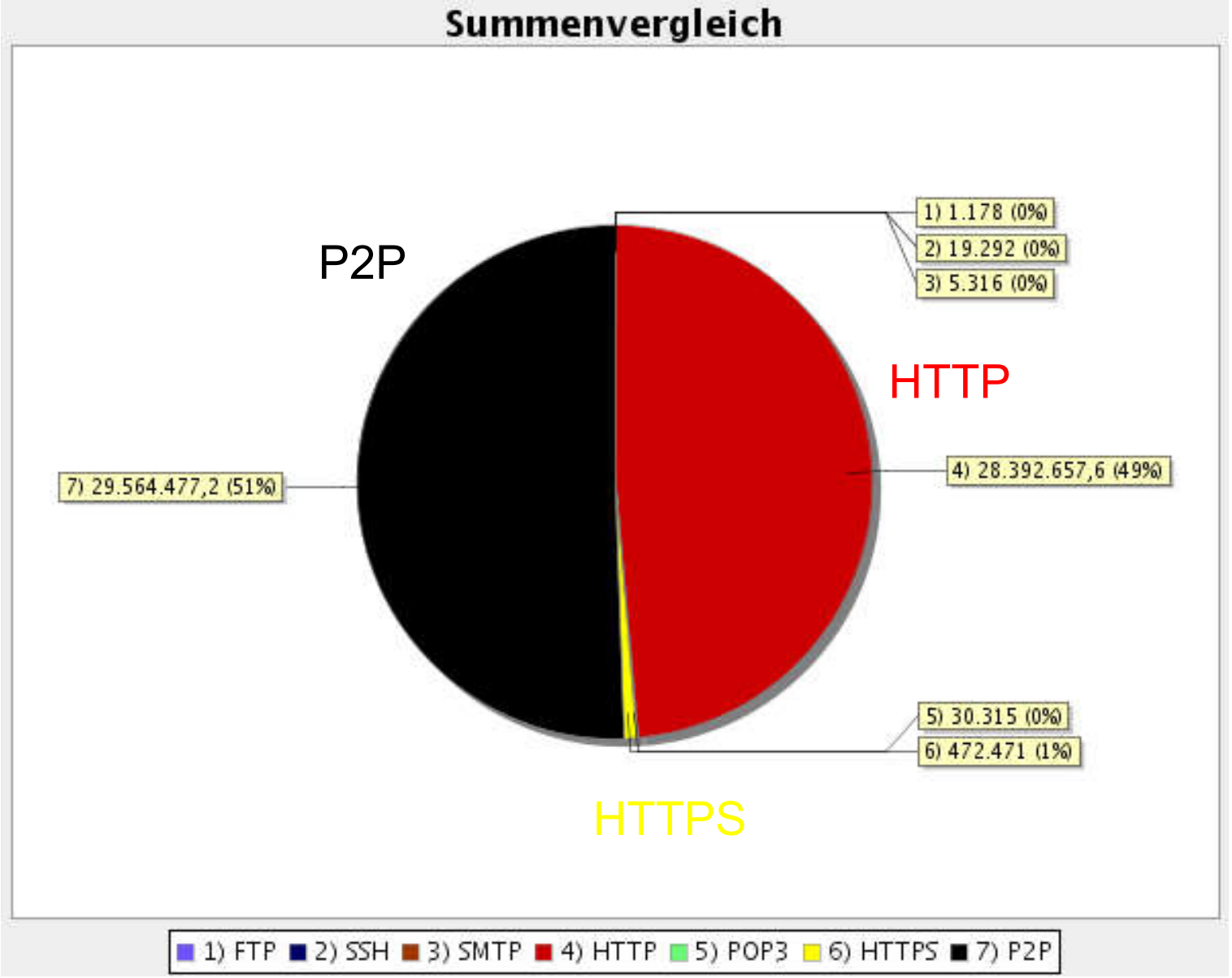
“Encryption Algorithm”
→ Result: Awareness (Crypto used TLS)

!! 0.1 %: RSA / Export (40) / SHA1 and 0.01 %: RSA / NULL / SHA1 !!



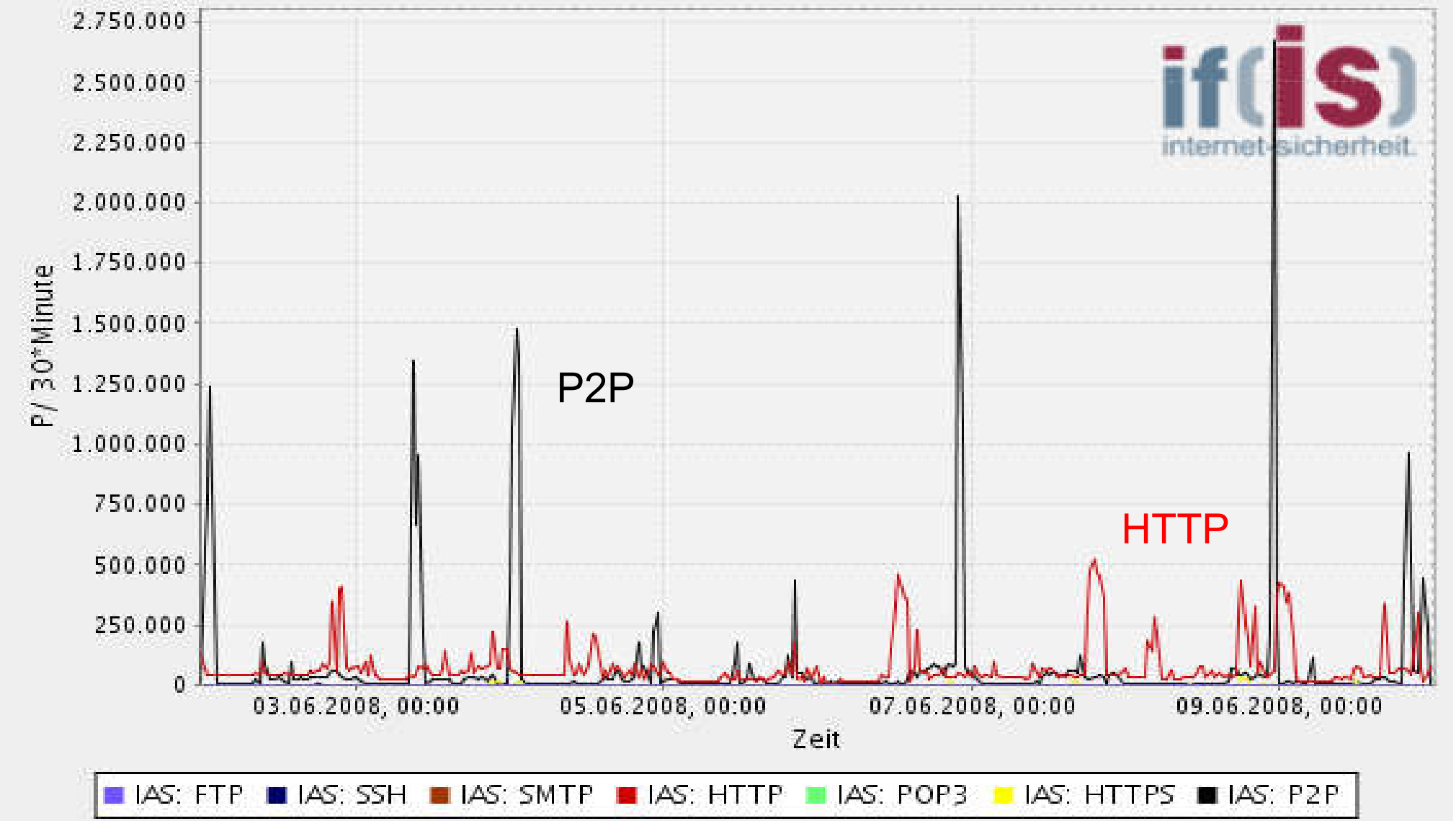
TCP-Header "Port Number"
→ Result: DSL-Connection (1/2)

- Distribution of protocols (sum)

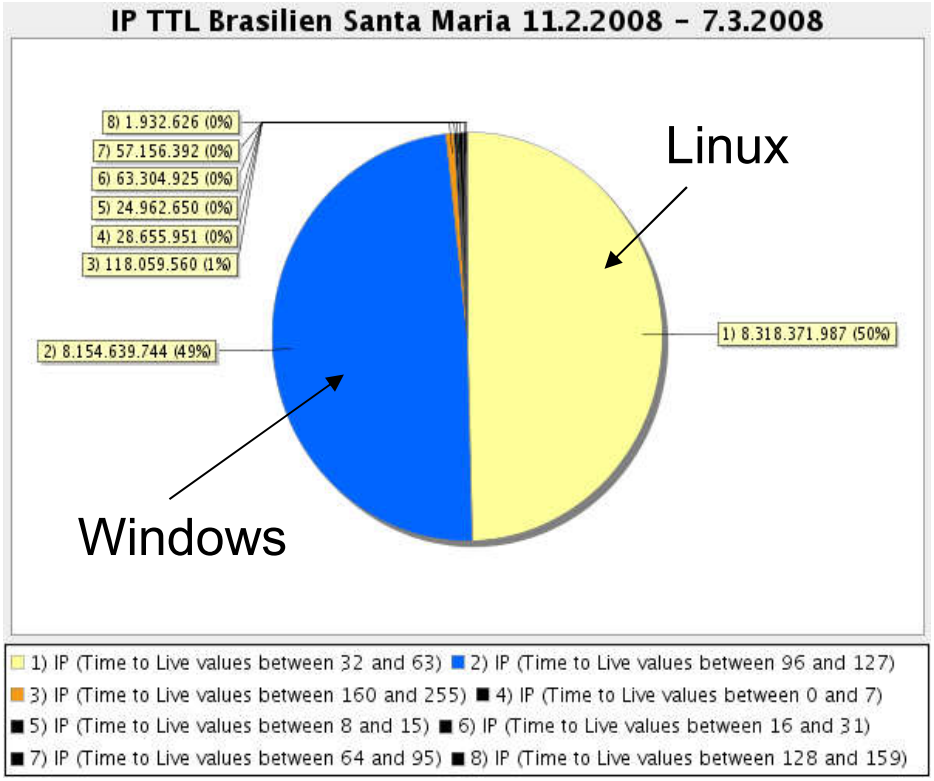


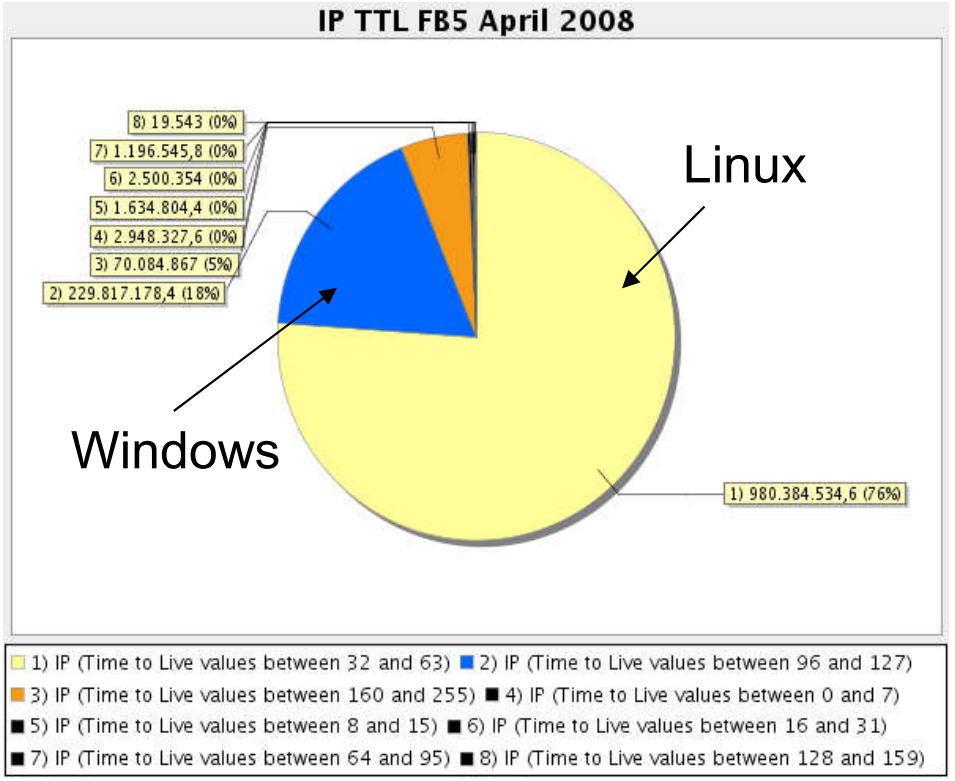
TCP-Header "Port Number"
→ Result: DSL-Connection (2/2)

- Distribution of protocols (over the time)



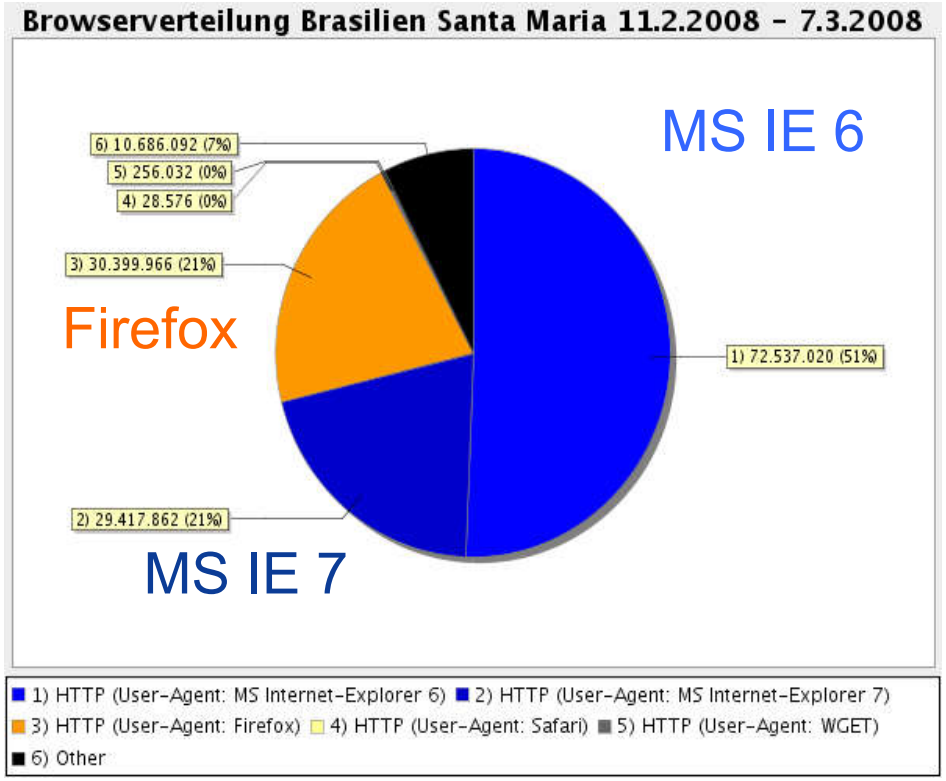
IPv4 Header „Time To Live"-field
→ Result: Different places (1/4)

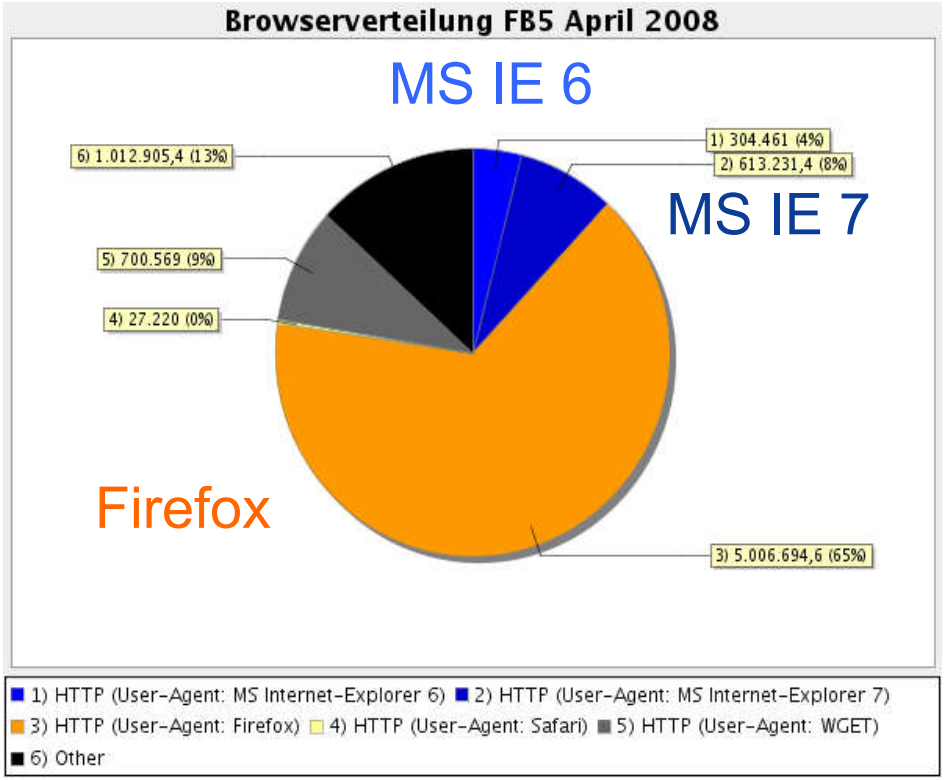




- TTL 64 value set by Linux.
- TTL 128 value set by Windows.
- TTL 255 value set by some Routers.
- Finding: a lot more Linux users at the department of computer science

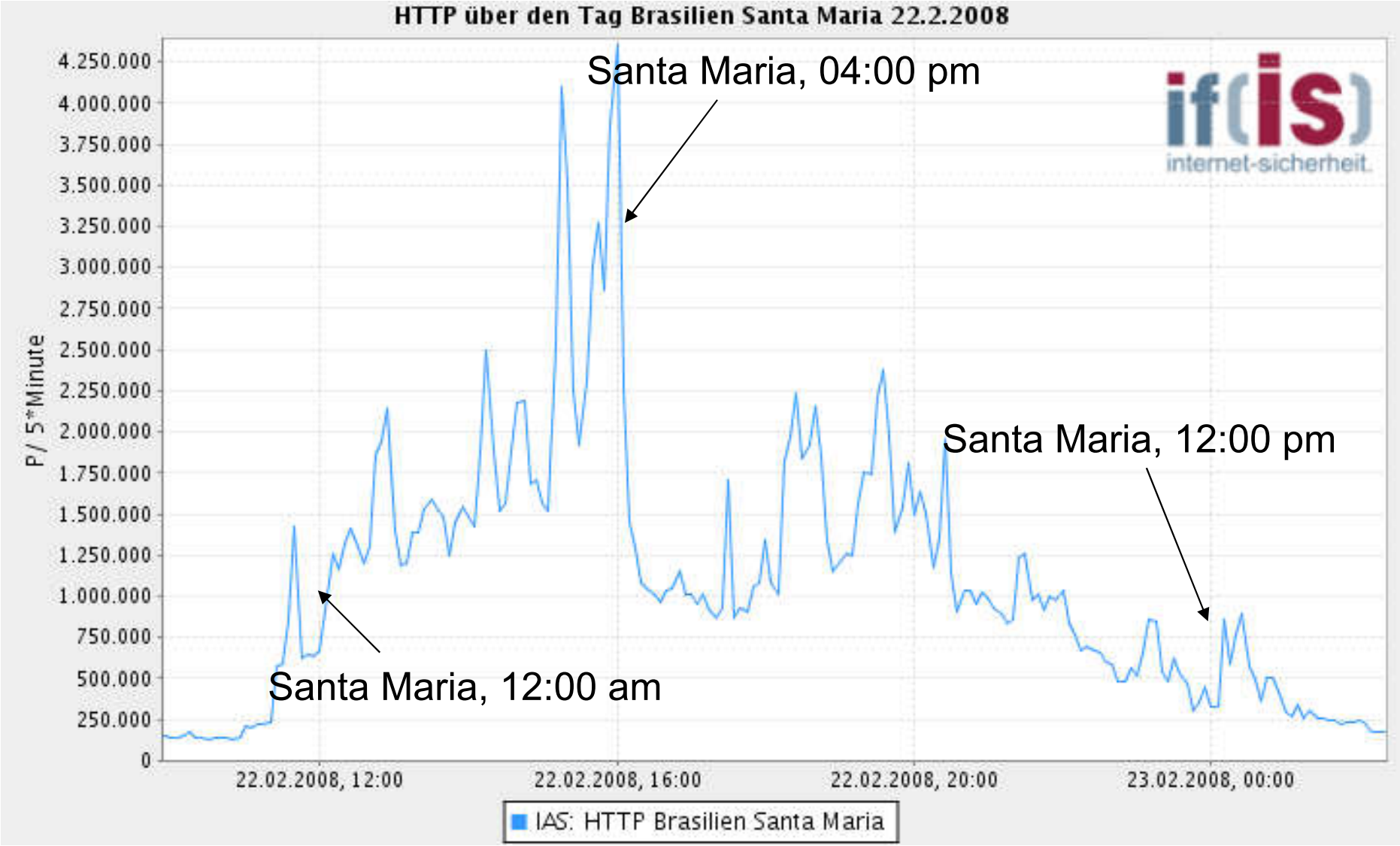
HTTP-Header "M: User Agent"
→ Result: Different places (2/4)



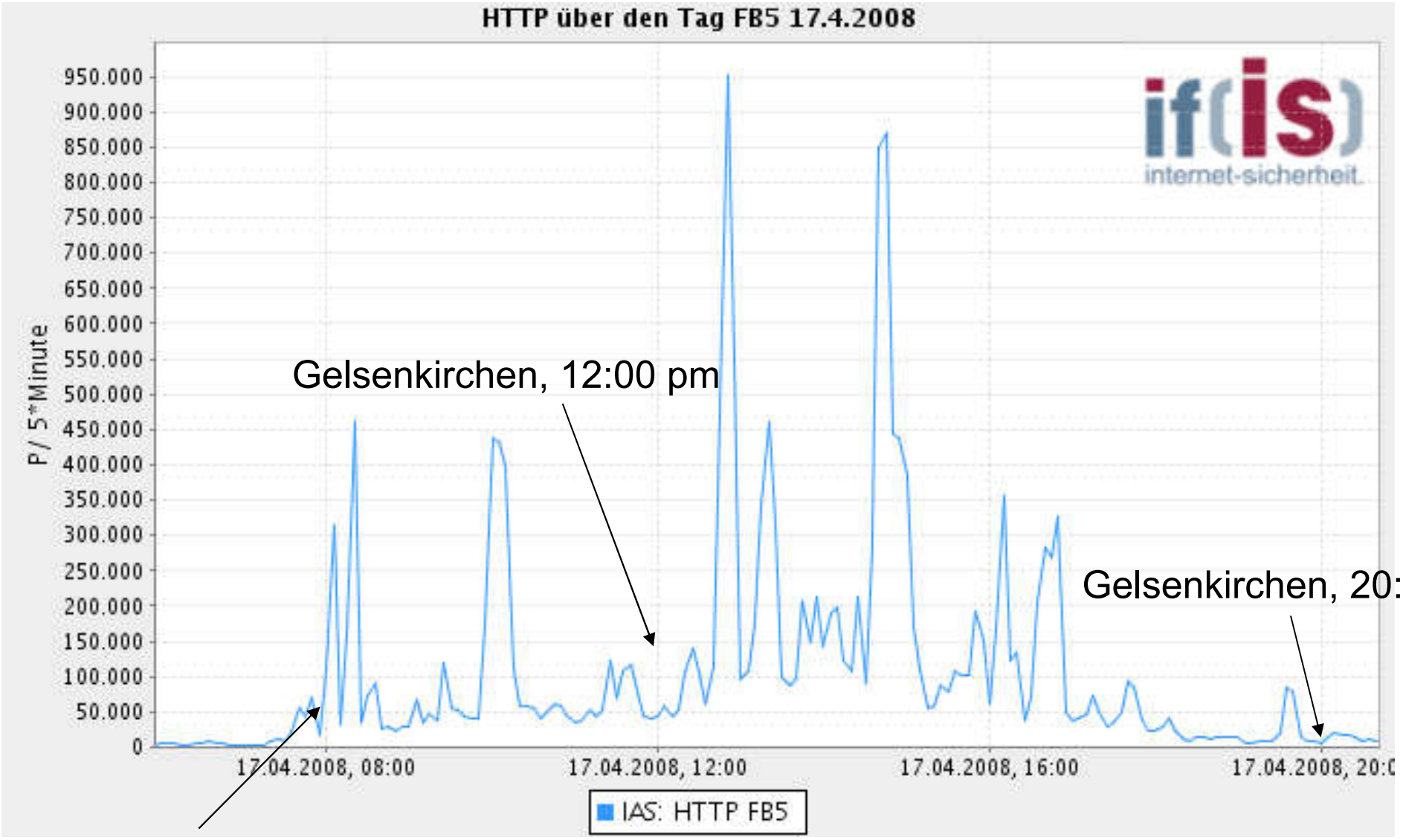


- Department of computer science: large portion of Firefox users, even though windows is used as an operating system in the computer lab.
- Brazil: a lot of Internet Explorer, what requires a Microsoft operating system.

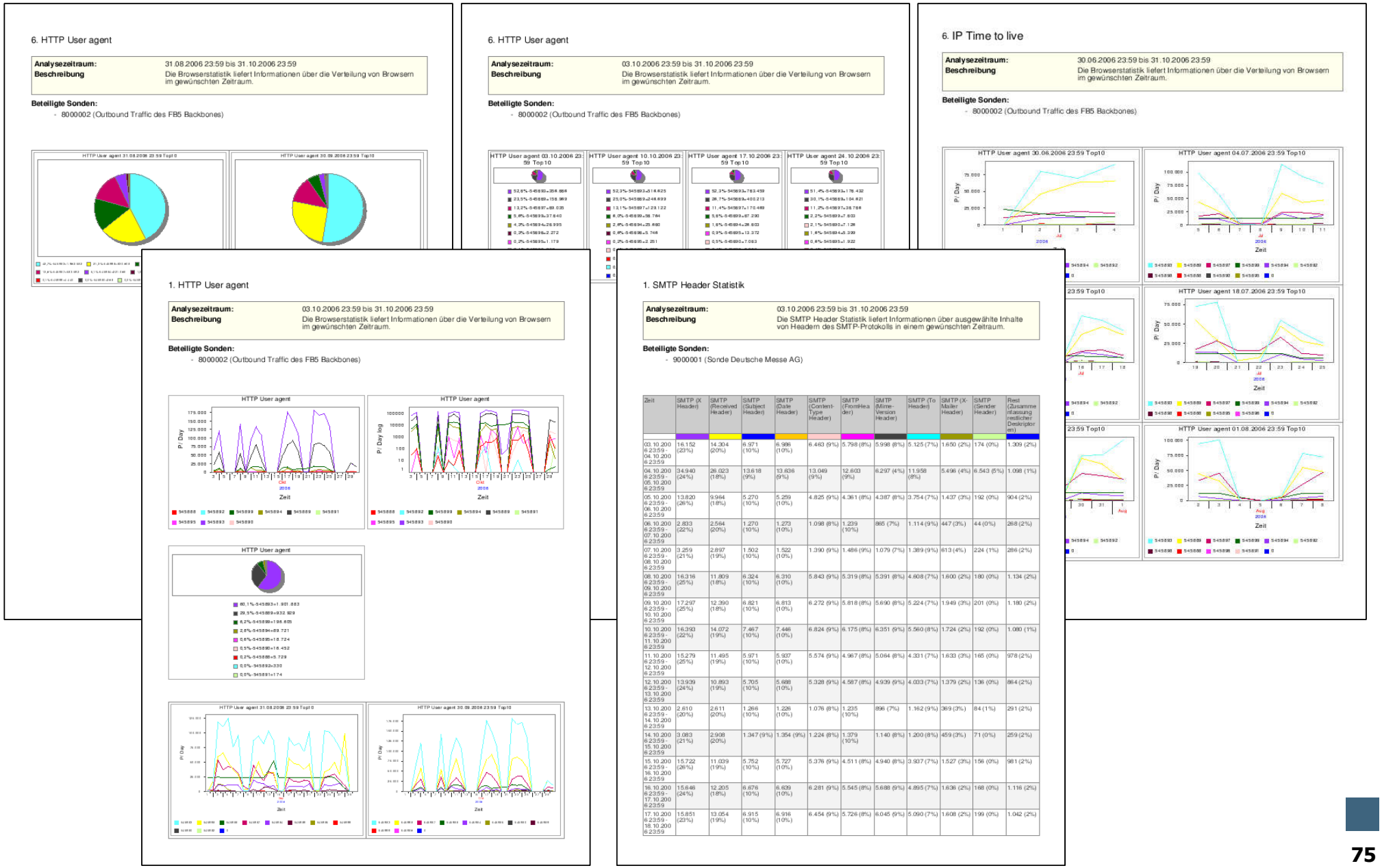
TCP Header "Port Number"
→ Result: Different places (3/4)



TCP Header "Port Number"
→ Result: Different places (4/4)



IAS: Current State of Development
→ Situation Reports



6. HTTP User agent

Analysezeitraum: 31.08.2006 23:59 bis 31.10.2006 23:59
Beschreibung Die Browserstatistik liefert Informationen über die Verteilung von Browsern im gewünschten Zeitraum.

Beteiligte Sonden:
- 8000002 (Outbound Traffic des FB5 Backbones)

6. HTTP User agent

Analysezeitraum: 03.10.2006 23:59 bis 31.10.2006 23:59
Beschreibung Die Browserstatistik liefert Informationen über die Verteilung von Browsern im gewünschten Zeitraum.

Beteiligte Sonden:
- 8000002 (Outbound Traffic des FB5 Backbones)

6. IP Time to live

Analysezeitraum: 30.06.2006 23:59 bis 31.10.2006 23:59
Beschreibung Die Browserstatistik liefert Informationen über die Verteilung von Browsern im gewünschten Zeitraum.

Beteiligte Sonden:
- 8000002 (Outbound Traffic des FB5 Backbones)

1. HTTP User agent

Analysezeitraum: 03.10.2006 23:59 bis 31.10.2006 23:59
Beschreibung Die Browserstatistik liefert Informationen über die Verteilung von Browsern im gewünschten Zeitraum.

Beteiligte Sonden:
- 8000002 (Outbound Traffic des FB5 Backbones)

1. SMTP Header Statistik

Analysezeitraum: 03.10.2006 23:59 bis 31.10.2006 23:59
Beschreibung Die SMTP Header Statistik liefert Informationen über ausgewählte Inhalte von Headern des SMTP-Protokolls in einem gewünschten Zeitraum.

Beteiligte Sonden:
- 9000001 (Sonde Deutsche Messe AG)

Zeit	SMTP (X Header)	SMTP (Received Header)	SMTP (Subject Header)	SMTP (Date Header)	SMTP (Content-Type Header)	SMTP (FromHeader)	SMTP (Mime-Version Header)	SMTP (To Header)	SMTP (X-Mailer Header)	SMTP (Sender Header)	Rest (Zusammenfassung restlicher Deskriptoren)
03.10.2006 23:59 - 04.10.2006 23:59	16.152 (23%)	14.304 (20%)	6.971 (10%)	6.986 (10%)	6.463 (9%)	5.798 (8%)	5.998 (8%)	5.125 (7%)	1.650 (2%)	174 (0%)	1.309 (2%)
04.10.2006 23:59 - 05.10.2006 23:59	34.940 (24%)	26.023 (18%)	13.618 (9%)	13.636 (9%)	13.049 (9%)	12.603 (9%)	6.297 (4%)	11.958 (8%)	5.496 (4%)	6.543 (5%)	1.098 (1%)
05.10.2006 23:59 - 06.10.2006 23:59	13.820 (26%)	9.984 (18%)	5.270 (10%)	5.259 (10%)	4.825 (9%)	4.361 (8%)	4.387 (8%)	3.754 (7%)	1.437 (3%)	192 (0%)	904 (2%)
06.10.2006 23:59 - 07.10.2006 23:59	2.833 (22%)	2.564 (20%)	1.270 (10%)	1.273 (10%)	1.098 (8%)	1.239 (10%)	865 (7%)	1.114 (9%)	447 (3%)	44 (0%)	268 (2%)
07.10.2006 23:59 - 08.10.2006 23:59	3.259 (21%)	2.897 (19%)	1.502 (10%)	1.522 (10%)	1.390 (9%)	1.486 (9%)	1.079 (7%)	1.389 (9%)	613 (4%)	224 (1%)	286 (2%)
08.10.2006 23:59 - 09.10.2006 23:59	16.316 (25%)	11.809 (18%)	6.324 (10%)	6.310 (10%)	5.843 (9%)	5.319 (8%)	5.391 (8%)	4.608 (7%)	1.600 (2%)	180 (0%)	1.134 (2%)
09.10.2006 23:59 - 10.10.2006 23:59	17.297 (25%)	12.390 (18%)	6.821 (10%)	6.813 (10%)	6.272 (9%)	5.818 (8%)	5.690 (8%)	5.224 (7%)	1.949 (3%)	201 (0%)	1.180 (2%)
10.10.2006 23:59 - 11.10.2006 23:59	16.393 (22%)	14.072 (19%)	7.467 (10%)	7.446 (10%)	6.824 (9%)	6.175 (8%)	6.351 (9%)	5.560 (8%)	1.724 (2%)	192 (0%)	1.080 (1%)
11.10.2006 23:59 - 12.10.2006 23:59	15.279 (25%)	11.495 (19%)	5.971 (10%)	5.937 (10%)	5.574 (9%)	4.967 (8%)	5.064 (8%)	4.331 (7%)	1.633 (3%)	165 (0%)	978 (2%)
12.10.2006 23:59 - 13.10.2006 23:59	13.939 (24%)	10.893 (19%)	5.705 (10%)	5.688 (10%)	5.328 (9%)	4.587 (8%)	4.939 (9%)	4.033 (7%)	1.379 (2%)	136 (0%)	864 (2%)
13.10.2006 23:59 - 14.10.2006 23:59	2.610 (20%)	2.611 (20%)	1.266 (10%)	1.226 (10%)	1.076 (8%)	1.235 (10%)	896 (7%)	1.162 (9%)	369 (3%)	84 (1%)	291 (2%)
14.10.2006 23:59 - 15.10.2006 23:59	3.083 (21%)	2.908 (20%)	1.347 (9%)	1.354 (9%)	1.224 (8%)	1.379 (10%)	1.140 (8%)	1.200 (8%)	459 (3%)	71 (0%)	259 (2%)
15.10.2006 23:59 - 16.10.2006 23:59	15.722 (26%)	11.039 (19%)	5.752 (10%)	5.727 (10%)	5.376 (9%)	4.511 (8%)	4.940 (8%)	3.937 (7%)	1.527 (3%)	156 (0%)	981 (2%)
16.10.2006 23:59 - 17.10.2006 23:59	15.646 (24%)	12.205 (18%)	6.676 (10%)	6.639 (10%)	6.281 (9%)	5.545 (8%)	5.688 (9%)	4.895 (7%)	1.636 (2%)	168 (0%)	1.116 (2%)
17.10.2006 23:59 - 18.10.2006 23:59	15.851 (23%)	13.054 (19%)	6.915 (10%)	6.916 (10%)	6.454 (9%)	5.726 (8%)	6.045 (9%)	5.090 (7%)	1.608 (2%)	199 (0%)	1.042 (2%)

Internet Analysis System (IAS)
→ Example for a fragment form the Reports

Internet Analysis System (IAS)
→ Status

- At the moment we can analyze more than 3.200.000 different parameters
- We have collected a lot of data in your knowledge base, which helps us to define what we consider the normal state.
- The statistics help us, to understand the actual traffic
- With the help of the reports, we can receive aggregations with the most important results on a regular basis
  - Gives a great overview
  - These are very good information, to understand the normal behavior of an environment
  - The communication behavior stays under monitoring
  - Trends can be recognized at this stage
  - Abnormal behavior, which were left out of perspective during the manual analysis, can be detected with the help of these summaries

Internet Analysis System (IAS)
→ Further proceeding

- **Validating of the communication parameters**
  - Which are the once that are really used?
  - Which once are redundant?
  - How can we further reduce the amount of the collected data, for instance by using aggregation?
- **Identifying of new communication parameters**
  - Which protocols will gain in importance?
  - Which data is necessary to give a complete description of the Internet?
- **Working with / Analyzing of the knowledge base**
  - Use of data mining to find correlations and to better understand what we are dealing with
- **Find more partner**
  - Have more sensor running at different places

Content

- Aim and outcomes of this lecture
- Idea of the Internet Analysis System
- Knowledge Base
- **Outline of the Current State**
- Detection of Attacks and Deflection
- Forecast of Patterns and Attacks
- Summary

Internet Analysis System (IAS)
→ Defined targets

Target 2

- **Outline of the current state of the internet.**

current state

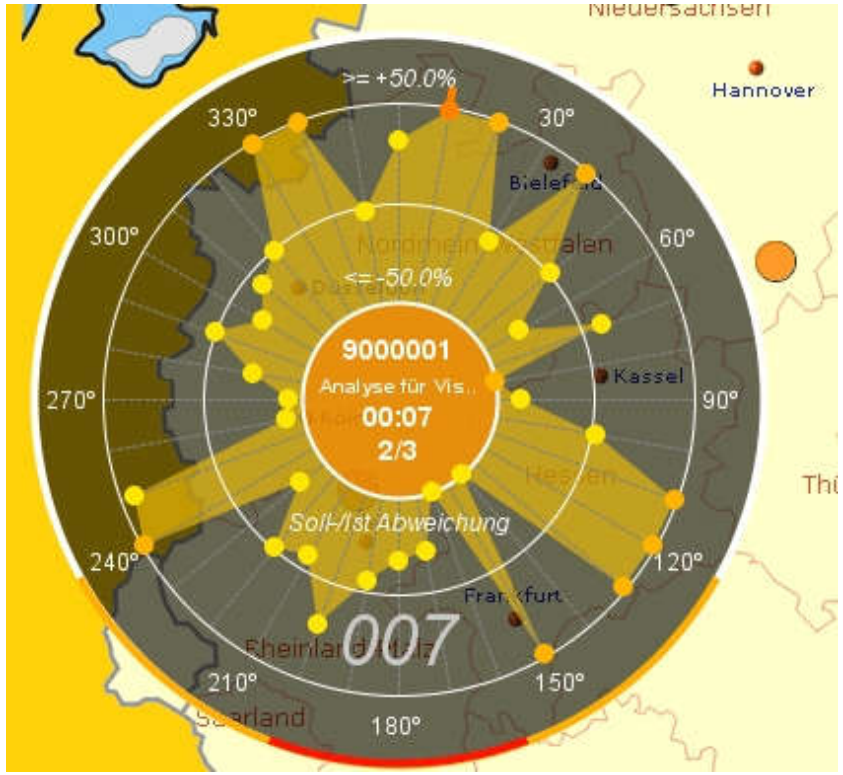
In this field an important function is the clear visual representation of the state of the internet, like traffic jam maps.

Current state - IAS
→ Target 2: Overview

- We need designs that help us determine the current state.

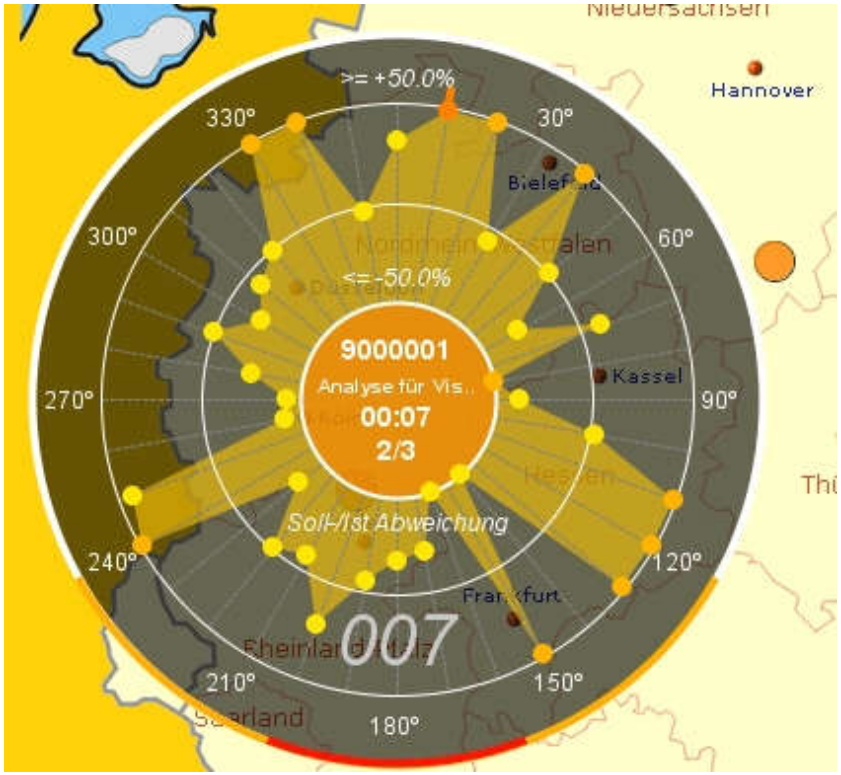
  Challenge: display enormous amount of data in an intuitive manner

- One example for a visualization tool we use to gain on experience is **VisiX**.
- VisiX: **V**isual **I**nternet **S**ensor **I**nformation
- Pre selected, **important Parameters**
- **Continuous updating**
- Alignment on the basis of fixed reference values
- selectable, **colored coding**

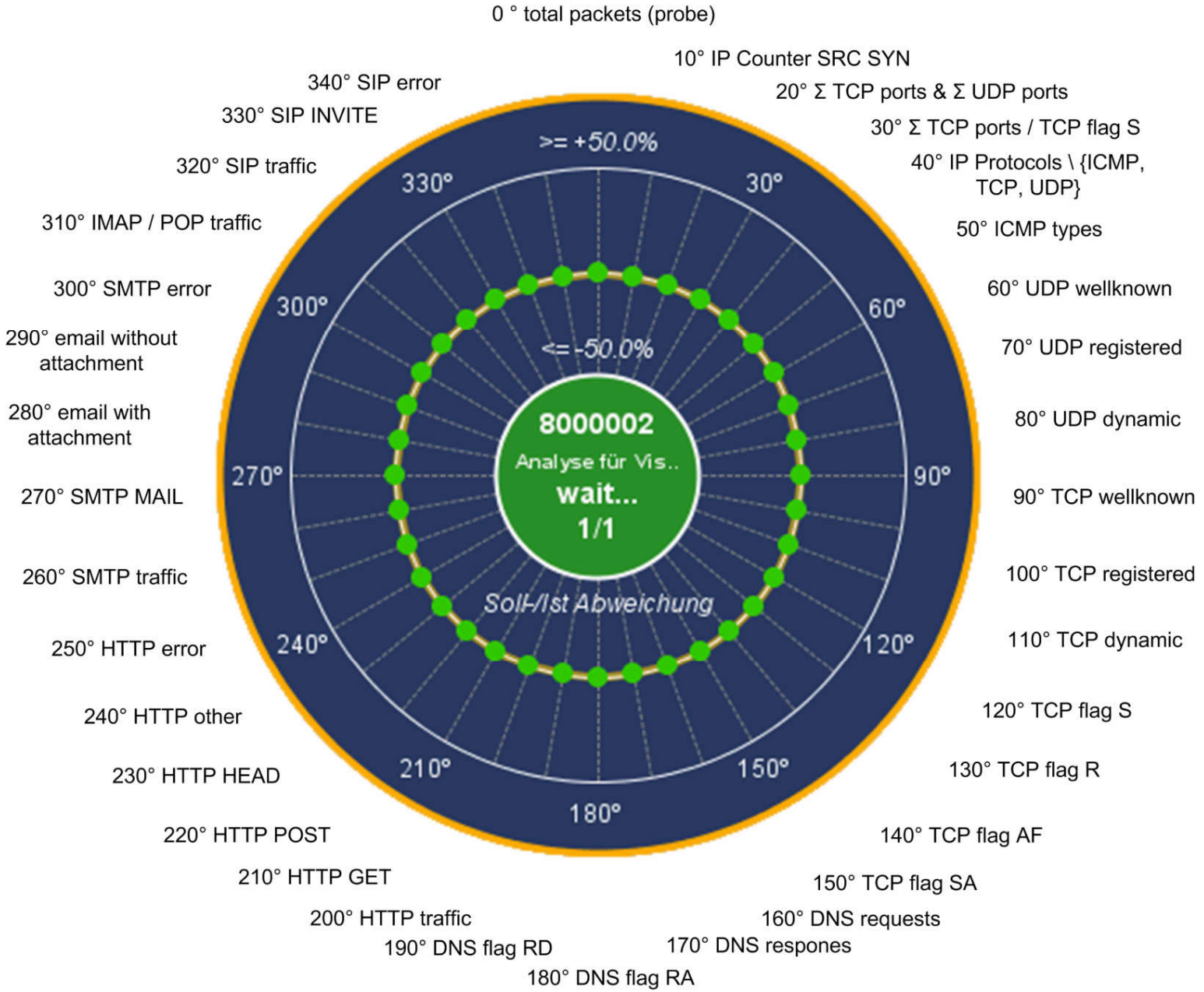


Current state - IAS
→ Visual Internet Sensor Information

- Visualization of the data of multiple probes at the same time using multiple diagrams in one visualization
- This allows to detect coherences between different probes
- For example:
  - sensor X: extremely high level of http traffic
  - sensor Y: extremely high as well
  - → external event like a Windows Update or a possible attack
- VisiX allows the user to get to know the communication behavior of a network
- Continuous monitoring in the case of an alert
- Helps the user to initiate further measures
- Procedure: (i) Alert → (ii) VisiX → (iii) EagleX, ...



Current state - IAS
→ Special parameter: Overview



Current state - IAS
→ Special communication parameter (1/3)

- 0 META: total Packets (probe)
- 10 IP Counter SRC + SynFlag
- 20 Total TCP and UDP Ports
- 30 Total Packet / (TCP Ports + UDP Ports) 7 / {na}
- 40 IP Protocols
  - ALL (without 1 , 6, 17)
- 50 ICMP Types
  - ICMP (Type 0 echo reply)
  - ICMP (Type 3 destination unreachable)
  - ICMP (Type 4 source quench)
  - ICMP (Type 5 redirect)
  - ICMP (Type 6 alternate host address)
  - ICMP (Type 8 echo request)
  - ICMP (Type 9 router advertisment)
  - ICMP (Type 10 router solicitation)
  - ICMP (Type 11 time exceeded)
  - ICMP (Type 12 parameter problem)
- 60 UDP wellknown Ports
  - UDP SRC Port wellknown
  - UDP DST Port wellknown
- 70 UDP registered Ports
  - UDP SRC Port registered
  - UDP DST Port registered
- 80 UDP dynamic Ports
  - UDP SRC Port dynamic
  - UDP DST Port dynamic
- 90 TCP wellknown Ports
  - TCP SRC Port wellknown
  - TCP DST Port wellknown
- 100 TCP registered Ports
  - TCP SRC Port registered
  - TCP DST Port registered
- 110 TCP dynamic Ports
  - TCP SRC Port dynamic
  - TCP DST Port dynamic

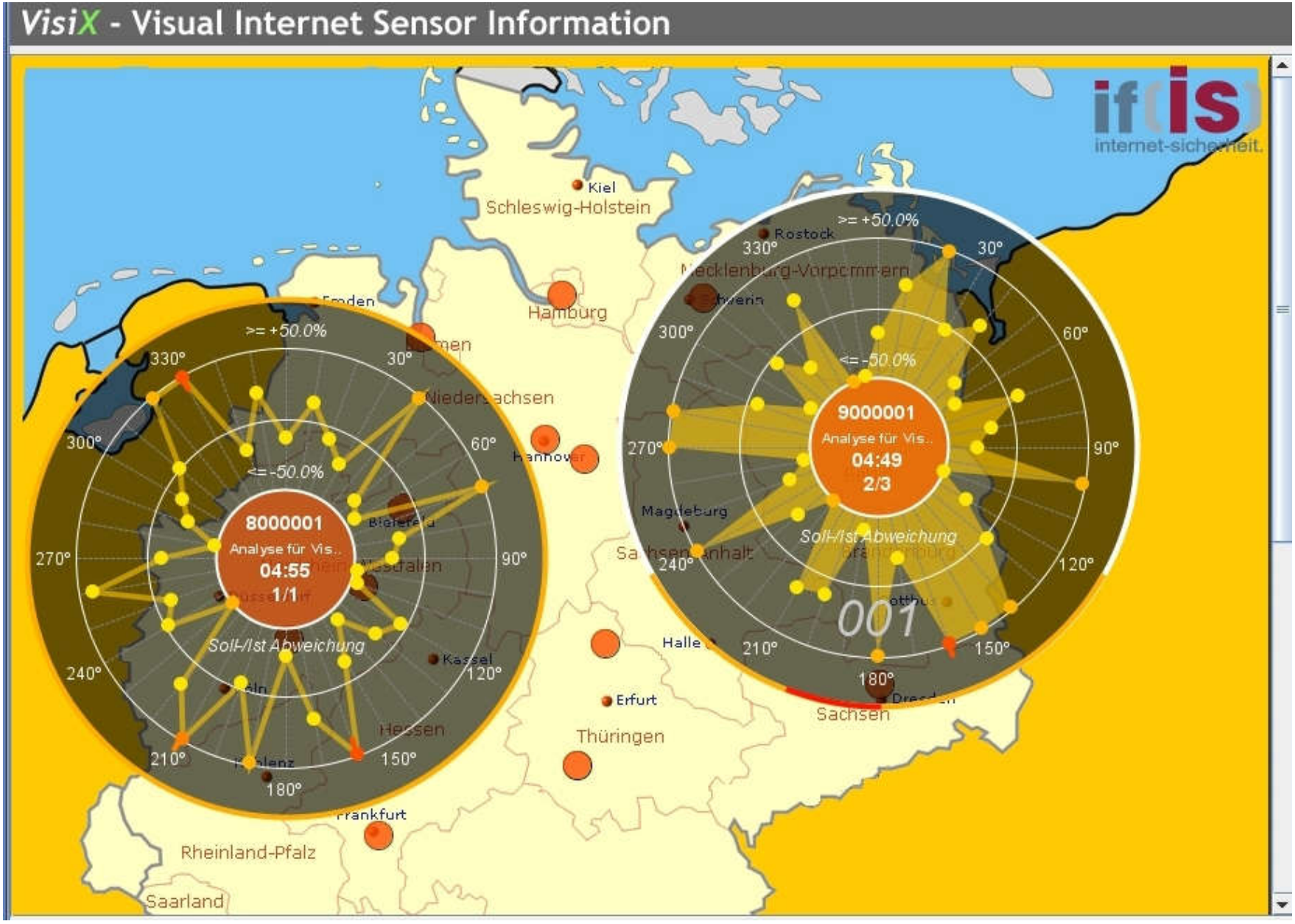
Current state - IAS
→ Special communication parameter (2/3)

- 120 TCP Flag S
- 130 TCP Flag R
- 140 TCP Flag AF
- 150 TCP Flag SA
- 160 DNS Requests
- 170 DNS Responses
- 180 DNS Flag RA
- 190 DNS Flag RD
- 200 HTTP / HTTPS Traffic
  - TCP (Source port 80)
  - TCP (Source port 443)
  - TCP (Destination port 80)
  - TCP (Destination port 443)
- 210 HTTP Request method GET
- 220 HTTP Request method POST
- 230 HTTP Request method HEAD
- 240 HTTP Request method OTHER
  - HTTP (Request Method PUT)
  - HTTP (Request Method DELETE)
  - HTTP (Request Method TRACE)
  - HTTP (Request Method OPTIONS)
  - HTTP (Request Method CONNECT)
- 250 HTTP Server response codes
  - 4xx
  - 5xx
- 260 SMTP / SMTPS Traffic
  - TCP (Source port 25)
  - TCP (Source port 465)
  - TCP (Source port 587)
  - TCP (Destination port 25)
  - TCP (Destination port 465)
  - TCP (Destination port 587)

Current state - IAS
→ Special communication parameter (3/3)

- 270 SMTP MAIL
- 280 SMTP E-Mail with attachment
  - Client header multipart/mixed
- 290 SMTP E-Mail without attachment
  - client header text/plain
  - client header text/html
  - client header multipart/alternative
  - client header multipart/report
- 300 SMTP Server response codes
  - 4xx
  - 5xx
- 310 POP / POPS / IMAP / IMAPS
  - TCP (Source port 110)
  - TCP (Source port 143)
  - TCP (Source port 993)
  - TCP (Source port 995)
  - TCP (Destination port 110)
  - TCP (Destination port 143)
  - TCP (Destination port 993)
  - TCP (Destination port 995)
- 320 SIP Invite
- 330 SIP Traffic
  - TCP (Source port 5060)
  - TCP (Destination port 5060)
  - UDP (Source port 5060)
  - UDP (Destination port 5060)
- 340 SIP error codes
  - 4xx
  - 5xx
  - 6xx

IAS: Current State of Development
→ Continuous situation awareness (1/2)



IAS: Current State of Development
→ Continuous situation awareness (2/2)

general state

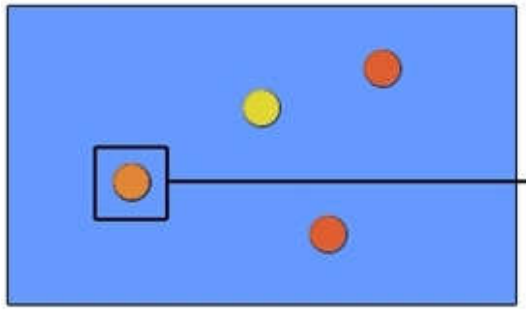


illustration in minimized view

general state and state of individual parameters

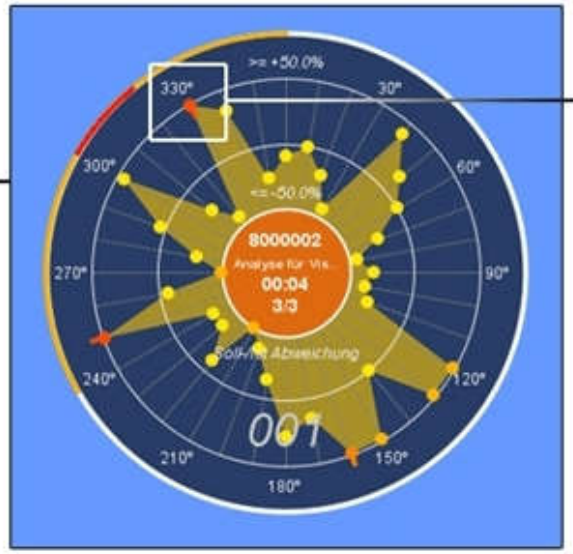


illustration in maximized view

state of individual parameters

detail view of a certain variance comparsion

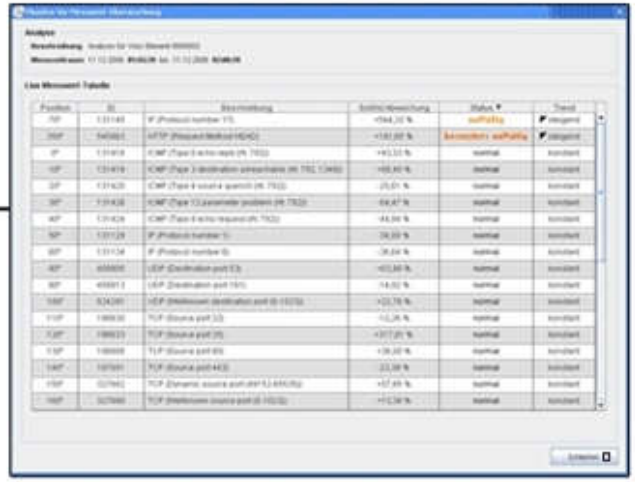
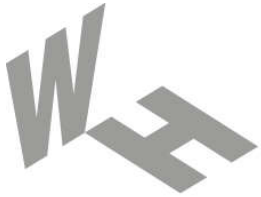


table view

level of detail



Westfälische Hochschule
Gelsenkirchen Bocholt Recklinghausen
University of Applied Sciences

Internet Analysis System
→ Part 1

Thank you for your attention!
Questions?

Prof. Dr. (TU NN)
Norbert Pohlmann

Institute for Internet Security - if(is)
University of Applied Sciences Gelsenkirchen
http://www.internet-sicherheit.de

if(is) internet security.

Internet Analysis System (IAS)
→ Literature (1/2)

- [1] N. Pohlmann: "Internetstatistik" (statistics of the internet), Proceedings of CIP Europe Publisher, B.M. Hämmerli, 2005.
- [2] N. Pohlmann, M. Proest: „Internet Early Warning System: The Global View", in "Securing Electronic Business Processes - Highlights of the Information Security Solutions Europe 2006 Conference", Hrsg.: S. Paulus, N. Pohlmann, H. Reimer, Vieweg-Verlag, Wiesbaden 2006
- [3] N. Pohlmann: “Probe-based Internet Early Warning System”, ENISA Quarterly Vol. 3, No. 1, Jan-Mar 2007
- [4] N. Pohlmann: „The global View of Security Situation in the Internet“, ECN - European CIIP Newsletter, Volume 3, Brüssel 12/2007
- [5] Sebastian Spooren, Entwicklung eines profilgestützten Visualisierungssystems zur Darstellung von raum- & zeitbezogenen Soll-/Ist-Abweichungen (development of a visualization tool for the IAS), Diploma Thesis, University of Applied Sciences Gelsenkirchen, 2007.
- [6] Gianfranco Ricci, Betrachtung der vom IAS gesammelten Kommunikationsparameter auf Relevanz zur Anomalie und Angriffserkennung (evaluation of the relevance for the detection of abnormalities and attacks of the communication parameters collected by the internet analysis system), Diploma Thesis, University of Applied Sciences Gelsenkirchen, 2008

Internet Analysis System (IAS)
→ Literature (1/2)

- [7] Uwe van Heesch: Entwicklung eines Plugin basierten Analyse-Frameworks für das Internet-Analyse-System (development of a plugin-based analyzing framework for the Internet Analysis System), Diploma Thesis, University of Applied Sciences Gelsenkirchen, 2006.
- [8] Sabyasachi Basu, Amarnath Mukherjee, Steve Klivansky: Time Series Models For Internet Traffic, 1996
- [9] Peter J. Brockwell, Richard A. Davis: Introduction To Time Series and Forecasting, Springer, 2002

Links:

Institute for Internet Security:
http://www.internet-sicherheit.de/forschung/aktuelle-projekte/internet-frhwarnsysteme/