



CMP

United Business Media

Network Computing®

09.16.2004 | WWW.NWC.COM

For IT By IT

[Product Analysis]

DON'T FEAR THE DATA

EII is expensive and complex, but its rewards can be great. We tested seven offerings; Composite's CIS set the bar high by making access to dissimilar data a snap **BY LORI MACVITTIE**

» Enterprise information integration's claim to fame is its ability to federate data—it provides a single point of access to disparate information sources. This reduces the complexity inherent in client applications attempting to join various data sources while offering another way to access that information. Sounds good, right? So why did more than 56 percent of those responding to our reader poll say they have no plans to deploy EII, even though 57 percent say their organizations would benefit from it?

There are a few reasons. If you're one of the 25 percent of respondents who cited complexity as a major barrier, you're right. These products were a bear to test. The one in four who said cost is a problem are correct, too—EII isn't cheap. But if you can get past those details, EII can deliver standardized access to heterogeneous data sources.

That's good news for the 30 percent of respondents who said access standardization is a reason to implement one of these monsters.



Composite Information Server's Composite Studio made it a breeze to configure and manage all of the EII functions provided by the program.

Because we feel your pain in these decisions, we took seven EII products for a spin. Our goals, we thought, were modest: To standardize access to our disparate data sources—Microsoft SQL Server, Oracle9i, SOAP and XML—while decreasing the complexity of custom application deployments and maintaining reliable access for our reporting tools, Excel 2003 and Crystal Reports 9.

We invited Actuate Corp. (which acquired EII vendor Nimble Technology in 2003), Avaki Corp., BEA Systems, Certive Corp., Cincom, Composite Software, Denodo Technologies, IBM, Ipedo, MetaMatrix, Snapbridge Software and XAware to submit products that would meet our needs. Seven of the 12—Cincom,

Composite, IBM, Ipedo, MetaMatrix, Snapbridge and XAware—agreed to participate. Avaki bowed out, claiming it wasn't a good fit. Denodo, citing a planned Q4 upgrade to its platform, declined to submit its current product. Certive agreed to participate but then mysteriously disappeared. BEA declined, then reconsidered, but finally declined without citing a reason. Actuate did not respond to our invitation. We ran products from the brave combatants through a gauntlet of integration challenges in our NWC Inc. business applications lab in Green Bay, Wis.

Our EII Aim

Our goal was to evaluate how well the EII products integrate data sources, manage the system and configure client access (see "How We Tested EII Suites," page 48, for

details on our test setup). We wanted a single access point for NWC Inc. data stored in Oracle9i (order details), SQL Server 2000 (shipping details) and XML files (inventory details). Out-of-the-box support for these sources means fewer deployment headaches. All were successful, though almost all required an RDBMS client on the EII server.

The ideal EII platform should also support legacy apps in the form of ODBC access. In addition, the EII platform must perform minimal data transformations, such as type conversion, and offer some form of caching. The platform also must be bidirectional: It should update data sources *and* read from them, and it must perform basic cost-based query optimizations.

We quickly discovered the products fall into two cat-



REAL-WORLD
LABS®

REPORT CARD

EII Suites

	Composite Software Composite Information Server 2.5	IBM DB2 Information Integrator 8.1	MetaMatrix Server	Snapbridge Software FDX Information Server 3.5	XAware XA-iServer 3.52	Ipedo XML Intelligence Platform 3.5	Cincom Systems Tiger
DATA INTEGRATION							
Algorithmic optimizations (10%)	5	4	3	0	0	1	0
Structured data sources (10%)	4	4	4	4	4	3	3
Unstructured data sources (10%)	4	5	2	3	2	1	2
Framework for extending data-source support (5%)	5	4	3	4	2	4	3
DATA MANAGEMENT							
Data modeling (10%)	5	4	4	1	2	2	1
Security (10%)	4	3	4	2	3	2	2
Transformation (5%)	4	2	3	3	3	3	1
INTERFACES							
Output (10%)	5	2	3	4	3	2	2
Standards support (10%)	5	4	3	3	2	2	3
ADMINISTRATION							
Management (4%)	4	4	3	3	2	2	2
Caching (2%)	5	2	1	3	2	3	1
Configuration (2%)	5	4	4	3	4	4	2
Platform deployment (2%)	4	5	4	5	5	5	2
PRICE (5%)	3	4	3	4	4	1	1
RELIABILITY (5%)	5	5	3	3	3	3	2
TOTAL SCORE (100%)	4.49	3.73	3.20	2.74	2.50	2.17	1.83

A≥4.3, B≥3.5, C≥2.5, D≥1.5, F<1.5 A-C GRADES INCLUDE + OR - IN THEIR RANGES. TOTAL SCORES AND WEIGHTED SCORES ARE BASED ON A SCALE OF 0-5.

A⁻

B

C⁺

C

C⁻

D

D

DATA MANAGEMENT rates support of data sources, ability to maintain the integrity of database security, availability of encryption of connections to data sources and the audit/logging capabilities that are paramount to achieving SOX compliance.

ADMINISTRATION covers connector configurations, alerts and notifications, data modeling techniques, caching capabilities and metadata management. This category is graded based on the administration and management of the product. Connector configuration, alerts and notifications, data-modeling techniques, caching capabilities and metadata management determine this grade.

PLATFORM DEVELOPMENT is based on the standards (interfaces) and languages supported for end-user/developer-facing applications. This includes support for reporting and BI tools as well as standard development environments such as .Net- and J2EE-based suites.

RELIABILITY evaluates our lab experience. Products that crashed, locked up, failed to respond to queries or otherwise caused us agita were deemed less reliable than those that ran consistently.

Customize the results of this report card using the Interactive Report Card®, a Java applet, at www.nwc.com.

egories—those that grew from an RDBMS background (the software from Cincom, Composite, IBM and MetaMatrix) and those that emerged from the XML world (the products from Ipedo, Snapbridge and XAware). This difference radically affected the products' capabilities on the server side, including their caching, data management, metadata and modeling, and query-optimization features, and on the client side, in regards to ODBC/JDBC connectivity.

No Free JDBC

Most of the products—Cincom's is the exception—are J2EE applications, and we discovered to our chagrin that Microsoft does not let vendors distribute its SQL Server JDBC client with their products. You can download it for free from Microsoft (go to www.microsoft.com/downloads and search for JDBC), but it can't be included with J2EE apps. Cincom provided a SQL Server adapter, but its Tiger failed miserably at dealing with SQL Server 2000. Seems its adapter is broken, and we needed a Cincom-provided work-around to integrate the Microsoft RDBMS into its system. We weren't impressed with the flubbing of such a core feature nor with the work-around, for which we had to construct a SQL query just to add connectivity to the system.

Metadata support levels vary and depend entirely on the JDBC driver used by the platform to access the RDBMS. Snapbridge's JDBC driver is not a true JDBC implementation; it's a CSV-based interface, which means that no metadata—information about columns, data types, keys and constraints, for instance—can be retrieved from RDBMS data sources. The CSV format gives you only field names and values, and doesn't tell you that FieldA is a string or FieldB is a decimal. That information is important when you join tables and for transforming data to

other formats. We could swap in a different, full-fledged JDBC driver.

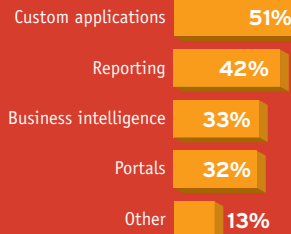
Nearly every product we tested—the exception was Ipedo's Information Hub—let us standardize on an XML interface, while supporting our legacy reporting applications, Excel 2003 and Crystal Reports 9. Ipedo does not provide ODBC connectivity, and discussions with the company on this subject yielded an XML-centric approach to solving the problem. Crystal Reports 10 and Excel 2003 support XML as a data source, and Ipedo suggested that the lack of ODBC support would be less of a problem as companies migrate to the latest versions of both products.

Although the other entries all supported ODBC, they did so with varying degrees of implementation ease and a variety of architectures. Composite, IBM and Cincom provide client software that make deployment a breeze. Administrators familiar with installation of ODBC client drivers and configuration of DSNs (data source names) will find the process familiar and simple. MetaMatrix and XAware use third-party OpenRDA ODBC drivers, and to make things more complicated (and pricey)—XAware requires a separately purchased JDBC server before client connectivity could be achieved. Configuration was tricky at times for both, requiring some extra work—not to mention extra cash—to provide client-side ODBC access. Snapbridge requires an ODBC-to-JDBC bridge and suggested we use EasyLink's product.

XML and SOAP support also varied widely, with the XML-centric products providing the most flexible feature sets. Composite offers the most comprehensive set of XML and SOAP features of the three RDBMS-based products, providing the point-and-click SQL-to-XML/SOAP abilities we craved. However, data transformation using Composite's modeling tool, Composite

READER POLL

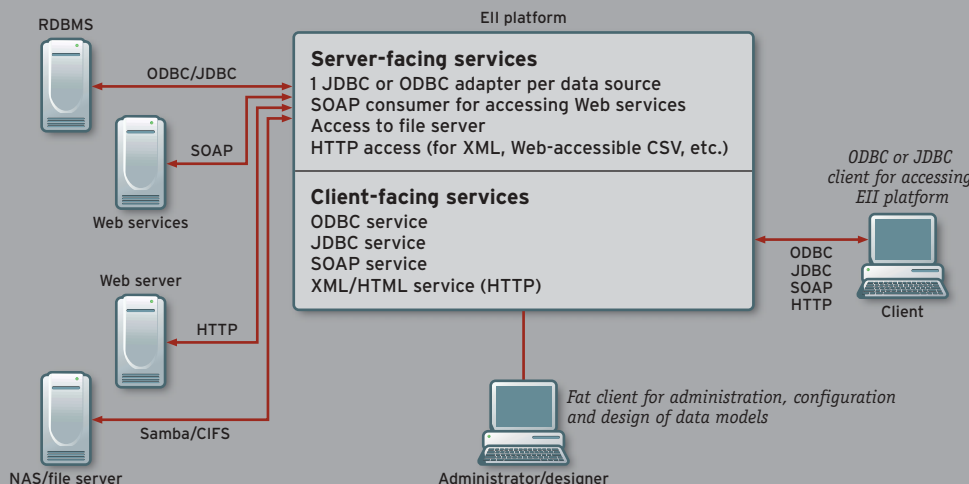
If your organization has an EII product, what types of applications primarily access data using it?



Multiple responses allowed.

Source: NETWORK COMPUTING Reader Poll, 504 respondents

How EII Works



- 1 Administrator configures the EII platform to connect to each data source.
- 2 Administrator or designer creates a data model from two or more accessible data sources.
- 3 Administrator publishes the data model for access via ODBC, JDBC, HTTP (XML), or Web services (SOAP). Access can be controlled using user and group access lists on the EII platform.
- 4 Client makes requests to the EII platform via one of the publishing methods.

Studio, was less flexible and required manual XSL creation, a process we dread.

Snapbridge's XML Studio is a sexy, easy-to-use tool that gave us everything necessary to achieve drag-and-drop XSL functionality, while Ipedo's XQuery Builder made creation and mapping of columns to XML a breeze. Cincom's solution to the XML and SOAP problem is to use its TotalXML product with NetBeans' IDE. Too developer-oriented, but effective.

Client access over SOAP was all over the map, with security provided through proprietary XML elements or HTTP Basic Authorization rather than via WSSE (Web Services Security Extensions) 1.0. IBM was the standards-based standout here, though its SOAP producer support of Web services is provided separately, over WebSphere Application Server, and a developer must take advantage of IBM's support for WSSE 1.0. We were disappointed by the other vendors' lack of standards support, but most said they would do better in future versions. We did enjoy the flexible SOAP

APIs provided by MetaMatrix, XAware and Ipedo, which let us create specific SOAP operations on a per-query basis and perform ad hoc queries using a standard SOAP operation. Ipedo required us to include XQuery, an XML-based human-readable query syntax derived from XPath 2.0, while the others let us submit straight SQL 92-compliant queries. We like SOAP's flexibility and hope to see support in other offerings sooner rather than later.

Caching capabilities also set products apart. We were most pleased with Composite's model, which includes automated invalidation on a user-configurable basis, and found IBM's use of MQTs (materialized query tables) simple and effective. MQTs are RDBMS tables that mirror a data source's structure but reside on the EII server, making queries and general access to that table faster. Caching support and maturity seemed to fall in line with the products' RDBMS versus XML origins: RDBMS products support caching far better than products on the XML side. XAware supports caching,

HOW WE TESTED EII SUITES

EII suites are all about federating data. To test the capabilities of these products, we put them to work in NWC Inc., our business application lab in Green Bay, Wis.

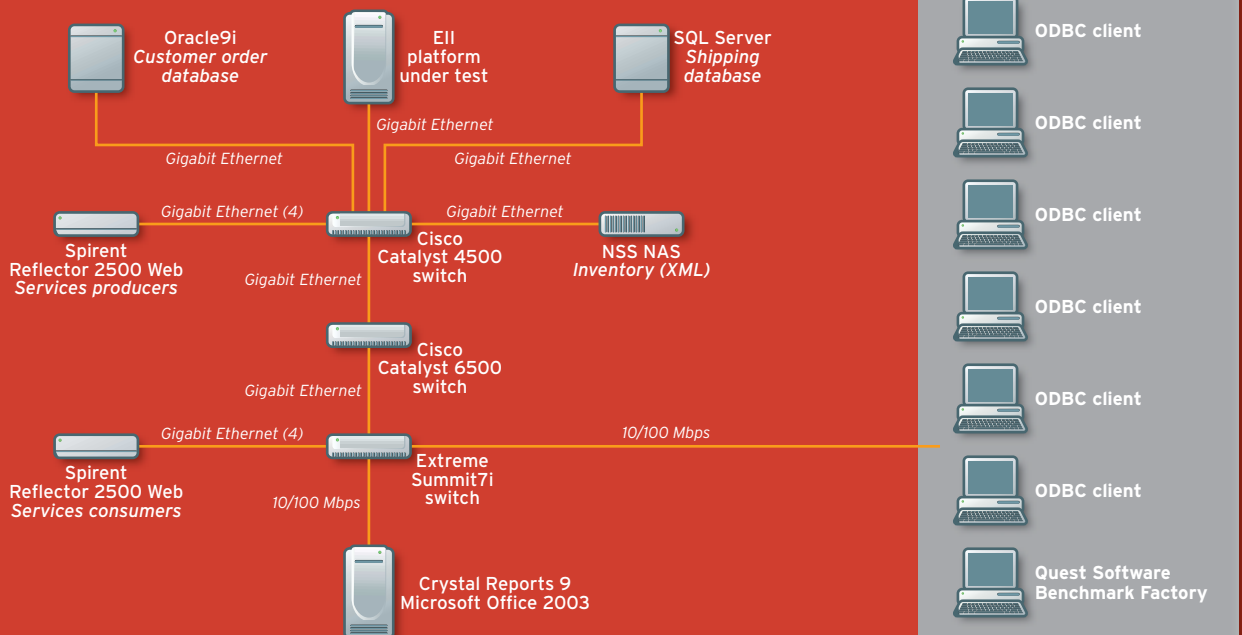
NWC Inc. is interested in consolidating access to its many disparate sources of data, including SQL Server, Oracle9i, flat files, documents and Web services. Our primary goal was to facilitate cross-database queries

for enterprise reporting (via Crystal Reports 9), while letting NWC Inc. standardize on an XML interface for custom development. NWC Inc. is a small enterprise, so the need to comply with Sarbanes-Oxley is looming, but not immediate. We took into consideration how each suite could assist us in addressing SOX compliance, but it was not a primary factor in our decision-making process.

We deployed each product on a Dell 2650 equipped with dual Xeon processors and 1 GB of RAM running Windows 2000 SP3. Each product had to communicate with SQL Server, Oracle9i, XML and flat files, as well as Web services (SOAP). We examined the configuration and management of the integration points, as well as management features, such as caching.

Each product had to connect to our

Performance Test Bed



for a fee (do you see a theme here?), but the configuration must be defined manually in a configuration file and then associated using an esoteric setting in a dialog cleverly hidden in the administrative console.

Ipedo's caching capabilities are similar to XAware's, using an XML collection or document in much the same way IBM or Composite use an RDBMS table. XAware provides caching of XML only and does not allow for direct caching of SQL result sets; this functionality is provided by shoving a SQL result set into an XML document and caching the XML. XAware says it will provide full support of SQL result set caching in a September release. MetaMatrix's caching is limited to small lookup tables, making caching larger tables a real pain, involving proprietary SQL extensions within queries. Cincom does not provide any caching, though the company says it could offer such capabilities with its professional services ... for a fee.

For some products, notably those from Composite and Snapbridge, caching was easy to manage and let us



www.nwc.com/locator

This article's ID# **1518f2**
 "Data Without Borders" is at ID# **1518f1**

automate cache invalidation, forcing a refresh of the data from its original source. Other products offer this capability but require manual creation of job schedules using cron or Windows scheduler. Snapbridge provides an interesting twist on caching that requires its proprietary XML-based XRAP language to indicate which cache a particular query result would be stored in—one minute, two minute, five minute, one hour. Composite and IBM simply cache results and let individual cache contents be invalidated.

Riddle Us This

We discovered that though some products (Composite's and IBM's) provide exemplary—and automated—

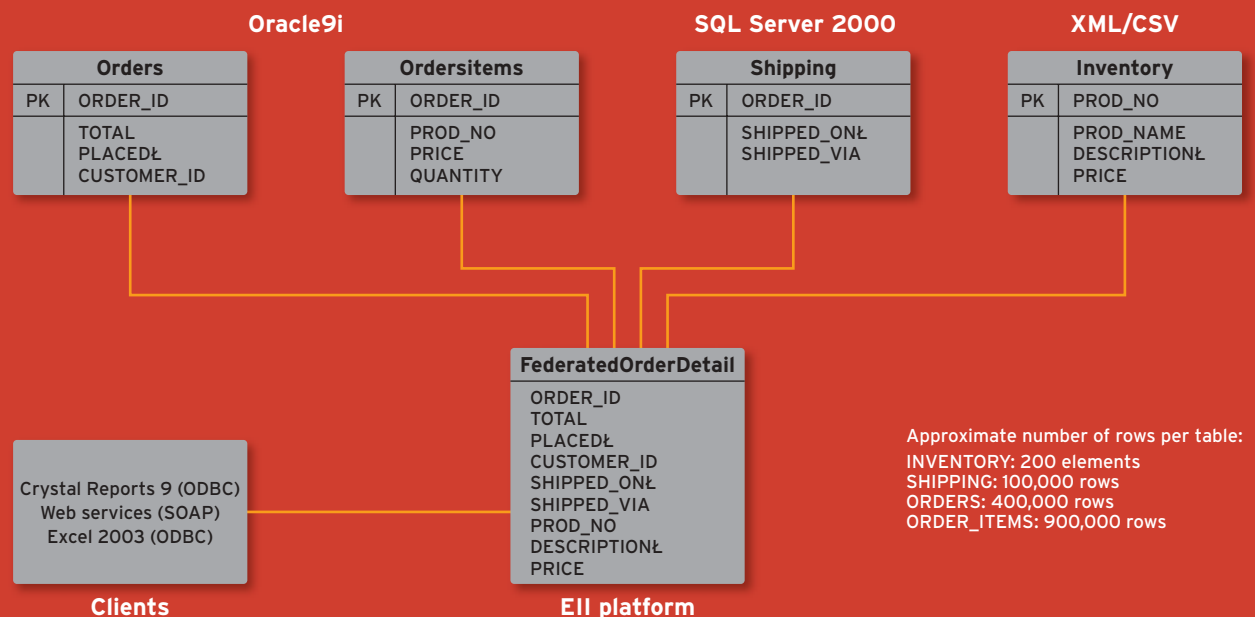
Oracle 9i and SQL Server databases, creating a singular view of three tables (orders, order items and shipping) and a CSV or XML file (inventory). The process of federating the data let us evaluate each product's data-modeling method and query optimization, down to the *join* algorithm used. To validate the federation, we used Crystal Reports 9 and Excel 2003 to pull data from the newly created data source.

Products were then required to integrate and federate data from Web services, CSV and XML files. With the inclusion of flat files, we could examine the mechanisms available for performing transformations on the data.

We tested each product's performance against a simple federation of two tables from Oracle and SQL Server. Where applicable, federated data was published as ODBC/JDBC, XML

and SOAP interfaces. Using BenchmarkFactory and five agent machines, we tested read-only transaction rates against the published ODBC data source. After the initial test runs, we enabled caching and ran the test again. Web services functionality was tested in the same manner—initial runs without caching enabled and subsequent runs taking advantage of the products' caching features.

Data Federation: Where Our Data is Located



support for query optimization, others (Ipedo's, Snapbridge's and XAware's) do not optimize queries at all, and MetaMatrix's architecture requires manual manipulation of queries to eke the best performance out of the system. Composite and IBM offer detailed query plans, and we were pleased with IBM's explanation of the *join* algorithms available. The architecture of the XML-based products makes it nearly impossible for them to take advantage of native database capabilities to join disparate sources, though XAware does provide basic optimizations, such as query-join blocking and push-down. Mechanisms for trimming down result sets in Snapbridge and Ipedo to reduce load on databases are somewhat painful, requiring the creation of XQuery code (Ipedo) and XSL (Snapbridge) to do what Composite and IBM do easily.

As you may have guessed, Composite Information Server (CIS) takes our Editor's Choice award, with IBM's DB2 Information Integrator (DB2II) coming in second.

MetaMatrix made a decent showing as well. See our pricing chart (page 57) for the product's list pricing as configured for our tests.

Composite Information Server (CIS) 2.5 CIS was the first product we tested, and it set the bar. We ran into problems now and then, such as with the SQL Server JDBC connector, but worked through them quickly—and small, irritating glitches were not unique to CIS.

Out of the box, CIS provides JDBC connectors for all products (with the exception of Microsoft, of course), and adding additional functionality through other drivers is a simple process; we added CSV during testing. Configuration is a breeze; our biggest complaint about the organization of projects and federated data is that it's easy to forget to create a federated data source in a shared folder. CIS' security model is role-based and flexible, but we were blown

EII VENDORS AT A GLANCE

PUBLIC COMPANIES

Company name (stock symbol)	Year founded	Product name	Product line launched	Market capitalization as of Sept. 2 \$000	Current assets \$000	Current liabilities \$000	Net income (loss) most recent quarter \$000	Net income (loss) year earlier \$000
IBM CORP. (IBM)	1911	DB2 Information Integrator 8.1	2003	\$141,670,000	\$41,409,000	\$34,357,000	\$1,988,000	\$1,705,000

PRIVATE COMPANIES

Company name	Year founded	Product name	Product launched	Employees	Amount raised \$000	Key investors
CINCOM	1968	Cincom Tiger	2002	750	None	None
COMPOSITE SOFTWARE	2002	Composite Information Server 2.5	2003	40	\$17,350	APAX Ventures, Clearstone Ventures, DOT EDU Ventures, Lehman Brothers, Palomar Ventures
IPEDO	2000	XML Intelligence Platform 3.5	2002	40	6,000	AsiaTech Management, Draper, Fisher, Jurvetson
METAMATRIX	1998	MetaMatrix Server	1999	75	Undisclosed	Gateway Associates, Invus, Kleiner Perkins Caufield & Byers
SNAPBRIDGE SOFTWARE	2001	FDX Information Server 3.5	2002	20	7,000	JP Morgan, Mission Ventures
XAWARE	1999	XA-iServer 3.52	2001	23	Undisclosed	Gefinor Ventures, ITU Ventures

away by MetaMatrix's extremely granular security model, unmatched by the other products.

CIS provides ODBC drivers for client connectivity, and we had no problem hooking up to our federated source over Excel and Crystal. The caching configuration was just as easy, involving a simple point-and-click operation. Both CIS and Snapbridge invalidate cached result sets based on your configurations, but we preferred CIS' method of specifying the interval on each data set we cached rather than having to choose an interval based on the options supported by Snapbridge's product.

Caching performance over ODBC was impressive. CIS served up 3.3 TPS (transactions per second) before caching was configured, and we were stunned by the 63 TPS it churned out when our clients queried the cached result sets.

Configuring CIS to provide a SOAP-based interface to our federated data also was pleasantly innocuous;

just a few clicks of the mouse and it was ready for use. After selecting the option to create a Web service, we just needed to specify which data source to publish. The sole drawback to CIS' SOAP services is its required use of HTTP Basic AUTH rather than support for WSSE 1.0. Composite says it is working to comply with such standards.

On the federation side, CIS will aggregate only DOC/LIT-encoded Web services, which complies with the WS-I Basic Profile but leaves a lot of enterprise services unable to be federated into CIS. "Parameterization" of Web services-based queries required us to use SQL scripting; we much preferred MetaMatrix's easy, GUI method of achieving this.

CIS was on par with DB2II in terms of query plan analysis and optimization of queries. We enjoyed the visual breakdown of atomic queries and CIS' ability to present performance information on subqueries, something the other product tested could not do. CIS automatically determined the most efficient method of joining our order and shipping data, querying Oracle first and then using the Order_ID values as part of the predicate to retrieve the appropriate data set from the shipping table in SQL Server.

Composite Information Server 2.5. Composite Software, (650) 227-8200. www.compositesw.com



Other Products	Key customers
DB2 Information Integrator Classic Federation for z/OS 8.1	Indiana University School of Medicine, Mazda North America, Merrill Lynch, Pioneer Display Products, Taikang Life Insurance, West Corp.

Source: SEC filings, company reports, Yahoo.com

Other Products	Key customers
Cincom Encompass, Cincom Smalltalk, Control, Intelligent Document Solution, Knowledge Builder, Mantis, Supra Synchrony	Undisclosed
None	EDS, Fathom Online, Life Time Fitness, Starbucks, Visa
None	BT, Hewlett-Packard, Navio, U.S. Navy
MetaBase, MetaBase Modeler, MetaMatrix XA Server	Credit Suisse, Defense Information Systems Agency, DoD Intelligence Agencies, First Boston, Government Communications Head Quarters (GCHQ) in the U.K., Lockheed Martin, Merrill Lynch, Motorola, SAP, U.S. Army
XStudio Stylemaker, XStudi Pro XML Tools Suite, FDX Cross Media Server	Acadient, Greenspun Media Group, Los Alamos National Laboratory, Rotech Healthcare, San Jose State University
XA-Xchange Hubs for Insurance and Financial Services, XA-iServe for Criminal Justice Exchange, XA-Node for Environmental Information Exchange	Merrill Lynch, Oppenheimer Funds, State of Nebraska, Unisys, U.S. Air Force, U.S. Department of State

Source: Company reports

IBM DB2 Information Integrator (DB2II) 8.1 DB2II was a snap to configure.

It installs on top of a DB2 database and capitalizes on many features available within DB2, such as client connectivity and caching functionality. Using Wrappers to federate data from Oracle9i and SQL Server 2000 was a pleasant experience that IBM has made exceedingly easy by automatically generating virtual views from the metadata it extracts via the JDBC drivers. Including hierarchal data, such as our XML-based inventory list, followed the same process, and we were up and querying our data sources in short order after installation.

DB2II presents all federated data as virtual tables, which let us query XML documents using SQL rather than XPath. This was a definite plus—database administrators are unlikely to be well-versed in XPath syntax. We quickly created a virtual view of our order data from Oracle and our shipping data from SQL Server in a visual design mode that included detailed descriptions of the different *join* algorithms available.

Once we set up the appropriate view, we easily configured an ODBC data source on our client machine and retrieved the data within both Excel 2003 and Crystal Reports 9. Upgrading client software from 7.2 to 8.1 provided a noticeable performance boost in data access, making it well worth the time investment.

After running our first set of ODBC-based perform-

ance tests, yielding 1.8 TPS under a load of 20 concurrent users, we configured DB2II to cache the result sets from both queries. DB2II uses MQTs for caching, essentially creating a local table within the required DB2 instance supporting DB2II, and then populating the table from the queries. By specifying “Data initially deferred refresh deferred” when creating the table, you can refresh the data from the source at any time by issuing the proper command. This makes it possible to schedule updates to the cache at any interval. Once we’d populated the cache tables, we reran the performance test and got 2.7 TPS using the cached tables. That may not sound like much of a performance gain, but considering the number of products we tested that do not support any caching for RDBMS data sources, we were pleased.

IBM says the forthcoming version of DB2II, 8.2, will

include significant caching enhancements and a Wrapper specifically for including Web services as a federated source of data. Version 8.1 offers a mechanism for federating Web services, but it’s not as neatly integrated as it will be in the next release.

DB2 Information Integrator 8.1. IBM, (800) IBM-CALL. www.ibm.com

MetaMatrix Server MetaMatrix Server deploys by default on a BEA application server at no additional cost. With the included JDBC drivers from DataDirect,



connecting to NWC Inc.’s relational database was a cinch. MetaMatrix Server uses a virtual database model to federate data sources, and we had no problem creating a virtual database that included our specified targets.

EII Suite Features

	Cincom Systems Tiger	Composite Software Composite Information Server 2.5	IBM DB2 Information Integrator 8.1	Ipedo XML Intelligence Platform 3.5	MetaMatrix Server	Snapbridge Software FDX Information Server 3.5	XAware XA-iServer 3.52
Platform support	Microsoft Windows, Sun Solaris	Linux, Microsoft Windows, Sun Solaris	IBM AIX, Linux, Microsoft Windows, Sun Solaris	IBM AIX, Linux, Microsoft Windows, Sun Solaris	Linux, Microsoft Windows, Sun Solaris	IBM AIX, Linux, Microsoft Windows, Sun Solaris	IBM AIX, Linux, Microsoft Windows, Sun Solaris
Data management							
Replication	N	N	Y	Y	N	Y	N
Transformation	Y	Y	Y	Y	Y	Y	Y
Output type	SQL Result Set, XML	CSV, e-mail, HTML, SQL Result Set, Web service (SOAP), XML	HTML, SQL Result Set, Web service (SOAP), XML	CSV, HTML, Web service (SOAP), XML	SQL Result Set, Web service (SOAP), XML	CSV, HTML, Web service (SOAP), XML	CSV, fixed length, HTML, user-defined separator format, SQL Result Set, Web service (SOAP), XML
Protocol access	DOM, JDBC, ODBC, SAX	Files (delimited or XML), LDAP (as data source), JDBC, ODBC, URLs, WSDL/SOAP	JDBC, ODBC, WDSL/SOAP	JDBC, ODBC, WDSL/SOAP, Xquery	JDBC, JMS, ODBC, WDSL/SOAP, Xquery	JDBC, ODBC, WDSL/SOAP	COM, EJB, JDBC, JMS, ODBC, Servlet, WDSL/SOAP
Metadata management	Y	Y	Y	Y	Y	Y	Y
Multisource update	Y	Y	N	N	Y	Y	Y
Stored-procedure support	Client	Client, server	Client	Client, server	Client, server	Client, server	Client, server
Structured data							
XML	Y	Y	Y	Y	N	Y	Y
Web services	None	Consumer, provider	Consumer, provider	Consumer, provider	Provider	Consumer, provider	Consumer, provider
LDAP	N	Y	Y	Y	Y	Y	Y
Relational (out of box)	IBM DB2, Microsoft SQL Server, MySQL, ODBC, Oracle, Sybase	IBM DB2, JDBC, Microsoft SQL Server, MySQL, ODBC, Oracle, Sybase	IBM DB2, JDBC, Microsoft SQL Server, MySQL, ODBC, Oracle, Sybase	IBM DB2, JDBC, Microsoft SQL Server, MySQL, Oracle, Sybase	IBM DB2, JDBC, Microsoft SQL Server, MySQL, ODBC, Oracle, Sybase	IBM DB2, JDBC, Microsoft SQL Server, MySQL, ODBC, Oracle, Sybase	IBM DB2, JDBC, Microsoft SQL Server, MySQL, ODBC, Oracle, Sybase
Message queues	None	None	MQ	MQ, JMS, Tibco RV	JMS	MQ, JMS, Tibco RV	MQ, JMS, Tibco RV
Unstructured data	Flat files (CSV)	Excel 2003, flat files (CSV), Extender API, Web pages	Excel 2003, flat files (CSV), Extender API, Web pages	Excel 2003, flat files (CSV), Extender API, Web pages	Excel 2003, flat files (CSV), Extender API, Web pages	Excel 2003, flat files (CSV), Extender API, Web pages	Excel 2003, flat files (CSV), Extender API, Web pages

Y=Yes, N=No

MetaMatrix Server processes queries in parallel but, like Composite and DB2II, can push down atomic queries to the appropriate RDBMSs. Query-plan analysis isn't visible to the administrator but can be logged and examined by a database administrator for optimization or sent to MetaMatrix for evaluation as part of the service contract.

MetaMatrix Server does not support federation of XML files or Web services, but it does support CSV files. We manually created the appropriate descriptor file, indicating details like which separator character was used, and then federated our data. Support for XML files is planned for an upcoming release.

One problem with IBM's and Composite's systems is that you need to know how sources are federated to have any hope of avoiding unwieldy references within SQL queries. In lieu of being very careful during initial configuration of the virtual database, queries exposed by MetaMatrix Server required us to craft our queries such that both catalog and schema, rather than merely the latter, were needed to reference the tables.

MetaMatrix Server offers a snazzy Web console for constructing queries and testing virtual databases. We liked the tool because it assisted in building queries and ensuring that our federation was correctly configured.

Configuration of the ODBC client was straightforward as well, though as with the products from Snapbridge and XAware, a third-party OpenRDA client is needed. Initially, we ran into problems trying to test

performance, but after shipping the log files to MetaMatrix engineers, we increased the thread pool and maximum threads within the MetaMatrix Server. Once we'd done that we could support our 20 concurrent user loads.

MetaMatrix Server. MetaMatrix, (212) 514-7112. www.metamatrix.com

Snapbridge Software FDX Information Server 3.5 FDX is an interesting combination of flexibility and confusion. Off-the-shelf support for major RDBMSs,



including IBM DB2, SQL Server, MySQL, Oracle and Sybase, is included and, as a bonus, the package let us configure data-source connection pooling within the included Tomcat application server. The confusion set in when we were presented with two different methods of federating data and only learned that one of them is not intended for "real" federation after experiencing difficulties during the configuration process.

FDX made it, well, a snap to publish federated data sources as XML, HTML, CSV and fixed column output via JDBC or SOAP protocols. HTML-published access to federated data included rudimentary formatting, and FDX's point-and-click method of transformation from row/column data into HTML was easier than DB2II's manual transformation requirements.

Federation of simple data sources (two or fewer) was simplified somewhat by FDX's builder, but we quickly discovered that this capability is limited. If sample queries didn't return data, the federated source was useless. Likewise, joins failed if key columns from both data sets did not yield matching results. Using FDX's design tool made it much easier to federate data sources but required that we have a working knowledge of XPATH, XSL and Snapbridge's proprietary XRAP language. XRAP provides great flexibility in terms of presentation, but the coding required to federate sources is unwieldy compared with the methods offered by Composite, IBM and MetaMatrix.

FDX performs joins on the server unless administrators have the foresight to craft the underlying queries properly, such as by querying the initial data source and using data returned to craft the second query, for example, using the Order-ID from a query to Oracle as part of the predicate in a query to SQL server. FDX does not provide detailed query analysis or assistance in this area; we recommend that a database administrator help you optimize joins across various data sources.

We accomplished all FDX administration using its Web console; this was unique among products tested. All the others required fat-client (Java) applications to fully administer and configure servers. We preferred the complete integration of all administrative and publishing capabilities in a single Web environment.

FDX Information Server 3.5. Snapbridge Software, (877) SNAPBRIDGE, (760) 431-3036. www.snapbridge.com »

EII Suite Pricing

Our scenario: Two-CPU server; two JDBC/ODBC data sources; one XML data source; one CSV data source

	Price as tested	Breakdown
Cincom Tiger	\$90,000	\$20,000 per CPU; \$10,000 per CPU for XML support; \$10,000 per data source type
Composite Software Composite Information Server 2.5	\$140,000	\$70,000 per CPU
IBM DB2 Information Integrator 8.1	\$100,000	\$20,000 per CPU; \$15,000 per data source
Ipedo XML Intelligence Platform 3.5	\$144,000	\$72,000 per CPU
MetaMatrix Server	\$90,000	\$40,000 for as many as 2 CPUs and 1 to 5 data sources; \$50,000 for MetaBase and the modeler
Snapbridge Software FDX Information Server 3.5	\$9,998	\$4,999 per CPU
XAware XA-iServer 3.52	\$52,000	\$30,000 per CPU; \$5,000 per data source type; \$2,500 per client per adapter type

XAware XA-iServer 3.52 XAware has coined the most amusing EII term: *functoids*. This is a comical name for a powerful concept that offers an easy mechanism via Java or JavaScript to extend the functionality of its XA-iServer.

XA-iServer is installed into JBoss by default but can be deployed to BEA, IBM and Oracle application servers. One advantage of being deployed into an application server is that XA-iServer can take advantage of JDBC connection pooling, which ostensibly means increased performance and less strain on the server. Under a load of 20 concurrent users, however, no product we tested could keep the CPU utilization under 100 percent, so we expect that using JDBC connection pooling would only increase the performance of XA-iServer, which managed 3.3 TPS via ODBC and 2 TPS via a SOAP request. Caching is not available for ODBC-based connections, but caching of SOAP requests increased performance to 5 TPS under the same load.

XA-iServer offers a unique reusable model for building a federated data source, comprising BizDocuments (what the end user sees), BizComponents (bits of data) and BizDrivers (connections to data sources). This model supports only nonstructured presentation; the SQL-based interface requires a virtual database that is separated from hierarchical data-source views, meaning if we wanted to support both SQL result sets and XML presentation of the same federated data, we had to build it twice.

A cool feature of XA-iServer is its ability to present streaming of XML both as input and output. Streaming XML means that the product can handle extremely large files while using minimal resources. This is unique and—after watching Ipedo choke on an 8-MB CSV file with an out-of-memory error—useful.

XA-iServer 3.52. XAware, (866) 607-7706, (719) 884-5400. www.xaware.com

Ipedo XML Intelligence Platform (XIP) 3.5 XIP is the only product to embrace the XQuery standard.

XQuery is an ugly beast—normally we like standards, but in this case we're making an exception. Still, we must admit Ipedo's XQuery Builder visual tool makes XQuery queries less of a chore to construct.

Ipedo has embraced the XML world and, given its origins as an XML database provider, this makes sense. XIP uses Oracle's thin JDBC client to federate data, meaning that no metadata can be returned. So it was up to us to build federated documents. Ipedo deploys by default on Tomcat/ Apache but supports implementation within the leading J2EE containers (IBM and BEA).

ODBC support is, however, nonexistent. To access data federated within XIP we had to use Crystal's

ODBC XML driver and XIP's WebDAV interface. It worked, but it was a distinct departure from the norm.

One thing we like about XIP is its easy mechanism for parameterizing queries (letting an end user pass a value). This is something Snapbridge couldn't do effectively and that was exceedingly difficult for most other products we tested.

Ipedo XML Intelligence Platform 3.5. Ipedo, (800) 601-3962, (650) 306-4000. www.ipedo.com

Cincom Tiger Tiger (short for **Totally Integrated Enterprise**) is an interesting mix of old and new technology, proving that data federation has been around

for some time. Cincom relies heavily on DOS batch files and textual configuration files to be modified in order to configure connectivity to data sources. It also uses a proprietary database for metadata storage and offers a unique view on federation through proxies.

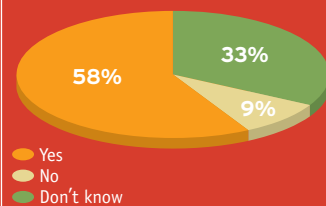
Each data source we federated required that we create a proxy. After digging through the manuals to find the required syntax, we determined that proxies are really Tiger terminology for virtual views, a concept used by almost every product we evaluated. We automatically generated a proxy for Oracle9i, but discovered that we could not do so for SQL Server. Cincom immediately responded to our queries on this, saying that its SQL Server adapter is known to have "issues." The work-around? Generate the proxy manually. In fact, most functionality within Tiger is configured and managed manually, with a heavy emphasis on understanding basic object-oriented design concepts. Cincom would do better to hide this complexity from administrators and designers; a full makeover of the system is in order, and Cincom says it is working on a long list of updates and enhancements.

We also had to fight with transformation of data types, as the Order-ID in our shipping table on SQL Server is not the same type as the Order_ID in our order table in Oracle9i. Finally, we were able to install the client software, configure an ODBC DSN and successfully query our federated data source from Excel 2003 and Crystal Reports 9. We couldn't get our performance test clients to do so reliably, though.

Cincom Tiger. Cincom Systems, (800) 2CINCOM, (513) 612-2300. www.cincom.com **NWC**

READER POLL

Would your organization benefit from EII?



Source: NETWORK COMPUTING Reader Poll, 504 respondents



LORI MACVITTIE is a NETWORK COMPUTING senior technology editor working in our Green Bay, Wis., labs. She has been a software developer, a network administrator and a member of the technical architecture team for a global transportation and logistics organization.

Write to her at lmacvittie@nwc.com. Post a comment or question on this story at www.nwc.com/go/ask.html.