ASERL's Virtual Storage/Preservation Concept

John Burger, Paul M. Gherman, and Flo Wilson

One strength of research libraries' current print collections is in the redundancy built into the system whereby multiple copies of most print books are held by many libraries in the U.S. and Canada, let alone in the rest of the world. This redundancy gives us comfort that information will not be lost unless the last copy is discarded, lost or destroyed. Yet few librarians are aware of which books in their collection are unique or last copies, and therefore we do not take any special precautions to assure the safety of our unique holdings. We do place our valuable and rare books under special conditions to safeguard them, but do we really know which of our books in our general collection are unique?

On the other hand, we could be keeping too many books on our shelves, ones that we could easily withdraw, because they are of little value or interest to our readers and are held in a number of libraries. How high a rate of redundancy do we need of books that are rarely read and of questionable value? For instance, how many copies of Thomas Register do we need collectively on our shelves?

Are we storing too many copies across our libraries that collectively we could discard, if a few libraries promised to keep them safe and in preservation conditions?

This two-sided issue may have a solution if we could create virtual storage collections among those libraries that maintain separate storage collections. This paper hypothesizes that there is significant duplication across our storage collections and proposes a rather simple and cost effective solution to both issues of saving unique titles and eliminating excess copies from our collections by creating a virtual storage system. In this system, participating libraries with dedicated storage collections would pledge to retain titles in their storage collections and lend them as needed, so that many other libraries can discard the same low use, low value titles from their general collections.

History
In the late 1990's, the Southeastern Library Network (SOLINET) investigated the idea of establishing a

John Burger is Executive Director of the Association of Southeastern Research Libraries, based in Atlanta, Georgia, email: jburger@solinet.net. Paul Gherman is University Librarian at Vanderbilt University, Nashville, Tennessee, email: paul.gherman@vanderbilt.edu. Flo Wilson is Deputy University Librarian at Vanderbilt University, Nashville, Tennessee..

regional storage facility to be shared by its 2300 member libraries. In 1998 SOLINET hired Ralph Russell and a consulting firm to conduct a feasibility study; the concept was also discussed at several SOLINET Board meetings that year. In 2000, the Association of Southeastern Research Libraries (ASERL – a subset of SOLINET members) surveyed its member to determine the need for off-site storage among its members. At the time, on average ASERL members had need to store 93,000 volumes immediately, and expected to need room to store an average of 300,000 volumes each by 2005.

During the course of these discussions within SOLINET and ASERL, many libraries expressed a need and interest in additional storage space; however, there were significant reservations from state-supported libraries concerning sending their materials out of state. The idea of discarding multiple copies of deposited material met even greater resistance because of the impact on comparative statistics and rankings among libraries. (Few libraries really want fewer books!) Additionally, the up-front capital cost to build such a facility would have been prohibitive. Instead, SOLINET decided at the time to serve as a broker for storage space in existing storage facilities in the region at a discounted cost. WRLC had available space to lease to SOLINET, so arrangements were made to lease space to SOLINET members. Very few (if any) SOLINET libraries took advantage of the offer. Today, WRLC no longer has space available to lease.

During that SOLINET Board retreat in 1999, Paul Willis, then Dean of Libraries at University of Kentucky, suggested that a virtual storage concept might make more sense than building a shared regional facility. He believed that new technologies, linked catalogs and improved delivery systems made such a system possible. Today, ASERL operates a shared catalog (Kudzu) and a delivery system (Lanter) serving many of its member libraries, so the timing is now right to re-consider this concept.

More recently, at an ARL/OCLC Institute workshop on library architecture in Las Vegas in 2002, Don Kelsey, from the University of Minnesota, made the comment that very few libraries will ever weed their storage collections. This comment led to a discussion at the meeting about how a virtual storage collection could be created, based on the idea that storage collections will never be weeded.

Storage facilities are generally better designed and managed for longer book life than our general libraries. Books are less apt to be stolen because our storage facilities have limited public access. The books are shifted less often mainly because they are shelved by accession date to the facility instead of call number. And the environmental conditions can be better adjusted to lengthen book life. Most shelving facilities have few windows so less ultraviolet light, lower light levels in general, and lower temperatures, since they are not used for study, and many are humidity controlled. A recent quick poll of ARL directors shows that there are about 50 storage facilities in operation in the U.S. with a combined collection in the range of 20 million volumes.

In the summer of 2003, the OCLC Office of Research conducted a study of unique items in WorldCat held in the Vanderbilt University Library. The results showed that Vanderbilt held approximately 23,000 unique titles. Of these about half were dissertations and theses leaving about 12,000 items that WorldCat indicated only Vanderbilt had in its collection. Many were rare books that were already housed in Special Collections, but a significant number of other materials were on the open shelves. Having this information has prompted Vanderbilt to consider identifying these titles in their bibliographic record and to begin examining each title for preservation purposes.

Overview of ASERL's Proposed System

The concept is to create a national system of virtual storage collections by developing inventories of current monographic storage collections, to discover the level of duplication in existing collections, so that libraries can make informed decisions about whether or not to move their copies of monographs to storage or discard them. (Monographs were selected as a starting point because they are more easily compared bibliographically than serials. However, it is understood that the proposed virtual storage system could work as well for serials, and indeed might even yield greater return on effort; however, they are more of a problem bibliographically.) The general principles of participation will be as follows; participating research libraries with collections in storage would pledge to:

- not weed their stored collections,
- maintain acceptable preservation-quality environmental conditions for their stored collections,

• lend material from their stored collections to other libraries via standard ILL processes,

• generate an inventory file of their storage holdings to be sent to OCLC.

OCLC would indicate on the WorldCat bibliographic record that a specific volume has been pledged to be preserved by x-number of libraries, and that these titles are available for loan. As participating libraries add new material to storage, they would need to periodically report these volumes to OCLC to be included in the national/regional preservation collection.

Such a system would allow other libraries to consider discarding an item known to be held in the national/regional preservation collections, thereby avoiding sending yet another copy of the item into storage. Each library could arrive at its own comfort level about how many copies need to be held in the national/regional preservation collection before it discarded its copy of the item. Initially no effort would be made to determine the condition of each pledged book that is retained, although this additional step could be added to the plan at a later date.

This plan would therefore accomplish two major objectives:

1) Assure that a number of copies of specific titles will be retained by some libraries, so that libraries do not inadvertently discard the last extant copy of a text,

2) Allow other libraries to discard their copies of these titles, thereby reducing the overall cost of maintaining low-use collections.

The Association of Research Libraries has recently agreed that member libraries can count volumes that they no longer house in their collections that they jointly own in a cooperative storage facility. We envision the next step to be counting those volumes a library has virtually, in partnership with a regional virtual storage collection such as the one we propose here.

Proof of Concept

To gain insight into the level of overlap or duplication in storage collections in research libraries, in the summer of 2004, nine ASERL members with storage collections containing more than 100,000 volumes each sent their holding information about those storage collections to OCLC for comparison. The nine libraries included Duke, Tulane, University of Alabama, University of Georgia, University of North Carolina, University of South Carolina, University of Virginia, Vanderbilt University, and Virginia Tech. The resulting data (Table 1 and Appendix 1) showed combined holdings of 2.3 million monographs and 1,783,854 individual titles in the nine storage collections (Figure 1 and Appendix 2). Of those, 1,426,825 titles are held by only one library; in other words, 80% of all storage holdings were unique. The number of unique titles proved to be much higher than was hypothesized. We postulated that most libraries' comfort levels, before being willing to discard a title held in other storage facilities, would be three copies in a regional storage facility. There were 117,650 distinct titles, or 7% of the total, held by three or more of the nine participating libraries. If four copies were required to reach a library's comfort level for discard, then only 43,097 titles, or 2%, were contained in these nine storage collections.

These finding are somewhat discouraging at first glance, since the level of duplication in the nine storage collections is rather low. However, it should be remembered that it is not the goal of this proposal to eliminate books stored in storage collections. Indeed one underlying principle is that no titles will be eliminated from existing storage collections. Rather the

Table 1. ASERL Storage Collections Overlap Study		
	Total #	Percentage
Records for stored monographs submitted by nine libraries	2,322,202	
Records submitted where 3 or more libraries hold the title in storage	417,619	18% of total records
Monograph titles in storage across the nine libraries (duplicates eliminated)	1,783,854	
Monograph titles in only one storage collection	1,426,825	80% of total titles
Monograph titles held in 3 or more storage collections	117,650	7% of total titles
Titles held in 4 or more storage collections	43,097	2% of total titles
Copies needed to insure preservation of 3 or fewer copies	2,257,533	

benefit of the proposal is that books currently housed in library's main collections can be discarded rather than deposited in their storage collection.

A next step that we have yet to complete at the time of the writing of this paper, is to match Vanderbilt's main collection against the combined holdings of eight of the storage collections in ASERL. Such a move would allow better estimation of the potential value to be gained by discarding titles rather than transferring them to storage. Vanderbilt's collection falls in the middle of ARL libraries in terms of size, and our collection is typical of a liberal arts collection with professional schools; we believe this will yield a good sample of how many copies might be discarded by matching against a regional virtual storage collection. Based on the overlap study data, there could be up to 640,000 copies of the 117,650 storage titles held in the main collections of the nine participating libraries that could be discarded and still leave three copies remaining in the regional virtual storage collection. However, there are 42 members in ASERL of varying sizes, including a number of state libraries, so the number of copies in the consortium to be possibly discarded from their main collections is much greater. And libraries who are not even members of ASERL but are located in the Southeastern United States could take advantage of those books housed in ASERL storage repositories.

More importantly, we envision a system of similar virtual libraries being established across the country by a number of consortia; this would permit any library within their region could discard books with the assurance that copies were being held in near preservation conditions, yet available to them for loan. The likelihood would be good that millions of volumes might be discards from libraries without a loss of accessibility.
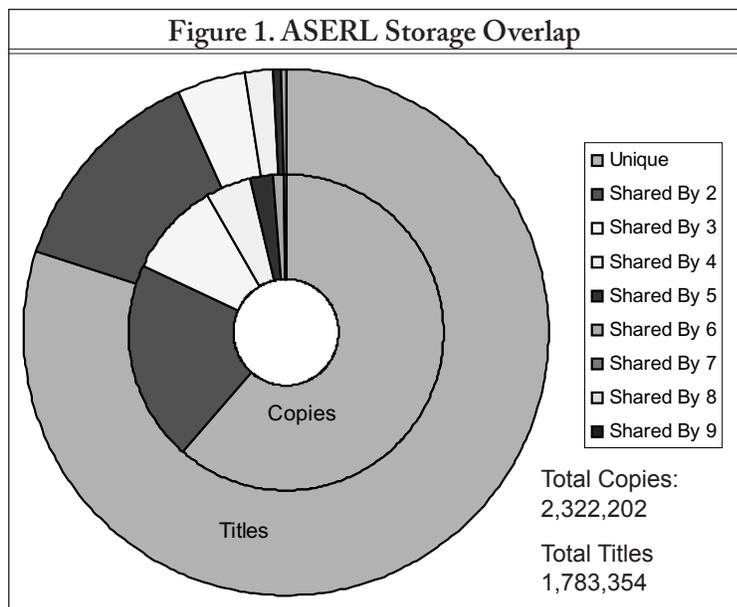
OCLC's Role

OCLC is key to creating a national virtual storage system. First each participating library with a storage collection will need to send OCLC an inventory of the monographs stored in their storage collection, with updates as additional titles and copies are added. OCLC will then need to indicate that a storage copy is pledged to be retained in the Digital Registry 583 field on the MARC record. Any other library can then check the MARC record for that title and determine if it is already pledged to be retained. If so, they can discard it if they determine it is not a high use item and that there are sufficient copies in their region or held nationally to satisfy their comfort level.

Using the storage collections data, OCLC' s new OCLC WorldCat Collection Analysis service could provide a report that will facilitate the weeding of library collections of material already pledged by host storage libraries. This report could compare each library's holdings against the regional or national storage collection based on the comfort level of each library and produce a pull list of titles and copies to be considered for discard. The library could set parameters by for the number of copies that must be in storage collections either regionally or nationally before they will weed a title from their collection. This service could eliminate the need to do title by title checks by library staff, and the report could be run against individual library circulation records on their local system to further refine the pull list.

Another important report that OCLC could generate for libraries would be a listing of their unique holdings in WorldCat. These materials can then be safeguarded by the host library. Such listings could also help establish important priorities for digitization projects.

Preservation

Although a national virtual storage system can not be called a pure preservation system, it is a

Figure 1. ASERL Storage Overlap

Legend:
□ Unique
■ Shared By 2
□ Shared By 3
□ Shared By 4
■ Shared By 5
■ Shared By 6
■ Shared By 7
□ Shared By 8
■ Shared By 9

Copies

Titles

Total Copies: 2,322,202

Total Titles 1,783,354

step toward developing a "last copy protection system." Once OCLC has an inventory of those titles held in storage collections, libraries could then begin examining the condition of their storage copies according to an agreed upon system of rating. The highest rated copy could then be designated a non-circulating copy, while the other storage copies could be circulating. This process would be time-consuming and expensive, and of course it will not account for those copies held in main collections. But the first step of simply inventorying the collective holdings of our storage collections is easy and inexpensive to accomplish, and it sets the library community on the path to a more comprehensive preservation strategy.

Conclusion

By implementing a regional or even national virtual storage system, the library community can reduce the size of our collections by weeding multiple copies of little used and low value material from our shelves while at the same time assuring that last copies will not be discarded inadvertently. The system is relatively easy to implement and low cost for the gains possible. Only about 50 libraries operating storage facilities need to work with OCLC to implement the system, yet virtually every library that cares to weed their collections on the basis of the virtual storage system can participate in the savings. It is not a perfect system, but it offers significant savings and safeguards for very little investment.

Appendix 1. ASERL Monograph Storage Overlap Study

ACAS Results

Storage Titles Submitted

Library	Unique	Shared By 2	Shared By 3	Shared By 4	Shared By 5	Shared By 6	Shared By 7	Shared By 8	Shared By 9	Total	3 or fewer	4 or more
Duke University	439,642	103,639	40,200	17,868	7,951	3,052	860	164	15	613,391		
Tulane University	98,759	46,024	25,085	13,191	6,518	2,738	855	166	15	193,351		
University of Alabama	34,235	19,795	13,911	9,191	5,247	2,432	788	161	15	85,775		
University of Georgia	109,210	41,416	19,903	10,116	5,105	2,096	721	144	15	188,726		
University of North Carolina	164,361	60,835	27,390	12,849	6,100	2,450	735	151	15	274,886		
University of South Carolina	79,188	42,740	24,728	13,232	6,376	2,645	784	150	15	169,858		
University of Virginia	274,303	69,617	26,819	12,266	5,555	2,287	731	149	15	391,742		
Vanderbilt University	117,495	56,544	28,293	14,032	6,701	2,681	812	165	15	226,738		
Virginia Tech University	109,632	37,148	17,330	7,931	3,642	1,429	490	118	15	177,735		
Total copies for all	1,426,825	477,758	223,659		53,195	21,810	6,776	1,368	135	2,322,202	2,128,242	193,960
											92%	8%
Actual titles	1,426,825	238,879	74,553	27,669	10,639	3,635	968	171	15	1,783,354		
Titles held by 3 or more										117,650	7% of actual titles	
Titles held by more than 3										43,097	2% of actual titles	

Copies

Library	Unique	Shared By 2	Shared By 3	Shared By 4	Shared By 5	Shared By 6	Shared By 7	Shared By 8	Shared By 9	Total	3 or fewer	4 or more
Preserve in aggregate	1,426,825	477,758	223,659	83,007	31,917	10,905	2,904	513	45	2,257,533	97% of copies	
Discard in aggregate	-	-	-	27,669	21,278	10,905	3,872	855	90	64,669	3% of copies	
										2,322,202		

Potential discards if we have them in non-storage collections (total actual copies less those held in storage)

Library	Unique	Shared By 2	Shared By 3	Shared By 4	Shared By 5	Shared By 6	Shared By 7	Shared By 8	Shared By 9	Total	3 or fewer	4 or more
Duke University			34,353	9,801	2,688	583	108	7	-	47,540		
Tulane University			49,468	14,478	4,121	897	113	5	-	69,082		
University of Alabama			60,642	18,478	5,392	1,203	180	10	-	85,905		
University of Georgia			54,650	17,553	5,534	1,539	247	27	-	79,550		
University of North Carolina			47,163	14,820	4,539	1,185	233	20	-	67,960		
University of South Carolina			49,825	14,437	4,263	990	184	21	-	69,720		
University of Virginia			47,734	15,403	5,084	1,348	237	22	-	69,828		

Appendix 1. ASERL Monograph Storage Overlap Study												
ACAS Results												
Storage Titles Submitted												
Library	Unique	Shared By 2	Shared By 3	Shared By 4	Shared By 5	Shared By 6	Shared By 7	Shared By 8	Shared By 9	Total	3 or fewer	4 or more
Vanderbilt University			46,260	13,637	3,938	954	156	6	–	64,951		
Virginia Tech University			57,223	19,738	6,997	2,206	478	53	–	86,695		
Vanderbilt, e.g.												
Preserve	117,495	56,544	28,293							202,332		
Shared preservation decision				10,524	4,021	1,341	348	62	5	16,301		
Discards				3,508	2,680	1,340	464	103	10	8,105	4% of VU titles	
										226,738		

Appendix 2. Overlap Proportions for Each Library

Legend:
- Shared By 9
- Shared By 8
- Shared By 7
- Shared By 6
- Shared By 5
- Shared By 4
- Shared By 3
- Shared By 2
- Unique

Libraries: Duke University, Tulane University, University of Alabama, University of Georgia, University of North Carolina, University of South Carolina, University of Virginia, Vanderbilt University, Virginia Tech University