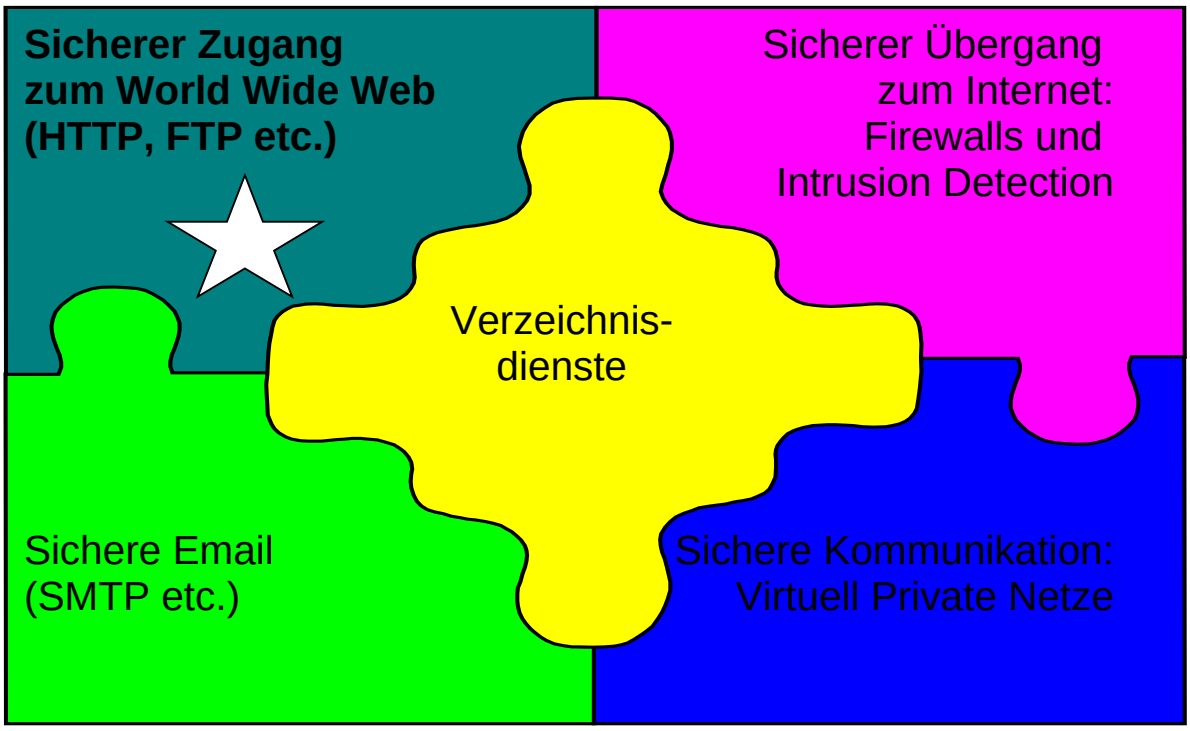
Design und Realisierung von E-Business- und Internet-Anwendungen

„Web-Zugang und Internet Sicherheit“

Dr. Stephen Heilbronner et al.
Prof. Dr. Heinz-Gerd Hegering

SoSe 2005

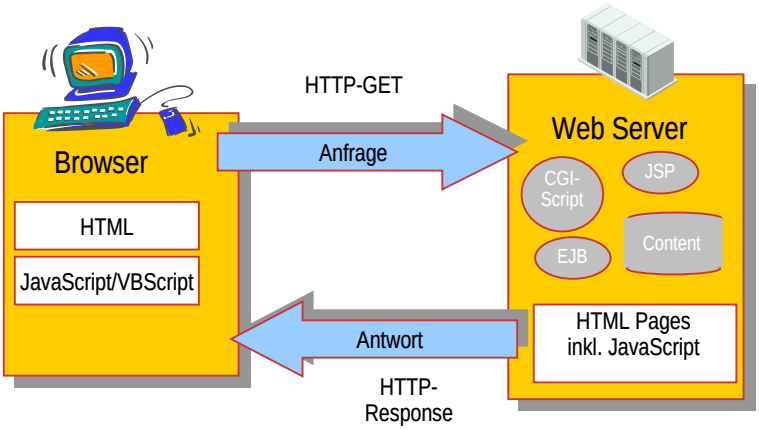
Sicherheitsdienste großer IT-Infrastrukturen => Überblick



Web-Zugang
Grundprinzip

Zugriff auf WWW-Server durch WWW-Clients:

- 1. Browser
- 2. andere, „automatische“ Programme



Formate:

- Nicht nur HTML !
- Auch: WML, XML, beliebige Dateiformate

Web-Zugriff
Veränderte Nutzuung des HTTP-Protokolls

Ursprüngliche Verwendung von HTTP:

- Übertragung statischer HTML-Seiten bzw. Dateien
- keine Unterscheidung zw. Anfragenden

Heutige Web-Zugriffe

- nicht nur mehr vom Typ „Request/Response“...
- häufig nicht anonym, sondern finden in einem längerdauernden Kontext statt, z.B. von
  - Individualisierung der ansonsten anonymem Anfragenden
    - session-spezifische Spracheinstellungen
  - Identifizierung / Authentisierung
    - „Zuordnung langlebiger Merkmale/Ressourcen“

Web-Zugang
Protokollentwicklung HTTP

Protokoll-Typ: „Request/Response“

Ursprünglich nur gedacht nur für „kurze“ Verbindungen:
HTTP 1.0 - IETF RFC

- Aufwendig: TCP-Handshake (3-way) beim Verbindungsaufbau

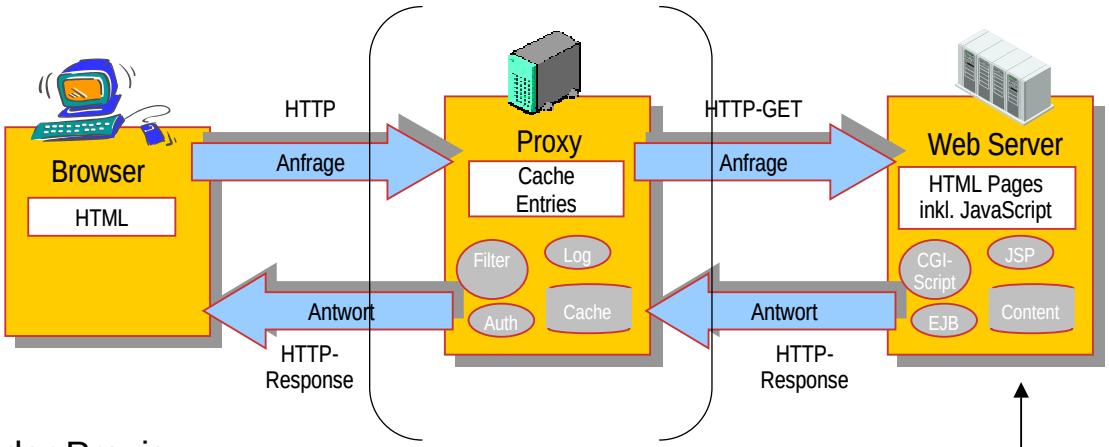
Heute: TCP-Verbindung bleibt nach Übertragung bestehen:
HTTP 1.1 - IETF RFC

- Kein Bedarf für Handshakes beim Abbau/Wiederaufbau
- Effizientere und schnellere Übertragung kleiner Informationsmengen (weniger Verzögerungen)

Integration in Proxies:

- Zieladresse nicht mehr direkt adressiert mit TCP/IP
- Alle Adressierungsinfos im HTTP-Header

Web-Zugang
Web-Architektur mit Proxies (HTTP)



In der Praxis:

Kette oft mehrfach wiederholt: „Proxy Chaining“

Großzügige Proxy-Auslegung wichtig für:

- Multimedia-Streaming in Echtzeit
- „Pre-pushed content“

Mehr zu dessen Architektur im Juni

Exkurs: AAA

Authentisierung
- Feststellung der Identität
- Implementierung:
  - Abfrage Benutzername/Passwort
  - Keycard

Autorisierung
- Festlegung der Nutzungsrechte
- hier: Welche weiteren Zugriffe sind erlaubt ?

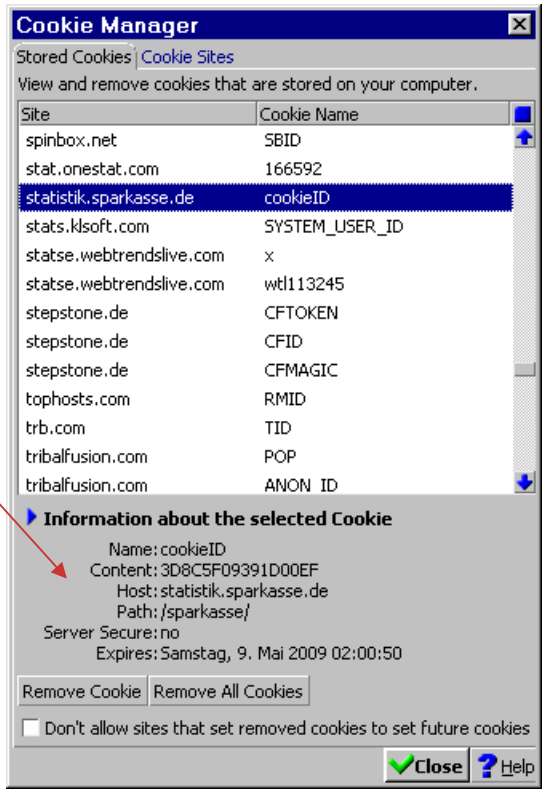
Accounting
- Aufzeichnung verrechnungsrelevanter Nutzungsdaten
- Aggregation der Daten
- Ziel: Verrechnung der Nutzen

AAA: Übertragung von Autorisierungs-Information

Cookies
- Kleine „Stücke von Information“
- Inhalte vom Server festgelegt
- Browser stellt sie bestimmten Servern zur Verfügung
- lange Lebensdauer

HTTP-Basic/Digest
- Autorisierunginformation im HTTP-Header

Oder: Im URL kodiert
- „Session-ID‘s“



Wie sieht so etwas aus ? ...

Web-Zugang
Proxies (für HTTP)

HTTP ist

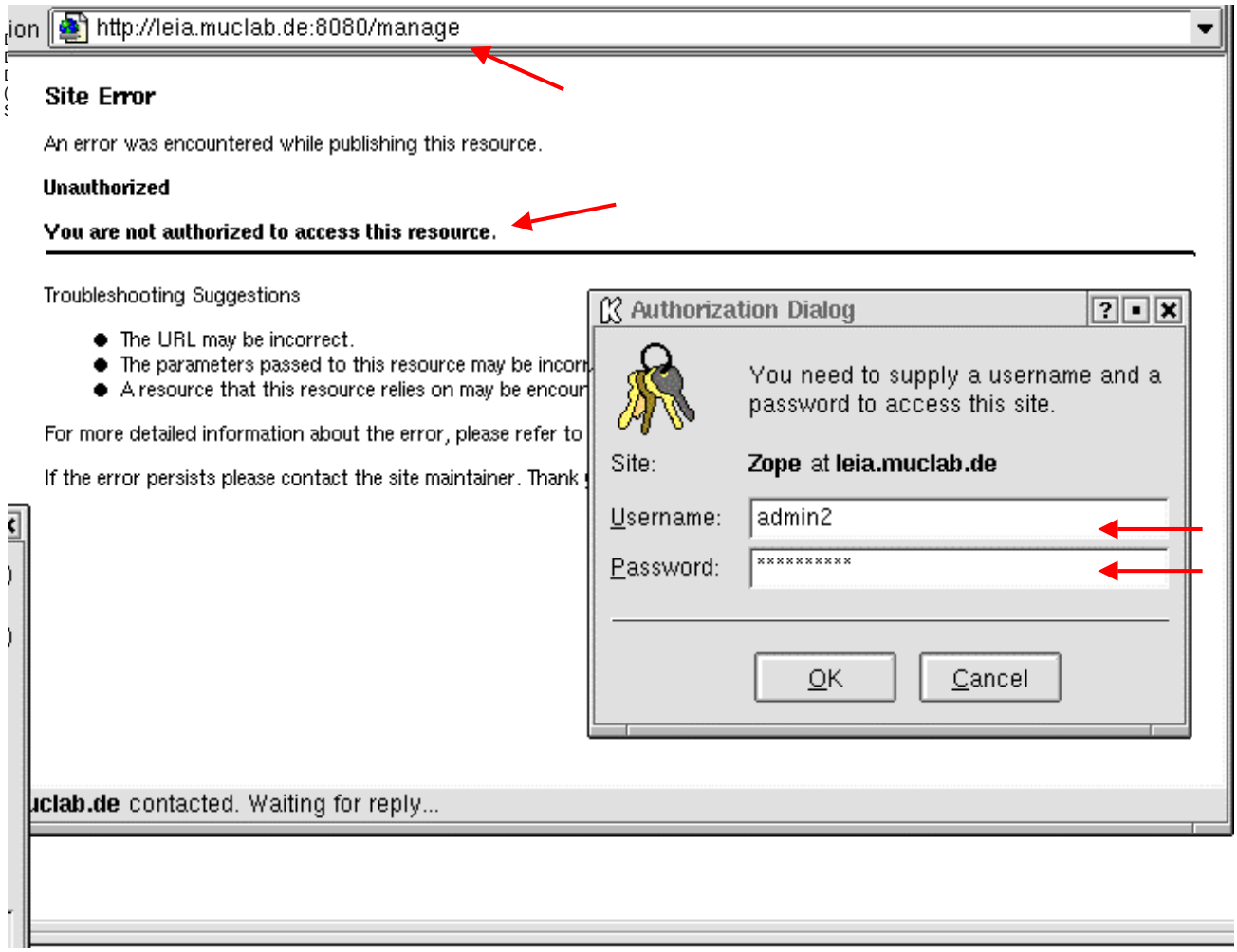
- entweder anonym, oder
- Authentisierungs-Information im Header

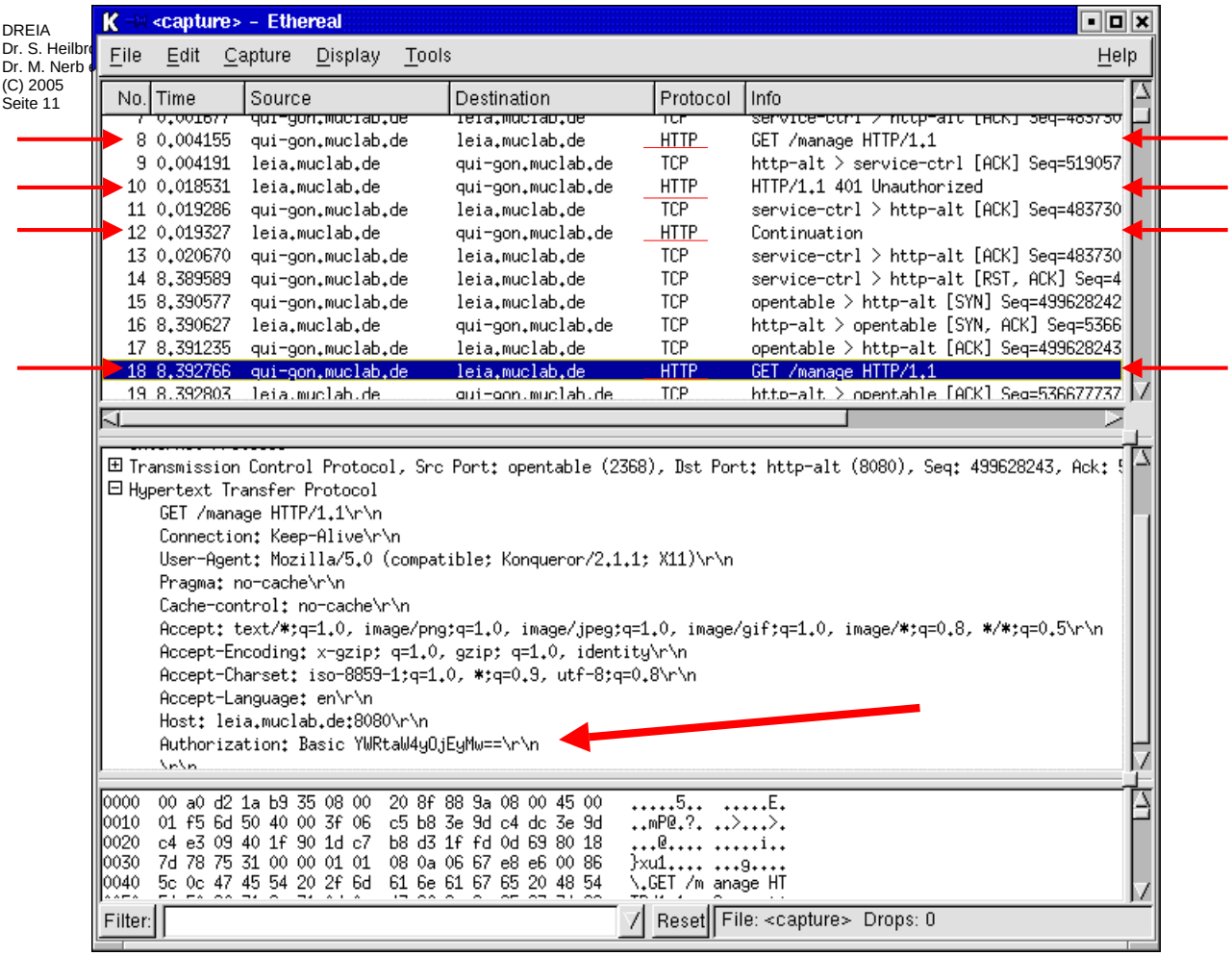
Autorisierende Proxies unterbrechen den Fluß zum Server

- erscheinen dem Client genauso wie geschützte Server
- verlangen Authentisierung vom Benutzer
- filtern diese vor Weitergabe wieder heraus

  echter Server benötigt dann eventuell weitere Autorisierung

Nachfolgendes Beispiel:
Zugriff von *qui-gon* auf *leia* ...





Web-Proxies
Backends

Abfrage von Informationen für Authentisierung und Autorisierung
- Benutzer
- Jeweilige Rechte
- Nutzungszeiten

Aufzeichnung der Nutzungsdaten
- Accounting
- Leistungsmanagement

Prüfung von Inhalten
- Angefragte URLs
- Empfangene Daten

Web-Proxies
Implementierung der Backends

Datenbank

- Vorzuhaltende Information
  - Autorisierung: Name/Passwort
  - Autorisierung: Welche Bereiche dürfen erreicht werden?
- Zugriffsprotokolle
  - ODBC für SQL-Datenbank
  - RADIUS (Remote Access and DialIn User Service
  - LDAP (Lightwait Directory Access Protocol)

Logging

- Logdatei aus Performanz-Gründen (KEINE DB)

Weitere Dienste

- Spezielle Protokolle

Typisches Nutzerverhalten 2001 bis 2004: Ziele aus großen IT-Infrastrukturen

2004

Top Websites Yesterday:

destination	request	%	Byte	%	hit-%
*.ebay.de	3032	2.44	99970K	13.43	0.66
*.uni-kl.de	143	0.11	91385968	11.99	0.00
*.	133	0.11	27819751	3.65	97.74
*.ebaystatic.com	25216	20.26	24799285	3.25	66.86
*.hp.com	66	0.05	22314733	2.93	34.85
*.ebayimg.com	2382	1.91	21111267	2.77	25.65
*.berkeley.edu	114	0.09	20421144	2.68	0.00
*.smc.com	349	0.28	17355549	2.28	59.03
*.comdirect.de	4300	3.45	17172862	2.25	1.00
*.t-online.de	3260	2.62	13278385	1.74	32.82
<error>	6481	5.21	12263399	1.61	13.56
*.gmx.net	1899	1.53	11962863	1.57	66.72
*.praline.de	1325	1.06	10141176	1.33	50.79
*.ebay.com	3103	2.49	8927071	1.17	58.07
*.mobile.de	978	0.79	8708235	1.14	35.99
*.web.de	1539	1.24	8012137	1.05	15.85
*.	1355	1.09	6998993	0.92	79.63
*.sportbilder.de	297	0.24	6690092	0.88	43.77

2001

Destination	Request	%	Bytes	%	hit-%
<error>	33683	12.84	39721881	3.41	5.38
*.t-online.de	8765	3.34	19083028	1.64	66.77
*.bild.de	8282	3.16	55732135	4.79	46.61
*.doubleclick.net	5061	1.93	7087115	0.61	9.27
*.web.de	4917	1.87	24289001	2.09	35.67
*.akamai.net	4884	1.86	7574273	0.65	87.24
*.xxxxxxxxxxxxxxxxx.de	4098	1.56	14681246	1.26	68.64
*.sueddeutsche.de	3768	1.44	13554644	1.16	38.96
*.consors.de	3133	1.19	6986643	0.60	69.04
*.boerse.de	2831	1.08	11167450	0.96	72.45
*.lycos.de	2796	1.07	22830497	1.96	65.24
*.microsoft.com	1898	0.72	14627786	1.26	54.74
*.br-online.de	1761	0.67	3545059	0.30	73.25
*.ebay.com	1623	0.62	2222164	0.19	91.44
*.yyyyyyyyyyyyyyy.de	1599	0.61	3464164	0.30	49.47
*.gmx.net	1483	0.57	11860845	1.02	0.20
other: 2691 2nd-level-domains	160337	61.13	854876K	75.20	48.87
Sum	262293	100.00	1136733K	100.00	43.35

Web-Proxies

Ein paar Gedanken zu Optimierungspotentialen....

Top Websites Yesterday:					
destination	request	%	Byte	%	hit-%
*.ebay.de	3032	2.44	99970K	13.43	0.66
*.uni-kl.de	143	0.11	91385968	11.99	0.00
*.	133	0.11	27819751	3.65	97.74
*.ebaystatic.com	25216	20.26	24799285	3.25	66.86
*.hp.com	66	0.05	22314733	2.93	34.85
*.ebayimg.com	2382	1.91	21111267	2.77	25.65
*.berkeley.edu	114	0.09	20421144	2.68	0.00
*.smc.com	349	0.28	17355549	2.28	59.03
*.comdirect.de	4300	3.45	17172862	2.25	1.00
*.t-online.de	3260	2.62	13278385	1.74	32.82
<error>	6481	5.21	12263399	1.61	13.56
*.gmx.net	1899	1.53	11962863	1.57	66.72
*.praline.de	1325	1.06	10141176	1.33	50.79
*.ebay.com	3103	2.49	8927071	1.17	58.07
*.mobile.de	978	0.79	8708235	1.14	35.99
*.web.de	1539	1.24	8012137	1.05	15.85
*.	1355	1.09	6998993	0.92	79.63
*.sportbilder.de	297	0.24	6690092	0.88	43.77

Hit/Miss-Rate:

- Anzahl: ca. 1/3 Treffer
- Größe: ca. 1/4 Treffer
- 2/3 aller Anfragen werden verlangsamt

Nutzung über Tageszeit

- Mittags NICHT weniger :-)

Server-seitige Optimierung der Übertragung?

- Header „LAST-MODIFIED“ mitschicken
- Explizite Informationen zu „EXPIRES“ (z.B. in 10 Minuten)
- Grafiken/Inhalte browser-spezifisch aufbereiten
- Inhalte komprimieren (GZ)

Caching Proxies

Statische Auslegung Caches

Plattenplatzbedarf:

- Statistische Fragen
  - Wie häufig wird auf welche Seiten zugegriffen?
  - Wie schnell veralten welche Seiten?
- Nutzen vs. Verwaltungsaufwand berücksichtigen
- Typische, sinnvolle Größe ??

Welche Schlüsse zieht man aus der Beobachtung:

- 80 % der Seiten im Cache veralten innerhalb eines Tages...
- Bringt ein großer Cache wirklich so viel ..... (nein, aber ....)

Caching Proxies
Dynamische Auslegung Caches

Anzahl Prozessoren

- Anzahl parallel laufender Zugriffe (HTTP 1.1 vs 1.0)
- Wieviel kann ein Prozessor davon abwickeln ?
  - Wie ist das Verhalten bei Überlast ?
  - Toleranz der Benutzer ?
- Entscheidend sind Zusatzdienste:
  - Virenscanning für HTTP/FTP => hohe Prozessorlast
  - Multimedia-Streaming

Weitere limitierende Faktoren

- Zugangsbandbreiten eingehend
- Zugangsbandbreiten abgehend
- Zugriffscharakteristik für Hintergrundspeicher

Virenscanning
Implementierung

Scanner im HTTP-Strom

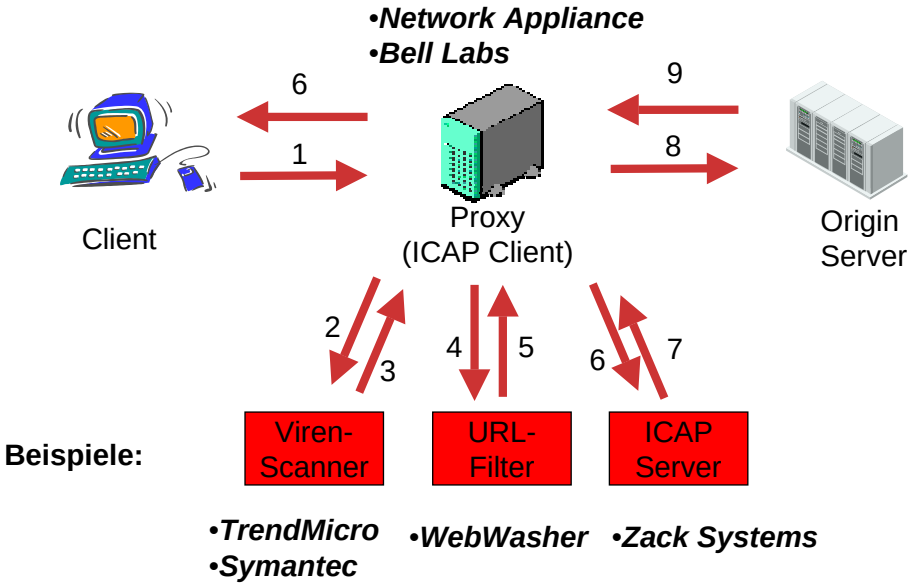
- unterbricht Zugriff bei
  - Erkennung Virenmuster
  - Zugriff auf bestimmte Seiten
- Problem:
  - Gesamter Strom muß gefiltert werden (auch HTML)

besser wäre:
Proxy „präsentiert“ dem Scanner nur Wichtiges....

Exkurs
Internet Content Adaptation Protocol

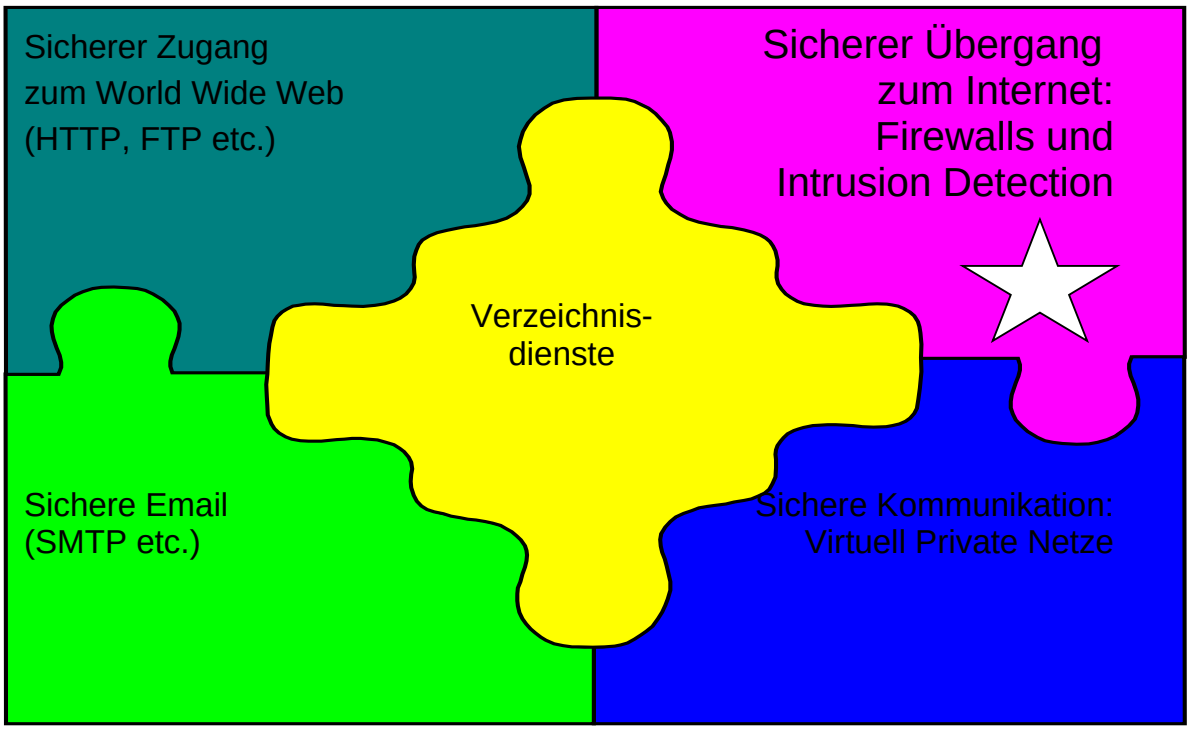
ICAP-Server nur für „bestimmten“ Content registriert

- somit: Kein Durchschleusen des gesamten HTTP-Stroms



Vgl. http://www.i-cap.org/docs/icap_whitepaper_v1-01.pdf

Sicherheitsdienste für große Firmen
=> Teil 2: Firewalls



Firewalls
Einsatzzweck

*„A Firewall helps you to keep **unauthorized** users from accessing your network **resources**. „*

- Zugriffsrechteverwaltung für Kommunikationsbeziehungen (*Access Control Policy)*

Grundprinzip:

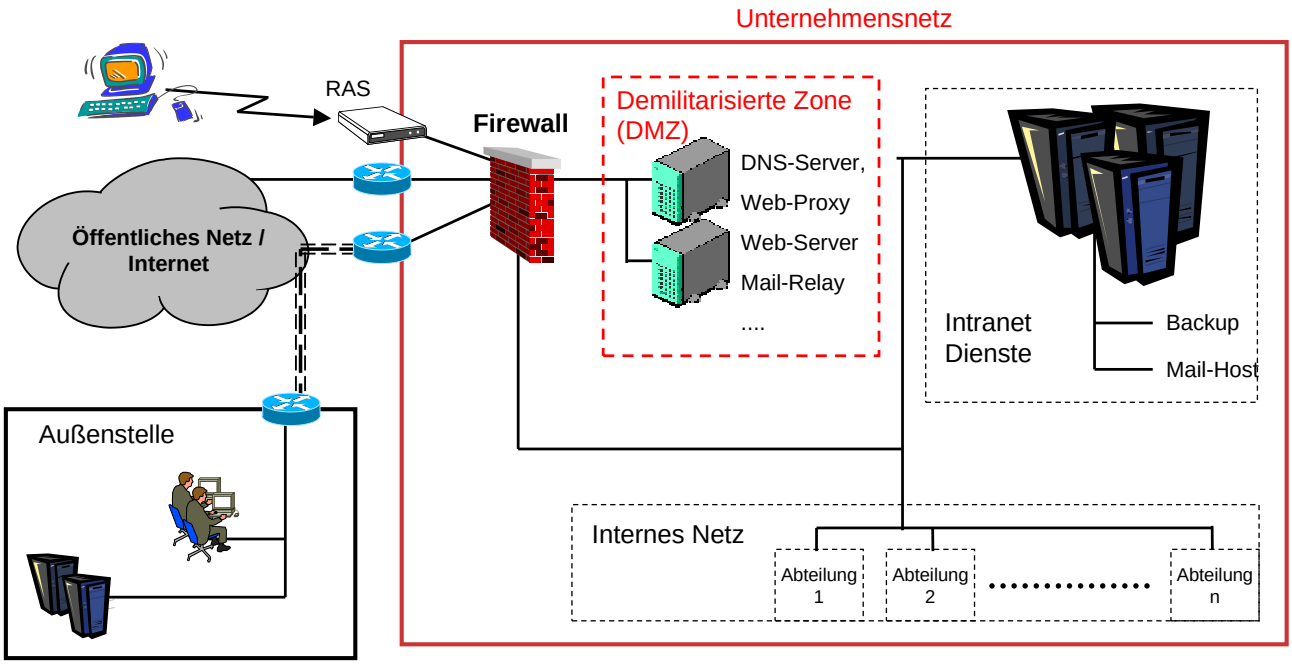
- Alles ist (zunächst) prinzipiell gesperrt.
- Kommunikationsbeziehungen werden einzeln erlaubt.

=> ALLE Bereiche des Netzzugangs werden tangiert!

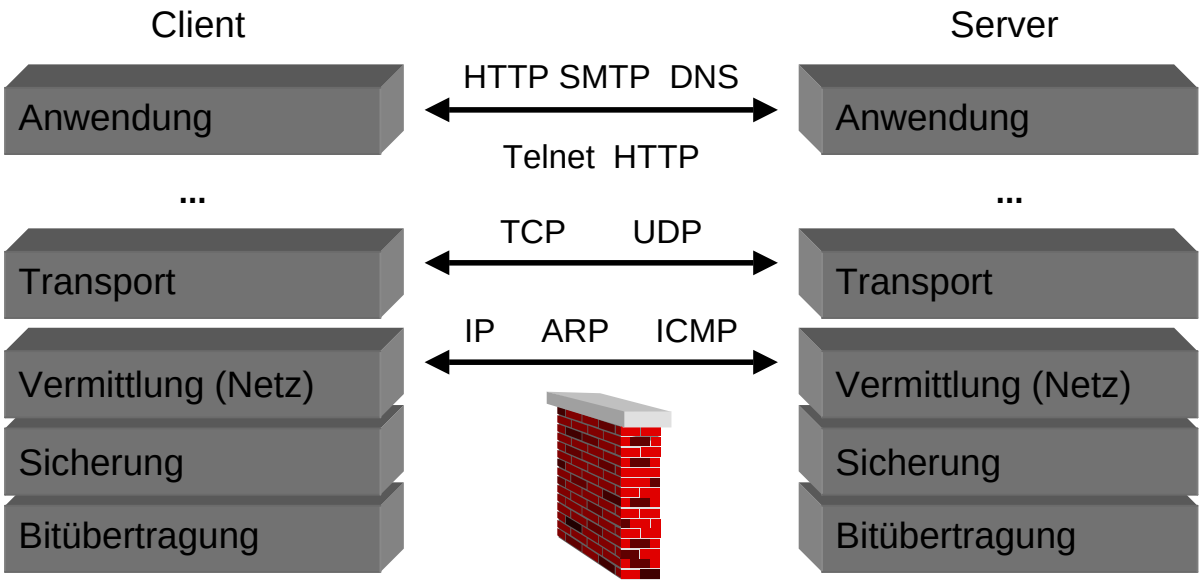
Festlegung der Konfiguration in großen IT-Infrastrukturen

- Iterativer Prozeß in Abstimmung mit vielen Beteiligten
- Unterliegt ständigem „Change Management“
- Umgehung durch Tunnelling vermeiden

Internet-Übergang
Architektur



Firewalls
Erster Überblick



Firewalls
Überblick Typen

Packet Filtering

- Untersuchung des Paket-Headers
- Keine Untersuchung über mehrere Pakete (*Stateless Inspection*)

Stateful Inspection

- Kontext einer Kommunikationsbeziehung wird untersucht
- TCP-Strom, UDP-Request/Response-Paare
- z.B. Email-SMTP: State der Protokollmaschine

Application Level Gateway

- Unterbrechung der Kommunikationsbeziehung
- Eigene Protokollmaschinen für jedes Anwendungsprotokoll

Das wärs für heute ...

Fragen / Diskussion

Verbesserungsvorschläge

Die Folien von heute kommen auf die Web-Seite der Vorlesung (zusammen mit einigen URLs).

Einen schönen Abend !!!