

ZAPTHINK ZAPNOTE™

DATA DIRECT SIMPLIFYING DATA INTEGRATION WITH XQUERY

Analyst: Ronald Schmelzer

Abstract

The more companies attempt to connect their disparate systems in the enterprise, and the more success they have doing so using emerging technologies and approaches such as XML, Web Services, and Service-Oriented Architecture (SOA), the more they realize that simply getting systems to communicate is only half the problem. The other half is getting those systems to understand what they are saying to each other. While many people pin their hopes on abstract concepts such as semantic webs and ontologies, they require a more practical short-term solution to enable humans to map data structures between dissimilar systems.

Fortunately, the W3C standard XQuery has recently emerged as a robust technology that is well suited for common data integration needs. Rather than using proprietary data adapters, tightly-coupled integration middleware, or complex mapping technologies, the XQuery approach applies standards-based documents and general-purpose transformation engines to the problem of data integration. In this ZapNote, we explore XQuery and DataDirect XQuery™ as an illustration of this powerful technique in practice.

All Contents Copyright © 2006 ZapThink, LLC. All rights reserved. Reproduction of this publication in any form without prior written permission is forbidden. The information contained herein has been obtained from sources believed to be reliable. ZapThink disclaims all warranties as to the accuracy, completeness or adequacy of such information. ZapThink shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The reader assumes sole responsibility for the selection of these materials to achieve its intended results. The opinions expressed herein are subject to change without notice. All trademarks, service marks, and trade names are trademarked by their respective owners and ZapThink makes no claims to these names.



Rethinking Data Integration in an XML-Centric World

One of the most appealing benefits of Service Oriented Architecture (SOA) is that organizations will be able to combine the capabilities and functionality of multiple, heterogeneous systems and change them at will as the business requires. Companies are increasingly realizing such benefits but are also finding that simply getting systems to talk to each other is a small part of the challenge of making distributed computing in a heterogeneous infrastructure work. Indeed, the bigger challenge comes from trying to piece together data from these multiple systems in a way that makes them easy to understand and process.

Traditionally, building such data integration logic is costly, time-intensive, complex, and error prone, because data integration specialists typically take a point-to-point approach, using a combination of brittle integration middleware, proprietary adapter interfaces, and human-intensive mapping. The net result is that businesses build one-time-use data integration projects that are difficult to manage and impossible to reuse or share across the organization, wasting a considerable portion of their IT budget in the process.

While the eXtensible Markup Language (XML) aimed to solve the problem of interoperable data via a standards-based language for data integration, it doesn't provide the actual mechanism for integrating with endpoints nor converting different representations of XML data into different formats, giving rise to XML Stylesheet Language Transformations (XSLT). But XSLT over time proved too cumbersome and too complex to manage for large-scale data integration. The XQuery language was designed to support this kind of data integration, as well as native XML queries, and it is now gaining traction for both purposes.

An XQuery program's primary functions are to select, filter, transform, join, and aggregate data across multiple XML-based data sources. XQuery simplifies data integration through XML querying and transformation by virtue of its concise and powerful syntax. Users develop XQueries that join or aggregate data from diverse data sources, and produce XML as output. The results can then be fed directly into applications or stored for use in further data transformation requirements. The XQuery API for Java (XQJ) adds onto these capabilities with a standard interface for easily embedding XQueries in any Java program in much the same way as the Java Database Connector (JDBC) does today for SQL.

However, while XQuery offers significant power for data integration, developers still require an easy way to develop and manage XQuery-based data integration. This means that developers need tools to build the XQueries as well as to transform data from original data sources into an XML format. The ability to abstract relational and other data formats into XML and the ability to aggregate both the inbound composite data from those sources is the key to success with XQuery. In the future, bi-directional insert and update support back into heterogeneous systems will also become important.

DataDirect XQuery™ the New Breed of Data Integration Product

DataDirect Technologies, a provider of XML development tools and XQuery processing components is readying the charge for XQuery-based data integration with their eponymous product *DataDirect XQuery™*. The product provides unified data access across XML and relational data, leveraging XQuery as the primary language for data manipulation, composition, and transformation.

DataDirect XQuery is an embeddable software product based on the XQuery and XQJ standards that enables Java developers to access both relational and XML data sources using

XQueries in their applications. The product supports integrating data originating from XML, relational databases, or a combination of the two using a single XQuery. DataDirect XQuery enables users to build high performance applications that process XML and relational data sources, by employing special query optimization and mediation techniques when accessing relational data. For example, DataDirect XQuery pushes down as much of the query burden (e.g., join operations) as possible onto the relational database server, rather than performing such expensive operations in memory or at the application server.

DataDirect built the product specifically to be an easily embeddable software solution, not requiring any additional product or application server. It has no server of its own, and is not tied to a specific vendor database or database version. As a result, DataDirect XQuery is appropriate for software developers and independent software vendors (ISVs) whenever they need to deal with XML either as part of a set of heterogeneous data sources or as the target format for a data integration application.

The product supports the latest working draft of the XQuery specification for both XML and relational sources, and has native direct data access to Oracle, Microsoft SQL Server, IBM DB2, and Sybase databases. The product provides XQuery API for Java (XQJ) 1.0 specification version 0.2.1 (early draft review) support, and includes *DataDirect Spy for XQJ*, a utility for logging XQJ calls at runtime. Developers can design, test, and deploy their XQuery-based applications using either Stylus Studio, DataDirect's award-winning XML IDE with advanced XQuery development support, or use the <oxygen/> XML Editor for Eclipse (DataDirect XQuery Edition), a tool for browsing data sources, writing and testing XQuery queries, and embedding XQuery queries in Java applications.

DataDirect Products

DataDirect (a division of Progress Software)

Availability: Now

Overview:

DataDirect Technologies offers a wide range of commercial-quality database drivers for major databases and a wide range of applications. The company provides a wide range of ODBC drivers, JDBC drivers and ADO.NET data providers. The most recent line of products includes solutions for applications that need to process relational and XML data in a consistent, standards-based way across databases.

Product Details:

The company's product suite includes the following offerings:

- *DataDirect XQuery* – DataDirect XQuery is an embeddable software component based on the XQuery and XQJ standards that enables Java applications to access both relational and XML data sources using a single query. DataDirect XQuery is embeddable, standards-based, database-independent and operating-system independent.
- *Stylus Studio XML IDE* – Stylus Studio provides an Integrated Development Environment (IDE) for XML and its related technologies: XSL, XSLT, XML Schema, DTD, SOAP, WSDL, SQL/XML, XML mapping and XQuery.
- *ODBC, JDBC, and ADO.NET drivers* – The wire protocol design of DataDirect's drivers eliminates the need for database client software and libraries, which simplifies installation and administration and improves performance.

- *DataDirect SequeLink* – DataDirect SequeLink is an end-to-end middleware solution that provides a universal client for ODBC, JDBC, ADO or .NET.

Key Differentiators:

- *Easily embeddable* – plugs into any architecture, does not require a server
- *Works with most databases and application servers*
- *Leverages the latest and most important standards for data access and XML*

The ZapThink Take

It's clear that data integration remains one of the greatest challenges for businesses to solve. Standards by themselves can't solve the problem, but good implementations that leverage the power of standards can. In particular, we've seen increasing adoption of XQuery, but many vendors have introduced proprietary XQuery extensions and application-specific implementations that further tightly couple the user to the particular implementation. DataDirect XQuery, however, is a better alternative. By leveraging widely accepted standards for data integration and leaving it up to the user to determine the runtime infrastructure for those data integration services, DataDirect XQuery offers a chance at solving long-lived data integration challenges. As companies look increasingly to XML and XQuery for their data integration solutions, approaches that depend on open standards and loosely coupled approaches to data integration will win out.

DataDirect Profile

Profile: DataDirect	March 2006
Funding:	Division of Progress Software
President:	Richard D. Reidy
Employees:	NA
Address:	14 Oak Park Drive Bedford, MA USA 01730
URL:	http://www.datadirect.com or http://www.stylusstudio.com
Phone:	800-876-3101
Contact:	Nancy.vodicka@datadirect.com

Related Research

- *Service Orientation Market Trends Report* (ZTR-WS110)
- *SOA Tools and Best Practices Report* (ZTR-WS107)
- *Business-to-Business Data Integration in a SOA World ZapForum Podcast* (ZTP-0208)
- *DataDirect Technologies: Deploying XML Data Services with Stylus Studio ZapNote* (ZTZN-1181)

About ZapThink, LLC

ZapThink is an IT advisory and analysis firm that provides trusted advice and critical insight into the architectural and organizational changes brought about by the movement to XML, Web Services, and Service Orientation. We provide our three target audiences of IT vendors, service providers and end-users a clear roadmap for standards-based, loosely coupled distributed computing – a vision of IT meeting the needs of the agile business.

ZapThink helps its customers in three ways: by helping companies understand IT products and services in the context of Service-Oriented Architecture (SOA) and the vision of Service Orientation, by providing guidance into emerging best practices for Web Services and SOA adoption, and by bringing together all our audiences into a network that provides business value and expertise to each member of the network.

ZapThink provides market intelligence to IT vendors and professional services firms that offer XML and Web Services-based products and services in order to help them understand their competitive landscape, plan their product roadmaps, and communicate their value proposition to their customers within the context of Service Orientation.

ZapThink provides guidance and expertise to professional services firms to help them grow and innovate their services as well as promote their capabilities to end-users and vendors looking to grow their businesses.

ZapThink also provides implementation intelligence to IT users who are seeking guidance and clarity into the best practices for planning and implementing SOA, including how to assemble the available products and services into a coherent plan.

ZapThink's senior analysts are widely regarded as the "go to analysts" for XML, Web Services, and SOA by vendors, end-users, and the press. Respected for their candid, insightful opinions, they are in great demand as speakers, and have presented at conferences and industry events around the world. They are among the most quoted industry analysts in the IT industry.

ZapThink was founded in October 2000 and is headquartered in Baltimore, Maryland. Its customers include Global 1000 firms and government organizations, as well as many emerging businesses. Its analysts have worked at such firms as IDC, marchFIRST, and ChannelWave, and have sat on the working group committees for standards bodies such as RosettaNet, UDDI, and ebXML.

Call, email, or visit the ZapThink Web site to learn more about how ZapThink can help you to better understand how SOA will impact your business or organization.

ZAPTHINK CONTACT:

ZapThink, LLC
108 Woodlawn Rd
Baltimore, MD 21210
Phone: +1 (781) 207 0203
Fax: +1 (786) 524 3186
info@zapthink.com
www.zapthink.com

