UNIVERSITY OF CALIFORNIA

OFFICE OF THE PRESIDENT

Review of Corporate Information Systems

February 13, 1997

TABLE Of CONTENTS

SECTION I – EXECUTIVE SUMMARY

Background

The Information Resources and Communications (IR&C) department within the Office of the President, collects common data from each of the campuses, validates the information and consolidates it into multiple databases, collectively referred to as Corporate Systems.   The University of California, Office of the President (UCOP) retained Deloitte & Touche LLP to independently review the Corporate Systems and answer the following questions:

- Is UCOP utilizing the most appropriate technical tools and architecture currently available?
- Is the data being stored in the most organized way possible?
- Is the ease of access to data sufficient for corporate users?
- Should UCOP make changes to improve efficiency and effectiveness?

Corporate Systems include the following systems:

1. Contracts and Grants
2. Equipment and Facilities
3. Financial
4. Personnel
5. Student
6. Staffing

At the request of UCOP, the scope of this project was limited to researching and analyzing technology and Corporate System data *after* it was received at UCOP. Another limitation of the project was a request to not review current chargeback mechanisms or consider the cost of implementing new technologies. At management's request, we did not conduct interviews with campus personnel responsible for compiling, accessing, and transmitting data to Corporate Systems.

Summary of Findings

Our project approach included:

- Corporate Systems documentation reviews
- Interviews with analysts and end users
- Focus Groups
- Analysis of current technology trends in data warehousing

Our findings are separated into three areas. First, findings specifically related to technology, second, findings pertaining to process improvements, and third, findings specific to an

individual Corporate System. Issues that were raised in numerous sessions were validated during the Focus Group review process.

Technology Findings

Six significant technology findings were identified, including:

1. *Corporate Systems lack an easy-to-use interface for querying and reporting*
2. *Shadow systems are prevalent**
3. *Ability to drill-up/down on Corporate System information is difficult or not possible*
4. *Not enough historical information is being stored online*
5. *Corporate Systems lack a multi-dimensional querying capability*
6. *There is a lack of integration between Corporate Systems databases*

*   The term "*shadow systems*" is defined as the replication of all or part of a database to a different platform.

Process Improvement Findings

Five significant process improvement findings were identified, including:

1. *Outdated and unclear user documentation for all Corporate Systems exists*
2. *The Corporate System chargeback mechanism is significantly altering the way and time in which work is performed*
1. *Users perceive that not all databases are accessible and policies for accessing Corporate Systems are not well understood*
4. *Data integrity is generally good, although questionable in some cases*
5. *Data is out of date, in certain Systems, by the time it is available online*

Process and policy issues were found to affect the user as much as the technological tools currently in use by UCOP.

Corporate System Specific Findings

Exhibit I: *Corporate Systems Databases*, in Section V, summarizes findings and characteristics relative to each Corporate System. Each of the Corporate Systems had specific data elements that users would like to see included if revisions were to be made, but overall, Corporate Systems seem to be capturing a majority of the information users require. The data elements that users did request could be characterized as "nice to know" vs. mandatory elements required to perform one's job. The ability to access data as soon as possible, stored and online for a longer period of time was also identified by users as a need. Additionally, users would like to be involved in discussions related to future direction of the databases.

Recommendations

We recommend UCOP address the above findings in two phases, Short-term and Long-term. The action plan for the short-term, Phase I tasks, include:

- Implement new IBI Business Intelligence tools and other front-end tools
- Increase user access, training, and education of current Corporate Systems and new Business Intelligence tool capabilities to potentially reduce shadow systems

The key success factors in implementing Phase I action items would include:

- Reduced implementation time because of familiarity with the existing environment and compatibility of proprietary software tools
- Minimal user disruption because the operating environment and database design are not changed
- Contained overall expense because the operating environment and database engine are not changed

Prior to implementing any changes, UCOP must develop evaluation criteria. The criteria should address the following at a minimum:

- Compliance with business/functional requirements
- Level of risk acceptable to UCOP
- Costs
  - software
  - maintenance and upgrades
  - hardware, if required
  - solution development and support
  - pilot testing
  - conversion support
  - procedure development/user documentation
  - training (technical and user)

Steps necessary to successfully accomplish Phase II action items include:

1. Evaluate and document data and user requirements in business terms and use as evaluation criteria
2. Establish additional evaluation criteria (vendor stability, market share, upgrade frequency, etc.), procurement strategy, and implementation strategy
3. Choose an integrated solution vs. a Best of Breed approach
4. Review and select a Database Management System (DBMS) that meets at least 80% of UCOP business requirements

These steps will require an in depth study of user requirements and specific technologies available at the time of analysis, as this technological arena is extremely dynamic. Other essential elements include a detailed work plan that is feasible and realistic to help insure successful implementation, transition and adequate training in order to maximize the desired objectives.

Conclusion

Based on our research and findings, UCOP is not achieving maximum utilization of Corporate Systems. Users are relying on ancillary and/or shadow systems or other workaround scenarios in order to answer requests for information or to produce analytical reports.

It is clear that advancements in data warehousing tools could be of significant benefit to UCOP's user community. All six of the technology findings can be addressed, but the time, money and effort to achieve results vary significantly.

The short-term phased approach enables UCOP to address immediate problems without a major overhaul of the existing environment. The appropriate short-term decisions should support the long-term goal of a more effective corporate information systems environment. Therefore, this report recommends a pragmatic, phased approach to accomplishing a more user-friendly, effective environment.

SECTION II – INTRODUCTION

The University of California system serves over 160,000 students at nine campuses and operates three research laboratories for the U.S. Department of Energy. Since 1978, each campus has been responsible for its own administrative processes and has developed information systems to support their own operations. While each campus has the authority to maintain their operations independently, each campus must submit required information according to schedules and in identical formats to the Corporate Information Systems. The exception is the University Payroll/Personnel System which is the single core administrative system common to all campuses.

The Office of the President has the critical role of overall operation of the Corporate Information Systems, including policy planning and evaluation, resource acquisition and allocations, and external reporting. The Information Resources and Communications (IR&C) department within the Office of the President, collects common data from each of the campuses, validates the information and consolidates it into multiple databases, collectively referred to as Corporate Systems. Corporate Systems include:

1. Contracts and Grants
2. Equipment and Facilities
3. Financial
4. Personnel
5. Student
6. Staffing

The Corporate Systems operate as separate data repositories or warehouses for similar types of information such as student, contracts and grants, and financial. This review focused on the needs of UCOP users of the systems, appropriate technologies, and advancements in data warehousing tools and techniques to meet those needs. Sophisticated warehousing tools and techniques were not available when the systems were initially placed into service. Newer technology examples include software suites from IBI, Red Brick and Oracle. These vendors provide various support tools such as Decision Support Systems (DSS) and Executive Information Systems (EIS) capabilities that can provide enhancements over the basic database and querying tools currently in use. Appendix A, Data Warehousing Technologies and Trends, discusses the applicability and availability of data warehousing solutions.

As the University of California System approaches the 21st century, several critical challenges exist:

- How to accommodate the tremendous growth in eligible students while not overtaxing the administrative system infrastructure

- How to maintain academic quality with limited financial support from the state
- How to continue to demonstrate leadership by helping to strengthen the California economy through existing programs and new initiatives

Corporate Systems data is used to analyze and address these types of issues. To address these challenges the Office of the President must have easy access to information and systems flexible enough to meet the ever-changing environment.

SECTION III – PROJECT SCOPE

The University of California, Office of the President (UCOP) retained Deloitte & Touche LLP to independently review the Corporate Systems and answer the following questions:

- Is UCOP utilizing the most appropriate technical tools and architecture currently available?
- Is the data being stored in the most organized way possible?
- Is the ease of access to data sufficient for corporate users?
- Should UCOP make changes to improve efficiency and effectiveness?

At the request of UCOP, the scope of this project was limited to researching and analyzing technology and Corporate System data *after* it was received at UCOP. Another limitation of the project was a request to not review current chargeback mechanisms or consider the cost of implementing new technologies. At management request, we did not conduct interviews with campus personnel responsible for compiling, accessing, and transmitting data to Corporate Systems

This report is segregated into the following sections:

- Methodology
- System Descriptions and Characteristics
- Corporate Systems Findings
- Conclusion
- Action Plan

SECTION IV – METHODOLOGY

To develop an understanding of Corporate Systems, we reviewed documentation and conducted interviews. Included in our documentation review was the report entitled "Migrating UCOP Corporate Systems from VM/CMS" prepared by Information Systems and Computing (IS&C). IS&C addressed the technical benefits of migrating from one IBM operating system to another (VM/CMS to MVS), the use of DB2 as the database engine and implementation of new front-end query tools.

Interviews were conducted with users of the Corporate Systems. Additionally, various IR&C staff responsible for system maintenance and technology, analysis, and general use were interviewed. The objective was to obtain the users' perspectives, issues and concerns regarding the existing systems ability to meet current business requirements. An interview questionnaire, with general and specific system questions, was used as a guide to ensure consistency in developing our analysis. A list of interviewees is included in Appendix B, Interview Listing.

Topics outlined in the questionnaire and addressed during the initial interviews included the following:

Ease of Use
- Reporting
- Ad-hoc querying
- General use

"Shadow" Systems
- Quantity
- Degree of reliance

Data Elements Captured
- Too many
- Too few
- Documentation

Timeliness of Data
- Provided by the campuses
- Available to the user community

Access
- To specific systems, including frequency
- To information needed
- Costs involved

Analytical Capabilities
- Reporting
- Decision Support Tools

Database Structure and Organization
- Meeting business requirements
- Flexible

Support from IR&C

Accuracy and Integrity of the Data

Documentation of Systems

Systems Availability

After the initial round of interviews, we conducted five Focus Group meetings. Each meeting focused on a single Corporate System and included IR&C personnel, current users and potential users in the same meeting. The purpose of the Focus Groups was to validate the information gathered during the individual interviews and to clarify issues affecting the different user groups. The involvement of IR&C personnel, current users and potential users was critical to the successful identification of issues with the current systems. A list of Focus Group attendees is included in Appendix C, Focus Group Meetings Listing.

A review of user service requests submitted to IR&C was also performed to determine which systems are generating the most requests and the nature of those requests. This analysis revealed the majority of service requests were for reports from the Corporate Personnel System. This information, coupled with the interview and Focus Group analysis, was used to develop a preliminary list of potential change opportunities that are included in the Corporate Systems Analysis.

We would like to note the frequent and high praise given for IR&C personnel. Users acknowledged the added value and unwavering support of IR&C personnel in helping make the most of systems placed into service more than 15 years ago.

SECTION V – SYSTEM DESCRIPTIONS AND CHARACTERISTICS

Corporate Systems Overview

The Corporate Systems are comprised of six main systems, with multiple databases within each system. Campuses send the data to UCOP with OP being the primary users of the information. The specific databases allow common data elements to be grouped together or to provide different "views" of the data. For example, the Student Longitudinal Database contains historical information separate from the current year information. This approach reduces the search time required and provides focused data for analysis. The systems are as follows:

1. *Corporate Contracts and Grants*
2. *Corporate Equipment and Facilities System*
3. *Corporate Financial Systems*
4. *Corporate Personnel System*
5. *Corporate Student System*
6. *Corporate Staffing System*

Corporate Systems were developed approximately 15 years ago using Information Builders Inc. (IBI) proprietary FOCUS hierarchical software database management system and associated scripting language for executables which are referred to as FOCEXECs. The COBOL and PL1 programming languages are also used in conjunction with FOCUS, for edit and validation routines of the data received from the campuses. This process ensures that the data received is reasonable and accurate. Data transmissions between UCOP and campuses are performed using File Transfer Protocol (FTP) standards. Finally, the databases reside on an IBM mainframe utilizing the VM/CMS operating system.

Corporate Systems Databases

Exhibit 1, Corporate Systems Databases, depicts detail associated with the Corporate Systems. It includes the purpose, statistics, characteristics and integration capabilities with other databases. Database records and sizes are approximate and are useful for comparing systems. The section: *Ancillary Systems Used,* should be considered a tool to analyze expansion of Corporate System capabilities and a method for identifying deficiencies.

Corporate Systems Databases
Exhibit I

System	Purpose	Size	Databases Comprising the System	Users	
Corporate Contracts and Grants (CGX)	To provide UCOP management and staff information about proposals and awards from UC Campuses	200,000 records, 50 Mb	• Awards • Proposal	• Costing Policy and Analysis • Research Administration • Office of Technology Transfer • Academic Affairs Research	• STAR Program Users • UC Biotechnology Program • Governmental Relations • Agriculture and Natural Resources
Corporate Equipment and Facilities System (EFA)	To provide UCOP management and staff an inventory and information on buildings, rooms and equipment at UC Campuses	900,000 records, 275 Mb	• Equipment • Buildings and Rooms	• Budget Office • Corporate Accounting Office • Costing, Policy and Analysis • Materiel Management	• Risk Management • Facilities Administration • IR&C
Corporate Financial Systems (CFS)	To provide UCOP management and staff with University budget and general ledger information for analytical and reporting purposes	1,200,000 records, 215 Mb	• Master (although four snapshots of different reporting periods are available)	• Corporate Accounting Office • Budget Office • Agriculture and Natural Resources • Costing Policy and Analysis • Research Administration	• Budget Administration • Audit • Academic Affairs: Budget Administration • Academic Affairs: UC Biotechnology Program
Corporate Personnel System (CPS)	To provide UCOP management and staff with demographic, personnel, and pay data	2,300,000 records, 2.2 Gb	• Fiscal Year • Four snapshot files are made available • History	• Human Resources • Academic Personnel • IR&C • Academic Affairs Budget Administration • Student Academic Services	• Risk Management • Benefits • Loan Programs • Real Estate Services • Budget Office • Costing Policy and Analysis
Corporate Staffing System (SLC)	To provide UCOP management and staff the capability to reconcile academic and staff salary commitments and provisions with budgeted appropriations	Not significant	• Staffing	• Budget Office	
Corporate Student System (CSS)	To provide UCOP management and staff with analytical and operation information regarding the student population	1,500,000 records, 1.3 Gb	• Registrant • Undergraduate Admission • Graduate Admission • Financial Aid • Undergraduate Longitudinal • Graduate Longitudinal • Post Doctoral	• Student Academic Services • Budget Office • Academic Advancement • Planning and Analysis • Education Abroad Program	

Corporate Systems Databases (continued)
Exhibit I

System	Database Owner	Frequency of Use	Characteristics	Ancillary Tools/Applications Used
Corporate Contracts and Grants (CGX)	Research Administration	Database(s) is/are accessed several times per week	• Significant Campus interest in data • System processes quarterly • Voluminous reports printed on cycle • Data is retained for 5 years from termination of activity (Awards DB) • Prior year data in the Proposal DB is overwritten annually • Campuses send the increase/decrease to current awards to change their balances, i.e. data is current as of the last reporting cycle • CGX-OPS1 creates Awards and Proposals	• Microsoft Access • Delphi: Office of Technology Transfer (OTT) extracts summary CGX data and enhances • FoxPro is used by Costing Policy and Analysis (CPA)
Corporate Equipment and Facilities System (EFA)	Facilities/Budget Office Material Management	Facilities: Daily	Capital Asset Parameter (Equipment database): • More than $500.00 • Life expectancy greater than 2 years • Free standing	• Real Estate Information Management System (REIMS) • Comprehensive Rate Information System (CRIS) • Lease Information Management (SLIM) • Space Tables: Setup in FoxPro database for reporting to the State • Risk Management: 4th Dimension database
Corporate Financial Systems (CFS)	Corporate Accounting Office	Database(s) is/are accessed between once a month to once a week.	• Budget data is merged with general ledger data 5 times a year. Oct., Dec., Mar., May, and Jun. are reporting periods • Reports being compiled are characterized as useful • End users run the Gateway product as a query tool • Many detailed budget reports being sent to the State	• FoxPro • FileMaker Pro
Corporate Personnel System (CPS)	IR&C: There's no one single owner	Range of use varies significantly from about once a week to a few times a year.	• Largest of all Corporate Systems databases	• FoxPro • Excel • Access
Corporate Staffing System (SLC)	Budget Office	Database(s) is/are accessed a few times a year and a very limited number of personnel accessing	• Not available	• Not applicable
Corporate Student System (CSS)	IR&C: There's no one single owner	Database(s) is/are accessed between about once a week to a few times a year.	• Reports sent to Federal government and the Board of Regents • Re-specifying the Financial Aid database in 1998 or 1999	• SAS • FoxPro • Excel • Access • PC Focus • KEdit (for sorting purposes)

SECTION VI – CORPORATE SYSTEMS FINDINGS

Multiple interviews and Focus Group sessions were conducted to address the questions posed by UCOP as the scope of this project. To reiterate, the questions regarding Corporate Systems include:

- Is UCOP utilizing the most appropriate technical tools and architecture currently available?
- Is the data being stored in the most organized way possible?
- Is the ease of access to data sufficient for corporate users?
- Should UCOP make changes to improve efficiency and effectiveness?

If similar responses to these questions occurred multiple times during the interviews, the responses were summarized into Technology Findings or Process Improvement Findings. Technology Findings are defined as issues or constraints with general system capabilities. Process Improvement Findings may be addressed through policy, procedure or practice changes. In addition, if findings were identified for a specific Corporate Systems, we have included these individual findings at the end of this section.

Technology Findings

1. Corporate Systems lack an easy-to-use interface for querying and reporting

A significant amount of effort and skill is required to extract information from the Corporate Systems. This theme was raised regardless of whether the user was a sophisticated analyst comfortable with FOCUS or a casual user. The team has identified this as the number one technology issue affecting end users.

This impediment is discouraging use of the Corporate Systems and the implications are significant. It results in users creating and maintaining a subset of system data (ancillary systems and/or shadow systems) because the tools used to access the subset of data are easier to use and faster than accessing the mainframe. Also they are frustrated because they cannot access the information they need without assistance by IR&C. This may impact the users ability to be responsive with analysis or information due to the need to turn to someone with more technically proficient skills. "Power-users" have become an important resource to assist less sophisticated users. This creates a bottleneck for information flow due to the availability of the power users time.

2. Shadow systems are prevalent

Periodically, situations warrant the use of secondary or shadow systems. Unfortunately, the situation for development of shadow systems at UCOP is generally related to the ease-of-use theme cited above, not enough historical information being made available on-line, and

the high-cost of accessing the systems. Numerous database products are deployed, e.g. FoxPro, Access and FileMaker Pro to overcome these issues.

Since users require information to perform their jobs, alternative methods for maintaining information are created if the main source of information is not adequate. The result is duplicate information, potentially jeopardizing data integrity, and overall increased total cost of operations to maintain multiple vendors and versions of software. Additional impacts are disparate levels of training and user proficiency and an inability to share information at the local user level.

3. *Ability to drill-up/down on Corporate System information is difficult or not possible*

The inability to summarize information and then drill-up/down through the detail is an ease-of-use issue for users. Most users are familiar with the graphical, point and click capabilities of other databases and want that timesaving flexibility within Corporate Systems.

The effect is that PC shadow systems with graphical user interfaces are implemented to increase the ability to easily summarize and manipulate data.

4. *Not enough historical information is being stored online*

The nature of a data warehousing environment connotes long term storage of information for organization-wide reporting, trend analysis, etc. The historically high cost of disk storage has minimized the amount of information stored on-line. The industry average is five years of data and this coincidentally was the period of time most often requested by users

The outcome is that users can not answer historical questions effectively and newer tools that could help identify previously unrecognized trends can not be deployed due to the limited historical data retained.

5. *Corporate Systems lack a multi-dimensional querying capability*

Currently, the tools necessary to perform multi-dimensional complex querying are available on the market. Specific database queries with minimum complexity or limited data elements are currently the only tools available to UCOP users.

From a user perspective, the Corporate Systems do not provide the ability to efficiently and effectively execute multi-dimensional queries to uncover trends and associations in campus activity. An example of a multi-dimensional query would be accessing student information, then identifying how many students are on a particular campus, and of those how many are of a certain ethnic background, in a specific class, anticipating graduation. This type of query provides analytical information from which business decisions can be made. Multi-dimensional queries are complex and currently unavailable to users.

6. Lack of integration between Corporate Systems databases increases staff workload

The lack of integration between databases within a Corporate System and lack of integration between Corporate Systems results in increased staff workload. For example, revenue and expenditures for awards would require linking the fund number between the Corporate Financial System and the Contract and Grants System. This combination of information is not readily available. Therefore, staff must develop a work around environment, possibly in the form of an ancillary database, to try and generate the desired information.

To obtain cross system information requires multiple queries and subsequent exporting and importing of data to produce the answer, which at times is unobtainable.

Process Improvement Findings

1. Outdated and unclear user documentation for all Corporate Systems exists

Although some of the databases have relatively few data elements, most have a substantial number of data elements. While the majority of the data elements were reported as being required, the multitude of data elements available in each of the systems causes sophisticated and casual users alike difficulty in deciphering the intent of the data element. For example, a consistent business-like definition is not available to explain what is included in the definition of data elements such as full-time equivalent (FTE). Various types of information can be included or excluded from this data element and without a clear and definitive definition, users are uncertain of the true meaning.

The result is that the user is apprehensive about the results generated by the element, i.e. what the results truly represent and are forced to request assistance from another source to answer the question. This is impacting how far staff will go outside of their *comfort zone* to retrieve information. This lack of system reliance by the end-user decreases the effectiveness of Corporate Systems.

2. The Corporate System chargeback mechanism is significantly altering the way and time in which work is performed

Many data centers use a chargeback mechanism to recoup system costs. UCOP also employs a chargeback methodology for accessing Corporate Systems. The current methodology appears to be causing user avoidance and the creation of shadow systems that may be exacerbating the chargeback intent. Users are copying vast amounts of data from the mainframe to their desktops in order to avoid CPU charges. This allows the users to frequently access and report on the information they have downloaded without further cost. Also, staff come in early, delay analysis by executing batch jobs during off-hours, or just do without information, rather than incur charges.

There also appears to be a lack of understanding regarding the costs, including queries, reports, programming, etc. and how to obtain the cost information prior to accessing the system. Alternative chargeback methods were not analyzed as part of this project.

3. *Users perceive that not all databases are accessible and policies for accessing Corporate Systems are not well understood*

User access to certain databases is not available. The Equipment database and Corporate Personnel fiscal year database were never designed for user access. If users request access they are denied for these reasons. For other databases, IR&C analysts will discuss with individuals their specific needs so that access to the appropriate databases is granted. Users have requested access to multiple databases but were granted access to less than the requested amount without understanding the reason.

There is a database access request process that does not appear to be well understood or consistently applied. As a consequence, a level of frustration exists because of the perceived lack of responsiveness and understanding by IR&C to actual user requirements. On the other hand, IR&C has provided information regarding the access request process so there may be a need to provide additional user training on this process.

4. *Data integrity is generally good, although questionable in some cases*

Data integrity is essential for any database to ensure accuracy, validity and consistency of information. Edit and validation routines are executed by UCOP during the load process, but organization of the Corporate System databases and the multiple stand-alone campus systems produce a less than ideal situation for ensuring data integrity.

Additional issues include some campuses not providing required data elements and some campuses using required fields to maintain non-standard information causing inconsistencies with other campus' information.

5. *Data is out of date by the time it is available online*

Each of the Corporate Systems has dates when data is or should be available, but many times the data is not available on that date. UCOP is dependent upon the campuses to forward accurate information in a timely fashion. If the campuses do not comply, UCOP is limited in its ability to run complete reports until the campus sends the necessary file. The situation is further hampered when the information sent is not usable, thereby requiring the campus to resubmit the data, further delaying the time when analysts can access the information.

This can result in delayed or incomplete analysis and reporting.

The following findings pertain to a specific Corporate System or database. These findings were not summarized into the Technology or Process Improvements Findings identified above.

Each specific system finding includes a *Data Elements Desired* listing that identifies data elements not currently available in that Corporate System.

Corporate Contracts and Grants System (CGX) Specific Findings

1. *The Proposal and Awards databases within the Contracts and Grants System lack integration that creates tracking problems*

   The natural evolution of a proposal is through award and funding. Some proposals may never be awarded but for those proposals that do receive funding, a method for identifying or linking the Proposal to the Award in the separate databases would be useful.

2. *Lack of integration between the Contracts and Grants System and the Corporate Personnel System (CPS) creates tracking problems*

   The ability to link with the CPS database would allow staff to link individuals to contracts and grants activities and/or revenues/expenses. Currently, it is difficult to gain this project level perspective.

3. *Lack of integration between the Contracts and Grants System and the Corporate Financial System (CFS) creates funding tracking problems*

   The ability to link with the CFS system would allow staff to determine the relationship between funding and/or expenses at the project level. However, there may be a problem in doing this because the campuses can re-use the fund number, thereby returning confusing information on a report, query etc.

4. *Campuses desire information from the Contracts and Grants Systems pertaining to other campuses*

   Although it was outside of the scope of this project to meet with campus personnel, the issue was raised that campuses would like to have access to public information regarding other campuses. Currently, campuses receive summary and hard copy reports of their campus. The format, level of detail, and scope of information does not satisfy campus requirements. For example, who has been awarded a grant for a specific project was requested. Campuses are also looking for easy to use interface and query tools.

5. *Entry of certain data elements into the Contracts and Grants System is not standardized*

   The ability to search for a specific entity is compromised due to a lack of consistency in the entry of certain elements. For example, principal investigator and co-investigator names are entered into the database using various formats. Since there can be multiple investigators involved and there is only one entry field for each type of investigator, the element makeup of the database would need to be altered to accommodate user needs.

Data Elements Desired:
- Link w/CPS for Multiple Co/Investigator entry
- Human/animal subject involvement
- Terms and conditions
- Material Transfers agreements
- Expand Project Type
- Gifts and Endowments

Corporate Equipment and Facilities (EFA) System Specific Findings

1. *Lack of integration between the Equipment and Facilities databases creates problems identifying equipment located at a specific facility*

   Linking equipment owned by the University to a specific facility within the University system is technically feasible, with programmer support, but from a practical point is not easily accomplished. It is not possible for an end user to summarize the value of equipment within a specific facility.

2. *Assets information is overwritten, affecting the analysis done on inventory*

   The equipment database is overwritten on an annual basis so necessary information may not be available in a subsequent period for analytical inquiry. The ability to ascertain if the end of an asset's useful life was reached, whether it was lost, stolen, or upgraded, etc. is not possible if the asset no longer appears on the report. As a result, the ability to fully analyze University assets is hampered due to lack of historical data.

Data Elements Desired:
- Financially related information
- Asset data with valuation
- Construction related information
- Use and age of building
  - Although the age and use of buildings are included in the database, multiple definitions are possible for these elements. For example, date of construction, date acquired, or date occupied. Likewise the use of a building can take on multiple forms and meanings.

Corporate Financial Systems (CFS) Specific Findings

1. *The Corporate Financial System is unable to link data from different years*

   The format of data from prior years has changed so that the end users do not have the ability to easily link information from various years. In addition, some of the data elements may have been eliminated from one year to the next. This increases the difficulty in comparing years.

   The ability to seamlessly link data from prior years is fundamental to understanding the financial environment and how/where it is changing. Use of data mining tools that could

aid the University in uncovering hidden trends, relationships, and analysis will also be hampered if querying data spanning multiple years is not possible.

2. *Unable to easily drill up or down on the data*

   This issue is related to ease-of-use brought forth in the overall themes but goes further in that CFS users would like to be able to create a reporting environment mirroring the reports which they are responsible for producing. The ability to review the summarized document and then *double-click* on a cell reference to view what that cell comprises, is intuitive and powerful.

Data Elements Desired:
- Insurable Value

Corporate Personnel System (CPS) Specific Findings

1. *Data required for the Corporate Personnel System is not timely from the Labs*

   It is the perception of the Benefits Department that data coming from the Labs is late or missing altogether. Obviously, access, reporting and analysis of the data would be affected.

2. *Leave and vacation data is not correct or reported inconsistently*

   Sick leave and vacation data received by UCOP from the campuses is not always accurate or reliable. Therefore, it is not useable by Personnel, Labor Relations or other departments for analysis. This appears to be due to a lack of importance to the campuses to forward this data to UCOP.

3. *Hard copy reports contain dated material*

   Due to the various pay methods and period cutoffs that take place in the middle of a reporting cycle, querying information on a complete cycle must be delayed until the next month. By the time standard reports are generated and distributed to the users, much of the information is outdated.

4. *Users lack control or input into the change process*

   Users feel a lack of decision making power in the ability to add/change/delete data elements within the system. There is no formal change control process for users to participate in the decision making process, prioritization of changes, or to communicate changes that have occurred. This process in essential if users are to continue to take advantage of the changes that are being implemented.

Data Elements Desired:

- Range adjustment information
- Termination date of health plan coverage
- UC-Paid Life and Employee-Paid Life Insurance amounts
- Medical Center designation
- Leave Accruals
- Life insurance beneficiary designations
- Percentage of time worked vs. amount paid
- Job code/description

Corporate Staffing System Specific Findings

Interview sessions did not reveal any issues with the Staffing System or any new data element requirements.

Corporate Student System (CSS) Specific Findings

1. *Hard copy report requirements for the Student databases need to be reviewed*

   Some of the other Corporate Systems have periodically reviewed the usefulness of the standard reports being generated by the system. It appears that the production reports from the Student databases have not been reviewed recently with end-users to determine their current value.

2. *Users are unable to determine the resources and income students are using to pay for their education*

   The ability to analyze a student's source of income would be useful in understanding students' financial needs. For example, income and resources include grants, loans, credit card debt, etc. This information would be used to identify trends in education spending.

Data Elements Desired:

- None, although the users would like to include high-school related information and graduate admission information on campus activity currently being captured

SECTION VII – CONCLUSION

UC Office of the President requested an independent study to address the following questions:

- Is UCOP utilizing the most appropriate technical tools and architecture currently available?
- Is the data being stored in the most organized way possible?
- Is the ease of access to data sufficient for corporate users?
- Should UCOP make changes to improve efficiency and effectiveness?

Based on our findings regarding these questions and research of data warehousing technologies, it appears UCOP is not achieving maximum utilization of Corporate Systems. While the appropriate data elements are being captured and IR&C staff are very knowledgeable, the general users are not relying on Corporate Systems as extensively as they could be. From D&T's perspective, significant amounts of time, energy and staff resources are consumed in the collection and maintenance of Corporate Systems while the return on investment and user reliance appears to be low.

We analyzed four options for resolving the Technology Findings. The options included:

- No change to Corporate Systems
- Migrating from VM/CMS to MVS
- Implementing new front-end user access tools
- Implementing a new data base and front-end user access tools

Each option was analyzed including the benefits, risks, issues and assumptions associated with implementing the option. This analysis is presented in Appendix D, Option Analysis.

We recommend UCOP modernize its technology to more fully utilize corporate information. In addition, to resolve the Process Improvement Findings, UCOP should increase the education and training provided to all UCOP and campus staff and develop user groups that are responsible for the development and modifications necessary to maintain the most effective data warehousing environment. The following Action Plan describes the steps UCOP should follow to achieve a more effective corporate information systems environment.

SECTION VIII – ACTION PLAN

The University of California is approaching the 21st century facing several technical challenges. If UCOP is to continue its long tradition of providing the best information systems available, system modifications need to occur. This will require implementing the tools and technologies of the 21st century, some of which are described in Appendix A, Data Warehousing Technologies and Trends. Empowering the end-user by providing easy access to the information they require, by providing the analyst with capabilities to perform sophisticated, multidimensional querying, and by providing a user-friendly, fast, flexible and scaleable environment, is the road to UCOP's continued success.

UCOP should continue to provide data repository or warehousing capabilities. The campuses should continue to maintain their transactional systems and submit information to UCOP in a standard format within a prescribed schedule and be held accountable for such.

UCOP should approach this situation in two phases. Phase I should provide easily and readily accessible data to the end users. Phase II should establish a new data warehouse that consolidates the data from the campuses into an integrated, reliable, and secure environment.

The Option Matrix, Table 1, Appendix D, summarizes the technical issues identified by the users and the options that resolve a majority of those issues. We recommend UCOP consider combining the "Implement New Front-End Access Tools" Option and the "Implement New DBMS and Front-End Access Tools" Option. Combining these two options creates a two-phase implementation.

The first phase resolves some user access issues within a short period of time. It provides immediate benefits with minimal risk and disruption. The second phase requires significant analysis and evaluation time to determine the most efficient and cost effective solution. Phase I steps are described in detail later in this document, Phase II steps include:

- developing requirements
- developing package and vendor evaluation criteria
- researching products
- conducting functional, risk, and cost analysis
- selecting and procuring solution
- pilot testing
- conversion
- user training
- implementation

To execute both phases we have organized the actions into short and long-term steps. The short-term actions continue to utilize the existing environment while providing increased access

2/13/98

to the users. The long-term actions replace existing Corporate Systems with newer technology. Implementing any change requires support and participation from management, IR&C, and the user community.

Our recommended actions include the following:

Short-term Actions - Phase I
- Evaluate and implement new IBI Business Intelligence and other front-end tools
- Increase user access, training, and education of current Corporate Systems and new Business Intelligence tool capabilities to potentially reduce shadow systems

Long-term Actions - Phase II
- Evaluate and document user requirements (data elements, timing of data, source, computations, standard reports and screens, etc.)
- Evaluate integrated solutions
- Replace existing systems with a new Database Management System and new Business Intelligence tools
- Evaluate the current chargeback mechanism

To accomplish the short-term actions, UCOP should proceed with the following steps:

1. Provide newer, easier to use tools for users to access Corporate Systems

The easiest, most beneficial change UCOP can make is to provide query tools that are easy for users to learn and use. Currently, the query tools available to users is limited. These tools include using FOCUS commands for all databases or entering query parameters, using the Gateway product, into the Corporate Financial System. Using these tools requires substantial user knowledge of the data and the technology. Therefore, UCOP staff are limited in their accessibility to the data warehouse and in analytical manipulations.

Query tools that are compatible with FOCUS and operate in an VM/CMS environment are currently available from IBI. Appendix H, IBI FOCUS Data Warehousing Products, contains some product offerings from IBI that UCOP should review to make user access easier and provide more functionality.

UCOP should identify tools that will resolve the short-term need for access to the current Corporate Systems. The tools should be evaluated for potential inclusion in the longer-term recommendation of replacing the existing systems with a new Database Management System (DBMS).

The Option Matrix, Table 1, Appendix D, clearly indicates that a majority of user technical issues would be resolved with new access tools. UCOP should analyze IBI's offerings to determine which products most closely meet user requirements. The users should be

involved in defining the requirements, reviewing vendor demonstrated products, and the selection process.

Appendix E, Integration of FOCUS and Business Intelligence Tools, contains an article describing the integration of FOCUS for data management with Business Intelligence tools to create an environment that enables users to access data more easily. The environment included in the article is very similar to the environment operating at UCOP.

2. *Increase user access, training, and education of current Corporate Systems capabilities and limitations*

The campuses forward vast amounts of information to Corporate Systems for use in analysis and business operations. Since the systems are read-only, thereby eliminating any chance of manipulation from the user community, access to data should be limited only in the rarest of cases where confidentiality is a concern. Of course, management policy and discretion on access should be retained.

UCOP should adopt an attitude of information and knowledge sharing. Users rarely stray from the comfort of elements they understand because of fear of the results. Periodic training and education of Corporate Systems strengths and weaknesses would result in a more sophisticated user community gaining the ability to answer more sophisticated questions, more completely, and in less time. This training and education of users would be enhanced by a completely standardized, documented environment. Users would then be more comfortable referencing data outside of their normal working set and have a reference point to come back to. The result should be an increased number of sophisticated and effective end users.

As indicated in the IS&C report, UCOP is maintaining two separate operating environments. If a short-term decision were made to migrate Corporate Systems from VM/CMS to an MVS operating environment, it would change the selection of data warehousing products available. Therefore, any change in the operating environment should be fully analyzed to determine the impact to acquiring and implementing new Business Intelligence tools. Understanding the options available must be incorporated into the evaluation process.

To accomplish Phase II, implementation of the long-term actions, UCOP should proceed with the following steps:

3. *Evaluate and document "data" in business terms*

Not knowing the name of an element, where it is located, or what it represents is not an uncommon situation in a data warehousing environment. Fortunately, products are available, including IBM's DataGuide, which aid all levels of users by describing data in business terms and then launching applications to access the element.

Documenting the "data" should include defining all data elements, identifying the sources of data (which databases), warehouse schema, and user views (screens, queries, and report formats). This type of documentation will help users to locate specific information because they will know exactly what the data element represents, where it is located, and how to access the data element.

One of the most important steps in building a data warehouse is not only defining the data element but also defining the environment in which it exists. This is known as metadata. There are three levels of metadata: data source, warehouse, and user. The metadata provides a "catalog" of what is in the warehouse and data sources that provide input to the warehouse. User metadata identifies computed fields, summaries, and detailed information.

Metadata details the structure of the data and also relationships within the data (both from a database view, and also as a result of business rules and data flow descriptions). Aliases, code tables, default values, archiving methodology, units of measure (dollars or pounds), algorithms and other related information should also be documented.

A detailed description of the business rules that apply to the metadata should be developed. For example, security rules and access limits to certain data elements should be defined. On-line messages should be clearly articulated and offer alternative methods or solutions when appropriate.

4. *Establish evaluation criteria and strategy*

Prior to implementing any changes, UCOP must develop evaluation criteria. The criteria should address the following at a minimum:

- compliance with business/functional requirements, defined in Step 4
- level of risk acceptable to UCOP
    - vendor performance, track record, client references
    - vendor financial stability
    - degree of change anticipated (impact to users, technical staff, analysts)
    - length of time required to achieve implementation
    - quantity of resources required
- cost
    - software
    - maintenance and upgrades
    - hardware, if required
    - solution development and support
    - pilot testing
    - conversion support
    - procedure development/user documentation
    - training (technical and user)

After development of evaluation criteria, UCOP should assign a point scheme for comparing various solutions. More points should be allocated to higher priority items. A consistent procedure for evaluating alternatives should be established prior to evaluation to ensure consistent scoring by the evaluation team.

5. *Choose an integrated solution vs. a Best of Breed approach*

   When researching the long-term recommendation of replacing the current FOCUS Corporate Systems with new DBMS and access (Business Intelligence) tools, UCOP should strongly consider a single vendor providing a fully integrated data warehouse and business intelligence tools.

   Integrating the various products discussed earlier (e.g., Decision Support Systems, Executive Information Systems and OnLine Analytical Processing tools) is not an insignificant issue. A best of breed approach is *not* recommended due to the degree of difficulty encountered during implementation. Procuring a single integrated solution vs. multiple products will maximize performance and reduce implementation problems. Additionally, with a lack of robust tools to assist in the implementation and management of the data warehouse, much of the total cost of ownership, according to Gartner Group: *Datawarehouses: Clarifying the Hype and Confusion*, comes from resources that "tie" the products together and create the architecture.

6. *Review and select a Database Management System (DBMS) that meets at least 80% of UCOP business requirements*

   The DBMS used for the data warehouse can be a standard relational database or a multidimensional database. The drawback of a relational database is that it is not well suited for performing multiple, cross-database queries typically required of a data warehouse. For example, a standard year-to-year comparison of students enrolled at a particular campus would require at least one query for each year, then another query to compare them. A multiyear trend analysis could require many relatively slow queries.

   For this reason, a Multi-Dimensional DBMS in a data mart arrangement is suggested. This would allow the diverse users within UCOP to effectively manipulate information specific to their department needs. In other words, the RDBMS is designed to optimize data sharing across a single business function, whereas the MDDBMS is designed to access information across multiple business functions. An example would include combining financial information with students, grants, or facilities information. In addition, MDDBMS's allow users to augment existing data with additional information without interfering with the other user department requirements.

   The decision to create a shadow system (or data mart) on a multi-dimensional DBMS *and/or* an RDBMS, depends on the strategic course taken by UCOP and the tactical decisions required for execution.

UCOP should thoroughly review the strengths and weaknesses of data warehouse solutions and Business Intelligence tools available. Refer to Appendix F, Integrated Warehouse Suite Vendors, and Appendix G, Data Warehousing Products, for information regarding various vendors.

Other Considerations

Implementation of the long-term recommendations may require the COBOL, PL1 and FOCUS routines be modified. Also, the routines may not be portable to the new software environment. If this occurs, UCOP should consider creating new routines in a more efficient or compatible language. It is probable that a new DBMS and any new data elements being captured would render the value of these programs less vital because significant changes will probably have to be made. The cost of migrating to a new operating system or creating new routines was not included in the scope of this project.

During the evaluation process, another option UCOP should consider is linking Corporate Systems databases to one of the most dynamic areas in the computer industry, i.e. the World Wide Web. This will provide power users with access to even greater resources and more information. UCOP is familiar with the power of this medium and is currently using it to communicate information on its home page.

Most of the major industry server vendors have introduced Web servers and products that directly support DBMSs. Sample include:

Web Servers	Server Products
IBM's Internet Connection Secure Server	IBM's DB2
Microsoft's Internet Information Server	Microsoft's SQL Server
Netscape's LiveWire Pro	Informix Software's DBMS

Appendix A – Data Warehousing Technologies and Trends

Data repositories or warehouses are comprehensive databases of information processed and forwarded from transactional systems. The transactional systems retain the day-to-day detailed activities and forward summary information to the data warehouse on a regular basis. UCOP was on the leading edge of this technology by summarizing campus transactional data and storing it in a data warehouse type environment 15 years ago.

Data warehouse databases are typically read-only and allow users to perform analysis and trending on consistent information unlike transactional system data that is constantly being updated or changed. Data warehouses consist of a database and code that facilitates data access and manipulation. Since neither UCOP, nor any other entity can anticipate the multitude of queries that may be executed against the databases, data warehouses are designed to be open architectures. Open architectures allow access to information stored in different software programs and hardware systems.

Access to information has been a long-standing need of decision-makers and users at all levels of any organization. However, this information is often not available in a format that can be used for meaningful analysis by multiple users. Technical constraints have historically been the cause of limited access and inflexible reports. As a result, information was primarily produced in standard hard copy reports.

In recent years, however, techniques have been developed to allow greater access to more information by more users. Traditionally users had to access information in one system then access information in another system as a second step. Today's new data integration tools make information accessible across the entire system rather than by program. These new techniques can be efficiently used in multiple applications and projects.

Making this information ubiquitously available often requires multilayered processing, with different software products to support each layer. For example, legacy data may require replication software in order to periodically extract data and map it to a relational Data Base Management System (DBMS). Data in the relational DBMS may, in turn, be summarized and aggregated before being imported into a data delivery vehicle, from where it may be accessed by client software that is designed to perform specific functions on behalf of specific users.

Creating and managing a well thought out environment requires an integrated approach rather than a single applications-based focus. A data warehousing process consists of five functional components, as follows:

1. Source of the data – Campus transactional systems (Outside the scope of this project)
2. Load – Loading, editing and validating data provided by the campuses
3. Storage – Storing the validated and consolidated information in the data warehouse

4. Query – Data access by the user community
5. Metadata – Foundation for the other four components including data element definitions, source of data, composition of data, use of data (calculations, report formats, etc.), processing criteria, restrictions and security

Data warehousing also requires five corresponding levels of support, including:

1. Operational data source
2. Data conversion and extraction
3. Data warehouse DBMS
4. Business Intelligence (BI) tools
5. Data warehouse administration

The first three components in both listings address the movement of data from the campuses to the mainframe. The Query and BI tool components support the decision making needs of UCOP and the Metadata and administration components provide information about the data, controls and interactions (i.e., business rules).

In a recent survey of industry leading Corporate Information Officers, conducted by Deloitte &Touche, data warehousing and its accompanying BI tools are seen as the *number one* most important technologies to be utilized within the next two years. Although UCOP established its' corporate environment long before the term "data warehouse" came into vogue, or comprehensive architectures were developed, it's important to note this increased usage because of the variety and maturation of tools available. Today, data warehousing technology and architectures are increasingly a mainstream activity.

In order to fully understand if and how newer technologies should be pursued in meeting user requirements, the following common foundation of data warehousing tools and techniques, trends, and types of products is provided. Specific vendor related tools and products are contained in Appendix F, Integrated Warehouse Suite Vendors.

Data Warehouse Data Base Management System (DBMS)

The database is the heart of any data warehouse. The database design for an online analytical processing (OLAP) system (i.e., Corporate Systems data warehouse) is as important as the design for an online transaction processing (OLTP) system (i.e., campus systems). However, each of these designs is intended to provide totally different results. OLTP tables are very detail-oriented, whereas OLAP typically uses aggregations, summaries, and time-oriented structures, which serve the intended purpose better.

The DBMS vendors are significantly improving support for data warehousing environments. Currently, databases are able to support complex Decision Support System schemas with databases approaching *several hundred terabytes*. By comparison, UCOP's databases are in the gigabyte range. Improved query algorithms used in the decision support tools also enhance the improved support.

A data warehouse environment is typically sized by the number of users accessing the system at a point in time and the size of the database. As an aid in determining specific vendor's products applicable to UCOP's environment, the following definitions are provided:

Data Warehouse Systems Definitions

Concurrent User Population:
Small: Fewer than 10
Medium: 10-30
Large: 30-100
Very Large: More than 100

Database Size (Gbytes):
Small: Fewer than 30
Medium: 30-100
Large: 100-500
Very Large: More than 500

Oracle, Sybase, and Informix have announced comparable versions of their relational data base management systems. These vendors support formal data warehouse programs and tout their ability to build and support terabyte-capacity warehouses. They have partnerships with tool and utility vendors to enrich their data warehouse portfolios. They have also acquired key data warehouse technologists to bolster their product offerings.

Oracle claims that Oracle Warehouse is the most widely used platform for building data warehouses. It also claims a 28 percent share of planned data warehouses, while Sybase estimates 15 percent and Informix only 5 percent. Over 1,000 Oracle Warehouse systems are in use around the world and hundreds are in production environments supporting data sets ranging from gigabytes to over a terabyte.

Another leader in the data warehouse market is Red Brick Systems. They claim to be the first relational DBMS company to focus exclusively on the warehouse market--a market that grows by 65 percent each year. Red Brick was recently honored with a first-place "Data Base Software" award by the World Class Solutions for Data Warehousing program, which is conducted by Data Management Review in conjunction with International Data Corp. (IDC).

An IDC study examined the financial impact of data warehouses on core processes contributing to an organization's success. The study was based on 62 case studies of organizations with successful data warehouses. Red Brick, IBM, Informix and SAS were finalists among the data warehouse vendors nominated by users in the study. The award was presented to Red Brick at a ceremony held in June 1996 in Santa Clara, CA.

Business Intelligence (BI) tools

Business Intelligence (BI) includes the exploration of data, establishment of data relationships and trend analysis. It's an iterative process of accessing data, analyzing data, deriving insights, drawing conclusions and communicating findings. Business Intelligence tools are comprised of four major product segments:

1. Interactive query and reporting tools
2. Data mining tools
3. Decision Support Systems (DSS)
4. Executive Information Systems (EIS)

Interactive query and reporting tools are standard methods for extracting and printing selected information. Data mining techniques are usually rules-based tools that identify patterns or exceptions in data. Data mining attempts to creatively analyze data (e.g. discovering unusual trends) while DSS allows manipulation of data for online query and report purposes. Executive Information Systems pre-establish summary statistics that are updated each time the data warehouse is updated. This view provides management with an online executive view of the organization.

Decision Support Systems (DSS)

Decision support software enables analysis of business information in the data warehouse with tools that are designed for forecasting, financial modeling, simulation and visualization. Solutions can be developed based on techniques such as statistical analysis, time series analysis, operations research, spreadsheet analysis and geographic information systems, e.g. buildings on a campus with related information.

Many of the exciting developments in the data warehouse industry are occurring in the DSS arena. Graphical report writers, multidimensional query engines and OLAP products are the keys to the sophisticated access, analysis and reporting being sought by corporate users. Online analytical processing takes DSS a step further. OLAP is an analytical processing technology that extracts information from databases using mathematical computations and data-processing techniques. This information can provide statistics, trend analysis, and forecasting in a flexible, interactive manner.

Executive Information Systems (EIS)

An Executive Information System (EIS) is an application program specifically designed for executive management. Presentation of material is structured in a "board briefing book" concept. Detailed information is available "behind" the management information and is accessible by using a concept known as "drilling." The ability to view corporate information at a glance and drill down through the detail used to compile the management information was expressed repeatedly by UCOP users.

Administration

Data warehousing creates many complex administrative issues that are very different from transactional applications issues. Because multiple subject areas (i.e., student, financial, equipment, etc.) and large volumes of historical data (i.e., more than five years) should be maintained, a data warehouse will require significant amounts of disk storage, and substantial CPU resources to execute applications. These new administration challenges require extensive planning to address the following activities:

- data usage auditing
- business data modeling
- directory management
- chargeback
- summary tables
- security
- service request prioritization
- query catalog and subscription services
- managing of operational data extracts

Market Trends and Techniques

Decision Support Systems, Executive Information System and RDBMS vendors are continuing to expand the market as corporations continue to implement data warehousing strategies. As corporate information becomes more easily accessible, vendors will continue to develop more user applications. Products with highly developed, multi-user, open architectures will continue to expand data warehousing capabilities.

A question remains, however, as to how much of this market will go to current vendors of OLAP and RDBMS products. These vendors, after all, are accustomed to dealing with relatively small-scale applications, with a small number of users accessing a limited amount of data. Even the most powerful of current-day products are designed for individual end users or small LANs. They are not equipped, for example, to perform parallel data loading operations that take advantage of current hardware technology. Similarly, the advanced indexing technologies (such as bit-mapped indexing) that are becoming available by the industry-leading RDBMSs are beyond the scope of these small-scale systems. DSS products must then be used as adjuncts to the larger RDBMS products that can take advantage of these technologies.

At the same time, business analysis modeling is becoming increasingly specialized, and new techniques for examining business data, such as data mining, are being developed. These new ways of viewing a business are often industry-specific. As a result, the coming years may bring to market DSS products that are customized to individual industries. This trend is well under way already. In addition, existing industry-specific vertical applications may increasingly incorporate DSS capabilities for the same reason.

Other Techniques

A *data mart* is a technique that allows users to augment existing corporate data without redundant data entry. A data mart is a subset of data from the data warehouse, designed to support the unique business requirements of a specific unit or application. This technique allows users to request a specific set of data and add additional information to the extracted subset. With a relational DBMS at the hub of a data warehouse architecture, copy management - or replication services - should be used to populate the data mart. Copy management allows the manipulation of data to tailor the data mart (i.e., time series data or summarized data) to the specific Decision Support Systems (DSS) application, without performing additional or unique extraction processes against the operational data stores.
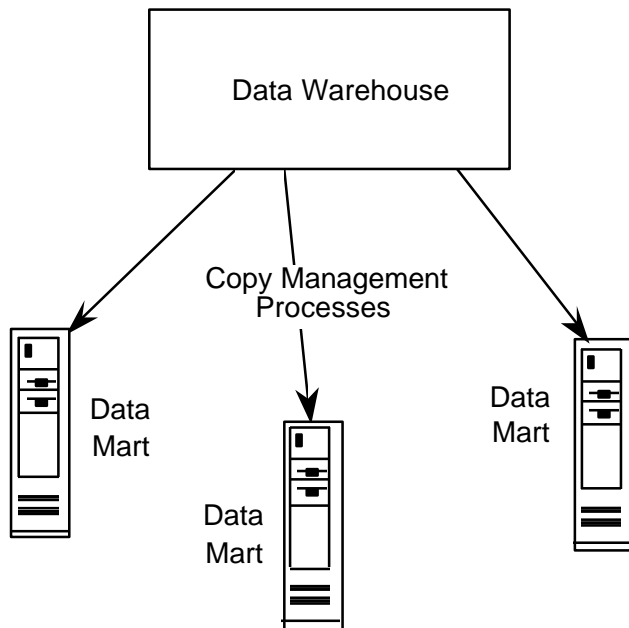
Source: Gartner Group

Appendix B – Interview Listing

Interview Listing	Interview Listing
Assily, Ruth	Litrownik, James
Beccar-Varela, Gabriel	Litrownik, Mona
Berman, Carol	Lynn, Stuart
Blaschczyk, Helmut	Marino, Marit
Cage, George	Nevel, Sergio
Cate, Joanne	Nixon, Dorothy
Cate, Mike	Ohy, Jorge
Colley, Greg	Pang, Lily
Coughlin, Kevin	Plebuch, Alice
Coy, Judy	Plotts, John
Cruz, Mark	Quick, Suzanne
Dolgonas, Jim	Raffetto, Carla
Estrada, Carmen	Rider, Sandra
Giacomi, Jon	Ross, Ellie
Gibson, Sam	Santee, Richard
Good, Jon	Scronce, David
Harrington, Connie	Skarakis, Eleanor
Harroun, Jeannette	Stanton, Mel
Hart, Kimberley	Strauman, Joan
Haskins, David	Strem, Charlotte
House, Suzanne	Trybulski, Jeremy
Hutton, Cate	Uemura, Catherine
Kodish, Jan	Uyeno, Katherine
Kowarsky, Judy	Vecchi, Karen
Krotzer, Carole	Wesner, Kathleen
Kunkle, Robert	White, David
Leong, Margaret	Wong, Carol
Lester, Barbara	Yastishak, Robert
Levin, Lubbe	Young, Randy

Appendix C – Focus Group Meetings Listing

Count	Equip. and Facilities	Contracts and Grants	Personnel	Student	Financial
1	Joanne Cate	Jeannette Harroun	Karen Vecchi	Richard Santee	Barbara Lester
2	Marsha Sato	Carol Berman	Margaret Leong	Mark Langberg	Joan Strauman
3	Suzanne Harrington-Cole	Jorge Ohy	David Scronce	Patti Mizuiri	Carol Berman
4	David Haskins	Gregory Colley	Lily Pang	Michael Cate	Helmut Blaschczyk
5	Carla Raffetto	Suzanne Quick	Helmut Blaschczyk	Judy Coy	Barbara Yoder
6	Judy Coy	Jeremy Trybulski	Barbara Yoder		Carole Krotzer
7		Barbara Yoder	Eleanor Skarakis		Richard Santee
8		Dorothy Nixon	Laura Martinez		Dorothy Nixon
9		Carla Raffetto	Richard Santee		Kevin Coughlin
10		Judy Coy	Suzanne Quick		Carol Wong
11			Joan Strauman		Susan Ohye
12			Jim Litrownick		Judy Coy
13			Marit Marino		
14			Susan Ohye		
15			Judy Coy		

Appendix D – Option Analysis

Based on our findings and the availability of newer technologies and techniques, several options were analyzed to address Corporate Systems data warehousing challenges. The most viable options include software changes or upgrades. Replacing existing hardware with a newer or higher capacity platform may provide more storage, may retain the existing knowledge base and may alter current chargeback processes, but it will not resolve the majority of the issues identified during the user interviews. For these reasons, hardware changes are not included as part of the options analysis even though it may be part of the ultimate solution. Viable options include:

- Continue to Operate Corporate Systems with No Changes
- Implement New Front-End Access Tools (e.g. Decision Support Systems, Executive Information System, Graphical User Interface front-end)
- Implement a New Data Warehouse and Front-End Access Tools

For each Option analyzed, the objective, benefits, risks and issues have been included. Remaining status quo or implementing changes provides both benefits to UCOP and includes varying amounts of risk, which must be considered. All Options considered require user education and training.

Prior to implementing any changes, UCOP must develop evaluation criteria and determine the financial impact of the changes. Performing a cost/benefit analysis of the changes was outside the scope of this project but should be conducted to identify the most viable solution. Additionally, the option of migrating to an MVS environment is included due to the effect it would have on vendor product offerings and because it was an issue recently reviewed by IS&C.

Continue to Operate Corporate Systems with No Changes

Objective: Continue to perform collection and consolidation of data from campuses while avoiding disruption to current operations

*Potential Benefits***:**
- No "battles" of change to be fought
- No operational or technical disruption
- Technical environment is well understood
- Operational environment is understood by key individuals
- No increased cost to continue operating as-is

*Potential Risks***:**
- UCOP technology will lag behind the technology utilized at the campuses which may be perceived as lack of efficiency or customer service

- As business requirements continue to change UCOP may encounter increased difficulty reporting information
- Continued maintenance of multiple operating system environments which stretch existing IR&C resources
- More "shadow" systems may be developed to avoid direct interaction with the Corporate Systems which may result in increased redundancies and greater total cost of ownership to UCOP

*Issues***:**
- Continuing to operate in the current hardware and software environments will not reduce the re-charge structure which is causing some users to avoid using the Corporate Systems
- Easy access to information for reporting and analysis is not achieved
- Perceived lack of customer service resulting in continued user frustrations

Assumptions:
- Users will receive education and training on existing Corporate Systems
- Users will receive increased access to Corporate Systems, if requested

Implement New Front-End Access Tools (e.g. Decision Support Systems, Executive Information System, Graphical User Interface front-end)

Objective: Increase accessibility to data, reporting capabilities and manipulation of data

*Potential Benefits***:**
- Resolves most of the issues identified by the users with a minimal change in existing Corporate Systems
- Minimal disruption to current operations
- Eases the difficulty in linking Corporate Systems together and internal databases together to increase reporting capabilities
- Retains existing load, edit and validation routines
- Continues to utilize existing knowledge of system operations

*Potential Risks***:**
- Increased usage by users may increase their chargeback costs which may backlash and cause more shadow systems to be created

*Issues***:**
- Increased cost for new software
- Increased cost for training technical staff and users
- May need to review the current chargeback mechanism based on increased user usage

Assumptions:
- Users will receive education and training on Corporate Systems and the Business Intelligence tools
- Users will receive increased access to Corporate Systems, at management's discretion

Implement a New Data Warehouse and Front-End Access Tools

Objective: Increase integration of Corporate Systems information, ease accessibility to data, and reduce user reliance upon "shadow" systems

*Potential Benefits***:**
- Resolves a majority of the issues identified by the users (refer to Corporate Systems Analysis section for issues)
- Provides an opportunity to develop an environment based on current business requirements
- Utilizes an environment that is flexible to meet business changes
- Provides an opportunity for users to define data organization and storage requirements (user ownership and empowerment)
- Reduces the total cost of ownership regarding PC systems and software packages as shadow systems decrease
- UCOP can take advantage of development and access tools and techniques that were not available 15 years ago
- UCOP may be in the forefront to guiding the implementation of technology at the campuses

*Potential Risks***:**
- User participation/disruption will be significant due to requirements development meetings, program piloting, testing, conversion, and implementation activities
- Users may not have the skillsets and willingness to change

*Issues***:**
- Increased cost for new software
- Increased cost for training technical staff and users
- Significant time is required to research, select and implement a new data warehouse and tools, if the time is excessive users may lose interest and momentum towards the new system

Assumptions:
- New load, edit and validation routines will be developed in the new environment
- User groups will be established to define requirements and evaluate warehouse and business intelligence tools
- Users will receive education and training on the new Corporate Systems and Business Intelligence tools
- Users will receive increased access to the new Corporate Systems, at management's discretion

Ancillary Issue: Migrate Corporate Systems from VM/CMS to MVS

Objective: Increase the number of product offerings by IBI, or other vendors, and reduce the level of technical support effort expended while maintaining the current mainframe investment

*Potential Benefits***:**
- Increase the number of data warehousing products available to end users
- Retains existing load, edit and validation routines
- MVS technical environment currently understood by IS&C
- Eliminates the use of an operating system receiving minimal support from the vendor
- Minimal investment cost
- Reduction of operating system license cost and support
- Ability to focus on a less heterogeneous operating system environment

*Potential Risks***:**
- Technical and operational disruption would occur during migration from VM/CMS to MVS

*Issues***:**
- Migrating to an MVS environment will minimally reduce the cost of operations by eliminating the VM/CMS costs
- User issues, such as easy access to information for reporting and analysis, will not be improved
- Lack of user-friendly query tools, graphical user interfaces and state-of-the-art technologies may continued to be perceived by users as a lack of commitment to providing efficient customer service

Assumptions:
- Users will receive education and training on Corporate Systems after the migration
- Users will receive increased access to Corporate Systems, if requested

Option Matrix
Table 1

The following Option Matrix, Table 1, was developed to illustrate the resolution of Technology Findings with the Options identified. Technology Findings are detailed in Section VI, Corporate System Findings. Again, the impact of hardware changes resolves few issues thus it is not included as an Option. In addition, the fiscal impact of any of the Options was not evaluated. UCOP must consider the software package cost as well as evaluation, selection, implementation, documentation and maintenance costs incurred during a new system implementation. If the box is "blank" the issue/finding is not resolved by that Option. Process Improvement Findings have not been included in this analysis matrix.

Item	Technology Finding	No Changes[1]	Migrate to MVS[2]	Implement New Access Tools[3]	Implement New DBMS and Tools[4]
1	System lacks an easy-to-use interface for querying and reporting			Resolved	Resolved
2	The existence of shadow systems is prevalent			Partial Resolution[5]	Partial Resolution[5]
3	Ability to drill-up/down on information is difficult or not possible			Resolved	Resolved
4	Not enough historical information is being stored online	Not Resolved[6]	Not Resolved[6]	Not Resolved[6]	Not Resolved[6]
5	Corporate Systems lack a multi-dimensional querying capability			Partial Resolution	Resolved
6	Lack of integration with internal and external databases increases staff workload			Not Resolved	Resolved

---

[1] Continue to Operate Corporate Systems with No Changes

[2] Migrate Corporate Systems from VM/CMS to MVS with No Other Changes

[3] Implement New Business Intelligence Tools (e.g. Decision Support Systems, Executive Information System, Graphical User Interface front-end)

[4] Implement a New Data Warehouse and New Business Intelligence Tools

[5] New tools will provide greater access but if the cost to access the system is not acceptable to the users the shadow system may persist

[6] Resolution of this Technology Findings may be dependent upon new hardware

Appendix E – Integration of FOCUS and Business Intelligence Tools

Microcontroller Product Group of Philips Semiconductors, in Sunnyvale, CA, established an integrated environment utilizing options available to UCOP. The following description of their situation is included here because of both the hardware and software similarities to UCOP's environment and the desire brought forth in the IR&C Migration from VM/CMS to MVS report.

P.J. Matarese, a programmer/analyst, and other members of the Microcontroller Product Group quickly developed a data warehouse using a warehouse product suite from Information Builders, Inc. They used mainframe FOCUS for data management, EDA/SQL as the middleware layer for data transfer, and FOCUS Executive Information System (EIS) for the front end.

"Considering that much of our data was in FOCUS, and we already had FOCUS expertise, we began by looking at front end products that use the FOCUS engine to extract data and display it on the desktop using a graphical user interface," Matarese recalls. "We used FOCUS EIS to build the user interface and FOCUS Reporter for Windows to handle the data extracts. The EDA/SQL middleware layer makes the physical location of global data totally transparent to the end user."

Since the division had a host system at its disposal, they decided to use the mainframe as a server. The data warehouse repository is currently on an IBM host, running FOCUS under the VM/CMS operating system. Matarese and his colleagues are evaluating UNIX servers from Hewlett-Packard as an eventual mainframe replacement to host the data warehouse. The FOCUS database is portable from one platform to the other.

With help from Information Builders consultants, the entire data warehouse and Decision Support Systems (DSS) was constructed in just four months. The results have been spectacular. "The system drastically reduced my paperwork burden," Matarese stresses. Before the data warehouse was installed, I typically generated 200 to 300 pounds of paper every month. Today, my need to print is virtually eliminated as most information is available on screen and individual end users can print their own reports as required."

Users can personally generate hundreds of reports by browsing through the data warehouse report menus within the Executive Information System. Even unique or ad hoc reports can be handled by end users, giving the corporate database new visibility.

"The benefits of giving end users direct access to reports goes beyond saving the time of programmers, freeing up programming resources for mission-critical development work is significant," Matarese says. "The most specific benefit of the data warehouse is that Philips Semiconductor decision makers can get immediate feedback on market trends."

Appendix F – Integrated Warehouse Suite Vendors

The following is a list of companies that offer total data warehousing solutions:

1. DEC
2. HP
3. IBI
4. IBM
5. INTERSOLV
6. Oracle
7. Platinum Technology
8. Praxis International, Inc.
9. Prism Solutions
10. Pyramid Technologies
11. Red Brick Systems
12. SAS Institute, Inc.
13. Software AG
14. Sybase

Appendix G – Data Warehousing Products

The following provides an overview of some of the major data warehousing vendors and their products. It is not meant as an endorsement of any firm or product line. The primary basis of the information is from a Faulkner article entitled *Data Warehousing Market Status,* 9/96.

IBM Information Warehouse

IBM announced the Information Warehouse in September 1991. Its goal is to provide open access across multiple hardware platforms and vendor products. The framework includes a published architecture that explains how IBM and third party products work together.

IBM offers many systems-oriented products designed to facilitate the development and management of large-scale data warehouses. Included are replication and transformation utilities (DataPropagator and DataRefresher), data gateways (DataJoiner), and metadata management (DataGuide).

In addition, IBM offers a parallel version of DB2 called DB2 Parallel Edition (PE), which runs on a MPP version of IBM's RS/6000 UNIX server called the SP2. DB2 PE also runs on LAN-connected RS/6000 servers. Though immature, DB2 PE offers excellent scalability potential.

POWERquery for SP2 is IBM's RISC- and AIX-based parallel data base server marketed for decision support. It includes the DB2 Parallel Edition for AIX and the RAID-based 7135 RAIDiant Array, as well many other IBM services. The S/390 Parallel Query Server is IBM's equivalent MVS offering.

In 1995, IBM's PC Company announced Visual Warehouse, a low-cost ($23,000) bundle of query tools, DB2/2 database, connectivity software, and data warehouse management tools. Since Visual Warehouse supports only 50G bytes, it is a starter package for end users and a base technology for VARs. Due to its low price and limited scalability, Visual Warehouse does not offer customers a seamless path for scaling a first-time data warehouse into the hundred gigabyte or terabyte range.

IBM markets several software products under its Information Warehouse framework:

- Data base management systems, including DB2, DB2/2, DB2/6000, SQL/DS, OS/400 Data Manager, IMS DB, and IMS Client Server/2.

- Data delivery products, including Data Propagator, Data Propagator Relational, Data Refresher (formerly DXT)--All products that support the Distributed Relational Data Architecture (DRDA), Information Builders' Enterprise Data Access (EDA)/SQL, and EDA/Dynamic Extender for OS/2.

- Warehouse Managers such as DataGuide/2, and CDF/MVS.

- Applications and DSS/EIS, including AIX Query/6000, Query Management Facility (QMF), Application System (AS), Personal Application System/2 (PAS/2), Lotus 1-2-3/M for MVS or VM, Metaphor Data Interpretation System, Comshare products, and Intelligent Office Company products.

- Management tools such as DataHub, which integrate multiple DBMSs into a single view.

IBM provides advanced warehousing tools that give end users an enterprise-level view of data and enable data sharing among DRDA-compliant DBMSs. The IBM offering is strong. Its DRDA standard offers high relational functionality compared to standards that supply the lowest common denominator (like ODBC). IBM also provides management tools such as end-user data views, conversion tools, data migration tools, and data replication tools that provide much of the functionality required for big information warehouses.

One well-known IBM product, Intelligent Miner, provides information analysts with a common framework for using all IBM's current and emerging data-mining techniques. Intelligent Miner's preprocessing capability is considered one of the best in the industry. It is scaleable and flexible, and adapts to data base size and type, and to computer architecture. It can be used with IBM's DB2 family, other relational data bases and flat files, or with extracts from data warehouses and data marts.

IBM offers Data Interpretation System (DIS) as part of its data-mining suite. DIS provides powerful multidimensional analysis to scaleable, open clients. With DIS, decision-making teams can create simple or complex dimensional views of business metrics, run business functions and statistics, and present data quickly.

Hewlett-Packard OpenWarehouse

OpenWarehouse is HP's framework for delivering IW products and services. HP uses the HP 9000 Series 800 Business Servers as the platform for housing the staged database, warehouse manager, and connectivity products. In traditional HP fashion, OpenWarehouse is the result of collaboration with several independent vendors of leading-technology software.

HP markets various software products under its OpenWarehouse framework:

- Data base management systems including HP ALLBASE/SQL, Informix RDBMS, Ingres RDBMS, Red Brick Warehouse RQS, and Sybase SQL Server.

- Warehouse managers such as Prism Warehouse Manager, Vality Integrity, and Evolutionary Technologies ET1 Extract.

- Connectivity software including HP Data Access Language (DAL), Red Brick DIS Gateway, Sybase Open Client/Open Server, Information Builders EDA/SQL, Metaphor DIS Data base Gateway, Microsoft ODBC, and the Open Distributed Environment (ODE) Remote Procedure Call (RPC).

HP supplements products from third-party vendors with its Open Warehouse Intelligent Warehouse Solution (IWS). IWS is a suite of tools designed to manage the warehouse across multiple databases and individual products. It provides such capabilities as replication, table partitioning (allowing tables to be distributed across several disks so they can be searched simultaneously), synonyms for table names (so it is easier for end users to locate information), security features, and query logs. IWS can reside on any TCP/IP-based server with a relational DBMS.

Digital Equipment ACCESSWORKS

Digital Equipment bases its data warehouse strategy on its hardware platforms and its ACCESSWORKS Integrator/DB middleware. Digital offers its 64-bit Alpha servers combined with RAID technology, personal computing, and clustering devices.

ACCESSWORKS DB Integrator is more than a mapping utility. It provides advanced features for distributed computing at the system level. These advanced features include a distributed query optimizer, a parallel query capability for queries across multiple platforms, and a global catalog for maintaining stored and multistatement procedures. Normally, these features are available only in high-end distributed RDBMSs like Oracle and Sybase. Their availability at the system level provides a mechanism for heterogeneous distributed data processing within the information warehouse and production environments.

External functions are a feature that allows the DBA to transform and verify data, and manipulate text strings through 3GL routines. It is also an optimizer for user-generated queries and applications, performing processes such as simplifying the nesting structure of 4GL-generated queries.

Digital Equipment offers consulting and third party solutions as part of its data warehousing strategy. Some partners are:

- DBMS--Informix, Oracle, Red Brick Systems, and Sybase.

- Decision Support Systems--Information Builders, Software AG, SAS Institute, IMRS/Pillar, Pilot Software, Holistic Systems, and Comshare.

- Specialized Data Warehouse Tools--Prism Solutions, Carleton, and ETI.

- Connectivity Software--Information Builders (EDA/SQL), and Digital Equipment (ACCESSWORKS/DB Integrator).

- Data Access Tools (via ODBC)--Microsoft, Brio, Clear Access, Business Objects, Andyne, and Trinzic.

Unisys DataCentral Server System

Unisys bases its information warehouse strategy on its DataCentral Server System. DataCentral, a parallel database machine, is suitable for decision support and transaction processing. To optimize transaction processing, DataCentral uses its large memory bank (up to

14G bytes, mirrored) as a virtual disk. This minimizes I/O operations and quickens response times. DataCentral uses disk striping and dedicated I/O processors to optimize query processing. Unisys emphasizes immediate access to current information as a product strength.

Unisys' client/server architecture is based on the UniAccess product suite provided by Applied Information Sciences. It is a source-code based, licensed implementation of the SYBASE Open Client/Open Server Architecture and ODBC APIs. DataCentral can be configured with relational, network, and index sequential/random access DBMSs.

NCR's Scaleable Data Warehouse

NCR (formerly AT&T GIS) will be an independent company in January 1997. The new management is currently developing marketing strategies and products. One such product is its Scaleable Data Warehouse. NCR's Computer Systems Group will be responsible for selling WorldMark servers, Teradata RDBMS, call center solutions, data warehouse programs, and other hardware and software.

The data warehousing industry began in 1985 when Teradata, which was acquired by NCR in 1991, shipped the first complete parallel query environments to large corporate customers. In 1994 and 1995, NCR updated its data warehousing program to embrace client/server tools, open systems platforms, a broad range of consulting services, nonproprietary pricing, and corporate demands for quicker, more cost-effective IT implementations.

The building blocks of NCR's Scaleable Data Warehouse framework are its multiprocessing WorldMark servers, the Teradata data base, and data warehouse consulting services. The NCR data warehouse strategy includes:

- The WorldMark server, available in SMP, cluster, or MPP configurations.

- The Teradata database, which scales from 10G to 10T bytes and now runs on UNIX and third party databases.

- Support of products from third-party vendors:
  - Data transformation--Prism, Carleton, CPM, Evolutionary Technologies, Apertus Praxis
  - Data base--Oracle, Informix, Sybase, CA-Ingres, Red Brick, Arbor, Microsoft
  - Data transfer--Praxis
  - Data mining--Angoss, HNC, SAS, HyperParallel, Cross/Z
  - Query/analysis--MicroStrategy, Pilot Software, Cognos, Andyne, PowerBuilder, Platinum, Holistics, Kenan, Comshare, Business Objects, Information Builders, SAS, Brio, IQ Software, Gentium, Visual Basic, Information Advantage, Software AG.

NCR relies on its data warehousing initiative to bolster revenues. To address the tremendous potential of data warehousing, NCR has two "starter" programs to help companies prototype and pilot their first data warehouses:

- First Step—This program helps customers build a data warehouse prototype (to justify a full-scale pilot) in 30 to 45 days. First Step costs $30,000 and provides evaluation hardware and software (relational data base and query/analysis tool) as well as 15 days of on-site consulting and training services. It is based on the WorldMark 4100 (evaluation copy) platform.

- RightSTART—RightSTART helps companies build a pilot data warehouse that can support up to 20G bytes and about 20 concurrent users. Unlike the old NCR QuickStart program, RightSTART does not restrict users to specific tools. RightSTART costs about $800,000 and delivers a working pilot in 90 days. A unique feature is that users can upgrade the pilot hardware to support more users and data. Compared to IBM's Visual Warehouse, which tops out at 50G bytes, customers can scale RightSTART data warehouses to the 100G-byte range if they use commercial data bases, or the terabyte range if they use Teradata DBS for UNIX.

Tandem Computers Data Warehouse

Tandem Computers pioneered the technology of a scaleable multiprocessing architecture and query parsing. Since the technology has gained popularity, Tandem can capitalize on its expertise.

Tandem is trying to change the market perception that it sells expensive, proprietary systems. It has ported its NonStop SQL/MP relational data base and its line of fault-tolerant MPP Himalaya servers to UNIX, and offers commercial relational data bases (e.g., Oracle, Sybase, and Informix) on SMP machines it OEMs from Silicon Graphics.
Tandem's data warehouse strategy revolves around the NonStop SQL/MP (Massively Parallel) RDBMS, the NonStop ODBC Server, and its storage architecture. DBMS features a parallel query capability that allows one query to be parsed and distributed among multiple processors. Combined with the scalability of Tandem's hardware platforms, this technique addresses a wide range of complex applications.

Theoretically, a query divided among 100 processors with Tandem's linear scalability can reduce processing time by 99 percent. Tandem emphasizes that it has delivered this level of performance and scalability for years, and is equally advanced in managing terabyte-sized databases.

Tandem has long provided storage technology with dynamic partitioning across thousands of disks. Unique performance algorithms reduce required disk space for sorting activities over competing SQL data base products.

Tandem's warehouse architecture combines its scaleable Himalaya hardware, NonStop Kernel operating system, NonStop SQL/RDBMS, and TorusNet interconnectivity technology in a fully integrated parallel system. The NonStop SQL/MP, an ANSI SQL RDBMS, supports a multiterabyte database.

Tandem provides open interfaces through its NonStop ODBC server, Sybase SQL server, and Apple's Data Access Language. Tandem also offers the Prism Warehouse Manager and Directory Manager, Evolutionary Technology's Extract Tool Suite, and Information Builder's Enterprise Data Access (EDA)/SQL. In addition, Tandem has certified many applications for use with its systems and DBMS. Certified systems include Clear Access, Microsoft Access, Visual Basic, Sybase Powerbuilder, Q+E Database Editor, and Trinzic Forest & Trees. Other systems, such as KBMS, have been certified for use with Tandem's NonStop ODBC Server and NonStop SQL/MP.

SAS Institute

SAS Institute is among the world's 10 largest independent software companies. In a recent survey conducted by Datamation (February 1996), SAS Institute's SAS System was chosen as 1996's "Product of the Year" in the category of data warehousing. According to the magazine, all products nominated met three criteria. They had to support real-world client/server computing, be technologically forward-looking, and solve real-world business problems.

Red Brick Systems

Red Brick's engine supports much faster data reading than a transactional database. Due to optimized query capabilities, data updates and insertions are less efficient when compared to OLTP engines. Unlike transactional databases, the data warehouse is not on-line, nor is it updated frequently. Red Brick does not permit end users to amend records; the System Administrator performs this function by means of the RISQL entry tool. Red Brick is intended to co-exist with relational servers optimized for OLTP. Its solution includes:

- STARIndex schema and corresponding join technology that accommodate complex query analyses.

- TARGETIndex (Version 4) based on bit-vector data representation, where attributes are set up as a bit index. Bit-mapped indexing helps relationships between facts and a limited number of attributes, including banded number ranges and tables with hundreds of columns and millions of records or rows.

- Table Management Utility (TMU), customized to allow fast bulk loading and indexing of operational data from production data bases. Version 4 supports low-cost mass storage devices such as Hewlett-Packard's OmniStorage.

- Query and reporting language provided by Intelligent SQL (RISQL), which formulates queries and control loading.

Red Brick's system runs on a range of UNIX systems, including NCR systems, Hewlett-Packard HP 9000 Series, Sequent, Sun SPARC machines, Digital Alpha platforms, Unisys U6000 and OPUS.

In April 1996, Red Brick announced Red Brick Warehouse for Windows NT, the only relational data base for data warehouse and data mart applications running on Intel-based Windows NT servers. Corresponding product offerings include:

- Red Brick Warehouse for Windows NT—A single server or standalone configuration designed for personal, workgroup, or departmental deployments up to 100G bytes.

- Builder's Edition—A full-featured, cost-effective development environment to pilot data warehouse or data mart applications up to 5G bytes and five named users.

- Enterprise Packs—For mass deployment of multiple data warehouse servers, subject- or function-oriented subsets of data warehouses or data mart applications.

Informix Data Base Server

The Informix data base server is a set of multithreaded processes that provide scaleable performance gains on multiresource systems (CPUs, disks, controllers). The internal implementation is a distinguishing feature. While Oracle and Sybase provide parallel architectures, Informix has the only multithreaded implementation.

Oracle

Oracle has put a tremendous amount of R&D effort into Oracle8 to dramatically increase the amount of data that can be loaded onto a database and then managed. However, data warehouses have complex data models and large numbers of users issuing complex queries to perform sophisticated analysis. With Oracle8, many of the improvements are more appropriate for large-scale data marts with denormalized data models than for data warehouses (normalized and complex data models). Here, we investigate those enhancements.

Query Performance: Oracle8 has several major enhancements targeted at improving query performance. As an early implementer of parallel technology among portable RDBMS vendors, Oracle has lagged the competition (e.g., Informix Software) in providing as thorough an implementation. Oracle8 will deliver parallel index scan and parallel execution plans, both of which improve query performance. Commonly, with Oracle7.3, users would artificially partition data into multiple tables (e.g., sales data partitioned by date) to gain additional parallelism with parallel table scans of smaller data volume. With a parallel index scan, Oracle8 can retrieve data in parallel when an index is specified in the selection criteria.

Strengths:

- Building on well-proven Oracle7 RDBMS technology with several enhancements to improve the performance of complex query processing

- Product enhancements respond squarely to the decision support demands of Oracle's leading users and the competitive marketplace

- Straightforward migration without database or application disruption for the majority of sites

- Extensive internal quality assurance geared to producing a quality product, not a market statement

- Market leadership and mind share

Challenges:

- Prove performance capabilities in the large and very large data warehouse market

- Prove workload management capabilities in a data mart/data warehouse environment with a large number of users

- Provide additional query processing algorithms for complex query processing

- Provide additional data partitioning capabilities (e.g. hashing) to reduce data skewing and performance bottlenecks

Consider This Product When:

- Planning to implement a star schema data mart and need to comfortably support database sizes of up to 1 Tbyte of raw detail data, and a large number of concurrent users issuing complex queries.

- Planning to implement a data warehouse and need to comfortably support database sizes of up to 500 Gbytes of raw detail data, and less than 30 concurrent users issuing complex queries.

Consider Alternatives When:

- Needing a proven RDBMS for a data mart implementation requiring support for a very large number of concurrent users, and a data warehouse implementation supporting more than 500 Gbytes of raw detail data and large concurrent user populations.

Oracle DSS-Related Enhancements

- Enhanced parallelism (index scan, insert, update, delete, analyze, optimization and execution)

- Enhanced star query optimization

- Improved bitmap indexes

- Enhanced cost-based optimization (IN list, view and subquery processing)

- Partitioned tables and indexes

- "Rolling window" partition operations

The previous Oracle analysis is part of the June 1997 Gartner Report: *Oracle8 and Data Warehousing: Are Big Databases Enough?*

Appendix H – IBI FOCUS Data Warehousing Products

The following is an overview of data warehousing products offered by IBI.

- **Focus For Windows (FFW):** A graphical interface for Microsoft Windows users.

- **FFW Report Painter:** Users can design mainframe report layouts on their PCs, using the GUI they're familiar with.

- **Focus Professional:** Includes an unrestricted runtime environment enabling creation and broad distribution of Focus Reports throughout an organization

- **SmartMart:** Combination of tools for building data marts. Provides data extraction and transformation capabilities, multi-dimensional data storage, reporting along with data mart management.

- **Managed Reporter**: A limited version of Focus Reporter, enables less sophisticated users to create and view reports and charts using predefined Focus domains created by an administrator. Here IBI intends to compete with market leaders Business Objects and Cognos, which each offer a managed environment through a semantic dictionary and highly intuitive interface.

- **Report Server:** A version of Focus Reporter (Microsoft Windows only) that accepts report requests, scheduled (deferred) execution and distribution of results via E-mail or HTML. Special features include "bursting" - the ability to create a single report, segmenting and distributing sections based on specified report groupings (e.g., geography, division, manager).

- **WebFocus**: Leveraging the Focus 4GL, IBI has retrofitted it with HTML generation and a Common Gateway Interface (CGI) link to any Web server. Combined with some Java applets, IBI has created a useful and appealing intranet option for reporting

BI Edition: IBI's Focus Six "BI Edition" includes a potpourri of decision-support tools loosely integrated with Focus Six.