

The ZapLetter

Monthly Research & Analysis on XML



Inside this issue:

RosettaNet Sharpens Focus	1
Chrome Systems Develops Automobile Configuration	1
A New Vision for the Accredited Standards Committee X12	4
TranXML Provides Cross-Industry Shipping and	5
Epicentric Develops User Interface Specification for Web Services	6
Jabber: A Real-Time, Streaming Protocol for XML	7
Asteria: Platform for XML B2B Infrastructure	8
Avinon Empower Business Users to Create Applications with Web Services	9
Cybermation Perks Up Enterprise Scheduling with Espresso	10
Storing Arbitrary XML Documents with XYZFind	11
SoftQuad Guides Content Developers in XML World	12
Altova Provides XML Developer Intelligence with XML Spy	14
Tamino Provides Enterprise Solution for XML Storage	15
Provisioning Internal Resources with ADPr and Business Layers	15
How to be Nimble at Querying Heterogeneous	17
Budgeting and Forecasting in XML: FRx	19

Tools & Technology

Tamino Provides Enterprise Solution for XML Storage and Retrieval



Software AG has entered the XML space with a bang. Long known as providers of enterprise-strength database technology in the form of the ADABAS database, Software AG has leveraged their knowledge and experience in developing data storage and integration systems to deliver the **Tamino** product – a robust platform for XML storage and retrieval.

The increasing emergence of XML is creating a demand for a means to reliably store and retrieve XML documents from a central repository or database. While relational database technology creates a context for data through tables, joins, and other relationships, they aren't capable of dealing with the hierarchical, metadata-enriched data that XML contains. In a relational system, data that has been extracted from the database loses its context without the relationships that give them meaning. RDBMS systems are simply not capable of representing XML information without inefficiency and loss of meaning. These systems would need to decompose an XML document into a number of interrelated tables. Queries, by necessity, would result in many relational retrieval and join operations, and would require intense processing before, during, and after data retrieval. In addition, most relational data stores cannot handle data with constantly changing structure, as is the case with XML. An RDBMS, however, needs a well-defined structure for each table, so any changes in document structure would require a new schema definition before the data can be stored and used. Similarly, object-oriented databases would require a new class definition. All this is not optimized for the normal use of XML.

This need for "native XML" data stores has sparked rapid growth in the space of metadata document storage. While there are many approaches to this problem, the general idea is that XML should be stored in a native format without being deconstructed into tables to fit

into an RDBMS system and without the inefficiency and slowness of Object-Oriented DBMS systems. XML databases can store and retrieve any well-formed XML document, even if the DTD or Schema of the document is not available.

First, the question should be posed: why would anyone want to store XML? In many instances, XML is transitory and not meant for long-term storage in a system. This is especially the case for messaging and many e-Business contexts where the XML message is intended to represent a transaction and not specific data to be archived. However, there are many areas where XML is being applied as a long-term storage format. In particular, content management, which requires the ability to retrieve heterogeneous data types, heavily utilizes XML as a final destination for content information. Even electronic transactions have a need for receipts and auditing to leave a "paper trail". All of this information needs to be stored in a permanent repository and retrieved by means of intelligent querying.

Enter the XML native database and Tamino. Software AG's Tamino is a native XML data store, or rather a "virtual DBMS" that includes a high-performance XML store. Not based on ADABAS, but using the same developers, Software AG developed a small and extremely fast XML engine. The Tamino system however, is not just a data store. It actually is comprised of five major parts that make up a system for XML storage and retrieval. Those five major components consist of the XML Engine, X-Node (a means for accessing non-XML data stores), Data Map (for getting data in and out), the XML store itself, and Xtension (a programmatic interface for accessing XML content.)

The built-in native XML data store is focused on speedy retrieval of XML-based data using HTTP and TCP/IP as its main access protocols. To

(Continued on page 2)

Tamino Provides Enterprise Solution for XML Storage and Retrieval (Cont.)

(Continued from page 1)

optimize the performance of what may potentially be a resource hog, Tamino makes use of XML-optimized technologies such as field compression, record caching, and physical disk I/O. Since XML itself is an extensible format, users can add information elements to XML documents without changing the underlying database structures. This dynamic extensibility of XML databases allows users to focus on making changes to documents, rather than relational database structures. Queries against the Tamino database can then be issued via the Internet using open standards such as Xpath and XQL. These queries can be stored in the browser interface as a simple caching mechanism. Data can be inserted into the database in XML format, and can likewise be retrieved in XML format, so there is a true "XML-in, XML-out" interface. The XML that is outputted can then be transformed via XSL into whatever final document format is required by a user.

Tamino's XML engine stores XML objects and retrieves them from their respective data sources using a multi-threaded process. It is designed for high concurrency, and leverages web protocols to integrate with existing web servers that support the Apache API, Microsoft's ISAPI, or Netscape's NSAPI. Tamino's built-in XML parser checks that documents follow the correct syntax of schema, and verifies the well-formedness of incoming XML objects. In order to optimize search time, the system performs indexing on elements and attributes, supports document compression, and allows full text retrieval. In addition to XQL and Xpath, Tamino supports full text search verbs additions to the Xpath specification and its own query language X-Query, which is based on Xpath. Storage and retrieval of international, multi-lingual documents is enabled by use of Unicode, which supports country-specific character sets.

The X-Node component of Tamino allows the system to integrate with external heterogeneous data residing in RDBMS and OODBMS systems. This integration with "legacy" data stores gives the system its virtual database moniker. By integrating these other stores, an XML output can be created by merging multiple sources from a single query. Access to these data stores is provided by ODBC and JDBC interfaces, which allows access most relational data stores such as Oracle, Microsoft SQL Server, and of course Software AG's ADABAS, as well as file systems and other document and hierarchical data formats. Tamino X-Node maps this data to XML structures, presenting heterogeneous data to the client application as if it were obtained from a single data source. Tamino thus acts as a central server for existing databases. The data that is accessed, however, remains in its original source. The information retrieved by the X-Node application is aggregated, but not moved, allowing information to retain its original structure. Individual objects can be stored in different data stores, based on the schemas defined in Tamino's Data Map.

Tamino Data Map provides a knowledge base for the Tamino system. Data Map contains XML metadata includ-

ing DTDs, schema, style sheets, relational data schemas, which define the rules for how XML objects are stored and composed. The Data Map also indicates how XML objects are mapped to actual data stores and data structures, whether they reside in the native XML store or in a relational database. Data Map schemas are created and configured by an administrator by means of a graphical tool integrated in the Tamino Manager GUI. Data Map supports storage and indexing of XML objects within Tamino XML Database, mapping of data to data structures, external databases, and X-Tension server extensions, and provision of user-defined, schema-based application logic.

Administration is offered through a central Tamino-Manager which is implemented as a client-server application and is integrated into the System Management Hub, Software AG's multi-platform environment for the unified management of Software AG products. The system also provides a web-based interface that allows

administrators to create databases, control server processes, perform back-ups, view and modify session parameters, monitor and maintain servers, and restore and load data. Data Map schema configuration for mapping of XML objects to data stores and X-Tension mappings are also performed through the Tamino-Manager. For basic administration functions, Tamino Manager uses an administration-agent component that must first be installed on Tamino XML Database servers.

Server and functionality extensions are defined through the Tamino X-Tension component. These functions handle data in application for situation specific capacities, and provide

seamless interchange with Tamino functions. Extensions can be programmed in COM-enabled languages such as Visual C++, and wizards and add-ins are provided for Microsoft's Visual Studio application. Incoming XML objects can be mapped to these functions using the Data Map service. These extension functions are installed in the Tamino XML Database server using Tamino Manager. In this manner, X-Tension allow users to write hooks or triggers to do what you want based on incoming XML data. Integration with the EntireX product line allows XML documents to be populated with calls from mainframe.

There is definitely a considerable amount of functionality contained in the Tamino product line. However, Software AG considers themselves to be an XML solutions company rather than XML storage or database company. Their primary goal is to XML enable systems. With these products, Software AG hopes to take any legacy system and can turn it into a web service, creating an end-to-end XML system. Software AG competes with many vendors in different ways. From an education point of view, they compete with relational database technology vendors who offer "XML Strap-ons". They also compete with a number of other XML database upstarts, some of which are also featured in this newsletter.

General availability for the product was in September 1999, and they now have about 400 customers worldwide. Tamino 2.3 was

(Continued on page 3)

Most relational data stores cannot handle data with constantly changing structure, as is the case with XML



Tamino Provides Enterprise Solution for XML Storage and Retrieval (Cont..)

(Continued from page 2)

released in May 2001, and greatly improved storage and footprint over previous versions. In previous versions, document validation really slowed down the product, but 2.3 raised performance back to acceptable levels. While there is currently no industry benchmark for XML database insert times, a client in Germany did an internal study that compared in speed for inserts and found Tamino to be quite speedy. A future November release will support XML Schema, XML Query, and further enhancements to the product.

Contact Information:

www.softwareagusa.com

Dedra Jones Dedra.Jones@softwareagusa.com

John Taylor john.taylor@softwareagusa.com



ZAPTHINK, LLC

681 Main Street
Suite 2-11
Waltham, MA 02451

Phone: 781-207-8534
Fax: 786-524-3186
Email: info@zapthink.com

*Find More XML
Research &
Analysis at
www.zapthink.com*



Continuous Innovation
through Continuous Thought

Subscription Form

Fee: ZapLetter is \$595 for 12 monthly issues. ZapLetter is distributed electronically. Please add \$95 for print version. Fill out and fax or mail this application to ZapThink. Order online at <http://www.zapthink.com>. Please call or email info@zapthink.com for reprint or back order information.

Name _____

Company / Organization _____

Address _____

Phone _____

Method of Payment

Check

Visa

MasterCard

American Express

Number of Issues _____

Credit Card # _____

Exp. date _____

Signature _____

ZAPTHINK, LLC

681 Main Street
Suite 2-11
Waltham, MA 02451

Phone: 781-207-8534
Fax: 786-524-3186
Email: info@zapthink.com

