Data Warehousing Puts New Life into Legacy System
Kathy L. Fisher, Information Services & Telecommunications
Ferris State University
Big Rapids
Michigan

The major focus of this presentation is to explain how Ferris State University (FSU) is developing a data warehouse application to deliver timely, accurate student information to college-specific customers. This presentation will benefit anyone who wishes to serve academic customers by providing an PC database environment for working with legacy data.

Two underlying factors led to the development of this application. A more timely method for delivering reports to college staff is needed at a time when Information Services (IS) staff resources are diminished. The mainframe-based Student Information System presents information on an individual student basis; whereas, colleges often need to view student information in an aggregate form by college, department, course, or advisor. This application is designed to empower the customer to design and create reports and queries based on their individual needs with the ultimate outcome being an improved service to students.
PROJECT OVERVIEW

In recent years, Ferris State University (FSU) has experienced the re-engineering of many procedures that were once a common practice in daily office routines. In the fall of 1993, the University converted all academic programs to a semester calendar which necessitated the implementation of the current Student Information System (SIS). This change in technology brought about a new way of doing “business” across campus. A large majority of the batch reports previously produced by the Information Services department were no longer valid or usable. Hundreds of new on-line screens were introduced causing office staff to revise their methods for accessing student information.

At the time the University was learning to use this new system, it was also experiencing a decline in student enrollment. This led to financial constraints and reduced staffing in many areas, including both the Information Services department and the college administrative offices. At a time when the University needs to improve student services to remain competitive, it is facing a critical challenge in meeting that need. College administrative staff need to be able to easily access and summarize student information in order to serve students in a timely manner.

In an effort to re-engineer the process by which college administrators and faculty acquire information, Ferris State University is developing a data warehouse approach to working with legacy SIS data. Data delivery tools placed on client workstations allow staff to extract data from the warehouse for local data analysis and reporting.

RATIONALE FOR IMPLEMENTATION

Two underlying factors led to the development of this application. A more timely method for delivering reports to college staff is needed at a time when Information Services (IS) staff resources are diminished. The mainframe-based Student Information System presents information on an individual student basis; whereas, colleges often need to view student information in an aggregate form by college, department, course, or advisor. The data warehouse application is designed to empower the customers to design and create reports and queries based on their individual needs with the ultimate outcome being an improved service to students.

The traditional approach to accessing student information takes either of two routes: 1) a request for information is sent to the IS department, describing the data needed and the format in which it is to be printed, or 2) the data is viewed on-line through the use of CICS screens. There are limitations to both approaches that hinder timely access to information.

Limitations in the Delivery of Information

Several factors lead to a less-than-satisfactory turn around time for delivering information to the department or person requesting it. One factor is the limited number of programmers in the IS department. There are currently six full-time programmers
available to respond to incoming report requests from over 16 functional departments, seven college administrative offices and various faculty and management personnel. In addition, fulfilling reporting needs constitutes only a portion of their workload. These report requests must be prioritized along with other application development, maintenance, and support tasks. This workload delays the delivery of information for days or sometimes weeks. If users need to query the IS department for assistance every time they require access to critical data, vital decision-making time is consumed and opportunities can be lost.

A second factor that often leads to frustration for the user/customer is the process of printing and delivering the needed report. The majority of reports currently generated are printed centrally through the Computer Operations department and delivered by the campus mail service. This type of delivery adds at least another day to the process. Recent implementation of an on-line Report Distribution System has alleviated some of this frustration. However, most users continue to request the information in hard copy form.

One last item that needs to be discussed is the effect budget limitations has had on this traditional approach. With the recent restructuring of academic programs and the drop in previous levels of enrollment, the university has had a difficult time in maintaining staffing levels -- not only in the IS department, but across campus. As other departments become lean and need to “do more with less”, they are turning to the already downsized IS staff for technological solutions to enhance the effectiveness of their workflow. This puts an additional strain on IS resources to meet user needs for timely information.

**Limitations of Mainframe Transaction-Processing System**

In 1993, the university purchased and implemented a mainframe-based Student Information System (SIS). This system offered the ability to process semester-based data and added greater functionality than our previous student record system which had been in use for over 15 years. While this new system meets the needs for processing admission, financial aid, academic program and student account information, it lacks the capabilities of providing information that is readily available or useful for management decision-making activities.

This SIS system is a basic transaction-processing system utilizing CICS screens for data entry and verification. The application code is written in traditional COBOL language which is used to process data stored in VSAM file structures. This structure prohibits easy access for the average user and requires programmer intervention to turn this data into useful information. In addition, all on-line screens are based on individual student name or identification number. This system works fine when specific student data is needed but does not provide executive information capabilities for summarizing the data into easy-to-digest formats, eye-catching graphics, or analytical trend data.

**The Data Warehouse Approach**

The initiative for the data warehouse approach resulted from several circumstances coming together at the right time. As an IS manager, responsible for application systems and
development, I was well aware of the current limitations of our present Student Information System and the ineffective process for delivering timely information. At the same time, I was enrolled in FSU’s graduate program for Information Systems Management and was contemplating ideas for a thesis project. The two situations converged during a meeting with the Assistant Deans. Representatives from the Information Services department, including myself, were presenting a PC-database approach for accessing university financial data. This meeting evolved into a discussion of student information reporting needs with the Assistant Deans expressing their desire for a similar PC-based system for accessing student data. Considering the current backlog of requests in the IS department, it was decided that the best approach would be to develop this student data warehouse as my graduate program thesis project. Development activity could occur after normal work hours and would not interfere with other priorities in the IS department.

**Implementation Highlights**

The main objective of the thesis project is to define and develop a student information data warehouse on a SQL server that will be easily accessible through the use of client/server technology and PC-based reporting tools (e.g. MS-ACCESS). These features will provide selected end-users in the college offices with the capability to create their own reports, perform search and sort functions, produce mail-merge documents, and view data on-line in an Windows environment -- all from their individual client workstations. Thus, end-users will be empowered with effective management tools for planning and decision making.

In the initial planning phase, several discussions with the Assistant Deans from each college demonstrated the need to have multiple types of information relating to student data (e.g. recruiting and admission data, degree requirement information, and alumni relationships). However, after considering the resources available and the timeframe in which a usable application was needed, it was determined that a subset of this information would be used as a pilot project. Therefore, the initial data warehouse will be populated with data pertaining to currently enrolled students only. An eight-step implementation plan was developed for creating the data warehouse application (see Table 1).

<table>
<thead>
<tr>
<th>Step</th>
<th>Activity</th>
<th>Target Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Determine immediate data needs for development of warehouse database.</td>
<td>April, 1995</td>
</tr>
<tr>
<td>2</td>
<td>Evaluate and select data warehouse platform.</td>
<td>June, 1995</td>
</tr>
<tr>
<td>3</td>
<td>Select PC-based reporting tools.</td>
<td>July, 1995</td>
</tr>
<tr>
<td>4</td>
<td>Determine workstation hardware requirements.</td>
<td>October, 1995</td>
</tr>
<tr>
<td>5</td>
<td>Extract mainframe student data for import to data warehouse.</td>
<td>October, 1995</td>
</tr>
<tr>
<td>6</td>
<td>Develop reporting applications for initial use and training activities.</td>
<td>December, 1995</td>
</tr>
<tr>
<td>7</td>
<td>Train pilot group and solicit feedback.</td>
<td>February, 1996</td>
</tr>
</tbody>
</table>
Step 1 involved a review of the IS department’s most commonly received requests for information. The outcomes of this review were discussed with the Assistant Deans and were refined to include only biographic and enrollment information for currently registered students. The initial data that will be stored in the data warehouse includes the following:

- student biographic information such as name, birth date, handicap, citizenship, gender, and ethnic origin
- local mailing address of student
- permanent mailing address of student
- enrolled program of study, including college, degree, department and major
- current enrolled courses, credit hours, etc.
- current GPA, attempted hours, etc.

Step 2 covers the evaluation and selection of an appropriate platform to use for storing the files associated with the data warehouse. A Windows NT server running Microsoft’s SQL Server software was selected for this purpose. This decision was based on a combination of factors. Previously, a Windows-NT server had been purchased and was controlled by the IS department for research and development purposes. Its only use at the time of the data warehouse implementation, was for a similar development project. Therefore, it could easily be allocated for the new project without conflicting with other university needs. In addition, the SQL Server software would provide a centralized relational database that could be queried to transmit only the requested data to the user workstation.

Step 3, the evaluation and selection of PC-based reporting tools, was based on three factors. These include its ease-of-use, integration to other Windows-based software, and its technical sophistication and potential for growth. The MS-ACCESS database product was selected as meeting these requirements and went on to become our university standard for PC databases.

Step 4 involved the selection of workstation hardware that would support the MS-ACCESS requirements for memory, disk capacity, and processor type. A specification sheet comparing these requirements with current workstation configurations was distributed to each of the seven colleges. This comparison chart was used to determine the upgrades needed for each workstation that would be equipped to access the data warehouse information.

Step 5 was the easiest step for me to accomplish. Being the manager of Application Systems, I was frequently involved in extracting data from our mainframe system. This same type of extract process is used daily to fulfill the many report requests received in IS. The only learning that needed to occur was how to the import the data into the SQL Server warehouse to be queried with MS-ACCESS.

On the other hand, Step 6 is probably the most training-intensive and time consuming piece of the development. Not having worked with either SQL Server or MS-ACCESS before, I was starting at ground zero. Intense self-training exercises are conducted in “off”
hours to prepare myself for development of the front-end, user-interface portion of the application.

Step 7 will involve training selected end-users from each college on the use of the delivered applications and the types of further reporting capabilities that can be accomplished. The goal is to develop four report/query applications and one mailing label application to deliver with the final product and to use for training purposes. This step is currently in the planning phase. The pilot group of users has been selected and includes one or two key users from each of the college administrative offices. Training sessions will be conducted by the application developer and will concentrate on two specific areas: 1) a demonstration on how to generate the delivered reports through the user interface, and 2) other report, query, and analysis functions available in MS-ACCESS.

Step 8 includes the development of database administration procedures along with a system reference manual. This process, although included as the last step of the project, will be ongoing throughout the development phase. Items contained in the reference manual will include the following:

- Schedule of data warehouse “refresh” procedures
- Procedures for adding new fields to the data tables
- Copy of extract programs used to transfer data from the mainframe SIS system
- Extract file definitions
- Copy of table layouts and definitions for each table in the data warehouse
- A data flow diagram of the system

The completion of the data warehouse along with the deliverables specified in each step is scheduled for March, 1996.

**Impact on Customer Functions**

The implementation of this data warehouse approach to student information will have a significant impact on customer functions related to data gathering and analysis. More timely information is provided for review by administration since information is available as soon as the data is queried. An improvement in data quality results from direct feeds of data to the warehouse, eliminating errors produced by manual data entry and the consequences of making decisions based on erroneous information. Hours of staff time that were involved in manual data collection can be freed up to add valuable interpretation and background information to reports. Data presentation can be tailored to the needs of different audiences.

The implementation of this proposed data warehouse application will enable the college administrative offices at Ferris State University to make effective management decisions regarding their enrolled student population. They will be able to readily access the information needed to verify and analyze college, program, department and course enrollment statistics.
The use of PC-based software on the client workstation adds the potential to access, manage, analyze, and present the warehouse data in various formats. Class lists, course registration audits, and departmental reports can easily be produced. Search and sort capabilities will result in faster data lookup and make specific student information more readily accessible. The ability to summarize, report, or project specific college or program data will be available. The integration with other Windows-based tools offers the potential to graphically analyze and view trends in enrollment data.

Local printing capabilities will be available to produce mail-merge documents, hard copy reports, or student mailing labels. This will relieve some of the need for the Information Services department to produce, print, and distribute ad hoc reports. The ability to quickly access and view student data and/or reports in an on-line Windows environment could result in a significant savings in paper costs.

Changes in Information Services Role

The implementation of this data warehouse approach to student information will significantly change the role in the Information Services department when fulfilling report requests and other information needs across campus. By empowering the college administrative staff and faculty with the ability to generate their own reports, mail-merge applications, and student mailing labels, the backlog of report requests received by IS can be significantly reduced. Technical staff time can be freed up to work on more complex programming issues. Other technological advances can be developed to meet university needs, resulting in a more challenging and rewarding work experience for the technical staff.

Reduced operational support will be needed for printing and mailing reports, resulting in less overtime pay to an already overloaded operations staff. Paper costs for the IS department can be greatly reduced while resulting in more attractive reports being produced locally in the colleges.

FUTURE PLANS

The primary objective of a data warehouse is to give end-users faster, easier, direct access to institutional data so that they can make quicker, more-informed and better decisions. This initial data warehouse project only targets a small percentage of the reporting and data analysis needs across campus. Used as a pilot research project concentrating on currently enrolled student data only, it is hoped that the outcomes will provide significant insight into the methodology, impact, and usefulness of this approach for meeting data access needs in other areas. These same techniques can be used to access data in the financial, human resource, or alumni/development systems.

Setting the Stage for Client/Server Architecture

As technological advances progressively move to client interfaces for ease-of-use and direct access to decision-making data, so, also, must the underlying architecture support these client needs. In general, the mainframe provides a great deal of reliable processing
power and storage space for large database systems. However, it does not provide adequate interfaces for easy access by the average end-user. On the other hand, microcomputers have relatively friendly and powerful user interfaces, but they lack the storage and processing power of larger computers.

In a true client-server system, computers share the processing load. The data warehouse approach allows each computer to do what it is best at. The administrative functions of the university are based on mainframe systems that are tuned for high-volume transaction processing (e.g. registering students, crediting and debiting university accounts, and issuing paychecks and W2 forms). By placing the data warehouse on a separate SQL server platform, the efficient processing of the University’s operational systems will not be compromised by ad hoc queries against the data they contain. In turn, the data warehouse can be tuned to provide optimal performance for ad hoc queries from the client workstations.

**Improved Customer Relations**

The growth of this data warehouse to include additional stores of data from multiple sources will lead to improved customer relations throughout the university. Staff will be motivated to re-tool, data accuracy will be improved, and many workflow and decision-making processes will be re-engineered for improved effectiveness. Several other factors relating to this potential include the following:

- **Decision Support.** FSU is currently engaged in developing an Executive Information System (EIS). Decision-support capabilities will provide all interested parties with the information necessary to make the best possible decisions regarding the status and progress of FSU. Included in the benchmarks relating to the development of this EIS are the considerations for existing databases, accurate and reliable data, and alignment with the university’s Information Technology Strategy. The development of this data warehouse will add to and strengthen our portfolio of EIS tools.

- **Decentralization of Ad Hoc Reporting.** With ad hoc reporting responsibilities currently placed in the IS department, customers are not always fully aware of where their requests fall in the prioritization process or what currently exists in the programming backlog. The development of the data warehouse will provide end-user departments with the ability to produce their own ad hoc reports. The “ownership” of the programming backlog will shift from the IS department to the colleges and customer departments. Prioritization of reporting needs will be done within individual departments and will not conflict with other campus needs.

- **LAN Server Consolidation.** The promotion of a central server for data warehouse processing lends itself to a streamlined process for data entry and administration. Many departments currently maintain data on local servers and individual workstations. This often results in the duplicate entry of data originating from the same source. By consolidating LAN servers into a central database server, this redundancy of effort can be eliminated. The central server can handle the database processing while the client
receives only the data that was requested. This model causes minimum network traffic.

- **Continuous Quality Improvement.** This data warehouse approach will provide for more accurate, reliable data. Duplicate departmental databases that may not be systematically maintained to assure accuracy will be no longer be needed. In addition, this approach will empower the customer to detect exceptions from the norm by providing new perspectives on production data. These exceptions can be noted and reported to the owner, building a continuous effort to provide quality at the source.

**CONCLUSION**

Ferris State University is committed to improving customer service. By providing easier methods for accessing institutional data, customers can have needed information at their fingertips for improved decision-making, analysis, and reporting. The implementation of this data warehouse approach for student enrollment information, will move the university one step closer to that goal. To date, the project is well underway, and is currently being evaluated for use in other areas.