Ausbau und Nutzung des DFN–E–Mail–Verzeichnisses

AMBIX – Statusbericht

X.500–AK der 28. DFN–Betriebstagung
18. März 1998

Dr. Ralf Schneider, DFN–Projekt AMBIX,
Zentrum für Datenverarbeitung, Universität Tübingen

Inhalt

Datenverteilung

•

Index–DSA

•

Statistik

•

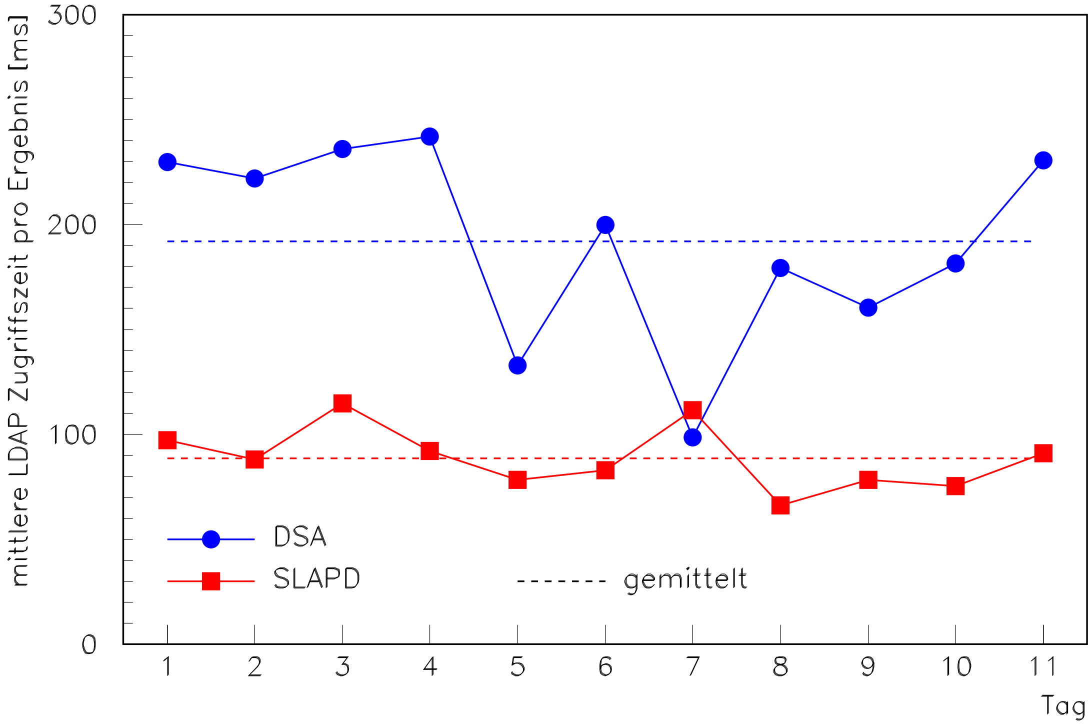
WWW–Proxies

•

LDAP–Bridge

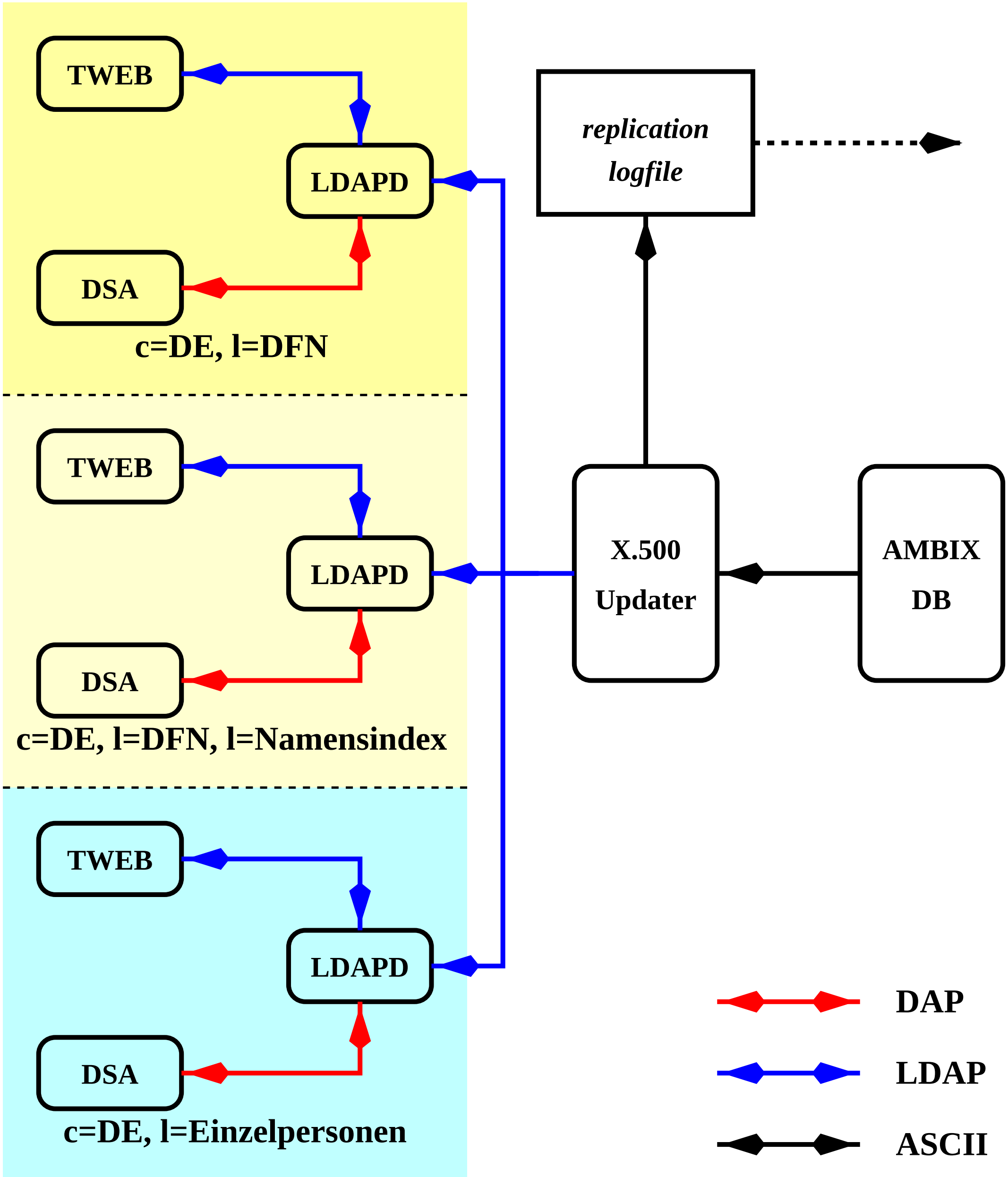
Datenverteilung

- Mai 1997: Test Datenhaltung mittels SLAPD *(Standalone LDAP Daemon)*
- Konvertierung DSA–Datenbestand für SLAPD–Technologie
- Datenbestand auf DSA und $2\times$ SLAPD, je 33% der Anfragen
- Zugriffszeiten

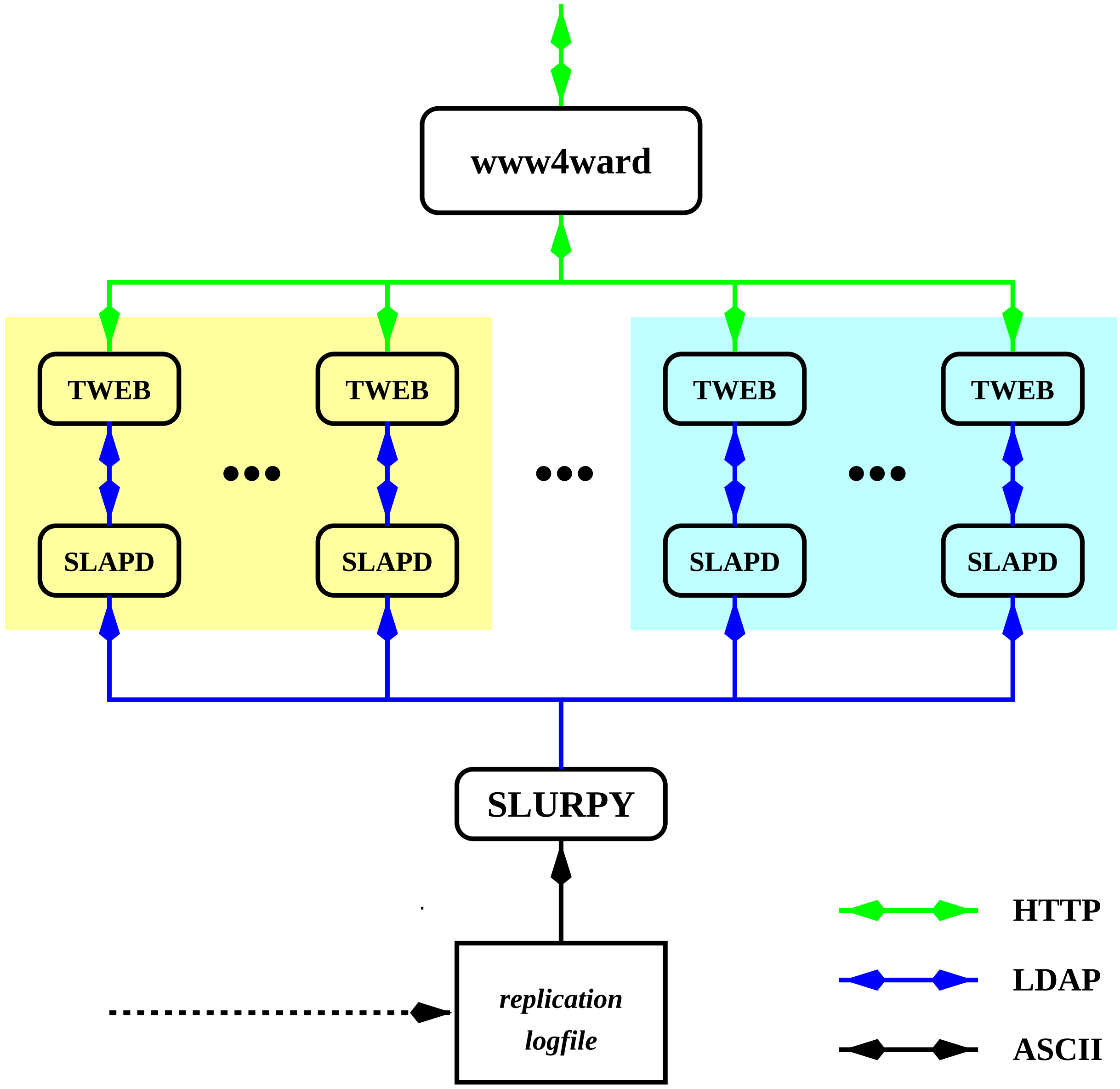


- DSA: 192 ms; SLAPD: 88 ms
- Produktionsbetrieb seit Woche 03/98

Referenzdatenbestand (DSAs)



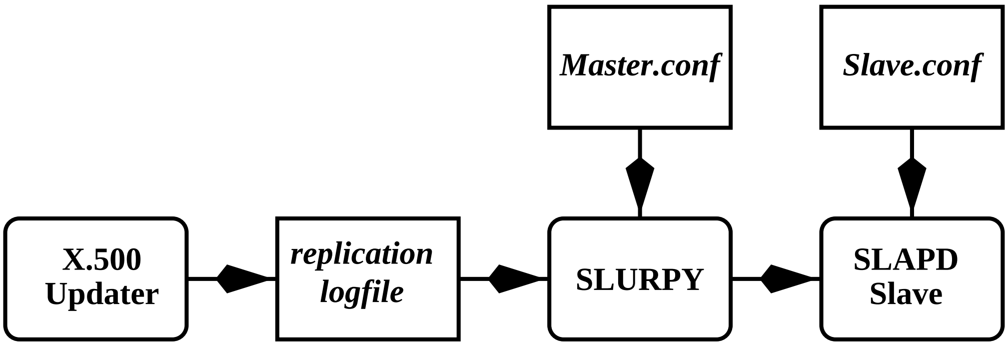
Benutzeranfragen (SLAPD & TWEB)



replication logfile

- Verbindung zwischen Referenzdatenbestand und benutzerzugänglichem Datenbestand
- Änderungen eines Datensatzes in der internen Datenbank (AMBIX-DB):
  - X.500Updater:
    - führt Datenänderung im zuständigen DSA aus
    - protokolliert Änderungen im *replication logfile*
  - SLURPY:
    - liest *replication logfile*
    - veranlaßt Datenänderung in den zugeordneten SLAPDs

SLURPY



- Betriebssystemunabhängige Neuimplementation (Tübingen) des SLURPD *(Standalone LDAP Update and Replication Daemon)* der University of Michigan
- *Master.conf:*

```
replogfile   /LDAP33/work/replication.data

replica      host=ambixhp1.uni-tuebingen.de:10310
             binddn="cn=DS-Manager, l=Einzelpersonen, c=DE"
             bindmethod=simple credentials=<password>
replica      host=ambixhp2.uni-tuebingen.de:20310
             binddn="cn=DS-Manager, l=Einzelpersonen, c=DE"
             bindmethod=simple credentials=<password>
```

Format des replication logfiles

- allgemeiner Header für jeden Datensatz

```
replica: host1:port1
...
replica: hostn:portn
time: 887731012
dn: cn="Vor Nachname",ou=Suborg,o=Org,l=DFN,c=DE
changetype: <[add|delete|modify|modrdn]>
```

- `changetype = delete` ⇒ fertig
- `changetype = modify`

```
<[add|delete|replace]>: <attributetype>
<attributetype>: <value1>
...
<attributetype>: <valuen>
-
```

- `changetype = modrdn`

```
newrdn: <newrdn>
deleteoldrdn: 0 | 1
```

- `changetype = add`

```
<attributetype1: <value1>
...
<attributetype1: <valuen>
...
<attributetypen: <value1>
...
<attributetypen: <valuen>
```

Datenverteilung in Worten

- AMBIX.DSA: c=DE, l=DFN
  - ◇ Datenbestand geteilt, je Bereich 4 SLAPD/TWEBs:
  - ◇ `o=\"?[A-T][^=]+,␣*l=\"?DFN␣|`
    `o=\"?Universitaet␣[A-F,a-z][^=]+,␣*l=\"?DFN`
  - ◇ `o=\"?[U-Z][^=]+,␣*l=\"?DFN`
- AMBIX3.DSA: c=DE, l=Einzelpersonen, (c=DE, l=Schulen)
  - ◇ 2 SLAPD/TWEBs, (2 SLAPD/TWEBs)
- AMBIX-IDX.DSA: c=DE, l=DFN, l=Namensindex
  - ◇ 2 SLAPD/TWEBs
- Verteilung der HTTP–Anfragen mittels www4ward
- zentrale Replikations–Konfigurationsdatei
- Replikationsbibliothek

www4ward

- Verteilung von Anfragen auf mehrere TWEBs
- HTTP–Redirect
- Vorgabe der Verteilung wird aus Datei gelesen (Laufzeit):

```
http://to.host1   port1   fix   from.host1[:from.host2 ...]
http://to.host2   port2   100-n1-n2  ["regular expression"]
http://to.host3   port3      100-n1  ["regular expression"]
http://to.host4   port4         100  ["regular expression"]
```

- Bsp:

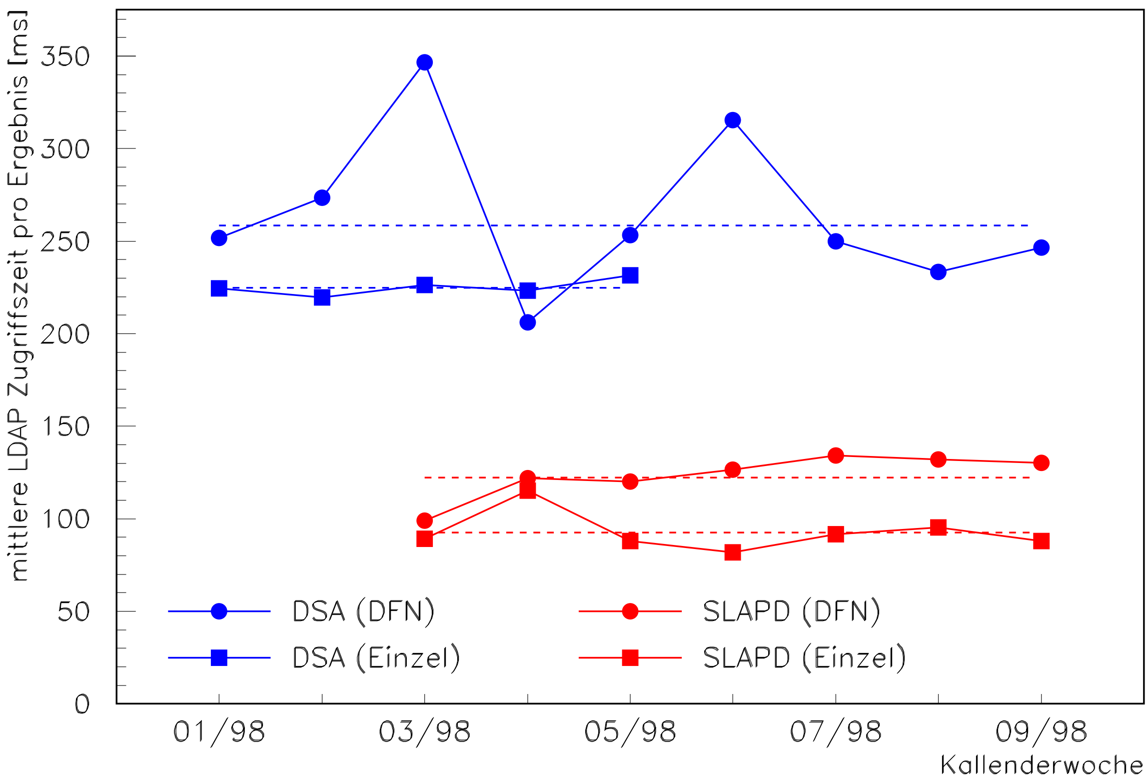
```
http://ambixhp1.uni-tuebingen.de 10101 fix x500.uni-tuebingen.de
http://ambixhp1.uni-tuebingen.de 10201  50 "o=\"?[U-Z][a-z][^=]+, *l=\"?DFN"
http://ambixhp2.uni-tuebingen.de 10201 100 "o=\"?[U-Z][a-z][^=]+, *l=\"?DFN"
http://ambixhp1.uni-tuebingen.de 10301 100 "l=Einzelpersonen"
http://ambixhp2.uni-tuebingen.de 10001 100
```

- Vorteil: weiterhin eine zentrale Basisadresse
- Datei kann jederzeit angepaßt werden

Replikation–Konfigurationsdatei

- Verteilung Datenbestand (DSA)
- Replikation des Datenbestands (SLAPD)
- allgemeine Informationen für die zu startenden Programme
- Allgemeiner Header
  - Pfade, Login IDs, Programmnamen, . . .
  - #*Master–Server* (DSAs)
- je *Master–Server*
  - Pfade, Programmnamen, . . .
  - Datenbereich *(regular expression)*
  - #*Slave–Server* mit Replikat des Datenbereichs (SLAPDs)
  - je *Slave–Server*
    - Pfade, Programmnamen, . . .

Zugriffszeiten



	l=DFN	l=Einzelpersonen	TEST
DSA:	258 ms	224 ms	192 ms
SLAPD:	122 ms	93 ms	88 ms

Organisationsunahängige Suche

- Finger–Service $\Rightarrow$ c=DE, l=Einzelpersonen

  `finger [vor.]nachname@ambix.uni-tuebingen.de`

  Untersuchung Akzeptanz und Performance

- Index (Replikationsbibliothek)
  - offene Definition, beliebig erweiterbar
  - SEARCH_OBJECTCLASS: Objektklasse die im Eintrag enthalten sein muß, damit dieser im Index aufgenommen wird
  - ATTRIBUTE: Attributname, nach dem indiziert wird

Namensindex

```
                    c=DE
                     |
                   l=DFN
                     |
              l=Namensindex
            /        |        \
        l=A        l=M        l=Z
                  /   \
         l=Maier       l=Mueller
                      /          \
    cn=Anton Mueller              cn=Hans Mueller
```

Eintrag im Namensindex

```
dn: cn="Hans Mueller",l=Mueller,l=M,l=Namensindex,l=DFN,c=DE
(...)
labeleduri: http://ambix.uni-tuebingen.de:8889/Mcn=Hans%20Mueller,
    o=AMBIX,l=DFN,c=DE AMBIX
labeleduri: http://ambix.uni-tuebingen.de:8889/Mcn=Hans%20Mueller,
    l=Tuebingen,l=Baden%20Wuerttemberg,l=Einzelpersonen,c=DE
    Einzelpersonen
labeleduri: http://ambix.uni-tuebingen.de:8889/Mcn=Hans%20Mueller,
    ou=Institut%20fuer%20Theoretische%20Physik%20A,ou=Physik,
    o=RWTH%20Aachen,l=DFN,c=DE RWTH Aachen
```

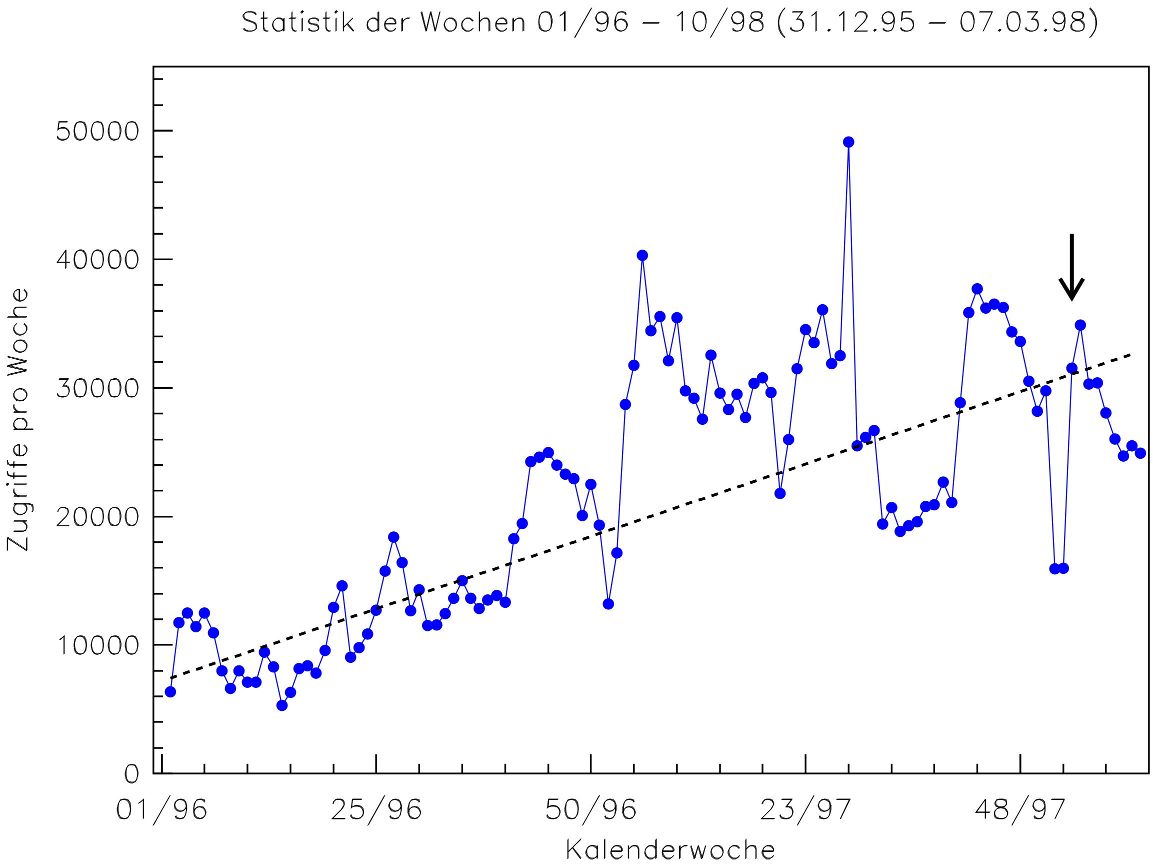
Statistik zum DFN–E–Mail–Verzeichnis*

- Bestand im X.500:
  - l=DFN
    - 116 Organisationen (davon 37 aktiv)
    - 4.366 Strukturdatensätze
    - bis zu 4 Hierarchieebenen
    - Verweis auf 25 Organisationen, die eigenen Datenbestand pflegen
    - 32.130 Personeneinträge
    - $\simeq$ 60% explizite Zustimmung
  - l=Einzelpersonen
    - 16 Bundesländer, 412 KFZ–Gebiete
    - 318 Personeneinträge
- ca. 23 Selbsteinträger pro Tag†
- Kontakt zu weiteren 5 Organisationen (aktive Beteiligung)

*Stand 14. März 1998

†gemittelt Januar/Februar 1998

Zugriffszahlen



- seit 1. Jan. 1996 (Stand 8. März 1998):
  - 2.795.845 Zugriffe, davon 284.574 keine Personendaten
  - 6.888 verschiedene Domains
- Woche 10/1998 (01.03 – 07.03):
  - 26.859 Zugriffe, davon 1.927 keine Personendaten
  - 685 verschiedene Domains

WWW–Proxies

- Datenschutzexpertise
- $\leq 17\%$ der Zugriffe über Proxies
- Problem Zugriffsbeschränkung $\Rightarrow$ keine Personendaten
- TWEB erlaubt Öffnung für „unkritische“ Proxies
- Bestätigung

  `Hiermit bestätige ich, daß auf den WWW-Proxy <Name> (IP-Adresse, alias <AliasName>) nur von den Domains *.<Domain>.de (IP <Nummer>.<Nummer>.*.* aus zugegriffen werden kann, und daß eine Änderung dieser Konfiguration dem Projekt AMBIX sofort mitgeteilt wird.`

- erlaubte Proxies (7):
  $2\times$ Humboldt–Universität zu Berlin, $1\times$ Universität Halle,
  $1\times$ Universität Kiel, $2\times$ Universität Roststock,
  $1\times$ Universität Tübingen
- DFN–Projekt „Aufbau eines nationalen Cache–Service“
  Cache–Verbund nicht „unkritisch“

TWEB: Abwehr Hacker

- REP (*robots exclusion protocol)*
- temporäre Sperrung von IP–Bereichen

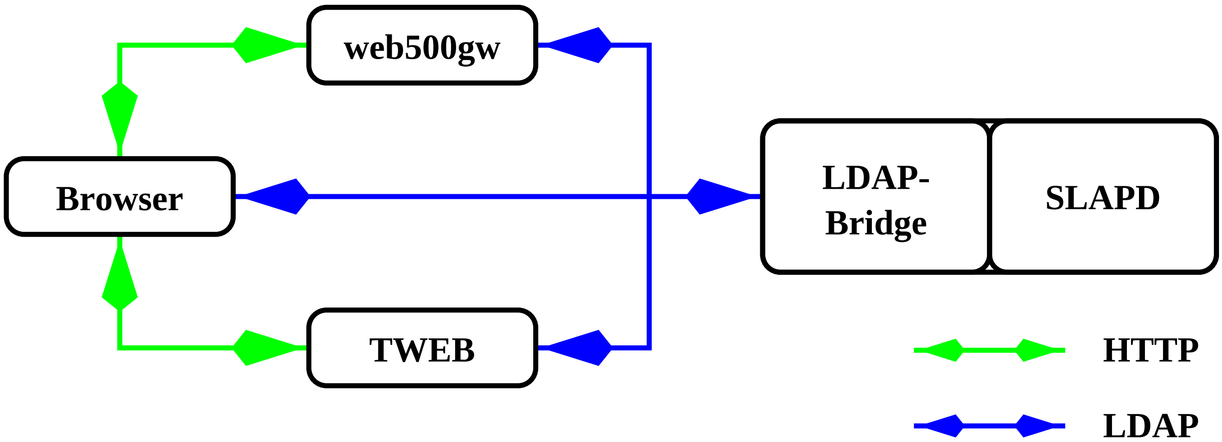
```
#COMREFUSE   TMIN TMAX MAX_ACCEPT SUSPEND_CYCLE STAT_CYCLE STAT_FILE
COMREFUSE    100  200     40      12  43200  /LDAP33/log/hackstat.log
```

`TMIN:`	minimaler Zeitzyklus in Sekunden
`TMAX:`	maximaler Zeitzyklus in Sekunden
`MAX_ACCEPT:`	maximale Anzahl akzeptierter Zugriffe
`SUSPEND_CYCLE:`	Dauer der Sperrung in Zeitzyklen
`STAT_CYCLE:`	Zeitspanne für Datei in Sekunden
`STAT_FILE:`	Datei

- IP–Adresse & Maske wählt 13 Bit aus $\Rightarrow$ 8192 IP–Bereiche
- Januar 1998: häufige Sperrung von IP–Bereichen
  - Anlaß: verschiedene Rechner derselben Domain (56 Stück)
  - Rechner A gesperrt $\Rightarrow$ Rechner B ...
  - Reaktion: Sperrung aller Rechner
  - Aufklärung
- Erweiterung TWEB *Regular Expressions*

LDAP–Bridge

- Wunsch: direkter LDAP–Zugriff auf AMBIX–Verzeichnis über WWW–Browsers
- Problem: Zugriffsbeschränkung
- Lösung: LDAP–Bridge oder Patch für SLAPD? NEIN!



Ausblick

?