Powerful Solutions for XML-Based Data Integration™

Architecture & Product Overview

TABLE OF CONTENTS

Executive Summary

Enterprise IT needs a new architecture for integrating and sharing heterogeneous data. Dynamic self-service value chain applications are pushing past the limits of existing integration technologies, showing them to be too cumbersome and inflexible to meet the challenge of eBusiness. The rapid development of new business processes, in response to market opportunities made possible by ubiquitous Internet connectivity, demand real-time, unified access to the corporate information assets, regardless of the source.

Open Internet standards are providing the building blocks for the next generation of integration technologies. XML, the XML Query Language, DOM, and the new Web Services standards; UDDI, WSDL, and SOAP, offer invaluable specifications that allow business partners, suppliers, customer and far flung parts of the organization to seamlessly share business data.

Leveraging the power of Internet standards, Enosys provide the first end-to-end XML-based data integration solution. The Enosys Information Integration Infrastructure offers state-of-the-art data integration software to support the development of the next generation of self-service value chain applications. Using sophisticated query, transformation and mediation technology, Enosys reconciles disparate data models in real-time, providing unified XML views of live, multi-source business data.

Integrating information at the data layer, Enosys eliminates the time consuming, costly and complex application coding associated with traditional integration technologies. The Enosys Information Integration architecture and product family incorporates advanced research into XML and database technology and offers a pure-XML solution for integrating and querying distributed information sources.

The Data Integration Problem

Value chain applications that support efficient, tightly integrated business processes demand access to real-time information from a multitude of information systems. Providing dynamic, unified access to these data sources is a significant problem as organizations increasingly represent, maintain, and export data using a variety of formats, data models, interfaces and semantics. In particular:

- Different types of information reside in different systems, have different structure, and are usually in heterogeneous formats;
- Data coming from different sources need to be joined or otherwise appropriately combined;
- Different types of data change at different rates, making data warehousing a logistical challenge;
- Different applications often need different views of the data.

Problems with EAI

Existing integration technologies have attempted to provide access to combined legacy, eBusiness and XML data, but have failed to satisfy the demands of value chain applications. Primarily designed to connect ERP systems and legacy applications, Enterprise Application Integration (EAI) cannot provide the dynamic flexibility and real-time access demanded by eBusiness. EAI linkages between applications are cumbersome, labor-intensive and require a significant amount of development time, and even then connect only a limited number of applications. Created for systems that sit behind the company firewall, under the watchful eye of IT, EAI does not respond well to the needs of open, shared Internet access to data.

Manually Intensive

EAI software has made significant advances commoditizing integration technology, but it still requires significant programming manpower. No two software applications look alike, and integrating the different architectures requires comprehensive technical understanding of the applications being connected, and the software used to make the connection. The result is a tightly bound, application-to-application, linkage that is unresponsive to modifications. Change at either end of the integrated connection results in the entire effort being manually reworked.

Restricted Scalability

Value chain applications invariably require data from many different sources in order to provide a unified view of the business process. Point-to-point integration solutions, like EAI, are limited in their ability to connect more than two or three applications. Tying more systems together quickly results in a complex, unwieldy mess of connections, stifling the dynamic scalability needed for eBusiness.

Time Consuming

A typical point-to-point integration project will frequently take months to implement. Each integration effort requires enormous amounts of complex, manual coding, which means

time spent in design, development and testing of software to ensure that systems being connected are not adversely affected.

No Cross-Firewall Connectivity

Existing application integration technologies work well in connecting systems inside the enterprise IT data center. But when communication is required beyond the corporate firewall, where processes being connected are independently owned and managed, the application integration framework fails. Proprietary data sharing protocols, and the need for comprehensive knowledge of both applications being integrated, create obstacles to smooth integration.

The Benefits of the Enosys Platform

Unified Legacy, eBusiness and XML Data

The Enosys Information Integration Infrastructure addresses these heterogeneous data challenges using a high-level, XML-based, declarative query language and sophisticated query mediation technology. Applications and end-users request information found in various data sources using SQL-like data access commands. The Enosys infrastructure returns consistent, integrated, XML result sets, regardless of where the information sources store the information or how the information is accessed.

Integration Within and Beyond the Firewall

Using World Wide Web Consortium (W3C) standards as a foundation, Enosys implements the XML, XML query language, and Document Object Model (DOM) standards for open, flexible sharing of data across the Internet. XML queries can be submitted from any browser-based, or Java value chain application inside or outside the corporate firewall. Queries return self-describing, well-formed XML result sets to the calling application.

Enosys fully supports the infrastructures and architectures driving the development of Web Service, including: Universal Description Discovery and Integration (UDDI), a directory lookup system for businesses and web services; Web Services Description Language (WSDL), a web services description standard; and Simple Object Access Protocol (SOAP), the remote procedure call standard.

Time to Market

Enosys bypasses costly, time-consuming and complex application integration by integrating at the data layer. Developers gain immediate access to data sources for use in applications and web services. A suite of integrated development software provides developers and end-users with rapid access to XML views.

The Enosys Platform Architecture

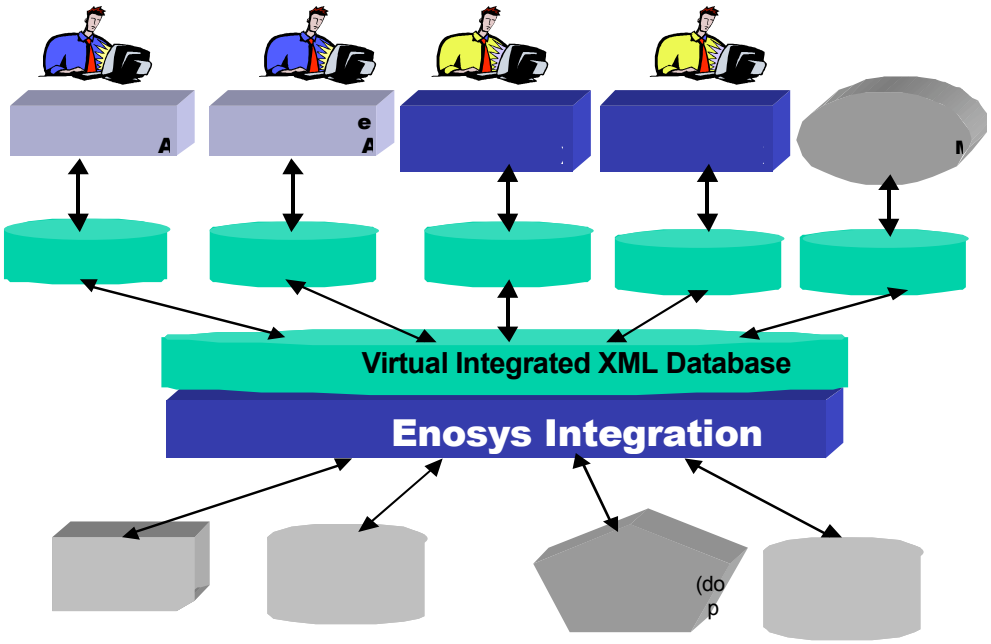

Figure 1 - Enosys Integration Platform

The Enosys Information Integration Infrastructure accesses multiple, distributed, heterogeneous data sources and exports a virtual integrated XML database and views (VIXDB™, VIXView™). The VIX database and views enable front-end applications to seamlessly access distributed heterogeneous information as if it were stored in a single XML database. Information from the distributed sources is organized into XML objects that conform to the application's needs.

For example, to a private trading exchange application the VIXView can provide front-end access to an integrated catalog, where the heterogeneities between the suppliers' products are resolved, and the products are integrated and classified according to the needs of the marketplace. Each product object contains catalog data along with attributes from the pricing, delivery, CRM, and other databases. The VIXView provides distribution transparency, that is, the originating sources and methods of access are transparent to the application.

The VIXView architecture addresses the three key pain points of accessing distributed and heterogeneous information:

- Discovering the source of each piece of relevant information
- Accessing the sources to retrieve the necessary information
- Putting together the accessed information

When the application issues a query to the VIXView, the integration modules of the Enosys platform decompose the query into requests that are directed to the dynamic and cached sources. The VIXView components that correspond to dynamic and quick-

responding sources are collected and computed on-demand, in real-time. The VIXView components that are derived from slow and static sources are collected, integrated, and cached in advance.

In addition to the integration platform, Enosys provides the XSDesign family of software, enabling the rapid development of Web front-end applications for rendering the integrated information. An initial application template for eCommerce product selection has been developed. This template enables the customer to receive product selection advice and seamlessly access, evaluate, and compare the suitability of manufacturer or multiple supplier offerings. The front-end applications are automatically generated and domain, or business, experts need only declare how products should be queried and what product selection advice should be provided.

Platform and Feature Overview

The Enosys Information Integration Infrastructure consists of a suite of components that work together to provide integrated access to distributed data sources. The Information Integration Server provides the core functionality, accessing and processing queries in real-time, and is accessible to applications through a query language API and a DOM (Document Object Model) API.

XSDesign offers a web-base front-end to easily build query forms and application templates to display views of the data that are easily customized to the needs of the viewer.

The Enosys platform also includes a full set of management tools that enable the user to easily create and manage front-end applications, view definitions, data transformations, and source connections.

XMLizers

The XMLizer family of tools turns structured and semi-structured data into virtual XML views. The XMLizer wrappers instantly encapsulate databases, HTML repositories and file servers creating XML data sources. Enosys offers:

- Relational wrappers, exporting a virtual XML view of any JDBC-compliant database;
- HTML wrappers, translating HTML pages into XML;
- And comma and tab-delimited file wrappers.

Additional wrappers can easily be written for other data sources. Existing investments in XML information exchange, such as adapters for WebMethods or STC, can provide excellent leverage for the development of wrappers.

Information Integration Server

The Enosys Information Integration Server accesses the virtual view exported by the wrappers and provides a virtual integrated XML view to applications. The unified view appropriately transforms and integrates data from information sources into XML conforming to the XML schemas of the target application.

The transformation and integration is rapidly and concisely specified using the XML Query Language.

The XML query language is a high-level, declarative query and view definition language for XML. It meets the requirements of the W3C workgroup on XML Query Languages and has additional features for processing XML data with a loose structure. Transformations expressed in a brief XML view definition can easily resolve important integration problems, including name, value, and classification heterogeneities. The XML query language cuts the cost of developing source-to-target transformations compared to using Java or XSLT.

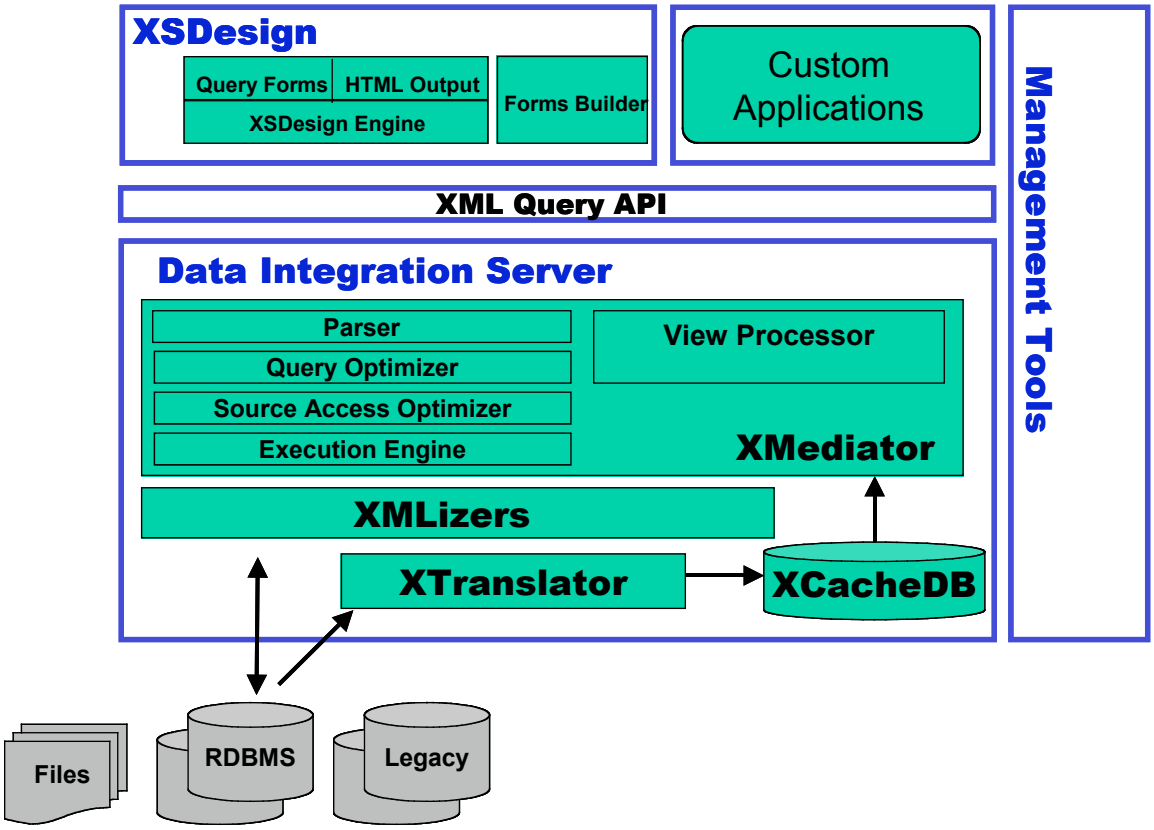

Figure 2 - Enosys Platform Features

The Enosys Information Integration Server processes XML query language queries against VIXViews. The view processor transforms the query to replace references and conditions on views with references and conditions on the actual data sources. The query is then translated into a query plan and is processed by the query optimizer. The source access optimizer chooses an efficient way of decomposing the optimized query plan into requests that are sent to the downstream information sources. The query plan is finally run by the execution engine, which sends the requests to the XML wrappers, collects the information, and composes it into an XML query result.

The Information Integration Server also provides easy customization of virtual views that provide gateways to other eBusiness applications. These views can easily convert XML

data from one XML standard to another in real time. Information Integration Server views can support applications within a business enterprise or be extended to work with applications that support business-to-business transactions.

XTranslator/XCacheDB

The Enosys Information Integration Infrastructure uses the XTranslator and the XML Cache Database (XCacheDB) to cache the components of the VIXView's that correspond to slow or static data sources. The transformations and caching can happen on demand or at regularly scheduled intervals, as directed by the XTranslator administration interface.

XCacheDB is a native XML database for storing XML documents. Using a JDBC compliant relational database for storage and query processing, XCacheDB responds to XML queries with well-formed XML document results. Automated tools load the database with XML data so the developer need not be aware of the underlying relational database. The XCacheDB architecture uses proprietary storage and query processing algorithms to deliver improvements in run-time efficiency, and reduced development and maintenance costs.

XSDesign

The XSDesign family of tools enables the rapid development of customized Web front-ends that use integrated XML views, made available through the Information Integration Server. XSDesign is used by the business analyst and provides:

- Forms for parametric querying of the data sources in the integrated view;
- Summarization and navigation of large query results;
- Query assistance in formulating and refining queries;
- Advice on product selection based on domain expertise.

Management Tools

The Enosys Management Tools provide a comprehensive system administration and development environment for managing the data integration platform.

Query Builder

A graphical query builder tool provides for the creation of XML view definitions and queries. After importing the XML schema of the data sources or existing views, users define joins, function invocations, and filtering conditions on the input data using wizards and a drag-and-drop user interface. The builder allows the user to graphically arrange output data into a different XML schema by specifying mappings to the input schemas, groupings, and creation of new elements.

XML Data Entry

Enosys provide an innovative XML editor, data entry, and data management tool for importing the XML schema of a data source. The tool can import a set of simple presentation annotations, and generate a data entry form based on the annotations, enabling rapid data entry, update, and validation driven by the XML schema.

Management Console

The management console is an integrated development environment for Web-based applications that access multiple information sources. XConsole integrates tools to allow for XML view and query construction and testing as well as for building, deployment and management of XSDesign-powered Web front-ends that access the integrated views.

Summary

Traditional data integration technologies are hampering the adoption of value chain applications. Proving inadequate in responding to the dynamic pace of eBusiness automation, technologies like EAI are cumbersome, time consuming and have limited scalability. Self-service value chain applications need integration technology that leverages open Internet standards for seamless data sharing inside and outside the enterprise firewall boundary.

XML's rapid move into the mainstream of enterprise computing is a testament to both the universal applicability of open standards, and the need within enterprise IT for technologies that can leverage the global connectedness of the internet. The self-describing ease-of-use provided by XML fits perfectly with the demands of new value chain initiatives that are speeding the flow of information between enterprises.

Leveraging open standard building blocks, the Enosys Information Integration Infrastructure provides a pure-XML platform for accessing real-time, unified heterogeneous data. Eliminating the manual coding necessary with EAI solutions, Enosys reduces long-term cost of ownership, and provides a scalable data integration platform for value chain computing. Providing significantly better time-to-market, scalability and portability than traditional point-to-point integration, the Enosys Information Integration Infrastructure is the only solution that delivers the flexibility needed by enterprise IT to respond quickly to new network economy business opportunities.