



Z39.50

An Overview

By Janifer Gatenby
Consultant OCLC|PICA

Protocol

- machine to machine
- querying a database
 - web search engines index web pages
 - don't index databases behind web pages
- stateful - "Z association"
 - searches in a session can create a continuity
 - web is stateless
- V1 1984; v2 1992; v3 1995 (+ISO 23950)

Search possibilities > a search engine

- Precise searching
 - e.g. title only
 - Boolean operations, nested etc.
- Results set handling
 - segmentation
 - sort & duplicate removal
 - refine search, no cookies
- Browse

Not like a search engine cont.

- Retrieval options
 - brief or full records
 - can specify content
- Choice of format
- Extended services
 - SDI - saved query and schedule
 - Saved results
 - Item order, and update
- Linked – electronic & physical

Advantages

- Multi-target searching
 - One user interface; one search
 - different platforms (UNIX, NT, IBM etc.)
 - different database systems (relational, network)
 - different database models
- Searching based on abstract concepts
 - Title, i.e. not database columns
- Can combine results from diverse dbases
 - Common record syntax (usually MARC)
- Reuse of results
- Follow on delivery – electronic & physical

Drawbacks

- Complex and difficult to implement
- Too many options
 - Interoperability problems without profiles
- Interoperates with web but.....
 - Not fully web compatible
- Has not achieved widespread acceptance
- Contenders
 - ZING SRW / SRU
 - XML Query / SQL like
 - OAI
 - OpenURL

Implementations

Domain

- Libraries (500+)
- Museums
- Government info
- Scientific dbases
- Thesauri

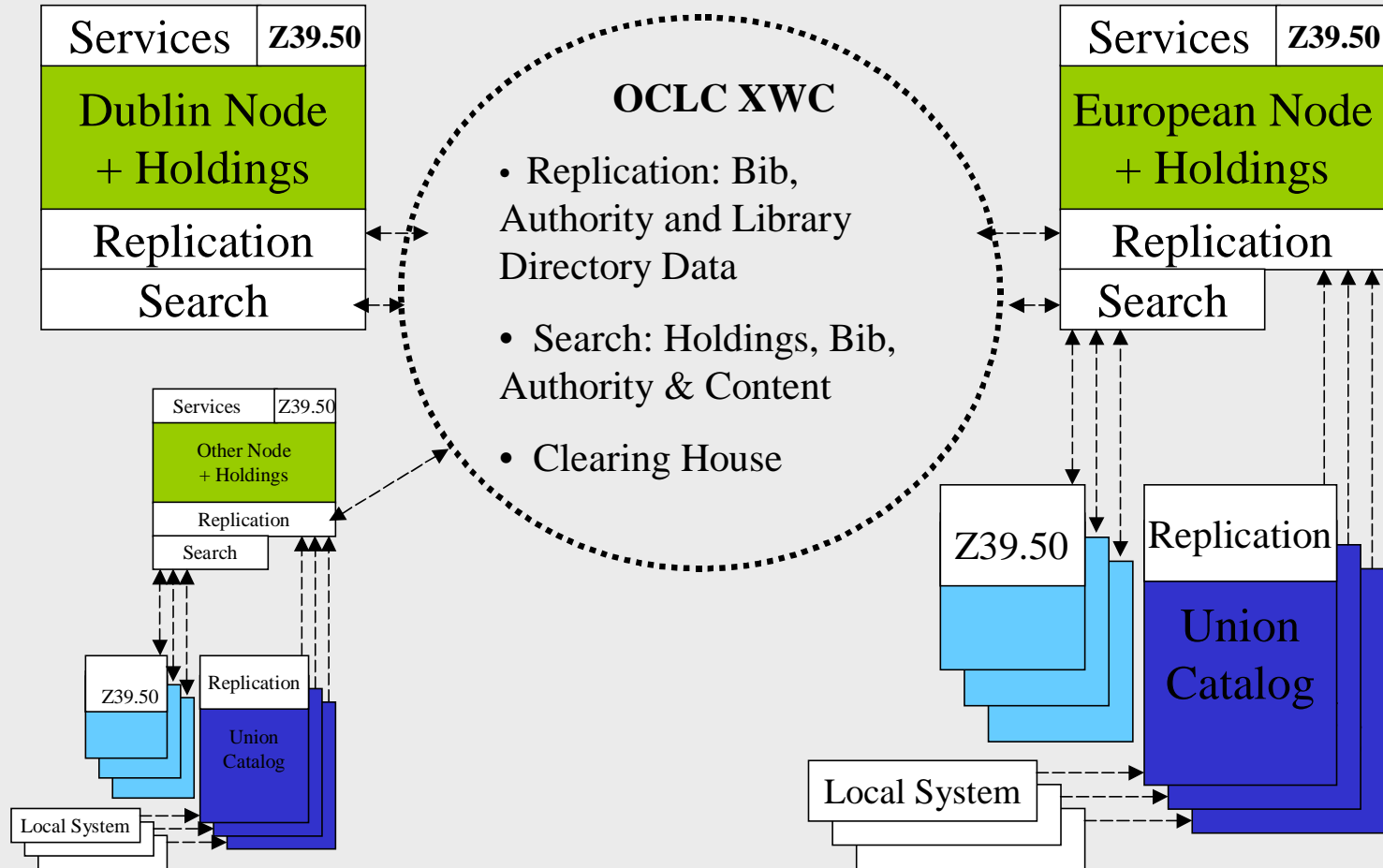
Recognized profile

- Bath
- CIMI
- GILS
- STAS
- Zthes

Z39.50 & OCLC

- WorldCat is a Z39.50 target
 - Local systems access, ILMS
 - Search and holdings update
 - Specialised portal access, e.g. ALEG
- /Port is a Z39.50 origin (client)
 - Accesses multiple targets
- Role in Extended WorldCat (XWC)
 - Optional tier of membership
 - Will support contending protocols too

Global Resource...Regional Autonomy



Union Catalogues

REAL

VIRTUAL

- Many accesses
- Speedy – bespoke srch
- Precision
 - duplicate detection
 - Consistent indexing
- Recall
 - Authority control
- Delivery
 - ILL; text links
- Multiple views
- Resource
 - Data mining
 - Collection development
 - Authentic source

Lowest common den.

- Slower
- Only small sets
- Dep. Profile adherence

Z39.50 magic

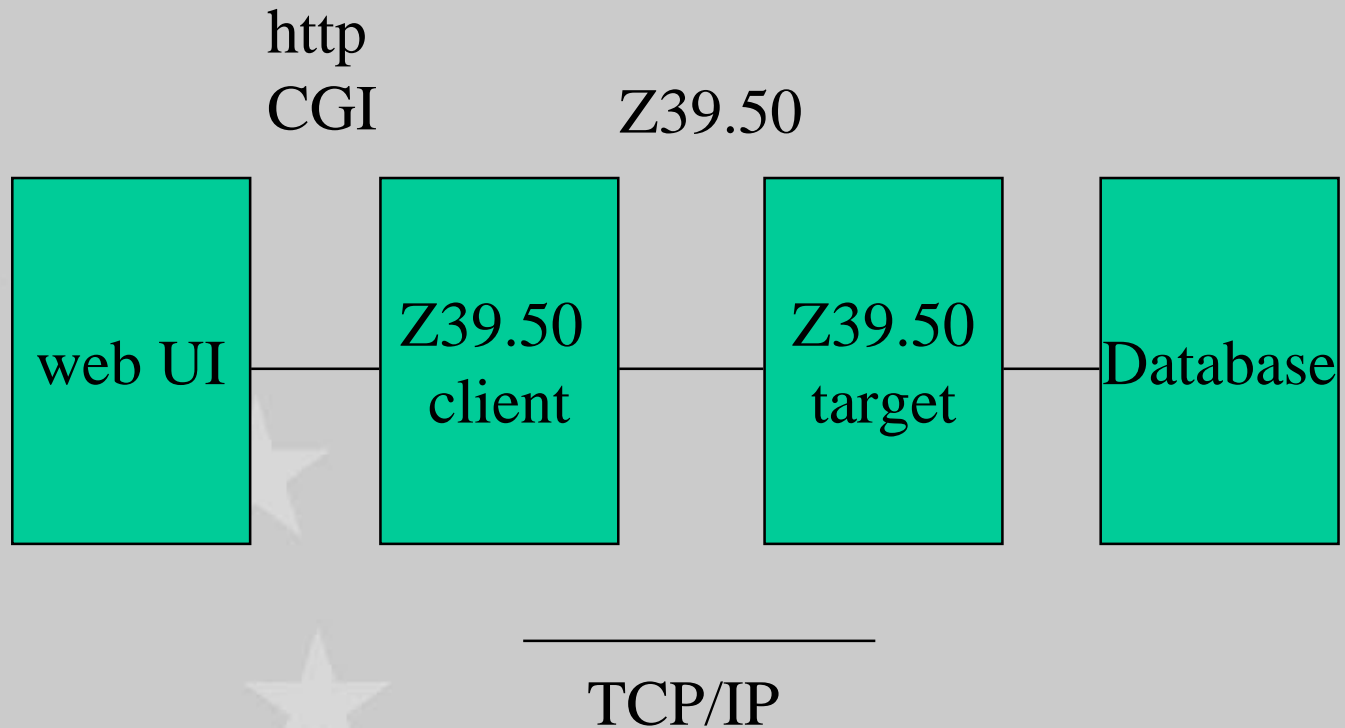


- Legacy systems client server
- Can access foreign databases with local UI
- Multi-target searches make virtual catalogues
- Databases without authority control appear authorised
 - client does it with mirrors
- Copy cataloguing world wide

Cataloguing steps

- Create MARC record in user interface
- Authorities checked against authority database
 - Field by field OR
 - During record validation
- Record checked for duplication
 - Title & ID elements, perhaps algorithm
- Record updates database

Most common architecture



Z39.50 Technicalities

- Specified in ASN.1
 - not easy to read... but
 - designed to minimise byte traffic
- Encoded in BER
- TCP/IP port 210 normal but not mandatory
- Minimum to connect
 - TCP/IP address & Port number

Basics

- Init request
 - requests options
 - password
- Search request
 - srch terms - attributes
 - syntax of results data
 - ESN - content
- Present request
- Scan request
- Init response
 - can reduce options
 - authenticates
- Search response
 - results count
 - can include results
- Present response
- Scan response

Search

- Attributes
- Bib 1 attribute set
 - use title / author / etc.
 - relation equal / LE / GT etc.
 - truncationright / none / left / I&r
 - completeness complete / incomplete fld
 - position anywhere / first etc.
 - structure word / phrase etc.

Results

- Requesting results - format (Syntax)
 - list of MARC (syntax is really ISO 2709)
 - MARC 21 (most commonly supported)
 - UNIMARC etc.
 - XML
 - Z39.50 inventions
 - GRS-1 (generic record syntax - tree)
 - SUTRS (simple unstructured record syntax)

Results options

- Requesting results - content (Schema)
 - Element set name (ESN)
 - f (full) and b (brief) *Server defined*
 - eSpec - complex request (*little supported*)

Historical mess

- Confused syntax and schema
 - USMARC, UNIMARC etc. called syntaxes
- Syntax is ISO 2709
 - Leader, directory, field, subfield
- Syntax could be XML
 - <200a> title proper data </200a> (UNIMARC)
 - <245a> title proper data </245a> (MARC21)

Other Z39.50 functions

- Sort
 - Origin asks target to pre-sort, e.g. date descending
 - De-duplicate can be part of sort
- Access Control

| Facility | Service |
|------------------------------|---|
| Initialization** | Init** |
| Search** | Search** |
| Retrieval** | Present** Segment* |
| Result-set-delete | Delete |
| Browse* | Scan* |
| Sort* | Sort* |
| Access Control | Access-control |
| Accounting/ Resource Control | Resource-control Trigger-resource-control Resource-report |
| Explain | uses the Services of Search and Retrieval |
| Extended Services* | Extended-services* |
| Termination* | Close* |

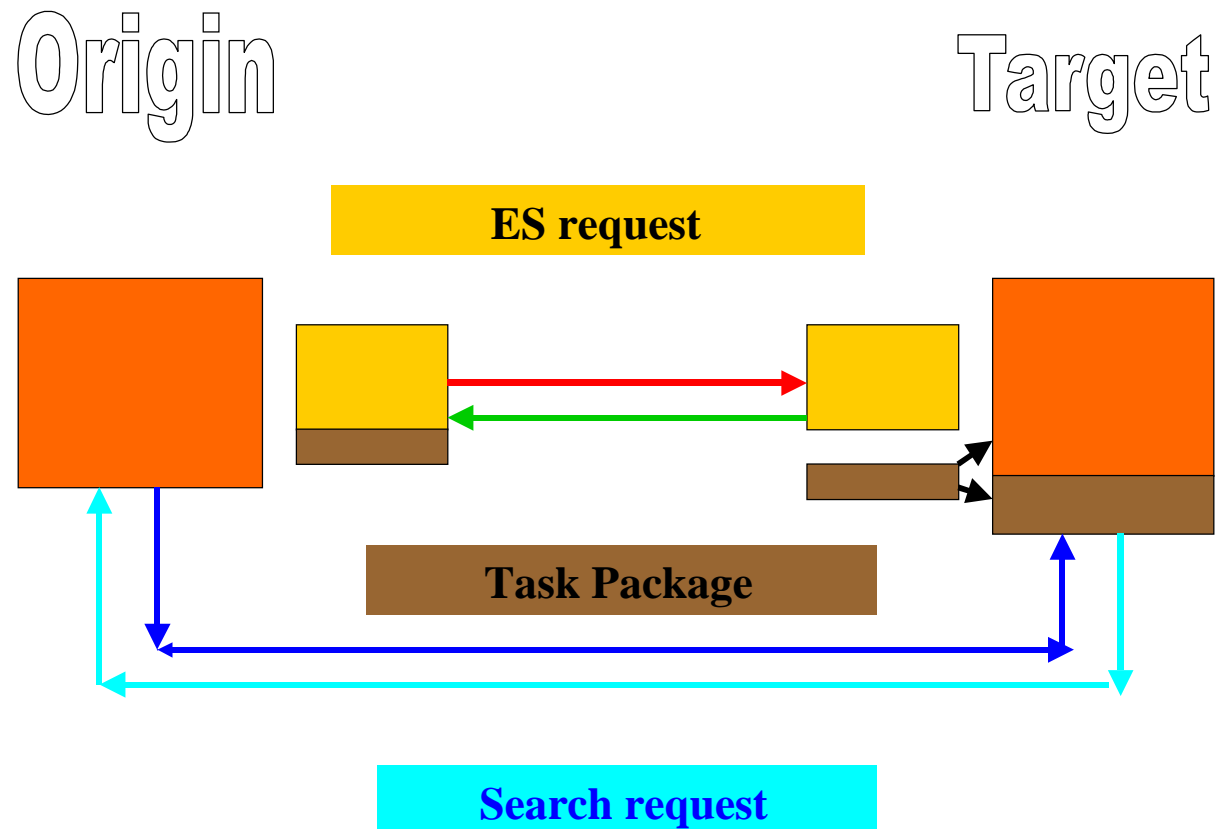
Browse

- Does not return records, just terms
- Follow on search necessary
- Bath profile indicates method

Extended Services

- Save a result set for later use:
- Save a query for later use
- Define a periodic search schedule
- Order an item*
- Update a database*
- Create an export specification

ES - Life beyond Z Association



ES update

- UCP - Union Catalogue Profile
 - Update cases
 - Diagnostics
 - Record locking / version control
 - Duplicate detection
 - Bibliographic / Authority / Holdings data
 - Global change
 - Merge

ES update & UCP

- Client can update >1 database
- Encourages union catalogue maintenance
- Other side of the coin
 - mirroring versus multi-target search
 - reduces de-duplication, merge
 - better quality results - authorised
- Distributed error resolution
- FTP pref. when update results not wanted

Current topics

- Attribute set architecture
- Holdings retrieval
- Better web integration

Attribute architecture

- Bib1 a mess
 - duplicates
 - similar attributes not grouped; choice unclear
- New architecture
 - can mix attributes from different sets
 - general attributes with
 - functional qualifiers - e.g. name as subject
 - semantic qualifiers - e.g. personal name (600 not 610 etc.)

New Attribute sets

- Bib 2
- Holdings
- Utility
 - record id
 - date of modification etc.
- Cross Domain

Holdings

- Holdings schema
 - network and local models
 - retrieval by bibliographic item
 - retrieval by location
- Holdings attribute set
- Mechanism for retrieving holdings
 - eSpec Q - search refinement as part of retrieval
 - searching by citation elements, Vol, date etc.

Better web interaction - ZING

- SRU = Simple URL..... in web browser
 - Target can include style sheet indicating display
 - Very thin client
- SRW = more complex
 - SOAP wrapper, remote procedure call
 - Send and receive XML documents
- Common elements
 - remove statefulness; no init
 - Only 1 record syntax = XML
 - text passed not encoded data
 - CQL based on CCL search
 - Inherits search modelling of Z39.50 & Bath

iPort

- Portal
- Multi-target broadcast search
- Multi-protocol
 - Z39.50
 - HTTP / XML; openURL like
 - LDAP etc. for authentication
- Based on YAZ tools
 - Search formatted in PQN

Demonstration



Maintenance Agency

- <http://www.loc.gov/z3950/>