

Z39.50 Broadcast Searching and Z-Server Response Times: perspectives from CC-interop

George Macgregor

George Macgregor (george.macgregor@strath.ac.uk), Centre for Digital Library Research (CDLR), Department of Computer and Information Sciences, University of Strathclyde, Glasgow, UK.

Keywords: Z39.50, online catalogues, union catalogues, distributed searching, information retrieval, SRW

Abstract

<i>Purpose of this paper</i>	<p>This paper begins by briefly outlining the evolution of Z39.50 and the current trends, including the work of the JISC CC-interop project.</p> <p>The research crux of the paper focuses on an investigation conducted with respect to testing Z39.50 server (Z-server) response times in a broadcast (parallel) searching environment.</p>
<i>Design/methodology/approach</i>	<p>Customised software was configured to broadcast a search to all test Z-servers once an hour, for eleven weeks. The results were logged for analysis.</p>
<i>Findings</i>	<p>Most Z-servers responded rapidly. 'Network congestion' and local OPAC usage were not found to significantly influence Z-server performance. Response time issues encountered by implementers may be the result of non-response by the Z-server and how Z-client software deals with this. The influence of 'quick and dirty' Z39.50 implementations is also identified as a potential cause of slow broadcast searching.</p>
<i>Research limitations/implications</i>	<p>The paper indicates various areas for further research, including setting shorter time-outs and greater end-user behavioural research to ascertain user requirements in this area. The influence more complex searches, such as Boolean, have on response times and suboptimal Z39.50 implementations are also emphasised for further study.</p>
<i>Practical implications</i>	<p>This paper informs the LIS research community and has practical implications for those establishing Z39.50 based distributed systems, as well as those in the Web Services community.</p>
<i>What is original/value of the paper?</i>	<p>The paper challenges popular LIS opinion that Z39.50 is inherently sluggish and thus unsuitable for the demands of the modern user.</p>

Introduction

It is often forgotten that Z39.50 protocol has existed, in one form or another, for almost 30 years. Still, it was only in 1995, with approval granted by the National Information Standards Organisation (NISO), that the standard attracted significant attention from the LIS community, as well as some minor acknowledgement from beyond the library world (Needleman, 2002, p.248). By the late-1990s, this attentiveness had spread internationally and had manifested itself in a flurry of Z-based research projects and activity, particularly in the UK where the third phase of the Electronic Libraries programme (eLib) stimulated the creation and evolution of several virtual union catalogues (or 'Clumps' as they became colloquially known) (Dovey, 2000).

Yet perhaps more incredibly, it is only now that deployment of Z39.50 within the library and information services sectors is truly reaching 'critical mass'. Z-enabled OPACs are, as Needleman (2002, p.249) notes, now commonplace within the academic and research library fraternities, an observation that could not have been made until recently. Indeed in the UK, as in many information rich countries, Z39.50 is now gaining prevalence within FE and public library sectors, thus facilitating the creation of ever larger, heterogeneous, virtual union catalogues and cracking open the possibilities for distributed searching by end-users (Dunsire & Macgregor, 2003). More intriguingly, it is predicted that with the next revision of Z39.50 scheduled for 2005, those sectors that have hitherto expressed tepid enthusiasm for the standard (museums, archives, and others) will edge closer to Z39.50 compliance (Taylor, 2003). Though this development would undoubtedly uncover a plethora of difficulties and interesting issues pertaining to the interoperability between, and distributed searching of, cross-domain catalogues, it underlines the pervasive nature of Z39.50 and further illustrates the confidence sought by others in a standard that is, by now, ubiquitous in the library community, as well being internationally recognised as the "global standard" for networked information search and retrieval (NISO, 2002, p.5).

Whilst the advantages of any standard are manifest in its original introduction and adoption, Z39.50 is not without its faults. Some of these have been widely documented (Gatenby, 2002; East, 2003) and examined (Moen, 2001a; Moen & Murray, 2002), whilst others have undergone thorough analysis under the auspices of the CC-interop project (Nicolaidis, 2003; Gilby & Sanders, 2003; Gilby *et al*, 2004; Dunsire & Macgregor, 2004). Nevertheless it remains true that despite whatever difficulties Z39.50 might present, it continues to rule distributed searching for the library world and will do for the foreseeable future. It constitutes a significant cornerstone in the technical architecture of the UK Joint Information Systems Committee (JISC) Information Environment (IE) (Powell, 2004), continues to be assiduously bandied by library system vendors, and remains a central component of many commercial content management systems (CMS), such as ENCompass (Dietz & Noerr, 2004).

Those technologies expected to eventually supersede Z39.50 entirely, Web Services Technologies (WST), are currently thought to fall short of providing the rich access already offered by Z39.50 (McDonald, 2003) and, as Yu and Chen (2003) note, there are limitations and barriers to be overcome by Web Services, many of which are similar to Z39.50. However, the 'Z39.50 International: Next Generation' initiative (ZING, 2004) have been spearheading a flood of immensely exciting experiments and developments, particularly Search/Retrieve Web Service (SRW) and Search/Retrieve URI (SRU). SRW/SRU is an attempt to conflate the powerful capabilities of Z39.50 by implementing them in tandem with updated Web-friendly protocols and technologies, such as HTTP (Hypertext Transfer Protocol) with SOAP (Simple Object Access Protocol), a protocol for XML (Extensible Markup Language) messaging, and by utilising WSDL (Web Services Definition Language) to define the Z39.50 messages. Although promising far greater functionality, developments remain tentative with the first official specification (Version 1.1) only released in early 2004, but coinciding with some tantalising 'real life' applications of the protocol via the European Library project (van Venn & Oldroyd, 2004).

Indeed, although ZING are aiming to "lower the barriers to implementation while preserving the existing intellectual contributions of Z39.50" (ZING, 2004) - a move that is hoped will eventually assist wide adoption in the larger Web-based community - it will be many years before it is as widely accepted as Z39.50 in the library community. In addition, and perhaps ultimately, SRW will not provide deliverance in respect to semantic interoperability and those variations in cataloguing and indexing practices that continue to blight

optimal performance of Z39.50 virtual union catalogues will linger. In any case it would appear that Z39.50 will retain, at least for some time yet, its crown as the "eminent enabling technology for distributed, parallel access to information sources" (Hammer & Andresen, 2002).

To this end JISC in the UK (<http://www.jisc.ac.uk/>) has been funding research via the CC-interop project (<http://cc-interop.cdrl.strath.ac.uk/>) into numerous issues, including testing the feasibility of inter-linking between union catalogues, both physical and virtual, as well as investigating the use of collection-level description schemas in relation to physical and virtual union catalogues. The crux of this paper, however, will focus on research and findings relating to Z-Server response times and the performance of Z39.50 for parallel searching.

Before discussing this, it is worth contextualising the said research within the remit of CC-interop. For those unacquainted with the technology, there is also some merit in briefly summarising how Z39.50 functions, however it is not the purpose of the authors to provide an exhaustive explanation of the technical operations of the protocol. For this refer to NISO (2002), Moen (2001b), Lynch (1997) and Taylor (2003).

Z39.50

ANSI/NISO Z39.50 is a communications protocol maintained by the Z39.50 Maintenance Agency at the Library of Congress (Z39.50, 2004), enabling standard messaging between a Z39.50 client (Z-client) and a Z39.50 server (Z-server), and supporting the searching and retrieval of information in all formats in a distributed networked environment. NISO defines Z39.50 yet more simply, as a "standard protocol used by networked computer systems for information retrieval" (NISO, 2002, p.3).

Essentially Z39.50 functions as a common language allowing interpretation by Z-enabled systems, irrespective of what software, systems, or platforms are in operation at the client or server. Most implementations use the standard TCP/IP Internet communications protocol to connect systems and Z39.50-compliant software in order to decipher messages between them for searching and retrieval. By normalizing the messages used by the client and the server, technical interoperability can be achieved. Thus, any search query initiated by the end-user (at the client interface) is immediately translated by the client software for sending to the remote 'Z-server' (or 'Z-target'). Once the server is in receipt of the search details, it utilises those rules dictated by Z39.50 to decode the search into a format recognised by the local database. These exchanges are defined by attribute sets, the most prevalent of which is the Bib-1 attribute set (Z39.50, 2003). The Bib-1 attribute set underpins the dominant library profiles, such as the Bath Profile (Bath Profile Maintenance Agency, 2004) and the Z-Texas Profile (TZIG, 2003). Once the remote server has decoded the search according to the aforementioned conventions, it initiates the search locally and then returns the results of that search to the client. The results will then be displayed to the user in a pre-determined format. This format will depend on the configuration adopted by the client. Increasingly Z-client software conducts this processing, but more often than not Z-client software either has to be customised or custom software has to be deployed in tandem with the Z-client to undertake this post-results processing.

Virtual Union Catalogues and Clumps

As Z-client software has developed, and as librarians have recognised the potential for distributed search and retrieval for the end-user, the protocol has made feasible the construction of complex distributed information environments whereupon it is possible for the Z-client to connect to multiple Z-targets. Such an approach allows the user to 'broadcast' a single search to multiple Z-enabled catalogues simultaneously and have the results from each catalogue returned and merged into a single result set, perhaps with duplicate records removed depending on Z-client configuration. As mentioned, the late-1990s witnessed a spate of Z39.50 activity as various LIS communities across the globe furiously set about developing virtual union catalogues. The UK was no exception and was the hub of significant activity.

Arising from the MODELS (Moving to Distributed Environments for Library Services) initiative, the JISC-funded electronic libraries programme (eLib), funded the creation of four virtual union catalogue services (or Clumps) in 1998 to conduct further research and develop Z39.50 for the purposes of expansive information retrieval in the UK (Stubley, 1998). A 'Clump' was defined as an aggregation of catalogues, including

physical union catalogues; this definition has subsequently been refined to refer to those aggregations that are inherently distributed only, and is now more commonly used to specifically describe aggregations based on Z39.50 (Dunsire & Macgregor, 2003). Although creating a service that would experience wide use by end-users was a tacit objective, the overarching purpose of the Clumps was to "kick start critical mass" in the use of Z39.50 and to generate model technical architectures and agreements to precipitate the subsequent growth of new clumps in their various permutations, perhaps even nationally (Whitelaw & Joy, 2001, p.2).

Of the four Clumps created, three were regionally orientated and existing library consortia provided the sure foundation for development:

- The Co-operative Academic Information Retrieval Network for Scotland (CAIRNS) (<http://cairns.lib.strath.ac.uk/>) included members of the Scottish Confederation of University and Research Libraries (SCURL) and is now developed and maintained by the Centre for Digital Library Research (CDLR) at the University of Strathclyde;
- M25 Link had six partners drawn from the M25 Consortium of Academic Libraries based in the London area (<http://www.m25lib.ac.uk>). The resulting distributed catalogue now forms part of the InforM25 service and is maintained for the consortium by the M25 Systems Team;
- RIDING included members from the Yorkshire and Humberside Universities Association (YHUA) (<http://www.riding.ac.uk/>)
- Music Libraries Online (MLO) was the only Clump not to be regionally focused. Comprising nine UK conservatoire libraries, MLO facilitated distributed access to scholarly music resources (<http://www.musiconline.ac.uk/>)

All these projects successfully established fully functioning clumps, each with common and peculiar features. CAIRNS, for instance, instantiated a 'dynamic clumping' mechanism - or 'landscaping mechanism' - based on Conspectus subject strength measurements conducted by the SCURL member libraries (Nicholson *et al*, 2001), whilst M25 Link investigated dynamic clumping by geographical zones of London and the availability of periodicals holdings via Z39.50 (Brack *et al*, 2001).

The CC-interop project

By 2002 JISC had provided a two-year funding grant to the Copac/Clumps Continuing Technical Cooperation Project (CC-interop), a collaborative project involving the M25 Systems Team, CDLR, Manchester Information and Associated Services (MIMAS), RIDING, and latterly the Centre for Research in Library and Information Management (CERLIM). Building on the results and findings of the JISC eLib programme, CC-interop enhanced the 'distributed' thread of the JISC Information Environment in that it "aims to bring together, in a virtual *modus operandi*, distributed catalogues to facilitate richer search and retrieval possibilities for users" (Gilby & Dunsire, 2004, p.4). The inclusion of the Copac service (<http://copac.ac.uk/>) at MIMAS - a physical union catalogue based on the consolidated bibliographic records of the Consortium of University and Research Libraries (CURL) and searching some 30 million bibliographic records - exemplified the cooperative nature of the project: true collaborative research emanating from both the virtual and physical union catalogues schools of thought.

Ending in the summer of 2004, CC-interop comprised three work packages, each investigating a plethora of issues, including:

- Inter-linking between very large physical union catalogues (i.e. Copac) and large virtual union catalogues (i.e. InforM25)
- The ability to 'clump the clumps' thus creating a 'hyper-clump'
- Thorough research of collection-level description requirements for such environments
- Improving interoperability in distributed and physical environments
- Investigating user requirements and behaviour for union catalogues.

For a greater discussion of the project outcomes and findings refer to Gilby and Dunsire (2004).

It was also within the remit of CC-interop to undertake some investigation of certain Z39.50 performance issues. Naturally, as in any research project, an abundance of noteworthy findings were accumulated in

relation to this topic alone. Yet within this, particularly interesting findings pertaining to Z-server responses times were gleaned and hereupon is a detailed exposition of that research and the results attained.

Research: Z39.50 Searching and Response Times

As noted earlier, Z39.50 is not without its faults. Conducting broadcast searches (or 'parallel searches') via Z39.50 is often considered to be sluggish and lacking robustness (Stubley *et al*, 2001). Such perceptions have been borne out by detailed user studies whereby current user expectations are increasingly influenced by Web searching tools such as Google, to such an extent that failure to achieve rapid retrieval often compels users to abandon searches altogether (Booth & Hartley, 2004). Whilst Web search engines have a long way to go before they can address their respective lack of precision, ponderous recall and retrieval of base quality information, the unfortunate fact remains that users increasingly appear to rank speed of delivery over quality. As Nicholas *et al* (2003) note, user behaviour is increasingly 'promiscuous', with users progressively surpassing traditional quality concerns and conforming to the so-called 'bouncer' paradigm. As Nicholas *et al* conclude, "time plainly is a rare commodity" (Ibid., p.28). Such developments should not be ignored. Rather, they should inform the subsequent improvement of those services that embrace metadata, as well as informing those pioneering the improvement or augmentation of information literacy orientation at colleges and universities.

Yet since the emergence of Z39.50 the precise cause of this anomaly in performance and the potential for broadcast searching has never undergone detailed scientific or exhaustive study. Instead the LIS community has been exposed to a variety of conclusions based on speculation or conjecture. It has become, as Hammer and Andresen (2002) pertinently note, "a 'folk wisdom' among Z39.50 implementers that the maximum, realistic number of servers to search in parallel was somewhere between 10 and 15" and that "Z39.50 is just inherently clumsy and slow to work with". At this juncture it is worth noting that this area of research is not without some contributions. Exciting, albeit 'informal', research conducted by Hammer and Andresen (Ibid.) under the auspices of Denmark's Electronic Research Library (DEF - Danmark's Elektroniske Forskningsbibliotek) have provided insights to some of the issues CC-interop wished to expose in a UK context. However this research, by their own admission, was not particularly 'scientific'. Rather, it was an "attempt to move the discussion of parallel Z39.50 applications away from guesswork and in the direction of hard information" (Ibid.), upon which other studies could construct further investigation. It was therefore this anomaly that CC-interop wished to address. Furthermore, it was also an opportunity to study any specific peculiarities within InforM25 - which would constitute the test-bed for investigation - and inform subsequent CC-interop and JISC IE developments.

Methodology

To enable investigation of the research question, Java Access to Electronic Resources (JAFER) software was configured to execute automated search tests across a number of the InforM25 member libraries, thus allowing the recording of search response times over a considerable period of time. JAFER is an open source software package that has recently been developed as part of the JISC 5/99 programme by staff at Oxford University and is described as a "Java based toolkit for building Z39.50 clients and servers" aimed at "programmers and web developers building resources for teaching and learning" (JAFER, 2003). As well as being freely available, it is built using industry standard tools that are themselves freely available for both Unix and Windows platforms. JAFER is also extremely flexible, supporting a broad selection of record syntaxes, including that of UKMARC and MARC21.

The decision to use JAFER for this experiment was dictated by two factors. Firstly, JAFER had already been deployed in CC-interop successfully to investigate the feasibility of transforming a clump into a Z-target (Gilby & Sanders, 2003) and was ergo readily available. Secondly, it was recognised that for this particular research task JAFER exemplified fitness for purpose and could be easily configured to achieve the desired research aims.

The JAFER client was therefore configured to broadcast a search to those InforM25 library Z-servers that were known to respond. This initially meant that 16 InforM25 libraries were included in the testing. However, by early December 2003, Buckinghamshire Chilterns University College (BCUC) appeared to be

responding to the search queries and was therefore added also. The libraries for which results are presented are available in Figure 1. Testing began on 6th October 2003 and concluded on 23rd December 2003.

Abbreviation	Institution	Library System
Birkbeck	Birkbeck, University of London	Horizon
Brunel	Brunel University	Unicorn
BCUC	Buckinghamshire Chilterns University College	Unicorn
City	City University	Innopac
Hertfordshire	University of Hertfordshire	Voyager
IOE	Institute of Education	Unicorn
Kent	University of Kent	Voyager
LBS	London Business School	Unicorn
Metro/LGU	London Metropolitan University (formerly London Guildhall University)	Innopac
Pharmacy	School of Pharmacy, University of London	Unicorn
Queen Mary	Queen Mary, University of London	Unicorn
St. George's	St. George's Hospital Medical School, University of London	Unicorn
St. Mary's	St. Mary's College, University of Surrey	Innopac
SAS	School of Advanced Study	Innopac
SOAS	School of Oriental & African Studies, University of London	Innopac
ULL/Heythrop	University of London Library and Heythrop College	Innopac
Wellcome Library	Wellcome Library for the History of Understanding of Medicine	Innopac

Figure 1 Libraries for which results are presented.

A simple author test search for 'Austen' was broadcasted, using Bib-1 'Use' attribute 1003. Exact attribute settings configured in JAFER for the individual Z-servers were the same as used earlier in the project (see *Ibid.*). JAFER was then configured to broadcast the search to all test Z-servers once an hour and the results were logged for analysis. The time recorded was the duration of initiating the Z39.50 connection between the JAFER client to the Z-servers, as well as the time taken to broadcast the query and receive a response from the Z-server giving the number of records in the result set. This specifically does *not* include the time needed to request and receive individual or groups of records from a Z-server, nor does it include any post-processing time or the time taken to display the records received via a user interface. The results give an indication of connection/database search times, wholly independent of the number of records, record type and any client specific processing.

Caveats

Whilst the authors are confident in the methodology deployed, there are several caveats that are worth noting:

- Given the large distributed nature of InforM25, the total data set did have the potential to greatly exceed 17. Regrettably though there were a number of Aleph and Talis libraries systems that did not function correctly when connected to JAFER and consequently these libraries had to be excluded from the tests and do not feature in the results. The exact cause for Aleph and Talis systems not connecting with JAFER is not yet known, but early tests indicated that it was attributable to the way in which the connection is requested by JAFER. This would necessitate further investigation but does not suggest a fundamental deficiency with the software. Institutional firewalls at some of the Talis sites were also identified. Such sites were removed from the data set to avoid the potentially lengthy negotiations required to have them opened for testing.
- Birkbeck, University of London was offline for significant periods during testing so data on Birkbeck does not appear in all the results. Also, as BCUC data was only recorded during December 2003, any

local problems that were present may have affected the results more than would have been the case if they had been recorded for a longer time period.

- As noted, testing was undertaken between 6th October 2003 and 23rd December 2003. However, data coverage during this period was not entirely comprehensive as JAFER occasionally runs out of system memory after a few days. When this occurred the software needed to be shut down and restarted. Obviously this marginally reduced the comprehensiveness of the recorded data, but this downtime did not always happen at the same time of day or for particularly long periods of time, so it is considered that this will not have significantly affected the results obtained, nor the observations that it is possible to draw.
- One final issue to note is that the tests were done with JAFER, installed on a PC at MIMAS and connected to the UK's education and research network, JANET. All the tested Z-servers were also connected to JANET, most via the London Metropolitan Area Network. Testing did not reveal any influences on response times due to the various network elements.

Results and Discussion

The test results are summarised in the graphs shown in Figures 2, 3 and 4. Figure 2 shows the frequency of response times for the tested Z-servers (rounded to the nearest 5 milliseconds). The second graph in Figure 3 depicts the way in which the response times varied during the day. Figure 4 has been included to illustrate the percentage (%) of searches per Z-server responding within categorised time periods (in seconds).

- Take in Figures 2, 3 -

As can be observed from the clustering of results shown in Figure 2, the majority of responses were received quickly, with c.91% of searches receiving a response within 1 second. This is what would be anticipated with a very simple query of the type used in the tests. By contrast, c.4% of all searches took between 4-27 seconds.

As Figure 3 reveals, some Z-servers were consistently fast in their response, indicated in the graph by an almost flat profile. For example, the City Z-server responded to almost 95% of searches within 0.125 seconds, with a small number of responses proving lengthier, up to 12.7 seconds in the slowest instance. Other libraries showed a much broader spread of response times, for example London Metropolitan University (LGU) responded to c.36% of queries within 1 second, c.33% 1-2 seconds, and c.27% in 2-4 seconds. BCUC and Pharmacy show a cluster of fast response times, then a cluster of slow ones, with over 34% of queries taking 4-27 seconds. In these examples the reasons for the cluster of markedly slower response times are worthy of further investigation, as the systems have revealed that they are perfectly capable of fast responses.

- Take in Figures 4 -

Figure 4 would also suggest that the response time does depend on the type of library system. In most cases the Innopac and Voyager sites (City, SOAS, Kent, SAS, St Mary's, Hertfordshire) have a very high percentage of response times under 0.25 seconds. Comparing London Met. with the other tested Innopac sites would suggest that there was something different about the Z-server installation at that institution as it constantly responded slowly when compared to other Innopac sites. It is entirely possible, as Moen and Murray (2003) have suggested in other cases, that this delay is attributable to sub-optimal Z-server implementations. Given the constant nature of testing variables, this would be a reasonable assumption to make, but would obviously be no substitution for further testing in our case study. Unicorn sites do generally appear to respond a little more slowly but it is unclear as to the cause of this. As East (2003) and Taylor (2003) have both noted, the implementation of Z39.50 at libraries can be an arduous task for even the most experienced librarians and information professionals (something that ZING developments hope to dissipate). Added to which, those 'quick and dirty' implementations favoured by systems vendors often engender yet further obstacles that the librarian has to overcome to ensure an optimal and smooth implementation.

Figure 3 illustrates the range of response times averaged for each hour of the day. As can be observed, the times vary from those Z-servers with a very consistent response time, to others showing large differences in average (mean) response time. For example City and Kent show relatively little variation in response times, whilst BCUC showed very obvious periods of slow response times, especially during the evening and overnight. It is noteworthy that where response time variations were prominent, the average slowest responses tended to occur early and late in the day, with the fastest responses occurring around mid-day and early afternoon. One probable cause for this is that library system databases often run jobs overnight such as re-indexing and back-ups which tend to take up processing capacity. Library OPACs generally experience higher usage from late morning to early evening so it can arguably be concluded that existing usage of the library system does not directly affect the Z39.50 response times as tested, and vice versa. This resonates with interpretations made by Moen and Murray (2003) and contradicts more popular assumptions that Z39.50 queries are more resource intensive than those queries delivered via the local OPAC (including remote OPAC interrogation over the Web). It would also suggest that so-called 'network congestion', reputed to occur from late morning to late afternoon, and reputed by the laws of 'folk wisdom' to diminish day-to-day Z39.50 performance, is not entirely valid. This latter finding confirms those obtained by Hammer and Andresen (2002). It is worth re-emphasizing, however, that testing carried out by CC-interop did not include the transfer of records, which as well as potentially increasing response times, may perhaps be influenced by the local usage of the library system.

The maximum search time of 27 seconds (Figure 4) is understood to reflect a time-out within JAFER that is initiated so as to avoid the user waiting for slowly responding or non-responding Z-servers. Most distributed systems have a time-out function and if this is too long, searches can appear slow to the user. System designers are presented with a dilemma in that sufficient time needs to be permitted for a slow Z-server to respond, but this is contrasted with the issue of what to do with a Z-server that is not responding at all.

Although there are potentially issues relating to perfunctory Z-server installations, the generally good performance of the Z-servers suggests that many of the response time problems, experienced by searchers conducting broadcast searching for uncomplicated searches, may be the result of non-response by the Z-server and how that is dealt with by the client software. For example, JAFER has a timeout of 27 seconds for non-responding Z-servers, but InforM25 has a cumulative timeout of 65 seconds. More complex searches may, of course, give somewhat different results. It is important to be aware that in InforM25, like many distributed searching environments, the overall searching time experienced by a user is only as fast as the slowest Z-server, so even where most searches are being performed quickly, one slow search is all that is needed to degrade the final response to the user. However it is also important to recognise that not all implementations take this approach and alternative systems, such as DScovary (Crossnet Systems, 2004) or Metalib (Ex Libris, 2004), can allow users to view results as they are received. This obviously means that the user receives result sets according to Z-target with little post-processing, as opposed to receiving a combined and definitive result set from a service like InforM25.

In an implementation such as InforM25, the test results do appear to suggest that where the slow response of a Z-server adversely affects the overall user query-to-results time, setting a short time-out for the initial Z39.50 connection and search response (e.g. 2 seconds) may help mitigate this. Of course the corollary dictates that those Z-servers that are slow to respond, or which are erratic in their behaviour, may usually be unavailable within a service for searching. Such decisions would have to be taken gingerly by service providers and be taken on a service by service basis. Inevitably such a decision would also require detailed analysis of users' requirements. Be that as it may, user behaviour in the searching of union catalogues, as found under CC-interop (Booth & Hartley, 2004), may perhaps suggest that most users would consider such a 'trade off' acceptable, especially if it meant that results sets were displayed more quickly.

Conclusions and Further Research

Whilst the true promise of ZING is afoot and is particularly alluring for the LIS community, it is quite clear that the deployment and uptake of Z39.50 by libraries will not abate for the immediate future. Indeed it is only now, in 2004, that Z-compliance is reaching decisive levels. Such decisive levels of compliance obviously render the creation of large heterogeneous distributed union catalogues ever more likely, and it is therefore imperative that issues pertaining to semantic interoperability and, in our case, performance are

addressed to ensure end-users do not consider such retrieval tools as irrelevant in the face of those 'low value' tools to which they cling bitterly.

As revealed by the crux of this paper, Z39.50 need not conform to the popular perception that it is 'dinosaurian' and too 'clunky' or bloated for deployment on the modern Web. As Hammer and Andresen (2002) are keen to indicate, Z39.50 is a lightweight protocol, optimised for good performance over large slothful networks. In point of fact, the lightweight genesis and subsequent development of the protocol was necessary as the 1980s imposed severe bandwidth limitations. The results of this study should hopefully inform further research in the area of Z-server response times. In particular, the community would benefit immeasurably from further research into the effect 'quick and dirty' implementations have on Z-server response times, as well as greater technical analysis as to why, in our study, certain library systems appear to greatly influence response rapidity. Moreover, although on site usage of local OPACs did not appear to influence Z39.50 response times as tested, and whilst Moen and Murray (2003) consider Z39.50 queries to be no more resource intensive on local OPACs than those queries delivered via the local OPAC or over the Web, further testing of intensive Z39.50 querying would be prudent so that conclusive data can be gathered on whether such querying could negatively influence Z-server response times and/or local OPAC performance. There would also be some merit in examining the influence more complex searches, such as Boolean, have on response times. With ever larger virtual union catalogue implementations probable in the future, such research is essential to avoid performance degradation of local OPACs for local user communities, and also to further our collective understanding of Z-server response times generally.

More importantly, greater end-user behavioural research has to be undertaken to ascertain user requirements with respect to the applicability of establishing short time-outs for Z39.50 connections and search responses. Such research would not only inform the Z community (including ZING), but would also inform those champions of Web Services, where the issue of 'transaction time' constitutes a significant obstacle for successful Web Services application (Yu & Chen, 2003). In a similar vain, further end-user behavioural studies are required in relation to those Z39.50 implementations like DScovary, as establishing short time-outs may perhaps be a preferable solution if users' necessity for post-processing is significant.

Ultimately though, semantic interoperability remains the single largest obstacle to improving the overall performance of virtual union catalogues based on Z39.50, an issue that CC-interop grappled with and one that will likely remain atop the LIS agenda even when SRW compliance reaches critical mass.

Acknowledgements

The author wishes to gratefully acknowledge **Ashley Sanders** and **Dr. Shirley Cousins** of Copac for their contribution to JAFER configuration, Z39.50 testing and results processing. Acknowledgement is also extended to **John Gilby** of the M25 Systems Team for his insightful comments and assistance throughout.

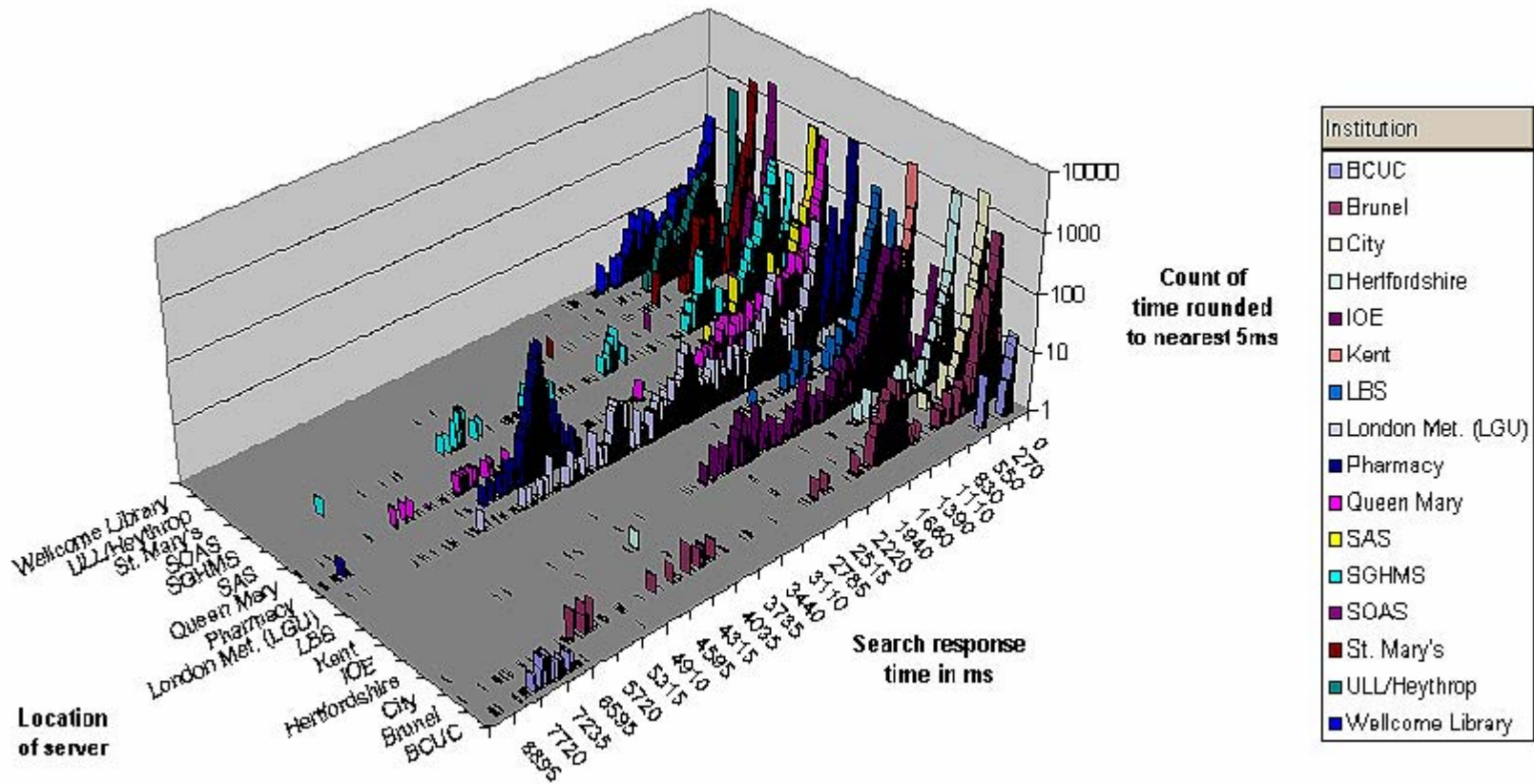


Figure 2: Frequency of Response Times for the Tested Z-Servers

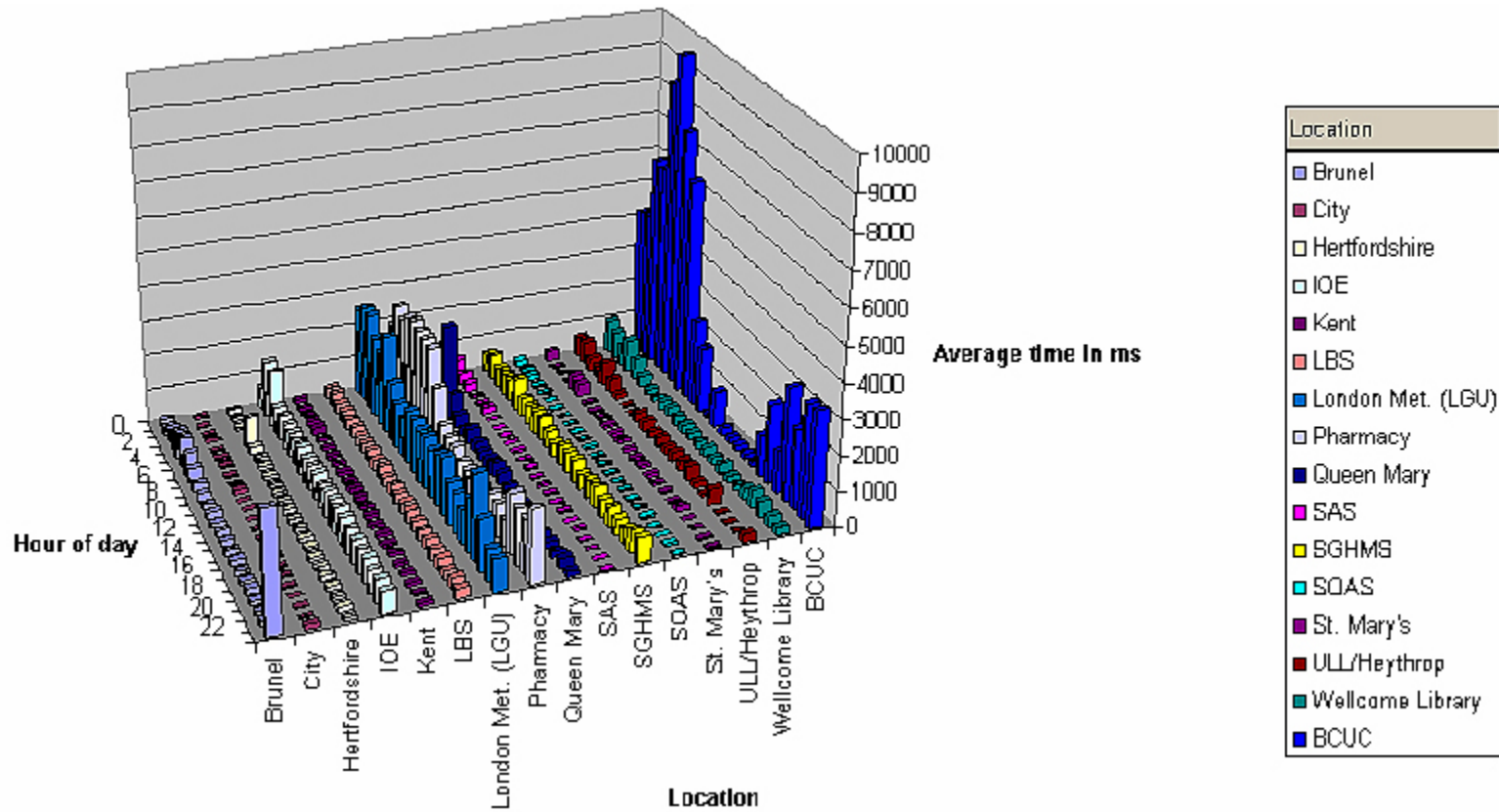


Figure 3: Average Hourly Response Times for Tested Z-Servers

Percentage of searches completing within a given time

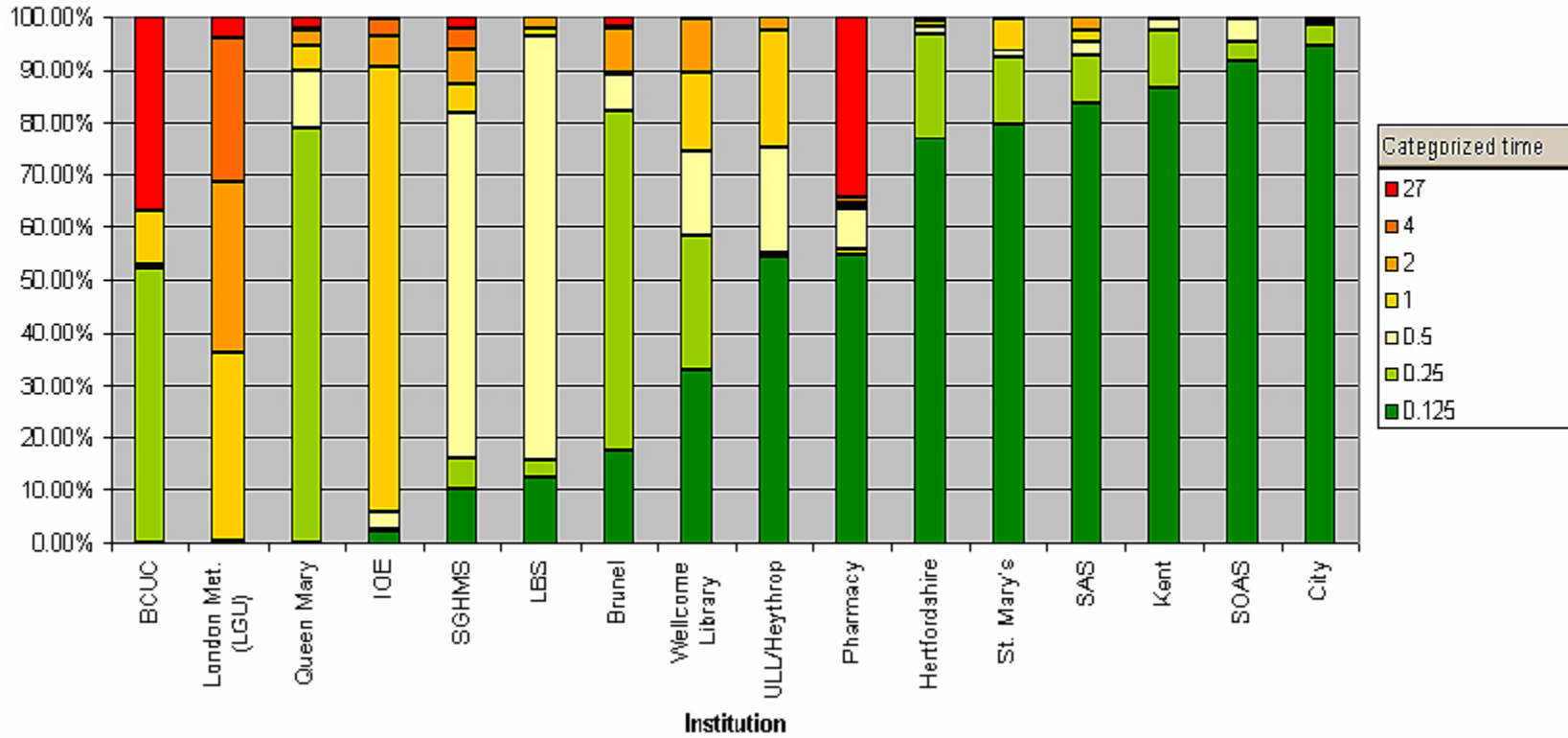


Figure 4: Percentage (%) of Searches per Z-server Responding in categorised time periods (seconds)

References

- Bath Profile Maintenance Agency. (2004). *The Bath Profile: An International Z39.50 Specification for Library Applications and Resource Discovery - Release 2.0*, Library and Archives of Canada, Ottawa. Available: <http://www.collectionscanada.ca/bath/tp-bath2-e.htm> (Visited 17 August 2004)
- Booth, H. & Hartley, R. J. (2004). *User behaviour in the Searching of Union Catalogues: An Investigation for Work Package C of CC-interop*, Centre for Digital Library Research, Glasgow. Available: <http://ccinterop.cdrl.strath.ac.uk/documents/finalreportWPC.pdf> (Visited 17 August 2004)
- Brack, V., Gilby, J., Gillis, H. & Hogg, M. (2001). Clumps Come Up Trumps, *Ariadne*, No.26. Available: <http://www.ariadne.ac.uk/issue26/clumps26/> (Available 17 August 2004)
- Crossnet Systems. (2004). *Crossnet Systems: DDiscovery*. Available: <http://www.crxnet.com/dscovery.php> (Visited 17 August 2004)
- Dempsey, L. & Russell, R. (1997). Clumps or ... Organised Access to Printed Scholarly Material: Outcomes from the Third MODELS Workshop, *Program: electronic library and information systems*, Vol. 31 No. 3, pp.239-249.
- Dietz, R. & Noerr, K. One-stop Searching Bridges the Digital Divide, *Information Today*, Vol. 21 No. 7, pp.24.
- Dovey, M. (2000). So You Want to Build a Union Catalogue?, *Ariadne*, No. 23. Available: <http://www.ariadne.ac.uk/issue23/dovey/intro.html> (Visited 15 August 2004)
- Dunsire, G. & Macgregor, G. (2003). Clumps and Collection Description in the Information Environment in the UK, with particular reference to Scotland, *Program: electronic library and information systems*, Vol. 37 No. 4, pp.218-225.
- Dunsire, G. & Macgregor, G. (2004). *Improving Interoperability in Distributed and Physical Union Catalogues Through Co-ordination of Cataloguing and Indexing Policies*, Centre for Digital Library Research. Available: <http://ccinterop.cdrl.strath.ac.uk/documents/CCICatInterop.pdf> (Visited 17 August 2004)
- East, J. W. (2003). Z39.50 and Personal Bibliographic Software, *Library Hi-Tech*, Vol. 21 No. 1, pp.34-43.
- Ex Libris. (2004). *Metalib: The Library Portal*. Available: <http://www.exlibrisgroup.com/metalib.htm> (Visited 17 August 2004)
- Gatenby, J. (2002). Aiming at Quality and Coverage Combined: Blending Physical and Virtual Union Catalogues, *Online Information Review*, Vol. 26 No. 5, pp.326-334.
- Gilby, J. & Sanders, A. (2003). *Transforming a Clump into a Z-target: A Feasibility Study*, Centre for Digital Library Research, Glasgow. Available: http://ccinterop.cdrl.strath.ac.uk/documents/Clump_ztarget_issue1_0.pdf (Visited 17 August 2004)
- Gilby, J., Sanders, A. & Cousins, S. (2004). *Bibliographic Union Catalogue Results and Display Issues*, Centre for Digital Library Research, Glasgow. Available: <http://ccinterop.cdrl.strath.ac.uk/documents/WPALastReportIssue1.pdf> (Visited 17 August 2004)
- Gilby, J. & Dunsire, G. (2004). *COPAC/Clumps Continuing Cooperation Project (CC-interop): Final Report*, Centre for Digital Library Research, Glasgow. Available: <http://ccinterop.cdrl.strath.ac.uk/documents/CCiFinalReportVersion1.pdf> (Visited 17 August 2004)

Hammer, S. & Andresen, L. (2002). *Issues in Z39.50 Parallel Searching*, Denmark's Electronic Research Library, Copenhagen. Available: http://www.deflink.dk/upload/doc_filer/doc_alle/895_Issues%20in%20Z39_50%20Paralle%20Searching.htm (Visited 17 August 2004)

JAFER. (2003). *JAFER Toolkit Website*, Oxford University Library Services, Oxford. Available: <http://www.jafer.org/> (Visited 17 August 2004)

Lynch, C, A. (1997). The Z39.50 Information Retrieval Standard - Part 1: A Strategic View of its Past, Present and Future, *D-Lib Magazine*, April 1997. Available: <http://www.dlib.org/dlib/april97/04lynch.html> (Visited 17 August 2004)

McDonald, D. (2003). *Web Services Technologies (WST) Report*, JISC, London. Available: http://www.jisc.ac.uk/uploaded_documents/tsw_03-04.pdf (Visited 17 August 2004)

Moen, W, E. (2001a). Improving Z39.50 Interoperability: Z39.50 Profiles and Testbeds for Library Applications, *In: Proceedings of the 67th International Federation of Library Associations Council and General Conference, Universal Dataflow and Telecommunications Workshop*, Aug. 16-25 2001, IFLA, The Hague. Available: <http://www.ifla.org/IV/ifla67/papers/050-203e.pdf> (Visited 17 August 2004)

Moen, W, E. (2001b). Resource Discovery Using Z39.50: Promise and Reality, *In: Sandberg-Fox, A, M., Proceedings of the Bicentennial Conference on Bibliographic Control in the New Millennium: Confronting the Challenge of Networked Resources and the Web*, Nov. 17-20 2000, Library of Congress Cataloging Directorate, Washington DC. Available: http://www.loc.gov/catdir/bibcontrol/moen_paper.html (Visited 17 August 2004)

Moen, W, E. & Murray, K, R. (2002). A Service Based Approach for Virtual Libraries / Designing and Demonstrating a Resource Discovery Service for the Library of Texas, *Texas Library Journal*, Vol.78 No.3, pp.4-14. Available: http://www.txla.org/pubs/tlj78/TLJ78_3.PDF (Visited 17 August 2004)

Moen, W, E. & Murray, K, R. (2003). *Z39.50 Server Implementation Issues and Recommendations*, Texas Center for Digital Knowledge, Denton. Available: http://www.unt.edu/zlot/Deliverables/WA4_del_z-server_v2_krm_23jun2003.doc (Visited 17 August 2004)

NISO. (2002). *Z39.50: A Primer on the Protocol*, NISO Press, Bethesda. Available: http://www.niso.org/standards/resources/Z3950_primer.pdf (Visited 18 August 2004)

Needleman, M. (2002). ZING - Z39.50 International: Next Generation, *Serials Review*, Vol. 28 No. 3, pp. 248-250.

Nicholas, D., Dobrowolski, T, Withey, R., Russell, C., Huntington, P. & Williams, P. (2003). Digital Information Consumers, Players and Purchasers: Information Seeking Behaviour in the New Digital Interactive Environment, *Aslib Proceedings: New Information Perspectives*, Vol.55 No.1/2, pp.23-31.

Nicholson, D., Denham, M., Dunsire, G. & Gillis, H. (2001). *CAIRNS Final Report: An Embryonic Cross-sectoral, Cross-domain National Networked Information Service for Scotland?*, Glasgow University, Glasgow. Available: <http://cairns.lib.gla.ac.uk/cairnsfinal.pdf> (Visited 17 August 2004)

Nicolaides, F. (2003). *A Comparative Study of the Performance of COPAC and Selected Independent Z39.50 Servers*, Centre for Digital Library Research, Glasgow. Available: http://ccinterop.cdlr.strath.ac.uk/documents/WPA_server_tests_issue1.pdf (Visited 17 August 2004)

Powell, A. (2004). *JISC Information Environment Architecture: Standards Framework. Version 1.1*, UKOLN, Bath. Available: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/standards/> (Visited 17 August 2004)

- PublicTechnology.net. (2004). *Research Libraries Network Promises UK Researchers Joined-up Services*, 30 July 2004. Available: <http://www.publictechnology.net/modules.php?op=modload&name=News&file=article&sid=1504&newlang=eng> (Visited 17 August 2004)
- RSLG. (2003). *Research Support Libraries Group: Final Report*, Higher Education Funding Council for England, London. Available: <http://www.rslg.ac.uk/final/final.pdf> (Visited 17 August 2004)
- Stubley, P. (1998). Clumping in the UK: Towards Virtual Union Catalogues, *New Library World*, Vol. 99 No. 1145, pp. 287-290.
- Stubley, P. , Bull, R. & Kidd, T. (2001). *Feasibility Study for a National Union Catalogue - Final Report*, University of Sheffield, Sheffield. Available: <http://www.uknuc.shef.ac.uk/NUCrep.pdf> (Visited 17 August 2004)
- Taylor, S. (2003), A Quick Guide to Z39.50, *Interlending & Document Supply*, Vol. 31 No. 1, pp.25-30.
- TZIG. (2003). *Z Texas Profile: A Z39.50 Specification for Library Systems Applications in Texas - Release 3.0*, Texas State Libraries and Archives Commission, Austin. Available: <http://www.tsl.state.tx.us/ld/projects/z3950/tzigprofilerelease30.html> (Visited 17 August 2004)
- van Veen, T. & Oldroyd, B. Search and Retrieval in the European Library: A New Approach, *D-Lib Magazine*, Vol. 10 No. 2. Available: <http://www.dlib.org/dlib/february04/vanveen/02vanveen.html> (Visited 17 August 2004)
- Whitelaw, A. & Joy, G. (2001). *Summative Evaluation of Phase 3 of the eLib Initiative: Final Report*, UKOLN, Bath. Available: <http://www.ukoln.ac.uk/services/elib/papers/other/summative-phase-3/elib-eval-main.pdf> (Visited 17 August 2004)
- Yu, Shien-Chiang & Chen, Ruey-Shun. (2003). Web Services: XML-based System Integrated Techniques, *Electronic Library*, Vol. 21, No. 4, pp.358-366.
- ZING. (2004). *ZING - Z39.50 International: Next Generation*, Library of Congress, Washington DC. Available: <http://www.loc.gov/z3950/agency/zing/zing-home.html> (Visited 17 August 2004)
- Z39.50. (2003). *Bib-1 Attribute Set*, Z39.50: International Standards Maintenance Agency, Library of Congress, Washington DC. Available: <http://www.loc.gov/z3950/agency/defns/bib1.html> (Visited 17 August 2004)
- Z39.50. (2004). *Z39.50: International Standards Maintenance Agency*, Library of Congress, Washington DC. Available: <http://www.loc.gov/z3950/agency/> (Visited 17 August 2004)