News: XML—The Center of Attention Up and Down the Stack

Greg Goth

Computing historians might someday refer to 2005 as the year of XML. Many have anticipated the promise of XML for almost a decade, but major industry players finally accepted in 2005 that the nuts and bolts of XML processing in the heart of their networks will demand new ways of thinking.

Furthermore, in a reprise of many an industry scuffle, Microsoft started touting its vision of XML-based applications for end users, while much the rest of the industry rallied under the banner of a competing standard written by the Organization for the Advancement of Structured Information Standards (http://www.oasis-open.org) (OASIS) consortium. The battle for the XML desktop might be played out in international standards discussions over the next two years.

"It looks like 2005 was definitely the year [XML] went from, 'It looks like it's a major trend' to 'There's no technology that's going to replace it,'" says Ron Schmelzer, senior analyst and founder of ZapThink. This analysis firm specializes in Web services and service-oriented architecture (SOA) issues. One ZapThink report estimates that XML traffic on corporate networks will grow from approximately 15 percent in 2004 to just under 48 percent by 2008. Back in 2002, the firm estimated XML traffic accounted for just 2 percent of network traffic.

"Already on some networks, one-quarter of all traffic is XML," Schmelzer says. This is due to both XML's pervasiveness and its inefficiency. "It's in security systems, enterprise applications, business process tooling, and document formats," says Schmelzer. The end result of this pervasiveness can be a network clogged with XML messages. Because XML is text-based and both human- and machine-readable, its messages are between 10
and 50 times larger than similar binary messages. Additionally, Schmelzer says completing communication between XML-capable endpoints requires increased security owing to flexible, loosely coupled service-oriented networks.

"The whole idea of Web services and SOA is that we're basically mixing and matching applications from different parts of the network," he says. "How do you make sure you have the right people accessing the right functionalities?" He also says it's not just about preventing hackers. If every message on the network needs ID control—decryption, validation, parsing, and so forth—it increases message size and the required processing and bogs down application servers. "It's a real challenge and a real problem," he says.

**Time to hit the accelerator**

In 2003, IBM engineers tested XML efficiency (ftp://ftp.software.ibm.com/software/websphere/partners/datapower_whitepaper.pdf)—or rather its inefficiency—through several configurations. They discovered that large latency times resulted in transforming XML to browser-ready HTML. One experiment, using Extensible Style Language Transformation and Xalan (also XSLT-based) transform engines, resulted in a 95 percent overhead rate—that is, only 5 percent of the server's capacity was left for the actual application. Another configuration, which used a compiled Java-to-bytecode transform strategy, resulted in an 87 percent overhead rate.

As the size and complexity of the engineers' benchmark XML processes increased, they discovered that even optimized approaches resulted in server meltdown. According to IBM, a calculation-intensive benchmark operating on a 1-Mbyte XML document reduced application throughput to the point where the application became unusable. The transaction rate was down to two transactions per second and the response time was pushed up to a full 23 seconds. IBM thus reported that the Xalan transform engine was too slow to make meaningful measurements.

Schmelzer says the quandary over how to improve XML processing made its way into the market at large from mid-2004 to early 2005, with the vanguard mulling over the processing-intensive problem of security in Web services and SOA-capable applications.

By mid-2005, a niche technology called the XML acceleration appliance started showing traction. The technology moves the load of XML processing from the application server to a dedicated plug-and-play piece of hardware. In June, Cisco announced its Application-Oriented Network technology, which included XML processing capabilities.
In August, Intel announced it was purchasing XML appliance maker Sarvega. In October, IBM announced it was purchasing DataPower, another leading XML acceleration vendor.

ZapThink's Schmelzer says the adoption curve for the dedicated appliance is still on the left side of the bell curve, but the major vendors' entry into the market should push the technology far into the mainstream. "We've proved that [the dedicated appliance] is a viable technology. Now we're asking how to mass-market it, how to mass-produce it in the most viable way, which is why you're seeing all the jockeying for position."

Marie Wieck, vice president for IBM's WebSphere platform, says the company's decision to buy DataPower hinged on the smaller company's strong position in the acceleration market. It also hinged on the large-scale recognition that standards-based SOAs might provide significant competitive advantage for companies that can master multiple implementations of common data.

"Whether you're modernizing legacy applications or leveraging new Web services in a composite application, XML does tend to form the basis for the data movement," Wieck says.

She also contends that enterprise developers and IT managers recognize that new standards-based SOA initiatives such as Service Component Architecture and Service Data Objects (http://www-128.ibm.com/developerworks/webservices/library/specification/ws-scascosum/), have led to an environment where customers are trying to simplify technology's adoption and increase its value.

**Time to bang the drum slowly for middleware?**

Since the accelerator concept has gained legitimacy through the entry of the industry giants, a major question remaining is what will happen to the middleware market, long thought of as the critical sweet spot in bringing Web services and SOAs to fruition. Schmelzer says the industry might be on the brink of a tectonic shift in conventional wisdom.

"Do people need application servers or a so-called ESB [enterprise service bus] infrastructure to make SOA work, or can they use intermediaries?" Schmelzer says. "The intermediary approach might be the architecturally more elegant way of doing it. It's simpler, more flexible, there's no single point of failure, and it doesn't require proprietary
middleware." He says that appliance vendors are beginning to realize they can provide a lot of the functionality of what the so-called ESB can do. "People are offloading a lot of functions from application servers onto network appliances."

"A lot of the middleware vendors are still basically [involved in] EAI (enterprise application integration); they're not really providing loosely coupled services," says Schmelzer, which is why he thinks the movement toward SOA jeopardizes many middleware vendors.

Eric Newcomer, chief technology officer at Iona Technologies and a Web services veteran, says humming a dirge for middleware would be premature. "Everybody's always predicting the death of some old technology, and we still have mainframes and Cobol years after it was predicted they would die," Newcomer says. "I don't see much merit in these kinds of sweeping predictions." He says middleware vendors will likely need to explore how they fit into these environments, and they'll need to figure out how to exploit them in extended architectures while still providing other value to customers. "Middleware is a very broad area, and at the moment, these hardware vendors are only able to focus on certain aspects of middleware and not the whole picture."

Newcomer says the rapidly evolving landscape of data formats and protocols mandates that savvy developers and end users remain cognizant of both software and hardware solutions. "One of the things we do is take XML formats and map them to multiple transport and data formats, and that kind of flexibility is going to be very hard to achieve in hardware anytime soon, because data formats change and protocols change," he says. "These guys have to take a fairly securable slice of the functionality and put it into the hardware, so it won't be outmoded six months after somebody deploys it."

**XML living large publicly**

How best to deploy the rapidly expanding XML-based network is gaining new importance well beyond the standards conferences and corporate boardrooms. The discussion over the ultimate XML endpoint—client applications and how to deploy them—has entered the public arena to an extent that XML-savvy technologists never imagined.

In the US, the discussion has revolved around a decision by the Massachusetts state CIO (http://www.mass.gov/Aitd/docs/policies_standards/etrm3dot5/ETRM_v3dot5draft_information.pdf), Peter Quinn, mandating that as of 1 January 2007, archival documents must be saved in either the OASIS -developed, XML-based Open
Document Format (ODF) or Adobe's PDF. The decision has resulted in a tempestuous argument between ODF backers, including IBM and Sun Microsystems, and Microsoft, which dominates the productivity application market with its Office suite. Since the Massachusetts decision, Microsoft has proposed a competing XML productivity application (http://www.microsoft.com/presspass/features/2005/nov05/11-21Ecma.mspx) standard through ECMA International, the European association for standardizing information and communication systems.

"If you live in the Boston area, it's kind of funny to sit here and look through the Boston Globe. Once a week, if not more, there's an article talking about Microsoft and ODF and what's happening, something you would never expect to see in the mainstream press," says Mary McRae, manager of the technical committee administration at OASIS.

Even in the short term—considering just traditional productivity applications and componentized versions of them such as IBM's Workplace—the specter of two competing standards could result in just the opposite of SOA evangelists' intentions. We could end up with incompatible formats and schemas that don't simplify integration but instead open up subtle and dangerous technical and legal minefields over an increasingly complex network. Competing XML schemas might also inhibit the broad adoption of intelligent XML-based modules capable of operations such as routing purchase orders to the proper department based on price or catalog category.

Building operations such as these lets us "use Web services business-to-business capabilities and extend them out to the last mile, where supply chains have had the hurdle of reaching small and medium enterprises that don't have HTTP servers to do B2B transactions," says OASIS CEO Patrick Gannon.

Industry watchers will have to monitor multiple battlefronts to discern how things might play out. Gannon says OASIS has no plans to enter formal discussions with ECMA at the moment, though the International Standards Organization might request some sort of liaison if it ends up evaluating two separate proposals for essentially the same functions. While ECMA has just begun to consider Microsoft's specification, the ODF format has been under ISO consideration since September, giving it a head start with the organization many European nations look to as the bellwether for a given standard's legitimacy.

No matter what transpires, Wieck says the public and private-sector enterprises expecting improved efficiency and service capability through XML-based networks want results, not just rhetoric.
"I think they know this is sort of the natural evolution, but I don't think they're particularly tolerant," she says. "We'll eventually work it out because it's in nobody's best interest to support two standards, and customers expect us to make things easier."

Making things easier was what XML was supposed to be about from the beginning. It seems people both inside and outside the industry are finally coming to terms with how difficult "making things easier" has proven to be.

Related Links


DS Online's Web Systems Community, http://dsonline.computer.org/was

Cite this article: Greg Goth, "XML: The Center of Attention Up and Down the Stack," IEEE Distributed Systems Online, vol. 7, no. 1, 2006, art. no. 0601-o1003.