X.500-Index-DSAs

Technologie, Standards und ihre Anwendungsmöglichkeiten

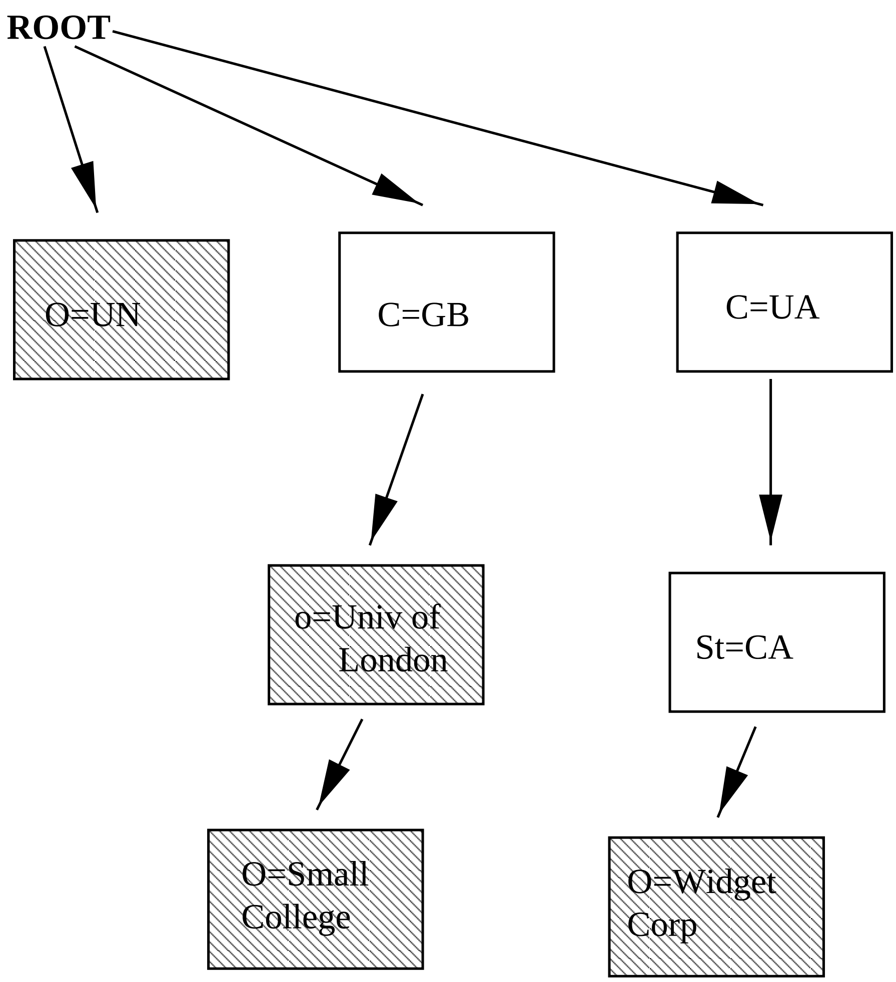
Vortrag gehalten im X.500-AK der 27. DFN-Betriebstagung
am 28. Oktober 1997

Karl-Peter Gietz, DFN-Projekt AMBIX,
Zentrum für Datenverarbeitung, Universität Tübingen

Das Problem

- X.500 ist eine weltweit auf Datenserver (DSA) verteilte Datenbank
- Eine weltweite Suche muß alle DSAs auf der Welt befragen
- Einzelne DSAs können hierbei ausfallen
- Die Suche kann nicht ab einer bestimmten Hierarchietiefe abgebrochen werden
- $\Longrightarrow$ Das alles kann dauern
- Die gängige Replikationspraxis (RFC 1276 bzw. X.525) kann Search-Problem nicht lösen

Verschiedene Ebenen für gleiche Objektarten



X.525 Replikation (1933)

- Einzeleinträge oder Teilbereiche des DIT
- Ausgewählte Attribute
- Nur DN auch möglich
- Objektklassen und Zugriffskontrollmechanismen automatisch repliziert
- Ausgewählte Einträge (ausgefilterte Einträge werden durch sog. Glue-Einträge ersetzt)
- Inkrementelles Replizieren
- Nach abgeschlossenen und eingetragenen Abkommen automatische Replikation

RFC 1588

(J.Postel; C.Andersen: White Pages Meeting Report, Feb. 1994)

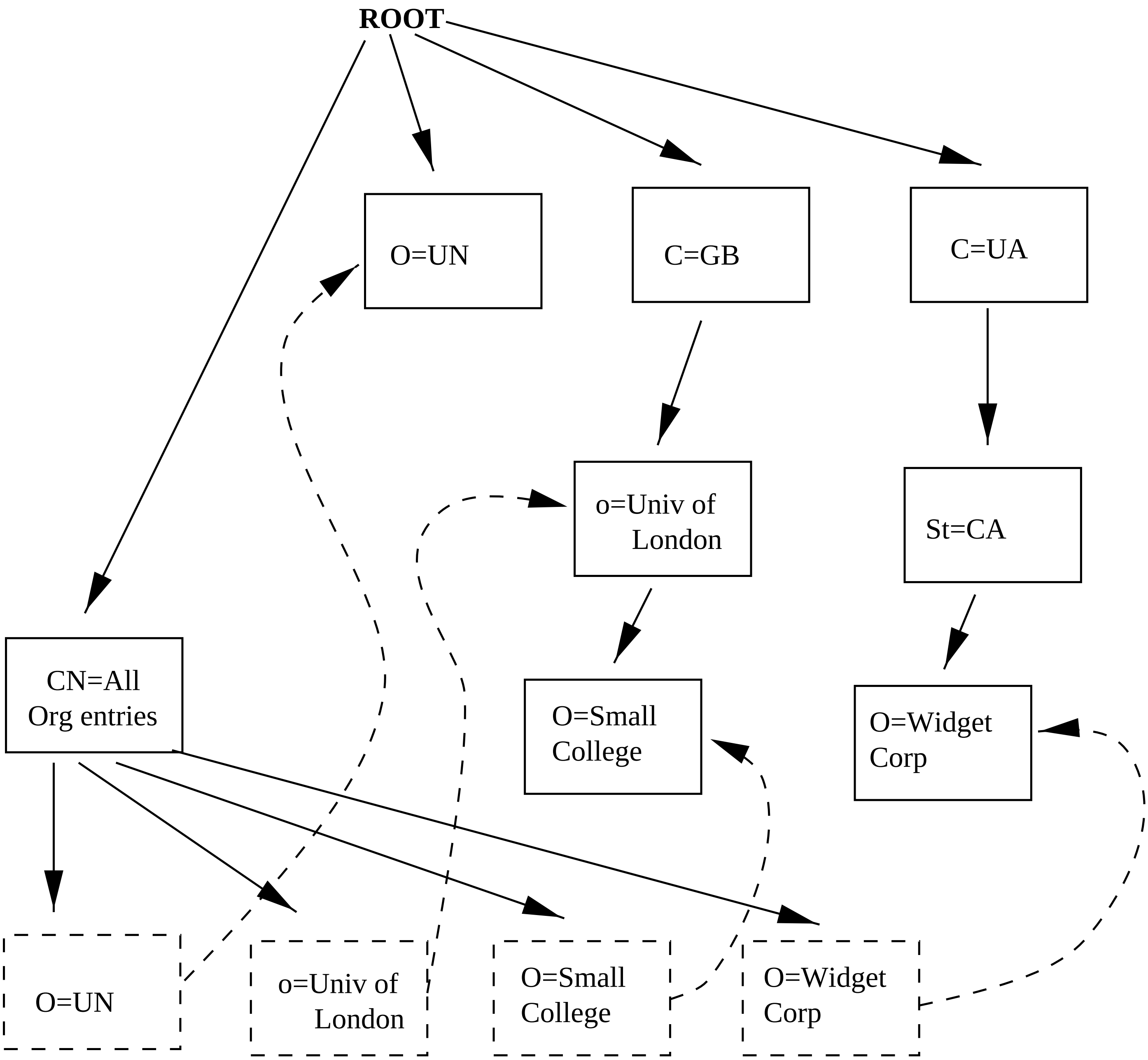
- Report der White Pages IETF, Nov. 1993
- Fragestellung: Wie erhalten wir einen WP-Dienst für das gesamte Internet
- Resultat: 4 Empfehlungen
  - Alle Protokolle (X.500, whois++, Netfind, finger, ...) sollen auf alle Daten zugreifen können durch ein standardisiertes „interoperational protokoll“
  - Entwicklung von Index-Servern mit verschiedenen Methoden des Updates und mit verschiedenen Abfragemethoden
  - Zentrales Management für X.500 (???)
  - Sicherheitszertifikate durch Speicherung im X.500 unterstützen
- X.500 beste Technologie, aber zu komplex

Barkers Index DSAs

(Paul Barker: X.500 Index DSAs, Paper vom 10.3.1995)

- X.500 wie andere Internetdienste indizieren (vgl. FTP ⇒ Archie; Gopher ⇒ Veronica)
- Aliase als Indizierung ungeeignet:
  - müssen per Hand gepflegt werden
  - keine sichere Vollständigkeit
  - Benutzer bzw. DUAs müssen speziellen Ausgangsknoten kennen
  - machen Suche langsam
- Replikation von selektierten Einträgen in einem zentralen Index-DSA
- Für verschiedene Indices jeweils ein DSA.
- DIT-Struktur bleibt im Index-DSA enthalten, Einträge sind unterschiedlich
- Index-DSA müssen „special purpose, high performance“ DSAs sein
- Restrukturierung des DIT in den Index-DSAs ???
- Nicht vollständig durchdacht, viel „future work“

Index durch Aliase

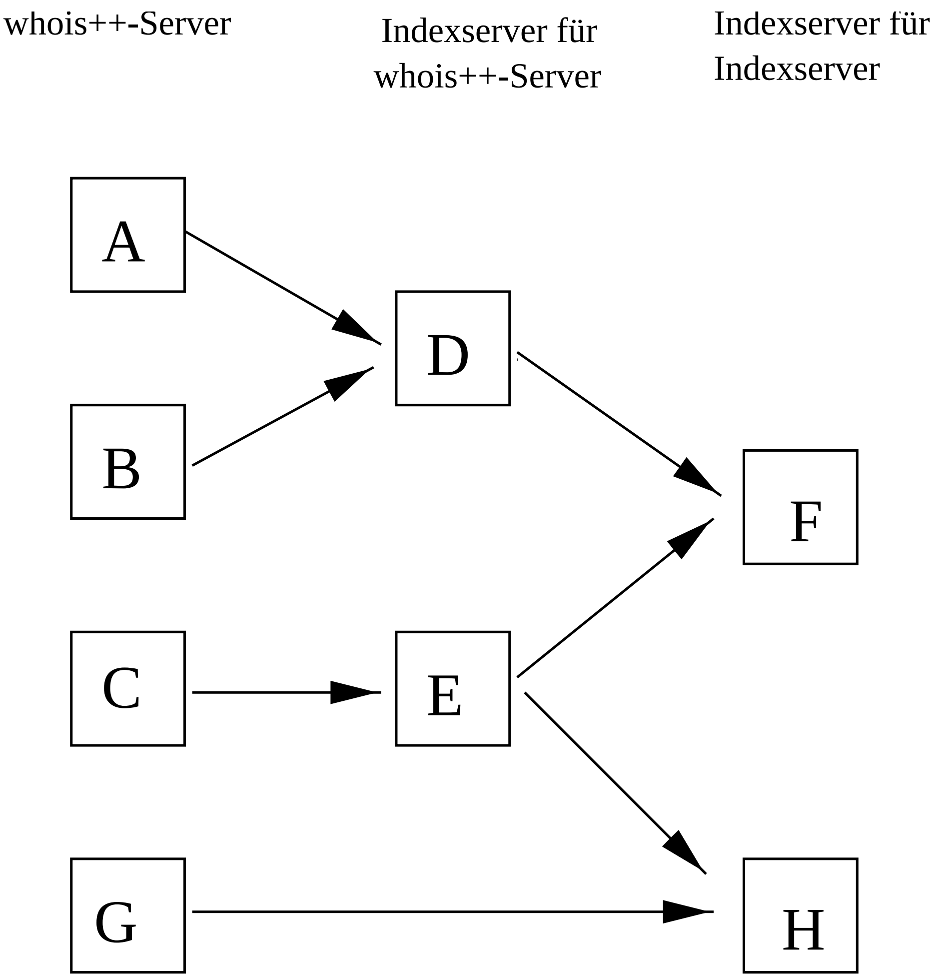


Whois++

(RFC 1835: P. Deutsch et.al.: Architecture of the WHOIS++ Service, Aug 1995)
(RFC 1913: C. Weidner et.al.: Architecture of the WHOIS++ Index Service, Feb. 1996)
(RFC 1914:P. Faltstrom et.al.: How to interact with a WHOIS++ Mesh, Feb. 1996)

- Einfaches Internetprotokoll für WP-Dienst
- Stammt von Whois (RFC 954, Okt. 1995) ab
- Datenmodell: Serverhandle, Datensatzhandle, Template, Attribut, Wert
- Client/Server-Protokoll: Verbindung öffnen, Abfrage schicken, Antwort erhalten, Verbindung schliessen.
- Antwort kann ein Refferal sein.
- Indexinformation via Centroide
- Hierarchische Index-Server-Struktur

Indexhierarchie in whois++



Beispiel eines Centroids

An example of a centroid is as follows: if a whois++ server contained exactly three records, as follows:

```
Record 1                          Record 2
Template: Person                  Template: Person
First-Name: John                  First-Name: Joe
Last-Name: Smith                  Last-Name: Smith
Favourite-Drink: Labatt Beer      Favourite-Drink: Molson Beer

Record 3
Template: Domain
Domain-Name: foo.edu
Contact-Name: Mike Foobar
```

the centroid for this server would be

```
Template:        Person
First-Name:      Joe
                 John
Last-Name:       Smith
Favourite-Drink:Beer
                 Labatt
                 Molson
Template:        Domain
Domain-Name:     foo.edu
Contact-Name:    Mike
                 Foobar
```

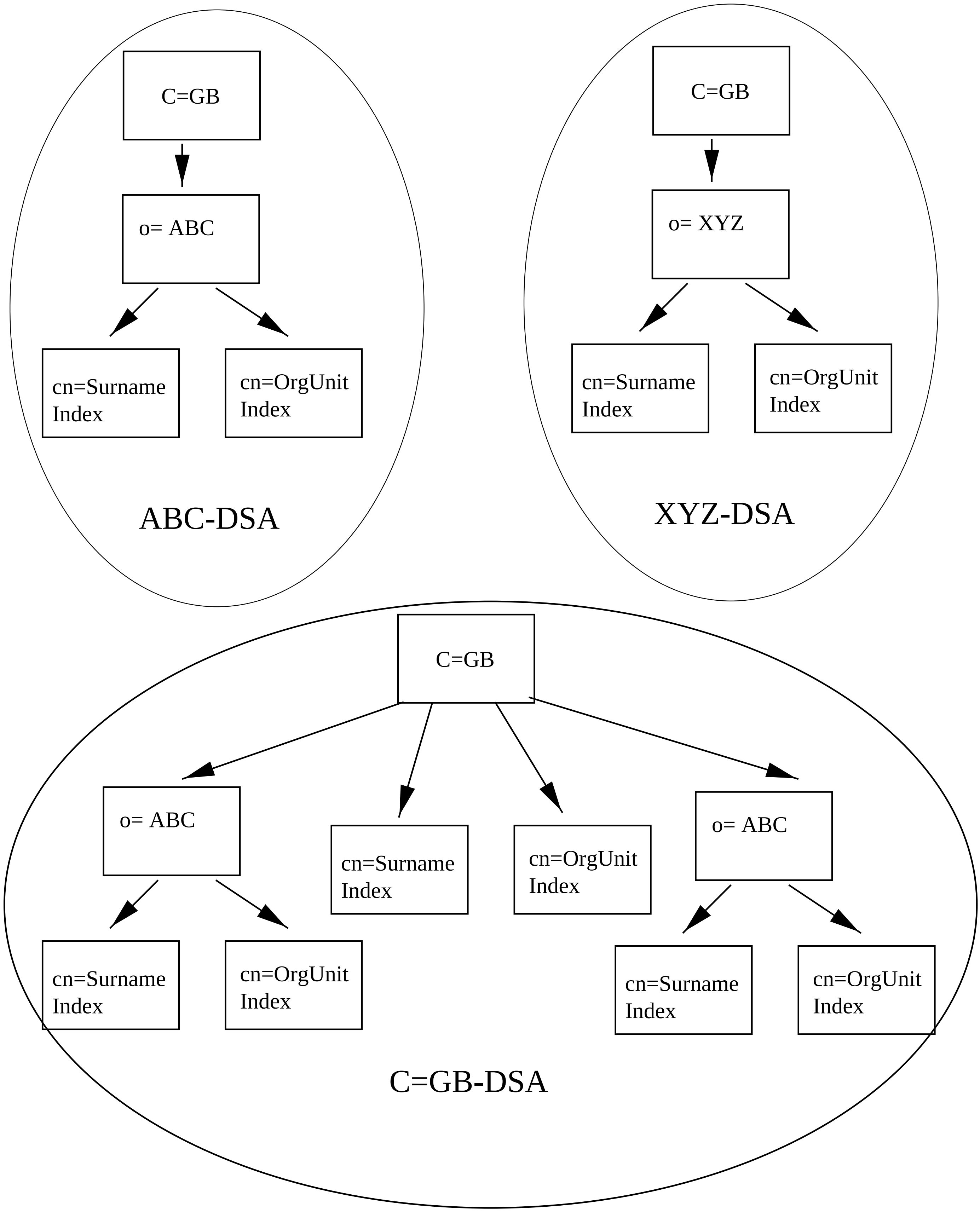
An index server would then collect this centroid for this server as forward knowledge.

Chadwicks IndeX.500

(David Chadwick: IndeX.500, Dante in Print #19, Mai 1996)

- Prinzip von whois++ auf X.500 übertragen
- Ein Centroid-DSA hält Access-Informationen von Server-DSAs und eine Liste der Attributwerte, die diese enthalten
- Objectclass index
- Eine Suche beim Centroid-DSA wird mit einem Referral beantwortet
- Die Centroide werden im Server-DSA gebildet und abgelegt
- Der Centroid-DSA repliziert oder sucht und kopiert die Daten
- Die Centroide werden im Centroid-DSA nicht zusammengefügt, da:
  - entweder die Attributsyntax komplex wird (Zusammenhalten von DSA- und Atributinformation)
  - oder eine Hierarchie von Centroid-DSAs aufgebaut werden muß die jeweils weiterverweisen
- Indexvorschläge:
  - Alle Organisationen weltweit
  - Alle Nachnamen, Länderweit
  - Alle Unterorganisationen Länderweit

Index-DSAs nach Chadwick



Meulemans Centroid500

(Nils Meulemans: Centroid500, an X.500 Indexing System, Paper 30.12.96)

- Whois++-Centroid-Technik
- Definition von Objektklassen für personen-, ou- und o-Centroid
- Verweis durch seeAlso-Attribut
- Centroide werden eine Ebene über ihren Gültigkeitsbereich gespeichert
- Problem wenn zwischen DSA-Grenzen
- Implementierung: 2 Programme
  - Centroid Builder: geht rekursiv durch den DIT und baut die Centroide auf
  - Centroid Search Engine: Führt onelevel-Searches in den Centroiden durch und verfolgt die seeAlso-Verweise

Common Indexing Protocol

(Im Augenblick 6 aktuelle IETF-Drafts)

- Allgemeines Protokoll für verschiedene WP-Dienste (X.500, LDAP, whois++, PH, ..)
- definiert:
  - Abfrageinterface mit Referral-Funktion
  - Datenformat
    - Anstelle des Centroid steht das Tagged Index Object
    - Relative Handle Identifier
    - DN mit aufgenommen
  - Verteilungsmechanismen via MIME
- CIP wird von Netsolutions und Microsoft unterstützt und mitdefiniert

Beispiel eines Tagged Index Object

4 Einträge im LDIF Format

```
dn: cn=Barbara Jensen, ou=Product Development, o=Ace Industry, c=US
objectclass: top
objectclass: person
objectclass: organizationalPerson
cn: Barbara Jensen
cn: Barbara J Jensen
cn: Babs Jensen
sn: Jensen
uid: bjensen
telephonenumber: +1 408 555 1212
description: A big sailing fan.
dn: cn=Bjorn Jensen, ou=Accounting, o=Ace Industry, c=US
objectclass: top
objectclass: person
objectclass: organizationalPerson
cn: Bjorn Jensen
sn: Jensen
telephonenumber: +1 408 555 1212
dn: cn=Gern Jensen, ou=Product Testing, o=Ace Industry, c=US
objectclass: top
objectclass: person
objectclass: organizationalPerson
cn: Gern Jensen
cn: Gern O Jensen
sn: Jensen
uid: gernj
telephonenumber: +1 408 555 1212
dn: cn=Horatio Jensen, ou=Product Testing, o=Ace Industry, c=US
...
```

The Tagged Index Object for this example would be:

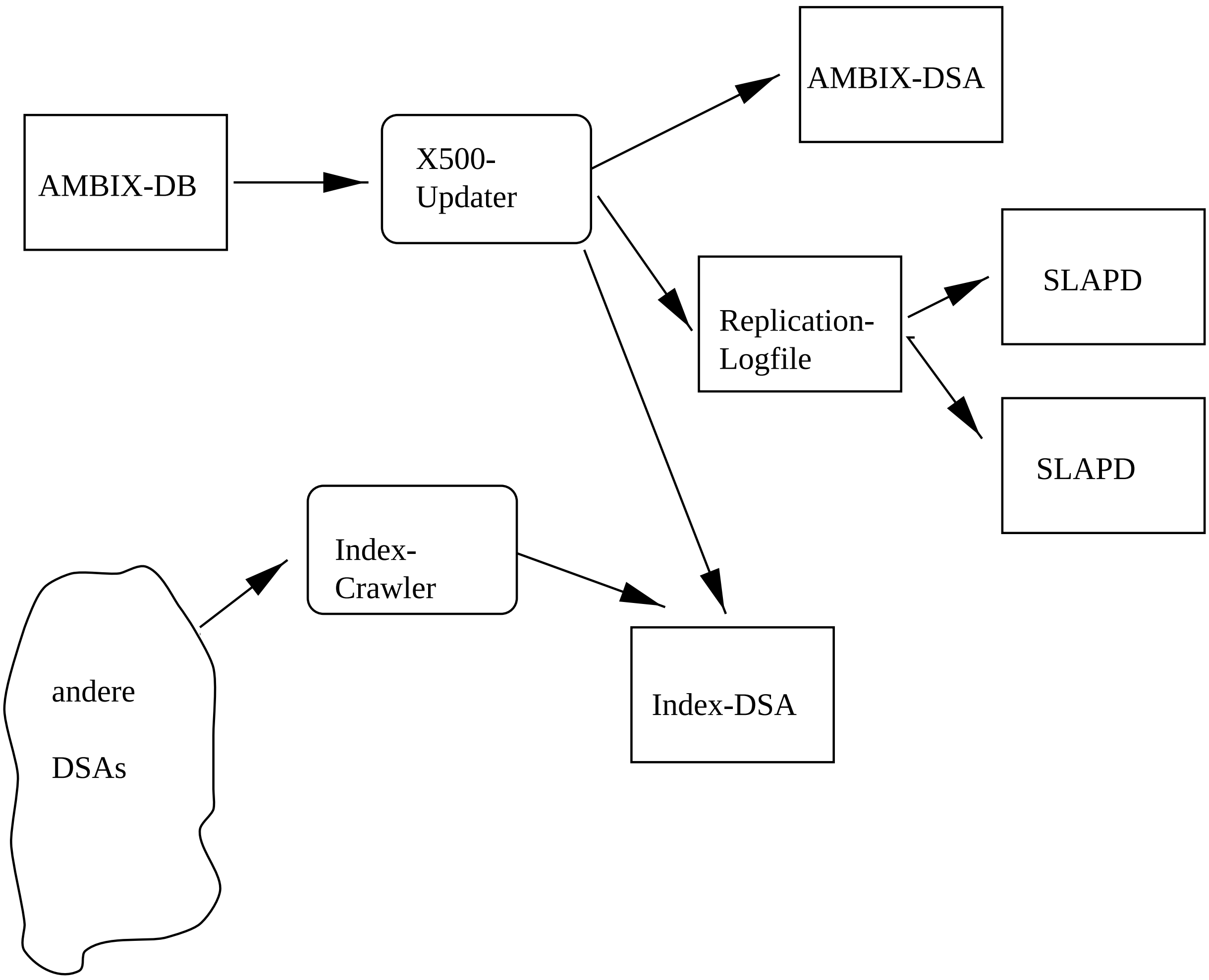
```
version: x-tagged-index-1
updatetype: total
thisupdate: 855938804
BEGIN IO-Schema
dn: FULL
ou: TOKEN
o: TOKEN
c: TOKEN
objectclass: FULL
cn: TOKEN
sn: FULL
uid: FULL
title: TOKEN
END IO-Schema
BEGIN Index-Info
dn: 1/cn=Barbara Jensen,ou=Product Development,o=Ace Industry,c=US
-2/cn=Bjorn Jensen,ou=Accounting,o=Ace Industry,c=US
-3/cn=Gern Jensen,ou=Product Testing,o=Ace Industry,c=US
-4/cn=Horatio Jensen,ou=Product Testing,o=Ace Industry,c=US
ou: 1,3-4/Product
-1/Development
-2/Accounting
-3-4/Testing
o: */Ace
-*/Industry
c: */US
```

Hedbergs Schweden-Index

(http://x500.umu.se/is.en.html)

- Hedberg massgeblich am CIP beteiligt
- keine Dokumentation
- Email-Anfrage erfolglos
- Widersprüchliche Aussagen über Technik, jedenfalls zentraler Index-Server (X.500 oder SLAPD?), offen für andere Protokolle (CIP?)
- Direkter Verweis auf die Einträge

Gesamtindex bei AMBIX



- Index-DSA braucht nicht unbedingt neue Objektklassen
- Verweise auf Daten mittels Gateway-Switches

Mögliche Indices für Deutschland

```
Vorschlag: 4 Index-DSAs

a. inhaltliche Kategorien, beliebig hierarchisierbar, z.B.
       Suchindex
      /    |   \
    ...   ...  Forschungseinrichtungen  ...
               /         |           \
            Unis        FHs         Forschungsabteilungen  ...


b. geographischer Index:
        /       |
     Bayern   BadenWuerttenberg
      /         |
 Muenchen       Stuttgart



c. Personennamen
    /     |     \
   A      B ...  M
                /|\
              ...  Mueller
                   /  |  \
                  ... ... Hans
```

```
d1. Organisationsnamen
     /     |     \
    A      B ...  M
                 /|\
               ...  Max Plank Gesellschaft

 Alternativ hierzu:

d2. Organisationsnamen
      /     |     \
     A      B ...  G
                  /|\
              GWD ... Gesellschaft
                     /    |    \
                    zur  ...   fuer
                               / | \
                      Informatik .. Wissenschaftliche
                                        Datenverarbeitung
```

- Umstrukturierung von C=DE verliert an Brisanz
- Definition von Index-Objekten sinnvoll
- CIP-Konformität