Z39.50 Implementation Team
Summary of responses to OCLC’s additional z39.50 questions; Linking in the UC Union Catalog
30 May 2007

The Task Force received additional questions from OCLC (via Patricia Martin) concerning z39.50 connectivity. A summary of responses for each local system follow each question.

In addition, the issues concerning linking in the UC Union Catalog have been compiled by Mary Heath and Michael Thwaites. Those issues follow the responses to the OCLC z39.50 questions.

The Task Force has additional questions for OCLC. Those questions are listed at the end of this document.

Additional OCLC queries:

- Can the current z39.50 service respond to an OCLC number in the query instead of a local system number (recognizing limitations they point out)? If more than one record matches, what will be returned?

ALEPH systems:

UC Davis built an index based on Bibliographic Utility Number, which contains oclc record numbers, rlin record numbers, and Ballots record numbers. UC Santa Barbara doesn't appear to have such an index, but could build one if they were to regenerate their direct indexes. The oclc index has been added to the UCD Z39.50 configuration with a Use attribute value of 1211 and is currently searchable in production.

Z39.50 server always returns a search result message indicating whether the search succeeded and including the number of matching records. It is up to the client to request the display of records in the set using the protocol.

Gladis system:

Pathfinder Z39.50 service doesn't support queries using an OCLC number. While the code could be modified to allow it, the programming involved is significant. In case it is decided to follow this route, the return value in case of multiple records could be negotiated.

III systems:

The III systems can modify their z39.50 server so that when it sees a particular incoming use attribute, the search goes to a specified index on our
systems. It appears that attribute 48 (No. nat'l biblio.) may be the correct attribute for oclc number. Each of the Ill user libraries would have to contact Ill to set up a config so the Server would search in the index containing our oclc numbers anytime an incoming search specified use attribute #48. Will have to confirm with oclc that attribute 48 is the correct one.

If a search by oclc number results in multiple hits, the client receives a 'browse screen'

**Voyager system:**

Yes... sort of. Z39.50 doesn't appear to define an OCLC# use attribute as such. Voyager does not support use attribute 1007 "Identifier-standard", which theoretically includes the 035 field where we store our OCLC numbers.

Our OCLC numbers *are* included in our keyword indexes, thus accessible via Z39.50. However, Voyager doesn't seem to allow us to create or assign a use attribute which is used *solely* for OCLC numbers. So, false positives can happen when some other number in our records matches the input string.

Regardless of the details, if multiple records match, it's up to the Z39.50 client to decide what to do. The server simply creates a result set containing all matching records; the client must request the record(s) and handle them as needed.

If we use OCLC number as the key into the local system, we'll probably need to do some custom programming to try to make the search more precise than Voyager's implementation currently allows.

- Can the current Z39.50 service return a human-readable location instead of a location code? What changes would be needed to allow that?

**ALEPH systems:**

For UC Davis, the OPAC record includes a localLocation field that is the human-readable name of the unit where the item belongs and there is another field, shelvingLocation, which is the human readable location in the unit if there is a location other than the regular stacks.

**Gladis system:**

Currently only a 4-character code is supplied, however a change to a more readable result could be done with only limited changes.
**III systems:**

Currently, III systems will return the 'display' version of a location, not the code itself (for example, code rstk in an item record will display as Rivera, not rstk). If true, no additional changes are needed.

**Voyager system:**

Using the Z39.50 OPAC format, our public human-readable location is in the localLocation attribute for each holdings block.

- **Can the current Z39.50 service return summary holdings information?** If multiple holdings statements exist in the record, what would be returned (eg we understand that Voyager returns only the first)? Are there size limits in what would be returned and what happens when the size limits are reached (e.g. we hear that Aleph won't return any holdings).

**ALEPH systems:**

UC Davis modified the ALEPH X-Server output to map holdings information into an 852 field for presentation that contains in subfield 3 a summary holdings field. It would probably be possible to do the same thing to produce a summary holdings field in the MARC21 record that is returned with the OPAC record. (I will try to test this shortly.) Please note that the summary holdings statements are not automatically produced but created through serials maintenance staff. The OPAC record contains one entry for each piece. There is a point at which there are too many for processing, something that normally occurs with the very few serials that have hundreds of volumes. Processing serials with large numbers of items can cause an overload which produces an error message. There is a fix for this known bug which we have not been able to install, since our maintenance protocol requires considerable staff testing on our trial server prior to moving maintenance patches into production.

With regard to question 3, my first attempt to get a summary holdings statement into the Marc record by reconfiguring the Z39.50 output didn't work, so I'll have to figure out what was wrong with my approach.

**Gladis system:**

Pathfinder Z39.50 service provides multiple holding, however, if a journal has many volume statements/holdings, it returns information about the volumes which are charged out and not available. So more of an "unavailability statement", rather than full availability for each item.

There is an upper limit of 7,440 characters, and beyond that it gets
truncated. This could probably be upped some, maybe to 15,000, but there are some structural limitations at play.

**III systems:**

III Systems can return summary holdings information if it receives requests in either OPAC format or SUTRS format.

If multiple holding statements exist, the system will return information from each of the linked records.

There are no size limits, the system will output holdings information for as many linked checkin or item records that exist for a particular title.

**Voyager system:**

Voyager returns only the *final* 866, 867, and 868 in each holdings record, concatenated in the enumAndChron attribute for each holdings block. Thus, incomplete summary holdings would be returned for roughly 25,000 of our holdings records. This mainly affects records for things like government documents in print or on microfiche - many numbers are missing or were never issued, and our holdings statements are really long because we try to list everything we do have.

I'm not aware of size limits for what's returned, other than those imposed by the MARC standard (fields: 9999 octets; records: 99999 octets). I don't know whether the Z39.50 standard imposes different limits. Since Voyager concatenates data in enumAndChron, I suppose the combination of multiple long-but-valid fields could result in a too-long result.

**Linking in the UC Union Catalog**

I. It is important that a UC Union catalog display holdings, local call numbers, shelving locations, and availability information for all items it contains. Currently, this is accomplished by a Z39.50 query from the bibliographic database server to the local campus ILS's in real time as user interface record displays are constructed. In any new implementation of a UC Union catalog, linking between the bibliographic database server and the local ILS’s must be enabled to provide this function. It is important the linking be effective for a very large percent of records displayed. Because a relatively small percent of the UC bibliographic records contain ISBN or ISSN numbers, there seem to be only two numbers that have the possibility of providing a close to 100% match between OCLC records and the UC campus ILS
records; these are the OCLC bibliographic number and the local system number in the campus ILS.

a. Use of the OCLC bibliographic number: currently, a percentage of ILS bibliographic records contain OCLC numbers. A study of the campus records in Melvyl shows that approximately 30%-40% have OCLC numbers. In order for the OCLC number to be useful as a linking device, it would be necessary for all campuses to update their ILS bibliographic files with OCLC numbers. This would presumably follow the reloading of their ILS bibliographic files into OCLC for the existing Reclamation Project. This poses a rather large initial workload for the campus libraries, in addition to an ongoing cost to maintain the numbers.

b. Use of the local system number: currently, this data is completely missing from the OCLC databases. The Reclamation Project reloads provide an opportunity for OCLC to capture this data for all records that are loaded. This would create an issue for campuses that do not send their full bibliographic file as part of the project; many of the records that are not included would probably lack match points in OCLC. It seems that the workload involved in having every campus send a full file for reloading would be less than that required to utilize the OCLC bibliographic number as a match point, and the ongoing maintenance cost would be less.

II. Any new UC Union catalog must also provide input for the UC Request service. And, it should also be noted that we would like for a user to be able to request multiple items with a single click of the Request button. Currently, the UC Request system uses Z39.50 to obtain required information for the request processes. The Request system performs a number of tasks before constructing the request that is sent to VDX; it authenticates the user and informs him/her if there is any blocking on his record, it alerts the user if there is an electronic copy of the document being requested, it alerts the user if there is a copy of the item available on his/her campus, it obtains full bibliographic data and holdings and circulation data, it filters out copies that are library use only, it sorts the rota based on net-lender/net-borrower status of owning libraries, geographic nearness to the patron, and the most recently used location. For these operations, the Request system must have both full bibliographic information as well as holdings and circulation level data.

The ideal scenario for communication between the OCLC UC Union catalog and the UC Request system is one that uses both OpenURL and Z39.50 protocols. There are number of variations of ways in which the Request system can obtain the information it requires, but some,
by requiring significant programming in each of the 3 systems involved, have greater overhead; and other methods have greater impact on the performance levels of the Request system or VDX. These factors must be weighed together to determine which is truly the most effective. The following are three examples of how the methods used might impact the existing services.

a. When the Request button is clicked, OCLC constructs an OpenURL containing the full bibliographic data that is sent to the Request service, including the OCLC number. The Request service queries OCLC through the Z39.50 protocol and OCLC returns a Z39.50 OPAC record containing holdings, shelving locations, call numbers, and availability information, for all UC libraries that own the item. Although the OPAC record contains the USMARC record, passing the full bibliographic data in the OpenURL would allow the request system to respond to the patron as quickly as possible by deferring the OCLC search until after the initial interaction with the patron, thus creating a better user experience. The Request system then builds the rota and passes the request into the UC VDX system. This seems the most efficient approach with the least performance impact on any of the systems involved.

b. OCLC constructs an OpenURL containing the full bibliographic data AND the holding symbols and linking numbers of UC libraries owning the item which it forwards to the Request service. The Request system builds the rota using the OCLC symbols and linking numbers to obtain, through Z39.50 calls to the owning UC libraries, holdings and circulation data, then passes the request to VDX. This has the advantage of a low impact on Request system performance.

c. OCLC sends only the OCLC bibliographic number and the OCLC symbols of the owning UC Libraries to the Request service. The Request system uses a Z39.50 query to obtain a USMARC record from OCLC and builds a rota containing only the OCLC symbols, forwarding the request to VDX. VDX then queries the local ILS of each library that it routes the request to using the linking number to obtain holdings, call numbers, shelving locations, and availability information for that location. By shifting the Z39.50 calls to VDX, this method has to potential to triple the number of such calls made by VDX, and would significantly affect that system’s performance. I have some hesitancy about putting the additional Z39.50 work into VDX.
Questions for OCLC:

- If each of the local systems allowed for searching on OCLC number, what attribute value should be used? Would it be helpful if each system used the same attribute value?
- Since a large percentage of records in Melvyl do not have an OCLC number, what immediate advantage would there be to use an OCLC number search?