

Selecting An Automated Library System for Finnish Research Libraries, Linnea2

Annu Jauhainen

Helsinki University Library Finland

annu.jauhainen@helsinki.fi

Abstract: Finnish university libraries have chosen a new integrated library system. This is continuation to the co-operation that started in the late 80's when these libraries got their first automated library system. The co-operation is now expanded, when the databases of all university libraries as well as the national and union databases will be assembled on a single server, which is outsourced. This means significant savings, a long with technical improvements, both in the purchase price of software and hardware as well as in the maintenance costs.

The Present Situation

All Finnish academic libraries and a number of other Finnish research libraries have used the VTLS software during the 90's. The contract with VTLS Inc. was signed in 1988 and implementation took place during the following years. A unified network called Linnea was created, consisting of the local installations and a common physical union catalogue which all were connected by the powerful academic data transmission network FUNET.

The VTLS-based network, now called Linnea1, was very advanced when built a decade ago, and it has served Finnish libraries well. VTLS Inc. has also been a trusty companion of Finnish academic libraries during these ten years.

Creation of the union catalogue Linda in early 90's was an ambitious project. Not only was data from all academic libraries loaded into a single database; software development was also needed. For example, a duplicate control algorithm was designed in Finland and implemented by VTLS. VTLS developed many unique consortium features which enabled the libraries to use the Linda database efficiently for copy cataloguing purposes. Depending on the library, 50-90% of MARC records can be copied. ILL localisation is also very efficient, because Linda contains summary-level serials holdings from about 400 Finnish libraries.

The Automation Unit of Finnish Research Libraries, created in the Ministry of Education in 1974, was instrumental in the implementation, development and running of the Linnea network. In 1993 the Unit, with all its tasks and resources, was moved to the National Library, where the Division of Library Network Services is now managing the Linnea1 network, functioning as a common agency for the academic libraries. In this capacity the National Library is also responsible for the new steps toward Linnea2, as the next generation network is called.

Selection of a New Automation System

To summarize the need for a new generation software we can say that all library system vendors are building so-called third generation library systems with relational database and Client/Server technology, graphical user interface and web gateways, the ability to search multiple databases simultaneously, multimedia support and support for internationally accepted standards such as Z39.50, Unicode, Edifact and ISO ILL, to meet the growing needs of the users. It was also evident that the classic VTLS system was coming to the end of its life-cycle and would not be developed further since VTLS Inc. is concentrating on their new system, which is called

Virtua.

The Finnish academic libraries have since early 90s enjoyed the benefits of being a consortium. The ten years of VTLS use have taught the libraries and all parties involved that co-operation is power, even if it is not always easy or simple.

Because of the great success of Linnea1, there was no need to revise the basic service philosophy when moving to a new system. Libraries were satisfied with the system and the work flows and with co-operation with one another.

When the present VTLS system was purchased, the Ministry of Education funded the acquisition of both software and hardware. This time the universities had to find the money out of their own budgets. Nevertheless, both the universities and their libraries wanted to ensure the benefits of the present common approach. Libraries also were open to totally new technical and organizational solutions if they should prove more favourable both functionally and economically. Libraries clearly wanted to avoid transplanting old patterns into a totally new environment. Everything had, therefore, to be looked at from a new perspective.

Three major issues had to be tackled: the selection of the software, the future database or network architecture and the maintenance of the hardware.

Selection of the Software

The Linnea libraries started to look for a new-generation library system about four years ago. The National Library was asked by the directors of the Finnish academic libraries to survey the systems either on the market or being developed at the time. A questionnaire was compiled and sent to the vendors who had recently been shortlisted in corresponding procurements in Europe or in the U.S. The vendors were asked about their database management system, database structure, standards, various functions and features, the user interface, languages and formats, training, support, prices and future plans.

Procuring a new library system for a large network is a major project which is regulated by European Union statutes. When the value of the contract exceeds the threshold, which is 200,000 euros at the moment, the procurement has to be advertised across the European Union. Of the three alternative types, the restricted procedure seemed to be the most suitable for the Linnea2 project.

When VTLS was selected in the late 80's, the selection process was handled by the Automation Unit of Finnish Research Libraries alone, without much involvement from the libraries themselves. This approach was quite natural at the time, because there was little experience of library automation in the libraries. More than ten years after, the situation was completely different. Libraries were well acquainted with at least one library system and, most importantly, they knew what their needs were and what they wanted of the new system. The resources of the libraries were welcomed by the National Library, which, as the service facility of the academic libraries, had the task of coordinating the process and pulling everything together.

The procedure started officially in April 1998 and the tenders were received in July. At this point, tenders were invited for software only, another procurement was planned for the hardware once the software had been chosen. During the fall the tenders were

evaluated thoroughly. Attention was paid to the technical structure and the technical solution of the system, references from present and future users of the system, the services and the support offered by the vendor and the quality and the completion of the various functions and modules. Four systems were shortlisted based on these criteria. They were Horizon, Innopac, Taos and Voyager. These four systems had been found to fulfil our requirements best in the first phase of the selection process.

At the beginning of the second phase the four short-listed systems were all on the same line. In nine months we had to find out which of the four was functionally the most suitable and economically the most advantageous for the local databases as well as the union and national databases. The systems were first demonstrated to a large group of library representatives. The next step was to get our hands on the applications. The National Library, together with the four vendors, organized the testing of these systems. This was the part of the evaluation in which the contribution of the libraries was most significant. Over 70 people from the libraries and the computing centres of the universities participated in testing, which took about three months. A number of testing groups, each specializing in different functions, i.e. cataloguing, circulation, acquisition, OPAC, etc. listed the merits of the systems, without knowing how the other groups ranked them. Objectivity was the main guideline here. In addition to the ranking list, the groups also produced lists of open questions. Answers to these questions were sought in two ways, through site visits and negotiations with the vendors. A group of six people, representing both the National Library and other libraries as well as the university computing centres, visited libraries using these systems, both in Europe and in the U.S. The site visits were essential in finding out how the systems worked in real life.

During these nine months of evaluation the National Library negotiated with the four vendors (Dynix GmbH, Innovative Interfaces Inc, Data Research Associates Inc and Endeavor Information Systems Inc) in several ways and on several occasions. The vendors came to Helsinki a number of times and we went to their headquarters once to talk with the development staff, support staff and the company management. There was also constant discussion via email whenever any questions about the functionality of the systems needed to be answered.

An essential feature in selection processes was a fair and objective treatment of all parties involved. Since every step was documented, we would have been able to reconstruct the process, should it have proved necessary.

We have been told both by many foreign colleagues and by the vendors that the Finnish library system selection process has been the most thorough ever carried out. It is clear that when purchasing a system for all major research libraries of a country we are dealing with a much more serious issue than satisfying the needs of just a single library.

When the different parts of the selection process were drawn together, Voyager, by Endeavor Information Systems Inc. proved to fulfil the criteria best. Voyager was found to be a complete, integrated system that was finished in the essential, traditional functions needed by the libraries, but which however is being further developed to meet the new needs and changing technologies. It fits both individual Linnea libraries and the Linnea network well. Local services can be streamlined and their scope extended. But centralised services will also benefit from Voyager via its consortium-driven functions. Increased efficiency is largely based on improved networking since Voyager supports both Z39.50 and ISO ILL.

The company, Endeavor Information Systems Inc. had also been thoroughly investigated by an economic expert and found to be sound and stable, with good prospects. An example of the difficulties in anticipating future changes is that Endeavor has since then been sold to Elsevier Science, raising a number of question marks.

The National Library proposed to the libraries that Voyager should be chosen, which was unanimously accepted. The National Library was asked to conclude the negotiations with the company, and was also empowered to sign the contract on behalf of all

universities and other bodies participating in the purchase. This happened on February 4, 2000.

The Network Architecture

One of the important decisions in Linnea2 was whether to merge existing databases or to keep the current structure. Discussions with Endeavor experts made it clear that although it is technically possible to merge databases, actually doing this would be time-consuming and expensive. The technical merits of such action would be limited, since Voyager databases can be merged into a virtual union catalogue by using the Z39.50 Information Retrieval protocol. Politically there was quite a lot of reluctance among libraries to merge databases, even though Voyager makes living with a shared database much easier than our present system. A decision was, therefore, made to retain the 24 databases in Linnea2.

The next question was how many servers an optimal solution for the Linnea2 network would require. In the present Linnea1 network there are 17 HP3000 servers for the 24 databases. The number of servers was never really discussed during the implementation of Linnea1 because of the limitations of the computer technology of the time.

How far can one go in centralisation? The answer depends on three factors, the available data transmission network, the capabilities of the software and the state of the computer technology.

The Finnish Academic and Research Network, FUNET, is already at present a key factor for the Linnea network. Without the infrastructure provided by FUNET it would not have been possible to use the Union Catalogue Linda as a cataloguing tool in a way we have done since the early 90's.

A shared server is not possible if there can only be one database on the server. The Voyager software allows an unlimited number of databases on a single server. However, practical experience from other Voyager consortia made it clear that there should not be more than about 5-7 databases on a single server, since a large number of databases means that much time may be needed for Oracle and Voyager updates: it may take several days to update many large databases, and during the process all the databases must be shut down.

More importantly, if all databases are dependent on the same hardware and operating system process, severe problems would have an impact on every library simultaneously. Fortunately, new server technologies make it possible to have a single server and still avoid this problem: there are servers that can be internally split into several logical (and physical) parts.

Both Sun and IBM, which are the platforms Voyager supports, can deliver cluster-like computers, which can be separated into logical parts called domains (Sun) or nodes (IBM). Each part has its own operating system process and dedicated hardware from network card to processors. To the operators and users, the server looks like a cluster of computers.

So there were no technical constraints on choosing the network architecture freely. Linnea libraries were eager to find out whether centralisation would save money. In the 90's the resources and budgets of the Finnish academic libraries have been cut; this is unfortunately a problem common to all kinds of libraries everywhere in the world.

At the request of the universities three scenarios were analysed:

- centralised model; all databases placed on a single machine
- semi-centralised model; 3-5 servers
- decentralised model; the current number of servers

Cost analysis was based on both purchase price and the total cost of ownership, calculated for five years.

After a thorough analysis of the various options, Sun E10000 was chosen as the server system. The decision to go for Sun was based on technical merit and price. Both Endeavor and Oracle use Sun machines as their development platforms; this fact was also taken

into account.

The Linnea2 server will initially have 28 400 MHz CPUs. According to Endeavor, this is enough for 1400 active users, or more than 5000 concurrent users, about twice as much as now. Both Endeavor and we felt that an ample safety margin is needed in order to avoid performance problems.

Of course buying a lot of CPUs is not enough; there may be other bottlenecks. The E10000 will have 24 GB of memory and 800 gigabytes of mirrored fiber disk dedicated to Voyager databases.

The universities had set an upper limit for the total purchase price of the software and hardware, including conversion of the databases. Because of the unfavourable exchange rate of the US dollar, the National Library felt increasing pressure to arrive at a low-price solution. We found out that even if list prices may tell you a different story, for a big customer like our consortium it was cheaper to purchase one big server system than a number of smaller ones. But bargain prices are not automatically offered. We managed to establish a competition between Sun and IBM in real terms because both companies saw Linnea2 as an important project.

After the server was chosen, the decision was made to outsource the maintenance of the new server to the Center for Scientific Computing, CSC, a non-profit company owned by the Ministry of Education. It hosts Finnish supercomputers and maintains the FUNET network.

In spite of better maintenance coverage and better support from the hardware vendor, maintenance costs will diminish a lot compared with Linnea1. Basic maintenance of the 17 HP3000 servers takes about three man-years, but we estimate that a single E10000 will require less than a man-year. If this estimate is correct, we will save about two man-years or even more because managing a UNIX system is generally believed to be more time-consuming than managing an HP3000 computer.

Thus we have good evidence for the claim that an unprejudiced approach to server architecture has enabled us to combine significant savings with important technical improvements. Being a consortium helps a lot: libraries buying systems only for themselves will not be able to utilise new technology with similar efficiency. It is easy to understand from this point of view why library consortia are becoming more common in the US and some European countries. Finland has been one of the pioneering countries in this area, and our experiences from such co-operation are very encouraging.

Implementation

At present we are in the middle of the implementation phase. Building Linnea1 and implementing VTLS took several years, but this time all 24 databases will migrate from VTLS to Voyager during a fairly short period of time, April-August 2001. This means that everything has to be scheduled very carefully and the schedules have to be kept. We have a joint national implementation project, and each library has its own project. There are three parties in all of these projects: the Linnea libraries, the core group in the National Library and Endeavor Information Systems Inc. and all of these

parties have to work together seamlessly.

Endeavor is doing some software development for us. In general we are buying the system off the shelf and didn't want as many customizations as in the VTLS time, for we have seen the problems raising from localization, but there are some things that could not be avoided.

Training is a vital part of implementation. We use the "train the trainer" method, so that Endeavor is training only the trainers. This way we get customized training for Finnish local needs, and also save quite a lot of money.

Endeavor has converted several VTLS databases before, but in spite of that, testing the loads is important. Early tests for some sites were carried through in the fall and at present we are doing test loads for all databases, to make sure that the production conversions will be successful.

Conclusion

The cornerstone of this process has been co-operation, the will to pull together. This is not enough nevertheless: there also has to be a workhorse, to pull everything together. This is important, especially when there is no higher authority to manage the process, as was the case when Linnea1 was built and the Ministry of Education took care of the negotiations and funded the whole process. This time university libraries felt the need to start the process of acquiring a new system together. They were willing to make an effort to find a new solution to improve the quality of their services, as well as to use their scarce resources for the evaluation, which was seen as benefiting all. They were also willing to find the money to pay for the new system, with everything included.

The Linnea2 consortium was build from below, the National Library acting as the workhorse but not as a higher authority. This was a successful approach. In order to continue this success, there must be a formal organisation for the consortium. That is why the Linnea2 consortium has just been established, with a formal organisational structure and bylaws.

The thorough selection process for a new automation system for the Finnish research libraries has not simply been a question of technology and technical expertise, which the National Library has been responsible for. It was even more a question of policy and co-operation. Many things may be possible technically, but politically they are not, unless you know how to handle them correctly and diplomatically.

Sometimes our neighbours in the Scandinavian countries say that libraries in Finland find it easier to co-operate than libraries in other countries. Of course, this is not true. Libraries in Finland are as individualistic as libraries everywhere. They also have their particular local needs. But there is obviously a will to co-operate, as difficult as it may be at times.