Internet-Frühwarnsysteme

Prof. Dr. Norbert Pohlmann

Institut für Internet-Sicherheit
Fachhochschule Gelsenkirchen
https://www.internet-sicherheit.de

Agenda

- Einführung
- Frühwarnsysteme
- Struktur für Internet-Frühwarnsysteme
- Verschiedene Realisierungsansätze
- Internet-Analyse-System
- Internet-Verfügbarkeits-System
- Zusammenfassung

Agenda

- **Einführung**
- Frühwarnsysteme
- Struktur für Internet-Frühwarnsysteme
- Verschiedene Realisierungsansätze
- Internet-Analyse-System
- Internet-Verfügbarkeits-System
- Zusammenfassung

Einführung (1/4)

- **Das Internet**
  - Grosse Vielfalt an Anwendungen, Zugangsarten und Nutzern
  - Veränderungen in der Bedeutung des Internets
  - Störungen, Ausfälle, mangelnde Vertrauenswürdigkeit und Verlässlichkeit sowie Sabotageakte nehmen zu
  - „Kritische Infrastruktur“ Internet

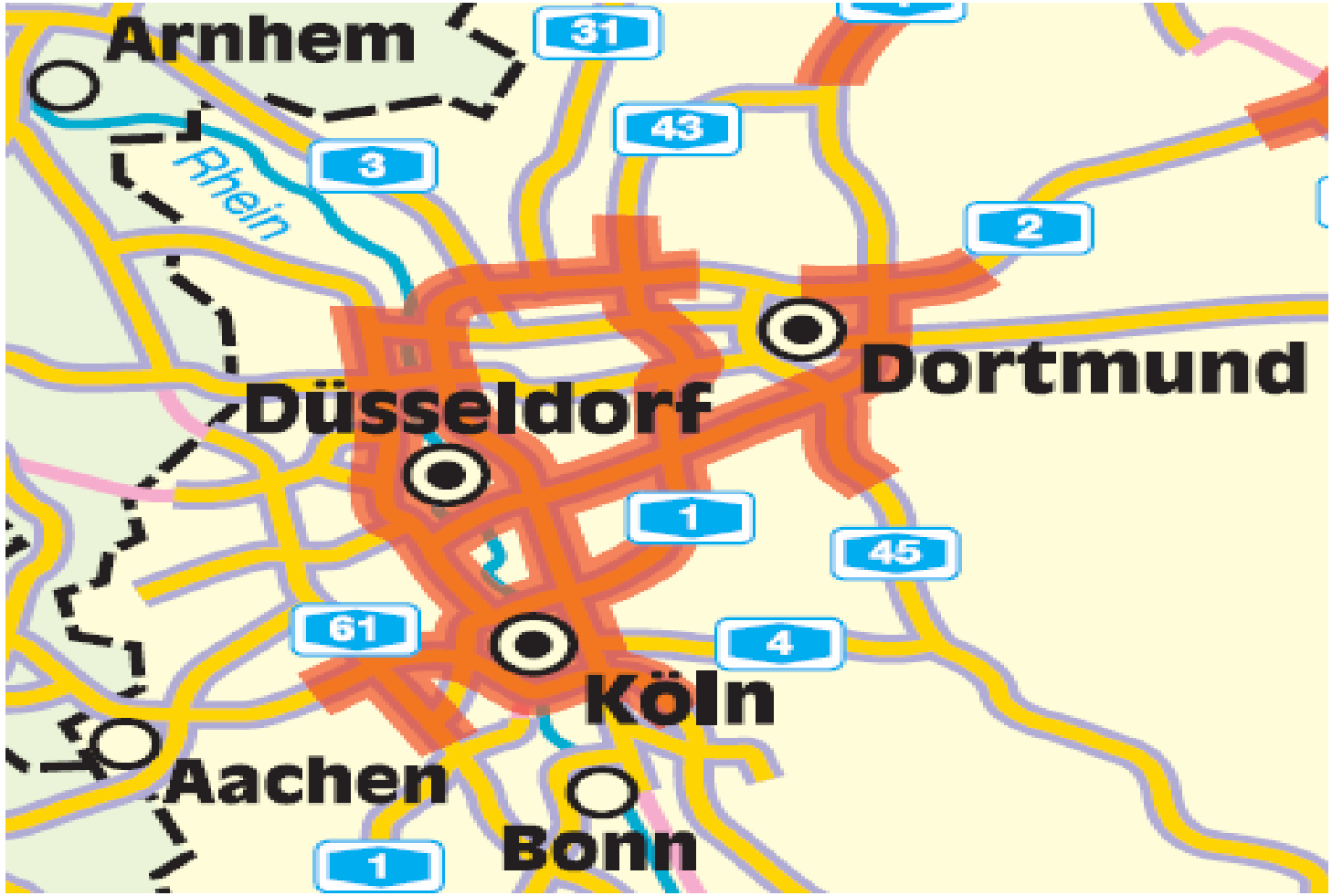
- **Konkrete Gefahren**
  - Viren, Würmer, trojanische Pferde
  - Kriminelle Aktivitäten mit gezielten Attacken
  - Phishing
  - Spam

Einführung (2/4)
→ Analogie

- **Lokale Sicht**

Einführung (3/4)
→ Analogie

- **Globale Sicht**



Einführung (4/4)

- **Konsequenzen**
  - Ohne Vertrauenswürdigkeit und Verfügbarkeit werden Dienste des Internets nicht mehr sinnvoll benutzbar sein
  - Bedrohungen erfordern angemessene Reaktionen und Gegenmaßnahmen
- **Ziele**
  - Globale Sicht auf aktuelle Situation und Zustand des Internets
  - Gefahren erkennen und Auswirkungen mildern/vermeiden
  - Erkennen von Trends und Ausgabe von Vorhersagen

Die Lösung: Internet-Frühwarnsysteme

Agenda

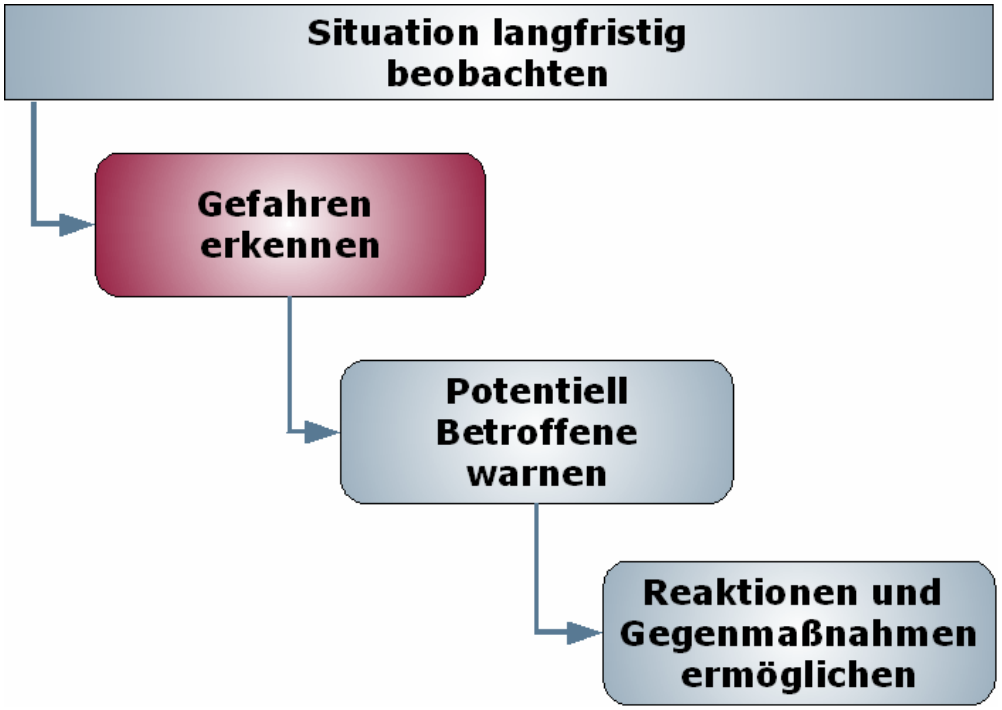
- Einführung
- **Frühwarnsysteme**
- Struktur für Internet-Frühwarnsysteme
- Verschiedene Realisierungsansätze
- Internet-Analyse-System
- Internet-Verfügbarkeits-System
- Zusammenfassung

Frühwarnsysteme (1/4)

- **Definition**

"...eine Einrichtung, welche aufkommende Gefahren frühzeitig als solche erkennt, und Gefährdete möglichst schnell darüber informiert. Es soll ermöglichen, durch eine rechtzeitige Reaktion die Gefahr abzuwenden oder zu mildern"[1]

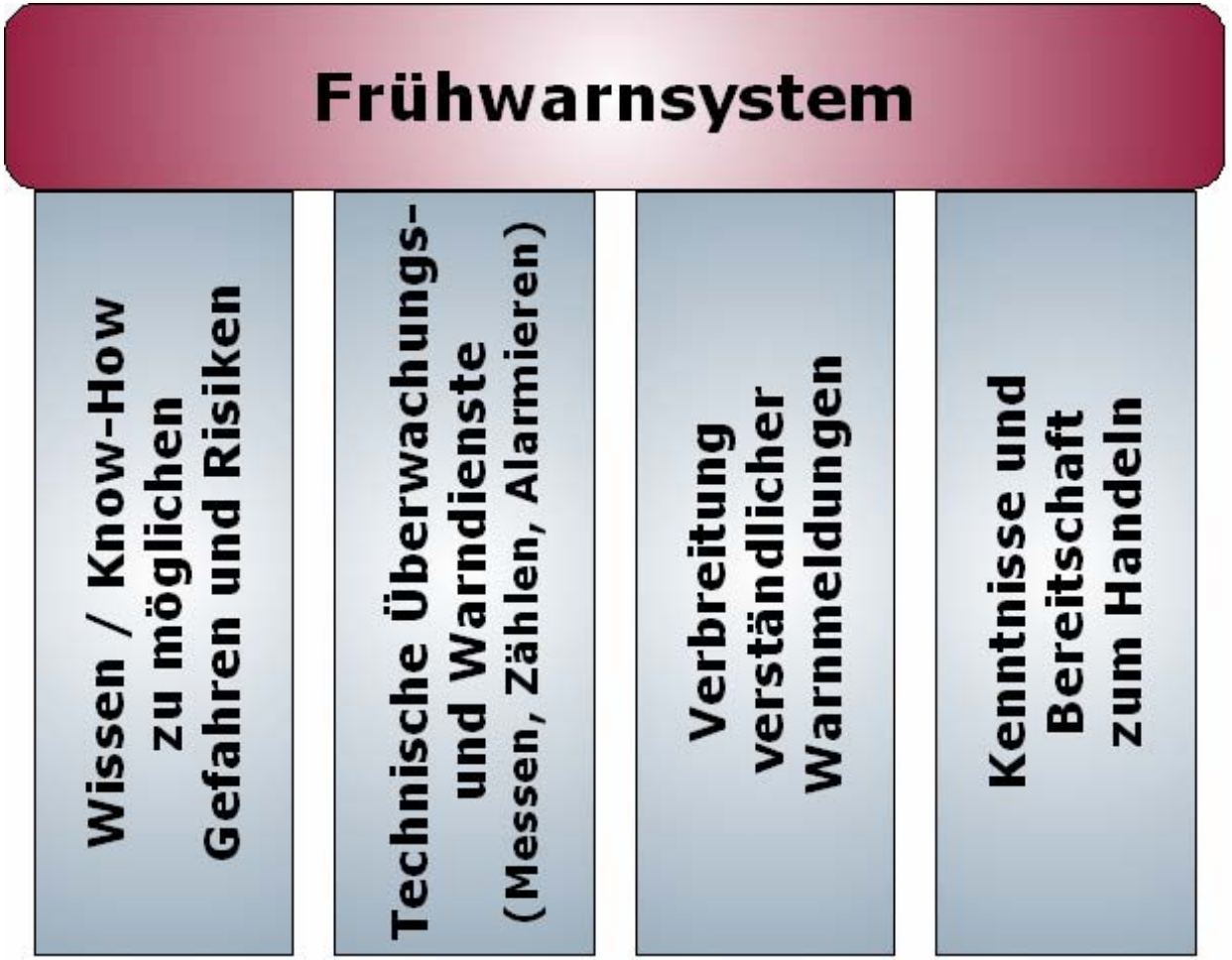
"...System weit reichender Radarstationen, mit dem feindliche Flugkörper frühzeitig erfasst werden."[2]



1) Quelle: Wikipedia
2) Quelle: Duden

Frühwarnsysteme (2/4)

- **Definition anhand eines 4-Säulen-Modells**

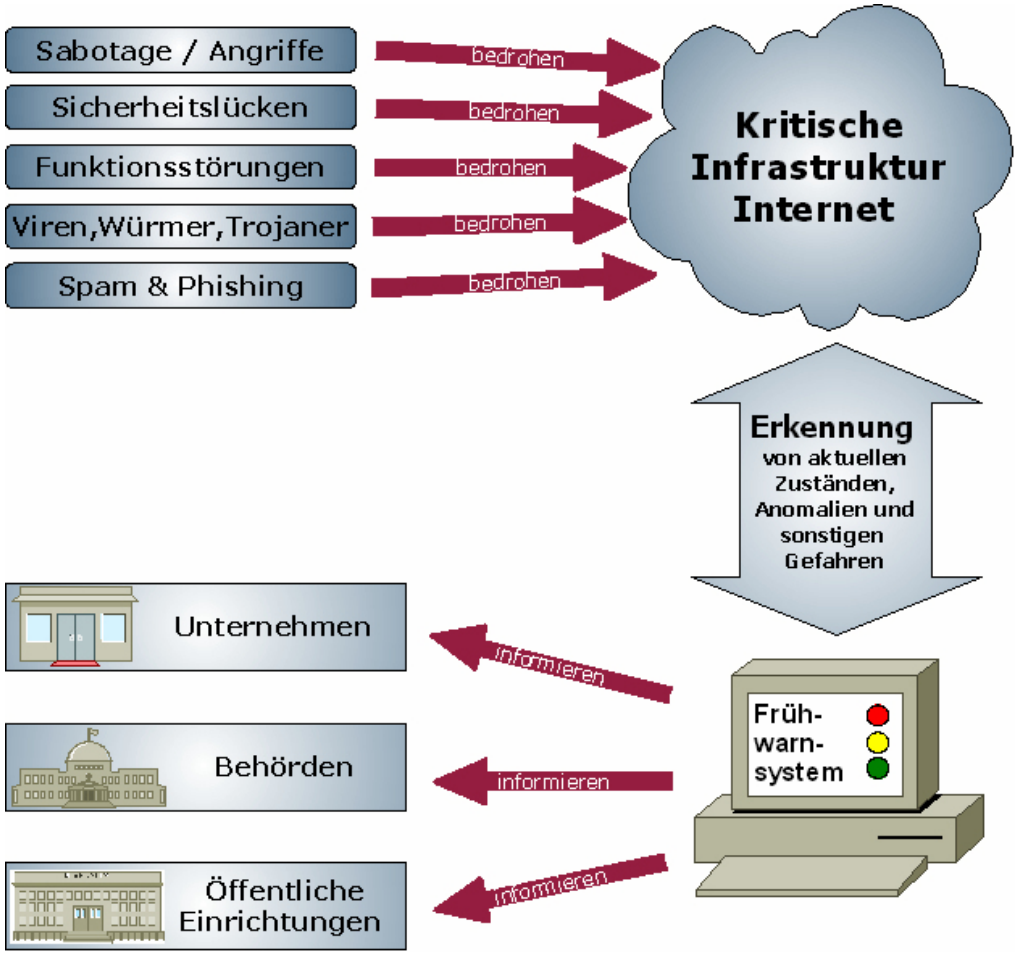


Frühwarnsysteme (3/4)

- **Bekannte Systeme aus anderen Bereichen**
  - Deutscher Wetterdienst
  - Erdbebenerkennung
  - Vulkanaktivitäts-Beobachtung
- **Übertragbarkeit auf IT-Systeme**
  - Schützenswerte Einrichtungen
    geistiges Eigentum, Privatsphäre, Unternehmenswerte
  - Bedrohungen
    Viren, Würmer, Trojanische Pferde, gezielte Attacken, Spam, Phishing, etc.
  - Zeitliche Aspekte
    Sekunden? Minuten? 3-Stunden-Lösung?

Frühwarnsysteme (4/4)

- **Konzeptioneller Überblick**

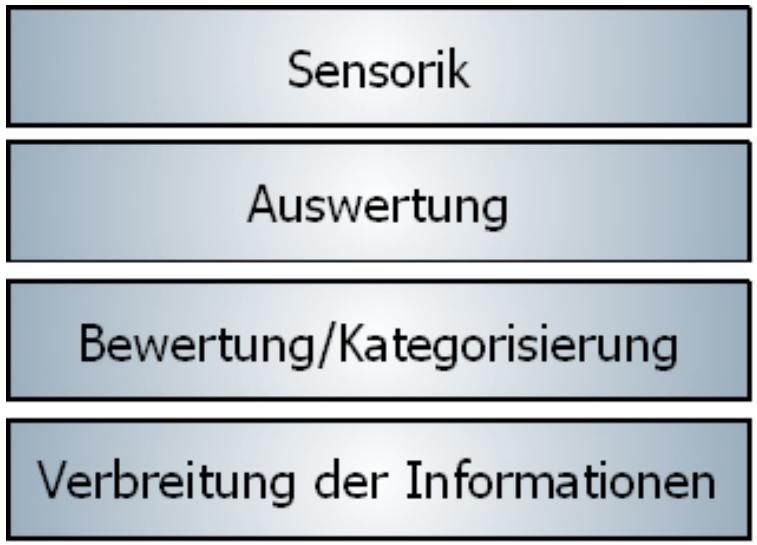


Agenda

- Einführung
- Frühwarnsysteme
- **Struktur für Internet-Frühwarnsysteme**
- Verschiedene Realisierungsansätze
- Internet-Analyse-System
- Internet-Verfügbarkeits-System
- Zusammenfassung

Struktur für Frühwarnsysteme
→ Ebenen - Modell (1 / 4)

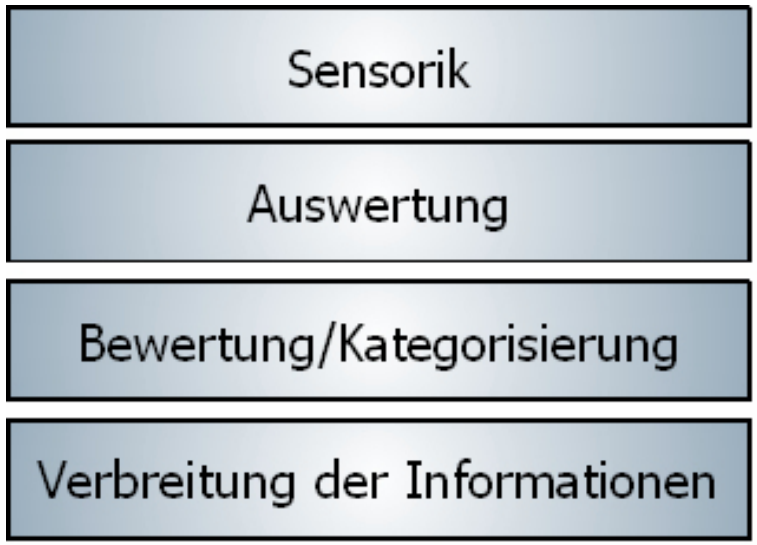
- Ebenen - Modell



- Komponenten und Techniken für die Erhebung der Rohdaten
- Berücksichtigung von
  - rechtlichen Aspekten
  - geographischer und logischer Verteilung

Struktur für Frühwarnsysteme
→ Ebenen - Modell (2 / 4)

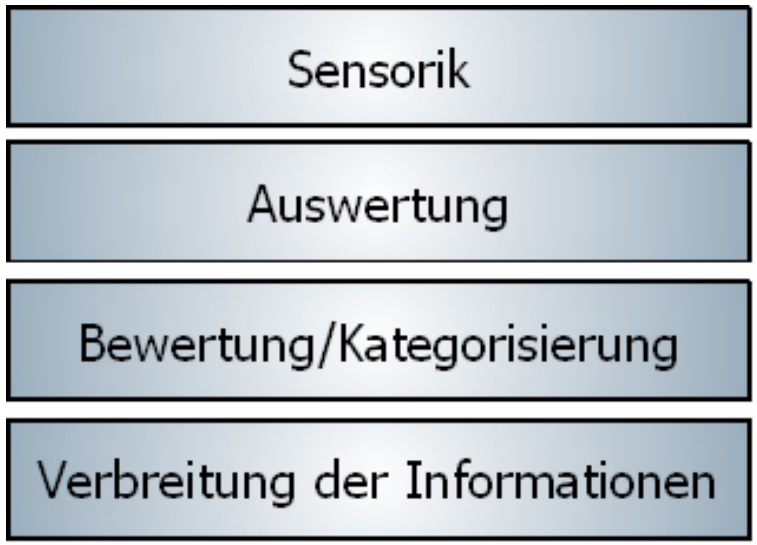
- Ebenen - Modell



- Weiterverarbeitung, Konsolidierung und Speicherung der Rohdaten
- Probleme bei
  - sehr großen Datenmengen
  - der Erkennung von unbekannten Anomalien und zukünftigen Trends

Struktur für Frühwarnsysteme
→ Ebenen - Modell (3 / 4)

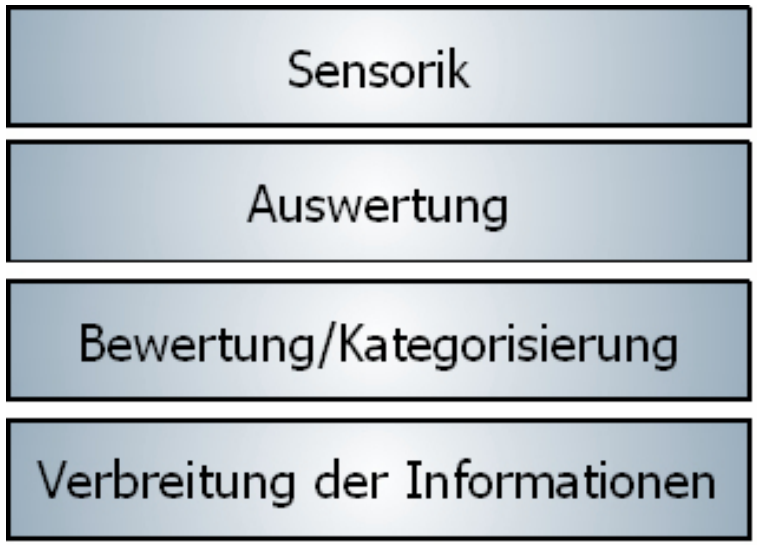
- Ebenen - Modell



- Einteilung der erkannten Gefahren in Schweregrade, Anomalie-klassen, Gefährdungskategorien oder Alarmierungsstufen
- Auswahl von Gegenmaßnahmen
- Probleme bei
  - automatischer Verarbeitung
  - Integration in best. Systeme

Struktur für Frühwarnsysteme
→ Ebenen - Modell (4/4)

- Ebenen - Modell



- Definition von Kommunikations-wegen und Adressaten
- Abgrenzung von Kompetenzen und Verantwortlichkeiten
- Definition von Austauschformaten
- Viele Besonderheiten und Probleme

Struktur für Frühwarnsysteme
→ Informationsverbreitungs-Ebene (1 / 2)

- CERTs (**C**omputer **E**mergency **R**esponse **T**eam)
  - Organisationen, die sich mit Computersicherheit befassen, Warnungen vor Sicherheitslücken herausgeben und Lösungsansätze bieten.
  - in Deutschland: CERT-Bund, Bürger-CERT, …
- CSIRTs (**C**omputer **S**ecurity **I**ncident **R**esponse **T**eam)
  - Organisationen, die für die Entgegennahme und Überprüfung von Computersicherheitsvorfällen verantwortlich sind und diese Aktivitäten in Berichten dokumentieren.
  - In Deutschland vertreten durch: DFN, CERT-Bund, …
  - In Europa: eCSIRT, TF-CSIRT, …

Struktur für Frühwarnsysteme
→ Informationsverbreitungs-Ebene (2/2)

- **Standard-Austausch-Formate**
  - Deutsches Advisory Format (DAF)
  - Intrusion Detection Message Exchange Format (IDMEF)
  - Incident Object Description and Exchange Format (IOMEF)
- Detaillierungsgrad, Form und Sprache der Warnungen müssen den Adressaten angepasst werden
- Sicherheitsbewusstsein und Bereitschaft zum Handeln

Struktur für Frühwarnsysteme
→ CERT-Bund (1/2)

- Ziel vom **CERT-Bund** ist die Bereitstellung einer zentralen Anlaufstelle für präventive und reaktive Maßnahmen in Bezug auf sicherheits- und verfügbarkeitsrelevante Vorfälle in Computer-Systemen.
- **Aufgaben:**
  - Erstellen und veröffentlichen von präventiven Handlungsempfehlungen zur Schadensvermeidung
  - Hinweisen auf Schwachstellen in Hardware- und Software-Produkten
  - Vorschlagen von Maßnahmen zur Behebung von bekannten Sicherheitslücken
  - Warnen/alarmieren bei besonderen Bedrohungslagen
  - Empfehlen von reaktiven Maßnahmen zur Schadensbegrenzung und/oder -beseitigung

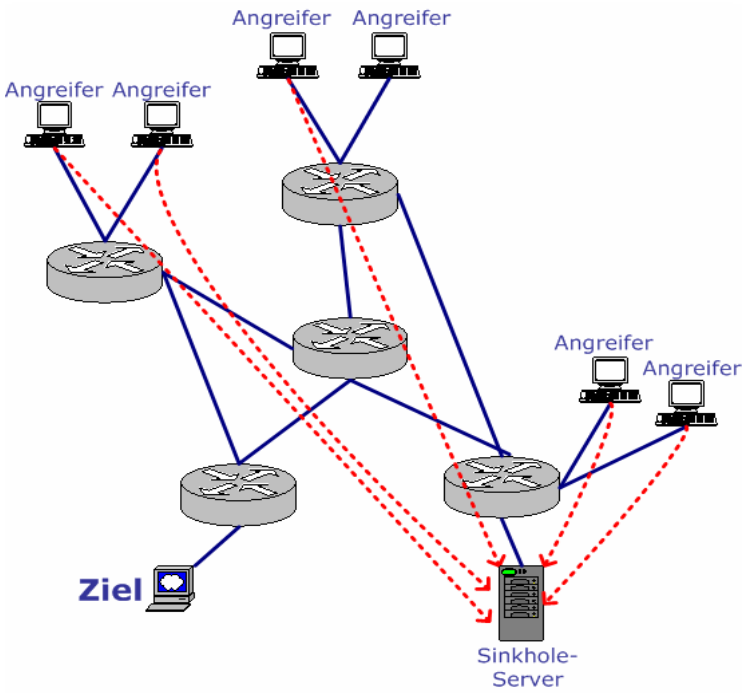
Struktur für Frühwarnsysteme
→ CERT-Bund (2/2)

- Die Dienstleistungen von CERT-Bund stehen in erster Linie den **Bundesbehörden** zur Verfügung und umfassen derzeit:
  - Eine 24-Stunden Rufbereitschaft
  - Den Betrieb eines Lagezentrums
  - Die Analyse eingehender Vorfallsmeldungen
  - Die Erstellung daraus abgeleiteter Empfehlungen
  - Das Betreiben eines Warn- und Informationsdienstes
  - Die aktive Alarmierung der Bundesverwaltung bei akuten Gefährdungen
- Benutzt wird das vom CERT-Bund entwickelte Vorfallsbearbeitungssystem **SIRIOS**.

Struktur für Frühwarnsysteme
→ Reaktionsmöglichkeiten

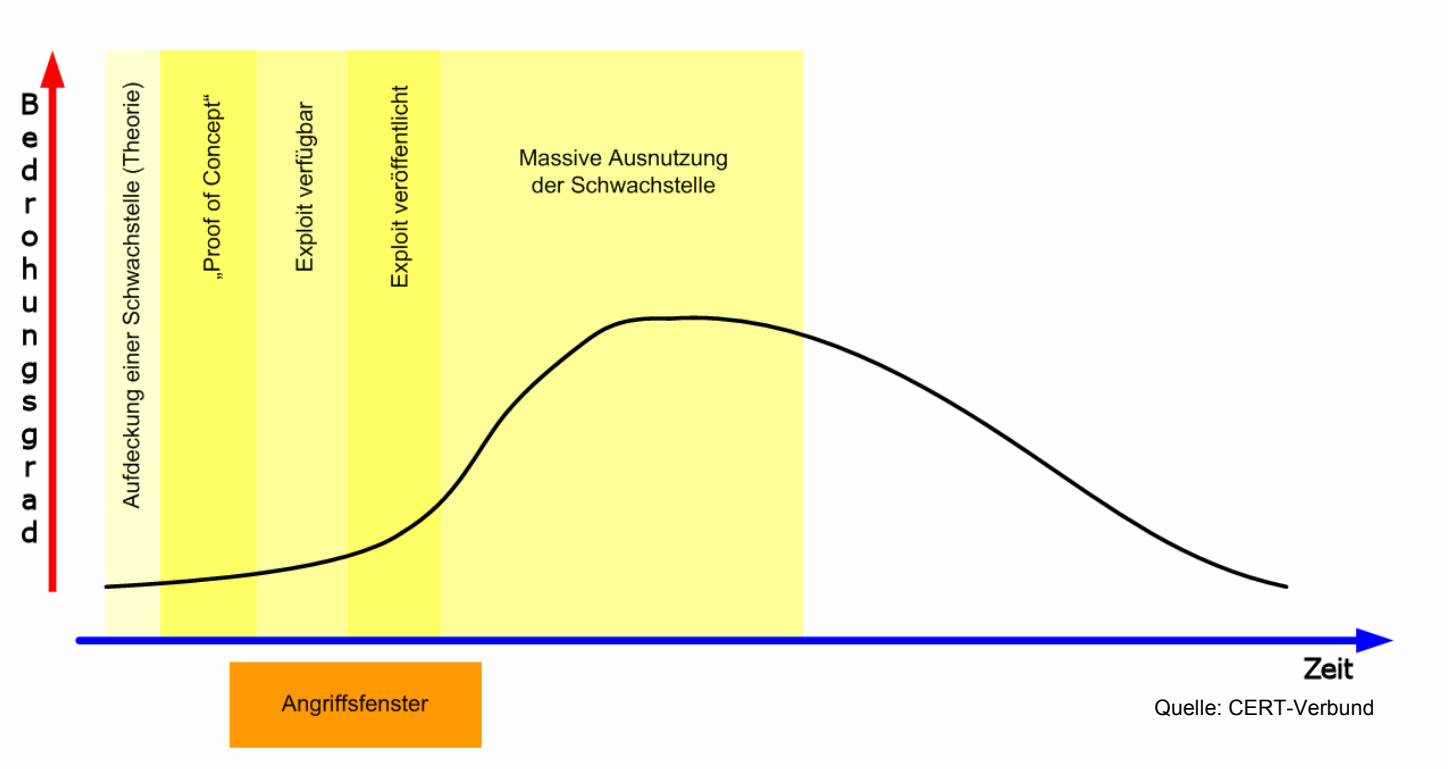
- **Privatnutzer / Unternehmen**
  - Erhöhen der Sicherheitsvorkehrungen
  - Selektives Abschalten betroffener Dienste
  - Komplette Deaktivierung des Internet-Zugangs

- **Internet Service Provider**
  - Access Controll Lists
  - Rate-Limiting
  - Blackholing
  - Off-Ramping / Sinkholing



Struktur für Frühwarnsysteme
→ Verkürzung der Reaktionszeiten

- Mehr Sicherheitslecks durch immer komplexere IT-Systeme
- Schnellere Entwicklung von Schadcode / Exploits

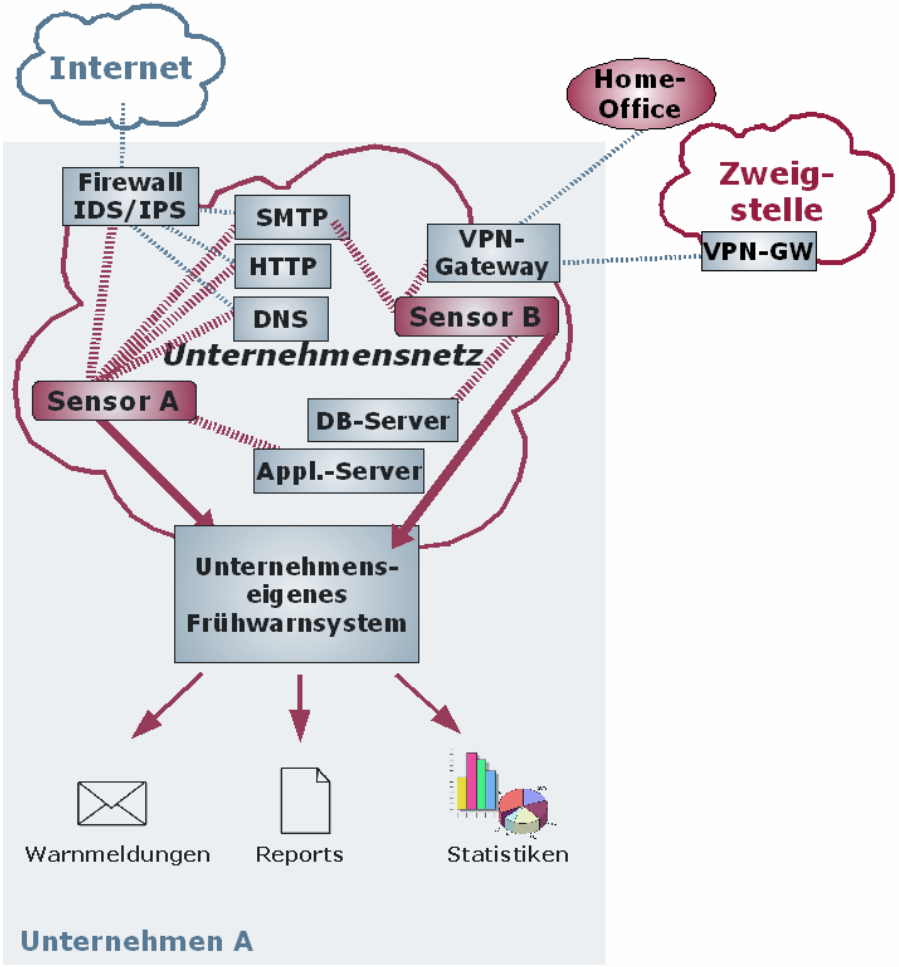


Agenda

- Einführung
- Frühwarnsysteme
- Struktur für Internet-Frühwarnsysteme
- **Verschiedene Realisierungsansätze**
- Internet-Analyse-System
- Internet-Verfügbarkeits-System
- Zusammenfassung

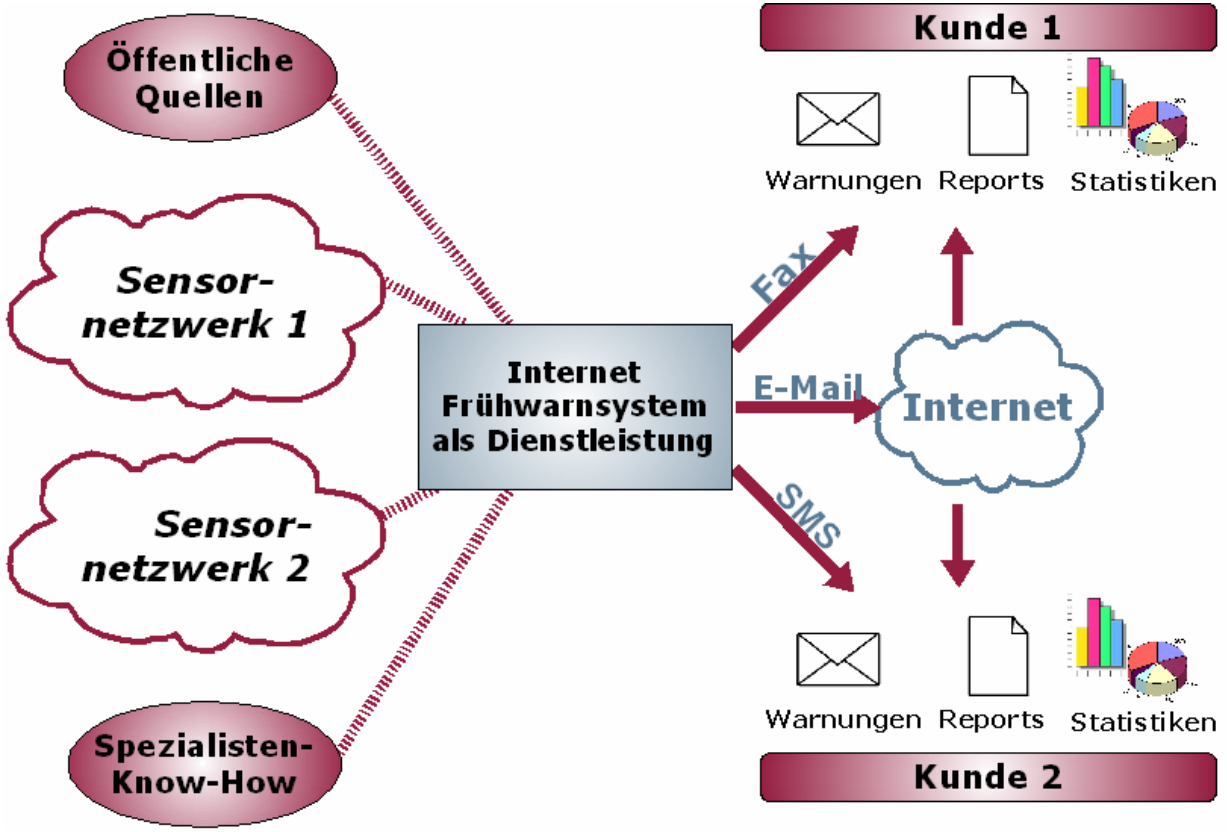
Realisierungsansätze
→ Lokale Installation

- **Lokale Installation**



Realisierungsansätze
→ Global verfügbares System

- Global verfügbares System



Realisierungsansätze
→ Logdaten-basierte Systeme

- **Rohdatengewinnung aus aktiven Komponenten im Internet**
  (Router, Switches, Intrusion-Detection- und –Prevention-Systemen, Firewalls, Webservern, Honeypots, Security-Appliances, etc.)
- **Auswertung von Logfiles und Protokolldaten**

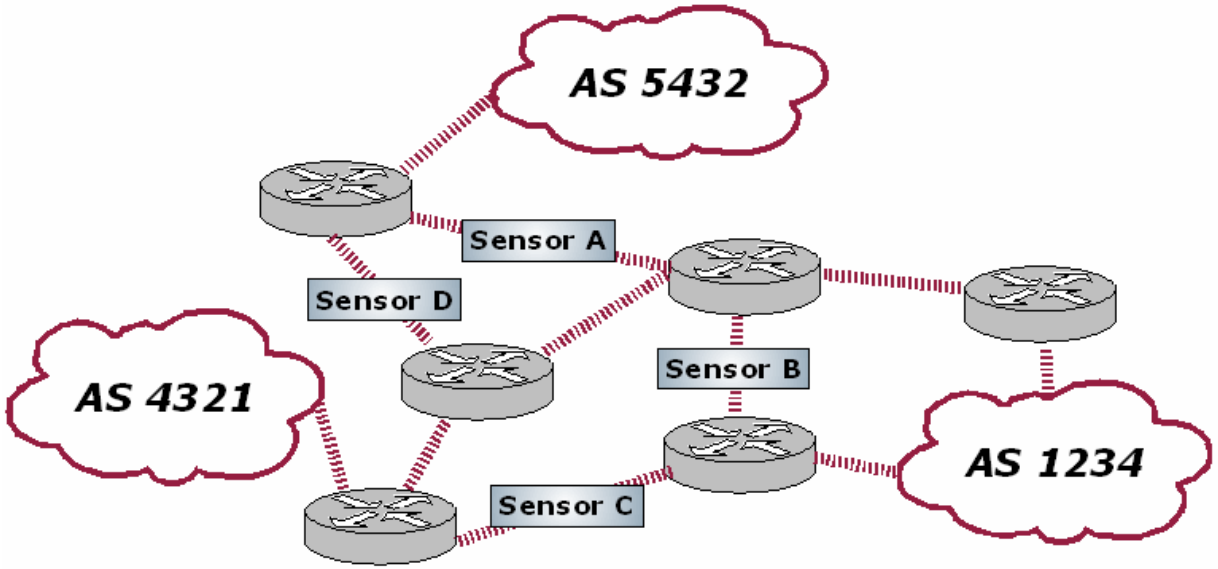
```
Feb 12 16:26:49 ipcop kernel: INPUT IN=ppp0 OUT= MAC=
SRC=X.X.X.X DST=Y.Y.Y.Y LEN=64 TOS=0x00 PREC=0x00 TTL=44
ID=18033 DF PROTO=TCP SPT=2285 DPT=139 WINDOW=53760
RES=0x00 SYN URGP=0
```

- **Auswertung von NetFlow- und SNMP-Daten**
  - Standard-Formate für Verkehrsfluss- und Statistik-Informationen

Realisierungsansätze
→ Datenerheb. durch Eigenentwicklungen

- **Speziell angepasste Sensor-Systeme**
  Paketdaten-Analyse, Dienstgüte-Beobachtung, Verfügbarkeitsprüfungen, etc.

- **Beispiel:**



Realisierungsansätze
→ Existierende Systeme

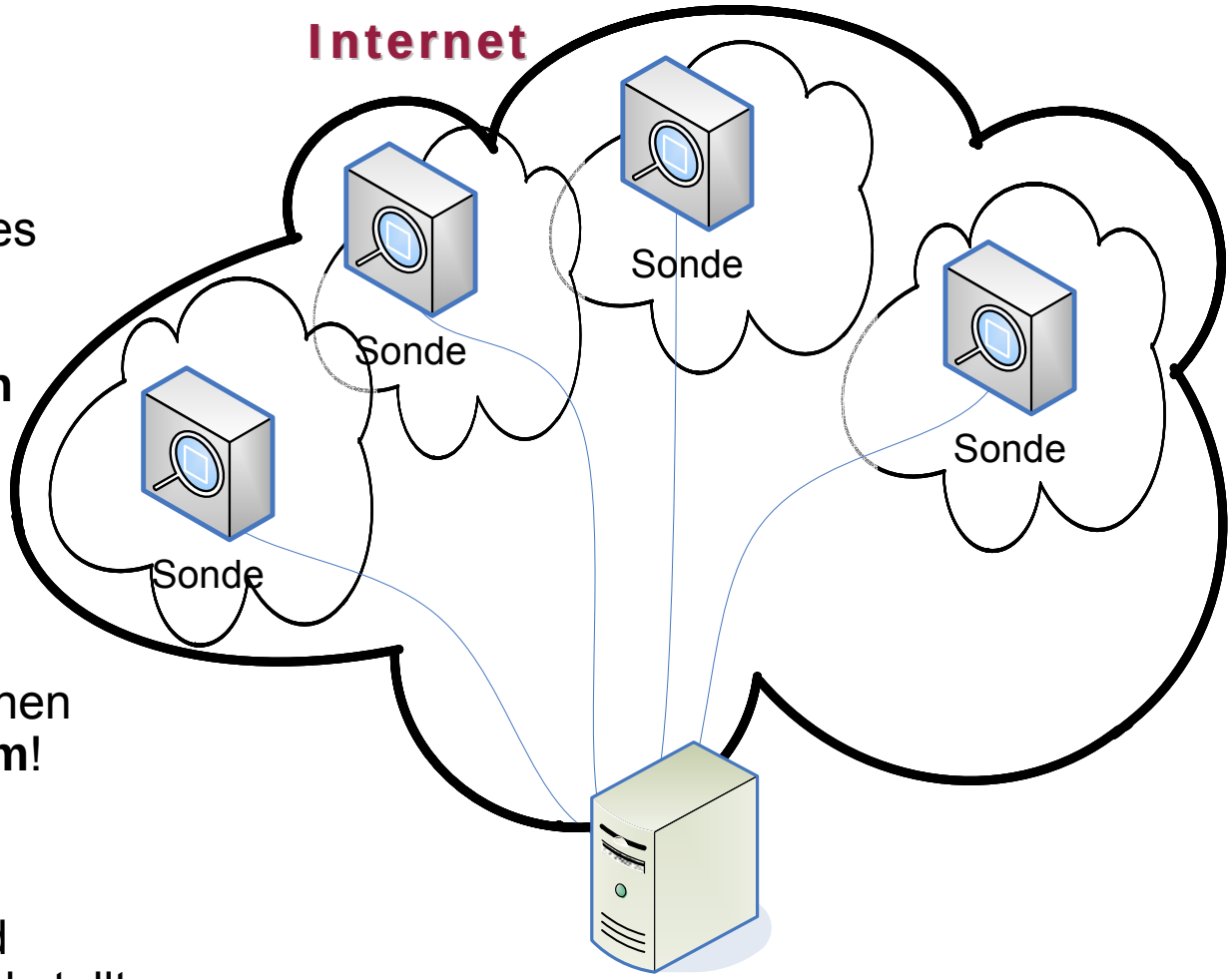
- Symantec DeepSight Threat Management System
- X-Force Threat Analysis Service von ISS
- Arbor Networks Peakflow X / SP
- Computer Associates – eTrust Network Forensics
- DShield.org – Distributed Intrusion Detection System
- **Internet-Analyse-System des ifis**
- **Internet-Verfügbarkeits-System des ifis**
- CarmentiS-Initiative des CERT-Verbunds

Agenda

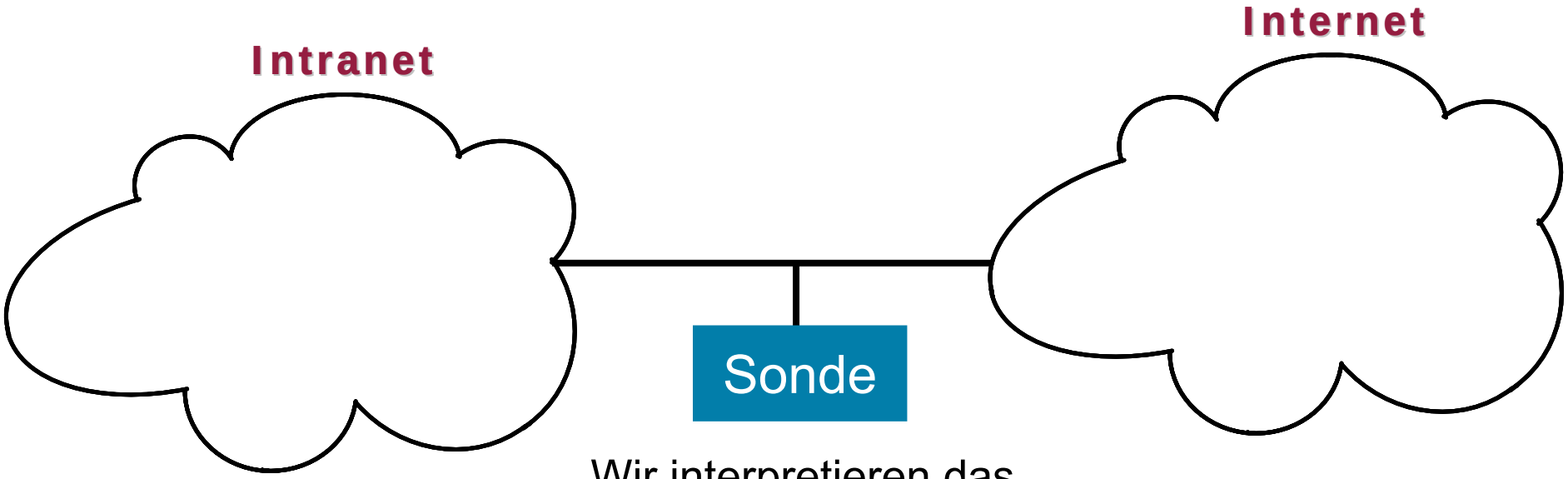
- Einführung
- Frühwarnsysteme
- Struktur für Internet-Frühwarnsysteme
- Verschiedene Realisierungsansätze
- **Internet-Analyse-System**
- Internet-Verfügbarkeits-System
- Zusammenfassung

Internet-Analyse-System (IAS)
→ Idee

- Beobachtung der kritischen Infrastruktur **„Internet“.**
- **Sonden** werden an ausgesuchten Positionen des Internets zur Erfassung von Rohdaten in die **Kommunikationsleitungen** eingebunden.
- Zählen von Header-Informationen, die **nicht datenschutzrelevant** sind.
- System sammelt Informationen über einen **großen Zeitraum**!
- Ein zentrales **Auswertungssystem** analysiert die Rohdaten und Auswertungsergebnisse und stellt diese **umfangreichen Ergebnisse** dar.

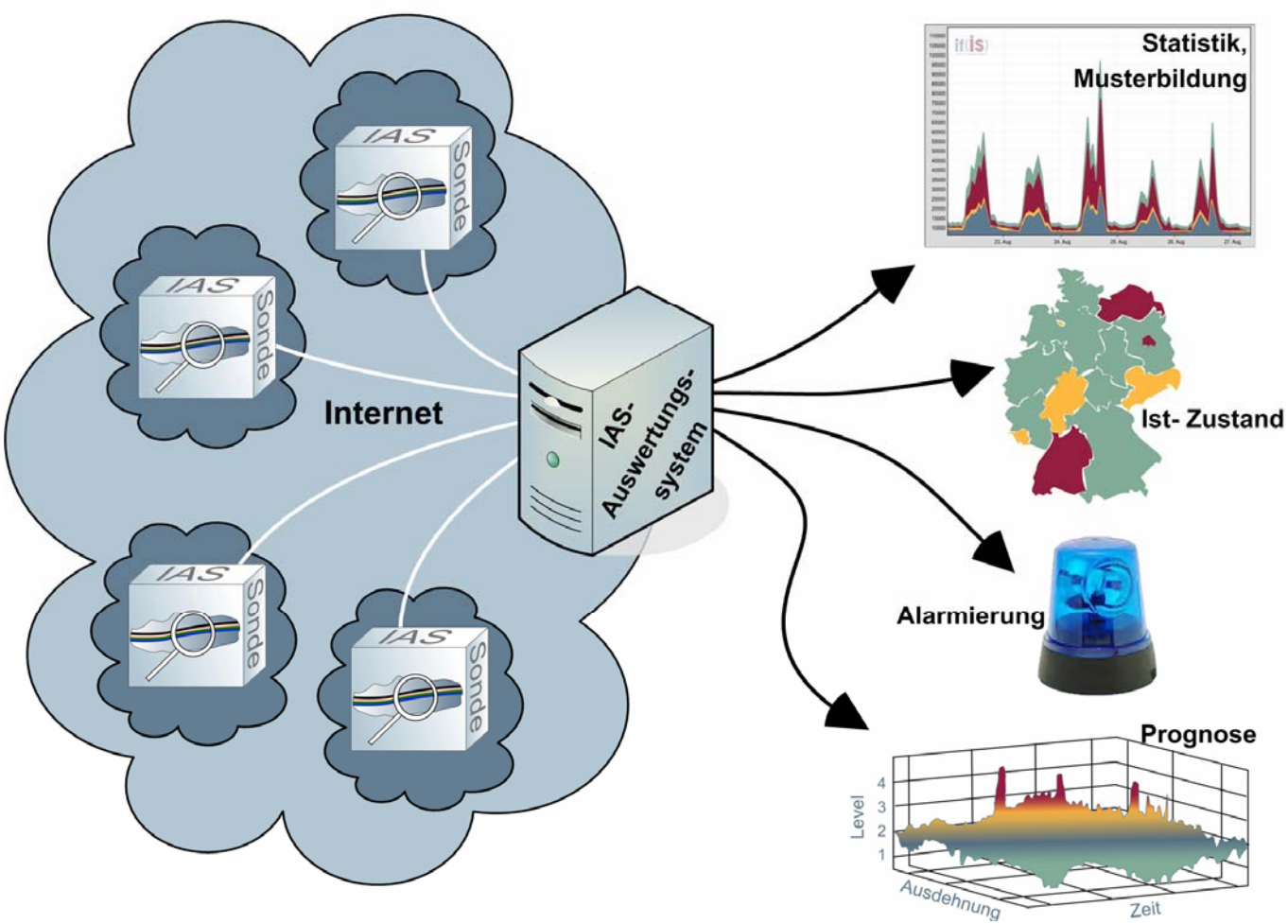


Sonde
→ Grundsätzlicher Ansatz



Wir interpretieren das
Kommunikationsverhalten zwischen
dem Intranet eines Unternehmens und dem Internet

Internet-Analyse-System (IAS)
→ Ziele



1) Beschreibung von Profilen, Mustern, Technologietrends und Zusammenhängen.

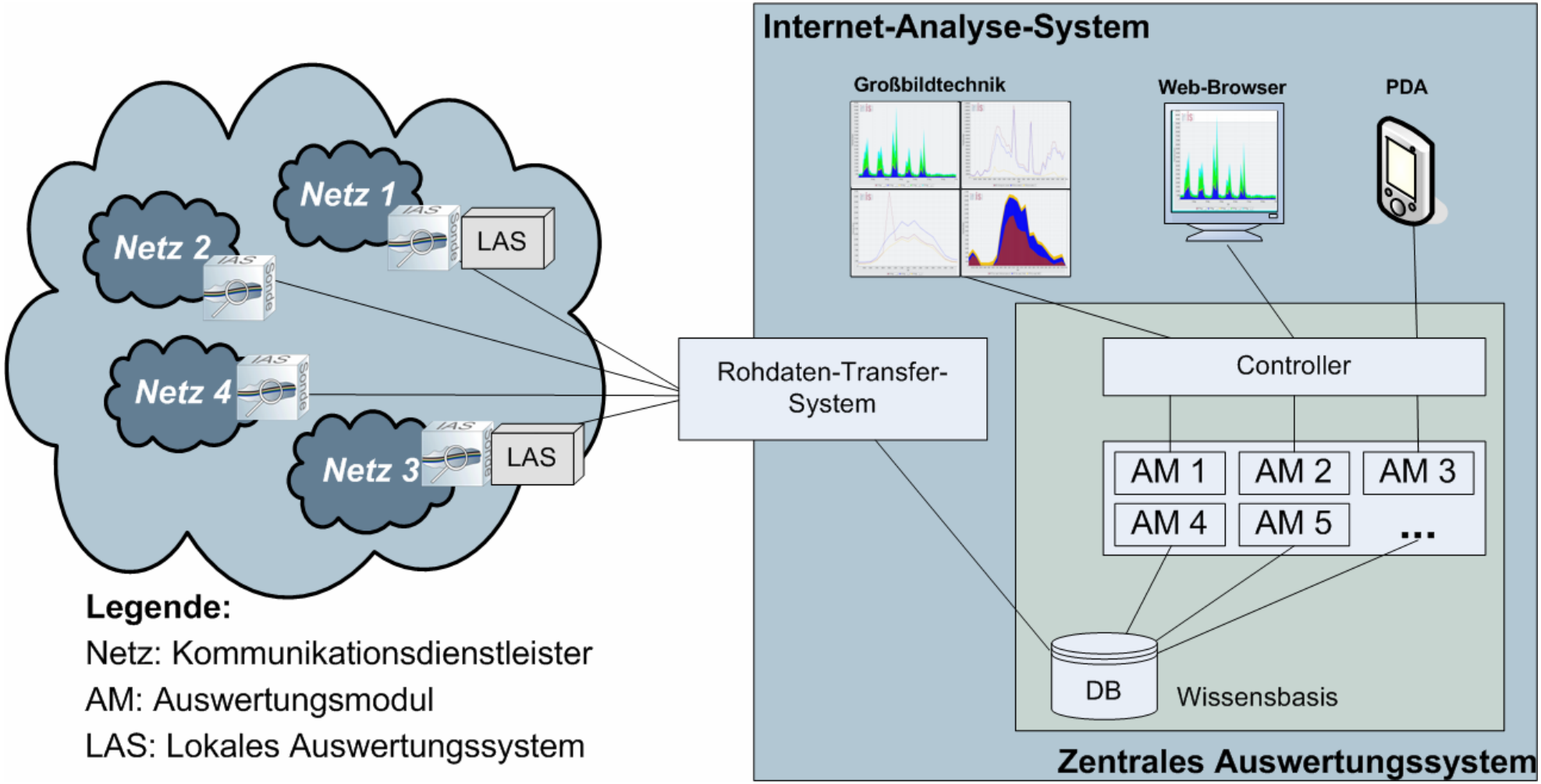
Schaffung einer **Wissensbasis.**

2) Überblick über den **aktuellen Zustand** des Internets

3) Erkennen von **Angriffssituationen** und Anomalien

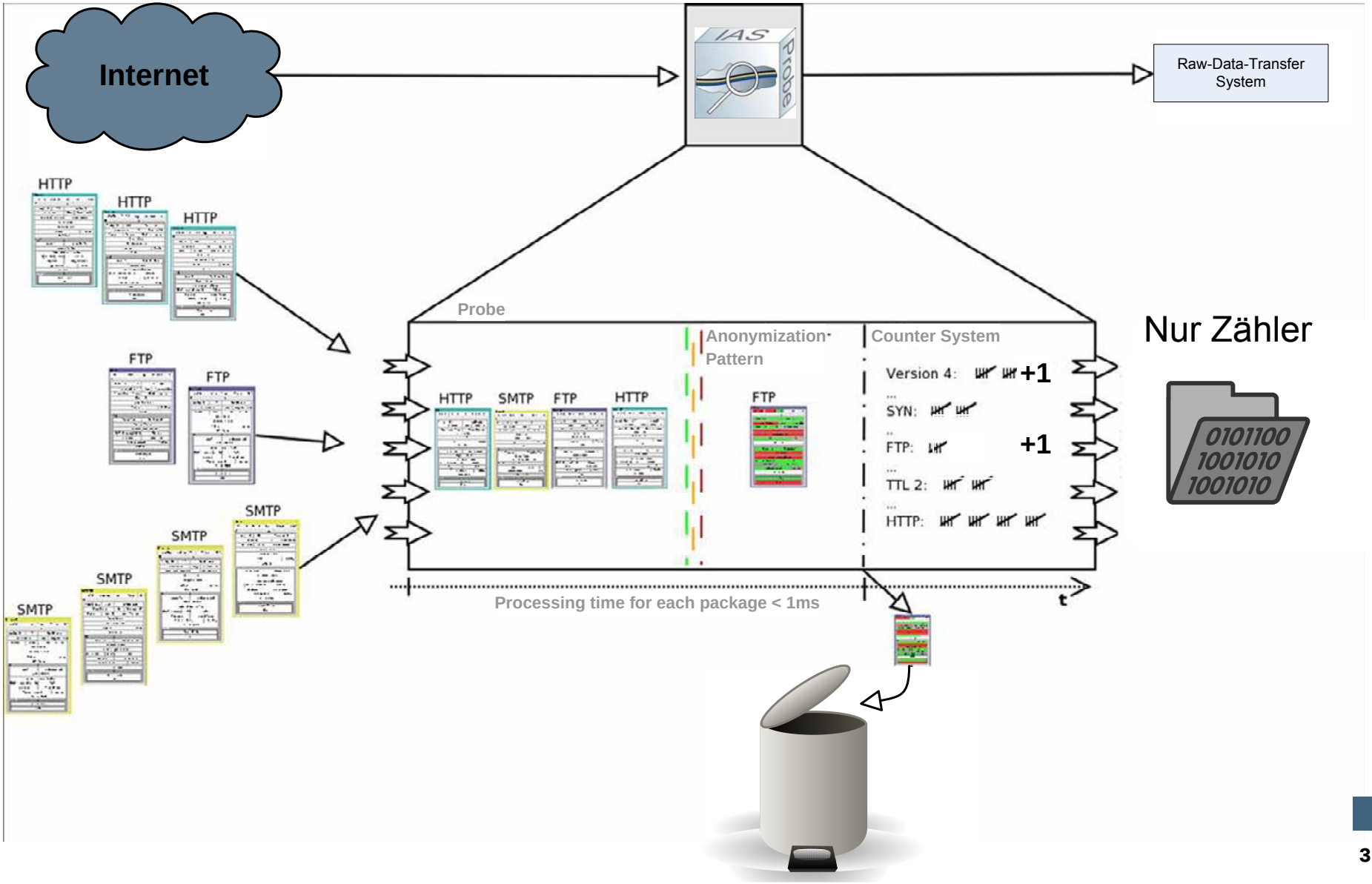
4) **Prognosen** von Mustern und Angriffen

Umsetzung des IAS
→ Übersicht



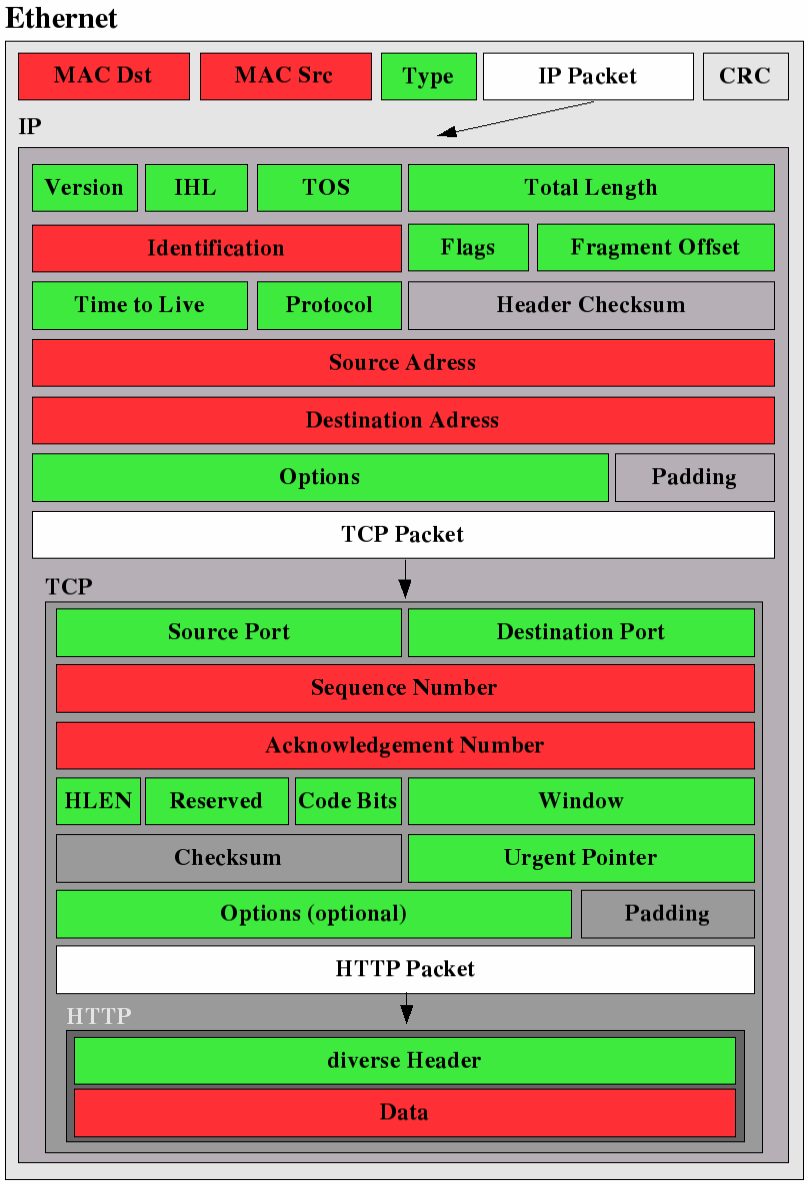
Umsetzung des IAS

→ Prinzip: Zählen von Header-Informationen



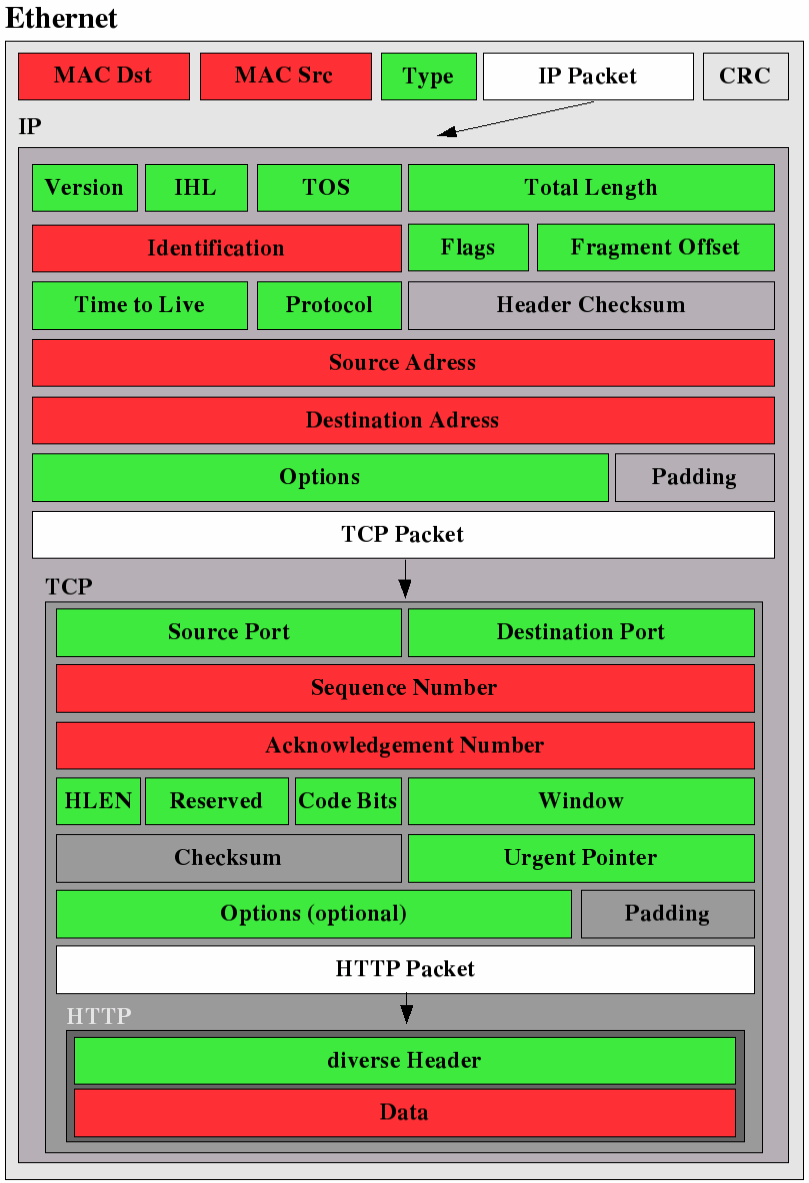
Prinzip der Rohdatengenerierung (1/2) → Protokollverschachtelung

- **Ethernet**
  - Type: Typ des geschachtelten Paketes, hier: 0x0800 (IP)
  - Checksumme (CRC) irrelevant
- **Internet Protocol**
  - z.B.: Total Length: Länge des gesamten Paketes
  - Protocol: Typ des geschachtelten Paketes, hier: 6 (TCP)
  - Quell- und Zieladressen datenschutzrechtlich bedenklich



Prinzip der Rohdatengenerierung (2/2) → Protokollverschachtelung

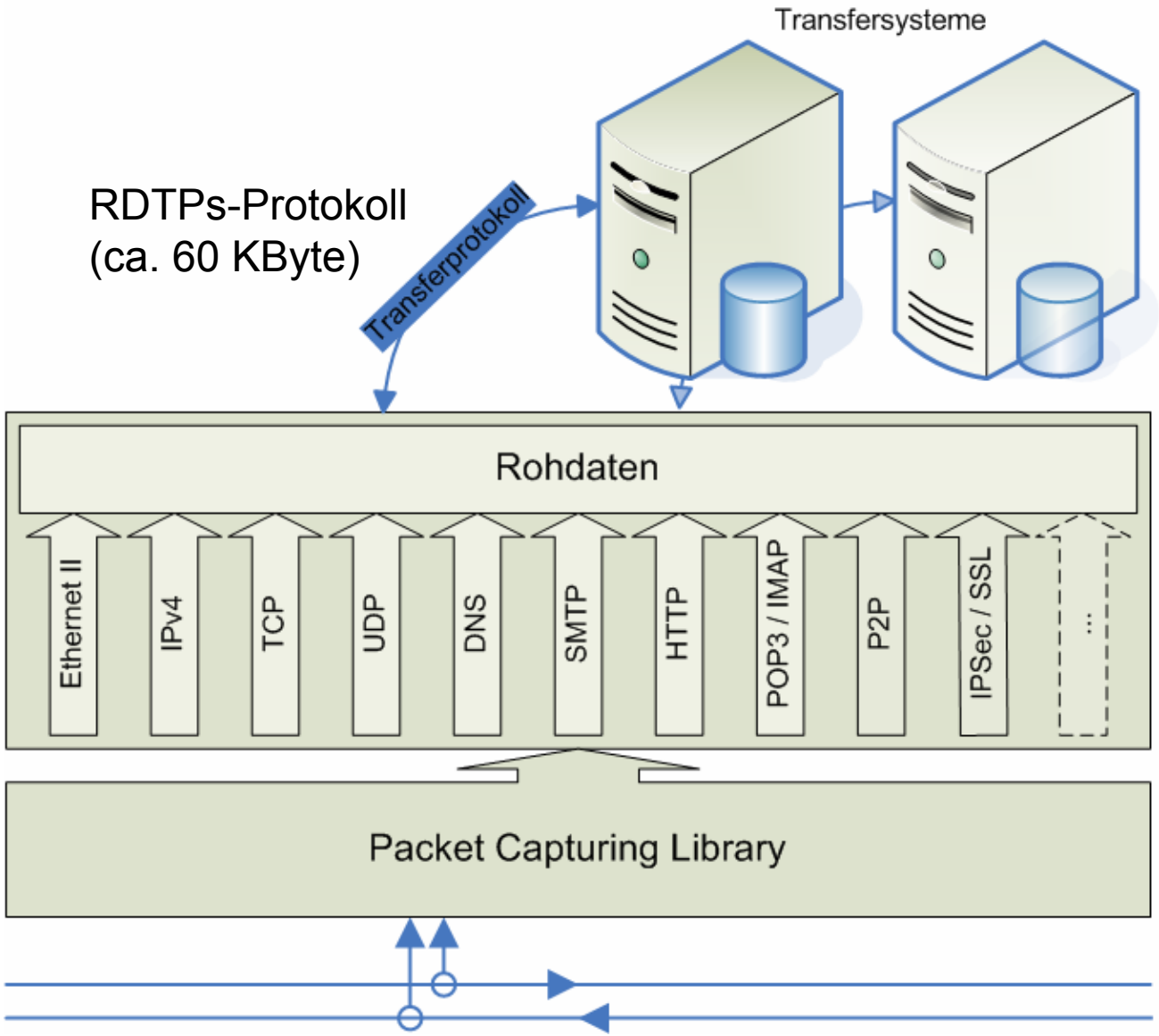
- **Transmission Control Protocol**
  - Port: Endpunkt einer Verbindung
    - HTTP: 80 (WWW)
    - Weitere z.B.: SMTP (25), HTTPS (443)
  - Code Bits
    - Informationen über Verbindungs-aufbau oder -abbau
- **Hypertext Transfer Protocol**
  - Header:
    - z.B.: User Agent: beschreibt den Browser des Users
  - Benutzerdaten (DATA) z.B.: Inhalt einer Website



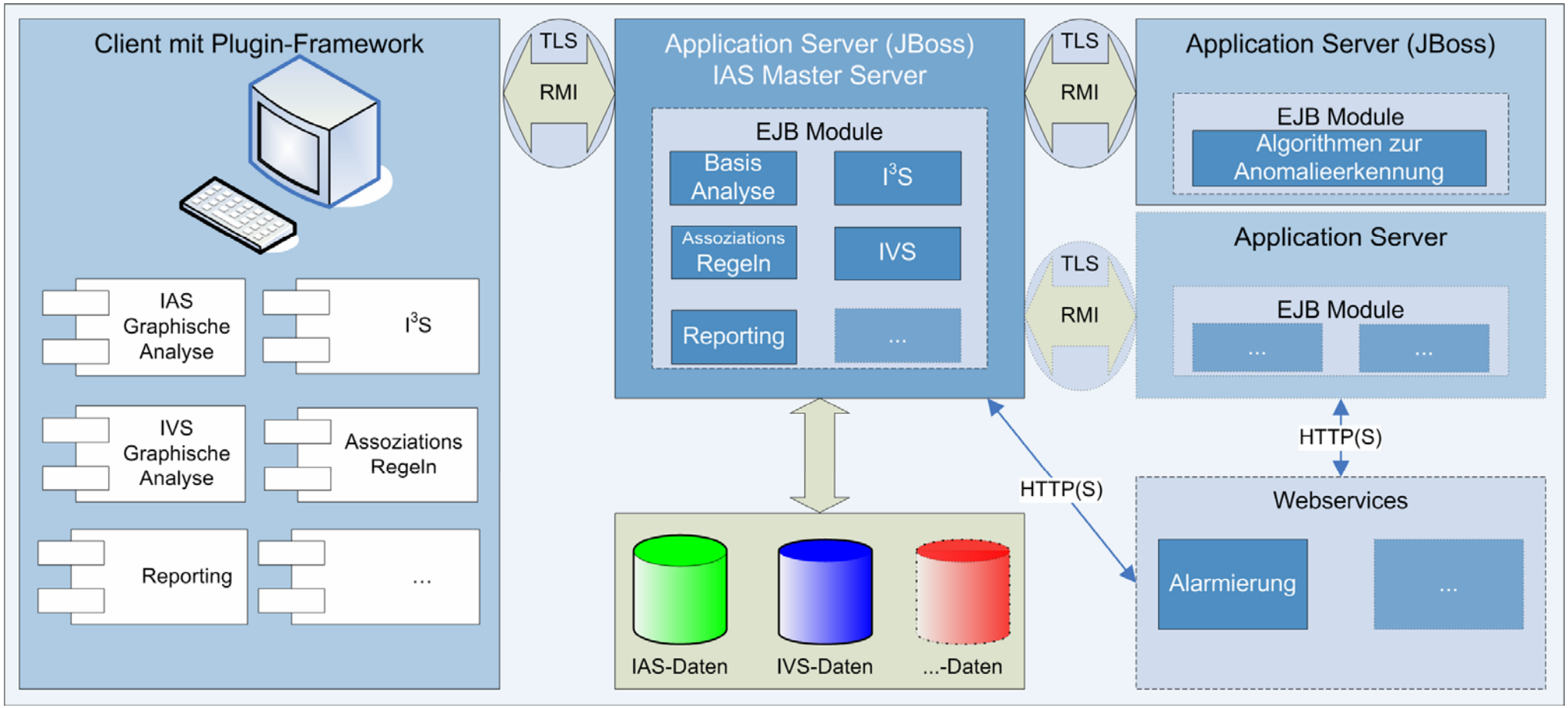
Prinzip der Rohdatengenerierung (4/4) → Warum nur „Strichlisten“

- **Erhöhte Performance**
  - Kein Connection / Session Tracking
  - Ignorieren irrelevanter Informationen
    - z.B.: Checksummen
- **Schutz schützenswerter Informationen**
  - Durch Nicht- Wiederherstellbarkeit der Connection / Session
  - Durch bewusstes „Weglassen“ der sensitiven Informationen
    - IP/ MAC Adressen
    - Userdaten
  - Anonymisierung per Konzept

Umsetzung des IAS
→ Sonde



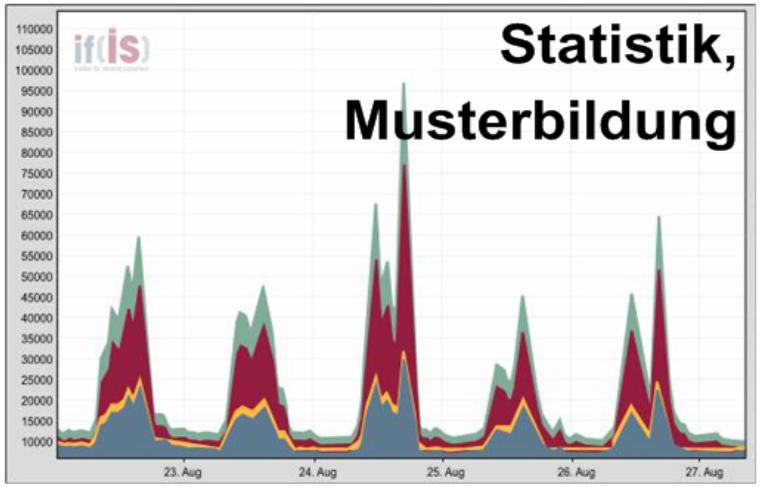
Umsetzung des IAS
→ Auswertungssystem



Internet-Analyse-System (IAS)
→ Definierte Ziele

Ziel 1

- **Beschreibung von Profilen, Mustern, Technologietrends und Zusammenhängen.**
- **Schaffung einer Wissensbasis**



Internet-Analyse-System (IAS)
→ Fortschritt Ziel 1

- Zählen von Kommunikationsparametern in den Sonden
- Senden der Zähler (Rohdaten) zum Transfersystem
- Langzeitspeicherung in einer Datenbank

Schaffung einer Wissensbasis

- Konservieren der Rohdaten in einer Datenbank
- Erfahrungsgewinn und Sammlung von Ergebnissen aus der Wissenbasis

Beschreibung von Profilen, Mustern, Technologietrends und Zusammenhängen.

- Analysieren der Rohdaten mit dem „***EagleX Analysis Client***“
  - Expertenwerkzeug, mit dem die Rohdaten „per Hand“ analysiert werden können.
- (Automatisierte) Generierung von Reporten

Internet-Analyse-System (IAS)
→ Zusammenhänge (1/2)

- **Architekturtechnisch Zusammenhänge**
  - Wenn http, dann auch TCP und IP
- **Protokolltechnisch Zusammenhänge**
  - Wenn http Request, dann http Response
- **Systemtechnisch Zusammenhänge**
  - Wenn http, dann i.d.R. auch eine DNS Anfrage
- **Verhaltensbedingt Zusammenhänge**
  - z.B. Wenn http, dann auch SMTP, d.h. wenn wir im Internet Surfen, schreiben wir i.d.R. auch E-Mails

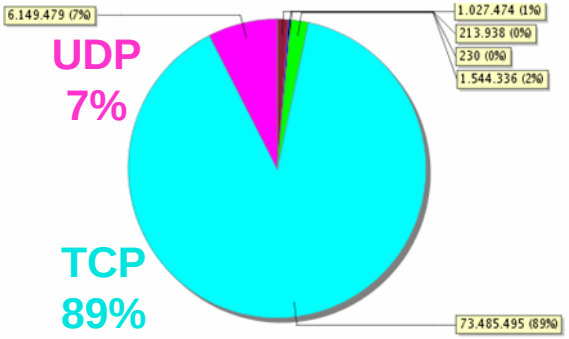
Internet-Analyse-System (IAS)
→ Zusammenhänge (2/2)

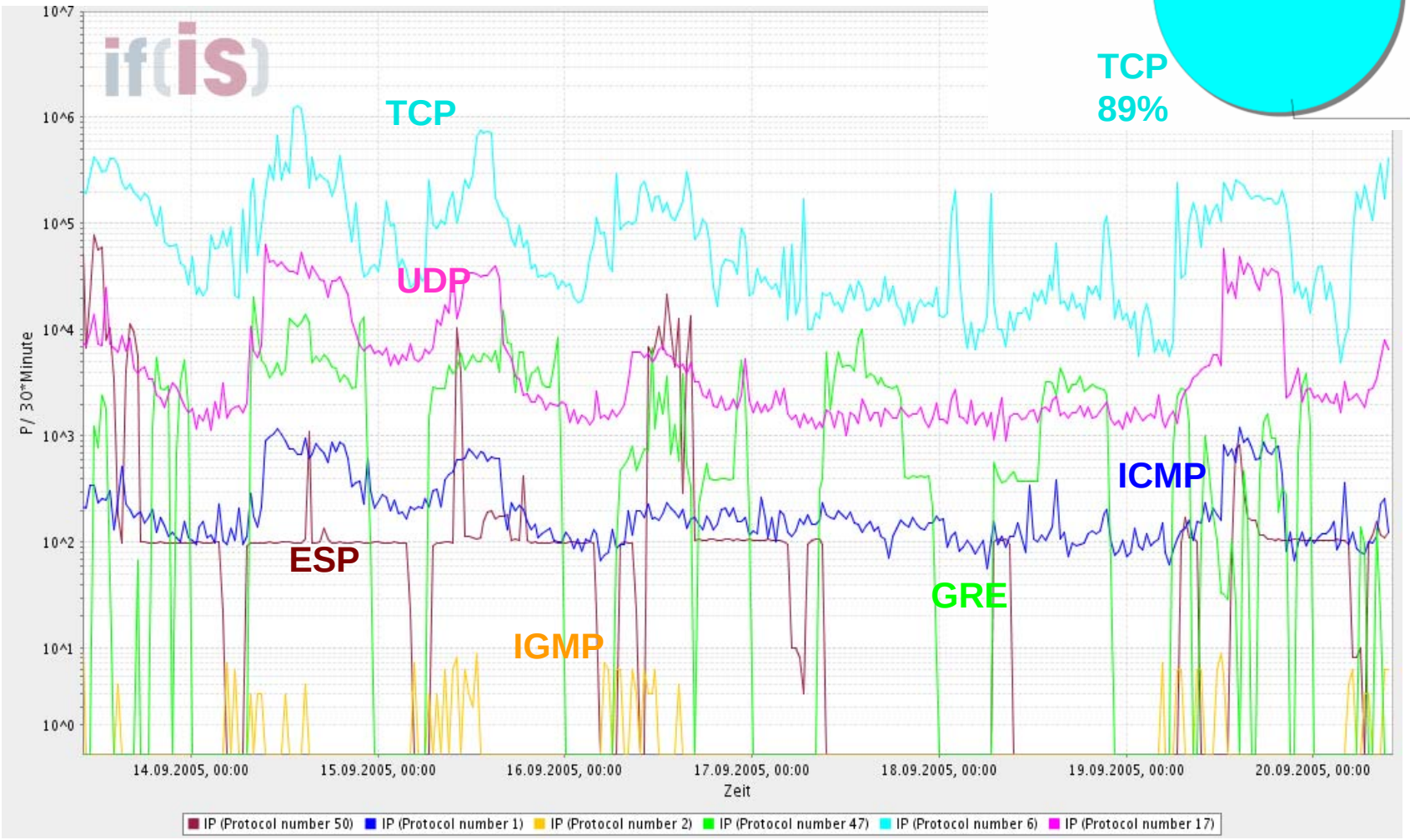
- **Situationsbedingt Zusammenhänge**
  - Wenn eine besondere Nachricht kommt, z.B. Anschlag in den USA, dann sehen wir bedeutend mehr Internet-Verkehr

- **Orts- und Anwendungsbedingt Zusammenhänge**
  - DSL-Provider, Content-Provider und Business-Anwender haben sehr unterschiedliche Verteilungen der Kommunikationsprotokolle.
  - Z.B. bei DSL-Provider haben wir sehr viel mehr p2p-Datenverkehr als Business-Anwender.

Beispiele von Ergebnissen des IAS
→ Wissensbasis: Erfahrungen

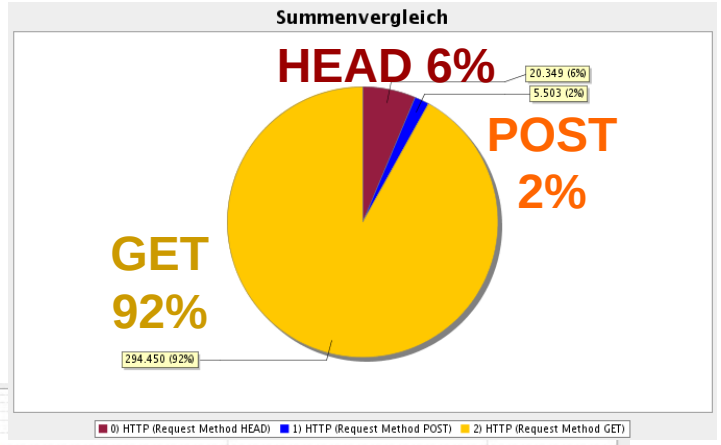
- **Transport-Protokoll Verteilung (Profil)**

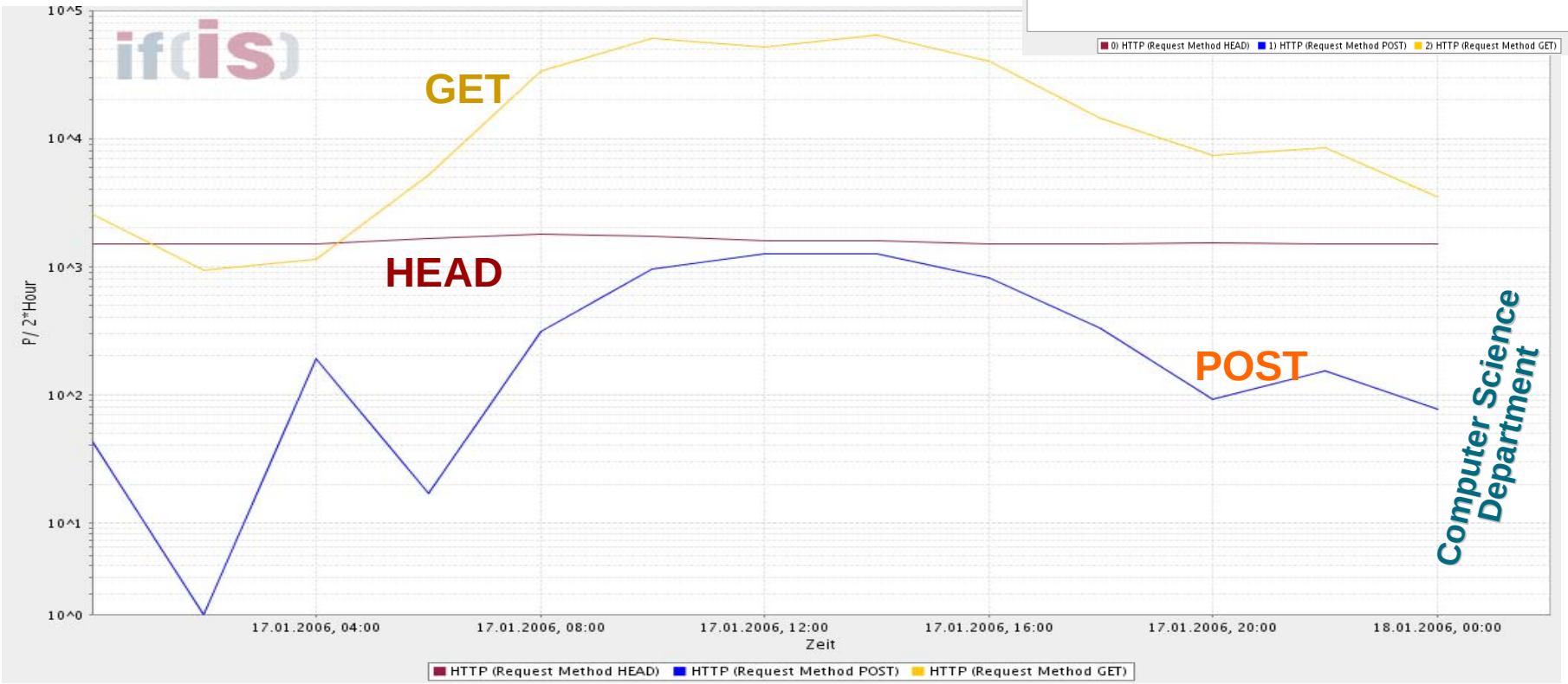




Beispiele von Ergebnissen des IAS
→ Wissensbasis: Erfahrungen

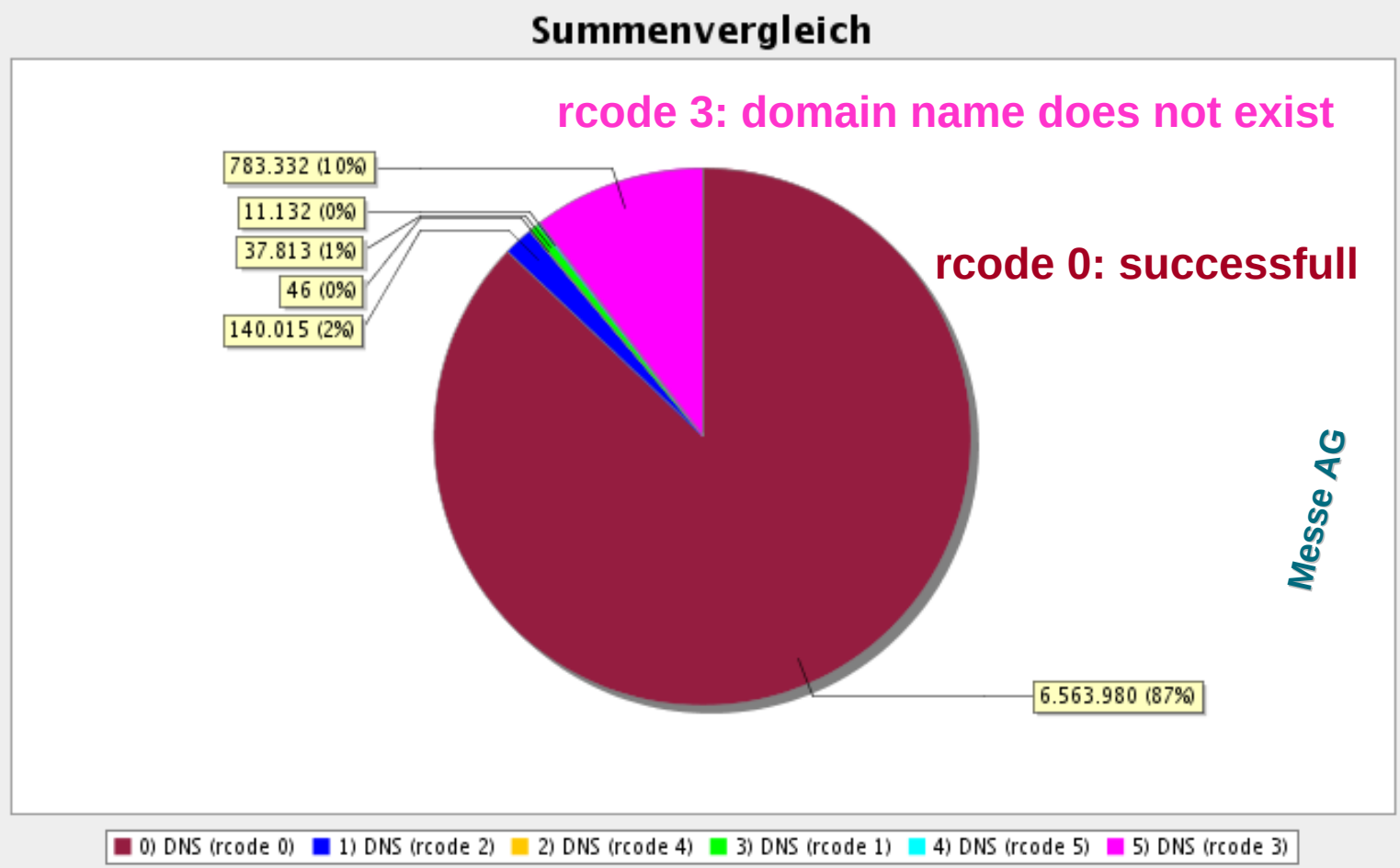
- **HTTP Methoden**
  - Tagesrhythmus
    - HEAD genutzt von automatischen Prozessen
    - GET und POST i.d.R. genutzt von menschlichen Nutzern





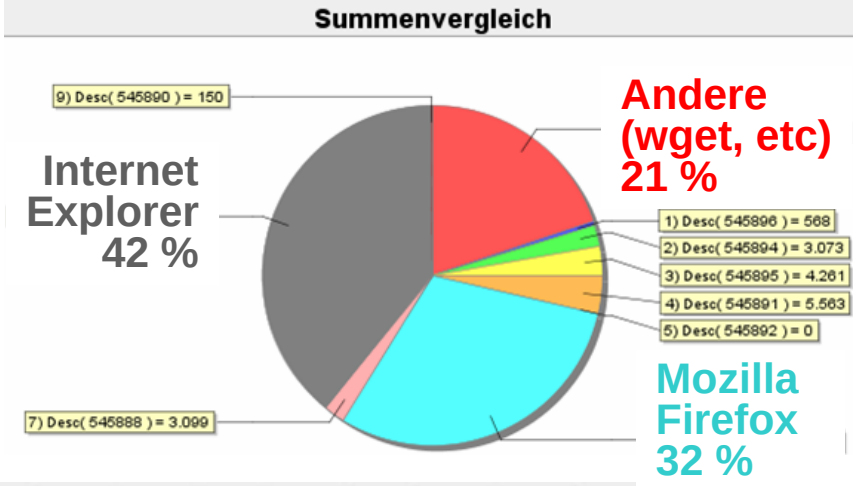
Beispiele von Ergebnissen des IAS
→ Wissensbasis: Erfahrung

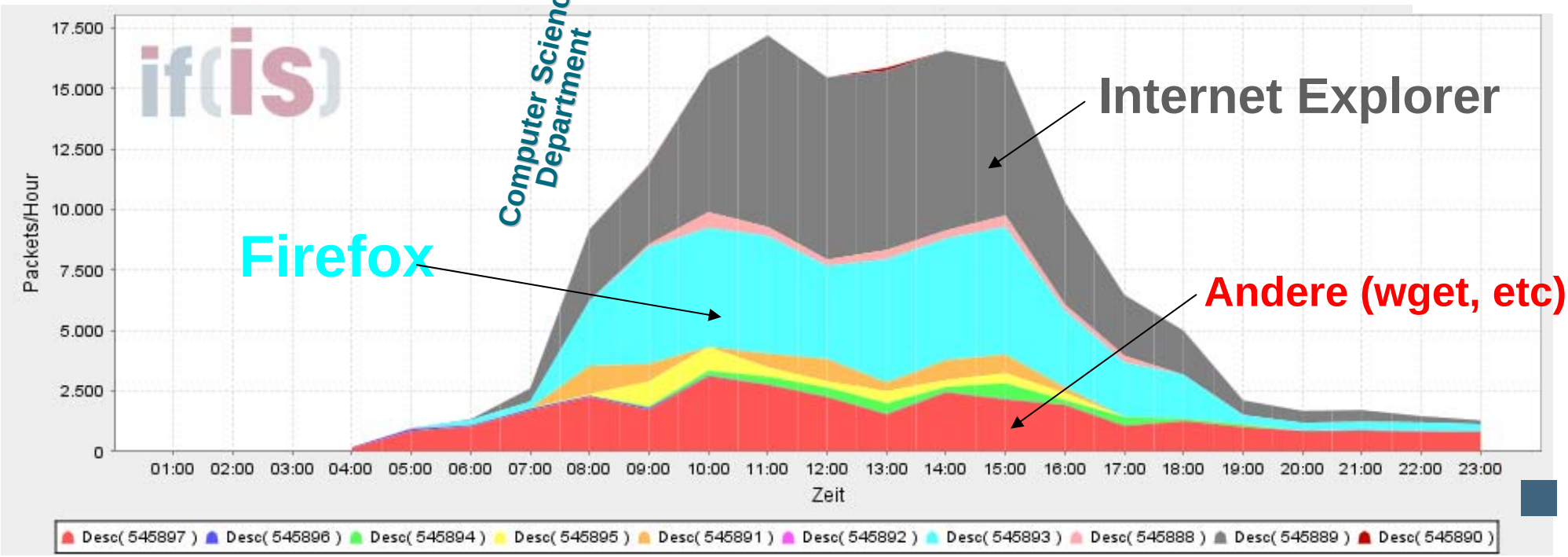
- **DNS Server Return Codes**
  - Normalerweise: Alles in Ordnung
  - Ca. 10%: Domain-Name nicht gefunden



Beispiele von Ergebnissen des IAS
→ Wissensbasis: Technologietrend

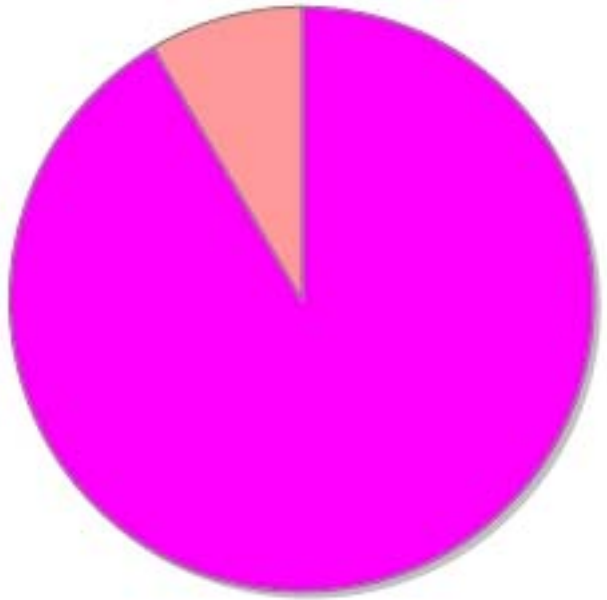
- **Browserverteilung (Technologietrend)**
  - Tagesprofile, Keine Serverstatistik, sondern „Leitungsstatistik“
  - Unterschied zwischen manueller Nutzung (z.B. Internet Explorer und Firefox) und automatischer Nutzung (z.B. wget) zu erkennen.





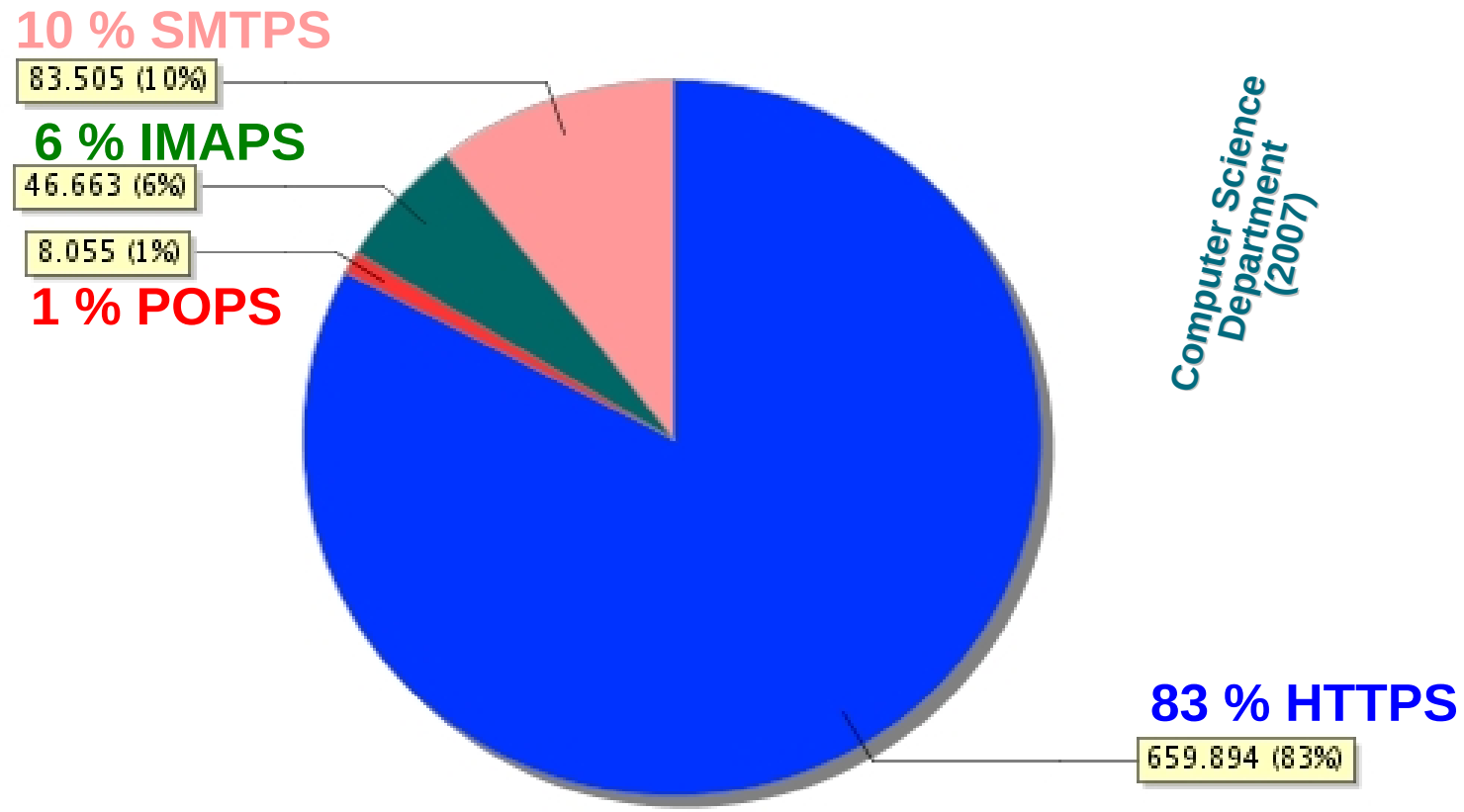
Verhältnis: SSL / TLS ⇔ ohne SSL / TLS

- Der Anteil an Protokollen, die SSL/TLS nutzen ist noch sehr gering.
- Bsp. Verhältnis http zu https

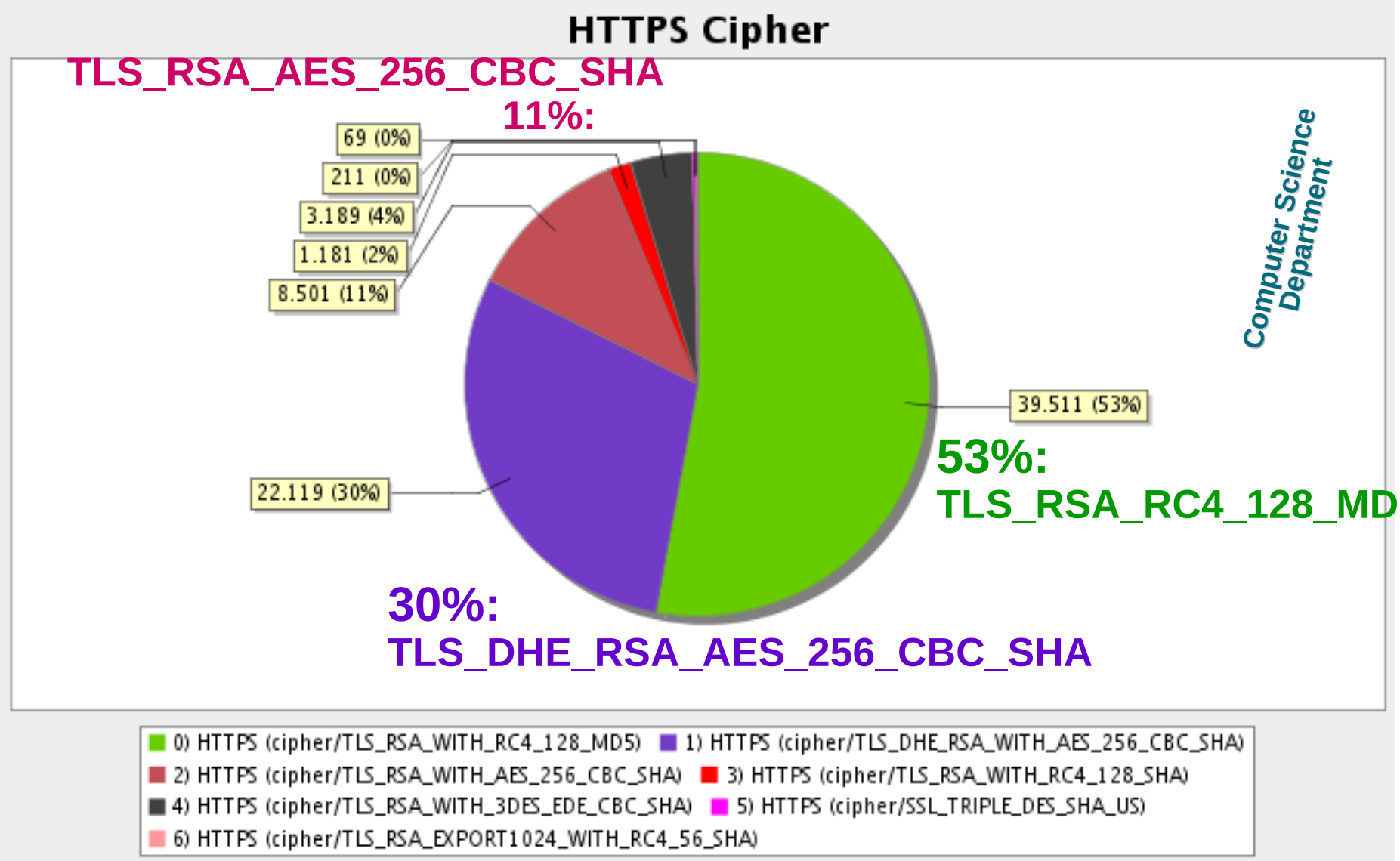


Bei anderen Protokollen ist das Verhältnis noch weit schlechter und liegt meist etwa bei 99:1

Verteilung gängigste Protokolle die SSL/TLS nutzen



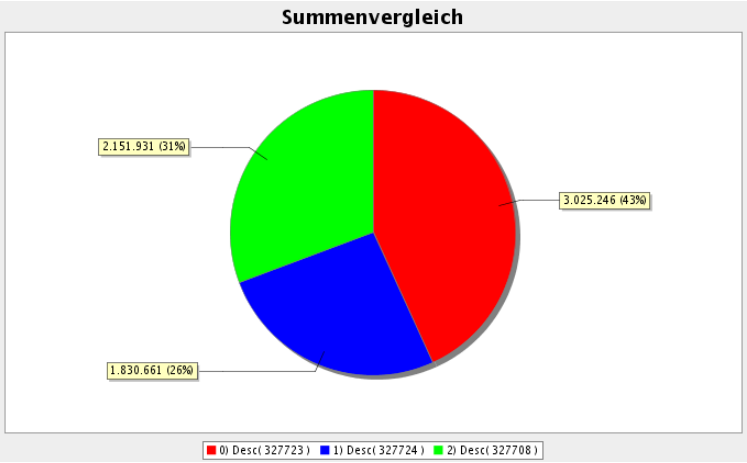
Beispiele von Ergebnissen des IAS
→ Wissensbasis: Technologietrend (TLS)



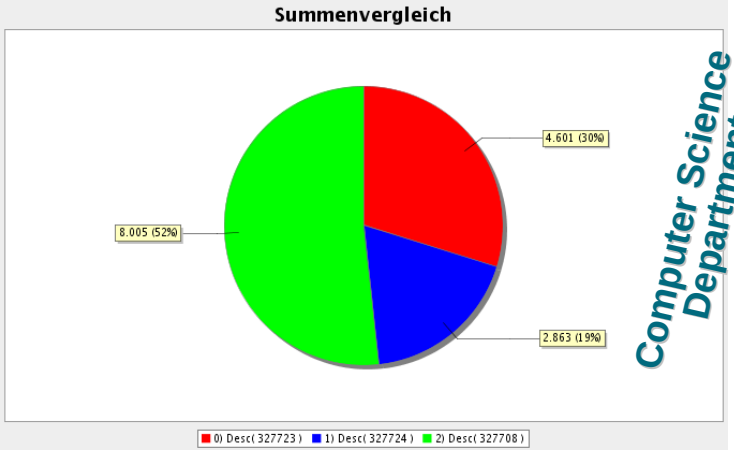
Computer Science Department

Beispiele von Ergebnissen des IAS
→ Wissensbasis: Zusammenhänge

- **SYN-Scan (Potentielle Angriffsituation)**
  - Vergleich unterschiedlicher Zeiträume
    - Normal: SYN > SYN/ACK > (FIN/ACK) / 2
  - Diskrepanz: Normalverteilung zu Verteilung in einer Angriffssituation
    → Angriffssituationserkennung



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

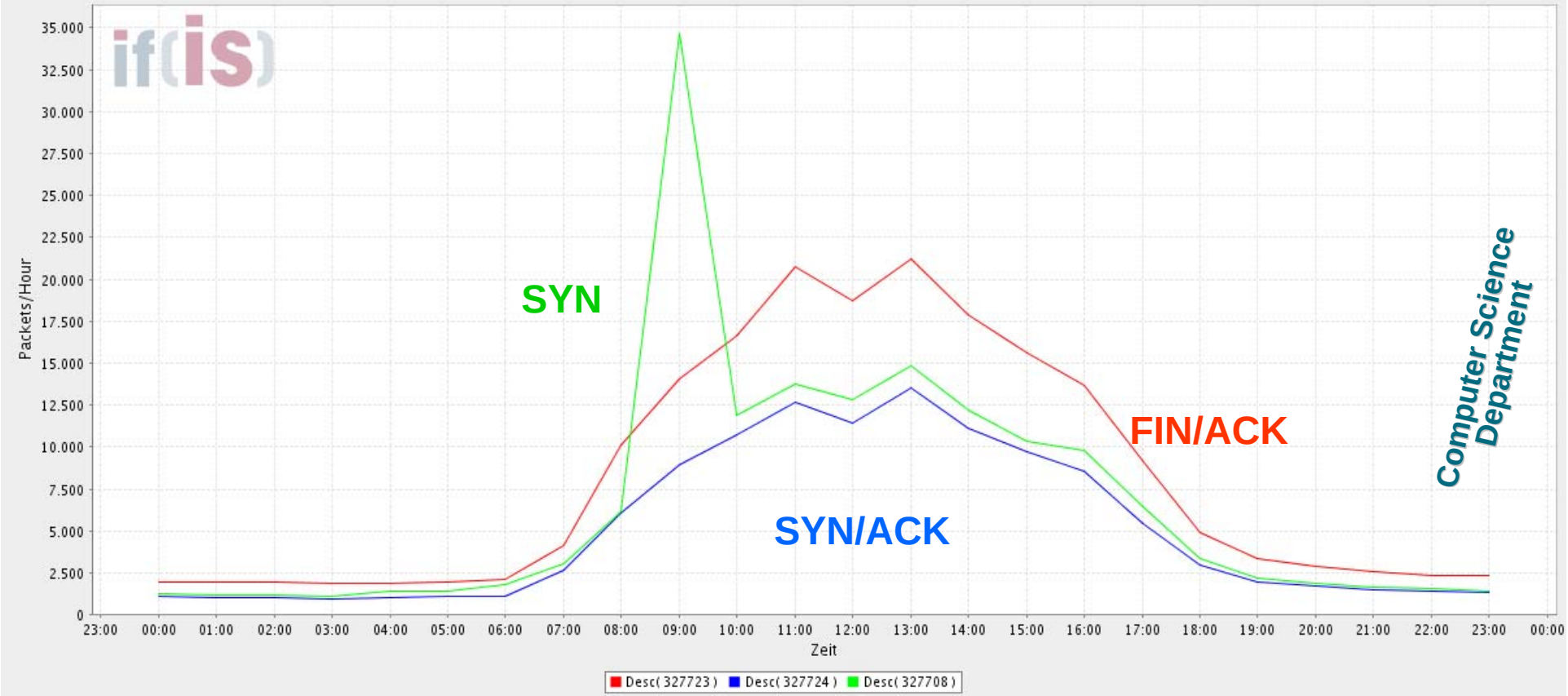


Normale Verteilung

Annormale Verteilung

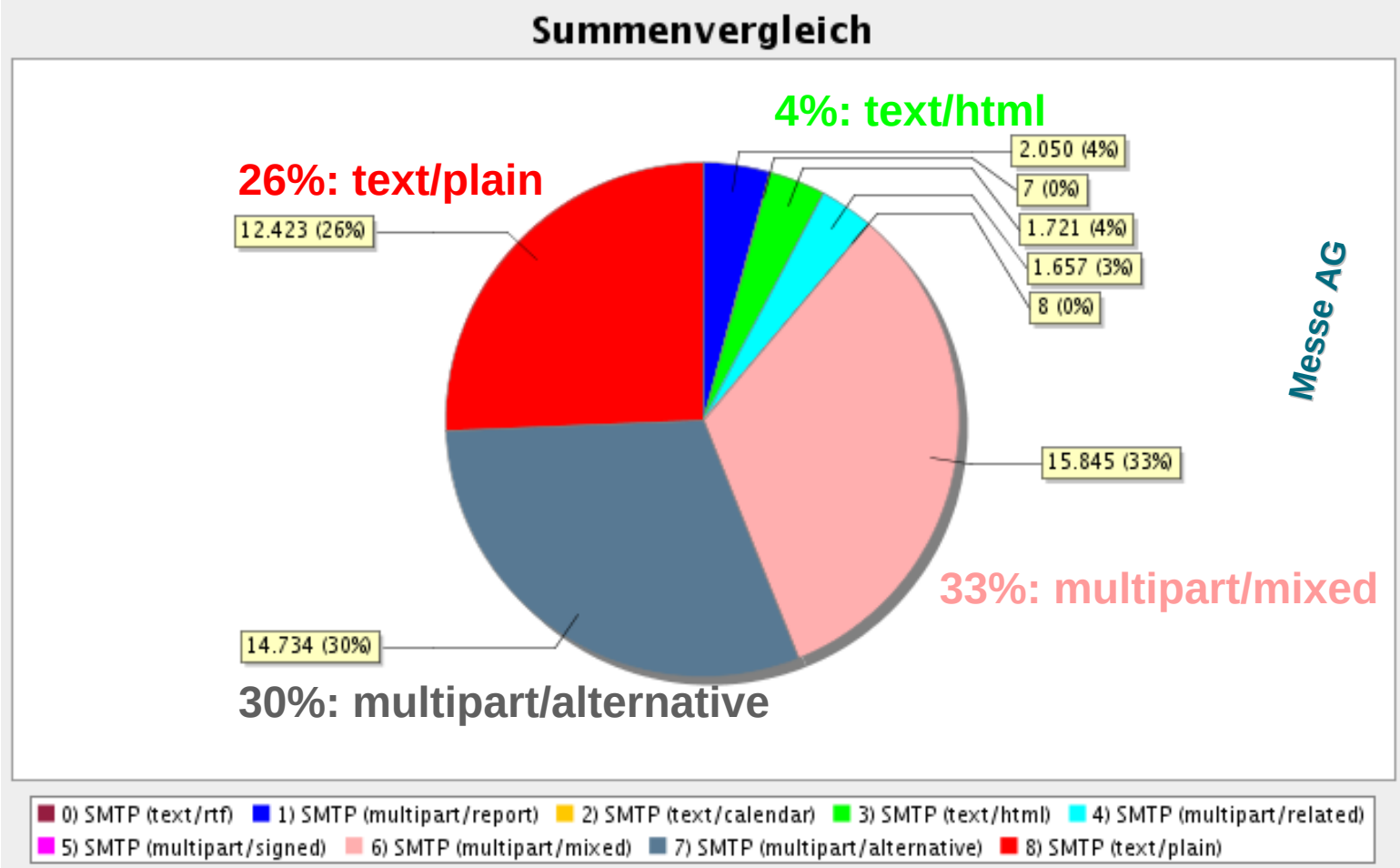
Beispiele von Ergebnissen des IAS
→ Wissensbasis: Profile

- **SYN-Scan (Potentielle Angriffsituation)**
  - Scan-Zeitraum deutlich zu erkennen



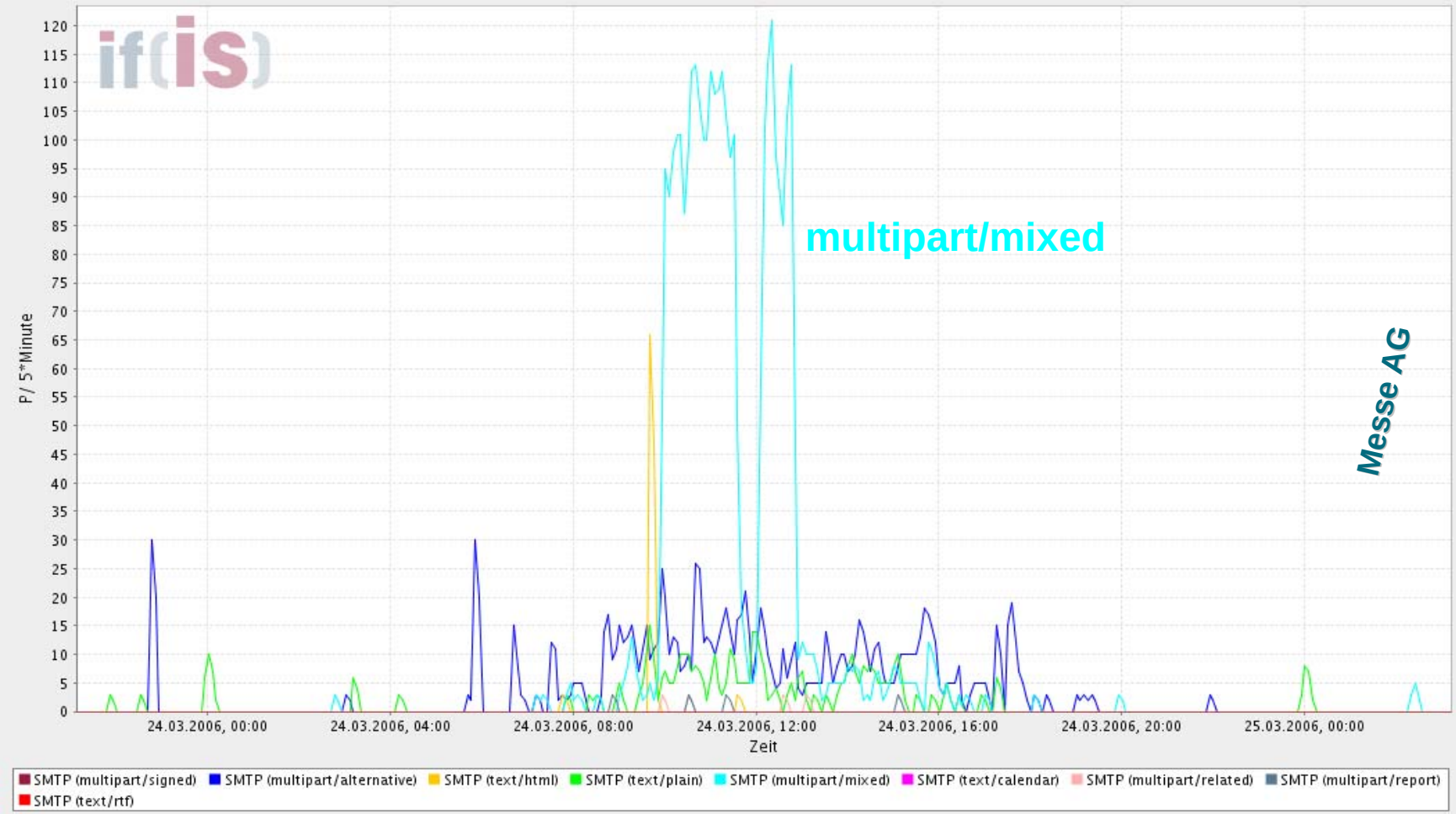
Beispiele von Ergebnissen des IAS
→ Wissensbasis: Erfahrungen

- **SMTP Content Type**
  - 60% “text” Mails
  - 33 % “attachments”

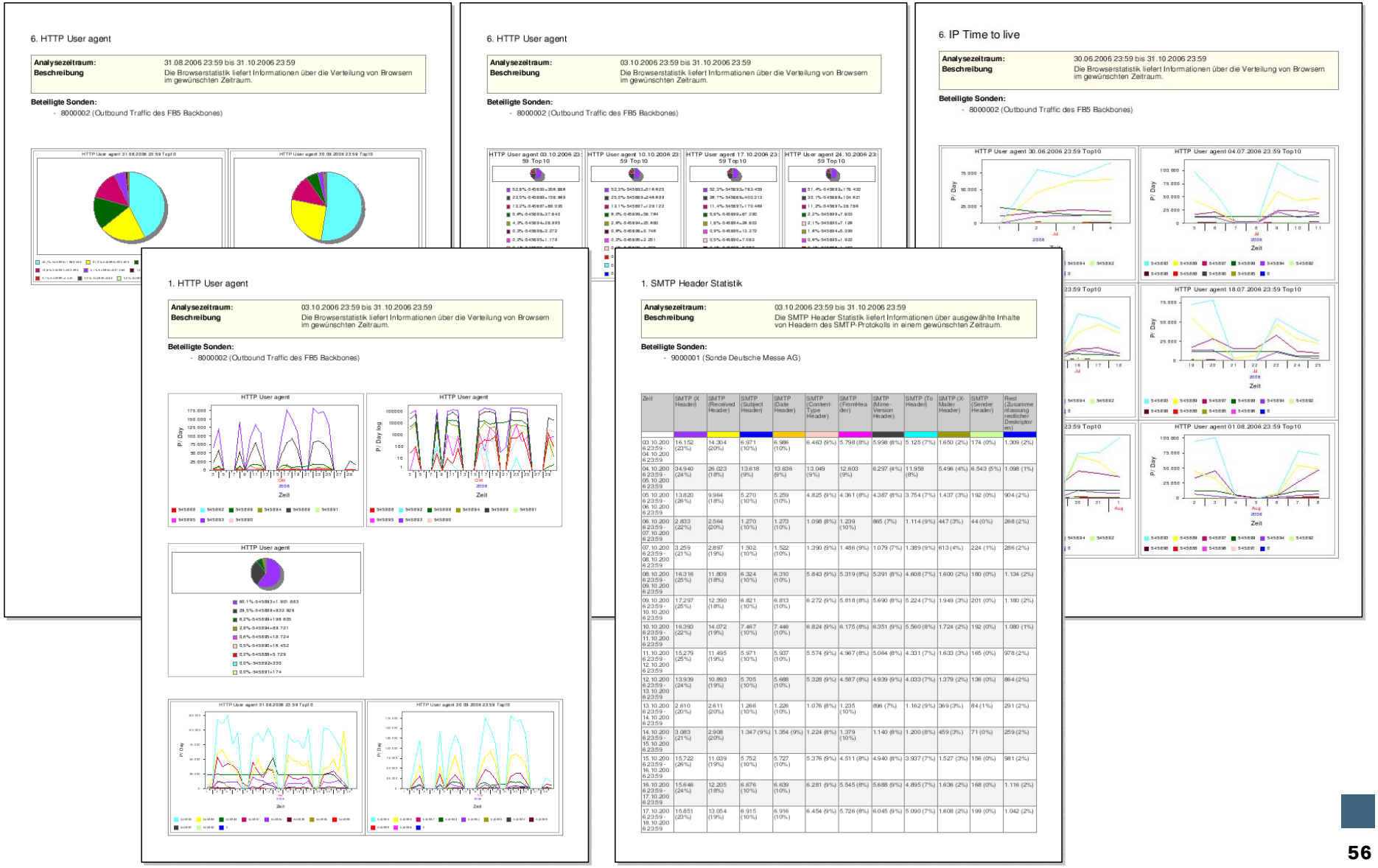


Beispiele von Ergebnissen des IAS
→ Angriffssituationserkennung

- **SMTP Content Type**
  - Zeitweise mehr E-Mail mit content-type multipart/mixed
    -> Mail-Virus?



Beispiele von Ergebnissen des IAS → 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 (From Header)	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%)	9.994 (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%)	6.629 (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-Analyse-System (IAS)
→ Ergebnis Ziel 1

- Wir können zurzeit ca. **600.000** Kommunikationsparameter zählen.
- Wir können mit dem „EagleX Analysis Client“ sehr gut **Analysen „per Hand“** durchführen
- Mit Hilfe der **Reporte** können wir uns regelmäßige Zusammenfassungen senden lassen.
  - Sehr gute Übersicht
  - Sehr gute Information, um das normale Verhalten zu verstehen.
  - Kommunikationsverhalten bleibt in Beobachtung
  - Trends können hier schon erkannt werden
  - Auffälligkeiten, die bei der manuellen Analyse ausser Acht gelassen wurden, treten hier auf.

Internet-Analyse-System (IAS)
→ Weiteres Vorgehen

- **Validierung der Kommunikationsparameter**
  - Welche werden wirklich benötigt?
  - Welche sind redundant?
  - Wie lässt sich durch geschickte Aggregation die Datenmenge weiter reduzieren?
- **Identifizieren neuer Kommunikationsparameter**
  - Welche Protokolle werden in Zukunft interessant sein?
  - Welche Daten werden für eine komplette Beschreibung des Internet benötigt?
- **Arbeiten mit /Analysieren der Wissenbasis**
  - Einsatz von Data Mining um Zusammenhänge zu finden und besser zu verstehen

Internet-Analyse-System (IAS)
→ Definierte Ziele

Ziel 2

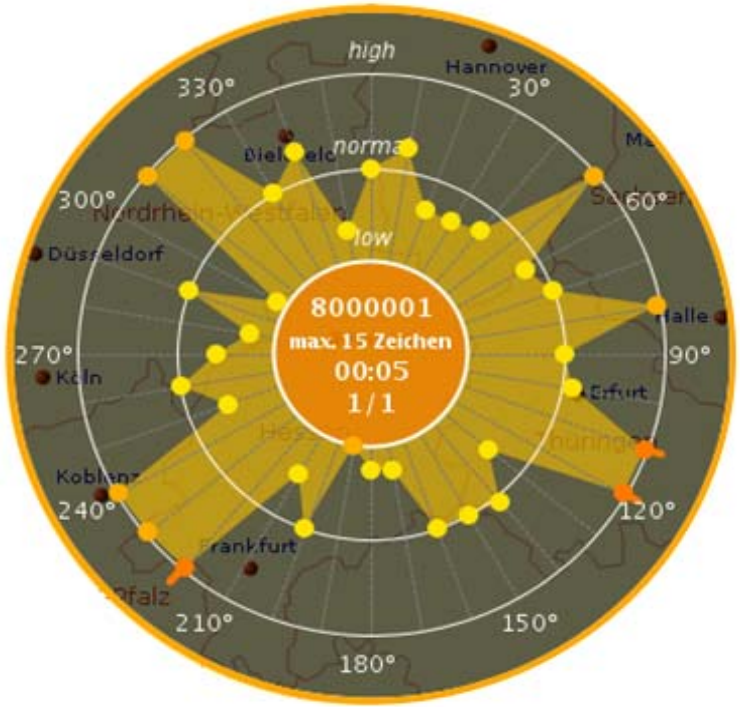
- **Überblick über den aktuellen Zustand des Internets**

Internet-Analyse-System (IAS)
→ Ergebnis Ziel 2

- Wir müssen Modelle haben, die uns helfen, den aktuellen Zustand zu ermitteln.

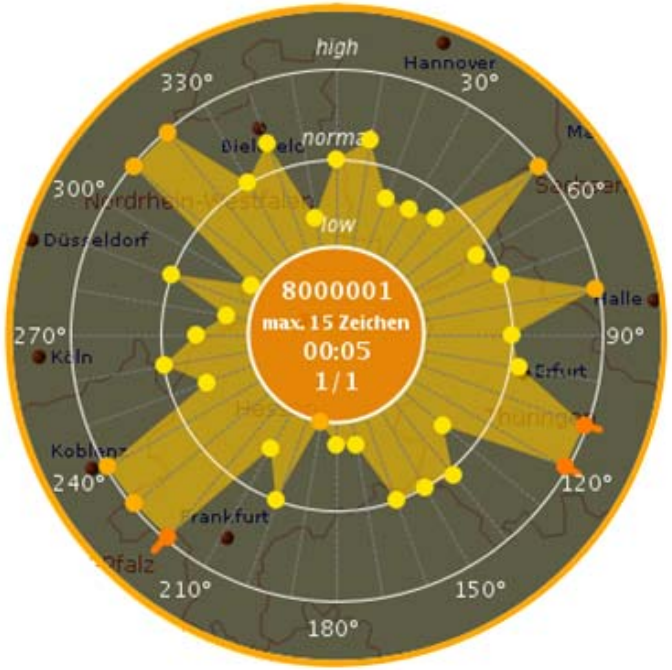
  Herausforderung: Große Datenmengen übersichtlich darstellen.

- Ein Beispiel, mit dem wir erste Erfahrungen sammeln ist **VisiX**.
- Ausgewählte, **wichtige Parameter**
- **Fortlaufende Aktualisierung**
- Ausrichtung an fixen Referenzwerten
- Wählbare, **farbliche Kodierung**

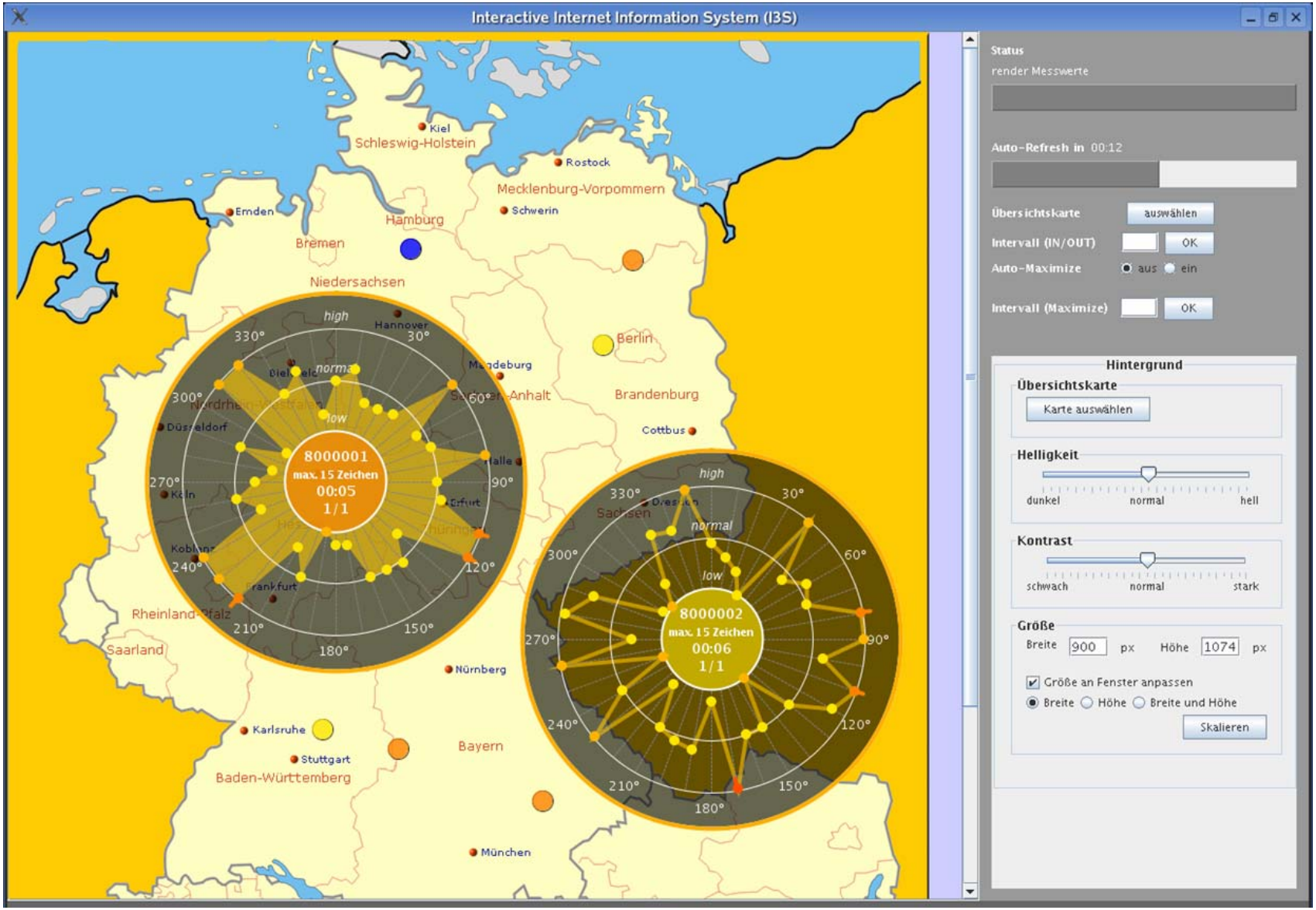


Internet-Analyse-System (IAS)
→ Ergebnis Ziel 2

- Visualisierung der Daten mehrerer Sonden zur selben Zeit
- Dadurch Erkennen von Abhängigkeiten zwischen unterschiedlichen Sonden
- z.B. Sonde X: extrem hoher HTTP-Traffic
  Sonde Y: auch extrem hoch
  ==> externes Ereignis (z.B. Windows Update)
- VisiX ermöglicht das Kennenlernen des Netzverhaltens
- Kontinuierliche Beobachtung bei Alarmen
- Veranlassung weiterer Maßnahmen
- Ablauf: Alarm, VisiX, EagleX, ...



Beispiele von Ergebnissen des IAS
→ Übersicht über den aktuellen Zustand



Internet-Analyse-System (IAS)
→ Definierte Ziele

Ziel 3

- **Erkennen von Angriffssituationen und Anomalien**

Internet-Analyse-System (IAS)

→ Erkennen von Angriffssituationen u. Anomalien

Angriffssituationen

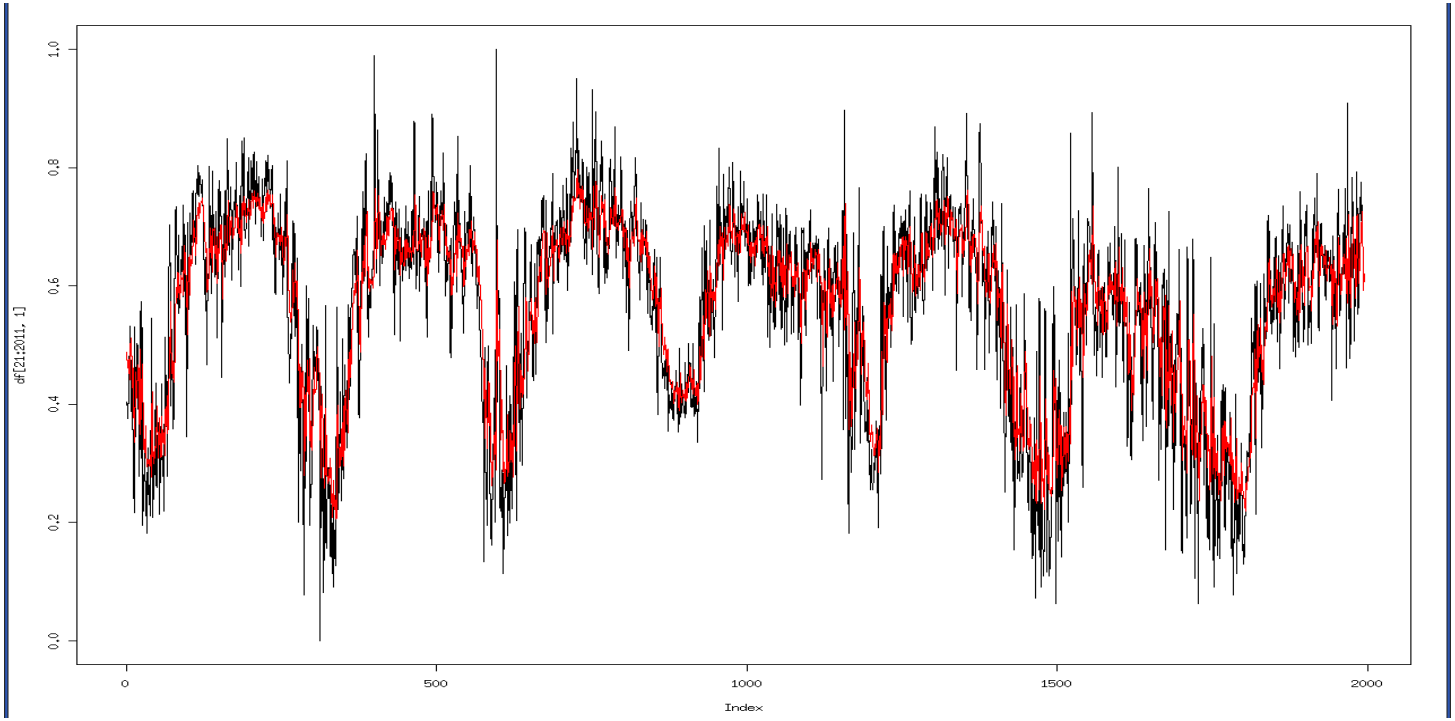
- Definition von **Signaturen**, die wir durch Analysen mit dem „**EagleX Analysis Client**“ ermittelt haben
- Zusammenarbeit mit UNI Mannheim
  - **CW-Sandbox**
    - **Austausch von Signaturen und „Nachlade-URLs“**
    - **Ermitteln von Kommunikationsprofilen**

Anomalien

- Kernbestandteil eines Frühwarnsystems
- Wir entwickeln und nutzen Modelle, mit denen wir eine Abweichung vom erwarteten Zustand feststellen können.

IAS: Current State of Development
→ Result Examples

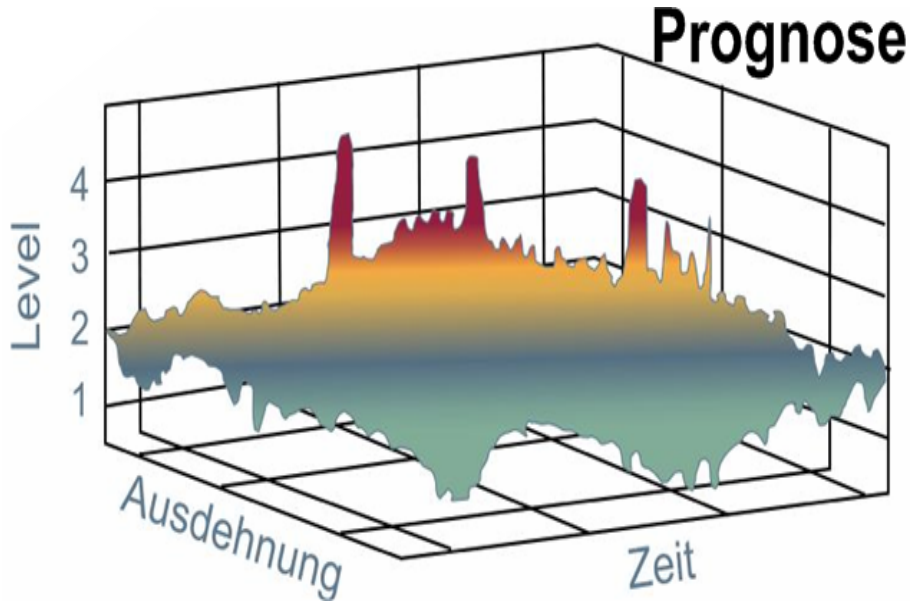
- **Predicition**
  - Neural network predicts next datapoint by analyzing previous datapoints
  - The results are promising, but more research is necessary



Internet-Analyse-System (IAS)
→ Definierte Ziele

Ziel 4

- **Prognosen von Mustern und Angriffen**



Internet-Analyse-System (IAS)
→ Prognosen von Mustern und Angriffen

- Zwei Aspekte sind für die Sicherstellung der Funktionsfähigkeit des Internet von Bedeutung:
  - Das Netz muss auf kommende Technologien vorbereitet sein
  - Angriffe müssen rechtzeitig erkannt und ihre Verbreitung effektiv unterbunden werden

=> 1. Technologietrends müssen rechtzeitig erkannt werden!

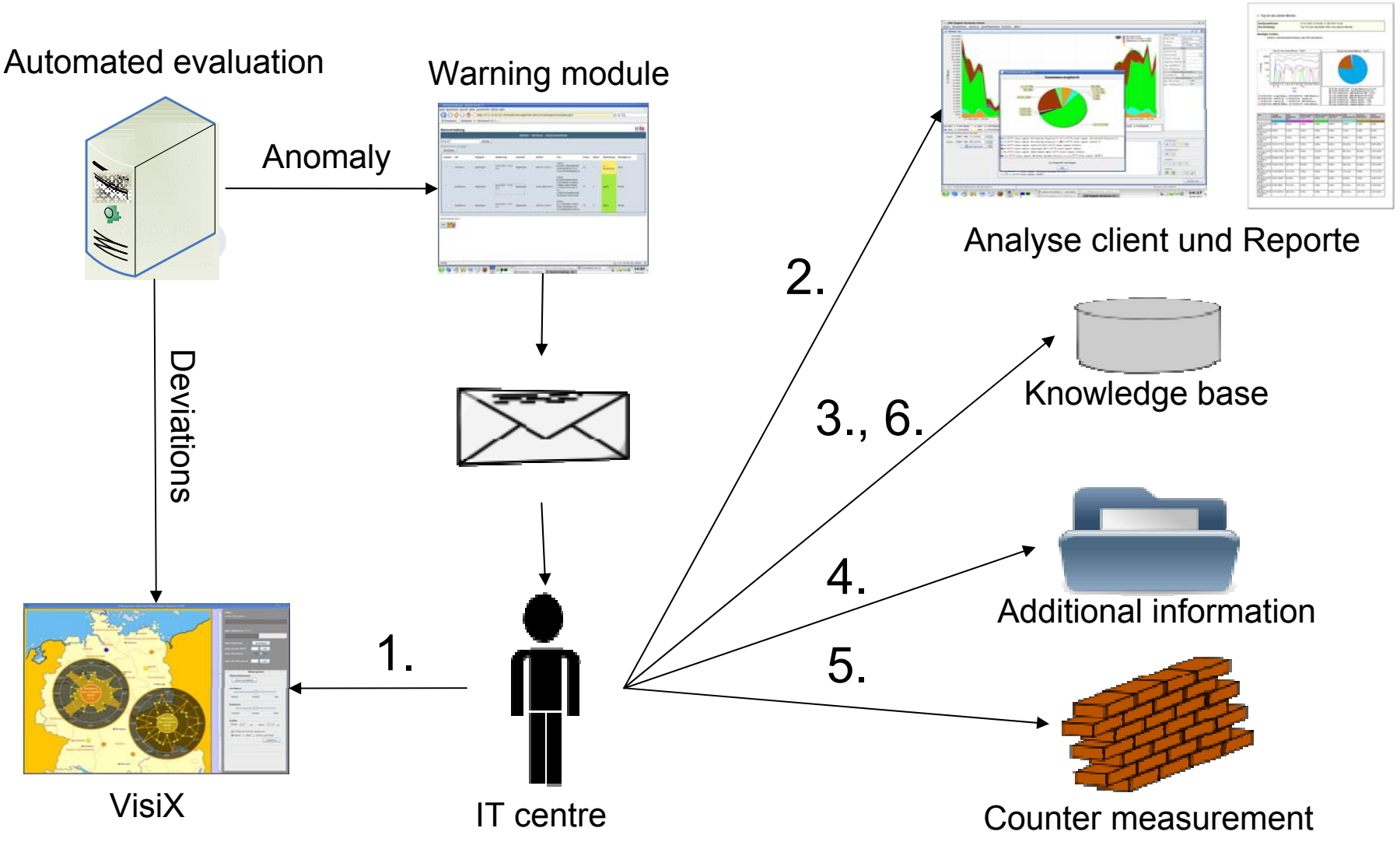
2. Die initiale Phase von Angriffen muss besser verstanden bzw. beschrieben werden!

3. Die Ausbreitung von Angriffen muss verstanden werden!

4. Sicherheitsmechanismen müssen sichere Kryptographieverfahren nutzen

=> Prognosefähigkeit und Identifikation von Mustern

Process: Early Warning



Agenda

- Einführung
- Frühwarnsysteme
- Struktur für Internet-Frühwarnsysteme
- Verschiedene Realisierungsansätze
- Internet-Analyse-System
- **Internet-Verfügbarkeits-System**
- Zusammenfassung

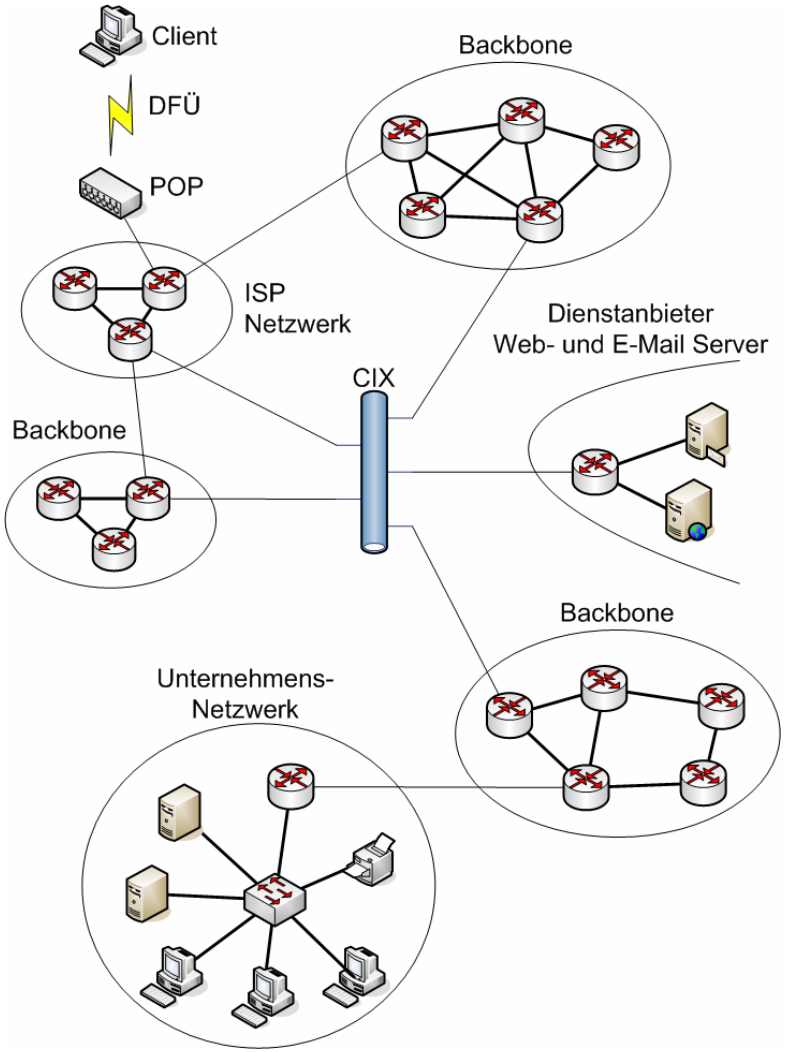
Internet-Verfügbarkeits-System
→ Einleitung (1/2)

- Es gibt immer mehr Benutzer und Möglichkeiten im Internet.
- Aufgrund der größer werdenden Bandbreiten entstehen **neue Dienste**.
- Geschäftsprozesse werden gemäß neuer Möglichkeiten optimiert.
- Wir begeben uns in eine **Abhängigkeit von diesen Technologien**.
- Die Dienste müssen sich auf einer **akzeptablen Qualitätsebene** bewegen:
  - **Quality of Service - QoS**
  - **Quality of Availability – QoA**
- Funktionsstörungen oder bösartige Angriffe stellen eine Gefahr dar.
- Neue Sicherungsmaßnahmen sind erforderlich.

If you can‘t measure it, you can‘t manage it!

Internet-Verfügbarkeits-System
→ Einleitung (2/2)

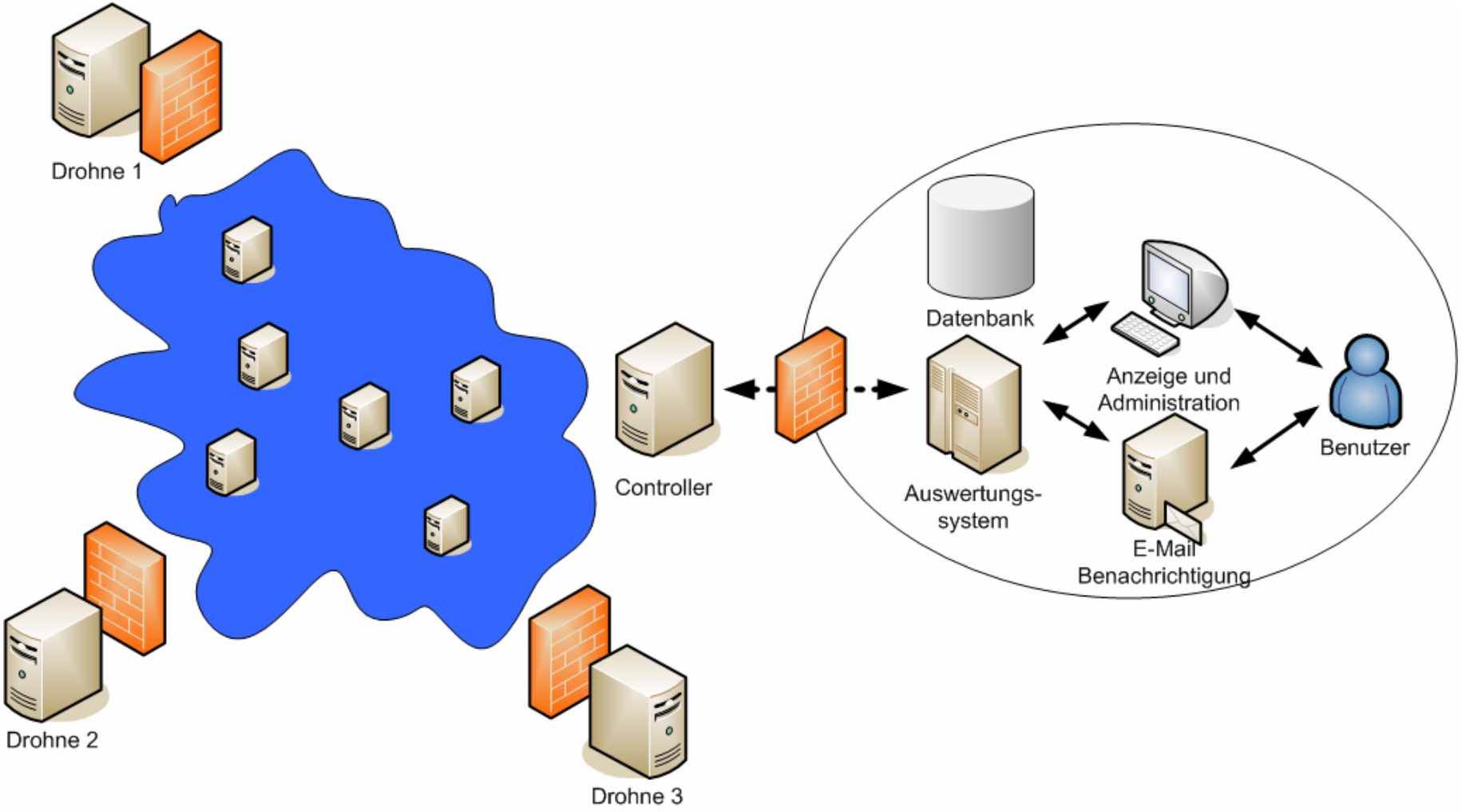
- Administratoren eines Dienstanbieters überwachen nur ihre eigenen Systeme!
- **Ein Benutzer ist aber von mehreren Faktoren abhängig:**
  - Provider
  - Router in den autonomen Systemen
  - DNS Server
  - Dienst-Server
  - Daten des Dienstes



Internet-Verfügbarkeits-System
→ Übersicht (1/2)

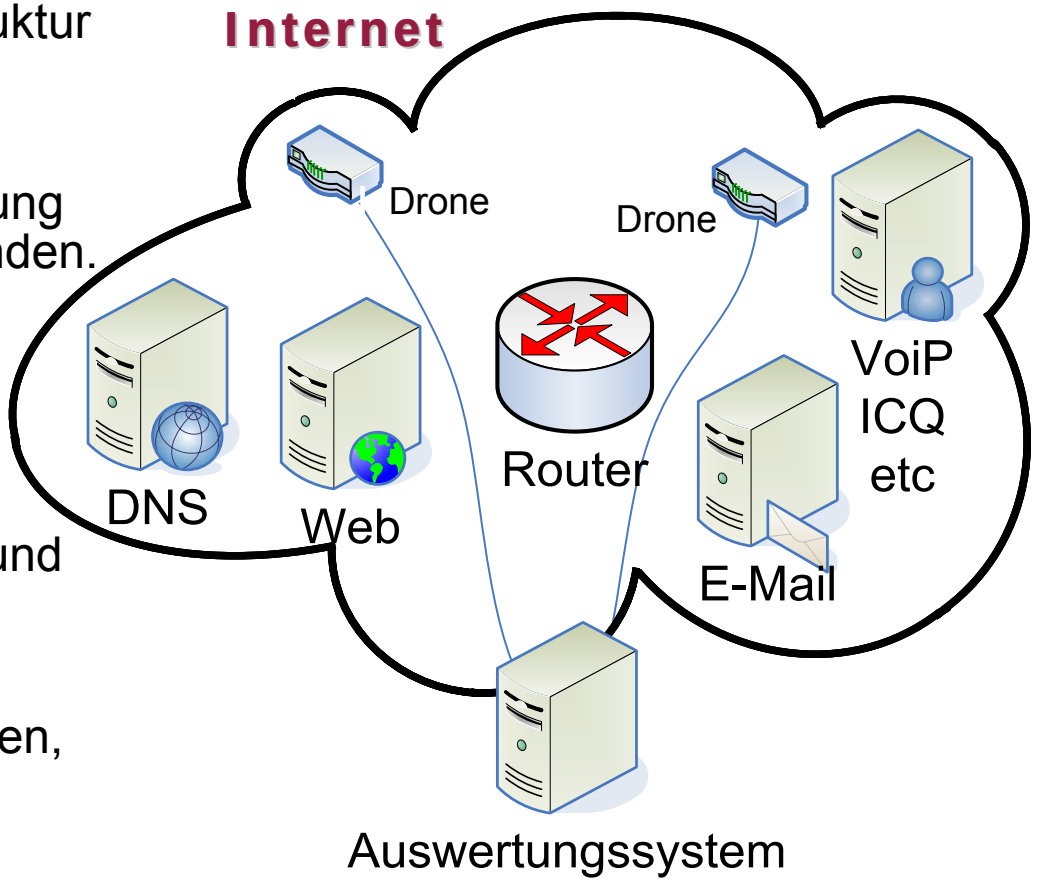
- Das IVS stellt die **Verfügbarkeit** und **Dienstgüte** von Internetdiensten aus Benutzersicht dar.
- Analyse wird auf Anwendungs- und Transportebene vorgenommen.
- Ein überlasteter Router kann Messwerte bezüglich eines Servers stark beeinflussen.
- Messsysteme, nachfolgend ***Drohnen***, werden an mehreren Stellen im Internet platziert und führen die Messungen durch.
- Ergebnisse können untereinander verglichen werden.
- Drohnen befinden sich in geschützten Netzwerken und sind nicht direkt zur Administration erreichbar.
- Als zentrale Vermittlungsstelle dient der ***Controller***.
- Auswertungssystem mit Datenbank und E-Mail-Benachrichtigungssystem.
- Grafische Anzeige und Administrationsoberfläche

Internet-Verfügbarkeits-System
→ Übersicht (2/2)



Internet-Verfügbarkeits-System
→ Idee

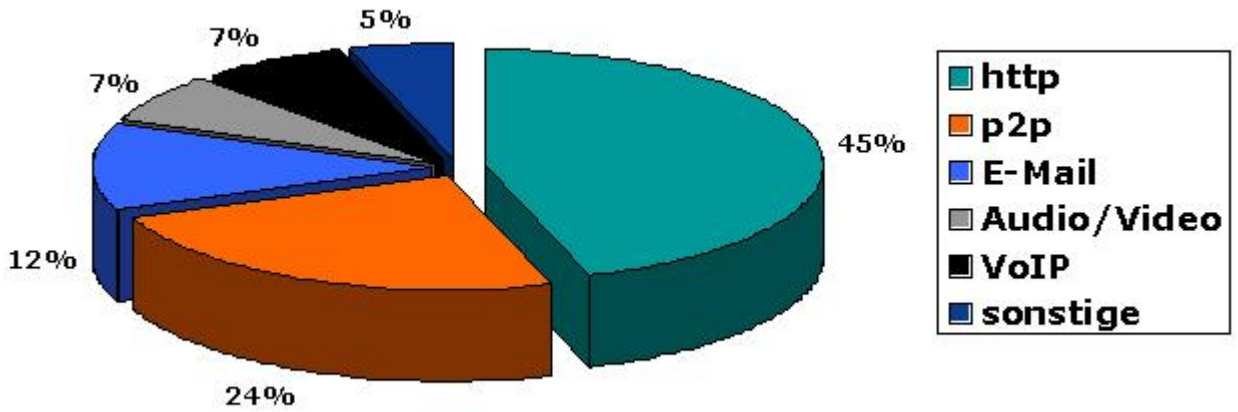
- Beobachtung der kritischen Infrastruktur **„Internet“.**
- **Dronen** werden an ausgesuchten Positionen des Internets zur Erfassung von **Verfügbarkeitsdaten** eingebunden.
- Es werden verschiedene Arten von Verfügbarkeiten gemessen
  - Wichtige Webdienste
  - DNS-Dienst
  - Kommunikationsverbindungen und Router
  - Mail-Dienste und –Server
- **Parameter:** Dienstgüte: Funktionen, Fehlerrate, Jitter, Verzögerung, Paketverlust
- Ein zentrales **Auswertungssystem** analysiert die **Verfügbarkeitsdaten** und Auswertungsergebnisse und stellt diese umfangreichen Ergebnisse intuitiv dar.



Drone: aktive Sonde
Platzierung unabhängig von Dritten!

Dienste, die überwacht werden können

- Web-Server HTTP-Protokoll
- Datei-Server FTP-Protokoll
- E-Mail-Server SMTP-, POP3-, IMAP-Protokoll
- SIP-Server SIP-Protokoll
- DNS-Server DNS-Protokoll (Dienst der Infrastruktur)



Quelle: http://www.heise.de/ct/05/07/088

Messmodule
→ Anwendungsebene (1/4)

- Für jedes Dienstprotokoll ein Messmodul
- Aktive Funktionstests durch Dienstanfragen
- Ist der Dienst nutzbar und stehen die Daten zur Verfügung?

- **HTTP-Modul**
  - Verbindungsaufbau
  - Aufruf der GET-Methode
  - Übertragung der Inhalte

Messmodule
→ Anwendungsebene (2 / 4)

- **FTP-Modul**
  - Verbindungsaufbau mit dem Steuerkanal
  - Benutzeranmeldung wenn möglich
  - Anforderung des aktuellen Verzeichnisses
  - Antwort erfolgt über den Datenkanal
- **SMTP**
  - Verbindungsaufbau
  - Benutzeranmeldung wenn möglich
  - Versand einer E-Mail

Messmodule
→ Anwendungsebene (3 / 4)

- **POP3-Modul**
  - Benutzeranmeldung wenn möglich
  - Abruf der E-Mails
  - Löschen der zuvor versendeten E-Mails
- **IMAP-Modul**
  - Gleicher Ablauf wie bei dem POP3 Modul

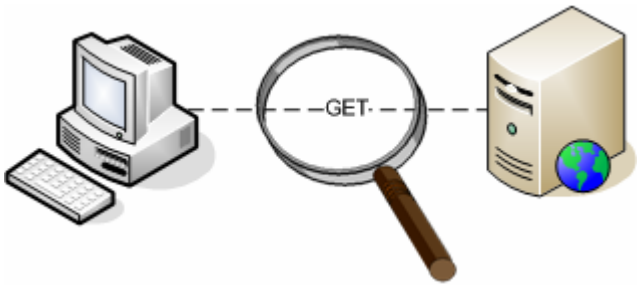
Messmodule
→ Anwendungsebene (4/4)

- **SIP**
  - OPTION-Nachricht an einen SIP Server
  - Antwort innerhalb einer Zeit
- **DNS**
  - Domäne wird zur IP-Auflösung an den Server geschickt
  - Antwort innerhalb einer Zeit

Messmodule
→ Vermittlungsebene

- **ICMP-Modul**
  - Echo-Request Nachrichten
  - **Traceroute**
    - 3 Pakete mit Time-To-Live 1
    - 3 Pakete mit Time-To-Live 2
    - ...
  - Router antworten mit einer **„Time-Exceeded“** Nachricht
  - Server antwortet mit einer **„Echo-Reply“** Nachricht

Analyse der übertragenen Daten



- **Packet-Capture (basiert auf libpcap)**
- **TCP-Analyse**
  - **Laufzeit:**
    - Die Dauer der Datenübertragung
    - Zeit zwischen dem ersten und letzten Paket
  - **Bandbreite:** Übertragene Datenvolumen / Laufzeit
  - **Paketverlustrate:**
    - Doppelte Pakete als Indiz für ein verloren gegangenes Paket
    - Aber: Mehr als zwei gleiche ACK-Pakete werden als ein Verlust gewertet
  - Anzahl der Pakete mit einer **Window-Size von 0**

Analyse der übertragenen Daten

- **UDP-Analyse**
  - **Serverreaktionszeit:** Zeit zwischen der Anfrage und dem Eintreffen eines Paketes

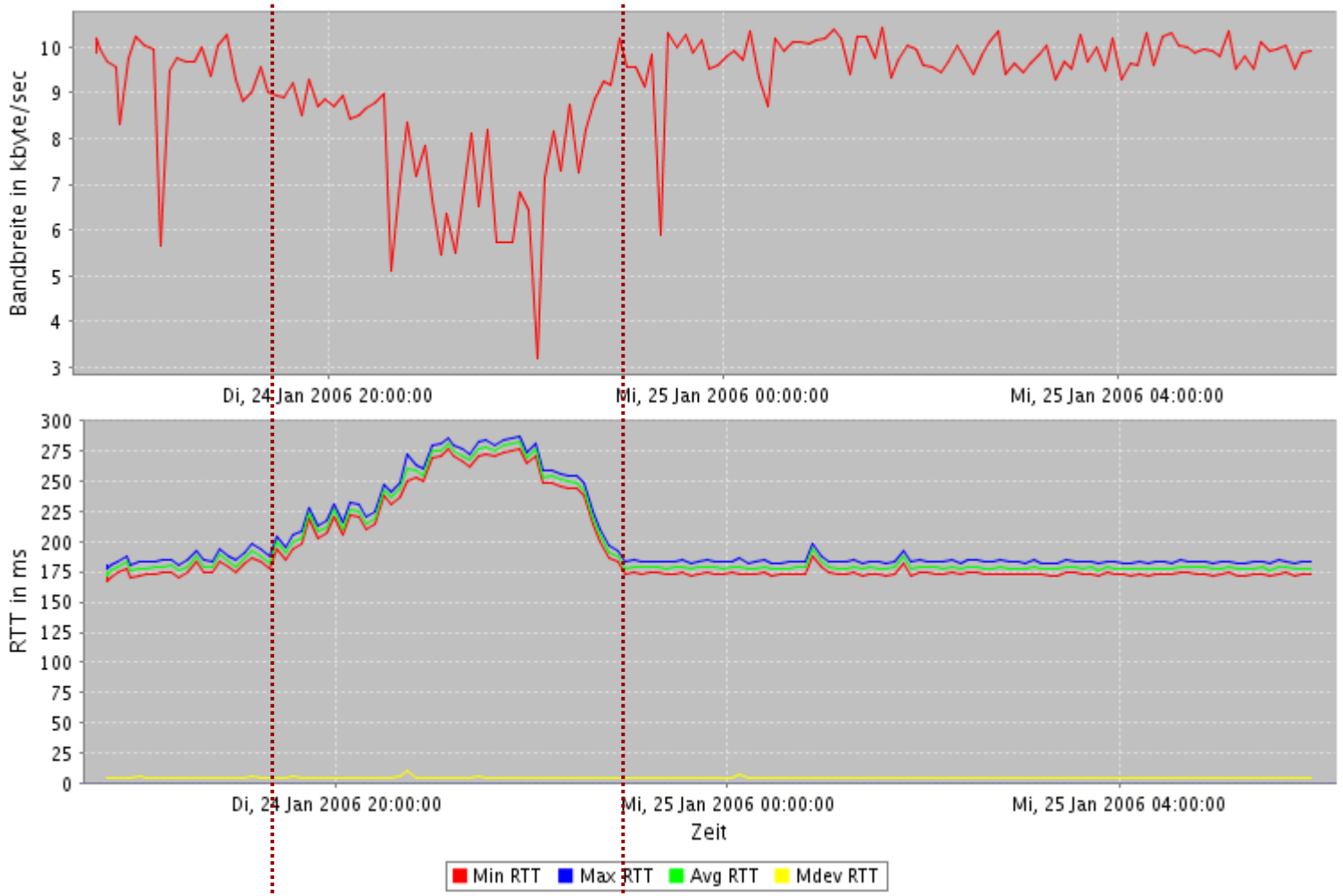
- **ICMP-Analyse**
  - Round-Trip-Times (RTT) von den Routern und des Servers (min/max/avg/mdev)
  - IP-Adressen der Router
  - Anzahl der Router

Arbeitsweise

- Messmodule werden zu Messprofilen zusammengeschlossen

Web-Server	HTTP-, ICMP-, Analysemodul
FTP-Server	FTP-, ICMP-, Analysemodul
DNS-Server	DNS-, ICMP-, Analysemodul
SIP-Server	SIP-, ICMP-, Analysemodul
E-Mail-Server	SMTP-, POP3-, IMAP-, ICMP-, Analysemodul

Internet-Verfügbarkeits-System
→ Beispiele von Ergebnissen (1 / 2)



rapidshare.de

File Sharing Portal

Der größte Datenverkehr und die geringste Bandbreite ist zwischen 18:00 und 23:00 Uhr!

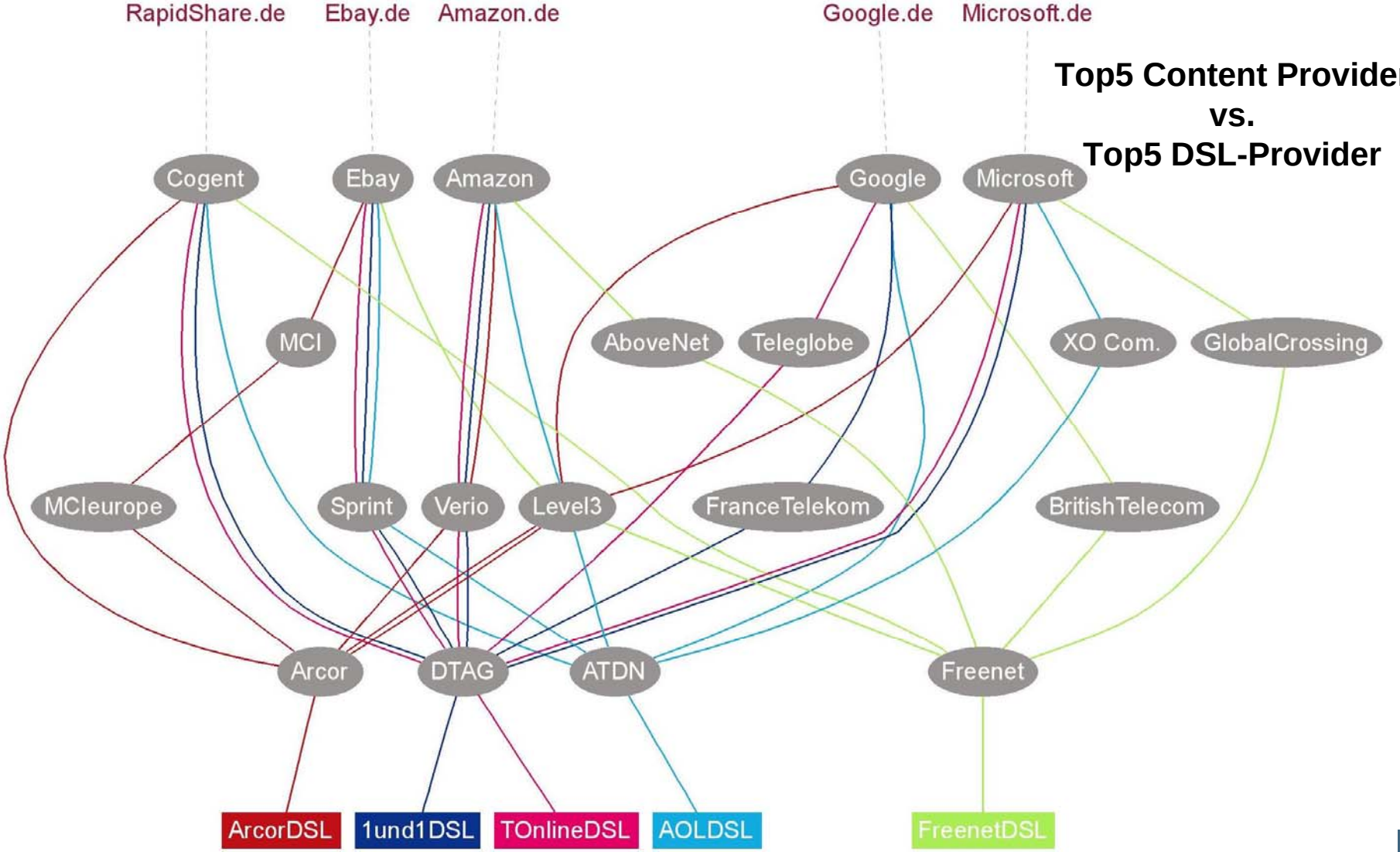
Internet-Verfügbarkeits-System
→ Beispiele von Ergebnissen (2 / 2)

t-online.net

Informations-Portal

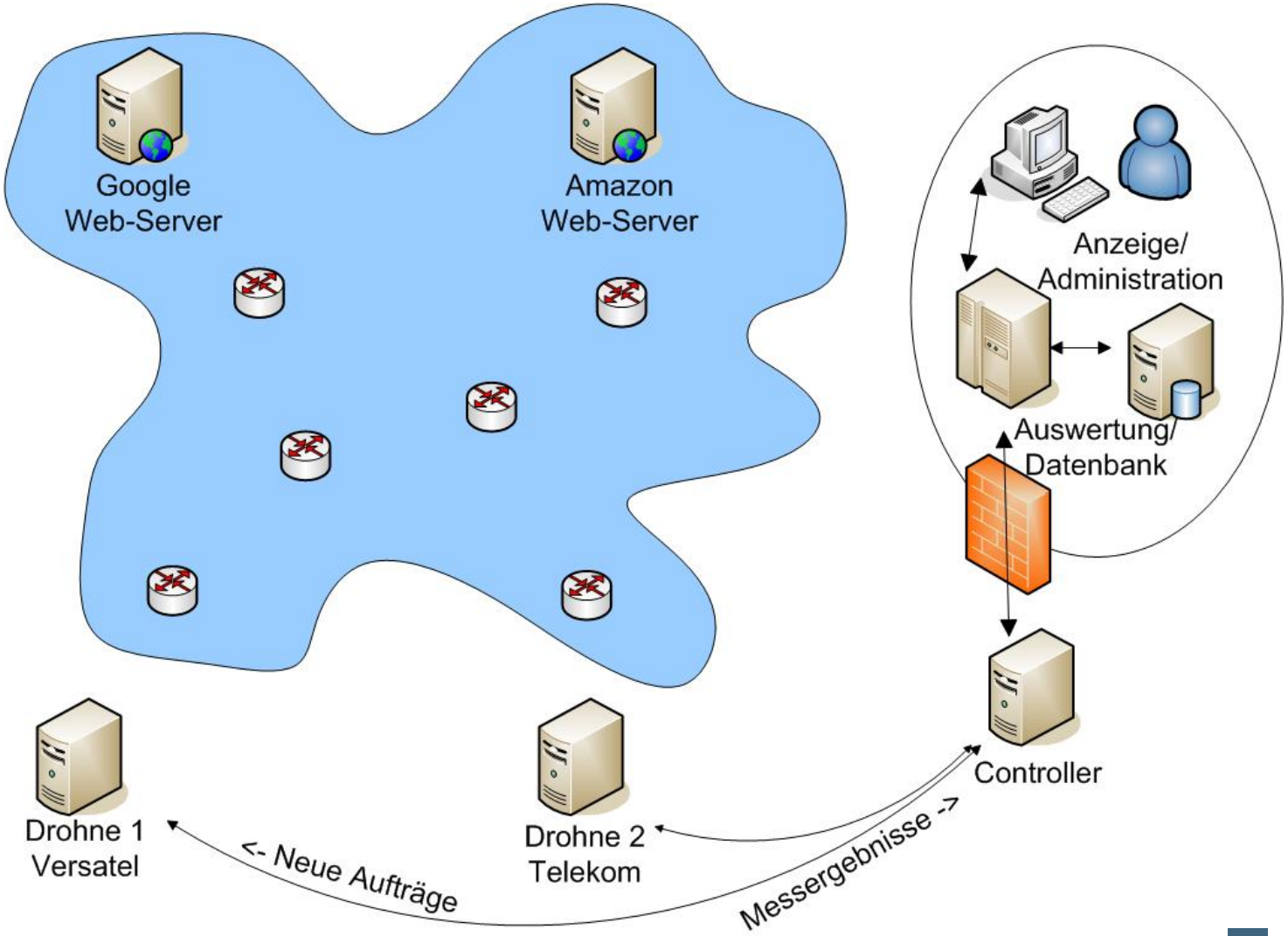
Verschiedene Routings haben einen Einfluss auf die Bandbreite

Leistungsfähigkeit des Internets
→ Idee



Leistungsfähigkeit des Internets
→ Methode

- Messung von:
  - Web-Server

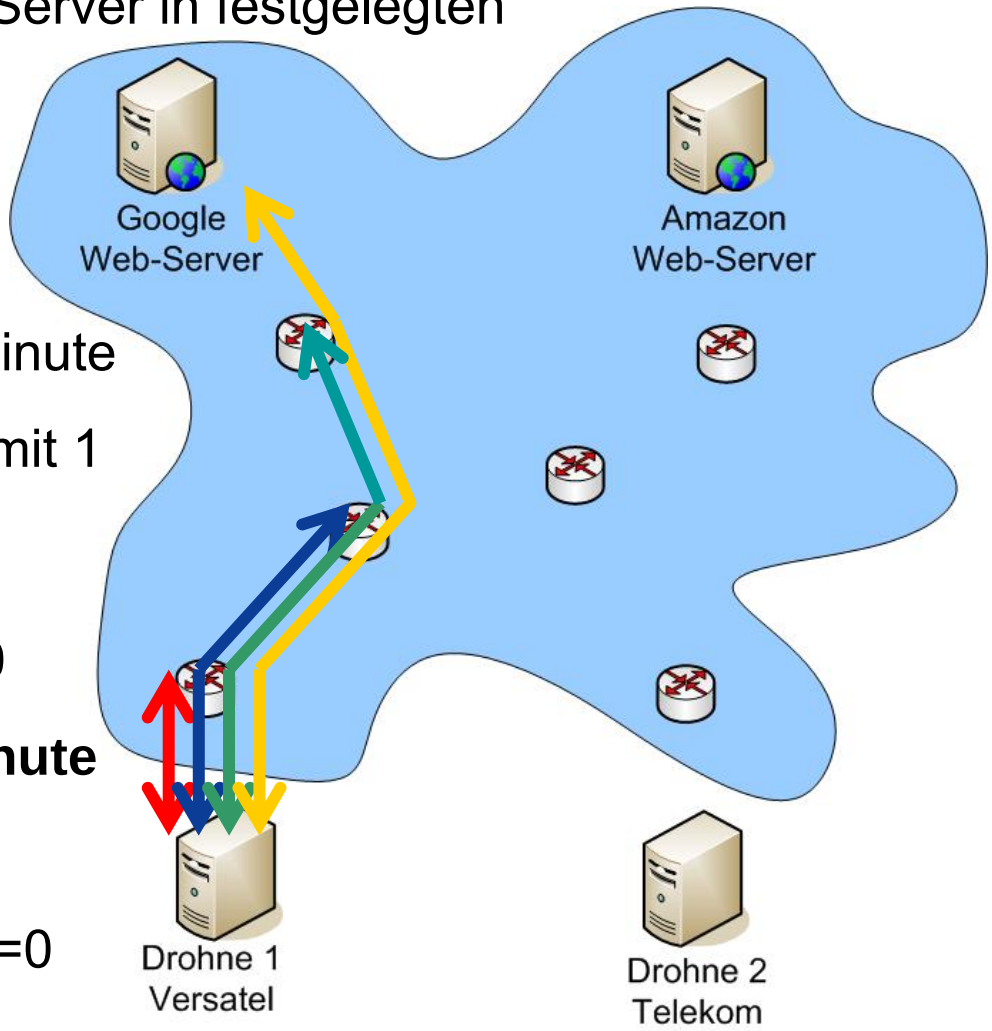


Leistungsfähigkeit des Internets
→ Messwerte

- **Messbare Eigenschaften (Quality of Service):**
  - **Paketverlust-Rate:** Wie viele Pakete gehen während der Übertragung verloren
  - **Laufzeit:** Wie lange dauert der Download
  - **Round Trip Time (RTT):** Zeit für ein Datenpaket von der Quelle zum Ziel und zurück – mehrere Messungen pro Hop
    - rtt – min
    - rtt – max
    - rtt – avg
    - rtt – mdev
  - **Hops:** Anzahl der Zwischenstationen
  - **Window Size = 0:** Wie ausgelastet ist die Gegenstelle
  - **Bandbreite:** Wie viele Daten können pro Zeit übertragen werden

Leistungsfähigkeit des Internets
→ Durchführung der Messungen

- Eine Drohne überprüft verschiedene Server in festgelegten Intervallen
- Beispiel Web-Server:
  - **Traceroute:**
    - 3 ICMP-Pakete pro Hop pro Minute
    - Beim ICMP-Paket beginnend mit 1 wird das TTL-Feld erhöht
    - Stationen senden ICMP-Info (Time-Exceeded) wenn TTL=0
  - **Download der Startseite pro Minute**
    - Paketverlustrate, Laufzeit,
    - Bandbreite (BW), Windowsize=0



Beispiel eins „Pings“

Ping

```
vmsuse80:/ # ping www.heise.de
PING www.heise.de (193.99.144.71) from 172.16.48.111 : 56(84) bytes of
data.
64 bytes from www.heise.de (193.99.144.71): icmp_seq=1 ttl=244 time=18.2 ms
64 bytes from www.heise.de (193.99.144.71): icmp_seq=2 ttl=244 time=16.8 ms
64 bytes from www.heise.de (193.99.144.71): icmp_seq=3 ttl=244 time=15.9 ms
64 bytes from www.heise.de (193.99.144.71): icmp_seq=4 ttl=244 time=16.8 ms
64 bytes from www.heise.de (193.99.144.71): icmp_seq=5 ttl=244 time=17.0 ms

--- www.heise.de ping statistics ---
6 packets transmitted, 5 received, 16% loss, time 5055ms
rtt min/avg/max/mdev = 15.910/16.955/18.201/0.746 ms
vmsuse80:/ #
```

Traceroute
→ Beispiel

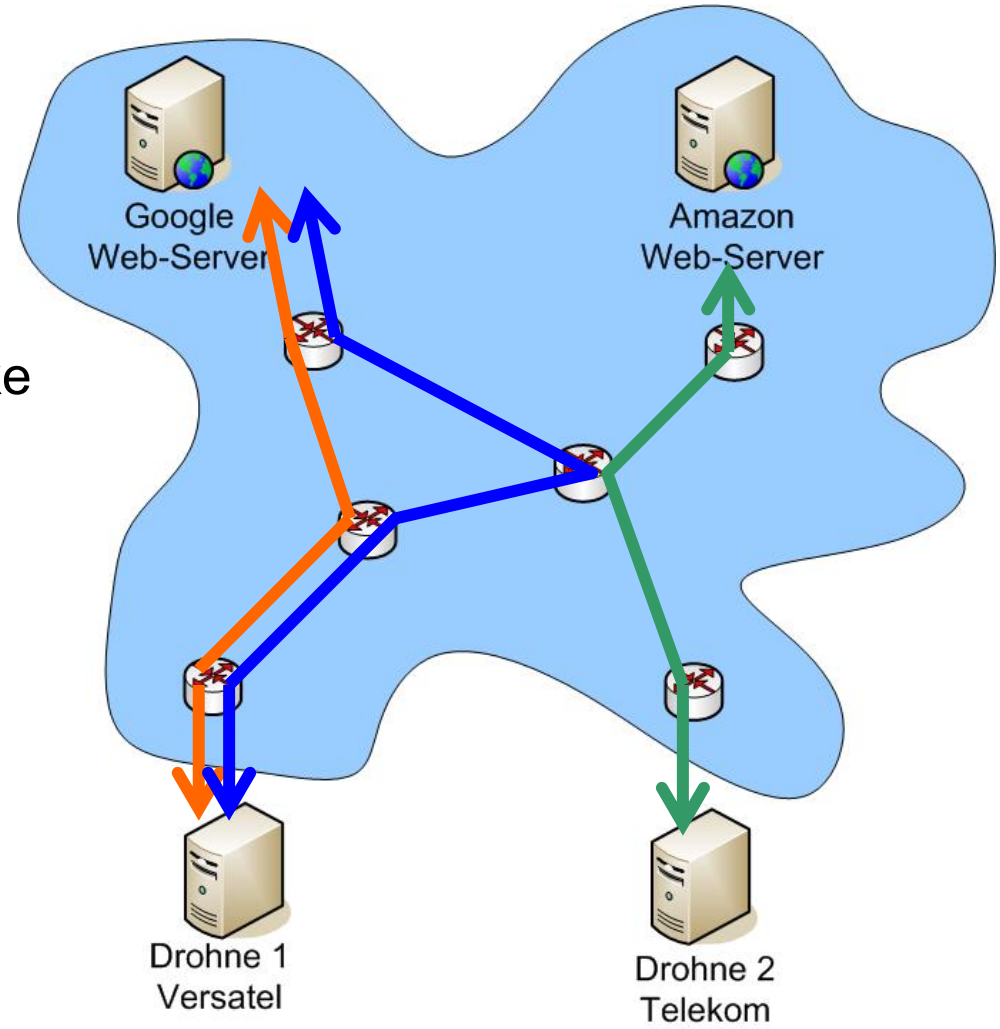
Traceroute

```
vmsuse80:/ # traceroute www.heise.de
traceroute to www.heise.de (193.99.144.71), 30 hops max, 40 byte packets
 1  gw502_48.informatik.fh-ge.de (172.16.48.2)  1 ms  1 ms  1 ms
 2  fb5gwint.informatik.fh-ge.de (172.16.0.5)  1 ms  1 ms  1 ms
 3  172.16.16.3 (172.16.16.3)  2 ms  2 ms  2 ms
 4  fb5gw.informatik.fh-gelsenkirchen.de (194.94.127.2)  3 ms  3 ms  3 ms
 5  193.175.172.2 (193.175.172.2)  3 ms  3 ms  3 ms
 6  ar-essen2.g-win.dfn.de (188.1.44.33)  6 ms  5 ms  5 ms
 7  cr-essen1-ge0-0.g-win.dfn.de (188.1.86.1)  5 ms  5 ms  5 ms
 8  cr-frankfurt1-po8-1.g-win.dfn.de (188.1.18.89)  16 ms  16 ms  16 ms
 9  ir-frankfurt2-po3-0.g-win.dfn.de (188.1.80.38)  15 ms  15 ms  15 ms
10  de-cix2.ffm.plusline.net (80.81.193.132)  16 ms  16 ms  15 ms
11  c22.f.de.plusline.net (213.83.57.53)  16 ms  16 ms  16 ms
12  www.heise.de (193.99.144.71)  16 ms  16 ms  17 ms
vmsuse80:/ #
```

- Zeitanalyse:
  - Verweildauer in einem Router: ca. 0.1 bis 1.5 ms ? **(12 * 1ms = 12ms)**
  - Übertragungszeit für ein Paket bei einer Übertragungsrate von 2 Mbits:
    1.500 Byte: ca. 6 ms; 100 Byte: ca. 0.4 ms **(ca. 5ms)**

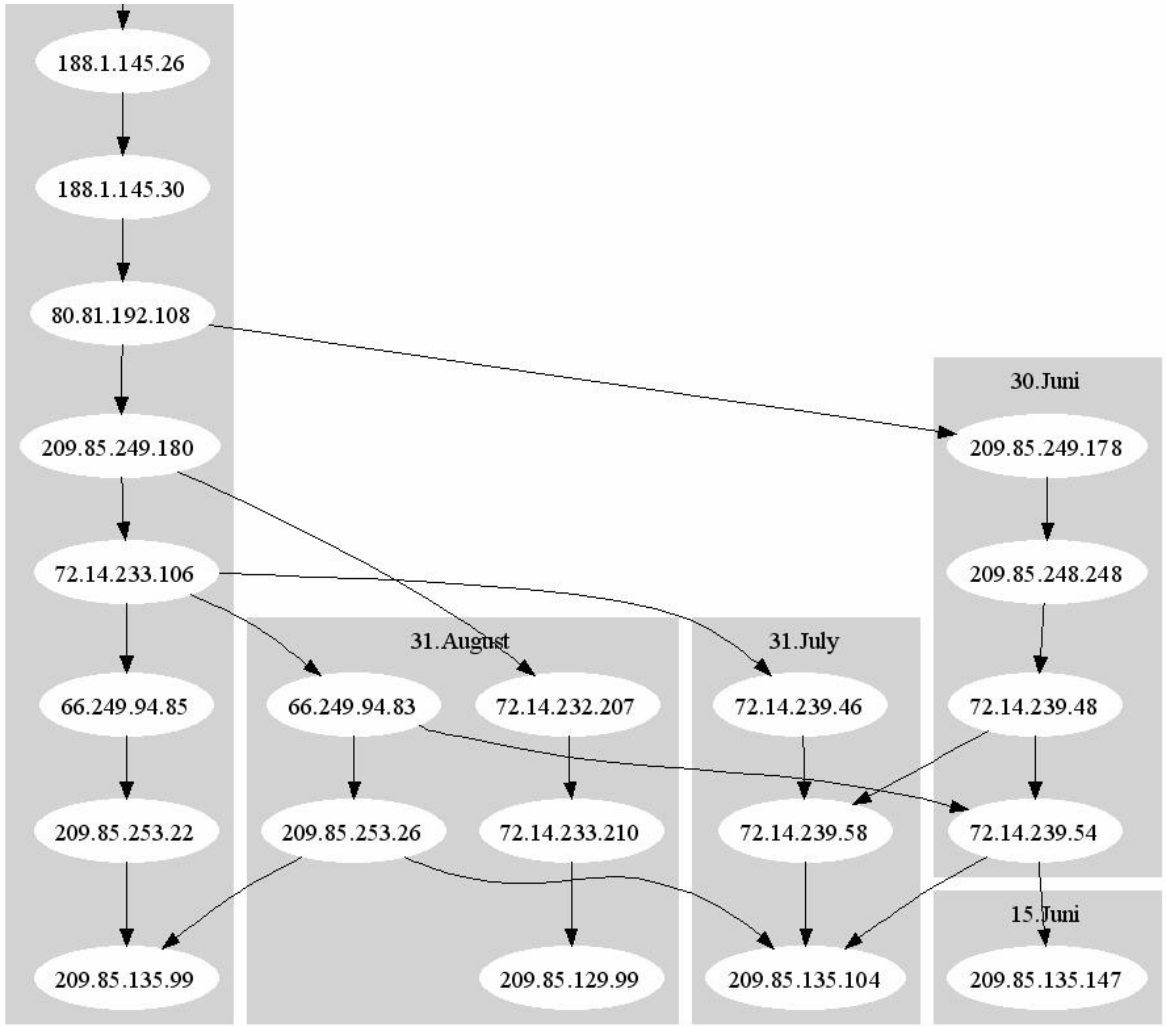
Leistungsfähigkeit des Internets
→ Routing

- Routing-Wege verändern sich mit der Zeit
- Wahl der Route zu Servern wird durch das Network-Management der Netzbetreiber beeinflusst
  - Ausfall eines Routers
  - Starke Belastung einer Teilstrecke



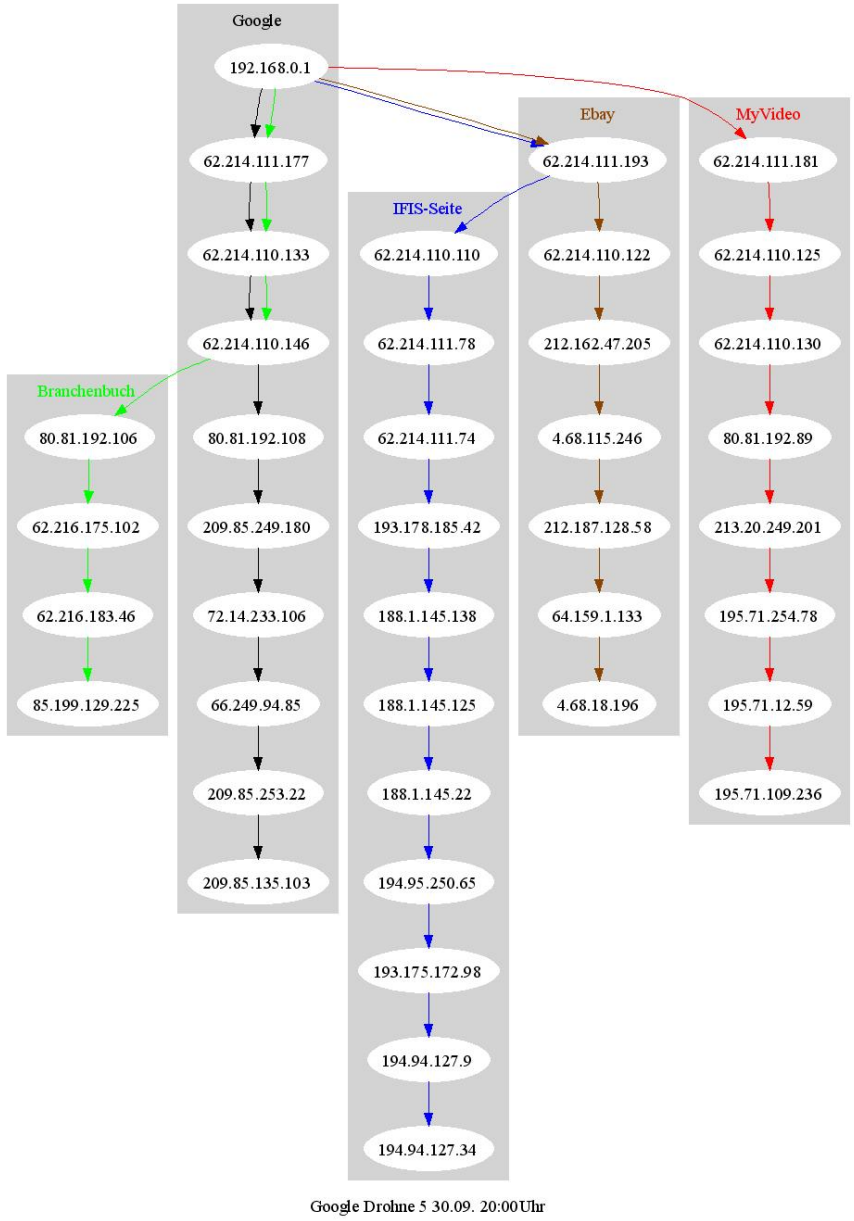
Leistungsfähigkeit des Internets
→ Routing Messung (Zeit)

- **Beispiel:** Route zu Google vom DFN-Netz (Drohne1) aus über die Zeit
- Änderungen nur im letzten Teilstück der Strecke bedingt durch Verteilung der Anfragen auf Mirrors
- positive Auswirkung auf Bandbreite
- In KW 36:
  - 31 verschiedene Hops
  - Ø 12 Hops/Route



Leistungsfähigkeit des Internets
→ R.-Messung (verschiedene Server)

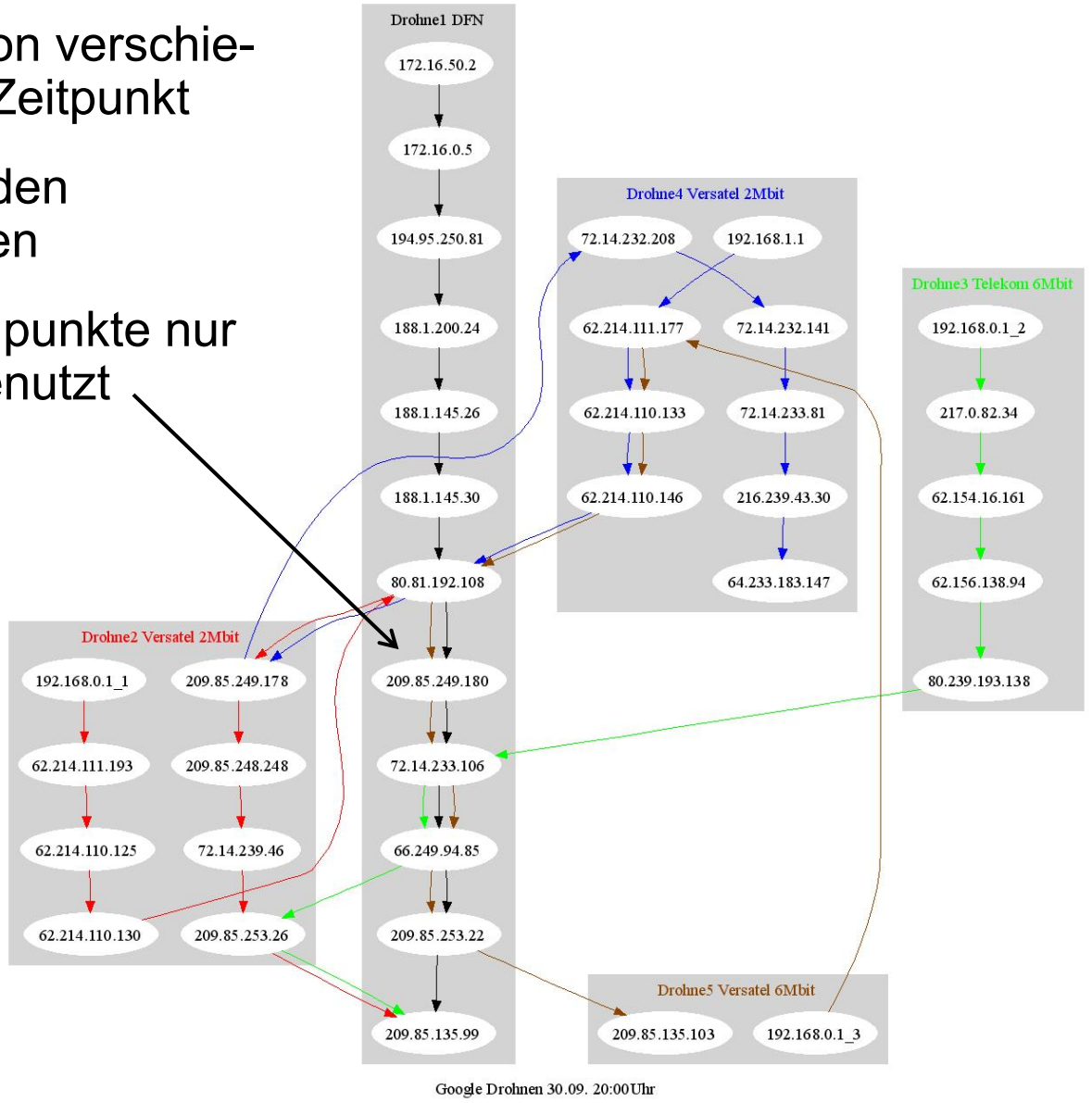
- **Beispiel:**
  Routing von einem Standort (Versatel) zu verschiedenen Servern zu einem Zeitpunkt
- Der überwiegende Teil der Route wird separat durchlaufen
- Verzögerungen durch längere oder ausgelastete Routen wirken sich nur auf einzelne Dienste aus



Google Drohne 5 30.09. 20:00Uhr

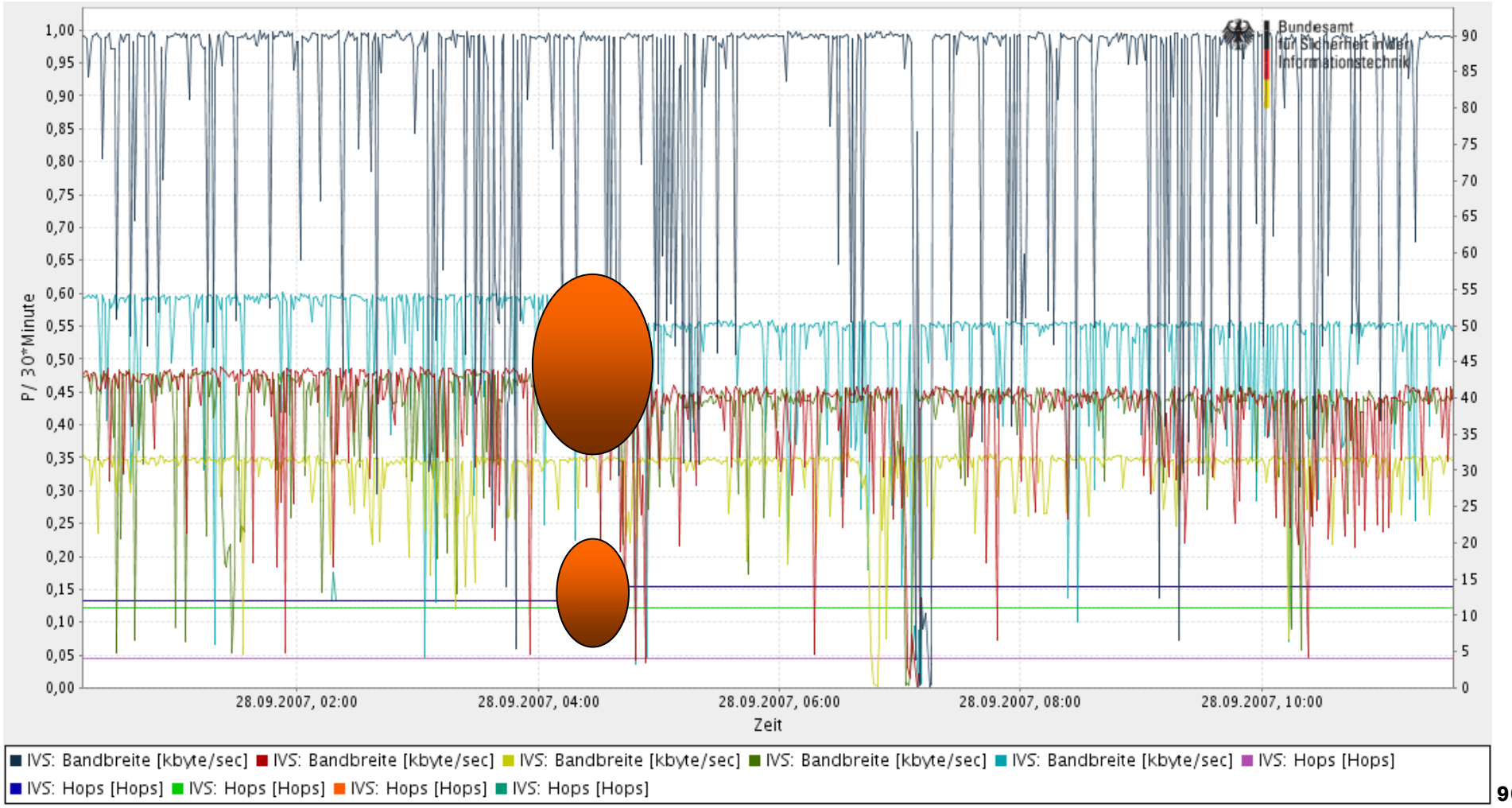
Leistungsfähigkeit des Internets
→ R.-Messung (verschiedene Netze)

- **Beispiel:** Route zu Google von verschiedenen Netzen aus zu einem Zeitpunkt
- Es gibt keinen Knotenpunkt, den alle Routen gemeinsam nutzen
- Es werden bestimmte Knotenpunkte nur von schnelleren Leitungen genutzt
- Selbst bei gleichen Anschlüssen und örtlicher Nähe werden unterschiedliche Routen benutzt



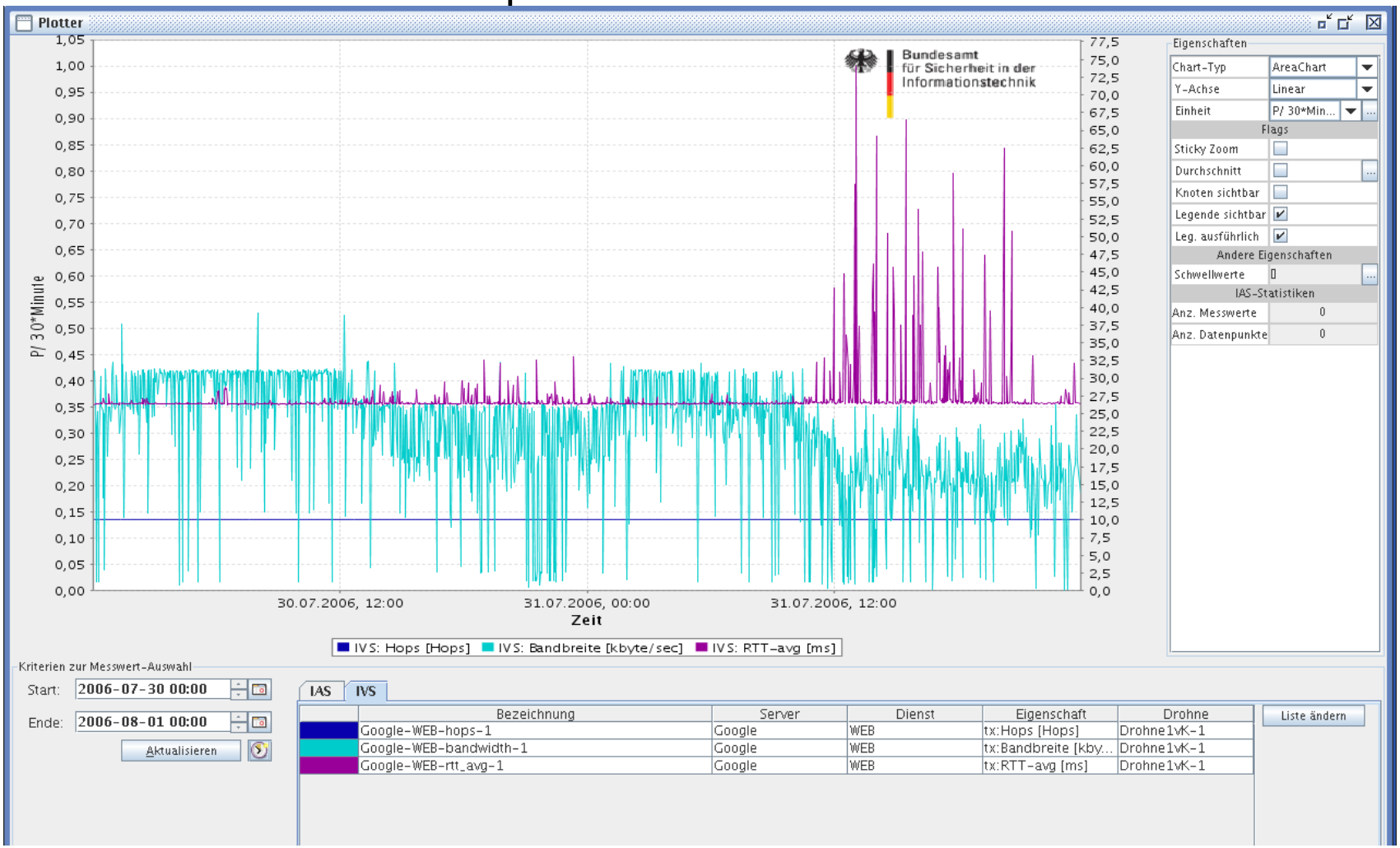
Messung if(is):
→ Beeinflussung BW durch mehrere Hops

- Route im Versatel-Netz wurde um 2 Hops erhöht: BW sinkt
- Telekom- und DFN-Netz werden nicht beeinflusst



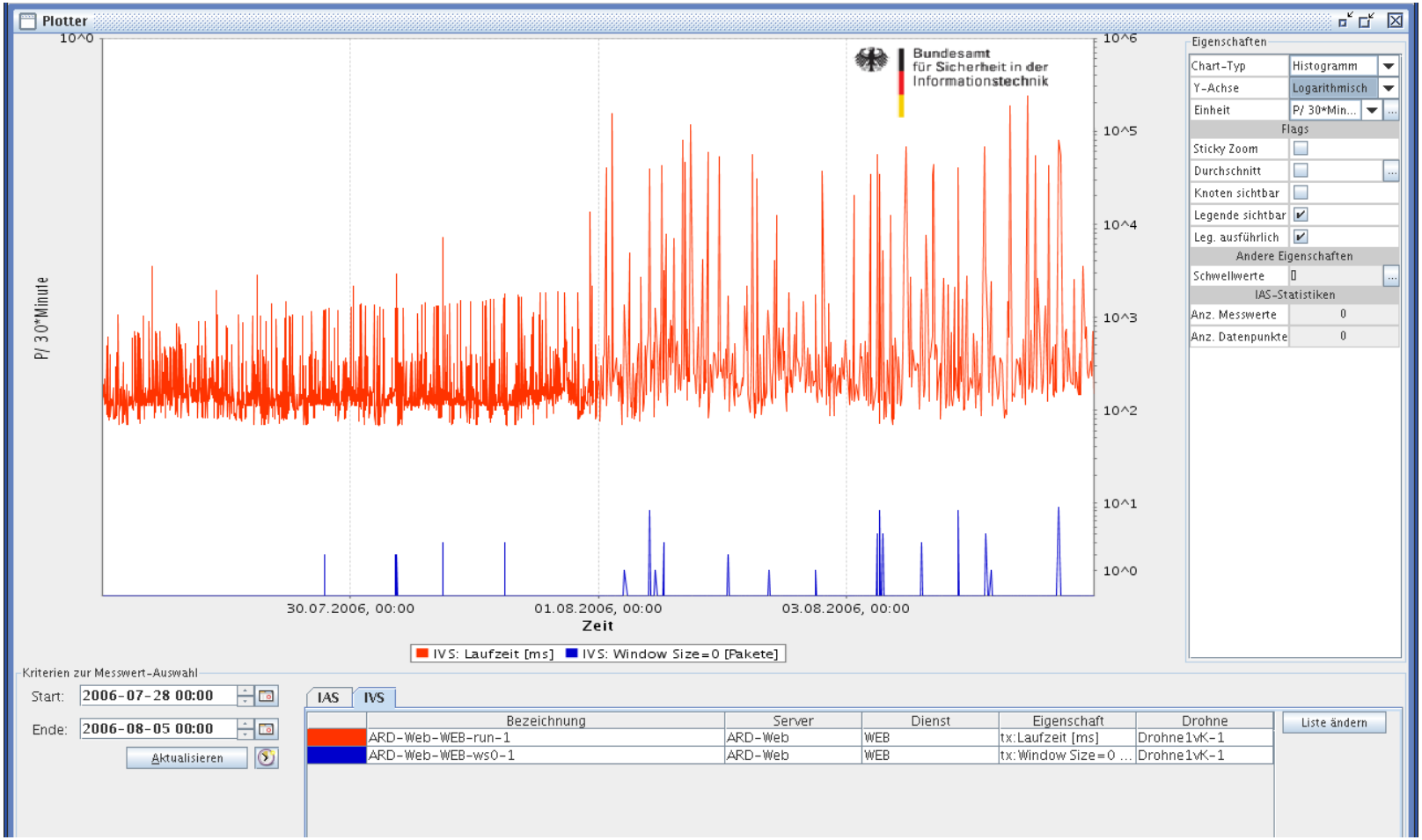
Messung Google:
→ Bandbreite – RTT-avg - Hops

- Hohe Auslastung der Routen wird durch RTT signalisiert
- Bandbreite sinkt – Hauptverkehrszeit: 12:00Uhr-0:00Uhr



Messung ARD:
→ Laufzeit - WS 0

- Höhere Auslastung eines Servers wird mit WS0 quittiert
- Die Laufzeit der Pakete (logarithmisch) erhöht sich stark, da Anfragen nicht oder nicht ausreichend beantwortet werden können



Leistungsfähigkeit des Internets
→ Auswertung Messung

- Drohne 1: FH-Gelsenkirchen, DFN-Netz
- Drohne 2: Arnsberg, Versatel, 2 Mbit
- Drohne 3: Datteln, Telekom, 6 Mbit
- Drohne 4: Arnsberg, Versatel, 2 Mbit
- Drohne 5: Arnsberg, Versatel, 6 Mbit
- Aussagen über Qualität/Geschwindigkeit der Netze bzw. Server
  - Ermittlung der durchschnittlichen Bandbreite über eine Woche(KW36: 03.09. - 10.09)

KW 36						
Bandbreite [kbyte/s]	Drohne1	Drohne2	Drohne3	Drohne4	Drohne5	ø
Amazon	69,53	55,88	49,56	55,57	-	57,64
Ebay	47,23	44,47	38,56	44,55	48,25	44,61
Google	57,13	29,82	21,31	30,01	52,47	38,15
IFIS	78,72	36,61	29,72	36,56	45,96	45,51
MyVideo	217,22	91	53,06	88,57	172,38	124,45
ø	93,97	51,56	38,44	51,05	79,77	

Auswertung Messung

- Über die Mittelung der Drohnen-Werte (waagerecht) kann ein ungefährer Vergleich der Server durchgeführt werden:
- Relativ stabil, Änderung der Werte über mehrere Wochen <10%

KW 36	
Server	øBW [kbyte/s]
T-Online	216,91
MyVideo	124,45
Heise	104,55
Amazon	57,64
Youtube	53,52
IFIS	45,51
Ebay	44,61
Google	38,15
Rapidshare	18,71

- Über die Mittelung der Server-Werte (senkrecht) kann ein Vergleich der Internet-Zugänge durchgeführt werden:
- Konstante Werte, Änderung über mehrere Wochen <3%

6 Mbit/s

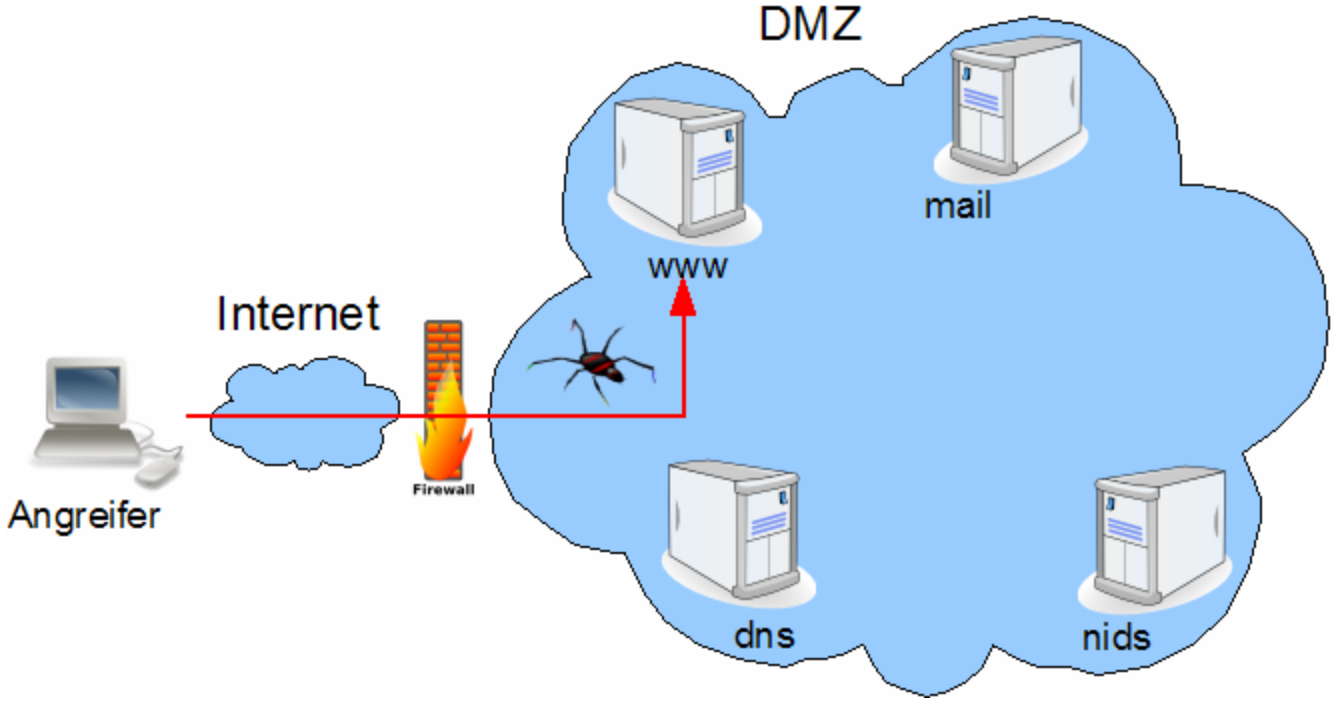
KW 36	
Drohnen	øBW [kbyte/s]
D1 DFN	126,95
D5 Versatel 6	79,77
D2 Versatel 2	64,92
D4 Versatel 2	64,61
D3 Telekom 6	57,87

Logdatenanalyse-System
→ Allgemeines über Logdaten

- werden oft vernachlässigt und unterschätzt
- enthalten zur Systemadministration wichtige Informationen über
  - Ressourcen-Engpässe
  - Hard- und Softwareprobleme
  - **Sicherheitsprobleme und Angriffe**
- können als Grundlage zur juristischen Verfolgung von Angreifern eingesetzt werden

Logdatenanalyse-System
→ Idee (1/4)

- Öffentliche Dienste wie www, E-Mail, etc. sind zunehmend Angriffen aus dem Internet ausgesetzt.
- Diese Angriffe verursachen, wie jeder andere Zugriff auf einen Dienst auch, Einträge in den Serverlogs.
- Angriffe erzeugen dabei markante Muster.

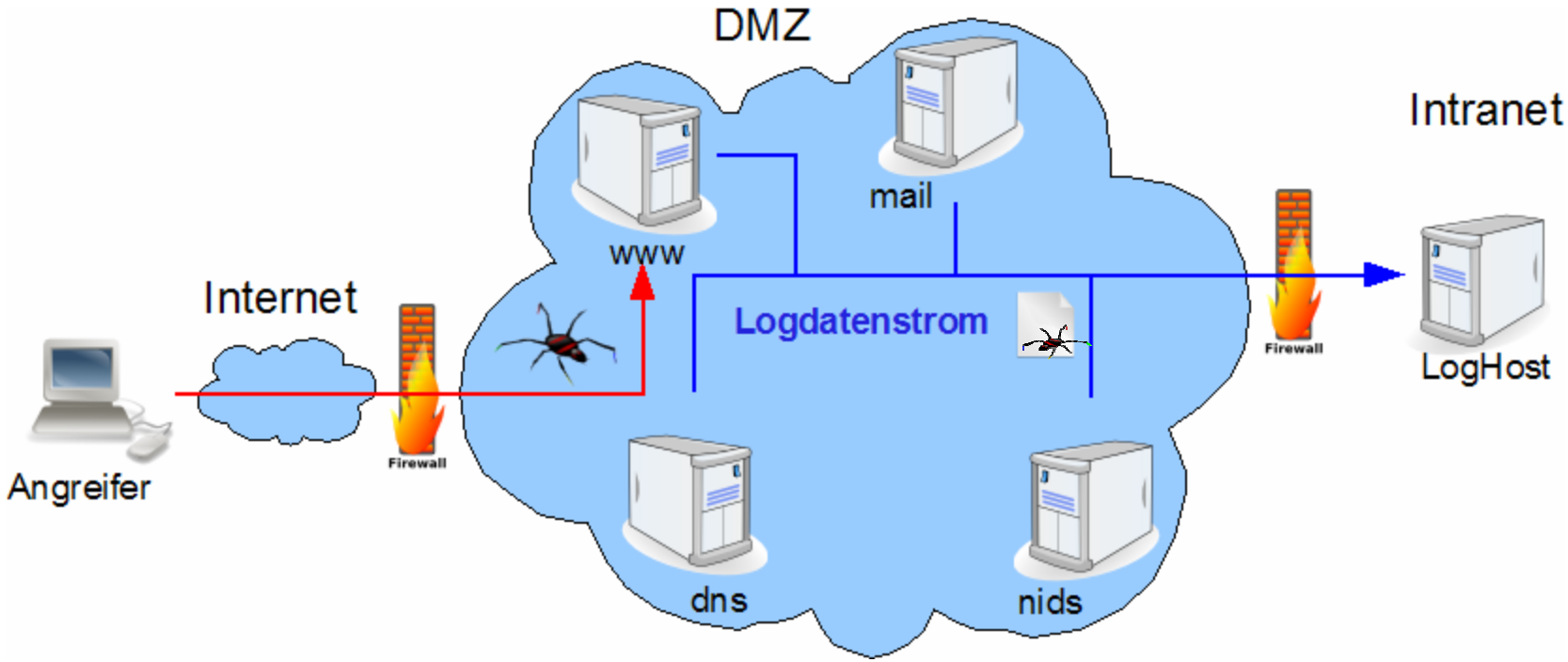


Logdatenanalyse-System
→ Mögliche Logquellen

- Zur Erkennung sicherheitsrelevanter Ereignisse eignen sich besonders Dienste mit Anbindung an das Internet, wie. z.B.
  - Firewall (z.B. Iptables)
  - Mail-Server (z.B. Sendmail, Postfix)
  - Web-Server (z.B. Apache)
  - VPN-Server (z.B. OpenVPN)
  - DNS-Server (z.B. BIND)
  - VoIP-Server (z.B. Asterisk)
  - NIDS (Netzwork Intrusion Detection Systems, z.B. Snort)
  - Remote Shell (z.B. sshd)
  - etc.

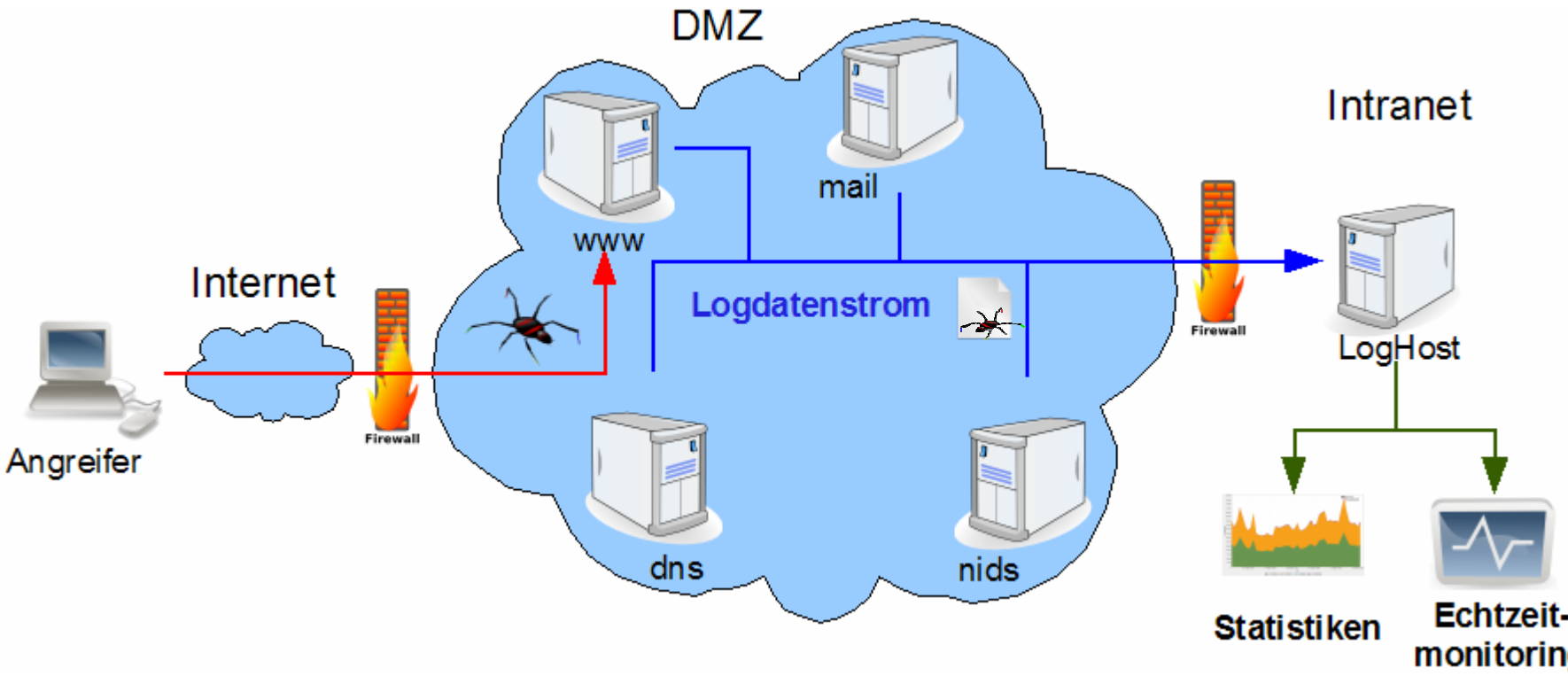
Logdatenanalyse-System
→ Idee (2/4)

- Die Logdaten aller Server werden als Live-Datenstrom an einem Log-Host / Log-Collector zusammengeführt
  - **Centralized Logging**
  - Transport über das **syslog**-Protokoll
  - Ermöglicht Korrelation der Daten



Logdatenanalyse-System
→ Idee (3/4)

- Der gebündelte Livedatenstrom wird auf Angriffsmuster untersucht.
  - **Echtzeitanalyse**
  - **Langzeitanalyse**

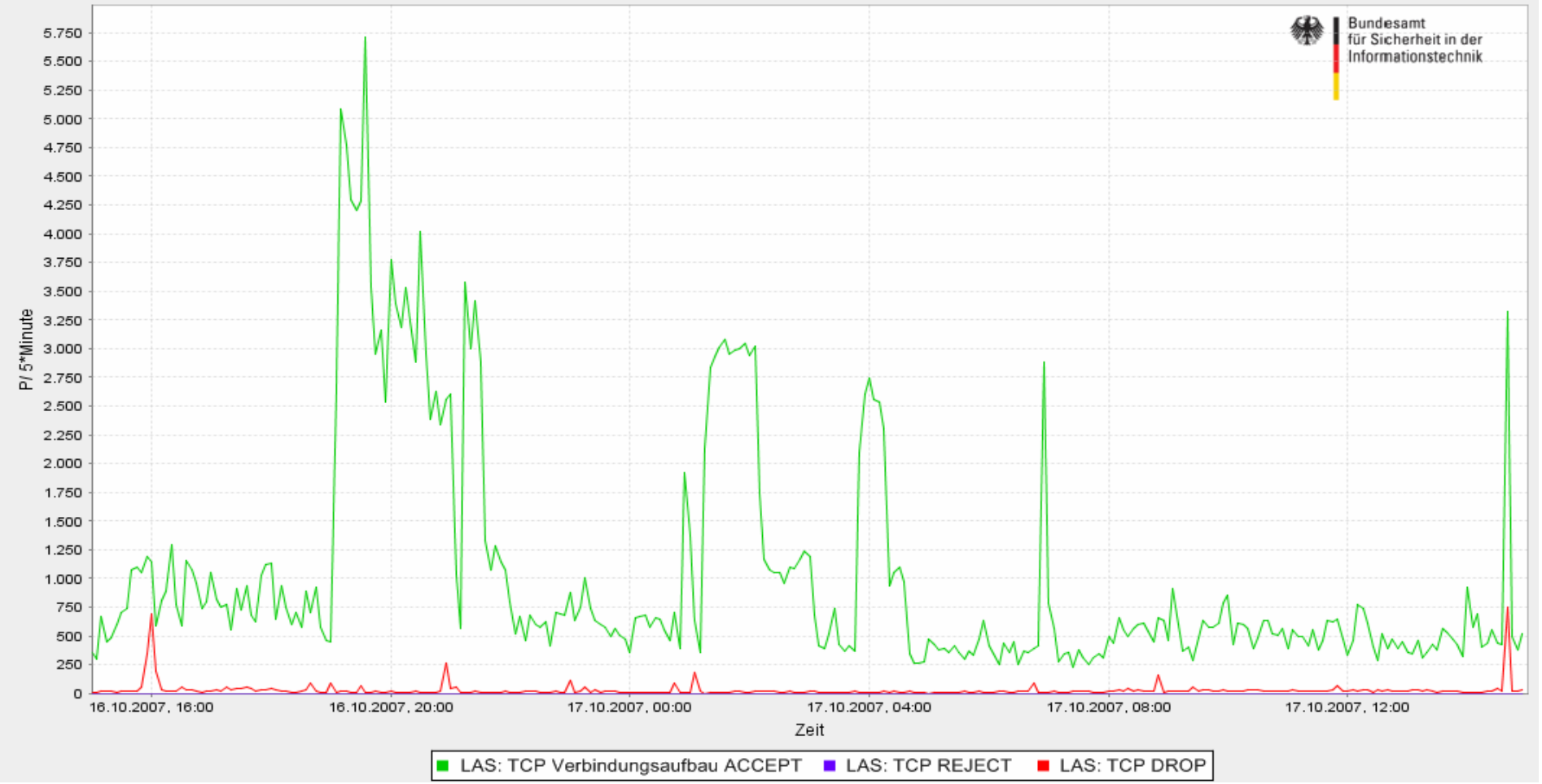


Logdatenanalyse-System
→ Idee (4/4)

- **Echtzeitanalyse**
  - Überprüfung des Logdatenstroms in nahezu Echtzeit auf Angriffssignaturen mit anschließender Alarmierung
  - Ermöglicht zeitnahe Reaktion auf einen Angriff
    - Dienste anhalten, Ports schließen, User-Accounts sperren, ...
- **Langzeitanalyse**
  - Statistische Auswertung der Daten
  - Anwendung des Prinzips der Deskriptoren auf geloggte Ereignisse
    - Anonymisierung der Daten
    - Graphische Darstellung zeitlicher Häufigkeitsverläufe
  - Zusammenführung der Daten mehrerer Netze lässt Aussagen über den Zustand des Internets zu

Logdatenanalyse-System

→ Langzeitanalyse - Logdaten als Deskriptoren



Logdatenanalyse-System
→ Echtzeitanalyse - Techniken

- **Baselining**
  - Erkennung einzelner ungünstiger Ereignisse
  - **Anomaly Detection**
    - Identifiziert alles, “was man bisher noch nicht gesehen hat“
  - **Thresholding**
    - Prüft Anzahl definierter Ereignisse gegen Schwellenwerte, die das System im normalen Betrieb aufweist
  - **Windowing**
    - Identifiziert Ereignisse, die über Parameter außerhalb eines erwarteten Wertebereichs liegen
- **Correlation**
  - Stellt Beziehungen zwischen scheinbar voneinander unabhängigen Ereignissen her, etwa zwischen Logdaten verschied. Anwendungen

Logdatenanalyse-System
→ Beispiele (1/4) – Anomaly Detection

- Buffer-Overflow-Versuch auf einen SSH-Deamon
- Logeintrag ist auf normalen Systemen unüblich

```
Oct 11 14:27:26 listserver sshd[6169]: fatal: Local: Corrupted check
bytes                                       on input.
Oct 11 14:27:28 listserver sshd[6253]: fatal: Local: crc32 compensation
                                           attack: network attack detected
```

Logdatenanalyse-System
→ Beispiele (2/4) - Thresholding

- Dictionary-Attack auf einen SSH-Deamon
- Anzahl so vieler ungültiger Anmeldeversuche ist unüblich

```
Oct  8 12:43:26 listserver sshd[7379]: Invalid user smo from ::ffff:210.188.206.248
Oct  8 12:43:28 listserver sshd[7381]: Invalid user sashroot from ::ffff:210.188.206.248
Oct  8 12:43:30 listserver sshd[7383]: Invalid user deddy from ::ffff:210.188.206.248
Oct  8 12:43:33 listserver sshd[7385]: Invalid user sysmin from ::ffff:210.188.206.248
Oct  8 12:43:35 listserver sshd[7387]: Invalid user clamav from ::ffff:210.188.206.248
Oct  8 12:43:37 listserver sshd[7389]: Invalid user chris from ::ffff:210.188.206.248
Oct  8 12:43:40 listserver sshd[7391]: Invalid user christia from ::ffff:210.188.206.248
Oct  8 12:43:42 listserver sshd[7393]: Invalid user fam from ::ffff:210.188.206.248
Oct  8 12:43:45 listserver sshd[7395]: Invalid user helma from ::ffff:210.188.206.248
Oct  8 12:43:47 listserver sshd[7397]: Invalid user manfred from ::ffff:210.188.206.248
Oct  8 12:43:49 listserver sshd[7399]: Invalid user spaadmin from ::ffff:210.188.206.248
Oct  8 12:43:52 listserver sshd[7422]: Invalid user grafik from ::ffff:210.188.206.248
Oct  8 12:43:54 listserver sshd[7424]: Invalid user martha from ::ffff:210.188.206.248
Oct  8 12:43:56 listserver sshd[7426]: Invalid user testie from ::ffff:210.188.206.248
Oct  8 12:43:59 listserver sshd[7428]: Invalid user walter from ::ffff:210.188.206.248

                    ...

Oct  8 20:07:52 listserver sshd[30886]: Invalid user nagios from ::ffff:210.188.206.248
Oct  8 20:07:54 listserver sshd[30890]: Invalid user victoria from ::ffff:210.188.206.248
Oct  8 20:07:56 listserver sshd[30892]: Invalid user schweitzer from ::ffff:210.188.206.248
Oct  8 20:07:59 listserver sshd[30894]: Invalid user finanzen from ::ffff:210.188.206.248
Oct  8 20:08:01 listserver sshd[30896]: Invalid user jonathan from ::ffff:210.188.206.248
Oct  8 20:08:03 listserver sshd[30898]: Invalid user bouncer from ::ffff:210.188.206.248
Oct  8 20:08:06 listserver sshd[30900]: Invalid user mywebeditde from ::ffff:210.188.206.248
Oct  8 20:08:08 listserver sshd[30902]: Invalid user schlesier from ::ffff:210.188.206.248
Oct  8 20:08:10 listserver sshd[30904]: Invalid user klein from ::ffff:210.188.206.248
```

Logdatenanalyse-System
→ Beispiele (3/4) - Windowing

- Ausschnitt aus den iptables-Logs der Fachbereichs-Firewall
- Wo ist der Angriff?

```
Oct 10 04:19:41 fb5gwint info kern kernel: forward Rule 157 - DENY IN=eth0 OUT=eth4
SRC=218.83.175.154 DST=194.94.127.90 LEN=44 TOS=0x00 PREC=0x00 TTL=242 ID=12233 PROTO=TCP
SPT=80 DPT=1116 WINDOW=8190 RES=0x00 ACK SYN URGP=0
Oct 10 04:19:44 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth2
SRC=61.134.60.146 DST=194.94.127.32 LEN=40 TOS=0x00 PREC=0x00 TTL=115 ID=54426 PROTO=TCP
SPT=2999 DPT=53783 WINDOW=0 RES=0x00 ACK RST URGP=0
Oct 10 04:20:08 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth2
SRC=58.221.28.199 DST=194.94.127.79 LEN=48 TOS=0x00 PREC=0x00 TTL=114 ID=14822 DF PROTO=TCP
SPT=6020 DPT=37105 WINDOW=65535 RES=0x00 ACK SYN URGP=0
Oct 10 04:20:08 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth0
SRC=218.83.175.154 DST=194.94.127.117 LEN=44 TOS=0x00 PREC=0x00 TTL=242 ID=29926 PROTO=TCP
SPT=80 DPT=1262 WINDOW=8190 RES=0x00 ACK SYN URGP=0
Oct 10 04:20:21 fb5gwint info kern kernel: forward Rule 157 - DENY IN=eth0 OUT=eth4
SRC=12.158.171.206 DST=194.94.127.84 LEN=1500 TOS=0x00 PREC=0x00 TTL=114 ID=5857 DF PROTO=TCP
SPT=3826 DPT=25 WINDOW=64011 RES=0x00 ACK URGP=0
Oct 10 04:20:21 fb5gwint info kern kernel: forward Rule 157 - DENY IN=eth0 OUT=eth4
SRC=12.158.171.206 DST=194.94.127.84 LEN=1500 TOS=0x00 PREC=0x00 TTL=114 ID=5859 DF PROTO=TCP
SPT=3826 DPT=25 WINDOW=64011 RES=0x00 ECE URG RST SYN FIN URGP=0
Oct 10 04:20:50 fb5gwint info kern kernel: Internet Rule 3 - REJECT IN= OUT=eth0
SRC=193.175.172.98 DST=58.221.246.21 LEN=76 TOS=0x00 PREC=0xC0 TTL=64 ID=53921 PROTO=ICMP
TYPE=3 CODE=1 [SRC=58.221.246.21 DST=194.94.127.10 LEN=48 TOS=0x00 PREC=0x00 TTL=114 ID=6194 DF
PROTO=TCP SPT=6021 DPT=7904 WINDOW=65535 RES=0x00 ACK SYN URGP=0 ]
Oct 10 04:21:38 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth2
SRC=67.15.83.36 DST=194.94.127.76 LEN=40 TOS=0x00 PREC=0x00 TTL=47 ID=27110 DF PROTO=TCP SPT=22
DPT=1024 WINDOW=0 RES=0x00 ACK RST URGP=0
```

Logdatenanalyse-System
→ Beispiele (3/4) - Windowing

- SYN/FIN-Portscan gegen den SMTP-Port des FB-Mailservers
- SYN/FIN-Flagkombination laut RFC 793 nicht zulässig

```
Oct 10 04:19:41 fb5gwint info kern kernel: forward Rule 157 - DENY IN=eth0 OUT=eth4
SRC=218.83.175.154 DST=194.94.127.90 LEN=44 TOS=0x00 PREC=0x00 TTL=242 ID=12233 PROTO=TCP
SPT=80 DPT=1116 WINDOW=8190 RES=0x00 ACK SYN URGP=0
Oct 10 04:19:44 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth2
SRC=61.134.60.146 DST=194.94.127.32 LEN=40 TOS=0x00 PREC=0x00 TTL=115 ID=54426 PROTO=TCP
SPT=2999 DPT=53783 WINDOW=0 RES=0x00 ACK RST URGP=0
Oct 10 04:20:08 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth2
SRC=58.221.28.199 DST=194.94.127.79 LEN=48 TOS=0x00 PREC=0x00 TTL=114 ID=14822 DF PROTO=TCP
SPT=6020 DPT=37105 WINDOW=65535 RES=0x00 ACK SYN URGP=0
Oct 10 04:20:08 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth0
SRC=218.83.175.154 DST=194.94.127.117 LEN=44 TOS=0x00 PREC=0x00 TTL=242 ID=29926 PROTO=TCP
SPT=80 DPT=1262 WINDOW=8190 RES=0x00 ACK SYN URGP=0
Oct 10 04:20:21 fb5gwint info kern kernel: forward Rule 157 - DENY IN=eth0 OUT=eth4
SRC=12.158.171.206 DST=194.94.127.84 LEN=1500 TOS=0x00 PREC=0x00 TTL=114 ID=5857 DF PROTO=TCP
SPT=3826 DPT=25 WINDOW=64011 RES=0x00 ACK URGP=0
Oct 10 04:20:21 fb5gwint info kern kernel: forward Rule 157 - DENY IN=eth0 OUT=eth4
SRC=12.158.171.206 DST=194.94.127.84 LEN=1500 TOS=0x00 PREC=0x00 TTL=114 ID=5859 DF PROTO=TCP
SPT=3826 DPT=25 WINDOW=64011 RES=0x00 ECE URG RST SYN FIN URGP=0
Oct 10 04:20:50 fb5gwint info kern kernel: Internet Rule 3 - REJECT IN= OUT=eth0
SRC=193.175.172.98 DST=58.221.246.21 LEN=76 TOS=0x00 PREC=0xC0 TTL=64 ID=53921 PROTO=ICMP
TYPE=3 CODE=1 [SRC=58.221.246.21 DST=194.94.127.10 LEN=48 TOS=0x00 PREC=0x00 TTL=114 ID=6194 DF
PROTO=TCP SPT=6021 DPT=7904 WINDOW=65535 RES=0x00 ACK SYN URGP=0 ]
Oct 10 04:21:38 fb5gwint info kern kernel: forward Rule 197 - DENY IN=eth0 OUT=eth2
SRC=67.15.83.36 DST=194.94.127.76 LEN=40 TOS=0x00 PREC=0x00 TTL=47 ID=27110 DF PROTO=TCP SPT=22
DPT=1024 WINDOW=0 RES=0x00 ACK RST URGP=0
```

Logdatenanalyse-System
→ Beispiele (4/4) - Correlation

- Auszug aus **/var/log/deamon.log**

```
Oct 09 17:47:03 host in.ftpd[16273]: connect from 202.10.30.49
```

- Auszug aus **/var/log/auth.log**

```
Oct 09 17:47:08 host PAM_unix[16273]: check pass; user unknown
Oct 09 17:47:08 host PAM_unix[16273]: authentication failure; (uid=0)
                                          -> **unknown** for ftp
service
Oct 09 17:47:13 host PAM_unix[16273]: check pass; user unknown
Oct 09 17:47:13 host PAM_unix[16273]: authentication failure; (uid=0)
                                          -> **unknown** for ftp
service
```

- Über die identische Prozess-ID und den zeitlichen Bezug lässt sich ein Bezug zwischen den Logeinträgen herstellen.
- So lässt sich die IP-Adresse des Angreifers ermitteln

Logdatenanalyse-System
→ Stärken

- NIDS schlagen Alarm, sobald ein Angriff auf der Leitung erkannt werden konnte
  - Keine Informationen über Ablauf und Ausgang der Attacke, da der Angriff bei Erkennung noch nicht ausgeführt war
- Logdaten beschreiben,
  - ob ein Angriff erfolgreich war
  - was der Angreifer auf dem Zielsystem getan hat
    - ermöglichen Rekonstruktion des Angriffsablaufs
  - wie ein System auf seine Eingangsdaten reagiert und diese interpretiert
- Korrelation der Logdaten ermöglicht es, verteilte Angriffe, die sich parallel gegen mehrere Rechner richten, zu erkennen

Logdatenanalyse-System
→ Schwächen

- Logdaten sind extrem umfangreich
  - sehr unübersichtlich (siehe Beispiel (3/4))
  - nützliche Infos sind versteckt
  - nur geringer Anteil sicherheitsrelevanter Infos < 5%
- Es gibt kein Standard-Logformat
  - Logs verschiedener Anwendungen können unterschiedlich sein, auch wenn sie das gleiche Ereignis beschreiben
  - Erschwert die maschinelle Verarbeitung und Interpretation
- Logdatenformate sind oft schlecht oder gar nicht dokumentiert
- Logdatenanalyse erfordert ein genaues Verständnis der Anwendungen und technischen Abläufe, aus denen die Daten hervor gehen (**Expertenwissen**)

- **Structure of the Internet**
  - **Connectivity of the Internet**

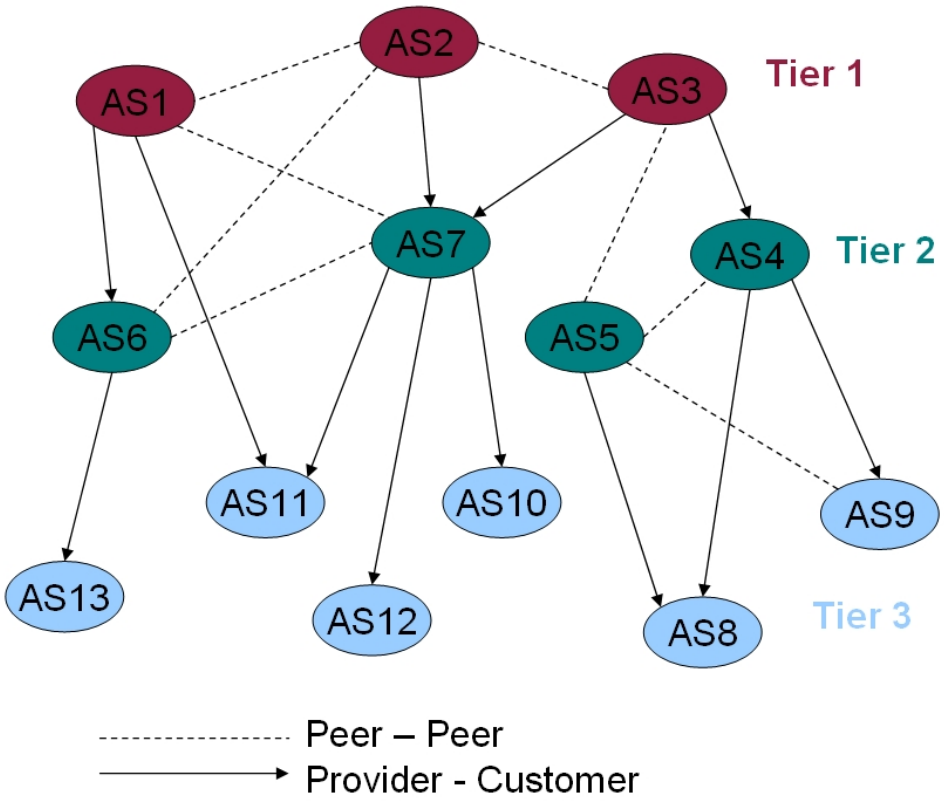
Structure of the Internet
→Autonomous Player

- **Autonomous Systems (AS)**
  - The global Internet consists of thousands of independent networks, the Autonomous Systems (AS)
  - Actually there are about 22.000 different ASs advertised in the global Routing table
  - The AS operators have different policies for the size and expansion of their network
  - An AS needs a strategy to connect with other ASs using upstreams, private or public peerings
  - There are more than 60.000 logical connections between ASs at the moment

- **Different types of Autonomous Systems**
  - Large Companies, e.g. business consumer (41 %)
  - Internet Service Providers, e.g. IP-carrier (35 %)
  - Universities (11 %)
  - Internet Exchange Points, e.g. public data exchange nodes (2 %)
  - …

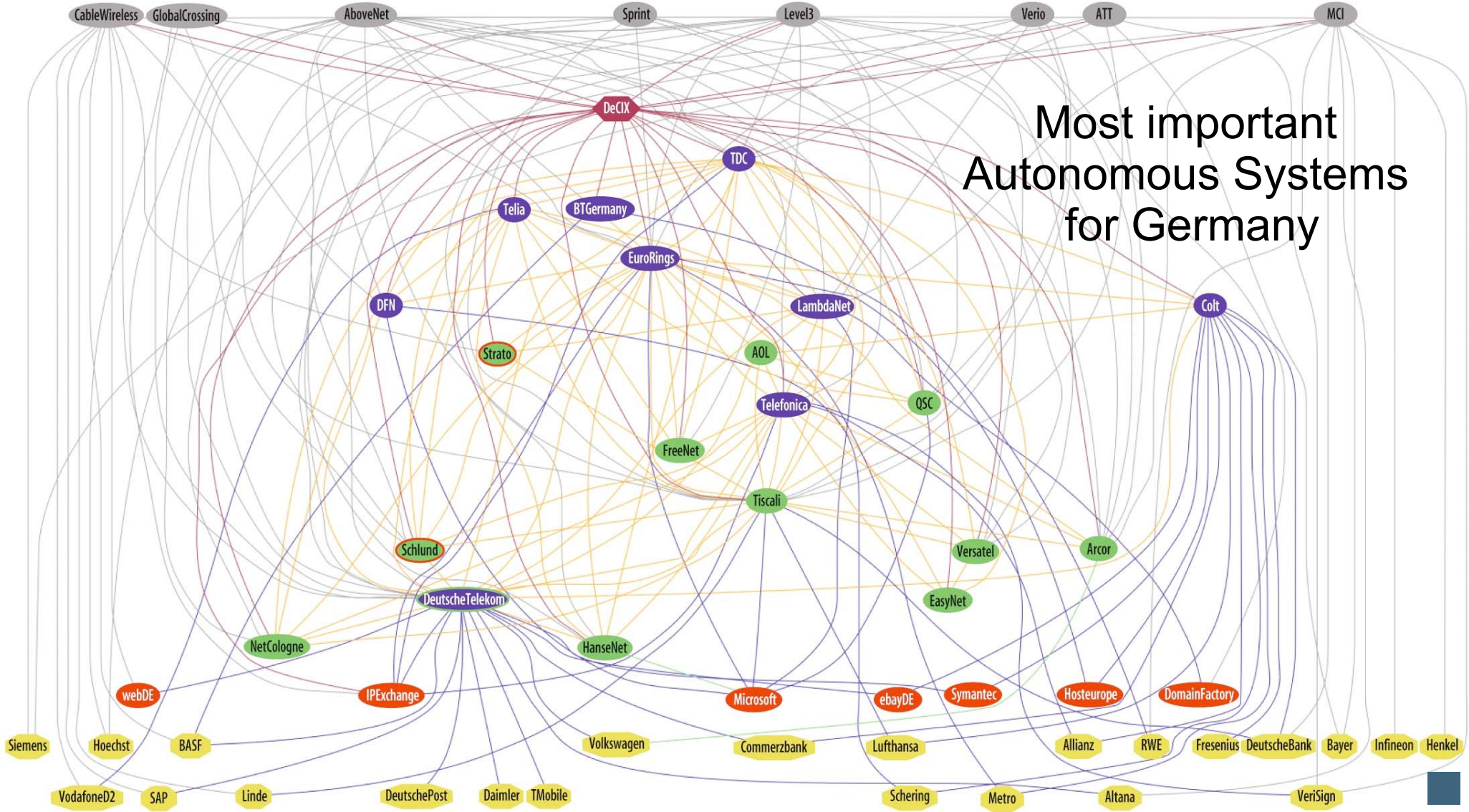
Structure of the Internet
→Connectivity of the Internet

- **Ongoing analysis on the Route Views Snapshot**
  - <= 2 = 63 %
    <=10 = 94 %
    >10 = 6 %
    >100 = 0,4 %
    >300 = 0,1 %
  - Economical necessities affect the carrier‘s proceeding
  - This yields to a destabilization of the internet infrastructure



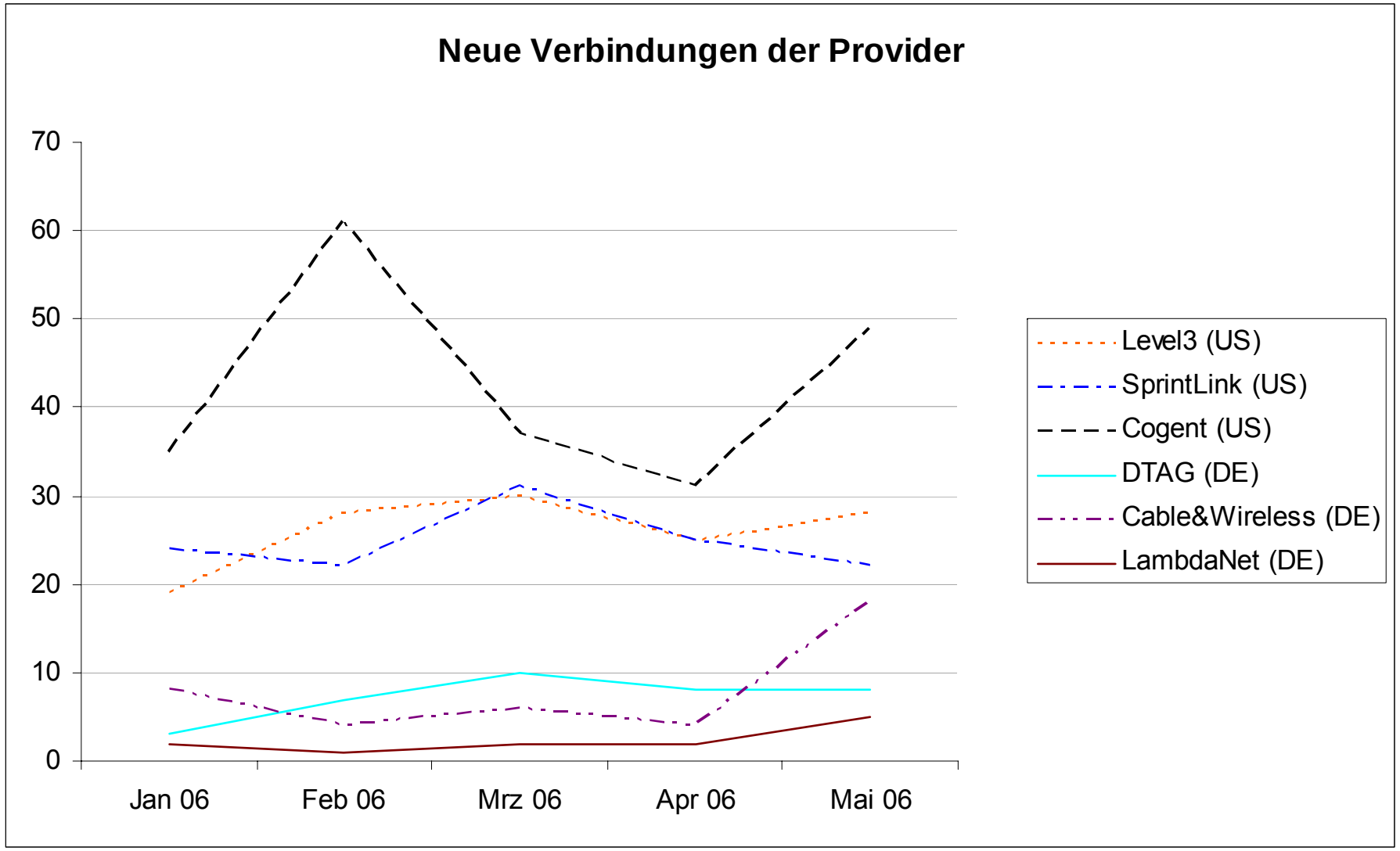
- **What is imported in this field?**
  - We need an entity which keeps an eye on the level of connection and the reliability of all ASs in the Internet

Structure of the Internet
→ Analysis of „Internet Germany“



Most important Autonomous Systems for Germany

Structure of the Internet
→ Analysis of Internet

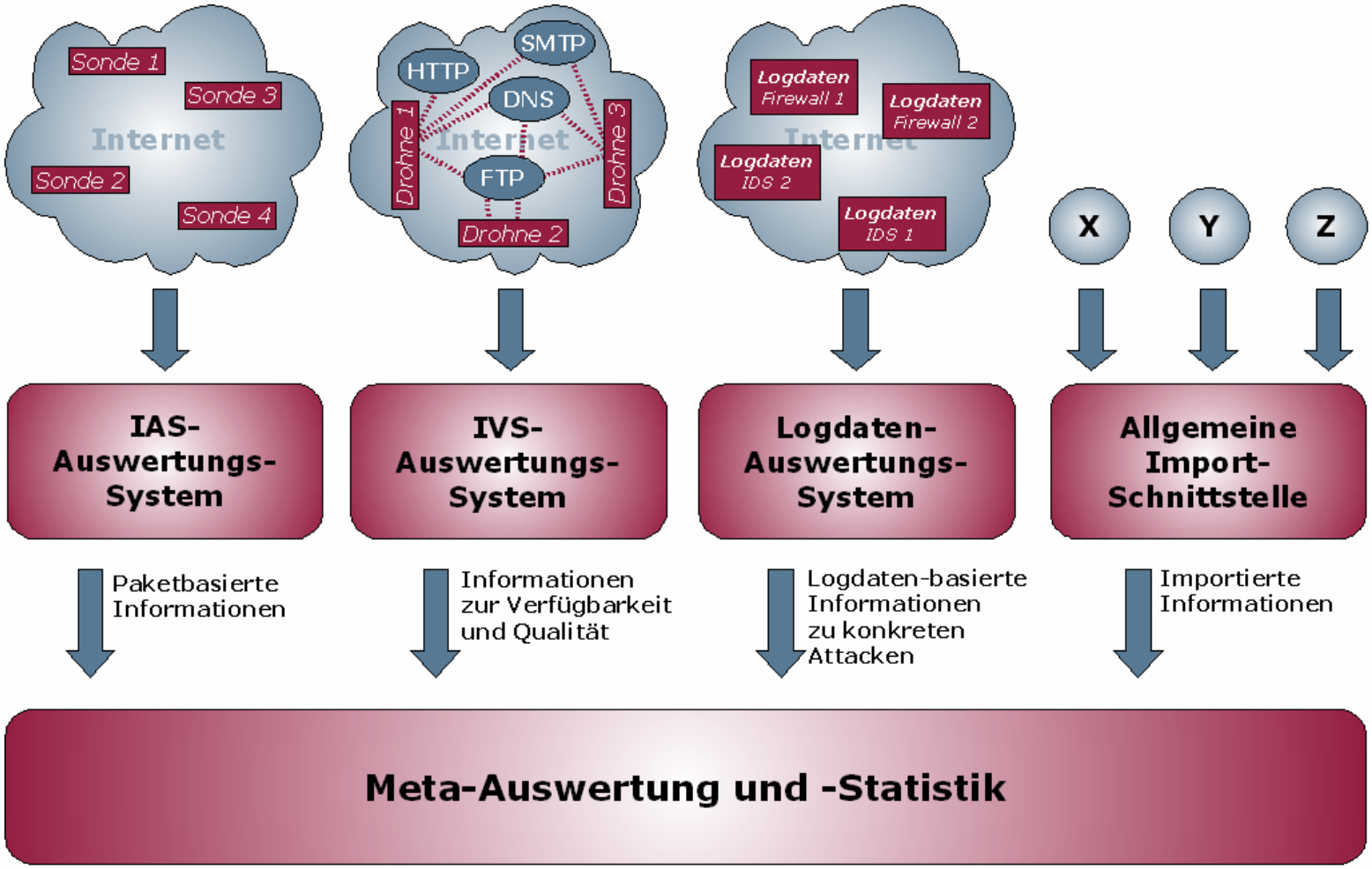


Agenda

- Einführung
- Frühwarnsysteme
- Struktur für Internet-Frühwarnsysteme
- Verschiedene Realisierungsansätze
- Internet-Analyse-System
- Internet-Verfügbarkeits-System
- **Zusammenfassung**

Zusammenfassung
→ Ausblick

- Erweiterungsmöglichkeiten des IAS / IVS



Zusammenfassung
→ Fazit

- Internet-Frühwarnsysteme sind nur unter Berücksichtigung vieler technischer, sozialer, organisatorischer und politischer Aspekte realisierbar
- Pilotphasen und Bereitschaft auch Beta-Versionen einzusetzen sind wichtige Voraussetzungen
- Ganzheitliche Ansätze für Internet-Frühwarnsysteme haben gute Chancen, flächendeckend eingeführt zu werden
- Einführung von wirkungsvollen, produktiven Internet-Frühwarnsystemen mit sinnvollen Funktionen und Schnittstellen können das Internet sicherer und vertrauenswürdiger machen

Fachhochschule Gelsenkirchen

Internet-Frühwarnsysteme

Vielen Dank für Ihre Aufmerksamkeit
Fragen ?

Prof. Dr. Norbert Pohlmann

Institut für Internet-Sicherheit
Fachhochschule Gelsenkirchen
https://www.internet-sicherheit.de

if(is) internet-sicherheit.