Track E1, Information Technology
Norris Peterson, Alstom ESCA
Thomas A. Green, P.E., ASCE Consultants
Arnold deVos, Alstom ESCA Corporation

Enterprise Integration via Data Federation

Distribution operations have interfaces with numerous other electric utility business processes in the Enterprise. An EMS/DMS might share boundaries with maintenance management, planning and engineering, customer information, financial and accounting, asset management and executive information systems. The ability to interface with these neighboring systems has become even more critical with deregulation and restructuring of the industry.

1) INTRODUCTION

At the heart of the requirement to integrate operations related applications is the need to reconcile the databases (or more generally, the data models) of these systems. The electric utility industry has addressed this need through a series of investigations sponsored by the Electric Power Research Institute (EPRI).

On a broader scale technology leaders, including Microsoft and Sun Microsystems, have introduced new enterprise integration technology. The world's leading software companies have built a multi-billion dollar business around integration utility business processes. Yet the utility industry is still looking for more efficient ways to establish integration of the operations related systems; Trouble Call Processing, AM/FM, Distribution SCADA, etc.

This paper describes an integration technology based on "Data Federation." Alstom ESCA has applied this technology to integrate multiple, independently developed, databases with each other. The Data Federation philosophy is that databases throughout an enterprise should be connected (federated). With Data Federation, databases are linked together so that the information continues to reside in their independent, native data environments. The paper traces the evolution of the electric utility enterprise, its vendors, and products being applied for enterprise integration. The value of the solutions offered by technologies is examined as a means of establishing the significance of the Data Federation approach.

2) EVOLUTION OF THE UTILITY TECHNICAL ENTERPRISE

The dominant vision for electric utility automation consists of integrated computing resources that allow information to be easily shared. The structure of the corporate-wide enterprise architecture is envisioned to consist of two classes of data processing systems sharing a common wide-area network.-- the Management Information System (MIS) and the Technical Information System (TIS). This architecture has evolved naturally and best represents the processing differences between business and operations. Now the demand to integrate the TIS and MIS to achieve greater staff efficiency and use of operational data will reshape traditional EMS applications.

Enterprise Integration In Electric Utilities

Electric utilities are beginning to embrace this technology in automating to meet the competitive market. Deregulation has not only reshaped the organizational structure, but it is also reshaping the implementation of automation. A Distribution Management System must now be viewed as part of a technical enterprise and must exchange data with a number of other operations-related systems. These include corporate data warehouses, customer information systems, trouble call systems, and substation automation systems. Utilities are faced with large-scale integration to provide new services and data for customer retention.

In response to industry restructuring, competitive threats, and regulatory reform, energy utilities have been shedding legacy system development in-house in favor of integrated suites of commercial software packages. These Enterprise Resource Management (ERM) solutions (also referred to as Energy Resource Planning –ERP packages) have traditionally provided functionality for finance, human resources, materials management, and, sometimes, maintenance management. The top-tier ERM solution providers have been SAP, Oracle, and PeopleSoft. They have responded to this shift by extending their functionality to include utility specific products.

Industry Sponsored Development

The Electric Power Research Institute (EPRI), at the urging of the utility industry, sponsored the work needed to develop standards for data definitions needed to define objects. The results have been presented to the IEEE and the International Electrotechnical Commission (IEC) which are the standard setting bodies that will ultimately issue and maintain these standards.

EPRI'S Utility Communication Architecture - Probably the first EPRI driven effort in this direction was the development of the Utility Communication Architecture (UCA), started in the late 80's, to define a 'framework' for communicating between systems used in various parts of the utility or between different utilities. The major outcome of this work was the EPRI UCA (now IEEE SC 36). This standard includes the ICCP protocol and the IEC TC57, Working Group 7, TASE.2 protocol.

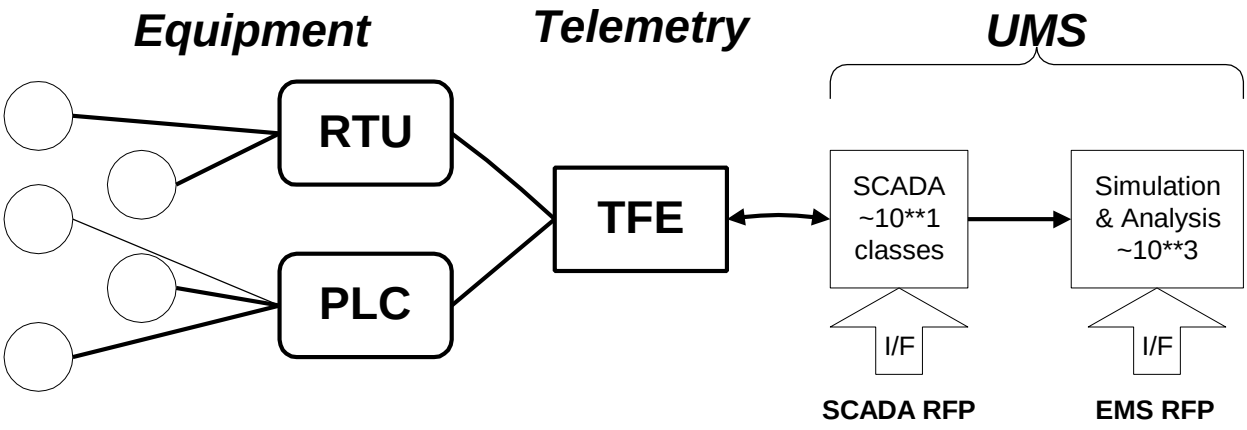
Although this standardized the communication protocol, it did not standardize the definitions of the data that needed to be communicated. This resulted in a second round of standard setting, known as UCA II, which can be traced back to the efforts of the IEEE/PES Substations Committee. The resulting standard, to be issued by IEEE and IEC, will have definitions of all hardware equipment in substations including all measurements. Most vendors of substation equipment and substation automation are already using these standards.

EMS vendors adopted the UCA communication protocols for applications to exchange information between the EMS of different utilities. However, UCA does not cover system data (line impedances, topology, operational limits, etc.); whereas a large portion of the EMS database is made up of this information. With vendors using proprietary database management systems to handle this data, such data could not be exchanged readily but more importantly, this prevented the use of third-party application software in existing EMS'.

EPRI'S CCAPI - A parallel development to UCA was sponsored by EPRI to define a common Control Center Applications Program Interface (CCAPI). The objective is to open up control center systems to integrate with software supplied from multiple sources. EPRI used the data definitions from their Operator Training Simulator (OTS), which is similar to an EMS database, as the starting point of this effort. This resulted in the standard called the Common Information Model (CIM). The CIM is a semantic description of EMS data to foster a standard for interoperation between EMS application components. The reference model has a common data access mechanism for obtaining CIM defined information and a message bus interface (MBI) for communicating between components via messages.

The small overlap between CCAPI and UCA has recently been brought into compliance and the CIM is on its way to becoming an IEEE and IEC standard. This work has been carried forward by the Object Management Group's Utility Domain Task Force which issued a request for proposal (RFP) to develop a standard data access facility that includes CIM defined data. A companion RFP is in preparation which will cover access to SCADA data. The latter is being developed cooperatively by the Utility and Manufacturing Task Forces. Between the two, a large number of new integration and interoperation possibilities will be made available.

Figure 1: OMG Utility Data Interfaces



In addition to the OMG work, the IEC TC57; Working Group 13 intends to standardize an EMS API, and Working Group 14 is defining a framework for DMS integration.

The CCAPI MBI has been slower than the CIM to take shape, perhaps because it tackles a larger problem. The software industry at large is addressing the same problem area with Message Oriented Middleware (MOM), application servers, and server-side component approaches as discussed below. The latest attempt to define MBI is being led by SISCO systems with a group of interested companies providing design review. The challenge is to define an architecture that does not conflict with broader industry software design trends.

The UIB is being offered as an alternative to the dominant integration approach that uses a centralized source database. In this latter approach data is exchanged through a large relational database, such as Oracle. Some success has been achieved with this approach but often the performance of real-time applications is unacceptable.

3) LEADERSHIP ENTERPRISE TECHNOLOGIES

Object technology is quickly moving from the lab and novelty applications to component based architectures suitable for mission-critical applications in the enterprise. Both CORBA and DCOM/NT are being promoted as competing architectures and have emerged as leadership enterprise technologies. This initial flowering of the technology is now being followed by an explosion of object technology applications being led by Microsoft's COM+, Sun's Enterprise JavaBeans, and OMG's CorbaBeans Enterprise frameworks. These developments will reshape corporate software deployment and are examined below.

Microsoft Transaction Server (COM+)

To help promote the concept of three-tier, or n-tier computing and server-side development, Microsoft's transaction coordinator, or MTS, was created. COM+. includes the Microsoft Transaction Server (MTS) development and deployment tools as well as a series of services, including events, a publish-and-subscribe event model, dynamic load balancing, and queued components via Microsoft Message Queue Server. MTS is comparable to Sun's Enterprise Java Beans (EJB) and Microsoft wants COM+ to go a step further, offering language choice, database connection pooling, and administration.

Microsoft may sometimes get off to a slow start, just as it did with Windows and the Internet, but leads the way with MTS and its capability to steer the third-party industry to MTS, and COM+ may again generate success. Another key to success is the number of desktops equipped with basic COM, which invites use of MTS on the server.

Sun's Enterprise JavaBeans Specification

JavaBeans is the standard Java component model promoted by Sun's JavaSoft Division and the Java supporters. The JavaBean is a software object that can be used as a component in building larger application programs. It is becoming the (non-Microsoft) standard for various object-oriented interfaces.

The age of building software application like Lego structures is just around the corner as Sun Microsystems gears up Enterprise JavaBeans (EJB). EJB is a specification for server-side Java components, which is designed to speed the development of portable server-based Java applications. EJB specifies new enterprise API's such as the Java Transaction Server, and the Java Messaging Server. This initiative includes the Java Naming and Directory Interface (JNDI) and a JavaBeans Bridge for ActiveX. JNDI is part of the Java enterprise APIs and supplies a standard method for corporate developers to crate Java-based applications while leveraging both existing and new standards among naming and directory services.

OMG's CORBAComponents
At the time of writing, the OMG is in the process of extending its Common Object Request Broker Architecture (CORBA) to provide component facilities. Using the upcoming OMG specification, an ORB vendor could provide containers for components written in various languages, as well as, an EJB container. Containers are software receptacles that know how to communicate with and manage the beans (components) in a computing environment.

The container provides a portability layer that lets anyone's components operate in anyone else's container. The components need only implement the proper semantics and naming conventions, and of course, they encapsulate business process logic.

Application To Utilities
Enterprise JavaBeans have taken root and are gaining favor with industry leaders. EJBs compete directly with Microsoft Transaction Server (MTS), and Microsoft's COM-based transaction manager. On a larger scale, however, EJB "container" model correlates closely with Microsoft's notion of "interception," a service-addition layer that the company is putting forth as a key element in its COM+ framework.

Some "early adopters" have employed these framework technologies to integrate the Technical Information System in demonstration projects. Some of the lessons learned include the difficulty of acceptance of such a complex technology by management who must approve the large budgets to apply these technologies. Also, staff retraining is considerable and the availability of experienced integrators is very limited.

4) LEADERSHIP ENTERPRISE APPLICATIONS

Who will provide the leadership for the continuation of the object software revolution? The answer can be found in the $10 billion business enterprise market. This market, driven by world market competition, Y2k, and the European momentary changes, is causing major companies to reinvent themselves. Microsoft, Borland/Inpris, and Compact/Digital provide very visible evidence of this business process reengineering. The following gives the market size in revenues and 1997 revenue growth of the four enterprise applications leaders (Oracle figures approximated for the Applications Group business):

? ? SAP; Walldorf, Germany ---------- $ 3.345 Mil. (67%)
? ? Oracle; Redwood City ---------- $ 1,245 Mil. (7%)
? ? PeopleSoft; Pleasanton, CA ---------- $ 1,210 Mil. (81%)
? ? Baan, The Netherlands ---------- $ 820 Mil. (75%).

The outcome of this titanic market share struggle will dictate the technologies that will be adopted for the corporate enterprise. Below, the products and object technologies of the leading enterprise application software vendors are examined. Utilities are embracing these enterprise integration solutions as the most viable and available solutions. The common characteristics of these products are that they are designed for large-scale vertically-integrated applications to replace existing systems at high purchase and integration costs.

SAP AG'S Enterprise Impact

SAP of Walldorf, Germany controls 1/3 of the $10B enterprise market, larger than the next seven suppliers, including Oracle's Applications Group business. This market leadership has been built around a family of software applications called R/3. SAP's R/3 integrated business software runs the back office of nearly half of the world's 500 largest companies.

R/3 was introduced in '92 using client/server technology and sales hit $1.5B by '95. It was later extended to use three-tier C/S. R/3 addresses the Management Information System side of the utility enterprise and consists of a set of modules that automate business processes. R/3 contains over 70 modules which include: Sales, Inventory, Production Scheduling, Human Resources, Purchasing, etc. R/3 has also been adapted to work with the Net so companies can share data with customers and suppliers. This could easily include real-time operating and production data. Adoption of standards and interfaces compatible with SAP will be requested of EMS vendors.

SAP's R/3 is a complex set of programs that can take several years to roll out. It requires all corporate sites to adhere to the same precise business process and often forces re-engineering. Installing R/3 involves large numbers of specialists that can increase the cost of an installation three to five times the software's price tag. A $20B industry has grown up around SAP.

Oracle's Traditional US Leadership

Oracle built its leadership in the enterprise market around its relational database technology. Relational Database Management Systems (RDBMS) provide the breakthrough technology for businesses that invest heavily in client/server systems. Oracle provides a means of centralizing operational and business data for many US electric utilities.

Oracle is now moving to expand their share of the enterprise market. Oracle is reengineering both its application products and its applications sales methodology to expand its enterprise market share. In the products area, Oracle is focusing on the process automation and front-office applications. In the sales engineering area, the focus is on reducing product cost, less complicated product packaging, channel sales, and faster approval of deals.

Oracle's flagship product is Oracle 8 which was introduced in June of '97 and has already shipped several thousand units. Oracle 8 marked Oracle's entry into the object-relational database realm with the capability to manage non-relational data types. It also provided large scale (over 1 terabyte) database support. With Oracle 8, users can move incrementally to the object model by accessing data either with object interfaces or through conventional tables. Oracle's Network Computing Architecture relies on "cartridges" to support different data types and the CORBA Internet InterORB Protocol (IIOP) to connect application servers with the database.

Oracle databases are widely used in electric utilities and have been successfully applied to exchange data between dissimilar systems. According to users, the major weakness of this approach is unacceptable response times for large-scale, real-time applications.

Peoplesoft Challenges Oracle

PeopleSoft is challenging Oracle as the number two supplier of enterprise resource planning software. Its latest product suite, PeopleSoft 7.5, was introduced in March '98. To achieve the number two position, it must prove that it can be a viable force in global markets and satisfy large customers with multi-national requirements. PeopleSoft must also continue to expand into vertical markets such as financial and health care, and increase its efforts to attract mid market customers. PeopleSoft must prove that it can play in the accelerating world of distributed computing.

PeopleSoft 7.5 offers financial, accounting, payroll, and even some manufacturing applications for international markets. Before this version, PeopleSoft only sold a Human Resources (HR) product outside of North America. In addition, the new version includes business-productivity applications and packaged software bundles for vertical industries such as health care, finance, and manufacturing. PeopleSoft has not, as yet, targeted the electric utility industry.

BAAN's Enterprise Approach

Baan has been aggressive with application suites aimed at high-profile vertical markets such as automotive and aerospace. They have now targeted the mid-size companies looking to implement enterprise applications quickly and easily with Baan-on-Board, and integrated software and services packages. This is the largely untapped middle market of companies from $50M to $500M. This effort was launched in early '98 with hardware partners (Compaq, Digital and HP). It is aimed again at vertical markets, starting with wholesalers and machine builders.

Baan also intends to introduce a series of Web based applications called Baan E-Enterprise. They will be designed to extend an enterprise's supply-chain by using mobile agents over the Web to search key suppliers for materials, check product inventory, lock in orders, notify accounting systems, and handle shipping. PeopleSoft has not, as yet, targeted the electric utility industry.

5) ALSTOM ESCA'S DATA FEDERATION TECHNOLOGY

More and more electric utility Information Technology Departments are faced with integrating islands of computing technology. As pointed out earlier, new subsystems are being integrated throughout the utility enterprise by utilizing a myriad of software technologies. Legacy functions are being replaced by new capabilities while more control center data must be integrated into enterprise-wide operations. These subsystems include energy accounting systems, customer information systems, and functions emerging from the deregulation initiatives.

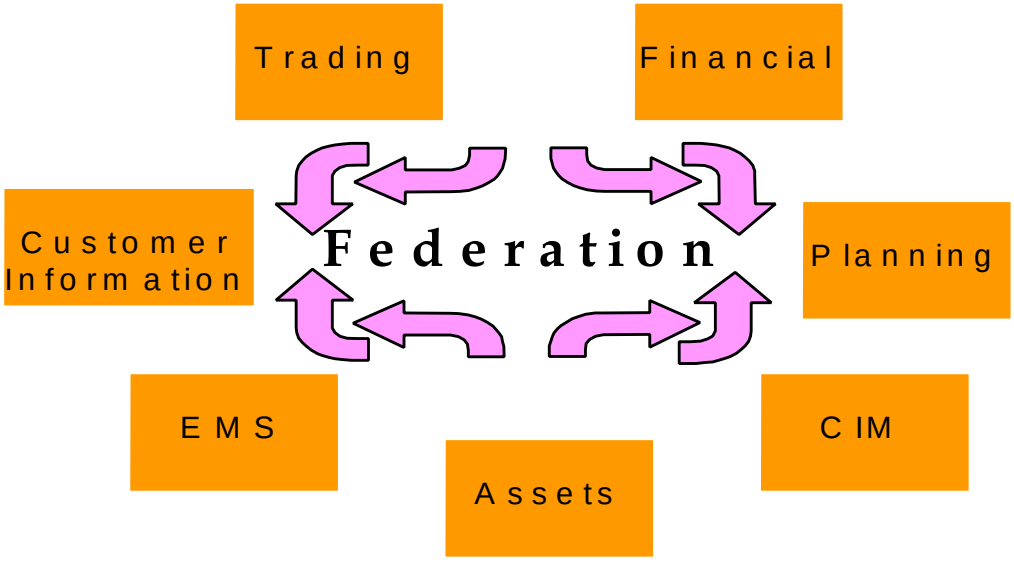
In addition, electric system operations segments have interfaces with numerous other business processes. A Distribution Management System might share databases with maintenance management, planning and engineering, customer information, financial and accounting, asset management and executive information systems. The ability to interface with these neighboring systems has become even more critical with the rapid restructuring of the industry.

The Data Federation Solution

At the heart of the requirement to integrate operations related software is the need to reconcile the databases (or more generally, the data models) of these systems. The term "Data Federation" has been adopted to describe the integration of multiple, independently developed, databases with each other.

The Data Federation philosophy is that databases throughout an enterprise should be connected (federated). With Data Federation, independent databases are linked together while information continues to resides in independent data environments. Data Federation links enterprise, legacy, and new application data as illustrated in the Figure below. This approach parallels that used for application integration where applications are made to work together as a whole instead of building new applications

INTEGRATION THROUGH FEDERATED DATABASES



A key component to support this integration is the coordination of the different information models residing in different data environments that make up these systems. In order to address these challenges, a better approach to the enterprise modeling is needed than that afforded by a central, "source database" design. The traditional approach to system-wide modeling has been to combine data from various databases into one large centralized database or "super model." With this approach, special converters must be developed, then modified every time models are revised or new applications and data are added to the system. This results in an ongoing maintenance chore and a continuous drain on programming resources and budgets.

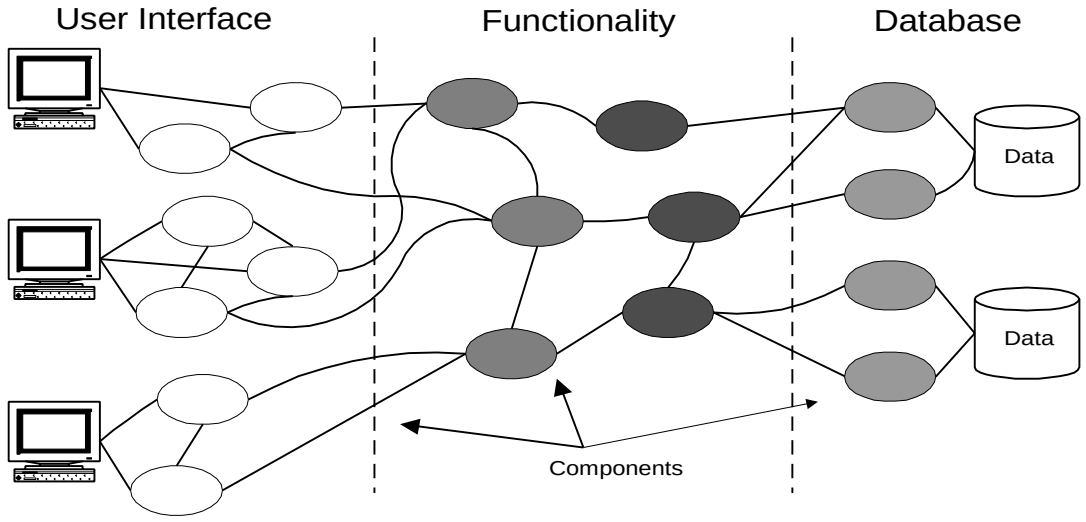
In contrast, the Data Federation strategy combines information from different databases residing in different islands of technology across the enterprise into one unified view, while keeping existing databases in tact. That is, instead of concentrating on one, "super model" database where all models reside, Alstom ESCA is taking the approach of constructing tools to support the

integration of information across different data environments. Data continues to reside in its "native" database and each database remains independent. Enterprise investments in existing systems and databases are preserved and the systems are administered, as they are now, independently with no special coordination required.

Data Federation is relevant to the typical organization that has developed multiple, overlapping, departmental databases. With business process re-engineering, the former departmental interfaces may not make sense anymore. There is also increased pressure to streamline business operations by merging and connecting databases. Data Federation's primary purpose is to provide services and infrastructure to build electric utility Technical Information Systems.

Data Federation Technology

The technology is based on a multi-tiered model for application integration, partitioning the application into the user client, a set of shared application services, and the data server as shown in the Figure below.



User Tier - The user interface is the client software which provides interfacing with the model validation functions and provides feedback messages from the middle layer servers. With the client software, data can be viewed (summary or detail form) and entered (including creation/deletion of objects and updating attributes). The user can also navigate between different parts of the model and between specific objects of interest. However, no validation code will reside on this layer.

Application Services Tier - The application services provided to apply Data Federation at the server level consist of the following:

? ? *Schema Browser* - Data Federation uses a schema browser to view the model and its objects and provides a unified view of all the model schema developed on the supporting database engines. This tool can be used to view and modify a part of the model schema.

? ? *Schema Mapping* - Data Federation makes use of the schema mapping service to build

mapping relationships and property maps between parts of schema residing in different database engines. The mapping service is a key federation facility that can be used to unify or segregate the model schema as necessary to suit the application.

? ? *Data Replication* - Data Federation makes use of a data replication service to populate data from one database environment to another database environment. The data transfer follows the guidance of the mapping relationships and property maps built between the databases. Converters are developed as necessary to perform data conversion.

? ? *Data Access* - Data access services are used to store data in the supporting database systems. The database adapters handle the peculiarities of each database and the user sees no difference as to where a piece of data resides.

Data Tier -- This represents the data persistent storage which is defined through a schema (characteristics of objects within a framework). The model schema describes the system model to the underlying database management system. Data Federation requires a model schema which describes the power system and associated control and communication equipment used to support power system operations. The database federation facility allows different sections of the model schema to reside in different database systems. However, the sections of model schema can be unified such that the model data can be presented to the end user as an integrated model.

This is accomplished by building software adapters that transform the special data representations and data access mechanisms in each data subsystem into a common, object oriented application programming interface (API). All databases across the enterprise can be displayed in graphical form, where they can be accessed interactively. Associations between like data in different subsystems can be established through this visual interface. These links are maintained in persistent storage and can be navigated by programs in the same manner as native associations in individual data subsystems. Built-in data synchronization software reads the schema and synchronizes data between the related data subsystems.

Data Federation's Promise

The Data Federation API provides a powerful mechanism for application development so one can view, select, and combine data from any data subsystem into a new application. To enable communications across multiple, heterogeneous hardware and operating system environments, Data Federation employs a distributed architecture based on industry-standard CORBA and Microsoft COM technologies. This enables integration of federated data in desktop applications, and the creation of application user interfaces using web page tools, Windows Visual Basic, JAVA, Delphi, and other popular development tools.

Alstom ESCA's Data Federation technology offers many advantages to the enterprise:

- ? ? Provides a single point for the entry of data that may need to be replicated in multiple subsystem databases.
- ? ? Provides a straightforward, consistent user interface for navigating to, viewing, and updating data models.
- ? ? Allows data to reside in its indigenous data environment minimizing copying and data synchronization issues that exist with centralized database approaches.
- ? ? Facilitates creating the maps required to bridge common data in different data environments.
- ? ? Uses industry standard technology (CORBA and OLE) for intra- and inter-process communications across heterogeneous hardware and operating system environments.

Alstom ESCA is introducing applications using this integration technology that will simplify enterprise integration and continue to provide for the automation needs of electric utilities. The Data Federation approach is particularly appropriate with the current level of rapid change that is underway. As companies restructure, disaggregate and are impacted by mergers/acquisitions, the needs to present an integrated view and access to data is growing more rapidly than automation can be implemented.