Business/enterprise modeling

by R. L. Katz

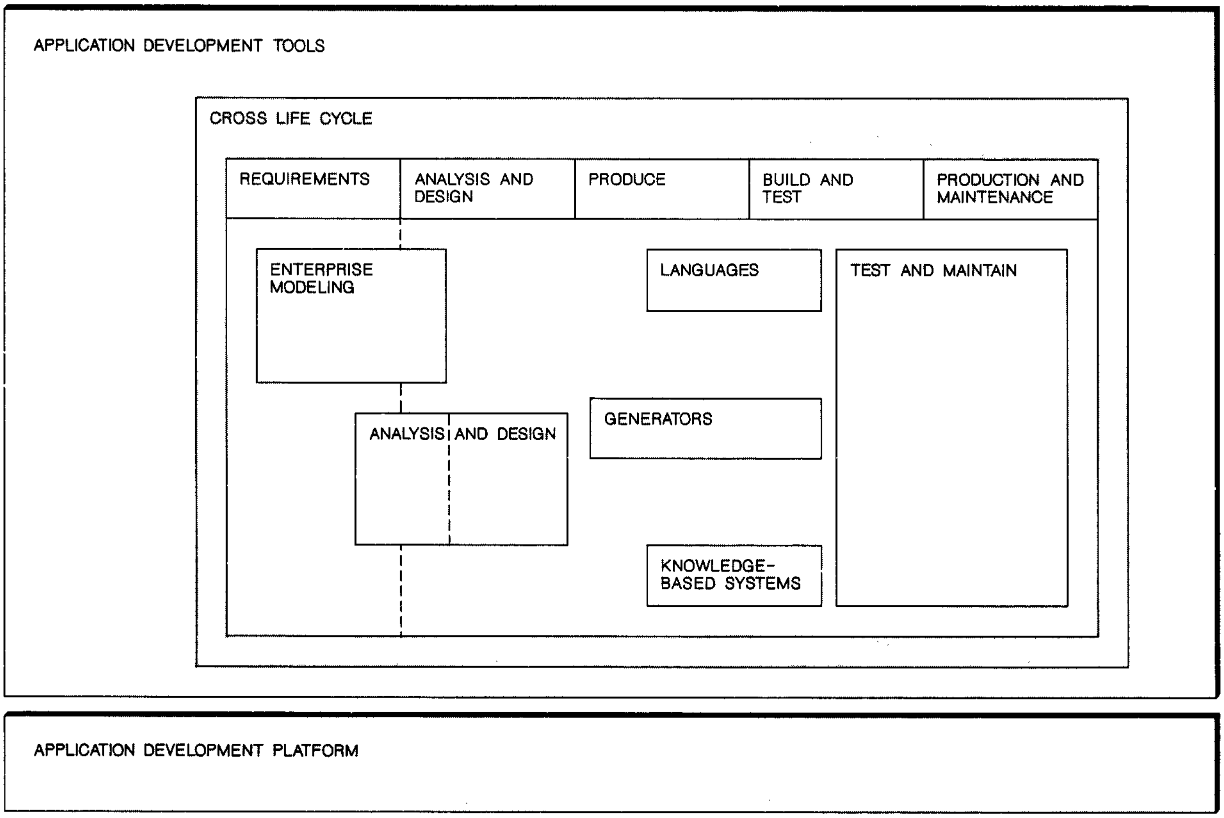
This paper reports on pertinent aspects of business/enterprise modeling studies that were conducted with nine IBM customers using what are now called computer-aided software engineering (CASE) tools. Coming shortly after the recent AD/Cycle™ announcement and the increased focus in IBM on tool-supported (CASE) business/enterprise modeling, this description of actual modeling studies should be especially germane. The model definitions (dimensions) used in the studies correspond exactly to many of the dimensions used by AD/Cycle, DevelopMate™, and the Repository Manager™. Compelling business reasons for conducting the studies are identified.

The term *business modeling* can be used to describe a number of different aspects of a business: financial, distribution, manufacturing lines, information and data requirements, flow of information and data across business processes, etc. This paper addresses data and information requirements and the flow of data and information across business processes, along with how business modeling has been used to develop requirements for new computer-based systems that in turn support the data and information needs of businesses. In addition, business modeling can be performed at both higher, or general, levels and at lower, more detailed levels. The focus of this paper is on the higher levels of modeling and can be seen in the context of AD/Cycle™ as enterprise modeling in Figure 1.

An additional valuable approach to understanding the context and role that business/enterprise modeling plays may be obtained from the framework for information systems architecture formulated by John Zachman,[1] as depicted in Figure 2. Business modeling would fit in the following components of Zachman's framework:

- Scope description (ballpark view)
  - List of entities
  - List of processes
- Model of the business (owner's view)
  - Entity-relationship diagram
  - Functional flow diagram
- Model of the information system (designer's view)
  - Data model (entity-relationship)
  - Data flow diagram

Figure 1 AD/Cycle framework



Higher-level modeling focuses on the scope description and model of the business perspectives, whereas a lower, or more detailed, level of modeling focuses on the model of the information system perspective.

The scope of this paper is to report on modeling experiences with which the author was personally involved. The methodologies and tools (now called computer-aided software engineering, or CASE, tools), especially the Information System Model and Architecture Generator (ISMOD),[2] have been successfully used with over 450 IBM customers during the past 10 years, ranging across the entire spectrum of large enterprises. Enterprises using this approach include all kinds; some examples are schools, aerospace manufacturers, airlines, governmental agencies, the military, oil companies, automobile manufacturers, steel manufacturers, insurance companies, and banks. The approach can therefore be said to be universal and not limited to any one kind of industrial genre.

Customer participants. The IBM customers who performed business/enterprise modeling studies, and with which the author was personally involved, follow:

- One of the three largest airlines in the U.S.
- One of the three largest commercial and retail banks in the U.S.
- One of the largest electric utilities in Arizona
- One of the largest electric utilities in Michigan
- A major electric utility providing most of the power to one of the largest cities in Pennsylvania, focusing on nuclear power sources
- One of the largest electric utilities in Pennsylvania, focusing on nuclear power sources
- One of the world's largest multinational grain shippers and traders
- A multinational, U.S.-based insurance company with the reputation of being one of the best-managed and aggressive companies of its kind
- A multinational U.S. oil company currently included in the top 50 of the *Fortune* 500 industrial listings

Figure 2 Framework for information systems architecture

	DATA DESCRIPTION ENTITY RELATION	PROCESS DESCRIPTION PROCESS INPUT/OUTPUT	NETWORK DESCRIPTION NODE LINE
SCOPE DESCRIPTION (BALLPARK VIEW)	**LIST OF ENTITIES IMPORTANT TO THE BUSINESS** ENTITY = CLASS OF BUSINESS ENTITY	**LIST OF PROCESSES THE BUSINESS PERFORMS** PROCESS = CLASS OF BUSINESS I/O = PROCESS	**LIST OF LOCATIONS IN WHICH THE BUSINESS OPERATES** NODE = BUSINESS LOCATION
MODEL OF THE BUSINESS (OWNER'S VIEW)	**E. G., ENTITY-RELATIONSHIP DIAGRAM** ENTITY = BUSINESS ENTITY RELN. = BUSINESS RULE	**E. G., FUNCTIONAL FLOW DIAGRAM** PROCESS = BUSINESS PROCESS I/O = BUSINESS RESOURCES	**E. G., LOGISTIC NETWORK** NODE = BUSINESS UNIT LINE = BUSINESS RELATIONSHIP/FLOW
MODEL OF THE INFORMATION SYSTEM (DESIGNER'S VIEW)	**E. G., DATA MODEL** ENTITY = DATA ENTITY RELN. = DATA RELATIONSHIP	**E. G., DATA FLOW DIAGRAM** PROCESS = APPLICATION FUNCTION I/O = USER VIEWS (SET OF DATA ELEMENTS)	**E. G., DISTRIBUTED SYSTEMS ARCHITECTURE** NODE = I/S FUNCTION (PROCESSOR, STORAGE, ACCESS, ETC.) LINE = LINE CHARACTERISTICS
TECHNOLOGY MODEL (BUILDER'S VIEW)	**E. G., DATA DESIGN** ENTITY = SEGMENT/ROW RELN. = POINTER/KEY	**E. G., STRUCTURE CHART** PROCESS = COMPUTER FUNCTION I/O = SCREEN/DEVICE FORMATS	**E. G., SYSTEM ARCHITECTURE** NODE = HARDWARE/SYSTEM SOFTWARE LINE = LINE SPECIFICATIONS
DETAILED DESCRIPTION (OUT-OF-CONTEXT VIEW)	**E. G., DATABASE DESCRIPTION** ENTITY = FIELDS RELN. = ADDRESSES	**E. G., PROGRAM** PROCESS = LANGUAGE STATEMENTS I/O = CONTROL BLOCKS	**E. G., NETWORK ARCHITECTURE** NODE = ADDRESSES LINE = PROTOCOLS
ACTUAL SYSTEM	**DATA**	**FUNCTION**	**COMMUNICATIONS**

Purposes of studies. The business reasons for seven of the customer studies related to the need for creating tactical and strategic computer support plans. All of these companies realized that their major systems were implemented at least five years earlier, and new applications were being installed with no thought given to the overall picture. They were sure that they needed to review their current status, assess any shortcomings, and put in place plans to ensure that they would have adequate support to handle future requirements in addition to capitalizing on new technology.

The results of these seven studies were, therefore, focused on proposed computer systems and improvements. Business/enterprise modeling was used as a disciplined way of identifying current computer support, lack of computer support, and potential new applications.

Two other studies were conducted for different reasons. One reason was computer availability. In this case, an IBM customer had most of their computer support located in a single physical location, and they were concerned about the effects on their operations if a natural disaster, such as an earthquake, were to occur. They were interested, therefore, in business/enterprise modeling and in doing simulations to determine whether or not they could distribute their business processes and the associated computer support and software applications to four or more disparate computer sites. The customer knew that the disparate computer sites would interact with each other with regard to applications and databases, and business/enterprise modeling was used to determine whether or not distributed data processing of applications and data would be feasible.

Another study involving business/enterprise modeling was triggered by an IBM customer's increasing already-high-volume workload and the fear that their requirements would eventually exceed the performance capabilities of the largest IBM computers. Their use of business/enterprise modeling was also geared toward determining the feasibility of distributing their processing from one mainframe computer to two or more mainframes, while still communicating with each other in real time to ensure that the data in all of the mainframes would be synchronized.

Key aspects of the studies

The cited business/enterprise modeling studies incorporated the following list of dimensions and concepts. The full analytical underpinnings of the methodologies utilized in the studies are contained in the documents listed at the end of this paper.

- Business processes
- Entities
- Attributes
- Data views
- Data qualifiers
- Data criticisms
- Organizations
- Physical locations
- Events or triggers
- Repositories
- Analytical methods for feasibility of distribution of processes and data

Compelling business reasons. Important aspects of these studies that cannot be overemphasized to potential business/enterprise modeling practitioners are dollars spent and dollars saved, urgent and identifiable business problems needing solutions, and the perspective of the business person regarding immediate and direct payback from any business/enterprise modeling activity.

The meaning and context of "dollars spent and dollars saved" have to do with the outlook of a business person (i.e., businesses exist solely to make money and profit). This statement may seem to some to be a platitude, but seven of the nine business/enterprise modeling studies that are reflected in this paper had this aspect as the first and foremost study goal. Each of the modeling studies produced a final report, and the most important section in each of these reports dealt with the cost to implement a new computer system and the direct dollar savings that could be attributed to these new systems.

Business/enterprise modeling activities *per se* were viewed only as adjuncts to the goal of dollars saved. The businesses had already intuitively sensed that new systems and applications were needed. They were at a point where they were ready to obtain a higher degree of rigor and discipline to validate their systems and application requirements. These businesses were, therefore, primed and receptive to some kind of outside review of their situation by a consultant, whether from IBM or elsewhere.

When IBM offered to provide a business/enterprise perspective of their current status, this kind of modeling approach was accepted. This approach might be compared to requesting and receiving a second medical opinion to ensure that the right decision is made.

The notion of an urgent and pressing business problem needing a solution relates to two of the IBM customers already mentioned. One, an airline, concerned that its transaction volume would exceed the performance capabilities of IBM's largest computers, wanted to investigate the feasibility of off-loading their reservations applications to one or more distributed processors. They therefore opted to build a business model of their reservations system to see if distributed processing was a viable approach.

Modeling was a support activity leading to the identification of another way to save money.

The second customer, a large bank, was concerned about the exposure of having all of its processing requirements located at a single physical site, in the event that a natural disaster should occur. They also used modeling to determine the feasibility of dispersing their nightly processing to two or more distributed processors.

Payback from modeling activities is directly related to the focus of the modeling studies on the return-on-investment aspect of the study recommendations (i.e., cost and benefit analysis). Modeling was viewed as a support activity leading to the identification of another way to save money for the business.

Bridge from business/enterprise modeling to implementation. Over time, it has become evident that it would be desirable to be able to use the information gathered in the modeling activity during the subsequent implementation phases. Until the advent of AD/Cycle it was difficult to do this. Using DevelopMate™, it is now possible to take the results of the business/enterprise modeling activity and use those results to prototype and validate the application and data architecture views. That is, one may take the following approach:

- Perform business/enterprise modeling at the highest level of an enterprise to obtain the information requirements at that perspective
- Prioritize the sequence of applications to be developed as determined by the enterprise-level study
- Use the results of the enterprise-level modeling study and DevelopMate to decompose the business process and data to the level where prototyping can be used to validate the rigor of design and user satisfaction

Noteworthy is the exact correspondence of many of the dimensions used in the cited business/enterprise modeling studies and the dimensions used by DevelopMate[3,4] and the Repository Manager™.[5] (See Figures 3, 4, and 5 for the ISMOD enterprise model—process and data, the enterprise process model portion of the IBM-supplied meta model, and the enterprise data model portion of the IBM-supplied meta model, respectively.)

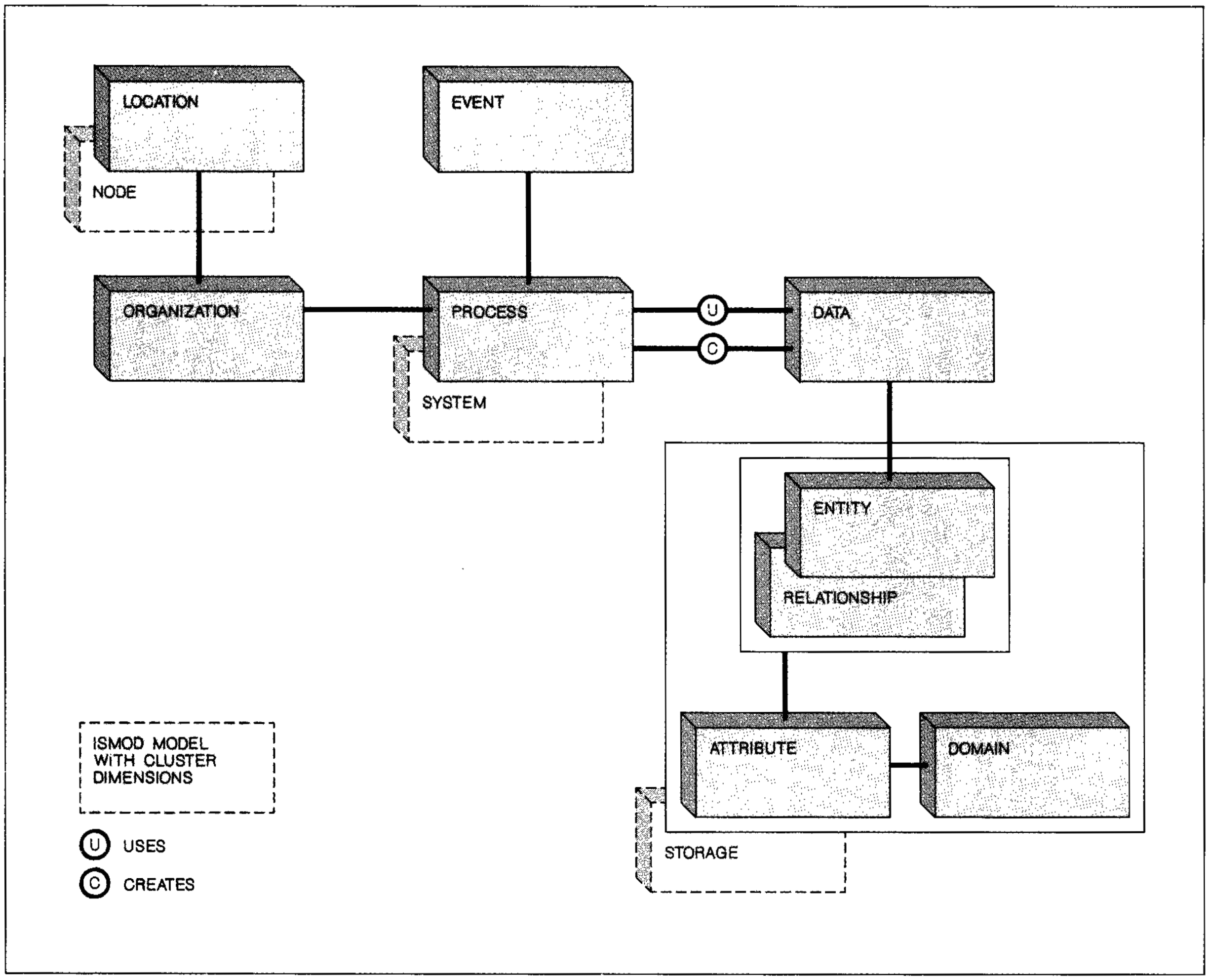
Excerpts from a customer study

Much of this section was excerpted from an actual customer business/enterprise modeling study to provide a realistic and pragmatic perspective of how IBM customers view business/enterprise modeling. Names have been omitted and changed to protect customer confidentiality. Especially noteworthy in this customer study are terms such as *integrated system architecture*, *business processes*, *data views*, *events*, *locations*, and *organizations*. These terms (dimensions) are equivalent to many of the terms (dimensions) used by DevelopMate. (See Figures 4 and 5 for the enterprise process model and the enterprise data model of the IBM-supplied AD/Cycle Repository Manager.)

Table of contents example. Below is listed an excerpted table of contents from the customer study. Note that the business/enterprise modeling study methodology is relegated to Appendix F (Item 12), and the real business aspects of the study are in the forefront of the document.

1. Executive Summary
2. Recommended Strategic Environment
3. Recommended Tactical Projects
4. Implementation Schedule
5. Cost and Benefit Analysis
6. Configuration
7. Appendix A: Improvement Opportunities, Major Needs, Individual Solutions

Figure 3 Schematic for the ISMOD business/enterprise model

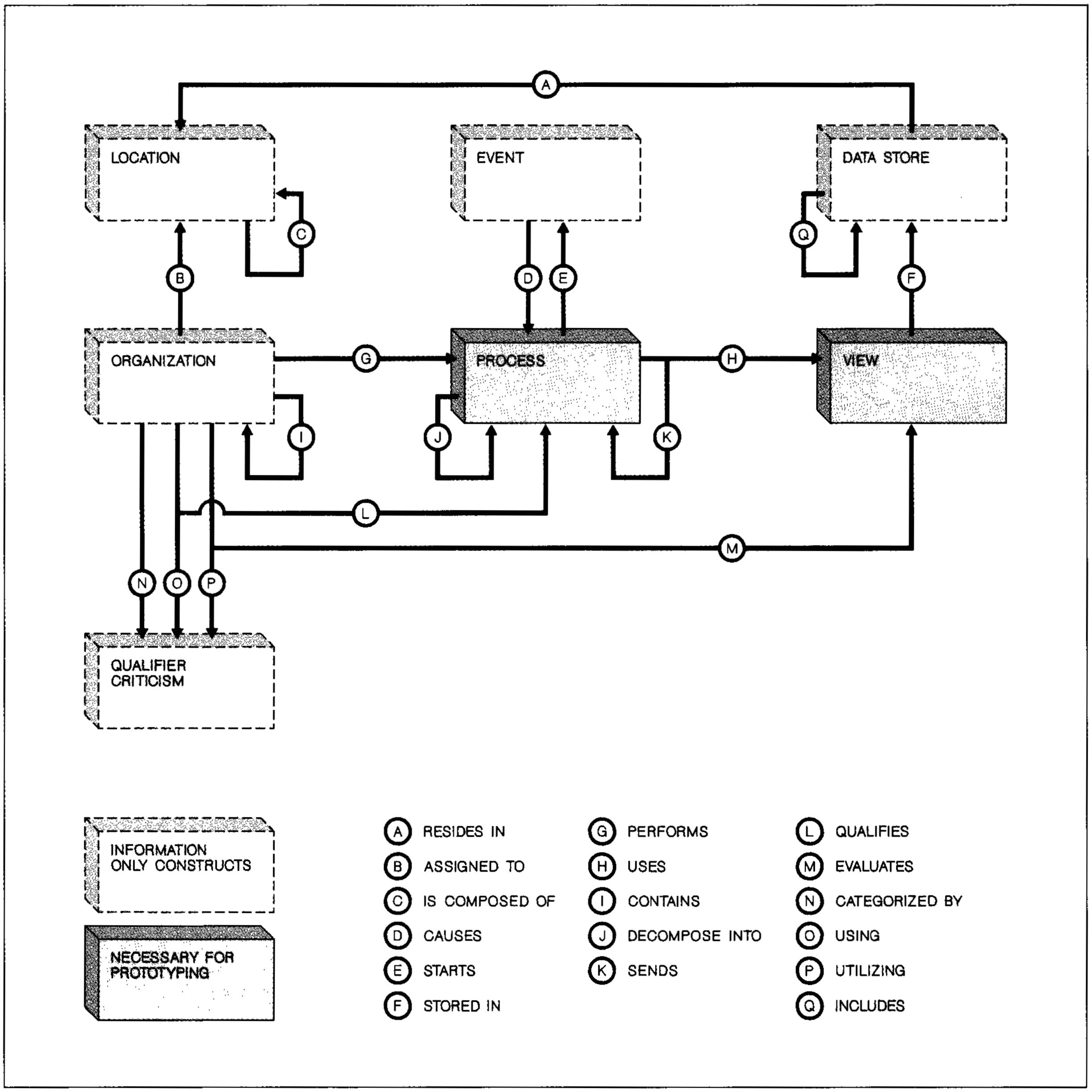


8. Appendix B: Tactical Clusters
9. Appendix C: Building Blocks/Individual Solutions Matrix
10. Appendix D: Tactical Projects
11. Appendix E: Commercial Product Description
12. Appendix F: Methodology

Executive summary. At the end of the study the following executive summary was provided:

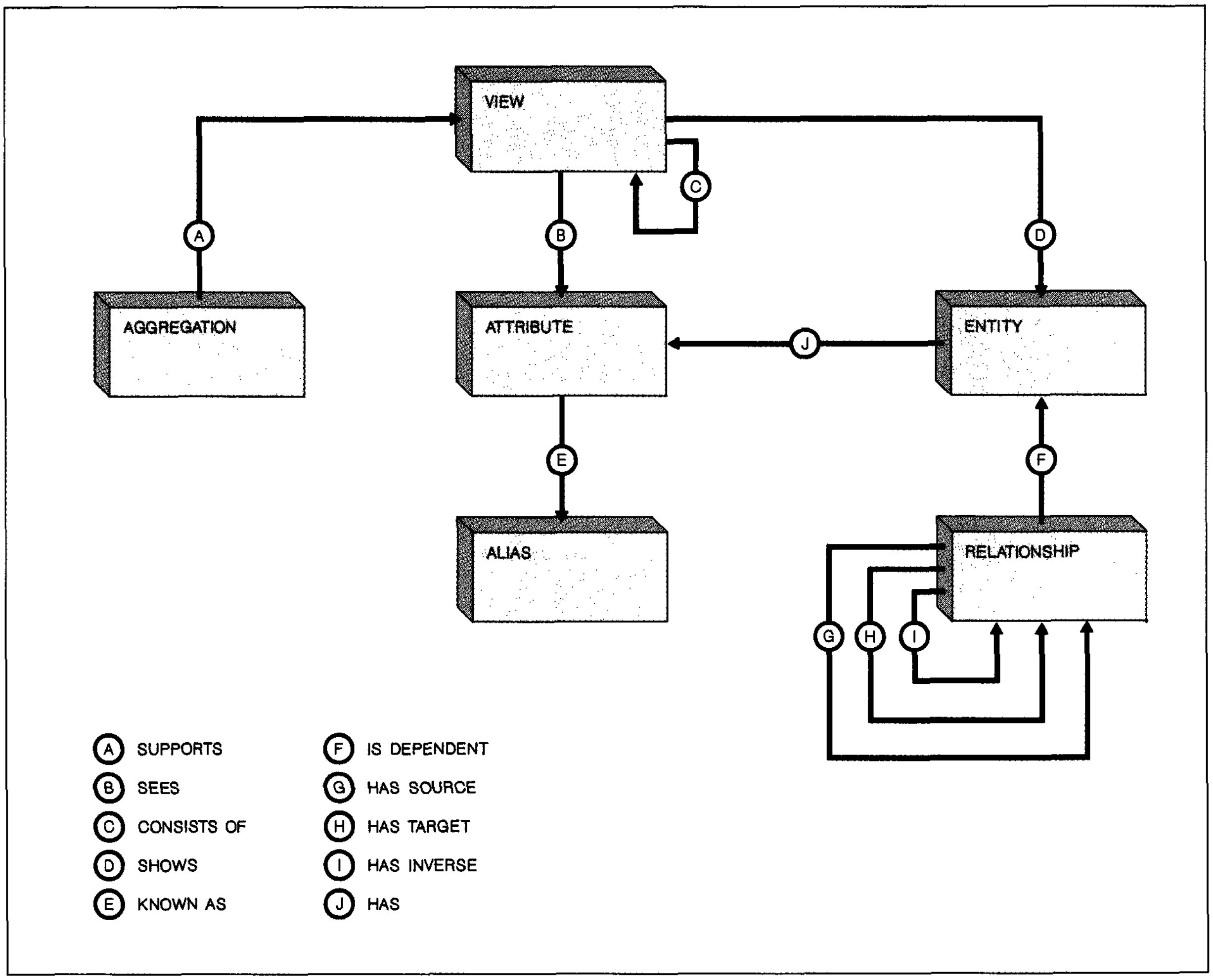
> Company President J. Jones commissioned an analysis of the enterprise's business information needs. This analysis was performed during the period of April 4 through June 3, 1988, by a study organization consisting of representatives of our company, IBM, and representatives from two other outside consulting organizations.
>
> IBM Information System Model and Architecture Generator (ISMOD) was the method used for the analysis. Through ISMOD, business processes performed to manage business resources and the data used by those business processes were identified. Simulation capabilities of the model allowed a measure of the impact a change in a business process could have in terms of user satisfaction. Existing management studies were factored into the analysis.

Figure 4 Schematic of the enterprise process model portion of the IBM-supplied AD/Cycle Repository Manager



The results and recommendations contained in the Plan are the professional opinions of the study team, based upon the research conducted during the nine-week study. The Plan provides the observed opportunities to improve business processes, recommendations to implement improvements, and a proposed integrated system architecture that enhances the information resources with associated costs, benefits, and schedule. Detailed design of the final system configuration with refinements to cost and schedule will need to be accomplished if the Plan is implemented. It is the

Figure 5 Schematic of the enterprise data model portion of the IBM-supplied AD/Cycle Repository Manager



team's belief, however, that the Plan's analysis bounds the cost and benefits resulting from the Plan's detailed design implementation.

Methodology. In order to analyze information flow, the processes within the company had to be clearly defined and understood. Since no one person is knowledgeable about all business processes, interviews with a number of key individuals were required to define all of the processes of the company. Prior to conducting these interviews, the study team reviewed and organized the business processes of the company. The resulting business process list would be modified by the interview phase of the study until a complete and accurate list had evolved.

Business processes were organized in the following format:

Business function
  BUSINESS SYSTEM
    Process
      Subprocess
        Sub-subprocess

This five-tier hierarchy was important because it allowed the processes to be organized by the nature of the information they contained rather than by the physical organization of a plant department or the personnel who performed in a process.

Business functions can be thought of as those resources that exist to satisfy the mission, goals, and objectives of a business. Seven business functions were identified for the company. Business systems are those systems that manage each particular business function. Twenty-four business systems were identified. Each of the business systems is supported by a number of processes, subprocesses, and subsubprocesses. Following is a list of the business function and business system organization:

Business functions are those resources that exist to satisfy the mission, goals, and objectives of a business.

1. *Production*
   - Operations
   - Maintenance
   - Materials and service
2. *Administration*
   - Payroll
   - Cashiering
   - Accounts payable
   - Document control
3. *Management*
   - Commitments
   - Quality control
   - Performance
4. *Legal*
   - Environment
   - Licensing
   - Security
   - Quality assurance
5. *Personnel*
6. *Engineering*
   - Modifications
   - Controls
   - Supplies
7. *Planning*
   - Budget
   - Schedule

The business systems named above are not physical departments of the company organization. They are business systems that are independent of company organization and that may cross several department boundaries. The business processes grouped by any business system may be performed by any company individual, regardless of his or her department affiliation. This means that any future company reorganization will have little effect on how the current information flow is defined, or on the final proposed integrated design.

Once the business processes were defined, the flow of information through these processes had to be examined. All business processes use raw data in some form to produce an output. This combination of data to represent something useful in the execution of a process is called a *data view.* A data view is a physical grouping of elementary data in a form that can be viewed by people. Examples of data views are invoices, reports, memos, computer screens, conversations, and forms. Business processes, in almost all cases, use data views as input, and they create output data views, which are sent to other business processes. Therefore, an evaluation of specific data views will provide a measure of the flow of information in business processes.

One of the greatest benefits from interviewing key individuals was their evaluation of the data views they use. Prior to the interviews, the study team prepared a list of data views to be used as a starting point for the interviewees. This list, like the process list, evolved through the interview phase.

Interviews. Interviews were conducted with key individuals knowledgeable about the business processes in which they are involved and the data needed to perform these processes. These individuals were selected to represent the processes of the 24 business systems. After this selection, a cross check was made of the company's organizational chart to ensure that all levels of management and all functional departments were represented.

Once selected, individuals to be interviewed were organized into groups of three or four so that they would both enhance and restrain group discussions. Interviews lasted from six to eight hours and consisted of the following three parts:

1. Identification—All interviewees were identified by a unique code, by their physical location within the company, and by their interview number.

2. Process identification—With use of a prepared business process list as a starting point, all interviewees were asked to identify all of the processes as they performed them. In most cases, several interviewees indicated that they used the same process. On a group basis, each interview session identified about 50 to 100 processes. Individually, each interviewee identified between 15 and 25 processes he or she performed. Each interviewee then coded each process he or she performed on a process form. The interviewees also indicated the frequency with which they performed the process by indicating an event code.
3. Data view identification—Using the prepared data view list, interviewees were asked to identify all of the data views that they used. After the prepared list was exhausted, the group of interviewees identified all other data views that they used by brainstorming. Typically, each interview group identified approximately 100 data views, although some groups identified over 200. Each interviewee then coded each data view that he or she used on a "data form."

Data evaluation. The data form allowed the interviewees to qualify and criticize each data view individually. Nine data qualifiers were developed by the study team and placed on the data form. These were:

1. Essential data: Is this data view vital to the business processes or tasks being performed?
2. Copied: Are the data duplicated material, rekeyed, or manually recorded from another source?
3. Text: Are the data written, typed, or printed material not from a computer?
4. Graphics: Is the data view a chart of any kind?
5. Drawings: Is the data view a blueprint or drawing of any type?
6. Microform: Is the data view reproduced on microfilm, slides, or photos?
7. Computer output: Is the information obtained from a computer printout or read from a terminal?
8. Real time required: Are the data rapidly changing, and must the data be kept current?
9. Not formalized: Is the information received by telephone, word of mouth, or any other informal source?

Seven data criticisms were developed by the study team. These criticisms were the measuring tool that the computer model used to determine the quality of specific data views. They are as follows:

1. Accessibility: Are the data difficult or time-consuming to access?
2. Clarity: Are the data difficult to understand or interpret?
3. Accuracy: Is there any doubt about the accuracy of the data?
4. Timeliness: Do you receive the data late?
5. Completeness: Are data missing?
6. Legibility: Are the data difficult to read?
7. Excessive: Are there useless or redundant data?

Business/enterprise model. After the interviews were completed, information contained in the process and in the data forms was coded and entered into ISMOD. ISMOD allows a number of reports to be generated and sorted by any one of these five dimensions: the interviewee, the hierarchy (department), the process, the data view, or the origin (process that creates the data view).

Many reports are useful in designing an integrated systems architecture. The statistics reports are very useful in determining problems associated with the current information system. They present the interview data in summarized form and show the satisfaction index.

The satisfaction index was computed by dividing the number of satisfied responses (no criticisms) by the total number of interviewee usages. Therefore, if a data view was never criticized, its satisfaction index would be 1.0. The statistics report also printed out any or all of the data qualifiers and criticisms. This information provided insight as to why a given business process or data view had a low satisfaction index.

Another ISMOD report useful in diagnosing the current information system was the initial matrix. This report presented the interview data in matrix form showing two dimensions on the axis. Each intersection on the matrix has four entries associated with it. It showed the number of satisfactory usages, the number of unsatisfactory usages, the total number of usages, and the satisfaction index. When this report was run with the dimensions business process and data view, it showed all of the data views used by a business process and all of the business processes using a particular data view, along with the usage and satisfaction calculations.

Since the statistics report was useful in acquiring an overall understanding of the usage of data in information systems, it was logically the first report executed. Statistics reports were generated with business processes, data views, and hierarchies as single dimensions on the row axis and qualifiers or criticisms on the column axis. As suspect business processes with low satisfaction indexes were identified, an initial matrix with business processes and data views was generated. This was the extent to which the model was used to identify problems.

Integrated solution. The earlier subsection on methodology describes how a computer model was used to analyze the information derived from the interviews of the company's personnel. The results of this analysis were a number of opportunities for improvement that were categorized into improvement opportunities. From the improvement categories, major needs and their corresponding individual solutions were defined. At this point, it was recognized that appreciably more benefits were possible by using an *integrated solution* rather than implementing each individual solution independently. The following describes how this was done.

Cluster formation. The ISMOD computer model was used to build a matrix that identified business processes that use data views versus those that create data views. This effort helped build 18 business, or tactical, clusters. A *tactical cluster* is a functional area that uses a large amount of the data that it creates. The flow of data from these clusters was then analyzed to identify the business process at the central site that feeds data to other business processes.

Once the flow of data was clear, simulations were run on the tactical clusters to identify data that, if improved, would have the greatest impact on the organization. From this diagnosis, ideas were discussed on how to improve the way these critical data could be presented for use with an integrated approach. The integrated architecture building blocks evolved from this process.

Needs and individual solutions were then grouped under the tactical clusters and checked against the building blocks to ensure that the proposed integrated solution provided the same functionality that had been provided by the individual solutions.

Tactical projects. To implement this integrated solution, the tactical clusters were divided into discrete projects. Upwards of 50 tactical projects were identified from 13 of the 18 tactical clusters. The selected tactical projects were deemed sufficient to develop an integrated information strategic plan, and, when implemented, would satisfy the major information needs of all 18 tactical clusters.

Data satisfaction metric

An aspect of these studies that was frequently cited as one of the most valuable findings has to do with

Evaluations highlighted areas where there were opportunities to improve the information needs of the business.

how satisfied or dissatisfied people are with the information and data they say they need to perform their jobs or business processes in their respective businesses or enterprises. An actual study determined that one of the greatest benefits from the interview of key individuals was their evaluation of the data views they use.

These evaluations used both individually and collectively highlighted areas where there were opportunities to improve the information needs of the business or enterprise. Among the powerful features of the ISMOD methodology and tool is a data satisfaction metric; this feature allows the data requirements for a business or enterprise to be clearly identified. This metric is described below.

The metric. ISMOD defines a data satisfaction metric as follows: if one has no criticisms of the information needed to perform a job or a business process, the data satisfaction index is said to be 1.0; if, in contrast, one has a criticism of the information needed to perform a job or business process, the data satisfaction index is said to be 0.0. ISMOD generally allows each customer to tailor the critiques to reflect their own organization so that a typical list of possible data critiques might include:

- Accuracy: Do you have doubts about the accuracy of the data?
- Completeness: Are all of the data that are needed present?

- Medium used: Are the data presented in an effective way?
- Missing data: Are the data missing?
- Reliability: Do you question the reliability of the data?
- Timeliness: Do you receive the data late?

The total satisfaction index is defined to be the ratio of INFORMATION WITHOUT CRITICISMS to ALL INFORMATION NEEDED.

Thus, if a person or a department indicates that they need 100 items of information to do a job and they are unhappy with 50 of these items, their satisfaction index is said to be 0.50 or 50/100. This index is used in three key ways:

1. To highlight data with the lowest satisfaction index
2. To highlight processes which have information and data requirements with the lowest satisfaction index
3. To highlight organizations having the lowest satisfaction index with regard to their information and data requirements

Data qualifiers. ISMOD provides another unique perspective that allows businesses or enterprises to classify and view their information requirements. This perspective is the identification of qualifiers, and they typically include the following kinds of categories:

- Essential data—Are these data key to the organization?
- Computerization—Are the data currently computerized?
- Text—Are the data presented in text form?
- Graphics—Are the data presented in graphics format?
- Office data—Do the data consist of office data?
- Required computerization—Are the data of a type that should be computerized?
- Batching—Are the data batch-oriented?
- Real-time access—Do the data require real-time access?

These categories may be reviewed individually and viewed with their individual satisfaction indices. This perspective allowed the IBM customers to quickly ascertain those areas that needed assistance in satisfying their data requirements. Comparisons across organizations and across processes are also provided to further analyze where and what the key areas are in which data and information requirements are not being met.

Data sharing

The significance of data sharing in these studies was related to the development of integrated solutions. That is, high-level process and application architectures were identified by use of the concept of data sharing. This action was possible because of the powerful feature of the ISMOD tool as described previously. Another perspective on how the cluster feature works is:

- Identify and store in the ISMOD computer model information about business processes, such as which business processes create which data views and which business processes use which data views.
- Run ISMOD reports and matrixes to ascertain the degree of data sharing among different business processes. How many data views do individual business processes use that are created by business processes other than themselves?
- Regroup business processes using ISMOD tool support to obtain optimal clusters of business processes focusing on minimization of data sharing across business processes.

As stated, the major way in which data sharing was utilized in most of the studies cited in this paper was in identifying tactical or major application clusters. That is, the criterion for establishing application clusters of business processes was that the respective business processes had the greatest degree of interaction with each other relative to using data views created by other business processes in the same cluster, versus using data views created by business processes in other clusters.

Data views were shared between clusters but not to the same degree as sharing between business processes in the same cluster. The degree of data sharing was identified with an analytic approach incorporated in ISMOD.

Data sharing, as defined by ISMOD, is meant to be the use of data or a data view by one business process of data or a data view that was created by another business process. Furthermore, ISMOD defines a metric that describes the degree of data sharing used by one or more business processes. This definition is:

> Assume a business process, say A, uses 100 elements of data or 100 data views to accomplish its functions. Furthermore, assume that business process A itself creates 50 elements of data or data

views that it needs to accomplish its functions. Then, ISMOD defines the degree of data sharing or isolation of A to be 0.50, or 50 percent. That is, A itself creates 50 percent of the information it needs to accomplish its functions. It gets the other 50 percent of the information it needs from other business processes. So it can be said that it shares data with other business processes for 50 percent of its information requirements.

Data sharing across processes or functions is represented by the data elements or data views created by one process and used by another process.

A metric for measuring data sharing. ISMOD defines a metric to measure the degree of data sharing and to identify clusters as follows:

Assume business process A uses 100 data views to accomplish its functions and it supplies 40 of the data views itself. Then its *isolation ratio* is defined to be 40/100 or 0.40.

Assume business process B uses 100 data views to accomplish its functions and it supplies 60 of the data views itself. Then its *isolation ratio* is defined to be 60/100 or 0.60.

The *average isolation ratio* (IM) between two business processes, say A and B, is defined to be the arithmetic average of the individual isolation ratios. For example, 0.40 for A and 0.60 for B so that

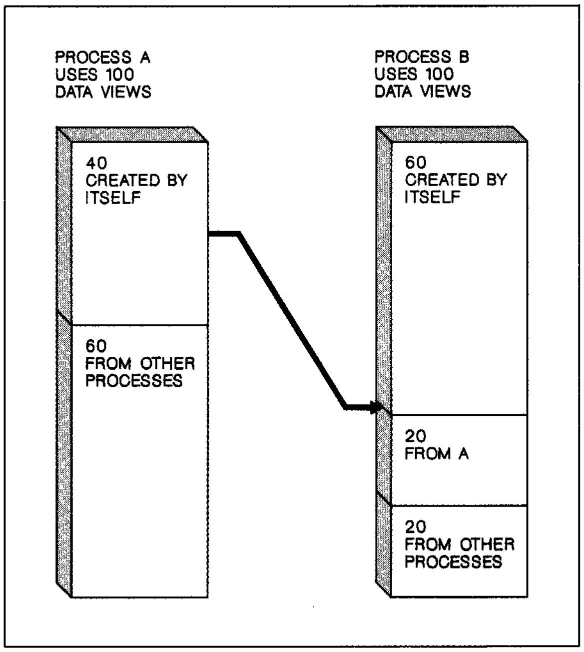
IM = 0.40 + 0.60 = 1.00 divided by 2 = 0.50

The *extended isolation ratio* (IE) between two business processes, say A and B, is defined to be the isolation ratio calculated, assuming the two processes (in this case A and B), were grouped together and formed one process. In this example we will assume that there is data sharing between business processes A and B to the extent that B gets 20 of the 100 data views it needs from A, but A gets none of the 100 data views it needs from B. (See Figure 6.) Then the extended isolation ratio between business processes A and B is defined to be:

IE = (40 + 60 + 20)/(100 + 100)
= 120/200 = 0.60

The *delta isolation ratio* (DI) between two business processes is defined to be the difference between the average and the extended isolation ratios for the two processes as follows:

Figure 6 Data sharing between two processes



DI = (IE − IM) = 0.60 − 0.50 = 0.10

The significance of these ratios relates to how one can identify where the greatest degree of data sharing takes place between business processes or clusters of business processes. ISMOD calculates all of the above ratios for every combination of business processes in its model and allows one to look for those that share the most data. The DI is the key to this perspective for it indicates where the biggest increase in data sharing will occur across all of the processes with which an individual process may interact. That is, the bigger the DI, the greater the degree of data sharing. This perspective is elaborated in the next section.

Data sharing between two processes is depicted in Figure 6.

Example of a data sharing matrix. Figure 7 represents a data sharing or isolation ratio matrix produced by ISMOD for an actual customer study. Its direct use to develop a process architecture can be seen in Figure 8. The interpretation of this matrix follows.

The rows and columns represent business processes; in fact, they are the the very same business processes

Figure 7 Data sharing isolation matrix

```
**********************
*  ISOLATION RATIOS  *        ALL DATA.
**********************

DI - MINIMUM :  .001
```

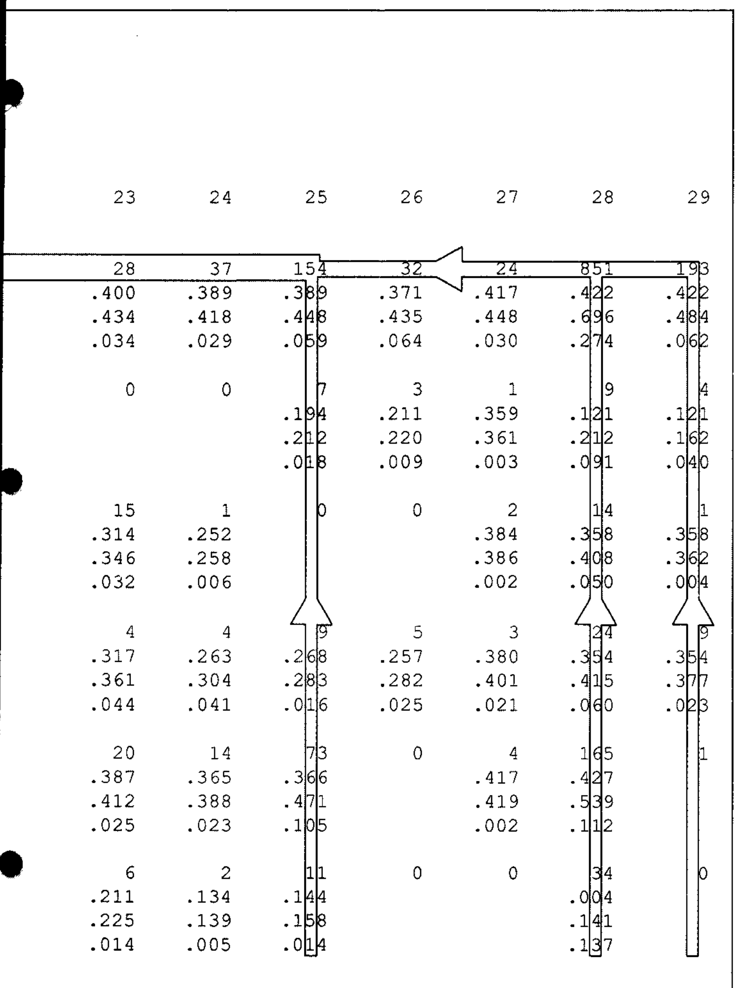
ORIG PROC		11	12	13	14	15	16	17	18	19	21	22
11	ABC	1309	69	12	181	132	4	7	14	0	14	43
	IM		.412	.416	.414	.423	.391	.392	.417		.357	.361
	IE		.447	.437	.479	.488	.403	.401	.422		.404	.417
	DI		.034	.020	.065	.065	.012	.009	.005		.047	.057
12	ABC	41	12	0	13	2	0	2	0	0	0	5
	IM	.412			.308	.407		.052				.215
	IE	.447			.404	.526		.186				.329
	DI	.034			.097	.119		.135				.114
13	ABC	56	19	101	64	4	0	0	2	0	2	1
	IM	.416	.297		.356	.416			.326		.207	.245
	IE	.437	.346		.532	.418			.331		.221	.247
	DI	.020	.050		.176	.002			.006		.014	.002
14	ABC	48	35	56	141	30	0	3	4	0	21	2
	IM	.414	.308	.356		.411		.227	.331		.218	.252
	IE	.479	.404	.532		.460		.242	.352		.269	.267
	DI	.065	.097	.176		.049		.015	.021		.050	.015
15	ABC	163	184	0	61	626	11	76	32	0	15	22
	IM	.423	.407		.411		.365	.368	.416		.318	.328
	IE	.488	.526		.460		.414	.467	.455		.363	.358
	DI	.065	.119		.049		.048	.098	.039		.045	.030
16	ABC	36	30	0	12	72	1	24	11	0	3	7
	IM	.391	.037		.219	.365		.014	.050		.136	.168
	IE	.403	.124		.238	.414		.066	.084		.141	.198
	DI	.012	.086		.019	.048		.052	.034		.005	.010

mapped against each other. The two-digit numbers from 11 to 29 refer to unique business clusters of business processes, such as trading, positions, futures, contracts, and log or plan. These processes can be interpreted differently depending on whether they are viewed from the row or column perspective.

The row perspective is that of a "using" business process. That is, one looks across a particular row to determine how many data elements or data views a particular process is using and is receiving from a particular business process, as viewed downward from a particular column. The interpretation of the business processes listed on the column headings is that of creating business processes. These processes create the data elements or data views used by the processes listed in the rows.

The entries in individual cells of the matrix show: number of data usages by using process from creating process (ABC); the average data sharing or isolation ratio of the using and creating processes at that cell (IM); the data sharing or isolation ratio of the using and creating processes if they were included in the same cluster (IE); and the improvement in data sharing or isolation ratio if the using and creating processes were indeed grouped together in the same cluster (DI).

At the intersection of row 11 with column 14 the respective values are: ABC = 181, IM = 0.414, IE = 0.479, and DI = 0.065. The meanings of these numbers are: process 11 has 181 usages of data or data views that are created by process 14; the average data sharing or isolation ratio of processes 11 and 14 is



0.414; the data sharing or isolation ratio, if processes 11 and 14 were grouped into the same cluster would be 0.479; and the increase in data sharing or isolation ratio for processes 11 and 14, if they indeed were put in the same cluster, would be 0.065.

Example of process architecture

Figure 8 is actually taken from an IBM-supported business/enterprise modeling study. It depicts the subsystems suggested by the clustering or process architecture facilities of ISMOD. Each of the rectangles in the figure represents a cluster of business processes that were identified as candidates for the same cluster using the data sharing analysis of ISMOD. Therefore, they can also be viewed as applications. The size of the rectangles is proportionate to the number of data usages identified in the business/enterprise modeling study. The size of the connectors is also proportionate to the number of data elements or data views being shared across clusters or applications. Each rectangle is subdivided into three sections: the leftmost section represents the data elements or data views created by that cluster or application and shared by another cluster; the middle section represents the data elements or data views created by that cluster or application and used by itself; and the rightmost section represents the data elements or data views used by that cluster or application and obtained from a different cluster or application. The size of the sections within the rectangles is also proportionate to the number of data views being created and used.

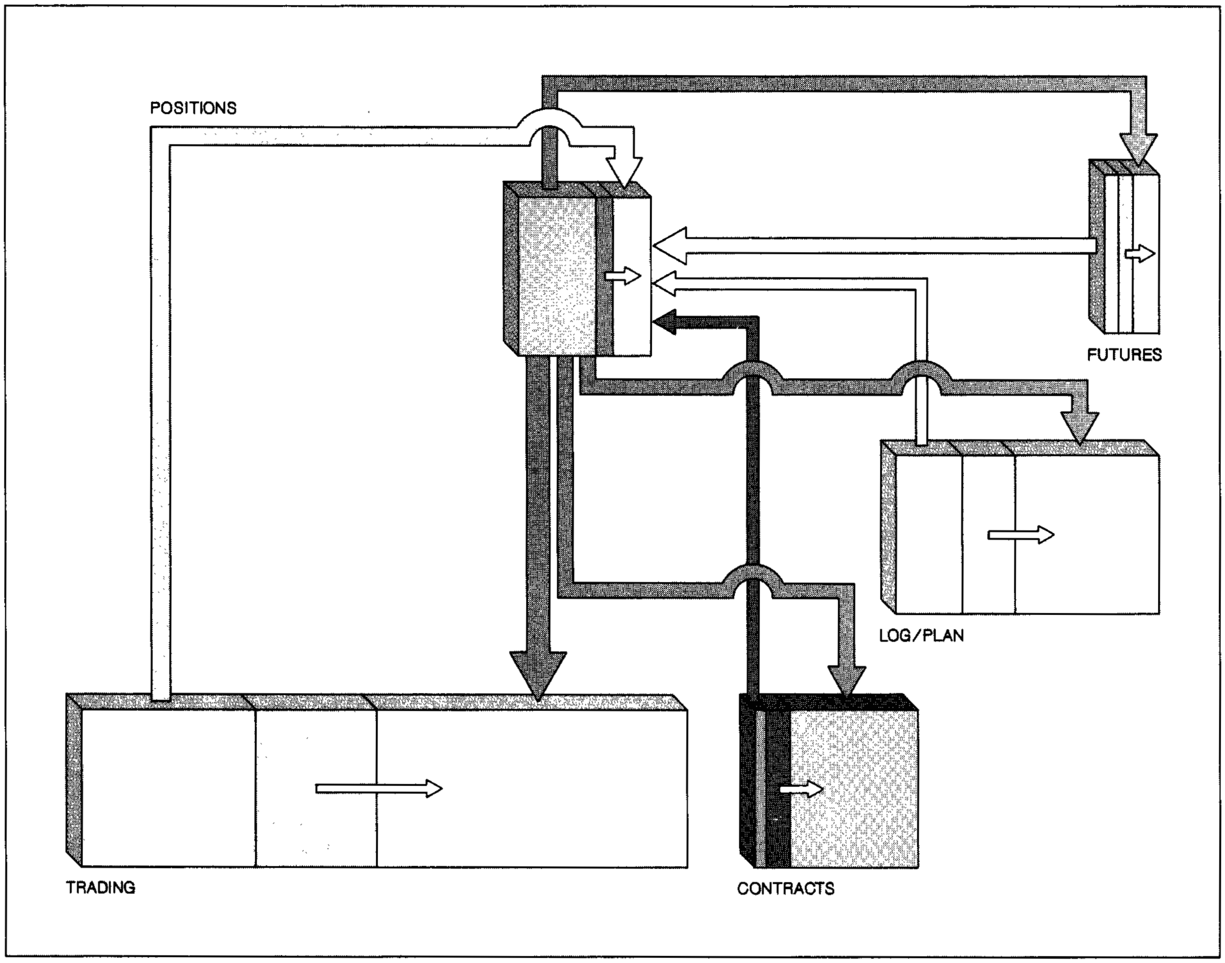
Tools used

ISMOD is a tool used to support a top-down information requirements study of a business or enterprise. It has been used in more than 450 studies of IBM customers worldwide and incorporates key quantitative analyses of the information requirements and flow throughout a company. In addition, it provides a simulation capability to determine the effect of implementing a new computer-based application. The dimensions of the business model include data and information, data criticism, data flow identification, data flow metrics, data qualifiers, event trigger, location, organization, and processes. Relationships are identified and maintained in the model among all of the above dimensions.

DevelopMate is a tool used to develop a business/enterprise model with dimensions that include business processes, business data, events, physical locations, and performing organizations. It incorporates an import function that will load from ISMOD business/enterprise modeling information.

The Logical Design Process (LDP) is a methodology with associated tool support that allows one to analyze an existing information management system (IMS) or related application and determine its propensity for distributed data processing. That is, one could define the business processes, databases, and physical locations and then take actual real-life data from IMS log tapes and build a business model showing processes, data and data flows, and locations where the processes were performed. This information comes from an existing centralized set of IMS applications. Simulations could then be performed to determine the feasibility of distributing or offloading the data to distributed processors so as to be

Figure 8 Example of process architecture



in physical proximity to the performers of the business processes. This tool was used for a number of customers from the banking, manufacturing, and transportation industries.

Concluding remarks

There was added value for every IBM customer cited here that utilized the approach of business/enterprise modeling as described in this paper. One customer, the oil company, indicated that there was no way in which they could have undertaken a worldwide requirements study of their corporation without the assistance of a tool such as ISMOD and that the study was one of the most important they had ever undertaken. Another, a power company, stated that the use of this approach to develop their strategic requirements for computer support allowed them significant cost savings in terms of shortening the study duration and avoiding the use of costly outside consultants.

A summary of the benefits of the value of business/enterprise modeling cited in this paper would include both quantitative and qualitative perspectives as follows:

- Qualitative—A greater degree of precision and enlarged scope is provided from the use of a rigorous methodology and tool support to analyze the business information needs of an enterprise. In addition, opportunities to improve business processes are identified, and an integrated systems architecture is produced.

- Quantitative—There are actual dollar savings associated with the approaches cited in this paper relating to cost avoidance of consultant services, less time devoted to these studies, and savings in time required for application and database design.

Acknowledgments

There are a number of individuals currently or previously employed by IBM that this author acknowledges as being the pioneers or pathfinders for business/enterprise modeling. They identified the need for tool support of these activities with the ultimate link into a repository and, ideally, the automatic generation of code. This concept dates back to at least 1977 in IBM. Included among these people are Hari Arora, Charley Baker, Dick Bergstresser, Ron Gale, Peter Hein, Hal Hichborn, Bernie Roth, Gene Sakamoto, Bob Tabory, Michel Veys, and John Zachman.

AD/Cycle, DevelopMate, and Repository Manager are trademarks of International Business Machines Corporation.

Cited references

1. J. A. Zachman, "A Framework for Information Systems Architecture," *IBM Systems Journal* **26,** No. 3, 276–292 (1987).
2. K. P. Hein, "Information System Model and Architecture Generator," *IBM Systems Journal* **24,** Nos. 3/4, 213–235 (1985).
3. K. P. Hein, "DevelopMate: A New Paradigm for Information System Enabling," *IBM Systems Journal* **29,** No. 2, 250–264 (1990).
4. *DevelopMate General Information, Version 1 Release 1,* GC26-4641-0, IBM Corporation; available through IBM branch offices.
5. J. M. Sagawa, "Repository Manager Technology," *IBM Systems Journal* **29,** No. 2, 209–227 (1990).

General references

C. T. Baker, "Logical Distribution of Applications and Data," *IBM Systems Journal* **19,** No. 2, 171–191 (1980).

Andre and Paul Blokdijk, *Planning and Design of Information Systems,* Academic Press (Harcourt Brace Jovanovich, Publisher), New York (1987).

BSP: Business Systems Planning, Information Systems Planning Guide, GE20-0527, IBM Corporation; available through IBM branch offices.

Enterprise Analysis, *IBM Systems Journal* **21,** No. 1 (1982, whole issue).

J. E. Dickson, *Information Quality Analysis Consulting Guidelines,* Document Number IQA-0001-4, IBM Systems Integration and Professional Services.

Information System Model and Architecture Generator (ISMOD), Study Guide, SB11-5989, IBM Corporation; available through IBM branch offices.

Information System Model and Architecture Generator (ISMOD), Systems Guide, LB11-5989, IBM Corporation; available through IBM branch offices.

J. G. Sakamoto and F. W. Ball, "Supporting Business Systems Planning Studies with the DB/DC Data Dictionary," *IBM Systems Journal* **21,** No. 1, 54–80 (1982).

J. A. Zachman, "Business Systems Planning and Business Information Control Study: A Comparison," *IBM Systems Journal* **21,** No. 1, 31–53 (1982).

Robert L. Katz *IBM Enterprise Systems, P.O. Box 700, Suffern, New York 10901.* Mr. Katz is a senior development analyst at IBM Sterling Forest. He joined IBM in 1962 working on the antiballistic missile system (Nike-Zeus) as a scientific programmer. Subsequent to that he has held a variety of positions in management, industry marketing, and consulting, as well has having been a project leader. His IBM experience in support of customers includes the following areas: banking, brokerage, defense systems, expert systems, federal regulatory agencies, hospitals, nuclear power, manufacturing, securities, stock exchanges, and transportation. He was the development manager of Securities Order Matching (SOM), an IBM program product for the brokerage industry, and represented IBM in Washington at the congressionally chartered Advisory Committee for a National Market System. His education includes an undergraduate degree with honors in physics from Brooklyn College.

Reprint Order No. G321-5415.